总体协方差矩阵不存在情形下的经验似然

刘鹏飞¹, 王绍臣^{2*}, 周望³

1. 江苏师范大学数学与统计学院、数学研究院、江苏省大数据科学与工程重点实验室, 徐州 221116;

2. 华南理工大学数学学院, 广州 510641;

3. Faculty of Science, Department of Statistics & Applied Probability, National University of Singapore, Singapore 117546, Singapore

E-mail: liupengfei@jsnu.edu.cn, mascwang@scut.edu.cn, stazw@nus.edu.sg

收稿日期: 2020-01-29; 接受日期: 2020-05-08; 网络出版日期: 2022-02-21; * 通信作者

国家自然科学基金 (批准号: 52034007, 11671178, 11501261 和 1801181)、江苏高校优势学科项目、广东省自然科学基金 (批准号: 2018A030310358) 和新加坡国立大学 (批准号: R-155-000-211-114) 资助项目

摘要 本文证明当总体分布处于多元正态分布的广义吸引域时, 总体均值的对数经验似然比统计量渐近服从卡方分布.

关键词 置信域 经验似然 正态分布的广义吸引域

MSC (2020) 主题分类 60F05

1 引言

设 $\mathbf{X}, \mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n \in \mathbb{R}^p$ 是来自总体分布函数为 F 的独立观测, 则 F 的非参数似然函数为

$$L(F) = \prod_{i=1}^n F\{\mathbf{x}_i\},$$

其中, $F\{\mathbf{x}_i\}$ 是总体分布函数为 F 时事件 $\{\mathbf{X}_i = \mathbf{x}_i\}$ 的概率 (若 F 是绝对连续的, 则它是密度函数), $\mathbf{x}_i$ 是 $\mathbf{X}_i$ 的观测值, $F \in \mathcal{F}$ 是 $\mathbb{R}^p$ 上的某个分布函数, $\mathcal{F}$ 是 $\mathbb{R}^p$ 上的分布函数全体. 容易看出, 经验分布函数 F_n , 即 $F_n(\mathbf{x}) = n^{-1} \sum_{i=1}^n I(\mathbf{X}_i \leq \mathbf{x})$ 关于所有可能的 F 使得 $L(F)$ 达到极大. 类似于参数似然比, 我们可以定义非参数似然比 $R(F) = L(F)/L(F_n)$. 由此可得轮廓似然比如下:

$$\mathcal{R}(\boldsymbol{\theta}) = \sup\{R(F) \mid T(F) = \boldsymbol{\theta}, F \in \mathcal{F}\},$$

其中 $T(\cdot)$ 是我们所关心的统计泛函. 在零假设 $H_0: \boldsymbol{\theta} = \boldsymbol{\theta}_0$ 下, $\mathcal{R}(\boldsymbol{\theta}_0)$ 应当接近于 1. 基于这个直观的考虑, 我们提出使用

$$\{\boldsymbol{\theta} \mid \mathcal{R}(\boldsymbol{\theta}) \geq r_0\} \quad (1.1)$$

英文引用格式: Liu P F, Wang S C, Zhou W. Empirical likelihood without population covariance matrix (in Chinese). Sci Sin Math, 2022, 52: 1089–1094, doi: 10.1360/SCM-2020-0073

作为 θ_0 的置信水平为 $1 - \alpha$ 的置信域, 其中 r_0 是由覆盖概率 $1 - \alpha$ 所决定的临界参数.

