arXiv:2602.04010v1 [stat.ME] 3 Feb 2026

基于扩展布雷格曼散度与互信息的稳健非参数双样本检验

Arijit Pyne 诺华公司，印度
海得拉巴
arijitpynestat@gmail.com

文章历程
编制于2026年2月5日

摘要
本文基于扩展布雷格曼散度提出互信息（MI）的一种广义形式；这一框架涵盖了广义S-Bregman（GSB）散度族。GSB散度统一了两类重要的统计距离，即S-散度和Bregman 指数散度（BED），由此涵盖了多个广泛使用的子族，包括幂散度（PD）、密度幂散度（DPD）和S-Hellinger 距离（S-HD）。在参数推断中，最小散度估计量以能够在稳健性与相对于最大似然估计量的高渐近效率之间取得平衡而著称。然而，基于此类统计距离的非参数检验相对而言研究较少。本文利用广义互信息构建了一类一致且稳健的非参数双样本检验，用于检验两个绝对连续分布是否相等。我们证明了所提出的检验统计量在原假设和邻近备择下的渐近正态性。我们通过影响函数和崩溃点严格研究广义互信息的稳健性，并表明广义互信息的稳定性会转化为相关检验的稳定性。大量模拟研究表明，超出PD族的散度在存在污染时往往具有更优的稳健性，同时保持较高的渐近功效。本文还提出了一种数据驱动方案，用于选择最优调节参数。最后，通过真实数据应用展示了该方法。

关键词
扩展布雷格曼散度、互信息、非参数双样本检验、相合性、稳健性、影响函数、渐近崩溃点

1. 引言

自二十世纪初以来，研究人员一直运用统计方法分析不同学科领域的科学问题。统计推断构成了统计方法的核心，并在发展过程中逐渐演变为如今的形式。此后，基于似然的方法在后续的每一次发展中都占据了核心地位。最大似然估计量（MLE）的早期应用可追溯至高斯、拉普拉斯等人的时代。然而，最大似然估计量在统计推断中的广泛重要性，主要是在 R. A. Fisher 爵士的推动下才逐渐显现；他在统计估计理论方面的开创性研究

为后续的所有突破奠定了坚实基础。在参数推断中，使用似然函数之前，唯一需要做的就是为所研究的数据集所依据的概率分布选择一个“恰当的”数学模型，并以未知参数（可能是一个向量）作为索引。随后，通过最大化似然函数来估计模型参数，并进一步利用该函数检验统计假设。最大化似然函数，也可以等价地通过最小化模型与数据之间的似然差异（LD）来实现。随着最大似然估计量的兴起，似然比检验（LRT）、沃尔德检验和得分检验等多种检验统计假设的方法也广受认可。参数模型在现实情境中往往难以构建；在此类情况下，非参数方法一次又一次地为统计学家解围。

基于似然的推断方法广受欢迎，主要是因为当数据确实由某个参数模型生成时，在某些正则条件下，这类方法具有完全的渐近最优性。然而，在大多数现实情境中，这些条件很少得到满足。在这种情况下，即使参数模型只有轻微的模型错设，基于似然的方法通常也会变得高度敏感。随着我们从日益多样化的研究来源中获取数据，统计方法的稳健性问题也愈发重要。这一需求推动了稳健推断方法的发展，并为M估计理论的发展奠定了基础（参见 Huber（2011） [1]; Hampel 等（1986） [2]）。尽管最大似然估计量（MLE）属于M估计量这一类，但许多其他著名的稳健替代方法也属于此类。稳健性提高往往要以渐近效率为代价。在最小距离估计方法中，先构造数据与模型之间的统计散度，再通过最小化该散度来选择估计量。最小距离估计量构成了M估计量中的一个重要类别。随后，人们基于最小距离估计量构造散度偏差检验（DDT）。由于在模型成立时具有较高（有时甚至完全）的渐近效率，同时还具备针对模型错设的强稳健性，最小距离方法在统计学文献中得到了广泛应用。

尽管卡尔·皮尔逊是首位通过引入 χ^2 度量来拟合参数模型而使用统计距离的人，但这一方法花了很长时间才被主流研究界接受，直到 70 年代末，Beran (1977) [3, 4] 提出了一种稳健且有效的估计量，即数据与参数模型之间海林格距离的最小化量。这类估计量称为最小海林格距离估计量。海林格距离（HD）是著名的幂散度（PD）族的一员，该族由 Cressie 和 Read (1984) [5]提出。然而，幂散度族属于更广义的距离类，即 Csiszár (1963) [6]; Ali 和 Silvey (1966) [7]提出的 ϕ-散度。ϕ-散度族的所有成员在真实模型下都能实现完全渐近有效的推断，其中一些成员还具有额外的稳健性性质。Bregman 散度（参见 Bregman (1967) [8]）是另一类重要的统计距离，其中包括 Basu 等人 (1998) [9]提出的密度幂散度（DPD）。后者可以说是 Bregman 散度类中最重要的一族，因为即使对于连续模型，它也无需进行核平滑即可估计参数。此外，它还在 LD（似然差异，可导出 MLE）与平方 L_2-距离（可产生非常稳健但效率略低的估计量）之间实现了平滑衔接。Mukherjee 等人 (2019)[10] 提出的 Bregman 指数散度（BED）构成了 Bregman 散度类中的另一重要族。显然，除似然差异外，ϕ-散度族的其他成员都不属于 Bregman 散度类。然而，Ghosh 等人 (2017) [11] 定义了 S-散度族，它将幂散度族与密度幂散度族平滑地衔接起来。

最后，Basak 和 Basu (2022) 提出的广义 S-Bregman 散度 [12] 通过扩展现称为扩展布雷格曼散度的概念，将 S-散度与 BED 族统一起来。关于最小距离统计推断的更多细节，请参见 Pardo (2006) [13] 以及 Basu 等人 (2011) [14] 。

如前所述，在现实情境中，往往无法建立参数模型。在这种情况下，最好使用非参数模型，因为它只要求满足最低限度的统计假设。经验似然在非参数推断领域的应用，主要因 Owen 的研究而广受关注。关于这一方法的精彩阐述，参见 Owen（2001） [15] 。此后，经验似然已拓展至 ϕ散度族；参见 Morales 等人（2001） [16]；Balakrishnan 等人（2015，2017） [17, 18]。Micheas 和 Zografos（2006） [19] 使用ϕ散度来衡量一组随机变量之间的随机依赖。

在双样本检验问题中，基于经验似然函数的 ϕ散度已被广泛应用，例如 Jing（1995） [20]; Qin 和 Zhao（2000） [21]；Liu 等人（2008） [22]；Wu 和 Yan（2012） [23]; Balakrishnan 等人（2017） [18], 等。在这些早期论文中，大多数作者考虑的是某些汇总统计量（如均值等）的相等性原假设。Guha 和 Chothia（2014） [24] 使用互信息（似然差异），基于核密度估计检验两个完全非结构化绝对连续分布是否相等。当底层分布为连续分布时，有若干知名的非参数检验可供使用，其中许多基于经验分布函数。常见的例子包括 Kolmogorov–Smirnov（KS）检验、威尔科克森检验、Anderson–Darling（AD）检验以及 Cram´ér–von Mises（CVM）检验。双样本 t 检验和威尔科克森检验常用于检测位置差异。对于位置相同但形状、偏度或峰度不同的分布，t 检验和威尔科克森检验的效果较差。尽管 AD、CVM 和 KS 检验在这些情况下表现尚可，Guha 和 Chothia（2014） [24] 证明，基于互信息（MI）的检验往往优于这些检验。Guha 等（2021） [25] 使用 ϕ散度，特别是幂散度族，提高了双样本非参数检验的稳健性。然而，我们进一步发现，基于幂散度族的经验互信息的渐近零分布要求连续随机变量的支撑集有界。这一严格要求在某种程度上严重妨碍了幂散度族的渐近零分布应用于更一般的情形，即连续随机变量的支撑集无界。为克服这一局限，Guha 等（2021）[25] 建议使用置换检验，但这种方法的计算成本相当高。

本文提出一种基于 ϕ 生成的扩展布雷格曼散度的双样本非参数检验，用于检验两个完全非结构化连续分布之间的相等性。本文的主要贡献概述如下。

(a) 首先，我们提出一种基于扩展 Bregman 散度，特别是广义 S-Bregman 散度的互信息一般定义。不言而喻，这一定义将互信息的应用范围扩展到了幂散度族以外的许多其他族。(b) 我们提出一类基于广义互信息的相合非参数检验，用于检验两个完全非结构化绝对连续分布之间的相等性。(c) 我们证明了原假设及其邻近备择下检验统计量的渐近正态性。GSB族中的成员除 $\alpha = 0$ 外，在使用渐近

零分布。

(d) 通过影响函数研究广义互信息泛函在某一点处的无穷小稳定性行为。此外，还计算了检验泛函的水平影响函数(LIF)和功效影响函数(PIF)。(e) 在一定的正则条件下，我们计算广义互信息泛函的渐近崩溃点。(f) 模拟结果从经验上表明，与GSB族的其他成员相比，基于幂散度族的检验在数据污染程度较高时通常稳定性较差。不过，除在 $\beta < 0$的情况下有时例外，所有这些检验都具有较高的渐近功效。尽管如此，进一步研究发现，对于 $\beta < 0$，收敛至渐近正态性的速度有时可能会变慢。(g) 我们提出了一种算法，可根据给定的数据集选择最优调节参数。我们还将这一方案应用于几个真实数据示例。

下文将介绍本文的结构。第 2 节，具体而言，第 2.1 小节，介绍由 ϕ生成的扩展布雷格曼散度及其最受欢迎的子族之一，即广义S-Bregman（GSB）散度。第 2.2 小节则基于由 ϕ生成的扩展布雷格曼散度，定义互信息（$\mathscr{B}$-MI）。尽管这一推广并未涵盖 ϕ-散度族的所有成员，但包含其中研究最为充分的一族，即幂散度（PD）族。此外，我们还讨论了它的一些重要性质。第 3 节从广义互信息的角度阐述双样本检验问题。我们讨论了利用核密度估计得到的广义互信息非参数估计量，并在第 3.1 小节研究其渐近性质。在第 3.2 小节，我们证明了原假设下检验统计量（广义互信息的标准化版本）的渐近正态性。第 3.3 小节研究了邻近备择下检验统计量的相合性与功效。第 4 节研究检验统计量的稳健性行为。特别是，我们在第 4.1 和 4.2 小节计算了检验泛函的影响函数（IF），以及其水平影响函数和功效影响函数。第 4.3 小节计算了互信息泛函的渐近崩溃点。第 5 节给出了大量模拟结果。我们在第 5.2 小节讨论调节参数的选择，并在第 5.3 小节将其应用于几个真实数据示例。最后，第 6 节总结全文。所有理论结果的证明均列于附录 7。

2. 基于扩展Bregman散度的广义互信息

2.1. 扩展*Bregman*散度类

由严格凸且可微函数 ϕ :S →R 生成的两点 $x, y \in$ S 之间的 Bregman 散度（参见 Bregman (1967) [8]），定义为 $\phi(x)$ 与其在 y 处的一阶泰勒级数近似之差

$$D_\phi(x, y) = \phi(x) - \phi(y) - \langle \phi'(y), x - y \rangle, \tag{2.1}$$

其中， $\phi'(y)$ 是在 y 处计算的 ϕ梯度，S 是 R^p 的一个凸子集。 ϕ 的凸性确保 D_ϕ 非负，并且 $D_\phi(x, y) = 0$ 当且仅当 $x = y$。Bregman 散度

与度量相似，但通常既不满足三角不等式，也不满足对称性。不过，它们满足勾股定理的一种推广形式。

当将点视为概率密度函数 g，f（定义在共同支撑集χ 上，并相对于共同的 σ-有限测度 μ）时，所得的 Bregman 散度为

$$D_\phi(g,f)=\int_\chi\Big\{\phi(g)-\phi(f)-(g-f)\phi'(f)\Big\}d\mu. \tag{2.2}$$

稍后，在上下文清楚时，符号 χ，μ 将省略不写。似然差异 (LD) 和平方 L_2-距离是属于此类的两种著名散度，它们分别由 $\phi(t)=t\log t$ 和 $\phi(t)=t^2$生成。

此类中的一个重要族是 Basu 等人 (1998年) 提出的密度幂散度 (DPD)[9]。它由 $\phi(t)=\frac{t^{1+\alpha}-t}{\alpha}$ 生成，其中 $\alpha>0$。该散度的形式如下：

$$d_\alpha(g,f)=\int\Big\{f^{1+\alpha}-\Big(1+\frac{1}{\alpha}\Big)f^\alpha g+\frac{1}{\alpha}g^{1+\alpha}\Big\}\text{ for }0<\alpha\le 1. \tag{2.3}$$

$\alpha=0$处的 DPD 定义为 $d_0(g,f)=\lim_{\alpha\downarrow 0+}d_\alpha(g,f)=\int g\ln\frac{g}{f}$。当 $\alpha=1$时，它就成为平方 L_2-距离：$d_1(g,f)=\int(g-f)^2$。因此，DPD 平滑地连接了两种著名的统计距离。在参数估计中，参数模型 $\{f_\theta;\theta\in\Theta\}$ 和经验密度 g_n 分别替代 f 和 g。已知最小密度幂散度估计量 (MDPDE) 在 α取值较小时具有很高的效率，并且其稳健性会随着调节参数 α 增大而提高。

Mukherjee 等人（2019年） [10] 定义了 Bregman 散度类中另一个重要的子族，称为 Bregman 指数散度（也称为 B-指数散度或 BED），其定义为

$$BED_\beta(g,f)=\frac{2}{\beta}\int\Big\{e^{\beta f}\Big(f-\frac{1}{\beta}\Big)-e^{\beta f}g+\frac{e^{\beta g}}{\beta}\Big\}\text{ for }\beta\ne 0. \tag{2.4}$$

该族由 $\phi(t)=\frac{2}{(e^{\beta t}-\beta t-1)\ \beta^2}$ 生成，其中 $\beta\ne 0$。当 $\beta=0$ 时，BED 定义为$BED_0(g,f)=\lim_{\beta\to 0}BED_\beta(g,f)$，即平方 L_2-距离。

另一方面，ϕ-散度是另一类重要的距离，其形式如下：

$$d_\phi(g,f)=\int\phi\Big(\frac{g}{f}\Big)f, \tag{2.5}$$

其中凸函数 ϕ 还需额外满足以下条件：$\phi(1)=\phi'(1)=0$，$0\phi(0/0)=0$，以及 $0\phi(f/0)=f\lim_{u\to+\infty}\frac{\phi(u)}{u}$。该类包含著名的幂散度族，其定义为

$$PD_\lambda(g,f)=\frac{1}{\lambda(1+\lambda)}\int g\Big\{\Big(\frac{g}{f}\Big)^\lambda-1\Big\}\text{ for }\lambda\ne -1,0. \tag{2.6}$$

在 $\lambda = -1$，0处的散度定义为 $PD_\lambda(g,\ f)$在 $\lambda \to -1$，0时的连续极限。LD是唯一同时属于Bregman散度类与 ϕ-散度类的公共成员。

Ghosh 等人（2017） [11] 定义了 S-散度，将幂散度族与密度幂散度整合到一个统一框架中。S-散度族涉及两个调节参数 $\alpha \in [0,1]$ 和 $\lambda \in \mathbb{R}$，其定义如下

$$SD_{\alpha,\lambda}(g,f) = \int \left\{ \frac{1}{B}(g^{A+B} - f^{A+B}) - (g^A - f^A)\frac{A+B}{AB} f^B \right\} \text{ for } AB \neq 0, \tag{2.7}$$

其中 $A = 1 + \lambda(1-\alpha)$，$B = \alpha - \lambda(1-\alpha)$。S-散度可由（2.2）通过$\phi(t) = \frac{t^{1+\frac{B}{A}}}{B}$ 生成，其中 $AB \neq 0$。在 $A = 0$ 和/或 $B = 0$ 处的散度定义为当 $A \to 0$和 $B \to 0$时的连续极限。当 $\lambda = 0$ 时，（2.7）中的 S-散度 $SD_{\alpha,\lambda}(g,f)$ 变为（2.3）中的密度幂散度 $d_\alpha(g,f)$；而在 $\alpha = 0$处，它变为（2.6）中的幂散度PD_λ 。S-散度族的一个重要子族是 S-海林格距离，其定义如下

$$SD_{\alpha,\lambda=-0.5}(g,f) = \frac{2}{1+\alpha}\int \left(g^{\frac{1+\alpha}{2}} - f^{\frac{1+\alpha}{2}} \right)^2 \text{ for } \alpha \in [0,1]. \tag{2.8}$$

尽管 S-散度族包含了 Bregman 散度的若干成员，但并未涵盖全部成员。不过，Basak 和 Basu (2022) [12] 通过以下扩展形式实现了这一点:

$$D_\phi^{(k)}(g,f) = \int \left\{ \phi(g^k) - \phi(f^k) - (g^k - f^k)\phi'(f^k) \right\} \text{ for some index } k > 0. \tag{2.9}$$

这称为由 ϕ生成的、具有正指标 k的扩展布雷格曼散度。在本文的讨论中，我们假设凸函数 ϕ 四次连续可微。Basak 和 Basu (2022) [12] 进一步通过选取以下凸函数定义广义 S-Bregman 散度 (GSB):

$$\phi(t) = e^{\beta t} + \frac{t^{+\frac{B}{A}}}{B} \text{ for } AB \neq 0, \tag{2.10}$$

其中 A，B如（2.7）中所定义，其中 $\alpha \geq -1$和 λ，$\beta \in \mathbb{R}$。式（ ϕ（2.10）中的函数与指数 $k = A > 0$共同生成以下形式的扩展布雷格曼散度:

$$D^*(g,f) = \int \left\{ e^{\beta f^A}(\beta f^A - \beta g^A - 1) + e^{\beta g^A} + \frac{1}{B}(g^{A+B} - f^{A+B}) - (g^A - f^A)\frac{A+B}{AB} f^B \right\} \tag{2.11}$$

当 $A, B \neq 0$ 时，GSB 散度在 $A = 0$ 和/或 $B = 0$ 处定义为式（2.11）中表达式在 $A, B \to 0$ 时的连续极限。注意，当 $\beta = 0$ 时，GSB 散度变为 S-散度。对于 $BED_\beta^s(g,f) = \frac{\beta^2}{2} \times BED_\beta(g,f)$，可为 $\beta \neq 0$生成一个缩放 BED 族$\alpha = -1, \lambda = 0$。GSB 散度不仅将 Bregman 散度的所有成员纳入幂散度族，还在参数推断中产生了许多其他稳定且高效的估计量，这些估计量不属于 BED 族或 S-散度族。

2.2. 广义互信息及其性质

定义 2.1。 （$\mathscr{B}$-互信息）基于由 ϕ 生成的扩展 *Bregman* 散度的互信息，对于一组随机变量 $\{X_1, \ldots, X_m\}$ 可以定义为映射 $I_{D_\phi^{(k)}} : \otimes_{i=1}^m \chi_i \longrightarrow [0, +\infty]$ ，使得

$$I_{D_\phi^{(k)}}(X_1, X_2, \ldots, X_m) = D_\phi^{(k)}(f_{X_1, X_2, \ldots, X_m}, f_{X_1} \cdots f_{X_m}), \tag{2.12}$$

其中， f 表示所指随机变量的密度， χ_i 为 X_i、 $i = 1$、2、.....、 m的样本空间。

注意，对于任意一组随机变量，式（2.12）在零测集意义下是唯一确定的。 $I_{D_\phi^{(k)}}$ 可解释为一组随机变量之间随机关联程度的度量$\{X_1$、.....、 $X_m\}$。将 ϕ函数（2.10）代入式（2.12），即可得到基于GSB散度的互信息显式形式。此外，当 $\alpha = \beta = 0$， $I_{D_\phi^{(k)}}$ 变为 I_λGuha等人（2021）提出的 [25]。类似地，对于 $\alpha \geq 0$ 且 $\lambda = \beta = 0$，可得到基于DPD（I_α）的互信息。对于其他若干散度，选择不同的 ϕ函数，也可以类似地得到相应的互信息表达式。

幂散度是同时属于 ϕ-散度和ϕ-生成的扩展布雷格曼散度族的唯一共同成员。这些共同成员记作$C_{B\phi} = \{\phi$对于某个 $k > 0 : d_\phi(g, f) = D_\phi^{(k)}(g, f) \forall f, g\}$。Micheas & Zografos（2006） [19] 利用随机变量的联合密度与其边缘分布乘积之间的 ϕ-散度，研究多元数据中的依赖性度量。

基于 ϕ-生成的扩展布雷格曼散度的互信息也具有类似性质。首先，我们定义 $\phi_0 = \phi(0)$、 $\phi_2 = \phi_0 + \lim_{u \to +\infty} {}^{\phi(u)}_{u}$ 以及 $c_L = \int_{f>0} {}^{k^2 f^{2k} \phi''}_{\phi''\ 1}(f^k)\ ()$，其中 $f = f_{X_1} f_{X_2}, \ldots, f_{X_m}$。通常，~~$\phi$-~~散度会进行标准化，使得 $\phi''(1) = 1$。广义互信息 $I_{D_\phi^{(k)}}$ 的一些数学性质将在以下命题中列出。

命题 2.1。 基于 ϕ 生成的扩展布雷格曼散度的互信息具有以下性质。

(P1) $I_{D_\phi^{(k)}}(X_1, \ldots, X_m) = 0$ 当且仅当 $f_{X_1, X_2, \ldots, X_m} = f_{X_1} f_{X_2} \cdots f_{X_m}$ 几乎处处成立。*(P2)* 对于 $(X_1, X_2, \ldots, X_m)$ 的任意排列 $(X_{i_1}, X_{i_2}, \ldots, X_{i_m})$， $I_{D_\phi^{(k)}}(X_1, \ldots, X_m) = I_{D_\phi^{(k)}}(X_{i_1}, X_{i_2}, \ldots, X_{i_m})$。*(P3)* 对于任意一一映上的变换 $(X_1, \ldots, X_m) \mapsto (T_1(X_1)\ T_2(X_2) \ldots, T_m(X_m))$， $I_{D_\phi^{(k)}}(T_1(X_1) \ldots, T_m(X_m)) = I_{D_\phi^{(k)}}(X_1, \ldots, X_m)$。因此， $I_{D_\phi^{(k)}}$ 在严格递增变换下保持不变。*(P4)* 设 $\phi_2 = \infty$ ，且联合分布相对于边缘分布的乘积是奇异的，则$I_{D_\phi^{(k)}}(X_1, \ldots, X_m) = \infty$。*(P5)* 假设 ϕ 是严格凸且二次连续可微的，并且满足 $\phi(1) = \phi'(1) = 0$

且 $\phi_2, c_L < \infty$。那么，对于所有 $\phi \notin C_{B\phi}$， $I_{D_\phi^{(k)}}(X_1, \ldots, X_m) = c_L\phi_2$ 当且仅当联合分布相对于边缘分布的乘积是奇异的。

(P1) 指出广义互信息为零当且仅当随机变量相互独立。在(P2) 我们证明广义互信息对其自变量的任意排列保持不变。我们在(P3) 中证明，广义互信息是不变的，并且不能通过$(X_1, \ldots, X_m)$的任何一一且满射变换而增大。在(P4) 中，我们研究当联合分布相对于其边缘分布的乘积为奇异时广义互信息的一个上界。在$(X_1, \ldots, X_m)$之间存在几乎必然关系，当$I_{D_\phi^{(k)}}$ 达到其上界时。Micheas & Zografos（2006） [19] 表明，基于 ϕ-散度的互信息满足许多其他重要性质，这些性质也可推广到幂散度族，即 $\phi \in C_{B\phi}$。在(P5) 我们证明(P4) 在某些条件下的一个逆结果。

假设 X是一个二元 $0-1$变量，且 Y是一个连续随机变量。则点（x， y） $\in \{0, 1\} \times \mathbb{R}$处的联合密度和条件密度定义为

$$f_{x,y}dy = \mathbb{P}\Big\{X = x, y < Y < y + dy\Big\}, \tag{2.13}$$

$$f_{y|x}dy = \mathbb{P}\Big\{y < Y < y + dy | X = x\Big\}. \tag{2.14}$$

X 和 Y的边缘密度分别记为 $f_x = \mathrm{P}[X = x]$ 和$f_y = \sum_{x=0}^{1} f_{x,y}$。为记号方便，我们将$f_{x_0}$ =P$[X = 0]$ 和 f_{x_1} =P$[X = 1]$记为。
在此混合设定中， $\mathscr{B}$-互信息变为

$$\begin{aligned} I_{D^*}(X,Y) &= D^*(f_{X,Y}, f_X f_Y) \\ &= \sum_{x=0}^{1} \int_{f_y>0} \Big\{ e^{\beta(f_x f_y)^A}(\beta(f_x f_y)^A - \beta f_{x,y}^A - 1) + e^{\beta(f_{x,y})^A} + \frac{1}{B}(f_{x,y}^{A+B} - (f_x f_y)^{A+B}) \\ &\quad -(f_{x,y}^A - (f_x f_y)^A)\frac{A+B}{AB}(f_x f_y)^B \Big\} dy \text{ for } AB \neq 0. \end{aligned} \tag{2.15}$$

若 $A = 0$ 但 $B \neq 0$，则（2.15）中的表达式定义为

$$\lim_{A \downarrow 0+} I_{D^*}(X,Y) = \sum_{x=0}^{1} \int_{f_y>0} \left\{ (f_x f_y)^{1+\alpha} \ln\left(\frac{f_x f_y}{f_{x,y}}\right) - \frac{(f_x f_y)^{1+\alpha} - f_{x,y}^{1+\alpha}}{1+\alpha} \right\} dy, \tag{2.16}$$

类似地，对于 $A \neq 0$， $B = 0$ ，其定义为

$$\begin{aligned} \lim_{B \to 0} I_{D^*}(X,Y) = \sum_{x=0}^{1} \int_{f_y>0} \Bigg\{ & e^{\beta(f_x f_y)^{1+\alpha}} \Big(\beta(f_x f_y)^{1+\alpha} - \beta f_{x,y}^{1+\alpha} - 1\Big) + e^{\beta(f_{x,y})^{1+\alpha}} \\ & + f_{x,y}^{1+\alpha} \ln\left(\frac{f_{x,y}}{f_x f_y}\right) - \frac{f_{x,y}^{1+\alpha} - (f_x f_y)^{1+\alpha}}{1+\alpha} \Bigg\} dy \end{aligned} \tag{2.17}$$

对于 $\alpha > -1$。在（2.16）和（2.17）中，情形 $\alpha = -1$ 也类似地定义为 $\alpha \to -1$。结果在两种情形下均为 0。

3. 基于 $\mathscr{B}-MI$ 的两样本检验

本节提出一类基于 ϕ-生成的扩展 Bregman 散度的检验，用于对两个独立随机样本 $Y_0=(Y_{01},\ldots,Y_{0n_0})$ 和 $Y_1=(Y_{11},Y_{12},\ldots,Y_{1n_1})$ 进行非结构化比较。我们假设 Y_0 和 Y_1 分别独立地取自绝对连续分布函数 F_0 和 F_1。F_0 与 F_1 之间的非结构化比较假设表述为

$$\mathbb{H}_0: F_0=F_1 \text{ against } \mathbb{H}_1: F_0\neq F_1. \tag{3.1}$$

由于 F_0 和 F_1 均为绝对连续的，其概率密度函数 f_0 和 f_1 存在。按照 Guha 和 Chotia（2014, 2021）的方法, [24, 25], 我们将这两个样本合并 Y_0 为 Y_1 $Y:=(Y_0,Y_1)$。进一步定义一个 0–1 二元向量 $X=(X_{01},\ldots,X_{0n_0},X_{11},\ldots,X_{1n_1})$。定义 X 的各个分量如下:

$$X_{ij}=\begin{cases}0 & \text{if } Y_{ij}\in\{Y_{01},\ldots,Y_{0n_0}\},\\ 1 & \text{if } Y_{ij}\in\{Y_{11},\ldots,Y_{1n_1}\}\end{cases} \tag{3.2}$$

对于 $j=1,\ldots,n_i$ 和 $i=0,1$。注意, X 是一个由 n_0 个零后接 n_1 个 1 组成的向量。然而, 下面的检验并不依赖于我们赋予这两组的 X 各分量的具体取值。当 H_0 为真时，合并向量 $Y=(Y_0,Y_1)$ 的所有分量均来自同一分布，因此，它们不应依赖于对 X 各分量的任何排列。因此，原假设 H_0 等价于：X 和 Y 相互独立。根据式（2.12）中定义的互信息 $I_{D_\phi^{(k)}}(X,Y)$，式（3.1）中的非结构化检验问题可改写为

$$\mathbb{H}: f_{x,y}=f_xf_y \text{ a.s.} \iff I_{D_\phi^{(k)}}(X,Y)=0, \tag{3.3}$$

$$\text{against } \mathbb{K}: f_{x,y}\neq f_xf_y \text{ for at least one } (x,y) \iff I_{D_\phi^{(k)}}(X,Y)>0. \tag{3.4}$$

表达式 $I_{D_\phi^{(k)}}(X,Y)$ 在混合设定下已在（2.15）中给出。为实施该检验问题，我们需要估计 $I_{D_\phi^{(k)}}(X,Y)$ 及其在独立性下的渐近分布。给定一个数据集，我们拒绝 $\mathbb{H}$，若估计的互信息超过临界值，否则接受。为简便起见，我们将讨论限制于两样本问题。向两个以上样本的推广可借助高阶核实现。为简洁起见，我们将互换使用记号 $I_{D_\phi^{(k)}}$ 表示 $I_{D_\phi^{(k)}}(X,Y)$。在下一小节中，我们给出 $I_{D_\phi^{(k)}}(X,Y)$ 基于核密度估计的插件估计。

3.1. $\mathscr{B}$-互信息的非参数估计

假设 $Y=(Y_0,Y_1)$ 是一个合并样本，并令 X 为其如前所述的 0–1 标签。为简单起见，我们将合并样本的观测值依次记为 $Y=(Y_1,\ldots,Y_{n_0},Y_{n_0+1},\ldots,Y_n)$，并对 X 也采用类似记法，仅使用一个下标，其中 $n=n_0+n_1$。广义互信息 $I_{D_\phi^{(k)}}$ 可通过点 $(x,y)\in\{0,1\}\times\mathrm{R}$ 处的密度估计来估计

as

$$\widehat{f}_X(x) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}\{X_i = x\}, \widehat{f}_Y(y) = \frac{1}{nh_n}\sum_{i=1}^{n} K\left(\frac{Y_i - y}{h_n}\right), \tag{3.5}$$

$$\widehat{f}_{X,Y}(x,y) = \frac{1}{nh_n}\sum_{i=1}^{n} K\left(\frac{Y_i - y}{h_n}\right)\mathbb{1}\{X_i = x\}, \tag{3.6}$$

其中 K 是一个合适的核函数，而 h_n 是带宽序列。估计得到的广义互信息记为 $I_{D_\phi^{(k)}}$。

遵循 Fernandes 和 N´eri（2009） [26], ，我们证明了 $\widehat{I}_{D_\phi^{(k)}}$ 的类似渐近表示，其误差项的阶适当。首先，我们作出以下假设。

（A1）凸函数 ϕ（t）在（2.9）中定义，且四次可微。这些导数均由可积函数一致有界，对所有 t 均成立。

（A2）概率密度函数 $f_{x、y}$、 f_y 关于 y 连续二次可微，其中 $x = 0$，**1**。假设这些导数有界。

（A3）（i）核 K 关于零对称且有界，即

$$K(u) = K(-u) \text{ for all } u \in \mathbb{R}, \text{ and } \|K\| := \sup_{u\in\mathbb{R}} |K(u)| < \infty. \tag{3.7}$$

（ii）核 K 具有有界支撑。

（A4）带宽序列 $\{h_n\}$ 满足以下条件：

$$h_n \longrightarrow 0, nh_n^2 \longrightarrow +\infty, \text{ and } nh_n^4 \longrightarrow 0 \text{ as } n \longrightarrow +\infty. \tag{3.8}$$

假设 (A1) 允许对 $\widehat{I}_{D_\phi^{(k)}}$ 进行二阶泰勒级数展开，并保留一个余项。假设（A2）要求 Y 的联合密度和边缘密度足够光滑，从而可以进行泛函展开。假设 (A1) 和 (A2) 共同意味着，$\widehat{I}_{D_\phi^{(k)}}$ 展开中的余项依概率有界。假设 (A3) 对核施加条件，以减少核密度估计中的偏差。假设 (A4) 是关于带宽序列的一组技术条件，是确立 $\widehat{I}_{D_\phi^{(k)}}$ 渐近正态性所必需的。假设 (A3)(i) 和 (A4) 共同确保收敛 $\widehat{f}_{x,y}$ P—→ $f_{x,y}$ and$\widehat{f}_y$ P—→ f_y 以适当的阶在 x, y 上一致。假设 (A3)(ii) 限制了对无界支撑集核的使用。由于核密度估计通常对核的选择具有稳健性，而且有界支撑集核的类别足够丰富，这在实际应用中或许不会造成严重问题。不过，通过对涉及核的复杂积分的有限性施加条件，可以放宽假设 (A3)(ii)；这一问题留待进一步研究。引理中的以下渐近展开给出了原假设下的渐近分布。

引理 3.1. 假设 *(A1)–(A3)(i)*和*(A4)* 均成立。那么在原假设 $\widehat{}$下， $\mathbb{H}$， $I_{D_\phi^{(k)}}$ 具有以下展开：

$$\widehat{I}_{D_\phi^{(k)}} = \frac{1}{2}\int_{f_y>0}\sum_{x=0}^{1}\left[k^2(f_xf_y)^{2k}\phi''(f_x^k f_y^k)\left(\frac{h_x}{f_x}+\frac{h_y}{f_y}-\frac{h_{x,y}}{f_xf_y}\right)^2\right]dy + o_{\mathbb{P}}\left(\frac{1}{nh_n^{1/2}}\right), \tag{3.9}$$

where $h_x = \widehat{f}_x - f_x$, $h_y = \widehat{f}_y - f_y$, $h_{x,y} = \widehat{f}_{x,y} - f_{x,y}$ *for all* x, y.

