

基于经验似然的流形条件Fréchet均值同时置信带与均值集

黄世锋

广西师范大学数学与统计学院, 广西 桂林

收稿日期: 2026年9月26日; 录用日期: 2026年10月17日; 发布日期: 2026年10月28日

摘 要

对流形数据的位置参数, 统计推断通常只给出单个点或者逐点意义下的置信区域。然而在许多实际问题中, 人们关心的并不是某一个点, 而是一族点或者一个集合: 例如沿某个协变量连续变化的整条条件均值曲线, 又或者当总体位置本身不唯一时由多个点共同组成的集合。对这类对象, 逐点区间无法控制整体的置信水平, 必须借助同时的或者集合型的推断工具。本文在流形数据的经验似然框架下处理两类这样的问题: 一是条件Fréchet均值在有限预测值网格上的同时置信带, 二是Fréchet均值不唯一时的均值集置信区域。之所以选用经验似然, 是因为其统计量只使用估计函数的矩协方差、并不涉及Fréchet函数的Hessian, 因而在Hessian退化时依然有效。文中引言先按流形统计推断、经验似然、条件Fréchet均值与Fréchet均值集四个方向综述已有进展。就方法而言, 前者由局部经验似然统计量的逐点极限经联合界得到, 不需要带宽收缩下的经验过程Donsker条件; 后者则以经验似然函数作为准则函数构造水平集估计, 并用乘子自助校准其覆盖概率。文中还说明了乘子自助阈值在此设定下渐近正确的理由, 并讨论了网格点数随样本量增长时所面临的技术困难。数值研究的结果与理论一致, 文中最后讨论了向连续预测带推广的方向。

关键词

集合型推断, 均值集, 水平集估计, 乘子自助, 条件Fréchet均值, 流形数据

Manifold Conditional Fréchet Mean Based on Empirical Likelihood, with Simultaneous Confidence Band and Mean Set

Shifeng Huang

School of Mathematics and Statistics, Guangxi Normal University, Guilin Guangxi

Received: September 26, 2026; accepted: October 17, 2026; published: October 28, 2026

文章引用: 黄世锋. 基于经验似然的流形条件 Fréchet 均值同时置信带与均值集[J]. 统计学与应用, 2026, 15(10): 157-168. DOI: 10.12677/sa.2026.1510236

Abstract

For the positional parameters of manifold data, statistical inference usually only provides confidence regions at a single point or point by point sense. However, in many practical problems, people are not concerned with a single point, but with a family of points or a set: for example, the entire conditional mean curve that continuously changes along a covariate, or a set composed of multiple points when the population position itself is not unique. For such objects, point by point intervals cannot control the overall confidence level, and simultaneous or set based inference tools must be used. This article deals with two types of problems in the empirical likelihood framework of manifold data: one is the simultaneous confidence band for the conditional Fréchet mean on a finite prediction-value grid; the other is the confidence region for the mean set when Fréchet means are not unique. The reason for choosing empirical likelihood is that its statistic only uses the moment covariance of the estimating function and does not involve the Hessian of the Fréchet function, so it is still effective in when the Hessian is degenerate. The introduction of the article first summarizes the progress in manifold statistical inference, empirical likelihood, conditional Fréchet mean, and Fréchet mean set. In terms of methodology, the former is obtained by the joint bound of the point by point limit of the local empirical likelihood statistic, and does not require the Donsker condition of the empirical process under bandwidth contraction; The latter uses empirical likelihood function as the criterion function to construct level set estimation, and self calibrates its coverage probability using multipliers. The article also explains the reason for the asymptotic correctness of the multiplier bootstrap threshold under this setting, and discusses the technical difficulties faced when the number of grid points increases with the sample size. The results of numerical research are consistent with theory, and the direction of extending to continuous prediction bands is discussed in the end of the article.

Keywords

Set-Based Inference, Mean Set, Level Set Estimation, Multiplier Bootstrap, Conditional Fréchet Mean, Manifold Data

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

设想我们沿时间或者某个协变量观察一批形状数据，希望了解这批对象的“平均形状”如何随该变量变化。需要回答的并不是某个单独时点上的平均形状，而是整条曲线；同样，当代替唯一均值的是一组点构成的均值集合时，需要给出的也是一个集合的置信区域。这两类同时的与集合型的推断问题，都无法靠逐点置信区间解决，曲线上的每一点各自以 $1-\alpha$ 的概率被覆盖，并不意味着整条曲线被同时覆盖的概率也是 $1-\alpha$ 。本文在流形数据的经验似然框架下处理这两类问题：条件 Fréchet 均值的有限网格同时置信带，以及 Fréchet 均值不唯一时的均值集置信区域。

取值于非欧度量空间的对象型数据在近年日益常见，球面方向、形状构型、协方差矩阵以及函数型对象都属于此类，其统计分析可参见 Marron 和 Alonso [1] 与 Patrangenaru 和 Ellingson [2] 的综述。在这类数据上，标准的位置概括是 Fréchet 均值，即 Fréchet 函数 $F(p) = E[d^2(p, X)]$ 的最小值点；它把欧氏空间中“平方距离的期望最小”这一均值刻画原封不动地搬到了度量空间上。Fréchet 均值的渐近理论由

Bhattacharya 和 Patrangenaru [3] 与 Bhattacharya 和 Patrangenaru [4] 建立, 他们给出了内在与外在样本均值的存在性、唯一性与相合性, 其中心极限定理的重要推广见 Bhattacharya 和 Lin [5]。基于该极限定理, 对 Fréchet 均值的推断长期以来以 Wald 型方法为主。这类方法把置信区域取为切空间中的一个椭球, 形式简洁、计算方便, 但其有效性有一个不容回避的前提: Fréchet 函数在均值处的 Hessian 必须非奇异。当 Hessian 退化时, 样本均值的收敛速率会偏离标准的 $\sqrt{n}$, 极限分布也不再是正态的, 这一现象称为 smeariness, 此时 Wald 型区域随之失去名义覆盖。Eltzner 和 Huckemann [6] 给出了这一情形下的一般中心极限定理, Hundrieser 等 [7] 则进一步指出, 即使渐近意义下是正则的, 有限样本下经典逼近仍可能明显不准。更值得注意的是, smeariness 在实际数据中往往难以事先判定, 因此它并不是一种罕见病态, 而是一个需要普遍防范的风险。

