Biostatistics (2016), **0**, 0, *pp.* 1–11
doi:10.1093/biostatistics/Supply

Supplementary material to: A non-randomized procedure for discrete multiple testing based on randomized tests

XIAOYU DAI

Department of Mathematics, Washington University in St. Louis,

One Brookings Drive, Saint Louis, MO 63130, USA

NAN LIN*

Department of Mathematics, Washington University in St. Louis,
(joint) Division of Biostatistics, Washington University in St. Louis, School of Medicine,

One Brookings Drive, Saint Louis, MO 63130, USA

nlin@wustl.edu

DAOFENG LI, TING WANG

Department of Genetics, Washington University in St. Louis,

One Brookings Drive, Saint Louis, MO 63130, USA

This supplementary material contains the proofs of Lemma 1, Theorem 3.1, Theorem 3.2, Theorem 3.3 and Theorem 3.4 in the main paper. And it also gives the details of merging nearby significant CpG sites into DMRs mentioned in Section 5 of the main paper.

*Corresponding author.

Appendix A: Proof of Theorem 3.1

Without loss of generality, assume that among the m tests, the null hypotheses from the first k tests are true and the alternative hypotheses from the other tests are true, i.e.

$$\Delta_i = 0, i = 1, ..., k, \qquad \text{and} \qquad \Delta_i = 1, i = k+1, ..., m.$$

Then for $i = 1, ..., k$, suppose that the p-value P_i from the ith test follows the discrete null distribution with cdf H_{0i} and support $\mathcal{P}_i$. Then $H_{0i}(p_i) = p_i$, for any $p_i \in \mathcal{P}_i$ (Westfall and Wolfinger, 1997). And for $i = k+1, ..., m$, suppose that P_i has the alternative cdf H_{1i} and support $\mathcal{P}_i$.

For the ith test, let $b_{i\lambda} = \min\{p : p \geqslant \lambda, p \in \mathcal{P}_i\}$ and $a_{i\lambda} = \max\{p : p < \lambda, p \in \mathcal{P}_i \cup \{0\}\}$, i.e. $a_{i\lambda} < \lambda \leqslant b_{i\lambda}$ are two consecutive values in $\mathcal{P}_i$ around λ. Then, based on the Definition 2.1 in the main paper, for $i = 1, ..., m$, the conditional MCF for the ith test, r_i, follows the discrete distribution with density,

$$g(r_i) = \begin{cases} Pr(p_i \leqslant a_{i\lambda}) = H_{\Delta_i i}(a_{i\lambda}), & \text{if } r_i = 1, \\ Pr(p_i = b_{i\lambda}) = H_{\Delta_i i}(b_{i\lambda}) - H_{\Delta_i i}(a_{i\lambda}), & \text{if } r_i = w_i(\lambda), \\ Pr(p_i > b_{i\lambda}) = 1 - H_{\Delta_i i}(b_{i\lambda}), & \text{if } r_i = 0, \end{cases}$$

where $w_i(\lambda) = \frac{\lambda - a_{i\lambda}}{b_{i\lambda} - a_{i\lambda}}$.

So marginally the MCF follows the average distribution G^m with p.m.f

$$g^m(r) = \begin{cases} \frac{1}{m}\sum_{i=1}^m H_{\Delta_i i}(a_{i\lambda}), & \text{if } r = 1, \\ \frac{1}{m}[H_{\Delta_i i}(b_{i\lambda}) - H_{\Delta_i i}(a_{i\lambda})], & \text{if } r = w_i(\lambda), \qquad i = 1, ..., m, \\ 1 - \frac{1}{m}\sum_{i=1}^m H_{\Delta_i i}(b_{i\lambda}), & \text{if } r = 0. \end{cases}$$

