arXiv:2611.32898v1 [math.ST] 24 Nov 2026

缓慢衰减的高斯相关性下 Benjamini–Hochberg 程序错误发现率的无维数依赖界

Zuofeng Shang[1] 和 Wenge Guo[1]

[1]新泽西理工学院数学科学系

2026年11月25日

摘要

已证明，在独立性和正依赖条件下，Benjamini–Hochberg（BH）程序能够控制错误发现率。然而，在任意依赖条件下，其保证会因一个随假设数量对数增长的因子而减弱。对于高斯检验统计量，目前尚不清楚依赖是否会导致错误发现率随检验数量增加。我们研究了这样一个问题：当两个检验统计量之间的相关系数随其索引间距离增大而衰减时，会出现什么情况。我们旨在针对广泛的相关结构类以及所有信号配置，给出统一成立的保证。我们证明，如果相关性至少以距离倒数对数的幂次速度衰减，且幂指数大于一，那么原始 BH 程序的错误发现率就有一个上界，即标称水平乘以一个与假设数量无关的因子。该因子仅随标称水平降低而按多重对数速度增长。此结果适用于单侧和双侧检验，且该程序不需要有关相关性的任何信息。所考虑的相关结构类允许相关系数取正值或负值。所需的衰减速度可以任意接近对数倒数。关键技术工具是对超越指示变量进行截断赫尔米特展开。该展开表明，其主要波动对 BH 可能采用的所有阈值都是共同的。这些波动可以得到良好控制，而余项过小，不会累积，因此所得上界不会随维数增大。

关键词： 错误发现率、Benjamini–Hochberg 程序、多重检验、缓慢衰减的相关性、高斯检验统计量、Hermite 展开。

1 引言

现代科学研究通常会同时检验数千个甚至数百万个假设，而所涉及的检验统计量很少彼此独立。例如，共调控基因的表达水平会同步变化，基因组上相邻的基因座以及脑图像中相邻的体素具有空间相关性，而沿时间序列计算的统计量则会继承其序列依赖性。Benjamini 和 Hochberg (1995年) 提出的 BH 程序已成为控制此类研究中错误发现率（FDR）的标准工具。当独立性不成立时，寻求相应的保证是很自然的。

对 Benjamini 和 Hochberg （1995年）的最初分析表明，对于相互独立的检验统计量，该程序的错误发现率等于名义水平乘以真实零假设所占的比例。 Benjamini 和 Yekutieli （2001年）以及Sarkar （2002年）将这一保证扩展到了正回归依赖；对于单侧高斯检验，这一条件相当于相关系数非负，但对于双侧检验而言则更难验证（Karlin 和 Rinott，1981年）。在没有任何

关于依赖的假设时，Benjamini 和 Yekutieli（2001年）表明，错误发现率超过名义水平的幅度至多为假设数量的对数阶。对于所有联合分布而言，这一因子无法进一步改进；但达到该上界的分布远非高斯分布，而在实践中，很少观察到该程序表现得如此糟糕。因此，大量研究转而考察弱依赖，结果表明，在混合条件或衰减条件下，错误发现比例的渐近行为与独立性情形相同（Farcomeni，2007；Wu，2008；Clarke 和 Hall，2009；Delattre 和 Roquain，2016；Datta 和 Dey，2026），并且依赖主要影响其变异性（Owen，2005；Efron，2007；Schwartzman 和 Lin，2011）。这些结果描述的是：在依赖结构固定的情况下，假设数量增长时的极限行为。它们并未给出对于每一个假设数量、并对可能的相关系数和信号均一致成立的错误发现率界，这恰恰是未知依赖下的保证。

近期有多项研究在仅对零假设 p-值施加条件的情况下，获得了不随 m增长的 BH 程序错误发现率界。Su（2018）研究了零假设 p-值之间的正回归依赖，Chi 等人（2025）则研究了弱负依赖。对于高斯检验统计量，这一问题已从多个方向展开研究。首先，高斯依赖可能会使错误发现率膨胀。Dobriban（2026）表明，当双侧高斯检验之间存在相关性时，BH 程序可能无法将错误发现率控制在名义水平以内。Lei（2026）表明，这种膨胀可以超过名义水平的任意固定倍数，并猜想最坏情形错误发现率不随 m增长。其次，针对特殊结构，人们已经获得了与 m 无关的界。Lei（2026）为共同因子模型证明了此类界。Zhang（2026）为双侧高斯检验建立了精确的 Simes 不等式；这些不等式在全局零假设下，对于若干类协方差矩阵控制了错误发现率。Lyu 和 Wang（2026）研究了固定维数下的单侧检验，并在全局零假设下、相关系数非正以及秩一模型的情形中，获得了与 m无关的界。其他近期工作针对给定协方差矩阵推导了 BH 程序错误发现率的有限样本界（Ghosh 和 Sarkar，2026），为双侧检验提出了平移程序（Sarkar，2023；Ghosh 和 Sarkar，2025），并在依赖结构已知时恢复了精确控制（Fithian 和 Lei，2022；Nguyen 和 Fithian，2025）。据我们所知，对于原始 BH 程序，在允许相关系数变号且缓慢衰减的广泛高斯相关结构类上，当信号强度及其位置任意时，Benjamini 和 Yekutieli（2001）中的对数因子是否确实是必需的，这一问题仍未解决。

在这篇简短的说明中，当两个检验统计量之间的相关系数随其索引之间的距离增大而衰减时，我们回答了这一问题。这包括平稳依赖情形，此时相关系数仅取决于该距离。我们证明，如果绝对相关系数至少以该距离的逆对数的幂次衰减，且幂次大于一，那么 BH 程序的错误发现率由名义水平乘以一个与假设数量无关的因子所界定。我们不对相关系数的符号施加任何条件，其符号可以任意变化。对于所有此类相关矩阵和所有信号配置，以及单侧检验和双侧检验，该界均一致成立；所需的衰减速度可以任意接近逆对数速度。一个主要挑战在于，BH 的阈值是随机的，因此必须控制所有可能阈值下的超越事件。截断赫尔米特展开表明，这些事件的主要波动对所有阈值都是共同的，因而可以轻松控制；而赫尔米特展开的余项过小，不会累积起来。这就得到一个不会随假设数量增加而膨胀的 FDR界。

本文其余部分的组织如下。第2 节介绍模型，定义最坏情形错误发现率，陈述主要结果，讨论我们结果所适用的依赖范围，并概述其证明。第3 节对 BH 的错误发现率与 BY界及我们的界进行数值比较。第4 节给出主要结果和初步结果的证明。

2 多重检验与错误发现率

假设我们观测到一个高斯向量 $T = (T_1、\ldots\ldots、T_m)^\top \sim N(\mu,\ \Sigma)$，其未知均值向量为$\mu = (\mu_1、\ldots\ldots、\mu_m)^\top$，未知协方差矩阵为 Σ，且其对角线元素均等于 1。我们检验零假设

$$H_i^0 : \mu_i = 0, \qquad i = 1, \ldots, m,$$

针对单侧备择假设 $H^1{}_i{}^a$: $\mu_i > 0$ 或双侧备择 $H^2{}_i{}^a$: $\mu_i \neq 0$。设$\mathcal{H}_0 = \{i : \mu_i = 0\}$ 为真实零假设的集合，$m_0 = |\mathcal{H}_0|$ 为真实零假设的数量，$\pi_0 = m_0/m$ 为真实零假设所占的比例。令 Φ为标准正态累积分布函数，则单侧和双侧 p值为

$$\begin{aligned} &\text{(one-sided)}: \qquad P_i^1 = 1 - \Phi(T_i), \\ &\text{(two-sided)}: \qquad P_i^2 = 2\big(1 - \Phi(|T_i|)\big), \end{aligned}$$

在 H_i^0下，两者均服从 $[0,1]$ 上的均匀分布。全文中，P_i表示对所有 i取 P_i^1（单侧情形），或对所有 i取 P_i^2（双侧情形）。令 $P_{(1)} \le \cdots \le P_{(m)}$为排序后的 p值。在显著性水平 $\alpha \in$（0，1）下，当且仅当 $P_i \le \alpha R/m$时，BH程序拒绝 H_i^0，其中

$$R = \max\{1 \le i \le m : P_{(i)} \le \alpha i/m\}$$

为拒绝数；若集合为空，则 $R = 0$ （不拒绝）。令 V为被拒绝的真实零假设的数量。错误发现率为

$$\mathrm{FDR}(\alpha; m, \mu, \Sigma) = \mathbb{E}\Big[\frac{V}{R \vee 1}\Big], \tag{1}$$

其中 $a \vee b = \max\{a, b\}$。此处的记号省略了所用的是单侧还是双侧p-值这一差别。我们的所有结果在这两种情况下均成立。

2.1 衰减相关性下的最坏情形错误发现率

令 $\mathbb{N}_0 = \{0, 1, 2, \ldots\}$。对于 $\epsilon > 0$和 $C > 0$，定义包络函数 $\rho_{\epsilon,C}$: $\mathbb{N}_0 \to [0,1]$，其定义为 $\rho_{\epsilon,C}(0) = 1$且

$$\rho_{\epsilon,C}(k) = \min\big\{1,\ C\log^{-(1+\epsilon)}(k + e)\big\}, \qquad k \ge 1.$$

考虑如下相关矩阵类

$$\mathcal{C}_m(\epsilon, C) = \big\{\Sigma \in \mathbb{R}^{m\times m} : \Sigma = \Sigma^\top,\ \Sigma \succeq 0,\ \Sigma_{ii} = 1,\ |\Sigma_{ij}| \le \rho_{\epsilon,C}(|i-j|) \text{ for all } i \neq j\big\}.$$

因此，两个检验统计量之间的相关系数，受包络函数在其索引间距处的取值所界定。一个重要的特殊情形是平稳依赖，此时，对于相关序列 ρ，有 $\Sigma_{ij} = \rho(|i-j|)$。于是条件可写为 $|\rho(k)| \le \rho_{\epsilon,C}(k)$，其中 $1 \le k \le m-1$。由于$\rho_{\epsilon,C}(k)\log k \to 0$，每个这样的 ρ都满足 Berman条件 $\rho(k)\log k \to 0$ (Berman，1964年)。当 ϵ趋近于零时，包络函数趋近于 Mittal和Ylvisaker（1975年）的临界速率 $\log^{-1} k$ 。为研究此类缓慢衰减相关性下 BH程序的错误发现率，我们考虑该类上的最坏情形错误发现率：

$$U(\alpha; \epsilon, C) = \sup_{m\ge 1} \sup_{\Sigma\in\mathcal{C}_m(\epsilon,C)} \sup_{\mu\in\mathbb{R}^m} \mathrm{FDR}(\alpha; m, \mu, \Sigma).$$

对于任意 μ和 Σ，Benjamini 和 Yekutieli (2001年）证明了

$$\mathrm{FDR}(\alpha; m, \mu, \Sigma) \le \alpha\pi_0 H_m, \tag{2}$$

其中 $H_m \triangleq \sum_{k=1}^m k^{-1} \sim \log m$。该界随 m 增长，并且在 T_i之间存在任意依赖时是紧的，但已知能够达到该界的分布都是非高斯分布。当 Σ 限制为 $\mathcal{C}_m$（ϵ，C）时，是否存在一个与 m 无关的界仍不清楚；也就是说，U（α；ϵ，C）是否 $< \infty$。

2.2 主要结果

在本节中，我们针对均值向量 μ 以及相关矩阵 $\mathcal{C}_m$（ϵ，C）这一类，建立最坏情形错误发现率的一个上界，并且该上界与假设的数量 m 无关。令 $\mathcal{S} = \{1, 2, 3, \dots\}$ 表示单侧 p 值，$\mathcal{S} = \{2, 4, 6, \dots\}$ 表示双侧 p 值，并固定一个元素 $k^* \in \mathcal{S}$。令 ϕ 表示标准正态密度，He_k 表示概率论学派厄米多项式，其定义为

$$He_0(x) \equiv 1,\ \phi^{(k)}(x) = (-1)^k He_k(x)\phi(x) \text{ for } k = 1, 2, \dots.$$

对于 $w \in (0,\ 1/2]$，令

$$z_w = \begin{cases} \Phi^{-1}(1-w), & \text{(one-sided)}, \\ \Phi^{-1}(1-w/2), & \text{(two-sided)}. \end{cases}$$

这里，z_w 是单侧情形下 w 的上侧 Φ 分位数，以及双侧情形下 Φ 的上侧 $w/2$ 分位数。对于 $z \geq 0$ 和 $r \in [-1, 1]$，令

$$r_z = \min\{1/2, z^{-2}\} \text{ and } R(r, z) = \mathrm{Cov}(\mathbf{1}\{X > z\}, \mathbf{1}\{Y > z\}),$$

其中，$(X,\ Y)$ 是相关系数为 r 的标准二元正态变量对。这里，$R(r,\ z)$ 是两个超越指示变量 $\mathbf{1}\{X > z\}$ 和 $\mathbf{1}\{Y > z\}$ 的协方差。令

$$\tau(r, z) = \begin{cases} 2, & |r| > 1/2, \\ 2\min\left\{1, \frac{R(|r|, z)}{1-\Phi(z)}\right\}, & |r| \leq 1/2. \end{cases}$$

令 $J = \lceil \log_2 m \rceil$，$w_j = 2^{-j}\alpha$ 对于 $j = 0,\ \dots\dots,\ J-1$，以及 $z_j = z_{w_j}$。对于 $k \geq 1$，令

$$G_{\epsilon,C,k}(m) = \sum_{l=1}^{m-1} \rho_{\epsilon,C}(l)^k, \quad b_k = \sup_{0 \leq z \leq z_{J-1}} \frac{\phi(z)|He_{k-1}(z)|}{k!\,(1-\Phi(z))}.$$

