引用格式:白璐,陈夏.固定和自适应设计下高维广义线性模型的经验似然检验[J].陕西师范大学学报(自然科学版),2018,46(4):6-11.[BAI L, CHEN X. Empirical likelihood test for high-dimensional generalized linear models with fixed and adaptive designs[J]. Journal of Shaanxi Normal University (Natural Science Edition), 2018,46(4):6-11.] DOI:10.15983/j.cnki.jsnu.2018.04.142

固定和自适应设计下高维广义线性模型的经验似然检验

白 璐,陈 夏*
(陕西师范大学 数学与信息科学学院,陕西 西安 710119)

摘 要:利用经验似然方法,讨论固定设计和自适应设计下高维广义线性模型中的参数检验问题。数值计算结果表明,所构造的经验对数似然比统计量渐近于标准卡方分布,具有较稳定的真实检验水平和功效。

关键词:广义线性模型;经验似然;高维数据;假设检验
中图分类号:O221.1 **文献标志码**:A **文章编号**:1672-4291(2018)04-0006-06

Empirical likelihood test for high-dimensional generalized linear models with fixed and adaptive designs

BAI Lu, CHEN Xia*
(School of Mathematics and Information Science, Shaanxi Normal University, Xi′an 710119, Shaanxi, China)

Abstract: An empirical likelihood method is considered to test the coefficients in a high-dimensional generalized linear model with fixed and adaptive designs. Numerical analysis show that the empirical log-likelihood ratio at the true parameter converges to the standard chi-square distribution, and the proposed empirical likelihood test has a very stable size with respect to the number of covariates and is powerful too.

Keywords: generalized linear models; empirical likelihood; high-dimensional data; hypothesis test
MR subject classification: 62J12

广义线性模型作为经典线性模型的重要推广,自从 Nelder 和 Wedderburn[1] 引入以来,它已被应用到许多领域。关于广义线性模型参数估计渐近理论的研究可参见文献[2-10]。

经验似然方法作为一种非参数统计推断方法由 Owen 在 1988 年提出[11]。该方法在构造置信域和假设检验方面有许多突出的优点,例如无需对渐近方差进行估计、置信域的形状由数据自行决定、域保持性、变换不变性、Bartlett 纠偏性以及无需构造枢轴统计量等。许多研究者将经验似然方法应用到处理各种数据的问题[12-20]。更多关于经验似然方法的介绍可参见文献[21-22]。

近年来,高维数据分析被广泛地应用于各个领域,如生物信息学、医学、金融分析和风险控制等。故当维数 p 发散时,即当样本量 $n\to\infty$ 时,p 也趋于无穷,甚至在 p 大于 n 的情况下,研究广义线性模型的统计推断问题有重要的理论意义和实用价值。但当 p 发散时,传统检验方法的计算会

收稿日期:2017-07-10
基金项目:国家自然科学基金(11201276);陕西省自然科学基础研究计划(2014JQ1042);中央高校基本科研业务费专项资金(GK201503015)
*通信作者:陈夏,男,副教授,博士,主要从事概率统计方面的研究。E-mail:xchen80@snnu.edu.cn

很困难，甚至不可行，具体讨论可参见文献[23]。更多关于高维数据经验似然方法的讨论见文献[24-26]。基于样本分割方法，Peng 等[27]在 2014 年提出了一种高维线性模型下的经验似然检验方法，该方法对线性模型参数的检验有较好的实际检验水平（size）和功效（power）。Zang 等[28]于 2017 年将这种方法应用于高维广义线性模型参数的检验。关于样本分割方法的其他研究可参见文献[29-30]。

本文利用样本分割法，在协变量为固定设计和自适应设计情形下，研究高维广义线性模型的经验似然检验问题。当协变量维数 p 是固定或发散时，对于已知的 $\boldsymbol{\beta}_0\in\mathbf{R}^p$，给出一种经验似然方法检验以下问题：

$$H_0:\boldsymbol{\beta}=\boldsymbol{\beta}_0 \text{ vs } H_a:\boldsymbol{\beta}\neq\boldsymbol{\beta}_0。\tag{1}$$

