STAT 689: Large-scale and high-dimensional statistical inference **Fall 2024**

Lecture 6

1 False discovery rate

Recall that there are four types of outcomes in multiple testing, as illustrated in Table 1. The rows indicate the true state of the world with respect to the null hypotheses. On the other hand, the columns indicate acceptance and rejection of the null hypotheses. The random variables U and V indicate the number of correctly accepted and falsely rejected hypotheses, i.e., the number of true negatives and false positives (often called false discoveries). The random variables T and S indicate the number of false negatives and true positives. The number of hypotheses under the null is denoted by n_0. The four random variables U, V, T, and S are unobserved. The random variable R indicates the total number of rejections by a given multiple testing procedure and is observed. Note that the quantities of primary interest are the number of false discoveries V and the number of discoveries R. It is desired to maximize the number of discoveries subject to the constraint that the number of false discoveries remains low.

	H_0 accepted	H_0 rejected	Total
H_0 true	U	V	n_0
H_0 false	T	S	$n - n_0$
	$n - R$	R	n

Table 1: Potential outcomes for testing multiple hypotheses.

1.1 Definition

In many applications (such as GWAS), controlling the FWER will result in very few discoveries. This created a need within many scientific communities to abandon FWER for other ways to highlight and rank in publications those variables showing marked effects across individuals or treatments that would otherwise be dismissed as non-significant after standard FWER correction for multiple tests.

The false discovery rate (FDR) is an alternative error measure proposed by Benjamini and Hochberg (1995) that leads to more discoveries. It is particularly useful when researchers are looking for "discoveries" that will give them follow-up work (E.g., detecting promising genes for follow-up studies) and are interested in controlling the proportion of "false leads" they are willing to accept.

To introduce FDR, we first define the so-called false discovery proportion (FDP):

$$\mathrm{FDP} = \frac{V}{1 \vee R},$$

where we define $a \vee b = \max\{a, b\}$. Note that when $R = 0$, FDP is equal to zero. The FDR is defined as

$$\mathrm{FDR} = \mathbb{E}[\mathrm{FDP}].$$

One important observation and a common objection to FDR is that controlling the FDR gives us security on average across many repetitions of an experiment, but unlike FWER, FDR does not guarantee anything about a particular study. So, we cannot make a statement about the FDP for the study. Hence, it is useful to think about FDR in the sense of the experiments done by the scientific community as a whole as opposed to each single experiment.

1.2 Comparison with FWER

FDR is also a weaker notion of control than FWER, which makes it a useful compromise in modern settings when the number of hypotheses is large enough that FWER control can be too stringent. We argue that

$$\mathbf{1}\{V \geq 1\} \geq \text{FDP}, \tag{1}$$

which implies that FDR $\leq$ FWER. To show (1), we note that when $V = 0$, FDP $= 0$. On the other hand, if $V \geq 1$, FDP is less or equal to one. Therefore, controlling the FDR is more liberal, leading to more discoveries compared to FWER control.

1.3 Benjamini-Hochberg procedure

The Benjamini-Hochberg procedure is perhaps the most commonly used multiple-testing procedure for controlling the FDR. Let $p_{(1)} \leq p_{(2)} \leq \cdots \leq p_{(n)}$. The BH procedure finds the largest i (denoted by i_0) such that

$$p_{(i)} \leq \frac{i\alpha}{n}$$

and rejects $H_{(1)}, \ldots, H_{(i_0)}$.

Recall that in the Hochberg's procedure, we find the largest i (denoted by j_0) such that

$$p_{(i)} \leq \frac{\alpha}{n - i + 1}$$

and reject $H_{(1)}, \ldots, H_{(j_0)}$. We now compare the two thresholds $i\alpha/n$ and $\alpha/(n - i + 1)$ for $p_{(i)}$. Suppose $i = \beta n$. We have

$$\frac{i\alpha}{n} = \beta\alpha$$

and

$$\frac{\alpha}{n - i + 1} = \frac{\alpha}{n(1 - \beta) + 1}.$$

The threshold for Hochberg's procedure can be much more conservative than that of the BH procedure.

