Confidence Intervals for the Amount of Heterogeneity Accounted for in Meta-Regression Models

European Congress of Methodology
Ghent, Belgium
July 11-13, 2023

Wolfgang Viechtbauer
Maastricht University

2023-06-13

Overview

- describe a pseudo-R^2 statistic for meta-regression models
- describe methods for computing CIs for this R^2 statistic
- present results from a simulation study comparing the methods
- non-parametric bootstrapping appears to work adequately
- apply this method to an example for illustration
- ...

Overview

- describe a pseudo-R^2 statistic for meta-regression models
- describe methods for computing CIs for this R^2 statistic
- present results from a simulation study comparing the methods
- non-parametric bootstrapping appears to work adequately
- apply this method to an example for illustration
- ...
- profit!

Standard Linear Regression Model

- the standard linear regression model:

$$y_i = \beta_0 + \beta_1 x_{1i} + ... + \beta_p x_{pi} + \varepsilon_i$$

  where $\varepsilon_i \sim N(0, \sigma^2)$
- coefficient of determination (R^2):

$$R^2 = \frac{\mathrm{Var}(y) - \mathrm{Var}(e)}{\mathrm{Var}(y)}$$

Example

```
# examine the mtcars dataset
mtcars
```

	mpg	cyl	hp	wt	am
Mazda RX4	21.0	6	110	2.620	1
Mazda RX4 Wag	21.0	6	110	2.875	1
Datsun 710	22.8	4	93	2.320	1
Hornet 4 Drive	21.4	6	110	3.215	0
Hornet Sportabout	18.7	8	175	3.440	0
...					
Ferrari Dino	19.7	6	175	2.770	1
Maserati Bora	15.0	8	335	3.570	1
Volvo 142E	21.4	4	109	2.780	1

Example

```
# linear regression model to predict mpg (miles per gallon) from:
# - cyl (number of cylinders)
# - hp  (gross horsepower)
# - wt  (weight in 1000 lbs)
# - am  (transmission: 0 = automatic, 1 = manual)
res <- lm(mpg ~ cyl + hp + wt + am, data=mtcars)
summary(res)
```

```
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 36.14654    3.10478  11.642 4.94e-12
## cyl         -0.74516    0.58279  -1.279   0.2119
## hp          -0.02495    0.01365  -1.828   0.0786
## wt          -2.60648    0.91984  -2.834   0.0086
## am           1.47805    1.44115   1.026   0.3142
##
## Residual standard error: 2.509 on 27 degrees of freedom
## Multiple R-squared:  0.849,  Adjusted R-squared:  0.8267
## F-statistic: 37.96 on 4 and 27 DF,  p-value: 1.025e-10
```

Example

```r
# extract the R^2 value from the model object
summary(res)$r.squared
```

```
## [1] 0.8490314
```

```r
# illustrate the computation of R^2
(var(mtcars$mpg) - var(resid(res))) / var(mtcars$mpg)
```

```
## [1] 0.8490314
```


Confidence Intervals for P^2 (Smithson, 2003)

- F-statistic follows a non-central F-distribution with $\mathrm{df}_1 = p$ and $\mathrm{df}_2 = n - p - 1$ degrees of freedom with non-centrality parameter $\Delta = f^2(\mathrm{df}_1 + \mathrm{df}_2 + 1)$, where $f^2 = \frac{\mathrm{P}^2}{1-\mathrm{P}^2}$
- find the two values of Δ such that

$$\left(\int_{-\infty}^{\Delta_L} F_{\mathrm{df}_1,\mathrm{df}_2,\Delta} d_F = .975 \; ; \; \int_{-\infty}^{\Delta_U} F_{\mathrm{df}_1,\mathrm{df}_2,\Delta} d_F = .025 \right)$$

- (Δ_L, Δ_U) is an exact 95% CI for Δ (test inversion principle)
- convert into an exact 95% CI for P^2 with

$$\Delta_L / (\Delta_L + \mathrm{df}_1 + \mathrm{df}_2 + 1)$$
$$\Delta_U / (\Delta_U + \mathrm{df}_1 + \mathrm{df}_2 + 1)$$