特别地，我们对备择假设有密集转变的情况感兴趣（即在许多维的细小转变而不是一些维的较大转变）。

1　方法与主要结论

1.1　固定设计的经验似然检验

考虑如下的广义线性模型：

$$y_i=\mu(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{\beta}_0)+\boldsymbol{\varepsilon}_i,i=1,2,\cdots,n。\tag{2}$$

其中 $y_i\in\mathbf{R}$ 为响应变量，$\boldsymbol{X}_i=(X_{i1},X_{i2},\cdots,X_{ip})^{\mathrm{T}}$ 为协变量。固定设计的情形，即指 X_i 为非随机的设计向量，或者当以 X_i 为条件时做讨论有效。$\boldsymbol{\beta}_0$ 为 p 维未知参数向量 $\boldsymbol{\beta}$ 的真值。$\mu(\cdot)$是光滑的连接函数，且满足对任意的 $t\in\mathbf{R}$，$\mu(t)$连续可微，$\dot{\mu}(t)=\mathrm{d}\mu(t)/\mathrm{d}t>0$。$\varepsilon_i$ 为独立同分布的随机误差项，$E(\varepsilon_i)=0$，$\mathrm{Var}(\varepsilon_i)=\sigma^2$。

基于模型(2)，文献[2，4，8，10]考虑了一类简单而重要的拟似然方程

$$\sum_{i=1}^{n}\boldsymbol{X}_i(\boldsymbol{y}_i-\boldsymbol{\mu}(\boldsymbol{X}_i^{\mathrm{T}}\boldsymbol{\beta}))=0,\tag{3}$$

其解定义为 $\boldsymbol{\beta}_0$ 的极大拟似然估计。在参数维数 p 固定的情形下，基于模型(2) 和方程(3)，文献[20] 提出如下关于 $\boldsymbol{\beta}$ 的经验对数似然比函数：

$$l_n^{(\mathrm{T})}(\boldsymbol{\beta})=-2\max\left\{\sum_{i=1}^{n}\log(n\omega_i):\omega_i\geqslant 0,\right.$$
$$\left.\sum_{i=1}^{n}\omega_i=1,\sum_{i=1}^{n}\omega_i(\boldsymbol{y}_i-\boldsymbol{\mu}(X_i^{\mathrm{T}}\boldsymbol{\beta}))X_i=0\right\}。$$

在一定的条件下，证明了 $l_n^{(\mathrm{T})}(\boldsymbol{\beta}_0)$ 收敛于自由度为 p 的卡方分布，即 Wilks 定理成立。因此，$l_n^{(\mathrm{T})}(\boldsymbol{\beta}_0)$ 可用于参数β的检验问题。然而，当参数维数 $p\geqslant n$ 时，这种检验方法却不可行。因此，本文基于样本分割方法提出一种新的检验统计量。

首先，我们将样本分为两组。令 $m=[n/2]$，表示 $n/2$ 的整数部分。定义 $\widetilde{\boldsymbol{X}}_i=\boldsymbol{X}_{m+i}$，$\widetilde{\boldsymbol{Y}}_i=\boldsymbol{Y}_{m+i}$，$\widetilde{\boldsymbol{\varepsilon}_i}=\boldsymbol{\varepsilon}_{m+i}$。利用模型(2) 和拟似然方程(3)，构造如下辅助随机向量：

$$\widetilde{\boldsymbol{\eta}}_i(\beta)=(\boldsymbol{X}_i(\boldsymbol{y}_i)-\mu(\boldsymbol{X}_i^{\mathrm{T}}\boldsymbol{\beta})))^{\mathrm{T}}(\widetilde{\boldsymbol{X}}_i(\widetilde{\boldsymbol{y}}_i-$$
$$\mu(\widetilde{\boldsymbol{X}}_i^{\mathrm{T}}\boldsymbol{\beta}))),i=1,2,\cdots,m。$$

