

# Permutations frequently don't work

## Nulls are not unique, some structure can't be permuted away

Behavioral Ecology and Sociobiology (2022) 76:151  
<https://doi.org/10.1007/s00265-022-03254-x>

### METHODS



## Common permutation methods in animal social network analysis do not control for non-independence

Jordan D. A. Hart<sup>1</sup> · Michael N. Weiss<sup>1,2</sup> · Lauren J. N. Brent<sup>1</sup> · Daniel W. Franks<sup>3</sup>

Received: 23 September 2022 / Revised: 4 October 2022 / Accepted: 10 October 2022 / Published online: 29 October 2022  
© The Author(s) 2022

### Abstract

The non-independence of social network data is a cause for concern among behavioural ecologists conducting social network analysis. This has led to the adoption of several permutation-based methods for testing common hypotheses. One of the most common types of analysis is nodal regression, where the relationships between node-level network metrics and nodal covariates are analysed using a permutation technique known as node-label permutations. We show that, contrary to accepted wisdom, node-label permutations do not automatically account for the non-independences assumed to exist in network data, because regression-based permutation tests still assume exchangeability of residuals. The same assumption also applies to the quadratic assignment procedure (QAP), a permutation-based method often used for conducting dyadic regression. We highlight that node-label permutations produce the same  $p$ -values as equivalent parametric regression models, but that in the presence of non-independence, parametric regression models can also produce accurate effect size estimates. We also note that QAP only controls for a specific type of non-independence between edges that are connected to the same nodes, and that appropriate parametric regression models are also able to account for this type of non-independence. Based on this, we suggest that standard parametric models could be used in the place of permutation-based methods. Moving away from permutation-based methods could have several benefits, including reducing over-reliance on  $p$ -values, generating more reliable effect size estimates, and facilitating the adoption of causal inference methods and alternative types of statistical analysis.

**Keywords** Animal social network analysis · Mixed models · Node-label permutations · Permutation tests

### BRIEF COMMUNICATION

doi:10.1111/j.1558-5646.2010.00973.x

## POOR STATISTICAL PERFORMANCE OF THE MANTEL TEST IN PHYLOGENETIC COMPARATIVE ANALYSES

Luke J. Harmon<sup>1,2</sup> and Richard E. Glor<sup>3</sup>

<sup>1</sup>Department of Biological Sciences, University of Idaho, Moscow, Idaho, 83844

<sup>2</sup>E-mail: [lukeh@uidaho.edu](mailto:lukeh@uidaho.edu)

<sup>3</sup>Department of Biology, RC Box 270211, University of Rochester, Rochester, New York, 14627

Received May 21, 2008

Accepted January 21, 2010

The Mantel test, based on comparisons of distance matrices, is commonly employed in comparative biology, but its statistical properties in this context are unknown. Here, we evaluate the performance of the Mantel test for two applications in comparative biology: testing for phylogenetic signal, and testing for an evolutionary correlation between two characters. We find that the Mantel test has poor performance compared to alternative methods, including low power and, under some circumstances, inflated type-I error. We identify a remedy for the inflated type-I error of three-way Mantel tests using phylogenetic permutations; however, this test still has considerably lower power than independent contrasts. We recommend that use of the Mantel test should be restricted to cases in which data can only be expressed as pairwise distances among taxa.

**KEYWORDS:** Comparative methods, independent contrasts, phylogenetic signal, statistical power, type-I error.



# Bayes is a practical choice

Also, it's just probability theory...

- Using Bayesian methods is simply easier in real life:
  - Simple models are the same, so why bother?
  - Because complicated models are possible:
    - Real data is messy: missing data, replicates, correlated observations, mark-recapture...
- Models are **generative**, easy to simulate from and easy(er?) to build using scientific knowledge

## Probability Theory

The Logic of Science

E. T. JAYNES

CAMBRIDGE