When an MCF r comes from a true null test, i.e. one of the first k tests, it follows the average distribution G_0^m with p.m.f,

$$g_0^m(r) = \begin{cases} \frac{1}{k}\sum_{i=1}^k H_{\Delta_i i}(a_{i\lambda}) = \frac{1}{k}\sum_{i=1}^k a_{i\lambda}, & \text{if } r = 1, \\ \frac{1}{k}[H_{\Delta_i i}(b_{i\lambda}) - H_{\Delta_i i}(a_{i\lambda})] = \frac{1}{k}[b_{i\lambda} - a_{i\lambda}], & \text{if } r = w_i(\lambda), \qquad i = 1, ..., k, \\ 1 - \frac{1}{k}\sum_{i=1}^k H_{\Delta_i i}(b_{i\lambda}) = 1 - \frac{1}{k}\sum_{i=1}^k b_{i\lambda}, & \text{if } r = 0. \end{cases}$$

Similarly, when r comes from a true non-null test, it follows G_1^m with p.m.f,

$$g_1^m(r) = \begin{cases} \frac{1}{m-k}\sum_{i=k+1}^m H_{1i}(a_{i\lambda}), & \text{if } r=1, \\ \frac{1}{m-k}[H_{1i}(b_{i\lambda}) - H_{1i}(a_{i\lambda})], & \text{if } r = w_i(\lambda), \quad i=k+1,...,m, \\ 1-\frac{1}{m-k}\sum_{i=k+1}^m H_{1i}(b_{i\lambda}), & \text{if } r=0. \end{cases}$$

In our method, among the total m tests, we reject the tests with the largest $n_m = \lfloor m(\pi_0\lambda + \pi_1 F_1(\lambda))\rfloor$ MCFs r_i. Here $F_1(\lambda) = \frac{1}{m}\sum_{i=1}^m F_{1i}$ and $\lim_{m\to\infty} F_1 \to F_1^*$. Define the $(1-n_m/m)$-th quantile of the empirical distribution of $(r_1,...,r_m)$ by $\tilde{q}_{n_m/m}(r_1,...,r_m)$. It is equivalent to rejecting the ith test if $r_i > \tilde{q}_{n_m/m}(r_1,...,r_m)$. Note that

$$\frac{n_m}{m} = \frac{\lfloor m(\pi_0\lambda + \pi_1 F_1(\lambda))\rfloor}{m} \to \pi_0\lambda + \pi_1 F_1^*(\lambda), \qquad m\to\infty.$$

So if we let $u = \pi_0\lambda + \pi_1 F_1^*(\lambda)$ and let q_u^m denote the $(1-u)$-th quantile of G^m, by the Glivenko-Cantelli theorem in Billingsley (2013), we have, as $m\to\infty$,

$$\tilde{q}_{n_m/m}(r_1,...,r_m) - q_u^m \xrightarrow{a.s.} 0.$$

Let pFDR(m,λ) be the pFDR of our method, then based on Theorem 1 in (Storey, 2003), we have, as $m\to\infty$,

$$\begin{aligned} &\mathrm{pFDR}(m,\lambda) - \frac{\pi_0 \Pr(r > q_u^m | r\sim G_0^m)}{u} \\ =&\frac{\pi_0\Pr(r > \tilde{q}_{n/m}(r_1,...,r_m)|r\sim G_0^m, r_i\sim G^m)}{n_m/m} - \frac{\pi_0\Pr(r>q_u^m|r\sim G_0^m)}{u} \xrightarrow{a.s.} 0, \end{aligned} \tag{1}$$

where $r\sim G$ means that random variable r follows distribution G.

Next we consider $\Pr(r > q_u^m | r\sim G_0^m)$. Without loss of generality, we first assume that

$$w_1(\lambda) \geqslant ... \geqslant w_t(\lambda) > q_u^m \geqslant w_{t+1}(\lambda) \geqslant ... \geqslant w_k(\lambda),$$

$$w_{k+1}(\lambda) \geqslant ... \geqslant w_s(\lambda) > q_u^m \geqslant w_{s+1}(\lambda) \geqslant ... \geqslant w_m(\lambda).$$

Thus,

$$\Pr(r > q_u^m | r\sim G_0^m) = \frac{1}{k}[\sum_{i=1}^k a_{i\lambda} + \sum_{i=1}^t (b_{i\lambda} - a_{i\lambda})]. \tag{2}$$