然而，注意到 $E[\widetilde{\boldsymbol{\eta}}_i(\boldsymbol{\beta})]=O(\|\boldsymbol{\beta}-\boldsymbol{\beta}_0\|^2)$ 而不是 $O(\|\boldsymbol{\beta}-\boldsymbol{\beta}_0\|)$，其中 $\|\boldsymbol{v}\|$ 表示向量 $\boldsymbol{v}\in\mathbf{R}^p$ 的 L_2 范数。当 $\|\boldsymbol{\beta}-\boldsymbol{\beta}_0\|$ 很小时，直接应用辅助随机向量 $\widetilde{\boldsymbol{\eta}}_i(\boldsymbol{\beta})$ 通常功效较差。这种现象在 Peng 等[27] 的论文中也有讨论，类似文献[27] 在线性模型中提到的方法，可再增加一个线性方程使其接近 $O(\|\boldsymbol{\beta}-\boldsymbol{\beta}_0\|_1)$ 来提高功效，其中 $\|\boldsymbol{\beta}-\boldsymbol{\beta}_0\|_1$ 是 L_1 范数。如此便能捕捉到 $\beta-\beta_0$ 的细微变化。更具体地，定义

$$\boldsymbol{\eta}_i^*(\boldsymbol{\beta})=(\boldsymbol{X}_i(y_i-\mu(\boldsymbol{X}_i^{\mathrm{T}}\boldsymbol{\beta})))^{\mathrm{T}}\mathbf{1}_p+(\widetilde{\boldsymbol{X}}_i(\widetilde{y}_i-$$
$$\mu(\widetilde{\boldsymbol{X}}_i^{\mathrm{T}}\boldsymbol{\beta})))\mathbf{1}_p,i=1,2,\cdots,m,$$

其中 $\mathbf{1}_p=(1,1,\cdots,1)^{\mathrm{T}}\in\mathbf{R}^p$。令 $\boldsymbol{\eta}_i(\boldsymbol{\beta})=(\widetilde{\boldsymbol{\eta}}_i(\boldsymbol{\beta}),\boldsymbol{\eta}_i^*(\boldsymbol{\beta}))^{\mathrm{T}}$，检验问题(3) 等价于检验

$$H_0:E[\boldsymbol{\eta}_i(\boldsymbol{\beta})]=0 \text{ vs } H_a:E[\boldsymbol{\eta}_i(\boldsymbol{\beta})]\neq 0。$$

同时，若 $E[\boldsymbol{\eta}_i(\boldsymbol{\beta})]=0$，我们可以构造一个估计方程 $\sum_{i=1}^{n}\boldsymbol{\eta}_i(\boldsymbol{\beta})=0$。由此，可定义关于 $\boldsymbol{\beta}$ 的经验对数似然比函数为

$$l_n(\boldsymbol{\beta})=-2\max\left\{\sum_{i=1}^{n}\log(n\omega_i):\omega_i\geqslant 0,\right.$$
$$\left.\sum_{i=1}^{n}\omega_i=1,\sum_{i=1}^{n}\omega_i\boldsymbol{\eta}_i(\boldsymbol{\beta})=0\right\},$$

由 Lagrange 乘子法得

$$\omega_i=\frac{1}{n[1+\boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{\eta}_i(\boldsymbol{\beta})]},$$

则 $l_n(\boldsymbol{\beta})$ 可表示为

$$l_n(\boldsymbol{\beta})-2\sum_{i=1}^{n}\log(1+\boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{\eta}_i(\boldsymbol{\beta})),$$

其中 $\boldsymbol{\lambda}=\boldsymbol{\lambda}(\boldsymbol{\beta})$ 满足

$$\frac{1}{n}\sum_{i=1}^{n}\frac{\boldsymbol{\eta}_i(\boldsymbol{\beta})}{1+\boldsymbol{\lambda}_i^{\mathrm{T}}\boldsymbol{\eta}_i(\boldsymbol{\beta})}=0。$$

