

Inference on multiverse meta-analysis

A multivariate permutation testing approach

Filippo Gambarota

In collaboration with: Anna Vesely, Livio Finos and Gianmarco Altoè

Department of Developmental Psychology and Socialization

Contents I

Multiverse

Multiverse meta-analysis

PIMA on meta-analysis

A simulated example

Guidelines for multiverse meta-analysis

Future steps

Multiverse

Multiverse (Steegen et al., 2016)

- ▶ Real-world data analysis involve **several choices at each step**
- ▶ There are **many plausible alternatives** to the chosen analysis
- ▶ The **impact of alternatives** is often neglected or strongly underrated

Inference on multiverse

- ▶ The increase in complexity after taking into account scenarios is usually handled only descriptively
- ▶ The specification curve (Simonsohn et al., 2020) is the only inferential method but is limited to standard linear models

Note

There is a lack of a general and valid inferential framework for multiverse analysis



POST-SELECTION INFERENCE IN MULTIVERSE ANALYSIS (PIMA): AN INFERENCE FRAMEWORK BASED ON THE SIGN FLIPPING SCORE TEST

PAOLO GIRARDI¹

CA' FOSCARI UNIVERSITY OF VENICE

ANNA VESELY²

UNIVERSITY OF BOLOGNA

DANIEL LAKENS³

EINDHOVEN UNIVERSITY OF TECHNOLOGY

GIANMARCO ALTOÈ⁴ AND MASSIMILIANO PASTORE⁵

UNIVERSITY OF PADOVA

ANTONIO CALCAGNI⁶

UNIVERSITY OF PADOVA

GNCS-INDAM RESEARCH GROUP

LIVIO FINOS⁷

UNIVERSITY OF PADOVA

PIMA

- ▶ Use a multivariate extension of the sign-flip score test (Hemerik et al., 2020)
- ▶ Works on generalized linear models (and meta-analysis)
- ▶ Controls the family-wise error rate
- ▶ Provides an overall multiverse p-value and corrected p-values for each included scenario

Multiverse meta-analysis

Meta-analysis

We can define a (random-effects) meta-analysis model as:

$$y_i = \mu_\theta + \delta_i + \epsilon_i$$

$$\delta_i \sim \mathcal{N}(0, \tau^2)$$

$$\epsilon_i \sim \mathcal{N}(0, \sigma_{\epsilon_i}^2)$$

Where μ_θ is the average true effect, δ_i is the random-effect of the study i ($\theta_i = \mu_\theta + \delta_i$) and ϵ_i is the sampling error of the study i . When $\tau^2 = 0$ we have an equal-effects (or fixed-effect) model.

Meta-analysis many choices

Meta-analysis is a prototypical case of multiverse analysis:

- ▶ Should the study x be excluded for theoretical or statistical (e.g., outliers) reasons?
- ▶ Should we use an equal or random-effects model?
- ▶ Which value should take the pre-post missing correlation?
- ▶ ...

Multiverse meta-analysis example

We collected k randomized controlled trials on the effectiveness of short-term memory training. The authors used multiple memory measures that are safe to combine.

- ▶ pre-post correlations are missing
- ▶ correlations between outcomes are missing
- ▶ we have n outliers
- ▶ we have n papers that we are not sure to include
- ▶ we could run an equal or random-effects model
- ▶ ...

Multiverse meta-analysis example

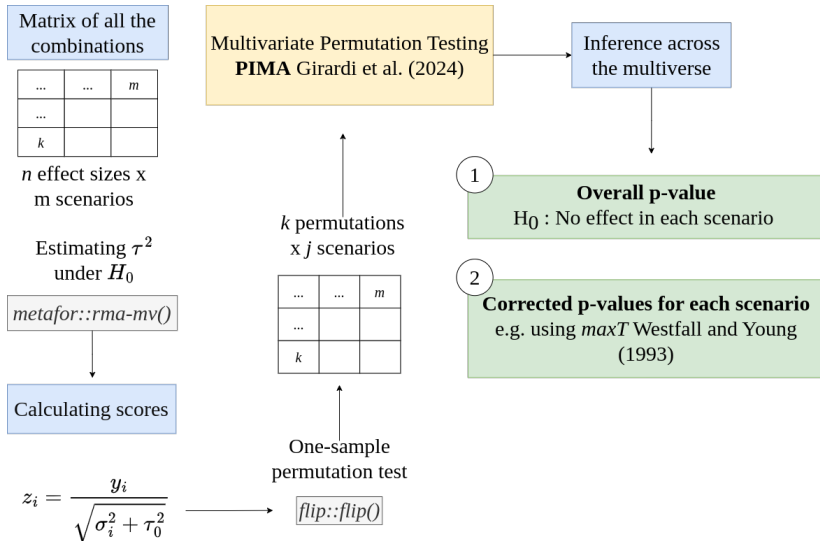
Even for the relatively simple previous example, we could have:

- ▶ pre-post $\rho \rightarrow 5$ values
- ▶ outcomes $\rho \rightarrow 5$ values
- ▶ 5 outliers
- ▶ 3 papers including/not including
- ▶ 2 equal vs random-effects

With all combinations, we have 750 **PLAUSIBLE** models with the same datasets.