这些项 h_x、 h_y、 $h_{x、y}$ 用于量化估计量中的偏差。由引理3.1 可知， $\widehat{I}_{D_\phi^{(k)}}$ （经过适当规范化后）的极限分布$\widehat{}$由 $I_{D_\phi^{(k)}}$ 的第一个非退化导数决定。事实证明，展开中的前两项是奇异的，而渐近分布仅由其二阶导数决定。

这一结果是证明独立性条件下检验统计量渐近正态性的关键。对 ϕ 要求四阶可微性相当严格；不过，广义 S-Bregman 散度中的成员可能满足这一条件。在这种情况下， ϕ的导数由一个指数项 $e^{\beta(f_xf_y)^A}$ 和一个多项式函数组成。在大多数情况下，对于任何有限的 β，指数项都保持有界。

注3.1。 对于 *GSB*散度，引理*3.1*给出

$$2\widehat{I}_{D^*}(X,Y)$$
$$= \int_{f_y>0}\sum_{x=0}^{1}\left[\left\{(f_xf_y)^{2A}(A\beta)^2e^{\beta(f_xf_y)^A} + \Big(1+\alpha\Big)(f_xf_y)^{1+\alpha}\right\}\left(\frac{h_x}{f_x}+\frac{h_y}{f_y}-\frac{h_{x,y}}{f_xf_y}\right)^2\right]dy + o_{\mathbb{P}}\left(\frac{1}{nh_n^{1/2}}\right). \tag{3.10}$$

$\widehat{}$这里需要注意的重要一点是，当 $\beta = 0$ 时， I_{D^*} 的渐近展开仅依赖于 $\alpha \geq 0$。此外，在 $\alpha = 0$ 下，它给出了幂散度族的表达式。稍后我们将看到，幂散度族的渐近均值和方差并不依赖于其调节参数 $\lambda \in \mathbb{R}$。

引理 3.2. 设假设 *(A1)* – *(A4)* 成立。则 $\widehat{I}_{D_\phi^{(k)}} \xrightarrow{\mathbb{P}} I_{D_\phi^{(k)}}$ 在真实联合密度 $f_{X,\,Y}$处。

注3.2。 *Parzen*（*1962*） *[27]* 表明，当满足 $nh_n \to +\infty$时，核密度估计是弱相合的。有关相合性更强形式的更多详情，请参见 *Wied* 和 *Weißbach*（*2012*） *[28]*。

3.2. 独立性下 $I_{D_\phi^{(k)}}$ 的渐近正态性

在本小节中，我们将正式陈述在原假设 H 下 $I_{D_\phi^{(k)}}$ 的渐近分布。为此，我们定义

$$\mu_\phi = \frac{1}{2}\int_u K^2(u)du \int_{f_y>0}\left[\sum_{x=0}^{1} k^2(f_x f_y)^{2k-1}\phi''(f_x^k f_y^k)(1-f_x)\right]dy, \tag{3.11}$$

$$\sigma_\phi^2 = \frac{1}{2}\int_u \left(\int_z K(z)K(z+u)dz\right)^2 du \int_{f_y>0}\left[\sum_{x=0}^{1} k^2(f_x f_y)^{2k-1}\phi''\left(f_x^k f_y^k\right)(1-f_x)\right]^2 dy. \tag{3.12}$$

为此，我们作出以下假设。

(A5) μ_ϕ 和σ_ϕ 均为有限值。此外，σ_ϕ 必须为正。

稍后我们将看到，$\frac{\mu_\phi}{nh_n}$ 和 $\frac{\sigma_\phi}{nh_n^{1/2}}$将分别用作中心化和缩放序列，以对$I_{D_\phi^{(k)}}$ 进行标准化，从而推导其独立性下的渐近分布。为说明假设(A5)并非多余，考虑以下情形：

$$\sum_{x=0}^{1} k^2(f_x f_y)^{2k-1}\phi''\left(f_x^k f_y^k\right)(1-f_x) \tag{3.13}$$

作为 y的函数保持不变。此外，考虑 Y 具有无界支撑集的情形。在这种情况下，仅凭假设(A1) 和(A3) 并不能推出 μ_ϕ 和σ_ϕ均为有限值。因此，还需单独作出假设(A5)。注意，对于幂散度族，(3.13)成立；因此，为了使原假设下的渐近正态性结果成立， μ_ϕ 和 σ_ϕ^2 都必须为有限值。在这种情况下，假设(A5)等价于假设连续随机变量具有有界支撑，从而限制了该结果在实际应用中的适用范围。不过，当 $\alpha \neq 0$时，即使支撑集无界，我们仍可将该结果用于广义 S-Bregman 散度族。现在我们从以下引理开始，它实质上验证了 Hall 和 Heyde（2014） [31] (引理 3.1，第 57 页）中的条件，并将在混合框架下的中心极限定理证明中进一步使用。定义 $T_i = \frac{2}{nh_n^{3/2}}\sum_{j<i} V_{ij}$ 其中

$$V_{ij} = \sum_{x=0}^{1}\int_{f_y>0} k^2(f_x f_y)^{2k-2}\phi''(f_x^k f_y^k)Z_{ix}Z_{jx}K_{h_n i}(y)K_{h_n j}(y)dy = V_{ji}. \tag{3.14}$$

令 $Z_n := \prod_{j=1}^{n}(1+itT_j)$ 定义在共同概率空间 $(\Omega, \mathrm{F}, \mathrm{P})$ 上，其中 $t \in$ 固定 $\mathbb{R}$.

引理 3.3。 将随机变量之和记为 $S_n = \sum_{i=1}^{n} T_i$，并令 $\mathcal{F}_m := \sigma(T_1, \ldots, T_m)$ 为由以下对象生成的 σ- 域的递增序列 $\{(T_1, \ldots, T_m) : m \leq n\}$。在原假设 $\mathbb{H}$ 及假设 *(A1)-(A5)* 下，以下结论成立：

(a) $\sum_{i=1}^{n} T_i^2 \xrightarrow{\mathbb{P}} 4\sigma_\phi^2$,**(b)** 最大值 $\max_{1\leq i\leq n}|T_i| \xrightarrow{\mathbb{P}} 0$,,

(c) $Z_n \longrightarrow 1$ *(weakly in* L_1*), i.e.,* $\mathbb{E}\Big[Z_n \mathbb{1}(E)\Big] \longrightarrow \mathbb{P}(E)$ *for all* $E \in \mathcal{F}$ *as* $n \to \infty$.

现在，我们可以陈述与检验统计量相关的渐近正态性。

定理 3.1. 假设 *(A1)-(A5)* 均成立。在原假设$\mathbb{H}$下，有

$$nh_n^{1/2}\Big(\widehat{I}_{D_\phi^{(k)}} - \frac{\mu_\phi}{nh_n}\Big) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma_\phi^2) \ \ as\ n \to \infty. \tag{3.15}$$

在 $h_n^{-1/2}$的混合框架中，偏差会随其阶数增加而增大，这是 MI 估计的已知问题之一。要应用定理 3.1，偏差估计至关重要。我们观察到，GSB 散度的正态收敛速度相当缓慢，尤其是在 $\beta < 0$ 时。不过，增大样本量可以改善这一点。接下来，我们计算 GSB 散度族不同特殊情形下不含因子 $\frac{1}{nh_n}$ 的渐近均值和渐近方差。 ——

推论 3.1. 对于 *GSB* 散度，其渐近均值和方差为

$$\mu_\phi = \frac{1}{2}\int_u K^2(u)du \int \Bigg[\sum_{x=0}^{1}(f_x f_y)^{2A-1}\Big[(A\beta)^2 e^{\beta(f_x f_y)^A} + (A+B)(f_x f_y)^{B-A}\Big](1-f_x)\Bigg]dy, \tag{3.16}$$

$$\sigma_\phi^2 = \frac{1}{2}\int \Bigg(\int K(u)K(u+z)dz\Bigg)^2 du \int \Bigg[\sum_{x=0}^{1}(f_x f_y)^{2A-1}\Big[(A\beta)^2 e^{\beta(f_x f_y)^A} + (A+B)(f_x f_y)^{B-A}\Big](1-f_x)\Bigg]^2 dy. \tag{3.17}$$

推论 3.2. *(*幂散度族*)* 这里 $\beta = 0$, $A + B = 1$ 和

$$\mu_\phi = \frac{1}{2}\int_u K^2(u)du \int_{f_y>0} dy, \tag{3.18}$$

$$\sigma_\phi^2 = \frac{1}{2}\int_u \Big(\int_z K(z)K(z+u)dz\Big)^2 du \int_{f_y>0} dy. \tag{3.19}$$

推论 3.3. （*S*-散度、*S*-海林格、*DPD*）这里有 $\beta = 0$、 $A + B = 1 + \alpha$ 和

$$\mu_\phi = \frac{(1+\alpha)}{2}\int_u K^2(u)du\Big(f_{x_0}^\alpha f_{x_1} + f_{x_1}^\alpha f_{x_0}\Big)\int_{f_y>0} f_y^\alpha dy, \tag{3.20}$$

$$\sigma_\phi^2 = \frac{(1+\alpha)^2}{2}\int_u \Big(\int_z K(z)K(z+u)dz\Big)^2 du\Big(f_{x_0}^\alpha f_{x_1} + f_{x_1}^\alpha f_{x_0}\Big)^2 \int_{f_y>0} f_y^{2\alpha} dy. \tag{3.21}$$

推论 3.4. （平方 L_2-散度）这里有 $\beta = 0$、 $A + B = 2$。因此

$$\mu_\phi = 2 f_{x_0} f_{x_1} \int_u K^2(u) du, \tag{3.22}$$

$$\sigma_\phi^2 = 8 (f_{x_0} f_{x_1})^2 \int_u \Big(\int_z K(z) K(z+u) dz \Big)^2 du \int_{f_y > 0} f_y^2 dy. \tag{3.23}$$

推论 3.5. *(BED)* 此处 $\beta \in \mathbb{R} \setminus \{0\}$， $A + B = 0, A = 1$. 最后，均值乘以 $\frac{2}{\beta^2}$，方差乘以 $\frac{4}{\beta^4}$。因此

$$\mu_\phi = (f_{x_0} f_{x_1}) \int_u K^2(u) du \int_{f_y > 0} \left(e^{\beta (f_{x_0} f_y)} + e^{\beta (f_{x_1} f_y)} \right) f_y dy \tag{3.24}$$

and

$$\sigma_\phi^2 = 2 (f_{x_0} f_{x_1})^2 \int_u \left(\int K(z) K(z+u) dz \right)^2 du \int_{f_y > 0} \left(e^{\beta (f_{x_0} f_y)} + e^{\beta (f_{x_1} f_y)} \right)^2 f_y^2 dy. \tag{3.25}$$

对于此族， $\beta = 0$ 定义为 $\beta \to 0$的连续极限。

我们发现， S-散度族的渐近均值和方差仅取决于调节参数 $\alpha \in [0, 1]$。经验性证据表明，随着 α 逐渐增至 1，渐近均值和方差会减小。基于 S- 散度的互信息也会随着 α 增大而减小。这一观察至关重要，或许能部分解释较高 α下检验统计量的稳定性行为。相反，随着 β 在实数轴上减小，BED 族的稳健性会增强。

如果 Y 的支撑集无界，我们就无法对幂散度族使用定理 3.1 。因此，当 Y 的支撑集无界时，我们需要调用置换算法来计算经验水平和幂。我们注意到，执行置换检验需要耗费大量计算资源；对于此族中的其他成员，即使 Y 的支撑集无界，也完全可以避免这种计算。此外还要注意，渐近正态性结果有时可能只有在样本量非常大时才有用。

这一事实显然是诸多重要亮点之一，足以证明除 $\alpha = \beta = 0$ 外使用 GSB 散度族的合理性。

注记 3.3。基于 *Itakura-Saito* 距离的互信息可由针对 $\phi(t) = -\frac{\log(t)}{2\pi}$ 的扩展 *Bregman* 散度得到，其中 $k = 1$。对于这一族，我们有

$$\mu_\phi = \frac{1}{4\pi} \int_u K^2(u) du \left[\frac{f_{x_0}}{f_{x_1}} + \frac{f_{x_1}}{f_{x_0}} \right] \int_{f_y > 0} \frac{dy}{f_y}, \tag{3.26}$$

$$\sigma_\phi^2 = \frac{1}{8\pi^2} \int_u \left(\int K(z) K(z+u) dz \right)^2 du \left[\frac{f_{x_1}}{f_{x_0}} + \frac{f_{x_0}}{f_{x_1}} \right]^2 \int_{f_y > 0} \frac{dy}{f_y^2}. \tag{3.27}$$

3.3. 连续备择下的一致性与检验功效

假设满足定理 3.1 的中心极限定理条件，且 μ_ϕ, σ_ϕ 已知。当 $\widehat{I}_{D_\phi^{(k)}} > t$ 满足 $\mathrm{P}_{\mathbb{H}}\{\widehat{I}_{D_\phi^{(k)}} > t\} = c$ 时，在 $100c\%$ 的名义显著性水平下，将拒绝式（3.3）中给出的原假设 H。简单计算可知，当 n足够大时，$t = \frac{\tau_c \sigma_\phi}{nh_n^{1/2}} + \frac{\mu_\phi}{nh_n}$，其中 τ_c是标准正态分布的上侧 $100c\%$ 分位点。利用引理 3.2，可以容易地证明，当 $n \to \infty$时，$\mathrm{P}_{\mathbb{K}}\{\widehat{I}_{D_\phi^{(k)}} \geq \frac{\tau_c \sigma_\phi}{nh_n^{1/2}} + \frac{\mu_\phi}{nh_n}\} \longrightarrow 1$ 。因此，该检验类是一致的。

在实践中，μ_ϕ，σ_ϕ 通常是未知的，需要对其进行估计才能执行检验程序。然而，可以使用 $\widehat{f}_X(x)$ 和 $\widehat{f}_Y(y)$ 估计 μ_ϕ 和 σ_ϕ。标准化检验统计量及其经验版本记为

$$\widehat{T}_{D_\phi^{(k)}} = \frac{nh_n^{1/2}(\widehat{I}_{D_\phi^{(k)}} - \frac{\widehat{\mu_\phi}}{nh_n})}{\widehat{\sigma}_\phi} \text{ and } T_{D_\phi^{(k)}} = \frac{nh_n^{1/2}(\widehat{I}_{D_\phi^{(k)}} - \frac{\mu_\phi}{nh_n})}{\sigma_\phi}. \tag{3.28}$$

在下一个结果中，我们证明，在原假设 H 下，$T_{D_\phi^{(k)}}$ 和 $T_{D_\phi^{(k)}}$ 都收敛于相同的极限分布；此外，基于 $T_{D_\phi^{(k)}}$ 的检验是相合的。

定理 3.2。 在假设 *(A1)-(A5)*下，当 $n \to \infty$ 时，以下结论成立。

(i) $|T_{D_\phi^{(k)}} - \widehat{T}_{D_\phi^{(k)}}| \mathbb{P} \longrightarrow 0$，因此，$\widehat{T}_{D_\phi^{(k)}} \mathcal{L} \longrightarrow \mathcal{N}(0,1)$ 在原假设 $\mathbb{H}$下成立。***(ii)*** 基于 $\widehat{T}_{D_\phi^{(k)}}$ 的检验是相合的。

当核密度估计一致收敛时，基于它们构造的检验也一致收敛。接下来，研究备择假设与原假设连续时其功效函数的表现便很有意义。假设我们要检验原假设 H 与以下连续或局部备择序列

$$\mathbb{K}_n : (X,Y) \sim f^{(n)} \text{ where } f_{x,y}^{(n)} = f_x f_y + \frac{d}{\sqrt{nh_n^{1/2}}}\Delta_{x,y} \text{ where } n = 1,2,\ldots, \tag{3.29}$$

对于已知常数 $d \geq 0$ 和 $\Delta_{x,\,y}$ 满足 $\sum_{x=0}^{1}\int_{f_y>0}\Delta_{x,\,y}dy = 0$。由于 $f_{x}^{(n),\,y}$ 是相对于 H 的连续序列，其边缘分布在相同的污染比例序列下也将如此。在此背景下，将边缘分布的污染序列记为

$$f_x^{(n)} = \int_{f_y>0} f_{x,y}^{(n)}dy = f_x + \frac{d}{\sqrt{nh_n^{1/2}}}\Delta_x \text{ and } f_y^{(n)} = \sum_{x=0}^{1} f_{x,y}^{(n)} = f_y + \frac{d}{\sqrt{nh_n^{1/2}}}\Delta_y, \tag{3.30}$$

其中 $\Delta_x = \int_{f_y>0}\Delta_{x,\,y}dy$，$\Delta_y = \sum_{x=0}^{1}\Delta_{x,\,y}$。注意，连续密度以比 $\frac{1}{nh_n^{1/2}}$ 更慢的速率 $\frac{1}{\sqrt{nh_n^{1/2}}}$ 收敛到原假设密度。这个速率恰好是稳定偏差项所必需的；当 $\mathbb{K}_n$ 成立时，该偏差项必然会使渐近零分布发生偏移。因此，选择这样的污染序列对于推导渐近分布起着至关重要的作用

$T_{D_\phi^{(k)}}$ 在 K$_n$ 下的情况。注意，当 $d=0$ 时，K$_n$ 下的解为 $f_{x,\ y}=f_xf_y$ 对所有 x，y 几乎必然成立，这意味着 K$_n$ =H。在陈述下一个结果之前，先建立一些记号会很有帮助。定义

$$I_{D_\phi^{(k)}}^{(n)}=\sum_{x=0}^{1}\int_{f_y>0}\left[\phi\Big((f_{x,y}^{(n)})^k\Big)-\phi\Big((f_x^{(n)}f_y^{(n)})^k\Big)-\Big\{\Big(f_{x,y}^{(n)}\Big)^k-\Big(f_x^{(n)}f_y^{(n)}\Big)^k\Big\}\phi'\Big((f_x^{(n)}f_y^{(n)})^k\Big)\right]dy. \tag{3.31}$$

接下来，我们将 $I_{D_\phi^{(k)}}^{(n)}$ 近似到一个二阶项，以便推导邻近备择下 $T_{D_\phi^{(k)}}$ 的渐近分布。

引理 3.4. 假设 *(A1)* 成立，则有

$$I_{D_\phi^{(k)}}^{(n)}=\frac{d^2}{2nh_n^{1/2}}\sum_{x=0}^{1}\int_{f_y>0}k^2(f_xf_y)^{2k}\phi''(f_x^kf_y^k)\left(\frac{\Delta_x}{f_x}+\frac{\Delta_y}{f_y}-\frac{\Delta_{x,y}}{f_xf_y}\right)^2dy+o\left(\frac{d^2}{nh_n^{1/2}}\right). \tag{3.32}$$

接下来，我们证明，对（3.32）中的二阶项进行适当缩放后，会使邻近备择$\widehat{\ }$下 $I_{D_\phi^{(k)}}$ 的零分布发生$\mathbb{K}_n$位置偏移，而渐近方差保持不变。

Now，我们给出邻近备择 K 下 $T_{D_\phi^{(k)}}$ 的渐近分布 n.

定理 3.3. 假设*(A1)-(A5)*均成立，且 $0\le d\le C$ ·上确界 x $y-,\ y|f_{x,}$ 其中 为常数。则以下结果成立：$f_xf_y|$ $C>0$

$$T_{D_\phi^{(k)}}\xrightarrow{\mathcal{L}}\frac{d^2}{2\sigma_\phi}\sum_{x=0}^{1}\int_{f_y>0}k^2(f_xf_y)^{2k}\phi''(f_x^kf_y^k)\left(\frac{\Delta_x}{f_x}+\frac{\Delta_y}{f_y}-\frac{\Delta_{x,y}}{f_xf_y}\right)^2dy+\mathcal{N}(0,1) \tag{3.33}$$

如邻近备择$\mathbb{K}_n$ 下

注 3.4. 特别地，假设我们取 $\Delta_{x,\ y}=(\delta_{x_0}(x)\delta_{y_0}(y)-f_xf_y)$，其中 $\delta_{x_0}(x)\delta_{y_0}(y)$是一个在点 $t_0=(x_0,y_0)$处退化的概率密度函数。于是我们得到

$$T_{D_\phi^{(k)}}\xrightarrow{\mathcal{L}}\frac{d^2}{2\sigma_\phi}\mathcal{IF}_2\Big(I_{D_\phi^{(k)}},f_Xf_Y,t_0\Big)+\mathcal{N}(0,1)\ \ under\ \mathbb{K}_n, \tag{3.34}$$

其中，影响函数$_2$ 是 $I_{D_\phi^{(k)}}$ 在原假设下、点 t_0处的二阶影响函数。稍后我们将看到，当原假设成立时，影响函数$_2$ 衡量了 $I_{D_\phi^{(k)}}$ 在点 t_0处受到无穷小污染时的稳定性行为。影响函数$_2$的计算将在下一节中给出。邻近备择$\mathbb{K}_n$ 下的渐近局部功效因此为

因此由下式给出

$$\lim_{n\to\infty}\mathbb{P}_{\mathbb{K}_n}\left[T_{D_\phi^{(k)}} > \tau_c\right] = 1 - \Phi_1(\tau_c - d^2 S_\phi) \ \textit{where}\ S_\phi := \frac{\mathcal{IF}_2(I_{D_\phi^{(k)}}, f_X f_Y, t_0)}{2\sigma_\phi}. \tag{3.35}$$

我们可以将 S_ϕ 称为与 $T_{D_\phi^{(k)}}$ 在局部备择下的渐近功效相关联的斜率函数。请注意，随着 S_ϕ本身随 ϕ-函数的变化， $T_{D_\phi^{(k)}}$ 的局部渐近功效也会发生变化。我们可以研究 S_ϕ ，以比较由不同 ϕ生成的扩展 ***Bregman*** 散度族成员在此类邻近备择下的标准化检验统计量 $T_{D_\phi^{(k)}}$ 的幂。此外，还要注意，对于所有严格凸且至少二次可微的 ϕ-函数， S_ϕ 均为正值。斜率 S_ϕ 越大，随着 d^2的增加，邻近备择下的渐近功效就越快地从名义水平*"c"*升至 *1*。因此，作为 $I_{D_\phi^{(k)}}$ 稳健性度量的二阶影响函数 影响函数$_2$，，以一种有趣的方式直接关联到 $T_{D_\phi^{(k)}}$ 在邻近备择 $\mathbb{K}n$ 下的局部渐近功效。

在下一节4中，我们将研究 $I_{D_\phi^{(k)}}$ 的稳健性行为。

4. 稳健性研究

分析泛函稳健性的一种方法，是考察底层数据生成概率密度函数受到扰动时的表现。影响函数（IF）是广泛使用的局部稳健性度量概念。然而，它有时可能无法捕捉稳健性特征。在这种情况下，就有必要考虑高阶影响函数。

4.1. 独立性下 $I_{D_\phi^{(k)}}$ 的影响函数

假设 $(X,\ Y)$ 的联合密度在点 $t_0 = (x_0,\ y_0) \in \{0,\ 1\} \times \mathrm{R}$ 处受到污染，形式如下：

$$f^\epsilon_{X,Y}(x,y) = (1-\epsilon) f_{x,y} + \epsilon\delta_{x_0}(x)\delta_{y_0}(y), \tag{4.1}$$

其中， $\delta_{z_0}(z) = \mathbb{1}\{z = z_0\}$ 是狄拉克 δ 函数，且 $0 \le \epsilon \le 1$。我们知道， $\int \delta_{z_0}(dz) = 1$；或者根据 z 是连续的还是离散的， $\sum \delta_{z_0}(z) = 1$ 。对于狄拉克 δ 函数的积分，一种方便但不严格的记号是 $\int \delta_{z_0}(z)dz = 1$，这里将其理解为前一种情况。与 ϵ-受污染的联合密度相对应的边缘分布为

$$f^\epsilon_X(x) = (1-\epsilon) f_x + \delta_{x_0}(x) \text{ and } f^\epsilon_Y(y) = (1-\epsilon) f_y + \epsilon\delta_{y_0}(y). \tag{4.2}$$

广义互信息泛函 $I_{D_\phi^{(k)}}$ 在点 $t_0 = (x_0,\ y_0)$ 处的（一阶）影响函数定义为

$$\mathcal{IF}_1(I_{D_\phi^{(k)}}, f_{X,Y}, t_0) = \left[\frac{\partial D_\phi^{(k)}(f^\epsilon_{X,Y}, f^\epsilon_X f^\epsilon_Y)}{\partial \epsilon}\right]_{\epsilon=0}. \tag{4.3}$$

简单计算可得

$$\mathcal{IF}_1(I_{D_\phi^{(k)}}, f_{XY}, t_0) = \sum_{x=0}^{1}\int_{f_y>0}\Bigg[k(f_{x,y})^{k-1}\Big(\delta_{x_0}(x)\delta_{y_0}(y) - f_{x,y}\Big)\Big\{\phi'\Big(f^k_{x,y}\Big) - \phi'\Big(f^k_x f^k_y\Big)\Big\}$$
$$-k(f_x f_y)^{k-1}\Big(f^k_{x,y} - (f_x f_y)^k\Big)(f_x\Big(\delta_Y(y_0) - f_y) + f_y(\delta_X(x_0) - f_x)\Big)\phi''\Big(f^k_x f^k_y\Big)\Bigg]dy. \tag{4.4}$$

在原假设 H 下，这可以写成

$$\mathcal{IF}_1(I_{D_\phi^{(k)}}, f_X f_Y, t_0) = 0 \text{ for all } t_0 \in \{0,1\}\times\mathbb{R}. \tag{4.5}$$

可以看出，一阶影响函数无法体现泛函 $I_{D_\phi^{(k)}}$ 在独立性下的任何稳健性特征，因为对于构成扩展 Bregman 散度族的所有 ϕ 函数，它都变为常数。因此，为了获得进一步的洞见，我们需要考察二阶影响函数，其定义方式类似：

$$\mathcal{IF}_2(I_{D_\phi^{(k)}}, f_{X,Y}, t_0) = \left[\frac{\partial^2 D_\phi^{(k)}(f^\epsilon_{X,Y}, f^\epsilon_X f^\epsilon_Y)}{\partial \epsilon^2}\right]_{\epsilon=0}. \tag{4.6}$$

下一个结果给出了原假设下二阶影响函数的显式表达式。

定理 4.1. 在原假设$\mathbb{H}$下，$I_{D_\phi^{(k)}}$ 在点 t_0 处的二阶影响函数为

$$\mathcal{IF}_2(I_{D_\phi^{(k)}}, f_X f_Y, t_0) = \sum_{x=0}^{1}\int_{f_y>0} k^2(f_x f_y)^{2k}\phi''\Big(f^k_x f^k_y\Big)\left[\frac{\Delta_x}{f_x} + \frac{\Delta_y}{f_y} - \frac{\Delta_{x,y}}{f_x f_y}\right]^2 dy \tag{4.7}$$

当 $t_0 \in \{0,1\}\times\mathbb{R}$ 时，其中 $\Delta_x = \delta_{x_0}(x) - f_x$，$\Delta_y = \delta_{y_0}(y) - f_y$ 和 $\Delta_{x,y} = \delta_{x_0}(x)\delta_{y_0}(y) - f_x f_y$。

该结果的证明较为直接，因此在附录中予以省略。由（4.7）可见，二阶影响函数的取值始终为正。影响函数的绝对值越大，检验统计量的稳定性越好。为简化计算，通常可利用狄拉克 δ 函数的如下性质：$\int_{-\infty}^{\infty} g(z)\,\delta_{z_0}(dz) = g(z_0)$ 对于“性态良好”的函数 g。随后，我们将绘制广义 S-Bregman 散度族中不同成员的影响函数，并更详细地讨论其性质。

接下来，我们计算检验泛函的影响函数。若不计速率 $nh_n^{1/2}$，该检验泛函可视为

$$W_{D_\phi^{(k)}} = \left[\frac{I_{D_\phi^{(k)}} - \frac{\widehat{\mu}_\phi}{nh_n}}{\widehat{\sigma}_\phi} \right]. \tag{4.8}$$

简单计算表明

$$\mathcal{IF}_1(W_{D_\phi^{(k)}}, f_X f_Y, t_0) = 0 \text{ and } \mathcal{IF}_2(W_{D_\phi^{(k)}}, f_X f_Y, t_0) = \frac{1}{\widehat{\sigma}_\phi} \mathcal{IF}_2(I_{D_\phi^{(k)}}, f_X f_Y, t_0) \tag{4.9}$$

在原假设 H 下。检验泛函 $W_{D_\phi^{(k)}}$ 的稳定性行为由 $I_{D_\phi^{(k)}}$ 决定，因为它们的二阶影响函数彼此成比例。

遗憾的是，当原假设不成立时，情况会困难得多。在这些情形下，由于备择假设 K 是复合假设，我们没有二阶影响函数的简化表达式。另一种方法是利用水平与功效影响函数，研究其在毗连备择处的稳定性，这将在下一小节中讨论。

对于 GSB族，定理4.1给出

$$\mathcal{IF}_2(I_{D_\phi^{(k)}}, f_X f_Y, t_0) = \sum_{x=0}^{1} \int_{f_y>0} \left\{ A^2\beta^2 e^{\beta(f_x f_y)^A} (f_x f_y)^{2A-1} + (A+B)(f_x f_y)^{A+B-1} \right\} U_{xy} dy, \tag{4.10}$$

where $U_{xy} = (f_x f_y) \left(\frac{h^X_{x_0}}{f_x} + \frac{h^Y_{y_0}}{f_y} - \frac{h^{X,Y}_{x_0,y_0}}{f_x f_y} \right)^2$ and $t_0 \in \{0,1\} \times \mathbb{R}$.

特别地，

$$\mathcal{IF}_2(I_{D_\phi^{(k)}}, f_X f_Y, t_0) = \begin{cases} \sum_{x=0}^{1} \int_{f_y>0} A^2\beta^2 e^{\beta(f_x f_y)^A} (f_x f_y)^{2A-1} U_{xy} dy & \text{if } A+B=0, A, \beta \neq 0, \\ (1+\alpha) \sum_{x=0}^{1} \int_{f_y>0} (f_x f_y)^\alpha U_{xy} dy & \text{if } \beta = 0. \end{cases} \tag{4.11}$$

注意，当 $A+B=0$，$A=1$ 时，我们得到缩放 BED 族的表达式，需将其乘以 $2/\beta^2$（其中 $\beta \neq 0$），才能得到 BED 族的影响函数[2]。

影响函数的有界性本质上取决于控制求和与积分中的项 U_{xy}（见4.10）。由于 f_y 是连续随机变量的概率密度函数，其有界性可能会成为一个问题。然而，一个充分条件是，例如 Y 的密度具有连续有界导数，这可以确保其密度一致有界。当 f_y 一致有界时，U_{xy} 仍可能无界，因为该密度仍位于其分母中。然而，在大多数情况下，我们会看到因子 $(f_x f_y)^\tau$会增强 U_{xy} 对于 $\tau > 0$的稳定性。取值越大，影响函数对极端离群值的稳定性越高 τ。

根据上述推测，可以看到，影响函数$_2$ 在（4.10）中的第一项对于$2A-1>0$，即 $A>\frac{1}{2}$，会变得更加稳定。同样地，当 $A+B-1>0$，即 $\alpha>0$时，第二项会更加稳定。此外，注意到 $e^{\beta(f_x f_y)}$ 在有限的 β下始终保持有界。

如 Ba sak 和 Basu (2022) [12], ，我们研究以下区域中的影响函数 .

