wildmeta: Cluster Wild Bootstrapping for Meta-Analysis

ESMAR 2022

Megha Joshi & James E. Pustejovsky

2022-02-13

wildmeta

Dependence in Meta-Analysis

- Typical meta-analytic techniques (like meta-regression) involves the assumption that effect sizes are independent
- However, common for each primary study to yield more than one effect size or studies to be nested in some way creating dependence
- Example: Tanner-Smith and Lipsey (2015) meta-analysis of the effects of brief alcohol interventions
    - 185 studies, 1446 effect sizes
    - Multiple correlated outcome measures: e.g., alcohol consumption measured by frequency of consumption, quantity consumed, blood alcohol concentration
    - Repeated measures
    - Multiple comparison groups

Handling Dependence

- Ignore dependence
    - Incorrect standard errors, incorrect inference from hypothesis tests
- Ad-hoc methods
    - Selecting one effect per study
    - Analysing subsets of data separately
    - Loss of information
- Standard multivariate methods - ideal
    - Require info on covariance between effect sizes
    - Primary studies often don't report

Robust Variance Estimation

- Robust variance estimation (CR0-type CRVE) (Hedges, Tipton, and Johnson, 2010)
    - Only works well when number of studies is large (> 40, Hedges Tipton, Johnson, 2010)
    - Meta-analysis in social science research typically have smaller number of studies
    - Small number of studies - CRVE - Type 1 error inflation - meta-analysts can conclude some effect is present when it is actually not
    - Tipton (2015) and Tipton and Pustejovsky (2015) examined small sample corrections - HTZ test - CR2 + Satterthwaite for single-coefficient tests and multiple-contrast hypothesis tests
- HTZ controls Type 1 error rates adequately but possibly low power especially for multiple-contrast hypothesis tests (Tipton and Pustejovsky, 2015)

Cluster Wild Bootstrapping (CWB)

- Alternative method - examined in the econometrics literature - not in meta-analytic framework
- Bootstrapping - estimate unknown quantities by re-sampling from original data many times (Boos et al., 2013)
- CWB - re-sampling residuals by multiplying them by cluster-level random weights (Cameron, Gelbach, and Miller 2008)

CWB Algorithm

1. Fit a null model and a full model on the original data
2. Obtain residuals from the null model
3. Generate an auxiliary random variable that has mean of 0 and variance of 1 and multiply the residuals by the random variable (e.g., Rademacher weights) set to be constant within clusters (CWB)
    - Can also multiply the residuals by CR2 matrices before multiplying by weights (CWB Adjusted)
4. Obtain new outcome scores by adding the transformed residuals to the predicted values from the null model fit on the original data
5. Re-estimate the full model with the new calculated outcome scores and obtain the test statistic
6. Repeat steps 3-5 R times. Calculate p-value:

$$p = \frac{1}{R}\sum_{r=1}^{R} I\left(F^{(r)} > F\right)$$

Simulation

- Ran two simulations
- Compared CWB against the HTZ test in terms of Type 1 error rates and power
- CWB maintained Type 1 error rates adequately and provided more power than HTZ

Recommendations

- Dependent effect sizes - common in meta-analyses in social sciences
- If we use RVE for meta-analyses with small number of studies - Type 1 error inflation - false discovery rate high
- If we use the HTZ test - low power - may miss effects that are present
- We recommend use of CWB - balances Type 1 error rates and also provides more power than existing corrections

wildmeta

- The main function in the package is `Wald_test_cwb()`
- Works with meta-regressions models fit using `robumeta::robu()` and `metafor::rma.mv()`

```
Wald_test_cwb(
  full_model,
  constraints,
  R,
  cluster = NULL,
  auxiliary_dist = "Rademacher",
  adjust = "CR0",
  type = "CR0",
  test = "Naive-F",
  seed = NULL
)
```