Example

```r
# load the confintr package
library(confintr)

# calculate 95% CI for R^2
fstat <- summary(res)$fstatistic
ci_rsquared(fstat["value"], df1=fstat["numdf"], df2=fstat["dendf"])
```

```
##  Two-sided 95% F confidence interval for the population R-squared
##
## Sample estimate: 0.8490314
## Confidence interval:
##      2.5%     97.5%
## 0.6833385 0.8889493
```


Mixed-Effects Meta-Regression Model

- the random-effects model:

$$y_i = \mu + u_i + \varepsilon_i$$

where $u_i \sim N(0, \tau^2)$ and $\varepsilon_i \sim N(0, v_i)$
- the mixed-effects meta-regression model:

$$y_i = \beta_0 + \beta_1 x_{1i} + \ldots + \beta_p x_{pi} + u_i + \varepsilon_i$$

where $u_i \sim N(0, \tau^2_{res})$ and $\varepsilon_i \sim N(0, v_i)$
- pseudo R^2 statistic for meta-regression (Raudenbush, 1994):

$$R^2 = \frac{\hat{\tau}^2 - \hat{\tau}^2_{res}}{\hat{\tau}^2}$$

- see also Aloe et al. (2010)

Example: CBT for Recidivism (Landenberger & Lipsey, 2005)

```r
# load the metafor package
library(metafor)

# results from 58 studies on the effectiveness of cognitive-behavioral
# therapy (CBT) for reducing recidivism (Landenberger & Lipsey, 2005)
dat <- dat.landenberger2005

# look at a subset of the data
dat[c(1:6,56:58),c(1,8:11,18)]
```

study	n.ctrl.rec	n.ctrl.non	n.cbt.rec	n.cbt.non	sessions
Guerra (1990)	23	29	10	19	1
Leeman (1993)	15	22	3	17	1
Kownacki (1995)	4	6	2	9	1
Larson (1989)	11	2	8	5	1
Finn (1998)	22	60	16	66	5
Goldstein (1989)	14	18	5	28	2
...					
Berry (1998)	31	51	20	62	10
Pelissier (2001)	340	554	237	526	5
Pelissier (2001)	36	192	27	144	5

Example: CBT for Recidivism (Landenberger & Lipsey, 2005)

```r
# calculate log odds ratios (for non-recidivism in CBT vs. control groups)
# and the corresponding sampling variances
dat <- escalc(measure="OR", ai=n.cbt.non,  bi=n.cbt.rec,
                            ci=n.ctrl.non, di=n.ctrl.rec, data=dat)
dat
```

study	n.ctrl.rec	n.ctrl.non	n.cbt.rec	n.cbt.non	sessions	yi	vi
Guerra (1990)	23	29	10	19	1	0.4101	0.2306
Leeman (1993)	15	22	3	17	1	1.3516	0.5043
Kownacki (1995)	4	6	2	9	1	1.0986	1.0278
Larson (1989)	11	2	8	5	1	1.2347	0.9159
Finn (1998)	22	60	16	66	5	0.4138	0.1398
Goldstein (1989)	14	18	5	28	2	1.4715	0.3627
...							
Berry (1998)	31	51	20	62	10	0.6336	0.1180
Pelissier (2001)	340	554	237	526	5	0.3090	0.0109
Pelissier (2001)	36	192	27	144	5	0.0000	0.0770

Example: CBT for Recidivism (Landenberger & Lipsey, 2005)

```
# fit a random-effects model (log odds ratios and variances as input)
res1 <- rma(yi, vi, data=dat)
res1
```

```
## Random-Effects Model (k = 58; tau^2 estimator: REML)
##
## tau^2 (estimated amount of total heterogeneity): 0.1046 (SE = 0.0352)
## tau (square root of estimated tau^2 value):      0.3234
## I^2 (total heterogeneity / total variability):   70.62%
## H^2 (total variability / sampling variability):  3.40
##
## Test for Heterogeneity:
## Q(df = 57) = 213.6898, p-val < .0001
##
## Model Results:
##
## estimate      se    zval    pval   ci.lb   ci.ub
##   0.4226  0.0605  6.9880  <.0001  0.3041  0.5411
```


