Teacher Expectancy Random-Effects Meta-Analysis and Meta-Regression using Bayes Modal Estimation to Avoid Boundary Estimates

Yeojin Chung
Kookmin University, Seoul, Korea

Sophia Rabe-Hesketh
University of California, Berkeley, CA

In-Hee Cho
University of California, Berkeley, CA

Abstract

We provide an example of random-effects meta-analysis and meta-regression where the maximum likelihood estimate of the between-study variance is zero and where more reasonable estimates are obtained using Bayes modal estimation. See Chung, Rabe-Hesketh, Dorie, et al. (2013); Chung, Rabe-Hesketh, & Choi (2013) for the theory and further examples. Please also cite the above papers, not the current document.

Introduction

According to Raudenbush and Bryk (2002, p.210-211), "the hypothesis that teachers' expectations influence pupils' intellectual development as measured by IQ (intelligence quotient) scores has been the source of sustained and acrimonious controversy for over 20 years."

Raudenbush (1984) and Raudenbush and Bryk (1985) discuss meta-analysis of 19 experiments testing this hypothesis. In each study, children were assigned either to an experimental group or a control group. Teachers were encouraged to have high expectations of the children in the experimental group, whereas no particular expectations were encouraged of the children in the control group. The data are listed in Table 1 of Raudenbush and Bryk (1985).

It is also hypothesized that the amount of teacher-pupil contact prior to the experiment influences the size of the expectancy effect. Following Raudenbush and Bryk (1985), we will therefore also consider a meta-regression with number of weeks of prior contact as explanatory variable. The model can be written as

$$y_i = \mu + \beta x_i + \theta_i + \epsilon_i,$$

The research reported here was supported by the Institute of Education Sciences (grant R305D100017)

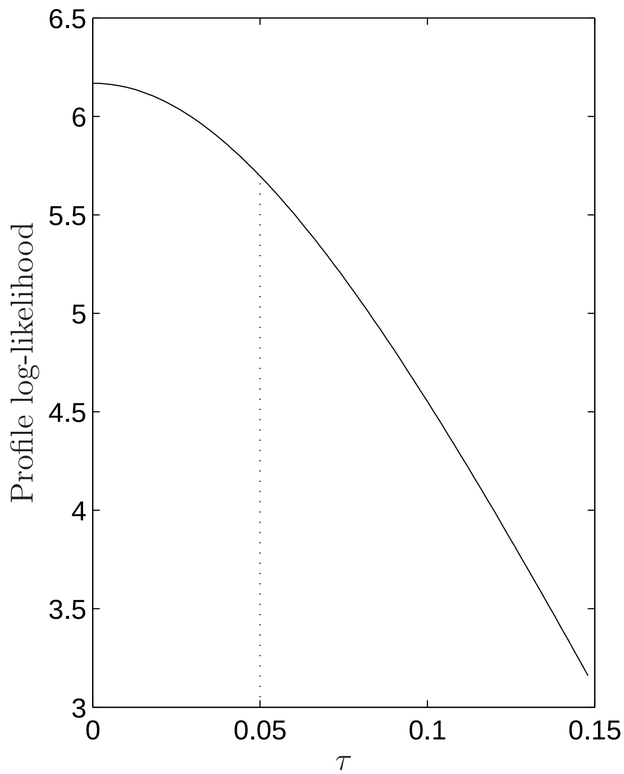

Figure 1. : Profile log-likelihood function

where x_i is the mean-centered number of weeks of prior contact, $\theta_i \sim N(0, \tau^2)$, and $\epsilon_i \sim N(0, s_i^2)$. The standard random-effects meta-analysis model corresponds to the model above with $\beta = 0$.

Maximum likelihood and restricted maximum likelihood estimates

The model with and without the covariate was estimated by maximum likelihood (ML) and restricted (or residualized) maximum likelihood (REML, Patterson & Thompson, 1971) using the `metaan` (Kontopantelis & Reeves, 2010) and `metareg` (Harbord & Higgins, 2008) commands in Stata. In the meta-analysis model without covariates, the ML estimate of μ is 0.078 with standard error 0.052 (see Table 1). The REML estimate is 0.084 with standard error 0.052. Both ML and REML estimates of the overall effect μ are not significant, but the between-study standard deviation τ is estimated as 0.112 and 0.137 by ML and REML, respectively. This suggests that there is important variability between studies.

