第 40 卷 第 5 期 广西师范大学学报(自然科学版) Vol. 40 No. 5
2022 年 9 月 Journal of Guangxi Normal University (Natural Science Edition) Sept. 2022

DOI: 10.16088/j.issn.1001-6600.2022021101 http://xuebao.gxnu.edu.cn

张军舰. 非参数似然方法及其应用研究进展[J]. 广西师范大学学报(自然科学版), 2022, 40(5): 150-159. ZHANG J J. Review on nonparametric likelihood and their applications[J]. Journal of Guangxi Normal University (Natural Science Edition), 2022, 40(5): 150-159.

非参数似然方法及其应用研究进展

张军舰

(广西师范大学 数学与统计学院, 广西 桂林 541006)

摘 要: 非参数似然方法是在参数似然方法基础上发展的一种非参数方法, 具有与传统似然方法类似的许多优良性质, 特别是其中的经验似然方法, 得到许多学者的重视和研究, 是目前统计学的一个重要研究方向。本文在介绍非参数似然方法相关思想基础上, 结合团队研究工作, 主要从估计、检验和复杂数据应用等方面分别对其研究进展进行较系统地综述, 分析相关研究思路和研究内容, 给出一些主要结论和研究的侧重点。

关键词: 非参数似然; 经验似然; 估计; 拟合优度检验; 复杂数据

中图分类号: O212.7 **文献标志码**: A **文章编号**: 1001-6600(2022)05-0150-10

似然方法是统计中非常重要的方法之一, 自 Fisher 提出后引起很多学者的重视, 得到非常丰富的研究结果, 统计中大部分论文或多或少要用到似然方法。从最开始的参数似然方法, 发展到如今的非参数似然和半参数似然方法, 其理论和应用研究还在不断深入, 详情可参见文献[1-3]。参数似然方法是要求总体的分布类型和形式已知, 仅依赖若干个未知参数, 接着通过似然函数对相应参数进行统计推断, 进而研究相应的统计性质; 非参数似然则不需要总体的分布类型或形式已知, 它仅利用样本点的观测信息对分布作出估计, 进而借助参数与分布函数的关系对参数进行估计; 半参数则是介于参数与非参数之间, 其部分信息需要知道分布结构, 部分信息对分布类型或形式不做要求, 在此基础上, 对分布或相应参数做出统计推断。参数似然方法理论相对来说比较成熟, 其应用非常广泛。但在实际中, 对所研究问题的背景所知甚少, 仅仅依靠经验或已有研究基础很难对分布类型或形式做出要求, 这就必须考虑非参数或半参数似然方法, 尽可能有效地利用部分已有信息, 从而进行统计推断。本文正是在此基础上, 结合团队多年研究内容, 对非参数似然方法做一综述。

论文安排如下: 第 1 章简述非参数似然的基本概念和思想, 进而探讨基于经验似然推广的非参数似然方法, 侧重点主要是估计方面; 第 2 章重点考虑非参数似然拟合优度检验问题; 第 3 章考虑更复杂数据的非参数似然方法; 最后是总结和展望。

1 非参数似然方法

1.1 参数与非参数似然

为便于理解, 先从参数似然开始。假定所研究的总体为 X, 其对应的分布函数为 F, 概率函数为 $f(x,\theta)$ (如果 X 是离散型随机变量, $f(x,\theta)$ 就是 X 的概率分布列; 如果 X 是连续型随机变量, $f(x,\theta)$ 就是 X 的密度函数), 其中 $f(x,\theta)$ 的分布类型或形式已知, 参数 $\theta\in\Theta$ 未知。现从总体 X 中抽取一个独立同分布(i.i.d.)样本 $X_1,X_2,\cdots,X_n$, 希望通过样本获得参数 θ 的一个合理估计, 极大似然估计(MLE)即是其中之一。给定 $X_1,X_2,\cdots,X_n$, θ 的似然函数定义为

收稿日期: 2022-02-11 修回日期: 2022-03-13
基金项目: 国家自然科学基金(11861017)
通信作者: 张军舰(1973—), 男, 河南内乡人, 广西师范大学教授, 博导。E-mail: jjzhang@gxnu.edu.cn

$$L(\theta) = k\prod_{i=1}^{n} f(X_i,\theta),$$

其中 $k>0$ 为不依赖于参数 θ 的一个常数。MLE $\hat{\theta}_n$ 则是使得 $L(\theta)$ 达到最大值时所对应的 θ，即

$$L(\hat{\theta}_n) = \sup_{\theta\in\Theta} L(\theta)\text{，或者}\hat{\theta}_n = \arg\max_{\theta\in\Theta} L(\theta)。$$

当总体分布未知时，概率函数 $f(x,\theta)$ 的分布类型或形式也是未知的，MLE 自然没有办法求出，这时需要借助非参数似然函数，定义分布 F 的非参数似然函数如下

$$L(F) = k\prod_{i=1}^{n}(F(X_i) - F(X_i-)) = k\prod_{i=1}^{n} F\{X_i\} = k\prod_{i=1}^{n} p_i,$$

式中：$k>0$ 为不依赖于 F 的一个数；$p_i = F\{X_i\} = F(X_i) - F(X_i-)$ 为 F 在 $X=X_i$ 处的质量，满足 $p_i\geqslant 0$，$\sum_{i=1}^{n} p_i = 1$。容易证明，在上述约束下，使 $L(F)$ 达到最大的 F 是经验分布函数 $F_n(x) = \frac{1}{n}\sum_{i=1}^{n} I[X_i\leqslant x]$，也即 $p_i = 1/n, i=1,\cdots,n$。相应地，$L(F_n) = k\prod_{i=1}^{n} n^{-1}$。这样就可以构造非参数似然比函数（消去数 k 且使该值在 $[0,1]$ 内）

$$R(F) = \frac{L(F)}{L(F_n)} = \prod_{i=1}^{n} np_i。\tag{1}$$

如果参数 θ 为 F 的一个泛函 $\theta=\theta(F)$，则可以得到 θ 的非参数 MLE 为 $\theta(F_n)$。例如总体均值 $\mu = \int x\mathrm{d}F$ 的非参数 MLE 为 $\hat{\mu}_n = \int x\mathrm{d}F_n(x) = \frac{1}{n}\sum_{i=1}^{n} X_i = \bar{X}_n$，也就是常说的样本均值。