对于总体均值 $T(F) = E[\mathbf{X}] =: \boldsymbol{\mu}$ 的情形, 其中 $\mathcal{R}(\boldsymbol{\mu})$ 可以表示为

$$\mathcal{R}(\boldsymbol{\mu}) = \sup \left\{ \prod_{i=1}^n n\omega_i \mid \sum_{i=1}^n \omega_i \mathbf{X}_i = \boldsymbol{\mu}, \omega_i \geq 0, \sum_{i=1}^n \omega_i = 1 \right\}, \quad (1.2)$$

Owen^[1,2] 给出了 (1.1) 中 r_0 的确切值. 换言之, Owen 证明了在零假设 $H_0: \boldsymbol{\mu} = \boldsymbol{\mu}_0$ 下且在 $\text{Var}(\mathbf{X}) = \boldsymbol{\Sigma}$ 是满秩的情形下, 当 $n \rightarrow \infty$ 时,

$$-2\log(\mathcal{R}(\boldsymbol{\mu}_0)) \rightarrow_d \chi_p^2, \quad (1.3)$$

这是非参数形式的 Wilks 定理. 统计量 $\log(\mathcal{R}(\boldsymbol{\mu}_0))$ 被称为 Owen 的对数经验似然 (empirical likelihood, EL) 比统计量. 与其他非参数方法相比, EL 方法具有简洁、变换不变性和 Bartlett 修正等优点, 自 1990 年以来, 该方法广泛应用于很多领域.

注意到在 (1.3) 中要求满足 $\text{Var}(\mathbf{X}) = \boldsymbol{\Sigma}$ 这一条件. 一个自然的问题是, 在总体协方差矩阵不存在的情形下, Wilks 定理是否依然成立. 直观而言, 由于对数经验似然比统计量渐近等价于 Hotelling 的 T^2 统计量, 那么上面的问题归结为考虑 Hotelling 的 T^2 统计量在没有二阶矩条件下的极限分布问题. 另一方面, 从 Giné 等^[3] 的文章可以看出, 当且仅当 $\mathbf{X}$ 的分布处于正态吸引域 (domain of attraction of normal law, DAN) 时, 学生化的 t 统计量渐近服从标准正态分布. 这一结果的充分性由 Sepanski^[4] 推广到多元的情形 (多元情形下的必要性目前还未被解决). 换言之, 当 $\mathbf{X}$ 落在多元正态广义吸引域 (generalized domain of attraction of a multivariate normal law, GDAN) 时, Hotelling 的 T^2 统计量渐近地服从卡方分布. 因此, 若 $\mathbf{X}$ 落入 GDAN 时, 我们猜测 Wilks 定理依然成立.

本文将证明这一猜测. 事实证明, 一维的情形非常简单 (参见文献 [5]). 然而高维的情形并不容易证明. 根据定义, $\mathbf{X}$ 在 GDAN 中当且仅当对任意单位向量 $\boldsymbol{\eta}$, $\langle \mathbf{X}, \boldsymbol{\eta} \rangle$ 在 DAN 中, 其中 $\langle \cdot, \cdot \rangle$ 表示通常的 Euclid 内积. 此时需要考虑无穷维单位向量 $\boldsymbol{\eta}$ 所起的作用.

在介绍主要结果之前, 先给出 GDAN 的正式定义, 详情可参见文献 [6, 7].

定义 1.1 假定存在一个可逆 $p \times p$ 矩阵 $\mathbf{T}_n$ 和一个向量 $\mathbf{v}_n$ 的序列, 使得当 $n \rightarrow \infty$ 时, $\mathbf{T}_n(\sum_{i=1}^n \mathbf{X}_i - \mathbf{v}_n) \rightarrow_d N(\mathbf{0}, \mathbf{I})$, 其中, $\mathbf{0}$ 是 p 维零向量, $\mathbf{I}$ 是 $p \times p$ 单位矩阵. 如果 p 维向量 $\mathbf{X}$ 的支撑集不在任何 $p-1$ 维空间中, 则称其具有全律.

定理 1.1 设 $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$ 是 GDAN 中的独立同分布 p 维随机向量且具有全律. 再记 $\boldsymbol{\mu}_0 = E[\mathbf{X}_1]$, 则当 $n \rightarrow \infty$ 时,

$$-2\log(\mathcal{R}(\boldsymbol{\mu}_0)) \rightarrow_d \chi_p^2.$$

注 1.1 根据定理 1.1, 可以构造一个 $\boldsymbol{\mu}_0$ 的近似置信水平为 $1 - \alpha$ 的置信域, 即

$$R_c = \{\boldsymbol{\mu} \mid -2\log(\mathcal{R}(\boldsymbol{\mu})) \leq c\},$$

其中, c 满足 $P(\chi_p^2 \leq c) = 1 - \alpha$, $p \geq 1$.