Data SATcoaching

Show 10 entries Search:

	study	d	V	study_type	hrs	test
1	Burke (A)	0.5	0.0825	Matched	50	Verbal
2	Burke (B)	0.74	0.0855	Matched	50	Verbal
3	Coffin	0.33	0.2534	Matched	18	Math
4	Coffin	-0.23	0.2517	Matched	18	Verbal
5	Davis	0.13	0.0933	Matched	15	Math
6	Davis	0.13	0.0933	Matched	15	Verbal
7	Frankel	0.34	0.0451	Matched	30	Math
8	Frankel	0.13	0.0445	Matched	30	Verbal
9	Kintisch	0.06	0.0527	Matched	20	Verbal
10	Whitla	-0.11	0.0401	Matched	10	Math

Showing 1 to 10 of 67 entries

robumeta Model

```r
library(wildmeta)
library(clubSandwich)
library(robumeta)

robu_model <- robu(d ~ 0 + study_type + hrs + test,
                   studynum = study,
                   var.eff.size = V,
                   small = FALSE,
                   data = SATcoaching)
```

```r
robu_model
```

```
## RVE: Correlated Effects Model
##
## Model: d ~ 0 + study_type + hrs + test
##
## Number of studies = 46
## Number of outcomes = 65 (min = 1 , mean = 1.41 , median = 1 , max = 2 )
## Rho = 0.8
## I.sq = 0
## Tau.sq = 0
##
##                             Estimate  StdErr t-value dfs P(|t|>)  95% CI.L
## 1       study_typeMatched  0.11888 0.08480   1.402  41   0.168 -0.052376
## 2 study_typeNonequivalent  0.02866 0.05630   0.509  41   0.613 -0.085046
## 3    study_typeRandomized  0.08229 0.05008   1.643  41   0.108 -0.018842
## 4                     hrs  0.00394 0.00173   2.278  41   0.028  0.000447
## 5             testVerbal -0.01111 0.04727  -0.235  41   0.815 -0.106577
##   95% CI.U Sig
## 1  0.29013
## 2  0.14237
## 3  0.18342
## 4  0.00744  **
## 5  0.08436
## ---
## Signif. codes: < .01 *** < .05 ** < .10 *
## ---
```

robumeta

```
robu_res <- Wald_test_cwb(full_model = robu_model,
                          constraints = constrain_equal(1:3),
                          R = 999,
                          seed = 20220209)

robu_res
```

```
##   Test Adjustment CR_type Statistic   R     p_val
## 1  CWB        CR0     CR0   Naive-F 999 0.3343343
```