这里，$G_{\epsilon,C,k}(m)$ 是包络函数 $\rho_{\epsilon,C}(l)$ 在滞后 $l = 1, \dots, m-1$ 上的 k 次幂之和。可以证明，超越指示变量 $\mathbf{1}\{X > z\}$ 的 k 阶厄米系数为 $\phi(z)He_{k-1}(z)/k!$；见引理 4。因此，b_k 相对于该指示变量的均值 $1 - \Phi(z)$，在阈值 $z_0, \dots, z_{J-1}$ 上一致地界定了该系数。对于 $z \geq 0$，令

$$\mathsf{A}(z) \;=\; \sum_{k \in \mathcal{S},\ k \geq k^*} \frac{|He_{k-1}(z)|^2}{k!}\, r_z^{k-k^*}, \quad e_{\epsilon,C,m}(z) = 2\Big(\frac{\phi(z)}{1-\Phi(z)}\Big)^2 \mathsf{A}(z)\, \frac{G_{\epsilon,C,k^*}(m)}{m}.$$

函数 $\mathsf{A}(z)$ 收集阈值 z 处超越指标的 $k \geq k^*$ 阶厄米系数的平方，并通过 $\phi(z)^2$ 以及 r_z 的幂进行归一化。函数 $e_{\epsilon,\ C,\ m}(z)$ 将其与滞后项上弱相关的平均值 $G_{\epsilon,\ C,\ k^*}(m)/m$ 结合起来，并界定它们对超越计数在 z 处的相对方差的贡献。定义

$$K_{\epsilon,C} = 1 + \sup_{z \geq 0} \sum_{k \geq 1:\, \rho_{\epsilon,C}(k) > r_z} \tau\big(\rho_{\epsilon,C}(k), z\big), \tag{3}$$

and

$$\Lambda_{\epsilon,C,m}(\alpha) = \begin{cases} \Big(\sum_{k \in \mathcal{S},\ k < k^*} b_k \Big[k!\Big(\frac{1}{m} + \frac{2G_{\epsilon,C,k}(m)}{m}\Big)\Big]^{1/2}\Big)^2, & k^* > \min \mathcal{S}, \\ 0, & k^* = \min \mathcal{S}. \end{cases} \tag{4}$$

常数 $K_{\epsilon,\ C}$ 汇集强相关滞后项，而常数 $\Lambda_{\epsilon,\ C,\ m}(\alpha)$ 涉及 $k < k^*$ 阶的 Hermite项。前者说明了少数具有较大相关系数的滞后项。后者说明了对所有阈值 z_0，……，z_{J-1} 都共有的随机波动（差别仅在于确定性系数）。通过 b_k，这些系数在上述阈值范围内是一致有界的，因此这一波动一经控制即可适用于所有阈值，不会随阈值数量累积。

在整个过程中，我们固定 $\alpha \in (0, 1/2]$。对于 $K \geq 1$、$E \geq 0$ 和 $\pi_0 \in (0, 1]$，令

$$g(t; K, E, \pi_0) = \frac{6\pi_0\sqrt{K}}{t - 2\pi_0} + \frac{4\pi_0 E}{(t - 2\pi_0)^2},\ t > 2\pi_0,$$

and

$$\Psi_{\pi_0}(K, E, \alpha) = 2\pi_0 + \int_{2\pi_0}^{1/\alpha} \min\{1, g(t; K, E, \pi_0)\}\, dt. \tag{5}$$

由于 $\pi_0 \leq 1$ 且 $\alpha \leq 1/2$，我们有 $2\pi_0 \leq 1/\alpha$，因此上述积分上下限是有效的。令$\pi_0 \to 0$ 代入（5），得到 $\Psi_0(K,\ E,\ \alpha) = 0$。对于 $m \geq 2$，令

$$E_{\epsilon,C,m}(\alpha) = \sum_{j=0}^{J-1} e_{\epsilon,C,m}(z_j).$$

数量 $E_{\epsilon,\ C,\ m}(\alpha)$ 在阈值 z_0、……、z_{J-1} 上加总 $e_{\epsilon,\ C,\ m}$。由于 BH程序的阈值是随机的，因此必须同时控制所有这些阈值处的超越计数。因此$E_{\epsilon,\ C,\ m}(\alpha)$ 衡量了弱相关滞后项对错误发现率的总影响。当 $\epsilon k^* \geq 1$ 时，它在 m 中保持有界。定义

$$F_{\epsilon,C,m}(\alpha; \pi_0) = \inf_{\theta \in (0,1)} \Psi_{\pi_0}\Big(\frac{K_{\epsilon,C}}{(1-\theta)^2},\ \frac{\Lambda_{\epsilon,C,m}(\alpha)}{\theta^2} + \frac{E_{\epsilon,C,m}(\alpha)}{(1-\theta)^2},\ \alpha\Big). \tag{6}$$

当 $m = 1$时，令 $F_{\epsilon,\ C,\ 1}(\alpha;\ \pi_0) = \pi_0$。如果 $k^* = \min \mathcal{S}$，则 $\Lambda_{\epsilon,\ C,\ m}(\alpha) = 0$，因此，（6）变为

$$F_{\epsilon,C,m}(\alpha; \pi_0) = \Psi_{\pi_0}(K_{\epsilon,C}, E_{\epsilon,C,m}(\alpha), \alpha). \tag{7}$$

Th以下定理给出了 HG 程序错误发现率的上界。

定理 1. 设 $\alpha \in (0、\ 1/2]$、$\pi_0 \in [0, 1]$、$\epsilon > 0$、$C > 0$ 以及 $k^* \in \mathcal{S}$。

(i). 对于 b 对于单侧和双侧 p 值，任取 $m \geq 1$、$\mu \in \mathbb{R}^m$ 和 $\Sigma \in \mathcal{C}_m(\epsilon、\ C\quad)$，

$$\mathrm{FDR}(\alpha; m, \mu, \Sigma) \leq \alpha \min\Big\{\pi_0 H_m,\ F_{\epsilon,C,m}(\alpha; \pi_0)\Big\}. \tag{8}$$

(ii). 此外，若 $\epsilon k^* \geq 1$，则

$$\sup_{m \geq 2} F_{\epsilon,C,m}(\alpha; \pi_0) \leq A(\epsilon, C, \pi_0) + B(\epsilon, C, \pi_0) \log^{\max\{1, k^*/2\}}(1/\alpha), \tag{9}$$

其中，对于 $\pi_0 > 0$,

$$A(\epsilon, C, \pi_0) = \pi_0\big(2 + 12\sqrt{\bar{K}_{C,k^*}}\big) + 12\pi_0\sqrt{\bar{K}_{C,k^*}}\Big(1 + \log_+ \frac{1}{12\pi_0\sqrt{\bar{K}_{C,k^*}}}\Big)$$
$$+ 4\max\{2, 2^{k^*/2}\}\sqrt{\pi_0\big(\bar{\Lambda}_{C,k^*} + \bar{E}_{C,k^*}\big)},$$
$$B(\epsilon, C, \pi_0) = 12\pi_0\sqrt{\bar{K}_{C,k^*}} + 4\max\{2, 2^{k^*/2}\}\sqrt{\pi_0\big(\bar{\Lambda}_{C,k^*} + \bar{E}_{C,k^*}\big)},$$

且 $A(\epsilon, C, 0) = B(\epsilon, C, 0) = 0$。这里 $\log_+ x = \max\{\log x, 0\}$。当 $p = 1 + 1/k^*$ 和 $N = \#\{k \in \mathcal{S}: \ k < k^*\}$ 时，常数 K_{C,k^*}, Λ_{C,k^*} 和 E_{C,k^*} 如下给出

$$\bar{K}_{C,k^*} = 1 + 2\big(e^{(2C)^{1/p}} - e\big)_+ + \frac{2}{\sqrt{6\pi}} \sup_{z>\sqrt{2}} (1+z)e^{-z^2/6}\Big(e^{(Cz^2)^{1/p}} - e\Big)_+, \tag{10}$$

$$\bar{E}_{C,k^*} = a_* C^{k^*} \Big(\frac{4}{\log 2}\Big)^{k^*+1} \Big[\Big(\frac{2k^*+2}{e}\Big)^{k^*+1} + 2^{k^*+1}\Big], \tag{11}$$

$$\bar{\Lambda}_{C,k^*} = \frac{N}{2} \sum_{k\in\mathcal{S},\ k<k^*} \frac{1}{k!}\Big(\frac{8k+8}{\log 2}\Big)^k \Big[\Big(\frac{k}{e}\Big)^k + 2C^k\Big(\Big(\frac{2k}{e}\Big)^k + \frac{2^{kp}}{\log 2}\Big)\Big], \tag{12}$$

其中，$a_* = 2/\sqrt{3}$ 在双侧情形下满足 $k^* = 2$，而 $a_* = 2(e-1)/\sqrt{3}$ 在其他情况下成立。若 $k^* =$ 最小值 $\mathcal{S}$，则 $N = 0$，因此可以设定 $\Lambda_{C,k^*} = 0$。

定理 1 的证明见第 4 节。定理 1(i) 给出了在高斯相关性衰减时 BH 程序的错误发现率上界。对于每个 $m \geq 1$, $\mu \in \mathbb{R}^m$ 和 $\Sigma \in \mathcal{C}_m(\epsilon, C)$，该结论均成立。该界取 BY界 $\alpha\pi_0 H_m$ 和 $\alpha F_{\epsilon,C,m}(\alpha;\pi_0)$ 中较小者，因此不会超过 BY界。给定 m, ϵ, C, π_0 和 α，精确计算量 $F_{\epsilon,C,m}(\alpha;\pi_0)$ 较为困难，但数值计算是可行的。表 1 报告了在若干 ϵ, C, k^*, π_0 和 α取值下，$m = 10^4$ 和 $m = 10^8$ 对应的 $F_{\epsilon,C,m}(\alpha;\pi_0)$ 数值。对于所考察的 ϵ 和 C 取值，这些数值似乎变化不大。随着 π_0 减小，这些数值迅速下降。这些数值似乎随 m变化不大，这说明该错误发现率界与 m无关。

表 1 ： F_ϵ 的取值、C、m（α； π_0）在 $m = 10^4$ 和 $m = 10^8$ 下的各种情况。

			$m = 10^4$			$m = 10^8$		
p-值	(ϵ, C, k^*)	π_0	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$
单侧	(0。5, 0。01, 2)	1	27.79	15.92	9.98	27.14	15.66	9.96
		0.5	16.73	10.85	8.10	16.25	10.60	7.96
		0.2	8.33	5.94	4.86	8.02	5.74	4.72
		0.1	4.85	3.63	3.08	4.64	3.48	2.97
单侧	(1, 0.01, 1)	1	24.83	14.63	9.74	24.81	14.62	9.74
		0.5	14.56	9.59	7.33	14.54	9.57	7.32
		0.2	6.98	5.01	4.14	6.95	4.99	4.13
		0.1	3.94	2.96	2.53	3.92	2.94	2.51
双侧	(1, 0。1, 2)	1	24.77	14.60	9.73	24.76	14.59	9.73
		0.5	14.50	9.54	7.30	14.49	9.54	7.30
		0.2	6.91	4.96	4.10	6.90	4.95	4.10
		0.1	3.88	2.91	2.48	3.87	2.90	2.48

定理 1(ii)描述了 $F_{\epsilon,C,m}(\alpha;\pi_0)$ 如何随 $\log(1/\alpha)$ 增长，且该结论对 m 一致成立。这有助于理解其在 $\alpha \to 0$ 时的表现。具体而言，结合 (8) 可知，BH程序的错误发现率至多为 $\alpha \log^{\max\{1,k^*/2\}}(1/\alpha)$ 阶，且对 m, μ 和 $\Sigma \in \mathcal{C}_m(\epsilon, C)$ 一致成立。特别地，当 $\alpha \to 0$ 时，它趋于零。当 $\epsilon \geq 1/2$ 时，可以取 $k^* \leq 2$，此时其阶为 $\alpha\log(1/\alpha)$。需要强调的是，常数 $A(\epsilon, C, \pi_0)$ 和 $B(\epsilon, C, \pi_0)$ 仅用于演示，但取值较为保守。实际应用中，我们建议使用 (7) 计算 $F_{\epsilon,C,m}(\alpha;\pi_0)$ 的值。

经典错误发现率界要么假设正回归依赖，要么按 log 增长 m （Ben- jamini 和 Yekutieli，2001年；Sarkar，2002年），而弱依赖下的结果是渐近的 m（Farcomeni，2007；Wu，2008；Clarke 和 Hall，2009；Delattre 和 Roquain，2016）。表2 将定理1 与近期不随 m增长的界进行比较，其速率表示为 $\alpha \to 0$。

表 2： *BH*程序的错误发现率上界不依赖于 m。

结果	设定	错误发现率上界
Su (2018)	真零假设 p之间的正回归依赖-值。	$O(\alpha \log(1/\alpha))$
Chi 等(2025)	真零假设 p 值之间的弱负依赖——值。	$O(\alpha \log(1/\alpha))$
Zhang (2026)	真零统计量之间协方差有界的双侧高斯检验-相关结构。	$O(\alpha \log(1/\alpha))$
Lei (2026)	一类共同因子中的双侧高斯检验模型。	$O(\alpha\sqrt{\text{对数}(1/\alpha)})$
Lyu 和 Wang (2026)	具有秩一依赖的单侧高斯检验以及单调的非零假设 p-值映射。	$2W(1)\alpha$ 对于 $\alpha < 1/2$
定理 1	具有 $\Sigma \in \mathcal{C}_m(\epsilon, C)$ 的单侧和双侧高斯检验对于某个 $\epsilon C > 0$, 其中 k^* 是.....中的最小元素 $\mathcal{S}$，其中 $\epsilon k^* \geq 1$。	$O\big(\alpha\, \text{对数}^{\max\{1,k^*/2\}}(1/\alpha)\big)$

