arXiv:2602.08486v1 [math.ST] 9 Feb 2026

AMP 框架中基于 Lasso统计量的经验贝叶斯变量选择

LinaHidmi Asaf Weinstein耶路撒冷
希伯来大学*

摘要

Lasso 是高维线性回归中最常用的变量选择方法之一，并已在不同设定下得到广泛研究。在一种包含 $n/p \to$常数、独立同分布高斯分布 X 矩阵和线性稀疏性的特定渐近设定中，Su等人 [2017] 分析了 Lasso选择路径，并给出了负面结果，表明若要将错误发现比例维持在较低水平，就必须付出检验功效大幅下降的代价。随后，Wang等人 [2020] 沿用同一框架，研究阈值化 Lasso 选择中第一类错误与检验功效之间的权衡。该方法依据 Lasso估计量的大小而非其在 Lasso路径上的出现顺序对变量排序。研究表明，若恰当选择正则化参数，便有可能取得显著改进。我们将这项研究进一步推进，旨在 AMP 框架下，在依据固定正则化项下 Lasso估计量的某个单变量函数对变量排序的程序中，寻找最优选择程序 λ。我们注意到，Lasso估计量的模型在渐近意义下实际上可化简为一种广为研究的双组模型，因此提出一种基于局部错误发现率估计值的经验贝叶斯变量选择程序。我们将 AMP 框架中的现有结果加以拓展，从而精确预测描述该程序中第一类错误与检验功效渐近权衡关系的曲线。此外，我们证明最优 λ 值是渐近均方误差的最小化点，因此建议使用通过 λ交叉验证估计所得参数的经验贝叶斯程序。理论预测表明，检验功效可以大幅提升；我们通过不同设定下的数值研究证实了这一点。

1 引言

考虑高斯线性回归模型，

$$Y = \boldsymbol{X}\beta + \zeta, \qquad \zeta \sim \mathcal{N}_n(0, \sigma^2 \boldsymbol{I}), \tag{1}$$

其中，$\boldsymbol{X} \in \mathbb{R}^{n\times p}$ 是针对 n 个受试对象中每个对象的 p 个预测变量的观测测量矩阵，系数向量 $\beta = (\beta_1, \beta_2, \ldots, \beta_p)^\top$ 和噪声水平 σ^2均未知。在基因组学、图像分析和文本挖掘等现代应用中，[Varoquaux 等，2012，Yamashita 等，2008，Gramfort 等，2011，Beer 等，2018] 可用预测变量的数量 p 通常很大，有时甚至超过样本量 n；但人们往往先验地假设其中大多数变量都无关紧要，即相应的 β_i 基本为零。在这些稀疏设定中，一项主要任务是进行

*电子邮箱：lina.hidmi@mail.huji.ac.il, asaf.weinstein@mail.huji.ac.il

变量选择，即估计真正影响响应变量（条件）分布的预测变量子集。

在噪声模型（1）中，对于有限的 n 和p，我们实际上并不期望任何变量选择程序都能完美恢复相关变量子集，但可以根据程序选择真正相关变量、同时避免选择过多无关变量的能力来评估其质量。Candes 等人提出的受控变量选择问题 [，2018] 将变量选择正式视为一个多重检验问题，具体而言，就是同时检验以下 p 原假设

$$H_i : \beta_i = 0, \qquad i \in \mathcal{H} \equiv \{1, 2, \ldots, p\},$$

因此，选择一个变量就相当于拒绝相应的原假设。令 $\mathcal{H}_0 \equiv \{i : \beta_i = 0\}$表示真实零假设的（未知）子集， $\mathcal{S} \equiv \mathcal{H} \setminus \mathcal{H}_0$ 表示真实非零假设的子集。对于任何输出 $\mathcal{S}$ 估计值 $\hat{}$ $\mathcal{S} \subseteq \{1, 2, \ldots, p\}$ 的变量选择程序，定义错误发现比例（FDP）为拒绝项中真实零假设所占的比例，定义真正例比例（TPP）为真实非零假设中被拒绝项所占的比例，

$$\mathrm{FDP} = \frac{|\hat{\mathcal{S}} \cap \mathcal{H}_0|}{|\hat{\mathcal{S}}|}, \qquad \mathrm{TPP} = \frac{|\hat{\mathcal{S}} \cap \mathcal{S}|}{|\mathcal{S}|},$$

其中 $|\mathcal{A}|$ 是集合 $\mathcal{A}$的基数，并且我们采用约定 $0/0 \equiv 0$。根据（1），对这两个随机分数取期望，得到错误发现率（FDR）和真阳性率（TPR），

$$\mathrm{FDR} = \mathbb{E}[\mathrm{FDP}], \qquad \mathrm{TPR} = \mathbb{E}[\mathrm{TPP}],$$

它们分别是衡量程序第一类错误和检验功效的特定指标。有效的选择程序能够在保持错误发现率较低的同时实现较高的真阳性率。与将第一类错误和第二类错误区别对待的经典多重检验观点一致，在受控变量选择框架中，我们称一个程序在$q \in [0, 1]$ 水平上有效，如果对于参数 β的所有取值，它的错误发现率均不超过 $\leq q$；通常，任务是在满足有效性的前提下，以某种适用的方式最大化真阳性率。

许多变量选择程序都可以表示为按如下方式运行。首先，对每个变量 $i = 1$，……， p，计算一个重要性统计量 T_i，其值越大，通常表示反对 H_i的证据越强。重要性统计量的选择由此确定了该程序的选择路径，这是一个集合值映射，它将每个固定阈值$t \in \mathbb{R}$ 与相应子集相关联

$$\hat{\mathcal{S}}(t) \ := \ \{i \leq p : \ T_i > t\}. \tag{2}$$

我们指出，尽管默认采用上述约定（除非另有说明），但针对统计量 T_i较大或较小取值进行选择的程序之间并无本质区别；因此，从现在起，每当提及形式为 (2) 的程序时，我们指的是对较小 T_i 或较大 T_i取值进行拒绝的程序，并会根据需要在两种约定之间切换。当阈值 t 在 (2) 中从大到小变化时，选择路径会形成一条由候选变量子集构成的递增轨迹。确定路径后，程序需要通过指定数据依赖阈值 $\hat{}t$来决定何时停止纳入更多变量，从而得到最终选定的子集 $\hat{}_{\mathcal{S}} = \hat{}\mathcal{S}(\hat{}t) \ = \{i \leq p : \ T_i \ > \hat{}t\}.$。在这种选择程序的视角下，$\hat{}t$ 的设定负责校准程序的第一类错误，而 T_i 的选择则负责变量的排序，因此控制着程序的检验功效。现有关于受控变量选择的文献大多聚焦于校准，这是一个重要且具有挑战性的问题，尤其是在要求具备上述一般性的情况下。事实上，直到近些年，诸如 Knockoff 方法 [Barber 和 Cand`es（2015）这样的程序才开始出现在文献中，它们可以围绕任意（自由选择的）重要性统计量构建，以实现有限样本错误发现率控制，

Candes 等人， 2018], 此类程序围绕任意（自由选择的）重要性统计量构建，以实现有限样本错误发现率控制，近年来已见诸文献。

不过，假设我们有一种方法，能够针对重要性统计量的任意选择，选取 $\hat{t}$ ，使得 FDR $\leq q$ ，那么整个过程——尤其是在任意目标 FDR 水平 q 下的检验功效——只取决于统计量 T_i 的选择。这促使我们撇开校准问题，研究选择路径 $\{\mathcal{S}(t) \quad t \in \mathbb{R}\}$ 本身。更具体地说，如果我们定义

$$\mathrm{FDP}(t) \equiv \frac{|\hat{\mathcal{S}}(t) \cap \mathcal{H}_0|}{|\hat{\mathcal{S}}(t)|}, \qquad \mathrm{TPP}(t) \equiv \frac{|\hat{\mathcal{S}}(t) \cap \mathcal{S}|}{|\mathcal{S}|}, \tag{3}$$

那么我们可以考虑参数曲线 $t \mapsto$ (𝔼[真阳性比例(t)], 𝔼[假发现比例(t)])，将其视为该过程的受试者工作特征（ROC）曲线（尽管在 ROC 的标准定义中，假发现比例的分母会替换为 $|\mathcal{H}_0|$）。由于这条参数曲线完全由重要性统计量 T_i 的选择决定，因此它实际上体现了与统计量 T_i 相关的第一类错误率与检验功效的权衡。一个自然的问题是，是否存在某种意义上的最优重要性统计量选择。

要回答这个问题，对于固定的 n和 p，也就是说，要在问题规模固定时比较不同的重要性统计量选择，需要计算给定 T_i时，在（1）下 FDP(t) 和 TPP(t) 的期望；即使 T_i超出普通最小二乘法（而且普通最小二乘法只适用于 $p < n$）的相对简单选择，这也难以实现。例如，在常见的T_i 是基于 Lasso 估计量的统计量 [Tibshirani, 1996],这一情形下，计算就已经变得不可行。

$$\hat{\beta}(\lambda) = \underset{b \in \mathbb{R}^p}{\operatorname{argmin}} \frac{1}{2}\|Y - \boldsymbol{X}b\|_2^2 + \lambda\|b\|_1, \tag{4}$$

^ ^ 因为对于固定的 n和 p， $\beta = \beta(\lambda)$ 的分布难以处理。尽管如此，在特定的渐近设定下，近似消息传递（AMP）理论的最新进展 [Bayati 和 Montanari, 2011, 2012]使得我们能够获得以下形式统计量的 FDP(t) 和 TPP(t) 的精确极限

$$T_i = \varphi(\hat{\beta}_i(\lambda)), \tag{5}$$

其中 φ 是一个（性质足够良好的）单变量函数。这些重大进展在后续研究中得到应用和拓展，用于分析（渐近）权衡曲线，

$$t \mapsto \big(\lim_{p\to\infty} \mathrm{TPP}(t), \lim_{p\to\infty} \mathrm{FDP}(t) \big),$$

针对不同的基于 Lasso 的变量重要性统计量。具体而言，Bogdan 等人 [2013a]首次为 Lasso选择推导出 FDP 和 TPP 的渐近预测；也就是说，该^程序会选取 (4) 中的活跃坐标 β ，其选择路径则通过改变 λ形成。Su等人 [2017] 随后强化了这些结果，并推导出一个普适的权衡图，适用于任意 Π 和 σ^2；这意味着 Lasso选择不可能同时实现较高的 TPP 和较低的 FDP。Wang等人 [2020] 采用相同的渐近框架，研究了基于阈值化 ℓ_γ正则化估计量大小的选择所对应的权衡曲线；当 $\gamma = 1$ 取特定值时，这种方法即为阈值化 Lasso 选择，也就是说，该^程序在变量 i超过某个阈值时将其选中。与 Lasso选择不同，在这里 λ被视为固定值，选择路径由阈值参数化。在他们的贡献中，Wang等人 [2020] 根据权衡曲线刻画了 λ 的最优选择。总体而言，他们的分析表明，与 Lasso选择相比，通过使用阈值化 Lasso 并适当选择 λ，可以取得显著改进。

上述研究建立在对 Lasso 估计经验分布的渐近简化之上，该简化由 Bayati 和 Montanari [2012], 提出，用于分析

两种广泛使用的程序所对应的权衡曲线，即 Lasso选择和阈值化 Lasso 选择。在本研究中，我们采取相反的视角，利用模型的渐近简化来指导选择程序——或者等价地，选择重要性统计量 T_i。核心思想是，在 AMP 框架下，识别 Lasso 估计量（4）渐近地有效简化为的简单模型，并将其视为一种双组模型 [Efron 等人，2001，Efron，2008]。借鉴双组模型下多重检验的丰富文献 [Sun和Cai，2007，Xie 等人，2011，Heller和Rosset，2021], 我们建议使用相应的局部错误发现率（lfdr）对系数估计进行排序，因此 lfdr 越小，反对零假设的证据越强。就选择路径而言，这等价于对较小的

$$T_i^* = \frac{f_0(\hat{\beta}_i)}{f(\hat{\beta}_i)} \tag{6}$$

其中 f_0 表示零假设密度，而 f 表示相应双组模型下的边际（混合）密度。然而，T_i^*取决于未知参数，因此这并非合法选择，而是一个预言机统计量选择，我们用上标星号表示。事实上，与标准双组模型不同，标准模型中 f 未知，但 f_0 假定已知；而在我们的情形中，f_0 也仅在一组参数已知的条件下才是已知的。采用经验贝叶斯方法，我们通过将 f 和 f_0 的估计值代入（6），力求模拟上述预言机程序。对于混合密度 f，我们使用核密度估计量；而零假设密度 f_0基于现有的渐近理论具有已知的参数形式，因此采用参数化方法进行估计。

从理论角度来看，要为所提出的经验贝叶斯程序获得渐近预测，比分析 Lasso 或阈值化 Lasso 困难得多，因为 f 和 f_0 依赖于所有的 β_1（λ）.....、β_p（λ），这使我们无法直接将 Bayati 和 Montanari [2012] 的基本结果应用于 T_i^* 的经验贝叶斯估计。尽管如此，直观上我们预期，若 f 和 f_0 的估计量一致——这在我们的方法中要求包括核密度估计量一致在内的若干条件——那么经验贝叶斯程序的极限假发现比例和真阳性比例将与预言机程序的相应指标一致。本文的主要理论结果表明，在适当条件下确实如此；我们的分析实质上将固定 t 下 lfdr(t) 的极限，与对 p 个系数估计值上的 lfdr(β_j) 平均值的极限区分开来，其中后者是真实的 lfdr 在系数估计值处的取值。在适当假设下，对于任意固定的 λ，我们证明，我们的经验贝叶斯程序渐近地达到（5）中所有函数 φ 的最优权衡曲线；这是渐近实例最优性的一种强意义，这一概念在经验贝叶斯文献中广为人知，并可追溯至 Robbins的原始思想 [Robbins，1951]。

与阈值化 Lasso 类似，λ 的选择会对所提 EB 程序的权衡曲线产生显著影响。对于阈值化 Lasso（更一般地说，对于阈值化桥回归选择），Wang等人 [2020]证明，使渐近均方误差（MSE）最小化的 λ取值在 FDP-TPP 权衡方面也渐近最优。在定理 4.10 中，我们为所提 EB 程序得到了类似结果，证明渐近 MSE 的 λ 最小化点再次给出最优权衡曲线；尽管这一结论与Wang等人 [2020] 关于阈值化桥回归选择的结论一致，但我们的证明需要采用不同的技巧。这一结果促使我们使用交叉验证来估计最优的 λ，该参数取决于未知参数。根据 Weinstein 等人 [2023] 的结果，我们推导出 AMP 框架下 K折交叉验证估计量的极限。总而言之，我们的结果表明，除交叉验证带来的轻微偏差外，使用所提 λ 估计量的 EB 选择程序，在通过允许不仅对5也对 φ 进行搜索而从式（ λ）得到的更大一类程序中，将具有最优的渐近权衡曲线。

To 展示所提出方法的优势，图1 给出了一个 n

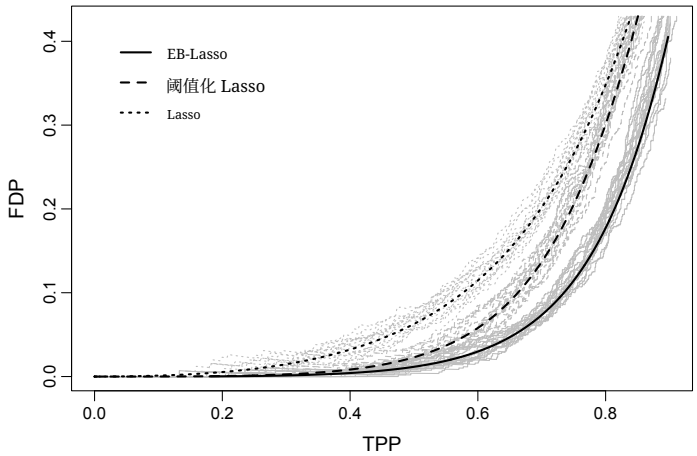

图 1： 在一个包含 $p = 10^4$、$n = 2p$ 和 $\beta_i \sim 0.9\delta_0 + 0.1\mathcal{N}(3.5,\ 1)$ 的示例中，所提出的 *EB* 方法（图例中标为 *"EB-Lasso"*）与 *Lasso* 及阈值化 *Lasso* 选择的权衡曲线。黑色曲线表示理论渐近预测，细灰线表示相应的模拟结果。

示例，其中 $p = 10^4$，$n = 2\cdot 10^4$，$\sigma = 1$，且 β_i 独立同分布地取自如下混合分布：以概率 $1-\epsilon = 0.9$ 取零点处点质量，以概率 $\mathcal{N}(3.5,\ 1)$取正态分布 ϵ =0.1。该图展示了所提出的经验贝叶斯程序对应的渐近FDP–TPP权衡曲线（由我们的理论预测），以及已有研究对 Lasso 选择和阈值化 Lasso 选择的渐近预测；其中，$\lambda = 1$ 用于计算经验贝叶斯程序和阈值化 Lasso 的 Lasso 估计量（4）。对于这三种方法，我们叠加了来自 17 次独立模拟运行的经验（实际观测）权衡曲线，结果显示它们与理论吻合良好。该图表明，检验功效的提升可能相当显著；后文还将给出采用不同信号分布和抽样比的更多示例。

本文其余部分安排如下。第 2 节回顾 AMP 框架及一些相关的已有结果。第 3节介绍一个预言机选择程序，以及旨在渐近地与之竞争的经验贝叶斯程序。第 4节给出对预言机程序和经验贝叶斯程序的理论分析，其中还包含关于 λ 最优选择的结果。第 5节报告数值研究结果，并评估其与理论的一致性。第 6 节采用近期研究中的一种观点，强调 lfdr 在定义多重检验问题的第一类错误率方面的作用；文中展示了如何利用 AMP 框架中的 lfdr 预测，以更具探索性的方式选择变量，同时保持对 lfdr（而非错误发现率）的渐近控制。第7 节最后讨论本文未涉及、值得进一步研究的一些问题。

2 AMP 框架及现有结果回顾

我们采用 Bayati 和 Montanari[2012], 所使用的近似消息传递（AMP）框架，该框架也被 Su等人 [2017]和Wang等人 [2020]采用。具体而言，在线性模型（1）的假设下，我们考虑高维渐近设定，其中 $p,\ n \to \infty$ 满足 $n/p \to \delta > 0$，且 $n \times p$ 矩阵 $\boldsymbol{X}$ 的元素是来自$\mathcal{N}(0, 1/n)$ 分布的独立同分布实现。此外，假设各个系数 β_i 是随机变量 Π 的独立同分布副本，该随机变量具有有界二阶矩，且在零点处具有点质量，

$$\beta_i \overset{iid}{\sim} \Pi = (1-\epsilon)\delta_0 + \epsilon\Pi_1, \qquad i = 1, ..., p, \tag{7}$$

独立于 $\boldsymbol{X}$ 和 ζ，其中 $\mathbb{P}(\Pi_1 = 0) = 0$。常数 $\epsilon \in (0,1)$ 以及 Π_1 的分布均假定未知。下文中，我们用 Π和 Π_1 表示随机

变量或其分布，只要不会引起混淆即可。我们指出，只要固定向量 $\beta^{(n)} = (\beta_1^{(n)}, ..., \beta_p^{(n)})$ 的各分量的经验分布 $\Pi^{(n)}$ 弱收敛到某个 Π具有有界二阶矩的、形式如7的分布，则以下结果同样适用于频率学派框架。

在这一框架下，Bayati 和 Montanari [2012]提出了以下结果，该结果刻画了 Lasso 估计的经验分布的渐近行为。这一结果是我们工作的核心，既为后文提出的新变量选择程序提供了动机，也构成了理论分析中的关键工具。

