The Journal of Open Source Software

{discord}: An R Package for Discordant-Kinship Regressions

Jonathan D. Trattner[1] **and S. Mason Garrison**[2]

1 Department of Neuroscience, Wake Forest School of Medicine **2** Department of Psychology, Wake Forest University

DOI:

Software

- Review
- Repository
- Archive

Submitted:
Published:

Summary

As a field, behavior genetics studies the genetic and environmental sources of individual differences in psychological traits and characteristics. More technically, the field focuses on decomposing the sources of phenotypic variation into genetic (Additive (A)+ Dominance (D)) and environmental (Shared Environment (C) + Non-Shared Environment (E)) variance components, by leveraging twin and family studies. These models can do more than merely describe sources of variance; they can be used to infer causation (Burt, Plaisance, & Hambrick, 2019). Here, we present software to facilitate genetically-informed quasi-experimental designs primarily for kinship modeling. Specifically, it facilitates discordant-kinship regressions by comparing kin, such as siblings. These designs account for genetic-and-environmental variance when examining causal links in the realm of 'nature vs. nurture.'

Statement of Need

Kin-comparison designs distinguish "within-family variance" from "between-family variance" (Chamberlain & Griliches, 1975). Within-family variance indicates how individuals of a specific family differ from one another; the between-family variance reflects sources that make family members more similar to one another (**garrison2016?**). By partitioning these sources of variance, scholars may greatly reduce confounds when testing hypotheses (Lahey & D'Onofrio, 2010). Our R package, {discord}, has customizable, efficient code for generating genetically-informed simulations and provides user-friendly functions to help researchers use kin-based quasi-experimental designs.

{discord} augments the NlsyLinks R package, which provides kinship links for the National Longitudinal Surveys of Youth – a series of cross-generational, nationally representative surveys of over 30,000 participants (Beasley et al., 2016; Rodgers et al., 2016). It has been used in thousands of studies [CITE this database https://nlsinfo.org/bibliography-start]

Mathematics

To facilitate kinship comparisons, {discord} implements a modified reciprocal standard dyad model (Kenny, Kashy, & Cook, 2006) known as the discordant-kinship model (see (**garrison2016?**) for an extension). Consider the simplified case where a behavioral outcome, Y, is predicted by variable, X. The discordant-kinship model relates the difference in the outcome, $Y_{i\Delta}$, for the ith kinship pair, where $\bar{Y}_i$ is the mean level of the outcome, $\bar{X}_i$ is the mean level of the predictor, and $X_{i\Delta}$ is the between-kin difference in the predictor.

$$Y_{i\Delta} = \beta_0 + \beta_1 \bar{Y}_i + \beta_2 \bar{X}_i + \beta_3 X_{i\Delta} + \epsilon_i$$

This model partitions variance in line with the above discussion to support causal inference. Specifically, the within-family variance is described by Y_Δ and X_Δ; between-family variance is captured by $\bar{Y}$ and $\bar{X}$ (Garrison & Rodgers, 2021).

A non-significant association between Y_Δ and X_Δ suggests that the relationship is not directly causal. Specifically, it means that kin differences in the outcome, after controlling for gene and shared-environmental factors, are not associated with kin differences in the outcome. In contrast, a significant association may provide support for a causal relationship between variables depending on the relatedness of each kin pair. That is, the discordant-kinship model is applicable for any set of kin: monozygotic twins who share 100% of their DNA; full-siblings who share 50%; half-siblings who share 25%; cousins who share 12.5%; etc. Thus, a significant relationship found with monozygotic twins would provide more compelling support for a causal claim than the same relationship between cousins.

Following (Garrison & Rodgers, 2021), we recommend interpreting significant associations as *not disproving a causal relationship*. Although this design controls for much (sibling) if not all (monozygotic twins) background heterogeneity, it is possible that a significant relationship between a phenotype and plausible covariates is possible due to non-shared environmental influences.

The next section illustrates one example of discordant-kinship regressions with the {discord} package.

Acknowledgments

We acknowledge contributions from Cermet Ream, Joe Rodgers, and support from Lucy D'Agostino McGowan on this project.

References

Beasley, W., Rodgers, J., Bard, D., Hunter, M., Garrison, S. M., & Meredith, K. (2016). *NlsyLinks: Utilities and kinship information for research with the NLSY*. Retrieved from https://CRAN.R-project.org/package=NlsyLinks

Burt, S. A., Plaisance, K. S., & Hambrick, D. Z. (2019). Understanding "What Could Be": A Call for 'Experimental Behavioral Genetics'. *Behavior Genetics*, *49*(2), 235–243. doi:10.1007/s10519-018-9918-y

Chamberlain, G., & Griliches, Z. (1975). Unobservables with a Variance-Components Structure: Ability, Schooling, and the Economic Success of Brothers. *International Economic Review*, *16*(2), 422–449.

Garrison, S. M., & Rodgers, J. L. (2021, June 28). *Using genetically-informed designs to test causal claims without experiments: Discordant-sibling designs and applications for differential psychology.*

Kenny, D. A., Kashy, D. A., & Cook, W. L. (2006). *Dyadic data analysis*. Dyadic data analysis. New York, NY, US: Guilford Press.

Lahey, B. B., & D'Onofrio, B. M. (2010). All in the Family: Comparing Siblings to Test Causal Hypotheses Regarding Environmental Influences on Behavior. *Current Directions in Psychological Science*, *19*(5), 319–323. doi:10.1177/0963721410383977

Rodgers, J. L., Beasley, W. H., Bard, D. E., Meredith, K. M., Hunter, M. D., Johnson, A. B., Buster, M., et al. (2016). The NLSY kinship links: Using the NLSY79 and NLSY-children data to conduct genetically-informed and family-oriented research. *Behavior genetics*, *46*(4), 538–551. doi:10.1007/s10519-016-9785-3