2.3 证明概要

~设 $\xi = T - \mu$，并设 $N(w)$ 为其 p值不超过 w的真实零假设数量。每个~被拒绝的真实零假设都会在 BH 阈值处由 N 计数，因此错误发现比例可被~的上确界所界定，该上确界取于 $N(u)/(mu)$ 在 BH 所使用的阈值范围内。该上确界仅通过真实零假设集合依赖于 μ （引理1），而错误发现率可由其尾概率的积分界定。因此，证明的主要目标是为这一标准化上确界建立一个不依赖于 m的尾界。对其积分后，即可得到一个形式为 α 乘以一个与 m无关的常数的错误发现率界。

为获得尾界，我们首先将阈值限制在二进层级上，这会带来一个因子 2 的代价（引理2）。在每个层级上，我们将由 ξ 生成的超越指示变量展开为厄米多项式，并将中心化计数拆分为低阶项和余项（引理 4）。低阶项除确定性系数外，在每个层级上都是相同的随机变量，因此只需控制一次，而不是每个层级控制一次，从而避免了对对数数量层级进行并集界。余项在粗层级上通过切比雪夫不等式控制，在细层级上则通过一阶矩界控制。在各层级上取二者中更优的界并求和，只需付出一个常数因子（引理 3）。余项的方差（命题 1）包含一个局部部分，它来自有限个强相关和中等相关的滞后项；还包含一个长程部分，它来自大量弱相关滞后项。这两部分仅通过相关系数的绝对值依赖于相关性，并且随相关性增大而增大，因此只需在包络函数处对其进行评估。当 m 且 $\epsilon k^* \geq 1$ 时，长程部分在各层级求和后仍保持有界（第 4.5节）。它在$1/\alpha$中按多重对数增长，这给出了定理 1(ii)中的速率。当 $\epsilon \geq 1$ 适用于单侧检验，或 $\epsilon \geq 1/2$ 适用于双侧检验时，可以取 $k^* = \min \mathcal{S}$。此时无需低阶项，该论证便简化为直接的二阶矩界。

2.4 依赖条件的适用范围

本节讨论若干满足定理 1 依赖条件的高斯依赖模型，以及若干不满足这些条件的模型。对于每个模型，我们先考察其相关性，再检验其相关矩阵是否属于 $\mathcal{C}_m(\epsilon, C)$。对于相关序列为 ρ 的平稳模型，只要对所有 $k \geq 1$ 都有 $|\rho(k)| \leq \rho_{\epsilon,C}(k)$，这一条件便成立。由于 Σ 是相

关矩阵 $T-\mu$，我们通过 $\xi_i = T_i - \mu_i$ 来描述每个模型。设 Z_i，$i \in \mathbb{Z}$ 为相互独立的标准正态变量。表 3 总结了这些模型。

表3：高斯依赖模型及类 $\mathcal{C}_m(\epsilon, C)$。

模型	相关性	涵盖	(ϵ, C) 的选择
AR(1)，$\lvert a\rvert < 1$	$\rho(k) = a^k$	Yes	$\epsilon > 0$, $C =$上确界$_{k\geq 1} \lvert a\rvert^k$ 对数$^{1+\epsilon}(k+e)$
MA(q)	$\rho(k) = 0$ 对于 $k > q$	Yes	$\epsilon > 0$, $C =$ 对数$^{1+\epsilon}(q+e)$
分数高斯噪声，$H \in (1/2,\ 1)$	$\rho(k) \sim H(2H-1)k^{2H-2}$	Yes	$\epsilon > 0$，一个合适的 C
大小为至多 B的独立区块	$\Sigma_{ij} = 0$ 对于 $\lvert i-j\rvert \geq B$	Yes	$\epsilon > 0$, $C = \log^{1+\epsilon}(B-1+e)$
等相关，固定 $r > 0$	$\Sigma_{ij} = r$ 对于所有 $i \neq j$	No	无，因为 $\Sigma_{1m} = r$ 而 $\rho_{\epsilon,\ C}(m-1) \to 0$
单因子，$\lvert\lambda_i\rvert \geq \lambda > 0$	$\Sigma_{ij} = \lambda_i\lambda_j$	No	无，因为 $\lvert\Sigma_{1m}\rvert \geq \lambda^2$ 而 $\rho_{\epsilon,\ C}(m-1) \to 0$
不断增大的区块，b 固定的 $r > 0$	$\Sigma_{ij} = r$ 在一个区块内	No	无，因为 $C \geq r\log^{1+\epsilon}(b-1+e) \to \infty$ 是必要的

涵盖的模型。(i). *AR(1)*模型。设 $\xi_i = a\xi_{i-1} + \sqrt{1-a^2}\,Z_i$, $i = \overline{1,2,\ldots}$ 为平稳的，且 $\lvert a\rvert < 1$。那么 $\rho(k) = a^k$。由于 $\lvert a\rvert^k \log^{1+\epsilon}(k+e) \to 0$，$\lvert\rho(k)\rvert\log^{1+\epsilon}(k+e)$ 在 $k \geq 1$ 上的上确界有限，因此对于每个 $\epsilon > 0$，存在适当的 C，使得 $\Sigma \in \mathcal{C}_m(\epsilon, C)$。对于 $a < 0$，相关系数的符号交替变化。

(ii). 移动平均模型。设 $\xi_i = \sum_{j=0}^{q} b_j Z_{i-j}$, $i = 1, 2, \ldots$ 为平稳的，且 $\sum_{j=0}^{q} b_j^2 = 1$。那么对于 $1 \leq k \leq q$，$\rho(k) = \sum_{j=0}^{q-k} b_j b_{j+k}$；对于 $k > q$，$\rho(k) = 0$。由于 $\lvert\rho(k)\rvert \leq 1$，对于每个 $\epsilon > 0$，当 $C = \log^{1+\epsilon}(q+e)$ 时，我们有$\Sigma \in \mathcal{C}_m(\epsilon, C)$。

(iii). 分数高斯噪声模型。令 $\xi_i = B_H(i) - B_H(i-1)$ 为一个平稳过程，其中 B_H 是具有赫斯特参数 $H \in (1/2,1)$ 的分数布朗运动。于是

$$\rho(k) = \frac{1}{2}[(k+1)^{2H} - 2k^{2H} + (k-1)^{2H}] \sim H(2H-1)k^{2H-2}.$$

由于 $k^{2H-2}\log^{1+\epsilon}(k+e) \to 0$，我们有 $\Sigma \in \mathcal{C}_m(\epsilon, C)$ 对每个 $\epsilon > 0$，其中适当的 C。这些相关性不可求和，因此定理1允许长程依赖。

(iv). 有界大小的独立区块。将 $\{1, \ldots, m\}$ 划分为大小至多为 B的连续区块。令 ξ 在不同区块上的子向量相互独立，而每个区块内具有任意的相关矩阵。则对于 $\lvert i-j\rvert \geq B$有 $\Sigma_{ij} = 0$，因为这样的 i 和 j 位于不同的区块中。对于 $1 \leq \lvert i-j\rvert \leq B-1$，我们有$\lvert\Sigma_{ij}\rvert \leq 1 = \rho_{\epsilon,C}(\lvert i-j\rvert)$，其中 $C = \log^{1+\epsilon}(B-1+e)$。因此$\Sigma \in \mathcal{C}_m(\epsilon, C)$。也就是说，该相关矩阵属于所考虑的类。

未涵盖的模型。(i). 等相关。令 $\xi_i = \sqrt{r}\,Z_0 + \sqrt{1-r}\,Z_i$ 具有固定的 $r \in (0,1]$。于是，对于所有 $k \geq 1$，$\rho(k) = r$。由于 $r\log^{1+\epsilon}(k+e) \to \infty$，因此不存在满足 $\Sigma \in \mathcal{C}_m(\epsilon, C)$ 的 C，使其对所有 m都成立。

(ii). 单因子模型。设 $\xi_i = \lambda_i Z_0 + \sqrt{1-\lambda_i^2}\,Z_i$with $\lvert\lambda_i\rvert \leq 1$。那么 Corr$(\xi_i, \xi_j) = \lambda_i\lambda_j$。如果对所有 i都有 $\lvert\lambda_i\rvert \geq \lambda > 0$，那么 $\lvert\Sigma_{1m}\rvert \geq \lambda^2$；而 $\rho_{\epsilon,C}(m-1) \to 0$。因此，当 m 足够大时，无论 C 是什么，相关矩阵都不属于 $\mathcal{C}_m(\epsilon, C)$。这包括 Lyu 和 Wang (2026) 的秩一模型，其中 $\lvert\lambda_i\rvert = 1$。Lei (2026) 的共同因子模型形式相同，其相关性也不会随距离增加而衰减。

(iii). 增长块。 设一个大小为 b 的块具有固定的块内相关系数 $r > 0$。其两个端点在滞后 $b-1$ 处的相关系数为 r。因此，界 $|\Sigma_{ij}| \le \rho_{\epsilon,\, C}(|i-j|)$ 要求$C \ge r\log^{1+\epsilon}(b-1+e)$，但当 $b \to \infty$ 时这一条件不成立。

最后需要指出的是，AR(1)无法由任何单因子模型生成。每个单因子相关矩阵都满足 $\Sigma_{12}\Sigma_{34} = \Sigma_{13}\Sigma_{24}$，而AR(1)给出 $a^2 \neq a^4$ 对于 $0 < |a| < 1$。从这个意义上说，定理1 和因子模型结果涵盖了不同的依赖结构。

3 实验

本节对我们的错误发现率界（8）进行数值评估，并将其与 Benjamini 和 Yekutieli（2001）的 BY界以及 BH 的经验错误发现率进行比较。我们取 $\alpha = 0.1$ 和 $\pi_0 = 1$，即最不利值。我们考虑了两种衰减率 $\epsilon = 0.5$ 和 $\epsilon = 1$，以及三种尺度$C = 0.01$、0.05和0.1。对于单侧情形，当 $\epsilon = 0.5$ 时我们使用 $k^* = 2$，当 $\epsilon = 1$时使用 $k^* = 1$。对于双侧 p值，仅偶数 Hermite阶数出现，因此在两种情形下我们都使用 $k^* = 2$。在每种情形下，$\epsilon k^* \ge 1$。对于每个 m，我们使用包络函数ρ_{ϵ}、C计算公式（8）中给出的错误发现率界。常数 K 根据 Plackett公式（Plackett，1954年）计算得到。量$E_{\rho_{\epsilon}}$、C、m（α）直接计算得到。其中的 θ 被选取为使常数最小。我们通过蒙特卡洛方法估计 BH 的错误发现率。我们取全局零假设 $\mu = 0$，使得 $\pi_0 = 1$。相关系数选为$\rho = \rho_{\epsilon}$、C。在全局零假设下，错误发现率等于至少一次拒绝的概率。单侧 p值为 $1 - \Phi(T_i)$，双侧值为2（$1 - \Phi(|T_i|)$）。我们考察了 $m = 2^a$ 在 $a = 2$、...、...、20时的情形。错误发现率通过20000次独立重复试验计算。

图1和图2展示了结果。Benjamini 和 Yekutieli（2001年）的 BY 界会像$\alpha\log\, m$ 那样不断增长，并在 $m = 2^{20}$时达到 1.44。我们的界仅在 m较小时与之重合，而在 m较大时保持平稳。在单侧情形下，对于 $\epsilon = 0.5$ 和 $\epsilon = 1$，它在 $m = 2^{13}$ 到 2^{14} 之间与 BY 界分离。此后它保持平稳：当 $\epsilon = 1$时，数值为 0.97 至 0.98。当 $\epsilon = 0.5$ 时，对于 $C = 0.01$，它保持在 0.996；对于 $C = 0.05$ 和 0.1，则达到平凡值 1。较大的 C 或较小的ϵ 意味着相关性更强，并会使界升高、分离点右移。在双侧情形下，六个界几乎相同。它们在$m = 2^{13}$ 到 2^{14} 之间与 BY 界分离，并保持在 0.97。所有情况下，经验错误发现率都介于 0.09 和 0.11 之间，在 m中没有增长。

致谢。 Claude Opus 5.5（Anthropic）协助完成了若干证明的开发工作，例如 Hermite 展开和协方差界（引理 4–6），并协助编辑了手稿。所有结果均已由作者核查，作者对内容承担全部责任。

4 附录：证明

本节证明定理1。全文使用第 2.2 节的记号。论证首先通过其滞后包络处理任意相关矩阵 ρ，该包络定义如下；随后专门处理包络 $\rho_{\epsilon,\, C}$。第 4.1 节将第 2.2 节中的量推广到任意ρ，并将问题化归为一个中心化高斯序列。第 4.2节汇总高斯与厄米特事实。第 4.3 节给出一个方差的界。第 4.4 节将这些结果结合起来，得到单个相关矩阵的错误发现率界（命题2）。第 4.5 节在该包络上计算此界。在第 4.1–4.4节中，$k^* \in \mathcal{S}$ 是任意的。条件 $\epsilon k^* \ge 1$ 仅在第 4.5节中使用。