定理 (Bayati 和 Montanari [2012] 的定理 1.5)。 在上述渐近设定ˆ ˆ下，令 $\beta = \beta(\lambda)$ 表示 *Lasso* 估计量(4)。令 $\psi : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ 为一个伪 *Lipschitz* 函数。则

$$\frac{1}{p}\sum_{i=1}^{p} \psi(\hat{\beta}_i, \beta_i) \xrightarrow{a.s.} \mathbb{E}\left[\psi\left(\eta_{\alpha\tau}(\Pi + \tau Z), \Pi\right)\right]$$

当 $p \to \infty$ 时， $Z \sim \mathcal{N}(0,\ 1)$ 与 Π 独立，并且参数 $\tau > 0$ 和 $\alpha > \alpha_{\min}$ 是下列方程组的唯一解:

$$\tau^2 = \sigma^2 + \frac{1}{\delta}\mathbb{E}\left[\left(\eta_{\alpha\tau}(\Pi + \tau Z) - \Pi\right)^2\right] \tag{8}$$

$$\lambda = \alpha\tau\left(1 - \frac{1}{\delta}\mathbb{P}(|\Pi + \tau Z| > \alpha\tau)\right). \tag{9}$$

这里，函数 $\eta_t(x) = \text{sign}(x) \cdot (|x| - t)+$ 是软阈值算子，而 $\alpha_{\min}$ 是下列方程的唯一非负解

$$(\alpha^2 + 1)\Phi(-\alpha) - \alpha\phi(\alpha) = \frac{\delta}{2}, \tag{10}$$

其中， Φ 和 ϕ 分别表示标准正态分布的累积分布函数和概率密度函数。

上述定理基本上表明，就在对这些对 (β_i, β_i) 的（随机）经验分布下的充分良态函数计算期望而言，我们可以渐近地将其视为

$$\left(\hat{\beta}_i, \beta_i\right) \sim \left(\eta_{\alpha\tau}(\Pi + \tau Z),\ \Pi\right), \qquad \text{independent for } i = 1, ..., p. \tag{11}$$

在 Bogdan 等人 [2013的] 这一BayatiandMontanari的基本结果 [2012] 中，借助一些必要的技术扩展，他们获得了路径由下式给出的过程的假发现比例和真阳性比例的精确极限

$$\hat{\mathcal{S}}^L(\lambda) = \{i : \hat{\beta}_i(\lambda) \neq 0\}, \qquad \lambda > 0. \tag{12}$$

形式化地，令 $\text{FDP}(\lambda) := |\hat{\mathcal{S}}^L(\lambda) \cap \mathcal{H}_0|/|\hat{\mathcal{S}}^L(\lambda)|$，以及 $\text{TPP}(\lambda) := |\hat{\mathcal{S}}^L(\lambda) \cap \mathcal{S}|/|\mathcal{S}|$，他们证明，在上述 AMP 框架下，

$$\begin{aligned} \text{FDP}(\lambda) &\xrightarrow{P} \frac{2(1-\epsilon)\Phi(-\alpha)}{2(1-\epsilon)\Phi(-\alpha) + \epsilon\mathbb{P}(|\Pi_1 + \tau Z| > \alpha\tau)} =: \text{fdp}^{\text{L}}(\lambda) \\ \text{TPP}(\lambda) &\xrightarrow{P} \mathbb{P}(|\Pi_1 + \tau Z| > \alpha\tau) =: \text{tpp}^{\text{L}}(\lambda), \end{aligned} \tag{13}$$

为 $p \to \infty$，其中（α，τ）满足方程（8）和（9）。Su等人 [2017] 随后对此进行了扩展，证明（13）中的收敛关于 λ是一致的。需要指出的是，严格来说，过程（12）并不属于（2）的形式；不过，正如 Weinstein 等人 [2017], 所指出的

与使用进入 Lasso路径的时刻 $T_i^{\mathrm{LM}} = \sup\{\lambda > 0 : \hat{}\ \beta_i(\lambda) \neq 0\}$ 作为重要性统计量的（2）形式的过程密切相关。实际上，只有当变量从 Lasso路径中退出时，这两种过程才会产生不同的选择路径；我们认为这种情况相当少见。因此，出于“实际”考虑，我们将（12）视为与该路径等价

$$\hat{\mathcal{S}}^{\mathrm{LM}}(t) = \{i \leq p : T_i^{\mathrm{LM}} \geq t\}, \qquad t > 0, \tag{14}$$

其中，上标表示“Lasso-max”，我们将14称为对应于$Lasso$ 选择的路径。

在Su 等人 [2017]中，研究者利用渐近预测（13）展示了 Lasso 选择的一个缺陷，表明在 AMP设定下，无论 Π_1 和 ϵ如何，高真阳性比例都必然要以假发现比例膨胀为代价。在 AMP 框架下的后续研究中，[Wang 等人、2020、Weinstein 等人， 2023] 表明，只选择幅度 $|\beta_i$（λ）$|$ 足够大的 Lasso估计量，即采用阈值化 *Lasso* 选择，可以显著缓解这一问题。对于固定的 $\lambda > 0$，相应的选择路径为

$$\hat{\mathcal{S}}_\lambda^{\mathrm{TL}}(t) = \{i : |\hat{\beta}_i(\lambda)| > t\}, \qquad t > 0. \tag{15}$$

定义 FDP(t； λ)： $= |\hat{}\mathcal{S}_\lambda^{\mathrm{TL}}$（$t$） $\cap \mathcal{H}_0|/|\hat{}\mathcal{S}_\lambda^{\mathrm{TL}}$（$t$）$|$，TPP($t$； λ)： $= |\hat{}\mathcal{S}_\lambda^{\mathrm{TL}}$（$t$） $\cap \mathcal{S}|/|\mathcal{S}|$，Wang 等人[2020] 证明了

$$\begin{aligned} \mathrm{FDP}(t;\lambda) &\xrightarrow{P} \frac{2(1-\epsilon)\Phi\left(-\alpha - \frac{t}{\tau}\right)}{2(1-\epsilon)\Phi\left(-\alpha - \frac{t}{\tau}\right) + \epsilon\,\mathbb{P}\left(|\Pi_1 + \tau Z| \geq \alpha\tau + t\right)} \ =: \ \mathrm{fdp}^{\mathrm{TL}}(t;\lambda) \\ \mathrm{TPP}(t;\lambda) &\xrightarrow{P} \mathbb{P}\left(|\Pi_1 + \tau Z| \geq \alpha\tau + t\right) \ =: \ \mathrm{tpp}^{\mathrm{TL}}(t;\lambda). \end{aligned} \tag{16}$$

渐近预测（16）被用于 Wang等人 [2020] 和Weinstein 等人 [2023]，以证明如果 λ 选择得当，这种策略——通过选择较小的 λ值，可以沿着 Lasso路径继续向下探索——与 Lasso选择相比，能够显著改善信号与噪声之间的区分。形式上，Wang等人 [2020, 定理 32] 证明了以下结论。令 λ^* 为渐近均方误差的最小化点 $\mathbb{E}$（$\eta_{\alpha\tau}$（$\Pi + \tau Z$） $- \Pi$）^{2}。那么，首先， λ^* 对于阈值化 Lasso 而言是最优的，因为对于任意固定的渐近真阳性比例水平，它都能使渐近错误发现比例最小。其次，他们证明，对于 Lasso 可实现的任意（渐近）真阳性比例水平，都有

$$\mathrm{tpp}^{\mathrm{TL}}(t;\lambda^*) = \mathrm{tpp}^{\mathrm{L}}(\lambda) \quad \Rightarrow \quad \mathrm{fdp}^{\mathrm{TL}}(t;\lambda^*) \leq \mathrm{fdp}^{\mathrm{L}}(\lambda). \tag{17}$$

换言之，如果选择 λ 作为 Lasso 的参数，并选择 $t = t(\lambda^*)$ 作为阈值化 Lasso 的参数，使得两种方法具有相同的 TPP，那么阈值化 Lasso 的 FDP 会低于 Lasso。

3 一种经验贝叶斯变量选择程序

回到上文对定理 1.5 所作的解释，Bayati 和 Montanari[2012], ，注意到（11）蕴含一个两组模型，

$$\begin{aligned} \mathbb{P}(\beta_i = 0) &= 1 - \epsilon, \qquad \hat{\beta}_i|\beta_i = 0 \ \sim \ \eta_{\alpha\tau}(\tau Z) \\ \mathbb{P}(\beta_i \neq 0) &= \epsilon, \qquad \hat{\beta}_i|\beta_i \neq 0 \ \sim \ \eta_{\alpha\tau}(\Pi_1 + \tau Z). \end{aligned} \tag{18}$$

这一观察促使我们借鉴双组模型下多重检验文献中的现有理论，并提出一种经验贝叶斯方法，将局部错误发现率的估计值用作检验统计量。因此，与通过对 Lasso 系数的绝对值进行阈值化来选择变量的阈值化 Lasso 相比，我们将提出一种经验贝叶斯程序，通过对估计的局部错误发现

ˆ率进行阈值化来选择 β_i （依据式（18））。本节中，我们在工作假设下构建经验贝叶斯程序：(β_i, β_i) 是来自双组模型（18）的独立同分布配对；需要说明的是，为了构建方法本身，假定式（18）中的模型为独立同分布完全没有问题，而下一节的分析当然只需依赖 Bayati 和 Montanari [2012] 结果中弱得多的性质。进一步阐述我们的动机：在真正的独立同分布设定下，经验贝叶斯选择程序有望在检验功效方面具有渐近最优性。更确切地说，可以证明，对真实局部错误发现率进行阈值化的预言机选择规则，能够在控制错误发现率的条件下渐近地最大化期望真阳性比例。事实上，这可由 Sun和Cai [2007] 关于边际错误发现率控制的非渐近结果，或 Heller和Rosset [2021] 关于错误发现率控制的结果推出 [；另请参见 Weinstein（2021），其给出了类似的非渐近结果，用于刻画β_i 为固定参数时的预言机规则], ，因为当阈值固定时，边际错误发现率与错误发现率渐近等价 [Benjamini, 2008]。反过来，经验贝叶斯规则旨在不使用问题参数的任何信息来模仿预言机程序。粗略地说，因此，在满足经验贝叶斯局部错误发现率估计量相合性的条件下，该方法有望渐近地达到预言机的性能。

在我们的工作假设下，记

$$f_0 = \text{density of } \eta_{\alpha\tau}(\tau Z), \qquad f_1 = \text{density of } \eta_{\alpha\tau}(\Pi_1 + \tau Z),$$

这里的“密度”应作广义理解，因为 f_0 和 f_1 在零点均有质量。与通常做法一样，我们将 f_0 称为零假设分布，将 f_1 称为非零假设ˆ分布。那么， β_i 的边际密度为以下混合分布

$$f = (1-\epsilon) f_0 + \epsilon f_1. \tag{19}$$

局部错误发现率 [（下文简称 lfdr）；Efron, 2004, 2008] 定义为在假定模型（18）下，给定 $\beta_i = x$时， β_i 属于零假设的后验ˆ概率，

$$\text{lfdr}(x) := \mathbb{P}(\Pi = 0 | \, \eta_{\alpha\tau}(\Pi + \tau Z) = x) = \frac{(1-\epsilon) f_0(x)}{f(x)}. \tag{20}$$

ˆ因此，lfdr(β_i)的值越小，表明反对零假设的证据越强， $\beta_i = 0$。注意，与标准双组模型一样， f_1 和 ϵ 均未知；但与通常的双组模型不同，此处 f_0 也未知，因为 α 和 τ 依赖于未知的 ϵ 和 Π_1 （以及 σ^2）。现在考虑预言机程序，该程序选择 lfdr（20）足够小的变量。它与拒绝密度比T_i^*（定义见（6））较小的变量的程序具有相同的选择路径；后者更便于我们处理，因为无需估计ˆϵ。我们始终规定，不选择满足 $\beta_i = 0$ 的变量；这一规定无关紧要，因为假设 ϵ 较小，则选择系数估计为零的规则通常会产生过高、因而不值得关注的假发现比例（FDP）。因此，预言机程序的选择路径为

$$\hat{\mathcal{S}}^*(t) = \left\{ i : \hat{\beta}_i \neq 0, \ \frac{f_0(\hat{\beta}_i)}{f(\hat{\beta}_i)} \leq t \right\}$$

相应的权衡曲线为

$$\text{FDP}^*(t) = \frac{|\hat{\mathcal{S}}^*(t) \cap \mathcal{H}_0|}{|\hat{\mathcal{S}}^*(t)|}, \qquad \text{TPP}^*(t) = \frac{|\hat{\mathcal{S}}^*(t) \cap \mathcal{S}|}{|\mathcal{S}|}.$$

为了用合法的选择规则（即不依赖问题未知参数的程序）来模拟预言机规则，我们提出从观测向量 f 中（非参数地）估计 ˆf_0 （参数化地）估计，并将它们代入（6），从而得到一个估计

$$T_i^{EB} = \frac{\hat{f}_0(\hat{\beta}_i)}{\hat{f}(\hat{\beta}_i)}$$

T_i^*。由此得到的经验贝叶斯 (EB) 程序的选择路径为

$$\hat{\mathcal{S}}^{EB}(t) = \left\{ i : \hat{\beta}_i \neq 0,\ T_i^{EB} \leq t \right\} = \left\{ i : \hat{\beta}_i \neq 0,\ \frac{\hat{f}_0(\hat{\beta}_i)}{\hat{f}(\hat{\beta}_i)} \leq t \right\},$$

其权衡曲线为

$$\mathrm{FDP}^{EB}(t) = \frac{|\hat{\mathcal{S}}^{EB}(t) \cap \mathcal{H}_0|}{|\hat{\mathcal{S}}^{EB}(t)|}, \qquad \mathrm{TPP}^{EB}(t) = \frac{|\hat{\mathcal{S}}^{EB}(t) \cap \mathcal{S}|}{|\mathcal{S}|}.$$

在描述我们为经验贝叶斯程序提出的具体实现之前，我们再给出一种表示$\hat{\mathcal{S}}^*$和$\hat{\mathcal{S}}^{EB}$的选择路径的方式，这对下一节的分析很有用。在双组模型（18）下，f_0 和 f_1 在零点处都有点质量，因此我们可以写成

$$f_0(\hat{\beta}_i) = (1 - w_0)\delta_0(\hat{\beta}_i) + w_0 g_0(\hat{\beta}_i), \tag{21}$$

其中

$$1 - w_0 = \mathbb{P}(|Z| \leq \alpha) = 2\Phi(\alpha) - 1, \qquad g_0 = \text{density of } \tau\eta_\alpha(Z) \big| |Z| > \alpha.$$

同理,

$$f_1(\hat{\beta}_i) = (1 - w_1)\delta_0(\hat{\beta}_i) + w_1 g_1(\hat{\beta}_i), \tag{22}$$

其中

$$1 - w_1 = \mathbb{P}(|\Pi_1 + \tau Z| \leq \alpha\tau), \qquad g_1 = \text{density of } \eta_{\alpha\tau}(\Pi_1 + \tau Z) \big| |\Pi_1 + \tau Z| > \alpha\tau.$$

ˆ 根据这些定义，式（ β_i 中的边际密度（19）可表示为

$$f(\hat{\beta}_i) = (1 - w)\delta_0(\hat{\beta}_i) + w g(\hat{\beta}_i), \tag{23}$$

其中

$$1 - w = (1 - w_0)(1 - \epsilon) + (1 - w_1)\epsilon$$

and

$$g = \frac{2(1-\epsilon)(1-\Phi(\alpha))}{w} g_0 + \left(1 - \frac{2(1-\epsilon)(1-\Phi(\alpha))}{w}\right) g_1.$$

ˆˆˆˆ 若进一步定义 $q(\beta_i) := wg(\beta_i)$ 和 $q_0(\beta_i) := w_0 g_0(\beta_i)$，则（21）和（23）分别变为

$$f_0(\hat{\beta}_i) = (1 - w_0)\delta_0(\hat{\beta}_i) + q_0(\hat{\beta}_i), \qquad f(\hat{\beta}_i) = (1 - w)\delta_0(\hat{\beta}_i) + q(\hat{\beta}_i), \tag{24}$$

分别。随后，我们建议将 f 估计为

$$\hat{f}(\hat{\beta}_i) = (1 - \hat{w})\delta_0(\hat{\beta}_i) + \hat{q}(\hat{\beta}_i), \tag{25}$$

with

$$\hat{w} = \frac{1}{p}\sum_{i=1}^{p} \mathbf{1}\{\hat{\beta}_i \neq 0\}, \qquad \hat{q}(x) = \frac{1}{ph}\sum_{i=1}^{p} \mathbf{1}\{\hat{\beta}_i \neq 0\}\, \mathcal{K}\left(\frac{x - \hat{\beta}_i}{h}\right), \tag{26}$$

全文中，我们均使用标准高斯密度 ϕ 作为核，即 $\mathcal{K} = \phi$。

为了估计 f_0，我们需要估计参数 $\alpha\tau$ 和 τ，因为 q_0 涉及密度 g_0的 $\eta_{\alpha\tau}(\tau Z) = \tau\eta_\alpha(Z)$ 条件分布于 $|Z| > \alpha$。在 Mousavi 等人 [2018] 中证明了，对于任意固定的 $\lambda > 0$,

$$\hat{\tau}^2 = \frac{\|Y - \boldsymbol{X}\hat{\beta}\|_2^2}{n(1 - \|\hat{\beta}\|_0/n)^2}, \qquad \widehat{\alpha\tau} = \frac{\lambda}{1 - \|\hat{\beta}\|_0/n}, \tag{27}$$

分别是 τ^2和 $\alpha\tau$的一致估计量。这自然导出了代入估计量

$$\hat{g}_0(x) = I(x > 0)\frac{1}{2\hat{\tau}(1-\Phi(\hat{\alpha}))}\phi\left(\frac{x+\widehat{\alpha\tau}}{\hat{\tau}}\right) + I(x < 0)\ \frac{1}{2\hat{\tau}(1-\Phi(\hat{\alpha}))}\phi\left(\frac{x-\widehat{\alpha\tau}}{\hat{\tau}}\right)$$

对于 $\hat{\tau} = \sqrt{\hat{\tau}^2}$和 $\hat{\alpha} = \widehat{\alpha\tau}/\hat{\tau}$。进一步定义 $\hat{w}_0 = 2(1-\Phi(\hat{\alpha}))$，可得

$$\hat{f}_0(\hat{\beta}_i) = (1-\hat{w}_0)\delta_0(\hat{\beta}_i) + \hat{q}_0(\hat{\beta}_i) \tag{28}$$

$\hat{}\ \hat{}\ \hat{}\ \hat{}\ \hat{}$ 作为 f_0 的估计量。最后，由于对于 $\beta_i \neq 0$，我们有 $f_0(\beta_i)/f(\beta_i) = q_0(\beta_i)/q(\beta_i)$，且 $\hat{}\ \hat{}\ \hat{}\ \hat{}\ \hat{}\ \hat{}f_0(\beta_i)/f(\beta_i) = \hat{q}_0(\beta_i)/\hat{q}(\beta_i)$，因此预言机的选择路径也可等价地写为

$$\hat{\mathcal{S}}^*(t) = \left\{i : \hat{\beta}_i \neq 0, \frac{q_0(\hat{\beta}_i)}{q(\hat{\beta}_i)} \leq t\right\}, \tag{29}$$

而经验贝叶斯程序的选择路径可写为

$$\hat{\mathcal{S}}^{EB}(t) = \left\{i : \hat{\beta}_i \neq 0, \frac{\hat{q}_0(\hat{\beta}_i)}{\hat{q}(\hat{\beta}_i)} \leq t\right\}. \tag{30}$$

4 理论结果

在上一节中，我们介绍了一种预言机程序，它根据变量的局部错误发现率对变量进行排序；我们还介绍了一种合法的经验贝叶斯程序，旨在模仿该预言机。本节将在 AMP 框架下分析这两种选择程序。最终目的是获得这两种程序的 FDP 和 TPP 的渐近预测，以便将它们的权衡曲线相互比较，并与 Lasso选择和阈值化 Lasso 选择等竞争方法进行比较。分析在技术上颇具挑战性，因为我们不能像大多数现有研究所采用的双组设定那样，假设 $(\beta_i,\ \beta_i)$ 是从 (18) 中独立同分布抽取的样本；在这里，我们只能依赖 Bayati 和 Montanari 的定理 1.5 中弱得多的性质。这不仅影响对局部错误发现率估计量本身极限的分析，尤其是在计算所得经验贝叶斯程序的渐近预测时也会造成问题，因为局部错误发现率是根据所有 β_i 估计出来的，这使我们无法直接应用 Bayati 和 Montanari 的定理来计算 FDP 和 TPP 的极限。