1.2 经验似然

从前述可以看出，非参数似然的基本思想就是把概率函数用样本观测值的质量来替换，也就是把 F 限制在支撑点 $X_1, X_2, \cdots, X_n$ 上，即 $F\ll F_n$，若用 $\mathscr{F}$ 表示某一分布族，则

$$\{F \mid F\ll F_n, F\in\mathscr{F}\} = \left\{(p_1,p_2,\cdots,p_n) \;\middle|\; p_i\geqslant 0, \sum_{i=1}^{n} p_i = 1\right\}。$$

现在的问题是上述考虑的分布族仅利用样本信息，忽略了参数与未知分布之间的关系信息 $\theta=\theta(F)$。如何利用 $\theta=\theta(F)$ 所带来的信息？直观想法就是把所考虑的分布限制到一个更小的分布族中，其中 $\theta=\theta(F)$ 用样本点上的质量表达出来，作为辅助信息使用，也就是考虑

$$\{F \mid \theta = T(F), F\ll F_n, F\in\mathscr{F}\} = \left\{(p_1,\cdots,p_n) \;\middle|\; p_i\geqslant 0, \sum_{i=1}^{n} p_i = 1, \theta = T(\boldsymbol{X}_1,\cdots,\boldsymbol{X}_n;p_1,\cdots,p_n)\right\}。$$

这便是 Owen 经验似然方法的思想。文献 [4-5] 提出关于均值参数 $\mu = \int x\mathrm{d}F$ 的 Profile 经验似然比函数

$$\mathrm{R}(\boldsymbol{\mu}) = \max\left\{\prod_{i=1}^{n} np_i \;\middle|\; p_i\geqslant 0, \sum_{i=1}^{n} p_i = 1, \sum_{i=1}^{n} p_i\boldsymbol{X}_i = \boldsymbol{\mu}\right\}。$$

设 $\boldsymbol{X}\in\mathbf{R}^p$，其协方差阵的秩为 $q>0$，文献 [5] 证明当 $n\to\infty$ 时，$-2\log\mathrm{R}(\boldsymbol{\mu})\to\chi_q^2$，从而可以得到均值参数 $\boldsymbol{\mu}$ 的渐近置信水平为 $1-\alpha$ 的置信区域

$$C_{\boldsymbol{\mu}} = \{\boldsymbol{\mu}\in\mathbf{R}^p \mid -2\log\mathrm{R}(\boldsymbol{\mu})\leqslant\chi_q^2(1-\alpha)\}。$$

其中 $\chi_q^2(1-\alpha)$ 表示自由度为 q 的 χ^2 分布的 $1-\alpha$ 分位数。具体计算时，

$$-2\log\mathrm{R}(\boldsymbol{\mu}) = 2\sum_{i=1}^{n}\log(1+\boldsymbol{\lambda}^{\mathrm{T}}(\boldsymbol{X}_i-\boldsymbol{\mu})),$$

其中 $\boldsymbol{\lambda}$ 满足

$$g(\boldsymbol{\lambda}) = \frac{1}{n}\sum_{i=1}^{n}\frac{\boldsymbol{X}_i-\boldsymbol{\mu}}{1+\boldsymbol{\lambda}^{\mathrm{T}}(\boldsymbol{X}_i-\boldsymbol{\mu})} = \boldsymbol{0}。$$

可以证明，在一定正则条件下，由经验似然方法得到的估计具有良好的统计性质：有类似于 bootstrap 的抽样特性，这一方法与经典非参数统计方法比较有很多突出优点，如构造的置信区间有域保持性、变换

不变性、置信域的形状由数据自行决定、有 Bartlett 纠偏性、无需构造枢轴量等,详情可参见文献[1,5]。文献[6]将其应用到广义估计方程模型中,给出了经验似然更一般的理论。

假定所考虑的是与 F 相关的 d 维参数 $\boldsymbol{\theta}=(\theta_1,\cdots,\theta_d)^{\mathrm{T}}$,$\boldsymbol{\theta}$ 与 F 的关联信息是通过 $r(\geqslant p)$ 个独立无偏的估计方程得到,也即

$$g(x,\boldsymbol{\theta})=(g_1(x,\boldsymbol{\theta}),g_2(x,\boldsymbol{\theta}),\cdots,g_r(x,\boldsymbol{\theta}))^{\mathrm{T}},r\geqslant d,$$

满足 $E_F\{g(\boldsymbol{X},\boldsymbol{\theta})\}=0$,则 Profile 经验似然比函数为

$$\mathrm{R}(\boldsymbol{\theta})=\max\left\{\prod_{i=1}^{n}np_i\,\middle|\,p_i\geqslant 0,\sum_{i=1}^{n}p_i=1,\sum_{i=1}^{n}p_ig(\boldsymbol{X}_i,\boldsymbol{\theta})=0\right\}。\tag{2}$$

若设 $q>0$ 为 $g(\boldsymbol{X},\boldsymbol{\theta})$ 的协方差阵所对应的秩,文献[6]证明:当 $n\to\infty$ 时,$-2\log\mathrm{R}(\boldsymbol{\theta})\to\chi_q^2$。类似于文献[5],可以得到参数 $\boldsymbol{\theta}$ 的渐近置信水平为 $1-\alpha$ 的置信区域

$$C_{\boldsymbol{\theta}}=\{\boldsymbol{\theta}\in\mathbf{R}^d\mid -2\log\mathrm{R}(\boldsymbol{\theta})\leqslant\chi_q^2(1-\alpha)\}。$$

具体计算时,

$$-2\log\mathrm{R}(\boldsymbol{\theta})=2\sum_{i=1}^{n}\log(1+\boldsymbol{\lambda}^{\mathrm{T}}g(\boldsymbol{X}_i-\boldsymbol{\theta})),$$

其中 $\boldsymbol{\lambda}$ 满足

$$g(\boldsymbol{\lambda})=\frac{1}{n}\sum_{i=1}^{n}\frac{g(\boldsymbol{X}_i,\boldsymbol{\theta})}{1+\boldsymbol{\lambda}^{\mathrm{T}}g(\boldsymbol{X}_i,\boldsymbol{\theta})}=\boldsymbol{0}。$$