经验似然为回避 Hessian 提供了一条自然的途径。经验似然由 Owen 系统发展, 它与一般估计方程之间的紧密联系由 Qin 和 Lawless 阐明, 详情参见文献 [8]–[10]; Hjort 等 [11] 把它的适用范围扩展到含讨厌参数的插入估计、慢于 $n^{-1/2}$ 的收敛速率以及估计方程个数较多的情形, Chen 和 Van Keilegom [12] 则对回归情形下的经验似然方法作了系统综述。在流形数据方面, Kurisu 和 Otsu [13] 建立了相应的经验似然: 他们利用流形的局部欧氏结构刻画样本 Fréchet 均值所满足的估计方程, 据此构造经验似然函数, 证明了 Wilks 定理, 并指出单样本的枢轴量对 smeariness 是稳健的。之所以如此, 原因在于该枢轴量的极限只依赖估计函数的矩协方差, 而完全不涉及 Fréchet 函数的 Hessian。此后, Bharath 等 [14] 把这一构造推广到 openbook 这样的分层度量空间。不过需要指出的是, 上述工作所给出的推断对象大多是单个或者逐点的参数, 这恰恰是本文要补上的两个空缺。

与本文所关心的两个问题直接相关的, 还有两篇文献。其一是条件 Fréchet 均值。它的估计由 Petersen 和 Müller [15] 提出的 Fréchet 回归给出, 该文把 Fréchet 均值的概念推广为条件 Fréchet 均值, 并建立了全局与局部加权两种回归形式; Chen 和 Müller [16] 借助经验过程的工具给出了局部 Fréchet 回归的一致收敛速率。这两项工作解决的是估计问题, 而如何为整条条件均值曲线给出同时的置信带, 则尚未涉及。其二是 Fréchet 均值集。Schötz [17] 给出了均值集的强大数律, Evans 和 Jaffe [18] 建立了相应的极限定理, Blanchard 和 Jaffe [19] 则在 Hausdorff 度量下构造了相合估计; 集合型参数置信区域的一般框架可参见 Chernozhukov 等 [20]。在这些结果之上, Kurisu 和 Otsu [13] 以经验似然函数作为准则函数给出了相合的水平集估计。本文第 4 节的均值集区域, 正是在这一构造的基础上补上覆盖概率的校准。

综合来看, 本文的工作可以概括为两点: (1) 条件 Fréchet 均值的同时置信带。针对有限个预测值构成的网格, 给出由逐点极限经联合界得到的同时带。该构造不依赖带宽收缩下的经验过程 Donsker 条件, 因而形式简单、易于实现; 同时给出一个乘子自助的实用阈值, 有关的讨论见注 2。(2) Fréchet 均值集的置信区域。以经验似然函数作为准则函数构造水平集估计, 给出乘子自助的覆盖校准, 并刻画该区域在总体平稳集之外的超出行为以及边界情形。

最后需要说明的是, 经验似然在流形上的基本构造、Wilks 定理及其对 smeariness 的稳健性均来自 Kurisu 和 Otsu [13]; 本文在此基础上处理上述两类集合型推断, 并不重复这些基础结果。在理论层面, 本文补充了三处细节: 逐点极限证明中核偏差项的处理(见定理 1 的证明与注 1)、乘子自助阈值渐近正确性的说明(见引理 1), 以及网格点数随样本量增长时联合界失效的机理(见注 2)。后文的安排如下: 第 3 节给出条件 Fréchet 均值的同时置信带, 第 4 节给出 Fréchet 均值集的置信区域, 第 5 节报告数值研究, 最后是结论。

2. 预备: 流形经验似然框架

本节回顾已有文献所建立的流形经验似然框架。本文后续的两类构造都以它为基础, 因此有必要先

把记号与基本事实交代清楚。

2.1. 记号与估计函数

设 $\mathcal{M}$ 为 m 维 Riemannian 流形, 其上的测地距离记为 d , 并设观测到取值于 $\mathcal{M}$ 的随机对象 X 。记 Fréchet 函数为 $F(p) = E[d^2(p, X)]$, 它的最小值点即 Fréchet 均值 μ 。对任意 $p \in \mathcal{M}$, 切空间记为 $T_p\mathcal{M}$, 指数映射与其逆分别记为 $\exp_p$ 与 $\log_p$; 后者把流形上的点送回切空间, 从而使得在均值附近可以像在欧氏空间中一样进行线性化处理。

估计函数由图表 Fréchet 函数在原点的导数给出。沿用 Kurisu 和 Otsu [13] 的构造, 当 $q=2$ 时该估计函数恰 $g(X, \mu) = -2\log_\mu X$, 于是 $\|g\| = 2d(\mu, X)$, 并且由 μ 的最小性可得 $\|g\| = 2d(\mu, X)$ 这一等式正是把 Fréchet 均值的求解转化为估计方程求解的关键。关于 $q=1$ (即中位数) 的情形以及相应的曲率引理, 亦参见 Kurisu 和 Otsu [13]。

2.2. 经验似然

以该估计方程为基础构造的经验似然统计量 $\ell(\mu)$ 取对偶形式, 它是一个关于乘子 λ 的低维凸优化问题, 其具体表达式、原假设下的极限 (即 Wilks 定理) 以及由此得到的置信区域, 分别参见 Owen [9] 与 Kurisu 和 Otsu [13], 此处不再重录。后文把相应的置信区域记为 $C_{1-\alpha} = \{\mu: \ell(\mu) \leq \chi_{m, 1-\alpha}^2\}$ 。

值得指出的是, 该统计量所依赖的只有估计函数的矩协方差, 而与 Fréchet 函数的 Hessian 无关。正因如此, 即使 Hessian 退化、Wald 型方法失效, 上述置信区域仍能保持名义覆盖, 这也是本文在流形数据上选用经验似然的根本原因。