将 λ 视为固定值，我们对模型作如下温和假设：所有满足 x 且 $f_0(x)/f(x) \leq t$ 的 Bogdan 等人 [2013b]可以表示为若干个互不相交区间的并，且这些区间均排除零附近的某个邻域。该条件使我们能够利用 Bogdan 等人$-\infty,\ -\Delta]$ 以及 $[\Delta, \infty)$ $\Delta > 0$上非零 Lasso 估计量密度的一致连续性，从而证明预言机程序的假发现比例（FDP）和真阳性比例（TPP）收敛至各自的极限；更具体地说，该条件使我们可以援引这些结果。

假设 1。 存在 $t_0 > 0$，使得对于所有 $t \leq t_0$，我们都有

$$\frac{f_0(x)}{f(x)} \leq t \iff x \in \bigcup_{m=1}^{M}(a_m(t), b_m(t)),$$

其中（a_m（t） b_m（t））$m = 1,\ \ldots\ldots,\ M$ 均为互不相交的区间（允许 a_m（t）$= -\infty$ 或 b_m（t）$= \infty$）。此外，每个区间都与零保持一定距离，即对于每个 m

$$a_m(t), b_m(t) > 0 \text{ and } a_m(t) > \Delta \qquad \text{or} \qquad a_m(t), b_m(t) < 0 \text{ and } b_m(t) < -\Delta$$

对某个较小的固定值 $\Delta > 0$ 成立。

假设1 保证了

$$\{x : \frac{f_0(x)}{f(x)} \le t\} = \{x : x \neq 0, \frac{f_0(x)}{f(x)} \le t\} = \{x : \frac{q_0(x)}{q(x)} \le t\}$$

举例来说，如果 $\Pi = 0.9\delta_0 + 0.02\mathcal{N}(-3.6,\ 1) + 0.08\mathcal{N}(4,\ 1)$，那么，例如，$f_0(x)/f(x) \le 0.6$当 $x \ge 1.413$ 或 $x \le -1.941$ 时。

我们首先陈述关于预言机程序的理论结果。固定 $t \le t_0$，其中 t_0 是假设1中的常数，并考虑选择子集（29）。令

$$V^*(t;\lambda) = \sum_{i=1}^{p} \mathbf{1}\{\hat{\beta}_i \neq 0, \frac{q_0(\hat{\beta}_i)}{q(\hat{\beta}_i)} \le t, \beta_i = 0\} = \sum_{i=1}^{p} \mathbf{1}\{\frac{q_0(\hat{\beta}_i)}{q(\hat{\beta}_i)} \le t, \beta_i = 0\},$$
$$R^*(t;\lambda) = \sum_{i=1}^{p} \mathbf{1}\{\hat{\beta}_i \neq 0, \frac{q_0(\hat{\beta}_i)}{q(\hat{\beta}_i)} \le t\} = \sum_{i=1}^{p} \mathbf{1}\{\frac{q_0(\hat{\beta}_i)}{q(\hat{\beta}_i)} \le t\},$$

分别表示错误拒绝数和拒绝总数；在该展示式的两行中，最后一个等号均由假设 1得出。定义

$$t_{V_m}(x,y) := \mathbf{1}\{x \in (a_m(t), b_m(t))\} \cdot \mathbf{1}\{y = 0\}, \qquad t_{R_m}(x,y) := \mathbf{1}\{x \in (a_m(t), b_m(t))\},$$

对于假设 a_m，b_m 中的函数 1，使得

$$V^*(t;\lambda) = \sum_{m=1}^{M}\sum_{i=1}^{p} t_{V_m}(\hat{\beta}_i, \beta_i), \qquad R^*(t;\lambda) = \sum_{m=1}^{M}\sum_{i=1}^{p} t_{R_m}(\hat{\beta}_i, \beta_i).$$

以下引理是获得预言机程序渐近预测的关键步骤。其证明用到了引理 A.1、A.2 和 A.3，这些引理收录于附录中。证明的主要难点在于，Lasso估计并非真正独立同分布；因此，无法直接应用大数定律来获得所需的极限，我们必须转而使用 2 中的定理 Bayati 和 Montanari [2012]。应用该定理时，我们采用与 Su等人 [2017]类似的方法，避开FDP和TPP表达式中涉及的指示函数的不连续性。

引理 4.1。 在假设 *1* 以及第 *2* 节的渐近设定下，我们有 *ve*

$$\mathrm{FDP}^*(t;\lambda) := \frac{V^*(t;\lambda)}{R^*(t;\lambda) \vee 1} \longrightarrow \frac{\mathbb{E}\sum_{m=1}^{M} t_{V_m}(\eta_{\alpha\tau}(\Pi + \tau Z), \Pi)}{\mathbb{E}\sum_{m=1}^{M} t_{R_m}(\eta_{\alpha\tau}(\Pi + \tau Z), \Pi)},$$

依概率收敛于 $p \to \infty$，其中 $\tau > 0$，$\alpha > \alpha_{\min}$ 是方程 (8)-(9)的唯一解。

下面的定理 4.2 给出了预言机程序（29）的假发现比例（FDP）和真阳性比例（TPP）的极限。结合引理（4.1），根据Bayati 和 Montanari [2012]的结果可得，这与Bogdan 等人 [2013为 Lasso选择所] 得到的结果类似。

定理 4.2。 在第 *2* 节的渐近框架下，固定 λ 和阈值 $t \le t_0$，其中 t_0 是假设 *1* 中的常数，假发现比例$^*(t;\lambda)$ 和 真阳性比例$^*(t;\lambda)$ 均依概率收敛于某个极限，当 $p \to \infty$时，我们将其记为

$$\mathrm{fdp}^*(t;\lambda) := \lim \mathrm{FDP}^*(t;\lambda) = \lim \frac{\sum_{i=1}^{p} \mathbf{1}\left\{\frac{q_0(\hat{\beta}_i)}{q(\hat{\beta}_i)} \le t,\ \beta_i = 0\right\}}{\sum_{i=1}^{p} \mathbf{1}\left\{\frac{q_0(\hat{\beta}_i)}{q(\hat{\beta}_i)} \le t\right\}},$$

$$\mathrm{tpp}^*(t;\lambda) := \lim \mathrm{TPP}^*(t;\lambda) = \lim \frac{\sum_{i=1}^{p} \mathbf{1}\left\{\frac{q_0(\hat{\beta}_i)}{q(\hat{\beta}_i)} \le t,\ \beta_i \neq 0\right\}}{\sum_{i=1}^{p} \mathbf{1}(\beta_i \neq 0)}.$$

此外，这些极限为

$$\mathrm{fdp}^*(t;\lambda) = \frac{(1-\varepsilon)\mathbb{P}(\frac{q_0(\eta_{\alpha\tau}(\tau Z))}{q(\eta_{\alpha\tau}(\tau Z))} \le t)}{(1-\varepsilon)\mathbb{P}(\frac{q_0(\eta_{\alpha\tau}(\tau Z))}{q(\eta_{\alpha\tau}(\tau Z))} \le t) + \varepsilon\mathbb{P}(\frac{q_0(\eta_{\alpha\tau}(\Pi_1+\tau Z))}{q(\eta_{\alpha\tau}(\Pi_1+\tau Z))} \le t)} \tag{31}$$

and

$$\mathrm{tpp}^*(t;\lambda) = \mathbb{P}\left(\frac{q_0(\eta_{\alpha\tau}(\Pi_1+\tau Z))}{q(\eta_{\alpha\tau}(\Pi_1+\tau Z))} \le t\right). \tag{32}$$

在分析经验贝叶斯程序之前，我们先陈述以下结果，以形式化说明预言机规则（29）在何种意义上是渐近最优的。该结果表明，程序（29）确实是一个预言机，不仅因为它依赖于未知参数，而且也因为它在通常意义上是类（5）中的（渐近）最优程序，其中任意 λ均如此。

定理 *4.3*。固定 $\lambda > 0$，并假设假设 *1* 对某个常数 t_0 成立。对于固定阈值 $t^* \le t_0$，记 $\xi = \mathit{tpp}^*(t^*;\lambda)$ 为预言机程序 *(29)* 的渐近真阳性比例。现在考虑在固定阈值 t 下按照 *(2)* 进行选择，其中 T_i 由某个函数 φ 按照 *(5)* 给出，并令 $\mathit{fdp}^\varphi(t;\lambda)$ 和 $\mathit{tpp}^\varphi(t;\lambda)$ 分别表示选择规则的假发现比例和真阳性比例的相应极限（若存在）。若 φ 满足以下条件：*Bayati* 和 *Montanari* 的定理 *1.5* *[2012]* 适用于 $\psi(x,y) = \mathbf{1}(\varphi(x) \ge t)\mathbf{1}(y=0)$ 和 $\psi(x,y) = \mathbf{1}(\varphi(x) \ge t)$，那么对于上述类中的任意程序，$\mathit{tpp}^\varphi(t;\lambda) = \xi$ 意味着 $\mathit{fdp}^*(t^*;\lambda) \le \mathit{fdp}^\varphi(t;\lambda)$。

我们转而分析完全由数据驱动的 EB 选择程序。为了在我们的渐近框架下建立假发现比例和真阳性比例的相应极限，必须证明估计量 $\hat{}q(x)$ 依概率一致收敛于 $q(x)$，其中 $[\Delta,\infty)$ 和 $(-\infty,-\Delta]$,Δ 是在假设 1中定义的。我们首先在命题4.5中证明逐点收敛，其证明关键依赖于以下引理。

引理 4.4。设

$$k(y,x) = \frac{1}{h}\mathcal{K}\left(\frac{x-y}{h}\right) = \frac{1}{\sqrt{2\pi h^2}}e^{-\frac{1}{2}(y-x)^2/h^2}, \quad h>0,$$

$\hat{}$ $\hat{}$由 $\mathcal{K}=\phi$ 标准高斯核得到。定义 $\psi(\beta_i,\beta_i) = k(\beta_i,x)$，其中 h,x 固定。则 ψ： $\mathbb{R}\times\mathbb{R}\to\mathbb{R}$ 是一个伪 *Lipschitz* 函数，并且

$$\frac{1}{p}\sum_{i=1}^{p}\mathbf{1}\{\hat{\beta}_i \ne 0\}\,\psi(\hat{\beta}_i,\beta_i) \longrightarrow \mathbb{E}\left[\mathbf{1}\{\eta_{\alpha\tau}(\Pi+\tau Z)\ne 0\}\,\psi(\eta_{\alpha\tau}(\Pi+\tau Z),\Pi)\right]$$

依概率收敛，当 $p\to\infty$时。

Pro命题 4.5。 在第2节的渐近框架下，并在假设1下，

$$\hat{q}(x) \longrightarrow q(x)$$

依概率收敛，当 $p\to\infty$时。

为了使这一结果扩展到 $\hat{}q(x)$ 的一致收敛，我们还需要一个额外假设。假设 2. 存在常数 $\mathcal{L}>0$，使得未归一化密度 $q(x)$ 在区间 $[-\mathcal{L},-\Delta]$ 和 $[\Delta,\mathcal{L}]$ 上均一致连续，其中 $\Delta>0$ 是假设 1中的常数。

命题 4.6. 在假设 *1*和 *2*下，核密度估计量 $\hat{}q(x)$ 在 (26) 中定义，且对于 $\mathcal{K}(x) = \phi(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}$，依概率一致收敛到 $q(x)$ 于 $[\Delta,\mathcal{L}]$上，即：

$$\sup_{x\in[\Delta,\mathcal{L}]}|\hat{q}(x)-q(x)| \longrightarrow 0$$

依概率收敛至 $p\to\infty$。

注 4.7. 命题 4.6 中的一致收敛在区间 $[-\mathcal{L}, -\Delta]$上同样成立。现在我们可以推导经验贝叶斯程序的渐近预测。

定理 4.8。在第 *2* 节的渐近框架下，且满足假设 *1* 和*2*，固定 λ 及阈值 $t \le t_0$，其中 t_0 是来自假设 *1*的常数。那么，当 假发现比例$^{EB}(t;\lambda)$和 真阳性比例$^{EB}(t;\lambda)$ 趋于无穷时，二者都依概率收敛于极限 $p \to \infty$，记为

$$\mathrm{fdp}^{EB}(t;\lambda) := \lim \mathrm{FDP}^{EB}(t;\lambda) = \lim \frac{\sum_{i=1}^{p} \mathbf{1}\left\{\hat{\beta}_i \neq 0,\ \frac{\hat{q}_0(\hat{\beta}_i)}{\hat{q}(\hat{\beta}_i)} \le t,\ \beta_i = 0\right\}}{\sum_{i=1}^{p} \mathbf{1}\left\{\hat{\beta}_i \neq 0,\ \frac{\hat{q}_0(\hat{\beta}_i)}{\hat{q}(\hat{\beta}_i)} \le t\right\}},$$

$$\mathrm{tpp}^{EB}(t;\lambda) := \lim \mathrm{TPP}^{EB}(t;\lambda) = \lim \frac{\sum_{i=1}^{p} \mathbf{1}\left\{\hat{\beta}_i \neq 0,\ \frac{\hat{q}_0(\hat{\beta}_i)}{\hat{q}(\hat{\beta}_i)} \le t,\ \beta_i \neq 0\right\}}{\sum_{i=1}^{p} \mathbf{1}(\beta_i \neq 0)}.$$

此外，还有

$$\mathrm{fdp}^{EB}(t;\lambda) = \mathrm{fdp}^{*}(t;\lambda), \qquad \mathrm{tpp}^{EB}(t;\lambda) = \mathrm{tpp}^{*}(t;\lambda), \tag{33}$$

其中 $\mathrm{fdp}^{*}(t;\lambda)$ 和 $\mathrm{fdp}^{*}(t;\lambda)$ 分别是公式 (31) 和 (32)。

注释 4.9。定理 4.8 对任意 Lipschitz 连续的核 $\mathcal{K}$ 也更一般地成立；也就是说，若存在常数 L 使得 $|\mathcal{K}(x) - \mathcal{K}(y)| \le L|x-y|$ 对所有 $x,\ y \in \mathbb{R}$，则此条件成立。上文使用的高斯核满足该条件。

定理4.8 表明，对于任意固定的 t，预言机规则（29）和 EB 竞争方法（30）在渐近意义下具有相同的假发现比例和真阳性比例；换言之，在我们的假设下，理论预测这两种方法的权衡曲线重合。这一结果令人鼓舞，因为它表明，对于任意固定的 λ，类别（5）中最优方法的性能可以由一个合理的方法渐近实现。不过，性能取决于正则化参数 λ的选择，而 λ的不同取值可能导致截然不同的权衡曲线（阈值化 Lasso 也是如此）。因此，对于每个分布 Π，自然会问 λ 取何值才是最优的，也就是在给定渐近真阳性比例水平下，使渐近错误发现比例最小。对于阈值化 Lasso 选择，Wang等人 [2020]证明，λ的最优取值是渐近均方误差的极小值点，

$$\lambda^{*} := \arg\min_{\lambda} \mathbb{E}[(\eta_{\alpha\tau}(\Pi + \tau Z) - \Pi)^2], \tag{34}$$

并且根据（8），可直接得出 λ^{*} 等价于以下优化问题的最小化解：τ 在 λ中。我们的下一个结果表明，上文定义的 λ^{*} 对于基于局部错误发现率进行选择的预言机（29）同样是最优的。在下述定理的陈述及其证明中，凡是出现星号（$*$）上标，均表示 $\lambda = \lambda^{*}$时的量；若没有上标，则表示任意的 $\lambda \neq \lambda^{*}$；例如，$\mathrm{lfdr}^{*}(x)$ 是在 $\lambda = \lambda^{*}$下计算的式（20）中的局部错误发现率，而 $\mathrm{lfdr}(x)$ 是任意 $\lambda \neq \lambda^{*}$对应的这一量。

定理 *4.10*。在假设 *1* 下，令 t_0^{*} 为 λ^{*}时 t_0 的取值。对于某个 $t^{*} \le t_0^{*}$ ˆ，考虑满足 $lfdr^{*}(\beta_i) \le t^{*}$ 时进行选择的预言机，并令 $\xi = tpp^{*}(t^{*};\lambda^{*})$ 为相应的渐近真阳性比例。假设存在另一个 λ 及相应的固定 t，使得满足 $lfdr(\beta_i) \le t$ 时进行选择的预言机具有 $tpp^{*}(t;\lambda) = \xi$。对于任意此类 λ，有

$$fdp^{*}(t^{*};\lambda^{*}) \le fdp^{*}(t;\lambda).$$

需要指出的是，定理4.10 所蕴含的预言机比定理4.3中的更强。具体而言，若 $\mathcal{P}$ 表示满足定理 φ 4.3 中条件的单变量函数集合，那么结合定理4.10 和定理 4.3，并利用以下事实：

根据定理 4.2 和定理 4.8，预言机程序与经验贝叶斯程序具有相同的渐近权衡曲线，因此我们得出结论：对于任意一对 (φ, λ)，只要 $\varphi \in \mathcal{P}$，若 $\mathrm{tpp}^*(t^*; \lambda^*) = \mathrm{tpp}^{\varphi}(t; \lambda)$，则 $\mathrm{fdp}^*(t^*; \lambda^*) \leq \mathrm{fdp}^{\varphi}(t; \lambda)$。也就是说，若固定渐近真阳性比例水平，那么采用 λ^* 的预言机程序可在式 (5) 中同时对 φ 和λ 的所有选择达到最小渐近错误发现比例；而定理 4.3 仅表明，对于 λ 的任意固定取值，该程序在 $\varphi \in \mathcal{P}$中达到最小值。这一点意义重大，因为扩展后的类别允许同时对 λ进行搜索，其中包含 Lasso 选择程序 (12)，这意味着在固定真阳性比例水平时，若取 λ^*，经验贝叶斯程序的渐近错误发现比例必然低于 Lasso（这并不排除在某些其他 λ 下 Lasso 优于经验贝叶斯程序的可能性）。在论文的最后一节中，我们将解释为何定理 4.3 中的最优性结果针对任意固定的 λ 而非最优值 λ^*来表述。

在上述定理4.10 的陈述中，结果对经验贝叶斯程序也成立这一断言，可直接由以下事实推出：在我们的假设下，这两种程序的权衡曲线相同。不过，最优值 λ^* 取决于问题中的未知量，也就是说，它仍是一个预言机值，在实践中无法使用。式（ λ^* 34）对 λ_{cv} 的刻画表明，可以通过交叉验证对其进行良好的估计。具体而言，令 λ 表示通过 K折交叉验证得到的 Weinstein 等人 [2023] 的引理 λ_{cv} 4.1 中，在 AMP 渐近框架下，推导了 $\boldsymbol{X}$ 的极限，其中 λ_{cv} 增添了一个 knockoffs 矩阵。将这一结果调整到当前设定（不涉及 knockoffs）后，我们立即得到以下结论。

命题 4.11 （Weinstein 等人 [2023]的引理 4.1 的推论）。 在我们的渐近框架下，当 $\lambda_{cv} \longrightarrow \lambda_{cv}^*$ 时， $p \to \infty$依概率成立，其中

$$\lambda_{cv}^* := \arg\min_{\lambda} \tau(\lambda; (K-1)\delta/K). \tag{35}$$

此外，通过调整Weinstein 等人提出的公式 (4.2) [2023], ，我们得出

$$\lambda_{cv}^* = \alpha_{\mathrm{cv}}\tau_{\mathrm{cv}}\left[1 - \frac{K}{(K-1)\delta}\mathbb{P}(|\Pi + \tau_{\mathrm{cv}}Z| > \alpha_{\mathrm{cv}}\tau_{\mathrm{cv}})\right],$$