当然也可由此分别得到参数 $\boldsymbol{\theta}$ 以及分布函数 F 相应的经验似然比估计

$$\tilde{\boldsymbol{\theta}}_n=\arg\max_{\boldsymbol{\theta}\in\Theta}\mathrm{R}(\boldsymbol{\theta}),\ \tilde{p}_i=\frac{1}{n}\frac{1}{1+\boldsymbol{\lambda}^{\mathrm{T}}(\boldsymbol{X}_i,\tilde{\boldsymbol{\theta}}_n)},\ \tilde{F}_n(x)=\sum_{i=1}^{n}\tilde{p}_iI[\boldsymbol{X}_i\leqslant x]。$$

式(2)给出的 Profile 经验似然比函数极大地拓宽了其应用范围,只要寻找到满足条件的估计方程 $g(x,\boldsymbol{\theta})$,就可以使用经验似然方法。Owen 的方法可以看做是 $g(x,\theta)=x-\theta$ 的特殊形式。经验似然提出后,引起很多学者的重视,形成了统计学的一个研究热点方向。如文献[7]将其应用到线性回归模型的统计推断中;文献[8]给出了密度函数的经验似然置信区间;文献[9]发展了非参数回归模型的经验似然;文献[10-12]发展了部分线性模型的经验似然;文献[13-15]将其应用于抽样数据的研究;文献[16]将其应用于经济模型的研究;文献[17-20]将其应用于相依样本;文献[2]综述了其在半参数模型中的应用;文献[3]综述了其在生存分析中的应用等。

1.3 经验似然的推广

尽管经验似然方法具有许多优势,得到广泛应用,但也存在一些不足,主要是 2 个方面:一是计算中需要迭代,计算比较复杂;二是均值参数 μ 有解需要在样本$\{X_1,X_2,\cdots,X_n\}$的凸包内,对一般情况,则是要求 0 位于$\{g(X_1,\boldsymbol{\theta}),\cdots,g(X_n,\boldsymbol{\theta})\}$的凸包内部等,统称为凸包问题。凸包问题在样本量较小时是不可以忽略的,文献[21]通过模拟显示,在一维或二维情况,有可能出现凸包问题(也就是经验似然无解情况)的比例达到 50%。为了克服经验似然中存在的 2 个主要不足:计算复杂和凸包问题,许多学者对此进行研究,提出一些改进办法。

首先考虑计算复杂性问题。为了克服计算的复杂性,文献[5]提出可以考虑用经验欧氏距离来代替对数经验似然,进而提出经验欧氏似然方法,也就是说,针对式(1)的对数经验似然

$$\log\left(\prod_{i=1}^{n}np_i\right)=\sum_{i=1}^{n}\left(\log p_i-\log\left(\frac{1}{n}\right)\right)$$

可以看做是 p_i 和 $1/n$ 的对数似然距离,因此用其欧氏距离

$$\frac{1}{2}\sum_{i=1}^{n}\left(p_i-\frac{1}{2}\right)^2$$

来代替对数经验似然距离。文献[22]进一步研究经验欧氏似然方法,得到与经验似然类似的统计性质,说明该方法计算的简便性。

事实上,可以考虑用更一般的与 p_i 与 $1/n$ 有关的函数来代替对数似然距离。在此思想基础上,文献

[23-25] 进一步提出用 L_q 度量 $\sum_{i=1}^{n}\frac{p_i^{1-q}-(1/n)^{1-q}}{1-q}$ 来代替对数经验似然距离，提出了经验 L_q 似然方法，得到了与经验似然类似的统计性质。当 $q=1$ 时，用 $q\to 1$ 时 L_q 的极限表达来替换，也就是对数经验似然方法。实施上，这种方法在参数情况，文献 [26] 已经进行过研究，文献 [23-25] 只是将其推广到非参数情况，并将其应用到纵向数据的单指标模型、Logistic 回归模型等。

针对凸包问题，也有很多学者进行探讨，文献 [27] 给出经验似然 t 方法；文献 [28-29] 给出惩罚经验似然方法；文献 [30-31] 给出调整经验似然方法；文献 [32-33] 给出平衡经验似然方法等，也进一步探讨了这些方法的统计性质。也有学者把计算复杂和凸包问题结合在一起，提出调整经验欧氏似然、平衡经验欧氏似然、惩罚经验欧氏似然等方法，详细内容可参见文献 [33-35] 等。也有学者从 Bayes 角度对经验似然进行研究，提出贝叶斯经验思想并探讨他们相应的性质，具体可参见文献 [36-40]。

2 非参数拟合优度检验

下面讨论拟合优度检验问题。设总体 $X\sim F$，分布函数 F 未知，从总体 X 中抽取 i.i.d. 样本 $X_1,\cdots,X_n$，感兴趣的问题是检验

$$H_0: F\equiv F_0 \leftrightarrow H_1: F\neq F_0, \tag{3}$$

$$H_0: F\in\mathscr{F}_{\theta} \leftrightarrow H_1: F\notin\mathscr{F}_{\theta}, \tag{4}$$

式中：F_0 为一个完全已知的分布函数；$\mathscr{F}_{\theta}=\{F_{\theta}:\theta\in\Theta\subset\mathbf{R}^d\}$ 为带有未知参数 θ 的分布族（分布类型或形式已知）。式(3)称为简单假设，式(4)称为复合假设。针对上述 2 种检验问题，常用的检验有 2 大类，一类是 χ^2 型检验，另一类是经验分布函数（EDF）型检验。EDF 型检验又可以分为上界型检验和积分型检验，经典的上界型检验包括 Kolmogorov 和 Smirnov（KS）检验、加权 KS 检验等；经典的积分型检验包括 Cramér 和 von-Mises（CV）检验、Anderson 和 Darling（AD）检验以及加权 CV 检验等。其后，EDF 型检验又出现 Berk 和 Jones（BJ）检验、Einmahl 和 Mckeague（EM）检验、修正的 BJ 检验以及本文作者所提出的广义非参数似然比检验（GNLRT）等。此处仅简单介绍一下非参数似然比检验，更具体的内容可见文献 [41-49]。