(1) 我们知道， $\beta=0$ 给出了 S-散度族。此外，当 $\alpha=0$时，我们得到幂散度族的影响函数$_2$ 表达式，而它与 λ无关。根据前面的讨论，我们知道，只有当 $\alpha>0$ 时，S-散度族的影响函数$_2$ 才会更加稳定（见 (4.11)）。当 $\alpha=0$时，它会变得非常不稳定，即对于幂散度族而言如此。该区域记为 $S_1=\{(\alpha,\lambda,\beta):\alpha>0,\lambda\in R\ \beta=0\}$。(2) 当 $\beta\neq 0, A=0$时，(4.10) 中的第一项消失，而非零的第二项在 $\alpha>0$时更加稳定。然而， $A=0$ 意味着 $\lambda=\frac{1}{\alpha-1}$，且整个表达式与 β无关。因此，可行区域为 $S_2=\{(\alpha,\lambda,\beta):\alpha>0,\lambda=\frac{1}{\alpha-1},\beta\neq 0\}$。由于$\alpha>0$ 在 S_2, 中，可行的调节参数 λ 应为负值。(3) 当 $(A+B)=0$，或者 $\alpha=-1$时，(4.10) 中的第二项消失。此时，第一项在 $A>\frac{1}{2}$时更加稳定，即 $\lambda>-\frac{1}{4}$和 $\beta\neq 0$。在这种情况下，影响函数$_2$在 $S_3=\{(\alpha,\lambda,\beta):\alpha=-1,\lambda>-\frac{1}{4},\beta\neq 0\}$ 中会更加稳定。(4) 在第四种情况下， $\beta, A,(A+B)$ 均应非零。因此，可行区域为 $S_4=\{(\alpha,\lambda,\beta);\ \alpha>0,\lambda(1-\alpha)>-\frac{1}{2},\beta\neq 0\}$。

综合这些互不相交的区域可知，当调节参数属于 $S:=\cup_{i=1}^{4}S_i$ 时，由影响函数 IF_2 衡量的离群值的无穷小影响会明显减弱。在此区域之外，影响函数会变得非常不稳定。PD族属于这一不稳定区域。

为了进一步探讨这一问题，我们取 $X\sim Bernoulli(0.5)$ 和 $Y\sim N(0,1)$。 X、Y 在原假设 H 下假定相互独立。在图4.1、4.2 和 4.3 中，我们绘制二阶总体误差敏感度（GES），其定义为 $GES_2:=\sup_y IF_2(\cdots,(0,y))$，其中固定 $x=0$。出于实际考虑，我们对 $[-20,20]$ 取上确界；对于从连续分布 N(0,1) 中添加离群值而言，这是一个非常大的范围。从图4.1 可明显看出，幂散度族（$\alpha=\beta=0$）的所有成员在稳健性方面都等效。因此，来自真实连续分布极端区域、甚至位于支撑集之外的离群值，都会使基于 PD族的检验变得不稳定。然而，采用 $\alpha>0$ 后，其稳健性会大幅提高。在其他图中，我们绘制了调节参数属于 S_i、$i=2$、3 时的有界 GES 曲线，以及这些区域之外的一些无界曲线。在图 4.4 中，我们绘制了 IF_2，并恰当地标出了该区域之外的一些无界曲线。

经验性证据表明，在点质量处发生极端数据污染时，PD族之外的成员通常具有更好的稳定性。对于固定的 $\lambda\in R$，较高的 α andnegative β 值通常能带来更好的稳定性。

注 4.1。 （本设置中的污染概念）在参数估计中，通常假设污染密度不应属于参数族，或者应在几何上与真实分布分离。在此情形下，也可以作出类似的解释。注意，原假设$\mathbb{H}$ 定义了密度类

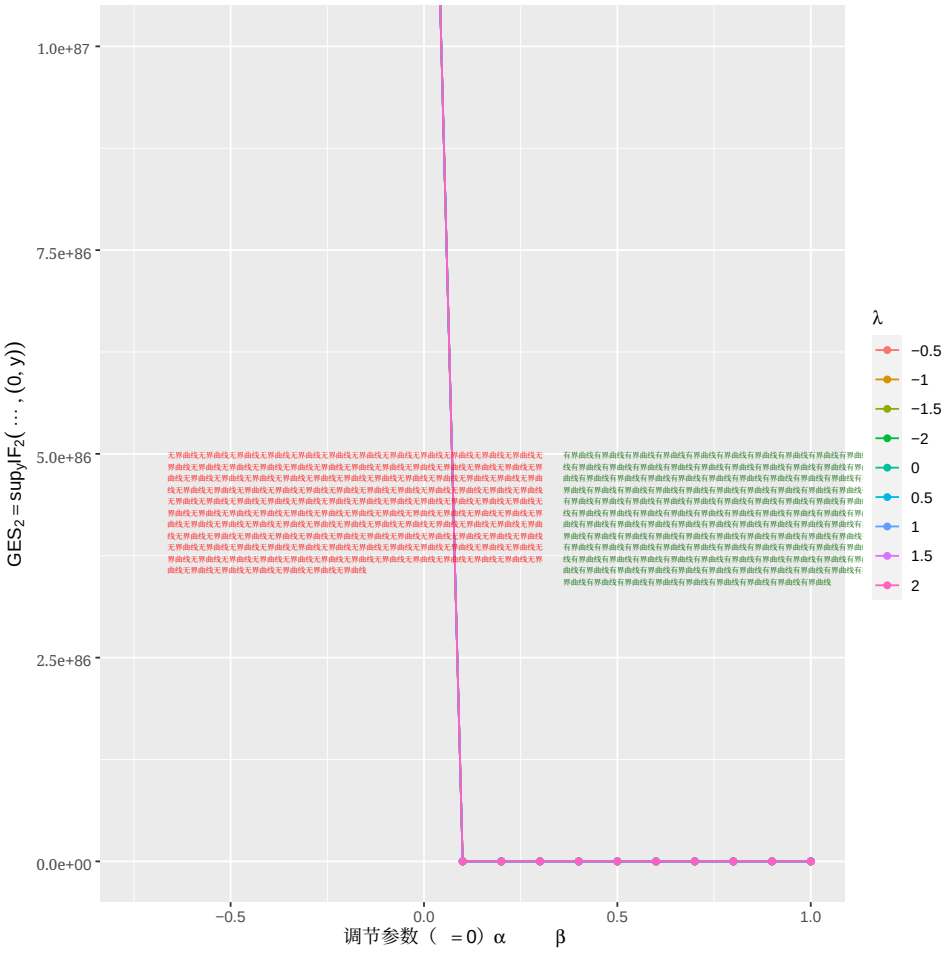

图4.1. 当 $(\alpha,\ \lambda,\ \beta)\ \epsilon S_1$，时的有界 GES 曲线，以及当 $\alpha \le 0$ 时的无界曲线。

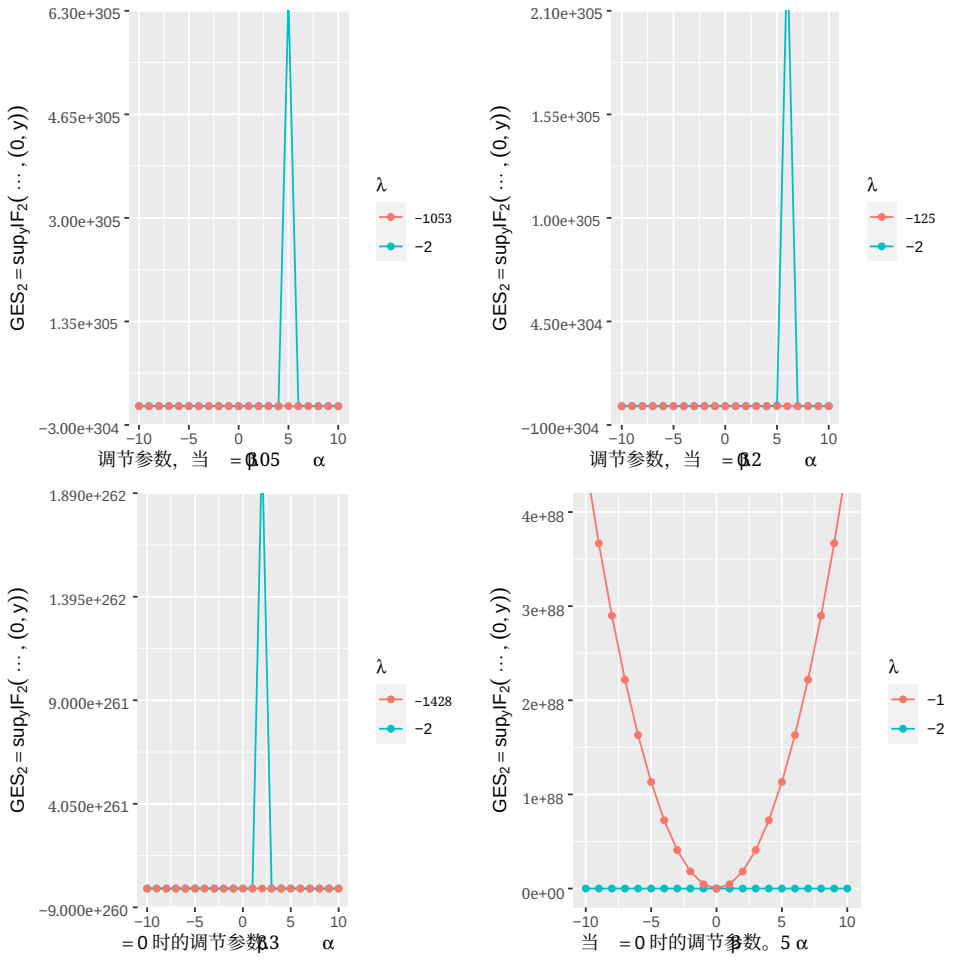

图 4.2。 当 $(\alpha,\ \lambda,\ \beta) \in S_2$, 时的有界 GES 曲线，以及当 $\lambda \neq \frac{1}{\alpha-1}$ 时的一些无界曲线。

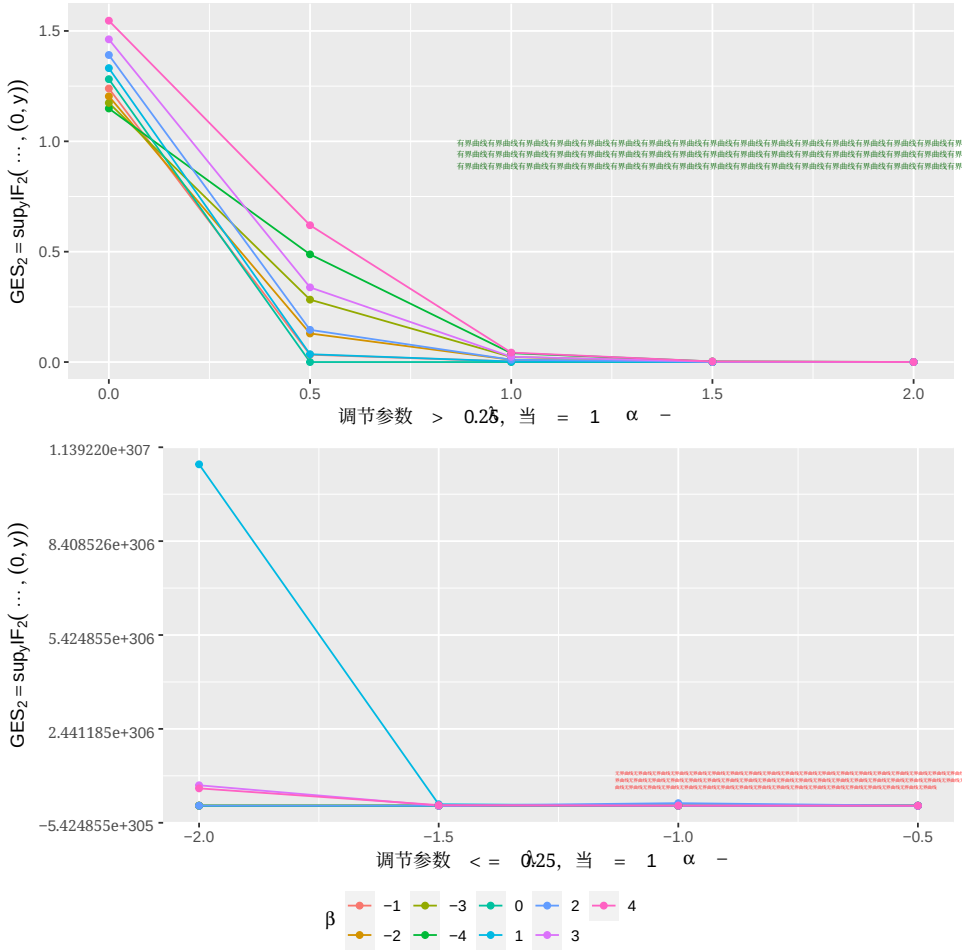

图4.3。 当 $(\alpha,\ \lambda,\ \beta) \in S_3$, 时的有界 GES 曲线，以及当 $\lambda \leq -0.25$ 时的一些无界曲线。

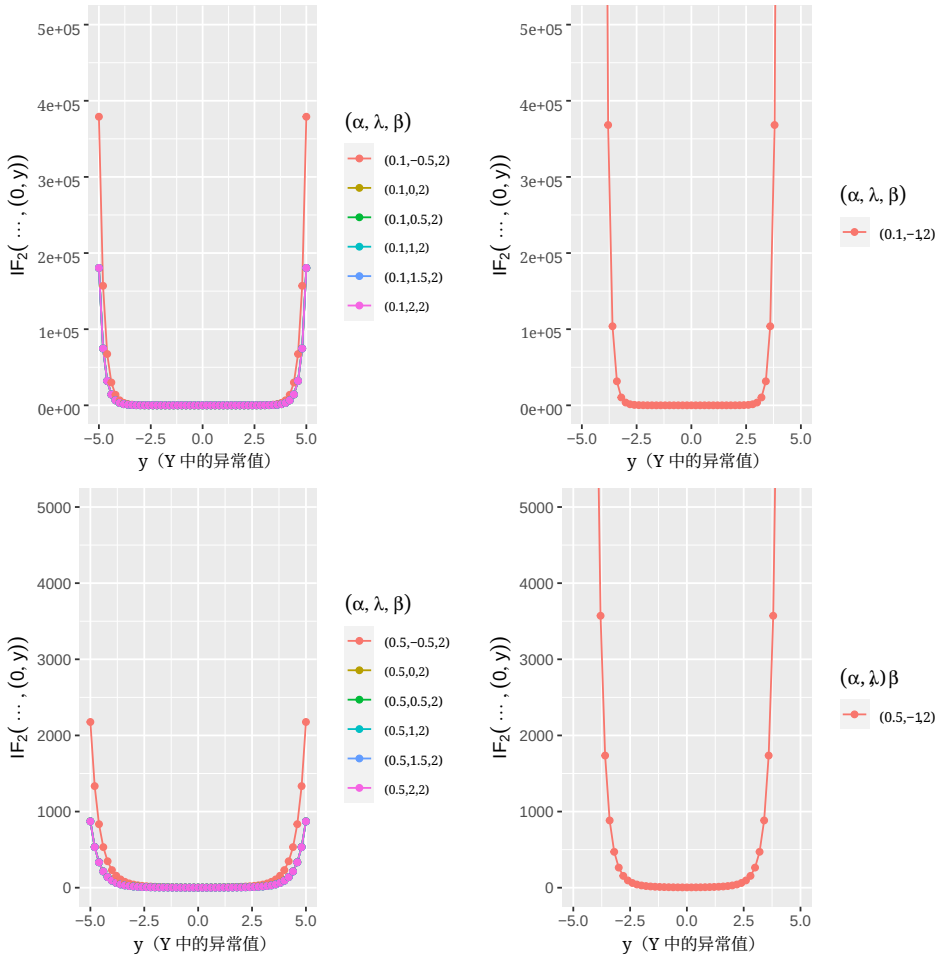

图 4.4. 当 $(\alpha, \lambda, \beta) \in S_4$, 时的有界影响函数$_2$， 以及当 $\lambda(1-\alpha) \leq -0.5$时的一些无界曲线。

$\mathcal{H} := \{f_{x,\ y} : f_{x,\ y} = f_x f_y$ 对所有 $x, y\}$均成立。原假设下的联合密度受到污染后变为

$$f^{\epsilon}_{x,y} = (1-\epsilon) f_x f_y + \epsilon \delta_{x_0}(x)\delta_{y_0}(y), \tag{4.12}$$

此时它不再属于检验类 $\mathcal{H}$。非稳健距离（例如 *Kullback-Leibler* 散度）对污染比例非常敏感，会扩大受污染密度与真实密度（原假设下）之间的差异。因此，检验水平会膨胀。然而，使用稳健统计距离可以减轻这种差异，从而使检验水平保持稳定。通过针对特定分布族的图示，可以进一步阐明影响函数在揭示这些非参数检验稳健性方面的作用。

4.2. 水平与功效影响函数

在上一小节中，我们在原假设下通过二阶影响函数分析考察了 $I_{D^{(k)}_{\phi}}$ 和 $T_{D^{(k)}_{\phi}}$ 的稳健性。然而，一个重要问题仍未解决：当原假设和邻近备择下的分布受到污染时，这类检验的水平和功效会受到怎样的影响？答案并不简单。一种可能的方法是评估与这类检验相对应的水平和功效泛函的影响函数。然而，主要挑战在于，在污染分布下，检验大小和功效通常难以得到闭式表达式。尽管如此，在某些正则条件下仍可作出适当的近似。本节将更详细地探讨这些方面。

假设原假设和邻近备择下的密度受到如下方式的污-染:

$$f^{\epsilon}_{x,y} = f_x f_y + \frac{\epsilon}{\sqrt{n h_n^{1/2}}}\Delta_{x,y} \text{ and } f^{(P_n)}_{x,y} = f^{(n)}_{x,y} + \frac{\epsilon}{\sqrt{n h_n^{1/2}}}\Delta'_{x,y} \text{ for all } x, y \tag{4.13}$$

分别为 $\epsilon \geq 0$ 和 $\sum_{x=0}^{1}\int_{f_y>0}\Delta'_{x,y}dy = 0$。这里允许 $\Delta'_{x,y}$ 依赖于样本量 n。邻近备择下的密度 $f^{(n)}_{x},_{y}$ 定义见（3.29）。当 $\epsilon = 0$时，两种情形下均不存在污染。特别地，当 $\Delta_{x,y} = \Delta'_{x,y}$ 对所有 x, y均成立，且 $d = 0$时，可得 $f(P_n)_{x,y} = f^{\epsilon}_{x,y}$对所有 x, y均成立。现在，考虑两种情形下的点质量污染，定义为 $\Delta_{x,y} = (\delta_{x_0}(x)\delta_{y_0}(y) - f_x f_y)$ 和 $\Delta'_{x,y} = (\delta_{x_1}(x)\delta_{y_1}(y) - f^{(n)}_{x},_{y})$，其中 $t_0 = (x_0, y_0)$ 和 $t_1 = (x_1, y_1)$。在这些污染下，水平和功效泛函分别为

$$\alpha(f^{\epsilon}_{X,Y}, t_0) = \mathbb{P}_{f^{\epsilon}_{X,Y}}\left\{T_{D^{(k)}_{\phi}} > \tau_c\right\} \text{ and } \pi(f^{(P_n)}_{X,Y}, t_1) = \mathbb{P}_{f^{(P_n)}_{X,Y}}\left\{T_{D^{(k)}_{\phi}} > \tau_c\right\}. \tag{4.14}$$

按照 Hampel（1986） [2],水平与功效影响函数被定义为 a s

$$\mathcal{LIF}(\alpha, t_0) = \lim_{n\longrightarrow\infty}\left[\frac{\partial\alpha(f^{(\epsilon)}_{X,Y}, t0)}{\partial\epsilon}\right]_{\epsilon=0}, \quad \mathcal{PIF}(\pi, t_1) = \lim_{n\longrightarrow\infty}\left[\frac{\partial\pi(f^{(P_n)}_{X,Y}, t_1)}{\partial\epsilon}\right]_{\epsilon=0}. \tag{4.15}$$

为了评估这些水平与功效影响函数，如（4.15）中所述，我们需要渐近

分布，对象为 $T_{D_\phi^{(k)}}$，且同时在污染零假设与受污染的邻接备择假设之下。前者已在定理 3.3 中导出。我们的下一个目标是求得 $T_{D_\phi^{(k)}}$ 在 $f^{(P_n)}{}_{X,\ Y}$ 下的渐近分布。为此，我们定义

$$I^{(P_n)}_{D_\phi^{(k)}} = D_\phi^{(k)}(f^{(P_n)}_{X,Y}, f^{(P_n)}_X f^{(P_n)}_Y), \tag{4.16}$$

其中 $f_{(P_n)X}$和 $f^{(P_n)}{}_Y$表示联合密度 $f_{(P_n)X,Y}$对应的边缘密度，具体含义由上下文确定。下面的引理给出了$I^{(P_n)}{}^{D_\phi^{(k)}}$ 的渐近表达式。

引理 4.1. 在假设 *(A1)* 成立时，有

$$I^{(P_n)}_{D_\phi^{(k)}} = \frac{k^2}{2nh_n^{1/2}} \sum_{x=0}^{1} \int (f_x f_y)^{2k} \left\{ d\left(\frac{\Delta_x}{f_x} + \frac{\Delta_y}{f_y} - \frac{\Delta_{x,y}}{f_x f_y} \right) + \epsilon \left(\frac{\Delta'_x}{f_x} + \frac{\Delta'_y}{f_y} - \frac{\Delta'_{x,y}}{f_x f_y} \right) \right\}^2 \phi''\left(f_x^k f_y^k \right) dy + o\left(\frac{(d+\epsilon)^2}{nh_n^{1/2}} \right). \tag{4.17}$$

现在，我们给出 $I^{(P_n)}{}^{D_\phi^{(k)}}$ 在 $f^{(P_n)}{}_{X,Y}$ 下的渐近分布。

定理 4.2. 假设 *(A1)-(A5)* 成立，且对某个 $C^* > 0$，有 $0 \le \max\{d, \epsilon\} \le C^* \sup_{x,y} |f_{x,y} - f_x f_y|$。则可得

$$T_{D_\phi^{(k)}} - \frac{k^2}{2\sigma_\phi} \sum_{x=0}^{1} \int (f_x f_y)^{2k} \left\{ d\left(\frac{\Delta_x}{f_x} + \frac{\Delta_y}{f_y} - \frac{\Delta_{x,y}}{f_x f_y} \right) + \epsilon \left(\frac{\Delta'_x}{f_x} + \frac{\Delta'_y}{f_y} - \frac{\Delta'_{x,y}}{f_x f_y} \right) \right\}^2 \phi''\left(f_x^k f_y^k \right) dy \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, 1\right) \tag{4.18}$$

在 $f^{(P_n)}{}_{X,Y}$下，当 $n \to \infty$时。

注 4.2. 检验统计量在 $f^{(P_n)}{}_{X,Y}$ 下的分布与定理 *3.3* 中的结果一致，当 $\epsilon = 0$.

我们现在可以推导 LIF 与 PIF 的表达式。在真实零假设和相邻分布的 ϵ-受污染版本下，水平和功效的精确值通常难以计算。然而，在适当的正则条件下，可以求得它们的极限值。为此，必须交换极限与求导的次序，而这仅在特定假设下才成立。这些结果在下一定理中正式陈述。为计算功效影响函数，我们取

$$\Delta'_{x,y} = \delta x_1(x) \delta_{y_1}(y) - f^{(P_n)}_{x,y}. \tag{4.19}$$

定理 4.3. 设序列 $\left\{\frac{\partial}{\partial\epsilon}\pi(f^{(P_n)}{}_{X,Y},t_1)\right\}$ 对所有 $\epsilon\in[0,1]$ 在固定的 $t_1=(x_1,y_1)$ 处一致收敛，且定理 *4.2* 的条件成立。则我们有如下结果。

(a) 功效影响函数为

$$\mathcal{PIF}(\pi,t_1)=\frac{d}{\sigma_\phi}\sum_{x=0}^{1}\int k^2(f_xf_y)^{2k}\left(\frac{\Delta_x^*}{f_x}+\frac{\Delta_y^*}{f_y}-\frac{\Delta_{x,y}^*}{f_xf_y}\right)\left(\frac{\Delta_x}{f_x}+\frac{\Delta_y}{f_y}-\frac{\Delta_{x,y}}{f_xf_y}\right)\phi''\left(f_x^kf_y^k\right)dy$$
$$\times\phi_1\left(\tau_c-\frac{d^2}{2\sigma_\phi}\mathcal{IF}_2(I_{D_\phi^{(k)}},f_Xf_Y,t_0)\right) \tag{4.20}$$

where $\Delta_{x,y}^*=(\delta_{x_1}(x)\delta_{y_1}(y)-f_xf_y)$, $\Delta_x^*=(\delta_{x_1}(x)-f_x)$ *and* $\Delta_y^*=(\delta_{y_1}(y)-f_y)$.

(b) 当 $t_0=t_1$ 时，我们得到

$$\mathcal{PIF}(\pi,t_1)=\frac{d}{\sigma_\phi}\mathcal{IF}_2(I_{D_\phi^{(k)}},f_Xf_Y,t_1)\cdot\phi_1\left(\tau_c-\frac{d^2}{2\sigma_\phi}\mathcal{IF}_2(I_{D_\phi^{(k)}},f_Xf_Y,t_1)\right). \tag{4.21}$$

(c) 水平影响函数为

$$\mathcal{LIF}(\pi,t_0)\equiv 0. \tag{4.22}$$

注 4.3. 功效影响函数的稳定性与原假设下 $I_{D_\phi^{(k)}}$ 的二阶影响函数成正比。然而，$\mathcal{LIF}$ 恒等于零。与影响函数类似，水平影响泛函也无法揭示第一类错误的稳健性，而实际情况并非如此。模拟研究表明，该检验类的水平稳健性也取决于 ϕ-函数的选择。

4.3. $I_{D_\phi^{(k)}}$ 的渐近崩溃点分析

此前，我们通过影响函数分析研究了 $I_{D_\phi^{(k)}}$ 的稳健性。影响函数定义为沿点质量方向的方向导数，描述了泛函在该点受到无穷小污染时的稳定性。尽管这是一个有用的概念，但有时它无法揭示泛函本应显而易见的稳健性。由于影响函数完全是一个局部概念，因此应辅以全局稳健性度量。崩溃点便是这样一种全局度量，它量化了数据集中离群观测所占的最大比例，也就是在统计泛函产生异常值之前所能容忍的比例。关于这一非参数设定中离群观测的概念，已在注 4.1。

有时，使用 Hampel（1986） [2] 给出的定义计算实际崩溃点难以实现。然而，在许多情况下，渐近崩溃点往往更容易计算。

在给出我们的工作定义之前，我们先介绍一些有用的记号。设 $k_{XY,m}(x,y):=k_{xy,m}$ 为一列污染密度，并设 $k_{x,m}=\int_{f_y>0}k_{xy,m}dy$, $k_{y,m}=\sum_{x=0}^{1}k_{xy,m}$

为其边缘分布。真实密度 $f_{x,y}$, f_x 和 f_y 分别受到如下污染

$$\begin{aligned} f_{x,y}^m &= (1-\epsilon) f_{x,y} + \epsilon k_{xy,m}, \\ f_x^m &= (1-\epsilon) f_x + \epsilon k_{x,m}, \\ f_y^m &= (1-\epsilon) f_y + \epsilon k_{y,m} \end{aligned} \tag{4.23}$$

其中 $0 \le \epsilon \le 1$。通常，对于所有 x，我们将 $f_{x,\ y}^m = f_x^m f_{y|x}^m$ 表示为 y，其中 $f_{y|x}^m$ 是给定 x 时 y 的条件密度，由 $f_{x,\ y}^m$ 得到。其他密度也采用类似的记号。回顾 $I_{D_\phi^{(k)}} = D_\phi^{(k)}(f_{X,\ Y}, f_X f_Y)$，并进一步定义

$$I_m = D_\phi^{(k)}(f_{X,Y}^m, f_X^m f_Y^m) \text{ and } I'_m = D_\phi^{(k)}(f_{X,Y}^m, f_X f_Y). \tag{4.24}$$

迄今在参数设定中使用的渐近崩溃点定义并不直接适用于此处，因为问题的性质有所不同。此处的目标是针对给定的联合密度计算互信息。在实践中，自然会使用核密度估计来近似互信息。当广义互信息对某一特定的污染密度序列具有稳健性时，沿该序列变化时，广义互信息不应发生剧烈波动。从这一角度来看，式（ I_m （4.24）中的受污染互信息不应在污染比例 ϵ 处发生崩溃，前提是 $\lim_{m\to\infty} I_m$ 存在。

在此背景下，我们现在给出当前语境中崩溃点的工作定义。

定义 4.1. 给定真实密度 $f_{X,\ Y}$，若广义互信息泛函$I_{D_\phi^{(k)}}$ 的受污染版本 I_m 在 $I_{D_\psi^{(k)}} < \infty$时不收敛，则称其在 ϵ 处渐近崩溃。若 ϵ^* 满足以下条件，则称其为 $I_{D_\phi^{(k)}}$ 的渐近崩溃点：

$$\epsilon^* = \inf\left\{\epsilon : I_m \ does\ not\ converge\ but\ I_{D_\psi^{(k)}} < \infty\right\}. \tag{4.25}$$

由定义4.1可知，$I_{D_\phi^{(k)}}$ 只有在有界时，才可能在 ϵ 处崩溃。排除满足 $I_{D_\phi^{(k)}} = \infty$ 的散度可能会在一定程度上限制这一定义的普适性，但不会显著影响分析，因为这只排除了一些在统计上无意义的特殊情况。

遵循 Park 和 Basu（2004） [33] 以及 Roy 等人（2023） [34], 的做法，我们采用类似的方法，利用定义 $I_{D_\phi^{(k)}}$ 4.1来确定其渐近崩溃点。下面列出以下假设：

(BP1) $\int \min\{f_{y|x}, k_{y|x,m}\} dy \longrightarrow 0$ 当 $m \to \infty$ 时，对于 $x = 0$, 1，(BP2) $\int \min\{f_{y|x}, f_y^m\} dy \longrightarrow 0$ 当 $m \to \infty$ 时，对于 $x = 0$, 1，(BP3) $\int \min\{f_y, k_{y|x,m}\} dy \longrightarrow 0$ 当 $m \to \infty$ 时，对于 $x = 0$, 1，(BP4) $\phi(0)$ $\phi'(0)$ 为有限值，且 Y 具有有界支撑集，(BP5) 对所有 $I_m \le I'_m$， $m \ge 1$,

(BP6) 存在 $\tilde{\epsilon} \in [0, \frac{1}{2}]$，使得对所有 $\epsilon < \tilde{\epsilon}$，

$$\liminf_{m\to\infty} D_\phi^{(k)}(\epsilon k_{XY,m}, f_X^m f_Y^m) > \limsup_{m\to\infty} \sum \int \left[\phi((\epsilon k_{xy,m})^k) - (\epsilon k_{xy,m})^k \phi'(0) \right] dy$$
$$+ \sum \int \left[((1-\epsilon) f_{xy})^k \phi'(0) - \phi((f_x f_y)^k) \right.$$
$$\left. - \left\{ ((1-\epsilon) f_{xy})^k - (f_x f_y)^k \right\} \phi'((f_x f_y)^k) \right] dy. \tag{4.26}$$

假设 (BP1) 确保条件污染密度 $k_{y|x,m}$ 关于条件密度 $f_{y|x}$ 在 $x = 0,1$ 时是渐近奇异的。假设 (BP2)确保污染密度 f_y^m对于所有 x 也关于 $f_{y|x}$ 是渐近奇异的。同样，假设 (BP3) 确保条件污染密度$k_{y|x,m}$ 对于所有 x 关于边缘密度 f_y 是渐近奇异的。

这三个条件共同构成可能导致 I_m 在足够大的 m 时发散的最坏情形。假设 (BP4) 对 ϕ-函数施加了某些技术条件，这些条件在大多数情形下通常都能得到满足。它还假设 Y 的支撑集是有界的，从而确保即使在散度度量中将某一密度替换为零的退化情形下，某些积分仍然保持有限。

假设(BP5) 是一项尚无正式证明的技术性要求。它指出，受污染的联合密度与未受污染的边缘分布之间的 $\mathscr{B}$-互信息，应至少与受污染联合分布及其自身边缘分布之间的互信息一样大。如果(BP5) 被违反，则当 X 与 Y 在受污染的联合密度 $f_{x,y}^m$下相互独立时，将产生矛盾。

最后，假设 (BP6)规定了后续结果仍然成立时污染程度的极限水平。注意，如果 X 和 Y 相互独立，则假设(BP1) 和(BP3)等价。

定理 4.4。 假设 *(BP1)–(BP6)* 在 $I_{D_\phi^{(k)}} < \infty$条件下成立。那么，广义互信息 $I_{D_\phi^{(k)}}$的渐近崩溃点至少为 $\min\left\{1/2, \tilde{\epsilon}\right\}$，其中$\tilde{\epsilon}$ 由假设 *(BP6)*定义。

定理 4.4 严重依赖假设 (BP6)，渐近崩溃点可以通过求解满足假设 (BP6)的 ϵ 的隐式解得到。这在实践中并不容易。不过，我们可以做得更好。尽管需要附加条件，我们仍能得到渐近崩溃点的闭式表达式。接下来，我们将仿照 Roy 等人（2023） [34]，给出一组满足假设 (BP6) 的充分条件。为此，首先我们定义$M_{g,f} = \sum \int g^k \phi'(f^k)$，并从以下引理开始，该引理为 ϕ生成的扩展布雷格曼散度提供一个下界（或上界）。随后，我们将进一步利用此引理给出 $I_{D_\phi^{(k)}}$ 渐近崩溃点的第二种形式。

引理 4.2. 假设 $M_{f,f} \gtreqless M_{g,f}$，其中 $g = f_{X,Y}$且 $f = f_X f_Y$。那么，对于 $\epsilon^k \lesseqgtr 1 + \frac{\sum\int[\phi(g^k - \phi f^k])(),,}{M_{f\ f} - M_{g\ g}}$，当 $g \neq f$时，$D_\phi^{(k)}(\epsilon g, f) \gtreqless D_\phi^{(k)}(\epsilon g, g)$。

当散度被定义为某些调节参数的极限时，在适当的假设下取相同的极限，例如，$M_{f,\ f} \leq M_{g,\ f}$ 以及 ϵ 的界亦然。现在，我们陈述假设(BP6)的一些额外充分条件。

(BP7) 设密度$g = f_{X,\ Y}$、 $f = f_X f_Y$、 $f^m = f_X^m f_Y^m$与$k_m = K_{XY,\ m}$满足

$$M_{f,f} \leq M_{g,f} \text{ and } \liminf_{m\to\infty} M_{f^m,f^m} \geq \limsup_{m\to\infty} M_{k_m,f^m}. \tag{4.27}$$

此外，对于所考虑的任何联合分布， X， Y 均不独立。

(BP8) 对所有 $\epsilon < \tilde{\epsilon}$，下列成立

$$\begin{aligned}\liminf_{m\to\infty} D_\phi^{(k)}(\epsilon k_m, k_m) > \limsup_{m\to\infty} D_\phi^{(k)}(\epsilon k_m, 0) \\ + D_\phi^{(k)}((1-\epsilon) f_{X,Y}, f_{X,Y}) - D_\phi^{(k)}((1-\epsilon) f_{X,Y}, 0).\end{aligned} \tag{4.28}$$

在下一个结果中，我们将看到，假设(BP7)与(BP8)将共同蕴含假设(BP6)。此外，渐近崩溃点的界将利用引理4.2导出。我们定义如下量：

$$\epsilon_1 = \left[1 + \limsup_{m\to\infty} \frac{\sum\int\left\{\phi[(k_{xy,m})^k] - \phi[(f_x^m f_y^m)^k]\right\}}{(M_{f^m,f^m} - M_{k_m,k_m})}\right]^{1/k}, \tag{4.29}$$

$$\epsilon_2 = 1 - \left[1 + \frac{\sum\int\left[\phi(f_{x,y}^k) - \phi(f_x^k f_y^k)\right]}{M_{f_X f_Y, f_X f_Y} - M_{f_{X,Y}, f_{X,Y}}}\right]^{1/k}. \tag{4.30}$$