注 1.2 当总体分布是重尾且尾部指数介于 1 和 2 之间 (方差无限) 时, 即总体处于稳定 (非正态) 分布的吸引域中时, Peng^[8] 基于经验似然方法考虑了均值的置信区间, 但并没有考虑方差无限的正态分布的吸引域. 此外, Peng^[8] 将极值理论应用于处理总体分布的两个尾部, 并对中间部分进行非参数化处理. 换言之, 文献 [8] 中的经验对数似然比统计量不是我们所考虑的 $\mathcal{R}(\boldsymbol{\mu})$.

2 证明

首先需要一些关于正态广义吸引域的基本事实. 对于 $\boldsymbol{\eta} \in S^{p-1} := \{\boldsymbol{\eta} \mid \boldsymbol{\eta} \in \mathbb{R}^p, \|\boldsymbol{\eta}\| = 1\}$, 令 $\mathbf{X}$ 在 GDAN 中具有全律, 设 $a_n(\boldsymbol{\eta})$ 为方程 $a_n^2(\boldsymbol{\eta}) = nE\langle \mathbf{X}, \boldsymbol{\eta} \rangle^2 I(|\langle \mathbf{X}_1, \boldsymbol{\eta} \rangle| < a_n^2(\boldsymbol{\eta}))$ 的唯一解. Hahn 和 Klass^[9] 证明了 $a_n(\boldsymbol{\eta})$ 的存在性和唯一性.

引理 2.1 设 $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$ 是 GDAN 中的独立同分布随机 p 维向量且具有全律. 则当 $n \rightarrow \infty$ 时, 以下结论成立:

- (i) 关于 $\boldsymbol{\eta} \in S^{p-1}$ 一致地有 $a_n^2(\boldsymbol{\eta}) \rightarrow \infty$;
- (ii) $\mathbf{T}_n = \mathbf{S}_n^{-\frac{1}{2}} \sum_{i=1}^n (\mathbf{X}_i - \boldsymbol{\mu}_0) \rightarrow^d N(\mathbf{0}, \mathbf{I})$, 其中 $\mathbf{S}_n = \sum_{i=1}^n (\mathbf{X}_i - \boldsymbol{\mu}_0)(\mathbf{X}_i - \boldsymbol{\mu}_0)^T$;
- (iii) $\sup_{\|\boldsymbol{\eta}\|=1} \left| \frac{\sum_{i=1}^n |\langle \mathbf{X}_i - \boldsymbol{\mu}_0, \boldsymbol{\eta} \rangle|^2}{a_n^2(\boldsymbol{\eta})} - 1 \right| \rightarrow_p 0$;
- (iv) $\sup_{\|\boldsymbol{\eta}\|=1} \frac{M_n(\boldsymbol{\eta})}{a_n(\boldsymbol{\eta})} \rightarrow_p 0$, 其中 $M_n(\boldsymbol{\eta}) = \max_{i \leq n} |\langle \mathbf{X}_i, \boldsymbol{\eta} \rangle|$.

证明 论断 (i) 来自于这样一个事实, 即序列 $\{nb_n^{-2}(\boldsymbol{\eta})\}$ 对 $\boldsymbol{\eta} \in S^{p-1}$ 来说是一致有界的. 其推导可以在文献 [6, 引理 4] 的证明中找到. 论断 (ii) 可以在文献 [4, 定理 1] 的证明中找到. 论断 (iii) 和 (iv) 可以由文献 [7, 定理 2] 的证明得出, 也就是文献 [10, 定理 3.8]. 事实上可以看到, 文献 [7] 的 (3.6) 和 (3.6) 之前的那一行就是我们的论断 (iv) 和 (iii). 我们也可参见文献 [10, (3.3.32)、(3.3.33) 和 (3.2.7)]. $\square$

下一个引理保证总体均值最终位于所有观测值的凸包.