When we include the covariate, the ML estimates $\hat{\mu}$ and $\hat{\beta}$ are both significant at the 5% level with estimates 0.134 and -0.157 and standard errors 0.040 and 0.036, respectively (also shown in Table 1). However, the residual between-study standard deviation is estimated as zero. This point estimate is unrealistic because it implies that the study-specific effect is perfectly predicted by the covariate.

Figure 1 shows the profile log-likelihood for τ for the meta-regression model (maximized with respect to μ and β.) Clearly the maximum of the profile log-likelihood is attained at

$\tau = 0$ but for slightly larger values of τ, e.g. $\tau = 0.05$ (=standard error of ML estimate $\hat{\tau}$), the profile log-likelihood does not decrease substantially. Comparing the change in profile log-likelihood with $\chi^2_{0.95}(1)/2 = 1.921$, we can infer that values for τ up to about 0.1 are reasonably supported by the data.

Bayes modal estimates

In this section we obtain Bayes modal estimates with a gamma$(2, 10^{-4})$ prior on τ using the Stata program `gllamm` (Rabe-Hesketh et al., 2005; Rabe-Hesketh & Skrondal, 2012); see Chung, Rabe-Hesketh, Dorie, et al. (2013) and Chung, Rabe-Hesketh, & Choi (2013) for full discussions of this approach. For the meta-analysis model (without covariates), estimates of μ and τ are also obtained using DL (DerSimonian and Laird) using `metaan` UMM (unweighted method of moments) using our own program (see Chung, Rabe-Hesketh, & Choi (2013) for details of these approaches).

For the meta-regression model, the BM estimate of τ is 0.054 while the ML and REML estimates are both zero. As expected, the BM estimate of τ is about one standard error (0.050) away from ML estimate of 0 and the log-likelihood decreases only 0.540 from the maximum value of 6.168, which shows that the BM estimate $\hat{\tau}_{\text{BM}} = 0.054$ is still reasonably supported by the data. As expected, the estimated standard errors of $\hat{\mu}$ and $\hat{\beta}$ are larger for BM than for ML and REML.

Table 1:: Parameter estimates in meta-analysis model and in meta-regression model with teachers' expectancy data.

	Coefficient estimates			
	$\hat{\mu}$ (se($\hat{\mu}$))	$\hat{\beta}$ (se($\hat{\beta}$))	$\hat{\tau}$	Log-likelihood
Meta-analysis model				
ML	0.078 (0.052)	·	0.112	-3.120
REML	0.084 (0.052)	·	0.137	-3.161
BM	0.089 (0.058)	·	0.160	-3.269
DL	0.089 (0.056)	·	0.161	-3.278
UMM	0.114 (0.079)	·	0.284	-4.856
Meta-regression model				
ML, REML	0.134 (0.040)	-0.157 (0.036)	0 (0.050*)	6.168
BM	0.135 (0.043)	-0.159 (0.038)	0.054 (0.041*)	5.628

*: Standard error of $\hat{\tau}$ based on the observed information.

Different types of confidence intervals

We now compare and visualize the different types of confidence intervals of μ that are discussed in Section 3.3 of Chung, Rabe-Hesketh, & Choi (2013). To do this, we consider the likelihood surface for the meta-analysis model without covariates shown as a contour plot in Figure 2. The "X" marks the maximum likelihood estimates $(\hat{\mu}_{\text{ML}}, \hat{\tau}_{\text{ML}}) = (0.078, 0.112)$.