其中， α_{cv} 和 τ_{cv} 是以下方程组的解：

$$\begin{aligned} &\tau_{\mathrm{cv}}^2 = \sigma^2 + \frac{K}{(K-1)\delta}\mathbb{E}[\eta_{\alpha_{\mathrm{cv}}\tau_{\mathrm{cv}}}(\Pi + \tau_{\mathrm{cv}}Z) - \Pi]^2 \\ &\mathbb{E}(Z + \alpha_{\mathrm{cv}}; \Pi + \tau_{\mathrm{cv}}Z < -\alpha_{\mathrm{cv}}\tau_{\mathrm{cv}}) - \mathbb{E}(Z - \alpha_{\mathrm{cv}}; \Pi + \tau_{\mathrm{cv}}Z > \alpha_{\mathrm{cv}}\tau_{\mathrm{cv}}) = 0. \end{aligned} \tag{36}$$

5 数值实验

现在，我们通过一些数值研究来展示上一节理论预测的充分性。首先，我们进行模拟，以考察预言机和经验贝叶斯程序的经验权衡曲线与理论预测的一致性。图2中的每个子图都展示了经验权衡曲线，即在不同设置下进行的 17 次独立模拟运行中，沿着实际选择路径得到的真实假发现比例和真阳性比例，其中我们固定了 $\lambda = 1$。所有子图中的数据（Y， $\boldsymbol{X}$）均由公式（1）和（7）给出的双层模型生成，其中 $\sigma = 1$ 且 $\epsilon = 0.1$。左列展示 $p = 5000$, $n = 2500$ （对应于 $\delta = 0.5$）时的结果，右列展示$p = 5000$， $n = 9000$ （对应于 $\delta = 1.8$）时的结果。图中的每一行对应非零分量 Π_1的一种不同取值：第一行中， Π_1 服从均值为 3.5、方差为 1 的高斯分布；第二行中， Π_1 服从两个高斯分布的混合，一个均值为 -3.6、方差为 1，另一个均值为 4、方差为 1，二者的概率分别为 0.2 和 0.8；第三行中， Π_1 在 -4.3 处具有点质量；第四行中，它是两个点质量分布的混合，一个位于 -2，另一个位于 3，其概率为

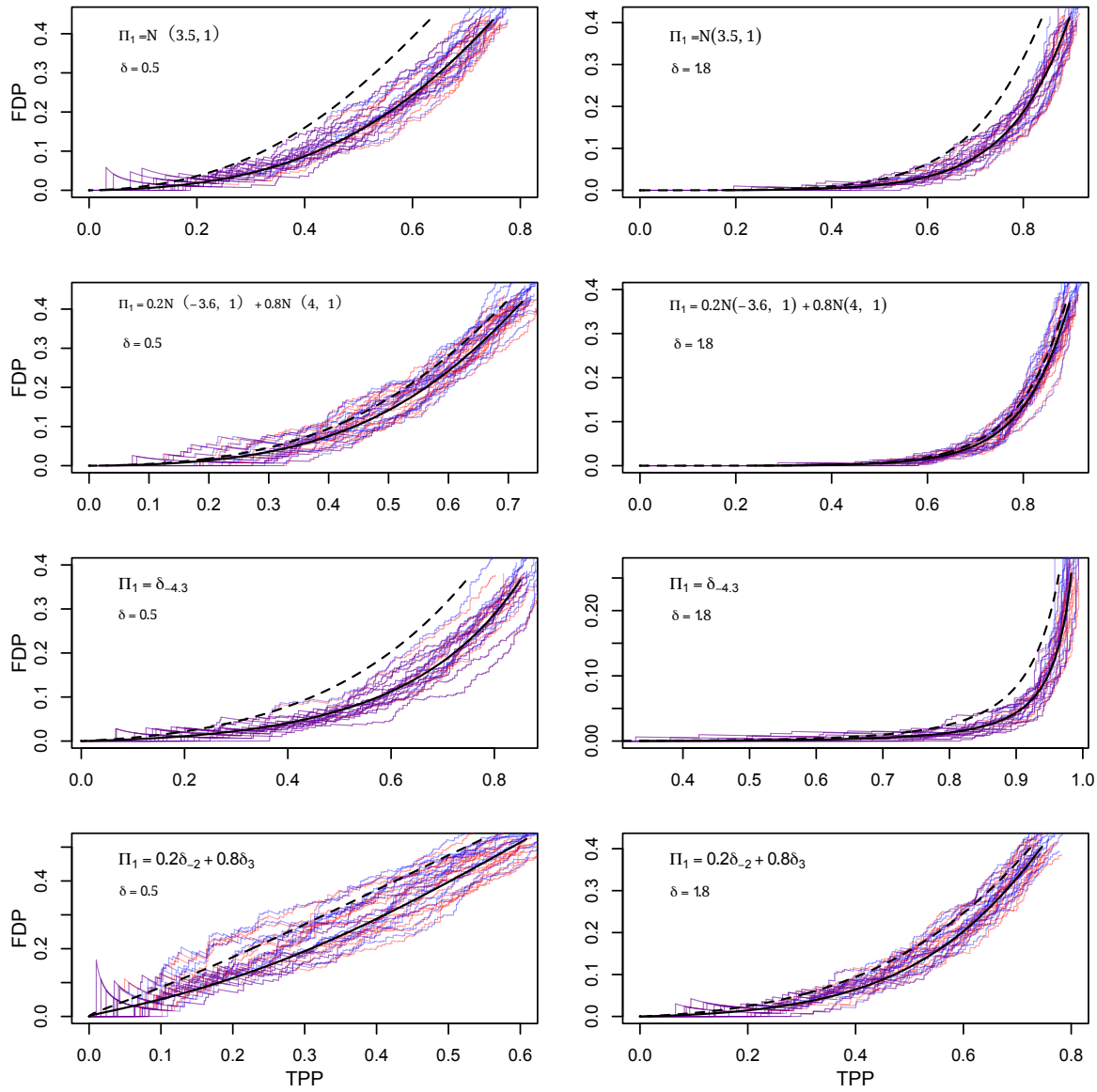

图2： 预言机和经验贝叶斯程序的经验权衡曲线与理论预测的对比。左列显示 $\delta = 0.5$ 的结果，右列显示 $\delta = 1.8$ 的结果。各行对应 Π_1 的不同取值。细线表示 *17* 次模拟运行中实际得到的假发现比例（*FDP*）和真阳性比例（*TPP*）；红色代表预言机，蓝色代表经验贝叶斯程序。黑色实线表示理论预测 $fdp^*(t)$、$tpp^*(t)$。黑色虚线表示阈值化 *Lasso* 的理论曲线。更多细节见正文。

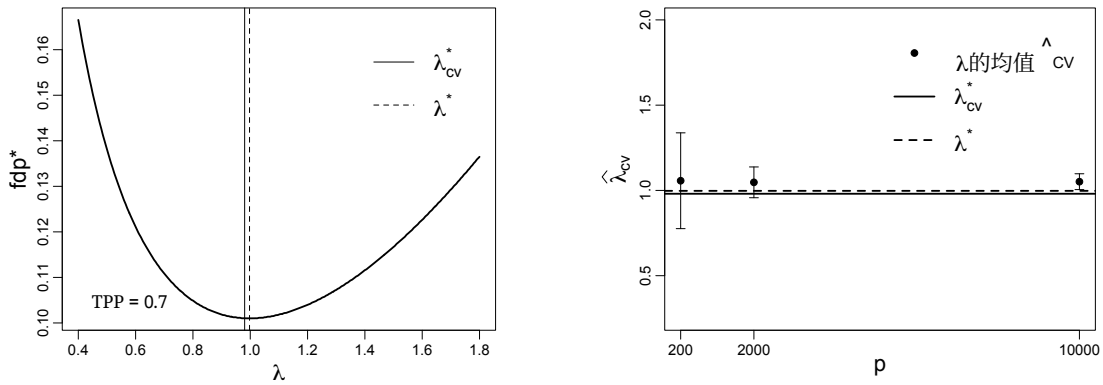

图3： 左图：固定渐近真阳性比例为0.7时，渐近错误发现比例与 λ 的关系；所用设置为$\epsilon = 0.1$、$\delta = 1$、$\sigma = 1$和 $\Pi_1 = .2\mathcal{N}(-3.6,\ 1) + .8\mathcal{N}(4,\ 1)$。右图：不同 p下 λ_{cv} 的模拟均值和标准误ˆ。

分别为 0.2 和 0.8。细蓝线表示经验贝叶斯程序的经验权衡曲线，细红线表示预言机程序的经验权衡曲线；该程序利用 Π_1, ϵ 和 σ ，根据真实的（渐近）局部错误发现率对变量进行排序。在每个子图中，黑线叠加显示理论权衡曲线，该曲线预测预言机程序和经验贝叶斯程序的渐近错误发现比例和渐近真阳性比例。作为比较，黑色虚线表示阈值化 Lasso 的理论权衡曲线。图中结果表明，理论与经验结果吻合良好。特别是，预言机程序和经验贝叶斯程序的经验曲线难以区分，这与理论预测一致：二者渐近地重合。

在图3中，我们考察 λ 对EB 选择规则检验功效的影响。左图显示经验贝叶斯程序的渐近错误发现比例随 λ的变化，此时渐近真阳性比例固定为0.7（一个任意选择），示例取 $\epsilon = 0.1$、$\delta = 1$、$\sigma = 1$和$\Pi_1 = .2\mathcal{N}(-3.6, 1) + .8\mathcal{N}(4, 1)$；也就是说，该图显示fdp*($t(\lambda)$; λ) 对 λ 的关系，其中 $t(\lambda)$ 使得 tpp*($t(\lambda)$; λ) $= 0.7$。虚竖线标出 (34) 中指定的最优值 λ^*，实竖线标出(35)中的 λ^*_{cv} 。该图与上一节的理论一致，该理论预测 λ^* 对任意固定的真阳性比例水平使渐近错误发现比例最小。图的右图给出相应ˆ模拟的结果，我们在其中考察命题 4.11所预测的 λ_{cv}向 λ^*_{cv}的收敛。对于 $p \in \{200, 2000, 10^4\}$，我们设定 $n = p$，并报告ˆ λ_{cv} 的均值和标准误，其中针对 $p = 200$、2000基于1000次独立运行，针对 $p = 10^4$基于100次运行。为考察ˆ因使用带 λ_{cv} 的经验贝叶斯程序而非 λ的最优选择而预期的渐近检验效能损失，图4将 λ^*_{cv} 的渐近错误发现比例与 λ^*的渐近错误发现比例进行比较，二者均校准到相同的渐近真阳性比例水平。也就是说，若对任意 λ，令 $t_\xi(\lambda)$使得tpp*($t_\xi(\lambda)$; λ) $= \xi$，则这些图显示fdp*($t_\xi(\lambda^*_{cv})$; λ^*_{cv})相对于fdp*($t_\xi(\lambda^*)$; λ^*)，随真阳性比例水平 ξ 变化。图中 $\delta = 1$，不同面板对应于图2中 Π_1 的四种不同选择。在所有面板中，这些点都紧密跟随恒等线，表明检验效能损失渐近地可忽略。

6 估计局部错误发现率：超越检验效能考量

到目前为止，我们对估计局部错误发现率（20）的兴趣，纯粹是出于检验效能考量，具体而言，是因为在所有属于该类的程序中，对于任意固定的渐近真阳性比例水平，它都是使渐近错误发现比例最小化的重要性统计量的（预言机）选择（5）。现在，我们将视角从检验效能考量拓展开来，说明为什么估计局部错误发现率本身就值得关注。这里采用的观点受到 Soloff 等人 [2024],Xiang 等人 [2025]所提出思想的启发。在较高

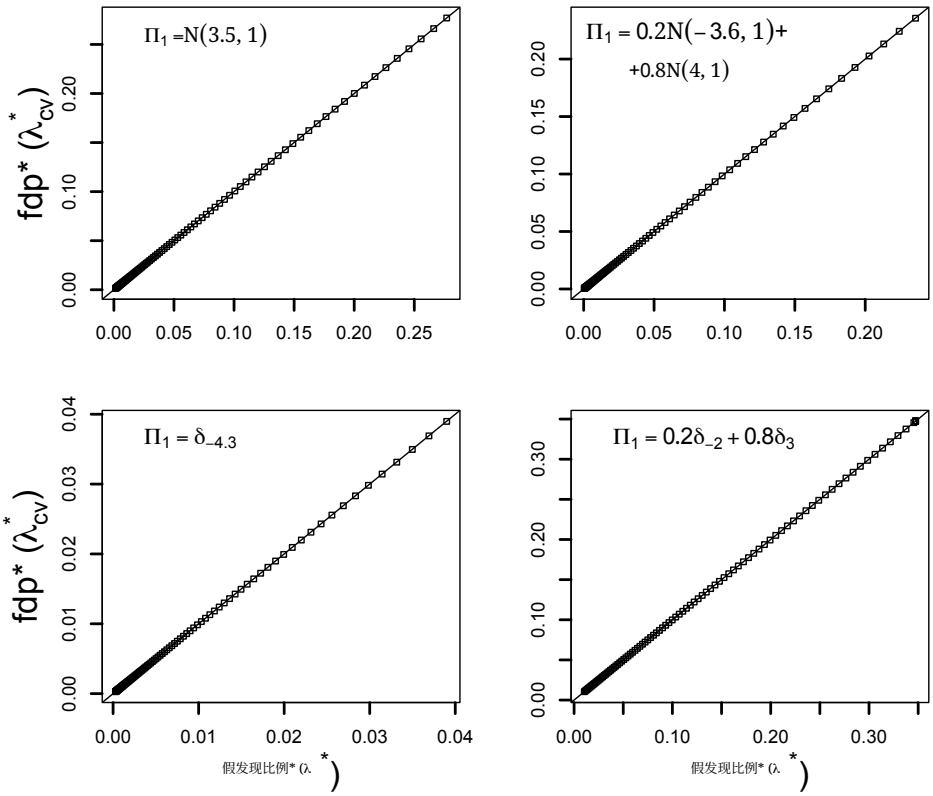

图4： 对 $\delta = 1$以及 Π_1的四种不同选择，比较使用 λ^*_{cv} 与最优值 λ^* 时经验贝叶斯程序的渐近错误发现比例。在每个面板中，每个点对应于（渐近）真阳性比例水平的某个固定值 ξ 。图中给出恒等线以供参考。

层面上，Soloff 等人 [2024]对在多重检验（双组模型下）中将错误发现率用作第一类错误判据提出了概念上的质疑，因为某个程序的（贝叶斯）错误发现率只考量了所有拒绝构成的整体中所犯错误的比例，而对每个单独拒绝的错误率作出保证或许更为恰当。具体而言，将这一观点应用于我们的设定，利用期望的迭代性质可以证明，对于某个统计量 i，若程序在 $T_i \leq t$ 时选择第 $T_i = \varphi(\beta_i)$ 个变量，且 t固定，则该程序的错误发现率渐近地等于在拒绝事件 [1] 条件下，lfdr(的期望)β_i。这意味着某些拒绝——通常是那些 $T_i \leq t$接近拒绝边界的 T_i ——其局部错误发现率可能远高于名义错误发现率水平（而其他拒绝的局部错误发现率可能低得多）。为弥补这一缺陷，Soloff 等人提出改用直接针对局部错误发现率的错误判据，例如，力求控制所有拒绝中的最大局部错误发现率的期望（“最大局部错误发现率”）。此外，他们还提出了 支持线(SL) 程序。该程序基于 p值，在假设数据独立同分布且服从双组模型、备择分布密度满足单调性条件的情况下，能够在有限样本中控制最大局部错误发现率（即使不满足单调性条件，也仍可在有限样本中得到稍弱的保证）。这些有限样本结果并不正式适用于回归模型下更复杂的受控变量选择问题（1）：即便可以计算出边际上服从双组模型的各个 p 值，它们通常也并不相互独立。事实上，要在变量选择问题中找到一种能够对最大局部错误发现率或类似的、基于局部错误发现率的错误度量提供有限样本控制保证的程序，似乎极具挑战性。

[1]如果配对（$\hat{}\beta_i$，β_i）实际上独立同分布于（18），则有pFDR ＝错误发现率/$\mathbb{P}$（$R > 0$） ＝$\mathbb{E}$[局部错误发现率（β_i）$|T_i \leq t]$ 对任意有限的 p [Efron， 2008]，其中 R 是拒绝总数；这在我们的设定中并不成立，但正文中的渐近断言仍然成立。

尽管如此，在本文的研究范围内，假设我们已经计算出给定某个固定 β_1、……、 β_p 时的 λLasso 估计。我们不考虑根据 β_i的某种函数对变量排序的变量选择程序的错误发现率，而是可以考虑 β_i 这些统计量本身，并且暂不作出任何决策，而是为每个统计量关联相应的局部错误发现率（6）。局部错误发现率是一个未知量，因为其计算需要知道 Π （以及 σ）；事实上，只有当（ P（$\beta_i = 0|\beta_i$）在边际上确实服从（β_i， β_i）时，它才具有后验概率的严格含义，而18）并非如此，正如我们前面所指出的。尽管如此，在特殊的 AMP 框架下，按照第4 节中的类似计算，可以渐近预测对于固定的 $\beta_i \in [s,t]$ ，在满足 $0 < s < t$的变量中真实零假设所占的比例，

$$\mathrm{FDP}\big([s,t]\big) := \frac{\#\{i : s \le \hat\beta_i \le t,\, \beta_i = 0\}}{\#\{i : s \le \hat\beta_i \le t\}} \;\longrightarrow\; \mathbb{P}\big(\Pi = 0 \big| s \le \eta_{\alpha\tau}(\Pi + \tau Z) \le t\big)$$

依概率收敛至 $p \to \infty$。反过来，当区间长度趋于零时，右侧表达式收敛至 t 处的局部错误发现率，即，

$$\mathbb{P}\big(\Pi = 0 \big| s \le \eta_{\alpha\tau}(\Pi + \tau Z) \le t\big) \longrightarrow \mathbb{P}\big(\Pi = 0 \big| \eta_{\alpha\tau}(\Pi + \tau Z) = t\big) = \frac{(1-\epsilon) f_0(\hat\beta_i)}{f(\hat\beta_i)} = \mathrm{lfdr}(t)$$

趋于 $s \to t^-$。因此，我们有

$$\lim_{s \to t^-} \lim_{p \to \infty} \mathrm{FDP}\big([s,t]\big) = \mathrm{lfdr}(t), \tag{37}$$

ˆ 也就是说，lfdr(t) 是一个极限比例：在 β_i 位于 t周围不断缩小的区间内的变量中，真实零假设所占的比例。注意，式 (20) 中的函数 lfdr(·) 并未被声称代表给定 β_i 时 $\beta_i = 0$ 的实际后验概率（同样，只有当 (β_i, β_i) 按照 (18) 独立同分布时，这种解释才成立）。尽管如此，(37) 仍赋予 AMP 框架下的式 (20) 中的局部错误发现率一种解释。与第 2 节开头的评论一致，我们还指出，即使 β_i 是固定的，只要其经验分布满足相应条件，(37) 仍有意义。

如果我们能够获得真实函数 lfdr(t)，就可以超越对变量选择程序的错误发现率评估，转而进行更具探索性的分析。事实上，假设我们观察到无标签向量 β，即数值 β_i ，但不知道它们的顺序。如果我们将这些数值中的一部分——完全事后——标记为“潜在感兴趣”的位置，那么 (37) 表明，在这些位置中的任意一处，局部错误发现率大致就是该位置对应真实零假设 $\beta_i = 0$ 的概率。局部错误发现率是未知的，但我们可以对其进行估计，例如

$$\widehat{\mathrm{lfdr}}(t) = \frac{\widehat{1-\epsilon}\,\hat f_0(t)}{\hat f(t)}, \qquad \widehat{1-\epsilon} = \frac{N(\lambda)}{p[1 - 2\Phi(-\hat\alpha)]},$$