由文献[28] 中定理 1 的结果知，在协变量为随机设计的情形下，$l_n(\boldsymbol{\beta}_0)$ 收敛于自由度为 2 的标准卡方分布。数值模拟结果表明，在协变量为固定和自适应设计下，本文所提方法具有相同的渐近理论性质和良好的检验效果。

1.2　自适应设计的经验似然检验

考虑自适应设计下的情况，自适应即指对每个 i，$\boldsymbol{X}_i$ 的值依赖于先前的观测值 $\{(y_j,\boldsymbol{X}_j),j=1,2,\cdots,i-1\}$。令 $\{F_i,i\geqslant 1\}$ 为一列递增的 σ -域，且满足

y_i 为 F_j 可测,$\boldsymbol{X}_i$ 为 F_{i-1} 可测,则自适应情形下的广义线性模型结构为

$$E(y_i \mid F_{i-1}) = \mu(\boldsymbol{X}_i^{\mathrm{T}}\boldsymbol{\beta}_0),i = 1,2,\cdots,n。$$

令 $\varepsilon_i = y_i - \mu(\boldsymbol{X}_i^{\mathrm{T}}\boldsymbol{\beta}_0) = y_i - E(y_i \mid F_{i-1})$,易知 $\{\varepsilon_i, i \geqslant 1\}$ 为关于 $\{F_i, i \geqslant 1\}$ 的鞅差序列。类似固定设计中的方法,可得到自适应设计下 $\boldsymbol{\beta}$ 的经验对数似然比函数,仍记为 $l_n(\boldsymbol{\beta})$。数值研究表明,自适应情形下的模拟结果和固定设计下类似。

2 数值研究

这里给出了一些数值模拟。通过两个模拟例子来分别说明所提出的经验似然检验方法对固定设计和自适应设计下广义线性模型的有限样本性质,并将它与传统经验似然方法在 $n < p$、$n = p$、$n > p$ 三种情况下的检验水平和功效进行比较。两例均模拟 10 000 组数据。

例 1 考虑如下固定设计下的广义线性模型:

$$Y_i = \mu(\boldsymbol{X}_i^{\mathrm{T}}\boldsymbol{\beta}) + \varepsilon_i, i = 1,2,\cdots,n。$$

其中,$\mu(x) = x^3$,协变量 $\boldsymbol{X}_i \sim N(0, \boldsymbol{\Sigma}_0)$,$\boldsymbol{\Sigma}_0 = (0.5^{|i-j|})_{1 \leqslant i,j \leqslant p}$,$\varepsilon_i \sim U(-\sqrt{3}, \sqrt{3})$,$\beta_0 = (1_{p/2}, 0_{p/2})^{\mathrm{T}}$ 为 p 维向量,对于 $\Delta > 0$,$\boldsymbol{\beta} = \boldsymbol{\beta}_0 + \Delta/10n$。注意到当 $\Delta > 0$ 时,备择假设在此模型中有密集的转变。

对于检验问题(1),利用基于检验统计量 $l_n^{(T)}(\boldsymbol{\beta})$ 的传统经验似然检验方法(TEL)和基于检验统计量 $l_n(\boldsymbol{\beta})$ 的新经验似然检验方法(NEL),分别在 $p = 5,10,20,\cdots,100$,显著性水平为 0.05 时,计算其实际检验水平($\Delta = 0$)和功效($\Delta = 0.2$)。

模拟结果分别列在表 1 和表 2 中。TEL 在 p 稍大时没有较好的实际检验水平,如表中结果所示。当 $p \geqslant n$ 时,该方法甚至失效(无法计算结果),因此其检验的功效没有实际意义。然而,不管 p 与 n 关系如何,本文给出的 NEL 方法相对于维数 p 具有稳定的实际检验水平,同时功效也很好。而且,当样本量 n 增大时,所提方法的实际检验水平更加接近于给定的显著性水平。

表 1 例 1 中 0.05 水平下 TEL 和 NEL 的实际检验水平($\boldsymbol{\Delta = 0}$)