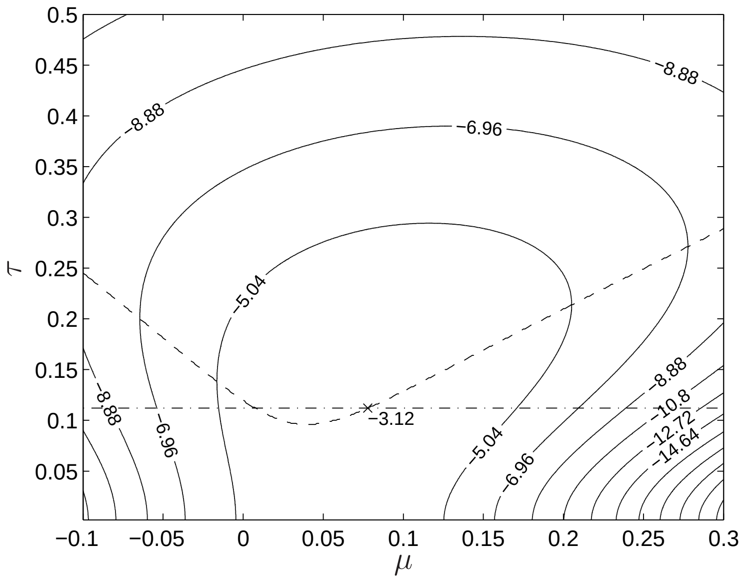

Figure 2. : Contour plot of the log-likelihood function for the meta-analysis without covariates applied to the teachers' expectancy dataset.

The dashed curve shows the values of τ that maximizes the log-likelihood for given values of μ. The log-likelihood along this path, as a function of μ, is the profile log-likelihood function $l_p(\mu)$. The dash-dotted line indicates where the conditional log-likelihood given the ML estimate of τ is located. The contours show decreases from the maximum (-3.12) in multiples of $\chi^2_{0.95}(1)/2 = 1.92$.

Figure 3 illustrates four different methods for constructing confidence intervals for μ. The top panel shows the conditional log-likelihood given $\tau = \hat{\tau}_{\mathrm{ML}}$, which is a two-dimensional view of Figure 2 along the dash-dotted line. Note that, for any given τ, $\hat{\mu}_{\mathrm{ML}}$ is a linear combination of the y_i and so it is exactly normally distributed. Therefore, the conditional log-likelihood function of μ given τ is always quadratic in μ, as observed in the top panel of Figure 3. The vertical solid line is at $\hat{\mu}_{\mathrm{ML}} = 0.078$ and the two dashed lines indicate the lower (-0.024) and upper bounds (0.179) of the 95% Wald-type CI of μ based on the expected information at $(\hat{\mu}_{\mathrm{ML}}, \hat{\tau}_{\mathrm{ML}})$. The horizontal solid line is 1.92 ($= \chi^2_{0.95}(1)/2$) lower than the maximum so that it crosses at the bounds of the confidence interval.

The plot in the middle panel of Figure 3 shows the profile log-likelihood $l_p(\mu)$. Again the solid horizontal line is 1.92 lower than the maximum of the profile log-likelihood function. Therefore, the limits of the 95% profile likelihood CI $(-0.016, 0.205)$ are where the profile log-likelihood curve and the horizontal line cross. Since the profile log-likelihood is slightly right-skewed, the confidence interval is not centered at $\hat{\mu}_{\mathrm{ML}}$. The dash-dot curve is a quadratic approximation of the profile-likelihood function at the mode. The curvature of this curve at the mode can be calculated from the observed information at $(\hat{\mu}_{\mathrm{ML}}, \hat{\tau}_{\mathrm{ML}})$. The limits of the 95% Wald-type CI $(-0.024, 0.179)$ based on the observed information are shown as vertical dash-dot lines, which coincide with the points where the quadratic curve and the horizontal line cross each other. Although this Wald-type CI based on the observed information is not as wide as the profile likelihood CI, it is wider than the Wald-type CI based on the expected