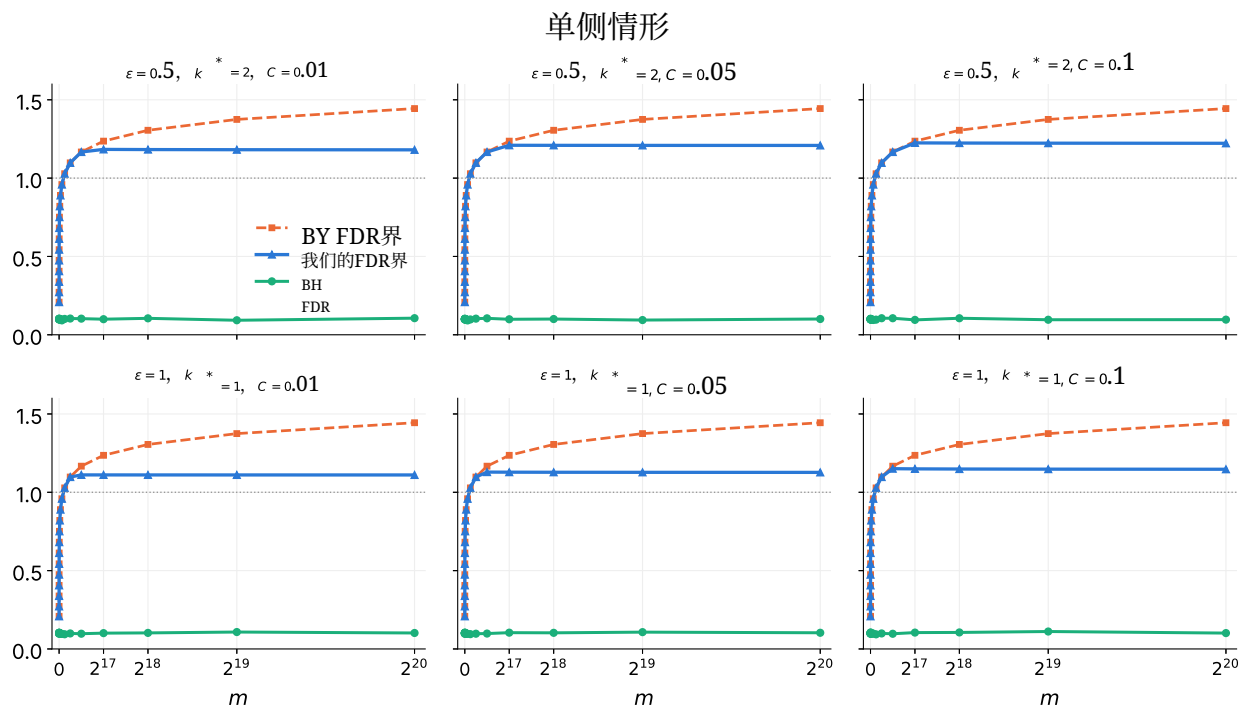

图1： 单侧情形。*BH*程序的错误发现率及其上界随 m 的变化。虚线：*BY FDR*界。实线：我们的*FDR*界。圆点：*BH FDR*。点线标示 *1*。

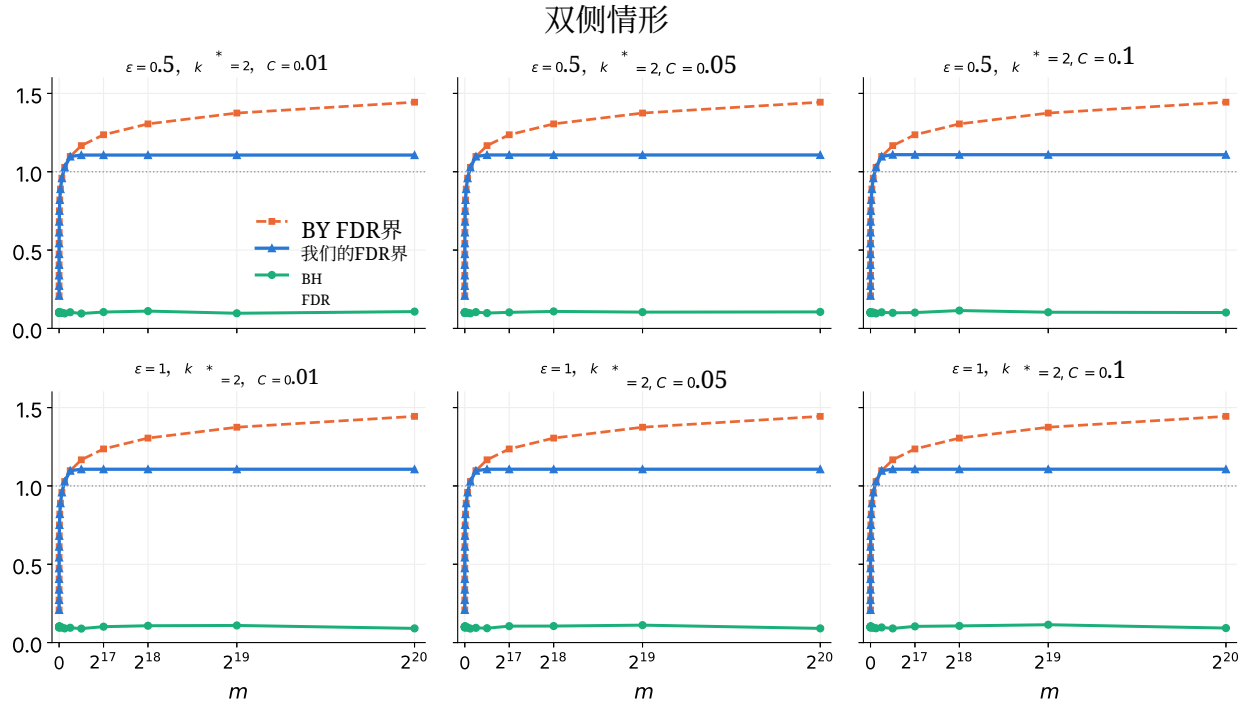

图2： 双侧情形。*BH*程序的错误发现率及其上界随 m 的变化。虚线：*BY FDR*界。实线：我们的*FDR*界。圆点：*BH FDR*。点线标示 *1*。

4.1 记号与约化为中心化序列

相关矩阵的滞后包络函数。对于一个 $m \times m$ 相关矩阵 Σ，定义其滞后包络函数ρ_Σ: $\mathbb{N}_0 \to [0,1]$ 令 $\rho_\Sigma(0) = 1$， $\rho_\Sigma(l) =$对 $1 \le l \le m-1$取最大值，以及 $\rho_\Sigma(l) = 0$ 对于 $l \ge m$。则对于所有 $i \ne j$，都有 $|\Sigma_{ij}| \le \rho_\Sigma(|i-j|)$。此外， $\Sigma \in \mathcal{C}_m(\epsilon, C)$ 当且仅当 $\Sigma \succeq 0$， $\Sigma_{ii} = 1$ 且$\rho_\Sigma \le \rho_{\epsilon, C}$ 逐点成立。

一般相关序列。对于 ρ: $\mathbb{N}_0 \to [-1,1]$ 满足 $\rho(0) = 1$，令 $D_\rho = \#\{l \ge 1 : |\rho(l)| > 1/2\}$，并令 K_ρ、G_ρ、$_k(m)$、e_ρ、$_m(z)$ 以及 Λ_ρ、$_m(\alpha)$ 分别定义为第 2.2 节中的 K_ϵ、$_C$、 G_ϵ、$_C(m)$、e_ϵ、$_C$、$_m(z)$ 以及Λ_ϵ、$_C$、$_m(\alpha)$，其中将 ρ_ϵ、$_C(l)$ 替换为 $|\rho(l)|$，对每个 $l \ge 1$ 均如此。也就是说，

$$K_\rho = 1 + \sup_{z \ge 0} \sum_{l \ge 1:\, |\rho(l)| > r_z} \tau(\rho(l), z), \qquad G_{\rho,k}(m) = \sum_{l=1}^{m-1} |\rho(l)|^k,$$

以及 $e_{\rho,\ m}$ 和 $\Lambda_{\rho,\ m}(\alpha)$ 可由$e_{\epsilon,\ C,\ m}$ 以及 $\Lambda_{\epsilon,\ C,\ m}(\alpha)$ 得到，方法是将 $G_{\epsilon,\ C,\ k}(m)$ 替换为 $G_{\rho,\ k}(m)$。具体而言，对于 $z \ge 0$,

$$\begin{aligned} e_{\rho,m}(z) &= 2\Big(\frac{\phi(z)}{1-\Phi(z)}\Big)^2 \mathsf{A}(z)\, \frac{G_{\rho,k^*}(m)}{m}, \\ \Lambda_{\rho,m}(\alpha) &= \begin{cases} \Big(\sum_{k \in \mathcal{S},\ k<k^*} b_k \Big[k!\Big(\frac{1}{m} + \frac{2G_{\rho,k}(m)}{m}\Big)\Big]^{1/2}\Big)^2, & k^* > \min \mathcal{S}, \\ 0, & k^* = \min \mathcal{S}. \end{cases} \end{aligned}$$

由于 $\rho_{\epsilon,\ C} \ge 0$，我们有 $K_{\rho_{\epsilon,\ C}} = K_{\epsilon,\ C}$, $e_{\rho_{\epsilon,\ C},\ m} = e_{\epsilon,\ C,\ m}$ 和 $\Lambda_{\rho_{\epsilon,\ C},\ m}(\alpha) = \Lambda_{\epsilon,\ C,\ m}(\alpha)$。对于 $z \le \sqrt{2}$，我们有 $r_z = 1/2$，因此 K_ρ 中的求和遍历满足 $|\rho(l)| > 1/2$ 的 D_ρ 个滞后，并且等于 $2D_\rho$。因此$K_\rho \ge 1 + 2D_\rho \ge 1$。

约化。始终， Σ是一个 $m \times m$ 相关矩阵， $\rho = \rho_\Sigma$，且 $\xi = T - \mu \sim N(0, \Sigma)$。令 $p(x) = 1 - \Phi(x)$（单侧）或 $p(x) = 2(1 - \Phi(|x|))$（双侧），从而对 $i \in \mathcal{H}_0$有 $P_i = p(\xi_i)$。对于 $w \in (0, 1/2]$，我们有 $z_w \ge 0$，并且 $p(\xi_i) \le w$ 当且仅当 $\xi_i \ge z_w$ （单侧），相应地$|\xi_i| \ge z_w$ （双侧）。置

$$A_i(w) = \mathbf{1}\{p(\xi_i) \le w\}, \qquad \widetilde{N}(w) = \sum_{i \in \mathcal{H}_0} A_i(w), \qquad \widetilde{S} = \sup_{u \in [\alpha/m, \alpha]} \frac{\widetilde{N}(u)}{mu},$$

~ ~使得 $\mathbb{E}A_i(w) = w$ 且 $\mathbb{E}N(w) = m_0 w = \pi_0 m w$。注意， N 仅统计真实零假设，而 ~S 仅通过 $\mathcal{H}_0$依赖于 μ。若 $m_0 = 0$，则 $V = 0$ 且 FDR $= 0$，因此每当使用 S时，我们都假设 $m_0 \ge 1$~。

引理 1. 错误发现比例 $\le$几乎处处$\{1, \alpha\widetilde{S}\}$。因此 错误发现率 $\le \alpha \int_0^{1/\alpha} \mathbb{P}(\widetilde{S} \ge t)\, dt$.

证明。 </style>若 $R = 0$，则错误发现比例 FDP $= 0$。否则 $u_R = \alpha R/m \in [\alpha/m, \alpha]$，且每个被拒绝的真实 ~ ~零假设都满足 $p(\xi_i) = P_i \le u_R$，因此 $V \le N(u_R)$ 且 FDP $= V/R = \alpha V/(mu_R) \le \alpha S$。最后，$\mathbb{E}$几乎处处为 $\{1, Y\} = \int_0^1 \mathbb{P}(Y \ge s)\, ds$，其中 $Y = \alpha\widetilde{S}$。证明完毕。 □

回顾 $J = \lceil \log_2 m \rceil$和 $w_j = 2^{-j}\alpha$，并令 $w_J = 2^{-J}\alpha$和 $a_j = m w_j$表示 $0 \le j \le J$。所有水平都满足 $w_j \le \alpha \le 1/2$。

引理 2. 对于每个 $t > 0$、 $\{\widetilde{S} \ge t\} \subseteq \bigcup_{j=0}^{J-1} \{\widetilde{N}(w_j) \ge \frac{t}{2} a_j\}$。

证明。由于 $w_J \le \alpha/m$，每个 $u \in [\alpha/m, \alpha]$ 都属于某个 $[w_{j+1}, w_j]$，且满足 $0 \le j \le J-1$，于是~ ~ ~$N(u)/(mu) \le N(w_j)/(m w_{j+1}) = 2N(w_j)/a_j$。证明完毕。 □

引理 3。 若 $a_0 > \cdots > a_{J-1} > 0$ 满足 $a_j = 2a_{j+1}$ 和 $b > 0$，则 $\sum_{j=0}^{J-1}$ 最小值$\{a_j, b/a_j\} \le 3\sqrt{b}$。

证明。 </style}带有 $a_j \le \sqrt{b}$ 的项构成公比为 $1/2$、首项为 $a_* \le \sqrt{b}$ 的几何级数，因此其和至多为 $2a_*$。带有 $a_j > \sqrt{b}$ 的项至多为 b/a_j，这些项构成公比为 $1/2$、首项为 b/a^* 的几何级数，其中 $a^* > \sqrt{b}$ 是满足条件的最小 a_j；它们的和至多为$2b/a^* < 2\sqrt{b}$。如果只有一组，合计至多为 $2\sqrt{b}$。如果两组都存在，$a^* = 2a_*$，则合计至多为 $2a_* + b/a_*$，且 $a_* \in [\sqrt{b}/2, \sqrt{b}]$；凸函数 $x \mapsto 2x + b/x$ 在两个端点处均等于 $3\sqrt{b}$。证明完毕。 □