3. 条件 Fréchet 均值的同时置信带

3.1. 局部经验似然

设 $Z \in \mathbb{R}^k$ 为欧氏预测变量, $X \in \mathcal{M}$ 为响应变量, 并观测到独立同分布的样本 $(X_i, Z_i), i=1, \dots, n$ 。为刻画 X 随 Z 变化的规律, 需要把 Fréchet 均值的概念局部化: 核加权局部 Fréchet 函数及其样本版本 $F(p; z)$ 与 $\hat{F}_n(p; z)$ 的构造见 Petersen 和 Müller [15] 与 Chen 和 Müller [16], 其基本思想是在预测值附近用核函数给样本加权, 使远离 z 的观测贡献较小, 此处不再重录。记条件 Fréchet 均值为 $\mu_z = \arg \min_p F(p; z)$, 它是曲线 $z \mapsto \mu_z$ 上的一个点。

相应的局部经验似然统计量为

$$\ell(\mu_z; z) = 2 \max_{\lambda} \sum_{i=1}^n \log \left(1 + \lambda' K \left(\frac{Z_i - z}{h} \right) g(X_i, \mu_z) \right) \quad (1)$$

即在经验似然统计量中以核权 $K((Z_i - z)/h)$ 加权各样本的贡献。直观地说, 与 z 较近的样本被赋予较大权重, 从而使得该统计量只反映 z 附近的局部信息。

3.2. 逐点极限

局部经验似然统计量的第一个性质, 是它在逐点意义下具有稳定的极限分布。

定理 1 (逐点极限)。在正则条件下 (样本独立、 μ_z 唯一且可识别、核函数有界对称 Lipschitz、条件光滑性等), 当 $h \rightarrow 0$ 、 $nh^k \rightarrow \infty$ 且欠光滑 $nh^{k+4} \rightarrow 0$ 时, 局部经验似然统计量依分布收敛到 χ_m^2 。

证明。记 $K_i(z) = K((Z_i - z)/h)$, $g_i = K_i(z)g(X_i, \mu_z)$ 。由核有界与条件光滑性,

$$\max_i \|g_i\| = o_p \left((nh^k)^{1/2} \right); \text{ 由 } f_z(z) > 0 \text{ 和 } \Sigma_z(z) > 0,$$

$$V_z = f_Z(z) \Sigma_z(z) \int K(u)^2 du > 0,$$

故对偶解存在且 $\hat{\lambda} = O_p\left((nh^k)^{-1/2}\right)$ 。由 $\log(1+x) = x - \frac{x^2}{2} + O(x^3)$ 与一阶条件 $\sum_i g_i / (1 + \hat{\lambda}' g_i) = 0$,

$$\ell(\mu_z; z) = 2 \left(\hat{\lambda}' \sum_i g_i - 12 \sum_i (\hat{\lambda}' g_i)^2 \right) + o_p(1) = \left(\frac{1}{\sqrt{nh^k}} \sum_i g_i \right)' \hat{V}_z^{-1} \left(\frac{1}{\sqrt{nh^k}} \sum_i g_i \right) + o_p(1),$$

其中 $\hat{V}_z = (nh^k)^{-1} \sum_i K_i^2 g_i g_i'$; 欠光滑 $nh^{k+4} \rightarrow 0$ 使核偏差在 $O(h^{k+2})$ 尺度上可忽略, 中心化与非中心化版本同极限。

把加权和分解为中心化部分与偏差部分

$$\frac{1}{\sqrt{nh^k}} \sum_i g_i = \frac{1}{\sqrt{nh^k}} \sum_i (g_i - E[g_i]) + \frac{1}{\sqrt{nh^k}} \sum_i E[g_i] := S_n + B_n,$$

则 $S_n = O_p(1)$ 。记 $b(u) = E[g(X, \mu_z) | Z = u]$, 由 μ_z 的最小性 $b(z) = 0$, 故

$b(z + hu) = h \partial b(z)[u] + O(h^2 \|u\|^2)$; 再由核的对称性与 Lipschitz 连续性, 线性项的主项积分为零、余项为 $O(h^2)$, 于是

$$E[g_i] = \int K(u) b(z + hu) f_Z(z + hu) du = O(h^{k+2}),$$

$$B_n = \frac{n}{\sqrt{nh^k}} O(h^{k+2}) = O\left((nh^{k+4})^{1/2}\right) \rightarrow 0.$$

这正是欠光滑条件的作用, 使核平滑带来的偏差在 $\sqrt{nh^k}$ 尺度上被完全压掉, 因此不需要任何偏差修正。

由核加权中心极限定理[16],

$$\frac{1}{\sqrt{nh^k}} \sum_i g_i \xrightarrow{d} N(0, V_z), \text{ 故 } \ell(\mu_z; z) \xrightarrow{d} \chi_m^2$$

□

注 1. 欠光滑条件 $nh^{k+4} \rightarrow 0$ 比带宽的最优阶 $h \approx n^{-1/(k+4)}$ 更严格。若只要求 $nh^{k+4} \rightarrow c \in (0, \infty)$, 则 B_n 不再消失, $\ell(\mu_z; z)$ 的极限变为非中心 $\chi_m^2 \left(\|c^{1/2} \delta_z\|_{V_z^{-1}}^2 \right)$, 其中 δ_z 由 $\partial b(z)$ 与核的二阶矩决定, 此时需要显式的偏差修正才能恢复 Wilks 定理。本文采用欠光滑路线, 因为它不必估计 ∂b , 实现最简单。

其中的欠光滑条件 $nh^{k+4} \rightarrow 0$ 用于压低核平滑所引入的偏差, 使其不影响一阶极限分布。于是对每个固定的 z , 逐点 $100(1-\alpha)\%$ 带为 $\{\mu: \ell(\mu; z) \leq \chi_{m, 1-\alpha}^2\}$, 即在切空间中给出一个以为中心的椭球状区域。

3.3. 有限网格同时带

逐点带虽然对每一个 z 都成立, 却不能同时覆盖整条曲线。若要在若干预测值上同时作出判断, 需把逐点结果组合起来。对有限网格 $\{z_c\}_{c=1}^N$ 有下述结果。

定理 2 (有限网格同时置信带)。在定理 1 的条件下,

$$\liminf_{n \rightarrow \infty} \mathbb{P} \left(\max_{1 \leq c \leq N} \ell(\mu_{z_c}; z_c) \leq \chi_{m, 1-\alpha/N}^2 \right) \geq 1 - \alpha. \quad (2)$$

证明. 由定理 1, 有 $\ell(\mu_{z_c}; z_c) \xrightarrow{d} \chi_m^2$, 故 $\forall \eta > 0$, $\limsup_n \mathbb{P} \left(\ell(\mu_{z_c}; z_c) > \chi_{m, 1-\alpha/N}^2 \right) \leq \alpha/N + \eta$.

