Median-unbiased estimators for the amount of heterogeneity in meta-analysis

European Association of Methodology
9th European Congress of Methodology

Wolfgang Viechtbauer
Maastricht University
2021-07-21

Meta-Analysis and the Random-Effects Model

- collect studies that have examined a phenomenon of interest
- quantify the results of each study in terms of an effect size / outcome measure (e.g., standardized mean difference, correlation coefficient, log odds/risk ratio)
- let y_i denote the observed outcome in the ith study ($i = 1, \ldots, k$) and v_i the corresponding sampling variance
- random-effects model:

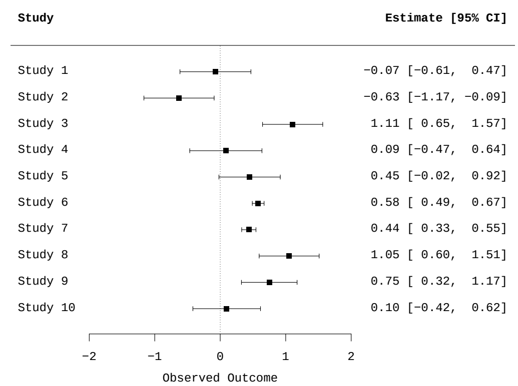
$$y_i = \theta_i + e_i$$

where $\theta_i \sim N(\mu, \tau^2)$ and $e_i \sim N(0, v_i)$
- τ^2 denotes the amount of 'heterogeneity' (i.e., between-study variance) in the underlying true effects/outcomes

Illustrative Example

study	y_i	v_i
1	-0.073	0.0764
2	-0.629	0.0749
3	1.106	0.0550
4	0.086	0.0792
5	0.449	0.0571
6	0.580	0.0021
7	0.437	0.0031
8	1.053	0.0546
9	0.749	0.0470
10	0.099	0.0694

Illustrative Example



Heterogeneity Estimators

- there has been a secret competition in the meta-analytic community to derive ever more estimators for τ^2
  - Hedges/Cochran estimator ([1], same as [2, 3])
  - Hunter-Schmidt estimator ([4, 5])
  - DerSimonian-Laird estimator ([6])
  - maximum-likelihood estimator ([6–8])
  - restricted maximum-likelihood estimator ([6, 9, 10])
  - Paule-Mandel estimator ([11])
  - empirical Bayes estimator ([12, 13])
  - Hartung-Makambi estimator ([14])
  - Sidik-Jonkman estimator ([15])
  - generalized Q-statistic estimator ([16])
  - various Bayesian estimators (e.g., [17–20])
  - ...

DerSimonian-Laird estimator ([6])

- let

$$Q = \sum w_i (y_i - \hat{\theta})^2$$

where

$$\hat{\theta} = \frac{\sum w_i y_i}{\sum w_i} \quad \text{and} \quad w_i = 1/v_i$$

- can show $\text{E}[Q] = c\tau^2 + (k-1)$ where $c = \sum w_i - \frac{\sum w_i^2}{\sum w_i}$
- hence

$$\hat{\tau}^2_{\text{DL}} = \frac{Q - (k-1)}{c}$$

is a method-of-moments estimator of τ^2

Illustrative Example

- $\hat{\theta} = 0.511$
- $Q = 43.665$
- $k = 10$
- $c = 568.562$

$$\hat{\tau}^2_{\text{DL}} = \frac{43.665 - (10 - 1)}{568.562} = 0.061$$

Restricted Maximum-Likelihood Estimator ([6, 9, 10])

- the restricted log likelihood function is given by

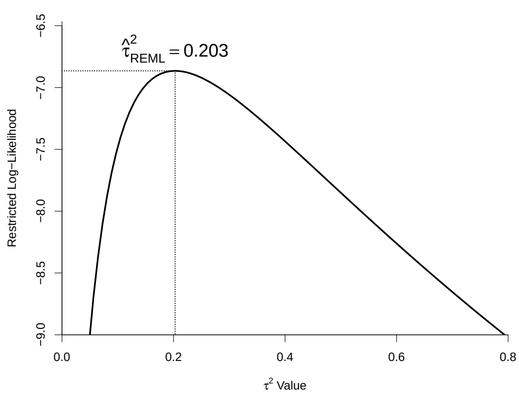
$$ll(\tau^2) = -\frac{1}{2}\sum \ln w_i^{-1} - \frac{1}{2}\ln \sum w_i - \frac{1}{2}\sum w_i (y_i - \hat{\mu})^2$$

where

$$\hat{\mu} = \frac{\sum w_i y_i}{\sum w_i} \quad \text{and} \quad w_i = 1/(v_i + \tau^2)$$