4.2 高斯与 Hermite 事实

回顾 He_k，$R(r, z)$，r_z 以及A（z）来自第2.2节。设 $c_k(z) = \phi(z)\, He_{k-1}(z)/k!$ 对于 $k \ge 1$。

引理 4。设 $z \ge 0$ 和 $|r| < 1$。

(*i*)。令 $X \sim N(0, 1)$，

$$\begin{aligned} \mathbf{1}\{X > z\} &= 1 - \Phi(z) + \sum_{k\ge 1} c_k(z) He_k(X), \\ \mathbf{1}\{|X| > z\} &= 2(1 - \Phi(z)) + \sum_{k\in 2\mathbb{N}} 2c_k(z) He_k(X), \end{aligned}$$

其中 $c_k(z) = \phi(z)\, He_{k-1}(z)\, k!$ 并且该级数在均方意义下收敛。例如，

$$\mathbb{E}\Big[\Big(\mathbf{1}\{X > z\} - \big(1 - \Phi(z)\big) - \sum_{k=1}^{n} c_k(z) He_k(X)\Big)^2\Big] \to 0 \ \ as\ n \to \infty.$$

(*ii*)。如果（X，Y）服从相关系数为 r 的标准二元正态分布，那么对于任意 k、$l \ge 1$，

$$\mathbb{E} He_k(X) He_l(Y) = \delta_{kl}\, k!\, r^k.$$

因此

$$\begin{aligned} R(r, z) &= \phi(z)^2 \sum_{k\ge 1} \frac{He_{k-1}(z)^2}{k!} r^k, \\ \mathrm{Cov}(\mathbf{1}\{|X| > z\}, \mathbf{1}\{|Y| > z\}) &= 4\phi(z)^2 \sum_{k\in 2\mathbb{N}} \frac{He_{k-1}(z)^2}{k!} r^k. \end{aligned}$$

(iii). $R(r, z) = \phi(z)^2 \int_0^r e^{z^2 s/(1+s)}(1 - s^2)^{-1/2}\, ds$. 因此，$R(r, z)$与 r同号，$R(\cdot, z)$递增，$|R(r, z)| \le R(|r|, z)$，且*(ii)*中的双侧协方差等于 $2[R(r, z) + R(-r, z)] \in [0, 2R(|r|, z)]$。

证明。*(i)*。对 $\phi^{(k-1)} = (-1)^{k-1} He_{k-1}\phi$ 求导得

$$\big(He_{k-1}\phi\big)' = (-1)^{k-1}\phi^{(k)} = -He_k\phi.$$

由于 $He_{k-1}(x)\phi(x) \to 0$ 作为 $x \to \infty$，对于 $k \ge 1$，

$$\mathbb{E}\big[\mathbf{1}\{X > z\} He_k(X)\big] = \int_z^\infty He_k(x)\phi(x)\, dx = -\Big[He_{k-1}(x)\phi(x)\Big]_z^\infty = \phi(z) He_{k-1}(z),$$

而 $\mathbb{E}[\mathbf{1}\{X>z\}He_0(X)]=1-\Phi(z)$。多项式 $He_k/\sqrt{k!}$、 $k\ge 0$ 构成函数空间 f（满足 $\mathbb{E}f(X)^2<\infty$）的标准正交基，该空间配备内积 $\mathbb{E}[f(X)g(X)]$；特别地， $\mathbb{E}He_k(X)^2=k!$ 由于 $\mathbf{1}\{x>z\}$ 有界，它在均方意义下等于其在该基底下的展开，而 He_k 的系数为

$$\frac{\mathbb{E}[\mathbf{1}\{X>z\}He_k(X)]}{\mathbb{E}He_k(X)^2}=\frac{\phi(z)He_{k-1}(z)}{k!}=c_k(z),\qquad k\ge 1.$$

这证明了第一个展开式。对于第二个， $\mathbf{1}\{|x|>z\}=\mathbf{1}\{x>z\}+\mathbf{1}\{-x>z\}$ 因为$z\ge 0$。由于 X 和 $-X$ 具有相同的分布，且 $He_k(-x)=(-1)^kHe_k(x)$,

$$\mathbb{E}\big[\mathbf{1}\{-X>z\}He_k(X)\big]=\mathbb{E}\big[\mathbf{1}\{X>z\}He_k(-X)\big]=(-1)^k\phi(z)He_{k-1}(z).$$

因此，在 $\mathbf{1}\{|x|>z\}$的展开式中， He_k的系数为 $(1+(-1)^k)c_k(z)$；当 k 为偶数时，该系数等于 $2c_k(z)$，当 k为奇数时则为 0，而常数项为 $2(1-\Phi(z))$。

(ii)。厄米多项式的生成函数为 $e^{sx-s^2/2}=\sum_{k\ge0}He_k(x)s^k/k!$ 令$Y=rX+\sqrt{1-r^2}\,W$，其中 $W\sim N(0,\mathbf{1})$与 X相互独立；这不会改变 $(X,\ Y)$ 的联合分布。对于实数 $s,\ t$,

$$\begin{aligned}\mathbb{E}\big[e^{sX-s^2/2}e^{tY-t^2/2}\big]&=e^{-(s^2+t^2)/2}\,\mathbb{E}e^{(s+rt)X}\,\mathbb{E}e^{t\sqrt{1-r^2}\,W}\\&=e^{-(s^2+t^2)/2}e^{(s+rt)^2/2}e^{t^2(1-r^2)/2}=e^{rst}.\end{aligned}$$

将两边按 s和 t 的幂展开，并比较 s^kt^l 的系数，得到

$$\frac{\mathbb{E}[He_k(X)He_l(Y)]}{k!\,l!}=\delta_{kl}\frac{r^k}{k!},$$

即,

$$\mathbb{E}He_k(X)He_l(Y)=\delta_{kl}\,k!\,r^k.\tag{13}$$

接下来令 $F_n(x)=\sum_{k=1}^n c_k(z)He_k(x)$。由 (i)， $F_n(X)\to\mathbf{1}\{X>z\}-(1-\Phi(z))$ 和 $F_n(Y)\to\mathbf{1}\{Y>z\}-(1-\Phi(z))$ 在均方意义下收敛，因此根据柯西–施瓦茨不等式， $\mathbb{E}[F_n(X)F_n(Y)]\to R(r,z)$。因此

$$R(r,z)=\lim_{n\to\infty}\sum_{k,l=1}^n c_k(z)c_l(z)\,\mathbb{E}\big[He_k(X)He_l(Y)\big]=\sum_{k\ge1}k!\,c_k(z)^2r^k=\phi(z)^2\sum_{k\ge1}\frac{He_{k-1}(z)^2}{k!}r^k.$$

双侧公式可由(i)中的第二个展开以相同方式推出，其中将 c_k 替换为 $2c_k$ （偶数 k 时），奇数 k时则替换为 0。

(iii)。令 $\phi_2(x,y;r)$ 表示相关系数为 r的标准二元正态密度。它满足$\partial_r\phi_2=\partial_x\partial_y\phi_2$ (Plackett, 1954年)，因此

$$\frac{\partial}{\partial r}\mathbb{P}(X>z,Y>z)=\int_z^\infty\int_z^\infty\partial_x\partial_y\phi_2(x,y;r)\,dx\,dy=\phi_2(z,z;r)=\frac{1}{2\pi\sqrt{1-r^2}}\exp\Big\{-\frac{z^2}{1+r}\Big\}.$$

由于 $-z^2/(1+r)=-z^2+z^2r/(1+r)$ 且 $\phi(z)^2=e^{-z^2}/(2\pi)$，右侧等于 $\phi(z)^2e^{z^2r/(1+r)}(1-r^2)^{-1/2}$。在 $r=0$ 时， $\mathbb{P}(X>z,Y>z)=(1-\Phi(z))^2,$，也就是说， $R(0,z)=0$。从 0 积分到 r，得到 $R(r,z)$ 的公式。被积函数为正，因此 $R(r,z)$ 的符号与 r 相同，且 $R(\cdot,z)$ 递增。对于 $r<0$，代入 $s\mapsto -s$,

$$|R(r,z)|=\phi(z)^2\int_0^{|r|}\frac{e^{-z^2s/(1-s)}}{\sqrt{1-s^2}}\,ds\le\phi(z)^2\int_0^{|r|}\frac{e^{z^2s/(1+s)}}{\sqrt{1-s^2}}\,ds=R(|r|,z).$$

对于双侧协方差，写作 $\mathbf{1}\{|X|>z\}=\mathbf{1}\{X>z\}+\mathbf{1}\{-X>z\}$，同理对于 Y，并将协方差展开为四项。数对 $(X,\ Y)$ 和 $(-X,\ -Y)$ 的相关系数为 r，而数对 $(X,\ -Y)$ 和 $(-X,\ Y)$ 的相关系数为 $-r$，因此

$$\mathrm{Cov}\left(\mathbf{1}\{|X|>z\},\mathbf{1}\{|Y|>z\}\right)=2\big[R(r,z)+R(-r,z)\big].$$

根据 (ii)，这是一个关于 r 的偶次幂级数，且系数均为非负数，因此结果非负。最后，$R(r,z)+R(-r,z)=R(|r|,z)+R(-|r|,z)\le R(|r|,z)$，因为 $R(-|r|,z)\le 0$。证明完成。 □

引理 5。令 $z\ge 0$ 和 $h(z)=(z+\sqrt{z^2+4})/2$。

(i)。 $\phi(z)/(1-\Phi(z))\le h(z)$（*Birnbaum，1942*），$h(z)^2\le z^2+2$，$h(z)\le 1+z$，以及 $1-\Phi(z)\le \frac{1}{2}e^{-z^2/2}$。因此，$z_w^2\le 2\log\frac{1}{2w}$（单侧）和 $z_w^2\le 2\log\frac{1}{w}$（双侧）。

(ii)。如果 $|r|\le 1/2$，则 $R(|r|,z)/(1-\Phi(z))\le (1+z)e^{-z^2/6}/\sqrt{6\pi}$。

(iii)。 $|He_n(z)|\le(z+\sqrt{n+1})^n$ 对于 $n\ge 0$。

证明。 (i)。 $h^2=\frac{1}{2}(z^2+2+z\sqrt{z^2+4})\le z^2+2$ 因为 $z^2(z^2+4)\le(z^2+2)^2$。其他结论是标准的。 (ii)。由引理 4 $s/\sqrt{1+s}\le 1/3$ $1-s^2{}^{-1/2}\le 2/\sqrt{3}$ $0\le s\le 1/2$ $R(|r|,z)\le$ (iii)，对于，（）和（），因此（，）$\phi(z)e^{z^2/3}/\sqrt{3}$ $1-\Phi(z)\ge\phi(z)/\sqrt{1+z}$（）；结合由 (i) 得到的（） （） （）。

（iii）。 $He_n(z)=\mathbb{E}(z+\mathrm{i}Y)^n$ 其中 $Y\sim N(0,1)$，因此根据闵可夫斯基不等式 $|He_n(z)|\le(z+\|Y\|_n)^n$；此外 $\|Y\|_q^q=(q-1)!!\le q^{q/2}$ 对于偶数 q，所以 $\|Y\|_n\le\sqrt{n+1}$（奇数 n时使用 $\|Y\|_n\le\|Y\|_{n+1}$）。证明完成。 □

根据 $\mathsf{A}(z)$ 的定义，逐项来看，对于 $|r|\le r_z$，

$$\sum_{k\in\mathcal{S},\ k\ge k^*}\frac{He_{k-1}(z)^2}{k!}|r|^k\le\mathsf{A}(z)\,|r|^{k^*}. \tag{14}$$

下面的引理给出了在不同 k 取值下 $\mathsf{A}(z)$ 的上界 *.

引理 6。设 $z\ge 0$。

(i). 在单侧情形下，若选择 $k^*=1$（$\mathcal{S}$ 中的最小元素），则 $\mathsf{A}(z)\le a_1:=2(e-1)/\sqrt{3}<1.985$。

(ii). 在双侧情形下，若选择 $k^*=2$（$\mathcal{S}$ 中的最小元素），则 $\mathsf{A}(z)\le a_2z^2$ 其中$a_2:=2/\sqrt{3}<1.155$。

(iii). 对任意 $k^*\in\mathcal{S}$，无论是单侧情形还是双侧情形，均有 $\mathsf{A}(z)\le a_1\max\{2,z^2\}^{k^*-1}$。

在(i) 和(ii) 中，定义$\mathsf{A}(z)$ 的级数是完整协方差级数，Plackett公式可将其求和为闭式。(iii) 部分将 $\mathsf{A}(z)$ 与同一级数进行比较，并且当(i) 时，它化简为 $k^*=1$。

证明。令 $x=z^2r_z\le 1$ 并对 $0\le s\le r_z\le 1/2$ 使用 $(1-s^2)^{-1/2}\le 2/\sqrt{3}$.
(i)。由引理4(ii)，(iii)以及 $(e^x-1)/x$ 是递增的，可得

$$\mathsf{A}(z)=R(r_z,z)/(\phi^2r_z)\le\frac{2}{\sqrt{3}}r_z^{-1}\int_0^{r_z}e^{z^2s}ds=\frac{2}{\sqrt{3}}\frac{e^x-1}{x}\le\frac{2(e-1)}{\sqrt{3}}.$$

(ii)。根据引理 4，$\mathsf{A}(z)=[R(r_z,\ z)+R(-r_z,\ z)]/(2\phi^2r_z^2)$ 以及

$$R(r,z)+R(-r,z)=\phi^2\int_0^r(1-s^2)^{-1/2}\big(e^{z^2s/(1+s)}-e^{-z^2s/(1-s)}\big)ds.$$