On the other hand, as $u = \pi_0\lambda + \pi_1 F_1^*(\lambda)$ and

$$
\begin{aligned}
u &= Pr(r > q_u^m | r \sim G^m) \\
&= \frac{1}{m}\Bigg[\sum_{i=1}^{k} a_{i\lambda} + \sum_{i=k+1}^{m} H_{1i}(a_{i\lambda}) + \sum_{i=1}^{t}(b_{i\lambda} - a_{i\lambda}) + \sum_{i=k+1}^{s}(H_{1i}(b_{i\lambda}) - H_{1i}(a_{i\lambda}))\Bigg]. \qquad (3)
\end{aligned}
$$

Also note that MCF represents the probability that the randomized p-value is less than λ, and F_1^* is the cdf of the limit alternative distribution of the randomized p-value when $m \to \infty$, so, as $m \to \infty$,

$$
\frac{1}{m-k}\Bigg[\sum_{i=k+1}^{m} H_{1i}(a_{i\lambda}) + \sum_{i=k+1}^{m} w_i(\lambda)(H_{1i}(b_{i\lambda}) - H_{1i}(a_{i\lambda}))\Bigg] = E(r | r \sim G_1^m) \to F_1^*(\lambda).
$$

Note that $\frac{k}{m} \to \pi_0$ and $\frac{m-k}{m} \to \pi_1$ as $m \to \infty$, then we have, as $m \to \infty$,

$$
\begin{aligned}
&\frac{k}{m}\lambda + \frac{1}{m}\sum_{i=k+1}^{m} H_{1i}(a_{i\lambda}) + \frac{1}{m}\sum_{i=k+1}^{m} w_i(\lambda)[H_{1i}(b_{i\lambda}) - H_{1i}(a_{i\lambda})] \\
&\to \frac{k}{m}\lambda + \frac{m-k}{m}F_1^*(\lambda) \\
&\to \pi_0\lambda + \pi_1 F_1^*(\lambda). \qquad (4)
\end{aligned}
$$

So combining (3) and (4), we have

$$\begin{aligned}
&\lim_{m\to\infty}\frac{\sum_{i=1}^k a_{i\lambda}+\sum_{i=1}^t(b_{i\lambda}-a_{i\lambda})}{k}\\
&=\lim_{m\to\infty}\left\{\frac{m}{k}(\pi_0\lambda+\pi_1F_1(\lambda))-\frac{1}{k}\left[\sum_{i=k+1}^m H_{1i}(a_{i\lambda})+\sum_{i=k+1}^s H_{1i}(b_{i\lambda})-H_{1i}(a_{i\lambda})\right]\right\}\\
&=\lim_{m\to\infty}\frac{m}{k}\left\{\frac{k}{m}\lambda+\frac{1}{m}\sum_{i=k+1}^m H_{1i}(a_{i\lambda})+\frac{1}{m}\sum_{i=k+1}^m w_i(\lambda)[H_{1i}(b_{i\lambda})-H_{1i}(a_{i\lambda})]\right\}\\
&\quad-\lim_{m\to\infty}\frac{1}{k}\left[\sum_{i=k+1}^m H_{1i}(a_{i\lambda})+\sum_{i=k+1}^s H_{1i}(b_{i\lambda})-H_{1i}(a_{i\lambda})\right]\\
&=\lambda+\lim_{m\to\infty}\left\{\frac{1}{k}\sum_{i=s+1}^m w_i(\lambda)(H_{1i}(b_{i\lambda})-H_{1i}(a_{i\lambda}))-\frac{1}{k}\sum_{i=k+1}^s(1-w_i(\lambda))(H_{1i}(b_{i\lambda})-H_{1i}(a_{i\lambda}))\right\}\\
&\leqslant\lambda+\lim_{m\to\infty}\frac{1}{k}\sum_{i=s+1}^m w_i(\lambda)(H_{1i}(b_{i\lambda})-H_{1i}(a_{i\lambda}))\\
&\leqslant\lambda+\lim_{m\to\infty}\frac{m-s}{k}q_u^m\cdot c\\
&\leqslant\lambda+\lim_{m\to\infty}\frac{m-k}{k}q_u^m\cdot c\\
&=\lambda+\frac{\pi_1}{\pi_0}q_uc, \qquad (5)
\end{aligned}$$