- can easily maximize this over τ^2 with standard algorithms (e.g., Fisher scoring) yielding $\hat{\tau}^2_{\text{REML}}$

Illustrative Example



Paule-Mandel estimator ([11])

- define

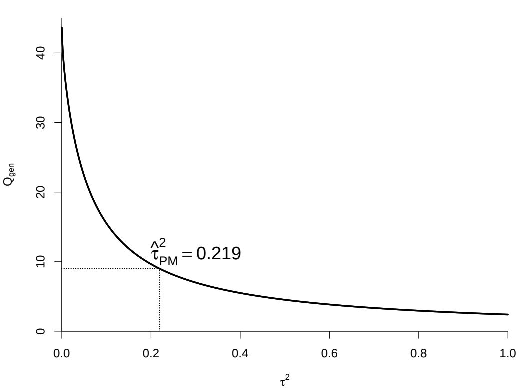
$$Q_{gen} = \sum w_i (y_i - \hat{\mu})^2$$

where

$$\hat{\mu} = \frac{\sum w_i y_i}{\sum w_i} \quad \text{and} \quad w_i = 1/(v_i + \tau^2)$$

- can show $Q_{gen} \sim \chi^2$ with $\text{df} = k - 1$
- hence $\text{E}[Q_{gen}] = k - 1$
- and Q_{gen} is a decreasing function of τ^2
- let $\hat{\tau}^2_{\text{PM}}$ be that value such that $Q_{gen} = k - 1$

Illustrative Example



Confidence Intervals for τ^2

- various methods for constructing confidence intervals for τ^2 have also been proposed
  - Wald-type CIs ([21, 22])
  - profile likelihood CIs ([8, 22])
  - Biggerstaff-Tweedie method ([21])
  - Hartung-Makambi method ([14])
  - Sidik-Jonkman method ([15])
  - bootstrap CIs ([22–24])
  - Q-profile method ([22])
  - generalized Q-statistic CIs ([25])
  - Bayesian credible intervals

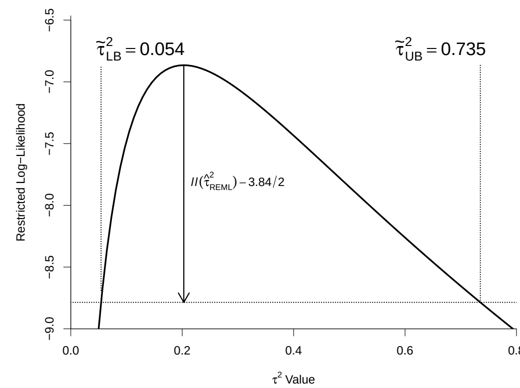
Profile Likelihood CIs ([8, 22])

- a 95% profile likelihood CI for τ^2 is given by the set of $\tilde{\tau}^2$ values satisfying

$$ll(\tilde{\tau}^2) > ll(\hat{\tau}^2_{\text{REML}}) - 3.84/2$$

- all values of $\tilde{\tau}^2$ not rejected by a likelihood ratio test
- bounds can be found via a root finding algorithm

Illustrative Example



Q-Profile Method [22]

- define

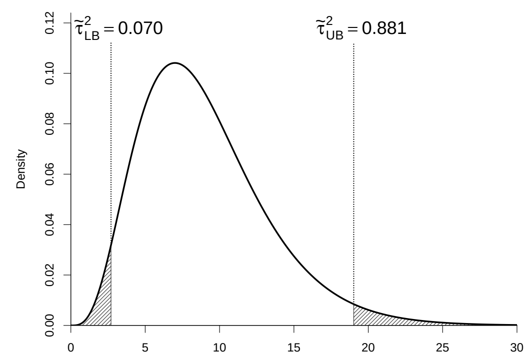
$$Q_{gen} = \sum w_i (y_i - \hat{\mu})^2$$