利用 $e^u - e^{-v} \le ue^u + v$ 以及 $u = z^2 s/(1+s) \le z^2 s$ 和 $v = z^2 s/(1-s) \le 2z^2 s$，我们有

$$\mathsf{A}(z) \le \frac{1}{\sqrt{3}\, r_z^2} \int_0^{r_z} z^2 s \left(e^{z^2 s} + 2\right) ds = \frac{z^2}{\sqrt{3}} \Big(\frac{x e^x - e^x + 1}{x^2} + 1 \Big) \le \frac{2z^2}{\sqrt{3}},$$

因为 $(xe^x - e^x + 1)/x^2$ 递增，并且在 $x = 1$处等于 1。

(iii)。固定 $k^* \in \mathcal{S}$。由于 $r_z^{k-k^*} = r_z^{-k^*} r_z^k$ 且下面所有项均非负，

$$\mathsf{A}(z) = r_z^{-k^*} \sum_{k \in \mathcal{S},\ k \ge k^*} \frac{He_{k-1}(z)^2}{k!}\, r_z^k \le r_z^{-k^*} \sum_{k \ge 1} \frac{He_{k-1}(z)^2}{k!}\, r_z^k = r_z^{-k^*} \frac{R(r_z, z)}{\phi(z)^2},$$

其中最后一步依据引理4(ii)。根据引理 4(iii)，如同(i)中一样，

$$\frac{R(r_z, z)}{\phi(z)^2} \le \frac{2}{\sqrt{3}} \int_0^{r_z} e^{z^2 s}\, ds = \frac{2}{\sqrt{3}}\, r_z\, \frac{e^x - 1}{x} \le \frac{2(e-1)}{\sqrt{3}}\, r_z = a_1 r_z.$$

因此 $\mathsf{A}(z) \le a_1 r_z^{1-k^*} = a_1 \max\{2, z^2\}^{k^*-1}$，因为 $r_z^{-1} = \max\{2, z^2\}$。证明完成。 □

4.3 方差界

固定 $k^* \in \mathcal{S}$ 以及一个水平 $w \in (0, 1/2]$，其阈值为 $z = z_w$。令 $\gamma_k = c_k(z)$（单侧）或 $\gamma_k = 2c_k(z)$（双侧），并且

$$Q_i(w) = A_i(w) - w - \sum_{k \in \mathcal{S},\ k < k^*} \gamma_k He_k(\xi_i).$$

对于 $k^* = \min \mathcal{S}$，该和为空，且 $Q_i(w) = A_i(w) - w$。

命题 1. 设 $z = z_w$，并设 $I \subseteq \{1, \dots, m\}$ 满足 $|I| = m_0 \ge 1$。则 方差 $\big(\sum_{i \in I} Q_i(w)\big) \le K_\rho\, m_0 w + e_{\rho,\ m}(z_w)\, m_0 m w^2$。

证明。 回顾 $w \le 1/2$，因此 $z = z_w \ge 0$，引理 4–6 适用。展开和的方差，

$$\operatorname{Var}\Big(\sum_{i \in I} Q_i(w) \Big) = \sum_{i \in I} \operatorname{Var} Q_i(w) + \sum_{i, i' \in I,\ i \ne i'} \operatorname{Cov}\big(Q_i(w), Q_{i'}(w)\big). \tag{15}$$

我们在步骤 1 中界定对角项，在步骤 2 中推导单个协方差的三个界，并在步骤 3 中按滞后逐项对协方差求和。

步骤 *1*：对角项。 由于 $He_0 \equiv 1$且（13），我们有 $\mathbb{E} H_k(\xi_i) = 0$当 $k \in \mathcal{S}$时，因此$\mathbb{E} Q_i(w) = 0$。由引理 4(i)，我们有

$$\operatorname{Var} Q_i(w) = \mathbb{E} Q_i(w)^2 \le \mathbb{E}\big(A_i(w) - w\big)^2 = w(1-w) \le w.$$

(15) 中的对角项因此至多贡献 $m_0 w$。

*St*第 *2* 步：单个协方差。固定 $i \ne i'$ 并令 $r = \Sigma_{ii'}$，于是 $|r| \le \rho(|i - i'|)$。S et

$$\sigma(r) = \sum_{k \in \mathcal{S},\ k \ge k^*} \frac{He_{k-1}(z)^2}{k!}\, r^k.$$

我们声称，在单侧情形和双侧情形下，

$$\operatorname{Cov}\big(Q_i(w), Q_{i'}(w)\big) = w^2 \Big(\frac{\phi(z)}{1 - \Phi(z)} \Big)^2 \sigma(r) \qquad \text{if } |r| < 1. \tag{16}$$

事实上，根据引理 4(i)，$Q_i(w) = \sum_{k\in\mathcal{S},\,k\ge k^*} \gamma_k He_k(\xi_i)$，$i'$ 也同理。数对 $(\xi_i, \xi_{i'})$ 服从相关系数为 r 的标准二元正态分布，因此引理 4(ii) 给出 $\mathrm{Cov}(Q_i, Q_{i'}) = \sum_{k\in\mathcal{S},\,k\ge k^*} k!\ \gamma_k^2 r^k$。在单侧情形下，$k!\ \gamma_k^2 = \phi(z)^2 He_{k-1}(z)^2/k!$ 以及 $w = 1-\Phi(z)$。在双侧情形下，$k!\ \gamma_k^2 = 4\phi(z)^2 He_{k-1}(z)^2/k!$ 以及 $w = 2(1-\Phi(z))$。在这两种情形下，这就是式 (16)。现在我们推导三个界。

(a). 平凡界。 由柯西–施瓦茨不等式和步骤 1，

$$|\operatorname{Cov}(Q_i, Q_{i'})| \le (\operatorname{Var} Q_i\ \operatorname{Var} Q_{i'})^{1/2} \le w \qquad \text{for every } r\in[-1,1].$$

(b). 与完整协方差的比较。 令 $|r|\le 1/2$。去掉限制条件 $k\ge k^*$ 并由引理 4(ii)，(iii)，可得

$$|\sigma(r)| \le \sum_{k\in\mathcal{S}} \frac{He_{k-1}(z)^2}{k!}|r|^k = \begin{cases} R(|r|,z)/\phi(z)^2 & \text{（单侧），} \\ [R(|r|,z)+R(-|r|,z)]/(2\phi(z)^2) \le R(|r|,z)/(2\phi(z)^2) & \text{（双侧）。} \end{cases}$$

将其分别代入带有 $w = 1-\Phi(z)$ 的（16）和 $w = 2(1-\Phi(z))$，两种情形均可得

$$|\operatorname{Cov}(Q_i, Q_{i'})| \le w\,\frac{R(|r|,z)}{1-\Phi(z)}.$$

结合 (a)，

$$\begin{aligned} |\operatorname{Cov}(Q_i, Q_{i'})| &\le w\min\Big\{1, \frac{R(|r|,z)}{1-\Phi(z)}\Big\}, \quad \text{if } |r|\le \tfrac12, \\ |\operatorname{Cov}(Q_i, Q_{i'})| &\le w, \quad \text{if } |r| > \tfrac12. \end{aligned}$$

根据 τ 的定义，两种情况可写为

$$|\operatorname{Cov}(Q_i, Q_{i'})| \le \frac{w}{2}\,\tau(r,z) \qquad \text{for all } r\in[-1,1]. \tag{17}$$

(c). 弱相关。 令 $|r|\le r_z$。则 $|\sigma(r)| \le \mathsf{A}(z)|r|^{k^*}$ 由 (14) 得到，而 (16) 给出

$$|\operatorname{Cov}(Q_i, Q_{i'})| \le w^2\Big(\frac{\phi(z)}{1-\Phi(z)}\Big)^2 \mathsf{A}(z)\,|r|^{k^*}. \tag{18}$$

界 (c) 关于 w 是二次的，且阶为 $|r|^{k^*}$，因此适用于大量弱相关配对。界 (b) 关于 w 是线性的，用于剩余的少数配对。

步骤 3：对配对求和。 按滞后 $l = |i-i'| \in \{1,\dots,m-1\}$ 对非对角线配对分组。每个 $i\in I$ 至多有两个满足 $|i-i'| = l$ 的配对对象 i'，因此对于每个 l，在 I 中至多有 $2m_0$ 个滞后为 l 的有序配对 (i,i')，且其绝对相关系数均不超过 $\rho(l)$。将滞后分为

$$L_1 = \{l : \rho(l) > r_z\}, \qquad L_2 = \{l \le m-1 : \rho(l) \le r_z\}.$$

对于 $l\in L_1$，我们使用（17）以及 $\tau(\Sigma_{ii'},\ z) \le \tau(\rho(l)\ \ z)$，这成立是因为 $\tau(r,\ z)$ 关于$|r|$非递减，由引理 4(iii)：

$$\sum_{l\in L_1}\ \sum_{i,i'\in I,\,|i-i'|=l} |\operatorname{Cov}(Q_i, Q_{i'})| \le \sum_{l\in L_1} 2m_0\cdot\frac{w}{2}\,\tau(\rho(l),z) = m_0 w\,S_\rho(z) \le m_0 w\,(K_\rho - 1),$$

其中 $S_\rho(z)=\sum_{l:\ |\rho(l)|>r_z}\tau(\rho(l)\ z)$，且 $K_\rho=1+\sup_{z'\ge 0}S_\rho(z')$。对于 $l\in L_2$，滞后为 l 的每一对均满足 $|\Sigma_{ii'}|\le\rho(l)\le r_z$，因此我们使用（18）以及 $|\Sigma_{ii'}|^{k^*}\le\rho(l)^{k^*}$：

$$\sum_{l\in L_2}\sum_{i,i'\in I,\,|i-i'|=l}|\operatorname{Cov}(Q_i,Q_{i'})|\le 2m_0\,w^2\Big(\frac{\phi(z)}{1-\Phi(z)}\Big)^2\mathsf{A}(z)\sum_{l=1}^{m-1}\rho(l)^{k^*}=e_{\rho,m}(z_w)\,m_0mw^2,$$

因为 $\sum_{l=1}^{m-1}\rho(l)^{k^*}=G_{\rho,\ k^*}(m)$ 和 $2m_0w^2G_{\rho,\ k^*}(m)=m_0mw^2\cdot 2G_{\rho,\ k^*}(m)/m$。

将步骤 1 和 3 代入（15）得

$$\operatorname{Var}\Big(\sum_{i\in I}Q_i(w)\Big)\le m_0w+(K_\rho-1)\,m_0w+e_{\rho,m}(z_w)\,m_0mw^2=K_\rho\,m_0w+e_{\rho,m}(z_w)\,m_0mw^2.$$

证明完毕。 □

4.4 一个有用的不等式

下面的引理总结了 Ψ_{π_0} 在（5）中的两个性质。记 $g(t)=g(t;\ K,\ E,\ \pi_0)$。

引理 7。设 $K\ge 1$、$E\ge 0$ 和 $\alpha\in(0,\ 1/2]$。

(i). $\Psi_{\pi_0}(K,E,\alpha)$ 关于 K、E 和 π_0 单调不减，且 $2\pi_0\le\Psi_{\pi_0}(K,E,\alpha)\le 1/\alpha$。

(ii). 令 $\pi_0\in(0,1]$。对每个 $s>0$,

$$\Psi_{\pi_0}(K,E,\alpha)\le 2\pi_0+s+6\pi_0\sqrt{K}\log_+\left(\frac{1/\alpha-2\pi_0}{s}\right)+\frac{4\pi_0E}{s}.\tag{19}$$

特别地，当 $s^*=3\pi_0\sqrt{K}+\sqrt{9\pi_0^2K+4\pi_0E}$ 时，

$$\begin{aligned}\Psi_{\pi_0}(K,E,\alpha)&\le 2\pi_0+2s^*-6\pi_0\sqrt{K}+6\pi_0\sqrt{K}\,\log_+\left(\frac{1/\alpha-2\pi_0}{s^*}\right)\\&\le\pi_0\big(2+6\sqrt{K}\big)+4\sqrt{\pi_0E}+6\pi_0\sqrt{K}\,\log_+\frac{1}{6\alpha\pi_0\sqrt{K}},\end{aligned}$$

其中 $\log_+x=\max\{\log\ x,0\}$.