由联合界,

$$\mathbb{P}\left(\max_{1 \leq c \leq N} \ell(\mu_{z_c}; z_c) > \chi_{m,1-\alpha/N}^2\right) \leq \sum_{c=1}^N \mathbb{P}\left(\ell(\mu_{z_c}; z_c) > \chi_{m,1-\alpha/N}^2\right) \leq \alpha + N\eta.$$

令 $\eta \rightarrow 0$ 取下极限即得(2).

□

定理 2 表明, $\bigcap_{c=1}^N \{\mu: \ell(\mu; z_c) \leq \chi_{m,1-\alpha/N}^2\}$ 构成了条件 Fréchet 均值 $\{\mu_{z_c}\}$ 的同时置信区域。这里有三点值得注意。第一, 证明只用到逐点收敛, 而没有用到网格上的经验过程一致 Donsker 条件, 这使得构造在理论上容易验证、在实现上也十分直接。第二, 代价是阈值由 $\chi_{m,1-\alpha/N}^2$ 放大到 $\chi_{m,1-\alpha/N}^2$, 因此当网格点较多时结果偏保守——这是联合界方法的固有代价。第三, 定理要求网格点数 N 固定, 关于网格随样本量增长的情形见注 2。

注 2. 定理 2 要求网格点数 N 固定这, 一点无法直接放松。若 $N = N_n \rightarrow \infty$, 则 $\chi_{m,1-\alpha/N_n}^2 \sim 2 \log(N_n/\alpha)$ 发散, 联合界给出的带虽然仍覆盖整条曲线, 宽度却随之发散而失去意义, 故必须改为直接研究网格上最大值 $\max_g \ell(\mu_{z_g}; z_g)$ 的分布。这需要三样工具: 一是核加权经验过程在网格上的一致逼近(即带宽收缩下的 Donsker 性质), 二是上确界分布的高斯逼近, 其误差随网格点数只以对数速度增长, 三是抗集中不等式, 用以把“分布收敛”提升为“自助分位数收敛”。相应的极限分布也不再是独立 χ_m^2 的最大值: 网格点弱相关时阈值约为 $2 \log N_n$, 而网格足够密、相邻点近似共线时有效独立点数退化为 $O(h^{-k})$, 阈值约为 $\chi_{m,1-\alpha/h^{-k}}^2$ 。这说明 Bonferroni 阈值的保守程度由网格点之间的相关性决定, 而乘子自助能自动适应这种相关性, 因而在网格加密时更值得推荐。本文不展开这一更精细的理论, 而给出有效而略显保守的有限网格带; 向连续预测带的推广是一个自然的后续方向。

3.4. 算法与调参

实践中可以用乘子自助阈值替代 Bonferroni 阈值, 以获得更紧的带宽。记 $K_i(z) = K((Z_i - z)/h)$, 具体步骤见算法 1。

Algorithm 1. Simultaneous confidence bands for the conditional Fréchet mean (multiplier bootstrap threshold)

算法 1. 条件 Fréchet 均值的同时置信带(乘子自助阈值)

输入: 样本: $\{(X_i, Z_i)\}_{i=1}^n$ 、网格 $\{z_g\}_{g=1}^G$ 、核 K 、带宽 h 、水平 α 、自助次数 B

输出: 同时带 $\{(z_g, \mu): \ell(\mu; z_g) \leq \hat{c}_{1-\alpha}\}$

1: 对每个 z_g , 计算 $\hat{\mu}_{z_g}$ 、统计量 $\ell(\hat{\mu}_{z_g}; z_g)$ 和 $\hat{V}_{z_g}$

2: for $b = 1$ to B

3: 抽取乘子 $\omega_i^{(b)}$, 计算 $\ell_n^{*(b)}(z_g)$, 并取 $s^{(b)} = \max_g \ell_n^{*(b)}(z_g)$

4: end for

5: 取 $\hat{c}_{1-\alpha} = \text{quantile}(\{s^{(b)}\}, 1-\alpha)$

6: 返回 $\{\mu: \ell(\mu; z_g) \leq \hat{c}_{1-\alpha}\}$

该算法的思路是: 先对每个网格点分别算出局部统计量, 再用乘子自助模拟这些统计量在网格上的最大值的分布, 取其分位数作为统一阈值。由于阈值同时作用于全部网格点, 这样得到的同时带通常比

Bonferroni 带更窄。

引理 1 (自助阈值的相合性). 设定理 1 的条件成立, $\{\omega_i\}$ 为与数据独立的乘子, $\mathbb{E}[\omega_i]=1$ 、 $\text{Var}(\omega_i)=1$ 、 $\mathbb{E}|\omega_i|^4 < \infty$, 且网格点数 G 固定。记 $W_z \sim N(0, V_z)$ 为逐点极限, 则**算法 1** 的阈值满足

$$\hat{c}_{1-\alpha} \xrightarrow{p} q_{1-\alpha} \left(\max_{1 \leq g \leq G} \|W_{z_g}\|_{V_z^{-1}}^2 \right),$$

并且

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(\max_{1 \leq g \leq G} \ell(\hat{\mu}_{z_g}; z_g) \leq \hat{c}_{1-\alpha} \right) = 1 - \alpha.$$