Example: CBT for Recidivism (Landenberger & Lipsey, 2005)

```
# fit a mixed-effects meta-regression model with number of sessions
# as a moderator (log odds ratios and variances as input)
res2 <- rma(yi, vi, mods = ~ sessions, data=dat)
res2
```

```
## Mixed-Effects Model (k = 58; tau^2 estimator: REML)
##
## tau^2 (estimated amount of residual heterogeneity):     0.0761 (SE = 0.0285)
## tau (square root of estimated tau^2 value):             0.2758
## I^2 (residual heterogeneity / unaccounted variability): 62.88%
## H^2 (unaccounted variability / sampling variability):   2.69
## R^2 (amount of heterogeneity accounted for):            27.27%
##
## Test for Residual Heterogeneity:
## QE(df = 56) = 146.6400, p-val < .0001
##
## Test of Moderators (coefficient 2):
## QM(df = 1) = 6.0695, p-val = 0.0138
##
## Model Results:
##
##           estimate      se    zval    pval   ci.lb   ci.ub
## intrcpt     0.2022  0.1015  1.9931  0.0463  0.0034  0.4010
## sessions    0.0695  0.0282  2.4636  0.0138  0.0142  0.1248
```


Example: CBT for Recidivism (Landenberger & Lipsey, 2005)

```
# extract the R^2 value from the model object
summary(res2)$R2
```

```
## [1] 27.26865
```

```
# illustrate the computation of R^2
100 * (res1$tau2 - res2$tau2) / res1$tau2
```

```
## [1] 27.26865
```


Example: CBT for Recidivism (Landenberger & Lipsey, 2005)

```
# fit a mixed-effects meta-regression model with number of sessions
# as a moderator (use the Knapp & Hartung method for inferences)
res3 <- rma(yi, vi, mods = ~ sessions, data=dat, test="knha")
res3
```

```
## Mixed-Effects Model (k = 58; tau^2 estimator: REML)
##
## tau^2 (estimated amount of residual heterogeneity):     0.0761 (SE = 0.0285)
## tau (square root of estimated tau^2 value):             0.2758
## I^2 (residual heterogeneity / unaccounted variability): 62.88%
## H^2 (unaccounted variability / sampling variability):   2.69
## R^2 (amount of heterogeneity accounted for):            27.27%
##
## Test for Residual Heterogeneity:
## QE(df = 56) = 146.6400, p-val < .0001
##
## Test of Moderators (coefficient 2):
## F(df1 = 1, df2 = 56) = 5.6305, p-val = 0.0211
##
## Model Results:
##
##           estimate      se    tval  df    pval    ci.lb   ci.ub
## intrcpt     0.2022  0.1053  1.9196  56  0.0600  -0.0088  0.4132
## sessions    0.0695  0.0293  2.3729  56  0.0211   0.0108  0.1281
```


Confidence Intervals for P^2 in Meta-Regression

- R^2 is inaccurate unless k is large (López-López et al., 2014)
- CIs for P^2 could be used to gauge the precision of the estimate
- no methods for calculating such CIs have been proposed so far
- try out method from standard linear regression
- also examine parametric and non-parametric bootstrapping

Simulation Study: Setup

- $y_i \sim \mathrm{N}(\theta_i, v_i)$ where $v_i \sim \mathrm{Unif}(.005, .1)$ for $i = 1, \ldots, k$
- mimics SMDs for studies with n between 20 and 400
- $\theta_i = \beta_0 + \beta_1 x_i + u_i$ where $x_i \sim N(0,1)$
- $\beta_1 = \sqrt{\tau^2 \times \mathrm{P}^2}$ and $u_i \sim \mathrm{N}(0, \tau^2 \times (1 - \mathrm{P}^2))$
- this way $\mathrm{Var}(\theta_i) = \tau^2$
- 176 conditions:

k	{20, 40, 80, 160}
τ^2	{0.005, 0.03, 0.1, 0.4}
P^2	{0, 0.1, 0.2, ..., 1}

- 1000 iterations per condition
- 1000 replicates for the (non-)parametric bootstrapping

