Journal of Statistical Planning and Inference 140 (2010) 961–970

Contents lists available at ScienceDirect

Journal of Statistical Planning and Inference

journal homepage: www.elsevier.com/locate/jspi

How does the DerSimonian and Laird procedure for random effects meta-analysis compare with its more efficient but harder to compute counterparts?

Dan Jackson [a,*], Jack Bowden [a], Rose Baker [b]

[a] MRC Biostatistics Unit, Institute of Public Health, University Forvie Site, Robinson Way, Cambridge CB2 0SR, UK
[b] Centre for OR and Applied Statistics, School of Business, University of Salford, Salford M5 4WT, UK

ARTICLE INFO

Article history:
Received 4 June 2009
Received in revised form
21 September 2009
Accepted 23 September 2009
Available online 30 September 2009

Keywords:
Confidence intervals
Efficiency
Meta-analysis
Random effects
Profile likelihood

ABSTRACT

The procedure suggested by DerSimonian and Laird is the simplest and most commonly used method for fitting the random effects model for meta-analysis. Here it is shown that, unless all studies are of similar size, this is inefficient when estimating the between-study variance, but is remarkably efficient when estimating the treatment effect. If formal inference is restricted to statements about the treatment effect, and the sample size is large, there is little point in implementing more sophisticated methodology. However, it is further demonstrated, for a simple special case, that use of the profile likelihood results in actual coverage probabilities for 95% confidence intervals that are closer to nominal levels for smaller sample sizes. Alternative methods for making inferences for the treatment effect may therefore be preferable if the sample size is small, but the DerSimonian and Laird procedure retains its usefulness for larger samples.

1. Introduction

Meta-analysis—the pooling of separate studies concerned with the same treatment or issue—is frequently used in medical and other applications. Although some debate concerning random versus fixed effects modelling continues, the random effects model has become a standard approach. Although the conventional random effects model is easily implemented, it has often been criticised. First of all, the studies must be large enough to use normal approximations, with known variances, for the within-study distributions. More recently, methods have been proposed using exact conditional distributions (van Houwelingen et al., 1993; Taye et al., 2008; Shi and Copas, 2002) and other developments recognise that the within-study variances are given in the form of estimates and that these are typically functions of the underlying treatment effect (Böhning et al., 2002; Malzahn et al., 2000). We will assume here, however, that studies are large enough to justify the standard within-study normal approximations.

The conventional random effects model also makes the assumption that the random effect is normally distributed, although alternative distributional assumptions have been considered (Lee and Thompson, 2008; Baker and Jackson, 2008). The naive application of the random effects model has also been questioned due to the suspicion that study results may have been distorted due to publication and related biases (Baker and Jackson, 2006). Whilst recognising all these issues and concerns, we will assume here that the random effects model is appropriate and hence that all issues relate to which

* Corresponding author. Tel.: +44 1223 330376; fax: +44 1223 330388.
E-mail addresses: daniel.jackson@mrc-bsu.cam.ac.uk (D. Jackson), jack.bowden@mrc-bsu.cam.ac.uk (J. Bowden), r.d.baker@salford.ac.uk (R. Baker).

0378-3758/$ - see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.jspi.2009.09.017

estimation and related procedures to use. In particular, a now standard method originally proposed by DerSimonian and Laird (1986) is widely used, although this idea has more recently been extended (DerSimonian and Kacker, 2007). The popularity of this procedure is no doubt partly due to its relative simplicity. The DerSimonian and Laird procedure also has the merit of not requiring the assumption of normality for the random effect, an assumption that is sometimes questioned (Hardy and Thompson, 1998). Hence the procedure is 'valid approximately in a distribution-free context when there are many studies' (Higgins et al., 2009). This statement does not reassure us that the DerSimonian and Laird procedure is effective compared to the alternatives, however.

The rest of the paper is set out as follows. In Section 2, the random effects model is described and we also provide a proof that estimates of treatment effect are unbiased under the assumptions of the model. In Section 3 the asymptotic (large number of studies) efficiency of the DerSimonian and Laird estimates is investigated and the small sample case is considered in Section 4. In Section 5, an investigation into using the profile likelihood suggests that this provides more suitable coverage probabilities of confidence intervals when the sample size is modest and we conclude with a discussion in Section 6.

2. The random effects model

The conventional fixed and random effects models (DerSimonian and Laird, 1986; Biggerstaff and Tweedie, 1997; Jackson, 2009; Hardy and Thompson, 1996) initially assume that the estimate of treatment effect from the i th study, Y_i, is distributed as $Y_i|\mu_i \sim \mathrm{N}(\mu_i, \sigma_i^2)$, where μ_i is the true underlying treatment effect of the i th study and σ_i^2 is the corresponding within-study variance. The variance σ_i^2 is unknown but is replaced by a consistent estimate in practice. The conventional random effects model further assumes that $\mu_i \sim \mathrm{N}(\mu, \tau^2)$, where μ and τ^2 denote the overall treatment effect and between-study variance, respectively, and that the studies are independent. This provides the marginal distributions $Y_i \sim \mathrm{N}(\mu, \sigma_i^2 + \tau^2)$.

2.1. The standard procedure

The usual procedure begins by estimating τ^2. Once $\hat{\tau}^2$ has been evaluated, irrespective of how this has been obtained, the standard inference for μ is straightforward, as $\hat{\tau}^2$ is effectively used or 'plugged in' as the true value. The simplest and most commonly used estimate of τ^2 is the DerSimonian and Laird (1986) estimate. This uses the Q statistic,

$$Q = \sum_{i=1}^{n} w_i (y_i - \overline{y})^2,$$

where $w_i = \sigma_i^{-2}$, $\overline{y} = \sum_{i=1}^{n} w_i y_i / \sum_{i=1}^{n} w_i$ and n denotes the number of studies. Under the assumptions of the random effects model it can be shown that the expectation of Q is

$$\mathrm{E}[Q] = (n-1) + \left(S_1 - \frac{S_2}{S_1}\right)\tau^2$$

where $S_r = \sum_{i=1}^{n} w_i^r$, which provides the DerSimonian and Laird estimate

$$\hat{\tau}^2_{DL} = \max\left(0, \frac{Q-(n-1)}{S_1 - \frac{S_2}{S_1}}\right). \tag{1}$$

The corresponding estimate of treatment effect is

$$\hat{\mu}_{DL} = \frac{\sum_{i=1}^{n} \frac{y_i}{\sigma_i^2 + \hat{\tau}^2_{DL}}}{\sum_{i=1}^{n} \frac{1}{\sigma_i^2 + \hat{\tau}^2_{DL}}}. \tag{2}$$

Confidence intervals are typically obtained using the approximation $\hat{\mu}_{DL} \sim \mathrm{N}(\mu, (\sum_{i=1}^{n} w_i^*)^{-1})$, where $w^* = 1/(\sigma_i^2 + \hat{\tau}^2_{DL})$, which is justified assuming that the studies are sufficiently large and also that there is at least a moderate number of these. For example, $100(1-\alpha)\%$ confidence intervals for μ are obtained as

$$\hat{\mu}_{DL} \pm Z_{\alpha/2} \left(\sum_{i=1}^{n} w_i^*\right)^{-1/2}, \tag{3}$$

where $Z_{\alpha/2}$ denotes the $\alpha/2$ quantile of the standard normal distribution. Quantiles from the t distribution, with $(n-1)$ degrees of freedom, rather than the standard normal, are sometimes used in (3) instead.