应用概率统计 第二十六卷
第二期 2010年4月
Chinese Journal of Applied Probability
and Statistics Vol.26 No.2 Apr. 2010

# 纵向数据单指标模型中参数的经验似然置信域 *

李高荣[1,2] 冯三营[3] 薛留根[1]

([1]北京工业大学应用数理学院, 北京, 100124; [2]华东师范大学金融与统计学院, 上海, 200241)

([3]洛阳师范学院数学科学学院, 洛阳, 471022)

**摘 要**

考虑纵向数据单指标模型, 针对纵向数据组间独立的特点, 提出了模型中未知参数的三种经验对数似然比统计量. 在适当条件下, 证明了所提出的统计量依分布收敛于$\chi^2$分布, 所得结果可以构造未知参数的置信域. 进一步证明了所提出的纠偏的经验对数似然比有许多优良的性质. 通过模拟研究对所提方法进行了说明.

**关键词**: 纵向数据, 单指标模型, 经验似然, 置信域.

**学科分类号**: O212.7.



## §1. 引 言

纵向数据广泛存在于医学、流行病学、经济学和社会科学等领域中, 已经成为统计学研究的热点问题. 考虑来自$n$个个体的数据, 其第$i$个个体具有$m_i$次观测$(i=1,\cdots,n)$, 总的观测数为$N=\sum\limits_{i=1}^{n}m_i$. 本文假定$m_i$是有界的, 该假定意味着总的样本容量$N$与个体数$n$是同价的量. 设$Y_{ij}$和$X_{ij}\in\mathbb{R}^p$分别是第$i$个个体的第$j$次观测$(j=1,\cdots,m_i)$的响应变量和协变量, 它们具有下面的关系形式:

$$Y_{ij}=g(X_{ij}^{\rm T}\beta)+e_{ij},\qquad i=1,\cdots,n,\ \ j=1,\cdots,m_i, \tag{1.1}$$

其中$\beta$是$p\times 1$的未知参数向量, $g(\cdot)$是一个未知的一元联系函数, $e_{ij}$是随机误差, 记第$i$个个体的随机误差向量为$e_i=(e_{i1},e_{i2},\cdots,e_{im_i})^{\rm T}$, $\{e_i,i=1,\cdots,n\}$相互独立, 且$\mathsf{E}(e_i|X_i)=0$, $\mathsf{Var}\,(e_i)=\Sigma_i$ (正定阵). 考虑到模型的可识别性, 我们要求$\|\beta\|=1$, 这里$\|\cdot\|$表示Euclidean模.

模型(1.1)是一个重要的统计模型, 它可以减少在拟合多元非参数回归函数时所谓的"维数祸根"问题. 关于模型(1.1)在独立数据情形下, 许多作者在不同的假设条件下对其做了一定研究, 并提出了一些研究方法. 例如, Ichimura[1]和Härdle等[2]分别研究了模型

*中国博士后科学基金(20080430633)、上海市博士后科研资助计划(08R214121)、国家自然科学基金(10871013)、北京市自然科学基金(1102008)、国家教育部博士点专项基金(20070005003)、北京市属市管高等学校人才强校计划、北京工业大学博士启动基金和河南省自然科学研究(2008B110009)资助.

本文2008年9月8日收到, 2008年10月30日收到修改稿.

中参数的最小二乘估计的渐近性质, 并且讨论了联系函数的非参数估计的窗宽选择问题; Stoker[3]和Härdle等[4]提出了平均导数方法, 相关文献还有[5, 6]; Powell[7]描述了密度加权平均导数估计方法; Xue and Zhu[8]构造了模型中指标参数向量的经验似然置信域. 对于纵向数据情形, Chiou and Müller[9]在研究响应变量$Y_{ij}$和协变量$X_{ij}$之间关系的统计推断时, 介绍了一种边缘建模方法, 并提出了估计的估计方程(EEE)方法, 进一步研究了广义线性混合模型.

本文采用Owen[10, 11]提出的经验似然方法构造了纵向数据单指标模型(1.1)中兴趣指标参数$\beta$的三种经验对数似然比统计量, 证明了所提出的统计量具有渐近$\chi^2$分布, 所得结果可以构造参数的置信域. 经验似然方法在构造置信域方面有许多优点, 首先无需对渐近方差进行估计; 其次所得到的置信域的形状由数据自行决定等. Kolaczyk[12]将经验似然方法应用到广义线性模型. Shi and Lau[13]和Wang and Jing[14]利用经验似然方法研究了具有固定设计的部分线性模型中参数的置信域. 最近, Xue and Zhu[15]和Zhu and Xue[16]研究了部分线性单指标模型中参数的经验似然置信域.

## §2.　方法与主要结果

### 2.1　参数的经验对数似然

假设记录数据$\{(X_{ij},Y_{ij}), i=1,\cdots,n, j=1,\cdots,m_i\}$由模型(1.1)产生. 本文主要是构造未知参数向量$\beta$的置信域. 由于$\|\beta\|=1$意味着真参数$\beta$是单位超球面上的点, 因此$g(X_{ij}^{\mathrm{T}}\beta)$在参数空间的边界点$\beta$处不存在偏导数. 然而, 在构造经验似然比时需要使用$g(X_{ij}^{\mathrm{T}}\beta)$关于$\beta$每一分量的偏导数. 为了解决这个问题, 我们使用了流行的“去一分量”方法, 其具体思路是: 由于$\|\beta\|=1$, 则$\beta$的某一分量可以用其余$p-1$个分量来表示, 这样可以将参数的维数减少1. 记$\beta=(\beta_1,\cdots,\beta_p)^{\mathrm{T}}$, 且$\beta^{(r)}=(\beta_1,\cdots,\beta_{r-1},\beta_{r+1},\cdots,\beta_p)^{\mathrm{T}}$是去掉$\beta$的第$r$个分量$\beta_r$后的$p-1$维参数向量. 不妨假定$\beta_r$是一个正的分量. 否则考虑$\beta_r=-(1-\|\beta^{(r)}\|^2)^{1/2}$, 则有

$$\beta=\beta(\beta^{(r)})=(\beta_1,\cdots,\beta_{r-1},(1-\|\beta^{(r)}\|^2)^{1/2},\beta_{r+1},\cdots,\beta_p)^{\mathrm{T}},$$

真参数$\beta^{(r)}$满足$\|\beta^{(r)}\|<1$, 因此$\beta$在真参数$\beta^{(r)}$的某个邻域内是有限可微的, 其Jacobian矩阵是

$$J_{\beta^{(r)}}=\frac{\partial\beta}{\partial\beta^{(r)}}=(\gamma_1,\cdots,\gamma_p)^{\mathrm{T}},$$

其中$\gamma_s\,(1\le s\le p, s\ne r)$是$p-1$维单位向量, 当$s<r$时, 其第$s$个元素为1, 当$s>r$时, 其第$s-1$个元素为1, 而$\gamma_r=-(1-\|\beta^{(r)}\|^2)^{-1/2}\beta^{(r)}$.

对于非纵向数据的单指标模型, Xue and Zhu[8]利用经验似然方法研究了未知参数$\beta$的置信域. 但是对于纵向数据单指标模型(1.1), 经验似然方法不能直接用来构造未知参数的

置信域. 因为由Owen[10]提出的经验似然方法是针对独立同分布的数据, 而纵向数据却不是独立同分布的, 其特点是不同个体之间的观测是独立的, 但由于对同一个体进行重复观测, 因此同一个体内的不同观测之间可能是相关的. 本文根据纵向数据不同个体间观测的独立性特点提出三种经验似然方法来构造模型(1.1)中未知参数的置信域. 令

$$X_i=(X_{i1},X_{i2},\cdots,X_{im_i})^{\mathrm{T}},\qquad Y_i=(Y_{i1},Y_{i2},\cdots,Y_{im_i})^{\mathrm{T}},$$
$$G(X_i\beta)=(g(X_{i1}^{\mathrm{T}}\beta),g(X_{i2}^{\mathrm{T}}\beta),\cdots,g(X_{im_i}^{\mathrm{T}}\beta))^{\mathrm{T}},$$
$$\Lambda_i(\beta^{(r)})=(g'(X_{i1}^{\mathrm{T}}\beta)J_{\beta^{(r)}}^{\mathrm{T}}X_{i1},\cdots,g'(X_{im_i}^{\mathrm{T}}\beta)J_{\beta^{(r)}}^{\mathrm{T}}X_{im_i}).$$

根据个体间的独立性特点构造下面的辅助随机向量

$$\eta_i(\beta^{(r)})=\Lambda_i(\beta^{(r)})[Y_i-G(X_i\beta)]=\sum_{j=1}^{m_i}[Y_{ij}-g(X_{ij}^{\mathrm{T}}\beta)]g'(X_{ij}^{\mathrm{T}}\beta)J_{\beta^{(r)}}^{\mathrm{T}}X_{ij},\tag{2.1}$$

其中$g'(\cdot)$是$g(\cdot)$关于$\beta^{(r)}$的导数. 当$\beta^{(r)}$是真实参数时, $\mathsf{E}(\eta_i(\beta^{(r)}))=0$. 对每个个体$i\ (i=1,\cdots,n)$是相互独立的, 因此我们提出经验对数似然比函数

$$l(\beta^{(r)})=-2\max\Big\{\sum_{i=1}^{n}\log(np_i)\Big|p_i\geq 0,\sum_{i=1}^{n}p_i=1,\sum_{i=1}^{n}p_i\eta_i(\beta^{(r)})=0\Big\}.\tag{2.2}$$