where $q_u=\lim_{m\to\infty}q_u^m$. The second inequality is due to that $q_u^m\geqslant w_i(\lambda)$ for $i=s+1,...,m$ and $H_{1i}(b_{i\lambda})-H_{1i}(a_{i\lambda})\leqslant c$, as H_{1i} satisfies Condition 3.1.

On the other hand, for $i=1,2,...,t$, $\frac{\lambda-a_{i\lambda}}{b_{i\lambda}-a_{i\lambda}}=w_i(\lambda)\geqslant q_u^m$, and $q_u^m\leqslant 1$. So

$$b_{i\lambda}\leqslant a_{i\lambda}+\frac{\lambda-a_{i\lambda}}{q_u^m}\leqslant\frac{\lambda}{q_u^m},\qquad i=1,2,...,t$$

So if we let $\rho=\lim_{m\to\infty}\Pr(r>q_u^m|r\sim G_0^m)=\lim_{m\to\infty}\frac{\sum_{i=1}^k a_{i\lambda}+\sum_{i=1}^t(b_{i\lambda}-a_{i\lambda})}{k}$, then

$$\rho=\lim_{m\to\infty}\frac{\sum_{i=k+1}^t a_{i\lambda}+\sum_{i=1}^t b_{i\lambda}}{k}\leqslant\lim_{m\to\infty}\frac{(t-k)\lambda+t\cdot\lambda/q_u^m}{k}\leqslant\lim_{m\to\infty}\frac{k\cdot\lambda/q_u^m}{k}=\frac{\lambda}{q_u}. \qquad (6)$$

From (5) and (6), we get

$$\rho\leqslant\lambda+\frac{\pi_1}{\pi_0}q_uc\leqslant\lambda+\frac{c\pi_1}{\pi_0}\frac{\lambda}{\rho}.$$

So,

$$\rho\leqslant\frac{\lambda+\sqrt{\lambda^2+\frac{4c\pi_1}{\pi_0}\lambda}}{2}. \qquad (7)$$

Combining (1), (2) and (7), we proved

$$\limsup_{m\to\infty} \mathrm{pFDR}(m,\lambda) \leqslant \frac{\pi_0\rho}{u} \leqslant \frac{\pi_0 \frac{\lambda+\sqrt{\lambda^2+\frac{4c\pi_1}{\pi_0}\lambda}}{2}}{\pi_0\lambda+\pi_1 F_1^*(\lambda)}.$$

Appendix B: Proof of Theorem 3.2

All the notations we used are the same as in the proof of Theorem 3.1.

From the fifth line of equation (5), let $w_i = w_i(\lambda)$ for simplicity, we have,

$$\begin{aligned}
\rho &= \lim_{m\to\infty} \frac{\sum_{i=1}^k a_{i\lambda} + \sum_{i=1}^t (b_{i\lambda}-a_{i\lambda})}{k} \\
&= \lambda + \lim_{m\to\infty} \left\{ \frac{1}{k}\sum_{i=s+1}^m w_i[H_{1i}(b_{i\lambda})-H_{1i}(a_{i\lambda})] - \frac{1}{k}\sum_{i=k+1}^s (1-w_i)[H_{1i}(b_{i\lambda})-H_{1i}(a_{i\lambda})] \right\} \\
&= \lambda + \lim_{m\to\infty} \frac{1}{k}\left\{ \sum_{i=k+1}^m w_i[H_{1i}(b_{i\lambda})-H_{1i}(a_{i\lambda})] - \sum_{i=k+1}^s [H_{1i}(b_{i\lambda})-H_{1i}(a_{i\lambda})] \right\} \\
&= \lambda + (A-B), \qquad (8)
\end{aligned}$$

where $A = \lim_{m\to\infty} \frac{1}{k}\sum_{i=k+1}^m w_i[H_{1i}(b_{i\lambda}) - H_{1i}(a_{i\lambda})]$ and $B = \lim_{m\to\infty} \frac{1}{k}\sum_{i=k+1}^s [H_{1i}(b_{i\lambda}) - H_{1i}(a_{i\lambda})]$.