引理 2.2 设 $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$ 是 GDAN 中的独立同分布随机 p 维向量且具有全律. 则 $\boldsymbol{\mu}_0$ 包含在 $\{\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n\}$ 凸壳内部的概率趋向于 1.

证明 根据文献 [2, 引理 2] 的证明, 可以证明存在一个正的常数 c_0 使得

$$c_1 := \inf_{\|\boldsymbol{\eta}\|=1} P(\langle \mathbf{X}_1 - \boldsymbol{\mu}_0, \boldsymbol{\eta} \rangle > c_0) > 0. \quad (2.1)$$

不难将 Glivenko-Cantelli 定理推广到在半空间上的一致收敛, 可以得到, 当 $n \rightarrow \infty$ 时,

$$\sup_{\|\boldsymbol{\eta}\|=1} \left| P(\langle \mathbf{X}_1 - \boldsymbol{\mu}_0, \boldsymbol{\eta} \rangle > c_0) - n^{-1} \sum_{i=1}^n I(\langle \mathbf{X}_i - \boldsymbol{\mu}_0, \boldsymbol{\eta} \rangle > c_0) \right| \rightarrow 0$$

几乎处处成立. 再与 (2.1) 相结合表明

$$\inf_{\|\boldsymbol{\eta}\|=1} n^{-1} \sum_{i=1}^n I(\langle \mathbf{X}_i - \boldsymbol{\mu}_0, \boldsymbol{\eta} \rangle > c_0) > \frac{c_1}{2}$$

且随着 $n \rightarrow \infty$ 上式成立的概率趋于 1. 这意味着 $\boldsymbol{\mu}_0$ 包含在 $\{\mathbf{X}_1, \dots, \mathbf{X}_n\}$ 凸包内部的概率趋于 1, 这是因为, 若上式不成立的话, 则必然存在 $\boldsymbol{\eta} \in S^{p-1}$, 使得 $\langle \mathbf{X}_i - \boldsymbol{\mu}_0, \boldsymbol{\eta} \rangle \leq 0, i = 1, \dots, n$. $\square$

定理 1.1 的证明 由引理 2.2, 可以假设 $\boldsymbol{\mu}_0$ 包含在由 $\{\mathbf{X}_1, \dots, \mathbf{X}_n\}$ 构成的凸包内部. 所以

$$\Omega := \left\{ (\omega_1, \dots, \omega_n) \mid \sum_{i=1}^n \omega_i \mathbf{X}_i = \boldsymbol{\mu}_0, \omega_i \geq 0, \sum_{i=1}^n \omega_i = 1 \right\}$$

是一个非空的紧集. 因此, $\mathcal{R}(\boldsymbol{\mu}_0)$ 必在 Ω 中的某一点取得且这个极值点是唯一的. 否则, 达到极值的两个不同点的凸组合将会导致此点处的似然比大于 $\mathcal{R}(\boldsymbol{\mu}_0)$.

可以用 Lagrange 乘子法继续寻找这个极值点. 记

$$G = \sum_{i=1}^n \log(n\omega_i) - n\boldsymbol{\lambda}^T \sum_{i=1}^n \omega_i (\mathbf{X}_i - \boldsymbol{\mu}_0) + \gamma \left(\sum_{i=1}^n \omega_i - 1 \right),$$

其中, $\boldsymbol{\lambda} \in \mathbb{R}^p$, $\gamma \in \mathbb{R}$ 为 Lagrange 乘子, $\mathbf{T}$ 表示转置. 令 G 对 ω_i 的偏导数为 0, 可以得到

$$\gamma = -n, \quad n\omega_i = \frac{1}{1 + \boldsymbol{\lambda}^T(\mathbf{X}_i - \boldsymbol{\mu}_0)}. \quad (2.2)$$