PIMA on meta-analysis

General Workflow



Fast meta-analysis using permutations

Follmann & Proschan (1999) proposed a standard permutation test by re-computing τ^2 and μ_θ after each permutation.

We proposed a faster approach, especially for multiverse analysis with several scenarios.

1. Estimating τ^2 when $H_0 : \mu_\theta = 0$ maximizing the log-likelihood fixing $\mu_\theta = 0$
2. Calculating observed scores as $z_i = \frac{y_i}{\sqrt{v_i + \tau_0^2}}$
3. Doing B permutations flipping the sign of scores z_i
4. Repeat for each scenario of the multiverse

A simulated example

Simulating a multiverse analysis

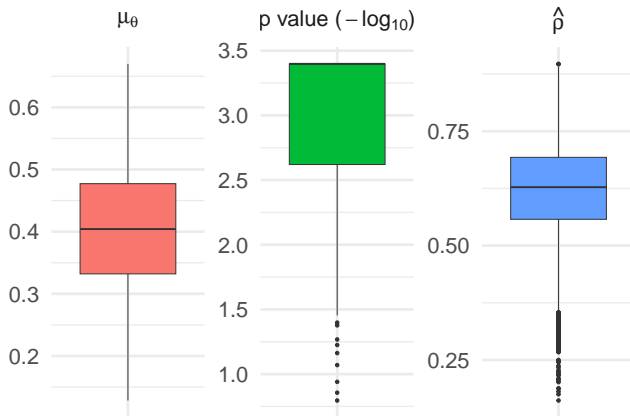
We simulated a meta-analysis with $k = 30$ and $m = 162$ sampling from a multivariate normal distribution.

We can describe a multiverse meta-analysis reporting summary statistics of the m estimated μ_θ and the correlation between scenarios.

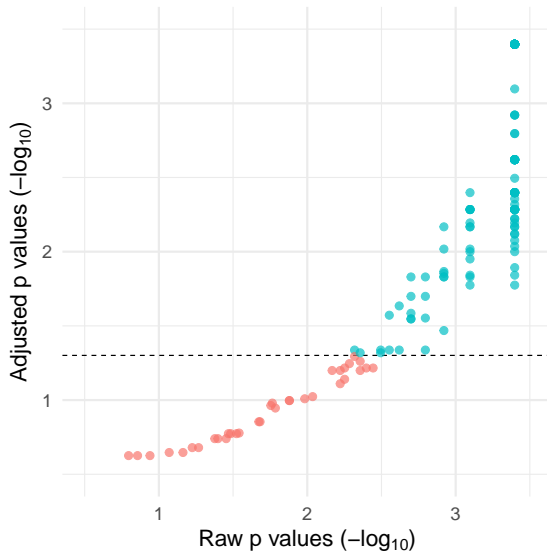
Overall inference

The multiverse is associated with an **overall p-value** < 0.001 , an **average effect of** 0.403 ($SD = 0.112$) and an **average correlation of** 0.620 ($SD = 0.102$)

Simulating a multiverse analysis



Impact of multiplicity correction



(valid) Post-hoc selective inference

Legal P-Hacking :)

After the overall test and p values correction, the survived scenarios (the blue dots) can be selectively commented, without inflating the type-1 error.

Guidelines for multiverse meta-analysis

Guidelines for multiverse meta-analysis

1. Multiverse meta-analysis must contain only **PLAUSIBLE** models. Including implausible models (e.g., assuming a pre-post correlation of 0.95) reduces the statistical power.
2. As with any other inferential test, multiverse analysis should be **PLANNED** otherwise no control of type-1 error.
3. Like in standard meta-analysis, the quality of the conclusions is related to the input data and the choice of multiverse scenarios.

Future steps

Future steps

- ▶ extending to multilevel and multivariate meta-analysis (the permutation approach is not straightforward)
- ▶ create an R package for multiverse meta-analyses with ad-hoc functions to analyze, report, and visualize the results
- ▶ create a data simulation framework for simulating a plausible multiverse for power and design analysis

References

- Follmann, D. A., & Proschan, M. A. (1999). Valid inference in random effects meta-analysis. *Biometrics*, 55(3), 732–737.
<https://doi.org/10.1111/j.0006-341x.1999.00732.x>
- Girardi, P., Vesely, A., Lakens, D., Altoè, G., Pastore, M., Calcagnì, A., & Finos, L. (2024). Post-selection inference in multiverse analysis (PIMA): An inferential framework based on the sign flipping score test. *Psychometrika*. <https://doi.org/10.1007/s11336-024-09973-6>
- Hemerik, J., Goeman, J. J., & Finos, L. (2020). Robust testing in generalized linear models by sign flipping score contributions. *J. R. Stat. Soc. Series B Stat. Methodol.*, 82(3), 841–864. <https://doi.org/10.1111/rssb.12369>
- Simonsohn, U., Simmons, J. P., & Nelson, L. D. (2020). Specification curve analysis. *Nat Hum Behav*, 4(11), 1208–1214.
<https://doi.org/10.1038/s41562-020-0912-z>
- Steenen, S., Tuerlinckx, F., Gelman, A., & Vanpaemel, W. (2016). Increasing transparency through a multiverse analysis. *Perspect. Psychol. Sci.*, 11(5), 702–712. <https://doi.org/10.1177/1745691616658637>

filippo.gambarota@unipd.it

Download the slides:

