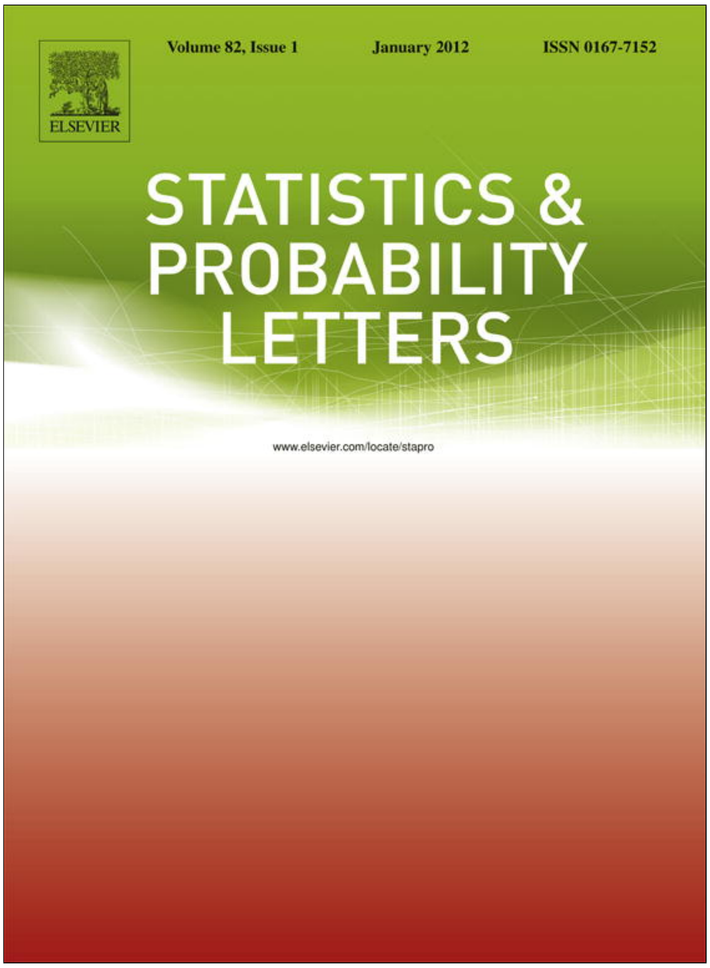

(This is a sample cover image for this issue. The actual cover is not yet available at this time.)

Statistics and Probability Letters 82 (2012) 1318–1325

Contents lists available at SciVerse ScienceDirect

Statistics and Probability Letters

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

Estimating common mean and heterogeneity variance in two study case meta-analysis

Andrew L. Rukhin *

Statistical Engineering Division, National Institute of Standards and Technology, Gaithersburg, MD 20899, USA

ARTICLE INFO

Article history:
Received 5 December 2011
Received in revised form 22 March 2012
Accepted 23 March 2012
Available online 1 April 2012

Keywords:
DerSimonian–Laird procedure
Mill's ratio
Quadratic risk
Random effects model
Restricted likelihood

ABSTRACT

The relative behavior of estimators of the common mean and of the heterogeneity variance in the simple random effects model of meta-analysis is explored. A new risk function relating these estimation problems is introduced. Bayes estimators for each of the parameters are derived.

Published by Elsevier B.V.

1. Parameter estimation in meta-analysis: random effects model

In a simple random effects model of meta-analysis involving, say, p studies the data is supposed to consist of normally distributed x_i, $i = 1, \ldots, p$, with an unknown mean μ and the variance $\tau^2 + s_i^2$. Here s_i^2 represents the reported uncertainty of the i-th study, and τ^2 is the variance of the between-study effect arising in the random effects model. In practice s_i^2 are often treated as given constants, and then the problem becomes one of estimating the common mean μ and the non-negative heterogeneity variance τ^2. This problem is considered here.

If τ^2 is known, then the best linear unbiased estimator of μ is the weighted means statistic, $\tilde{\mu} = \sum \omega_i^0 x_i$, with the normalized weights,

$$\omega_i^0 = \frac{1}{\tau^2 + s_i^2}\left[\sum_j \frac{1}{\tau^2 + s_j^2}\right]^{-1}, \qquad \sum \omega_i^0 = 1.$$

Under the normality assumption and also the maximum likelihood estimator, the best unbiased statistic is minimax and admissible. In order to estimate μ by the traditionally used plug-in version of $\tilde{\mu}$, say, $\tilde{x} = \sum_i x_i(\tilde{\tau}^2 + s_i^2)^{-1}[\sum_i (\tilde{\tau}^2 + s_i^2)^{-1}]^{-1}$, one needs an estimate $\tilde{\tau}^2$, $\tilde{\tau}^2 \geq 0$.

DerSimonian and Laird (1986) have suggested such a procedure with estimators of τ^2 and of μ. The latter has become very popular in meta-analysis but the estimator of τ^2 is known to have some undesirable features (e.g. Jackson et al., 2010). One

* Correspondence to: Statistical Engineering Division, Information Technology Laboratory, National Institute of Standards and Technology, U.S. Department of Commerce, Gaithersburg, MD 20899, USA. Tel.: +1 3019752951; fax: +1 3019753144.
E-mail addresses: rukhin@math.umbc.edu, andrew.rukhin@nist.gov.

0167-7152/$ – see front matter. Published by Elsevier B.V.
doi:10.1016/j.spl.2012.03.031

of the goals of this note is to explain this phenomenon by investigating the relationship between the mean squared error of μ-estimators and a special risk function for τ^2-estimation when $p = 2$. Another goal is to discuss admissible (Bayes) estimators for each of the parameters.

