Correction note on Roquain and Van De Wiel (2009)

Guillermo Durand

June 8, 2018

In Roquain and Van De Wiel (2009), the authors presented an important contribution to the p-value weighting literature, introducing for the first time (to our knowledge) the concept of multi-weighting, and using this new notion to derive optimal weights maximizing the power while controlling the False Discovery Rate (FDR). In the following we refer to the original paper for the notation and the numerotation.

The novelty of their approach can be summarized as follows: instead of assigning to each p-value p_i, $1 \leq i \leq m$ a weight w_i, assign a function $W_i : u \mapsto W_i(u)$ depending on a rejection threshold $u \in [0,1]$. This provides a new multiple testing class of procedures they call *multi-weighted step-up procedures*, denoted $\mathbf{SU}(\mathbf{W})$ (where $\mathbf{W} = (W_i)_i$), which consists in rejecting all $p_i \leq \alpha \hat{u} W_i(\hat{u})$ where

$$\begin{aligned} \hat{u} &= \mathcal{I}\left(\widehat{\mathbb{G}}_{\mathbf{W}}\right) \\ &= \max\left\{u \in [0,1] \middle| \widehat{\mathbb{G}}_{\mathbf{W}}(u) \geq u\right\} \end{aligned}$$

and

$$\widehat{\mathbb{G}}_{\mathbf{W}}(u) = m^{-1} \sum_{i=1}^{m} \mathbf{1}\{p_i \leq \alpha u W_i(u)\}.$$

They then define the optimal weight function $\mathbf{W}^\star$ by maximizing the power at each rejection threshold u:

$$\mathbf{W}^\star(u) \in \underset{\left\{\mathbf{w} \mid \sum_{i=1}^{m} w_i = m\right\}}{\arg\max} \operatorname{Pow}_u(\mathbf{w})$$

where $\operatorname{Pow}_u(\mathbf{w}) = \operatorname{Pow}\left(\{i | p_i \leq \alpha u w_i\}\right)$. Of course, maximizing the power requires the knowledge of the distribution of the p-values under the alternative hypothesis, that is, the c.d.f. F_i.

The two main theorems of the paper give statements about the power optimality of the resulting procedure $\mathbf{SU}(\mathbf{W}^\star)$ for finite sample (Theorem 4.2), and asymptotically (Theorem 4.3). However, the proof of these two theorems rely on a technical result, Proposition 8.3, and there is an error in the proof of the statement *(i)* of this proposition. In this correction note, we suggest to state a weaker statement than *(i)* which still allows to prove Theorems 4.2 and 4.3, up to a slightly modification of Theorem 4.3: in equation (16), the lim should be replaced by a lim sup.

The error Take the proof of statement *(i)*, page 702, in the second block of equations, when the authors go from

$$\mathbb{E}\left[\pi_1 m^{-1} \sum_{i=1}^{m} F_i \circ \Delta_i(\phi(\hat{u}'_{-i})) - (1 - \alpha\pi_0)\bar{u}\right]$$

to

$$\pi_1 \mathbb{P}(\Omega_1^c) + G_{\mathbf{W}}(\bar{u} + \lambda) - \alpha\pi_0(\bar{u} + \lambda) - (1 - \alpha\pi_0)\bar{u}.$$

The $(1-\alpha\pi_0)\bar{u}$ can be ignored because it does not change between the two lines. The upper-bounding made by the authors consists in writing

$$\mathbb{E}\left[\pi_1 m^{-1}\sum_{i=1}^m F_i\circ\Delta_i(\phi(\hat{u}'_{-i}))\right] = \pi_1\mathbb{E}\left[m^{-1}\sum_{i=1}^m F_i\circ\Delta_i(\phi(\hat{u}'_{-i}))\left(\mathbb{1}_{\Omega_1^c}+\mathbb{1}_{\Omega_1}\right)\right]$$
$$\leq \pi_1\mathbb{P}(\Omega_1^c) + \pi_1 m^{-1}\sum_{i=1}^m F_i\circ\Delta_i(\bar{u}+\lambda),$$

and then using that

$$\pi_1 m^{-1}\sum_{i=1}^m F_i\circ\Delta_i(\bar{u}+\lambda) = G_{\mathbf{W}}(\bar{u}+\lambda) - \alpha\pi_0(\bar{u}+\lambda). \tag{E}$$

But (E) is not generally true because the c.d.f. of a uniform distribution over $[0,1]$ is $t\geq 0\mapsto t\wedge 1$, not $t\mapsto t$, which means that:

$$\pi_1 m^{-1}\sum_{i=1}^m F_i\circ\Delta_i(u) = G_{\mathbf{W}}(u) - \pi_0 m^{-1}\sum_{i=1}^m\left(\Delta_i(u)\wedge 1\right),\ \forall u\in[0,1]. \tag{Pw}$$

Since $\Delta_i(\bar{u}+\lambda)$ can be >1 in general, (E) fails and the proof cannot be completed.