先考虑简单假设情况，文献 [41] 定义上界型检验统计量

$$R_{n,\lambda}(q)=\sup_x\frac{K_{\lambda}(F_n(x),F_0(x))}{q(F_0(x))}, \tag{5}$$

式中 $F_n(x)$ 为经验分布函数，

$$K_{\lambda}(t,s)=\frac{1}{\lambda(\lambda+1)}\left[\frac{t^{\lambda+1}}{s^{\lambda}}+\frac{(1+t)^{\lambda+1}}{(1-s)^{\lambda}}-1\right]\quad\forall t,s\in(0,1),$$

$\lambda\in(-\infty,\infty)$，$q(\cdot)$ 为一个权函数。当 $\lambda=0,-1$ 时，$K_{\lambda}(t,s)$ 分别取 $\lambda\to 0,-1$ 的极限形式。类似地，定义积分型检验统计量为

$$T_{n,\lambda}(q)=\int_x\frac{K_{\lambda}(F_n(x),F_0(x))}{q(F_0(x))}\mathrm{d}F_0(x)。 \tag{6}$$

事实上，$K_{\lambda}(t,s)$ 是 $(t,1-t)$ 和 $(s,1-s)$ 之间的幂偏差度量（详见文献 [50]），但它不是常规意义上所说的距离（除 $\lambda=1/2$ 外，它不满足对称性和三角不等式）。此外，它还满足：

$$K_{\lambda-1}(t,s)=K_{-\lambda}(s,t),\quad K_0(t,s)=t\log\frac{t}{s}+(1-t)\log\frac{1-t}{1-s},$$

$$K_{-1}(t,s)=s\log\frac{s}{t}+(1-s)\log\frac{1-s}{1-t},\quad K_1(t,s)=\frac{1}{2}\frac{(t-s)^2}{s(1-s)}=K_{-2}(s,t)。$$

$K_0(F_n(x),F_0(x))$ 即为文献 [51] 中所研究的非参数似然比检验统计量（经验似然比形式）：

$$K_0(F_n(x),F_0(x))=\frac{\sup\{L(F):F\in\mathscr{F}\}}{\sup\{L(F):F(x)=F_0(x),F\in\mathscr{F}\}},$$

式中:$L(F)=k\prod_{i=1}^{n}F\{X_i\}$ 为前述所说的非参似然函数;$\mathscr{F}$ 为某一研究的分布族。当 $q\equiv 1$,$R_{n,\lambda}(q)$ 被称为简单上界型检验函数,简记为 $R_{n,\lambda}$;$T_{n,\lambda}(q)$ 被称为简单积分型检验函数,简记为 $T_{n,\lambda}$。一般可以称 $R_{n,\lambda}(q)$($T_{n,\lambda}(q)$)为加权上界型(积分型)检验函数。

当 λ 取一些特殊值时,前面所提到检验都可以包括在此检验类中。例如 $2nR_{n,1}(q)$ 为加权 KS 检验函数的平方;$2nT_{n,1}(q)$ 为加权 CV 检验函数;$R_{n,0}(1)$ 为 BJ 检验函数;$T_{n,0}(1)$ 为 EM 检验函数;$R_{n,-1}(1)$ 为修正 BJ 检验函数等。

如果把式(5)、(6)中的 $F_0(x)$ 替换成 $F(x,\hat{\theta}_n)$,其中 $\hat{\theta}_n$ 为参数 θ 的一个合理估计,例如 MLE,就得到复合零假设(4)的检验函数。当然也可以利用式(5)和(6)的最小化求参数 θ 的最小幂偏差估计。事实上,非参似然比可以看做是理论真实分布与经验分布的某一泛函,如果理论分布包含未知参数,可以直接让此泛函达到最小获得最小非参数似然估计,一般情况下,这个估计比 MLE 估计稳健有效,而且还会使得拟合优度检验更加有效。文献[49,52]讨论了此问题并做简单的比较。

如果 $X\sim F_0$ 为连续型随机变量的分布函数,相应的 $F_0(X)$ 服从(0,1)上均匀分布 $U(0,1)$,文献[41]和文献[53]得到如下理论结果:在简单零假设成立时,$R_{n,\lambda}\triangleq R_{n,\lambda}(1)$($n\geqslant 2$)的精确分布为

$$P\{R_{n,\lambda}\leqslant z\}=P\{\tilde{a}_i\leqslant X_{(i)}\leqslant\tilde{b}_i,1\leqslant i\leqslant n\},$$

式中:

$$\tilde{a}_i=\min\left\{x\ \middle|\ K_\lambda\left(\frac{i}{n},x\right)\leqslant z\right\},i=1,\cdots,n-1;\tilde{a}_n=\tilde{a}_{n-1};$$

$$\tilde{b}_i=1-\tilde{a}_{n-i+1},i=1,\cdots,n。$$

n 较小时可以给出精确分布的显式表达;否则可以利用文献[54]的迭代算法求出其精确分布。由于权函数不依赖于样本和 λ,加权上界型检验函数也有类似结论,这些结论包含文献[45]($\lambda=0$ 和 -1)和文献[43]($\lambda=1$)的结论。当 n 较大时,用精确分布计算比较复杂,这时可以考虑用其极限分布计算。文献[41,55]给出如下结果:在简单零假设成立时,

$$\lim_{n\to\infty}P(nR_{n,\lambda}-e_n\leqslant x)=\exp(-4\exp(-x)),$$

$$\lim_{n\to\infty}P\left(nR_{n,\lambda}-\frac{1}{2}\frac{c_n^2}{b_n^2}\leqslant x\right)=\exp(-4\exp(-x)),$$

式中:

$$b_n=\sqrt{2\log\log n},\ c_n=2\log\log n+\frac{1}{2}\log\log\log n-\frac{1}{2}\log(4\pi),\ e_n=\log\log n+\frac{1}{2}\log\log\log n-\frac{1}{2}\log(4\pi)。$$

通过模拟比较可以发现:$nR_{n,\lambda}-\frac{1}{2}\frac{c_n^2}{b_n^2}$ 比 $nR_{n,\lambda}-e_n$ 收敛到其极限分布的速度稍快一些。