证明. 把定理 1 证明中的二次展开搬到自助版本(以 ω_i 替换常数权、以 $\hat{\mu}_{z_g}$ 替换 μ_{z_g}), 并把 $g(X_i, \hat{\mu}_{z_g})$ 在 μ_{z_g} 处展开: 由 $\hat{\mu}_{z_g} - \mu_{z_g} = O_p((nh^k)^{-1/2})$ 与 $\sum_i |\omega_i| K_i(z_g) = O_p(nh^k)$, 这一替换只带来 $o_p(1)$ 的差别。于是自助统计量的条件分布与

$\left(\frac{1}{\sqrt{nh^k}} \sum_i \omega_i K_i(z_g) g(X_i, \mu_{z_g}) \right)' V_{z_g}^{-1} \left(\frac{1}{\sqrt{nh^k}} \sum_i \omega_i K_i(z_g) g(X_i, \mu_{z_g}) \right)$, $g=1, \dots, G$ 的联合分布相同。视数据给定, $\{\omega_i - 1\}$ 为独立的中心化随机变量, 由定理 1 的矩条件满足 Lindeberg 条件, 故上式在给定数据下依概率收敛到 $\{W_{z_g}\}_{g=1}^G$ 的联合分布。由于 G 固定, 由连续映射定理, $\mathcal{L} \left(\max_g \ell_n^*(z_g) | \mathcal{D}_n \right) \xrightarrow{p} \mathcal{L} \left(\max_g \|W_{z_g}\|_{V_z^{-1}}^2 \right)$, 该极限的分布函数连续, 故 $\hat{c}_{1-\alpha}$ 依概率收敛到它的 $1-\alpha$ 分位数。另一方面, 由定理 2 与连续映射定理, $\max_g \ell(\hat{\mu}_{z_g}; z_g)$ 依分布收敛到同一个极限。两者相减即得覆盖概率收敛到 $1-\alpha$ 。证毕。

注 3. 引理 1 给出的是等式而非定理 2 中的下界。网格点之间的相关性使 $\max_g \|W_{z_g}\|_{V_z^{-1}}^2$ 的 $(1-\alpha)$ 分位数通常严格小于独立情形的 $\chi_{m, 1-\alpha/G}^2$, 这正是**算法 1** 比 Bonferroni 带更紧的原因。需要注意, 该结论依赖 G 固定: 当 $G \rightarrow \infty$ 时, 上确界分布的局部增量需要额外的抗集中条件才能控制, 分位数相合性不再自动成立。

关于调参, 有两点需要注意。其一是带宽: 逐点极限要求 $nh^k \rightarrow \infty$ 与 $nh^{k+4} \rightarrow 0$, 前者保证方差可控, 后者保证偏差可忽略; 实践中常取以在偏差 $O(h^{k+2})$ 与方差之间取得平衡。其二是网格: 应保证每个网格点的核支撑内都有不可忽略的样本量, 例如取间距与 h 同阶的正则网格。对同时带的阈值, 式(2)中的 Bonferroni 阈值是保守的, 而**算法 1** 给出的乘子自助阈值在实践中更紧(有关讨论见注 1)。

4. Fréchet 均值集的置信区域

4.1. 平稳集与水平集估计

前面的讨论都假定 Fréchet 均值唯一。然而当 Fréchet 函数的最小值点不唯一时, 总体的位置参数就不再是一个点, 而是均值集合 $\mathcal{E} = \arg \min_p F(p)$ 。此时逐点的置信区间不再适用, 需要为整个集合给出置信区域。

处理这一情形的一个自然思路是利用经验似然函数本身作为准则函数。记平稳集 $U = \{p: E[g(X, p)] = 0\}$, 它包含 $\mathcal{E}$; 在 U 上估计方程成立, 经验似然统计量 $\ell(p) = O_p(1)$; 而在之外, 由于估计方程不成立, $\ell(p) \rightarrow \infty$ 。这一对比说明, 经验似然函数的水平集恰好可以用来估计 U ; Kurisu 和 Otsu [13] 正是据此给出了相合的集合估计 $\hat{U}_n = \{p: \ell(p) \leq C \log n\}$ 。由于 $U \supseteq \mathcal{E}$, 对 U 的覆盖结论自动

覆盖均值集；当然，当二者的包含关系严格时，该区域并非 E 的精确区域，这一点在使用时需要留意。以下讨论水平集

$$C_{1-\alpha} = \{p \in K : \ell_n(p) \leq \hat{c}_{1-\alpha}\}, \quad (3)$$

其中 $\hat{c}_{1-\alpha}$ 为乘子自助上确界 $\hat{c}_{1-\alpha}$ 的条件 $1-\alpha$ 分位数。

4.2. 覆盖定理

为保证上述水平集具有名义覆盖，需要对总体与核函数作若干正则性假设。

假设 1. 设 $K \supseteq U$ 紧。(A1)类为 Donsker 且各阶矩连续有界；(A2)局部可识别：在 U 附近 $\|E[g(X, p)]\| \geq cd(p, U)^\beta$ (A3)水平集 $\hat{U}_n$ 依 Hausdorff 距离相合；(A4)高斯上确界律 T 在其 $1-\alpha$ 分位数处无原子；(A5)一致固定备择相合性：在远离 U 的紧集上 $\ell_n(p)/n \xrightarrow{P} D(p) > 0$ ；(A6)经验似然统计量的乘子自助一致展开。

其中(A1)保证经验似然函数作为随机过程具有良好的一致行为，(A2)保证平稳集可以被识别，(A3)保证水平集不会跑到平稳集之外，(A4)是分位数推断的常规技术条件，(A5)与(A6)分别控制区域在平稳集之外的扩张与自助校准的精度。更具体地说，(A1)是经验过程的常规条件，可用函数类的复杂度或逐点 Lipschitz 条件验证；(A2)刻画估计方程在平稳集附近的分离速度， $\beta=1$ 对应正则情形，Hessian 退化时 β 相应变大；(A4)在极限高斯过程轨道连续、协方差非退化时自动成立；(A6)是唯一直接涉及自助的条件，其可验证的版本即下文算法中的自助相合性。