Note that this does not impact the proof of statement *(ii)* because we still have

$$\pi_1 m^{-1}\sum_{i=1}^m F_i\circ\Delta_i(u) \geq G_{\mathbf{W}}(u) - \alpha\pi_0 u,\ \forall u\in[0,1].$$

Weaker statement We propose to replace previous statement *(i)* of Proposition 8.3 by:

$$\mathrm{Pow}\,\mathbf{SU}(\mathbf{W}) - (1-\alpha\pi_0)u^\star \leq \pi_1 m^2\exp\{-2m(\mathcal{I}_\lambda^+(G_{\mathbf{W}})-m^{-1})_+^2\} + \lambda(1-\alpha\pi_0). \tag{WeakState}$$

Two distinct changes have been made. The less notable one is that we dropped the "$-\mathcal{I}_\lambda^+(G_{\mathbf{W}})$"which turns out to be useless in the rest of the paper (the authors dropped it too as soon as they used the statement in their proofs of each theorem). The second change is that we switched $\bar{u}$ for $u^\star$.

It is easy to see that these changes do not modify at all the proof of Theorem 4.2, see the bottom of page 701 where the statement is used. The proof was already using that $u^\star\geq u_{\mathbf{w}}$ to switch $u_{\mathbf{w}}$ for $u^\star$. Nevertheless the old version of the statement was necessary to show equation (29) page 704, inside the proof of Theorem 4.3. Now, to prove Theorem 4.3, we have to first show (28) as in the original paper, then use (15) for a fixed $\mathbf{w}$, making $m\to\infty$ and then $\lambda\to 0$. This explains why the $\lim_m \mathrm{Pow}\,(\mathbf{LSU}(\mathbf{w}))$ should be replaced by a lim sup.

Proof of (WeakState) First note that equation (Pw) is also the equation of $\mathrm{Pow}_u(\mathbf{W}(u))$ for all $u\in[0,1]$. Note also that for $\mathbf{W}^\star$ instead of a general $\mathbf{W}$, equation (E) is valid because $\Delta_i^\star(u)\leq 1$ for all u (easy to see on the closed formula given by Proposition 3.2). So we start by proceeding as in the original paper until we get (26). Then we write:

$$\begin{aligned}
\mathrm{Pow}(\mathbf{SU}(\mathbf{W})) - (1-\alpha\pi_0)u^\star &\leq \mathbb{E}\left[\pi_1 m^{-1}\sum_{i=1}^m F_i\circ\Delta_i(\phi(\hat{u}'_{-i}))\right] - (1-\alpha\pi_0)u^\star \\
&\leq \pi_1\mathbb{P}(\Omega_1^c) + \pi_1 m^{-1}\sum_{i=1}^m F_i\circ\Delta_i(\bar{u}+\lambda) - (1-\alpha\pi_0)u^\star \\
&\leq \pi_1\mathbb{P}(\Omega_1^c) + \mathrm{Pow}_{\bar{u}+\lambda}(\mathbf{W}(\bar{u}+\lambda)) - (1-\alpha\pi_0)u^\star \\
&\leq \pi_1\mathbb{P}(\Omega_1^c) + \mathrm{Pow}_{u^\star+\lambda}(\mathbf{W}(\bar{u}+\lambda)) - (1-\alpha\pi_0)u^\star && \text{(a)} \\
&\leq \pi_1\mathbb{P}(\Omega_1^c) + \mathrm{Pow}_{u^\star+\lambda}(\mathbf{W}^\star(u^\star+\lambda)) - (1-\alpha\pi_0)u^\star && \text{(b)} \\
&\leq \pi_1\mathbb{P}(\Omega_1^c) + G_{\mathbf{W}^\star}(u^\star+\lambda) - \alpha\pi_0(u^\star+\lambda) - (1-\alpha\pi_0)u^\star && \text{(c)} \\
&\leq \pi_1\mathbb{P}(\Omega_1^c) + G_{\mathbf{W}^\star}(u^\star+\lambda) - (u^\star+\lambda) + \lambda(1-\alpha\pi_0) \\
&\leq \pi_1\mathbb{P}(\Omega_1^c) + \lambda(1-\alpha\pi_0). && \text{(d)}
\end{aligned}$$

Above, we used that $u^\star \geq \bar{u}$ along with that F_i is nondecreasing in (a) and the definition of $\mathbf{W}^\star$ as an arg max in (b). (E) applied to $\Delta^\star$ provides (c), and (d) comes from $G_{\mathbf{W}^\star}(u^\star + \lambda) \leq u^\star + \lambda$. The end of the proof, that is the upper-bounding of $\mathbb{P}(\Omega_1^c)$, is the same as in the original paper.

References

Etienne Roquain and Mark A. Van De Wiel. Optimal weighting for false discovery rate control. *Electronic Journal of Statistics*, 3:678–711, 2009.