因为$l(\beta^{(r)})$中包含未知函数$g(\cdot)$和$g'(\cdot)$, 所以还不能直接用$l(\beta^{(r)})$来构造未知参数$\beta^{(r)}$的置信域. 一个自然的办法就是用两个估计量来替换$g(\cdot)$和$g'(\cdot)$. 下面利用由Fan[17]提出的局部线性拟合的方法得到$g(\cdot)$和$g'(\cdot)$的估计量. 在模型(1.1)中, 假定$\beta$已知, 使得极小化下面的加权最小二乘问题, 求$a$和$b$

$$\sum_{i=1}^{n}\sum_{j=1}^{m_i}(Y_{ij}-a-b(X_{ij}^{\mathrm{T}}\beta-t))^2K_h(X_{ij}^{\mathrm{T}}\beta-t),\tag{2.3}$$

其中$K_h(\cdot)=h^{-1}K(\cdot/h)$, $K(\cdot)$是一个核函数, $h$是收敛于零的正数, 称之为窗宽. 用$\widehat{a}$和$\widehat{b}$分别作为$g(\cdot)$和$g'(\cdot)$的估计. 经过简单计算可得

$$\widehat{a}=\sum_{i=1}^{n}\sum_{j=1}^{m_i}W_{nij}(t;\beta,h)Y_{ij},\tag{2.4}$$
$$\widehat{b}=\sum_{i=1}^{n}\sum_{j=1}^{m_i}\widetilde{W}_{nij}(t;\beta,h)Y_{ij},\tag{2.5}$$

其中

$$W_{nij}(t;\beta,h)=\frac{N^{-1}K_h(X_{ij}^{\mathrm{T}}\beta-t)[S_{n,2}(t;\beta,h)-(X_{ij}^{\mathrm{T}}\beta-t)S_{n,1}(t;\beta,h)]}{S_{n,0}(t;\beta,h)S_{n,2}(t;\beta,h)-S_{n,1}^2(t;\beta,h)},$$
$$\widetilde{W}_{nij}(t;\beta,h)=\frac{N^{-1}K_h(X_{ij}^{\mathrm{T}}\beta-t)[(X_{ij}^{\mathrm{T}}\beta-t)S_{n,0}(t;\beta,h)-S_{n,1}(t;\beta,h)]}{S_{n,0}(t;\beta,h)S_{n,2}(t;\beta,h)-S_{n,1}^2(t;\beta,h)},$$
$$S_{n,l}(t;\beta)=\frac{1}{N}\sum_{i=1}^{n}\sum_{j=1}^{m_i}K_h(X_{ij}^{\mathrm{T}}\beta-t)(X_{ij}^{\mathrm{T}}\beta-t)^l,\qquad l=0,1,2.$$

用$\widehat{g}(t;\beta)$和$\widehat{g}'(t;\beta)$分别记$g(t)$和$g'(t)$的估计量, 其窗宽分别为$h$和$h_1$, 即有

$$\widehat{g}(t;\beta)=\sum_{i=1}^{n}\sum_{j=1}^{m_i}W_{nij}(t;\beta,h)Y_{ij},\qquad \widehat{g}'(t;\beta)=\sum_{i=1}^{n}\sum_{j=1}^{m_i}\widetilde{W}_{nij}(t;\beta,h_1)Y_{ij}. \tag{2.6}$$

在$\eta_i(\beta^{(r)})$中分别用$\widehat{g}(X_{ij}^{\mathrm{T}}\beta;\beta)$和$\widehat{g}'(X_{ij}^{\mathrm{T}}\beta;\beta)$代替$g(X_{ij}^{\mathrm{T}}\beta)$和$g'(X_{ij}^{\mathrm{T}}\beta)$, 得到一个估计的辅助随机变量, 记之为$\widehat{\eta}_i(\beta^{(r)})$, 则可以定义一个估计的经验对数似然比函数

$$\widehat{l}(\beta^{(r)})=-2\max\Big\{\sum_{i=1}^{n}\log(np_i)\Big|p_i\geq 0,\sum_{i=1}^{n}p_i=1,\sum_{i=1}^{n}p_i\widehat{\eta}_i(\beta^{(r)})=0\Big\}, \tag{2.7}$$

由Lagrange乘子法, 有

$$p_i=\frac{1}{n}\frac{1}{1+\lambda^{\mathrm{T}}\widehat{\eta}_i(\beta^{(r)})},\qquad i=1,\cdots,n. \tag{2.8}$$

由(2.8)式, $\widehat{l}(\beta^{(r)})$可以重写为

$$\widehat{l}(\beta^{(r)})=2\sum_{i=1}^{n}\log(1+\lambda^{\mathrm{T}}\widehat{\eta}_i(\beta^{(r)})), \tag{2.9}$$

其中$\lambda=\lambda(\beta^{(r)})$为Lagrange乘子, 满足

$$\frac{1}{n}\sum_{i=1}^{n}\frac{\widehat{\eta}_i(\beta^{(r)})}{1+\lambda^{\mathrm{T}}\widehat{\eta}_i(\beta^{(r)})}=0. \tag{2.10}$$

为了得到本文的结果, 需要作如下的假定:

C1　对任意的$i=1,\cdots,n$, $j=1,\cdots,m_i$, $X_{ij}^{\mathrm{T}}\beta$的密度函数$f_\beta(t)$在$\mathcal{T}$上满足1阶Lipschitz条件, 这里$\mathcal{T}=\{t=X_{ij}^{\mathrm{T}}\beta:X_{ij}\in A,i=1,\cdots,n,j=1,\cdots,m_i\}$, $A$是$X_{ij}$的有界支撑集. $0<\inf\limits_{t\in\mathcal{T}}f_\beta(t)\leq\sup\limits_{t\in\mathcal{T}}f_\beta(t)<\infty$.

C2　$g(t)$在$\mathcal{T}$上具有有界的二阶连续偏导数; $\mu_s(t)$满足一阶局部Lipschitz条件, 此处$\mu_s(t)$表示$\mu(t)=\mathsf{E}(X_{ij}|X_{ij}^{\mathrm{T}}\beta=t)$的第$s$个分量, 其中$1\leq s\leq p$.

C3　核$K(u)$是一个有界且对称的概率密度函数, 并且满足

$$\int_{-\infty}^{\infty}u^2K(u)\mathrm{d}u\neq 0,\qquad \int_{-\infty}^{\infty}|u|^iK(u)\mathrm{d}u<\infty,\qquad i=1,2,\cdots,8.$$

C4　存在正常数$M$, 使得

$$\max_{1\leq i\leq n,1\leq j\leq m_i}\sup_x\mathsf{E}(e_{ij}^4|X_{ij}=x)\leq M<\infty.$$

C5　存在正常数$c_1$和$c_2$, 使得

$$0<c_1\leq\min_{1\leq i\leq n}\lambda_{i1}\leq\max_{1\leq i\leq n}\lambda_{im_i}\leq c_2<\infty,$$

其中$\lambda_{i1}$和$\lambda_{im_i}$分别表示$\Sigma_i$的最小和最大特征根.

C6 $V(\beta^{(r)})$和$V_0(\beta^{(r)})$是两个正定矩阵, 其中

$$V(\beta^{(r)})=\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\mathsf{E}\{[\Lambda_i(\beta^{(r)})-\mathsf{E}(\Lambda_i(\beta^{(r)})|X_i\beta)](e_ie_i^{\mathrm{T}})[\Lambda_i(\beta^{(r)})-\mathsf{E}(\Lambda_i(\beta^{(r)})|X_i\beta)]^{\mathrm{T}}\},$$
$$V_0(\beta^{(r)})=\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\mathsf{E}\{\Lambda_i(\beta^{(r)})(e_ie_i^{\mathrm{T}})\Lambda_i^{\mathrm{T}}(\beta^{(r)})\}.$$

C7 当$n\to\infty$时, 窗宽$h$和$h_1$满足$nh^2\to\infty$, $nh^4\to 0$; $nhh_1^3\to\infty$, $\limsup\limits_{n\to\infty} nh_1^5<\infty$.

**定理 2.1** 假设条件C1–C7成立. 如果$\beta$是参数真值, 并且$\beta$的第$r$个分量是正数, 则

$$\widehat{l}(\beta^{(r)})\xrightarrow{L}\omega_1\chi_{1,1}^2+\cdots+\omega_{p-1}\chi_{1,p-1}^2,$$

其中"$\xrightarrow{L}$"表示依分布收敛, $\chi_{1,i}^2\,(1\le i\le p-1)$是相互独立的$\chi_1^2$变量, $w_i\,(1\le i\le p-1)$是$D(\beta^{(r)})=V_0^{-1}(\beta^{(r)})V(\beta^{(r)})$的特征根.