Also we have,

$$\begin{aligned}
\rho &= \lim_{m\to\infty} \frac{\sum_{i=1}^k a_{i\lambda} + \sum_{i=1}^t (b_{i\lambda}-a_{i\lambda})}{k} \\
&= \lambda + \lim_{m\to\infty} \frac{1}{k}\left[\sum_{i=1}^k a_{i\lambda} + \sum_{i=1}^t (b_{i\lambda}-a_{i\lambda}) - k\lambda \right] \\
&= \lambda + \lim_{m\to\infty} \frac{1}{k}\left[\sum_{i=1}^t (b_{i\lambda}-a_{i\lambda}) - \sum_{i=1}^k \frac{\lambda - a_{i\lambda}}{b_{i\lambda}-a_{i\lambda}}(b_{i\lambda}-a_{i\lambda}) \right] \\
&= \lambda + \lim_{m\to\infty} \frac{1}{k}\left[\sum_{i=1}^t (b_{i\lambda}-a_{i\lambda}) - \sum_{i=1}^k w_i(b_{i\lambda}-a_{i\lambda}) \right] \\
&= \lambda + (C-D), \qquad (9)
\end{aligned}$$

where $C = \lim_{m\to\infty} \frac{1}{k}\sum_{i=1}^t (b_{i\lambda}-a_{i\lambda})$ and $D = \lim_{m\to\infty} \frac{1}{k}\sum_{i=1}^k w_i(b_{i\lambda}-a_{i\lambda})$.

From equation (8) and (9), we know $A - B = C - D$, where A, B, C, D are non-negative.

Also, under the assumption of the Theorem 3.2 in the main paper,

$$
\begin{aligned}
AC - BD &= \lim_{m\to\infty} \left\{ \frac{1}{k} \sum_{i=k+1}^{m} w_i [H_{1i}(b_{i\lambda}) - H_{1i}(a_{i\lambda})] \frac{1}{k} \sum_{i=1}^{t} (b_{i\lambda} - a_{i\lambda}) \right. \\
&\qquad \left. - \frac{1}{k} \sum_{i=k+1}^{s} [H_{1i}(b_{i\lambda}) - H_{1i}(a_{i\lambda})] \frac{1}{k} \sum_{i=1}^{k} w_i (b_{i\lambda} - a_{i\lambda}) \right\} \\
&= \lim_{m\to\infty} \left\{ \frac{m-k}{k} \frac{1}{m-k} \sum_{i=k+1}^{m} w_i [H_{1i}(b_{i\lambda}) - H_{1i}(a_{i\lambda})] \frac{1}{k} \sum_{i=1}^{k} I_{\{w_i > q\}} [H_{0i}(b_{i\lambda}) - H_{0i}(a_{i\lambda})] \right. \\
&\qquad \left. - \frac{m-k}{k} \frac{1}{m-k} \sum_{i=k+1}^{m} I_{\{w_i > q\}} [H_{1i}(b_{i\lambda}) - H_{1i}(a_{i\lambda})] \frac{1}{k} \sum_{i=1}^{k} w_i [H_{0i}(b_{i\lambda}) - H_{0i}(a_{i\lambda})] \right\} \\
&= \frac{\pi_1}{\pi_0} \lim_{m\to\infty} \left\{ \frac{1}{|\mathcal{H}_1|} \sum_{i\in\mathcal{H}_1} w_i [H_i(b_{i\lambda}) - H_i(a_{i\lambda})] \frac{1}{|\mathcal{H}_0|} \sum_{i\in\mathcal{H}_0} I_{\{w_i > q\}} [H_i(b_{i\lambda}) - H_i(a_{i\lambda})] \right. \\
&\qquad \left. - \frac{1}{|\mathcal{H}_0|} \sum_{i\in\mathcal{H}_0} w_i [H_i(b_{i\lambda}) - H_i(a_{i\lambda})] \frac{1}{|\mathcal{H}_1|} \sum_{i\in\mathcal{H}_1} I_{\{w_i > q\}} [H_i(b_{i\lambda}) - H_i(a_{i\lambda})] \right\} \\
&= 0.
\end{aligned}
$$