ˆ ˆ ˆ其中 $N(\lambda) = \{i \le p : \beta_i(\lambda) = 0\}$； f 并且 f_0 分别由式（25）和式（28）给出；ˆα 由式（27）得到。在我们的渐近设定以及第 4 节的假设下，我们有

$$\begin{aligned}\lim_{p\to\infty} \frac{N(\lambda)}{p} &= \mathbb{P}\left(|\Pi + \tau Z| \le \alpha\tau\right) = (1-\epsilon)\mathbb{P}(|Z| \le \alpha) + \epsilon\mathbb{P}\left(|\Pi_1 + \tau Z| \le \alpha\tau\right) \\ &= (1-\epsilon)[1 - 2\Phi(-\alpha)] + \epsilon\mathbb{P}\left(|\Pi_1 + \tau Z| \le \alpha\tau\right)\end{aligned}$$

and

$$\lim_{p\to\infty} \widehat{\mathrm{lfdr}}(t) \;=\; \mathrm{lfdr}(t) \;+\; \frac{\epsilon\mathbb{P}\left(|\Pi_1 + \tau Z| \le \alpha\tau\right)}{1 - 2\Phi(-\alpha)} \frac{f_0(t)}{f(t)}, \tag{38}$$

预计这只是对 lfdr(t) 的轻微高估，因为 $\mathbb{P}\left(|\Pi_1 + \tau Z| \le \alpha\tau\right)$ 通常很小。作为说明，图 5 对应于一个 $\epsilon = 0.1$ 的示例， $\delta =$

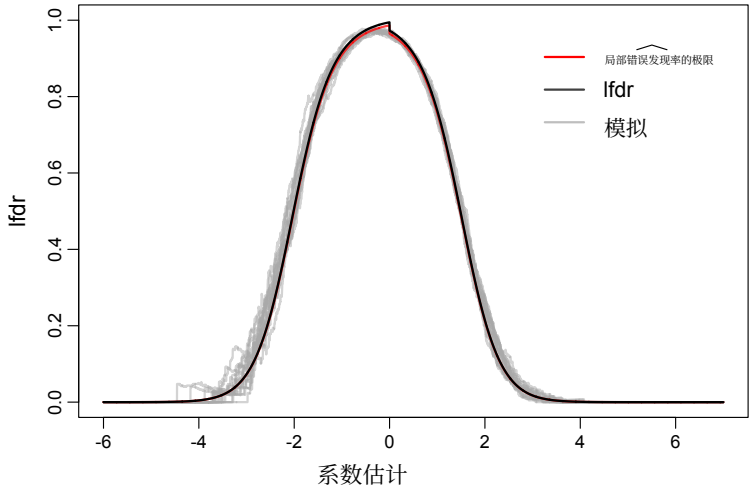

图5： 在一个示例中对局部错误发现率的估计，其中使用了 $\epsilon = 0.1, \delta = 1, \sigma = 1$, $\Pi_1 = .2\mathcal{N}(-3.6, 1) + .8\mathcal{N}(4, 1)$, 以及 $\lambda = 1$ 来计算 *Lasso*估计。实线曲线表示真实局部错误发现率及其估计值的理论渐近预测。灰线表示 *17* 次独立模拟运行中，在一个较小滑动窗口内实现的假发现比例；n=p= *14000*，具体说明见正文。

1、$\sigma = 1$、$\Pi_1 = .2\mathcal{N}(-3.6、1) + .8\mathcal{N}(4、1)$ 以及 $\lambda = 1$。实线曲线表示 lfdr(t)（黑色）以及所提估计量 $\widehat{\text{lfdr}}$(38) 的极限（红色）；在该示例中，两者确实非常接近。图中的灰线是模拟结果，模拟取 $n = p = 14000$；在 17 次独立运行中的每一次里，我们都计算了 $\text{FDP}([t - \Delta, t + \Delta])$，其中我们使用 $\Delta = 0.4$，即 t的小邻域内的实际 FDP，并针对 t 的一组网格值进行计算。理论预测与经验结果高度一致；特别是，它们捕捉到了左肩部和右肩部的不同行为。

7 讨论

最后，我们对当前工作未涉及的几个方面进行讨论。关于预言机的定义，请注意，定理 4.3 断言，对于固定的 λ，在 φ 上具有最优性，尽管我们随后也确定了 λ 的最优取值（定理 4.10）。之所以这样表述，是因为对于固定的 λ，我们能够证明预言机的性能渐近地可由一个合规的程序达到，即我们的 EB 选择规则。一个更具雄心的结果是，配备 λ_{cv} （而非固定的 λ）的经验贝叶斯程序能够渐近地达到使用 λ^*_{cv} 的预言机的性能；根据式（34）和第 5 节的数值示例，预期其表现几乎与使用 λ^* 一样好。尽管在我们的模拟中，经验贝叶斯程序在 λ_{cv} 处的 $\hat{}$ 收敛似乎确实成立，但要正式证明这一点，需要确立 $\hat{}q_0/\hat{}q$ 在 λ 中的一致收敛；在我们的方法中，这要求将关于 KDE $\hat{}q$ 的结果推广到 λ 上的一致收敛，并证明估计量 $\widehat{\alpha\tau}$ 和 $\hat{}\tau$ 在 λ 中的一致收敛。这种分析需要大量额外的技术工作，因此本文不予涉及。

其次，回到开篇部分讨论的一点，我们指出，在实践中，统计学家通常会希望在选择重要性统计量的同时，再辅以选择一个数据依赖阈值 $\hat{}t$ ，用以校准（或近似校准）针对第一类错误的选择。例如，Weinstein 等人 [2017] 使用了一种简单的 Knockoff 方法实现，在我们的特殊设定下（$\boldsymbol{X}$ 具有独立同分布的元素）对任意 n，p 控制错误发现率；对于所提出的经验贝叶斯程序，也可以考虑类似的 Knockoff 方法类比。在这种情况下，本文中与预言机的比较不仅需要针对预言机作用于原始 $\boldsymbol{X}$ 而 Knockoff 方法将作用于增广矩阵这一事实进行调整，还需要针对 Knockoff 方法程序潜在的保守性，在

选择数据依赖阈值方面。更具体地说，假设具有较大 T_i的变量被选择，在 Weinstein 等人 [2017] 的策略中，数据驱动阈值 $\hat{t}_q$ 是最宽松的 t，使得假发现比例$(t) \leq q$，对于一个（Knockoff 方法）估计假发现比例(t)作为对假发现比例(t)的估计。由于假发现比例(t)向上有偏，$\hat{t}_q$ 通常会收敛到某个值 t_q，使得假发现比例(t_q)，即实际的假发现比例，略小于名义水平 q。这本身通常会导致一定的检验功效损失，在与预言机进行诚实比较时需要加以考虑。尽管如此，差值预期会渐近地较小，尤其是当假发现比例(t)纳入了对 ϵ的一个相当好的估计时。

致谢

T本研究得到以色列科学基金会资助，资助编号为 2679/24。

参考文献

R. F. Barber 和 E. J. Cand`es。通过 Knockoff 方法控制错误发现率。2015 年 .

M. Bayati 和 A. Montanari。稠密图上的消息传递动态及其在压缩感知中的应用。*IEEE* 信息论汇刊，57(2): 764–785，2011。M. Bayati 和 A. Montanari。高斯矩阵的套索风险。*IEEE* 信息论汇刊，58(4):1997–2017，2012。doi: 10.1109/TIT.2011.2177385。J. C. Beer、H. J. Aizenstein、S. J. Anderson 和 R. T. Krafty。利用融合稀疏组套索纳入先验信息：将神经影像用于临床指标预测。《神经影像》，183:549–562，2018。doi: 10.1016/j.neuroimage.2018.08.040。Y. Benjamini。评论：微阵列、经验贝叶斯与双组模型。《统计科学》，23(1):23–28，2008。doi: 10.1214/07-STS236B。M. Bogdan、E. van den Berg、W. Su 和 E. J. Cand`es。通过排序 $\ell_1$ 范数进行统计估计与检验。*arXiv* 预印本 *arXiv:1310.1969*，2013a。M. Bogdan、E. van den Berg、W. Su 和 E. J. Cand`es。通过排序 ℓ_1 范数进行统计估计与检验的补充材料，2013b。可向作者索取或从期刊网站获取。E. Candes、Y. Fan、L. Janson 和 J. Lv。淘金：用于高维受控变量选择的“model-x” knockoffs。皇家统计学会会刊 *B* 辑：统计方法学，80(3):551–577，2018。B. Efron。大规模同时假设检验：零假设的选择。美国统计学会会刊，99(465):96–104，2004。B. Efron。微阵列、经验贝叶斯与双组模型。《统计科学》，23(1): 1–22，2008。B. Efron、R. Tibshirani、J. D. Storey 和 V. Tusher。微阵列实验的经验贝叶斯分析。美国统计学会会刊，96(456):1151–1160，2001。A. Gramfort、D. Strohmeier、G. Varoquaux、J.-B. Poline 和 B. Thirion。广义稀疏正则化及其在功能性磁共振成像脑解码中的应用。《神经影像》，56(2):456–465，2011。doi: 10.1016/j.neuroimage.2011.02.014。

R. Heller 和 S. Rosset。双组模型中假发现准则的最优控制。皇家统计学会会刊 *B* 辑：统计方法学，83(1):133–155，2021。A. Mousavi、A. Maleki 和 R. G. Baraniuk。Lasso 与近似消息传递的一致参数估计。统计学年鉴，46(1):119–148，2018。doi: 10.1214/17-AOS1544。H. Robbins。复合统计决策问题的渐近次极小极大解。载于第二届伯克利数学统计与概率研讨会论文集，第131–149页。加利福尼亚大学出版社，1951。J. A. Soloff、D. Xiang 和 W. Fithian。发现边界：在临界处控制局部错误发现率。统计学年鉴，52(2):580–601，2024。W. Su、M. Bogdan 和 E. J. Cand`es。假发现会在 Lasso路径早期出现。统计学年鉴，45(5):2133–2150，2017。doi: 10.1214/16-AOS1501。W. Sun 和 T. T. Cai。用于错误发现率控制的预言机与自适应复合决策规则。美国统计学会会刊，102(479):901–912，2007。R. Tibshirani。通过 Lasso 实现回归收缩与选择。皇家统计学会会刊：*B* 辑（方法学），58(1):267–288，1996。G. Varoquaux、A. Gramfort 和 B. Thirion。小样本脑图谱绘制：通过空间相关设计中的随机化与聚类实现稀疏恢复。*arXiv* 预印本 *arXiv:1206.6447*，2012。S. Wang、H. Weng 和 A. Maleki。哪种 Bridge 估计量最适合变量选择？统计学年鉴，48(5):2791–2823，2020。A. Weinstein。大规模推断中的置换不变问题。*arXiv* 预印本 *arXiv:2110.06250*，2021。A. Weinstein、R. Barber 和 E. Candes。基于 Lasso统计量的 knockoffs 的检验功效与预测分析。*arXiv* 预印本 *arXiv:1712.06465*，2017。A. Weinstein、W. J. Su、M. Bogdan、R. Foygel Barber 和 E. J. Cand`es。基于 ℓ_p 正则化统计量的 model-X knockoffs 的功效分析。统计学年鉴，51(3):1005–1029，2023。doi: 10.1214/23-AOS2274。D. Xiang、J. A. Soloff 和 W. Fithian。频率学派的局部错误发现率。《生物统计学》，页码 asaf083，2025。J. Xie、T. T. Cai、J. Maris 和 H. Li。相关数据的错误发现率最优控制。《统计学及其应用》，4(4):417，2011。O. Yamashita、M.-a. Sato、T. Yoshioka、F. Tong 和 Y. Kamitani。稀疏估计会自动选择与 fMRI 活动模式解码相关的体素。《神经影像》，42 (4):1414–1429，2008。doi: 10.1016/j.neuroimage.2008.05.050。

A 证明

在给出引理 4.1的证明之前，我们先给出另一个辅助结果。下面是 Bogdan 等人 [2013b], 针对当前设置略作修改的引理 2 版本。

引理 A.1。 设 Y是一个服从 $\mathcal{N}(\mu, \Sigma)$分布的 p维向量，其中 $\Sigma \geq \sigma^2 I$。对于任意 $\epsilon \in (0,\ 1)$，存在常数 $c = c(\epsilon) > 0$，使得对于任意 $h > 0$,

$$\mathbb{P}(at\ least\ \epsilon p\ components\ of\ Y\ are\ in\ (b, b+h)) \leq \left(\min\left(1, ch^{\epsilon}\sigma^{-\epsilon}\right)\right)^{p}.$$

引理 *A.1*的证明。 根据 Bogdan 等人 [2013b], 中的引理 1, 的方差（其中下标 $-i$ 表示除 i 之外的所有分量）大于或等于 σ^2。因此有

$$\mathbb{P}(Y_i \in (b, b+h) \mid Y_{(-i)}) \leq \min\left(1, \frac{h}{\sqrt{2\pi}\sigma}\right) \quad \text{a.s.}$$

因为正态密度函数被 $1/\sqrt{2\pi}$ 所界定。记 ξ_i 为与 Y 独立同分布的伯努利随机变量，且满足

$$\mathbb{P}(\xi_i = 1) = \tilde{p} = \min\left(1, \frac{h}{\sqrt{2\pi}\sigma}\right).$$

定义 $\zeta_i = \mathbf{1}(Y_i \in (b,\ \ b+h))$。由于

$$\mathbb{P}(\zeta_1 = 1 \mid \zeta_2, \ldots, \zeta_p) \overset{\text{a.s}}{\leq} \mathbb{P}(\xi_1 = 1 \mid \zeta_2, \ldots, \zeta_p),$$

我们得到:

$$\mathbb{P}(\zeta_1 + \zeta_2 + \cdots + \zeta_p \geq \varepsilon p) \leq \mathbb{P}(\xi_1 + \zeta_2 + \cdots + \zeta_p \geq \varepsilon p).$$

通过类似的论证，我们可以得出:

$$\mathbb{P}(\xi_1 + \xi_2 + \cdots + \xi_k + \zeta_{k+1} + \cdots + \zeta_p \geq \varepsilon p) \leq \mathbb{P}(\xi_1 + \xi_2 + \cdots + \xi_{k+1} + \zeta_{k+2} + \cdots + \zeta_p \geq \varepsilon p)$$

对于 $k = 1,\ \ldots\ldots,\ p-1$。因此,

$$\begin{aligned}
\mathbb{P}(\text{at least } \varepsilon p \text{ components of } Y \text{ are in } (b, b+h)) &= \mathbb{P}(\zeta_1 + \cdots + \zeta_p \geq \varepsilon p) \\
&\leq \mathbb{P}(\xi_1 + \zeta_2 + \cdots + \zeta_p \geq \varepsilon p) \\
&\leq \mathbb{P}(\xi_1 + \xi_2 + \zeta_3 + \cdots + \zeta_p \geq \varepsilon p) \\
&\ \ \vdots \\
&\leq \mathbb{P}(\xi_1 + \xi_2 + \cdots + \xi_p \geq \varepsilon p) \\
&= \mathbb{P}\left(\sum_{i=1}^{p} \xi_i \geq \varepsilon p\right).
\end{aligned}$$

因此,

$$\mathbb{P}\left(\sum_{i=1}^{p} \xi_i \geq \varepsilon p\right) \leq \sum_{i \geq \varepsilon p} \binom{p}{i} \tilde{p}^i \leq \tilde{p}^{\varepsilon p} \sum_{i \geq \varepsilon p} \binom{p}{i} \leq \tilde{p}^{\varepsilon p} \cdot 2^p \leq \left(\frac{2^{1+\varepsilon/2}}{\pi^{\varepsilon/2}}\right)^{p} h^{\varepsilon p} \sigma^{-\varepsilon p}.$$

□

引理 A.2。

$$\lim_{p\to\infty} \frac{R^*(t;\lambda)}{p} = \lim_{p\to\infty} \sum_{m=1}^{M} \left(\frac{1}{p} \sum_{i=1}^{p} t_{R_m}(\hat{\beta}_i, \beta_i)\right) = \sum_{m=1}^{M} \mathbb{E}\left[t_{R_m}\left(\eta_{\alpha\tau}(\Pi + \tau Z), \Pi\right)\right].$$

引理 *A.2* 的证明。 注意， $t_{R_m}(x,y)$ 不是伪 Lipschitz 函数，因此我们无法应用 Bayati 和 Montanari [2012] 的定理 1.5 来绕过 $t_{R_m}(x,y)$ 的不连续性，定义函数

$$t_{R_{m,h}}(x,y) = \begin{cases} 1, & a_m(t) < x < b_m(t) \\ 0, & x \geq b_m(t) \text{ or } x \leq a_m(t) \\ 1 - \frac{x - b_m(t)}{h}, & b_m(t) < x < b_m(t) + h \\ 1 + \frac{x - a_m(t)}{h}, & a_m(t) - h < x < a_m(t) \end{cases},$$

它是伪利普希茨函数。显然， $t_{R_{m,\ h}}(x,\ y) \to t_{R_m}(x,\ y)$ 随着 $h \to 0$， 对于固定的 x， y。根据 Bayati 和 Montanari [2012],的定理 1.5，我们有

$$\lim_{p\to\infty} \frac{1}{p} \sum_{m=1}^{M} \sum_{i=1}^{p} t_{R_{m,h}}(\hat{\beta}_i, \beta_i) \xrightarrow{P} \sum_{m=1}^{M} \mathbb{E}\left[t_{R_{m,h}}(\eta_{\alpha\tau}(\Pi + \tau Z), \Pi)\right].$$

现在，对于所有 $\epsilon > 0$,

$$\begin{aligned} &\mathbb{P}\left(\left|\frac{1}{p}\sum_{i=1}^{p} t_{R_{m,h}}(\hat{\beta}_i,\beta_i) - \frac{1}{p}\sum_{i=1}^{p} t_{R_m}(\hat{\beta}_i,\beta_i)\right| > \epsilon\right) \leq \\ &\quad \mathbb{P}\left(\frac{1}{p}\sum_{i=1}^{p}\left|t_{R_{m,h}}(\hat{\beta}_i,\beta_i) - t_{R_m}(\hat{\beta}_i,\beta_i)\right| > \epsilon\right) \leq \\ &\quad\quad \mathbb{P}\left(\frac{1}{p}\sum_{i=1}^{p} \mathbf{1}\{b_m < \hat{\beta}_i < b_m + h\} > \epsilon\right) + \mathbb{P}\left(\frac{1}{p}\sum_{i=1}^{p} \mathbf{1}\{a_m - h < \hat{\beta}_i < a_m\} > \epsilon\right) \end{aligned}$$

如果我们证明:

$$\lim_{h\to 0}\lim_{p\to\infty} \sup \mathbb{P}\left(\frac{1}{p}\sum_{i=1}^{p} \mathbf{1}\{b_m < \hat{\beta}_i < b_m + h\} > \epsilon\right) = 0 \tag{39}$$

and

$$\lim_{h\to 0}\lim_{p\to\infty} \sup \mathbb{P}\left(\frac{1}{p}\sum_{i=1}^{p} \mathbf{1}\{a_m - h < \hat{\beta}_i < a_m\} > \epsilon\right) = 0, \tag{40}$$

应用控制收敛定理，可得:

$$\begin{aligned} \lim_{p\to\infty} \frac{1}{p} \sum_{m=1}^{M} \sum_{i=1}^{p} t_{R_m}(\hat{\beta}_i, \beta_i) &\xrightarrow{P} \sum_{m=1}^{M} \lim_{h\to 0} \mathbb{E}\left[t_{R_{m,h}}(\eta_{\alpha\tau}(\Pi + \tau Z), \Pi)\right] \\ &= \sum_{m=1}^{M} \mathbb{E}\left[t_{R_m}(\eta_{\alpha\tau}(\Pi + \tau Z), \Pi)\right]. \end{aligned}$$

为证明 (39)，令 A 为 Lasso 解的活跃集。根据 KKT 条件 ns,

$$\boldsymbol{X}_A^T(Y - \boldsymbol{X}_A\hat{\beta}_A) = \tilde{\lambda}_A \quad \text{with } \tilde{\lambda}_A = \lambda \cdot \operatorname{sign}(\hat{\beta}_A).$$