对一般权函数,文献[41,55-56]分别给出如下结论:设 a_n 为一正数序列,$a_n\to 0$,对某一 $k>1$,$na_n^k\to 0$,$na_n/((\log_2 n)^3)\to\infty$;$FC_{0,1}\triangleq\{h(\cdot):\forall\delta\in(0,1/2),\inf_{\delta\leqslant t\leqslant 1-\delta}h(t)>0,h$ 在 0 附近非降,在 1 附近非增$\}$,

$$I(h(\cdot),c)\triangleq\int_0^1\frac{1}{t(1-t)}\exp\left(-c\frac{h^2(t)}{t(1-t)}\right)\mathrm{d}t。$$

假定 $q(t)$ 满足:(i) $t(1-t)q(t)\in FC_{0,1}$;(ii) $\sup_{0<x\leqslant a_n}\frac{t}{t(1-t)q(t)}=o((na_n)^{-1})$,$\sup_{1-a_n\leqslant x<1}\frac{1-t}{t(1-t)q(t)}=o((na_n)^{-1})$;(iii) 对某一 $c>0$,$I(\sqrt{t(1-t)q(t)},c)<\infty$。在简单零假设成立时有

$$2nR_{n,\lambda}(q)\xrightarrow{\mathscr{D}}\sup_{0<t<1}\frac{|B(t)|^2}{t(1-t)q(t)},$$

式中$\{B(t):0\leqslant t\leqslant 1\}$ 为 Brownian 桥过程(详见文献[57]中解释)。

设 a_n 为一正数序列,$a_n\to 0$,对 $\forall k>1$,$na_n^k\to 0$,$na_n/((\log_2 n)^3)\to\infty$,$FIC_{0,1}=\{f(\cdot):\forall\delta\in(0,1/2),\inf_{\delta\leqslant t\leqslant 1-\delta}f(t)>0,\int_0^1[1/f(t)]\,\mathrm{d}t<\infty\}$。进一步假定 $q(t)$ 满足:(i) $q(\cdot)\in FIC_{0,1}$;(ii) $\int_0^{a_n}1/q(t)\,\mathrm{d}t=$

$o((na_n)^{-1})$，$\int_{1-a_n}^{1} 1/q(t)\,\mathrm{d}t = o((na_n)^{-1})$。在简单零假设成立时有

$$2nT_{n,\lambda}(q) \xrightarrow{\mathscr{D}} \int_0^1 \frac{B^2(x)}{x(1-x)q(x)}\,\mathrm{d}x。$$

如果样本容量 n 有限，尽管可以给出其精确分布，但随着 n 的增加，需要迭代的次数增加，计算时间较长。文献[53]从计算精度和时间长度综合考虑，并通过模拟显示：如果 $n\leqslant 110$，分位数建议利用精确分布计算；如果 $n\geqslant 150$，分位数建议利用模拟或其极限分布计算。不同 λ 和不同权函数 $q(\cdot)$，功效较高的检验也不同。文献[53]发现，对给定的 $q(\cdot)$，不存在对任何情况都一致最优的检验，在小样本情况，λ 对检验函数的影响要比权函数的影响大一些，对于较大的 $|\lambda|$，$R_{n,\lambda}(q)$ ($T_{n,\lambda}(q)$) 使得 $F(t)$ 和 $F_0(t)$ 之间的差异变得更大，再加上误差等因素，检验函数基本无用，故一般情况，文献[53]建议 $\lambda\in[-2,2]$。

是否存在一个与 λ 无关的一致较优的检验统计量呢？文献[42]发现，在给定权函数 $q(\cdot)$ 时，功效一致较优的上界型非参似然比检验函数为

$$R_n(q) \triangleq \sup_{\lambda\in I} \sup_{X_{(1)}\leqslant x<X_{(n)}} K_\lambda(F_n(x),F_0(x))/q(F_0(x))，$$

其中 I 为一个区间，一般情况下，如果 $q\equiv 1$，经验上可取 $\lambda\in I=[-2/3,0]$；如果 $q\equiv[t(1-t)]^{-1}$，经验上可取 $\lambda\in I=[-1,2]$。在给定权函数 $q(\cdot)$ 时，功效一致较优的积分型非参似然比检验函数为

$$T_n(q) \triangleq \sup_{\lambda\in I}\int_{X_{(1)}}^{X_{(n)}} K_\lambda(F_n(x),F_0(x))/q(F_0(x))\,\mathrm{d}F_0(x)。$$

一般情况下，如果 $q\equiv 1$ 和 $q\equiv[t(1-t)]^{-1}$，经验上可取 $\lambda\in I=[-2,2]$。文献[42,48]中给出较为详细的模拟说明。

对于高维数据的拟合优度检验，由于“维数祸根(Huber,1985)”的存在，在一维上比较有效的检验，在高维上未必还有效，这时一般采用投影或者变换的思想去处理。文献[42]借助投影追踪技术，提出投影追踪非参数似然比检验。当然，也有借助经验欧氏似然讨论检验问题，例如文献[58-60]等，其基本思想与前述讨论类似，感兴趣的读者可参见相关文献。

3 复杂数据的非参数似然方法

文献[61]指出，21 世纪统计学主要面临 3 大问题：相依、高维和不完全数据的统计理论。针对这些复杂数据，本章简单综述在复杂数据下的非参数似然方法。

相依数据在实际工作中非常普遍，例如医学、生物学、金融学等许多领域都会出现相依数据。理论研究中，随机变量的函数、时间序列数据、马氏链和 Copulas 结构等都是相依数据。目前在相依序列数据中研究比较多的是鞅差序列、混合序列、相伴序列等。一般来说，此时数据结构比较复杂，样本内部不再是 i.i.d.，传统统计方法的前提条件往往很难满足，理论性质讨论相对比较困难。目前的主要解决思想就是分组(Blockwise)，把相依数据转化为独立或近似独立情况，然后再借助相关理论进行研究。针对混合相依的经验似然，文献[17]研究 α 混合相依序列；文献[18]研究 ϕ 混合相依序列；文献[62]研究一般混合相依序列等。针对相伴相依的经验似然，文献[19]研究负相伴相依序列；文献[20]研究正相伴相依序列等。还有较多文献也探讨相依序列下的经验似然并应用到各种模型中，具体可参见文献[62-71]等。