2. Estimating heterogeneity variance

2.1. Quadratic estimators

In this section we introduce estimators of the form

$$\tilde{\tau}^2_{(\alpha\beta)} = \max[0, \alpha(x_2 - x_1)^2/2 - \beta s^2], \tag{1}$$

where α and β are non-negative constants, $s^2 = (s_1^2 + s_2^2)/2$. By using the fact that $(x_2 - x_1)^2 \sim 2(\tau^2 + s^2)\chi_1^2$, the expression for their quadratic risk is derived next. Although the quadratic loss $(\tilde{\tau}^2 - \tau^2)^2$ may not be the most appropriate when estimating a non-negative τ^2, we use it here mainly because many other loss functions which depend on the ratio $\tilde{\tau}^2/\tau^2$ lead to infinite risks at $\tau^2 = 0$.

Besides its simplicity, the class (1) can be motivated by the fact that it includes the restricted maximum likelihood estimator,

$$\tilde{\tau}^2_{(1\,1)} = \max[0, (x_2 - x_1)^2/2 - s^2],$$

which coincides with the DerSimonian–Laird procedure. Indeed, the negative logarithm of the restricted likelihood function, say, $\mathcal{L} = (x_2 - x_1)^2/[2(\tau^2 + s^2)] + \log 2(\tau^2 + s^2)$, is maximized by $\tilde{\tau}^2_{(1\,1)}$.

We denote $\gamma = \sqrt{\beta/\alpha}$, $u^2 = s^2/(\tau^2 + s^2)$, $0 < u \le 1$, and by Φ and φ the standard normal distribution function and density respectively. Then since $\tilde{\tau}^2_{(\alpha\beta)} \sim (\tau^2 + s^2)\max[\alpha\chi_1^2 - \beta u^2, 0]$,

$$\begin{aligned} E(\tilde{\tau}^2_{(\alpha\beta)} - \tau^2)^2 &= \tau^4 Pr(\chi_1^2 \le \gamma^2 u^2) + (\tau^2 + s^2)^2 E[\alpha\chi_1^2 - 1 + (1-\beta)u^2]^2 \mathbf{1}_{\{\chi_1^2 > \gamma^2 u^2\}} \\ &= \tau^4[2\Phi(\gamma u) - 1] + 2(\tau^2 + s^2)^2 \int_{\gamma u}^{\infty} [\alpha z^2 - 1 + (1-\beta)u^2]^2 \varphi(z)dz \\ &= \tau^4[2\Phi(\gamma u) - 1] + 2(\tau^2 + s^2)^2 \times \{\alpha\gamma u[(2-\beta)u^2 + 3\alpha - 2]\varphi(\gamma u) \\ &\quad + [(1 - \alpha - (1-\beta)u^2)^2 + 2\alpha^2][1 - \Phi(\gamma u)]\}. \end{aligned} \tag{2}$$

If $\beta = 1$, $\gamma = 1/\sqrt{\alpha}$, and

$$E(\tilde{\tau}^2_{(\alpha 1)} - \tau^2)^2 = \tau^4[2\Phi(\gamma u) - 1] + 2(\tau^2 + s^2)^2\{\sqrt{\alpha}u(u^2 + 3\alpha - 2)\varphi(\gamma u) + (1 - 2\alpha + 3\alpha^2)[1 - \Phi(\gamma u)]\}.$$

In particular, when $\alpha = \beta = 1$, $\gamma = 1$, corresponding to the DerSimonian–Laird procedure,

$$E(\tilde{\tau}^2_{(11)} - \tau^2)^2 = \tau^4[2\Phi(u) - 1] + 2(\tau^2 + s^2)^2[u(u^2 + 1)\varphi(u) + 2(1 - \Phi(u))].$$

This fact is confirmed by the formula for $\alpha = \beta$, $\gamma = 1$,

$$\begin{aligned} E(\tilde{\tau}^2_{(\alpha\alpha)} - \tau^2)^2 &= \tau^4[2\Phi(u) - 1] + 2(\tau^2 + s^2)^2\{\alpha u[(2-\alpha)u^2 + 3\alpha - 2]\varphi(u) \\ &\quad + [(1-\alpha)^2(1-u^2)^2 + 2\alpha^2][1 - \Phi(u)]\}. \end{aligned}$$

If $\beta = 0$, $\gamma = 0$, so that

$$E(\tilde{\tau}^2_{(\alpha 0)} - \tau^2)^2 = (\tau^2 + s^2)^2[(1 - \alpha - u^2)^2 + 2\alpha^2] = (1 - 2\alpha + 3\alpha^2)\tau^4 - 2\alpha(1 - 3\alpha)s^2\tau^2 + 3\alpha^2 s^4.$$

When $\tau^2 = 0$, $u = 1$, (2) gives

$$E\tilde{\tau}^4_{(\alpha\beta)}/s^4 = 2\alpha^2\{\gamma(3 - \gamma^2)\varphi(\gamma) + [(\gamma^2 - 1)^2 + 2][1 - \Phi(\gamma)]\}.$$

The function, $\gamma(3 - \gamma^2)\varphi(\gamma) + [(\gamma^2 - 1)^2 + 2][1 - \Phi(\gamma)]$, of non-negative γ monotonically decreases from 1.5 to zero. Thus, unsurprisingly, $\alpha = 0$ is optimal for small τ^2, and for a fixed α, a larger β gives a smaller value of the quadratic risk at the origin.

