同济大学
TONGJI UNIVERSITY
地址：中国上海市四平路1239号 邮编：200092
1239 SIPING ROAD SHANGHAI CHINA 200092
电话（TEL）：+86 21- 传真（FAX）：+86 21-
网址（WEB）：www.tongji.edu.cn

B-H procedure:

1. 为FDR control的首先提出者.

2. 前提：检验统计量相互独立，B-Y(2001)证明了在特殊的相关关系(PDR)下B-H方法也有效控制FDR.

3. 理论结果：

没要求独立!!

引理一：对独立的 m_0 个 p 值(对应于 H_0)和 m_1 个 p 值(对应于 H_1)，B-H方法使下式成立.

$$E(Q \mid P_{m_0+1}=p_1, \cdots, P_m=p_{m_1}) \le \frac{m_0}{m} q^*$$

其中 q^* 事先给定. $Q=\frac{V}{V+S}$.

Pf: . 采用第二数学归纳法. $m=1$ 显然成立，设 $\forall m' \le m$ 均有结论成立，现当 $m=m+1$ 时. ① 当 $m_0=0$ 时，$Q=0$，$\therefore E(Q \mid P_1=p_1, \cdots, P_m=p_m)=0 \le \frac{m_0}{m+1} q^*$

② 当 $m_0>0$ 时，用 P_i', $i=1,\cdots,m_0$ 表示原假设 H_0 下的 p 值. 定义最大次序统计量 $P'_{(m_0)}$. 由 p 值定义及独立性，有 $P_i' \overset{iid}{\sim} U(0,1)$，不妨设 H_1 下 m_1 个 p 值为 $p_1 \le p_2 \le \cdots \le p_{m_1}$. 设 $y_0=\max\{y: p_y \le \frac{m_0+y}{m+1} q^*, 0 \le y \le m_1\}$. 设 $p_{y_0} \triangleq p''$

在 $P'_{(m_0)}=p$ 的条件下.

$$E(Q \mid P_{m_0+1}=p_1, \cdots, P_m=p_{m_1}) = \int_0^{p''} E(Q \mid P'_{(m_0)}=p, P_{m_0+1}=p_1, \cdots, P_m=p_{m_1}) f_{P'_{(m_0)}}(p)\,dp$$
$$+ \int_{p''}^1 E(Q \mid P'_{(m_0)}=p, P_{m_0+1}=p_1, \cdots, P_m=p_{m_1}) f_{P'_{(m_0)}}(p)\,dp$$
$$\triangleq I_1 + I_2$$

易求得 $f_{P'_{(m_0)}}(p)=m_0 p^{m_0-1} I_{(0,1)}(p)$

对于 I_1，$\because p \le p''$，故会拒绝 m_0+y_0 个假设，$Q=\frac{m_0}{m_0+y_0}$，再利用 y_0 的定义.

$$I_1 = \frac{m_0}{m_0+y_0}(p'')^{m_0} \le \frac{m_0}{m_0+y_0} \frac{m_0+y_0}{m+1} q^* (p'')^{m_0-1} = \frac{m_0}{m+1}(p'')^{m_0-1} q^*$$

对于 I_2，对每个满足 $p_{y_0} < p_y \le P'_{(m_0)} = p < p_{y+1}$，先考虑 $p_{y_0} \le p'' < P'_{(m_0)} = p < p_{y_0+1}$，由 p'' 定义，$p, p_{y+1}, \cdots, p_{m_1}$ 一定不会被拒绝，

做分割. $p_1, p_2, \cdots\cdots, p_{m_1}$. $\Sigma \int_{p_y}^{p_{y+1}} \bigcirc\, dp$.

故可能被拒绝的个数为 m_0-1+y ($P'_{(1)}, \cdots, P'_{(m_0-1)}$ 和 $p_1, \cdots, p_y$)，考虑 $H_{(1)}, \cdots, H_{(m+1)}$. $H_{(i)}$ 被拒绝当且仅当 $\exists k$. s.t. $1 \le k \le m_0-1+y$，$p_{(k)} \le \frac{k}{m+1} q^*$，i.e.，$\frac{p_{(k)}}{p} \le \frac{k}{m_0-1+y} \cdot \frac{m_0-1+y}{(m+1)p} q^*$. 在 $P'_{(m_0)}=p$ 的条件下，$\frac{P_i'}{p} \overset{iid}{\sim} U(0,1)$，$i=1,2,\cdots,m_0-1$. $\frac{p_i}{p} \in (0,1)$，$i=1,\cdots,y$. 因为 $m'=m_0-1+y \le m$，考虑 m' 个假设，$q^{*\prime}=\frac{m_0-1+y}{(m+1)p} q^*$ 的B-H方法. 由归纳假设可得

$$E(Q \mid P'_{(m_0)}=p, P_{m_0+1}=p_1, \cdots, P_m=p_{m_1}) \le \frac{m_0-1}{m_0+y-1} \frac{m_0-1+y}{(m+1)p} q^* = \frac{m_0-1}{(m+1)p} q^*$$

对每一部分应用该结果，可得

$$\int_{p''}^1 E(Q \mid P'_{(m_0)}=p, P_{m_0+1}=p_1, \cdots, P_m=p_{m_1}) f_{P'_{(m_0)}}(p)\,dp \le \int_{p''}^1 \frac{m_0-1}{(m+1)p} q^* m_0 p^{(m_0-1)}\,dp = \frac{m_0}{m+1} q^* (1-p''^{(m_0-1)})$$

综合 I_1, I_2，得证.

无论 $P_1'', \cdots, P_{m_0}''$ 的联合分布如何，只需在在引理一两端对 $P_1, \cdots, P_{m_0}$ 积分即得下述定理

m_0未知，即管不需要 H_i 独立，但全独立保证一定对.

定理一：对独立的检验统计量，B-H方法控制 FDR 在水平 q^* 上.

4. 步骤：Step 1：计算各检验下的 p 值

Step 2：对 p 值进行排序.

Step 3：计算 $j_0 = \max\{j: 1 \le j \le m,\ p_{(j)} \le \frac{j}{m} q^*\}$

Step 4：拒绝 $H_{(1)}, \cdots, H_{(j_0)}$.