高维数据也是近年来学者们非常关注的研究方向，它在金融、生物信息、经济等多个领域都很普遍。前面简单介绍了高维数据的拟合优度检验问题，文献[42]利用投影追踪技术进行研究。此外，文献[29,72-74]利用前面介绍的非参数似然方法探讨高维数据的变量选择和模型建立等方法，也获得了一些重要的理论结果。

不完全数据主要包括缺失或删失数据、四舍五入数据等，是常见的数据类型。针对缺失数据，文献[75]给出部分线性模型的经验似然方法；针对删失数据，文献[2]总结了生存分析中的经验似然方法；针对舍入数据，文献[76-77]研究舍入数据的经验似然方法，文献[78]研究舍入数据的经验欧氏似然等。

除以上比较困难的 3 大数据外，现实生活中还有许多现象都可以借助非参数似然方法进行研究。例

如当总体是离散分布时,如何合理利用非参数似然方法也是一个值得探讨的问题,文献[79-80]对此进行探讨。在风险分析中,VaR 和 ES 的估计也是一个比较重要的问题,文献[39-40,81-82]分别利用经验似然、经验欧氏似然和贝叶斯经验似然对其进行探讨;文献[83]研究具有测量误差的纵向部分线性模型的经验欧氏似然方法;文献[84]研究半参数位置偏移混合模型的最小似然距离估计;文献[85]研究基于平滑经验欧氏似然的分位回归分布式估计等。

近年来,变点问题也得到越来越多学者的重视,它在经济、地质、气候、生物医学、图像处理以及信号去噪等方面都有较广泛的应用。针对参数模型的变点检测与估计,文献[86-87]给出较为详细的综述;针对非参数变点检测,文献[88-91]分别探讨经验似然的变点检测方法。近几年,把变点检测看成一个稀疏问题,借助变量选择方法,从不同角度对变点检测和估计的非参数方法重新进行讨论,取得一定成果。由于问题内容和解决办法相对比较复杂,在此不再赘述,感兴趣的读者可参见文献[92-98]。

4 总结与展望

非参数似然比是近年来的一个研究热点,不同于模型研究,它是一种方法上的创新,可以适用于多种模型。本文简单介绍非参数似然的发展思路,从估计、检验和复杂数据运用 3 个方面进行简单综述,给出一些基本的研究内容和结论,指出其在复杂数据研究中的应用,希望这些内容能提供一定的借鉴意义和参考价值,尤其是那些对非参数似然方法有兴趣的读者。

目前,非参数似然方法还有许多方面可以进一步考虑,例如:可以考虑构造 p_i 与 $1/n$ 之间的合理函数,进一步说明它们在估计和检验中的优势;如何充分利用辅助信息,提高估计或检验的精度;对于一些复杂数据或模型,如何把现有工具合理科学地应用好等。特别地,一些实际问题并不需要很复杂的方法或者模型,关键是尽可能地利用好已有相关信息,得到简单实用的模型或方法,较好地发现规律或问题,为决策者制定正确科学的决策服务。

参 考 文 献

[1] OWEN A B. Empirical likelihood [M]. London: Chapman and Hall, 2001.

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

[3] ZHOU M. Empirical likelihood method in survival analysis [M]. Boca Raton: CRC Press, 2016.

[4] OWEN A B. Empirical likelihood ratio confidence intervals for a single functional [J]. Biometrika, 1988, 75(2): 237-249.

[5] OWEN A B. Empirical likelihood ratio confidence regions [J]. The Annals of Statistics, 1990, 18(1): 90-120.

[6] QIN J, LAWLESS J. Empirical likelihood and general estimating equations [J]. The Annals of Statistics, 1994, 22(1): 300-325.

[7] OWEN A B. Empirical likelihood for linear models [J]. The Annals of Statistics, 1991, 19(4): 1725-1747.

[8] CHEN S X. Empirical likelihood confidence intervals for nonparametric density estimation [J]. Biometrika, 1996, 83(2): 329-341.

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

[10] WANG Q H, JING B Y. Empirical likelihood for partially linear models with fixed design [J]. Statistics & Probability Letters, 1999, 41(4): 425-433.

[11] 秦永松. 部分线性模型参数的经验似然比置信域 [J]. 应用概率统计. 1999, 15(4): 363-369.

[12] SHI J, LAU T S. Empirical likelihood for partially linear models [J]. Journal of Multivariate Analysis, 2000, 72(1): 132-148.

[13] CHEN J, QIN J. Empirical likelihood estimation for finite populations and the effective usage of auxiliary information [J]. Biometrika, 1993, 80(1): 107-116.