为了应用定理2.1构造$\beta^{(r)}$的置信域或置信区间, 必须给出未知权$w_i$的相合估计. 利用Plug-in方法, $V_0(\beta^{(r)})$和$V(\beta^{(r)})$的相合估计可分别定义为

$$\widehat{V}_0(\widehat{\beta}^{(r)})=\frac{1}{n}\sum_{i=1}^{n}\widehat{\eta}_i(\widehat{\beta}^{(r)})\widehat{\eta}_i^{\mathrm{T}}(\widehat{\beta}^{(r)}),$$
$$\widehat{V}(\widehat{\beta}^{(r)})=\frac{1}{n}\sum_{i=1}^{n}[\widehat{\Lambda}_i(\widehat{\beta}^{(r)})-\widehat{\Phi}_i(\widehat{\beta}^{(r)})](\widehat{e}_i\widehat{e}_i^{\mathrm{T}})[\widehat{\Lambda}_i(\widehat{\beta}^{(r)})-\widehat{\Phi}_i(\widehat{\beta}^{(r)})]^{\mathrm{T}},$$

这里

$$\widehat{\beta}=(\widehat{\beta}_1,\cdots,\widehat{\beta}_p)^{\mathrm{T}},\qquad \widehat{\beta}^{(r)}=(\widehat{\beta}_1,\cdots,\widehat{\beta}_{r-1},\widehat{\beta}_{r+1},\cdots,\widehat{\beta}_p)^{\mathrm{T}},$$
$$\widehat{e}_i=Y_i-\widehat{G}(X_i\widehat{\beta}),$$
$$\widehat{\Lambda}_i(\widehat{\beta}^{(r)})=\big(\widehat{g}'(X_{i1}^{\mathrm{T}}\widehat{\beta};\widehat{\beta})J_{\widehat{\beta}^{(r)}}^{\mathrm{T}}X_{i1},\cdots,\widehat{g}'(X_{im_i}^{\mathrm{T}}\widehat{\beta};\widehat{\beta})J_{\widehat{\beta}^{(r)}}^{\mathrm{T}}X_{im_i}\big),$$
$$\widehat{\Phi}_i(\widehat{\beta}^{(r)})=\big(\widehat{g}'(X_{i1}^{\mathrm{T}}\widehat{\beta};\widehat{\beta})J_{\widehat{\beta}^{(r)}}^{\mathrm{T}}\widehat{\mu}(X_{i1}^{\mathrm{T}}\widehat{\beta};\widehat{\beta}),\cdots,\widehat{g}'(X_{im_i}^{\mathrm{T}}\widehat{\beta};\widehat{\beta})J_{\widehat{\beta}^{(r)}}^{\mathrm{T}}\widehat{\mu}(X_{im_i}^{\mathrm{T}}\widehat{\beta};\widehat{\beta})\big),$$

此处$\widehat{\beta}$是基于模型(1.1)的$\beta$的最小二乘估计, 即

$$\widehat{\beta}=\mathrm{Arg}\min_{\beta}\Big\{\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{m_i}[Y_{ij}-\widehat{g}(X_{ij}^{\mathrm{T}}\beta;\beta)]^2\Big\},\tag{2.11}$$

$\widehat{\mu}(t;\beta)$是$\mu(t)=\mathsf{E}(X_{ij}|X_{ij}^{\mathrm{T}}\beta=t)$的估计量, 定义为

$$\widehat{\mu}(t;\beta)=\sum_{i=1}^{n}\sum_{j=1}^{m_i}W_{nij}(t;\beta,h)X_{ij}.\tag{2.12}$$

由此可以推出$\widehat{D}(\widehat{\beta}^{(r)})=\widehat{V}_0^{-1}(\widehat{\beta}^{(r)})\widehat{V}(\widehat{\beta}^{(r)})$的特征根$\widehat{w}_i$是$w_i\,(i=1,\cdots,p-1)$的相合估计. 记$\widehat{S}=\widehat{w}_1\chi_{1,1}^2+\cdots+\widehat{w}_{p-1}\chi_{1,p-1}^2$. 用$H(\cdot)$表示在给定$\{(X_{ij},Y_{ij}),1\le i\le n,1\le j\le m_i\}$下$\widehat{S}$的条件分布. 设$\widehat{c}_\alpha$是$H(\cdot)$的$1-\alpha$分位点, 则$\beta^{(r)}$的置信度为$1-\alpha$的置信域为

$$\widehat{I}_\alpha(\widetilde{\beta}^{(r)})=\{\widetilde{\beta}^{(r)}:\widehat{l}(\widetilde{\beta}^{(r)})\le\widehat{c}_\alpha,\|\widetilde{\beta}^{(r)}\|<1\}.$$

实际上, 要得到分布$H(\cdot)$是很方便的. 它可以通过重复产生$\chi_1^2$样本$\chi_{1,1}^2,\cdots,\chi_{1,p-1}^2$, 然后使用Monte Carlo模拟而得到.

## 2.2 调整的经验对数似然

由2.1节可知, 在应用定理2.1构造参数$\beta^{(r)}$的置信域时, 需要估计权$w_i$, 还需要用到Monte Carlo算法, 这不仅会降低置信域的精度, 而且又将增大计算量. 本节我们对估计的经验对数似然比进行调整, 提出一种调整的经验对数似然比, 该方法避免了上面的两个问题.

按照Rao and Scott[18]的结果, $\rho(\beta^{(r)})\sum\limits_{i=1}^{p-1}w_i\chi_{1,i}^2$的分布可用自由度为$p-1$的$\chi^2$分布逼近, 这里$\rho(\beta^{(r)})=(p-1)/\mathrm{tr}\{D(\beta^{(r)})\}$. Wang and Rao[19]首次在研究线性EV模型时使用了这一方法, 提出调整后的经验对数似然比能够渐近到中心卡方分布. 由定理2.1以及$\widehat{V}_0(\widehat{\beta}^{(r)})$和$\widehat{V}(\widehat{\beta}^{(r)})$的相合性, 可以推出$\widehat{\rho}(\widehat{\beta}^{(r)})\widehat{l}(\beta^{(r)})$的渐近分布为$\chi_{p-1}^2$分布, 此处$\widehat{\rho}(\widehat{\beta}^{(r)})=(p-1)/\mathrm{tr}\{\widehat{D}(\widehat{\beta}^{(r)})\}$. 为了提高逼近精度, 在$\widehat{\rho}(\widehat{\beta}^{(r)})$中用$\beta^{(r)}$代替$\widehat{\beta}^{(r)}$. 因此, 改良Rao & Scott调整的经验对数似然比可定义为$\widetilde{l}(\beta^{(r)})=\widehat{\rho}(\beta^{(r)})\widehat{l}(\beta^{(r)})$, 然而, 这种逼近的精度依赖于$w_i$的值. 下面给出一个调整的经验对数似然比. 注意到

$$\widehat{\rho}(\beta^{(r)})=\frac{\mathrm{tr}\{\widehat{V}^{-1}(\beta^{(r)})\widehat{V}(\beta^{(r)})\}}{\mathrm{tr}\{\widehat{V}_0^{-1}(\beta^{(r)})\widehat{V}(\beta^{(r)})\}}.$$

检查$\widehat{l}(\beta^{(r)})$渐近表达式, 在$\widehat{\rho}(\beta^{(r)})$中用$\widehat{B}(\beta^{(r)})=\Big\{\sum\limits_{i=1}^{n}\widehat{\eta}_i(\beta^{(r)})\Big\}\Big\{\sum\limits_{i=1}^{n}\widehat{\eta}_i(\beta^{(r)})\Big\}^{\mathrm{T}}$代替$\widehat{V}(\beta^{(r)})$, 就得到一个不同的调整因子$\widehat{A}(\beta^{(r)})=\mathrm{tr}\{\widehat{V}^{-1}(\beta^{(r)})\widehat{B}(\beta^{(r)})\}/\mathrm{tr}\{\widehat{V}_0^{-1}(\beta^{(r)})\widehat{B}(\beta^{(r)})\}$. 因此定义调整的经验对数似然比为

$$\widehat{l}_{\mathrm{ad}}(\beta^{(r)})=\widehat{A}(\beta^{(r)})\widehat{l}(\beta^{(r)}).$$

**定理 2.2** 假设条件C1–C7成立. 如果$\beta$是参数真值, 并且$\beta$的第$r$个分量是正数, 则$\widehat{l}_{\mathrm{ad}}(\beta^{(r)})\xrightarrow{L}\chi_{p-1}^2$.

基于定理2.2, 可以使用$\widehat{l}_{\mathrm{ad}}(\beta^{(r)})$构造$\beta^{(r)}$的置信域

$$\widehat{I}_{\mathrm{ad},\alpha}(\widetilde{\beta}^{(r)})=\{\widetilde{\beta}^{(r)}:\widehat{l}_{\mathrm{ad}}(\widetilde{\beta}^{(r)})\le c_\alpha,\|\widetilde{\beta}^{(r)}\|<1\},$$

这里$\mathsf{P}(\chi_{p-1}^2\le c_\alpha)=1-\alpha$. 并且有$\mathsf{P}(\beta^{(r)}\in\widehat{I}_{\mathrm{ad},\alpha}(\widetilde{\beta}^{(r)}))=1-\alpha+o(1)$.

## 2.3 纠偏的经验对数似然

虽然调整的经验对数似然较估计的经验对数似然有良好的性质, 但还是需要估计调整因子, 而调整因子估计的好坏将直接影响参数置信域的精度. 本节我们对辅助随机向量进

行纠偏, 提出一种纠偏的经验对数似然, 该方法具有良好的性质: (1) 无需对参数进行估计; (2) 使所构造的纠偏的经验对数似然比函数能够渐近收敛到一个自由度为$p-1$的标准卡方分布, 避免了估计未知权$\omega_i$和调整因子; (3) 放宽了窗宽的选取范围, 在估计$g(\cdot)$和$g'(\cdot)$时只需使用同一个窗宽; (4) 该方法计算简单, 而且收敛速度快.