定理 4.5. 设假设*(BP1)*-*(BP5)*与*(BP7)*-*(BP8)* 成立。则$I_{D_{\underline{\phi}}^{(k)}}$的渐近崩溃点至少为$\min\left\{\epsilon_1,\ \epsilon_2,\ {}_{1\,2}\right\}$。

定理 4.4 可以很容易地推广到 GSB散度。由于它很大程度上依赖于假设(BP6)，我们给出一组蕴含假设(BP6)的充分条件。根据引理 4.2 我们发现 $D_*(\epsilon g, f) \geq D_*(\epsilon g, g)$ 对于

$$\epsilon^A \leq 1 + \frac{\sum\int\left[AB(e^{\beta g^A} - e^{\beta f^A}) + A(g^{1+\alpha} - f^{1+\alpha})\right]}{\sum\int\left[AB\{(\beta f^A)e^{\beta f^A} - (\beta g^A)e^{\beta g^A}\} + (1+\alpha)(f^{1+\alpha} - g^{1+\alpha})\right]}, \tag{4.31}$$

前提是

$$\sum\int\left[AB\beta e^{\beta f^A}(f^A - g^A) + (1+\alpha)(f^A - g^A)f^B\right] \geq 0 \text{ with } B > 0, \tag{4.32}$$

其中 $g = f_{X、Y}$、 $f = f_X f_Y$，以及 X、Y并不独立。当散度定义为极限时，(4.31) 和 (4.32) 中的所有条件也将使用相应的极限重新表述。如果 $A = 0$，则 (4.31) 中的表达式将与 ϵ 无关，因此可以舍去。因此，不失一般性，我们将默认假设 $A > 0$，而不再考虑退化情形。另一种退化情形发生在 $B \downarrow 0+$ 时。此外，如果 $B \downarrow 0+$，但 $\alpha > -1$，那么 $A > 0$，这意味着对于 $\epsilon = 0$，当 $\sum\int(f^{1+\alpha} - g^{1+\alpha}) \geq 0$ 且 X、Y 并不独立时，$\lim_{B\downarrow 0+} D_*(\epsilon g, f) \geq \lim_{B\downarrow 0+} D_*(\epsilon g, g)$。在这种情况下，调节参数属于一个足够小的邻域，该邻域包含 $S_5 = \{(\alpha, \frac{\alpha}{1-\alpha}, \beta) : \alpha > -1, \beta \in R\}$。同样，$\alpha = -1$ 和 $B \downarrow 0+$ 意味着 $A \downarrow 0+$。因此，(4.31) 中的表达式将与 ϵ 无关，故将被舍去。

推论4.1。 当 $\beta = 0, \alpha \geq 0$ 时，*GSB*散度就变成了 *S*-散度。在这种情况下$D_*(\epsilon g, f) \geq D_*(\epsilon g, g)$ 对于 $\epsilon^A \leq \frac{B}{1+\alpha}$，前提是

$$\sum\int \left\{f^{1+\alpha} - g^A f^B\right\} \geq 0 \ \ with\ B > 0, \tag{4.33}$$

其中 $g = f_{X,Y}, f = f_X f_Y$和 X, Y并非相互独立。若假设 $\sum\int f^{1+\alpha} \geq \sum\int g^{1+\alpha}$，则（*4.33*）可由简单应用 ***Hölder*** 不等式推出。*S*-散度的$B \downarrow 0+$ 退化情形也可以像之前那样进行刻画。

推论4.2。 对于密度幂散度， $A = 1, B = \alpha$。因此 $D_*(\epsilon g, f) \geq D_*(\epsilon g, g)$ 对于$\epsilon \leq \frac{\alpha}{1+\alpha}$ ，前提是

$$\sum\int \left\{f^{1+\alpha} - g^{1+\alpha}\right\} \geq 0, \tag{4.34}$$

其中 $g = f_{X,\ Y}, f = f_X f_Y$和 X, Y并不独立。

推论4.3。 当 $A = 1 + \lambda, B = -\lambda$ 时，它就成为幂散度族。由于 $A > 0$，即 $\lambda > -1$。结合退化情形 $\lambda = -1$，我们得到 $D_*(\epsilon g, f) \geq D_*(\epsilon g, g)$ 当

$$\epsilon \leq (-\lambda)^{\frac{1}{1+\lambda}} \ \ for\ -1 \leq \lambda \leq 0, \tag{4.35}$$

使得 X，Y 并不独立。由于引理 *4.2* 中的条件显然成立，因此无需进一步限制。

推论 4.4。 当 $\alpha = -1, \lambda = 0$ 时，它就成为调节参数为 β 的缩放 *BED* 族。这里 $A = 1, B = -1$。假设 $\beta \neq 0$，则 $D_*(\epsilon g, f) \geq D_*(\epsilon f, f)$ 使得

$$\epsilon \leq 1 + \frac{\sum\int\left(e^{\beta g} - e^{\beta f}\right)}{\beta\sum\int\left(fe^{\beta f} - ge^{\beta g}\right)} \ for\ \beta\sum\int\left(e^{\beta f}(f-g)\right) \geq 0, \tag{4.36}$$

且 X，Y不相互独立。当 $\beta = 0$时，引理*4.2* 在 $\beta \to 0$ 的极限下，对 $\epsilon \leq \frac{1}{2}$ 满足 $\sum\int f^2 > \sum\int g^2$ 且 X的情况Y不相互独立。

在定理4.5中，我们发现 I_{D_*} 的渐近崩溃点至少为 $\min\{\epsilon_1, \epsilon_2, _{12}\}$，其中 –

$$\epsilon_1 = \left(1 + \limsup_{m\to\infty} \frac{\sum\int\left[AB\{e^{\beta k_{xy,m}^A} - e^{\beta(f_x^m f_y^m)^A}\} + A(k_{xy,m}^{1+\alpha} - (f_x^m f_y^m)^{1+\alpha})\right]}{\sum\int\left[AB\beta\{(f_x^m f_y^m)^A e^{\beta(f_x^m f_y^m)^A} - (k_{xy,m}^A)e^{\beta k_{xy,m}^A}\} + (1+\alpha)((f_x^m f_y^m)^{1+\alpha} - k_{xy,m}^{1+\alpha})\right]}\right)^{1/A}, \tag{4.37}$$

$$\epsilon_2 = 1 - \left(1 + \frac{\sum\int\left[AB\{e^{\beta f_{x,y}^A} - e^{\beta(f_x f_y)^A}\} + A(f_{x,y}^{1+\alpha} - (f_x f_y)^{1+\alpha})\right]}{\sum\int\left[AB\{\beta(f_x f_y)^A e^{\beta(f_x f_y)^A} - (\beta f_{x,y}^A)e^{\beta f_{x,y}^A}\} + (1+\alpha)((f_x f_y)^{1+\alpha} - f_{x,y}^{1+\alpha})\right]}\right)^{1/A}. \tag{4.38}$$

事实证明，GSB 散度族的渐近崩溃点结果适用于 $A > 0$ 和 $B > 0$，即 $-\frac{1}{1-\alpha} \le \lambda \le \frac{1}{1-\alpha}$ 和 $\alpha \ge -1$。

推论 4.5。 对于 *S*-散度族 $\beta = 0$，且

$$\epsilon_1 = \left(\frac{B}{1+\alpha}\right)^{1/A} \; and \; \epsilon_2 = 1 - \left(\frac{B}{1+\alpha}\right)^{1/A}. \tag{4.39}$$

因此， I_{D_*} 的渐近崩溃点将为

$$\min\left\{\left(\frac{B}{1+\alpha}\right)^{1/A}, 1 - \left(\frac{B}{1+\alpha}\right)^{1/A}, \frac{1}{2}\right\} \; with \; A > 0, B > 0, \tag{4.40}$$

在定理*4.5*的假设下。

推论 4.6。 接下来，我们考虑幂散度族，即 $\alpha = \beta = 0$。那么，在定理 *4.5*的假设下， I_{D_*} 的渐近崩溃点为

$$\min\left\{\left(-\lambda\right)^{1/(1+\lambda)}, 1 - \left(-\lambda\right)^{1/(1+\lambda)}, \frac{1}{2}\right\} \; for \; -1 \le \lambda \le 0. \tag{4.41}$$

推论 4.7。 当 $\lambda = \beta = 0$时，它就是 *DPD*。那么，在定理*4.5* 的假设下， I_{D_*} 的渐近崩溃点变为

$$\min\left\{\frac{\alpha}{1+\alpha}, \frac{1}{1+\alpha}, \frac{1}{2}\right\}. \tag{4.42}$$

显然，随着 α 增大，基于 DPD 的广义互信息的渐近崩溃点会显著提高；而增大 λ 则会产生相反的效果，使渐近崩溃点降低。Roy 等人（2023）的图 1 [34] 展示了范围广泛的 (α, λ)，在此范围内，基于 S-散度得到的最小散度估计具有很强的稳健性。注意，如果 $\beta \neq 0$， I_{D_*} 的渐近崩溃点将同时取决于

Y 的密度及其污染序列。因此，在一般情况下，除非考虑具体示例，否则很难确定能产生较高渐近崩溃点的调节参数。这里需要注意的是，渐近崩溃点不取决于 Y 的维数。

5. 数值研究

5.1. 模拟结果

（A）未污染零假设：本节报告以下几组模型的模拟结果。

模型 0： （零模型） $Y_0 \sim \mathrm{N}(0,\ 1)$,

模型 1： $Y_1 \sim \mathrm{N}(0,\ (1+a)^2)$，其中 $a = 0.75$,

模型 2： $Y_1 \sim (1-a)\mathrm{N}(-1,\ 1) + a\mathrm{N}(1,\ 1)$，其中 $a = 0.6$。

(B) 污染零假设：为研究第一类错误的稳健性，对第一个样本进行污染，使其服从 $(1-\epsilon)\mathrm{N}(0,\ 1) + \epsilon\mathrm{N}(5,\ 2)$，其中 $\epsilon = 0.05$，0.10，0.12，0.15；第二个样本则来自 N(0，1)。

从两个分布中分别抽取大小为 $n_0 = 100$ 和 $n_1 = 100$ 的样本，进行 500 次重复。检验在 5% 的名义显著性水平下进行，调节参数选为 $\alpha = (0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1)^T$，$\beta = (0, -0.05)^T$和 $\lambda = (-0.5, -0.3, -0.2, -0.1, 0, 0.25, 0.50, 1)$。

由于此设置中的分布具有无界支撑集，我们无法使用渐近零分布来模拟幂散度族的检验功效。在这种情况下，我们需要使用排列检验，并采用 Guha 等人（2021）的算法 [25]。在这些情况下，我们考虑 500 次排列。不过，连续分布的无界支撑集不应妨碍我们使用渐近零分布来处理广义 S-Bregman 散度族中的其他成员，除了 $\alpha = 0$。使用渐近零分布可以大幅减轻计算负担。在数值研究中，我们选择的核和带宽序列如下

$$K(u) = \frac{3}{4}(1-u^2)\mathbb{1}\left\{|u| \le 1\right\} \text{ and } h_n = 1.06 \times sd(Y)n^{-1/5}, \tag{5.1}$$

其中 $n = n_0 + n_1$ 为合并样本量。该核称为 Epanechnikov 核，并满足假设(A3)。众所周知，核密度估计通常对核的选择具有稳健性，因此有时采用有界且对称的核会很有用。核密度估计的最优带宽序列取决于具体问题。不过，此处考虑的经验法则最优带宽序列 h_n 最适用于原始分布为高斯分布或对称分布的情形；即使分布并非高度偏斜，其表现通常也相当不错。更多详情请参见 Silverman（2018） [30] 。

在表 5.1 和表 5.2中，报告了纯模型下各检验的观测水平。我们看到，当 $\beta = -0.05$ 时，有时检验会变得保守，即产生非常低的观测水平。当零模型受到污染时，观测水平会随污染程度增加而上升。不过，除了 $\alpha = 0$ 之外，GSB 散度族的所有成员都表现出更好的稳健性。这些数值列于表 5.3 至表 5.10。如果第一个样本来自 模型 0 ，而第二个样本来自模型 1 或 模型 2，则拒绝原假设 $\mathbb{H}$ 。此类情况下的拒绝比例（即观测功效）列于表 5.11 至表5.14。有时我们会看到，对于 $\beta < 0$，观测功效相对较低，这可能是由于分布的收敛速度较慢。不过，随着样本量增加，观测功效会有所提高。

表 5.1。 在 模型 0 下的拒绝比例，使用 $\beta = 0$

λ \ α	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
−0.5	0.058	0.054	0.040	0.034	0.034	0.036	0.032	0.026	0.026	0.022	0.024
−0.3	0.058	0.040	0.034	0.034	0.032	0.034	0.030	0.026	0.026	0.022	0.024
−0.2	0.058	0.036	0.030	0.032	0.032	0.034	0.030	0.026	0.026	0.022	0.024
−0.1	0.058	0.034	0.030	0.032	0.032	0.034	0.030	0.026	0.026	0.022	0.024
0.0	0.040	0.036	0.030	0.030	0.032	0.032	0.030	0.026	0.024	0.022	0.024
0.25	0.042	0.098	0.096	0.028	0.032	0.030	0.028	0.026	0.024	0.022	0.024
0.50	0.042	0.092	0.084	0.078	0.032	0.026	0.028	0.026	0.022	0.022	0.024
1.0	0.042	0.092	0.082	0.076	0.076	0.098	0.024	0.026	0.022	0.022	0.024

表 5.2。 在 模型 0 下的拒绝比例，使用 $\beta = -0.05$

λ \ α	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
−0.5	0.028	0.054	0.040	0.034	0.034	0.036	0.032	0.026	0.026	0.022	0.024
−0.3	0.018	0.012	0.008	0.008	0.006	0.004	0.004	0.004	0.004	0.004	0.024
−0.2	0.012	0.008	0.008	0.006	0.004	0.004	0.004	0.004	0.004	0.004	0.024
−0.1	0.012	0.008	0.006	0.006	0.004	0.004	0.004	0.004	0.004	0.004	0.024
0.0	0.036	0.010	0.006	0.006	0.004	0.004	0.004	0.004	0.004	0.004	0.024
0.25	0.034	0.036	0.042	0.006	0.004	0.004	0.004	0.004	0.004	0.004	0.024
0.50	0.034	0.036	0.036	0.040	0.004	0.004	0.004	0.004	0.004	0.004	0.024
1.0	0.054	0.040	0.036	0.036	0.036	0.052	0.004	0.004	0.004	0.004	0.024

表 5.3。 在 模型 0 下的拒绝比例，使用 $\beta = 0$ 当第一个样本中有 5% 的观测值来自 $\mathcal{N}(5, 2)$

λ \ α	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
−0.5	1	0.148	0.094	0.078	0.056	0.048	0.044	0.040	0.038	0.036	0.038
−0.3	1	0.094	0.076	0.058	0.052	0.046	0.040	0.040	0.038	0.036	0.038
−0.2	1	0.084	0.072	0.056	0.048	0.046	0.040	0.040	0.038	0.036	0.038
−0.1	1	0.084	0.070	0.056	0.048	0.046	0.040	0.040	0.038	0.036	0.038
0.0	0.996	0.124	0.064	0.054	0.046	0.044	0.040	0.036	0.036	0.036	0.038
0.25	0.998	0.226	0.222	0.056	0.046	0.044	0.038	0.036	0.036	0.036	0.038
0.50	0.998	0.208	0.184	0.164	0.048	0.044	0.038	0.036	0.036	0.036	0.038
1.0	0.998	0.260	0.188	0.162	0.146	0.258	0.038	0.036	0.036	0.036	0.038

表 5.4。 在 模型 0 下的拒绝比例，使用 $\beta = -0.05$ 当第一个样本中有 5% 的观测值来自 $\mathcal{N}(5, 2)$

λ \ α	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
−0.5	1	0.148	0.094	0.078	0.056	0.048	0.044	0.040	0.038	0.036	0.038
−0.3	1	0.024	0.016	0.014	0.014	0.008	0.006	0.004	0.004	0.004	0.038
−0.2	1	0.020	0.016	0.014	0.010	0.008	0.006	0.004	0.004	0.004	0.038
−0.1	1	0.018	0.014	0.012	0.010	0.008	0.006	0.004	0.004	0.004	0.038
0.0	0.996	0.024	0.012	0.012	0.010	0.008	0.006	0.004	0.004	0.004	0.038
0.25	0.998	0.074	0.102	0.012	0.010	0.008	0.004	0.004	0.004	0.004	0.038
0.50	0.998	0.070	0.062	0.060	0.010	0.008	0.004	0.004	0.004	0.004	0.038
1.0	0.998	0.086	0.060	0.058	0.056	0.140	0.004	0.004	0.004	0.004	0.038

表 5. **5.** 在 模型 0 下，当第一个样本中有 10% 的观测值来自 $\beta = 0$ 时的拒绝比例（$\mathcal{N}(5$,2)

λ \ α	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
−0.5	1	0.770	0.510	0.308	0.214	0.148	0.122	0.106	0.100	0.092	0.088
−0.3	1	0.578	0.390	0.250	0.180	0.136	0.120	0.106	0.100	0.092	0.088
−0.2	1	0.522	0.338	0.236	0.168	0.134	0.118	0.106	0.100	0.092	0.088
−0.1	1	0.488	0.310	0.220	0.162	0.132	0.116	0.106	0.100	0.092	0.088
0.0	1	0.522	0.300	0.212	0.146	0.130	0.116	0.104	0.098	0.092	0.088
0.25	1	0.664	0.402	0.220	0.146	0.122	0.114	0.104	0.098	0.092	0.088
0.50	1	0.632	0.520	0.428	0.158	0.120	0.110	0.104	0.096	0.092	0.088
1.0	1	0.668	0.522	0.424	0.336	0.502	0.108	0.100	0.096	0.090	0.088

表5.6. 在 模型 0 下，当第一个样本中有 10% 的观测值来自 $\beta = -0.05$ 时的拒绝比例（$\mathcal{N}(5, 2)$

λ \ α	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
−0.5	1	0.770	0.510	0.308	0.214	0.148	0.122	0.106	0.100	0.092	0.088
−0.3	1	0.214	0.130	0.090	0.064	0.054	0.042	0.036	0.036	0.032	0.088
−0.2	1	0.182	0.114	0.084	0.060	0.052	0.040	0.036	0.036	0.032	0.088
−0.1	1	0.172	0.100	0.076	0.058	0.048	0.040	0.036	0.036	0.032	0.088
0.0	1	0.198	0.096	0.066	0.058	0.048	0.040	0.036	0.036	0.032	0.088
0.25	1	0.356	0.246	0.070	0.054	0.046	0.040	0.036	0.036	0.032	0.088
0.50	1	0.316	0.248	0.204	0.054	0.040	0.038	0.036	0.034	0.032	0.088
1.0	1	0.354	0.254	0.196	0.156	0.308	0.036	0.034	0.034	0.032	0.088

表5.7. 在 模型 0 下，当第一个样本中有 12% 的观测值来自 $\beta = 0$ 时的拒绝比例（$\mathcal{N}(5, 2)$

λ \ α	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
−0.5	1	0.932	0.792	0.540	0.360	0.260	0.206	0.168	0.148	0.134	0.128
−0.3	1	0.848	0.648	0.434	0.300	0.236	0.194	0.164	0.146	0.134	0.128
−0.2	1	0.792	0.588	0.408	0.292	0.230	0.192	0.164	0.142	0.134	0.128
−0.1	1	0.750	0.540	0.392	0.280	0.228	0.184	0.164	0.142	0.134	0.128
0.0	1	0.754	0.524	0.364	0.268	0.226	0.184	0.162	0.142	0.134	0.128
0.25	1	0.882	0.516	0.350	0.246	0.210	0.182	0.160	0.140	0.134	0.128
0.50	1	0.858	0.756	0.650	0.260	0.200	0.176	0.160	0.140	0.134	0.128
1.0	1	0.844	0.748	0.618	0.506	0.606	0.170	0.154	0.140	0.134	0.128

表5.8. 在 模型 0 下，当第一个样本中有 12% 的观测值来自 $\beta = -0.05$ 时的拒绝比例（$\mathcal{N}(5, 2)$

λ \ α	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
−0.5	1	0.932	0.792	0.540	0.360	0.260	0.206	0.168	0.148	0.134	0.128
−0.3	1	0.492	0.290	0.168	0.124	0.094	0.076	0.070	0.068	0.064	0.128
−0.2	1	0.414	0.236	0.156	0.118	0.094	0.074	0.070	0.066	0.064	0.128
−0.1	1	0.364	0.214	0.150	0.114	0.092	0.072	0.070	0.066	0.064	0.128
0.0	1	0.392	0.200	0.140	0.104	0.090	0.072	0.070	0.066	0.064	0.128
0.25	1	0.610	0.318	0.130	0.096	0.088	0.070	0.068	0.066	0.062	0.128
0.50	1	0.546	0.428	0.332	0.104	0.076	0.066	0.066	0.066	0.062	0.128
1.0	1	0.582	0.430	0.318	0.262	0.446	0.068	0.066	0.066	0.062	0.128

表 5.　9. 在模型 0 下，使用 $\beta = 0$ 且第一个样本中有 15% 的观测值来自$\mathcal{N}(5$,2)

λ \ α	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
−0.5	1	0.994	0.972	0.892	0.710	0.530	0.394	0.300	0.254	0.234	0.232
−0.3	1	0.978	0.934	0.802	0.626	0.472	0.350	0.290	0.252	0.234	0.232
−0.2	1	0.972	0.906	0.778	0.600	0.446	0.344	0.284	0.252	0.234	0.232
−0.1	1	0.962	0.880	0.740	0.590	0.436	0.334	0.280	0.250	0.234	0.232
0.0	1	0.954	0.854	0.708	0.560	0.424	0.334	0.276	0.248	0.234	0.232
0.25	1	0.980	0.652	0.654	0.522	0.404	0.314	0.272	0.244	0.234	0.232
0.50	1	0.974	0.954	0.906	0.498	0.384	0.306	0.268	0.240	0.234	0.232
1.0	1	0.974	0.952	0.892	0.796	0.764	0.296	0.266	0.240	0.234	0.232

表 5.10. 在模型 0 下，使用 $\beta = -0.05$ 且第一个样本中有 15% 的观测值来自$\mathcal{N}(5,\ 2)$

λ \ α	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
−0.5	1	0.994	0.972	0.892	0.710	0.530	0.394	0.300	0.254	0.234	0.232
−0.3	1	0.872	0.694	0.456	0.288	0.216	0.184	0.158	0.142	0.132	0.232
−0.2	1	0.810	0.604	0.388	0.262	0.208	0.182	0.158	0.142	0.132	0.232
−0.1	1	0.760	0.554	0.352	0.248	0.200	0.180	0.154	0.142	0.132	0.232
0.0	1	0.742	0.506	0.332	0.232	0.194	0.174	0.152	0.142	0.132	0.232
0.25	1	0.910	0.472	0.290	0.220	0.186	0.172	0.148	0.140	0.132	0.232
0.50	1	0.886	0.776	0.662	0.216	0.180	0.166	0.148	0.138	0.132	0.232
1.0	1	0.858	0.758	0.618	0.538	0.608	0.166	0.146	0.134	0.132	0.232

表 5.11. 在模型 1 下，使用 $\beta = 0$

λ \ α	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
−0.5	1	0.992	0.980	0.958	0.936	0.914	0.886	0.862	0.846	0.822	0.806
−0.3	1	0.986	0.966	0.946	0.928	0.904	0.886	0.858	0.838	0.822	0.806
−0.2	1	0.980	0.960	0.942	0.924	0.898	0.878	0.858	0.838	0.822	0.806
−0.1	1	0.978	0.954	0.938	0.924	0.896	0.874	0.858	0.838	0.822	0.806
0.0	1	0.984	0.948	0.932	0.918	0.894	0.874	0.854	0.838	0.822	0.806
0.25	1	0.986	0.964	0.932	0.912	0.888	0.866	0.848	0.836	0.822	0.806
0.50	1	0.982	0.974	0.964	0.910	0.884	0.864	0.846	0.832	0.822	0.806
1.0	1	0.986	0.968	0.960	0.942	0.964	0.860	0.842	0.830	0.816	0.806

表 5.12. 在模型 1 下，使用 $\beta = -0.05$

λ \ α	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
−0.5	1	0.992	0.980	0.958	0.936	0.914	0.886	0.862	0.846	0.822	0.806
−0.3	1	0.952	0.924	0.902	0.866	0.840	0.792	0.768	0.736	0.708	0.806
−0.2	1	0.944	0.922	0.892	0.860	0.832	0.792	0.764	0.736	0.708	0.806
−0.1	1	0.940	0.912	0.888	0.856	0.826	0.792	0.764	0.734	0.708	0.806
0.0	1	0.948	0.906	0.880	0.854	0.818	0.790	0.762	0.734	0.708	0.806
0.25	1	0.958	0.938	0.870	0.840	0.802	0.786	0.750	0.732	0.708	0.806
0.50	1	0.950	0.942	0.926	0.840	0.792	0.772	0.744	0.732	0.706	0.806
1.0	1	0.950	0.940	0.924	0.910	0.942	0.762	0.740	0.724	0.702	0.806

表 5.13。 模型 2 下的拒绝比例 模型 2 采用 $\beta = 0$

λ \ α	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
−0.5	1	0.868	0.822	0.780	0.720	0.674	0.634	0.604	0.586	0.568	0.544
−0.3	1	0.830	0.796	0.754	0.694	0.668	0.628	0.600	0.586	0.562	0.544
−0.2	1	0.820	0.786	0.738	0.690	0.664	0.628	0.600	0.586	0.562	0.544
−0.1	1	0.816	0.774	0.724	0.680	0.652	0.622	0.596	0.586	0.562	0.544
0.0	0.962	0.882	0.766	0.718	0.672	0.650	0.618	0.592	0.584	0.562	0.544
0.25	0.980	0.914	0.898	0.710	0.666	0.638	0.608	0.590	0.580	0.562	0.544
0.50	0.980	0.910	0.896	0.886	0.674	0.630	0.604	0.584	0.578	0.562	0.544
1.0	0.980	0.914	0.896	0.880	0.864	0.900	0.594	0.582	0.572	0.560	0.544

表 5.14。 模型 2 下的拒绝比例 模型 2 采用 $\beta = -0.05$

λ \ α	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
−0.5	1	0.868	0.822	0.780	0.720	0.674	0.634	0.604	0.586	0.568	0.544
−0.3	1	0.614	0.552	0.500	0.454	0.420	0.400	0.382	0.372	0.360	0.544
−0.2	1	0.598	0.532	0.488	0.442	0.414	0.398	0.380	0.372	0.358	0.544
−0.1	1	0.586	0.516	0.476	0.432	0.414	0.398	0.376	0.372	0.356	0.544
0.0	0.962	0.594	0.506	0.462	0.426	0.410	0.396	0.376	0.370	0.356	0.544
0.25	0.980	0.796	0.756	0.450	0.416	0.394	0.382	0.374	0.366	0.354	0.544
0.50	0.980	0.780	0.760	0.724	0.422	0.392	0.374	0.368	0.366	0.354	0.544
1.0	0.980	0.778	0.746	0.718	0.694	0.764	0.374	0.366	0.364	0.350	0.544

5.2. 调优参数选择

在模拟研究中，我们已经看到，调节参数的选择对检验的水平和幂起着重要作用。因此，有必要采用数据驱动方法来进行最优调节参数选择。我们从检验函数开始，其定义如下

$$\psi(Y) = \begin{cases} 1 & \text{if } \widehat{T}_{D_\phi^{(k)}} > \tau_c, \\ 0 & \text{otherwise }, \end{cases} \tag{5.2}$$

其中 $Y = (Y_0,\ Y_1)$ 为合并样本。决策规则 $\psi(\cdot)$ 的损失函数定义为

$$L(I_{D_\phi^{(k)}}, \psi(Y)) = \begin{cases} 1 & \text{if true } I_{D_\phi^{(k)}} = 0 \text{ and } \widehat{T}_{D_\phi^{(k)}} > \tau_c, \\ 1 & \text{if true } I_{D_\phi^{(k)}} > 0 \text{ and } \widehat{T}_{D_\phi^{(k)}} \le \tau_c, \\ 0 & \text{otherwise.} \end{cases} \tag{5.3}$$

与损失函数相关的风险函数定义如下：

$$R(I_{D_\phi^{(k)}}, \psi) = \mathbb{E}_Y\Big[L(I_{D_\phi^{(k)}}, \psi(Y))\Big] = \begin{cases} \mathbb{P}\Big\{\widehat{T}_{D_\phi^{(k)}} > \tau_c\Big\} & \text{when true } I_{D_\phi^{(k)}} = 0, \\ 1 - \mathbb{P}\Big\{\widehat{T}_{D_\phi^{(k)}} > \tau_c\Big\} & \text{when true } I_{D_\phi^{(k)}} > 0. \end{cases} \tag{5.4}$$

风险函数 $R(I_{D_\phi^{(k)}}, \psi)$ 在假设检验问题中发挥着类似的作用，正如均方误差 （MSE）在估计中的作用一样。考虑 GSB散度的情形。从决策理论的角度来看，一组最优（可容许的）调节参数应为（α_*， λ_*， β_*），如果

$$(\alpha_*, \lambda_*, \beta_*) := \arg\min_{\alpha,\lambda,\beta} R(I_{D_\phi^{(k)}}, \psi) \text{ for all true } I_{D_\phi^{(k)}}. \tag{5.5}$$

True $I_{D_\phi^{(k)}}$ 对我们来说是未知的。请注意，最优调节参数可能并不唯一。若存在多个最优解，则所有检验都称为风险等价。作为调节参数（α、 λ、 β）的函数，风险函数定义为一个映射，使得 R（$I_{D_\phi^{(k)}}$、 ψ） $:[-1,1]\times \mathbf{R}^2 \mapsto [0,1]$，其最小值点可能并不总是存在。给定一个数据集，我们需要对式（5.4）中定义的风险函数进行‘合适的’估计。利用从合并数据 Y中进行的重抽样来估计该概率。可以根据使用真实调节参数计算的p值来决定是否接受原假设，但这些参数对我们来说是未知的。为此，我们需要先从一些稳健试点调节参数入手。根据模拟研究的启示，我们知道，即使数据集是纯净的或受污染的，稳健初始值（α、 λ、 β）也可能带来相当低的第一类和第二类错误。调节参数选择算法如下所述。

算法5.1。

步骤1 将大小为 Y_0 的第一个样本 n_0 与大小为 Y_1 的第二个样本 n_1 合并为$Y = (Y_0, Y_1)$。定义一个 $0-1$ 值哑变量 X，如式（*3.2*）所示。

步骤2 取一个稳健初始值 $(\alpha_1, \lambda_1, \beta_1)$，并使用定理 *3.2*计算*p*值（$p_1$）。选择初始值时应使*p*值与数据相符。

步骤3 若 $p_1 \le c$，则在 $100c\%$ 的名义水平下拒绝原假设$\mathbb{H}$；否则接受原假设。

步骤4 如果 $p_1 \le c$，则分别从 Y_0 和Y_1中抽取独立的随机重抽样样本 Y_{0b} 和 Y_{1b}；否则，从 $Y :=$ 中同时抽取 Y_{0b} 和 $Y_{1b}(Y_0, Y_1)$，其中 $b = 1$、 2、 ..... 、 B。利用这些重抽样样本，计算检验统计量： $T^{(b)}_{D_\phi^{(k)}}$、 $b = 1$、.....、 B。

步骤 5 计算

$$\hat{P} = \frac{1}{B}\sum_{b=1}^{B} \mathbb{1}\left\{T^{(b)}_{D_\phi^{(k)}} > \tau_c\right\}, \tag{5.6}$$

其中 τ_c 为$100(1-c)\%\mathcal{N}(0,\ 1)$的分位点。

步骤 6 定义

$$\widehat{R}(I_{D_\phi^{(k)}}, \psi) = \begin{cases} \hat{P} & if\ p_1 > c, \\ 1-\hat{P} & if\ p_1 \le c. \end{cases} \tag{5.7}$$

步骤 7 计算最优调节参数 $(\alpha_2, \lambda_2, \beta_2) := \arg\min_{\alpha,\lambda,\beta} \widehat{R}(I_{D_\phi^{(k)}}, \psi)$。

容易看出，当合并样本 Y 固定时

$$\widehat{R}(I_{D_\phi^{(k)}},\psi) \xrightarrow{\mathbb{P}} R(I_{D_\phi^{(k)}},\psi) \text{ as } B\to\infty \text{ at a true pilot } (\alpha_1,\lambda_1,\beta_1). \tag{5.8}$$

因此，$\widehat{R}(I_{D_\phi^{(k)}},\psi)$ 是 $R(I_{D_\phi^{(k)}},\psi)$的良好近似。风险函数$R(I_{D_\phi^{(k)}},\psi)$ 的经验版本会在由合适的稳健试点调节参数确定的细网格上进行最小化。尽管算法5.1 在很大程度上取决于稳健试点调节参数的选择，但只要这些参数能得出与数据集相符的合理最优参数集，就不会造成严重问题。

5.3. 真实数据示例

这里我们讨论两个真实数据示例，并考虑使用算法5.1选择“最优”调节参数的问题。这里的主要目的是说明，算法5.1 得出的调节参数集至少不逊于幂散度族，甚至可能更好。重抽样次数 B 设为200。出于实际考虑，以下示例中的连续随机变量不可能无界。因此，对于调节参数的所有组合，使用渐近零分布都是安全的。