Tab. 1 Sizes of the TEL and NEL at 0.05 level are given for the case of $\boldsymbol{\Delta = 0}$ in example 1

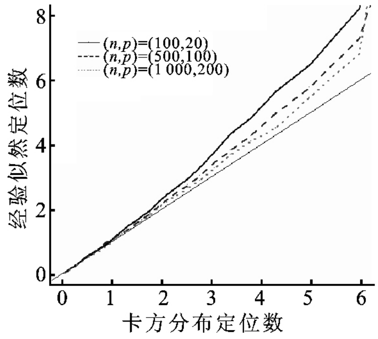
p	$n = 50$		$n = 100$		$n = 200$		$n = 500$	
	TEL	NEL	TEL	NEL	TEL	NEL	TEL	NEL
5	0.108 8	0.107 5	0.110 2	0.104 3	0.076 9	0.081 5	0.070 3	0.058 3
10	0.343 6	0.109 8	0.178 6	0.106 5	0.127 3	0.083 0	0.117 2	0.056 9
20	0.798 3	0.114 1	0.399 0	0.105 2	0.294 5	0.083 7	0.175 8	0.055 2
30	1	0.113 2	0.696 3	0.105 4	0.517 5	0.084 9	0.281 6	0.059 1
40	1	10.110 9	0.984 7	0.103 2	0.736 8	0.082 3	0.405 5	0.053 7
50	—	0.112 5	0.997 9	0.101 9	0.934 1	0.082 7	0.573 4	0.061 8
60	—	0.116 4	1	0.102 9	0.987 2	0.084 3	0.729 1	0.052 4
70	—	0.115 5	1	0.107 6	0.999 1	0.079 2	0.882 7	0.054 7
80	—	0.115 2	1	0.099 8	1	0.084 4	0.943 5	0.060 3
90	—	0.117 1	1	0.102 1	1	0.085 0	0.997 9	0.057 0
100	—	0.111 9	—	0.103 7	1	0.081 6	1	0.051 2

表 2 例 1 中 0.05 水平下 TEL 和 NEL 的功效($\boldsymbol{\Delta = 0.2}$)

Tab. 2 Powers of the TEL and NEL at 0.05 level are given for the case of $\boldsymbol{\Delta = 0.2}$ in example 1

p	$n = 50$		$n = 100$		$n = 200$	
	TEL	NEL	TEL	NEL	TEL	NEL
5	0.186 1	0.136 4	0.108 9	0.113 7	0.097 2	0.101 7
10	0.379 0	0.179 3	0.243 7	0.154 6	0.158 4	0.207 8
20	0.895 4	0.306 8	0.706 5	0.432 0	0.476 2	0.438 1
30	1	0.619 5	0.984 2	0.763 9	0.708 0	0.700 9
40	1	0.803 2	1	0.928 3	0.896 3	0.812 0
50	—	0.986 5	1	0.987 1	0.938 7	0.897 5
60	—	1	1	0.994 9	0.995 3	0.962 3
70	—	1	1	1	1	0.980 4
80	—	1	1	1	1	0.989 2
90	—	1	1	1	1	0.992 1
100	—	1	—	1	1	0.994 0

针对本例，图 1 给出了 $n > p$ 和 $p \gg n$ 两种情况下的 QQ 图，表明所提出的经验似然检验统计量在固定设计下渐近于自由度为 2 的标准卡方分布。同时，不论 p 与 n 大小关系如何，所提出的经验似然检验统计量均表现出优良的渐近性质，且随着样本量的增大，模拟结果更加接近理论值。