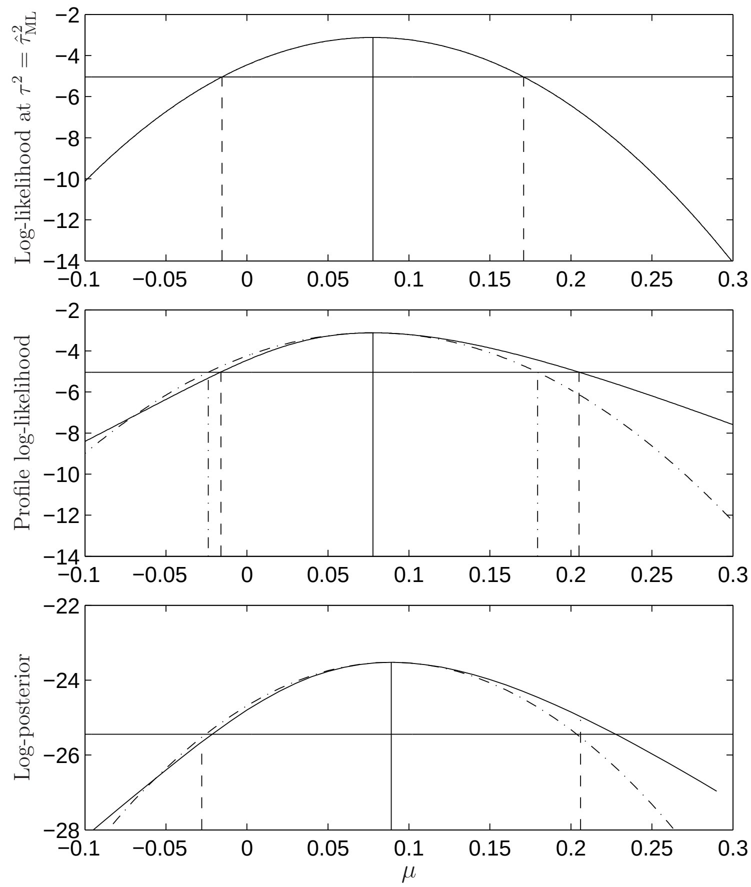

Figure 3. : Comparison of the Wald-type CI based on the expected information with ML estimates (top), profile likelihood CI and the Wald-type CI based on the observed information with ML estimates (middle), and a Wald-type CI with BM estimate (bottom).

information since it takes into account the uncertainty of $\hat{\tau}$.

The bottom panel shows the log-posterior at $\tau = \hat{\tau}_{\mathrm{BM}}$ with a gamma$(2, 10^{-4})$ prior on τ. The limits of the 95% Wald-type CI $(-0.026, 0.204)$ based on the "observed information" (the Hessian of the log-posterior function) are shown as dashed vertical lines. This confidence interval is close to the profile likelihood CI but is the widest among all the intervals considered here.

References

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

Chung, Y., Rabe-Hesketh, S., Dorie, V., Gelman, A., & Liu, J. (2013). A non-degenerate penalized likelihood estimator for variance parameters in multilevel models. *Psychometrika*, *78*, 685-709.

Harbord, R., & Higgins, J. P. T. (2008). Meta-regression in Stata. *The Stata Journal*, *8*(4), 493-519.

Kontopantelis, E., & Reeves, D. (2010). metaan: Random-effects meta-analysis. *The Stata Journal*, *10*(3), 395-407.

Patterson, H. D., & Thompson, R. (1971). Recovery of inter-block information when block sizes are unequal. *Biometrika*, *58*(3), 545-554.

Rabe-Hesketh, S., & Skrondal, A. (2012). *Multilevel and longitudinal modeling using Stata (third edition)*. College Station: Stata Press.

Rabe-Hesketh, S., Skrondal, A., & Pickles, A. (2005). Maximum likelihood estimation of limited and discrete dependent variable models with nested random effects. *Journal of Econometrics*, *128*(2), 301-323.

Raudenbush, S. W. (1984). Magnitude of teacher expectancy effects on pupil IQ as a function of the credibility of expectancy induction: A synthesis of findings from 18 experiments. *International Journal of Educational Research*, *76*(1), 85-97.

Raudenbush, S. W., & Bryk, A. S. (1985). Empirical Bayes meta-analysis. *Journal of Educational Statistics*, *10*(2), 75-98.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models (second edition)*. Thousand Oaks, CA: Sage.