Adjusted p-value. Similar to the Bonferroni correction on p-values (where we set $\tilde{p}_i = np_i$ and compare $\tilde{p}_i$ with α to decide if H_i should be rejected), we can define the BH-adjusted p-value for each hypothesis. Note that $H_{(i)}$ will be rejected in the BH procedure, if there exists some $j \geq i$ such that

$$\frac{np_{(j)}}{j} \leq \alpha.$$

This motivates us to define the adjusted p-value

$$\tilde{p}_{(i)} = \min\left(1, \min_{j \geq i} \frac{np_{(j)}}{j}\right).$$

It is not hard to see that $H_{(i)}$ is rejected in the BH procedure if and only if $\tilde{p}_{(i)} \leq \alpha$.

1.4 An alternative formulation for the BH procedure

The BH procedure is equivalent to rejecting all H_i with $p_i \leq T$, where T is defined as

$$T = \sup\left\{0 < t \leq 1\colon \frac{nt}{1 \vee R(t)} \leq \alpha\right\},$$

with $R(t) = \sum_{i=1}^n \mathbf{1}\{p_i \le t\}$ being the number of rejections given the threshold t. Let $\mathcal{N}_0$ be the set of true nulls. Note that by the law of large numbers

$$\frac{1}{|\mathcal{N}_0|} \sum_{i \in \mathcal{N}_0} \mathbf{1}\{p_i \le t\} \approx t$$

which implies that

$$\sum_{i \in \mathcal{N}_0} \mathbf{1}\{p_i \le t\} \approx t|\mathcal{N}_0| \le tn.$$

Thus, nt can be viewed as a conservative estimate for the number of false discoveries given the threshold t.

Exercise 6.1: Show that the two descriptions of the BH procedure in Sections 1.3 and 1.4 are equivalent.

1.5 FDR control theory

Theorem. Suppose the null p-values are mutually independent and are independent of the alternative p-values. The BH procedure controls the FDR at level $n_0\alpha/n$, where n_0 is the number of true nulls.

Proof. The proof is based on the so-called leave-one-out argument. Let $V_i = \mathbf{1}\{H_i \text{ is rejected}\}$. We have

$$\text{FDP}(T) = \sum_{i \in \mathcal{N}_0} \frac{V_i}{R(T) \vee 1} = \sum_{i \in \mathcal{N}_0} \frac{V_i}{nT} \frac{nT}{R(T) \vee 1} \le \alpha \sum_{i \in \mathcal{N}_0} \frac{V_i}{nT}.$$

Therefore, we need to bound

$$\mathbb{E}\left[\sum_{i \in \mathcal{N}_0} \frac{V_i}{nT}\right]$$

from above. Observing that, for a given R, $T = T(R)$ is a deterministic function of R, we have:

$$\mathbb{E}\left[\sum_{i \in \mathcal{N}_0} \frac{V_i}{nT}\right] = \sum_{i \in \mathcal{N}_0} \sum_{k=1}^n \mathbb{E}\left[\frac{V_i \mathbf{1}\{R = k\}}{nT(k)}\right] = \sum_{i \in \mathcal{N}_0} \sum_{k=1}^n \mathbb{E}\left[\frac{V_i \mathbf{1}\{R(p_i \to 0) = k\}}{nT(k)}\right],$$

where $R(p_i \to 0)$ is the number of rejections obtained by replacing the p-value p_i with 0. To clarify the second equality, note that if $V_i = 0$, the equation is trivially true. When $V_i = 1$, setting p_i to 0 does not change the number of rejections (Why?). By direct calculation,

$$\begin{aligned}
\mathbb{E}\left[\sum_{i \in \mathcal{N}_0} \frac{V_i}{nT}\right] &= \sum_{i \in \mathcal{N}_0} \sum_{k=1}^n \frac{1}{nT(k)} \mathbb{E}[\mathbf{1}\{p_i \le T(k)\}] \mathbb{E}[\mathbf{1}\{R(p_i \to 0) = k\}] \\
&\le \sum_{i \in \mathcal{N}_0} \sum_{k=1}^n \frac{T(k)}{nT(k)} \mathbb{E}[\mathbf{1}\{R(p_i \to 0) = k\}] \\
&\le \frac{1}{n} \sum_{i \in \mathcal{N}_0} \sum_{k=1}^n \mathbb{E}[\mathbf{1}\{R(p_i \to 0) = k\}] \\
&\le \frac{n_0}{n},
\end{aligned}$$

where $n_0 = |\mathcal{N}_0|$. Hence, $\text{FDR} = \mathbb{E}[\text{FDP}] \le n_0\alpha/n$.

1.6 BH procedure based on z-scores

Let $\mathbf{z} = (z_1, \ldots, z_n)$ be a set of z-statistics with $\mathbb{E}[z_i] = \mu_i$. We are interested in testing the two-sided hypothesis:

$$H_{0,i} : \mu_i = 0 \text{ versus } H_{a,i} : \mu_i \ne 0, \quad i = 1, 2, \ldots, n.$$

Define the FDP estimate as

$$\mathrm{FDP}(t) = \frac{2n(1-\Phi(t))}{1 \vee \sum_{i=1}^n \mathbf{1}\{|z_i| \geq t\}}.$$

The BH procedure works as follows: (1) Find the smallest $t^* \geq 0$ such that $\mathrm{FDP}(t^*) \leq \alpha$; (2) Reject all the hypotheses such that $|z_i| \geq t^*$. The above procedure is exactly equivalent to the BH procedure based on the two-sided p-values $p_i = 2(1-\Phi(|z_i|))$ (Why?).

1.7 Generalized BH procedure

Exercise 6.2: Consider the following generalized version of the BH procedure. Let $\psi_i : \mathbb{R} \to \mathbb{R}$ be a strictly increasing function and can differ for each i. Define the threshold

$$T = \sup\left\{0 < t \leq 1 \colon \frac{ng(t)}{1 \vee R(t)} \leq \alpha\right\},$$

where g is a strictly increasing function and $R(t) = \sum_{i=1}^n \mathbf{1}\{\psi_i(p_i) \leq t\}$. We reject H_i whenever $\psi_i(p_i) \leq T$. Using the leave-one-out technique, show that the above procedure controls the FDR at the level $C\alpha$, where

$$C = \sum_{i \in \mathcal{N}_0} \sup_{t \in \mathcal{C}_\alpha} \frac{\psi_i^{-1}(t)}{ng(t)}, \quad \mathcal{C}_\alpha = \{0 < t \leq 1 : g(t) \leq \alpha\}.$$

As a special case of Exercise 6.2, we consider the so-called weighted BH procedure, which was originally proposed by Genovese, Roeder, and Wasserman (2006, Biometrika). Let w_i be a sequence of positive weights such that $\sum_{i=1}^n w_i = n$. Define $\psi_i(p) = p/w_i$. We reject H_i if

$$\psi_i(p_i) = p_i/w_i \leq T$$

for the data-dependent threshold T. Note that $\psi_i^{-1}(p) = w_i p$. Let us set $g(t) = \sum_{i=1}^n \psi_i^{-1}(t)/n = t$. Therefore,

$$T = \sup\left\{0 < t \leq 1 \colon \frac{nt}{1 \vee R(t)} \leq \alpha\right\},$$

where $R(t) = \sum_{i=1}^n \mathbf{1}\{p_i/w_i \leq t\}$. In this case, notice that

$$C = \sum_{i \in \mathcal{N}_0} \sup_{t \in \mathcal{C}_\alpha} \frac{\psi_i^{-1}(t)}{ng(t)} = \sum_{i \in \mathcal{N}_0} \sup_{t \in \mathcal{C}_\alpha} \frac{w_i t}{nt} \leq \frac{1}{n} \sum_{i \in \mathcal{N}_0} w_i \leq 1.$$

Therefore, as a consequence of Exercise 6.2, the weighted BH procedure controls the FDR at level α. The weighted BH procedure allows each p-value to be compared with different thresholds. For properly chosen weights, the weighted BH procedure is expected to deliver higher power.