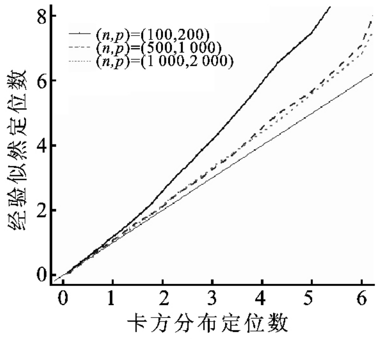


图 1　例 1 中在 $n > p$ 和 $p \gg n$ 两种情况下 NEL 检验统计量的 QQ 图

Fig. 1　QQ-plots of the NEL ratio in the case $n > p$ and $p \gg n$ for example 1

例 2　考虑如下自适应设计下的广义线性模型：

$$y_i = \mu(\boldsymbol{X}_i^{\mathrm{T}}\boldsymbol{\beta}_0) + v(\boldsymbol{X}_i^{\mathrm{T}}\boldsymbol{\beta}_0)\varepsilon_i, i = 1, 2, \cdots, n。$$

其中 $\mu(t) = 0.1\cos t, v(t) = 0.1\tanh t; \boldsymbol{X}_i = (1, y_{i-1}, y_{i-2}, \cdots, y_{i-(p-1)})$，初值 $y_0 = y_{-1} = \cdots = y_{-(p-2)} = 0$；$\varepsilon_i \sim N(0,1)$；$\boldsymbol{\beta}_0 = (1,1,\cdots,1)^{\mathrm{T}}$ 为 p 维向量，对于 $\Delta \geqslant 0$，$\boldsymbol{\beta} = \boldsymbol{\beta}_0 + \Delta$。当 $n = 50$ 时，$\Delta = 0.3$；$n = 100$ 时，$\Delta = 0.25$；$n = 200$ 时，$\Delta = 0.2$。其余变量参数的设定与例 1 相同。模拟结果列在表 3 和表 4 中，通过该例的结果可以得到与例 1 类似的结论。

表 3　例 2 中 0.05 水平下 TEL 和 NEL 的实际检验水平($\Delta = 0$)

Tab. 3　Sizes of the TEL and NEL at 0.05 level are given for the case of $\Delta = 0$ in example 2

p	$n = 50$		$n = 100$		$n = 200$		$n = 500$	
	TEL	NEL	TEL	NEL	TEL	NEL	TEL	NEL
5	0.181 1	0.127 5	0.132 0	0.083 2	0.088 2	0.073 4	0.058 1	0.054 2
10	0.432 7	0.121 6	0.286 0	0.089 1	0.155 4	0.073 9	0.062 8	0.055 3
20	0.810 4	0.123 5	0.497 3	0.082 0	0.305 7	0.075 4	0.098 3	0.050 1
30	0.993 5	0.122 0	0.739 2	0.086 3	0.491 7	0.070 9	0.168 4	0.049 7
40	1	0.122 1	0.851 7	0.087 5	0.695 2	0.071 3	0.255 7	0.051 3
50	—	0.125 9	0.917 2	0.063 7	0.870 6	0.077 8	0.354 2	0.051 0
60	—	0.120 2	0.999 3	0.085 9	0.958 9	0.072 5	0.433 6	0.048 7
70	—	0.106 7	1	0.080 2	0.987 7	0.074 4	0.575 8	0.051 1
80	—	0.109 1	1	0.089 7	1	0.072 1	0.673 9	0.050 4
90	—	0.102 8	1	0.082 6	1	0.070 5	0.748 5	0.052 5
100	—	0.101 7	—	0.080 7	1	0.071 6	0.839 0	0.053 1

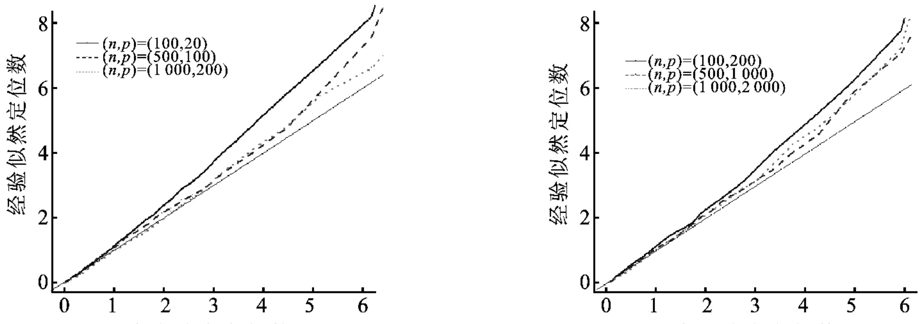
表 4　例 2 中 0.05 水平下 TEL 和 NEL 的功效($\Delta \neq 0$)

