UNIVERSITY OF TWENTE.

DEPARTMENT OF APPLIED MATHEMATICS

PRESENTED AT
SAVI 2026

PERMUTATION-BASED FDR CONTROL VIA THE E-CLOSURE PRINCIPLE

Rovanos T. Nzanguim, Aurèle Mingam,
Jelle Goeman, Rianne de Heide

PROBLEM STATEMENT

- Permutation methods provide exact p-values under some dependence and are widely used in modern applications such as genomics.
- Existing permutation-based multiple testing procedures primarily control the *family-wise error rate (FWER)*, which is often too conservative when thousands of hypotheses are tested simultaneously.
- There is currently no general framework for controlling the *false discovery rate (FDR)* using permutation methods without imposing assumptions on the dependence structure of the p-values.

Goal. Develop a permutation-based procedure that provably controls the FDR at level α under arbitrary dependence.

PROPOSED APPROACH

- Construct valid permutation p-values for every intersection hypothesis.
- Transform these p-values into an e-collection using the Su p-to-e calibrator.
- Apply the e-closure principle. The resulting procedure controls the FDR at level α.

PERMUTATION SETUP

Let $H_1, \dots, H_m$ be the hypotheses and let G be a finite set of transformations $\sigma : \Omega \to \Omega$, such that G is a group with respect to the operation of composition of transformations.

Permutation invariance assumption

For the set N of true null hypotheses, the joint distribution of the test statistics $\big(T_i(\sigma D) : i \in N, \sigma \in G\big)$ is invariant under all transformation in G of D.

Permutation p-value matrix

Let $\Sigma = \{\sigma_1 = \mathrm{id}, \sigma_2, \dots, \sigma_\omega\}$ be a vector of randomly sampled elements from G. For each element, compute the vector of p-values $P^\sigma = (p_1^\sigma, \dots, p_m^\sigma)$ and store in a $(\omega \times m)$ matrix

$$P = \begin{pmatrix} p_1^{\sigma_1} & p_1^{\sigma_2} & \cdots & p_1^{\sigma_\omega} \\ p_2^{\sigma_1} & p_2^{\sigma_2} & \cdots & p_2^{\sigma_\omega} \\ \vdots & \vdots & \ddots & \vdots \\ p_m^{\sigma_1} & p_m^{\sigma_2} & \cdots & p_m^{\sigma_\omega} \end{pmatrix}$$

KEY IDEA

For every intersection hypothesis $H_S = \bigcap_{i \in S} H_i, S \in 2^{[m]}$, we construct a valid permutation p-value, transform it into an e-value, and then apply e-closure.

MERGING P-VALUES FOR INTERSECTIONS

To apply closure principle, each intersection hypothesis H_S needs one p-value. For $S \in 2^{[m]}$, define the Simes merged p-value

$$Q_S^\sigma = \mathrm{Simes}(p_i^\sigma : i \in S)$$

If $p_{(1)}^\sigma \le \cdots \le p_{(|S|)}^\sigma$ are the ordered p-values in S, then

$$Q_S^\sigma = \min_{1 \le j \le |S|} \frac{|S|\, p_{(j)}^\sigma}{j}$$

The observed merged p-value is

$$Q_S^{\mathrm{id}} = \mathrm{Simes}(p_i^{\mathrm{id}} : i \in S)$$

GLOBAL PERMUTATION P-VALUES

For every $S \in 2^{[m]}$, define

$$p_S^\sigma = \frac{1 + \sum_{b=1}^{\omega} \mathbf{1}\left\{Q_S^{\sigma,b} \le Q_S^{\mathrm{id}}\right\}}{\omega + 1}.$$

This compares the observed merged statistic to its full permutation distribution.

> ***Theorem (Validity of permutation p-values)***
> *If D is exchangeable under the group of transformation G, then*
> $$p_S^\sigma$$
> *is a valid p-value for the intersection hypothesis H_S.*

FROM P-VALUES TO E-VALUES

For each $S \in 2^{[m]}$ use the Su p-to-e calibrator

$$e_S^\sigma = (l_\alpha p_S^\sigma \vee \alpha)^{-1} \tag{1}$$

where

$$l_\alpha = -W_{-1}(-\alpha/e),$$

and W_{-1} is the negative branch of the Lambert W function. Then

$$(e_S^\sigma)_{S \in 2^{[m]}}$$

is an an e-collection.

E-CLOSURE PRINCIPLE

Let $R \in 2^{[m]}$ be the rejected set and let $\mathcal{N} \in 2^{[m]}$ be the set of true nulls. The false discovery proportion is

$$\mathrm{FDP} = \frac{|R \cap \mathcal{N}|}{|R| \vee 1}$$

The e-closure framework uses an e-collection

$$\mathbf{e} = (e_S)_{S \in 2^{[m]}}$$

to define the set of candidate sets for the FDR error rate function at level α as

$$\mathcal{R}_\alpha^{FDR} := \left\{ R \in 2^{[m]} : \mathbf{e}_S \ge \frac{FDP_S(R)}{\alpha} \ \forall S \in 2^{[m]} \right\} \tag{2}$$

$\mathcal{R}$ is an FDR controlling procedure at level α if and only if there exists an e-collection $\mathbf{E}$ such that $\mathcal{R} \subseteq \mathcal{R}_\alpha^{FDR}(\mathbf{E})$.

MAIN RESULT

> ***Theorem***
> *The e-closure principle applied to the e-collection $\mathbf{e}_S^\sigma$ controls the FDR at level α.*

PROCEDURE SUMMARY

1. Compute observed and permuted p-values.
2. For each $S \in 2^{[m]}$, merge p-values using Simes:
$$Q_S^\sigma = \mathrm{Simes}(p_i^\sigma : i \in S).$$
3. Compute global permutation p-values:
$$p_S^\sigma = \frac{1 + \sum_b \mathbf{1}\{Q_S^{\sigma,b} \le Q_S^{\mathrm{id}}\}}{\omega + 1}.$$
4. Calibrate:
$$e_S^\sigma = (l_\alpha p_S^\sigma \vee \alpha)^{-1}.$$
5. Apply e-closure to obtain FDR control.

REFERENCES

[1] Z. Xu et al (2025).
Bringing closure to false discovery rate control: A general principle for multiple testing. arXiv preprint.

[2] Hemerik Jesse, and Goeman Jelle (2018).
Exact testing with random permutations.
Test, 27(4), 811-825, .

r.tsafacknzanguim@utwente.nl