When $\beta = 0$, $\tau^2 = 0$, $\gamma = 0$, and

$$E\tilde{\tau}^4_{(\alpha 0)}/s^4 = 3\alpha^2.$$

The risk at zero of the DerSimonian–Laird estimator is $4[\varphi(1) + 1 - \Phi(1)]s^4 \approx 1.6025s^4$. The quadratic risk of $\tilde{\tau}^2_{(1/3\ 1/3)}$ at $\tau^2 = 0$ is 9 times smaller, $4[\varphi(1) + 1 - \Phi(1)]s^4/9 \approx 0.1781s^4$. Under the quadratic loss the latter estimator as well as the estimator $\tilde{\tau}^2_{(1/3\ 0)} = (x_2 - x_1)^2/6$, whose risk is $(2\tau^4 + s^4)/3$, are substantially better than the DerSimonian–Laird estimator for all τ^2. The estimator $\tilde{\tau}^2_{(1/2\ 0)} = (x_2 - x_1)^2/4$, with risk $(3\tau^4 - 2s^2\tau^2 + s^4)/4$, is less competitive under this criterion, being worse than $\tilde{\tau}^2_{(1/3\ 0)}$, but providing an improvement over $\tilde{\tau}^2_{(1\ 1)}$.

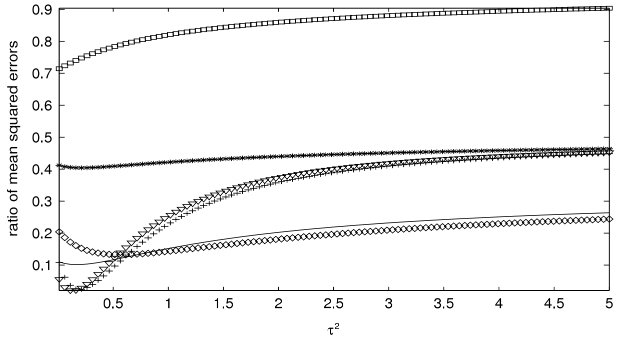

Fig. 1. Plots of ratios of quadratic risk functions of estimators based on $\tilde{\tau}^2_{(1/2\ 0)}$ (line marked by squares), $\tilde{\tau}^2_{(1/3\ 1/3)}$ (continuous line), $\tilde{\tau}^2_{(1/3\ 0)}$ (line marked by diamonds), $\tilde{\tau}^2_0$ (line marked by $*$), $\tilde{\tau}^2_1$ (line marked by triangles), $\tilde{\tau}^2_2$ (line marked by $+$) to the mean squared error of $\tilde{\tau}^2_{(1\ 1)}$.

According to (2), when $\tau^2 \to \infty$,

$$E(\tilde{\tau}^2_{(\alpha\beta)}/\tau^2 - 1)^2 \sim 1 - 2\alpha + 3\alpha^2, \tag{3}$$

which shows that the asymptotically optimal choice is $\alpha = 1/3$. This fact suggested to look at $\tilde{\tau}^2_{(1/3\ 0)}$ and $\tilde{\tau}^2_{(1/3\ 1/3)}$.

Fig. 1 plots the ratios of the mean squared errors of these estimators and $\tilde{\tau}^2_{(1/2\ 0)}$ to the mean squared error of the DerSimonian–Laird procedure $\tilde{\tau}^2_{(1\ 1)}$ The estimator $\tilde{\tau}^2_{(1/3\ 1/3)}$ is slightly better than $\tilde{\tau}^2_{(1/3\ 0)}$ for small τ^2. For large τ^2 the situation is reversed.

2.2. Bayes estimators

Under the uniform (non-informative) prior for μ, and a prior distribution Π for τ^2, the Bayes estimator of τ^2 has the form,

$$\tilde{\tau}^2 = \tilde{\tau}^2(x_1, x_2) = \frac{\int_0^\infty \int_{-\infty}^\infty \tau^2 \prod \frac{e^{-(x_i-\mu)^2/[2(\tau^2+s_i^2)]}}{\sqrt{\tau^2+s_i^2}} d\mu d\Pi(\tau^2)}{\int_0^\infty \int_{-\infty}^\infty \prod \frac{e^{-(x_i-\mu)^2/[2(\tau^2+s_i^2)]}}{\sqrt{\tau^2+s_i^2}} d\mu d\Pi(\tau^2)}$$

$$= \frac{\int_0^\infty \tau^2 e^{-(x_2-x_1)^2/[4(\tau^2+s^2)]} \frac{d\Pi(\tau^2)}{\sqrt{\tau^2+s^2}}}{\int_0^\infty e^{-(x_2-x_1)^2/[4(\tau^2+s^2)]} \frac{d\Pi(\tau^2)}{\sqrt{\tau^2+s^2}}}. \tag{4}$$

In our situation the Bayes estimators corresponding to the uniform prior for μ can be interpreted as the solutions based on the restricted likelihood function.

The prior density

$$\pi(\tau^2) = \frac{e^{-\beta/[4(\tau^2+s^2)]}}{(\tau^2+s^2)^{\rho+3/2}} \tag{5}$$

with hyper-parameters β and ρ provides a tractable estimator. The case when $\beta = 0$, $\rho = -1/2$ in (5) corresponds to the Jeffreys prior evaluated from the mentioned restricted likelihood. Indeed $E\mathcal{L}'' = -(\tau^2+s^2)^{-2}$.

Let $P(x, a) = \int_0^x e^{-t} t^{a-1}\, dt/\Gamma(a)$ denote the incomplete gamma-function, $v = [(x_2 - x_1)^2 + \beta]/2$. Then for $\beta = 0$,

$$\tilde{\tau}^2_B = \frac{vP(v/(2s^2), \rho)}{2\rho P(v/(2s^2), \rho+1)} - s^2.$$

A choice of the hyper-parameter ρ can be motivated by the asymptotic risk behavior of $\tilde{\tau}^2_B$ when $\tau^2 \to \infty$. Indeed if $v \to \infty$, $\tilde{\tau}^2_B \sim v/(2\rho)$, so that (3) with $\rho = 1/(2\alpha)$ gives the asymptotically optimal choice, $\rho = 3/2$.

When $\rho = 3/2$, we put

$$\tilde{\tau}^2_0 = \frac{vP(v/(2s^2), 1.5)}{3P(v/(2s^2), 2.5)} - s^2. \tag{6}$$

The quadratic risk of $\tilde{\tau}^2_0$ at $\tau^2 = 0$ can be readily found,

$$E\tilde{\tau}^4_0 = \frac{2s^4}{3}.$$

Indeed, when $\tau^2 = 0$, the random variable v/s^2 has the distribution χ_1^2, so that

$$E\tilde{\tau}_0^2 = \frac{2s^2}{3\sqrt{\pi}} \int_0^\infty \frac{\sqrt{y}P(y, 1.5)e^{-y}dy}{P(y, 2.5)} - s^2.$$

Recognizing $\sqrt{y}e^{-y}P(y, 1.5)/\Gamma(1.5)$ as the derivative of $P^2(y, 1.5)/2$ and integrating by parts, we obtain

$$\frac{1}{\Gamma(1.5)} \int_0^\infty \frac{\sqrt{y}P(y, 1.5)e^{-y}dy}{P(y, 2.5)} = \frac{1}{2} + \frac{1}{2\Gamma(2.5)} \int_0^\infty \frac{y^{3/2}P^2(y, 1.5)e^{-y}dy}{P^2(y, 2.5)}.$$

Therefore,

$$\begin{aligned} E\tilde{\tau}_0^4 &= \frac{4s^4}{9\sqrt{\pi}} \int_0^\infty \frac{y^{3/2}P^2(y, 1.5)e^{-y}dy}{P^2(y, 2.5)} - 2s^2(E\tilde{\tau}_a^2 + s^2) + s^4 \\ &= \frac{s^4}{3}\left[\frac{4}{\sqrt{\pi}} \int_0^\infty \frac{\sqrt{y}P(y, 1.5)e^{-y}dy}{P(y, 2.5)} - 1\right] - \frac{4s^4}{3\sqrt{\pi}} \int_0^\infty \frac{\sqrt{y}P(y, 1.5)e^{-y}dy}{P(y, 2.5)} + s^4 \\ &= \frac{2s^4}{3} \end{aligned}$$

which is smaller than the risk at zero of the DerSimonian–Laird estimator or of $\tilde{\tau}^2_{(1/2\ 0)}$, but larger than that of $\tilde{\tau}^2_{(1/3\ 1/3)}$ or $\tilde{\tau}^2_{(1/3\ 0)}$.

To remedy this fact, one may be interested in prior distributions $\Pi(\tau^2)$ with a possible atom at 0, and a density $\pi(\tau^2)$ for $\tau^2 > 0$. If similar previous studies are available, the probability of the zero value of τ^2 can be taken to be the proportion of cases when τ^2 was estimated by 0.

We denote by λ the odds ratio, $\lambda = Pr(\tau^2 = 0)/[1 - Pr(\tau^2 = 0)]$, and put $\xi = \lambda/\Gamma(\rho)$. Then

$$\tilde{\tau}_B^2 = \frac{vP(v/(2s^2), \rho)/2 + \xi[v/(2s^2)]^{\rho+1}e^{-v/(2s^2)}}{\rho P(v/(2s^2), \rho + 1) + \xi[v/(2s^2)]^{\rho+1}e^{-v/(2s^2)}/s^2} - s^2. \tag{7}$$

The ratios of the quadratic risk functions of estimators $\tilde{\tau}_i^2$ in (7) with $i = 0, 1, 2$ corresponding to $\lambda = 0$, $\lambda = 0.5$, and $\lambda = 1$ respectively to that of $\tilde{\tau}^2_{(1\ 1)}$ are also depicted in Fig. 1. Remarkably, both Bayes estimators $\tilde{\tau}_1^2$ and $\tilde{\tau}_2^2$ with a mass point at $\tau^2 = 0$ have a smaller mean squared error than the Bayes rule $\tilde{\tau}_0^2$ in the considered range, $0 \le \tau^2 \le 5$. (Actually, dominance of $\tilde{\tau}_1^2$ and $\tilde{\tau}_2^2$ holds for $\tau^2 \le 15$.) In this Figure $s^2 = 1$. Similar results hold for other loss functions like the absolute value loss.

3. Estimating the common mean

3.1. New risk function for τ^2

The conclusions reached at in Section 2 are to be contrasted with the quadratic risk behavior of μ-estimators. Let Λ be a prior distribution for τ^2 so that the Bayes estimator of μ has the form

$$\begin{aligned} \tilde{x}_B = \tilde{x}_B(x_1, x_2) &= \frac{\int_0^\infty \int_{-\infty}^\infty \mu \prod \frac{e^{-(x_i-\mu)^2/[2(\tau^2+s_i^2)]}}{\sqrt{\tau^2+s_i^2}} d\mu d\Lambda(\tau^2)}{\int_0^\infty \int_{-\infty}^\infty \prod \frac{e^{-(x_i-\mu)^2/[2(\tau^2+s_i^2)]}}{\sqrt{\tau^2+s_i^2}} d\mu d\Lambda(\tau^2)} \\ &= \frac{\int_0^\infty [(\tau^2 + s_2^2)x_1 + (\tau^2 + s_1^2)x_2]e^{-(x_2-x_1)^2/[4(\tau^2+s^2)]} \frac{d\Lambda(\tau^2)}{(\tau^2+s^2)^{3/2}}}{2\int_0^\infty e^{-(x_2-x_1)^2/[4(\tau^2+s^2)]} \frac{d\Lambda(\tau^2)}{(\tau^2+s^2)^{1/2}}}. \end{aligned} \tag{8}$$

With $d\Lambda(\tau^2) = (\tau^2 + s^2)d\Pi(\tau^2)$,

$$\tilde{x}_B = \frac{(\tilde{\tau}_B^2 + s_2^2)x_1 + (\tilde{\tau}_B^2 + s_1^2)x_2}{2(\tilde{\tau}_B^2 + s^2)},$$

i.e., the Bayes estimator of μ is the weighted mean with weights inversely proportional to $\tilde{\tau}_B^2 + s_i^2$. These Bayes weights ω_1, ω_2 are invariant functions of x_1, x_2, depending only on $x_2 - x_1$.

For any such estimator, say, $\tilde{x} = \sum \omega_i x_i$,

$$\begin{aligned} E(\tilde{x}-\mu)^2 &= \left[\sum_i \frac{1}{\tau^2+s_i^2}\right]^{-1} + E\left[\sum_i \omega_i(x_i-\tilde{\mu})\right]^2 \\ &= \frac{(\tau^2+s_1^2)(\tau^2+s_2^2)}{2(\tau^2+s^2)} + \frac{(s_2^2-s_1^2)^2}{16(\tau^2+s^2)^2} E\frac{(x_2-x_1)^2(\tilde{\tau}^2-\tau^2)^2}{(\tilde{\tau}^2+s^2)^2} \\ &= \frac{\tau^2+s^2}{2} + \frac{(s_2^2-s_1^2)^2}{8(\tau^2+s^2)}[R(\tilde{\tau}^2,\tau^2)-1]. \end{aligned} \tag{9}$$

Here

$$R(\tilde{\tau}^2,\tau^2) = E\frac{(x_2-x_1)^2(\tilde{\tau}^2-\tau^2)^2}{2(\tau^2+s^2)(\tilde{\tau}^2+s^2)^2} = E\frac{(x_2-x_1)^2}{2(\tau^2+s^2)}\left(1-\frac{\tau^2+s^2}{\tilde{\tau}^2+s^2}\right)^2,$$

is the new risk of the corresponding $\tilde{\tau}^2$ estimator which completely determines the variance of $\tilde{x}$. The resulting random loss function,

$$\frac{(x_2-x_1)^2(\tau^2+s^2)}{2}\left(\frac{1}{\tilde{\tau}^2+s^2}-\frac{1}{\tau^2+s^2}\right)^2,$$

is very different from the quadratic loss. Indeed it is designed to estimate $(\tau^2+s^2)^{-1}$ rather than τ^2 itself. Arguably this loss is most relevant for τ^2-estimators if their purpose is to provide the weights for the weighted means statistics for μ-estimation. It explains why τ^2-estimators which give reasonably good weights for $\tilde{x}$ may have a large mean squared error (or other risk) which discourages large values of such estimators.

Under notation of Section 2.1, when $0 < \beta \le 1$, one has for an estimator of the form (1),

$$\begin{aligned} R(\tilde{\tau}^2_{(\alpha\beta)},\tau^2) &= \frac{\tau^4}{s^4}E\chi_1^2 \mathbf{1}_{\{\chi_1^2\le\gamma^2u^2\}} + E\chi_1^2\left[1-\frac{1}{\alpha\chi_1^2+(1-\beta)u^2}\right]^2 \mathbf{1}_{\{\chi_1^2>\gamma^2u^2\}} \\ &= \frac{\tau^4}{s^4}[2\Phi(\gamma u)-1-2\gamma u\varphi(\gamma u)] + 2\left(1-\frac{2}{\alpha}\right)[1-\Phi(\gamma u)+\gamma u\varphi(\gamma u)] \\ &\quad + \frac{2(1-\beta)u^2+1}{\alpha^2}E\frac{\mathbf{1}_{\{\chi_1^2>\gamma^2u^2\}}}{\chi_1^2+(1-\beta)u^2/\alpha} - \frac{(1-\beta)u^2}{\alpha^3}E\frac{\mathbf{1}_{\{\chi_1^2>\gamma^2u^2\}}}{[\chi_1^2+(1-\beta)u^2/\alpha]^2}. \end{aligned} \tag{10}$$

If $\beta > 1$, the R-risk is infinite. When $\beta = 1$,

$$\begin{aligned} R(\tilde{\tau}^2_{(\alpha 1)},\tau^2) &= \frac{\tau^4}{s^4}[2\Phi(\gamma u)-1-2\gamma u\varphi(\gamma u)] + 2\left(1+\frac{\gamma^2}{u^2}\right)\gamma u\varphi(\gamma u) \\ &\quad + 2(1-2\gamma^2-\gamma^4)[1-\Phi(\gamma u)]. \end{aligned} \tag{11}$$