where

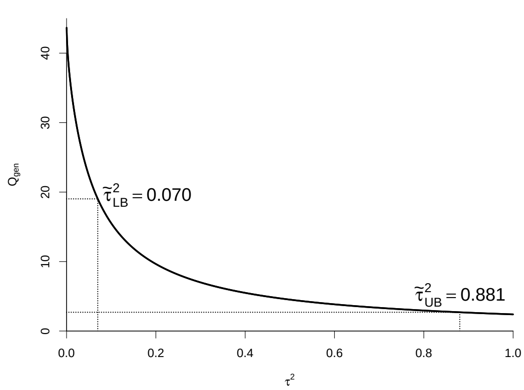
$$\hat{\mu} = \frac{\sum w_i y_i}{\sum w_i} \text{ and } w_i = 1/(v_i + \tau^2)$$

- can show $Q_{gen} \sim \chi^2$ with $\text{df} = k - 1$
- find $\tilde{\tau}^2_{\text{LB}}$ such that $Q_{gen} = \chi^2_{k-1;.975}$
- find $\tilde{\tau}^2_{\text{UB}}$ such that $Q_{gen} = \chi^2_{k-1;.025}$
- then $(\tilde{\tau}^2_{\text{LB}}, \tilde{\tau}^2_{\text{UB}})$ is a 95% CI for τ^2
- the CI is exact (under the assumptions of the model)

Illustrative Example



Illustrative Example



Illustrative Example

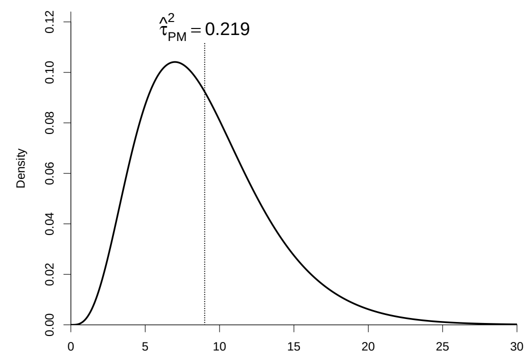
Estimator	Value	CI	Bounds
DL	0.061		
REML	0.203	PL	0.054 to 0.735
PM	0.219		
		QP	0.070 to 0.881

When the CI Does not Contain τ^2

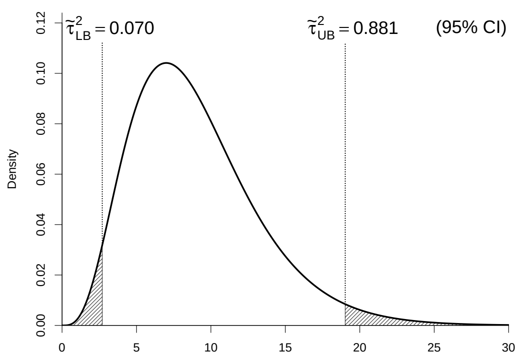
- can happen when the method for constructing the CI does not 'match up' with the method for estimating τ^2
- ML/REML estimation matches up with profile likelihood CIs
- clearly, the DL estimator does not match up with the Q-profile method (but what does?)
- does the PM estimator match up with the Q-profile method?

Estimator	Value	CI	Bounds
DL	0.061		
REML	0.203	PL	0.054 to 0.735
PM	0.219		
		QP	0.070 to 0.881