由于 $|A| \leq n$, $\boldsymbol{X}_A^T\boldsymbol{X}_A$ 可逆，因此

$$\hat{\beta}_A = (\boldsymbol{X}_A^T\boldsymbol{X}_A)^{-1}(\boldsymbol{X}_A^T Y - \tilde{\lambda}_A).$$

˜现在，对于任意满足 $|D| \leq n$ 的子集 $D \subset \{1, 2, \ldots, p\}$，以及任意长度为 $|D|$ 且各分量均满足 $\pm\lambda$ 的 λ，定义

$$\hat{\beta}_D^{\tilde{\lambda}} = (\boldsymbol{X}_D^T\boldsymbol{X}_D)^{-1}(\boldsymbol{X}_D^T\boldsymbol{X}\beta - \tilde{\lambda}) + (\boldsymbol{X}_D^T\boldsymbol{X}_D)^{-1}\boldsymbol{X}_D^T\zeta.$$

根据 Bogdan 等人 [2013b] , 存在 $\rho > 0$ 使得 $\mathbb{P}(|A| < \rho p) = o(1)$。对于任意 ϵ', $\epsilon > 0$, 当 n 足够大时, p, 有

$$\mathbb{P}\left(\frac{1}{p}\sum_{i=1}^{p}\mathbf{1}\{b_m < \hat{\beta}_i < b_m + h\} > \epsilon\right) \leq \sum_{|D|=\rho p}^{min(p,n)} \sum_{|\tilde{\lambda}|=|D|} \Bigg[$$

$$\mathbb{P}\left(\frac{1}{p}\sum_{i=1}^{|D|}\mathbf{1}\{b_m < \hat{\beta}_D^{\tilde{\lambda}}(i) < b_m + h\},\ \sigma_{\max}(\boldsymbol{X}) < 1 + \delta^{-1/2} + \epsilon'\right) + o(1) + \mathbb{P}(|A| < \rho p)\Bigg]$$

由于 $\hat{}\beta_D^{\tilde{\lambda}}$ 在这些条件下服从协方差有界的多元正态分布，应用引理A.1 可得

$$\mathbb{P}\left(\frac{1}{p}\sum_{i=1}^{|D|}\mathbf{1}\{b_m < \hat{\beta}_D^{\tilde{\lambda}}(i) < b_m + h\} > \epsilon\right) \leq (ch^{\epsilon}\sigma^{-\epsilon})^{\rho p}.$$

因此,

$$\mathbb{P}\left(\frac{1}{p}\sum_{i=1}^{p}\mathbf{1}\{b_m < \hat{\beta}_i < b_m + h\} > \epsilon\right) \leq (4c^{\rho}h^{\rho\epsilon}\sigma^{-\rho\epsilon})^{p} + o(1) + o(1).$$

这意味着

$$\lim_{h\to 0}\lim_{p\to\infty}\sup\mathbb{P}\left(\frac{1}{p}\sum_{i=1}^{p}\mathbf{1}\{b_m < \hat{\beta}_i < b_m + h\} > \epsilon\right) = 0$$

类似地，我们可以证明（40），从而得到

$$\lim_{p\to\infty}\frac{R^*(t;\lambda)}{p} = \sum_{m=1}^{M}\mathbb{E}\left[t_{R_m}(\eta_{\alpha\tau}(\Pi + \tau Z), \Pi)\right].$$

□

引理 A.3。

$$\lim_{p\to\infty}\frac{V^*(t;\lambda)}{p} = \lim_{p\to\infty}\sum_{m=1}^{M}\left(\frac{1}{p}\sum_{i=1}^{p}t_{V_m}(\hat{\beta}_i, \beta_i)\right) \xrightarrow{P} \sum_{m=1}^{M}\mathbb{E}[t_{V_m}(\eta_{\alpha\tau}(\Pi + \tau Z), \Pi)].$$

引理 *A.3* 的证明。证明与引理A.2 的证明类似 □

引理的证明 *4.1*。 由引理A.2 和A.3可得

$$\mathrm{FDP}^*(t;\lambda) = \frac{V^*(t;\lambda)}{R^*(t;\lambda)\vee 1} \xrightarrow{P} \frac{\mathbb{E}\left[\sum_{m=1}^{M}t_{V_m}(\eta_{\alpha\tau}(\Pi + \tau Z), \Pi)\right]}{\mathbb{E}\left[\sum_{m=1}^{M}t_{R_m}(\eta_{\alpha\tau}(\Pi + \tau Z), \Pi)\right]}$$

$$= \frac{\sum_{m=1}^{M}\mathbb{E}\left[t_{V_m}(\eta_{\alpha\tau}(\Pi + \tau Z), \Pi)\right]}{\sum_{m=1}^{M}\mathbb{E}\left[t_{R_m}(\eta_{\alpha\tau}(\Pi + \tau Z), \Pi)\right]}.$$

□

定理的证明 *4.2*。 首先证明陈述（31）。根据假设1，可得

$$\frac{q_0(x)}{q(x)} \le t \iff x \in \bigcup_{m=1}^{M} (a_m(t), b_m(t)), \tag{41}$$

其中，区间（a_m（t） b_m（t））互不相交。因此我们得到

$$\mathrm{FDP}^*(t;\lambda) = \frac{\sum_{i=1}^{p} \mathbf{1}\left\{\hat{\beta}_i \in \bigcup_{m=1}^{M}(a_m(t), b_m(t)), \beta_i = 0\right\}}{\sum_{i=1}^{p} \mathbf{1}\left\{\hat{\beta}_i \in \bigcup_{m=1}^{M}(a_m(t), b_m(t))\right\}}.$$

我们需要证明

$$\mathrm{FDP}^*(t;\lambda) \xrightarrow{P} \mathbb{P}\left(\Pi = 0 \mid \frac{q_0(\eta_{\alpha\tau}(\Pi + \tau Z))}{q(\eta_{\alpha\tau}(\Pi + \tau Z))} \le t\right).$$

注意

$$\mathrm{FDP}^*(t;\lambda) = \frac{\sum_{m=1}^{M}\sum_{i=1}^{p} \mathbf{1}\left\{\hat{\beta}_i \in (a_m(t), b_m(t)), \beta_i = 0\right\}}{\sum_{m=1}^{M}\sum_{i=1}^{p} \mathbf{1}\left\{\hat{\beta}_i \in (a_m(t), b_m(t))\right\}}.$$

我们可以将其写为

$$\frac{\sum_{m=1}^{M} \frac{1}{p}\sum_{i=1}^{p} \mathbf{1}\{\hat{\beta}_i \in (a_m(t), b_m(t)), \beta_i = 0\}}{\sum_{m=1}^{M} \frac{1}{p}\sum_{i=1}^{p} \mathbf{1}\{\hat{\beta}_i \in (a_m(t), b_m(t))\}}$$

$$= \mathrm{FDP}^{\mathrm{OR}}(t;\lambda)$$

根据引理4.1,

$$\xrightarrow{P} \frac{\sum_{m=1}^{M} \mathbb{E}\left[\mathbf{1}\{\eta_{\alpha\tau}(\Pi + \tau Z) \in (a_m(t), b_m(t)), \Pi = 0\}\right]}{\sum_{m=1}^{M} \mathbb{E}\left[\mathbf{1}\{\eta_{\alpha\tau}(\Pi + \tau Z) \in (a_m(t), b_m(t))\}\right]}$$

$$= \frac{\mathbb{E}\left[\sum_{m=1}^{M} \mathbf{1}\{\eta_{\alpha\tau}(\Pi + \tau Z) \in (a_m(t), b_m(t)), \Pi = 0\}\right]}{\mathbb{E}\left[\sum_{m=1}^{M} \mathbf{1}\{\eta_{\alpha\tau}(\Pi + \tau Z) \in (a_m(t), b_m(t))\}\right]}$$

$$= \frac{\mathbb{E}\left[\mathbf{1}\left\{\eta_{\alpha\tau}(\Pi + \tau Z) \in \bigcup_{m=1}^{M}(a_m(t), b_m(t)), \Pi = 0\right\}\right]}{\mathbb{E}\left[\mathbf{1}\left\{\eta_{\alpha\tau}(\Pi + \tau Z) \in \bigcup_{m=1}^{M}(a_m(t), b_m(t))\right\}\right]},$$

其中最后一个等号利用了并集中各区间互不相交这一事实。根据假设（41），我们得到以下等式

$$= \frac{\mathbb{E}\left[\mathbf{1}\left\{\frac{q_0(\eta_{\alpha\tau}(\Pi+\tau Z))}{q(\eta_{\alpha\tau}(\Pi+\tau Z))} \le t, \Pi = 0\right\}\right]}{\mathbb{E}\left[\mathbf{1}\left\{\frac{q_0(\eta_{\alpha\tau}(\Pi+\tau Z))}{q(\eta_{\alpha\tau}(\Pi+\tau Z))} \le t\right\}\right]}$$

$$= \mathbb{P}\left(\Pi = 0 \mid \frac{q_0(\eta_{\alpha\tau}(\Pi + \tau Z))}{q(\eta_{\alpha\tau}(\Pi + \tau Z))} \le t\right)$$

$$\equiv \mathrm{fdp}^*(t;\lambda).$$

下面证明命题（32）。根据引理 A.2 和引理 A.3 以及假设（41），我们有

$$
\begin{aligned}
\mathrm{TPP}^*(t;\lambda) &= \frac{\frac{1}{p}(R^*(t;\lambda) - V^*(t;\lambda))}{\frac{1}{p}\sum_{i=1}^{p}\mathbf{1}(\beta_i \neq 0)} \\
&\xrightarrow{P} \frac{\sum_{m=1}^{M}\mathbb{E}\left[t_{R_m}(\eta_{\alpha\tau}(\Pi+\tau Z),\Pi)\right] - \sum_{m=1}^{M}\mathbb{E}\left[t_{V_m}(\eta_{\alpha\tau}(\Pi+\tau Z),\Pi)\right]}{\mathbb{E}\left[\mathbf{1}(\Pi\neq 0)\right]} \\
&= \frac{\mathbb{E}\left[\mathbf{1}\left\{\frac{q_0(\eta_{\alpha\tau}(\Pi+\tau Z))}{q(\eta_{\alpha\tau}(\Pi+\tau Z))} \leq t\right\} - \mathbf{1}\left\{\frac{q_0(\eta_{\alpha\tau}(\Pi+\tau Z))}{q(\eta_{\alpha\tau}(\pi+\tau Z))} \leq t, \Pi = 0\right\}\right]}{\mathbb{E}\left[\mathbf{1}(\Pi\neq 0)\right]} \\
&= \frac{\mathbb{E}\left[\mathbf{1}\left\{\frac{q_0(\eta_{\alpha\tau}(\Pi+\tau Z))}{q(\eta_{\alpha\tau}(\Pi+\tau Z))} \leq t, \Pi \neq 0\right\}\right]}{\mathbb{E}\left[\mathbf{1}(\Pi\neq 0)\right]} \\
&= \mathbb{P}\Big(\frac{q_0(\eta_{\alpha\tau}(\Pi+\tau Z))}{q(\eta_{\alpha\tau}(\Pi+\tau Z))} \leq t \mid \Pi \neq 0\Big) \\
&= \mathbb{P}\Big(\frac{q_0(\eta_{\alpha\tau}(\Pi_1+\tau Z))}{q(\eta_{\alpha\tau}(\Pi_1+\tau Z))} \leq t\Big) \\
&\equiv \mathrm{tpp}^*(t;\lambda).
\end{aligned}
$$

□

定理 *4.3* 的证明。 在假设 1 以及对 φ 所作假设的条件下，问题归结为证明以下结论。令 $X = \eta_{\alpha\tau}(\Pi + \tau Z)$，其中 Π 如（7）所示，并考虑检验 H_0： $\Pi = 0$ 对 H_1： $\Pi \neq 0$。由于

$$
\frac{q_0(x)}{q(x)} = \frac{q_0(x)}{(1-\epsilon)q_0(x) + \epsilon q_1(x)},
$$

对于 $\frac{q_0(x)}{q(x)} \leq t^*$，拒绝检验等价于在 $\frac{q_1(x)}{q_0(x)} > s(t^*)$ 时拒绝，其中相应的临界值为 $s(t^*)$。因此，定义如下（似然比）检验

$$
\mathcal{R}^*_\lambda(x) = \mathbf{1}\left\{\frac{q_1(x)}{q_0(x)} > s^*\right\},
$$

其中， q_0 是 $\eta_{\alpha\tau}(\tau Z)||Z| > \alpha$ 的未归一化密度， q_1 是 $\eta_{\alpha\tau}(\Pi_1 + \tau Z)||\Pi_1 + \tau Z| > \alpha\tau$ 的未归一化密度；为简化记号，记 $s^* =: s(t^*)$。现在也定义检验

$$
\mathcal{R}^\varphi_\lambda(x) = \mathbf{1}\left\{\varphi(x) > t\right\}.
$$

对于某个其他的 φ。我们需要证明，若常数 s^* 和 t 满足

$$
\mathbb{P}(\mathcal{R}^*_\lambda(X) = 1 \mid \Pi \neq 0) = \mathbb{P}(\mathcal{R}^\varphi_\lambda(X) = 1 \mid \Pi \neq 0) = \xi,
$$

then

$$
\mathbb{P}(\Pi = 0 | \mathcal{R}^*_\lambda(X) = 1) \leq \mathbb{P}(\Pi = 0 | \mathcal{R}^\varphi_\lambda(X) = 1). \tag{42}
$$

由于检验功效固定为 ξ，证明（42）等价于证明频率学派（固定 Π）检验问题中的第一类错误满足

$$
\mathbb{P}(\mathcal{R}^*_\lambda(X) = 1 | \Pi = 0) \leq \mathbb{P}(\mathcal{R}^\varphi_\lambda(X) = 1 | \Pi = 0).
$$

因此，只需调整奈曼-皮尔逊引理的证明：不再证明在具有相同第一类错误概率的所有检验中，似然比检验具有最大的检验功效，而是证明在具有相同检验功效的所有检验中，它使第一类错误的概率最小。

考虑积分

$$\int \big(\mathcal{R}^*_\lambda(x) - \mathcal{R}^\varphi_\lambda(x)\big)\big(q_1(x) - s^*\, q_0(x)\big)\, dx.$$

根据 $\mathcal{R}^*_\lambda(x)$ 的构造，当 $q_1(x) - s^*\, q_0(x) > 0$ 时，$\mathcal{R}^*_\lambda(x) = 1$；而当 $q_1(x) - s^*\, q_0(x) < 0$ 时，$\mathcal{R}^*_\lambda(x) = 0$。由于 $\mathcal{R}^\varphi_\lambda(x) \in \{0, 1\}$，我们得到该积分非负。展开积分可得

$$\begin{aligned}
0 &\le \int \big(\mathcal{R}^*_\lambda(x) - \mathcal{R}_\lambda(x)\big) q_1(x)\, dx - s^* \int \big(\mathcal{R}^*_\lambda(x) - \mathcal{R}^\varphi_\lambda(x)\big) q_0(x)\, dx \\
&= \mathbb{P}(\mathcal{R}^*_\lambda(X) = 1 \mid \Pi \neq 0) - \mathbb{P}(\mathcal{R}^\varphi_\lambda(X) = 1 \mid \Pi \neq 0) \\
&\quad - s^* \big(\mathbb{P}(\mathcal{R}^*_\lambda(X) = 1 \mid \Pi = 0) - \mathbb{P}(\mathcal{R}^\varphi_\lambda(X) = 1 \mid \Pi = 0)\big).
\end{aligned}$$

利用两个检验具有相同检验功效 ξ 这一假设，我们得到

$$\mathbb{P}(\mathcal{R}^*_\lambda(X) = 1 \mid \Pi = 0) \le \mathbb{P}(\mathcal{R}^\varphi_\lambda(X) = 1 \mid \Pi = 0).$$

□

为证明定理 4.8，必须证明估计量 $\hat{}q(x)$ 一致收敛于真实密度 $q(x)$。为此，我们首先证明 $\hat{}q(x)$ 依概率逐点收敛于 $q(x)$。下文证明的引理 4.4 为这一中间结果奠定了基础。

˜引理 *4.4* 的证明。令 $\psi(x, y) = \mathbf{1}\{x \neq 0\}\, \psi(x, y)$，我们得到

$$\frac{1}{p}\sum_{i=1}^{p} \mathbf{1}\{\hat{\beta}_i \neq 0\}\, \psi(\hat{\beta}_i, \beta_i) = \frac{1}{p}\sum_{i=1}^{p} \tilde{\psi}(\hat{\beta}_i, \beta_i).$$

˜注意到 ψ 不是伪利普希茨的，因此我们不能应用定理2（Bayati 和 Montanari[2012]）。我们将证明，尽管该定理不能直接应用，引理4.4 仍然正确。因此，定义一个伪 Lipschitz 函数

$$\tilde{\psi}_{t'}(x, y) = \left(1 - Q\left(\frac{x}{t'}\right)\right) \psi(x, y),$$

˜其中 $Q(x) = \max(1 - |x|, 0)$，且 $t' > 0$。由于 $\psi_{t'}$ 是伪利普希茨函数，根据 Bayati 和 Montanari的定理 [2012] 2，我们有

$$\lim_{p\to\infty} \frac{1}{p}\sum_{i=1}^{p} \tilde{\psi}_{t'}(\hat{\beta}_i, \beta_i) = \mathbb{E}[\tilde{\psi}_{t'}(\eta_{\alpha\tau}(\Pi + \tau Z), \Pi)].$$

Now,

$$\begin{aligned}
&\left|\tilde{\psi}_{t'}(x, y) - \tilde{\psi}(x, y)\right| \\
&= \left|\left(1 - Q\left(\frac{x}{t'}\right) - \mathbf{1}\{x \neq 0\}\right) \psi(x, y)\right| \\
&\le \mathbf{1}\left\{0 < |x| < t'\right\} |\psi(x, y)| \\
&\le \mathbf{1}\left\{0 < |x| < t'\right\} \sup_{x\in\mathbb{R}} |\psi(x, y)| \\
&= \mathbf{1}\left\{0 < |x| < t'\right\} \|\psi\|_\infty.
\end{aligned}$$

现在我们来界定该概率:

$$\begin{aligned}
&\mathbb{P}\left(\left|\frac{1}{p}\sum_{i=1}^{p}\tilde{\psi}_{t'}(\hat{\beta}_i,\beta_i)-\frac{1}{p}\sum_{i=1}^{p}\tilde{\psi}(\hat{\beta}_i,\beta_i)\right|>\epsilon\right)\\
&=\mathbb{P}\left(\frac{1}{p}\left|\sum_{i=1}^{p}\left(\tilde{\psi}_{t'}(\hat{\beta}_i,\beta_i)-\tilde{\psi}(\hat{\beta}_i,\beta_i)\right)\right|>\epsilon\right)\\
&\leq\mathbb{P}\left(\frac{1}{p}\sum_{i=1}^{p}\left|\tilde{\psi}_{t'}(\hat{\beta}_i,\beta_i)-\tilde{\psi}(\hat{\beta}_i,\beta_i)\right|>\epsilon\right)\\
&\leq\mathbb{P}\left(\frac{1}{p}\sum_{i=1}^{p}\mathbf{1}\left\{0<|\hat{\beta}_i|<t'\right\}\|\psi\|_\infty>\epsilon\right)\\
&=\mathbb{P}\left(\frac{1}{p}\sum_{i=1}^{p}\mathbf{1}\left\{0<|\hat{\beta}_i|<t'\right\}>\tilde{\epsilon}\right),\qquad\tilde{\epsilon}=\frac{\epsilon}{\|\psi\|_\infty}.
\end{aligned}$$

在Bogdan 等人 [2013b] 证明了

$$\lim_{t'\to 0}\limsup_{p\to\infty}\mathbb{P}\left(\frac{1}{p}\sum_{i=1}^{p}\mathbf{1}\left\{0<|\hat{\beta}_i|<t'\right\}>\tilde{\epsilon}\right)=0,$$