Indeed integration by parts easily shows that

$$E\frac{\mathbf{1}_{\{\chi_1^2>u^2\}}}{\chi_1^2} = 2\left[\frac{\varphi(u)}{u}-1+\Phi(u)\right].$$

For the DerSimonian–Laird procedure, $\gamma = 1$, so that

$$R(\tilde{\tau}^2_{(11)},\tau^2) = \tau^4 s^{-4}[2\Phi(u)-1-2u\varphi(u)] + 2(u^{-1}+u)\varphi(u) - 4[1-\Phi(u)].$$

When $\beta = 0$, $\gamma = 0$,

$$\begin{aligned} R(\tilde{\tau}^2_{(\alpha 0)},\tau^2) &= E\chi_1^2\left(1-\frac{1}{\alpha\chi_1^2+u^2}\right)^2 \\ &= 1-\frac{2}{\alpha}-\frac{1}{2\alpha^2}+\frac{[u^2(4\alpha+1)+\alpha]}{2\alpha^{5/2}u}M\left(\frac{u}{\sqrt{\alpha}}\right). \end{aligned} \tag{12}$$

Here $M(u) = [1-\Phi(u)]/\varphi(u)$ is *Mill*'s ratio, which appears because of the formulas,

$$E\frac{u}{\chi_1^2+u^2} = M(u),$$

$$E\frac{2u^2}{(\chi_1^2+u^2)^2} = 1+\frac{(1-u^2)M(u)}{u}.$$

The first well-known identity is a consequence of the Parceval theorem and can be found in Erdelyi et al. (1953, Sec. 9.3 (3)). The second follows from the first one by differentiation in u.

For $\tau^2 = 0, u = 1$,

$$R(\tilde{\tau}^2_{(\alpha 0)}, 0) = 1 - \frac{2}{\alpha} - \frac{1}{2\alpha^2} + \frac{5\alpha+1}{2\alpha^{5/2}} M\left(\frac{1}{\sqrt{\alpha}}\right),$$

which is an increasing function of α. Thus as for the quadratic loss, smaller values of α are preferable to keep the risk at the origin small. When $\alpha < 0.567\ldots, R(\tilde{\tau}^2_{(\alpha\,0)}, 0) < R(\tilde{\tau}^2_{(11)}, 0) = 4[\varphi(1) - 1 + \Phi(1)] = 0.333\ldots$.

An explicit expression through functions Φ and φ can be also obtained when $\beta = 1/2$ by using the formulas,

$$E\frac{a\mathbf{1}_{\{\chi_1^2 > a^2\}}}{\chi_1^2 + a^2} = \frac{[1-\Phi(a)]^2}{\varphi(a)}, \quad a > 0,$$

$$E\frac{a^2\mathbf{1}_{\{\chi_1^2 > a^2\}}}{(\chi_1^2 + a^2)^2} = \frac{(1-a^2)[1-\Phi(a)]^2}{2a\varphi(a)} - \frac{\varphi(a)}{2a} + 1 - \Phi(a).$$

The first of these equalities follows from Erdelyi et al. (1953, Sec. 9.9 (15)). Their application shows that with $a^2 = u^2/(2\alpha)$,

$$R(\tilde{\tau}^2_{(\alpha 1/2)}, \tau^2) = \frac{\tau^4}{s^4}[2\Phi(a) - 1 - 2a\varphi(a)] + \left(2a - \frac{4a}{\alpha} + \frac{1}{2a\alpha^2}\right)\varphi(a)$$
$$+ \left\{2 - \frac{4}{\alpha} - \frac{1}{\alpha^2} + \frac{[(4\alpha+1)a^2 + 1]M(a)}{2\alpha^2 a}\right\}[1 - \Phi(a)]. \quad (13)$$

When $\tau^2 \to \infty, u \to 0$, for $0 \le \beta < 1$,

$$R(\tilde{\tau}^2_{(\alpha\beta)}, \tau^2) \approx \frac{2\gamma^3 u^3 \tau^4}{3\sqrt{2\pi}s^4} + E\frac{\chi_1^2 \mathbf{1}_{\{\chi_1^2 > (\gamma u)^2\}}}{[\alpha\chi_1^2 + (1-\beta)u^2]^2}$$
$$\approx \frac{\tau}{\sqrt{2\pi}\alpha^{3/2}s}\left[\frac{2\beta^{3/2}}{3} + \frac{1}{\sqrt{1-\beta}}\int_{\beta/(1-\beta)}^{\infty}\frac{\sqrt{t}dt}{(1+t)^2}\right]$$
$$= \frac{\tau}{\sqrt{2\pi}\alpha^{3/2}s}\left[\frac{2\beta^{3/2}}{3} + \sqrt{\beta} + \frac{\arcsin\sqrt{1-\beta}}{\sqrt{1-\beta}}\right].$$

Thus, there is no optimal choice of α for large τ^2: the larger α, the smaller is the risk of μ-estimator. For a fixed α, $\lim_{\tau^2\to\infty}\sqrt{2\pi}\alpha^{3/2}sR(\tilde{\tau}^2, \tau^2)/\tau$ as a function of β, $0 \le \beta \le 1$, is monotonically increasing from $\pi/2$ to $8/3$. Indeed for $\beta = 1$,

$$R(\tilde{\tau}^2_{(\alpha 1)}, \tau^2) \approx \frac{8\tau}{3\sqrt{2\pi}\alpha^{3/2}s}.$$