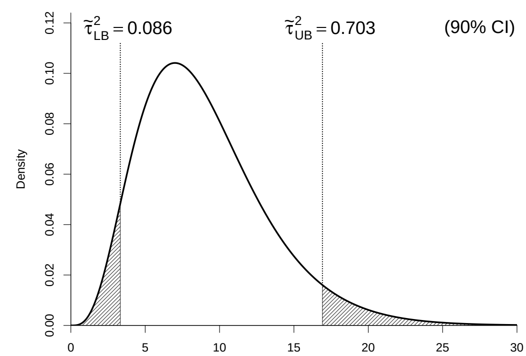
Illustrative Example



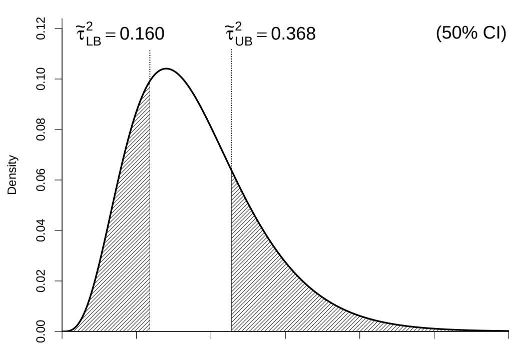
Illustrative Example



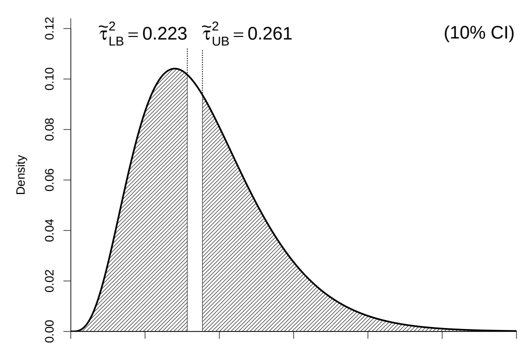
Illustrative Example



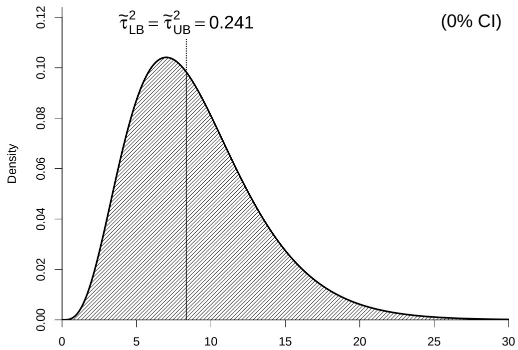
Illustrative Example



Illustrative Example



Illustrative Example



Median Unbiased Estimators

- define

$$Q_{gen} = \sum w_i (y_i - \hat{\mu})^2$$

where

$$\hat{\mu} = \frac{\sum w_i y_i}{\sum w_i} \text{ and } w_i = 1/(v_i + \tau^2)$$

- let $\hat{\tau}^2_{\text{PMM}}$ be that value such that $Q_{gen} = \chi^2_{k-1;.5}$ (i.e., the median of a chi-square distribution with $\mathrm{df} = k - 1$)
- the DL estimator is a special case of the generalized Q-statistic estimator where $w_i = 1/v_i$ ([16])
- can compute a CI based on the exact distribution of Q_{gen} ([25])
- 0% CI gives the respective median unbiased estimator $\hat{\tau}^2_{\text{DLM}}$

Illustrative Example

Estimator	Value	CI	Bounds
DL	0.061		
REML	0.203	PL	0.054 to 0.735
PM	0.219		
PMM	0.241	QP	0.070 to 0.881
DLM	0.090	GENQ	0.014 to 0.499

Final Notes / Conclusions

- median unbiased estimators ([26]) $\hat{\tau}^2_{\text{PMM}}$ and $\hat{\tau}^2_{\text{DLM}}$ are just as likely to overestimate τ^2 as to underestimate it
- median unbiased estimators are invariant under one-to-one transformations, so $\hat{\tau}_{\text{PMM}}$ and $\hat{\tau}_{\text{DLM}}$ are median unbiased for τ
- all estimators require truncation when they are negative which this introduces positive bias into $\hat{\tau}^2_{\text{PM}}$, $\hat{\tau}^2_{\text{DL}}$, and $\hat{\tau}^2_{\text{REML}}$
- $\hat{\tau}^2_{\text{PMM}}$ and $\hat{\tau}^2_{\text{DLM}}$ remain median unbiased even when $\tau^2 = 0$
- but the median unbiased estimators are less efficient
- in practice, a 95% QP CI will always encompass $\hat{\tau}^2_{\text{PM}}$ (and probably a 95% GENQ CI will always encompass $\hat{\tau}^2_{\text{DL}}$)
- more problematic: using $\hat{\tau}^2_{\text{REML}}$ with a QP CI (should use PL CI)
- this work arose from discussions with Theo Stijnen

References [1]