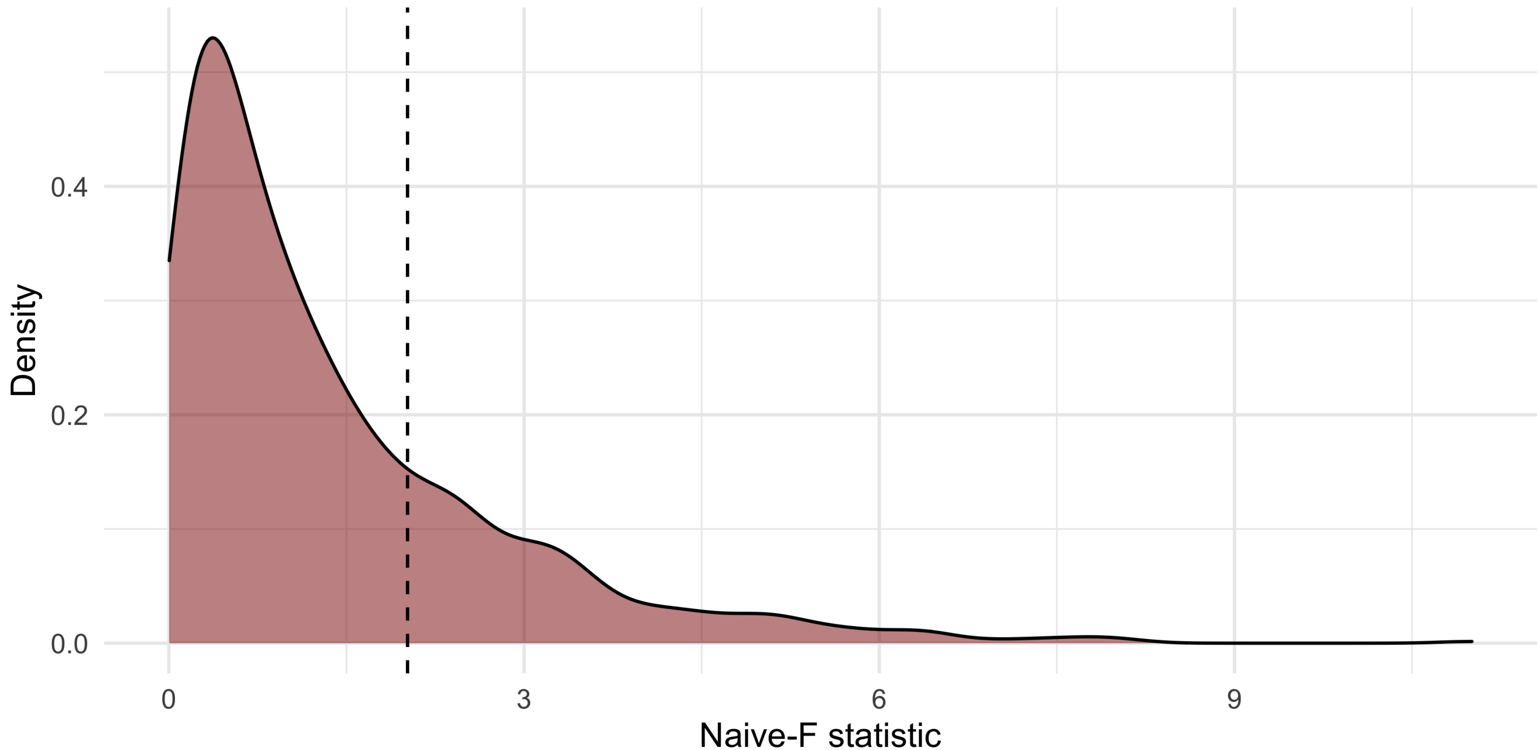
metafor

```r
library(metafor)

rma_model <- rma.mv(yi = d ~ 0 + study_type + hrs + test,
                    V = V,
                    random = ~ study_type | study,
                    data = SATcoaching)

rma_res <- Wald_test_cwb(full_model = rma_model,
                         constraints = constrain_equal(1:3),
                         R = 999,
                         seed = 20210314)

rma_res
```

```
##   Test Adjustment CR_type Statistic   R    p_val
## 1  CWB        CR0     CR0   Naive-F 999 0.242485
```

Bootstrap Distribution Plot

```
plot(rma_res,
     fill = "darkred",
     alpha = 0.5)
```



THANK YOU!

References

Boos, D. D., & others. (2003). Introduction to the bootstrap world. Statistical Science, 18(2), 168–174.

Cameron, A. C., Gelbach, J. B., & Miller, D. L. (2008). Bootstrap-Based Improvements for Inference with Clustered Errors. The Review of Economics and Statistics, 47.

Garrett, R., Citkowicz, M., & Williams, R. (2019). How responsive is a teacher's classroom practice to intervention? A meta-analysis of randomized field studies. Review of research in education, 43(1), 106-137.

Hedges, L. V., Tipton, E., & Johnson, M. C. (2010). Robust variance estimation in meta-regression with dependent effect size estimates. Research Synthesis Methods, 1(1), 39–65.

References

Joshi, M., Pustejovsky, J. E., & Beretvas, S. N. (2022). Cluster wild bootstrapping to handle dependent effect sizes in meta-analysis with small number of studies. Research Synthesis Methods. (Accepted).

Tipton, E., & Pustejovsky, J. E. (2015). Small-Sample Adjustments for Tests of Moderators and Model Fit Using Robust Variance Estimation in Meta-Regression. Journal of Educational and Behavioral Statistics, 40 (6), 604–634.

Tipton, E. (2015). Small sample adjustments for robust variance estimation with meta-regression. Psychological Methods, 20(3), 375–393.