2 BH procedure under arbitrary dependence

The previous theorem relies on the assumption that the null p-values are mutually independent. In many real applications, one would expect the p-values to be dependent. We show that the BH procedure at level α can control the FDR at the level $S_n\alpha$ for

$$S_n = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} \approx \log(n) + 0.577.$$

Theorem. Under (arbitrary) dependence among the p-values, the FDR of the BH procedure is controlled at the level $n_0 S_n \alpha/n$.

Proof. As in Section 1.5, we write

$$\mathrm{FDP} = \sum_{i\in\mathcal{N}_0} \frac{V_i}{R\vee 1}.$$

It suffices to show that

$$\mathbb{E}\left[\frac{V_i}{R\vee 1}\right] \le \frac{S_n\alpha}{n}.$$

We note that when $R = k$, the threshold $T = k\alpha/n$. Letting $a_k = k\alpha/n$, we have

$$\begin{aligned}
\frac{V_i}{R\vee 1} &= \sum_{k=1}^{n} \frac{V_i}{R\vee 1}\mathbf{1}\{R=k\} \\
&= \sum_{k=1}^{n} \frac{\mathbf{1}\{p_i \le a_k, R=k\}}{k} \\
&= \sum_{k=1}^{n}\sum_{j=1}^{k} \frac{\mathbf{1}\{a_{j-1} < p_i \le a_j, R=k\}}{k} \\
&= \sum_{j=1}^{n}\sum_{k\ge j} \frac{\mathbf{1}\{a_{j-1} < p_i \le a_j, R=k\}}{k} \\
&\le \sum_{j=1}^{n} \frac{\mathbf{1}\{a_{j-1} < p_i \le a_j, R\ge j\}}{j} \\
&\le \sum_{j=1}^{n} \frac{\mathbf{1}\{a_{j-1} < p_i \le a_j\}}{j}.
\end{aligned}$$

Taking expectations on both sides, we obtain

$$\mathbb{E}\left[\frac{V_i}{R\vee 1}\right] \le \sum_{j=1}^{n} \frac{P(a_{j-1} < p_i \le a_j)}{j} = \frac{S_n\alpha}{n}.$$

3 Barber and Candès procedure

Barber and Candès (2015) proposed a model-free multiple testing procedure (BC procedure) that exploits the symmetry of the null p-values (or test statistics) to estimate the number of false rejections. More precisely, the BC procedure specifies a data-dependent threshold, denoted by T, which is determined as follows:

$$T = \sup\left\{0 < t < 0.5 : \frac{1+\sum_{i=1}^{n}\mathbf{1}\{p_i \ge 1-t\}}{1\vee R(t)} \le \alpha\right\}$$

with $R(t) = \sum_{i=1}^{n}\mathbf{1}\{p_i \le t\}$ and it rejects all H_i with $p_i \le T$. Here, $1+\sum_{i=1}^{n}\mathbf{1}\{p_i \ge 1-t\}$ serves as an estimate of the number of false discoveries as

$$\sum_{i\in\mathcal{N}_0}\mathbf{1}\{p_i \le t\} \approx \sum_{i\in\mathcal{N}_0}\mathbf{1}\{p_i \ge 1-t\} \le 1 + \sum_{i=1}^{n}\mathbf{1}\{p_i \ge 1-t\}$$

where the approximation is due to the symmetry about 0.5 when p-values are under the null and the constant one is important for achieving the finite sample FDR control.

Exercise 6.3: Generate a sequence of independent z-statistics $(Z_1, \ldots, Z_n)$ with $n = 1000$. Let the number of true nulls $n_0 = 50$. Under the null, $Z_i \sim N(0,1)$ while under the alternative, $Z_i \sim N(\mu\log(n), 1)$, where μ controls the signal strength. Subsequently, the p-values are computed as $p_i = 1-\Phi(X_i)$, where Φ denotes the CDF of the standard normal distribution. We define the empirical power of a multiple testing procedure as

$$\mathrm{POW} = \frac{S}{n-n_0}.$$

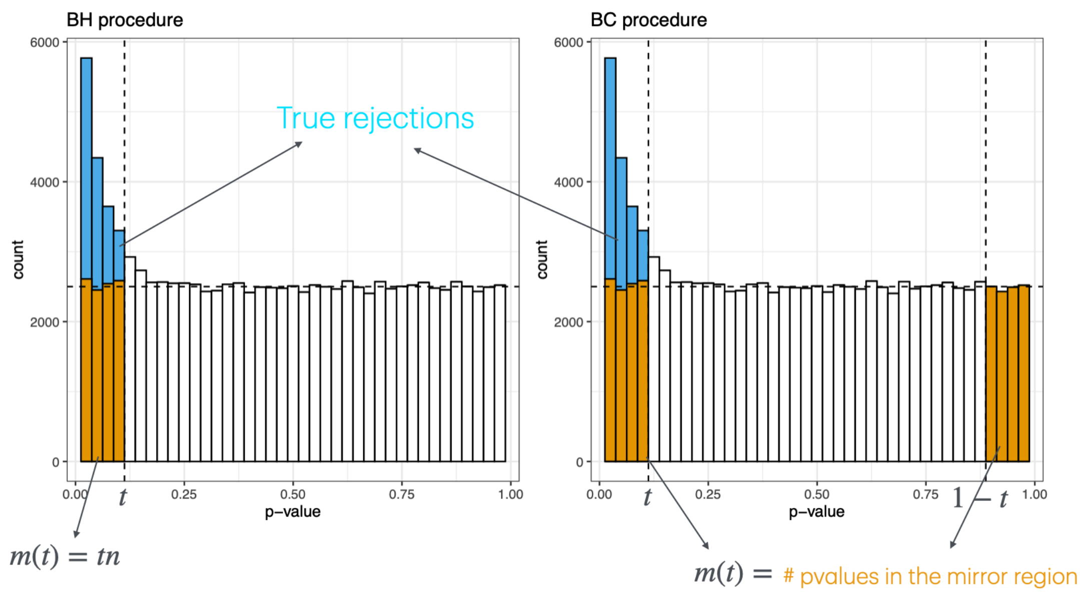

Figure 1: A comparison between the BH and BC procedures

Compute the FDP and power of the BH and BC procedures. Repeat the above simulation 100 times and report the average FDP and power for μ within a certain range.

3.1 Extra reading: FDR control theory

The BC procedure has been shown to provide finite sample FDR control under suitable assumptions in Barber and Candès (2015). Here, we provide proof based on the leave-one-out argument.

Theorem. Suppose the null p-values are mutually independent and are independent of the alternative p-values. The BC procedure controls the FDR at level α.

The proof relies on the following lemma.

Lemma. Let T_i be the threshold for the BC procedure when p_i is replaced with $\min\{p_i, 1-p_i\}$. For any i, j, if $\min(p_i, p_j) \geq 1 - \max\{T_i, T_j\}$, then we have $T_i = T_j$.

Proof of the Lemma. First, given a p-value vector $\mathbf{p} = (p_1, \cdots, p_n)$, recall that the threshold T is defined as

$$T = \max\left\{0 < t < 0.5\colon \underbrace{\frac{1 + \sum_{l=1}^{n} \mathbf{1}\{1 - p_l \leq t\}}{\sum_{l=1}^{n} \mathbf{1}\{p_l \leq t\}}}_{g(\mathbf{p},t)} \leq \alpha\right\}.$$

Without loss of generality, let us assume $T_i \geq T_j$. By the assumption that $\max\{1-p_i, 1-p_j\} \leq \max\{T_i, T_j\}$, we have $1 - p_i \leq T_i < 0.5$ and $1 - p_j \leq T_i < 0.5$. Thus $p_i > 0.5 > T_i$. The same discussion for p_j leads to $p_j > T_i$.

Denote $\tilde{p}_i = \min\{p_i, 1-p_i\}$ and $\mathbf{p}_{-i} = (p_1, \cdots, p_{i-1}, \tilde{p}_i, p_{i+1}, \cdots, p_n)$ for all i. Consider the function

$$g(\mathbf{p}_{-j}, T_i) = \frac{1 + \sum_{l=1}^{n} \mathbf{1}\{1 - p_{-j,l} \leq T_i\}}{\sum_{l=1}^{n} \mathbf{1}\{p_{-j,l} \leq T_i\}},$$

where $p_{-j,l}$ is the lth entry of p_{-j}. For the denominator, we have

$$\begin{aligned}
&\sum_{l=1}^{n} \mathbf{1}\{p_{-j,l} \le T_i\} \\
=&\sum_{l=1}^{n} \mathbf{1}\{p_{-i,l} \le T_i\} + \underbrace{\mathbf{1}\{p_{-j,j} \le T_i\}}_{=1} + \underbrace{\mathbf{1}\{p_{-j,i} \le T_i\}}_{=0} \\
&\quad - \underbrace{\mathbf{1}\{p_{-i,j} \le T_i\}}_{=0} - \underbrace{\mathbf{1}\{p_{-i,i} \le T_i\}}_{=1} \\
=&\sum_{l=1}^{n} \mathbf{1}\{p_{-i,l} \le T_i\}.
\end{aligned}$$

Similarly, for the numerator, we have

$$\begin{aligned}
&\sum_{l=1}^{n} \mathbf{1}\{1 - p_{-j,l} \le T_i\} \\
=&\sum_{l=1}^{n} \mathbf{1}\{1 - p_{-i,l} \le T_i\} + \underbrace{\mathbf{1}\{1 - p_{-j,j} \le T_i\}}_{=0} \\
&\quad + \underbrace{\mathbf{1}\{1 - p_{-j,i} \le T_i\}}_{=1} - \underbrace{\mathbf{1}\{1 - p_{-i,j} \le T_i\}}_{=1} - \underbrace{\mathbf{1}\{1 - p_{-i,i} \le T_i\}}_{0} \\
=&\sum_{l=1}^{n} \mathbf{1}\{1 - p_{-i,l} \le T_i\}.
\end{aligned}$$

Hence, $g(\mathbf{p}_{-j}, T_i) = g(\mathbf{p}_{-i}, T_i) \le \alpha$. By the definition of T_j, we must have $T_i \le T_j$. Similarly, we get $T_j \le T_i$ and hence $T_i = T_j$.

Proof of the theorem. First, note that

$$\begin{aligned}
&\mathbb{E}\left[\sum_{i \in \mathcal{N}_0} \frac{\mathbf{1}\{p_i \le T\}}{1 \vee \sum_{j=1}^{n} \mathbf{1}\{p_j \le T\}}\right] \\
=&\sum_{i \in \mathcal{N}_0} \mathbb{E}\left[\frac{\mathbf{1}\{p_i \le T\}}{1 \vee \sum_{j=1}^{n} \mathbf{1}\{p_j \le T\}} \frac{1 + \sum_{j=1}^{n} \mathbf{1}\{1 - p_j \le T\}}{1 + \sum_{j=1}^{n} \mathbf{1}\{1 - p_j \le T\}}\right] \\
\le&\alpha \sum_{i \in \mathcal{N}_0} \mathbb{E}\left[\frac{\mathbf{1}\{p_i \le T\}}{1 + \sum_{j=1}^{n} \mathbf{1}\{1 - p_j \le T\}}\right].
\end{aligned}$$

Hence, we only need to show that

$$\sum_{i \in \mathcal{N}_0} \mathbb{E}\left[\frac{\mathbf{1}\{p_i \le T\}}{1 + \sum_{j=1}^{n} \mathbf{1}\{1 - p_j \le T\}}\right] \le 1.$$