例5.1。（2023年数据科学薪资）该数据集包含不同公司和国家中，2023 年从事“数据科学”职业的员工薪资（单位：美元）。我们感兴趣的是，了解“研究科学家”与“数据工程师”的薪资分布是否存在显著差异。我们在图*5.1*右侧面板中绘制了该数据集。数据集中有1040名数据工程师和82 名研究科学家。我们从以下初始调节参数开始(α,λ,β) = (0.1,0.5,0) 得到的p值为 $0.01739 < 0.05$。因此，在5% 显著性水平下，这两个分布存在差异；这一结论显然与图*5.1*一致。表*5.15*中以绿色标出最低风险，对应的调节参数以蓝色标示。这里，最优调节参数集合不属于*PD*族，而是在其之外。

表5.15。 比较 示例中的数据集采用不同方法估算的风险值 n am题 5.1

方法	$(\alpha,\ \lambda,\ \beta)$	Risk	幂
PD族	$(0,-0.5,0)$	0.640	0.360
	$(0,\ -0.3,\ 0)$	0.690	0.310
	$(0,\ -0.2,\ 0)$	0.665	0.335
	$(0,\ -0.1,\ 0)$	0.580	0.420
	$(0,\ 0,\ 0)$	0.425	0.575
	$(0,0.25,0)$	0.380	0.620
	$(0,0.5,0)$	0.285	0.715
	$(0,\ 1,\ 0)$	0.105	0.895
GSB族	$(-0.1718,1.0348,-0.0204)$	0.040	0.960

例5.2。（美国健康保险数据集）该数据集包含不同类别的保险费用信息，例如年龄、性别和身体质量指数等。一个重要问题是检验购买某些保险单的男性与女性的身体质量指数是否相等。左面板中绘制了男性和女性的身体质量指数分布，

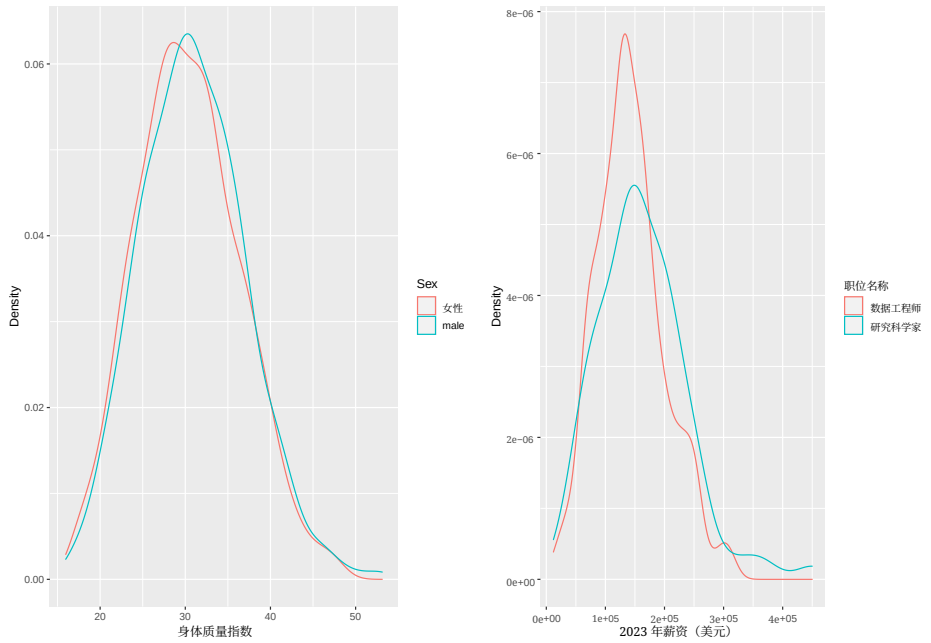

图5.1。数据集图

表5.16。示例中数据集的不同方法风险比较5.2

方法	(α, λ, β)	Risk	真实接受 H 的概率
PD族	$(0, -0.5, 0)$	0.020	0.080
	$(0, -0.3, 0)$	0.010	0.090
	$(0, -0.2, 0)$	0.010	0.090
	$(0, -0.1, 0)$	0.015	0.985
	$(0, 0, 0)$	0.020	0.980
	$(0, 0.25, 0)$	0.020	0.980
	$(0, 0.5, 0)$	0.020	0.980
	$(0, 1, 0)$	0.050	0.950
GSB族	$(0.5968, 1.2306, -0.5153)$	0.005	0.995
	$(0.1248, 0.0802, -0.4720)$	0.005	0.995

图*5.1*。这里，*Pilot*取为$(\alpha, \lambda, \beta) = (0, 0.8, 0)$，得到的$p$值为 $0.86274 > 0.05$。表*5.16*中以绿色标出最低风险，对应的调节参数以蓝色标示。这里，最优调节参数集合不属于*PD*族，而是在其之外。

6. 总结

从本文的讨论可以看出，所有基于互信息的广义定义的检验统计量，都能对两个完全非结构化绝对连续分布之间的相等性给出相合检验。然而，一个数据集中的污染可能会把观测水平（第一类错误）推高到不可接受的程度。面对污染时，GSB 散度族的成员通常表现出好得多的稳健性，当选择

于幂散度族之外时。此外，还可以看到，在调节参数的不同取值下，观测功效变化不大。最后，本文给出了一种针对给定数据集选择最优调节参数的算法。这对应用科学家可能相当有帮助。

致谢

本研究是作者提交给印度加尔各答印度统计研究所的博士论文的一部分。作者谨此感谢导师 Ayanendranath Basu 教授和 Abhik Ghosh 教授给予的指导和宝贵意见。作者也感谢诺华公司提供的良好工作环境。

利益披露声明

作者未报告任何潜在利益冲突。

数据可用性声明

支撑本研究结论的数据集可在 Kaggle 存储库中公开获取，网址为 https://www.kaggle.com/datasets/arnabchaki/data-science-salaries-2023?resource=下载以及 https://www.kaggle.com/code/teertha/us-health-insurance-eda/data。

7. 附录

本节包含这些结果所需的证明。

命题 *2.1*

(P1) 假设这些随机变量相互独立。根据定义 2.12 可得$I_{D_\phi^{(k)}} = 0$。反过来，由 ϕ的严格凸性可得。

(P2) 考虑将随机向量 $(X_1, \ldots, X_m)$ 的任意排列 π 作为映射 $\pi(X_1, \ldots, X_m) = (X_{i_1}, \ldots, X_{i_m})$。由于 $\pi(X_1, \ldots, X_m) = \pi(Y_1, \ldots, Y_m)$ 蕴含 $X_{i_j} = Y_{i_j}$ 对于 $j = 1, 2, \ldots, m$ 成立，因此映射 π 是一一映射。此外，π 是满射。因此 π 是双射，故 π^{-1} 存在。于是我们得到

$$\begin{aligned} f_{\pi(X_1,\ldots,X_m)}(\pi(x_1,\ldots,x_m)) &= f_{X_1,\ldots,X_m}(\pi^{-1}\pi(x_1,\ldots,x_m))|\pi| \\ &= f_{X_1,\ldots,X_m}(x_1,\ldots,x_m)|\pi|, \end{aligned} \tag{7.1}$$

其中 $|\pi|$ 表示该映射的雅可比行列式，其值为 1。因此我们有

$$I_{D_\phi^{(k)}}(X_{i_1},\ldots,X_{i_m}) = I_{D_\phi^{(k)}}(\pi(X_1,\ldots,X_m), \tag{7.2}$$

对于任意排列 π。

(P3) 由于 $(X_1, \ldots, X_m) \mapsto (T_1(X_1), \ldots, T_m(X_m))$ 是一一且满射变换，因此它是一个排列。于是，我们可以使用与之前相同的论证来得到该结果。

(P4) 对所有 $c\phi'(t) \neq k^2 t^{2k-1}\phi''(t)$，取一个有限常数 t。定义

$$R(s,t) = c\phi\Big(\frac{s}{t}\Big)t - \Big\{\phi(s^k) - \phi(t^k) - (s^k - t^k)\phi'(t^k)\Big\} \text{ for } s,t>0. \tag{7.3}$$

注意到 $R(s,s) = 0$。可见

$$\frac{\partial}{\partial s}R(s,t) = c\phi'(s/t) - ks^{k-1}\Big\{\phi'(s^k) - \phi'(t^k)\Big\}, \tag{7.4}$$

$$\frac{\partial^2}{\partial s^2}R(s,t) = c\phi''(s/t)\frac{1}{t} - k(k-1)s^{k-2}\Big\{\phi'(s^k) - \phi'(t^k)\Big\} - (ks^{k-1})^2\phi''(s^k) \neq 0 \text{ at } s \neq t. \tag{7.5}$$

当 t 固定时，$\frac{\partial}{\partial s}R(s,t) = 0$ 在 $s = t$ 处给出一个驻点。可见 $R(s,t)$ 根据以下条件在 $s = t$ 处取得最小值或最大值：

$$\Big[\frac{\partial^2}{\partial s^2}R(s,t)\Big]_{s=t} \gtrless 0 \iff c \gtrless \frac{k^2 t^{2k-1}\phi''(t^k)}{\phi''(1)}. \tag{7.6}$$

由于 $s = t$ 是一个驻点，对于所有 s,t，$R(s,t) \gtrless R(t,t) = 0$。分别将函数 $g(x)$ 和 $f(x)$ 代入 s，t，可得当以下条件成立时，$R(g(x)\ f(x)) \gtrless 0$：

$$cf(x) \gtrless \frac{k^2 f(x)^{2k}\phi''(f(x)^k)}{\phi''(1)} = \tilde{c}_L(x) \text{ (say)} \tag{7.7}$$

对于满足 $f(x)\ g(x) > 0$的每个点 x。由于 ϕ是二次连续可微且严格凸的，因此对所有 $t > 0$，$\phi''(t) > 0$。因此，$\tilde{c}_L(x) > 0$ 在 g, f的共同支撑集上。定义 $c_L := \int_{f>0}\tilde{c}_L(x)dx$。容易看出，当对所有满足$f, g > 0$的 x，$c \leqslant \tilde{c}_L(x)$ 时，$c \leqslant c_L$。此外，将 $R(g(x)\ f(x))$ 在 f, g的共同支撑集上积分，可得

$$c_1 d_\phi(g,f) \leq D_\phi^{(k)}(g,f) \leq c_2 d_\phi(g,f) \text{ for all } 0 < c_1 < c_L < c_2. \tag{7.8}$$

取 $g = f_{X_1,\ X_2,\ \ldots\ldots X_m}$ 和 $f = f_{X_1}\cdots f_{X_m}$。注意，$c_L$ 可能是 ∞。

假设 $\phi_2 = \infty$，Vajda（[32]）表明 $d_\phi(g,\ f) = \infty$ 当 $f,\ g$为奇异时。因此（7.8）意味着 $D_\phi^{(k)}(g,\ f) = \infty$。

（P5）假设 $\phi_2,\ c_L < \infty$和 $\phi \notin C_{B\phi}$。那么 $d_\phi(g,\ f) = \phi_2$（Vajda，[32]）也 $D_\phi^{(k)}(g,\ f) < \infty$。给定有限常数 $c_1,\ c_2$，选取两个收敛序列 $\{c_{1n}\}$ 和$\{c_{2n}\}$ 使得

$$0 < c_1 \leq c_{1n} < c_{1(n+1)} < \cdots < c_L < \cdots < c_{2(n+1)} < c_{2n} \leq c_2. \tag{7.9}$$

这是显而易见的 可知（7.8）对这两个序列成立。定义一个由闭且有界的 d

区间 $C_n := \left[D_\phi^{(k)}/c_{2n},\ D_\phi^{(k)}/c_{1n}\right]$，其中 $d_\phi := d_\phi(g,\ f)$ 和 $D_\phi^{(k)} := D_\phi^{(k)}(g,\ f)$。显然

$$d_\phi, \frac{D_\phi^{(k)}}{c_L} \in C_n \text{ and } C_{n+1} \subset C_n \text{ for all } n. \tag{7.10}$$

因此，$C := \cap_{n\geq 1} C_n$ 将是一个非空的单元素集合。由于 $d_\phi,\ \frac{D_\phi^{(k)}}{c_L} \in C$，它们必定相同。因此我们有 $D_\phi^{(k)}(g,\ f) = c_L d_\phi(g,\ f)$。在这种情况下，Vajda（[32]）表明

$$f \perp g \iff d_\phi(g,f) = \phi_2 \iff D_\phi^{(k)}(g,f) = c_L\phi_2. \tag{7.11}$$

因此，结论得证。

引理 ***3.1***

将 $g_{x,\ y}(\lambda) = (f_x + \lambda h_x)(f_y + \lambda h_y)$ 定义为 $\lambda \in [0,1]$ 的函数。注意到

$$g_{x,y}(0) = f_x f_y, \text{ and } g_{x,y}(1) = \widehat{f}_x \widehat{f}_y. \tag{7.12}$$

考虑以下参数下由 ϕ 生成的扩展布雷格曼散度

$$\Lambda(\lambda) = \int_{f_y>0} \sum_{x=0}^{1} \left\{ \phi\Big((f_{x,y} + \lambda h_{x,y})^k\Big) - \phi\Big(g_{x,y}(\lambda)^k\Big) - \Big((f_{x,y} + \lambda h_{x,y})^k - g_{x,y}(\lambda)^k\Big)\phi'\Big(g_{x,y}(\lambda)^k\Big) \right\} dy. \tag{7.13}$$

注意到

$$\Lambda(1) = D_\phi^{(k)}(\widehat{f}_{XY}, \widehat{f}_x\widehat{f}_y) = \widehat{I}_{D_\phi^{(k)}}(X,Y), \tag{7.14}$$

$$\Lambda(0) = D_\phi^{(k)}(f_{XY}, f_X f_Y) = I_{D_\phi^{(k)}}(X,Y). \tag{7.15}$$

由于假设 ϕ 四次可微，我们可以将 $\Lambda(1)$ 在 $\lambda = 0$ 处展开至二阶项，并采用拉格朗日余项形式，得到

$$\Lambda(1) = \Lambda(0) + \frac{\partial\Lambda(0)}{\partial\lambda} + \frac{1}{2}\cdot\frac{\partial^2\Lambda(0)}{\partial\lambda^2} + \underbrace{\frac{1}{6}\cdot\frac{\partial^3\Lambda(\lambda^*)}{\partial\lambda^3}}_{R_n} \text{ where } \lambda^* \in (0,1). \tag{7.16}$$

这里隐含地假设，关于 λ 的偏导数可以与对 y 的积分互换。很容易验证，

$$\frac{\partial g_{x,y}(0)}{\partial\lambda} = f_x h_y + f_y h_x, \tag{7.17}$$

$$\frac{\partial^2 g_{x,y}(0)}{\partial\lambda^2} = 2h_x h_y. \tag{7.18}$$

当 H 为真时，对于所有 $(x,\ y)$，$f_{x,\ y} = f_x f_y$。由此可知 $I_{D_\phi^{(k)}} = 0$,

$$\begin{aligned}\frac{\partial \Lambda(\lambda)}{\partial \lambda} = \int_{f_y>0} \sum_{x=0}^{1} \Bigg[& k(f_{x,y} + \lambda h_{x,y})^{k-1} h_{x,y} \phi'\Big((f_{x,y} + \lambda h_{x,y})^k\Big) - k g_{x,y}(\lambda)^{k-1} \frac{\partial g_{x,y}(\lambda)}{\partial \lambda} \phi'\Big(g_{x,y}(\lambda)^k\Big) \\ & - \Big((f_{x,y} + \lambda h_{x,y})^k - (g_{x,y}(\lambda))^k\Big) k g_{x,y}(\lambda)^{k-1} \frac{\partial g_{x,y}(\lambda)}{\partial \lambda} \phi''\Big(g_{x,y}(\lambda)^k\Big) \\ & - \Big(k(f_{x,y} + \lambda h_{x,y})^{k-1} h_{x,y} - k(g_{x,y}(\lambda))^{k-1} \frac{\partial g_{x,y}(\lambda)}{\partial \lambda}\Big) \phi'\Big(g_{x,y}(\lambda)^k\Big) \Bigg] dy = 0 \text{ at } \lambda = 0,\end{aligned} \tag{7.19}$$

$$\begin{aligned}\frac{\partial^2 \Lambda(0)}{\partial \lambda^2} = \int_{f_y>0} \sum_{x=0}^{1} \Bigg[& k(k-1) f_{x,y}^{k-2} h_{x,y}^2 \phi'(f_{x,y}^k) + k^2 f_{x,y}^{2k-2} h_{x,y}^2 \phi''(f_{x,y}^k) \\ & - k(k-1)(f_x f_y)^{k-2} (f_x h_y + f_y h_x)^2 \phi'\Big(f_x^k f_y^k\Big) - k(f_x f_y)^{k-1} (2 h_x h_y) \phi'\Big(f_x^k f_y^k\Big) \\ & - k^2 (f_x f_y)^{2k-2} (f_x h_y + f_y h_x)^2 \phi''\Big(f_x^k f_y^k\Big) \\ & - \Big\{ k f_{x,y}^{k-1} h_{x,y} - k(f_x f_y)^{k-1} (f_x h_y + f_y h_x) \Big\} k (f_x f_y)^{k-1} (f_x h_y + f_y h_x) \phi''\Big(f_x^k f_y^k\Big) \\ & - \Big\{ k f_{x,y}^{k-1} h_{x,y} - k(f_x f_y)^{k-1} (f_x h_y + f_y h_x) \Big\} k (f_x f_y)^{k-1} (f_x h_y + f_y h_x) \phi''\Big(f_x^k f_y^k\Big) \\ & - \Big\{ k(k-1) f_{x,y}^{k-2} h_{x,y}^2 - k(f_x f_y)^{k-1} (2 h_x h_y) - k(k-1)(f_x f_y)^{k-2} (f_x h_y + f_y h_x)^2 \Big\} \phi'\Big(f_x^k f_y^k\Big) \Bigg] dy.\end{aligned} \tag{7.20}$$

经代数化简，我们得到

$$\frac{\partial^2 \Lambda(0)}{\partial \lambda^2} = \int_{f_y>0} \sum_{x=0}^{1} k^2 (f_x f_y)^{2k} \phi''(f_x^k f_y^k) \left(\frac{h_x}{f_x} + \frac{h_y}{f_y} - \frac{h_{x,y}}{f_x f_y} \right)^2 dy. \tag{7.21}$$

余项的表达式 R_n 可由式（7.92）得到，其中将该表达式中的$\Delta_{x,\ y}$、Δ_x、Δ_y、θ、t 分别替换为$h_{x,\ y}$、h_x、h_y、λ和1。由于 f_x 是概率质量函数，$|h_x|$ 始终有界于 x。当 $\sup_y f_y < \infty$时，我们得到

$$\begin{aligned}|g'_{x,y}(\lambda)| &< \sup_y |h_y| (1 + 2 \sup_x |h_x|) + \sup_y (f_y) \times \sup_x |h_x| = \mathcal{O}_{\mathbb{P}}(\sup_y |h_y|), \\ |g''_{x,y}(\lambda)| &= \mathcal{O}_{\mathbb{P}}(\sup_y |h_y|^2)\end{aligned}$$

对所有 $0 < \lambda < 1$。稍后我们将看到，$\sup_x |h_x|$、$\sup_y |h_y|$ 和 $\sup_{x,\ y} |h_{x,\ y}|$ 具有相同的随机阶。由于 f_x、f_y、f_y 都假设具有有界的一阶导数，因此它们是利普希茨函数，进而一致连续。根据（7.92）和假设 (A1)，

我们看到， $|R_n|$ 可以由 $O_{\mathbb{P}}(\sup_y |h_y|^3)$ 与一个被积函数有限的积分之积来界定。因此得到 $R_n = O_{\mathbb{P}}(\sup_y |h_y|^3)$。

根据 Fernandes 和 Ne´ri（2006） [26]，我们知道 $\sup_y |h_y|^3 = o_{\mathbb{P}}(\frac{1}{nh_n^{1/2}})$ 和 $\sup_{x,\ y} |h_{x,\ y}|^3 = o_{\mathbb{P}}(\frac{1}{nh_n^{1/2}})$。我们已经知道 $h_x = n^{-1}\sum_{i=1}^n Z_{ix}$。利用马尔可夫不等式，可得

$$\mathbb{P}\Big\{(nh_n^{1/2})^{1/3}|h_x| > \epsilon\Big\} \leq \frac{1}{\epsilon^4}\mathbb{E}\Big\{(nh_n^{1/2})^{1/3}|h_x|\Big\}^4 = \frac{1}{\epsilon^4}\frac{h_n^{2/3}}{n^{5/3}}\mathbb{E}(Z_{1x})^4 \longrightarrow 0 \text{ as } n \to \infty \tag{7.22}$$

对固定的 $\epsilon > 0$。由此可得，对于所有 x， $|h_x|^3 = o_{\mathbb{P}}(\frac{1}{nh_n^{1/2}})$ 成立。由于 x 只取 $0-1$ 个值，$\sup_x |h_x| = \max\{|h_0|, |h_1|\} = o_{\mathbb{P}}(\frac{1}{nh_n^{1/2}})$。这意味着 $R_n = o_{\mathbb{P}}(\frac{1}{nh_n^{1/2}})$。因此，证明完成。

引理 *3.2*

回顾一下，点 $(x,\ y)$ 处的真实联合密度为 $f_{X,\ Y}(x,\ y) := f_{x,\ y}$。与引理 3.1的证明类似，我们将 $\Lambda(1)$ 在 $\lambda = 0$ 附近展开至一阶项，得到

$$\Lambda(1) = \Lambda(0) + \Lambda'(0) + \underbrace{\frac{1}{2}\Lambda''(\lambda^*)}_{R_n} \text{ for } \lambda^* \in (0,1) \tag{7.23}$$

其中 R_n 为余项。我们知道， $\Lambda(1) = \widehat{I}_{D_\phi^{(k)}}$， $\Lambda(0) = I_{D_\phi^{(k)}}$ 以及

$$\begin{aligned}\frac{\partial}{\partial\lambda}\Lambda(0) = \int_{f_y>0}\sum_{x=0}^{1}\Bigg[&k(f_{x,y})^k\Big\{\phi'\Big(f_{x,y}^k\Big) - \phi'\Big((f_x f_y)^k\Big\}\frac{h_{x,y}}{f_{x,y}} \\ &- \Big(f_{x,y}^k - (f_x f_y)^k\Big)k(f_x f_y)^k\left(\frac{h_x}{f_x} + \frac{h_y}{f_y}\right)\phi''\Big((f_x f_y)^k\Big)\Bigg]dy.\end{aligned} \tag{7.24}$$

写 $\widehat{I}_{D_\phi^{(k)}} - I_{D_\phi^{(k)}} = \frac{1}{n}\sum_{i=1}^n V_i + R_n = V_n + R_n$ 哪里

$$\begin{aligned}V_i = \frac{1}{2}\int_{f_y>0}\sum_{x=0}^{1}\Bigg[&k(f_{x,y})^k\Big\{\phi'(f_{x,y}^k) - \phi'\Big((f_x f_y)^k\Big)\Big\}\left(\frac{K_{h_n i}(y)\mathbb{1}_{ix}}{h_n f_{x,y}} - 1\right) \\ &- \Big(f_{x,y}^k - (f_x f_y)^k\Big)k(f_x f_y)^k\phi''\Big((f_x f_y)^k\Big)\left(\frac{\mathbb{1}_{ix}}{f_x} + \frac{K_{h_n i}(y)}{h_n f_y} - 2\right)\Bigg]dy,.\end{aligned} \tag{7.25}$$

看到那个

$$\begin{aligned}\mathbb{E}_{f_{X,Y}}\left(\frac{K_{h_n i}(y)\mathbb{1}_{ix}}{h_n f_{x,y}}\right) &= \frac{1}{h_n f_{x,y}}\sum_{t=0}^{1}\int_w K\Big(\frac{w-y}{h_n}\Big)\mathbb{1}(t=x)f_{X,Y}(t,w)dw \\ &= \frac{1}{h_n f_{x,y}}\int_u K(u)f_{X,Y}(x,y+uh_n)h_n du \\ &= \frac{1}{f_{x,y}}\int K(u)\big[f_{x,y}+\mathcal{O}(h_n)\big]du \ \Big(\text{Assumption (A2)}\Big) \\ &= 1+\mathcal{O}(h_n)\ \Big(\text{Assumption (A3)}\Big)\ . \end{aligned} \tag{7.26}$$

同样地，我们有

$$\mathbb{E}_{f_{X,Y}}\left(\frac{K_{h_n i}(y)}{h_n f_y}\right) = 1+\mathcal{O}(h_n) \tag{7.27}$$

由此得到 $\mathrm{E}_{f_{X,\,Y}}(V_i) = \mathrm{O}(h_n)$。注意到，$V_i$s 是相互独立的，因为成对的 $(Y_i,\ X_i)$s 相互独立。因此，大数弱律意味着 $V_n \xrightarrow{\mathbb{P}} 0$。余项 R_n包含这样的项:¯它们是 $h_x,\ h_y,\ h_{x,\,y}$ 的二阶交叉乘积，作为因子相乘，并与积分中的ϕ''，ϕ'''相乘。这意味着 $R_n = o_{\mathbb{P}}(\frac{1}{(nh_n^{1/2})^{2/3}})$，最终

$$\widehat{I}_{D_\phi^{(k)}} = I_{D_\phi^{(k)}} + o_{\mathbb{P}}(1) + o_{\mathbb{P}}\left(\frac{1}{(nh_n^{1/2})^{2/3}}\right) \longrightarrow I_{D_\phi^{(k)}} \text{ as } n\to\infty. \tag{7.28}$$

证明完毕。

引理 *3.3*

在整个证明过程中，都隐含地假设原假设 H 成立。我们看到，对于所有 n，$\mathrm{E}(S_n) = 0$；并且对于所有 $m < n$，$\mathrm{E}(S_n|\mathrm{F}m) = S_m$ a.s.-Fm。因此，$\{S_n,\ \mathrm{F}_n : n \geq 1\}$ 是一个均值为零且平方可积的鞅，而 $\{T_i : i \geq 1\}$ 是鞅差。条件方差$v_n^2 = \sum_{i=1}^n \mathrm{E}(T_i^2|\mathrm{F}_{i-1})$ 与 S_n 的行为密切相关。在第一个结果中，我们证明条件方差 v_n^2 可以通过平方变差 $u_n^2 := \sum_{i=1}^n T_i^2$ 得到良好近似。第二个结果表明，S_n 的尾部渐近地可忽略。

(a) 首先我们将证明 $\mathbb{E}\big(\sum_{i=1}^n T_i^2\big) \to 4\sigma_\phi^2$。可见

$$\sum_{i=1}^n \mathbb{E}(T_i^2) = \frac{4}{n^2h_n^3}\left[\sum_{i=1}^n\sum_{j<i}\mathbb{E}(V_{ij}^2) + 2\sum_{i=1}^n\sum_{j_1<j_2<i}\mathbb{E}(V_{ij_1}V_{ij_2})\right] = \frac{4}{n^2h_n^3}\sum_{i=1}^n\sum_{j<i}\mathbb{E}(V_{ij}^2), \tag{7.29}$$

由于在 H 下独立性成立，第二项为 0。因此

$$\sum_{i=1}^n \mathbb{E}(T_i^2) = \frac{4}{n^2h_n^3}\sum_{i=1}^n\sum_{j<i}\mathbb{E}(V_{ij}^2) = \frac{4}{n^2h_n^3}\binom{n}{2}\mathbb{E}(V_{ij}^2) = \frac{\mathbb{E}(C_n^2)}{n^2h_n^3} = 4\sigma_\phi^2 + o(1), \tag{7.30}$$

其中 $C_n = nh_n^{3/2}S_n$ 如定理 3.1所示。接下来我们证明 $Var(\sum_{i=1}^n T_i^2) \longrightarrow 0$。注意到

$$Var(\sum_{i=1}^{n} T_i^2) = \sum_{i=1}^{n} Var(T_i^2) + 2\sum_{j<k} Cov(T_j^2, T_k^2), \tag{7.31}$$

其中

$$\begin{aligned} Cov(T_j^2, T_k^2) &= \frac{2^4}{n^4 h_n^6} Cov\Big(\sum_{l<j} V_{lj}^2, \sum_{m<k} V_{mk}^2\Big) \\ &= \frac{2^4}{n^4 h_n^6}\Bigg[4\sum_{l_1<l_2<j}\sum_{m_1<m_2<k} Cov(V_{l_1 j}V_{l_2 j}, V_{m_1 k}V_{m_2 k}) + \sum_{l<j}\sum_{m<k} Cov(V_{lj}^2, V_{mk}^2)\Bigg]. \end{aligned} \tag{7.32}$$

由（7.73）可知

$$\mathbb{E}V_{ij}^2 = \mathbb{E}\Bigg[\sum_{x=0}^{1}\int_{f_y>0} k^2 (f_x f_y)^{2k-2}\phi''(f_x^k f_y^k) Z_{ix}Z_{jx}K_{h_n i}(y)K_{h_n j}(y)dy\Bigg]^2 = \mathcal{O}(h_n^3), \tag{7.33}$$

and

$$\begin{aligned} &\mathbb{E}(V_{lj}V_{mk})^2 \\ &= \mathbb{E}\Bigg[\sum_x \int C_{x,y} Z_{lx}Z_{jx}K_{h_n l}(y)K_{h_n j}(y)dy \times \sum_x \int C_{x,y} Z_{mx}Z_{kx}K_{h_n m}(y)K_{h_n k}(y)dy\Bigg]^2 \\ &= \mathbb{E}\Bigg[\sum_{x_1,x_2}\int C_{x_1,y_1}C_{x_2,y_2}(Z_{lx_1}Z_{jx_1}Z_{kx_2}Z_{mx_2})K_{h_n l}(y_1)K_{h_n j}(y_1)K_{h_n m}(y_2)K_{h_n k}(y_2)dy_1dy_2\Bigg]^2 \\ &= \mathbb{E}\Bigg[\sum_{x_1,x_2,x_3,x_4}\int C_{x_1,y_1}C_{x_2,y_2}C_{x_3,y_3}C_{x_4,y_4}(Z_{lx_1}Z_{lx_3})(Z_{jx_1}Z_{jx_3})(Z_{kx_2}Z_{kx_4})(Z_{mx_2}Z_{mx_4}) \\ &\times(K_{h_n l}(y_1)K_{h_n l}(y_3))(K_{h_n j}(y_1)K_{h_n j}(y_3))(K_{h_n m}(y_2)K_{h_n m}(y_4))(K_{h_n k}(y_2)K_{h_n k}(y_4))dy_1dy_2dy_3dy_4\Bigg] \\ &= \Bigg[\sum_{x_1,x_2,x_3,x_4}\int C_{x_1,y_1}C_{x_2,y_2}C_{x_3,y_3}C_{x_4,y_4}\mathbb{E}^2(Z_{lx_1}Z_{lx_3})\mathbb{E}^2(Z_{kx_2}Z_{kx_4}) \\ &\times\mathbb{E}^2\big(K_{h_n l}(y_1)K_{h_n l}(y_3)\big)\mathbb{E}^2\big(K_{h_n m}(y_2)K_{h_n m}(y_4)\big)dy_1dy_2dy_3dy_4\Bigg], \end{aligned} \tag{7.34}$$

其中$C_{x,y} = k^2(f_x f_y)^{2k-2}\phi''(f_x^k f_y^k)$。可见

$$
\begin{aligned}
dy_3\mathbb{E}^2\Big(K_{h_n l}(y_1)K_{h_n l}(y_3)\Big) &= dy_3\left[\int K\Big(\frac{w-y_1}{h_n}\Big)K\Big(\frac{w-y_3}{h_n}\Big)f_Y(w)dw\right]^2 \\
&= -h_n dz\left[\int K(u)K(u+z)f_Y(y_1+uh_n)h_n du\right]^2 \\
&= -h_n^3 dz\left[\int K(u)K(u+z)f_Y(y_1+uh_n)du\right]^2 \\
&= -f_{y_1}^2 h_n^3 dz\left[\int K(u)K(u+z)du\right]^2 + o(h_n^3)\ \Big\{\text{Assumptions (A2) and (A3)}\Big\}
\end{aligned}
\tag{7.35}
$$

通过变换 $u = \frac{w-y_1}{h_n}$，~~$u+z = \frac{w-y_3}{h_n}$~~，~~$y_3 =$~~ $y_1 - zh_n$。同样，

$$
dy_2\left[\int K\Big(\frac{w-y_2}{h_n}\Big)K\Big(\frac{w-y_4}{h_n}\Big)f_Y(w)dw\right]^2 = -f_{y_4}^2 h_n^3 dz\left[\int K(u)K(u+z)du\right]^2 + o(h_n^3). \tag{7.36}
$$

另外，注意到

$$
\begin{aligned}
\mathbb{E}(V_{lj}V_{mk})^2 &= \sum_{x_1,x_2,x_3x_4}\int C_{x_1,y_1}C_{x_2y_4+zh_n}C_{x_3y_1-zh_n}C_{x_4,y_4}\mathbb{E}^2(Z_{lx_1}Z_{lx_3})\mathbb{E}^2(Z_{kx_2}Z_{kx_4}) \\
&\quad\times f_{y_1}^2 f_{y_4}^2 h_n^6\left(\int\Big[\int K(u)K(u+z)du\Big]^2 dz\right)^2 dy_1 dy_4 \\
&= h_n^6\sum_{x_1,x_2,x_3x_4}\int C_{x_1,y_1}C_{x_2y_4}C_{x_3y_1}C_{x_4,y_4}f_{y_1}^2 f_{y_4}^2 dy_1 dy_4 \\
&\quad\times\mathbb{E}^2(Z_{lx_1}Z_{lx_3})\mathbb{E}^2(Z_{kx_2}Z_{kx_4})\left(\int\Big[\int K(u)K(u+z)du\Big]^2 dz\right)^2 + \mathcal{O}(h_n^8),
\end{aligned}
\tag{7.37}
$$