我们把Zhu and Xue[16]的纠偏思想推广到纵向数据单指标模型(1.1), 构造下面的纠偏的辅助随机向量

$$\begin{aligned}\widehat{\eta}_{\mathrm{bc},i}(\beta^{(r)}) &= \widehat{\Lambda}_{\mathrm{bc},i}(\beta^{(r)})[Y_i-\widehat{G}(X_i\beta)]\\ &= \sum_{j=1}^{m_i}[Y_{ij}-\widehat{g}(X_{ij}^{\mathrm{T}}\beta;\beta,h)]\widehat{g}'(X_{ij}^{\mathrm{T}}\beta;\beta,h)J_{\beta^{(r)}}^{\mathrm{T}}(X_{ij}-\widehat{\mu}(X_{ij}^{\mathrm{T}}\beta;\beta,h)),\end{aligned}\tag{2.13}$$

其中$\widehat{\Lambda}_{\mathrm{bc},i}(\beta^{(r)})=(\widehat{g}'(X_{i1}^{\mathrm{T}}\beta;\beta,h)J_{\beta^{(r)}}^{\mathrm{T}}[X_{i1}-\widehat{\mu}(X_{i1}^{\mathrm{T}}\beta;\beta,h)],\cdots,\widehat{g}'(X_{im_i}^{\mathrm{T}}\beta;\beta,h)J_{\beta^{(r)}}^{\mathrm{T}}[X_{im_i}-\widehat{\mu}(X_{im_i}^{\mathrm{T}}\beta;\beta,h)])$, $\widehat{\mu}(t;\beta)$在(2.12)中定义. 由于使用了纠偏的方法, 所以在(2.13)式中$g(\cdot)$和$g'(\cdot)$的估计可以使用同一个窗宽$h$. 更重要的是使得$\mathsf{E}(\widehat{\eta}_{\mathrm{bc},i}(\beta^{(r)}))=o(1)$, 因此我们可以定义一个纠偏的经验对数似然比函数

$$\widehat{l}_{\mathrm{bc}}(\beta^{(r)})=-2\max\Big\{\sum_{i=1}^n\log(np_i)\Big|p_i\geq 0,\sum_{i=1}^n p_i=1,\sum_{i=1}^n p_i\widehat{\eta}_{\mathrm{bc},i}(\beta^{(r)})=0\Big\}.\tag{2.14}$$

由Lagrange乘子法, $\widehat{l}_{\mathrm{bc}}(\beta^{(r)})$可以重写为

$$\widehat{l}_{\mathrm{bc}}(\beta^{(r)})=2\sum_{i=1}^n\log(1+\lambda^{\mathrm{T}}\widehat{\eta}_{\mathrm{bc},i}(\beta^{(r)})),\tag{2.15}$$

其中$\lambda=\lambda(\beta^{(r)})$为Lagrange乘子, 满足

$$\frac{1}{n}\sum_{i=1}^n\frac{\widehat{\eta}_{\mathrm{bc},i}(\beta^{(r)})}{1+\lambda^{\mathrm{T}}\widehat{\eta}_{\mathrm{bc},i}(\beta^{(r)})}=0.\tag{2.16}$$

为了得到$\widehat{l}_{\mathrm{bc}}(\beta^{(r)})$的渐近结果, 对窗宽$h$需要作如下的假定:

C8 当$n\to\infty$时, 窗宽$h$满足$h\to 0$, $nh^3\to\infty$, $nh^8\to 0$.

**注记 1** 条件C7中要求窗宽为$O(n^{-1/3})$, 在估计$g(\cdot)$时需要undersmothing, 而估计$g'(\cdot)$则要选取不同的窗宽; 与条件C7相比, 条件C8放宽了窗宽的选取范围, 使得窗宽可以从$O(n^{-1/4})$到$O(n^{-1/7})$中选取, 其中包含了最优窗宽. 同时, 条件C8意味着在估计$g(\cdot)$和$g'(\cdot)$时可以使用同一个窗宽, 从而大大降低计算过程中窗宽选取的工作量, 并可以简化定理的证明.

**定理 2.3** 假设条件C1–C6和条件C8成立. 如果$\beta$是参数真值, 并且$\beta$的第$r$个分量是正数, 则

$$\widehat{l}_{\mathrm{bc}}(\beta^{(r)})\xrightarrow{L}\chi^2_{p-1}.$$

由定理2.3, 对任给的$0<\alpha<1$, 存在$c_\alpha$使得$\mathrm{P}(\chi^2_{p-1}>c_\alpha)=\alpha$, 则

$$\widehat{I}_{\mathrm{bc},\alpha}(\widetilde{\beta}^{(r)})=\{\widetilde{\beta}^{(r)}:\widehat{l}_{\mathrm{bc}}(\widetilde{\beta}^{(r)})\le c_\alpha,\|\widetilde{\beta}^{(r)}\|<1\}$$

为参数向量$\beta^{(r)}$的具有渐近置信水平$1-\alpha$的置信域.

**注记 2** 　由定理2.1–2.3, 一旦我们得到了$\beta^{(r)}$的置信域, 根据$\beta_r=(1-\|\beta^{(r)}\|^2)^{1/2}$就能获得$\beta$的置信域.

## §3. 模拟研究

我们考虑$p=3$的情况. 通过模拟研究比较最小二乘(LS), 估计的经验似然(EEL), 调整的经验似然(AEL)和纠偏的经验似然(BCEL)这四种方法得到的参数置信域的精确度. 模拟中通过计算置信域及其覆盖概率的大小比较上面四种方法的优劣. 其中由文献[9]可知, 类似可以得到由(2.11)所定义的最小二乘估计有渐近正态性, 均值为0, 渐近方差矩阵的估计使用plug-in方法获得, 为

$$\widehat{V}_{\mathrm{LS}}=n^{-1}\widehat{\Sigma}(\widehat{\beta}^{(r)})^{-1}\widehat{V}(\widehat{\beta}^{(r)})\widehat{\Sigma}(\widehat{\beta}^{(r)})^{-1},$$

其中$\widehat{V}(\widehat{\beta}^{(r)})$在定理2.1中定义, 并且

$$\widehat{\Sigma}(\widehat{\beta}^{(r)})=\frac{1}{N}\sum_{i=1}^{n}\sum_{j=1}^{m_i}\widehat{g}'(X_{ij}^{\mathrm{T}}\widehat{\beta};\widehat{\beta})^2\widehat{J}_{\widehat{\beta}^{(r)}}^{\mathrm{T}}X_{ij}X_{ij}^{\mathrm{T}}\widehat{J}_{\widehat{\beta}^{(r)}}.$$

考虑纵向数据下单指标模型

$$Y_{ij}=\exp(X_{ij}^{\mathrm{T}}\beta)+e_{ij},\qquad i=1,\cdots,n,\quad j=1,\cdots,4,\tag{3.1}$$

其中$\beta=(\beta_1,\beta_2,\beta_3)^{\mathrm{T}}=(1/3,2/3,2/3)^{\mathrm{T}}$, 协变量$X_{ij}$是3维的随机变量, 设置为$X_{ij}=(X_{1ij},X_{2ij},X_{3ij})^{\mathrm{T}}$, 其中$X_{1ij}=U_i+\varepsilon_{ij}$产生$X_{1ij}$, 这里$U_i$服从标准正态分布, $\varepsilon_{ij}$服从具有方差为1、内相关系数为0.4的正态分布. $X_{2ij}$服从$(-1,1)$上的均匀分布, $X_{3ij}$是个体水平协变量, 其个体以概率1/2取值1或0. $Y_{ij}$由模型(3.1)产生, 模型误差$e_{ij}$服从具有方差1、内相关系数为0.6的正态分布.

模拟中取个体大小$n=50,100,150,200$, 为模拟方便个体的重复观测取$m_i\equiv m=3$. 非参数权函数取Epanechnikov核$K(u)=0.75(1-u^2)_+$. 对最小二乘, 估计的经验似然和调整的经验似然方法, 窗宽的选取首先采用交错证实法得到最优窗宽$h_{\mathrm{opt}}$, 然后取$h=h_{\mathrm{opt}}\times n^{-2/15}$和$h_1=h_{\mathrm{opt}}$满足条件C7. 对纠偏的经验似然方法, 由条件C8可知, 窗宽从$O(n^{-1/4})$到$O(n^{-1/7})$范围变化, 故我们取最优窗宽$h_{\mathrm{opt}}$进行模拟计算. 我们做$M=1000$次重复运算来计算$(\beta_1,\beta_3)$和$(\beta_2,\beta_3)$的置信水平为0.95的置信域及其覆盖概率. 模拟结果报告在表1, 图1和图2中.