For an estimator $\tilde{\tau}^2_{(\alpha\beta)}$ to improve upon the restricted maximum likelihood estimator for large τ^2, one must have

$$\alpha^{3/2} \ge \frac{3}{8}\left[\frac{2\beta^{3/2}}{3} + \sqrt{\beta} + \frac{\arcsin\sqrt{1-\beta}}{\sqrt{1-\beta}}\right].$$

However the values of α and β satisfying this condition cannot give a smaller value of the risk at $\tau^2 = 0$. Thus there are no uniform improvements in the class (1) upon the DerSimonian–Laird estimator. This fact and the asymptotics of $R(\tilde{\tau}^2, \tau^2)$ for more general (e.g. Bayes) estimators, are discussed in the next section.

3.2. *Permissible estimators*

To simplify the expression for R-risk,

$$R(\tilde{\tau}^2, \tau^2) = (\tau^2 + s^2)E\frac{(x_2 - x_1)^2}{2}\left(\frac{1}{\tilde{\tau}^2 + s^2} - \frac{1}{\tau^2 + s^2}\right)^2,$$

we use integration by parts formula according to which

$$E\frac{vg(v)}{\tau^2 + s^2} = 2Evg'(v) + Eg(v),$$

$v = (x_2 - x_1)^2/2$, $v \sim (\tau^2 + s^2)\chi_1^2$. Thus if $g(v) = [\tilde{\tau}^2(v) + s^2]^{-1}$ is continuous and piecewise differentiable, one has

$$R(\tilde{\tau}^2, \tau^2) = 1 + (\tau^2 + s^2)Ev\left[g^2(v) - 4g'(v) - \frac{2g(v)}{v}\right].$$

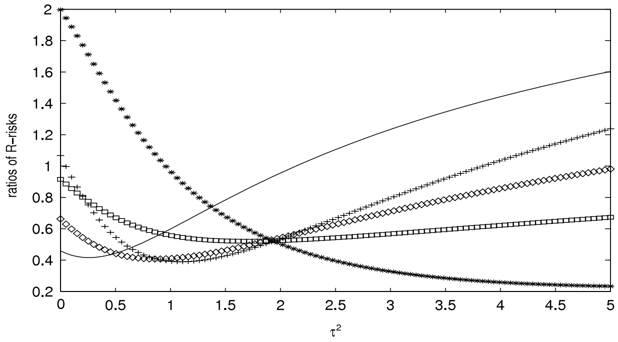

Fig. 2. Plots of ratios of risk functions $R(\tilde{\tau}^2, \tau^2)$ for $\tilde{\tau}^2_{(1/2\ 0)}$ (line marked by squares), $\tilde{\tau}^2_{(1/31/3)}$ (continuous line), $\tilde{\tau}^2_{(1/30)}$ (line marked by diamonds), $\tilde{\tau}^2_0$ (line marked by *), $\tilde{\tau}^2_1$ (line marked by +), to the risk of the DerSimonian–Laird estimator based on $\tilde{\tau}^2_{(1\ 1)}$.

We seek conditions under which the estimator $\tilde{x}$ cannot be improved in terms of the risk above, namely, when there is no estimator $\hat{x}$ with the corresponding function $h(v) = (\hat{\tau}^2 + s^2)^{-1}$ such that for all $v > 0$,

$$g^2(v) - 4g'(v) - \frac{2g(v)}{v} \geq h^2(v) - 4h'(v) - \frac{2h(v)}{v}$$

with a strict inequality for some v_0. Rukhin (1995) calls a function g *permissible* if this inequality does not have any continuous, piecewise differentiable solutions h. In our situation the class of positive functions h is restricted to those which are bounded by s^{-2}.

By putting $f(v) = h - g$, $|f| \leq s^{-2}$, one obtains a differential inequality,

$$f^2 + 2f\left(g - \frac{1}{v}\right) - 4f' \leq 0,$$

which is more conveniently written for $y = 1/f$ as

$$y' + \frac{y}{2}\left(g - \frac{1}{v}\right) + \frac{1}{4} \leq 0.$$

In our situation an estimator $\tilde{x}$ (or a function g) is permissible if for any v_0,

$$\int_{v_0}^{\infty} \exp\left\{-\frac{1}{2}\int_0^{v_1} g(v)dv\right\} \frac{dv_1}{\sqrt{v_1}} = \infty,$$

(cf. Ghosh and Sinha, 1981), and discussion in Section 5, (Rukhin, 1995). This shows that all estimators $\tilde{\tau}^2_{(\alpha\beta)}$ for $\alpha \geq 0, 0 \leq \beta \leq 1$ lead to permissible functions g. Thus it is difficult to find an explicit improvement over the DerSimonian–Laird estimator and other quadratic estimators with $\beta > 0$.