[14] ZHONG B, RAO J N K. Empirical likelihood inference under stratified random sampling using auxiliary population information [J]. Biometrika, 2000, 87(4): 929-938.
[15] CHEN J H, WU C B. Estimation of distribution function and quantiles using the model calibrated pseudo empirical likelihood method [J]. Statistica Sinica, 2002, 12(4): 1223-1239.
[16] KITAMURA Y. Empirical likelihood methods in econometrics: theory and practice [J]. Cowles Foundation Discussion Paper, 2006: 1569.
[17] KITAMURA Y. Empirical likelihood methods with weakly dependent processes [J]. The Annals of Statistics, 1997, 25(5): 2084-2102.
[18] 张军舰, 王成名, 王炜忻. 相依样本下的经验似然比置信区间 [J]. 高校应用数学学报, 1999, 14A(1): 63-72.
[19] ZHANG J J. Empirical likelihood for NA series [J]. Statistics & Probability Letters, 2006, 76(2): 153-160.
[20] ZHANG J J. Empirical likelihood ratio confidence interval for positively associated series [J]. Acta Mathematicae Applicatae Sinica, 2007, 23(2): 245-254.
[21] TSAO M. Bounds on coverage probabilities of the empirical likelihood ratio confidence regions [J]. The Annals of Statistics, 2004, 32(3): 1215-1221.
[22] 茆诗松, 罗旭. 可加模型中参数的经验欧氏似然估计 [J]. 应用概率统计, 1996, 12(4): 383-392.
[23] 张玉芳. 经验 Lq 似然 [D]. 桂林: 广西师范大学, 2012.
[24] 黄锦联. 二分类 Logistic 回归模型的 Lq 似然估计 [D]. 桂林: 广西师范大学, 2013.
[25] 罗巧玲. 纵向数据单指标模型的经验 Lq 似然方法 [D]. 桂林: 广西师范大学, 2014.
[26] FERRARI D, YANG Y. Maximum Lq-likelihood estimation [J]. The Annals of Statistics, 2010, 38(2): 753-783.
[27] BAGGERLY K A. Studentized empirical likelihood and maximum entropy [R]. Technical Report, Department of Statistics, Rice University, 1999.
[28] BARTOLUCCI F. A penalized version of the empirical likelihood ratio for the population mean [J]. Statistics & Probability Letters, 2007, 77(1): 104-110.
[29] 李吉妮. 单指标模型的高维惩罚经验似然 [D]. 桂林: 广西师范大学, 2014.
[30] CHEN J H, VARIYATH A M, ABRAHAM B. Adjusted empirical likelihood and its properties [J]. Journal of Computational and Graphical Statistics, 2008, 17(2): 426-443.
[31] VARIYATH A M, CHEN J H, ABRAHAM B. Empirical likelihood based variable selection [J]. Journal of Statistical Planning & Inference, 2010, 140(4): 971-981.
[32] EMERSON S C, OWEN A B. Calibration of the empirical likelihood method for a vector mean [J]. Electronic Journal of Statistics. 2009, 3: 1161-1192.
[33] 申群海, 黄运生, 张军舰. 平衡增加的经验欧氏似然 [J]. 应用概率统计, 2014, 30(5): 527-536.
[34] 胡志明. 调整经验欧氏似然及其性质 [D]. 桂林: 广西师范大学, 2011.
[35] 胡志明, 晏振, 张军舰. 调整经验欧氏似然及其性质 [J]. 应用数学, 2017, 30(2): 299-312.
[36] LAZAR N A. Bayesian empirical likelihood [J]. Biometrika, 2003, 90(2): 319-326.
[37] YANG Y W, HE X M. Bayeian empirical likelihood for quantile regression [J]. The Annals of Statistics, 2012, 40(2): 1102-1131.
[38] WU C B, Rao J N K. Bayesian pseudo-empirical-likelihood intervals for complex surveys [J]. Journal of the Royal Statistical Society(B), 2010, 72(4): 533-544.
[39] 赖廷煜. VaR 和 ES 的 Bayes 经验似然估计 [D]. 桂林: 广西师范大学, 2015.
[40] 张军舰, 赖廷煜, 杨晓伟. VaR 和 ES 的贝叶斯经验似然估计 [J]. 广西师范大学学报(自然科学版), 2016, 34(4): 38-45.
[41] 张军舰. 广义非参数似然比拟合优度检验 [D]. 北京: 中国科学院数学与系统科学研究院, 2006.
[42] 张军舰. 非参数似然比拟合优度检验 [D]. 北京: 北京工业大学, 2008.
[43] BERK R H, JONES D H. Goodness-of-fit statistics that dominate the Kolmogorov statistics [J]. Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete, 1979, 47: 47-59.
[44] EINMAHL J H J, MCKEAGUE I W. Empirical likelihood based hypothesis testing [J]. Bernoulli, 2003, 9(2): 267-290.
[45] JAGER L, WELLNER J A. A new goodness of fit test: the reversed Berk-Jones statistic [R]. Technical Report 443, Department of Statistics, University of Washington, 2004.

[46] JAGER L, WELLNER J A. Goodness-of-fit test via phi-divergence [J]. The Annals of Statistics, 2007, 35(5): 2018-2053.
[47] ZHANG J. Powerful goodness-of-fit tests based on the likelihood ratio [J]. Journal of the Royal Statistical Society (B), 2002, 64(2): 281-294.
[48] 杨振海, 程维虎, 张军舰. 拟合优度检验 [M]. 北京: 科学出版社, 2011.
[49] 李海锋. 最小非参似然比拟合优度检验 [D]. 桂林: 广西师范大学, 2010.
[50] CRESSIE N A C, READ T R C. Multinominal goodness-of-fit tests [J]. Journal of the Royal Statistical Society (B), 1984, 46(3): 440-464.
[51] OWEN A B. Nonparametric likelihood confidence bands for a distribution function [J]. Publications of the American Statistical Association, 1995, 90(430): 516-521.
[52] 张军舰, 黄运生. 复合零假设下的修正 Berk-Jones 检验 [J]. 应用数学, 2010, 23(1): 27-31.
[53] 张军舰, 李国英, 赵志源. 一类上界型拟合优度检验统计量的精确分布 [J]. 应用概率统计, 2009, 25(4): 421-432.
[54] NOE M. The calculation of distributions of two-sided Kolmogorov-Smirnov type statistics [J]. Annals of Mathematical Statistics, 1972, 43: 58-64.
[55] 张军舰, 李国英. 上界型拟合优度检验 [J]. 数学物理学报, 2010, 30A(2): 344-357.
[56] ZHANG J J, LI G. Integral-type tests for goodness of fit [J]. Journal of Systems Science and Complexity, 2010, 23(4): 784-795.
[57] SHORACK G R, WELLNER J A. Empirical processes with applications to statistics [M]. New York: John Wiley & Sons, 1986.
[58] 詹欢. 欧氏似然拟合优度检验 [D]. 桂林: 广西师范大学, 2012.
[59] 张军舰, 詹欢, 晏振. 基于经验欧氏似然的拟合优度检验 [J]. 广西师范大学学报(自然科学版), 2012, 30(3): 30-35.
[60] 张新成, 张军舰, 詹欢. 基于垂直密度表示的经验欧氏拟合优度检验 [J]. 广西师范大学学报(自然科学版), 2013, 31(4): 60-65.
[61] 李国英. 关于高维、相依和不完全数据的统计分析 [J]. 数学进展, 2002, 31(3): 193-199.
[62] LIN L, ZHANG R C. Blockwise empirical Euclidean likelihood for weakly dependent processes [J]. Statistics & Probability Letters, 2001, 53(2): 143-152.
[63] BANDYOPADHYAY S, LAHIRI S N, NORDMAN D J. A frequency domain empirical likelihood method for irregularly spaced spatial data [J]. The Annals of Statistics, 2015, 43(2): 519-545.
[64] CHENG J, PENG L, ZHAO Y. Empirical likelihood based confidence intervals for copulas [J]. Journal of Multivariate Analysis, 2009, 100(1): 137-151.
[65] 胡小凤. Copula 的经验欧氏似然估计 [D]. 桂林: 广西师范大学, 2013.
[66] QIN Y S, LEI Q Z. Empirical likelihood for linear models under linear process errors [J]. Communications in Statistics-Theory and Methods, 2015, 44: 3218-3233.
[67] LI Y H, QIN Y S, LEI Q Z. Empirical likelihood for quantiles under associated samples [J]. Acta Mathematicae Applicatae Sinica (English Series), 2015, 31(1): 71-80.
[68] LEI Q Z, QIN Y S. Confidence intervals for probability density functions under strong mixing samples [J]. Journal of Nonparametric Statistics, 2015, 27(2): 181-193.
[69] LEI Q Z, QIN Y S. Empirical likelihood for partially linear models under negatively associated errors [J]. Journal of Systems Science and Complexity, 2016, 29: 1145-1159.
[70] LI Y H, QIN Y S, LEI Q Z, et al. Quantile estimation with auxiliary information under positively associated samples [J]. Acta Mathematica Scientia, 2016, 36(2): 453-468.
[71] LI Y H, QIN Y S, LI Y. Empirical likelihood for nonparametric regression models with spatial autoregressive errors [J]. Journal of the Korean Statistical Society, 2020, 50: 447-478.
[72] 李玲玲. 高维线性模型的变量选择 [D]. 桂林: 广西师范大学, 2010.
[73] 田镇滔. 超高维删失数据的条件分位数特征筛选 [D]. 桂林: 广西师范大学, 2021.
[74] 田镇滔, 张军舰. 基于分位数方法的超高维删失数据的特征筛选 [J]. 广西师范大学学报(自然科学版), 2021, 39(6): 99-111.
[75] QIN Y S, LI Y H. Empirical likelihood for partially linear models with missing responses: the fixed design case [J]. Journal