Tab. 4　Powers of the TEL and NEL at 0.05 level are given for the case of $\Delta \neq 0$ in example 2

p	$n = 50$		$n = 100$		$n = 200$	
	TEL	NEL	TEL	NEL	TEL	NEL
5	0.616 1	0.585 0	0.623 8	0.678 6	0.732 4	0.858 3
10	0.805 0	0.773 7	0.775 5	0.798 0	0.854 1	0.932 7
20	0.950 4	0.809 6	0.862 5	0.832 7	0.872 3	0.941 2
30	0.998 4	0.820 6	0.951 9	0.904 8	0.908 0	0.948 0
40	1	0.831 1	0.992 6	0.919 3	0.924 0	0.953 4
50	—	0.840 9	0.998 9	0.928 7	0.930 2	0.956 8
60	—	0.842 3	1	0.940 1	0.948 7	0.960 9
70	—	0.843 1	1	0.952 3	0.954 9	0.968 1
80	—	0.840 7	1	0.958 8	0.965 3	0.972 9
90	—	0.842 9	1	0.960 5	0.970 3	0.978 9
100	—	0.845 6	—	0.971 2	0.976 9	0.980 2

同样地,针对本例,图2给出了$n>p$和$p\gg n$两种情况下的QQ图,表明所提出的经验似然检验统计量在自适应设计下渐近于自由度为2的标准卡方分布。

综上,所提出的经验似然检验方法相对于协变量的维数有比较稳定的检验水平,当n或p较大时,功效也较高。不论$n>p$或$p\gg n$,所提出的经验似然方法均具有较好的性质。同时,该检验方法不需要计算高维协方差阵的逆矩阵,可通过使用R程序包emplik实现。本文结论没有区分p是固定或发散的情况,并且在备择假设的密集转变情形下依然有效。



a.$n>p$　　b.$p\gg n$

图2　例2中在$n>p$和$p\gg n$两种情况下NEL检验统计量的QQ图

Fig. 2　QQ-plots of the NEL ratio in the case $n>p$ and $p\gg n$ for example 2

3　结论

本文主要讨论了经验似然方法在固定设计和自适应设计下的高维广义线性模型假在设检验问题中的应用。当参数维数p是固定或者发散时,我们给出的经验似然检验方法要优于传统方法,相对于协变量的维数有一个稳定的实际检验水平,同时功效表现也较好。值得注意的是,该方法没有区分参数维数趋于无穷大的速度是否比样本量趋于无穷大的速度快。

参考文献:

[1] NELDER J A, WEDDERBURN R W M. Generalized linear models[J]. Journal of the Royal Statistical Society Series A, 1972, 135(3): 370-384.

[2] CHEN K, HU Y, YING Z L. Strong consistency of maximum quasi-likelihood estimate in generalized linear models with fixed and adaptive designs[J]. The Annals of Statistics, 1999, 27: 1155-1163.

[3] CHIOU J M, MULLER H G. Nonparametric quasi-likelihood function[J]. The Annals of Statistics, 1999, 27: 36-34.

[4] CHANG Y. Strong consistency of maximum quasi-likelihood estimate in generalized models via a last time[J]. Statistics and Probability Letters, 1999, 45(3): 237-246.

[5] DING J L, CHEN X R. Asymptotic properties of the maximum likelihood estimateingeneralined linear models with stochastic regressors[J]. Acta Mathematica Sinica-English Series, 2012, 46: 833-846.

[6] YIN C M, ZHAO L C. Asymptotic normality and strong consistency of maximum quasi-likelihood in generalized linear models[J]. Science in China Series, A Mathematics, 2006, 49: 145-157.