证明。(i). 根据 (5)，$\Psi_{\pi_0}(K,E,\alpha)=\int_0^{1/\alpha}\psi(t)\,dt$，其中 $\psi(t)=1$ 对 $t\le 2\pi_0$ 成立，且 $\psi(t)=\min\{1,g(t)\}$ 对 $t>2\pi_0$ 成立。对每个 t，$\psi(t)$ 关于 K 和 E 单调不减。它关于 π_0 也单调不减，因为 $\pi_0/(t-2\pi_0)$ 和 $\pi_0/(t-2\pi_0)^2$ 随着 π_0 增大而增加（对 $t>2\pi_0$），且一旦 $\pi_0\ge t/2$，$\psi(t)=1$。最后，$0\le\psi\le 1$ 和 $\psi=1$ 在 $[0,2\pi_0]$ 上成立，由此得到这两个界。

(ii). 对每个 $s>0$，在 (5) 中，当 $2\pi_0<t\le 2\pi_0+s$ 时用 1、当 $t>2\pi_0+s$ 时用 $g(t)$ 对 $\min\{1,g(t)\}$ 进行上界估计，即得 (19)。这里的 $\log_+$ 用于处理 $2\pi_0+s\ge 1/\alpha$ 的情形，此时 g 的积分区间为空。取 $s=s^*$，即 $s^2-6\pi_0\sqrt{K}s-4\pi_0E=0$ 的正根。则 $4\pi_0E/s^*=s^*-6\pi_0\sqrt{K}$，从而得到第一个不等式。对于第二个不等式，$\sqrt{9\pi_0^2K+4\pi_0E}\le 3\pi_0\sqrt{K}+2\sqrt{\pi_0E}$ 得出 $s^*\le 6\pi_0\sqrt{K}+2\sqrt{\pi_0E}$，而 $s^*\ge 6\pi_0\sqrt{K}$ 和 $1/\alpha-2\pi_0\le 1/\alpha$ 给出 $\log_+((1/\alpha-2\pi_0)/s^*)\le\log_+(1/(6\alpha\pi_0\sqrt{K}))$。证明完毕。 □

对 $\theta\in(0,\ 1)$ 定义

$$E_{\rho,m}(\alpha)=\frac{\Lambda_{\rho,m}(\alpha)}{\theta^2}+\frac{1}{(1-\theta)^2}\sum_{j=0}^{J-1}e_{\rho,m}(z_j),$$

来自第4.1节的 $\Lambda_{\rho,m}(\alpha)$ 和 $e_{\rho,m}$。回顾，b_k 在 $\Lambda_{\rho,m}(\alpha)$ 中是关于 $0\le z\le z_{J-1}$ 的上确界，其中 z_{J-1} 是最大阈值。

命题 2。 设 Σ 是一个 $m \times m$ 相关矩阵，且 $\rho = \rho_\Sigma$。设 $\alpha \in (0, 1/2]$, $m \geq 2$ 和$\theta \in (0,1)$。对于单侧和双侧 p 值，以及任意 $\mu \in \mathbb{R}^m$,

$$\mathrm{FDR}(\alpha; m, \mu, \Sigma) \leq \alpha \min\Big\{\pi_0 H_m,\ \Psi_{\pi_0}\Big(\frac{K_\rho}{(1-\theta)^2},\, E_{\rho,m}(\alpha),\, \alpha\Big)\Big\}, \tag{20}$$

其中 $H_m = \sum_{k=1}^m k^{-1} \sim$ 对数 m。

证明。第一项由 Benjamini 和 Yekutieli（2001年）给出上界；见（2）。对于第二项，我们可以假设 $m_0 \geq 1$，否则错误发现率 $= 0$。记 $z_j = z_{w_j}$, $K' = K_\rho/(1-\theta)^2$ 且$\bar{H}_k = m_0^{-1} \sum_{i \in \mathcal{H}_0} He_k(\xi_i)$。

步骤 *1*：拆分中心化计数。根据 Q_i的定义，

$$\widetilde{N}(w_j) - \pi_0 a_j = m_0 \sum_{k \in \mathcal{S},\ k < k^*} \gamma_k(z_j) \bar{H}_k + \sum_{i \in \mathcal{H}_0} Q_i(w_j).$$

由于 $|\gamma_k(z_j)|/w_j = |c_k(z_j)|/(1 - \Phi(z_j)) \leq b_k$，第一项至多为 $m_0 w_j \Lambda^* = \pi_0 a_j \Lambda^*$ 其中 $\Lambda^* = \sum_{k<k^*} b_k |H_k|$，它不依赖于 j。根据引理 4(ii)，

$$\mathbb{E}\bar{H}_k^2 = k!\, m_0^{-2} \sum_{i,i' \in \mathcal{H}_0} \Sigma_{ii'}^k \leq k!(1/m_0 + 2G_{\rho,k}(m)/m_0),$$

因为 $|\Sigma_{ii'}| \leq \rho(|i - i'|)$，且每个 i 在每个滞后下至多有两个配对项。因此 $\pi_0 \mathbb{E}\bar{H}_k^2 \leq k!(1/m + 2G_{\rho,k}(m)/m)$，而闵可夫斯基不等式给出 $\pi_0 \mathbb{E}(\Lambda^*)^2 \leq \Lambda_{\rho,m}(\alpha)$。

步骤 2 $\widetilde{S}$ $t > 2\pi_0$ $B_j = \{\widetilde{N}(w_j) \geq \frac{t}{2} a_j\}$ B_j $\widetilde{N}(w_j) - \pi_0 a_j \geq$: 的尾界。固定并令 ( )。在 上，$t - 2\pi_0$ $\frac{t-2\pi_0}{2} a_j$ $B \cup B_j^2$ ( )。因此 成立，其中

$$B = \Big\{\pi_0 \Lambda^* \geq \theta \frac{t - 2\pi_0}{2}\Big\}, \qquad B_j^2 = \Big\{\sum_{i \in \mathcal{H}_0} Q_i(w_j) \geq (1-\theta)\frac{t - 2\pi_0}{2} a_j\Big\}.$$

此外 $\widetilde{N}(w_j) > 0$ 在 B_j上成立，因此 $B_j^1 \triangleq \{\widetilde{N}(w_j) \geq 1\}$ 成立。根据引理 2,

$$\{\widetilde{S} \geq t\} \subseteq \bigcup_{j=0}^{J-1} \big(B_j^1 \cap (B \cup B_j^2)\big) \subseteq B \cup \bigcup_{j=0}^{J-1} \big(B_j^1 \cap B_j^2\big),$$

因此

$$\mathbb{P}(\widetilde{S} \geq t) \leq \mathbb{P}(B) + \sum_{j=0}^{J-1} \min\{\mathbb{P}(B_j^1), \mathbb{P}(B_j^2)\}.$$

根据马尔可夫不等式， $\mathbb{P}(B_j^1) \leq \mathbb{E}\widetilde{N}(w_j) = \pi_0 a_j$。根据切比雪夫不等式和步骤 **1**,

$$\mathbb{P}(B) \leq \frac{4\pi_0^2 \mathbb{E}(\Lambda^*)^2}{\theta^2 (t - 2\pi_0)^2} \leq \frac{4\pi_0 \Lambda_{\rho,m}(\alpha)}{\theta^2 (t - 2\pi_0)^2}.$$

根据切比雪夫不等式和命题 1 ，有 $I = \mathcal{H}_0$,

$$\mathbb{P}(B_j^2) \leq \frac{4(K_\rho m_0 w_j + e_{\rho,m}(z_j) m_0 m w_j^2)}{(1-\theta)^2 (t - 2\pi_0)^2 a_j^2} = \frac{b}{a_j} + c_j,$$

其中 $b = 4\pi_0 K'/(t-2\pi_0)^2$ 且 $c_j = 4\pi_0 e_{\rho,\,m}(z_j)/((1-\theta)^2(t-2\pi_0)^2)$。由于 $\min\{x, y+c\} \le \min\{x,y\}+c$，对于 $c \ge 0$,

$$\min\{\mathbb{P}(B_j^1), \mathbb{P}(B_j^2)\} \le \min\left\{\pi_0 a_j, \frac{b}{a_j}\right\} + c_j = \min\left\{\pi_0 a_j, \frac{\pi_0 b}{\pi_0 a_j}\right\} + c_j.$$

这些数值$\pi_0 a_j$ 从一个层级到下一个层级减半，因此引理 3 给出

$$\sum_{j=0}^{J-1} \min\left\{\pi_0 a_j, \frac{\pi_0 b}{\pi_0 a_j}\right\} \le 3\sqrt{\pi_0 b} = \frac{6\pi_0\sqrt{K'}}{t-2\pi_0}.$$

根据E_ρ的定义，$m(\alpha)$,

$$\mathbb{P}(B) + \sum_{j=0}^{J-1} c_j \le \frac{4\pi_0}{(t-2\pi_0)^2}\Big(\frac{\Lambda_{\rho,m}(\alpha)}{\theta^2} + \frac{1}{(1-\theta)^2}\sum_{j=0}^{J-1} e_{\rho,m}(z_j)\Big) = \frac{4\pi_0 E_{\rho,m}(\alpha)}{(t-2\pi_0)^2}.$$

因此，对于每个 $t > 2\pi_0$,

$$\mathbb{P}(\widetilde{S} \ge t) \le \frac{6\pi_0\sqrt{K'}}{t-2\pi_0} + \frac{4\pi_0 E_{\rho,m}(\alpha)}{(t-2\pi_0)^2}.$$

步骤 *3*：积分。由于 $2\pi_0 \le 1/\alpha$，我们在引理 1的 $2\pi_0$处拆分积分。我们对 $t \le 2\pi_0$使用$\mathbb{P}(S \ge t) \le 1$，并对 $t > 2\pi_0$使用此界和步骤 2。由 (5),

$$\begin{aligned}\frac{\mathrm{FDR}(\alpha; m, \mu, \Sigma)}{\alpha} &\le \int_0^{2\pi_0} 1\,dt + \int_{2\pi_0}^{1/\alpha} \mathbb{P}(\widetilde{S} \ge t)\,dt \\ &\le 2\pi_0 + \int_{2\pi_0}^{1/\alpha} \min\left\{1, \frac{6\pi_0\sqrt{K'}}{t-2\pi_0} + \frac{4\pi_0 E_{\rho,m}(\alpha)}{(t-2\pi_0)^2}\right\} dt = \Psi_{\pi_0}\big(K', E_{\rho,m}(\alpha), \alpha\big).\end{aligned}$$

证明完毕。 □

4.5 定理1的证明

相关性单调性。量 $G_{\rho,k}(m)$、$e_{\rho,m}(z)$、$\Lambda_{\rho,m}(\alpha)$ 和$E_{\rho,m}(\alpha)$仅通过 $|\rho(l)|$, $l \ge 1$依赖于ρ，且对每个 $|\rho(l)|$均为非减。同样的结论也适用于 K_ρ。事实上，由引理 4(iii)，$R(\cdot, z)$递增，因此 $\tau(r, z)$关于 $|r|$非减；增大 $|\rho(l)|$ 只会使 K_ρ中的求和包含更多滞后。这些量在包络函数 $\rho_{\epsilon,C}$处均有定义，尽管元素为 $\rho_{\epsilon,C}(|i-j|)$的矩阵未必是半正定的。

(i) 的证明。对于 $m = 1$，我们有 $H_1 = 1$ and $F_{\epsilon,C,1}(\alpha; \pi_0) = \pi_0$，因此 (8) 表示 FDR $\le \alpha\pi_0$，这就是界 (2)。令 $m \ge 2$, $\theta \in (0, 1)$, $\Sigma \in \mathcal{C}_m(\epsilon, C)$ 和 $\rho = \rho_\Sigma$。由于 $\rho \le \rho_{\epsilon,C}$ 逐点成立，上述单调性给出 $K_\rho \le K_{\epsilon,C}$, $\Lambda_{\rho,m}(\alpha) \le \Lambda_{\epsilon,C,m}(\alpha)$ 以及 $\sum_j e_{\rho,m}(z_j) \le E_{\epsilon,C,m}(\alpha)$。因此

$$E_{\rho,m}(\alpha) \le E_m(\theta) := \Lambda_{\epsilon,C,m}(\alpha)/\theta^2 + E_{\epsilon,C,m}(\alpha)/(1-\theta)^2.$$

由于 Ψ_{π_0} 关于其前两个参数单调不减（引理 7(i)），命题 2 可得，对任意$\mu \in \mathbb{R}^m$,

$$\mathrm{FDR}(\alpha; m, \mu, \Sigma) \le \alpha \min\left\{\pi_0 H_m,\ \Psi_{\pi_0}\Big(\frac{K_{\epsilon,C}}{(1-\theta)^2}, E_m(\theta), \alpha\Big)\right\}. \tag{21}$$

对 $\theta \in (0,1)$ 取下确界，得到 (8)。若 $k^* = \min \mathcal{S}$，则 $\Lambda_{\epsilon,C,m}(\alpha) = 0$。式 (6) 中 Ψ_{π_0} 的自变量随着 $\theta \downarrow 0$ 趋近于 $(K_{\epsilon,C}, E_{\epsilon,C,m}(\alpha))$。由于 Ψ_{π_0} 关于 (K, E) 单调不减且连续，下确界等于 $\Psi_{\pi_0}(K_{\epsilon,C}, E_{\epsilon,C,m}(\alpha)\ \alpha)$。

(9) 的证明。若 $\pi_0 = 0$，则 $\Psi_0 = 0$，因此 $F_{\epsilon,C,m}(\alpha;0) = 0$，(9) 成立。因此，我们只考虑$\pi_0 \in (0,1]$。若 $k^* = \min \mathcal{S}$，则 $\Lambda_{\epsilon,C,m}(\alpha) = 0$，从而得到 $F_{\epsilon,C,m}(\alpha;\, \pi_0) = \Psi_{\pi_0}(K_{\epsilon,C}, E_{\epsilon,C,m}(\alpha)\ \alpha)$。否则，在 (6) 中取 $\theta = 1/2$，得到

$$F_{\epsilon,C,m}(\alpha;\pi_0) \le \Psi_{\pi_0}\big(4K_{\epsilon,C}, 4\Lambda_{\epsilon,C,m}(\alpha) + 4E_{\epsilon,C,m}(\alpha), \alpha\big).$$

根据引理 7(i)，因此只需对 $m \ge 2$ 上的 $K_{\epsilon,C}$, $\Lambda_{\epsilon,C,m}(\alpha)$ 和 $E_{\epsilon,C,m}(\alpha)$ 作一致上界。我们在步骤 1–3 中完成这一工作，并在步骤 4 中得出结论。

约化为 $\epsilon = 1/k^*$。由于 $\log(l+e) \ge 1$，包络函数 $\rho_{\epsilon,C}$关于 ϵ 单调不增。由于 $\epsilon \ge 1/k^*$，我们有 $\rho_{\epsilon,C} \le \rho_{1/k^*,C}$。令 $\rho = \rho_{1/k^*,C}$和 $p = 1 + 1/k^*$，使得对 $l \ge 1$和 $k^*p = k^* + 1$，$|\rho(l)| \le C\log^{-p}(l+e)$。根据上述单调性，

$$K_{\epsilon,C} \le K_\rho, \qquad \Lambda_{\epsilon,C,m}(\alpha) \le \Lambda_{\rho,m}(\alpha), \qquad E_{\epsilon,C,m}(\alpha) \le \sum_{j=0}^{J-1} e_{\rho,m}(z_j).$$