因此有

$$\frac{1}{n} \sum_{i=1}^n \frac{\mathbf{X}_i - \boldsymbol{\mu}_0}{1 + \boldsymbol{\lambda}^T(\mathbf{X}_i - \boldsymbol{\mu}_0)} = \mathbf{0}.$$

(如果 $\boldsymbol{\mu}_0$ 包含在 $\{\mathbf{X}_1, \dots, \mathbf{X}_n\}$ 的凸包内部, 则上式方程的解总是存在的, 可参见文献 [11, 第 3.14 小节].)

令

$$g(\boldsymbol{\lambda}) \equiv \sum_{i=1}^n \frac{\mathbf{X}_i - \boldsymbol{\mu}_0}{1 + \boldsymbol{\lambda}^T(\mathbf{X}_i - \boldsymbol{\mu}_0)}$$

以及

$$\boldsymbol{\eta} = \frac{\boldsymbol{\lambda}}{\|\boldsymbol{\lambda}\|}, \quad Y_i = \boldsymbol{\lambda}^T(\mathbf{X}_i - \boldsymbol{\mu}_0), \quad Z_n^* = \max_{1 \leq i \leq n} |\boldsymbol{\eta}^T(\mathbf{X}_i - \boldsymbol{\mu}_0)|.$$

将 $\frac{1}{1+Y_i} = 1 - \frac{Y_i}{1+Y_i}$ 代入 $g(\boldsymbol{\lambda}) = 0$, 得到

$$\begin{aligned} 0 &= \|g(\boldsymbol{\lambda})\| \geq |\boldsymbol{\eta}^T g(\boldsymbol{\lambda})| \\ &= \left| \boldsymbol{\eta}^T \left(\sum_{i=1}^n (\mathbf{X}_i - \boldsymbol{\mu}_0) - \|\boldsymbol{\lambda}\| \sum_{i=1}^n \frac{(\mathbf{X}_i - \boldsymbol{\mu}_0)(\mathbf{X}_i - \boldsymbol{\mu}_0)^T}{1 + Y_i} \boldsymbol{\eta} \right) \right| \\ &\geq \|\boldsymbol{\lambda}\| \boldsymbol{\eta}^T \sum_{i=1}^n \frac{(\mathbf{X}_i - \boldsymbol{\mu}_0)(\mathbf{X}_i - \boldsymbol{\mu}_0)^T}{1 + Y_i} \boldsymbol{\eta} - \left| \sum_{i=1}^n \boldsymbol{\eta}^T(\mathbf{X}_i - \boldsymbol{\mu}_0) \right|. \end{aligned} \quad (2.3)$$

由 $n\omega_i = (1 + y_i)^{-1} > 0$, 有

$$(2.3) \text{ 的右边} \geq \|\boldsymbol{\lambda}\| \frac{\boldsymbol{\eta}^T S_n \boldsymbol{\eta}}{1 + \|\boldsymbol{\lambda}\| Z_n^*} - \left| \sum_{i=1}^n \boldsymbol{\eta}^T(\mathbf{X}_i - \boldsymbol{\mu}_0) \right|.$$

因此由引理 2.1(iii) 和 Cauchy-Schwarz 不等式可知,

$$\frac{\|\boldsymbol{\lambda}\|}{1 + \|\boldsymbol{\lambda}\| Z_n^*} \leq \frac{|\sum_{i=1}^n \boldsymbol{\eta}^T(\mathbf{X}_i - \boldsymbol{\mu}_0)|}{\boldsymbol{\eta}^T S_n \boldsymbol{\eta}} = O_p(a_n^{-1}(\boldsymbol{\eta})). \quad (2.4)$$