So we get, $A - B = C - D = 0$. Then $\rho = \lambda$, and combining with equation (1) and (2) in the proof of Theorem 3.1, we have,

$$\limsup_{m\to\infty} \mathrm{pFDR}(m,\lambda) \leqslant \frac{\pi_0 \rho}{u} = \frac{\pi_0 \lambda}{\pi_0 \lambda + \pi_1 F_1^*(\lambda)}.$$

Appendix C: Proof of Theorem 3.3

All the notations we used are the same as in the proof of Theorem 3.1.

Similarly to the proof of Theorem 3.1 and based on Theorem 5 in (Storey, 2003), we have, as $m \to \infty$,

$$
\begin{aligned}
&\mathrm{pFNR}(m,\lambda) - \frac{\pi_1 \Pr(r < q_u^m | r \sim G_1^m)}{1-u} \\
&= \frac{\pi_1 \Pr(r < \tilde{q}_{n/m}(r_1, \dots, r_m) | r \sim G_1^m, r_i \sim G^m)}{1 - n_m/m} - \frac{\pi_1 \Pr(r < q_u^m | r \sim G_1^m)}{1-u} \xrightarrow{a.s.} 0. \quad (10)
\end{aligned}
$$

And

$$\begin{aligned}
\Pr(r < q_u^m | r \sim G_1^m) &= \sum_{i=s+1}^{m} \frac{1}{m-k}[H_{1i}(b_{i\lambda}) - H_{1i}(a_{i\lambda})] + 1 - \frac{1}{m-k}\sum_{i=k+1}^{m} H_{1i}(b_{i\lambda}) \\
&= 1 - \frac{1}{m-k}\Bigg[\sum_{i=s+1}^{m} H_{1i}(a_{i\lambda}) + \sum_{i=k+1}^{s} H_{1i}(b_{i\lambda})\Bigg].
\end{aligned}$$

Let $\gamma = \sum_{i=s+1}^{m} H_{1i}(a_{i\lambda}) + \sum_{i=k+1}^{s} H_{1i}(b_{i\lambda})$, then based on equation (3) from the proof of Theorem 3.1, we have

$$\begin{aligned}
\pi_0\lambda + \pi_1 F_1^*(\lambda) &= \frac{1}{m}\Bigg[\sum_{i=1}^{k} a_{i\lambda} + \sum_{i=k+1}^{m} H_{1i}(a_{i\lambda}) \\
&\quad + \sum_{i=1}^{t}(b_{i\lambda} - a_{i\lambda}) + \sum_{i=k+1}^{s}(H_{1i}(b_{i\lambda}) - H_{1i}(a_{i\lambda}))\Bigg] \\
&= \frac{1}{m}\Bigg(\sum_{i=t+1}^{k} a_{i\lambda} + \sum_{i=1}^{t} b_{i\lambda} + \gamma\Bigg).
\end{aligned}$$

So

$$\begin{aligned}
\Pr(r < q_u^m | r \sim G_1^m) &= 1 - \frac{1}{m-k}\gamma \\
&= 1 - \frac{1}{m-k}\Bigg[m(\pi_0\lambda + \pi_1 F_1^*(\lambda)) - \Bigg(\sum_{i=t+1}^{k} a_{i\lambda} + \sum_{i=1}^{t} b_{i\lambda}\Bigg)\Bigg] \\
&= 1 - \frac{m}{m-k}[\pi_0\lambda + \pi_1 F_1^*(\lambda)] + \frac{k}{m-k}\frac{\sum_{i=t+1}^{k} a_{i\lambda} + \sum_{i=1}^{t} b_{i\lambda}}{k}.
\end{aligned}$$