of Nonparametric Statistics, 2011, 23(1/2): 497-511.
[76] 唐美燕. 四舍五入数据对统计推断的影响分析[D]. 桂林: 广西师范大学, 2011.
[77] 吕婷. 舍入数据的经验似然方法[D]. 桂林: 广西师范大学, 2012.
[78] 何伦春. 基于舍入数据的经验欧氏似然估计[D]. 桂林: 广西师范大学, 2013.
[79] 李自军. 离散分布分位数的经验欧氏似然估计[D]. 桂林: 广西师范大学, 2013.
[80] MOLANES E M L, VAN KEILEGOM I, VERAVERBEBK N. Empirical likelihood for non-smooth criterion functions[J]. Scandinavian Journal of Statistics, 2009, 36(3): 413-432.
[81] 晏振. VaR 和 ES 的调整经验似然估计[D]. 桂林: 广西师范大学, 2012.
[82] YAN Z, ZHANG J J. Adjusted empirical likelihood for value at risk and expected shortfall[J]. Communications in Statistics-Theory and Methods, 2017, 46(5): 2580-2591.
[83] 易家英. 具有测量误差的纵向部分线性模型的经验欧氏似然方法[D]. 桂林: 广西师范大学, 2014.
[84] 马岱君. 半参数位置偏移混合模型的最小似然距离估计[D]. 桂林: 广西师范大学, 2020.
[85] 彭博. 基于平滑经验欧氏似然的分位回归分布式估计[D]. 桂林: 广西师范大学, 2021.
[86] 陈希孺. 变点统计分析简介[J]. 数理统计与管理, 1991, 1: 55-88.
[87] CSÖRGŐ M, HORVÁTH L. Limit theorem in change-point analysis[M]. New York: John Wiley & Sons, 1997.
[88] ZOU C L, LIU Y K, PENG Q, WANG Z J. Empirical likelihood ratio test for the change-point problem[J]. Statistics & Probability Letters, 2007, 77: 374-382.
[89] SHEN G. On empirical likelihood inference of a change-point[J]. Statistics & Probability Letters, 2013, 83: 1662-1668.
[90] ZOU C L, YIN G, FENG L, at al. Nonparametric maximum likelihood approach to multiple change-point problems[J]. The Annals of Statistics, 2014, 42(3): 970-1002.
[91] GAMAGE R D P, NING W. Empirical likelihood for change point detection in autoregressive models[EB/OL]. (2017-11-27)[2022-02-11]. https://arXiv:1711.09533vl.
[92] 覃志俭. 基于修正 BJ 方法的多变点检测研究[D]. 桂林: 广西师范大学, 2016.
[93] 杨晓伟. 混合回归模型的变点检测[D]. 桂林: 广西师范大学, 2017
[94] 卢振法. 基于自适应 LASSO 的均值变点检测[D]. 桂林: 广西师范大学, 2018
[95] 黄新程. 均值变点的两阶段估计[D]. 桂林: 广西师范大学, 2019
[96] 李智航. 基于经验欧氏似然的变点检测[D]. 桂林: 广西师范大学, 2020
[97] 黄雄琪. 混合模型实时变点的期望分位数检测[D]. 桂林: 广西师范大学, 2021
[98] 张军舰, 李智航. 基于经验欧氏似然的均值变点检测[J]. 应用概率统计, 2021, 37(1): 47-58.

Review on Nonparametric Likelihood and Their Applications

ZHANG Junjian

(School of Mathematics and Statistics, Guangxi Normal University, Guilin Guangxi 541006, China)

Abstract: Nonparametric likelihood, especially empirical likelihood, is one of the nonparametric methods, which has many similar properties with traditional parametric likelihood. This fascinating methodology has attracted the attention of many researchers, and become one of the important research fields in statistics. Numerous papers have been published and the list is getting longer every day. This review summarizes the nonparametric likelihood method from estimation, test and the applications in complex data, based on the works of the research group. The research approach and content are analyzed, some major findings are provided in the review.

Keywords: nonparametric likelihood; empirical likelihood; estimation; goodness of fit; complex data

(责任编辑 吴佃华)