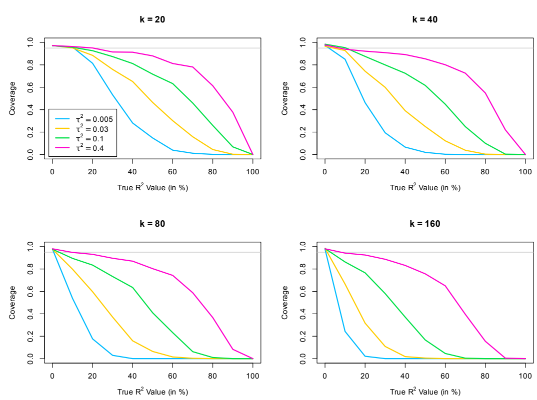
Simulation Study: Methods Examined

- F-based method (as in standard regression)
- parametric bootstrapping
  - normal CI
  - basic CI
  - percentile CI
- non-parametric bootstrapping
  - normal CI
  - basic CI
  - percentile CI
  - BCA CI
  - BCA/percentile CI

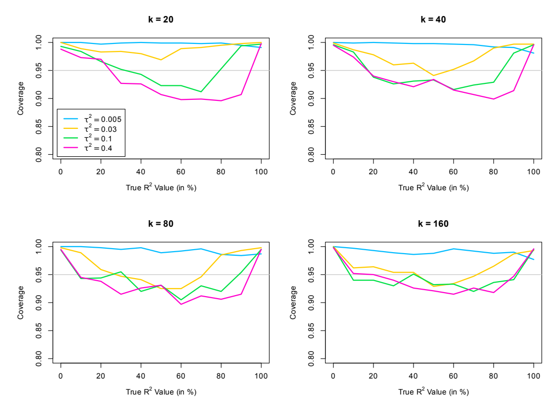
Results

- F-based method performs terribly
- for parametric bootstrapping, the percentile CI works best but is conservative when τ^2 is small and/or P^2 is close to 0%/100%
- for non-parametric bootstrapping, the BCA CI works best (with percentile CI substituted when BCA CI cannot be obtained)

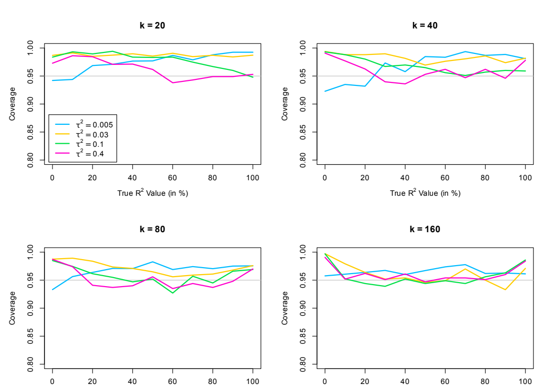
Results: F-Based Method



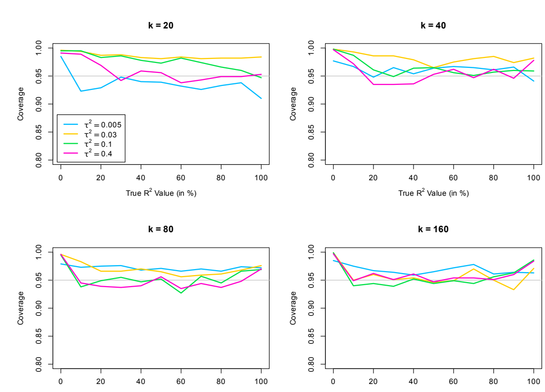
Results: Parametric Bootstrap with Percentile CI



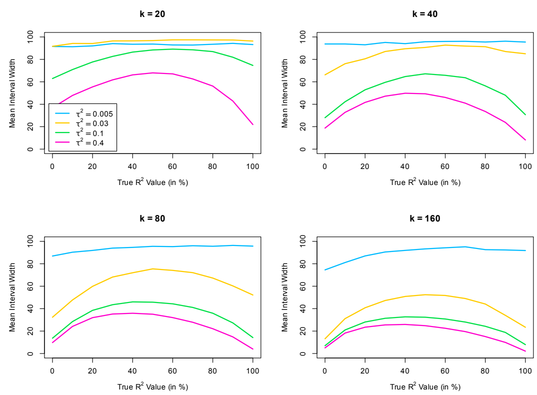
Results: Non-Parametric Bootstrap with BCA CI



Results: Non-Parametric Bootstrap with BCA/Percentile CI



Results: Non-Parametric Bootstrap with BCA/Percentile CI



Example: CBT for Recidivism (Landenberger & Lipsey, 2005)

```
# fit a mixed-effects meta-regression model with number of sessions
# as a moderator (log odds ratios and variances as input)
res2 <- rma(yi, vi, mods = ~ sessions, data=dat)
res2
```