定理 3 (均值集覆盖). 在假设 1 下，式(3)的水平集满足单侧覆盖

$$\liminf_{n \rightarrow \infty} \mathbb{P}(U \subseteq C_{1-\alpha}) \geq 1 - \alpha. \quad (4)$$

证明. 记 $\lambda(p) = E[g(X, p)]$ 、 $\bar{\ell}_n = \ell_n/n$ ，并记 $W(p)$ 为 $\sqrt{n} \left\{ \bar{\ell}_n(p) - \lambda(p)' V_p^{-1} \lambda(p) \right\}$ 的极限高斯过程， $T = \sup_{p \in U} \|W(p)\|_{V_p^{-1}}^2$ ；由假设 1 的(A1)， W 存在且的分布连续可微。由假设 1 的(A6)， ℓ_n^* 与 ℓ_n 在 $\hat{U}_n$ 上同模连续，故

$$\hat{c}_{1-\alpha} - q_{1-\alpha}(T) = o_p(1)$$

由假设 1 的(A2)与(A3)， $\|\lambda(p)\| \geq cd(p, U)^\beta$ 且 $d_H(\hat{U}_n, U) = o_p(1)$ ；由假设 1 的(A5)，对有 $\bar{\ell}_n(p) \geq D(p)(1 + o_p(1))$ 而 $\hat{c}_{1-\alpha}/n = o_p(1)$ ，故

$$\sup_{p \in \hat{U}_n} \ell_n^*(p) = \sup_{p \in U} \ell_n^*(p) + o_p(1)$$

由假设 1 的(A4)， $\mathbb{P}(T \leq q_{1-\alpha}(T)) = 1 - \alpha$ ；在 U 上 $E[g] = 0$ 故 $\ell_n(p) = O_p(1)$ ，于是 $\{U \subseteq C_{1-\alpha}\} = \{T \leq \hat{c}_{1-\alpha}\}$ 至 $o_p(1)$ 。综合即得(4)。□

注 4. 定理 3 给出的是单侧覆盖，即平稳集被区域包含的概率。与此相关的另一个问题是区域是否会超出平稳集。事实上，在 K 中任一与分离的紧子集上， $\ell_n(p)/n \rightarrow \lambda(p)' V_p^{-1} \lambda(p) > 0$ 而 $\hat{c}_{1-\alpha} = O_p(1)$ ，故 $C_{1-\alpha}$ 落入该子集的概率趋于零。由此可知，当水平满足时 $d_H(C_{1-\alpha}, U) \xrightarrow{P} 0$ ；而对固定的 $\alpha \in (0, 1)$ ，式(4)的单侧覆盖才是恰当的陈述，此时不宜宣称 Hausdorff 收敛。

5. 数值研究

本节通过两个数值实验考察前述构造的有限样本表现。所有实验在 S^2 上进行, 随机数由固定种子生成。

5.1. 同时带的校准

第一个实验考察有限网格同时带的覆盖。在 S^2 上取如下设定:

$$Z \in [0, 1], \quad \mu_z = (\sin(0.4 + 1.2z), 0, \cos(0.4 + 1.2z)), \quad X_i | Z_i \sim \text{vMF}(\mu_{Z_i}, \kappa)$$

核函数取 Epanechnikov 核, $m = 2$ 。表 tab:cov 报告了有限网格上 $\max_c \ell(\mu_z; z_c)$ 分别被逐点阈值、Bonferroni 阈值与乘子自助阈值覆盖的频率, 名义水平为 0.95, 每行的重复次数为 200。

Table 1. Simultaneous coverage of local empirical likelihood over a finite grid (nominal 0.95)

表 1. 有限网格上局部经验似然的同时覆盖(名义 0.95)

设定	自助	逐点	Bonferroni	配对自助
$\kappa = 2, n = 600, h = 0.25$	0.96	0.85	0.97	0.94
$\kappa = 5, n = 600, h = 0.25$	0.95	0.87	0.96	0.95
$\kappa = 2, n = 600, h = 0.35$	0.97	0.85	0.98	0.94

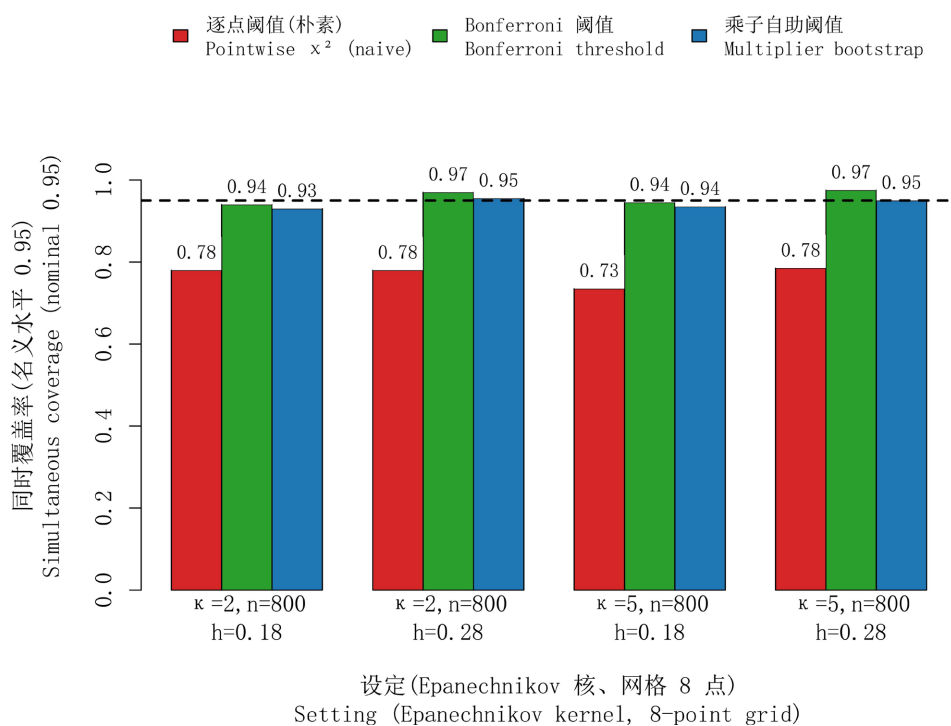


Figure 1. Simultaneous coverage of local empirical likelihood on a finite grid (nominal 0.95). Red represents the point by point threshold, green represents the Bonferroni threshold, and blue represents the multiplier bootstrap threshold; The horizontal dashed line represents the nominal level