我们对右侧的三个表达式作上界。回顾 $J = \lceil \log_2 m \rceil$，并令 $m \ge 2$。

步骤 1：关于 K_ϵ、C 的界。对于$z \le \sqrt{2}$，我们有 $r_z = 1/2$，而每个满足 $|\rho(l)| > 1/2$ 的滞后项都会对 K_ρ 中的和贡献 2。对于 $z > \sqrt{2}$，我们有 $r_z = z^{-2} < 1/2$。于是，每个满足 $|\rho(l)| > 1/2$ 的滞后项仍贡献 2，而每个满足 $z^{-2} < |\rho(l)| \le 1/2$ 的滞后项根据引理 5(ii) 至多贡献 $2(1+z)e^{-z^2/6}/\sqrt{6\pi}$。由于 $|\rho(l)| > c$ 蕴含 $l < \exp\{(C/c)^{1/p}\} - e$，第一类滞后项至多有 $(e^{(2C)^{1/p}} - e)_+$ 个，第二类滞后项至多有 $(\exp\{(Cz^2)^{1/p}\} - e)_+$ 个。因此

$$K_{\epsilon,C} \le \bar{K}_{C,k^*} := 1 + 2\big(e^{(2C)^{1/p}} - e\big)_+ + \frac{2}{\sqrt{6\pi}} \sup_{z>\sqrt{2}} (1+z)e^{-z^2/6}\Big(\exp\big\{(Cz^2)^{1/p}\big\} - e\Big)_+. \tag{22}$$

由于 $p > 1$，我们有 $(Cz^2)^{1/p} = o(z^2)$，如 $z \to \infty$ 所示，因此 $\bar{K}_{C,k^*} < \infty$。它只通过 $p = 1 + 1/k^*$ 依赖于 C 和 k^*。

两个初等事实。由于 $m \ge 2$，我们有 $1 \le J \le (2/\log 2)\log m$ $m^{-1}\sum_{l=1}^{m-1} (l+e)^{-q} \le G_{\rho,k}(m)/m \le C^k g_{kp}(m)$ $g_q(m) = \max_{l \le \sqrt{m}}$。对于 $q > 0$，令 $\bar{l} \ge \sqrt{m}$，$\log(l+e) \ge \frac{1}{2}\log m$。因此 $g_q(m) \le m^{-1/2} + 2^q \log^{-q} m$。利用 $\sup_{x>0} x^a e^{-x/c} = (ac/e)^a$，其中 $x = \log m$，可得

$$J^{k^*+1} g_{k^*+1}(m) \le M_{k^*} := \Big(\frac{2}{\log 2}\Big)^{k^*+1}\Big[\Big(\frac{2k^*+2}{e}\Big)^{k^*+1} + 2^{k^*+1}\Big], \tag{23}$$

$$J^k\Big(\frac{1}{m} + 2C^k g_{kp}(m)\Big) \le N_k := \Big(\frac{2}{\log 2}\Big)^k\Big[\Big(\frac{k}{e}\Big)^k + 2C^k\Big(\Big(\frac{2k}{e}\Big)^k + \frac{2^{kp}}{\log 2}\Big)\Big], \quad k \in \mathcal{S},\ k < k^*. \tag{24}$$

在 (24) 中，我们使用了 $\log^{k-kp} m = \log^{-k/k^*} m \le 1/\log 2$。此外，根据引理 5(i)，$z_w^2 \le 2\log(1/w)$ 在两种情况下均成立，因此

$$z_j^2 \le 2\log(1/\alpha) + 2j\log 2, \qquad 0 \le j \le J-1. \tag{25}$$

步骤 2：对 E_ϵ、C、$m(\alpha)$ 的界。根据引理 5(i)，$(\phi(z)/(1-\Phi(z)))^2 \le z^2 + 2$。根据引理 6(iii)，$A(z) \le a_1(z^2+2)^{k^*-1}$。在 $k^* = 2$ 的双侧情形下，引理 6(ii) 给出 $A(z) \le a_2(z^2+2)$。令$a_* = a_2$ 在此情形下取值，并在其他情形下令 $a_* = a_1$。于是 $e_{\rho,m}(z) \le 2a_* C^{k^*}(z^2+2)^{k^*} g_{k^*+1}(m)$。由 (25)，$z_j^2 + 2 \le 2(1 + \log(1/\alpha))(1+j)$。由于 $\sum_{j=0}^{J-1}(1+j)^{k^*} \le J^{k^*+1}$，由 (23) 得

$$E_{\epsilon,C,m}(\alpha) \le \bar{E}_{C,k^*}\big(1 + \log(1/\alpha)\big)^{k^*}, \qquad \bar{E}_{C,k^*} := 2^{k^*+1} a_* C^{k^*} M_{k^*}. \tag{26}$$

步骤 3：对 $\Lambda_{\epsilon,C,m}(\alpha)$ 的界。若 $k^* = \min\mathcal{S}$，则 $\Lambda_{\epsilon,C,m}(\alpha) = 0$。否则，令 $k \in \mathcal{S}$ with $k < k^*$，并写作 $Z = z_{J-1}$。根据引理 5(i),(iii)，$\phi(z)/(1-\Phi(z)) \le h(z)$ 且 $|He_{k-1}(z)| \le$

$(z+\sqrt{k})^{k-1}$，且二者关于 z 均为非递减。利用 $h(Z)^2 \le Z^2+2$ 和 $(Z+\sqrt{k})^2 \le 2(Z^2+k)$，可得 $b_k^2\,k! \le 2^{k-1}(Z^2+k+1)^k/k!$。由 (25) 和 $J \ge 1$，$Z^2+k+1 \le (k+1)(1+\log(1/\alpha))+2J \le 2(k+1)(1+\log(1/\alpha))J$。结合 (24)，

$$b_k^2\,k!\Big(\frac{1}{m}+\frac{2G_{\rho,k}(m)}{m}\Big) \le \frac{2^{k-1}(2k+2)^k}{k!}N_k\big(1+\log(1/\alpha)\big)^{k^*}.$$

令 $N=\#\{k\in\mathcal{S}: k<k^*\}$。根据柯西–施瓦茨不等式，

$$\Lambda_{\epsilon,C,m}(\alpha) \le \bar{\Lambda}_{C,k^*}\big(1+\log(1/\alpha)\big)^{k^*}, \qquad \bar{\Lambda}_{C,k^*} := N\sum_{k\in\mathcal{S},\ k<k^*}\frac{2^{k-1}(2k+2)^k}{k!}N_k. \tag{27}$$

如果 $k^*=\min\mathcal{S}$，则 $N=0$ 且 $\Lambda_{\epsilon,\,C,\,m}(\alpha)=0$，因此（12）在两种情况下均成立。

步骤 *4*：总结。根据步骤 1–3 和引理 7(i)，在两种情况下 $k^*=\min\mathcal{S}$ 且 $k^*>\min\mathcal{S}$，对每个 $m\ge 2$，

$$F_{\epsilon,C,m}(\alpha;\pi_0) \le \Psi_{\pi_0}\Big(4\bar{K}_{C,k^*},\ 4\big(\bar{\Lambda}_{C,k^*}+\bar{E}_{C,k^*}\big)\big(1+\log(1/\alpha)\big)^{k^*},\ \alpha\Big),$$

其中$K_{C、k^*}$、$\Lambda_{C、k^*}$ 和 $E_{C、k^*}$ 分别由（10）、（12）和（11）给出。将引理 7(ii) 的第二个不等式应用于 $K=4\bar{K}_{C、k^*}$ 和 $E=4(\bar{\Lambda}_{C、k^*}+\bar{E}_{C、k^*})(1+\log(1/\alpha))^{k^*}$。于是 $6\pi_0\sqrt{K}=12\pi_0\sqrt{\bar{K}_{C、k^*}}$ 和 $4\sqrt{\pi_0 E}=8\sqrt{\pi_0(\bar{\Lambda}_{C、k^*}+\bar{E}_{C、k^*})}\,(1+\log(1/\alpha))^{k^*/2}$。由于 $\log_+(xy) \le \log_+ x+\log_+ y$，我们有 $\log_+(1/(12\alpha\pi_0\sqrt{\bar{K}_{C、k^*}})) \le \log(1/\alpha)+\log_+(1/(12\pi_0\sqrt{\bar{K}_{C、k^*}}))$。因此

$$\sup_{m\ge 2}F_{\epsilon,C,m}(\alpha;\pi_0) \le \pi_0\big(2+12\sqrt{\bar{K}_{C,k^*}}\big)+12\pi_0\sqrt{\bar{K}_{C,k^*}}\Big(\log(1/\alpha)+\log_+\frac{1}{12\pi_0\sqrt{\bar{K}_{C,k^*}}}\Big)$$
$$+8\sqrt{\pi_0\big(\bar{\Lambda}_{C,k^*}+\bar{E}_{C,k^*}\big)}\,\big(1+\log(1/\alpha)\big)^{k^*/2}.$$

最后，令 $M=$最大值$\{1,k^*/2\}$ 且 $x=\log(1/\alpha)\ge 0$。由于 $1+x\ge 1$ 且 $k^*/2\le M$，由凸性可得 $(1+x)^{k^*/2} \le (1+x)^M \le 2^{M-1}(1+x^M)$。结合 $x\le 1+x^M$，可得（9）。定理1的证明至此完成。

□

参考文献

Yoav Benjamini 和 Yosef Hochberg。《控制错误发现率：一种实用且有效的多重检验方法》。皇家统计学会杂志：*B*辑，57(1):289–300，1995年。doi: 10.1111/j.2517-6161.1995.tb02031.x。

Yoav Benjamini 和 Daniel Yekutieli。《依赖情形下多重检验中错误发现率的控制》。《统计年刊》，29(4):1165–1188，2001年。doi: 10.1214/aos/1013699998。

Simeon M. Berman。平稳序列中最大项的极限定理。《数学统计年刊》，35(2)：502–516，1964年。

Z. W. Birnbaum。关于米尔斯比率的一个不等式。《数学统计年刊》，13(2)：245–246，1942。

Ziyu Chi、Aaditya Ramdas 和 Ruodu Wang。负依赖下的多重检验。《伯努利》，31(2):1230–1255，2025。doi: 10.3150/24-BEJ1768。

Sandy Clarke 和 Peter Hall。多重检验程序对依赖的稳健性。《统计年刊》，37(1):332–358，2009。doi: 10.1214/07-AOS557。Swarnadeep Datta 和 Monitirtha Dey。弱依赖下多重检验的一些渐近结果，2026。arXiv:2602.21359。Sylvain Delattre 和 Etienne Roquain。高维高斯向量分量的经验分布函数及其在多重检验中的应用。《伯努利》，22(1):302–324，2016。Edgar Dobriban。Benjamini–Hochberg 程序可能无法控制相关的双侧高斯检验的错误发现率。美国国家科学院院刊，123(41):e2626267123，2026。doi: 10.1073/pnas.2626267123。Bradley Efron。相关系数与大规模同时显著性检验。*Journal of the Amer-ican Statistical Association*，102(477):93–103，2007。Alessio Farcomeni。依赖条件下错误发现率控制的一些结果。斯堪的纳维亚统计学杂志，34(2):275–297，2007。doi: 10.1111/j.1467-9469.2006.00530.x。William Fithian 和 Lihua Lei。依赖条件下错误发现率控制的条件校准。《统计年刊》，50(6):3091–3118，2022。doi: 10.1214/21-AOS2137。Deepra Ghosh 和 Sanat K. Sarkar。双侧高斯均值检验中考虑依赖的错误发现率控制，2025。arXiv:2511.19960。Deepra Ghosh 和 Sanat K. Sarkar。双侧高斯均值检验中控制错误发现率的进一步结果，2026。arXiv:2608.21267。Samuel Karlin 和 Yosef Rinott。绝对值多元正态变量的全正性及其在置信区间估计和相关概率不等式中的应用。《统计年刊》，9(5):1035–1049，1981。doi: 10.1214/aos/1176345583。Lihua Lei。高斯依赖最多会使 Benjamini–Hochberg 程序的错误发现率膨胀多少？，2026。arXiv:2607.14812。Chenxin Lyu 和 Ruodu Wang。高斯依赖下单侧 p-值的错误发现率。*arXiv* 预印本 *arXiv:2611.19589*，2026。Yash Mittal 和 Donald Ylvisaker。平稳高斯过程极大值的极限分布。随机过程及其应用，3(1):1–18，1975。Drew T. Nguyen 和 William Fithian。非参数图形依赖模型下的错误发现率控制，2025。arXiv:2506.24126。Art B. Owen。错误发现数量的方差。皇家统计学会杂志：*B*辑，67(3):411–426，2005。R. L. Plackett。多元正态积分的约化公式。*Biometrika*，41(3/4):351–360，1954。Sanat K. Sarkar。逐步多重检验程序中错误发现率的一些结果。《统计年刊》，30(1):239–257，2002。

Sanat K. Sarkar. 关于对相关正态变量的均值进行双侧备择多重检验时如何控制错误发现率，2023。arXiv:2304.05261。Armin Schwartzman 和 Xihong Lin。相关系数对错误发现率估计的影响。*Biometrika*，98(1):199–214，2011。Weijie J. Su。错误发现率联系定理。*arXiv* 预印本 *arXiv:1812.08965*，2018。Wei Biao Wu。依赖情形下的错误发现控制。《统计年刊》，36(1):364–380，2008。Xianyang Zhang。双侧高斯检验的精确 Simes 不等式，2026。arXiv:2611.02513。