Outline

**Agenda:**  Multiple Testing Problems

3. Improving on BHq?

The material for this lecture is taken from Storey, Siegmund and Taylor (2004) [1].

1 The Empirical Process Viewpoint of BH(q)

In the previous lecture, we introduced the BHq step-up procedure by relating it to Simes’. We plot the sorted p-values and the critical line as shown in Figure 1(a) and look for the first time (going from large to small p-values) a p-value falls below the critical line. An equivalent representation can be obtained by swapping the $x$ and the $y$-axes, as in Figure 1(b). The latter best allows to describe the BH procedure in terms of an empirical process. In particular, the coordinated on the $y$-axis of Figure 1(b) are the values of the empirical CDF $\hat{F}_n(t)$ of the p-values, which is defined as

$$\hat{F}_n(t) = \frac{\# \{ i : p_i \leq t \}}{n}.$$  

Assuming that the p-values and the hypotheses are ordered according to

$$p(1) \leq \cdots \leq p(n), \quad H(1) \leq \cdots \leq H(n),$$

we defined the BH(q) as rejecting $H(1), \ldots, H(i_0)$ where

$$i_0 = \max \left\{ i : p(i) \leq q \frac{i}{n} \right\}.$$  

Using the definition of the empirical CDF, the critical $p$-value $p^* = p(i_0)$ can be rewritten as

$$p^* = \max \left\{ p(i) : p(i) \leq q \frac{i}{n} \right\} = \max \left\{ p(i) : p(i) \leq q \hat{F}_n(p(i)) \right\},$$

1
with the convention that $p^* = q/n$ if the above set is empty. Hence,

$$p^* = \max \left\{ t \in \{p_1, \ldots, p_n\} : t \leq q \hat{F}_n(t) \right\},$$

with the same convention as above if the set is empty. Now set

$$\tau_{BH} = \max \left\{ t : \frac{t}{\hat{F}_n(t) \lor 1/n} \leq q \right\} \quad (1)$$

and note that $\tau_{BH} \geq q/n$. By construction, the BH procedure rejects all hypotheses with $p_i \leq \tau_{BH}$. That is, when $k$ rejections are made, all $p$-values less or equal to $qk/n$ are rejected.

The BH procedure can be justified with the following simple argument due to [1]. Take a fixed value of $t \in (0, 1)$ and consider the rule that rejects $H_{0,i}$ if and only if $p(i) \leq t$. The entries in the table of outcomes (shown below) will depend on the value of $t$.

<table>
<thead>
<tr>
<th></th>
<th>$H_0$ accepted</th>
<th>$H_0$ rejected</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H_0$ true</td>
<td>$U(t)$</td>
<td>$V(t)$</td>
<td>$n_0$</td>
</tr>
<tr>
<td>$H_0$ false</td>
<td>$T(t)$</td>
<td>$S(t)$</td>
<td>$n - n_0$</td>
</tr>
<tr>
<td></td>
<td>$n - R(t)$</td>
<td>$R(t)$</td>
<td>$n$</td>
</tr>
</tbody>
</table>

The false discovery proportion and false discovery rate, defined as before, will also depend on $t$:

$$Fdp(t) = \frac{V(t)}{\max(R(t), 1)}.$$  

The threshold $t$ should then be chosen as large as possible, while controlling the FDR at level $q$. Observe that by definition, we have

$$FDR(t) = \mathbb{E}[Fdp(t)] = \mathbb{E} \left[ \frac{V(t)}{\max(R(t), 1)} \right].$$

Estimates of FDR process can be inverted to give a FDR-controlling thresholding procedure. If we have an estimate $\hat{FDR}(t)$ of $FDR(t)$, then we can take the threshold

$$\tau = \sup \{ t \leq 1 : \hat{FDR}(t) \leq q \}.$$
This defines the largest (most liberal) thresholding cut-off, s.t. our estimated FDR is controlled. Note that, since we choose \( \tau \) by looking at all the \( p \)-values, this is a data-dependent thresholding procedure.

How can we estimate

\[
\text{FDR}(t) = \frac{V(t)}{\max(R(t), 1)}
\]

Recall that:

- The number of rejections \( R(t) \) is known.
- The number of false rejections \( V(t) \) is not known.