因此，对于任意 $\tilde{\epsilon}>0$,

$$\begin{aligned}
\lim_{p\to\infty}\frac{1}{p}\sum_{i=1}^{p}\tilde{\psi}(\hat{\beta}_i,\beta_i)&=\lim_{t'\to 0}\lim_{p\to\infty}\frac{1}{p}\sum_{i=1}^{p}\tilde{\psi}_{t'}(\hat{\beta}_i,\beta_i)\\
&=\lim_{t'\to 0}\mathbb{E}\left[\tilde{\psi}_{t'}(\eta_{\alpha\tau}(\Pi+\tau Z),\Pi)\right]\\
&=\mathbb{E}[\tilde{\psi}(\eta_{\alpha\tau}(\Pi+\tau Z),\Pi)],
\end{aligned}$$

~其中，最后一个等式是对 $\psi_{t'}\to\tilde{}\psi$ 应用控制收敛定理得出的。

□

命题 *4.5* 的证明。 令

$$k(\hat{\beta}_i,x)=\frac{1}{h}\phi\left(\frac{x-\hat{\beta}_i}{h}\right)=\frac{1}{\sqrt{2\pi h^2}}\exp\left(-\frac{1}{2}\cdot\frac{(\hat{\beta}_i-x)^2}{h^2}\right),\quad h>0\text{ (fixed)},\ i=1,...,p.$$

定义:

$$\psi(\hat{\beta}_i,\beta_i)=k(\hat{\beta}_i,x),\quad\text{where } x\neq 0 \text{ is fixed.}$$

则 ψ： $\mathbb{R}\times\mathbb{R}\to\mathbb{R}$ 是一个伪 Lipschitz 函数。现在定义

$$\tilde{\psi}(\hat{\beta}_i,\beta_i)=\mathbf{1}\left\{\hat{\beta}_i\neq 0\right\}\cdot\psi(\hat{\beta}_i,\beta_i).$$

于是得到

$$\frac{1}{p}\sum_{i=1}^{p}\mathbf{1}\left\{\hat{\beta}_i\neq 0\right\}\cdot\psi(\hat{\beta}_i,\beta_i)=\frac{1}{p}\sum_{i=1}^{p}\tilde{\psi}(\hat{\beta}_i,\beta_i).$$

定义

$$\hat{q}(x)=\hat{q}_h(x)=\frac{1}{p}\sum_{i=1}^{p}\mathbf{1}\left\{\hat{\beta}_i\neq 0\right\}\cdot\frac{1}{h}\phi\left(\frac{x-\hat{\beta}_i}{h}\right)$$

So,

$$\begin{aligned}\hat{q}(x) &= \frac{1}{p}\sum_{i=1}^{p}\mathbf{1}\left\{\hat{\beta}_i \neq 0\right\}\cdot k(\hat{\beta}_i, x)\\ &= \frac{1}{p}\sum_{i=1}^{p}\mathbf{1}\left\{\hat{\beta}_i \neq 0\right\}\cdot \psi(\hat{\beta}_i, \beta_i)\\ &\xrightarrow{P} \mathbb{E}[\tilde{\psi}(\eta_{\alpha\tau}(\Pi+\tau Z), \Pi)] \quad \text{(by Lemma 4.4)}\\ &= \mathbb{E}\left[\mathbf{1}\left\{\eta_{\alpha\tau}(\Pi+\tau Z)\neq 0\right\}\cdot\psi\left(\eta_{\alpha\tau}(\Pi+\tau Z), \Pi\right)\right]\end{aligned}$$

现在考虑

$$\mathbb{E}\left[\mathbf{1}\left\{\eta_{\alpha\tau}(\Pi+\tau Z)\neq 0\right\}\cdot\psi\left(\eta_{\alpha\tau}(\Pi+\tau Z), \Pi\right)\right].$$

设 $Y \sim \eta_{\alpha\tau}(\Pi+\tau Z)$ 的密度为 $f(y) = (1-w)\delta_0(y) + wg(y) = (1-w)\delta_0(y) + q(y)$。则

$$\begin{aligned}&\mathbb{E}\left[\mathbf{1}\left\{\eta_{\alpha\tau}(\Pi+\tau Z)\neq 0\right\}\cdot\psi\left(\eta_{\alpha\tau}(\Pi+\tau Z), \Pi\right)\right]\\ &= \int_{-\infty}^{\infty}\mathbf{1}(y\neq 0)\cdot\frac{1}{h}\phi\left(\frac{x-y}{h}\right)f(y)\,dy\\ &= \int_{-\infty}^{\infty}\frac{1}{h}\phi\left(\frac{x-y}{h}\right)q(y)\,dy\\ &= \int_{-\infty}^{0}\frac{1}{h}\phi\left(\frac{x-y}{h}\right)q(y)\,dy + \int_{0}^{\infty}\frac{1}{h}\phi\left(\frac{x-y}{h}\right)q(y)\,dy\end{aligned}$$

令 $u = \frac{x-y}{h} \Rightarrow y = x - uh$ 进行变量替换后，我们可以继续

$$= \int_{-\infty}^{\frac{x}{h}}\phi(u)\cdot q(x-uh)\,du + \int_{\frac{x}{h}}^{\infty}\phi(u)\cdot q(x-uh)\,du$$

现在对 $q(x-uh)$ 在 x 处作泰勒展开，假设 q 至少有一个导数，

$$q(x-hu) = q(x) - hu\cdot q'(x) + o(h).$$

由于

$$\lim_{h\to 0^+} q(x-hu) = q(x) \; and \lim_{h\to 0^-} q(x-hu) = q(x),$$

我们得到，

$$\lim_{h\to 0} q(x-hu) = q(x)$$

另外，假设 $\sup_{x\in\mathbb{R}\setminus\{0\}} |q(x)| = M < \infty$，则根据控制收敛定理，

$$\begin{aligned}\lim_{h\to 0}\lim_{p\to\infty}\hat{q}(x) &= \lim_{h\to 0}\left(\lim_{p\to\infty}\frac{1}{p}\sum_{i=1}^{p}\mathbf{1}(\hat{\beta}_i\neq 0)\cdot\frac{1}{h}\phi\left(\frac{x-\hat{\beta}_i}{h}\right)\right)\\ &= \lim_{h\to 0}\mathbb{E}\left[\mathbf{1}\left\{\eta_{\alpha\tau}(\Pi+\tau Z)\neq 0\right\}\cdot\psi\left(\eta_{\alpha\tau}(\Pi+\tau Z), \Pi\right)\right] \quad \text{(by Lemma 4.4)}\\ &= q(x),\end{aligned}$$

这里我们用到了 $\int\phi(u)\,du = 1$ 这一事实。

□

命题 *4.6* 的证明。 我们要证明

$$\sup_{x\in[\Delta,\mathcal{L}]} |\hat{q}(x) - q(x)| \xrightarrow{P} 0.$$

定义等距点 $\Delta = x_0 < x_1 < x_2 < \cdots < x_l = \mathcal{L}$，其中

$$\Delta' = x_{i+1} - x_i = \frac{1}{l}, \quad i = 0, 1, \ldots, l-1,$$

其中 l 将在下文指定。
根据假设， 2 $q(x)$ 在 $[\Delta, \mathcal{L}]$ 上一致连续，因此对任意 $\epsilon' > 0$，存在充分大的 l，使得

$$|q(x) - q(x_i)| < \frac{\epsilon'}{2} \tag{43}$$

对所有满足 $|x - x_i| \le \Delta' = \frac{1}{l}$ 的 x， $x_i \in [\Delta, \mathcal{L}]$。
根据命题 4.5，我们知道，对固定的 $q(x) \xrightarrow{P} q(x)$ 逐点成立，即 $x \in [\Delta, \mathcal{L}]$,

$$|\hat{q}(x) - q(x)| \xrightarrow{P} 0.$$

由并集界可得

$$\max_{0\le i\le l} |\hat{q}(x_i) - q(x_i)| \xrightarrow{P} 0. \tag{44}$$

也就是说，对任意 $\epsilon > 0$,

$$\begin{aligned}
&\mathbb{P}\left(\max_{0\le i\le l} |\hat{q}(x_i) - q(x_i)| > \epsilon\right) \\
&= \mathbb{P}\left(\bigcup_{0\le i\le l} \{|\hat{q}(x_i) - q(x_i)| > \epsilon\}\right) \\
&\le \sum_{i=0}^{l} \mathbb{P}\left(|\hat{q}(x_i) - q(x_i)| > \epsilon\right) \xrightarrow{P} 0.
\end{aligned}$$

现在我们要证明，对于充分大（但固定的） l,

$$\max_{0\le i<l} \sup_{x_i\le x\le x_{i+1}} |\hat{q}(x) - \hat{q}(x_i)| \le \frac{\epsilon'}{2}.$$

对于任意 $x \in [\Delta, \mathcal{L}]$，都存在 i 使得 $x \in [x_i, x_{i+1}]$，因此 $|x - x_i| \le \Delta' = \frac{1}{l}$。接下来,

$$\begin{aligned}
|\hat{q}(x) - \hat{q}(x_i)| &= \left|\frac{1}{ph}\sum_{j=1}^{p} \phi\left(\frac{x - \hat{\beta}_j}{h}\right) \mathbf{1}\{\hat{\beta}_j \ne 0\} - \frac{1}{ph}\sum_{j=1}^{p} \phi\left(\frac{x_i - \hat{\beta}_j}{h}\right) \mathbf{1}\{\hat{\beta}_j \ne 0\}\right| \\
&= \left|\frac{1}{ph}\sum_{j=1}^{p} \left[\phi\left(\frac{x - \hat{\beta}_j}{h}\right) - \phi\left(\frac{x_i - \hat{\beta}_j}{h}\right)\right] \mathbf{1}\{\hat{\beta}_j \ne 0\}\right| \\
&\le \frac{1}{ph}\sum_{j=1}^{p} \left|\phi\left(\frac{x - \hat{\beta}_j}{h}\right) - \phi\left(\frac{x_i - \hat{\beta}_j}{h}\right)\right| \\
&\le \frac{1}{ph}\sum_{j=1}^{p} L\frac{|x - x_i|}{h} \le \frac{L\Delta'}{h^2},
\end{aligned} \tag{45}$$

其中 L 是连续 Lipschitz 核函数 ϕ 的 Lipschitz 常数。现在，对于任意 $x \in [\Delta, \mathcal{L}]$，都存在 i 使得 $x \in [x_i, x_{i+1}]$，并且我们有

$$\begin{aligned} |\hat{q}(x) - q(x)| &= |\hat{q}(x) - \hat{q}(x_i) + \hat{q}(x_i) - q(x_i) + q(x_i) - q(x)| \\ &\leq |\hat{q}(x) - \hat{q}(x_i)| + |\hat{q}(x_i) - q(x_i)| + |q(x_i) - q(x)| \\ &\overset{(43),(45)}{\leq} \frac{L\Delta'}{h^2} + \max_{0 \leq i \leq l} |\hat{q}(x_i) - q(x_i)| + \frac{\epsilon'}{2}. \end{aligned} \tag{46}$$

如果我们选择 l 足够大，使得

$$\Delta' = \frac{1}{l} \leq \frac{\epsilon' h^2}{2L},$$

那么该项 $\frac{L\Delta'}{h^2} \leq \frac{\epsilon'}{2}$。因此，式（46）中的界变为

$$|\hat{q}(x) - q(x)| \leq \frac{\epsilon'}{2} + \max_{0 \leq i \leq l} |\hat{q}(x_i) - q(x_i)| + \frac{\epsilon'}{2}. \tag{47}$$

根据式（43）、式（44）和式（47），并且由于 $\epsilon' > 0$ 可以任意小，我们得出

$$\sup_{x \in [\Delta, \mathcal{L}]} |\hat{q}(x) - q(x)| \xrightarrow{P} 0.$$

□

引理 A.4。 设 $\hat{}q$ 在 $[-\mathcal{L}, -\Delta]$ 上以及在 $[\Delta, \mathcal{L}]$上依概率一致收敛于 q。则 $\hat{}q \to q$ 在 $[-\mathcal{L}, -\Delta] \cup [\Delta, \mathcal{L}]$上依概率一致。

引理 *A.4*的证明。 由于 $\hat{}q \to q$ 在 $[-\mathcal{L}, -\Delta]$ 上以及在 $[\Delta, \mathcal{L}]$ 上依概率一致，我们发现，对于任意 $\epsilon > 0$，存在 P_1，使得对于所有 $p \geq P_1$，

$$\sup_{x \in [-\mathcal{L}, -\Delta]} |\hat{q}(x) - q(x)| < \epsilon$$

以高概率成立。类似地，对于任意 $\epsilon > 0$，存在 P_2，使得对于所有 $p \geq P_2$，

$$\sup_{x \in [\Delta, \mathcal{L}]} |\hat{q}(x) - q(x)| < \epsilon$$

以高概率成立。
令 $N = \max(P_1, P_2)$。那么对于所有 $p \geq N$ 以及每个 $\epsilon > 0$，都有

$$\sup_{x \in [-\mathcal{L}, -\Delta] \cup [\Delta, \mathcal{L}]} |\hat{q}(x) - q(x)| < \epsilon$$

以高概率成立。或者，我们也可以将其表示为

$$\mathbb{P}\left(\sup_{x \in [-\mathcal{L}, -\Delta] \cup [\Delta, \mathcal{L}]} |\hat{q}(x) - q(x)| > \epsilon \right) < \delta, \quad \text{for all small } \delta > 0.$$

□

命题 A.5. $\forall \epsilon > 0, \quad \mathbb{P}\left(\text{上确界}_{x \in I} |\hat{} q_0(x) - q_0(x)| > \epsilon \right) \to 0$，其中 $I = [-\mathcal{L}, -\Delta] \cup [\Delta, \mathcal{L}]$，且 $\mathcal{L}$，Δ 在假设 (1) 和 (2)中定义。

命题 *A.5* 的证明。

$$\sup_{x\in I}|\hat{q}_0(x)-q_0(x)| \xrightarrow{P} 0,$$

其中

$$q_0(x)=\frac{1}{\tau}\phi\left(\frac{x+\alpha\tau\cdot\operatorname{sign}(x)}{\tau}\right),\quad \hat{q}_0(x)=\frac{1}{\hat{\tau}}\phi\left(\frac{x+\widehat{\alpha\tau}\cdot\operatorname{sign}(x)}{\hat{\tau}}\right),$$

$\phi(x)=\frac{1}{\sqrt{2\pi}}e^{-x^2/2}$是标准正态密度，且 $\hat{\tau}\xrightarrow{P}\tau$, $\widehat{\alpha\tau}\xrightarrow{P}\alpha\tau$。
令 $\epsilon>0$。由于 ϕ 在任意紧区间上是一致连续的，且 $[-\mathcal{L},-\Delta]$ 是紧的，因此存在 $\delta_1>0$，使得对任意 $x,\ y\in[-\mathcal{L},-\Delta]$

$$|x-y|<\delta_1 \quad\Rightarrow\quad |\phi(x)-\phi(y)|<\epsilon.$$

定义：

$$x_1:=\frac{x+\widehat{\alpha\tau}\cdot\operatorname{sign}(x)}{\hat{\tau}},\quad x_2:=\frac{x+\alpha\tau\cdot\operatorname{sign}(x)}{\tau}.$$

Then

$$|\hat{q}_0(x)-q_0(x)|=\left|\frac{1}{\hat{\tau}}\cdot\phi(x_1)-\frac{1}{\tau}\phi(x_2)\right|=\left|\frac{1}{\hat{\tau}}\cdot\phi(x_1)-\frac{1}{\tau}\phi(x_2)-\frac{1}{\tau}\phi(x_1)+\frac{1}{\tau}\phi(x_1)\right|.$$

利用三角不等式，

$$|\hat{q}_0(x)-q_0(x)|\le\left|\frac{1}{\hat{\tau}}\cdot\phi(x_1)-\frac{1}{\tau}\phi(x_1)\right|+\left|\frac{1}{\tau}\phi(x_1)-\frac{1}{\tau}\phi(x_2)\right|$$

$$\le\phi(x_1)\left|\frac{1}{\hat{\tau}}-\frac{1}{\tau}\right|+\frac{1}{\tau}|\phi(x_1)-\phi(x_2)|$$

由于 ϕ 在 $[-\mathcal{L},-\Delta]$上有界，存在上界 $\mathcal{M}>0$，使得 $\phi(x_1)\le\mathcal{M}$；又由于$\hat{\tau}\xrightarrow{P}\tau$和 $\widehat{\alpha\tau}\xrightarrow{P}\alpha\tau$, $\exists P_1\ large\ enough\ s.t\ \forall p\ge P_1$，以高概率有：

$$|x_1-x_2|<\delta_1,\quad |\hat{\tau}-\tau|<\epsilon.$$

So

$$|\phi(x_1)-\phi(x_2)|\le\epsilon.$$

因此，以高概率，

$$\sup_{x\in[-\mathcal{L},-\Delta]}|\hat{q}_0(x)-q_0(x)|\le\mathcal{M}\cdot\frac{\epsilon}{\tau(\tau-\epsilon)}+\frac{\epsilon}{\tau}.$$

同理，我们可以证明，当 $\exists P_2$ 足够大，使得 $\forall p\ge P_2$ 时，以高概率
$\sup_x|\hat{q}_0(x)-q_0(x)|\le\epsilon$. 现在令 $P=max(P_1,P_2)$，则对于所有 $p\ge P$ 和 $\forall\epsilon>0$，我们有$\sup_{x\in[\Delta,\mathcal{L}]}|\hat{q}_0(x)-q_0(x)|\le\epsilon$ 以高概率。$\in I$
我们得出结论，当 $\epsilon\to 0$ 时，

$$\sup_{x\in I}|\hat{q}_0(x)-q_0(x)|\xrightarrow{P}0.$$

□

引理 A.6。

$$\frac{1}{p}\sum_{i=1}^{p}\mathbf{1}\left\{\hat{\beta}_i\neq 0,\frac{\hat{q}_0(\hat{\beta}_i)}{\hat{q}(\hat{\beta}_i)}\le t\right\}\xrightarrow{P}\mathbb{P}\left(|\Pi+\tau Z|>\alpha\tau,\frac{q_0(\eta_{\alpha\tau}(\Pi+\tau Z))}{q(\eta_{\alpha\tau}(\Pi+\tau Z))}\le t\right)$$

证明引理 A.6。 我们需要证明 $\forall \delta > 0$,

$$\mathbb{P}(|\frac{1}{p}\sum_{i=1}^{p}\mathbf{1}\{\hat{\beta}_i \neq 0, \frac{\hat{q}_0(\hat{\beta}_i)}{\hat{q}(\hat{\beta}_i)} \leq t\} - \mathbb{P}(|\Pi + \tau Z| > \alpha\tau, \frac{q_0(\eta_{\alpha\tau}(\Pi + \tau Z))}{q(\eta_{\alpha\tau}(\Pi + \tau Z))} \leq t)| > \delta) \to 0.$$

根据三角不等式,

$$\mathbb{P}(|\frac{1}{p}\sum_{i=1}^{p}\mathbf{1}\{\hat{\beta}_i \neq 0, \frac{\hat{q}_0(\hat{\beta}_i)}{\hat{q}(\hat{\beta}_i)} \leq t\} - \mathbb{P}(|\Pi + \tau Z| > \alpha\tau, \frac{q_0(\eta_{\alpha\tau}(\Pi + \tau Z))}{q(\eta_{\alpha\tau}(\Pi + \tau Z))} \leq t)| > \delta)$$

$$\leq \mathbb{P}(|\frac{1}{p}\sum_{i=1}^{p}\mathbf{1}\{\hat{\beta}_i \neq 0, \frac{\hat{q}_0(\hat{\beta}_i)}{\hat{q}(\hat{\beta}_i)} \leq t\} - \frac{1}{p}\sum_{i=1}^{p}1\{\hat{\beta}_i \neq 0, \frac{q_0(\hat{\beta}_i)}{q(\hat{\beta}_i)} \leq t\}| > \frac{\delta}{2})$$

$$+ \mathbb{P}(|\frac{1}{p}\sum_{i=1}^{p}\mathbf{1}\{\hat{\beta}_i \neq 0, \frac{q_0(\hat{\beta}_i)}{q(\hat{\beta}_i)} \leq t\} - \mathbb{P}(|\Pi + \tau Z| > \alpha\tau, \frac{q_0(\eta_{\alpha\tau}(\Pi + \tau Z))}{q(\eta_{\alpha\tau}(\Pi + \tau Z))} \leq t)| > \frac{\delta}{2})$$

由于在定理 4.2 中我们证明了

$$\frac{1}{p}\sum_{i=1}^{p}\mathbf{1}\{\hat{\beta}_i \neq 0, \frac{q_0(\hat{\beta}_i)}{q(\hat{\beta}_i)} \leq t\} \to \mathbb{P}(|\Pi + \tau Z| > \alpha\tau, \frac{q_0(\eta_{\alpha\tau}(\Pi + \tau Z))}{q(\eta_{\alpha\tau}(\Pi + \tau Z))} \leq t),$$

我们有

$$\lim_{p\to\infty} \mathbb{P}(|\frac{1}{p}\sum_{i=1}^{p}\mathbf{1}\{\hat{\beta}_i \neq 0, \frac{q_0(\hat{\beta}_i)}{q(\hat{\beta}_i)} \leq t\} - \mathbb{P}(|\Pi + \tau Z| > \alpha\tau, \frac{q_0(\eta_{\alpha\tau}(\Pi + \tau Z))}{q(\eta_{\alpha\tau}(\Pi + \tau Z))} \leq t)| > \frac{\delta}{2}) = 0.$$

若能进一步证明以下结论，证明便告完成:

$$\lim_{p\to\infty} \mathbb{P}(|\frac{1}{p}\sum_{i=1}^{p}\mathbf{1}\{\hat{\beta}_i \neq 0, \frac{\hat{q}_0(\hat{\beta}_i)}{\hat{q}(\hat{\beta}_i)} \leq t\} - \frac{1}{p}\sum_{i=1}^{p}\mathbf{1}\{\hat{\beta}_i \neq 0, \frac{q_0(\hat{\beta}_i)}{q(\hat{\beta}_i)} \leq t\}| > \frac{\delta}{2}) = 0.$$

Sinc ˆe 我们证明了$-^{q_0}_{\hat{q}} \rightrightarrows ^{q_0}_{q}$ 依概率一致，因此当 p 足够大时，偶数 t

$$A = \{\sup_{1\leq i\leq p} |\frac{\hat{q}_0(\hat{\beta}_i)}{\hat{q}(\hat{\beta}_i)} - \frac{q_0(\hat{\beta}_i)}{q(\hat{\beta}_i)}| \leq \delta'\}$$

以高概率发生。现在，当 p足够大且在事件 A发生时，
若$\frac{q_0(\hat{\beta}_i)}{q(\hat{\beta}_i)} < t - \delta'$，则

$$\frac{\hat{q}_0(\hat{\beta}_i)}{\hat{q}(\hat{\beta}_i)} = \frac{q_0(\hat{\beta}_i)}{q(\hat{\beta}_i)} + \frac{\hat{q}_0(\hat{\beta}_i)}{\hat{q}(\hat{\beta}_i)} - \frac{q_0(\hat{\beta}_i)}{q(\hat{\beta}_i)}$$

$$\leq \frac{q_0(\hat{\beta}_i)}{q(\hat{\beta}_i)} + |\frac{\hat{q}_0(\hat{\beta}_i)}{\hat{q}(\hat{\beta}_i)} - \frac{q_0(\hat{\beta}_i)}{q(\hat{\beta}_i)}|$$

$$\leq t - \delta' + \delta' = t,$$

因此，在事件 A 上,

$$\mathbf{1}\{\hat{\beta}_i \neq 0, \frac{q_0(\hat{\beta}_i)}{q(\hat{\beta}_i)} < t - \delta'\} \leq \mathbf{1}\{\hat{\beta}_i \neq 0, \frac{\hat{q}_0(\hat{\beta}_i)}{\hat{q}(\hat{\beta}_i)} < t\}. \tag{48}$$

类似地，如果 $\frac{\hat{q}_0(\hat{\beta}_i)}{\hat{q}(\hat{\beta}_i)} < t$, 则

$$\begin{aligned}\frac{q_0(\hat{\beta}_i)}{q(\hat{\beta}_i)} &= \frac{\hat{q}_0(\hat{\beta}_i)}{\hat{q}(\hat{\beta}_i)} + \frac{q_0(\hat{\beta}_i)}{q(\hat{\beta}_i)} - \frac{\hat{q}_0(\hat{\beta}_i)}{\hat{q}(\hat{\beta}_i)} \\ &\leq \frac{\hat{q}_0(\hat{\beta}_i)}{\hat{q}(\hat{\beta}_i)} + |\frac{q_0(\hat{\beta}_i)}{q(\hat{\beta}_i)} - \frac{\hat{q}_0(\hat{\beta}_i)}{\hat{q}(\hat{\beta}_i)}| \\ &\leq t + \delta'.\end{aligned}$$

因此，在事件 A 上，

$$\mathbf{1}\{\hat{\beta}_i \neq 0, \frac{\hat{q}_0(\hat{\beta}_i)}{\hat{q}(\hat{\beta}_i)} < t\} \leq \mathbf{1}\{\hat{\beta}_i \neq 0, \frac{q_0(\hat{\beta}_i)}{q(\hat{\beta}_i)} < t + \delta'\}. \tag{49}$$

方程（48）和（49表明，当 p 足够大时，以高概率有

$$\mathbf{1}\{\hat{\beta}_i \neq 0, \frac{q_0(\hat{\beta}_i)}{q(\hat{\beta}_i)} < t - \delta'\} \leq \mathbf{1}\{\hat{\beta}_i \neq 0, \frac{\hat{q}_0(\hat{\beta}_i)}{\hat{q}(\hat{\beta}_i)} < t\} \leq \mathbf{1}\{\hat{\beta}_i \neq 0, \frac{q_0(\hat{\beta}_i)}{q(\hat{\beta}_i)} < t + \delta'\},$$

这意味着

$$\frac{1}{p}\sum_{i=1}^{p}\mathbf{1}\{\hat{\beta}_i \neq 0, \frac{q_0(\hat{\beta}_i)}{q(\hat{\beta}_i)} < t - \delta'\} \leq \frac{1}{p}\sum_{i=1}^{p}\mathbf{1}\{\hat{\beta}_i \neq 0, \frac{\hat{q}_0(\hat{\beta}_i)}{\hat{q}(\hat{\beta}_i)} < t\} \leq \frac{1}{p}\sum_{i=1}^{p}\mathbf{1}\{\hat{\beta}_i \neq 0, \frac{q_0(\hat{\beta}_i)}{q(\hat{\beta}_i)} < t + \delta'\}$$

如果从各项中减去 $\frac{1}{p}\sum_{i=1}^{p}\mathbf{1}\{\hat{\beta}_i \neq 0, \frac{q_0(\hat{\beta}_i)}{q(\hat{\beta}_i)} < t\}$，则得到，

$$\begin{aligned}&\frac{1}{p}\sum_{i=1}^{p}\mathbf{1}\{\hat{\beta}_i \neq 0, \frac{q_0(\hat{\beta}_i)}{q(\hat{\beta}_i)} < t - \delta'\} - \frac{1}{p}\sum_{i=1}^{p}\mathbf{1}\{\hat{\beta}_i \neq 0, \frac{q_0(\hat{\beta}_i)}{q(\hat{\beta}_i)} < t\} \\ &\leq \frac{1}{p}\sum_{i=1}^{p}\mathbf{1}\{\hat{\beta}_i \neq 0, \frac{\hat{q}_0(\hat{\beta}_i)}{\hat{q}(\hat{\beta}_i)} < t\} - \frac{1}{p}\sum_{i=1}^{p}\mathbf{1}\{\hat{\beta}_i \neq 0, \frac{q_0(\hat{\beta}_i)}{q(\hat{\beta}_i)} < t\} \\ &\leq \frac{1}{p}\sum_{i=1}^{p}\mathbf{1}\{\hat{\beta}_i \neq 0, \frac{q_0(\hat{\beta}_i)}{q(\hat{\beta}_i)} < t + \delta'\} - \frac{1}{p}\sum_{i=1}^{p}\mathbf{1}\{\hat{\beta}_i \neq 0, \frac{q_0(\hat{\beta}_i)}{q(\hat{\beta}_i)} < t\}\end{aligned}$$

现在来看

$$\begin{aligned}&\frac{1}{p}\sum_{i=1}^{p}\mathbf{1}\{\hat{\beta}_i \neq 0, \frac{q_0(\hat{\beta}_i)}{q(\hat{\beta}_i)} < t - \delta'\} - \frac{1}{p}\sum_{i=1}^{p}\mathbf{1}\{\hat{\beta}_i \neq 0, \frac{q_0(\hat{\beta}_i)}{q(\hat{\beta}_i)} < t\} \\ &= -(\frac{1}{p}\sum_{i=1}^{p}\mathbf{1}\{\hat{\beta}_i \neq 0, t - \delta' < \frac{q_0(\hat{\beta}_i)}{q(\hat{\beta}_i)} < t\})\end{aligned}$$

根据定理4.2 可得

$$\frac{1}{p}\sum_{i=1}^{p}\mathbf{1}\{\hat{\beta}_i \neq 0, t - \delta' < \frac{q_0(\hat{\beta}_i)}{q(\hat{\beta}_i)} < t\} \to \mathbb{P}(|\Pi + \tau Z| > \alpha\tau, t - \delta' < \frac{q_0(\eta_{\alpha\tau}(\Pi + \tau Z))}{q(\eta_{\alpha\tau}(\Pi + \tau Z))} < t)$$

So,

$$\lim_{\delta'\to 0}\lim_{p\to\infty}\frac{1}{p}\sum_{i=1}^{p}\mathbf{1}\{\hat{\beta}_i\neq 0, t-\delta'<\frac{q_0(\hat{\beta}_i)}{q(\hat{\beta}_i)}<t\}$$

$$=\lim_{\delta'\to 0}\mathbb{P}(|\Pi+\tau Z|>\alpha\tau, t-\delta'<\frac{q_0(\eta_{\alpha\tau}(\Pi+\tau Z))}{q(\eta_{\alpha\tau}(\Pi+\tau Z))}<t).$$

随机变量

$$Y=\frac{q_0(\eta_{\alpha\tau}(\Pi+\tau Z))}{q(\eta_{\alpha\tau}(\Pi+\tau Z))}$$

连续，且在 t 处具有连续的概率密度函数，因此，如果 δ' 足够小，那么

$$\lim_{\delta'\to 0}\mathbb{P}(|\Pi+\tau Z|>\alpha\tau,,t-\delta'<\frac{q_0(\eta_{\alpha\tau}(\Pi+\tau Z))}{q(\eta_{\alpha\tau}(\Pi+\tau Z))}<t)=0,$$

对于 $\frac{1}{p}\sum_{i=1}^{p}\mathbf{1}\{\hat{}\beta_i\neq 0,\ \ t<\frac{q_0(\hat{}\beta_i)}{q(\hat{}\beta_i)}<\underline{t+\delta'}\}$，同样也成立。因此我们有

$$\lim_{\delta'\to 0}\lim_{p\to\infty}\frac{1}{p}\sum_{i=1}^{p}\mathbf{1}\{\hat{\beta}_i\neq 0, t-\delta'<\frac{q_0(\hat{\beta}_i)}{q(\hat{\beta}_i)}<t\}$$

$$=\lim_{\delta'\to 0}\mathbb{P}(|\Pi+\tau Z|>\alpha\tau, t-\delta'<\frac{q_0(\eta_{\alpha\tau}(\Pi+\tau Z))}{q(\eta_{\alpha\tau}(\Pi+\tau Z))}<t)=0,$$

and

$$\lim_{\delta'\to 0}\lim_{p\to\infty}\frac{1}{p}\sum_{i=1}^{p}\mathbf{1}\{\hat{\beta}_i\neq 0, t<\frac{q_0(\hat{\beta}_i)}{q(\hat{\beta}_i)}<t+\delta'\}$$

$$=\lim_{\delta'\to 0}\mathbb{P}(|\Pi+\tau Z|>\alpha\tau, t<\frac{q_0(\eta_{\alpha\tau}(\Pi+\tau Z))}{q(\eta_{\alpha\tau}(\Pi+\tau Z))}<t+\delta')=0$$

因此，根据夹逼定理，我们得出

$$\frac{1}{p}\sum_{i=1}^{p}\mathbf{1}\{\hat{\beta}_i\neq 0, \frac{\hat{q}_0(\hat{\beta}_i)}{\hat{q}(\hat{\beta}_i)}<t\}-\frac{1}{p}\sum_{i=1}^{p}\mathbf{1}\{\hat{\beta}_i\neq 0, \frac{q_0(\hat{\beta}_i)}{q(\hat{\beta}_i)}<t\}\to 0.$$

□

引理 A.7。

$$\frac{1}{p}\sum_{i=1}^{p}\mathbf{1}\left\{\hat{\beta}_i\neq 0, \frac{\hat{q}_0(\hat{\beta}_i)}{\hat{q}(\hat{\beta}_i)}\leq t, \beta_i=0\right\}\xrightarrow{P}\mathbb{P}\left(|\Pi+\tau Z|>\alpha\tau, \frac{q_0(\eta_{\alpha\tau}(\Pi+\tau Z))}{q(\eta_{\alpha\tau}(\Pi+\tau Z))}\leq t, \Pi=0\right)$$

引理的证明 *A.7*。 与A.6的证明类似。

□

*Pro*定理*4.8*的证明。陈述（32）由引理A.6 和引理A.7得出，即 e

$$\mathrm{FDP}^{EB}(t;\lambda)\xrightarrow{P}\frac{\mathbb{P}\left(|\Pi+\tau Z|>\alpha\tau, \frac{q_0(\eta_{\alpha\tau}(\Pi+\tau Z))}{q(\eta_{\alpha\tau}(\Pi+\tau Z))}\leq t, \Pi=0\right)}{\mathbb{P}\left(|\Pi+\tau Z|>\alpha\tau, \frac{q_0(\eta_{\alpha\tau}(\Pi+\tau Z))}{q(\eta_{\alpha\tau}(\Pi+\tau Z))}\leq t\right)}\equiv \mathrm{fdp}^*(t;\lambda)$$

and

$$\begin{aligned}\mathrm{TPP}^{\mathrm{EB}}(t;\lambda) &= \frac{\frac{1}{p}\sum_{i=1}^{p}\mathbf{1}\left\{\hat{\beta}_i \neq 0, \frac{\hat{q}_0(\hat{\beta}_i)}{\hat{q}(\hat{\beta}_i)} \leq t, \beta_i \neq 0\right\}}{\frac{1}{p}\sum_{i=1}^{p}\mathbf{1}(\beta_i \neq 0)} \\ &= \frac{\frac{1}{p}\sum_{i=1}^{p}\mathbf{1}\left\{\hat{\beta}_i \neq 0, \frac{\hat{q}_0(\hat{\beta}_i)}{\hat{q}(\hat{\beta}_i)} \leq t\right\} - \frac{1}{p}\sum_{i=1}^{p}\mathbf{1}\left\{\hat{\beta}_i \neq 0, \frac{\hat{q}_0(\hat{\beta}_i)}{\hat{q}(\hat{\beta}_i)} \leq t, \beta_i = 0\right\}}{\frac{1}{p}\sum_{i=1}^{p}\mathbf{1}(\beta_i \neq 0)} \\ &\xrightarrow{P} \mathbb{P}(|\Pi + \tau Z| > \alpha\tau, \frac{q_0(\eta_{\alpha\tau}(\Pi + \tau Z))}{q(\eta_{\alpha\tau}(\Pi + \tau Z))} \leq t | \Pi \neq 0) \\ &\equiv \mathrm{tpp}^*(t;\lambda).\end{aligned}$$

□

定理*4.10*的证明。定义随机变量

$$X = \eta_{\alpha\tau}(\tilde{X}), \qquad \tilde{X} = \Pi + \tau Z, \tag{50}$$

其中，α, τ 由定理陈述中 $\lambda \neq \lambda^*$ 的取值确定，并考虑如下检验问题

$$H_0 : \Pi = 0 \tag{51}$$

~假设只能观测到 X（而不能观测到 X）。注意，一般而言，无法从 X 中恢复 X（仅当 $X \neq 0$ 时可以）。对于这个单变量检验问题（由于 Π 是随机的，因此是贝叶斯检验问题），令 FPR $= \mathbb{P}(\mathrm{lfdr}(X) \leq t | \Pi = 0)$ 和 TPR $= \mathbb{P}(\mathrm{lfdr}(X) \leq t | \Pi \neq 0)$ 分别表示在 $\mathrm{lfdr}(X) \leq t$ 时拒绝原假设的检验的第一类错误概率和检验功效，其中 lfdr(·) 是式（20）中的函数。另记 FPR*、FPR* 为在 $\mathrm{lfdr}^*(X) \leq t^*$ 时拒绝原假设的检验所对应的量。在整个证明过程中，如果一个检验的检验功效等于 ξ，我们就称其为水平-ξ 检验。

我们需要证明，如果 TPR $=$ FPR* $= \xi$，那么 FPR* $\leq$ FPR。首先，我们声称，在一个（假设的）问题中，若观测到的是 X 而不是 X，则这一点~成立：

~**命题 A.8.** 考虑在(51) 中检验 X 而非 X的情形，并令~ ~*FPR* 和 *TPR* 分别为检验的第一类错误概率和检验功效，其中当(51)满足~局部错误发现率($\tilde{X}$) $\leq \tilde{t}$时拒绝原假设。如果 ~*TPR* $=$ ~*TPR**，那么 ~*FPR** $\leq$ ~*FPR*。

~ ~*为证明命题 A.8，反设存在 λ 使得 FPR $<$ FPR。定义

$$\tilde{Y}^* := \tilde{X}^* + \sqrt{\tau^2 - (\tau^*)^2}\tilde{Z}, \tag{52}$$

~其中 $Z \sim \mathcal{N}$ 与其他所有变量独立，且服从 (0, 1)，这里我们用到了 $\tau^* < \tau$ 这一事实。于是~ ~ ~*由于 (Y^*, Π) 与 (X, Π) 具有相同的分布，假设 FPR $<$FPR 意味着~存在一个基于 (52) 的检验，它是基于 X^* 的合法（随机化）检验，且其第一类错误概率小于 X^* 中水平为-ξ 的似然比检验（LR 检验）；该检验在 ~局部错误发现率 $(\tilde{X}^*) \leq \tilde{t}^*$ 时拒绝原假设（因为，基于 $\tilde{X}$ 的似然比检验中，~局部错误发现率$(\tilde{x})$ 是单调的）。这与奈曼-皮尔逊引理相矛盾（准确地说，是定理 4.3 证明中所包含的版本）。

回到主要结论的证明，由于 $t^* \leq t_0^*$，我们有

$$\mathrm{lfdr}^*(X) \leq t^* \iff \widetilde{\mathrm{lfdr}}^*(\tilde{X}) \leq t^*,$$

因此，基于 X^* 的 level-ξ 预言机检验与基于 X^* 的 level-ξ 预言机~检验具有相同的第一类错误概率。现在，我们已经证明，后者的第一类错误概率更小

$\tilde{}$ $\tilde{}$比基于 X 的 level-ξ 似然比检验更小。另一方面，基于 X 的 level-ξ 似然比检验$\tilde{}$更优（假阳性率更低），优于基于 X 的任何其他 level-ξ 检验（即非似然比检验）；尤其优于基于 X 的 level-ξ 似然比检验（因为任何基于 X 的检验也都是基于 X 的检验，因为 $X = \eta_{\alpha\tau}(X)$ 是 X 的函数）。$\tilde{}$ $\tilde{}$ □