通过分别在 y_4 和 y_1 附近展开 $C_{x_2,\ y_4+zh_n}$，$C_{x_3,\ y_1-zh_n}$。因此

$$
\begin{aligned}
\frac{1}{n^4h_n^6}\sum_{j<k}\sum_{l<j}\sum_{m<k}Cov(V_{lj}^2,V_{mk}^2) &= \frac{1}{n^4h_n^6}\binom{n}{3}Cov(V_{lj}^2,V_{mk}^2) \\
&= \frac{1}{n^4h_n^6}\binom{n}{3}\big[\mathbb{E}(V_{lj}V_{mk})^2 - \mathbb{E}^2(V_{lj}^2)\big] \\
&= \frac{1}{nh_n^6}\Big(\mathcal{O}(h_n^6)+\mathcal{O}(h_n^6)\Big)\longrightarrow 0.
\end{aligned}
\tag{7.38}
$$

阶数 $Cov(V_{l_1j}V_{l_2j},\ V_{m_1k}V_{m_2k})$ 的值取决于 $\mathrm{E}(V_{l_1j}V_{l_2j}V_{m_1k}V_{m_2k})$ 的顺序，而 ch

取决于以下项

$$
\begin{aligned}
&= dy_1dy_2dy_3dy_4\mathbb{E}\Big[K_{h_nl_1}(y_1)K_{h_nj}(y_1)K_{h_nl_2}(y_2)K_{h_nj}(y_2)K_{h_nm_1}(y_3)K_{h_nk}(y_3)K_{h_nm_2}(y_4)K_{h_nk}(y_4)\Big]\\
&= dy_1\mathbb{E}\Big[K_{h_nj}(y_1)K_{h_nj}(y_2)\Big]\times dy_3\mathbb{E}\Big[K_{h_nk}(y_3)K_{h_nk}(y_4)\Big]\\
&\times dy_2dy_4\mathbb{E}\Big[K_{h_nl_1}(y_1)\Big]\mathbb{E}\Big[K_{h_nl_2}(y_2)\Big]\mathbb{E}\Big[K_{h_nm_1}(y_3)\Big]\mathbb{E}\Big[K_{h_nm_2}(y_4)\Big]\Big]\\
&= h_n^4f_{y_1}f_{y_3}\Big[\iint K(u)K(u+z)dudz\Big]^2\times f_{y_1}f_{y_2}f_{y_3}f_{y_4}h_n^4\\
&= h_n^8(f_{y_2+zh_n}f_{y_4+zh_n})^2f_{y_2}f_{y_4}\Big[\iint K(u)K(u+z)dudz\Big]^2dy_2dy_4\\
&= h_n^8(f_{y_2}f_{y_4})^3\Big[\iint K(u)K(u+z)dudz\Big]^2dy_2dy_4+\mathcal{O}(h_n^{10}).
\end{aligned}
\tag{7.39}
$$

因此

$$
\begin{aligned}
\frac{2^4}{n^4h_n^6}\sum_{l_1<l_2<j}\sum_{m_1<m_2<k}Cov(V_{l_1j}V_{l_2j},V_{m_1k}V_{m_2k}) &= \frac{2^4}{n^4h^6}Cov(V_{l_1j}V_{l_2j},V_{m_1k}V_{m_2k})\sum_{l_1<l_2<j}\sum_{m_1<m_2<k}1\\
&= \frac{2^4}{n^4h^6}\binom{n}{4}Cov(V_{l_1j}V_{l_2j},V_{m_1k}V_{m_2k})\\
&= \mathcal{O}(h_n^2).
\end{aligned}
\tag{7.40}
$$

因此，当 $n\to\infty$ 时，我们有 $\sum_{j<k}Cov(T_j^2,T_k^2)\to 0$。对于 $j<i$，注意到

$$
\begin{aligned}
&\mathbb{E}(V_{ij}^4) = \mathbb{E}\Big[\sum_x\int C_{x,y}Z_{ix}Z_{jx}K_{h_ni}(y)K_{h_nj}dy\Big]^4\\
&= \mathbb{E}\Big[\sum_{x_1,x_2,x_3,x_4}\int C_{x_1,y_1}C_{x_2,y_2}C_{x_3,y_3}C_{x_4,y_4}(Z_{ix_1}Z_{jx_1})(Z_{ix_2}Z_{jx_2})(Z_{ix_3}Z_{jx_3})(Z_{ix_4}Z_{jx_4})\\
&\times\big(K_{h_ni}(y_1)K_{h_nj}(y_1)\big)\big(K_{h_ni}(y_2)K_{h_nj}(y_2)\big)\big(K_{h_ni}(y_3)K_{h_nj}(y_3)\big)\big(K_{h_ni}(y_4)K_{h_nj}(y_4)\big)\Big]dy_1dy_2dy_3dy_4\\
&= \sum_{x_1,x_2,x_3,x_4}\int C_{x_1,y_1}C_{x_2,y_2}C_{x_3,y_3}C_{x_4,y_4}\mathbb{E}^2\big(Z_{ix_1}Z_{ix_2}Z_{ix_3}Z_{ix_4}\big)\\
&\quad\times\mathbb{E}^2\big[K_{h_ni}(y_1)K_{h_ni}(y_2)K_{h_ni}(y_3)K_{h_ni}(y_4)\big]dy_1dy_2dy_3dy_4,
\end{aligned}
\tag{7.41}
$$

其中

$$
\begin{aligned}
&dy_1dy_2dy_3dy_4\mathbb{E}^2\big[K_{h_ni}(y_1)K_{h_ni}(y_2)K_{h_ni}(y_3)K_{h_ni}(y_4)\big]\\
&= dy_1dy_2dy_3dy_4\Big[\int K\Big(\frac{y_1-w}{h_n}\Big)K\Big(\frac{y_2-w}{h_n}\Big)K\Big(\frac{y_3-w}{h_n}\Big)K\Big(\frac{y_4-w}{h_n}\Big)f_Y(w)dw\Big]^2.
\end{aligned}
\tag{7.42}
$$

对 $u_1 = \frac{y_1-w}{h_n}$， $u_2 = \frac{y_2-w}{h_n}$， $u_3 = \frac{y_3-w}{h_n}$ 和 $u_4 = \frac{y_4-w}{h_n}$ 进行变换后，上述表达式可简化为

$$\begin{aligned}
&h_n^3 dy_1 du_2 du_3 du_4 \left[\int K(u_1)K(u_2)K(u_3)K(u_4) f_Y(y_1 - u_1 h_n) h_n du_1 \right]^2 \\
&= h_n^5 \left[\int K(u_1)K(u_2)K(u_3)K(u_4)(f_{y_1} + \mathcal{O}(h_n)) du_1 \right]^2 du_2 du_3 dy_4 dy_1 \\
&= h_n^5 \left[\int K(u_1)K(u_2)K(u_3)K(u_4) du_1 \right]^2 du_2 du_2 dy_3 \times f_{y_1}^2 dy_1 + \mathcal{O}(h_n^6). \qquad (7.43)
\end{aligned}$$

因此，对所有 $j < i$，我们得到 $\mathrm{E}(V_{ij}^4) = \mathrm{O}(h_n^5)$。类似地，当 $j_1 < j_2 < i$ 时，我们得到

$$\begin{aligned}
&\mathbb{E}\big(V_{ij_1}^3 V_{ij_2}\big) \\
&= \mathbb{E}\left[\sum_x \int C_{x,y} Z_{ix} Z_{j_1 x} K_{h_n i}(y) K_{h_n j_1}(y) \right]^3 \left[\sum_x \int C_{x,y} Z_{ix} Z_{j_2 x} K_{h_n i}(y) K_{h_n j_2}(y) \right] \\
&= \mathbb{E}\Bigg[\sum_{x_1,x_2,x_3} \int C_{x_1,y_1} C_{x_2,y_2} C_{x_3,y_3} (Z_{ix_1} Z_{j_1 x_1})(Z_{ix_2} Z_{j_1 x_2})(Z_{ix_3} Z_{j_1 x_3}) \\
&\times \Big(K_{h_n i}(y_1) K_{h_n j_1}(y_1)\Big)\Big(K_{h_n i}(y_2) K_{h_n j_1}(y_2)\Big)\Big(K_{h_n i}(y_3) K_{h_n j_1}(y_3)\Big) \Bigg] \\
&\times \left[\sum_x \int C_{x,y} Z_{ix} Z_{j_2 x} K_{h_n i}(y) K_{h_n j_2}(y) \right] \\
&= \mathbb{E}\Bigg[\sum_{x_1,x_2,x_3,x_4} \int C_{x_1,y_1} C_{x_2,y_2} C_{x_3,y_3} C_{x_4 y_4} (Z_{ix_1} Z_{j_1 x_1})(Z_{ix_2} Z_{j_1 x_2})(Z_{ix_3} Z_{j_1 x_3})(Z_{ix_4} Z_{j_2 x_4}) \\
&\times \Big(K_{h_n i}(y_1) K_{h_n j_1}(y_1)\Big)\Big(K_{h_n i}(y_2) K_{h_n j_1}(y_2)\Big)\Big(K_{h_n i}(y_3) K_{h_n j_1}(y_3)\Big)\Big(K_{h_n i}(y_4) K_{h_n j_2}(y_4)\Big) \Bigg]. \qquad (7.44)
\end{aligned}$$

这就变成了

$$\begin{aligned}
&\sum_{x_1,x_2,x_3,x_4} \int C_{x_1,y_1} C_{x_2,y_2} C_{x_3,y_3} C_{x_4 y_4} \mathbb{E}(Z_{ix_1} Z_{ix_2} Z_{ix_3} Z_{ix_4}) \mathbb{E}(Z_{j_1 x_1} Z_{j_1 x_2} Z_{j_1 x_3}) \underbrace{\mathbb{E}(Z_{j_2 x_4})}_{=0} \\
&\times \mathbb{E}\Big(K_{h_n i}(y_1) K_{h_n i}(y_2) K_{h_n i}(y_3) K_{h_n i}(y_4)\Big) \mathbb{E}\Big(K_{h_n j_1}(y_1) K_{h_n j_1}(y_2) K_{h_n j_1}(y_3)\Big) \mathbb{E}\Big(K_{h_n j_2}(y_4)\Big) \Bigg] = 0. \qquad (7.45)
\end{aligned}$$

其余交叉乘积项同样为 0。由此得到

$$\sum_{i=1}^{n}\mathbb{E}(T_i^4)=\frac{2^4}{n^4h_n^6}\sum_{i=1}^{n}\sum_{j<i}\mathbb{E}(V_{ij}^4)=\frac{2^4}{n^4h_n^6}\times\binom{n}{2}\mathcal{O}(h_n^5)=\mathcal{O}\Big(\frac{1}{n^2h_n}\Big)=o(1), \tag{7.46}$$

当 $nh_n^{1/2}\to\infty$ 时。因此，我们得到 $\sum_{i=1}^{n}Var(T_i^2)\le\sum_{i=1}^{n}\mathbb{E}(T_i^4)\to 0$，当 $n\to\infty$ 时。最后，应用切比雪夫不等式，我们得到

$$\mathbb{P}\Big\{\Big|\sum_{i=1}^{n}T_i^2-4\sigma_\phi^2\Big|>\epsilon\Big\}\le\frac{Var\Big(\sum_{i=1}^{n}T_i^2\Big)}{\epsilon^2}\longrightarrow 0 \text{ for any } \epsilon>0. \tag{7.47}$$

(b) 对于任意 $\epsilon>0$,

$$\mathbb{P}\Big\{\max_{1\le i\le n}|T_i|>\epsilon\Big\}\le\frac{\mathbb{E}\big(\max_{1\le i\le n}|T_i|^4\big)}{\epsilon^4}\le\frac{\sum_{i=1}^{n}\mathbb{E}(T_i^4)}{\epsilon^4}=\mathcal{O}\Big(\frac{1}{\epsilon^4n^2h_n}\Big)\longrightarrow 0, \tag{7.48}$$

当 $nh_n^{1/2}\to\infty$ 时，对于 $n\to\infty$。

(c) 我们已经知道 $\sum_{i=1}^{n}T_i^2\xrightarrow{\mathbb{P}}4\sigma_\phi^2$，且 $\max_{1\le i\le n}|T_i|\xrightarrow{\mathbb{P}}0$。为证明第三部分，首先我们将证明 $Z_n=\prod_{j=1}^{n}(1+itT_j)$ 一致可积。

任取满足 $0<4\sigma_\phi^2<M$ 的 M，并令 $t_n:=\sum_{i=1}^{n}T_i^2$。那么存在一个子序列 $\{t_{k_n}:k_n\ge n\}$，使得 $t_{k_n}\xrightarrow{a.s.}4\sigma_\phi^2<M$。定义 $A_n=\{\omega:t_n\le M\}$。注意，对于所有 $n\ge 1$，当$t_n\le t_{n+1}$ 时 $A_n^c\subseteq A_{n+1}^c$。请注意,

$$\mathbb{P}(A_n^c)\le\mathbb{P}\Big\{t_n>M \text{ infinitely often }\Big\}\le\mathbb{P}\Big\{t_{k_n}>M \text{ infinitely often }\Big\}=0$$

对于所有 $n\ge 1$。利用柯西-施瓦茨不等式，我们得到

$$R_n:=\mathbb{E}\Big(\prod_{j=1}^{n}(1+t^2T_j^2)\mathbb{1}(A_n^c)\Big)\le\sqrt{\mathbb{E}\Big(\prod_{j=1}^{n}(1+t^2T_j^2)^2\Big)\mathbb{P}(A_n^c)}=0. \tag{7.49}$$

现在，注意到

$$\begin{aligned}
\mathbb{E}|Z_n|^2 &= \mathbb{E}\left(\prod_{j=1}^{n}(1+t^2T_j^2)\mathbb{1}(A_n)\right) + \underbrace{\mathbb{E}\left(\prod_{j=1}^{n}(1+t^2T_j^2)\mathbb{1}(A_n^c)\right)}_{R_n} \\
&\leq \mathbb{E}\Big[\Big\{e^{t^2\sum_{j=1}^{n-1}T_j^2}\Big\}\Big(1+t^2T_n^2\Big)\mathbb{1}(A_n)\Big] \ (\text{ as } e^x \geq 1+x) \\
&\leq \Big\{e^{t^2M}\Big\}(1+t^2\mathbb{E}T_n^2) \\
&< \Big\{e^{t^2M}\Big\}\Big(1+t^2\sum_{n=1}^{n}\mathbb{E}T_n^2\Big) \\
&= \Big\{e^{t^2M}\Big\}\Big(1+t^24\sigma_\phi^2\Big) + o(1) < \infty \text{ uniformly in } n
\end{aligned} \tag{7.50}$$

由（7.30）。因此 Z_n 是一致可积的（UI）。接下来我们将证明，对于任意 $\mathbb{E}$（$Z_n\mathbb{1}$（E））$\to\mathbb{P}$（E）。定义

$$J_n = \begin{cases} \min\left\{m \leq n : \sum_{i=1}^{m} T_i^2 > 2M\right\} & \text{if } \sum_{i=1}^{n} T_i^2 > 2M, \\ n & \text{Otherwise.} \end{cases}$$

注意到 $J_n \leq n$。回顾一下，对于所有 $m \leq n$，$\mathbb{E}(S_n|\mathbb{F}m) = S_m$，因此

$$\mathbb{E}(T_j|\mathcal{F}_{j-1}) = \mathbb{E}(S_j - S_{j-1}|\mathcal{F}_{j-1}) \ = \mathbb{E}(S_j|\mathcal{F}_{j-1}) - S_{j-1} = S_{j-1} - S_{j-1} = 0.$$

注意到

$$\begin{aligned}
\mathbb{E}\Big[Z_n\mathbb{1}(E)\Big] &= \mathbb{E}\left\{\mathbb{1}(E)\prod_{j=1}^{J_n}(1+itT_j)\prod_{j=J_n+1}^{n}\Big(1+it\mathbb{E}\Big(T_j|\mathcal{F}_{j-1}\Big)\Big)\right\} \\
&= \mathbb{E}\left\{\mathbb{1}(E)\prod_{j=1}^{J_n}(1+itT_j)\right\} \\
&= \mathbb{P}(E) + R_n',
\end{aligned} \tag{7.51}$$

whe余项 R_n' 至多由 $(2^{J_n}-1)$ 个下列形式的项组成

$$\mathbb{E}\Big[\mathbb{1}(E)(it)^r T_{i_1}T_{i_2}\cdots T_{i_r}\Big] \tag{7.52}$$

使得 $1 \le r \le J_n$ 且 $1 \le i_1 \le i_2 \le \cdots \le i_r \le J_n$。由于

$$\begin{aligned}
&\left|T_{i_1}^2 T_{i_2}^2 \cdots T_{i_{r-1}}^2\right|^{\frac{1}{r-1}} \le \frac{1}{r-1}\sum_{t=1}^{r-1} T_{i_t}^2 \le \sum_{t=1}^{r-1} T_{i_t}^2 \le \sum_{t=1}^{J_n-1} T_{i_t}^2 \\
\Longrightarrow\ &\left|T_{i_1}^2 T_{i_2}^2 \cdots T_{i_{r-1}}^2\right| \le \left(\sum_{t=1}^{J_n-1} T_{i_t}^2\right)^{r-1} \\
\Longrightarrow\ &\left|T_{i_1}^2 T_{i_2}^2 \cdots T_{i_r}^2\right| \le \left(\sum_{t=1}^{J_n-1} T_{i_t}^2\right)^{r-1}\left(\max_{1\le i\le n} T_i^2\right) \le (2M)^{r-1}\left(\max_{1\le i\le n} T_i^2\right) \\
\Longrightarrow\ &\mathbb{E}\left|T_{i_1} T_{i_2} \cdots T_{i_r}\right| \le (2M)^{\frac{r-1}{2}}\mathbb{E}\left(\max_{1\le i\le n}|T_i|\right),
\end{aligned} \tag{7.53}$$

由此可得

$$|R_n'| \le (2^{J_n}-1)(2M)^{\frac{J_n}{2}}\mathbb{E}\Big(\max_{1\le i\le n}|T_i|\Big) \tag{7.54}$$

因为余项至多包含 $(2^{J_n}-1)$ 项。但是，对任意 $\epsilon > 0$,

$$\begin{aligned}
\mathbb{E}\left(\max_{1\le i\le n}|T_i|\right) &\le \epsilon + \mathbb{E}\left[\max_{1\le i\le n}|T_i|\mathbb{1}\big(|T_i|>\epsilon\big)\right] \\
&\le \epsilon + \left[\mathbb{E}\left(\max_{1\le i\le n} T_i^2\right)\mathbb{P}\left(\max_{1\le i\le n}|T_i|>\epsilon\right)\right]^{\frac{1}{2}} \text{ (Cauchy-schwarz inequality)} \\
&\longrightarrow \epsilon \text{ when } n\to\infty,
\end{aligned} \tag{7.55}$$

因为 (b) 成立，且对所有 i，$\mathrm{E}(T_i^2)$ 有限。由于 $\epsilon > 0$ 是任意的，$\mathrm{E}(\max_{1\le i\le n}|T_i|) \to 0$。因此我们得到 $|R_n'| \le (2^{J_n}-1)(2M)^{\frac{J_n}{2}}\epsilon$，其中 $J_n \le n$，从而在 $n\to\infty$ 时得到 $R_n' \to 0$。故对任意 $E \in \mathsf{F}$，我们得到 $\mathrm{E}(Z_n\mathbb{1}(E)) \longrightarrow \mathrm{P}(E)$，因此 $Z_n \longrightarrow 1$ 在 L_1 中弱收敛。

定理 *3.1*

回顾记号 $K_{h_n i}(y)$、Z_{ix} 和 $\mathbb{1}_{ix}$。我们定义

$$I_n = \frac{1}{2}\sum_{x=0}^{1}\int_{f_y>0} k^2(f_x f_y)^{2k}\phi''(f_x^k f_y^k)\left(\frac{h_x}{f_x}+\frac{h_y}{f_y}-\frac{h_{x,y}}{f_x f_y}\right)^2 dy, \tag{7.56}$$

$$\widetilde{I}_n = \frac{1}{2}\sum_{x=0}^{1}\int_{f_y>0} k^2(f_x f_y)^{2k}\phi''(f_x^k f_y^k)\left(\frac{h_{x,y}}{f_x f_y}-\frac{h_y}{f_y}\right)^2 dy. \tag{7.57}$$

由引理3.1 可知

$$\widehat{I}_{D_\phi^{(k)}} = I_n + o_{\mathbb{P}}\Big(\frac{1}{nh_n^{1/2}}\Big) \text{ under } \mathbb{H}. \tag{7.58}$$

首先我们将证明 $nh_n^{1/2}|I_n - \widetilde{I}_n| \xrightarrow{\mathbb{P}} 0$，这意味着在原假设 H 下，$I_n$ 与 $\widetilde{I}_n$ 具有相同的渐近分布。可见

$$\begin{aligned}
|I_n - \widetilde{I}_n| &\le \frac{1}{2}\sum_{x=0}^{1}\int_{f_y>0} k^2(f_x f_y)^{2k}\left|\phi''(f_x^k f_y^k)\right|\left|\left(\frac{h_{x,y}}{f_x f_y} - \frac{h_x}{f_x} - \frac{h_y}{f_y}\right)^2 - \left(\frac{h_{x,y}}{f_x f_y} - \frac{h_y}{f_y}\right)^2\right| dy \\
&\le \sum_{x=0}^{1}\int_{f_y>0} k^2(f_x f_y)^{2k-1}\left|\phi''(f_x^k f_y^k)\right|\frac{|h_x|}{f_x}\left|2h_{x,y} - h_x f_y - 2h_y f_x\right| dy \\
&\le \frac{\sup_x |h_x|}{\min_x f_x}\sup_{x,y}\left|2h_{x,y} - h_x f_y - 2h_y f_x\right| \times \sum_{x=0}^{1}\int_{f_y>0} k^2(f_x f_y)^{2k-1}\left|\phi''(f_x^k f_y^k)\right| dy \\
&\le \frac{2\sup_x |h_x|}{\min_x f_x} \times \left\{\sup_{x,y}\left|h_{x,y}\right| + \sup_x |h_x|\sup_y |f_y| + \sup_y |h_y|\sup_x |f_x|\right\} \\
&\quad \times \sum_{x=0}^{1}\int_{f_y>0} k^2(f_x f_y)^{2k-1}\left|\phi''(f_x^k f_y^k)\right| dy. \qquad (7.59)
\end{aligned}$$

我们知道，$\sup_y |h_y|^3 = o_{\mathbb{P}}\left(\frac{1}{nh_n^{1/2}}\right)$ 和 $\sup_{x,y}|h_{x,y}|^3 = o_{\mathbb{P}}\left(\frac{1}{nh_n^{1/2}}\right)$ 均一致。进一步，根据假设(A1)，$\sum_{x=0}^{1} k^2(f_x f_y)^{2k-1}\phi''(f_x^k f_y^k)$ 被一个可积函数一致地有界。此外，注意到

$$\mathbb{P}\left\{(nh_n^{1/2})^{2/3}|h_x| \ge \epsilon\right\} \le \frac{1}{\epsilon^4}\mathbb{E}\left((nh_n^{1/2})^{2/3}h_x\right)^4 = \frac{1}{\epsilon^4}\cdot\frac{h_n^{4/3}}{n^{1/3}}\cdot\mathbb{E}[Z_{1x}]^4 \to 0 \text{ as } n\to\infty, \qquad (7.60)$$

对任意固定的 $\epsilon > 0$。因此，我们得到 $(nh_n^{1/2})^{2/3}|h_x| = o_{\mathbb{P}}(1)$ 和 $nh_n^{1/2}|I_n - \widetilde{I}_n| \xrightarrow{\mathbb{P}} 0$。

接下来，我们将求出 I_n 的渐近分布。注意到

$$\begin{aligned}
\widetilde{I}_n &= \frac{1}{2}\sum_{x=0}^{1}\int_{f_y>0} k^2(f_x f_y)^{2k}\phi''(f_x^k f_y^k)\left[\frac{1}{nh_n}\sum_{i=1}^{n}\left\{\frac{K_{h_n i}(y)1_{ix}}{f_x f_y} - \frac{K_{h_n i}(y)}{f_y}\right\}\right]^2 dy \\
&= \frac{1}{2(nh_n)^2}\sum_{x=0}^{1}\int_{f_y>0} k^2(f_x f_y)^{2k-2}\phi''(f_x^k f_y^k)\left[\sum_{i=1}^{n} Z_{ix}^2 K_{h_n i}^2(y) + 2\sum_{i=1}^{n}\sum_{j<i} Z_{ix}Z_{jx}K_{h_n i}(y)K_{h_n j}(y)\right] dy \\
&= \frac{B_n + C_n}{2(nh_n)^2}, \qquad (7.61)
\end{aligned}$$

其中

$$B_n = \sum_{i=1}^{n}\sum_{x=0}^{1}\int_{f_y>0} k^2(f_x f_y)^{2k-2}\phi''(f_x^k f_y^k)Z_{ix}^2 K_{h_n i}^2(y)dy, \qquad (7.62)$$

$$C_n = 2\sum_{i=1}^{n}\sum_{j<i}\sum_{x=0}^{1}\int_{f_y>0} k^2(f_x f_y)^{2k-2}\phi''(f_x^k f_y^k)Z_{ix}Z_{jx}K_{h_n i}(y)K_{h_n j}(y)dy. \qquad (7.63)$$

注意，在原假设下，$\mathbb{E}(Z_{ix}^2) = f_x(1-f_x)$，并且

$$\begin{aligned}\mathbb{E}(K_{h_n i}^2(y)) &= \int_w K^2\Big(\frac{w-y}{h_n}\Big) f_Y(w)dw \\ &= h_n \int_u K^2(u) f_Y(y+uh_n)du \\ &= h_n\Big\{f_y \int_u K^2(u)du + \mathcal{O}(h_n)\Big\} \Big\{ \text{ Assumptions (A2) and (A3) } \Big\} \\ &= h_n f_y \int_u K^2(u)du + \mathcal{O}(h_n^2). \end{aligned} \tag{7.64}$$

因此，我们发现

$$\mathbb{E}(B_n) = nh_n \int_u K^2(u)du \int_{f_y>0} \sum_{x=0}^{1} k^2(f_x f_y)^{2k-1}\phi''(f_x^k f_y^k)(1-f_x)dy + \mathcal{O}(nh_n^2) \tag{7.65}$$

$$= 2nh_n\mu_\phi + o(nh_n). \tag{7.66}$$

类似地，在独立性条件下，

$$\begin{aligned} Var(B_n) &= \sum_{i=1}^{n} Var\Bigg[\sum_{x=0}^{1}\int_{f_y>0} \underbrace{k^2(f_x f_y)^{2k-2}\phi''(f_x^k f_y^k)}_{C_{x,y}} Z_{ix}^2 K_{h_n i}^2(y)dy\Bigg] \\ &\le \sum_{i=1}^{n}\mathbb{E}\Bigg[\sum_{x=0}^{1}\int C_{x,y} Z_{ix}^2 K_{h_n i}^2(y)dy\Bigg]^2 \\ &= n\mathbb{E}\Bigg[\sum_{x=0}^{1}\int C_{x,y} Z_{ix}^2 K_{h_n i}^2(y)dy\Bigg]^2 \\ &= n\mathbb{E}\Bigg[\sum_{x_1,x_2}\iint C_{x_1,y_1}C_{x_2,y_2} Z_{ix_1}^2 K_{h_n i}^2(y_1) Z_{ix_2}^2 K_{h_n i}^2(y_2)dy_1dy_2\Bigg] \\ &= n\sum_{x_1,x_2}\iint C_{x_1,y_1}C_{x_2,y_2}\mathbb{E}\Big(Z_{ix_1}^2 Z_{ix_2}^2\Big)\mathbb{E}\Big(K_{h_n i}^2(y_1)K_{h_n i}^2(y_2)\Big)dy_1dy_2. \end{aligned} \tag{7.67}$$

此外，成立

$$\begin{aligned} dy_1dy_2\mathbb{E}\Big(K_{h_n i}^2(y_1)K_{h_n y_2}^2\Big) &= dy_1dy_2\int\Bigg[K^2\Big(\frac{w-y_1}{h_n}\Big)K^2\Big(\frac{y_2-w}{h_n}\Big)f_Y(w)dw\Bigg] \\ &= h_n dz dy_1\Big[\int K^2(u)K^2(u+z)f_Y(y_1+uh_n)h_n du\Big] \\ &= h_n^2(f_{y_1}dy_1)\times\Big(\int K^2(u)K^2(u+z)du\Big)dz + \mathcal{O}(h_n^3) \end{aligned} \tag{7.68}$$

根据假设(A3)。因此我们有

$$Var(B_n) \le nh_n^2 \sum_{x_1,x_2} \mathbb{E}(Z_{ix_1}^2 Z_{ix_2}^2) \int C_{x_1,y_1} C_{x_2,y_1+h_n(u+z)} f_{y_1} dy_1 \times \iint K^2(u)K^2(u+z)dudz = \mathcal{O}(nh_n^2). \tag{7.69}$$

接下来注意到 $\mathrm{E}(C_n) = 0$ 和

$$\begin{aligned}\mathbb{E}(C_n^2) &= 4\binom{n}{2}\mathbb{E}\left[\sum_{x=0}^{1}\int k^2 (f_x f_y)^{2k-2}\phi''(f_x^k f_y^k) Z_{ix} Z_{jx} K_{h_n i}(y) K_{h_n j}(y) dy\right]^2 \\ &= 4\binom{n}{2}k^4 \iint \left[\sum_{x=0}^{1}(f_x f_{y_1})^{2k-2}(f_x f_{y_2})^{2k-2}\phi''(f_x^k f_{y_1}^k)\phi''(f_x^k f_{y_2}^k)\mathbb{E}^2(Z_{ix}^2)\right. \\ &\left. +2(f_{x_0} f_{y_1})^{2k-2}(f_{x_1} f_{y_2})^{2k-2}\phi''(f_{x_0}^k f_{y_1}^k)\phi''(f_{x_1}^k f_{y_2}^k)\mathbb{E}^2(Z_{ix_0} Z_{ix_1})\right]\mathbb{E}^2\Big(K_{h_n i}(y_1) K_{h_n i}(y_2)\Big) dy_1 dy_2.\end{aligned} \tag{7.70}$$

注意到，当 $x_1 \neq x_2$ 时， $Z_{ix_1} = -Z_{ix_2}$ ，并且

$$\mathbb{E}\Big(Z_{ix_1} Z_{ix_2}\Big) = \begin{cases} -f_{x_1}(1-f_{x_1}) & \text{if } x_1 \neq x_2, \\ f_{x_1}(1-f_{x_1}) & \text{if } x_1 = x_2. \end{cases} \tag{7.71}$$

因此 $\mathrm{E}^2\Big[Z_{ix_1} Z_{ix_2}\Big] = f_{x_1}^2(1-f_{x_1})^2$。注意到

$$dy_2 \mathbb{E}^2\Big(K_{h_n i}(y_1) K_{h_n i}(y_2)\Big) = dy_2 \left[\int_w K\Big(\frac{y_1 - w}{h_n}\Big) K\Big(\frac{y_2 - w}{h_n}\Big) f_Y(w) dw\right]^2. \tag{7.72}$$

代入 $\frac{y_1-w}{h_n} = \cancel{z,} \quad \frac{y_2-w}{h_n} = \cancel{z+} u$ 和 $y_2 = y_1 + uh_n$ 于（7.72）中，我们得到

$$\begin{aligned} & h_n du \left[h_n \int_w K(z)K(z+u) f_Y(y_1 - zh_n) dz\right]^2 \\ &= h_n^3 du \left[\int_w K(z)K(z+u)(f_{y_1} + \mathcal{O}(h_n)) dz\right]^2 \left\{ \text{Assumption (A3)} \right\} \\ &= f_{y_1}^2 h_n^3 du \left[\int_w K(z)K(z+u) dz\right]^2 + o(h_n^3). \end{aligned} \tag{7.73}$$

由此可得

$$\mathbb{E}(C_n^2) = 2n^2k^4h_n^3 \int \Bigg[\sum_{x=0}^{1} (f_x f_{y_1})^{2k-2} (f_x f_{y_1+uh_n})^{2k-2} \phi''(f_x^k f_{y_1}^k) \phi''(f_x^k f_{y_1+uh_n}^k) f_x^2 (1-f_x)^2 f_{y_1}^2$$
$$+2(f_{x_0} f_{y_1})^{2k-2} (f_{x_1} f_{y_1+uh_n})^{2k-2} \phi''(f_{x_0}^k f_{y_1}^k) \phi''(f_{x_1}^k f_{y_1+uh_n}^k) f_{x_0}^2 f_{x_1}^2 f_{y_1}^2 \Bigg] dy_1$$
$$\times \int \Big(\int K(z)K(z+u)dz \Big)^2 du + o(n^2h_n^3). \tag{7.74}$$

由于$f_{y_1+uh_h} = f_{y_1} + \mathcal{O}(h_n)$ 根据假设(A2)，我们可以近似表示 $f_{y_1+uh_n} \approx f_{y_1}$ 对于充分大的 n，余项为 $o(n^2h_n^3)$。利用这一近似，我们得到

$$\mathbb{E}(C_n^2) = 2n^2h_n^3 \int \Bigg[\sum_{x=0}^{1} k^2 (f_x f_y)^{2k-1} \phi''\Big(f_x^k f_y^k\Big)(1-f_x) \Bigg]^2 dy \int \Big(\int K(z)K(z+u)dz \Big)^2 du + o(n^2h_n^3)$$
$$= 4n^2h_n^3\sigma_\phi^2 + o(n^2h_n^3). \tag{7.75}$$