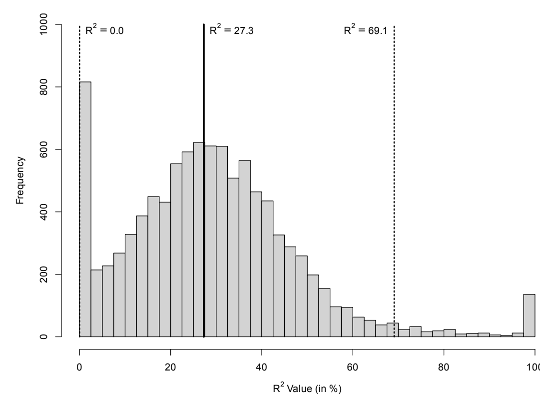
```
## Mixed-Effects Model (k = 58; tau^2 estimator: REML)
##
## tau^2 (estimated amount of residual heterogeneity):     0.0761 (SE = 0.0285)
## tau (square root of estimated tau^2 value):             0.2758
## I^2 (residual heterogeneity / unaccounted variability): 62.88%
## H^2 (unaccounted variability / sampling variability):   2.69
## R^2 (amount of heterogeneity accounted for):            27.27%
##
## Test for Residual Heterogeneity:
## QE(df = 56) = 146.6400, p-val < .0001
##
## Test of Moderators (coefficient 2):
## QM(df = 1) = 6.0695, p-val = 0.0138
##
## Model Results:
##
##           estimate      se    zval    pval   ci.lb   ci.ub
## intrcpt     0.2022  0.1015  1.9931  0.0463  0.0034  0.4010
## sessions    0.0695  0.0282  2.4636  0.0138  0.0142  0.1248
```


Example: CBT for Recidivism (Landenberger & Lipsey, 2005)

```
# load the boot package
library(boot)

# function to draw bootstrap replicates, fit the model, and return R^2
boot.func.np <- function(dat, indices, formula)
   rma(yi, vi, mods = formula, data=dat[indices,], method=c("REML","DL"))$R2

# run the bootstrapping (10000 replicates)
set.seed(1234)
res.boot <- boot(dat, boot.func.np, R=10000, formula=formula(res2))

# obtain the percentile and BCA CIs
boot.ci(res.boot, type=c("perc","bca"))
```

```
## BOOTSTRAP CONFIDENCE INTERVAL CALCULATIONS
## Based on 10000 bootstrap replicates
## Intervals :
## Level      Percentile            BCa
## 95%   ( 0.00, 74.98 )   ( 0.00, 69.07 )
```


Example: CBT for Recidivism (Landenberger & Lipsey, 2005)



Conclusion

- CIs for R^2 should be reported for meta-regression models
- non-parametric bootstrapping (with the BCA/percentile method) works adequately
- intervals are often quite wide (which reflects the high degree of uncertainty in the results unless k is large)

References

Aloe, A. M., Becker, B. J., & Pigott, T. D. (2010). An alternative to R^2 for assessing linear models of effect size. *Research Synthesis Methods*, 1(3-4), 272–283. https://doi.org/10.1002/jrsm.23

Landenberger, N. A., & Lipsey, M. W. (2005). The positive effects of cognitive-behavioral programs for offenders: A meta-analysis of factors associated with effective treatment. *Journal of Experimental Criminology*, 1, 451–476. https://doi.org/10.1007/s11292-005-3541-7

López-López, J. A., Marín-Martínez, F., Sánchez-Meca, J., Van den Noortgate, W., & Viechtbauer, W. (2014). Estimation of the predictive power of the model in mixed-effects meta-regression: A simulation study. *British Journal of Mathematical and Statistical Psychology*, *67*(1), 30–48. https://doi.org/10.1111/bmsp.12002

Raudenbush, S. W. (1994). Random-effects models. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis and meta-analysis* (pp. 301–321). Russell Sage Foundation.

Smithson, M. (2003). *Confidence intervals*. Sage.

Thank You for Your Attention!

Questions, Comments, Suggestions?

wolfgang.viechtbauer@maastrichtuniversity.nl

https://www.wvbauer.com

https://www.metafor-project.org

@wviechtb @wviechtb