

A negative result on shrinkage estimators in small-N replication

Example Author A

2026-05-13

Demonstration paper in the rrxiv reference corpus. The canonical machine-readable version lives at rrxiv.com/papers/rrxiv:2605.00004.

Abstract

We attempt to replicate the shrinkage-estimator results of [Brown 2021] under small-N conditions and report a negative finding. Our analysis identifies two likely sources of the original effect: data leakage in the holdout split and a misspecified prior. We release replication code and a corrected reanalysis.

1 Introduction

We attempt to replicate the shrinkage-estimator results of [Brown 2021] under small-N conditions and report a negative finding. Our analysis identifies two likely sources of the original effect: data leakage in the holdout split and a misspecified prior. We release replication code and a corrected reanalysis.

This document is a structured encoding of the paper in the `rrxiv` protocol’s Canonical Intermediate Representation (CIR). It engages with the topic `stat.ME`. The encoding registers 2 formal claims (2 untested). Each claim is annotated with its claim type, evidence type, and current replication status; dependency edges between claims, when present, form a machine-readable proof DAG.

2 Methodology

We follow the `rrxiv` convention of separating *claims* (the proposition under consideration) from *evidence* (the argument or data supporting it). Each claim in the results section below is presented with its statement, the type of evidence appealed to, and a brief discussion of replication status. Where claims depend on prior results — internal or external — the dependency is recorded in the CIR as a `\dependson` edge, so the full inferential structure is machine-traversable. Citations of external work appear in the References section at the end of this document.

3 Results: registered claims

Claim 1

Claim 1 (Claim 1). Brown 2021’s reported shrinkage gains do not survive a leak-free reanalysis.
Replication status: untested.

This claim is an empirical observation supported by data. As of the encoding date, it has not yet been independently tested.

Claim 2

Claim 2 (Claim 2). The misspecified prior accounts for ~60% of the reported effect size.

Replication status: untested.

This claim is an empirical observation supported by data. As of the encoding date, it has not yet been independently tested.

4 Discussion

The claim graph above is the primary product of this paper. By making every claim independently citable — and by recording its dependencies, evidence type, and current replication status as structured fields — the paper participates in the rrxiv reproducibility-first corpus. Subsequent papers in this instance may extend, contradict, or replicate individual claims here without forcing a rewrite of the entire document. See the canonical version online for the live discourse layer.

5 References

- Hierarchical shrinkage for meta-analysis