此外，还可以看到

$$\begin{aligned} \frac{1}{(nh_n)^2} |Cov(B_n, C_n)| &\le \frac{1}{(nh_n)^2} \sqrt{Var(B_n)Var(C_n)} \\ &= \frac{1}{(nh_n)^2} \sqrt{\mathcal{O}(nh_n^2) \times \mathcal{O}(n^2h_n^3)} \\ &= \mathcal{O}(n^{-1/2}h_n^{1/2}) \to 0 \text{ as } n \to \infty. \end{aligned} \tag{7.76}$$

因此可得

$$\mathbb{E}(\widetilde{I}_n) \approx \frac{\mu_\phi}{nh_n} \text{ and } Var(\widetilde{I}_n) \approx \frac{\sigma_\phi^2}{n^2h_n}. \tag{7.77}$$

~我们看到， I_n 的渐近分布由 C_n决定。我们已经知道$S_n = \sum_{i=1}^n T_i$，其中

$$T_i = \frac{2}{nh_n^{3/2}} \sum_{j<i} \sum_{x=0}^{1} \int_{f_y>0} k^2 (f_x f_y)^{2k-2} \phi''(f_x^k f_y^k) Z_{ix} Z_{jx} K_{h_n i}(y) K_{h_n j}(y) dy. \tag{7.78}$$

引理3.3 验证了 Hall 和 Heyde（2014）的条件 [31] (定理 3.2) 给出

$$\frac{C_n}{nh_n^{3/2}} = S_n = \sum_{i=1}^{n} T_i \xrightarrow{\mathcal{L}} \mathcal{N}\Big(0, 4\sigma_\phi^2\Big) \text{ as } n \to \infty. \tag{7.79}$$

将 $C_n = nh_n^{3/2} S_n$ 代入 $\widetilde{I}_n$ 的表达式中，得到

$$nh_n^{1/2}\Big(\widetilde{I}_n - \frac{\mu_\phi}{nh_n}\Big) \overset{\mathcal{L}}{=} nh_n^{1/2}\Big(\widetilde{I}_n - \frac{B_n}{2(nh_n)^2}\Big) = \frac{S_n}{2} \xrightarrow{\mathcal{L}} \mathcal{N}\Big(0, \sigma_\phi^2\Big) \text{ as } n \to \infty. \tag{7.80}$$

证明至此完成。

7.1. 定理*3.2*

(i) 由于核密度估计具有一致性，因此 $\widehat{\mu}_\phi, \widehat{\sigma}_\phi$ 也具有一致性。定义

$$T(\lambda) = \frac{nh_n^{1/2}\Big(\widehat{I}_{D_\phi^{(k)}} - \frac{\mu_\phi + \lambda(\hat{\mu}_\phi - \mu_\phi)}{nh_n}\Big)}{\sigma_\phi + \lambda(\widehat{\sigma}_\phi - \sigma_\phi)} \text{ for } \lambda \in [0,1]. \tag{7.81}$$

注意到 $T(1) = \widehat{T}_{D_\phi^{(k)}}$ 和 $T(0) = T_{D_\phi^{(k)}}$。将 $T(1)$ 在 $\lambda = 0$ 处展开，可得

$$T(1) = T(0) + \frac{1}{2}T'(\lambda^{**}) \text{ where } \lambda^{**} \in (0,1), \tag{7.82}$$

其中

$$T'(\lambda^{**}) = \frac{\mu_\phi - \hat{\mu}_\phi}{h_n^{1/2}(\sigma_\phi + \lambda^{**}(\widehat{\sigma}_\phi - \sigma_\phi))} - \underbrace{T(\lambda^{**})}_{\mathcal{O}_{\mathbb{P}}(1)} \times \underbrace{\frac{(\widehat{\sigma}_\phi - \sigma_\phi)}{\sigma_\phi + \lambda^{**}(\widehat{\sigma}_\phi - \sigma_\phi)}}_{o_{\mathbb{P}}(1)}. \tag{7.83}$$

定义

$$\mu_\phi(\lambda) = \frac{1}{2}\int_u K^2(u)du \int_{f_y>0} \sum_{x=0}^{1} k^2 \xi_{xy}^{2k-1} \phi''\Big(\xi_{xy}^k\Big)(1 - f_x - \lambda h_x)dy, \tag{7.84}$$

其中 $\xi_{xy} = (f_x + \lambda h_x)(f_y + \lambda h_y)$。注意，$\mu_\phi(0) = \mu_\phi$ 且 $\mu_\phi(1) = \hat{}\mu_\phi$。将 $\mu_\phi(1)$ 在 $\lambda = 0$ 附近展开至一阶项，得到 $\mu_\phi(1) = \mu_\phi(0) + {}^{\mu'_\phi}(\lambda^*)_2$，其中 $\lambda^* \in$ ~~$(0, 1)$~~。可见

$$\begin{aligned}
\left|\frac{\mu'_\phi(\lambda^*)}{0.5\int_u K^2(u)du}\right| &= \left| \int \sum_{x=0}^{1} k^2 \xi_{xy}^{2k-1} \Bigg\{ -\phi''(\xi_{xy}^k)h_x \right. \\
&\quad + k\xi^k \phi'''(\xi_{xy}^k)\Big(\frac{h_y}{f_y + \lambda^* h_y} + \frac{h_x}{f_x + \lambda^* h_x}\Big)(1 - f_x - \lambda^* h_x) \\
&\quad \left. + (2k-1)\Big(\frac{h_y}{f_y + \lambda^* h_y} + \frac{h_x}{f_x + \lambda^* h_x}\Big)\phi''(\xi_{xy}^k)(1 - f_x - \lambda^* h_x) \Bigg\} dy \right| \\
&\le \sup_x |h_x| \int \sum_{x=0}^{1} \left| 3k^2 \xi_{xy}^{2k-1} \left\{ -\frac{\phi''(\xi_{xy}^k)}{3} + \frac{k\xi^k \phi'''(\xi_{xy}^k) + (2k-1)\phi''(\xi_{xy}^k)}{f_x + \lambda^* h_x} \right\} \right| dy \\
&\quad + \sup_y |h_y| \int \sum_{x=0}^{1} \left| 3k^2 \xi_{xy}^{2k-1} \left\{ \frac{k\xi^k \phi'''(\xi_{xy}^k) + (2k-1)\phi''(\xi_{xy}^k)}{f_y + \lambda^* h_y} \right\} \right| dy \\
&\le \mathcal{O}_{\mathbb{P}}\Big(\sup_x |h_x|\Big) + \mathcal{O}_{\mathbb{P}}\Big(\sup_y |h_y|\Big).
\end{aligned} \tag{7.85}$$

由于 f_x 介于 $[0,1]$, 项 $|1-f_x-\lambda^* h_x|$ 被 3 所界定，结合假设(A1) 和(A2) 可知，上述积分有界。我们还知道 $\sup_x |h_x|$ 和 $\sup_y |h_y|$ 为 $o_{\mathbb{P}}(1/(nh_n^{1/2})^{1/3})$，因此 $|\mu'_\phi(\lambda)| = o_{\mathbb{P}}(1/(nh_n^{1/2})^{1/3})$ 关于 x， y 一致成立。这给出 $\hat{\mu}_\phi = \mu_\phi + o_{\mathbb{P}}(1/(nh_n^{1/2})^{1/3})$。于是我们得到

$$|\widehat{T}_{D_\phi^{(k)}} - T_{D_\phi^{(k)}}| = \frac{1}{2}|T'(\lambda^{**})| \leq \frac{1}{|\sigma_\phi + \lambda^{**}(\widehat{\sigma}_\phi - \sigma_\phi)|} o_{\mathbb{P}}\Big(\frac{1}{(nh_n^2)^{1/3}}\Big) + o_{\mathbb{P}}(1) \tag{7.86}$$

根据假设(A4)。因此， $\widehat{T}^{(n)}_{D_\phi^{(k)}}$ 和 $T_{D_\phi^{(k)}}$ 在分布上收敛到同一个极限。因此，在原假设H下，$\widehat{T}^{(n)}_{D_\phi^{(k)}} \xrightarrow{\mathcal{L}} \mathrm{N}(0,1)$。

(ii) 引理 3.2 以及 $\widehat{\mu_\phi}, \widehat{\sigma}_\phi$ 的一致性，意味着

$$(nh_n^{1/2})^{-1}\widehat{T}_{D_\phi^{(k)}} = \frac{\widehat{I}_{D_\phi^{(k)}} - \frac{\widehat{\mu_\phi}}{nh_n}}{\widehat{\sigma}_\phi} \xrightarrow{\mathbb{P}} \frac{I_{D_\phi^{(k)}}}{\sigma_\phi} = \begin{cases} 0 & \text{if null hypothesis is true ,} \\ t & \text{if alternative hypothesis is true} \end{cases} \tag{7.87}$$

对于某个固定的 $t>0$。因此，当备择假设K成立时，$T_{D_\phi^{(k)}}$ 依概率发散。这证明了 $T_{D_\phi^{(k)}}$ 的相合性。

引理 *3.4*

考虑

$$\zeta(t) = \sum_{x=0}^{1} \int_{f_y>0} \Big[\phi\Big((f_x f_y + t\Delta_{x,y})^k\Big) - \phi\Big(g_{x,y}(t)^k\Big) - \Big\{(f_x f_y + t\Delta_{x,y})^k - g_{x,y}(t)^k\Big\}\phi'\Big(g_{x,y}(t)^k\Big)\Big] dy \tag{7.88}$$

其中 $g_{x,y}(t) = (f_x + t\Delta_x)(f_y + t\Delta_y)$，且 $t = \frac{d}{\sqrt{nh_n^{1/2}}}$注意到

$$f_{x,y}^{(n)} = f_x f_y + t\Delta_{x,y},\ f_x^{(n)} = f_x + t\Delta_x \text{ and } f_y^{(n)} = f_y + t\Delta_y. \tag{7.89}$$

注意， $\zeta(0) = \zeta'(0) = 0$。将 $\zeta(t)$ 在 $t=0$ 附近展开至二阶，得到

$$\zeta(t) = \zeta(0) + t\zeta'(0) + \frac{t^2}{2}\zeta''(0) + \frac{t^3}{6}\zeta'''(\theta) \text{ for } 0<\theta<t. \tag{7.90}$$

经简单计算可得

$$
\begin{aligned}
\zeta''(t) = \sum_{x=0}^{1} \int_{f_y>0} \Bigg[& k^2 (f_x f_y + t\Delta_{x,y})^{2k-2} \Delta_{x,y}^2 \phi''\Big((f_x f_y + t\Delta_{x,y})^k\Big) \\
& + k(k-1)(f_x f_y + t\Delta_{x,y})^{k-2} \Delta_{x,y}^2 \phi'\Big((f_x f_y + t\Delta_{x,y})^k\Big) \\
& - k^2 (f_x f_y + t\Delta_{x,y})^{k-1} \Delta_{x,y} g_{x,y}^{k-1}(t) g'_{x,y}(t) \phi''\Big(g_{x,y}^k(t)\Big) \\
& - k(k-1)(f_x f_y + t\Delta_{x,y})^{k-2} \Delta_{x,y}^2 \phi'\Big(g_{x,y}^k(t)\Big) \\
& - \Big\{(f_x f_y + t\Delta_{x,y})^k - g_{x,y}^k(t)\Big\} k^2 g_{x,y}^{2k-2}(t) (g'_{x,y}(t))^2 \phi'''\Big(g_{x,y}^k(t)\Big) \\
& - \Big\{(f_x f_y + t\Delta_{x,y})^k - g_{x,y}^k(t)\Big\} k g_{x,y}^{k-1}(t) g''_{x,y}(t) \phi''\Big(g_{x,y}^k(t)\Big) \\
& - \Big\{(f_x f_y + t\Delta_{x,y})^k - g_{x,y}^k(t)\Big\} k(k-1) g_{x,y}^{k-2}(t) (g'_{x,y}(t))^2 \phi''\Big(g_{x,y}^k(t)\Big) \\
& - \Big\{k(f_x f_y + t\Delta_{x,y})^{k-1} \Delta_{x,y} - k g_{x,y}^{k-1}(t) g'_{x,y}(t)\Big\} k g_{x,y}^{k-1}(t) g'_{x,y}(t) \phi''\Big(g_{x,y}^k(t)\Big) \Bigg] dy \\
= & \sum_{x=0}^{1} \int_{f_y>0} k^2 (f_x f_y)^{2k} \phi''(f_x^k f_y^k) \left(\frac{\Delta_x}{f_x} + \frac{\Delta_y}{f_y} - \frac{\Delta_{x,y}}{f_x f_y} \right)^2 dy \text{ at } t = 0, \qquad (7.91)
\end{aligned}
$$

$$
\begin{aligned}
\zeta'''(\theta) =& \sum_{x=0}^{1} \int_{f_y>0} \Bigg[k^3 (f_x f_y + \theta \Delta_{x,y})^{3k-3} \Delta_{x,y}^3 \phi''' \Big((f_x f_y + \theta \Delta_{x,y})^k \Big) \\
&+ k^2 (2k-2) (f_x f_y + \theta \Delta_{x,y})^{2k-3} \Delta_{x,y}^3 \phi'' \Big((f_x f_y + \theta \Delta_{x,y})^k \Big) \\
&+ k^2 (k-1) (f_x f_y + \theta \Delta_{x,y})^{2k-3} \Delta_{x,y}^3 \phi'' \Big((f_x f_y + \theta \Delta_{x,y})^k \Big) \\
&+ k(k-1)(k-2) (f_x f_y + \theta \Delta_{x,y})^{k-3} \Delta_{x,y}^3 \phi' \Big((f_x f_y + \theta \Delta_{x,y})^k \Big) \\
&- k^3 (f_x f_y + \theta \Delta_{x,y})^{k-1} \Delta_{x,y} g_{x,y}^{2k-2}(\theta) (g'_{x,y}(\theta))^2 \phi''' \Big(g_{x,y}^k(\theta) \Big) \\
&- k^2 (f_x f_y + \theta \Delta_{x,y})^{k-1} \Delta_{x,y} g_{x,y}^{k-1}(\theta) g''_{x,y}(\theta) \phi'' \Big(g_{x,y}^k(\theta) \Big) \\
&- k^2 (k-1) (f_x f_y + \theta \Delta_{x,y})^{k-1} \Delta_{x,y} g_{x,y}^{k-2}(\theta) (g'_{x,y}(\theta))^2 \phi'' \Big(g_{x,y}^k(\theta) \Big) \\
&- k^2 (k-1) (f_x f_y + \theta \Delta_{x,y})^{k-2} \Delta_{x,y}^2 g_{x,y}^{k-1}(\theta) g'_{x,y}(\theta) \phi'' \Big(g_{x,y}^k(\theta) \Big) \\
&- k^2 (k-1) (f_x f_y + \theta \Delta_{x,y})^{k-2} \Delta_{x,y}^2 g_{x,y}^{k-1}(\theta) g'_{x,y}(\theta) \phi'' \Big(g_{x,y}^k(\theta) \Big) \\
&- k(k-1)(k-2) (f_x f_y + \theta \Delta_{x,y})^{k-3} \Delta_{x,y}^3 \phi' \Big(g_{x,y}^k(\theta) \Big) \\
&- \Big\{ (f_x f_y + \theta \Delta_{x,y})^k - g_{x,y}^k(\theta) \Big\} k^3 g_{x,y}^{3k-3}(\theta) (g'_{x,y}(\theta))^3 \phi'''' \Big(g_{x,y}^k(\theta) \Big) \\
&- \Big\{ (f_x f_y + \theta \Delta_{x,y})^k - g_{x,y}^k(\theta) \Big\} k^2 g_{x,y}^{2k-2}(\theta) 2 g'_{x,y}(\theta) g''_{x,y}(\theta) \phi''' \Big(g_{x,y}^k(\theta) \Big) \\
&- \Big\{ (f_x f_y + \theta \Delta_{x,y})^k - g_{x,y}^k(\theta) \Big\} k^2 (2k-2) g_{x,y}^{2k-3}(\theta) (g'_{x,y}(\theta))^3 \phi''' \Big(g_{x,y}^k(\theta) \Big) \\
&- \Big\{ k (f_x f_y + \theta \Delta_{x,y})^{k-1} \Delta_{x,y} - k g_{x,y}^{k-1}(\theta) g'_{x,y}(\theta) \Big\} k^2 g_{x,y}^{2k-2}(\theta) (g'_{x,y}(\theta))^2 \phi''' \Big(g_{x,y}^k(\theta) \Big) \\
&- \Big\{ (f_x f_y + \theta \Delta_{x,y})^k - g_{x,y}^k(\theta) \Big\} k^2 g_{x,y}^{2k-2}(\theta) g'_{x,y}(\theta) g''_{x,y}(\theta) \phi''' \Big(g_{x,y}^k(\theta) \Big) \\
&- \Big\{ (f_x f_y + \theta \Delta_{x,y})^k - g_{x,y}^k(\theta) \Big\} k(k-1) g_{x,y}^{k-2}(\theta) g'_{x,y}(\theta) g''_{x,y}(\theta) \phi'' \Big(g_{x,y}^k(\theta) \Big) \\
&- \Big\{ k (f_x f_y + \theta \Delta_{x,y})^{k-1} \Delta_{x,y} - k g_{x,y}^{k-1}(\theta) g'_{x,y}(\theta) \Big\} k g_{x,y}^{k-1}(\theta) g''_{x,y}(\theta) \phi'' \Big(g_{x,y}^k(\theta) \Big) \\
&- \Big\{ (f_x f_y + \theta \Delta_{x,y})^k - g_{x,y}^k(\theta) \Big\} k^2 (k-1) g_{x,y}^{2k-3}(\theta) (g'_{x,y}(\theta))^3 \phi''' \Big(g_{x,y}^k(\theta) \Big) \\
&- \Big\{ (f_x f_y + \theta \Delta_{x,y})^k - g_{x,y}^k(\theta) \Big\} k(k-1) g_{x,y}^{k-2}(\theta) 2 g'_{x,y}(\theta) g''_{x,y}(\theta) \phi'' \Big(g_{x,y}^k(\theta) \Big) \\
&- \Big\{ (f_x f_y + \theta \Delta_{x,y})^k - g_{x,y}^k(\theta) \Big\} k(k-1)(k-2) g_{x,y}^{k-3}(\theta) (g'_{x,y}(\theta))^3 \phi \Big(g_{x,y}^k(\theta) \Big) \\
&- \Big\{ k (f_x f_y + \theta \Delta_{x,y})^{k-1} \Delta_{x,y} - k g_{x,y}^{k-1}(\theta) g'_{x,y}(\theta) \Big\} k(k-1) g_{x,y}^{k-2}(\theta) (g'_{x,y}(\theta))^2 \phi'' \Big(g_{x,y}^k(\theta) \Big) \\
&- \Big\{ k (f_x f_y + \theta \Delta_{x,y})^{k-1} \Delta_{x,y} - k g_{x,y}^{k-1}(\theta) g'_{x,y}(\theta) \Big\} k^2 g_{x,y}^{2k-2}(\theta) (g'_{x,y}(\theta))^2 \phi \Big(g_{x,y}^k(\theta) \Big) \\
&- \Big\{ k (f_x f_y + \theta \Delta_{x,y})^{k-1} \Delta_{x,y} - k g_{x,y}^{k-1}(\theta) g'_{x,y}(\theta) \Big\} k g_{x,y}^{k-1}(\theta) g''_{x,y}(\theta) \phi'' \Big(g_{x,y}^k(\theta) \Big) \\
&- \Big\{ k (f_x f_y + \theta \Delta_{x,y})^{k-1} \Delta_{x,y} - k g_{x,y}^{k-1}(\theta) g'_{x,y}(\theta) \Big\} k(k-1) g_{x,y}^{k-2}(\theta) (g'_{x,y}(\theta))^2 \phi'' \Big(g_{x,y}^k(\theta) \Big) \\
&- \Big\{ k(k-1) (f_x f_y + \theta \Delta_{x,y})^{k-2} \Delta_{x,y}^2 - k g_{x,y}^{k-1}(\theta) g''_{x,y}(\theta) - k(k-1) g_{x,y}^{k-2}(\theta) (g'_{x,y}(\theta))^2 \Big\} \\
&\times k g_{x,y}^{k-1}(\theta) g'_{x,y}(\theta) \phi'' \Big(g_{x,y}^k(\theta) \Big) \Bigg] dy. \qquad (7.92)
\end{aligned}
$$

假设 (A1) 确保 $|\zeta'''(\theta)| = \mathcal{O}(1)$。因此，$\zeta(t)$ 在（7.90）中的展开余项为 $\mathcal{O}(t^3) = o(t^2)$ 当 $t \to 0$ 时。将 $t = \sqrt[d]{nh_n^{1/2}}$ 代入（7.90）得

$$I_{D_\phi^{(k)}}^{(n)} = \zeta(t) = \frac{d^2}{2(nh_n^{1/2})} \sum_{x=0}^{1} \int_{f_y>0} k^2 (f_x f_y)^{2k} \phi''(f_x^k f_y^k) \left(\frac{\Delta_x}{f_x} + \frac{\Delta_y}{f_y} - \frac{\Delta_{x,y}}{f_x f_y} \right)^2 dy + o\left(\frac{d^2}{nh_n^{1/2}} \right). \tag{7.93}$$

至此，证明完成。

定理 *3.3*

现在可以看到

$$\begin{aligned} T_{D_\phi^{(k)}} &= nh_n^{1/2} \sigma_\phi^{-1} \left(\widehat{I}_{D_\phi^{(k)}} - \frac{\mu_\phi}{nh_n} \right) \\ &= nh_n^{1/2} \sigma_\phi^{-1} \left(\widehat{I}_{D_\phi^{(k)}} - I_{D_\phi^{(k)}}^{(n)} - \frac{\mu_\phi}{nh_n} \right) + nh_n^{1/2} \sigma_\phi^{-1} I_{D_\phi^{(k)}}^{(n)} \\ &= nh_n^{1/2} \sigma_\phi^{-1} \left(\widehat{I}_{D_\phi^{(k)}} - I_{D_\phi^{(k)}}^{(n)} - \frac{\mu_\phi}{nh_n} \right) \\ &\quad + \frac{d^2}{2\sigma_\phi} \sum_{x=0}^{1} \int_{f_y>0} k^2 (f_x f_y)^{2k} \phi''(f_x^k f_y^k) \left(\frac{\Delta_x}{f_x} + \frac{\Delta_y}{f_y} - \frac{\Delta_{x,y}}{f_x f_y} \right)^2 dy + o(1). \end{aligned} \tag{7.94}$$

定义

$$U_n = nh_n^{1/2} \sigma_\phi^{-1} \left(\widehat{I}_{D_\phi^{(k)}} - I_{D_\phi^{(k)}}^{(n)} - \frac{\mu_\phi}{nh_n} \right), \; W_n(t) = \mathbb{1}\{U_n \le t\} \text{ and } L_n = \prod_{i=1}^{n} \frac{f_{X_i,Y_i}^{(n)}(x_i, y_i)}{f_{X_i}(x_i) f_{Y_i}(y_i)} \tag{7.95}$$

对任意固定的 $t \in \mathbb{R}$。原假设下的概率密度函数为 $f_{x,\ y} = f_x f_y$ 对所有 x，y均成立。令 $F^{(n)}$ 和 F_0 分别为与邻近备择和原假设相对应的分布函数。由于 $f_{x}^{(n),\ y} \to f_x f_y$ 逐点成立，根据 Scheff`e 定理可得$F^{(n)} \xrightarrow{\mathcal{L}} F_0$。然后应用 Portmanteau 引理可得

$$\limsup_{n\to\infty} \mathbb{P}_{\mathbb{K}_n}\{U_n \le t\} \le \limsup_{n\to\infty} \mathbb{P}_{\mathbb{H}}\{U_n \le t\} = \Phi_1(t), \tag{7.96}$$

如在原假设 H 下 $d = 0$。反过来，可见

$$\begin{aligned} \mathbb{P}_{\mathbb{K}_n}\{U_n \le t\} = \mathbb{E}_{\mathbb{K}_n}\left[W_n(t)\right] &\ge \sum_{\prod_{i=1}^{n} f_{x_i} f_{y_i} > 0} \int W_n(t) \prod_{i=1}^{n} f_{X_i Y_i}^{(n)}(x_i, y_i) dy_i \\ &= \mathbb{E}_{\mathbb{H}}\left[W_n(t) L_n\right] \text{ for all } n. \end{aligned} \tag{7.97}$$

因此，我们有 $\liminf_{n\to\infty} \mathbb{P}_{\mathbb{K}_n}\{U_n \le t\} \ge \liminf_{n\to\infty} \mathbb{E}_{\mathbb{H}}\Big[W_n(t)L_n\Big]$。在原假设 H 下，可见

$$L_n = 1 \text{ and } W_n(t) = \mathbb{1}\left\{nh_n^{1/2}\sigma_\phi^{-1}\left(\widehat{I}_{D_\phi^{(k)}} - \frac{\mu_\phi}{nh_n^{1/2}}\right) \le t\right\} \xrightarrow{\mathcal{L}} T, \tag{7.98}$$

其中 $T \sim Bernoulli(\Phi_1(t))$。由此可得，对于原假设下的任意 $t \in \mathbb{R}$，$W_n(t)L_n \xrightarrow{\mathcal{L}} T$。由于 W_n 有界，因此它一致可积。于是得到 $\mathbb{E}_{\mathbb{H}}(W_n(t)L_n) \longrightarrow \mathbb{E}(T) = \Phi_1(t)$，从而我们得到

$$\mathbb{P}_{\mathbb{K}_n}\left\{nh_n^{1/2}\sigma_\phi^{-1}\Big(\widehat{I}_{D_\phi^{(k)}} - I^{(n)}_{D_\phi^{(k)}} - \frac{\mu_\phi}{nh_n}\Big) \le t\right\} \longrightarrow \Phi_1(t) \text{ for all } t \in \mathbb{R} \tag{7.99}$$

当 $n \to \infty$ 时。综合以上各部分可得

$T_{D_\phi^{(k)}} \xrightarrow{\mathcal{L}} \frac{d^2}{2\sigma_\phi}\sum_{x=0}^{1}\int_{f_y>0} k^2(f_xf_y)^{2k}\phi''(f_x^kf_y^k)\Big(\frac{\Delta_x}{f_x} + \frac{\Delta_y}{f_y} - \frac{\Delta_{x,y}}{f_xf_y}\Big)^2 dy + \mathcal{N}(0,1)$ under $\mathbb{K}_n$. This completes the proof.

定理 *3.3*

我们知道

$$\begin{aligned}
T_{D_\phi^{(k)}} &= nh_n^{1/2}\sigma_\phi^{-1}\Big(\widehat{I}_{D_\phi^{(k)}} - \frac{\mu_\phi}{nh_n}\Big) \\
&= nh_n^{1/2}\sigma_\phi^{-1}\Big(\widehat{I}_{D_\phi^{(k)}} - I^{(n)}_{D_\phi^{(k)}} - \frac{\mu_\phi}{nh_n}\Big) + nh_n^{1/2}\sigma_\phi^{-1}I^{(n)}_{D_\phi^{(k)}} \\
&= nh_n^{1/2}\sigma_\phi^{-1}\Big(\widehat{I}_{D_\phi^{(k)}} - I^{(n)}_{D_\phi^{(k)}} - \frac{\mu_\phi}{nh_n}\Big) + \frac{d^2}{2\sigma_\phi}\sum_{x=0}^{1}\int_{f_y>0} k^2(f_xf_y)^{2k}\phi''(f_x^kf_y^k)\left(\frac{\Delta_x}{f_x} + \frac{\Delta_y}{f_y} - \frac{\Delta_{x,y}}{f_xf_y}\right)^2 dy + o(1).
\end{aligned} \tag{7.100}$$

对于固定的 $t \in \mathbb{R}$，定义

$$W_n(t) = \mathbb{1}\left\{nh_n^{1/2}\sigma_\phi^{-1}\Big(\widehat{I}_{D_\phi^{(k)}} - I^{(n)}_{D_\phi^{(k)}} - \frac{\mu_\phi}{nh_n}\Big) \le t\right\} \text{ and } L_n = \prod_{i=1}^{n}\frac{f^{(n)}_{X_i,Y_i}}{f_{X_i}f_{Y_i}}. \tag{7.101}$$

注意到

$$\begin{aligned}
\mathbb{P}_{\mathbb{K}_n}\left[nh_n^{1/2}\sigma_\phi^{-1}\Big(\widehat{I}_{D_\phi^{(k)}} - I^{(n)}_{D_\phi^{(k)}} - \frac{\mu_\phi}{nh_n}\Big) \le t\right] &= \mathbb{E}_{\mathbb{K}_n}\Big(W_n(t)\Big) \\
&= \mathbb{E}_{\mathbb{H}}\Big(W_n(t)L_n\Big).
\end{aligned} \tag{7.102}$$

上述表达式成立，因为 $\mathbb{K}_n$ 下的密度与 H 下的概率密度连续（形式上记为 $\mathbb{K}_n \lhd \mathbb{H}$；参见 Van der vaart（2000） [29], 第 87 页）。若 H 为真，

则 $d = 0$ 和 $I_{D_\phi^{(k)}}^{(n)} = 0$。这意味着

$$L_n \overset{\mathbb{H}}{=} 1 \text{ and } W_n(t) \overset{\mathbb{H}}{=} \mathbb{1}\left\{ nh_n^{1/2}\sigma_\phi^{-1}\left(\widehat{I}_{D_\phi^{(k)}} - \frac{\mu_\phi}{nh_n^{1/2}}\right) \le t \right\} \xrightarrow{\mathcal{L}} T, \tag{7.103}$$

其中 $T \sim Bernoulli(\Phi(t))$。因此 $W_n(t)L_n \xrightarrow{\mathcal{L}} T$ 对任意 $t \in \mathbb{R}$，在原假设 H 下成立。由于W_n 有界，因此它是一致可积的，从而 $\mathbb{E}_{\mathbb{H}}(W_n(t)L_n) \longrightarrow \mathbb{E}(T) = \Phi(t)$。因此，对于$n \to \infty$,

$$\mathbb{P}_{\mathbb{K}_n}\left[nh_n^{1/2}\sigma_\phi^{-1}\left(\widehat{I}_{D_\phi^{(k)}} - I_{D_\phi^{(k)}}^{(n)} - \frac{\mu_\phi}{nh_n}\right) \le t \right] = \mathbb{E}_{\mathbb{H}}\Big(W_n(t)L_n\Big) \longrightarrow \Phi(t), \text{ for all } t \in \mathbb{R}. \tag{7.104}$$

因此 $T_{D_\phi^{(k)}} \xrightarrow{\mathcal{L}} \frac{d^2}{2\sigma_\phi} \sum_{x=0}^{1} \int_{f_y>0} k^2 (f_x f_y)^{2k} \phi''(f_x^k f_y^k)\left(\frac{\Delta_x}{f_x} + \frac{\Delta_y}{f_y} - \frac{\Delta_x}{f_x f_y} \right)^2 dy + \mathrm{N}(0,\ 1)$ 在 $\mathbb{K}_n$ 下成立。至此，证明完成。

引理 *4.1*

记

$$\Gamma(s,t) = \sum_{x=0}^{1} \int \left\{ \phi\Big((f_{x,y}^{(P_n)})^k\Big) - \phi\Big((f_x^{(P_n)})^k (f_y^{(P_n)})^k\Big) - \Big((f_{x,y}^{(P_n)})^k - (f_x^{(P_n)})^k (f_y^{(P_n)})^k\Big)\phi'\Big((f_x^{(P_n)})^k (f_y^{(P_n)})^k\Big) \right\} dy$$

其中 $s = \frac{\epsilon}{\sqrt{nh_n^{1/2}}}$ 且 $t = \frac{d}{\sqrt{nh_n^{1/2}}}$ 将 $\Gamma(s,\ t)$ 在 $s = 0$ 处展开至二阶，得到

$$\Gamma(s,t) = \Gamma(0,t) + s\Gamma^{(10)}(0,t) + \frac{s^2}{2}\Gamma^{(20)}(0,t) + \frac{s^3}{3}\Gamma^{(30)}(s^*,t) \text{ with } 0 < s^* < s. \tag{7.105}$$

这里，偏导数记作 $\Gamma^{(r_1 r_2)}(s_1,t_1) = \frac{\partial^{r_1+r_2}}{\partial s^{r_1}\partial s^{r_2}}\Gamma(s,t)\mathrm{RRRRRRRRRRR}_{s=s_1,t=t_1}$, $r_1, r_2 = 1, 2, \ldots$. 我们得到 $f_x^{(P_n)},_y = f(n)_{x,y}$，对于所有 x, y，当 $s = 0$ 时。这给出 $\Gamma(0,t) = I_{D_\phi^{(k)}}^{(n)}$，如式 (3.31) 所示。根据引理 3.4，我们知道

$$\Gamma(0,t) = I_{D_\phi^{(k)}}^{(n)} = \frac{d^2}{2nh_n^{1/2}} \sum_{x=0}^{1} \int k^2 (f_x f_y)^{2k} \left(\frac{\Delta_{x,y}}{f_x f_y} - \frac{\Delta_x}{f_x} - \frac{\Delta_y}{f_y} \right)^2 \phi''\Big(f_x^k f_y^k\Big) dy + o\left(\frac{d^2}{nh_n^{1/2}}\right). \tag{7.106}$$

通过简单计算可得

$$\Gamma^{(10)}(0,t) = \sum_{x=0}^{1} \int_{f_y>0} \Bigg[k\Big(f_{x,y}^{(n)}\Big)^k \frac{\Delta'_{x,y}}{f_{x,y}^{(n)}} \Big\{ \phi'\Big((f_{x,y}^{(n)})^k\Big) - \phi'\Big((f_x^{(n)} f_y^{(n)})^k\Big) \Big\} \\ - k\left(\frac{\Delta'_x}{f_x^{(n)}} + \frac{\Delta'_y}{f_y^{(n)}} \right) \Big((f_{x,y}^{(n)})^k - (f_x^{(n)} f_y^{(n)})^k \Big) \Big(f_x^{(n)} f_y^{(n)} \Big)^k \phi''\Big((f_x^{(n)} f_y^{(n)})^k\Big) \Bigg] dy. \quad (7.107)$$

将 $\Gamma^{(10)}(0,\ t)$ 在 $t = 0$ 附近展开至 给出

$$\Gamma^{(10)}(0,t) = \Gamma^{(10)}(0,0) + t\Gamma^{(11)}(0,0) + \frac{t^2}{2}\Gamma^{(12)}(0,t^*) \text{ where } 0 < t^* < t. \quad (7.108)$$

简单计算可得 $\Gamma^{(10)}(0,0) = 0$ 以及

$$\Gamma^{(11)}(0,0) = \sum_{x=0}^{1} \int k^2 (f_x f_y)^{2k} \left(\frac{\Delta_{x,y}}{f_x f_y} - \frac{\Delta_x}{f_x} - \frac{\Delta_y}{f_y} \right) \left(\frac{\Delta'_{x,y}}{f_x f_y} - \frac{\Delta'_x}{f_x} - \frac{\Delta'_y}{f_y} \right) \phi''\Big(f_x^k f_y^k\Big) dy. \quad (7.109)$$

采用与之前相同的论证，可以证明余项为 $o(t)$。由此我们得到

$$s\Gamma^{(10)}(0,t) = \frac{d\epsilon}{n h_n^{1/2}} \sum_{x=0}^{1} \int k^2 (f_x f_y)^{2k} \left(\frac{\Delta_{x,y}}{f_x f_y} - \frac{\Delta_x}{f_x} - \frac{\Delta_y}{f_y} \right) \left(\frac{\Delta'_{x,y}}{f_x f_y} - \frac{\Delta'_x}{f_x} - \frac{\Delta'_y}{f_y} \right) \phi''\Big(f_x^k f_y^k\Big) dy + o\left(\frac{d\epsilon}{n h_n^{1/2}} \right). \quad (7.110)$$

接下来，注意到

$$\begin{aligned}\Gamma^{(20)}(0,t) = \sum_{x=0}^{1}\int \Bigg\{ & k(k-1)(f_{x,y}^{(n)})^k\left(\frac{\Delta'_{x,y}}{f_{x,y}^{(n)}}\right)^2\left\{\phi'\left(f_{x,y}^{(n)}\right)^k - \phi'\left(f_x^{(n)}f_y^{(n)}\right)^k\right\} \\ & + k^2(f_{x,y}^{(n)})^{2k}\left(\frac{\Delta'_{x,y}}{f_{x,y}^{(n)}}\right)^2\phi''\left((f_{x,y}^{(n)})^k\right) \\ & - k^2(f_{x,y}^{(n)}f_x^{(n)}f_y^{(n)})^k\frac{\Delta'_{x,y}}{f_{x,y}^{(n)}}\left(\frac{\Delta'_x}{f_x^{(n)}}+\frac{\Delta'_y}{f_y^{(n)}}\right)\phi''\left((f_x^{(n)}f_y^{(n)})^k\right) \\ & - k^2(f_x^{(n)}f_y^{(n)})^{2k}\left\{\left(\frac{f_{x,y}^{(n)}}{f_x^{(n)}f_y^{(n)}}\right)^k\frac{\Delta'_{x,y}}{f_{x,y}^{(n)}} - \frac{\Delta'_x}{f_x^{(n)}} - \frac{\Delta'_y}{f_y^{(n)}}\right\}\left(\frac{\Delta'_x}{f_x^{(n)}}+\frac{\Delta'_y}{f_y^{(n)}}\right)\phi''\left((f_x^{(n)}f_y^{(n)})^k\right) \\ & - \left((f_{x,y}^{(n)})^k - (f_x^{(n)}f_y^{(n)})^k\right)(f_x^{(n)}f_y^{(n)})^k\Bigg\{k(k-1)\left(\frac{\Delta'_x}{f_x^{(n)}}\right)^2 + 2k^2\left(\frac{\Delta'_x\Delta'_y}{f_x^{(n)}f_y^{(n)}}\right) \\ & + k(k-1)\left(\frac{\Delta'_y}{f_y}\right)^2\Bigg\}\phi''\left(\left(f_x^{(n)}f_y^{(n)}\right)^k\right) \\ & - k\left\{(f_{x,y}^{(n)})^k - (f_x^{(n)}f_y^{(n)})^k\right\}\left(\frac{\Delta'_x}{f_x^{(n)}}+\frac{\Delta'_y}{f_y^{(n)}}\right)^2\phi'''\left((f_x^{(n)}f_y^{(n)})^k\right)\Bigg\}dy. \end{aligned} \tag{7.111}$$

与之前一样，我们将 $\Gamma^{(20)}(0,t)$在 $t=0$ 附近展开，并得到

$$\Gamma^{(20)}(0,t) = \Gamma^{(20)}(0,0) + t\Gamma^{(21)}(0,t^{**}) \text{ where } 0 < t^{**} < t, \tag{7.112}$$

其中

$$\Gamma^{(20)}(0,0) = \sum_{x=0}^{1}\int k^2(f_xf_y)^{2k}\left(\frac{\Delta'_{x,y}}{f_xf_y} - \frac{\Delta'_x}{f_x} - \frac{\Delta'_y}{f_y}\right)^2\phi''\left(f_x^kf_y^k\right)dy. \tag{7.113}$$

因此我们得到

$$\frac{s^2}{2}\Gamma^{(20)}(0,t) = \frac{\epsilon^2}{2nh_n^{1/2}}\sum_{x=0}^{1}\int k^2(f_xf_y)^{2k}\left(\frac{\Delta'_{x,y}}{f_xf_y} - \frac{\Delta'_x}{f_x} - \frac{\Delta'_y}{f_y}\right)^2\phi''\left(f_x^kf_y^k\right)dy + O\left(\frac{s^2t}{2}\right). \tag{7.114}$$

结合式（7.106）、（7.110）和（7.114），可得

$$I_{D_\phi^{(k)}}^{(P_n)} = \frac{k^2}{2nh_n^{1/2}}\sum_{x=0}^{1}\int (f_xf_y)^{2k}\left\{d\left(\frac{\Delta_{x,y}}{f_xf_y} - \frac{\Delta_x}{f_x} - \frac{\Delta_y}{f_y}\right) + \epsilon\left(\frac{\Delta'_{x,y}}{f_xf_y} - \frac{\Delta'_x}{f_x} - \frac{\Delta'_y}{f_y}\right)\right\}^2\phi''\left(f_x^kf_y^k\right)dy + o\left(\frac{(d+\epsilon)^2}{nh_n^{1/2}}\right).$$

至此证明完毕。

定理 *4.2*

利用引理 4.1可知

$$T_{D_\phi^{(k)}} = nh_n^{1/2}\sigma_\phi^{-1}\Big(\widehat{I}_{D_\phi^{(k)}} - \frac{\mu_\phi}{nh_n^{1/2}} - I_{D_\phi^{(k)}}^{(P_n)}\Big) + nh_n^{1/2}\sigma_\phi^{-1} I_{D_\phi^{(k)}}^{(P_n)}$$
$$= U_n + \frac{k^2}{2\sigma_\phi}\sum_{x=0}^{1}\int (f_x f_y)^{2k}\left\{d\left(\frac{\Delta_{x,y}}{f_x f_y} - \frac{\Delta_x}{f_x} - \frac{\Delta_y}{f_y}\right) + \epsilon\left(\frac{\Delta'_{x,y}}{f_x f_y} - \frac{\Delta'_x}{f_x} - \frac{\Delta'_y}{f_y}\right)\right\}^2 \phi''\Big(f_x^k f_y^k\Big)dy + o(1),$$

其中 $U_n = nh_n^{1/2}\sigma_\phi^{-1}\Big(\widehat{I}_{D_\phi^{(k)}} - \frac{\mu_\phi}{nh_n^{1/2}} - I^{(P_n)}{}_{D_\phi^{(k)}}\Big)$。注意到 $f_{x\ y}^{(P_n)}{}_{,} \longrightarrow f_x f_y$ 对所有 x， y，当 $n \to \infty$时。因此 $f^{(P_n)}{}_{X,\ Y}$ 相对于原假设下的联合密度是连续的。应用定理 3.3中 Portmanteau 的结果，我们得到

$$\lim_{n\to\infty}\mathbb{P}_{f_{X,Y}^{(P_n)}}\Big[U_n \le t\Big] = \lim_{n\to\infty}\mathbb{E}_{f_{X,Y}^{(P_n)}}\Big(\mathbb{1}\{U_n \le t\}\Big) = \lim_{n\to\infty}\mathbb{E}_{\mathbb{H}}\Big(\mathbb{1}\{U_n \le t\}L_n\Big) \text{ for fixed } t \in \mathbb{R}, \quad (7.115)$$

其中 $L_n = \prod_{i=1}^n \frac{f_x^{(P_n)}}{f_x f_y}{}_{:\ y}$。根据条件 $0 \le \max\{d,\ \epsilon\} \le C^* \sup_{x,\ y}|f_{x,\ y} - f_x f_y|$ 我们得到 $d = \epsilon = 0$，因此 $f_x^{(P_n)}{}_{,\ y} = f_x f_y$ 对所有 x， y 当原假设成立时。因此我们得到 $L_n = 1$ 并且

$$\mathbb{1}\Big\{U_n \le t\Big\} = \mathbb{1}\left\{nh_n^{1/2}\Big(\widehat{I}_{D_\phi^{(k)}} - \frac{\mu_\phi}{nh_n}\Big) \le t\right\} \xrightarrow{L} T \text{ under } \mathbb{H}, \quad (7.116)$$

其中 $T \sim Bernoulli(\Phi_1(t))$。由于 $\mathbb{1}\{U_n \le t\}$ 有界，因此它是一致可积的。于是可得

$$\mathbb{P}_{f_{X,Y}^{(P_n)}}\Big[U_n \le t\Big] \longrightarrow \mathbb{E}_{\mathbb{H}}(T) = \Phi_1(t) \quad (7.117)$$

对所有 $t \in \mathbb{R}$。因此可得

$$T_{D_\phi^{(k)}} - \frac{k^2}{2\sigma_\phi}\sum_{x=0}^{1}\int (f_x f_y)^{2k}\left\{d\left(\frac{\Delta_{x,y}}{f_x f_y} - \frac{\Delta_x}{f_x} - \frac{\Delta_y}{f_y}\right) + \epsilon\left(\frac{\Delta'_{x,y}}{f_x f_y} - \frac{\Delta'_x}{f_x} - \frac{\Delta'_y}{f_y}\right)\right\}^2 \phi''\Big(f_x^k f_y^k\Big)dy \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)$$

在 $f^{(P_n)}{}_{X,Y}$ 当 $n \to \infty$时。这就完成了证明。

定理 *4.3*

(a) 记作

$$r_n = \frac{k^2}{2\sigma_\phi}\sum_{x=0}^{1}\int (f_x f_y)^{2k}\left\{d\left(\frac{\Delta_{x,y}}{f_x f_y} - \frac{\Delta_x}{f_x} - \frac{\Delta_y}{f_y}\right) + \epsilon\left(\frac{\Delta'_{x,y}}{f_x f_y} - \frac{\Delta'_x}{f_x} - \frac{\Delta'_y}{f_y}\right)\right\}^2 \phi''\Big(f_x^k f_y^k\Big)dy. \quad (7.118)$$

注意， $r_n \to r_*$ 其中

$$r_* = \frac{k^2}{2\sigma_\phi}\sum_{x=0}^{1}\int (f_x f_y)^{2k}\left\{d\left(\frac{\Delta_{x,y}}{f_x f_y}-\frac{\Delta_x}{f_x}-\frac{\Delta_y}{f_y}\right)+\epsilon\left(\frac{\Delta^*_{x,y}}{f_x f_y}-\frac{\Delta^*_x}{f_x}-\frac{\Delta^*_y}{f_y}\right)\right\}^2 \phi''\left(f_x^k f_y^k\right)dy. \tag{7.119}$$

由定理4.2 可知

$$\left|\mathbb{P}_{f_{X,Y}^{(P_n)}}\left\{T_{D_\phi^{(k)}}-t_n>\tau_c-r_n\right\}-(1-\Phi_1(\tau_c-r_n))\right|\longrightarrow 0 \text{ when } n\to\infty. \tag{7.120}$$

一阶泰勒级数展开给出

$$\Phi_1(\tau_c-r_n)=\Phi_1(\tau_c-r_*)+(r_*-r_n)\phi_1(\tau_c-r_*') \tag{7.121}$$

其中 $r_*^{'}$ 是介于 r_n， r_*之间的一个中间点。由于 ϕ_1 有界，

$$\mathbb{P}_{f_{X,Y}^{(P_n)}}\left\{T_{D_\phi^{(k)}}-r_n>\tau_c-r_n\right\}\longrightarrow 1-\Phi_1(\tau_c-r_*) \text{ when } n\to\infty. \tag{7.122}$$

关于{ $\frac{\partial}{\partial\epsilon}\pi(f^{(P_n)}{}_{X,\ Y},\ t_1)$}一致收敛的假设意味着求导与极限可以交换。这进而给出功效影响函数为

$$\mathcal{PIF}(\pi,t_1)=\left[\frac{\partial}{\partial\epsilon}\left\{\lim_{n\longrightarrow\infty}\pi(f_{X,Y}^{(P_n)},t_1)\right\}\right]_{\epsilon=0}=\left[\frac{\partial r_*}{\partial\epsilon}\cdot\phi_1(\tau_c-r_*)\right]_{\epsilon=0}. \tag{7.123}$$

当 $\epsilon=0$时，我们发现 $r_*=\frac{d^2}{2\sigma_\phi}$影响函数$_2(I_{D_\phi^{(k)}},\ f_X f_Y,\ t_1)$以及

$$\left[\frac{\partial r_*}{\partial\epsilon}\right]_{\epsilon=0}=\frac{d}{\sigma_\phi}\sum_{x=0}^{1}\int k^2(f_x f_y)^{2k}\left(\frac{\Delta^*_{x,y}}{f_x f_y}-\frac{\Delta^*_x}{f_x}-\frac{\Delta^*_y}{f_y}\right)\left(\frac{\Delta_{x,y}}{f_x f_y}-\frac{\Delta_x}{f_x}-\frac{\Delta_y}{f_y}\right)\phi''\left(f_x^k f_y^k\right)dy. \tag{7.124}$$

由此得到所需结果。

(b) 当 $t_0=t_1$ 时，我们得到 $\Delta_{x,\ y}=\Delta^*_{x,\ y}$ 对所有 x， y均成立。因此，结果得证。(c) 注意到，对于所有 x， y，当 $d=0$且 $t_0=t_1$ 时， $f_x^{(P_n),\ y}=f_{(n)}$成立。将 $d=0$ 和 $t_0=t_1$ 代入 (b) 中的 PIF 表达式可得， LIF$(\alpha,\ t_0)\equiv 0$。

证明至此完成。

定理 *4.4*

见

$$I_m=\int\sum\left\{\phi((f_{x,y}^m)^k)-\phi((f_x^m f_y^m)^k)-\left((f_{x,y}^m)^k-(f_x^m f_y^m)^k\right)\phi'((f_x^m f_y^m)^k)\right\}. \tag{7.125}$$

定义 $A_m = \{(x,y) : f_{x,y} > \max\{k_{xy,m}, f_x^m f_y^m\}\}$。为简洁起见，省略对 "$x$" 的求和以及对 "$y$" 的积分。注意，由于 $m \to \infty$,

$$\begin{aligned}\int \sum_{A_m} k_{xy,m} = \int \sum_{A_m} \min\{f_{x,y}, k_{xy,m}\} &\le \int \sum \min\{f_{x,y}, k_{xy,m}\} \\ &= \int \sum \min\{f_x f_{y|x}, k_{x,m} k_{y|x,m}\} \\ &\le \sum \max\{f_x, k_{x,m}\} \int \min\{f_{y|x}, k_{y|x,m}\} \to 0 \end{aligned} \tag{7.126}$$

根据假设 (BP1)，由于 f_x、k_x、$_m$ 在 $[0,1]$ 中是离散且有界的。因此我们得到 k_{xy}、$_m \to 0$，进而得到 f_x^m、$_y \to (1-\epsilon) f_x$、$_y$ 在 A_m 上满足 $m \to \infty$。类似地，我们还有

$$\begin{aligned}\int \sum_{A_m} f_x^m f_y^m = \int \sum_{A_m} \min\{f_{x,y}, f_x^m f_y^m\} &\le \int \sum_{A_m} \max\{f_x, f_x^m\} \min\{f_{y|x}, f_y^m\} \\ &\le \sum \max\{f_x, f_x^m\} \int \min\{f_{y|x}, f_y^m\} \to 0 \text{ as } m \to \infty, \end{aligned} \tag{7.127}$$

根据假设 (BP2)。因此我们得到 $f_x^m f_y^m \to 0$ 在 A_m 上，对 $m \to \infty$。现在注意到

$$\begin{aligned}\max\{k_{xy,m}, f_x^m f_y^m\} &\ge \min\{k_{x,m}, f_x^m\} \min\{k_{y|x,m}, f_y^m\} \\ &\ge \min\{k_{x,m}, f_x^m\} \min\{k_{y|x,m}, f_{y|x}, f_y^m\} \\ &= \min\{k_{x,m}, f_x^m\} \min\Big\{\min\{k_{y|x,m}, f_{y|x}\}, \min\{f_{y|x}, f_y^m\}\Big\}. \end{aligned} \tag{7.128}$$

由假设 (BP1) 和(BP2) 可知，当 $m \to \infty$ 时，最后一个因子趋于 0。这意味着$A_m \to \{(x, y) : f_{x,y} > 0\}$。因此我们得到

$$\begin{aligned}&\left| \int \sum_{A_m} \Big\{\phi((f_{x,y}^m)^k) - \phi((f_x^m f_y^m)^k) - \Big((f_{x,y}^m)^k - (f_x^m f_y^m)^k\Big)\phi'((f_x^m f_y^m)^k)\Big\} \right. \\ &\qquad \left. - \int \sum \Big\{\phi\Big((1-\epsilon)^k f_{x,y}^k\Big) - \phi(0) - \Big((1-\epsilon)^k f_{x,y}^k\Big)\phi'(0)\Big\} \right| \\ = &\left| \int \sum_{A_m} \Big\{\phi((f_{x,y}^m)^k) - \phi((f_x^m f_y^m)^k) - \Big((f_{x,y}^m)^k - (f_x^m f_y^m)^k\Big)\phi'((f_x^m f_y^m)^k)\Big\} - D_\phi^{(k)}\Big((1-\epsilon) f_{X,Y}, 0\Big) \right| \to 0 \end{aligned} \tag{7.129}$$

当 $m \to \infty$，因为 $\phi(0)$ $\phi'(0)$ 在假设 (BP4)中被假定为有限。接下来，注意到

$$\begin{aligned}\int \sum_{A_m^c} f_{x,y} &\le \int \sum \min\{f_{x,y}, k_{xy,m}\} + \int \sum \min\{f_{x,y}, f_x^m f_y^m\} \\ &\le \sum \max\{f_x, k_{x,m}\} \int \min\{f_{y|x}, k_{y|x,m}\} \\ &+ \sum \max\{f_x, f_x^m\} \int \min\{f_{y|x}, f_y^m\}. \end{aligned} \tag{7.130}$$

每个积分中的第一个因子都有界。因此，假设(BP1)和(BP2)共同意味着上述积分趋于 0，当 $m \to \infty$。因此，在 $f_{X,\,Y}$ 下， A_m^c 渐近地是一个零测集。因此我们得到

$$\left| \int \sum_{A_m^c} \left\{ \phi((f_{x,y}^m)^k) - \phi((f_x^m f_y^m)^k) - \left((f_{x,y}^m)^k - (f_x^m f_y^m)^k\right)\phi'((f_x^m f_y^m)^k) \right\} \right.$$
$$\left. - \underbrace{\sum \int \left\{ \phi((\epsilon k_{xy,m})^k) - \phi((f_x^m f_y^m)^k) - \left((\epsilon k_{xy,m})^k - (f_x^m f_y^m)^k\right)\phi'((f_x^m f_y^m)^k) \right\}}_{D_\phi^{(k)}(\epsilon k_{XY,m}, f_X^m f_Y^m)} \right| \to 0 \quad (7.131)$$

对于 $m \to \infty$。结合（7.129）和（7.131）可得

$$\left| I_m - D_\phi^{(k)}\Big((1-\epsilon) f_{X,Y}, 0\Big) - D_\phi^{(k)}\Big(\epsilon k_{XY,m}, f_X^m f_Y^m\Big) \right| \to 0 \text{ as } m \to \infty. \quad (7.132)$$

定义

$$a_1(\epsilon) = D_\phi^{(k)}\Big((1-\epsilon) f_{X,Y}, 0\Big) + \liminf_{m\to\infty} D_\phi^{(k)}\Big(\epsilon k_{XY,m}, f_X^m f_Y^m\Big). \quad (7.133)$$

接下来，注意到

$$\begin{aligned} \liminf_{m\to\infty} I_m &= \liminf_{m\to\infty} \left[D_\phi^{(k)}\Big((1-\epsilon) f_{X,Y}, 0\Big) + D_\phi^{(k)}\Big(\epsilon k_{XY,m}, f_X^m f_Y^m\Big) \right] \\ &\geq D_\phi^{(k)}((1-\epsilon) f_{X,Y}, 0) + \liminf_{m\to\infty} D_\phi^{(k)}(\epsilon k_{XY,m}, f_X^m f_Y^m) = a_1(\epsilon). \end{aligned} \quad (7.134)$$

此外，注意到 $\limsup I_m \leq \limsup I'_m$ 当 $I_m \leq I'_m$ 对所有 $m \geq 1$ 根据假设(BP5)。考虑$_{m\to\infty}$ $_{m\to\infty}$

$$B_m = \Big\{(x,y) : k_{xy,m} > \max\{f_{x,y}, f_x f_y\}\Big\}. \quad (7.135)$$

与之前类似， B_m 在$f_{X、\,Y}$和 $f_X f_Y$下渐近地为零测集。还要注意， B_m^c 在$k_{XY、\,m}$下渐近地为零测集。使用与之前相同的论证，我们得到

$$\left| I'_m - D_\phi^{(k)}(\epsilon k_{XY,m}, 0) - D_\phi^{(k)}\Big((1-\epsilon) f_{X,Y}, f_X f_Y\Big) \right| \to 0 \text{ for } m \to \infty. \quad (7.136)$$

接下来，可以看到

$$\begin{aligned} \limsup_{m\to\infty} I_m \leq \limsup_{m\to\infty} I'_m &= \limsup_{m\to\infty} \left[D_\phi^{(k)}(\epsilon k_{XY,m}, 0) + D_\phi^{(k)}\Big((1-\epsilon) f_{X,Y}, f_X f_Y\Big) \right] \\ &\leq \limsup_{m\to\infty} D_\phi^{(k)}(\epsilon k_{XY,m}, 0) + D_\phi^{(k)}\Big((1-\epsilon) f_{X,Y}, f_X f_Y\Big) = a_2(\epsilon) \text{ (say)}. \end{aligned} \quad (7.137)$$

假设 $a_2(\epsilon) < a_1(\epsilon)$，则 $\limsup I_m < \liminf I_{m\circ}$ 因此 $\lim I_m$ 存在，即当 $a_2(\epsilon) < a_1(\epsilon)$ 时，渐近地 $\text{there}_{m\to\infty}$ $_{m\to\infty}$ $_{m\to\infty}$将不会崩溃，这等同于

$$D_\phi^{(k)}(\epsilon k_{XY,m}, 0) + D_\phi^{(k)}\Big((1-\epsilon)f_{X,Y}, f_X f_Y\Big) < D_\phi^{(k)}(\epsilon k_{XY,m}, f_X^m f_Y^m) + D_\phi^{(k)}\Big((1-\epsilon)f_{X,Y}, 0\Big) \tag{7.138}$$

对于充分大的 $m_\circ$ 该条件等价于

$$\begin{aligned}
\liminf_{m\to\infty} D_\phi^{(k)}(\epsilon k_{XY,m}, f_X^m f_Y^m) &> \limsup_{m\to\infty} D_\phi^{(k)}(\epsilon k_{XY,m}, 0) + D_\phi^{(k)}\Big((1-\epsilon)f_{X,Y}, f_X f_Y\Big) - D_\phi^{(k)}\Big((1-\epsilon)f_{X,Y}, 0\Big) \\
&= \limsup_{m\to\infty} \sum \int \Big[\phi((\epsilon k_{xy,m})^k) - (\epsilon k_{xy,m})^k \phi'(0)\Big] \\
&\quad + \sum \int ((1-\epsilon)f_{xy})^k \phi'(0) - \phi(f_x^k f_y^k) \\
&\quad - \Big\{((1-\epsilon)f_{xy})^k - (f_x f_y)^k\Big\}\phi'(f_x^k f_y^k)\Big] dy.
\end{aligned} \tag{7.139}$$

显然， 假设 (BP6) 确保（7.139）对 $\epsilon < \tilde{\epsilon}$成立。证明完毕 .

引理 *4.2*

设 $M_{f,\ f} \geq M_{g,\ f\circ}$ 注意到

$$\begin{aligned}
&D_\phi^{(k)}(\epsilon g, f) - D_\phi^{(k)}(\epsilon g, g) \\
&= \sum \int \Big[\phi((\epsilon g)^k) - \phi(f^k) - \Big\{(\epsilon g)^k - f^k\Big\}\phi'(f^k) - \phi((\epsilon g)^k) + \phi(g^k) + \Big\{(\epsilon g)^k - g^k\Big\}\phi'(g^k)\Big] \\
&= -\sum \int \phi(f^k) - \epsilon^k M_{g,f} + M_{f,f} + \sum \int \phi(g^k) + (\epsilon^k - 1)M_{g,g} \\
&\geq \sum \int [\phi(g^k) - \phi(f^k)] - \epsilon^k M_{f,f} + M_{f,f} + (\epsilon^k - 1)M_{g,g} \\
&= \sum \int [\phi(g^k) - \phi(f^k)] - (\epsilon^k - 1)M_{f,f} + (\epsilon^k - 1)M_{g,g} \\
&= \sum \int [\phi(g^k) - \phi(f^k)] - (\epsilon^k - 1)(M_{f,f} - M_{g,g}) \geq 0
\end{aligned} \tag{7.140}$$

when

$$\epsilon^k \leq 1 + \frac{\sum \int [\phi(g^k) - \phi(f^k)]}{(M_{f,f} - M_{g,g})} \text{ for } g \neq f. \tag{7.141}$$

当 $M_{f、\ f} \leq M_{g、\ f}$ 时，反向不等式可用类似方法证明。至此，证明完成。

定理 *4.5*

引理 4.2 结合假设 (BP7) 可得

$$D_\phi^{(k)}(\epsilon k_m, f_X^m f_Y^m) \geq D_\phi^{(k)}(\epsilon k_m, k_m) \text{ for } \epsilon \leq \left[1 + \frac{\sum\int \left[\phi[(k_{xy,m})^k\right] - \phi\left[(f_x^m f_y^m)^k\right]}{(M_{f^m,f^m} - M_{k_m,k_m})}\right]^{1/k} \leq \epsilon_1. \tag{7.142}$$

同样，我们也发现

$$D_\phi^{(k)}((1-\epsilon)f_{X,Y}, f_{X,Y}) \geq D_\phi^{(k)}((1-\epsilon)f_{X,Y}, f_X f_Y) \text{ for } \epsilon \leq 1 - \left[1 + \frac{\sum\int[\phi(f_{xy}^k) - \phi(f_x^k f_y^k)]}{M_{f_X f_Y, f_X f_Y} - M_{f_{X,Y}, f_{X,Y}}}\right]^{1/k} = \epsilon_2. \tag{7.143}$$

结合这些结果，我们进一步得到，对于足够大的 m,

$$\begin{aligned} D_\phi^{(k)}(\epsilon k_m, f_X^m f_Y^m) &\geq D_\phi^{(k)}(\epsilon k_m, k_m) \left[\text{ by (7.142) }\right] \\ &> D_\phi^{(k)}(\epsilon k_m, 0) + D_\phi^{(k)}((1-\epsilon)f_{X,Y}, f_{X,Y}) - D_\phi^{(k)}((1-\epsilon)f_{X,Y}, 0) \\ &\left[\text{ by Assumption (BP8) and then (7.143) }\right] \\ &\geq D_\phi^{(k)}(\epsilon k_m, 0) + D_\phi^{(k)}((1-\epsilon)f_{X,Y}, f_X f_Y) - D_\phi^{(k)}((1-\epsilon)f_{X,Y}, 0) \end{aligned} \tag{7.144}$$

当 $\epsilon \leq \min\{\epsilon_1, \epsilon_{2,1\,2}\}$时。从（7.144）也可以清楚看出，在这样的 ϵ 选择下，假设 (BP6) 也始终成立。因此，假设 (BP7) 和 (BP8) 共同构成假设 (BP6)的充分条件，故定理 4.4适用。因此，渐近崩溃点至少为 $\min\{\epsilon_1, \epsilon_{2,1\,2}\}$。

参考文献

[1] Huber, P.J.稳健统计，Springer，2011年。[2] Hampel, F.R.、Ronchetti, E.M.、Rousseeuw, P.和Stahel, W.A. 稳健统计：基于影响函数的方法，Wiley-Interscience；纽约，1986年。[3] Beran, R. 参数模型的最小*Hellinger*距离估计，《统计学年刊》，445–463，1977年。[4] Beran, R.稳健位置估计，《统计学年刊》，431–444，1977年。[5] Cressie, N.和Read, T.R.C. 多项拟合优度检验，《皇家统计学会期刊：B辑（方法论）》，46(3)，440–464，1984年。[6] Csisz´ar,I.一个信息论不等式及其在证明马尔可夫链遍历性中的应用，《匈牙利科学院数学研究所公报》，8，85–108，1963年。[7] Ali, S.M.和Silvey, S.D. 一个分布相对于另一个分布的散度系数的一般类，《皇家统计学会期刊：B辑（方法论）》，1，131–142，1966年。[8] Bregman, L.M.寻找凸集公共点的松弛法及其在求解凸规划问题中的应用，《苏联计算数学与数学物理》，7(3)，200–217，1967年。[9] Basu, A.、Harris, I.R.、Hjort, N.L.和Jones, M.C. 通过最小化密度幂散度进行稳健且高效的估计，《Biometrika》，85(3)，549–559，1998年。

[10] Mukherjee, T.、Mandal, A.和Basu, A. B-指数散度及其推广与参数估计中的应用，《统计方法与应用》，28(2)，241–257，2019年。[11] Ghosh, A.、Harris, I.R.、Maji, A.、Basu, A.和Pardo, L. 统计推断中的广义散度，《伯努利分布》，23(4A)，2746–2783，2017年。[12] Basak, S.和Basu, A. 扩展布雷格曼散度与参数估计，《统计学》，30(4)，1–20，2022年。[13] Pardo, L. 基于散度度量的统计推断，CRC出版社，2018年。[14] Basu, A.、Shioya, H.和Park, C. 统计推断：最小距离方法，CRC出版社，2011年。[15] Owen, A.B. 经验似然，Chapman and Hall/CRC出版社，180–202，2001年。[16] Morales, D.、Pardo, L.和Pardo, M.C. 用于检验多个总体假设的似然散度统计量，《统计学通信：模拟与计算》，30(4)，867–884，2001年。[17] Balakrishnan, N.、Mart´ín, N.和Pardo, L. 用于检验简单和复合原假设的经验phi散度检验统计量，《统计学》，49(5)，951–977，2015年。[18] Balakrishnan, N.、Mart´ín, N.和Pardo, L. 用于检验两个总体均值差的经验phi散度检验统计量，《AStA统计分析进展》，101(2)，199–226，2017年。[19] Micheas, A.C.和Zografos, K. 使用 ϕ-散度度量随机依赖，《多元分析杂志》，97(3)，765–784，2006年。[20] Jing, B.Y. 双样本经验似然方法，《统计学与概率通讯》，24(4)，315–319，1995年。[21] Yong Song, Q.和Zhao, L.C. 各种两个总体差异的经验似然比置信区间，《系统科学与复杂性杂志》，13(1)，23，2000年。[22] Liu, Y.、Zou, C.和Zhang, R. 双样本均值问题的经验似然，《统计学与概率通讯》，75(5)，548–556，2008年。[23] Wu, C.和Yan, Y. 两样本问题的经验似然推断，《统计学及其接口》，5(3)，345–354，2012年。[24] Guha, A.和Chothia, T. 基于互信息的两样本检验，《加尔各答统计学会公报》，66(1-2)，39–54，2014年。[25] Guha, A.、Biswas, A.和Ghosh, A. 使用基于一般 φ-散度的互信息的非参数双样本检验，《荷兰统计学》，75(2)，180–202，2021年。[26] Fernandes, M.和N´eri, B. 随机过程之间独立性的非参数熵检验，《计量经济学评论》，29(3)，276–306，2009年。[27] Parzen, E. 概率密度函数及众数的估计，《数学统计年刊》，33(3)，1065–1076，1962年。[28] Wied, D.和Weißbach, R. 核密度估计量相合性综述，《统计学论文集》，53(1)，1–21，2012年。[29] Van der Vaart, A.W. 渐近统计，剑桥大学出版社，2000年。[30] Silverman, B.W. 统计学与数据分析中的密度估计，CRC出版社，2018年。[31] Hall, P.和Heyde, C.C. 鞅极限定理及其应用，学术出版社，2014年。[32] Vajda, I. 关于f-散度与概率测度奇异性，《匈牙利数学学报》，2(1–4)，223–234，1972年。[33] Park, C.和Basu, A. 最小差异估计：渐近正态性与崩溃点结果，统计科学研究协会，53(1)，19–33，2004年[34] Roy, S.、Sarkar, A.、Ghosh, A.和Basu, A. 一般类最小散度估计量的渐近崩溃点分析，《伯努利分布》，32(1)，698–722，2026年