1. Hedges, L. V. (1983). A random effects model for effect sizes. *Psychological Bulletin*, *93*(2), 388–395.
2. Cochran, W. G. (1954). The combination of estimates from different experiments. *Biometrics*, *10*(1), 101–129.
3. Hanushek, E. A. (1974). Efficient estimators for regressing regression coefficients. *American Statistician*, *28*(2), 66–67.
4. Schmidt, F. L., Gast-Rosenberg, I., & Hunter, J. E. (1980). Validity generalization results for computer programmers. *Journal of Applied Psychology*, *65*(6), 643–661.
5. Hunter, J. E., & Schmidt, F. L. (1990). *Methods of meta-analysis: Correcting error and bias in research findings*. Newbury Park, CA: Sage.
6. DerSimonian, R., & Laird, N. (1986). Meta-analysis in clinical trials. *Controlled Clinical Trials*, *7*(3), 177–188.
7. Council, N. R. (1992). *Combining information: Statistical issues and opportunities*. Washington, DC: National Academic Press.
8. Hardy, R. J., & Thompson, S. G. (1996). A likelihood approach to meta-analysis with random effects. *Statistics in Medicine*, *15*(6), 619–629.
9. Raudenbush, S. W., & Bryk, A. S. (1985). Empirical bayes meta-analysis. *Journal of Educational Statistics*, *10*(2), 75–98.
10. Viechtbauer, W. (2005). Bias and efficiency of meta-analytic variance estimators in the random-effects model. *Journal of Educational and Behavioral Statistics*, *30*(3), 261–293.

References [2]

11. Paule, R. C., & Mandel, J. (1982). Consensus values and weighting factors. *Journal of Research of the National Bureau of Standards*, *87*(5), 377–385.
12. Morris, C. N. (1983). Parametric empirical bayes inference: Theory and applications (with discussion). *Journal of the American Statistical Association*, *78*(381), 47–65.
13. Berkey, C. S., Hoaglin, D. C., Mosteller, F., & Colditz, G. A. (1995). A random-effects regression model for meta-analysis. *Statistics in Medicine*, 14(4), 395–411.
14. Hartung, J., & Makambi, K. H. (2002). Positive estimation of the between-study variance. *South African Statistical Journal*, *36*, 55–76.
15. Sidik, K., & Jonkman, J. N. (2005). Simple heterogeneity variance estimation for meta-analysis. *Applied Statistics*, 54(2), 367–384.
16. DerSimonian, R., & Kacker, R. (2007). Random-effects model for meta-analysis of clinical trials: An update. *Contemporary Clinical Trials*, *28*(2), 105–114.
17. Morris, C. N., & Normand, S. L. (1992). Hierarchical models for combining information and for meta-analysis. In J. M. Bernardo, J. O. Berger, A. P. Dawid, & A. F. M. Smith (Eds.), *Bayesian statistics* 4 (pp. 321–344). Oxford: Oxford University Press.
18. Smith, T. C., Spiegelhalter, D. J., & Thomas, A. (1995). Bayesian approaches to random-effects meta-analysis: A comparative study. *Statistics in Medicine*, 14(24), 2685–2699.

References [3]

19. Rukhin, A. L. (2013). Estimating heterogeneity variance in meta-analysis. *Journal of the Royal Statistical Society, Series B, 75*(3), 451–469.

20. Chung, Y., Rabe-Hesketh, S., & Choi, I. H. (2013). Avoiding zero between-study variance estimates in random-effects meta-analysis. *Statistics in Medicine, 32*(23), 4071–4089.

21. Biggerstaff, B. J., & Tweedie, R. L. (1997). Incorporating variability in estimates of heterogeneity in the random effects model in meta-analysis. *Statistics in Medicine, 16*(7), 753–768.

22. Viechtbauer, W. (2007). Confidence intervals for the amount of heterogeneity in meta-analysis. *Statistics in Medicine, 26*(1), 37–52.

23. Switzer III, F. S., Paese, P. W., & Drasgow, F. (1992). Bootstrap estimates of standard errors in validity generalization. *Journal of Applied Psychology, 77*(2), 123–129.

24. Turner, R. M., Omar, R. Z., Yang, M., Goldstein, H., & Thompson, S. G. (2000). A multilevel model framework for meta-analysis of clinical trials with binary outcomes. *Statistics in Medicine, 19*(24), 3417–3432.

25. Jackson, D. (2013). Confidence intervals for the between-study variance in random effects meta-analysis using generalised cochran heterogeneity statistics. *Research Synthesis Methods, 4*(3), 220–229.

26. Brown, G. W. (1947). On small-sample estimation. *Annals of Mathematical Statistics, 18*(4), 582–585.

Thank You for Your Attention!

Questions, Comments, Suggestions?

✉ wolfgang.viechtbauer@maastrichtuniversity.nl

https://www.wvbauer.com/

https://www.metafor-project.org/

@wviechtb