



### Oracle and Plug-in p-value 方法.

- 改进目标：设  $G(t)$  为 nonnull case 的 cdf. null case 为均匀分布，即  $G_0(t) = t$ . 考虑模型  $G(t) = (1 - \varepsilon_n)t + \varepsilon_n G(t)$ . 本意为研究 B-H 方法的原理，后提出一种基于  $p$  值，min FNR, s.t.  $mFDR \leq \alpha$  的最优程序， $mFDR$  在 cutoff 为  $t$  时， $Q(t) = \frac{(1 - \varepsilon_n)t}{G(t)}$ . 但  $\varepsilon_n$  和  $G(t)$  未知.  $\varepsilon_n$  有很多估计方法.  $\hat{\varepsilon}_n = 1 - \frac{W(\lambda)}{n(1-\lambda)}$ ，或 Jin, Cai (2007) 等.  $G(t)$  则用经验 cdf  $\hat{G}(t) = \frac{1}{m} \sum I(p_i \leq t)$ .
- 假设前提： $G(t)$  凹函数 (确保  $Q(t) = \alpha$  的唯一性)，
- 方法步骤. Step 1: 计算  $p$  值  
Step 2: 估计  $\hat{\varepsilon}_n$  和  $\hat{G}(t)$   
Step 3:  $t(p, \hat{G}) = \sup \{t : Q(t) \leq \alpha\}$ . 拒绝  $H_0$ . 若  $p_i < t(p, \hat{G})$   
在所有的  $p$  值方法中是最佳的. 应该可以使用 bootstrap 方法提高精度

### q 值方法.

- 改进目标：在单个假设检验问题中的  $p$  值作为检验显著性的度量，由此启发，在多重检验 FDR 方法中，提出  $q$  值的概念. 粗略的说，一个检验的  $q$  值为当该检验拒绝时错误发现的比例度量.  $q$  值为  $H_0$  被拒绝时的最小 FPR. 即  

$$q(p_i) = \inf_{t \geq p_i} \{PFDR(t)\} = \inf_{t \geq p_i} \left\{ \frac{(1 - \hat{\varepsilon}_n)t}{\hat{G}(t)} \right\}$$
 $q$  值虽然是对每一个检验独立计算的，但其考虑了与其它  $p$  值相比的相对重要性.
- 方法步骤. Step 1: 计算  $p$  值，并排序  
Step 2: 计算  $k = \max \{i : q(p_{(i)}) = \inf_{t \geq p_{(i)}} \frac{(1 - \hat{\varepsilon}_n)t}{\hat{G}(t)} \leq \alpha\}$ .  
Step 3: 拒绝  $H_{(1)}, \dots, H_{(k)}$ .