One has for $\tau^2 \to \infty$,

$$R(\tilde{\tau}^2, \tau^2) \sim (\tau^2 + s^2)E\frac{v}{(\tilde{\tau}^2 + s^2)^2} \sim \frac{\sqrt{2(\tau^2 + s^2)}}{\sqrt{\pi}s}\int_0^{\infty} \sqrt{v}g^2(v)dv.$$

This formula can be used to find behavior of $R(\tilde{\tau}^2, \tau^2)$ for the Bayes estimators $\tilde{\tau}^2_0$ and $\tilde{\tau}^2_1$ when τ^2 is large. In this case with ξ defined as in (7),

$$g(v) = \frac{\rho P(v/(2s^2), \rho + 1) + \xi[v/(2s^2)]^{\rho+1}e^{-v/(2s^2)}/s^2}{vP(v/(2s^2), \rho)/2 + \xi[v/(2s^2)]^{\rho+1}e^{-v/(2s^2)}}.$$

The integral, $\int g^2(v)v^{1/2}\, dy$, is an increasing function of ξ. Values of ρ smaller than 3/2 for large τ^2 give smaller values of $R(\tilde{\tau}^2, \tau^2)$, but larger risk $R(\tilde{\tau}^2, 0)$. Fig. 2 depicts the ratios of $R(\tilde{\tau}^2, \tau^2)$ for the estimators $\tilde{\tau}^2$ considered above to $R(\tilde{\tau}^2_{(1,1)}, \tau^2)$ (i.e., to the risk of the DerSimonian–Laird estimator). The Bayes μ-estimators based on $\tilde{\tau}^2_1$ and $\tilde{\tau}^2_2$ (not shown in Fig. 2) for large τ^2 demonstrate poor performance.

The explicit formulas (11)–(13) enable estimation of the quadratic risk of the corresponding μ-estimators, which is required in some applications.

Table 1
Summary of risk function values.

Estimator	$E(\tilde{\tau}/s)^4$	$\lim_{\tau^2\to\infty} E(\tilde{\tau}^2/\tau^2-1)^2$	$R(\tilde{\tau}^2, 0)$	$\lim_{\tau^2\to\infty} sR(\tilde{\tau}^2, \tau^2)/\tau$
$\tilde{\tau}^2_{(1\ 1)}$	1.602	2	0.333	1.064
$\tilde{\tau}^2_{(1/2\ 0)}$	0.750	0.750	0.314	1.772
$\tilde{\tau}^2_{(1/3\ 0)}$	0.334	0.667	0.221	3.256
$\tilde{\tau}^2_{(1/3\ 1/3)}$	0.178	0.667	0.155	3.888
$\tilde{\tau}^2_0$	0.667	0.667	0.665	3.628
$\tilde{\tau}^2_1$	0.193	0.667	0.142	4.411
$\tilde{\tau}^2_2$	0.101	0.667	0.068	5.013

3.3. Admissibility results

We discuss here some admissibility results referring to this concept understood within the class of all invariant procedures. The estimator $\bar{\tau}^2 = \infty$ has a constant risk, $R(\bar{\tau}^2, \tau^2) \equiv 1$, is admissible for this risk and is minimax which implies admissibility under the quadratic loss of the corresponding μ-estimator $\bar{x} = (x_1 + x_2)/2$. This fact can be proven by the Blyth method considering the Bayes estimators for the prior densities (5) when $\beta \to \infty$, (e.g. Lehmann and Casella, 1998, Ex 2.8, p 325.)

The Bayes estimator for the prior density $(\tau^2 + s^2)^{-3}\, d\tau^2$, i.e., when $\tilde{\tau}^2 = \tilde{\tau}^2_0$, is admissible. Indeed, $\tilde{\tau}^2_0$ is admissible for both risk functions: the quadratic in Section 2.2 and $R(\tilde{\tau}, \tau^2)$. It has finite Bayes risk in the second case, and in the first case its risk is well approximated by that of the Bayes rules against (5) with $\beta = 0$ and $\rho \downarrow 3/2$ (which have finite Bayes risks.) As a matter of fact, under $R(\tilde{\tau}, \tau^2)$ the densities (5) lead to admissible estimators when $\beta \geq 0$ and $\rho \geq 1$.

Another classical admissible procedure is the Graybill-Deal estimator, $\tilde{\mu}_{GD} = 0.5(s_2^2 x_1 + s_1^2 x_2)/s^2$, which corresponds to the prior distribution concentrated at $\tau^2 = 0$. Its admissibility in the setting with random $s_1^2, s_2^2, \tau^2 = 0$, remains an open problem despite a body of work (Sinha and Mouquadem, 1982; Kubokawa, 1987).

4. Conclusions

We summarize our main findings in the form of Table 1.

References

DerSimonian, R., Laird, N., 1986. Meta-analysis in clinical trials. Control. Clin. Trials 7, 177–188.
Erdelyi, A., et al., 1953. Higher Transcendental Functions, vol. 2. McGraw Hill Book Co., New York, NY.
Ghosh, J.K., Sinha, B.K., 1981. A necessary and sufficient condition for second order admissibility with applications to Berkson's bioassay problem. Ann. Statist. 9, 1334–1338.
Jackson, D., Bowden, J., Baker, R., 2010. How does the DerSimonian and Laird procedure for random effects meta-analysis compare with its more efficient but harder to compute counterparts? J. Statist. Plann. Inf. 140, 961–970.
Kubokawa, T., 1987. Admissible minimax estimation of a common mean of two normal populations. Ann. Statist. 15, 1245–1256.
Lehmann, E., Casella, G., 1998. Theory of Point Estimation, second ed. Springer, New York.
Rukhin, A.L., 1995. Admissibility: survey of a concept in progress. Int. Statist. Rev. 63, 95–115.
Sinha, B.K., Mouquadem, O., 1982. Estimation of a common mean of two univariate normal populations. Commun. Statist. A- Theory Methods 11, 1603–1614.