表1 $(\beta_1,\beta_3)$和$(\beta_2,\beta_3)$的置信水平为95%的置信域的覆盖概率

| $n$ | $(\beta_1,\beta_3)$ LS | EEL | AEL | BCEL | $(\beta_2,\beta_3)$ LS | EEL | AEL | BCEL |
|---|---|---|---|---|---|---|---|---|
| 50 | 0.8980 | 0.8910 | 0.9070 | 0.9110 | 0.8970 | 0.9010 | 0.9030 | 0.9120 |
| 100 | 0.9180 | 0.9160 | 0.9210 | 0.9270 | 0.9180 | 0.9150 | 0.9250 | 0.9290 |
| 150 | 0.9310 | 0.9320 | 0.9340 | 0.9400 | 0.9320 | 0.9340 | 0.9360 | 0.9380 |
| 200 | 0.9430 | 0.9420 | 0.9450 | 0.9480 | 0.9420 | 0.9410 | 0.9470 | 0.9490 |

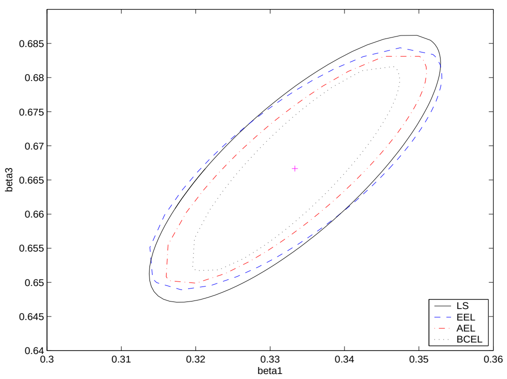


图1 $(\beta_1,\beta_3)$的置信水平为0.95的置信域

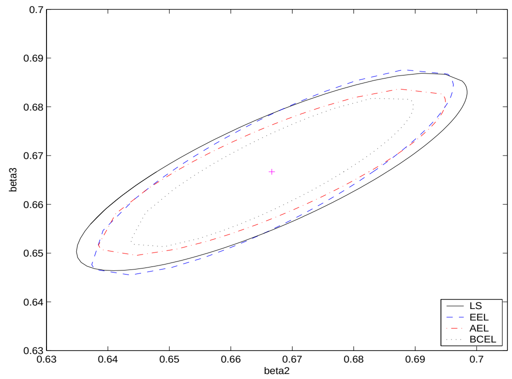


图2 $(\beta_2,\beta_3)$的置信水平为0.95的置信域

图1和图2给出的$(\beta_1,\beta_3)$和$(\beta_2,\beta_3)$的置信域可以看出, BCEL明显比其他三种方法好, 它得到的置信域比其他三种情况都小. 其次, LS和EEL的置信域差不多, 而且这两种方法表现最差, 因为最小二乘方法需要估计渐近方差矩阵, 估计的经验似然方法需要估计未知权重$\omega_i$, 此外这两种方法的计算量较大. AEL虽然优于EEL和LS, 但是明显比BCEL的置信域要大点. 从表1可以看出, 四种方法所得的覆盖概率随着个体大小$n$的增大而增大, 并且都趋近于置信水平0.95, BCEL方法得到的覆盖概率比其它三种都大, 而LS和EEL得到的覆盖概率基本接近.

从模拟结果可知, 纠偏的经验似然(BCEL)是一个优良的方法, 它得到$\beta$的置信域具有较高的精确度, 而且在模拟计算时计算量小.

## §4. 定理的证明

为叙述方便, 以下始终假设$c$表示任一不依赖于$n$和$N$的正常数, $c$每次出现可以取不同的值.

**引理 4.1** 若条件C1–C3成立. 如果$h=cn^{-\alpha}$, $0<\alpha<1/2$, 且$n=O(N)$, 则对

$i_1=1,\cdots,n$, $j_1=1,\cdots,m_{i_1}$, 有

$$\mathsf{E}\Big[\sum_{i=1}^{n}\sum_{j=1}^{m_i}W_{nij}(X_{i_1j_1}^{\mathrm T}\beta;\beta,h)g(X_{ij}^{\mathrm T}\beta)-g(X_{i_1j_1}^{\mathrm T}\beta)\Big]^2=O(h^4), \tag{4.1}$$

$$\mathsf{E}\Big[\sum_{i=1}^{n}\sum_{j=1}^{m_i}\widetilde{W}_{nij}(X_{i_1j_1}^{\mathrm T}\beta;\beta,h_1)g(X_{ij}^{\mathrm T}\beta)-g'(X_{i_1j_1}^{\mathrm T}\beta)\Big]^2=O(h_1^2). \tag{4.2}$$

**证明:** 此处仅证明(4.2)式, (4.1)式的证明类似于(4.2)式的证明. 记$\widetilde{W}_{nij}(t;\beta,h_1)=\widetilde{U}_{nij}(t;\beta,h_1)/U_n(t;\beta,h_1)$和$t=X_{i_1j_1}^{\mathrm T}\beta$, $i_1=1,\cdots,n$, $j_1=1,\cdots,m_{i_1}$, 其中

$$\widetilde{U}_{nij}(t;\beta,h_1)=\frac{1}{N}K_{h_1}(X_{ij}^{\mathrm T}\beta-t)[(X_{ij}^{\mathrm T}\beta-t)S_{n,0}(t;\beta,h_1)-S_{n,1}(t;\beta,h_1)],$$
$$U_n(t;\beta,h_1)=S_{n,0}(t;\beta,h_1)S_{n,2}(t;\beta,h_1)-S_{n,1}^2(t;\beta,h_1).$$

可知

$$\sum_{i=1}^{n}\sum_{j=1}^{m_i}\widetilde{W}_{nij}(t;\beta,h_1)(X_{ij}^{\mathrm T}\beta-t)=1,\qquad \sum_{i=1}^{n}\sum_{j=1}^{m_i}\widetilde{U}_{nij}(t;\beta,h_1)=0.$$

记$H(X_{ij}^{\mathrm T}\beta,t)=g(X_{ij}^{\mathrm T}\beta)-g(t)-g'(t)(X_{ij}^{\mathrm T}\beta-t)$. 简单计算得

$$\sum_{i=1}^{n}\sum_{j=1}^{m_i}\widetilde{W}_{nij}(t;\beta,h_1)g(X_{ij}^{\mathrm T}\beta)-g'(t)=\frac{\sum\limits_{i=1}^{n}\sum\limits_{j=1}^{m_i}\widetilde{U}_{nij}(t;\beta,h_1)H(X_{ij}^{\mathrm T}\beta,t)}{U_n(t;\beta,h_1)}. \tag{4.3}$$

用$\mathsf{E}_{t\in\mathcal{T}}[\xi_n]$表示给定$t\in\mathcal{T}$下$\xi_n$的条件期望. 如果$\mathsf{E}(|\xi_n|^r)=O(a_n^r)$, 则记$\xi_n=O_r(a_n)$. 利用Cauchy-Schwarz不等式, 容易得到

$$O_r(a_n)O_r(b_n)=O_{r/2}(a_nb_n). \tag{4.4}$$

对$S_{n,l}(t;\beta,h_1)$ $(l=0,1,2)$, 由$n=O(N)$, 并利用条件C1和对核函数的假设条件, 可证得

$$\mathsf{E}_{t\in\mathcal{T}}[S_{n,l}(t;\beta,h_1)]=a_lf(t)h_1^l(1+O(h_1)),$$
$$\begin{aligned}S_{n,l}(t;\beta,h_1)&=\mathsf{E}_{t\in\mathcal{T}}[S_{n,l}(t;\beta,h_1)]+O_4((\mathsf{E}|S_{n,l}(t;\beta,h_1)-\mathsf{E}_{t\in\mathcal{T}}[S_{n,l}(t;\beta,h_1)]|^4)^{1/4})\\&=a_lf(t)h_1^l(1+O_4(h_1+(nh_1)^{-1/2})),\qquad l=0,1,2,\end{aligned} \tag{4.5}$$

其中$a_l=\int_{-\infty}^{\infty}u^lK(u)\mathrm{d}u$, $l=0,1,2$, $a_0=1$, $a_1=0$. 由(4.4)和(4.5)式可得

$$\begin{aligned}U_n(t;\beta,h_1)&=S_{n,0}(t;\beta,h_1)S_{n,2}(t;\beta,h_1)-S_{n,1}^2(t;\beta,h_1)\\&=a_2f^2(t)h_1^2(1+O_2(h_1+(nh_1)^{-1/2})).\end{aligned} \tag{4.6}$$

记$U(t)=a_2f^2(t)$, 利用(4.6)式和Borel-Cantelli引理, 使用密度核估计的论证方法可以证得

$$\sup_{t\in\mathcal{T}}\Big|\frac{U_n(t;\beta,h_1)}{h_1^2}-U(t)\Big|\longrightarrow 0,\qquad \text{a.s..}$$

由条件C1知$\inf\limits_{t\in\mathcal{T}} f(t)\geq c>0$, 因此当$n$充分大时, 有

$$\begin{aligned}\inf_{t\in\mathcal{T}}|U_n(t;\beta,h_1)/h_1^2| &\geq \inf_{t\in\mathcal{T}}|U(t)|-\sup_{t\in\mathcal{T}}|U_n(t;\beta,h_1)/h_1^2-U(t)|\\ &> a_2c^2/2>0, \qquad \text{a.s..}\end{aligned}\tag{4.7}$$