另外, 再由引理 2.1(iv), 有

$$Z_n^* = o_p(a_n(\boldsymbol{\eta})). \quad (2.5)$$

因此有 $\|\boldsymbol{\lambda}\|(1 + o_p(1)) = O_p(a_n^{-1}(\boldsymbol{\eta}))$ 成立, 故

$$\|\boldsymbol{\lambda}\| = O_p(a_n^{-1}(\boldsymbol{\eta})). \quad (2.6)$$

利用 $\frac{1}{1+x} = 1 - x + \frac{x^2}{1+x}$, 将 $g(\boldsymbol{\lambda})$ 展开可以得到

$$\begin{aligned} 0 &= \sum_{i=1}^n (\mathbf{X}_i - \boldsymbol{\mu}_0) \left(1 - Y_i + \frac{Y_i^2}{1 + Y_i} \right) \\ &= \sum_{i=1}^n (\mathbf{X}_i - \boldsymbol{\mu}_0) - S_n \boldsymbol{\lambda} + \sum_{i=1}^n (\mathbf{X}_i - \boldsymbol{\mu}_0)(\mathbf{X}_i - \boldsymbol{\mu}_0)^T \frac{Y_i}{1 + Y_i} \boldsymbol{\lambda}. \end{aligned} \quad (2.7)$$

由 (2.5) 和 (2.6) 可知, 当 $n \rightarrow \infty$ 时, $\max_{1 \leq i \leq n} |Y_i| = o_p(1)$. 这意味着

$$\lim_{n \rightarrow \infty} P\left(-\epsilon S_n \leq \sum_{i=1}^n (\mathbf{X}_i - \boldsymbol{\mu}_0)(\mathbf{X}_i - \boldsymbol{\mu}_0)^T \frac{Y_i}{1 + Y_i} \leq \epsilon S_n\right) = 1$$

对于任意给定的 $\epsilon > 0$ 成立. 因此 $-\epsilon S_n \leq \sum_{i=1}^n (\mathbf{X}_i - \boldsymbol{\mu}_0)(\mathbf{X}_i - \boldsymbol{\mu}_0)^T \frac{Y_i}{1 + Y_i} \leq \epsilon S_n$ 意味着矩阵 $\epsilon S_n + \sum_{i=1}^n (\mathbf{X}_i - \boldsymbol{\mu}_0)(\mathbf{X}_i - \boldsymbol{\mu}_0)^T \frac{Y_i}{1 + Y_i}$ 和 $\epsilon S_n - \sum_{i=1}^n (\mathbf{X}_i - \boldsymbol{\mu}_0)(\mathbf{X}_i - \boldsymbol{\mu}_0)^T \frac{Y_i}{1 + Y_i}$ 是正定的. 因此 $\epsilon \mathbf{I} + S_n^{-1/2} \sum_{i=1}^n (\mathbf{X}_i - \boldsymbol{\mu}_0)(\mathbf{X}_i - \boldsymbol{\mu}_0)^T \frac{Y_i}{1 + Y_i} S_n^{-1/2}$ 和 $\epsilon \mathbf{I} - S_n^{-1/2} \sum_{i=1}^n (\mathbf{X}_i - \boldsymbol{\mu}_0)(\mathbf{X}_i - \boldsymbol{\mu}_0)^T \frac{Y_i}{1 + Y_i} S_n^{-1/2}$ 依概率正定. 由于 ϵ 是任意的, 所以矩阵 $S_n^{-1/2} \sum_{i=1}^n (\mathbf{X}_i - \boldsymbol{\mu}_0)(\mathbf{X}_i - \boldsymbol{\mu}_0)^T \frac{Y_i}{1 + Y_i} S_n^{-1/2}$ 的 p 个特征值 $\rho_1, \dots, \rho_p$ 在 $n \rightarrow \infty$ 时依概率一致趋于 0. 现在假设 $\mathbf{u}_i$ ($i = 1, \dots, p$) 是 $S_n^{-1/2} \sum_{i=1}^n (\mathbf{X}_i - \boldsymbol{\mu}_0)(\mathbf{X}_i - \boldsymbol{\mu}_0)^T \frac{Y_i}{1 + Y_i} S_n^{-1/2}$ 对应 ρ_i 的单位特征向量. 这意味着

$$S_n^{-1} \sum_{i=1}^n (\mathbf{X}_i - \boldsymbol{\mu}_0)(\mathbf{X}_i - \boldsymbol{\mu}_0)^T \frac{Y_i}{1 + Y_i} S_n^{-1/2} \mathbf{u}_i = \rho_i S_n^{-1/2} \mathbf{u}_i.$$