图 1. 有限网格上局部经验似然的同时覆盖率(名义 0.95)。红色为逐点阈值, 绿色为 Bonferroni 阈值, 蓝色为乘子自助阈值; 水平虚线为名义水平

由表 1 和图 1 可以看出三点。第一, 若直接沿用逐点阈值, 则整条曲线的实际覆盖率只有 0.85 至

0.87, 明显低于名义水平; 这正说明逐点带不能替代同时带。第二, Bonferroni 阈值的覆盖率为 0.96 至 0.98, 略高于名义水平, 与定理 2 所预言的保守性一致。第三, 乘子自助阈值最接近名义水平, 覆盖率在 0.94 至 0.97 之间, 说明算法 1 在实践中确能给出更紧的同时带。

5.2. 均值集的覆盖

第二个实验考察均值集置信区域的表现。取 S^2 上 $\text{vMF}(N)$ 与 $\text{vMF}(S)$ 的等权混合, 可以验证其总体 Fréchet 均值集恰为赤道——因为两个分量关于赤道对称, 且赤道上的任意一点都使 Fréchet 函数取到相同的全局最小值。由于该集合包含无穷多个点, 它是检验集合型推断的合适例子。

数值结果显示, 乘子自助区域在几乎全部重复中都完整地包含了整条赤道, 覆盖率不低于 0.97, 与定理 3 的结论相符; 相比之下, 若改用朴素的 χ^2 阈值, 则在部分设定下无法覆盖整个均值集。图 2 的左侧给出了一次典型重复中的置信区域, 右侧给出两种阈值下覆盖率随重复次数的变化。

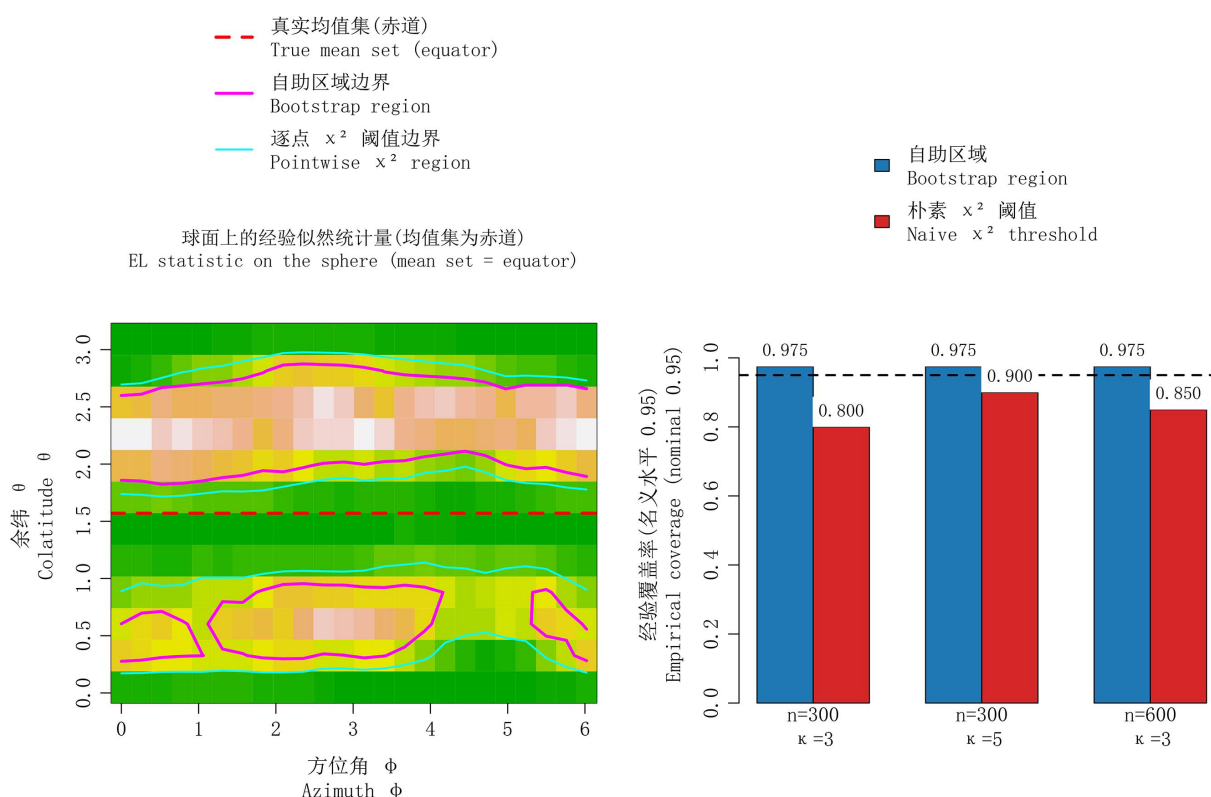


Figure 2. The confidence region of the non unique Fréchet mean set (equatorial). Left: The contour map of the empirical likelihood statistic $\ell(\mu)$ on the sphere, with the horizontal axis representing azimuth ϕ and the vertical axis representing colatitude θ . The more green the color, the smaller θ is.