Note that $\frac{k}{m} \to \pi_0$ and $\frac{m-k}{m} \to \pi_1$ as $m \to \infty$, then we have,

$$\begin{aligned}
\limsup_{m\to\infty} \Pr(r < q_u^m | r \sim G_1^m) &= 1 - \frac{1}{\pi_1}[\pi_0\lambda + \pi_1 F_1^*(\lambda)] + \frac{\pi_0}{\pi_1}\rho \\
&\leqslant 1 - \frac{1}{\pi_1}[\pi_0\lambda + \pi_1 F_1^*(\lambda)] + \frac{\pi_0}{\pi_1}\frac{\lambda + \sqrt{\lambda^2 + \frac{4c\pi_1}{\pi_0}\lambda}}{2},
\end{aligned}$$

where $\rho = \lim_{m\to\infty} \frac{\sum_{i=1}^{k} a_{i\lambda} + \sum_{i=1}^{t}(b_{i\lambda} - a_{i\lambda})}{k}$, and the inequality is due to (7) from the proof of

Theorem 3.1. Then from (10) we have,

$$\limsup_{m\to\infty} \mathrm{pFNR}(m,\lambda) \leqslant \frac{\pi_1\left\{1-\frac{1}{\pi_1}[\pi_0\lambda+\pi_1F_1^*(\lambda)]+\frac{\pi_0}{\pi_1}\frac{\lambda+\sqrt{\lambda^2+\frac{4c\pi_1}{\pi_0}\lambda}}{2}\right\}}{1-[\pi_0\lambda+\pi_1F_1(\lambda)]}$$
$$= \frac{\pi_1(1-F_1^*(\lambda))+\pi_0\frac{\sqrt{\lambda^2+\frac{4c\pi_1}{\pi_0}\lambda}-\lambda}{2}}{1-[\pi_0\lambda+\pi_1F_1^*(\lambda)]}.$$

Appendix D: Proof of Theorem 3.3

Assume that, for the ith test, the discrete p-value follows the two-component model

$$p_i|\Delta_i \sim (1-\Delta_i)G_{0i}+\Delta_i G_{1i},$$

where $\Delta_i \sim \mathit{Bernoulli}(\pi_1)$, and G_{0i}, G_{1i} are the null and alternative distribution.

Denote the corresponding alternative distribution for randomized p-value by F_{1i}, then

$$\tilde{p}_i|\Delta_i \sim (1-\Delta_i)U[0,1]+\Delta_i F_{1i}.$$

So marginally,

$$P \sim \pi_0 G_0+\pi_1 G_1,$$

$$\tilde{P} \sim \pi_0 U[0,1]+\pi_1 F_1,$$

where $G_0=\frac{1}{m}\sum_{i=1}^m G_{0i}$, $G_1=\frac{1}{m}\sum_{i=1}^m G_{1i}$ and $F_1=\frac{1}{m}\sum_{i=1}^m F_{1i}$.

The q-value method defines q-value for the ith test as

$$q(p_i)=\inf\left\{\frac{\pi_0\lambda}{\pi_0\lambda+\pi_1G_1(\lambda)}:\lambda\geqslant p_i\right\},$$

and for a nominal FDR level α, identify the ith test as significant if $q(p_i)\leqslant\alpha$.

This rejection rule is equivalent to identifying the ith test as significant if $p_i\leqslant\lambda^*$, where

$$\lambda^*=\sup\left\{\lambda:\frac{\pi_0\lambda}{\pi_0\lambda+\pi_1G_1(\lambda)}\leqslant\alpha\right\}.$$

Next we consider the MCF-based method. For the same nominal FDR level α, we define

$$\tilde{\lambda}^*=\sup\left\{\lambda:\frac{\pi_0\lambda}{\pi_0\lambda+\pi_1F_1(\lambda)}\leqslant\alpha\right\}.$$

and then calculate MCF values for each test corresponding to $\tilde{\lambda}^*$.