A solution starts by noting that \( \mathbb{E}V(t) = n_0 t \). Still, \( n_0 \) is not known. However, one can make a “conservative estimate” equal to \( nt \). This leads to

\[
\hat{\text{FDR}}(t) = \frac{nt}{\max(R(t), 1)} = \frac{t}{F_n(t) \lor 1/n}.
\]

This choice leads us to BH \((q)\) procedure since

\[
\tau_{BH} = \sup \left\{ t \leq 1 : \frac{nt}{R(t) \lor 1} \leq q \right\}.
\]

[1] proved that the FDR estimate is biased upward.

**Theorem 1.** Under independence, \( \mathbb{E}[\hat{\text{FDR}}(t)] \geq \text{FDR}(t) \).

A good thing/exercise for you would be to verify this inequality.

## 2 Martingale Proof of FDR (BH \((q)\))

The estimate of FDR can also be inverted to yield FDR control, giving another proof of the result of Benjamini and Hochberg [1].

**Theorem 2** (BH (1995)).

\[
\mathbb{E} \left[ \text{FDR}(\tau_{BH}) \right] = qn_0/n;
\]

i.e. the procedure rejecting all hypotheses with \( p_i \leq \tau_{BH} \) controls the FDR.

**Proof.** We use Martingale Theory.

1. **Filtration:** \( \mathcal{F}_t = \sigma\{V(s), R(s), t \leq s \leq 1\} \).

2. \( \tau_{BH} \) is a stopping time w.r.t. backward filtration \( \{\mathcal{F}_t\} \) since \( \{\tau_{BH} \leq t\} \in \mathcal{F}_t \). Indeed, knowledge of \( R(s) = n\tilde{F}_n(s) \) for \( s \geq t \) determines whether \( \tau_{BH} \leq t \) or not.
3. \( \left\{ \frac{V(t)}{t} \right\} \) is a martingale running backward in time:

\[
\mathbb{E} \left[ \frac{V(t)}{t} \mid \mathcal{F}_s \right] = \frac{1}{t} \mathbb{E} [V(t) \mid \mathcal{F}_s] = \frac{1}{t} V(s) \frac{t}{s} = \frac{V(s)}{s}
\]

where we used that under \( \mathcal{F}_s \), \( V(s) = \# \{ p_i^0 : p_i^0 \leq s \} \) and these \( p_i^0 \) are uniformly distributed on \([0, s]\) and are independent.

4. Optional Stopping Time Theorem:

\[
\max(R(\tau_{BH}), 1) = n \max(\hat{F}_n(\tau_{BH}), 1/n) = n \tau_{BH}/q.
\]

Therefore,

\[
\text{FDR}(\tau_{BH}) = \mathbb{E} \left( \frac{V(\tau_{BH})}{\max(R(\tau_{BH}), 1)} \right) = \frac{q}{n} \mathbb{E} \left( \frac{V(\tau_{BH})}{\tau_{BH}} \right) = \frac{q}{n} \mathbb{E} V(1) = \frac{q n_0}{n}.
\]

\[
\square
\]

3 Improving on BHq?

Consider the interpretation of BHq introduced in section (1). Can the distributions of p-values be used to improve the simple conservative estimate of \( \pi_0 = \frac{n_0}{n} \)? Fix \( \lambda \in [0, 1) \) and define

\[
\hat{\pi}_0^\lambda = \frac{n - R(\lambda)}{(1 - \lambda)n}.
\]

Based on this estimate for \( \pi_0 = \frac{n_0}{n} \), one obtains the following estimate for FDR:

\[
\hat{\text{FDR}}^\lambda(t) = \frac{\hat{\pi}_0^\lambda n t}{\max(R(t), 1)}.
\] (3)

For \( \lambda = 0 \), BHq is recovered. For a general \( \lambda \),

\[
\hat{\pi}_0^\lambda = \frac{n_0 - V(\lambda) + n_1 - S(\lambda)}{(1 - \lambda)n} \geq \frac{n_0 - V(\lambda)}{(1 - \lambda)n},
\]

and hence

\[
\mathbb{E} \left[ \hat{\pi}_0^\lambda \right] \geq \frac{n_0}{n} = \pi_0.
\]

The idea is that if non null p-values are small, then \( n_1 - S(\lambda) \approx 0 \) and \( \hat{\pi}_0^\lambda \) gives an accurate estimate of \( \pi_0 \). The goal would be to show that the threshold

\[
\tau = \sup \left\{ t \leq 1 : \hat{\text{FDR}}(t) = \frac{\hat{\pi}_0^\lambda n t}{\max(R(t), 1)} \leq q \right\}
\]

provides FDR control. While this may actually not be the case, [1] proves such a thing for a modified version of (3).
References