[7] XUE D, XUE L G, CHENG W H. Empirical likelihood for generalized linear models with missing responses[J]. Journal of the Statistical Planning and Inference, 2011, 141: 2007-2020.

[8] 高启兵,吴耀华. 广义线性回归拟似然估计的渐近正态性[J]. 系统科学与数学, 2004, 25: 738-745.

[9] CHEN X, CHEN X R. Adaptive quasi-likelihood estimate in generalized linear models[J]. Science China Mathematics, 2005, 48(6): 829-846.

[10] GAO Q B, LIN J G, ZHUN C H, et al. Asymptotic properties of maximum quasi-likelihood estimators in generalized linear models with adaptive designs[J]. Statistics, 2012, 46: 833-846.

[11] OWEN A B. Empirical likelihood ratio confidence intervals for a single function[J]. Biometrika, 1988, 75: 237-325.

[12] OWEN A B. Empirical likelihood for linear models[J]. The Annals of Statistics, 1991, 19:1725-1747.

[13] QIN J, LAWLESS J. Empirical likelihood and general estimating equations[J]. The Annals of Statistics,

1994, 22(1):300-325.

[14] CHEN S X, QIN Y S. Empirical likelihood confidence intervals for local linear smoothers[J]. Biometrika, 2000, 87(4):946-953.

[15] WANG Q H , RAO J N K. Empirical likelihood-based inference under imputation for missing response data [J]. The Annals of Statistics, 2002, 30(3):896-924.

[16] ZHU L X, XUE L G. Empirical likelihood confidence regions in a partially linear single-index model[J]. Journal of the Royal Statistical Society, Series B, 2006, 68(3):549-570.

[17] XUE L G, ZHU L X. Empirical likelihood for a varying coefficient model with longitudinal data[J]. Journal of the American Statistical Association, 2007, 102 (478):642-654.

[18] QIN J, ZHANG B. Empirical-likelihood-based inference in missing response problems and its application in observational studies[J]. Journal of the Royal Statistical Society, Series B, 2007, 69(1):101-122.

[19] WANG D, CHEN S X. Empirical likelihood for estimating equations with missing values[J]. The Annals of Statistics, 2009, 37(1):490-517.

[20] YAN L, CHEN X. Empirical likelihood for partly linear models with errors in all variables[J]. Journal of Multivariate Analysis, 2014, 130: 275-288.

[21] OWEN A B. Empirical likelihood[M]. New York: Chapman & Hall, 2001.

[22] XUE L G, ZHU L X. Empirical likelihood in nonparametric and semiparametric models[M]. Beijing: Science Press, 2010.

[23] CHEN J B, VARIYATH A M, ABRAHAM B. Adjusted empirical likelihood and its properties[J]. Journal of Computational and Graphical Statistics, 1972, 17 (2): 426-443.

[24] CHEN S, PENG L, QIN Y. Effects of data dimension on empirical likelihood [J]. Biometrika, 2009, 96:711-722.

[25] HJORT N L, MCKEAGUE I W, VAN KEILEGOM I. Extending the scope of empirical likelihood[J]. The Annals of Statistics, 1999, 37: 1079-1111.

[26] PENG L, SCHICK A. Empirical likelihood approach to goodness of fit testing[J]. Bernoulli, 2013, 19(3): 954-981.

[27] PENG L, QI Y C, WANG R D. Empirical likelihood test for high dimensional linear models[J]. Statistics and Probability Letters, 2014, 86: 85-90.

[28] ZANG Y G, ZHANG Q Z, ZHANG S G, et al. Empirical likelihood test for high dimensional generalized linear models[M]// AHMED S. Big and complex data analysis. contributions to statistics. New York: Springer, 2017.

[29] WASSERMAN L, ROEDER K. High dimensional variable selection[J]. The Annals of Statistics, 2009, 37: 2178-2201.

[30] MEINSHAUSEN N, MEIER L, BHLMANN P. *P*-values for high-dimensional regression[J]. Journal of the American Statistical Association, 2009, 104: 1671-1681.

〔责任编辑　宋轶文〕