图 2. 非唯一 Fréchet 均值集(赤道)的置信区域。左: 球面上经验似然统计量 $\ell(\mu)$ 的等值图, 横轴为方位角 ϕ 、纵轴为余纬 θ , 颜色越偏绿表示 θ 越小

图 2 中红色虚线为真实均值集即赤道。经验似然函数以平稳集为准则, 而 U 除赤道外还包含两极(两极位于旋转对称轴上, 故 Fréchet 函数在该处的梯度必为零), 于是图中赤道与两极附近都呈绿色带; 洋红色等值线为乘子自助阈值所对应的区域边界, 它完整地包含整条赤道, 青色等值线为逐点阈值所对应的边界, 范围偏小。右: 两种阈值对整条均值集的覆盖率随 (n, κ) 的变化。

6. 结论

本文在流形经验似然框架下处理了两类集合型推断问题。对于条件 Fréchet 均值, 本文给出了有限预测值网格上的同时置信带: 它由逐点极限配合联合界得到, 构造简单、理论条件弱, 不需要带宽收缩下的经验过程 Donsker 条件; 文中同时给出了一个乘子自助的实用阈值, 并说明了该阈值渐近正确、且比 Bonferroni 阈值更紧的理由(引理 1), 数值结果显示它比 Bonferroni 阈值更接近名义水平。对于 Fréchet 均值集, 本文以经验似然函数作为准则函数构造水平集估计, 并用乘子自助校准其覆盖概率, 同时刻画了区域在平稳集之外的超出行为。数值实验与理论分析的结果一致。后续工作可以在两个方向上进一步展开。一是连续预测带, 即让网格随带宽一同增长。此时每点的水平趋于零, 需要一致的中偏差逼近或者相关高斯过程上确界的极值理论, 这是把本文的有限网格结果推广到连续情形的关键。二是把上述构造推广到条件中位数(即 $q=1$ 的情形), 这需要建立相应的局部中位数理论。

参考文献

- [1] Marron, J.S. and Alonso, A.M. (2014) Overview of Object Oriented Data Analysis. *Biometrical Journal*, **56**, 732-753. <https://doi.org/10.1002/bimj.201300072>
- [2] Patrangenaru, V. and Ellingson, L. (2015) Nonparametric Statistics on Manifolds and Their Applications to Object Data Analysis. CRC Press. <https://doi.org/10.1201/b18969>
- [3] Bhattacharya, R. and Patrangenaru, V. (2003) Large Sample Theory of Intrinsic and Extrinsic Sample Means on Manifolds. *The Annals of Statistics*, **31**, 1-29. <https://doi.org/10.1214/aos/1046294456>
- [4] Bhattacharya, R. and Patrangenaru, V. (2005) Large Sample Theory of Intrinsic and Extrinsic Sample Means on Manifolds—II. *The Annals of Statistics*, **33**, 1225-1259. <https://doi.org/10.1214/009053605000000093>
- [5] Bhattacharya, R. and Lin, L. (2017) Omnibus CLTs for Fréchet Means and Nonparametric Inference on Non-Euclidean Spaces. *Proceedings of the American Mathematical Society*, **145**, 413-428. <https://doi.org/10.1090/proc/13216>
- [6] Eltzner, B. and Huckemann, S.F. (2019) A Smearly Central Limit Theorem for Manifolds with Application to High-Dimensional Spheres. *The Annals of Statistics*, **47**, 3360-3381. <https://doi.org/10.1214/18-aos1781>
- [7] Hundrieser, S., Eltzner, B. and Huckemann, S. (2024) Finite Sample Smeariness of Fréchet Means with Application to Climate. *Electronic Journal of Statistics*, **18**, 3274-3309. <https://doi.org/10.1214/24-ejs2276>
- [8] Owen, A.B. (1988) Empirical Likelihood Ratio Confidence Intervals for a Single Functional. *Biometrika*, **75**, 237-249. <https://doi.org/10.1093/biomet/75.2.237>
- [9] Owen, A.B. (2001) Empirical Likelihood. Chapman & Hall/CRC.
- [10] Qin, J. and Lawless, J. (1994) Empirical Likelihood and General Estimating Equations. *The Annals of Statistics*, **22**, 300-325. <https://doi.org/10.1214/aos/1176325370>
- [11] Hjort, N.L., McKeague, I.W. and Van Keilegom, I. (2009) Extending the Scope of Empirical Likelihood. *The Annals of Statistics*, **37**, 1079-1111. <https://doi.org/10.1214/07-aos555>
- [12] Chen, S.X. and Van Keilegom, I. (2009) A Review on Empirical Likelihood Methods for Regression. *TEST*, **18**, 415-447. <https://doi.org/10.1007/s11749-009-0159-5>
- [13] Kurisu, D. and Otsu, T. (2026) Empirical Likelihood for Manifolds. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, **88**, 91-119. <https://doi.org/10.1093/rssb/qkaf043>
- [14] Bharath, K., Le, H., Yan, X. and Wood, A.T.A. (2024) Empirical Likelihood for Frechet Means on Open Books. arXiv: 2412.18818.
- [15] Petersen, A. and Müller, H. (2019) Fréchet Regression for Random Objects with Euclidean Predictors. *The Annals of Statistics*, **47**, 691-719. <https://doi.org/10.1214/17-aos1624>
- [16] Chen, Y. and Müller, H. (2022) Uniform Convergence of Local Fréchet Regression with Applications to Locating Extrema and Time Warping for Metric Space Valued Trajectories. *The Annals of Statistics*, **50**, 1573-1592. <https://doi.org/10.1214/21-aos2163>
- [17] Schötz, C. (2022) Strong Laws of Large Numbers for Generalizations of Fréchet Mean Sets. *Statistics*, **56**, 34-52. <https://doi.org/10.1080/02331888.2022.2032063>
- [18] Evans, S.N. and Jaffe, A.Q. (2024) Limit Theorems for Fréchet Mean Sets. *Bernoulli*, **30**, 419-447. <https://doi.org/10.3150/23-bej1603>

- [19] Blanchard, M. and Jaffe, A.Q. (2025) Fréchet Mean Set Estimation in the Hausdorff Metric, via Relaxation. *Bernoulli*, **31**, 432-456. <https://doi.org/10.3150/24-bej1734>
- [20] Chernozhukov, V., Hong, H. and Tamer, E. (2007) Estimation and Confidence Regions for Parameter Sets in Econometric Models. *Econometrica*, **75**, 1243-1284. <https://doi.org/10.1111/j.1468-0262.2007.00794.x>