Efficient variable importance testing in predictive models through permutations

Eduardo García-Portugués • Department of Statistics, UC3M

Variable importance measures (VIMPs) quantify the relevance of predictors in black-box prediction models, supporting interpretation and variable selection. A widely used example is permutation VIMP (pVIMP), which evaluates the change in predictive performance when a variable is permuted, thereby breaking its relation to the outcome. While conceptually simple and broadly applicable, inference on VIMPs remains challenging. Analytical tests often rely on restrictive assumptions and can be computationally infeasible. Permutation tests are attractive since they are distribution-free and model-agnostic, but their computational cost can be prohibitive because they require many model refittings. To address this, Hapfelmeier et al. (2023) propose sequential permutation tests and sequential Monte Carlo p-value estimation (Besag & Clifford, 1991; Lock, 1991; Wald, 1945), which preserve type-I error control while substantially reducing the number of permutations needed by exploiting early termination rules. Although random forests and their pVIMP serve as the main application example, the framework applies more generally to other prediction models and variable importance measures. The R package `rfvimptest` (Hapfelmeier & Hornung, 2025) implements these sequential methods for random forests.

The thesis will: (1) review the theoretical foundations and methodology of sequential permutation testing for VIMPs as developed in Hapfelmeier et al. (2023); (2) conduct numerical experiments to compare sequential tests with standard permutation tests in terms of statistical accuracy and computational efficiency for random forests; and (3) explore the potential applicability of sequential testing to other black-box predictive models, such as gradient boosting or neural networks.

References

Besag, J., & Clifford, P. (1991). Sequential monte carlo p-values. *Biometrika*, *78*(2), 301–304. https://doi.org/10.1093/biomet/78.2.301

Hapfelmeier, A., & Hornung, R. (2025). *Rfvimptest: Sequential permutation testing of random forest variable importance measures*. https://CRAN.R-project.org/package=rfvimptest

Hapfelmeier, A., Hornung, R., & Haller, B. (2023). Efficient permutation testing of variable importance measures by the example of random forests. *Computational Statistics & Data Analysis*, *181*, 107689. https://doi.org/10.1016/j.csda.2022.107689

Lock, R. H. (1991). A sequential approximation to a permutation test. *Communications in Statistics — Simulation and Computation*, *20*(2-3), 341–363. https://doi.org/10.1080/03610919108812984

Wald, A. (1945). Sequential tests of statistical hypotheses. *Annals of Mathematical Statistics*, *16*(2), 117–186. https://doi.org/10.1214/aoms/1177731118