经过计算有

$$\begin{aligned}&\sum_{i=1}^n\sum_{j=1}^{m_i}\widetilde{U}_{nij}(t;\beta,h_1)H(X_{ij}^{\mathrm{T}}\beta,t)\\ =\;& S_{n,0}(t;\beta,h_1)\frac{1}{N}\sum_{i=1}^n\sum_{j=1}^{m_i}K_{h_1}(X_{ij}^{\mathrm{T}}\beta-t)(X_{ij}^{\mathrm{T}}\beta-t)H(X_{ij}^{\mathrm{T}}\beta,t)\\ &-S_{n,1}(t;\beta,h_1)\frac{1}{N}\sum_{i=1}^n\sum_{j=1}^{m_i}K_{h_1}(X_{ij}^{\mathrm{T}}\beta-t)H(X_{ij}^{\mathrm{T}}\beta,t),\end{aligned}\tag{4.8}$$

类似于(4.5)式的证明, 对$l=0,1$, 有

$$\begin{aligned}&\frac{1}{Nh_1^{2+l}}\sum_{i=1}^n\sum_{j=1}^{m_i}H(X_{ij}^{\mathrm{T}}\beta,t)(X_{ij}^{\mathrm{T}}\beta-t)^lK_{h_1}(X_{ij}^{\mathrm{T}}\beta-t)\\ =\;& h_1^{-2-l}\mathsf{E}_{t\in\mathcal{T}}[H(X_{11}^{\mathrm{T}}\beta,t)(X_{11}^{\mathrm{T}}\beta-t)^lK_{h_1}(X_{11}^{\mathrm{T}}\beta-t)]+O_4((nh_1)^{-1/2})\\ \widehat{=}\;& d_{nl}+O_4((nh_1)^{-1/2}),\end{aligned}\tag{4.9}$$

$$|d_{nl}|\leq c\int_{-\infty}^{\infty}|u|^{l+2}K(u)\mathrm{d}uf(t)+O(h_1)=O(1).$$

由(4.5), (4.8)和(4.9)式, 得

$$\sum_{i=1}^n\sum_{j=1}^{m_i}\widetilde{U}_{nij}(t;\beta,h_1)H(X_{ij}^{\mathrm{T}}\beta,t)=h_1^3[f(t)d_{n1}+O_2(h_1+(nh_1)^{-1/2})].\tag{4.10}$$

故由(4.7)和(4.10)式得

$$\mathsf{E}\Bigg[\frac{\sum\limits_{i=1}^n\sum\limits_{j=1}^{m_i}\widetilde{U}_{nij}(t;\beta,h_1)H(X_{ij}^{\mathrm{T}}\beta,t)}{U_n(t;\beta,h_1)}\Bigg]^2\leq ch_1^{-4}\mathsf{E}\Big[\sum_{i=1}^n\sum_{j=1}^{m_i}\widetilde{U}_{nij}(t;\beta,h_1)H(X_{ij}^{\mathrm{T}}\beta,t)\Big]^2=O(h_1^2).$$

□

**引理 4.2** 在引理4.1的条件下, 有

$$\begin{cases}\mathsf{E}[W_{nij}^2(X_{i_1j_1}^{\mathrm{T}}\beta;\beta,h)]=O((nh)^{-2}),\\ \mathsf{E}\Big[\sum\limits_{i=1}^n\sum\limits_{j=1}^{m_i}W_{nij}^2(X_{i_1j_1}^{\mathrm{T}}\beta;\beta,h)\Big]=O((nh)^{-1}),\end{cases}$$

$$\begin{cases}\mathsf{E}[\widetilde{W}_{nij}^2(X_{i_1j_1}^{\mathrm{T}}\beta;\beta,h_1)]=O((nh_1)^{-2})+O((n^3h_1^5)^{-1}),\\ \mathsf{E}\Big[\sum\limits_{i=1}^n\sum\limits_{j=1}^{m_i}\widetilde{W}_{nij}^2(X_{i_1j_1}^{\mathrm{T}}\beta;\beta,h_1)\Big]=O((nh_1^3)^{-1}).\end{cases}$$

**证明:** 类似于(4.2)式的证明方法可以证得引理4.2的结果, 此处省略. □

**引理 4.3** 在引理4.1的条件下, 对任意$i=1,\cdots,n$, $j=1,\cdots,m_i$, 有

$$\mathsf{E}[\widehat{g}(X_{ij}^{\mathrm T}\beta;\beta)-g(X_{ij}^{\mathrm T}\beta)]^2=O(h^4)+O((nh)^{-1}), \tag{4.11}$$

$$\mathsf{E}[\widehat{g}'(X_{ij}^{\mathrm T}\beta;\beta)-g'(X_{ij}^{\mathrm T}\beta)]^2=O(h_1^2)+O((nh_1^3)^{-1}). \tag{4.12}$$

**证明:** 这里仅证明(4.11)式, (4.12)式可类似证明. 利用引理4.1, 引理4.2和条件C4, 有

$$\begin{aligned}\mathsf{E}[\widehat{g}(X_{ij}^{\mathrm T}\beta;\beta)-g(X_{ij}^{\mathrm T}\beta)]^2 &= \mathsf{E}\Big[\sum_{i_1=1}^{n}\sum_{j_1=1}^{m_{i_1}}W_{ni_1j_1}(X_{ij}^{\mathrm T}\beta;\beta)g(X_{i_1j_1}^{\mathrm T}\beta)-g(X_{ij}^{\mathrm T}\beta)\Big]^2\\ &\quad+\mathsf{E}\Big[\sum_{i_1=1}^{n}\sum_{j_1=1}^{m_{i_1}}W_{ni_1j_1}(X_{ij}^{\mathrm T}\beta;\beta)e_{i_1j_1}\Big]^2\\ &\le ch^4+c(nh)^{-1}.\end{aligned}$$

□

**引理 4.4** 假设条件C1–C7成立. 如果$\beta$是参数真值, 并且第$r$个分量$\beta_r>0$, 则

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\widehat{\eta}_i(\beta^{(r)})\xrightarrow{L}N(0,V(\beta^{(r)})),$$

其中$V(\beta^{(r)})$在条件C6中定义.

**证明:** 经过简单计算得

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\widehat{\eta}_i(\beta^{(r)})=\frac{1}{\sqrt{n}}\sum_{i=1}^{n}[\Lambda_i(\beta^{(r)})-\mathsf{E}(\Lambda_i(\beta^{(r)})|X_i\beta)]e_i+J_{\beta^{(r)}}^{\mathrm T}\sum_{\nu=1}^{4}M_\nu, \tag{4.13}$$

其中

$$M_1=\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\sum_{j=1}^{m_i}[g(X_{ij}^{\mathrm T}\beta)-\widehat{g}(X_{ij}^{\mathrm T}\beta;\beta)][\widehat{g}'(X_{ij}^{\mathrm T}\beta;\beta)-g'(X_{ij}^{\mathrm T}\beta)]X_{ij},$$

$$M_2=\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\sum_{j=1}^{m_i}e_{ij}[\widehat{g}'(X_{ij}^{\mathrm T}\beta;\beta)-g'(X_{ij}^{\mathrm T}\beta)]X_{ij},$$

$$M_3=\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\sum_{j=1}^{m_i}[g(X_{ij}^{\mathrm T}\beta)-\widehat{g}(X_{ij}^{\mathrm T}\beta;\beta)]g'(X_{ij}^{\mathrm T}\beta)X_{ij},$$

$$M_4=\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\sum_{j=1}^{m_i}e_{ij}g'(X_{ij}^{\mathrm T}\beta)\mathsf{E}(X_{ij}|X_{ij}^{\mathrm T}\beta).$$

首先需要证明:

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}[\Lambda_i(\beta^{(r)})-\mathsf{E}(\Lambda_i(\beta^{(r)})|X_i\beta)]e_i\xrightarrow{L}N(0,V(\beta^{(r)})). \tag{4.14}$$

为证明方便, 令

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}[\Lambda_i(\beta^{(r)})-\mathsf{E}(\Lambda_i(\beta^{(r)})|X_i\beta)]e_i\,\widehat{=}\,\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\xi_{in},$$

由$\mathsf{E}(e_i|X_i)=0$和条件C4, 以及$\mathsf{Cov}\,(\xi_{in})=\mathsf{E}\{[\Lambda_i(\beta^{(r)})-\mathsf{E}(\Lambda_i(\beta^{(r)})|X_i\beta)](e_ie_i^{\mathrm{T}})[\Lambda_i(\beta^{(r)})-\mathsf{E}(\Lambda_i(\beta^{(r)})|X_i\beta)]^{\mathrm{T}}\}$. 令

$$V(\beta^{(r)})=\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\mathsf{Cov}\,(\xi_{in}),$$

和$\{\xi_{in}\}$的第$k$个元素列为$\{\xi_{in}^{(k)}\}$ $(k=1,\cdots,p-1)$, 对任给的$\varepsilon>0$, 容易证明

$$\frac{1}{n}\sum_{i=1}^{n}\mathsf{E}\{\xi_{in}^{(k)^2}I(|\xi_{in}^{(k)}|>\varepsilon n^{1/2})\}\longrightarrow 0\qquad(n\to\infty),$$

则由Lindeberg中心极限定理可得(4.14)式成立.

为了证明引理4.4, 只需证明: $M_l\xrightarrow{P}0$ $(l=1,2,3,4)$. 用$M_{1,s}$表示$M_1$的第$s$个分量, 由引理4.3和Cauchy-Schwarz不等式, 有

$$\begin{aligned}\mathsf{E}(|M_{1,s}|)&\leq c\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\sum_{j=1}^{m_i}\mathsf{E}^{1/2}[g(X_{ij}^{\mathrm{T}}\beta)-\widehat{g}(X_{ij}^{\mathrm{T}}\beta;\beta)]^2\mathsf{E}^{1/2}[\widehat{g}'(X_{ij}^{\mathrm{T}}\beta;\beta)-g'(X_{ij}^{\mathrm{T}}\beta)]^2\\&\leq c\sqrt{n}h^2h_1+c(nhh_1^3)^{-1/2}\to 0.\end{aligned}$$

这就证得$M_1\xrightarrow{P}0$. 用$M_{2,s}$表示$M_2$的第$s$个分量, 那么

$$\begin{aligned}M_{2,s}&=\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\sum_{j=1}^{m_i}\Big[\sum_{i_1=1}^{n}\sum_{j_1=1}^{m_{i_1}}\widetilde{W}_{ni_1j_1}(X_{ij}^{\mathrm{T}}\beta;\beta,h_1)g(X_{ij}^{\mathrm{T}}\beta)-g'(X_{ij}^{\mathrm{T}}\beta)\Big]X_{ijs}e_{ij}\\&\quad+\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\sum_{j=1}^{m_i}\sum_{i_1=1}^{n}\sum_{j_1=1}^{m_{i_1}}\widetilde{W}_{ni_1j_1}(X_{ij}^{\mathrm{T}}\beta;\beta,h_1)X_{ijs}e_{ij}e_{i_1j_1}\\&\triangleq M_{21,s}+M_{22,s}.\end{aligned}$$

利用引理4.1和引理4.2得

$$\begin{aligned}\mathsf{E}(M_{21,s}^2)&\leq cn^{-1}\sum_{i=1}^{n}\sum_{j=1}^{m_i}\mathsf{E}\Big[\sum_{i_1=1}^{n}\sum_{j_1=1}^{m_{i_1}}\widetilde{W}_{ni_1j_1}(X_{ij}^{\mathrm{T}}\beta;\beta,h_1)g(X_{i_1j_1}^{\mathrm{T}}\beta)-g'(X_{ij}^{\mathrm{T}}\beta)\Big]^2\\&\leq ch_1^2\to 0,\\\mathsf{E}(M_{22,s}^2)&=n^{-1}\mathsf{E}\Big[\sum_{i=1}^{n}\sum_{j=1}^{m_i}\sum_{i_1=1}^{n}\sum_{j_1=1}^{m_{i_1}}\widetilde{W}_{ni_1j_1}(X_{ij}^{\mathrm{T}}\beta;\beta,h_1)X_{ijs}e_{ij}e_{i_1j_1}\Big]^2\\&=n^{-1}\mathsf{E}\Big[\sum_{i=1}^{n}\sum_{j=1}^{m_i}\widetilde{W}_{nij}(X_{ij}^{\mathrm{T}}\beta;\beta,h_1)X_{ijs}e_{ij}^2\Big]^2\\&\quad+n^{-1}\mathsf{E}\Big[\sum_{i=1}^{n}\sum_{j=1}^{m_i}\sum_{i_1\neq i}\sum_{j_1\neq j}\widetilde{W}_{ni_1j_1}(X_{ij}^{\mathrm{T}}\beta;\beta,h_1)X_{ijs}e_{ij}e_{i_1j_1}\Big]^2\\&\leq c(nh_1^3)^{-1}+c(nh_1^2)^{-1}+c(n^2h_1^5)^{-1}\to 0.\end{aligned}$$

因此

$$\mathsf{E}(M_{2,s}^2)\leq 2\mathsf{E}(M_{21,s}^2)+2\mathsf{E}(M_{22,s}^2)\to 0,$$

即有$M_2\xrightarrow{P}0$. 类似的方法可证得$\mathsf{E}(M_{3,s}^2)\leq ch^4\to 0$, 即有$M_3\xrightarrow{P}0$, 而$M_4\xrightarrow{P}0$是显然的. 故由$M_l\xrightarrow{P}0$ $(l=1,2,3,4)$, 结合(4.13) 和(4.14)式证得引理4.4. $\square$

**引理 4.5** 在引理4.4的条件下, 则

$$\frac{1}{n}\sum_{i=1}^{n}\widehat{\eta}_i(\beta^{(r)})\widehat{\eta}_i^{\mathrm{T}}(\beta^{(r)}) \xrightarrow{P} V_0(\beta^{(r)}),$$

其中$V_0(\beta^{(r)})$在条件C6中定义.

**证明:** 令

$$\begin{aligned} R_{ni} &= \sum_{j=1}^{m_i}[g(X_{ij}^{\mathrm{T}}\beta)-\widehat{g}(X_{ij}^{\mathrm{T}}\beta;\beta)][\widehat{g}'(X_{ij}^{\mathrm{T}}\beta;\beta)-g'(X_{ij}^{\mathrm{T}}\beta)]J_{\beta^{(r)}}^{\mathrm{T}}X_{ij} \\ &\quad+\sum_{j=1}^{m_i}[g(X_{ij}^{\mathrm{T}}\beta)-\widehat{g}(X_{ij}^{\mathrm{T}}\beta;\beta)]g'(X_{ij}^{\mathrm{T}}\beta)J_{\beta^{(r)}}^{\mathrm{T}}X_{ij} \\ &\quad+\sum_{j=1}^{m_i}[\widehat{g}'(X_{ij}^{\mathrm{T}}\beta;\beta)-g'(X_{ij}^{\mathrm{T}}\beta)]J_{\beta^{(r)}}^{\mathrm{T}}X_{ij}e_{ij}. \end{aligned}$$

因此有

$$\begin{aligned} \frac{1}{n}\sum_{i=1}^{n}\widehat{\eta}_i(\beta^{(r)})\widehat{\eta}_i^{\mathrm{T}}(\beta^{(r)}) &= \frac{1}{n}\sum_{i=1}^{n}\Lambda_i(\beta^{(r)})e_ie_i^{\mathrm{T}}\Lambda_i^{\mathrm{T}}(\beta^{(r)})+\frac{1}{n}\sum_{i=1}^{n}R_{ni}R_{ni}^{\mathrm{T}} \\ &\quad+\frac{1}{n}\sum_{i=1}^{n}\Lambda_i(\beta^{(r)})e_iR_{ni}^{\mathrm{T}}+\frac{1}{n}\sum_{i=1}^{n}R_{ni}e_i^{\mathrm{T}}\Lambda_i^{\mathrm{T}}(\beta^{(r)}) \\ &\,\widehat{=}\, R_1+R_2+R_3+R_4. \end{aligned}$$

由大数定律可知: $R_1\xrightarrow{P}V_0(\beta^{(r)})$. 因此只需证明$R_l\xrightarrow{P}0$, $l=2,3,4$, 便可证得引理4.5. 首先证明$R_2\xrightarrow{P}0$, 其他可类似证得. $R_{2,st}$表示$R_2$的$(s,t)$元素, $R_{ni,s}$表示$R_{ni}$的第$s$个分量. 则由Cauchy-Schwarz不等式, 有

$$|R_{2,st}|\le\Big(\frac{1}{n}\sum_{i=1}^{n}R_{ni,s}^2\Big)^{1/2}\Big(\frac{1}{n}\sum_{i=1}^{n}R_{ni,t}^2\Big)^{1/2}.$$

由引理4.1和引理4.3, 类似引理4.4的证明可以证得

$$\frac{1}{n}\sum_{i=1}^{n}R_{ni,s}^2\xrightarrow{P}0.$$

因此证得$R_2\xrightarrow{P}0$. 类似的方法可以证明$R_3\xrightarrow{P}0$和$R_4\xrightarrow{P}0$. □

**引理 4.6** 假设条件C1–C7成立, 如果$\beta$是参数真值, 则

$$\max_{1\le i\le n}\|\widehat{\eta}_i(\beta^{(r)})\|=o_P(n^{1/2}).$$

**证明:** 由引理4.1–引理4.5, 可用文献[16]中引理A.6的证明方法容易证得本引理, 在此省略. □

**定理2.1的证明:** 对(2.9)式应用Taylor公式, 得

$$\widehat{l}(\beta^{(r)})=2\sum_{i=1}^{n}\Big[\lambda^{\mathrm{T}}\widehat{\eta}_i(\beta^{(r)})-\frac{1}{2}(\lambda^{\mathrm{T}}\widehat{\eta}_i(\beta^{(r)})^2\Big]+o_P(1),\tag{4.15}$$

由(2.10)式, 可以推出

$$\begin{aligned} 0 &= \frac{1}{n}\sum_{i=1}^{n}\frac{\widehat{\eta}_i(\beta^{(r)})}{1+\lambda^{\rm T}\widehat{\eta}_i(\beta^{(r)})} \\ &= \frac{1}{n}\sum_{i=1}^{n}\widehat{\eta}_i(\beta^{(r)})-\frac{1}{n}\sum_{i=1}^{n}\widehat{\eta}_i(\beta^{(r)})\widehat{\eta}_i^{\rm T}(\beta^{(r)})\lambda+\frac{1}{n}\sum_{i=1}^{n}\frac{\widehat{\eta}_i(\beta^{(r)})[\lambda^{\rm T}\widehat{\eta}_i(\beta^{(r)})]^2}{1+\lambda^{\rm T}\widehat{\eta}_i(\beta^{(r)})}. \end{aligned}$$

应用文献[10]的证明方法, 类似可得: $\lambda=O_P(n^{-1/2})$. 结合引理4.5和引理4.6得

$$\sum_{i=1}^{n}\lambda^{\rm T}\widehat{\eta}_i(\beta^{(r)})=\sum_{i=1}^{n}(\lambda^{\rm T}\widehat{\eta}_i(\beta^{(r)}))^2+o_P(1), \tag{4.16}$$

$$\lambda=\Big(\sum_{i=1}^{n}\widehat{\eta}_i(\beta^{(r)})\widehat{\eta}_i^{\rm T}(\beta^{(r)})\Big)^{-1}\sum_{i=1}^{n}\widehat{\eta}_i(\beta^{(r)})+o_P(n^{-1/2}). \tag{4.17}$$

结合(4.15)–(4.17)式, 即证明了

$$\widehat{l}(\beta^{(r)})=\Big\{\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\widehat{\eta}_i(\beta^{(r)})\Big\}^{\rm T}\Big(\frac{1}{n}\sum_{i=1}^{n}\widehat{\eta}_i(\beta^{(r)})\widehat{\eta}_i^{\rm T}(\beta^{(r)})\Big)^{-1}\Big\{\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\widehat{\eta}_i(\beta^{(r)})\Big\}+o_P(1), \tag{4.18}$$

由(4.18)式, 引理4.4和引理4.5, 得

$$\widehat{l}(\beta^{(r)})=\Big\{\frac{1}{\sqrt{n}}V^{-1/2}(\beta^{(r)})\sum_{i=1}^{n}\widehat{\eta}_i(\beta^{(r)})\Big\}^{\rm T}D(\beta^{(r)})\Big\{\frac{1}{\sqrt{n}}V^{-1/2}(\beta^{(r)})\sum_{i=1}^{n}\widehat{\eta}_i(\beta^{(r)})\Big\}+o_P(1),$$

这里$D(\beta^{(r)})=V^{1/2}(\beta^{(r)})V_0^{-1}(\beta^{(r)})V^{1/2}(\beta^{(r)})$. 记$\widetilde{D}={\rm diag}(w_1,\cdots,w_{p-1})$, 此处$w_i$ $(i=1,\cdots,p-1)$在定理2.1中定义. 则存在一个正交阵$Q$, 使得$Q^{\rm T}\widetilde{D}Q=D(\beta^{(r)})$, 于是

$$\widehat{l}(\beta^{(r)})=\Big\{\frac{1}{\sqrt{n}}QV^{-1/2}(\beta^{(r)})\sum_{i=1}^{n}\widehat{\eta}_i(\beta^{(r)})\Big\}^{\rm T}\widetilde{D}\Big\{\frac{1}{\sqrt{n}}QV^{-1/2}(\beta^{(r)})\sum_{i=1}^{n}\widehat{\eta}_i(\beta^{(r)})\Big\}+o_P(1). \tag{4.19}$$

由引理4.4, 可得

$$\frac{1}{\sqrt{n}}QV^{-1/2}(\beta^{(r)})\sum_{i=1}^{n}\widehat{\eta}_i(\beta^{(r)})\xrightarrow{L}N(0,I_{p-1}), \tag{4.20}$$

其中$I_{p-1}$是一个$(p-1)\times(p-1)$的单位矩阵. 由(4.19)和(4.20)式, 即完成了定理2.1的证明. □

**定理2.2的证明:** 由$\widehat{l}_{\rm ad}(\beta^{(r)})$的定义, 利用(4.15)式, 可以推出

$$\widehat{l}_{\rm ad}(\beta^{(r)})=\Big\{\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\widehat{\eta}_i(\beta^{(r)})\Big\}^{\rm T}\widehat{V}^{-1}(\beta^{(r)})\Big\{\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\widehat{\eta}_i(\beta^{(r)})\Big\}+o_P(1).$$

类似于引理4.5的证明, 可以证明$\widehat{V}_n(\beta^{(r)})\xrightarrow{P}V(\beta^{(r)})$. 利用引理4.4, 即可完成定理2.2的证明. □

对于定理2.3的证明, 只需要在引理4.1–引理4.3中使用相同的窗宽$h$, 此外还需要下面的引理4.7. 由引理4.7, 类似于定理2.1的证明容易证得定理2.3, 此处省略.

**引理 4.7**　在定理2.3条件下, 如果$\beta$是参数真值, 并且$\beta$的第$r$个分量是正数, 则

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\widehat{\eta}_{\mathrm{bc},i}(\beta^{(r)}) \xrightarrow{L} N(0,V(\beta^{(r)})),$$
$$\frac{1}{n}\sum_{i=1}^{n}\widehat{\eta}_{\mathrm{bc},i}(\beta^{(r)})\widehat{\eta}_{\mathrm{bc},i}^{\mathrm{T}}(\beta^{(r)}) \xrightarrow{P} V(\beta^{(r)}),$$
$$\max_{1\le i\le n}\|\widehat{\eta}_{\mathrm{bc},i}(\beta^{(r)})\| = o_P(n^{1/2}).$$

**证明:**　利用引理4.1–引理4.3, 类似引理4.4–引理4.6的证明方法容易证得本引理, 在此省略.　□

**致谢**　感谢编委和审稿人, 他们提出了一些非常有益的修改意见, 使得本文有了进一步的提高.

## 参 考 文 献

[1] Ichimura, H., Semiparametric least squares (SLS) and weighted SLS estimation of single-index models, *J. Econometrics*, **58**(1993), 71–120.

[2] Härdle, W., Hall, P. and Ichimura, H., Optimal smoothing in single-index models, *Ann. Statist.*, **21**(1993), 157–178.

[3] Stoker, T.M., Consistent estimation of scaled coefficients, *Econometrica*, **54(6)**(1986), 1461–1481.

[4] Härdle, W. and Stoker, T.M., Investigating smooth multiple regression by the method of average derivatives, *J. Amer. Statist. Assoc.*, **84**(1989), 986–995.

[5] Härdle, W. and Tsybakov, A.B., How sensitive are average derivative, *J. Econometrics*, **58**(1993), 31–48.

[6] Hristache, M., Juditsky, A. and Spokoiny, V., Direct estimation of the index coefficient in a single-index model, *Ann. Statist.*, **29**(2001), 595–623.

[7] Powell, J.L., Stock, J.H. and Stoker, T.M., Semiparametric estimation of index coefficients, *Econometrica*, **57**(1989), 474–523.

[8] Xue, L.G. and Zhu, L.X., Empirical likelihood for single-index models, *J. Multivariate Anal.*, **97**(2006), 1295–1312.

[9] Chiou, J.M. and Müller, H.G., Estimated estimating equations: semiparametric inference for clustered and longitudinal data, *J. Roy. Statist. Soc. B*, **67(4)**(2005), 531–553.

[10] Owen, A., Empirical likelihood ratio confidence intervals for a single function, *Biometrika*, **75(2)** (1988), 237–249.

[11] Owen, A., Empirical likelihood for linear models, *Ann. Statist.*, **19(4)**(1991), 1725–1747.

[12] Kolaczyk, E.D., Empirical likelihood and generalized linear models, *Statistica Sinica*, **4**(1994), 199–218.

[13] Shi, J. and Lau, T.S., Empirical likelihood for partially linear models, *J. Multivariate Anal.*, **72(1)** (2000), 132–148.

[14] Wang, Q.H. and Jing, B.Y., Empirical likelihood for partially linear models with fixed design, *Statist. Probab. Lett.*, **41(4)**(1999), 425–433.

[15] Xue, L.G. and Zhu, L.X., Empirical likelihood confidence regions of the parameters in a partially linear single-index model, *Sci. China, Ser. A*, **48(10)**(2005), 1333–1348.

[16] Zhu, L.X. and Xue, L.G., Empirical likelihood confidence regions in a partially linear single-index model, *J. Roy. Statist. Soc. B*, **68**(2006), 549–570.

[17] Fan, J.Q. and Gijbels, I., *Local Polynomial Modeling and Its Applications*, Chapman and Hall, London, 1996.

[18] Rao, J.N.K. and Scott, A.J., The analysis of categorical data from complex sample surveys: chi-squared tests for goodness of fit and independence in two-way tables, *J. Amer. Statist. Assoc.*, **76** (1981), 221–230.

[19] Wang, Q.H. and Rao, J.N.K., Empirical likelihood-based inference in linear errors-in-covariables models with validation data, *Biometrika*, **89(2)**(2002), 345–357.

# Empirical Likelihood Confidence Regions of the Parameters in Single-Index Models for Longitudinal Data

Li Gaorong[1,2] Feng Sanying[3] Xue Liugen[1]

([1] *College of Applied Sciences, Beijing University of Technology, Beijing, 100124*)

([2] *School of Finance and Statistics, East China Normal University, Shanghai, 200241*)

([3] *College of Mathematics and Science, Luoyang Normal University, Luoyang, 471022*)

In this paper, the single-index model for longitudinal data is considered, three empirical log-likelihood ratio statistics for the unknown parameters in the model are suggested to accommodate the independence of different subjects for longitudinal data. It is proved that the proposed statistics are asymptotically chi-square distribution under some suitable conditions, and hence they can be used to construct the confidence regions of the parameters. Furthermore, it is shown that the bias corrected empirical log-likelihood ratio shares some desired features. A simulation study is conducted to illustrate our methods.

**Keywords:** Longitudinal data, single-index model, empirical likelihood, confidence region.

**AMS Subject Classification:** 62G05.