

Metodologi for replikasjonsprosjekt

Ole Fredrik Borgundvåg Berg
NTNU

Dette er et replikasjonsprosjekt som er basert på et tilsvarende prosjekt som ble gjort av Ulrich Schimmack, der man prøvde å finne i hvilken grad vitenskaplige funn ved ulike amerikanske universiteter kunne replikeres, inkludert estimert replikasjonsrate på enkeltforskere (Schimmack, 2022). Siden faktisk replikasjon er dyrt, så laget heller han med flere en modell som finner estimert replikasjonsrate basert på p-verdier i artiklene til forskerene, som er implementert i biblioteket «z-curve» til R (Bartoš & Schimmack, 2022; Brunner & Schimmack, 2020).

Finne forskere og artikler

For å finne artikler og forskere ble Cristin APIet brukt (“Cristin API | Cristin REST API”, udatert). APIet gjør det enkelt både å finne ut hvilke forskere som er tilknyttet hvilke institutt og hvilke artikler som er skrevet av hvilke forskere. Artiklene som er brukt i dette prosjektet er de publikasjonene som har `category.code = "ARTICLE"` i APIet. Selve nedlastingen av artiklene skjedde også ved programmerisk, men ikke ved hjelp av Cristin. Dette kan føre til at noen av artiklene i Cristin ikke kom med i dette prosjektet pga. manglende tilgang.

Ekstrahering av z-verdier

Siden p-verdier ofte er på formen $p < 0.05$ er det vanskelig å lage et histogram ut fra disse. Derfor brukes kun statistiske tester der man har test-statistikken tilgjengelig. Deretter konverteres den til en z-verdi med tilsvarende p-verdi. Selve ekstraheringen av de statistiske testene ble gjort ved hjelp av verktøyet JATSdecoder og funksjonen `get.stats()` (Böschén, 2021). Skal sies at z-curve også har denne funksjonaliteten, men ved egen testing og i følge Böschén (2021) sine forsøk så vil JATSdecoder hente ut en større andel av de statistiske testene i artiklene.

Modellering av replikasjonsrate og visualisering

For å finne forventet replikasjonsrate og for å lage grafer for visualisering ble de statistiske testene som ble hentet ut ved hjelp av JATSdecoder puttet inn i z-curve. Den både regnet ut forventet replikasjonsrate og laget graf over de tilsvarende z-verdiene for de ulike statistikkene i artiklene. For mer informasjon om hvordan z-curve fungerer, se Brunner og Schimmack (2020) og Bartoš og Schimmack (2022).

Referanser

- Bartoš, F., & Schimmack, U. (2022). Z-Curve 2.0: Estimating Replication Rates and Discovery Rates. *Meta-Psychology*, 6. <https://doi.org/10.15626/MP.2021.2720>
- Brunner, J., & Schimmack, U. (2020). Estimating Population Mean Power Under Conditions of Heterogeneity and Selection for Significance. *Meta-Psychology*, 4. <https://doi.org/10.15626/MP.2018.874>
- Böschén, I. (2021). Evaluation of JATSdecoder as an automated text extraction tool for statistical results in scientific reports. *Scientific Reports*, 11(1), 19525. <https://doi.org/10.1038/s41598-021-98782-3>
- Cristin API | Cristin REST API. (udatert). Hentet 22. januar 2023, fra <https://api.cristin.no/>
- Schimmack, U. (2022). Replicability Rankings of Psychology Departments. Hentet 8. januar 2023, fra <https://replicationindex.com/2022/03/15/rr22-psy-dept/>