Let $\tilde{p}_i = \min\{p_i, 1 - p_i\}$ and $\mathbf{p}_{-i} = (p_1, \ldots, p_{i-1}, \tilde{p}_i, p_{i+1}, \ldots, p_n)$. Define $T_i = T(\mathbf{p}_{-i})$, where we view T as a function of the p-values. Notice that if $p_i \le T$, then we have $p_i \le T < 0.5$. Hence, if the ith hypothesis is rejected, then $p_i = \tilde{p}_i$. Thus, $\mathbf{1}\{p_i \le T\} = \mathbf{1}\{p_i \le T_i\}$, which further implies that

$$\mathbb{E}\left[\frac{\mathbf{1}\{p_i \le T\}}{1 + \sum_{j=1}^{n} \mathbf{1}\{1 - p_j \le T\}}\right] = \mathbb{E}\left[\frac{\mathbf{1}\{p_i \le T_i\}}{1 + \sum_{j \ne i} \mathbf{1}\{1 - p_j \le T_i\}}\right],$$

where we use the fact that if $p_i < 0.5$, then $1 - p_i \geq 0.5 > T_i$. Let $\mathcal{F}_i$ be the sigma algebra generated by $\mathbf{p}_{-i}$. For $i \in \mathcal{H}_0$, we have

$$\begin{aligned}
\mathbb{E}\left[\frac{\mathbf{1}\{p_i \leq T\}}{1+\sum_{j=1}^n \mathbf{1}\{1-p_j \leq T\}}\right] =& \mathbb{E}\left[\mathbb{E}\left[\frac{\mathbf{1}\{p_i \leq T_i\}}{1+\sum_{j\neq i} \mathbf{1}\{1-p_j \leq T_i\}}\middle|\mathcal{F}_i\right]\right] \\
=& \mathbb{E}\left[\frac{1}{1+\sum_{j\neq i} \mathbf{1}\{1-p_j \leq T_i\}}\mathbb{E}\left[\mathbf{1}\{p_i \leq T_i\}|\mathcal{F}_i\right]\right] \\
\leq& \mathbb{E}\left[\frac{1}{1+\sum_{j\neq i} \mathbf{1}\{1-p_j \leq T_i\}}\mathbb{E}\left[\mathbf{1}\{1-p_j \leq T_i\}|\mathcal{F}_i\right]\right] \\
=& \mathbb{E}\left[\frac{\mathbf{1}\{1-p_j \leq T_i\}}{1+\sum_{j\neq i} \mathbf{1}\{1-p_j \leq T_i\}}\right],
\end{aligned}$$

where we use the symmetry of the distribution of p_i under the null to get the inequality. By the Lemma, we have

$$\frac{\mathbf{1}\{1-p_j \leq T_i\}}{1+\sum_{j\neq i} \mathbf{1}\{1-p_j \leq T_i\}} = \frac{\mathbf{1}\{1-p_j \leq T_i\}}{1+\sum_{j\neq i} \mathbf{1}\{1-p_j \leq T_j\}} = \frac{\mathbf{1}\{1-p_j \leq T_i\}}{\sum_{j=1}^n \mathbf{1}\{1-p_j \leq T_j\}}.$$

If $1 - p_j > T_i$, both sides are equal to 0. If $1 - p_j \leq T_i$, we claim that $\mathbf{1}\{1 - p_j \leq T_i\} = \mathbf{1}\{1 - p_j \leq T_j\}$. Indeed, if $1 - p_j > T_i$ but $1 - p_j \leq T_j$, then we have $T_i < T_j$. Hence, $1 - p_j \leq T_i < T_j$. By the Lemma, we have $T_i = T_j$, which contradicts with the assumption $T_i < T_j$. The other direction can be proved similarly.

Hence,

$$\sum_{i\in\mathcal{N}_0} \mathbb{E}\left[\frac{\mathbf{1}\{p_i \leq T\}}{1+\sum_{j=1}^n \mathbf{1}\{1-p_j \leq T\}}\right] \leq \mathbb{E}\left[\frac{\sum_{i\in\mathcal{N}_0} \mathbf{1}\{1-p_j \leq T_i\}}{\sum_{j=1}^n \mathbf{1}\{1-p_j \leq T_j\}}\right] \leq 1, \tag{2}$$

which finishes the proof.