因此 ρ_i ($i = 1, \dots, p$) 也是 $S_n^{-1} \sum_{i=1}^n (\mathbf{X}_i - \boldsymbol{\mu}_0)(\mathbf{X}_i - \boldsymbol{\mu}_0)^T \frac{Y_i}{1 + Y_i}$ 的特征值, 其相应的特征向量是 $S_n^{-1/2} \mathbf{u}_i$. 记 $\mathbf{v}_i = S_n^{-1/2} \mathbf{u}_i / \|S_n^{-1/2} \mathbf{u}_i\|$ ($i = 1, \dots, p$). 我们知道 $\{\mathbf{u}_1, \dots, \mathbf{u}_p\}$ 张成 p 维空间 $\mathbb{R}^p$, 根据文献 [6, 引理 5], 由于 S_n 依概率 1 可逆, 故 $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ 依概率 1 张成 $\mathbb{R}^p$. 所以对于任意单位向量 $\mathbf{v}$, 在忽略依概率为 0 的事件后, 可以假设 $\mathbf{v} = \sum_{i=1}^p h_i \mathbf{v}_i$, 其中 $|h_i| \leq 1$ ($i = 1, \dots, p$). 由此可以得到

$$S_n^{-1} \sum_{i=1}^n (\mathbf{X}_i - \boldsymbol{\mu}_0)(\mathbf{X}_i - \boldsymbol{\mu}_0)^T \frac{Y_i}{1 + Y_i} \mathbf{v} = \sum_{i=1}^p h_i \rho_i \mathbf{v}_i.$$

由于当 $n \rightarrow \infty$ 时, 所有的 ρ_i 依概率一致趋于 0, 故

$$\left\| S_n^{-1} \sum_{i=1}^n (\mathbf{X}_i - \boldsymbol{\mu}_0)(\mathbf{X}_i - \boldsymbol{\mu}_0)^T \frac{Y_i}{1 + Y_i} \mathbf{v} \right\| \rightarrow_P 0$$

对 $\mathbf{v} \in S^{p-1}$ 一致成立. 将上述结果与 (2.6) 和 (2.7) 结合起来, 可以得到

$$\boldsymbol{\lambda} = S_n^{-1} \sum_{i=1}^n (\mathbf{X}_i - \boldsymbol{\mu}_0) + \beta(\boldsymbol{\eta}), \quad (2.8)$$

其中 $\beta(\boldsymbol{\eta}) = o_p(a_n^{-1}(\boldsymbol{\eta}))$.

注意到 $\max_{1 \leq i \leq n} |Y_i| = o_p(1)$. 将 $-2 \log \mathcal{R}(\boldsymbol{\mu}_0)$ 展开并将 $\boldsymbol{\lambda}$ 代入, 可以得到

$$\begin{aligned} -2 \log \mathcal{R}(\boldsymbol{\mu}_0) &= 2 \sum_{i=1}^n \log(1 + Y_i) = 2 \sum_{i=1}^n Y_i - \sum_{i=1}^n Y_i^2 + 2 \sum_{i=1}^n \xi_i \\ &= \sum_{i=1}^n (\mathbf{X}_i - \boldsymbol{\mu}_0)^T S_n^{-1} (\mathbf{X}_i - \boldsymbol{\mu}_0) - \beta(\boldsymbol{\eta})^T S_n^{-1} \beta(\boldsymbol{\eta}) - 2\beta(\boldsymbol{\eta})^T S_n^{-1} \boldsymbol{\lambda} + 2 \sum_{i=1}^n \xi_i. \end{aligned}$$