For the ith individual test, following Lehmann and Romano (2006), F_{1i} is stochastically smaller than G_{1i}. So the average distribution F_1 is stochastically smaller than G_1. Then $\frac{\pi_0\lambda^*}{\pi_0\lambda^*+\pi_1 F_1(\lambda^*)} \leqslant \frac{\pi_0\lambda^*}{\pi_0\lambda^*+\pi_1 G_1(\lambda^*)} \leqslant \alpha$. Therefore $\lambda^* \leqslant \tilde{\lambda}^*$.

Suppose that the ith test is identified as significant by the q-value method, i.e. $p_i \leqslant \lambda^*$. Then, we know,

$$p_i^- < \tilde{p}_i < p_i \leqslant \lambda^* \leqslant \tilde{\lambda}^*.$$

As a result, the MCF of the ith test is always 1 and rejected by the MCF-based method.

Appendix E: Lemma 1

Lemma 1 If F_1 is concave, then $\frac{\pi_0\lambda}{\pi_0\lambda+\pi_1 F_1(\lambda)}$ is non-decreasing in terms of λ.

Proof:

Showing that $\frac{\pi_0\lambda}{\pi_0\lambda+\pi_1 F_1(\lambda)}$ is non-decreasing in λ is equivalent to showing that $\frac{\partial}{\partial\lambda}(\frac{\pi_0\lambda}{\pi_0\lambda+\pi_1 F_1(\lambda)}) \geqslant 0 \iff F_1(\lambda) \geqslant \lambda F_1'(\lambda)$. So it suffices to show $F_1(\lambda) \geqslant \lambda F_1'(\lambda)$, i.e. $\int_0^\lambda f_1(t)dt \geqslant \lambda f_1(\lambda)$, where f_1 is the density of F_1. And it suffices to show f_1 is non-increasing. And this is a direct result given that F_1 is concave.

Appendix F: merge significant CpG sites into DMRs

We used a merging technique in Feng *and others* (2014) to combine nearby CpG sites into DMRs. First we grouped nearby signifiant CpG sites with minimal distance of 100 bps, i.e. if the distance between two adjacent significant CpG sites is less than 100 bps they are in the same group. Then we defined each group of significant CpG sites along with the non-significant CpG sites between them as one potential DMR window. For CD4 versus ES cells and CD8 versus ES cells, a window is identified as a DMR if its length is larger than 1000bps, it contains more than 20 CpG sites and of which more than 50% are significant. For CD4 versus CD8 we use a less stringent criterion 50% to ensure that we got enough DMRs to conduct a meaningful enrichment analysis later. And a window is identified as a DMR if its length is larger than 500bps, it contains more than 3 CpG sites and of which more than 50% are significant. Since our MCF-based method gave more significant CpG sites, by using this technique we may get wider windows which includes more significant CpG sites but also more non-significant CpG sites. So if this window is identified as a DMR based on the results from the q-value method, it is not necessarily identified as significant based on the results from our MCF-based method due to its possibly lower proportion of significant CpGs.

References

Billingsley, P. (2013). *Convergence of Probability Measures*. John Wiley & Sons.

Feng, H., Conneely, K.N. and Wu, H. (2014). A Bayesian hierarchical model to detect differentially methylated loci from single nucleotide resolution sequencing data. *Nucleic Acids Research* **42**(8), e69–e69.

Lehmann, E.L. and Romano, J.P. (2006). *Testing Statistical Hypotheses*. Springer Science & Business Media.

Storey, J.D. (2003). The positive false discovery rate: a Bayesian interpretation and the q-value. *Annals of Statistics* **31**, 2013–2035.

Westfall, P.H. and Wolfinger, R.D. (1997). Multiple tests with discrete distributions. *The American Statistician* **51**(1), 3–8.