由 (2.6) 和引理 2.1(iii) 可知, 当 $n \rightarrow \infty$ 时,

$$\sum_{i=1}^n |\xi_i| = \sum_{i=1}^n \left| \frac{Y_i^3}{3} - \frac{Y_i^4}{4} + \frac{Y_i^5}{5} + \dots \right| \leq \max_{1 \leq i \leq n} |Y_i| \|\boldsymbol{\lambda}\|^2 \sum_{i=1}^n \|\boldsymbol{\eta}^T (\mathbf{X}_i - \boldsymbol{\mu}_0)\|^2 = o_p(1).$$

另一方面, 引理 2.1(i) 意味着, 当 $n \rightarrow \infty$ 时, $\|S_n^{-1}\| \rightarrow 0$, 因此当 $n \rightarrow \infty$ 时,

$$\beta(\boldsymbol{\eta})^T S_n^{-1} \beta(\boldsymbol{\eta}) = o_p(1), \quad \beta(\boldsymbol{\eta})^T S_n^{-1} \boldsymbol{\lambda} = o_p(1).$$

最后由引理 2.1(ii) 可知, 当 $n \rightarrow \infty$ 时,

$$\sum_{i=1}^n (\mathbf{X}_i - \boldsymbol{\mu}_0)^T S_n^{-1} (\mathbf{X}_i - \boldsymbol{\mu}_0) \rightarrow_d \chi_{(p)}^2,$$

因此得到, 当 $n \rightarrow \infty$ 时, $-2 \log \mathcal{R}(\boldsymbol{\mu}_0) \rightarrow_d \chi_{(p)}^2$. 证毕. $\square$

致谢 非常感谢审稿人给予我们的意见.

参考文献

- 1 Owen A B. Empirical likelihood ratio confidence intervals for a single functional. *Biometrika*, 1988, 75: 237–249
- 2 Owen A B. Empirical likelihood ratio confidence regions. *Ann Statist*, 1990, 18: 90–120
- 3 Giné E, Götze F, Mason D M. When is the student t -statistic asymptotically standard normal? *Ann Probab*, 1997, 25: 1514–1531
- 4 Sepanski S J. Asymptotics for multivariate t -statistic for random vectors in the generalized domain of attraction of the multivariate normal law. *Statist Probab Lett*, 1996, 30: 179–188
- 5 Cheng C H, Liu Y M, Liu Z. Empirical likelihood ratio under infinite second moment. *Comm Statist Theory Methods*, 2017, 46: 6909–6915
- 6 Sepanski S J. Asymptotics for multivariate t -statistic and Hotelling's T^2 -statistic under infinite second moments via bootstrapping. *J Multivariate Anal*, 1994, 49: 41–54
- 7 Sepanski S J. Necessary and sufficient conditions for the multivariate bootstrap of the mean. *Statist Probab Lett*, 1994, 19: 205–216
- 8 Peng L. Empirical-likelihood-based confidence interval for the mean with a heavy-tailed distribution. *Ann Statist*, 2004, 32: 1192–1214
- 9 Hahn M G, Klass M J. Matrix normalization of sums of random vectors in the domain of attraction of the multivariate normal. *Ann Probab*, 1980, 8: 262–280
- 10 Sepanski S J. Normalization methods in the bootstrap central limit theorem. PhD Thesis. Vita: Texas A & M University, 1991
- 11 Owen A B. *Empirical Likelihood*. London: Chapman & Hall/CRC, 2001

Empirical likelihood without population covariance matrix

Pengfei Liu, Shaochen Wang & Wang Zhou

Abstract In this paper, we show that the log empirical likelihood ratio statistic for the population mean converges in distribution to a chi-squared distribution when the population is in the generalized domain of attraction of multivariate normal law.

Keywords confidence region, empirical likelihood, generalized domain of attraction of normal law

MSC(2020) 60F05

doi: 10.1360/SCM-2020-0073