Some Corrections of the Significance Level in Meta - Analysis *

ANNETTE BÖCKENHOFF and JOACHIM HARTUNG

Department of Statistics, SFB 475
University of Dortmund, Germany

Summary

In many applications we obtain test statistics by combining estimates from different experiments or studies. The usual combined estimator of the overall effect in independent studies leads to systematic overestimates of the significance level, see LI et al. (1994). This results in a great number of unjustified significant evidences. By examination of the convexity of composed functions involved and application of higher and inverse moments of the chi-square distribution we propose corrections for the estimated standard deviation of the overall effect estimator. Analytical results and simulations show that we improve the estimated significance level in such models.

Key words: meta-analysis; significance level; combined estimator

Zusammenfassung

In vielen Bereichen werden Teststatistiken dadurch erhalten, daß Schätzer aus verschiedenen Experimenten oder Studien kombiniert werden. Der üblicherweise verwendete Schätzer des Gesamteffektes aus unabhängigen Studien führt zu systematischen Überschätzungen des Signifikanzniveaus, vgl. LI et al. (1994), und daher zu vielen ungerechtfertigten Signifikanzen. Mit Hilfe der Konvexität zusammengesetzter Funktionen und Anwendung höherer sowie inverser Momente der χ^2-Verteilung stellen wir Korrekturen für die geschätzte Standardabweichung des Gesamteffektschätzers vor. Analytische Ergebnisse und Simulationen zeigen, daß das nominelle Niveau durch diese Korrekturen besser eingehalten wird.

1. Introduction

The problem of estimating an overall effect from different independent experiments occurs in a variety of application fields. Not only in meta-analysis but also in analysing multicenter trials for instance we have different samples with heterogeneous error variances to assess the overall effect.

In such cases of k so called homogeneous (for the common mean μ) independent studies let y_{ij} be an observation in study i, $i = 1, \ldots, k$, at subject j from a normal distribution, $j = 1, \ldots, n_i$, with mean μ and variance $\mathrm{var}(y_{ij}) = \sigma_{e_i}^2$, i.e. $y_{ij} \sim \mathcal{N}(\mu, \sigma_{e_i}^2)$, $\sigma_{e_i}^2 > 0$, $i = 1, \ldots, k$. The best estimator for μ in each study is known as the individual sample mean $\bar{y}_{i.}$ with variance $\mathrm{var}(\bar{y}_{i.}) = \sigma_{e_i}^2 / n_i =: \sigma_i^2$, $i = 1, \ldots, k$.

---
*to appear in: *Biometrical Journal*

The standard procedure to estimate the common mean μ in the present case of meta-analysis is given by

$$\hat{\mu} = \sum_{i=1}^{k} \frac{\frac{1}{\hat{\sigma}_i^2}}{\sum_{i=1}^{k} \frac{1}{\hat{\sigma}_i^2}} \bar{y}_{i.} \quad . \tag{1.1}$$

The estimate of the associated variance which is commonly used for the computation of confidence intervals or tests can be represented by

$$\hat{\sigma}_{\hat{\mu}}^2 = \frac{1}{\sum_{i=1}^{k} \frac{1}{\hat{\sigma}_i^2}} \quad . \tag{1.2}$$

This is known as the variance weight method, see Cochran (1937), Whitehead and Whitehead (1991), Li et al. (1994). Here $\hat{\sigma}_i^2$ is given by

$$\hat{\sigma}_i^2 = \frac{1}{n_i} \cdot \frac{1}{n_i - 1} \cdot \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{i.})^2 = \frac{1}{n_i} \cdot \hat{\sigma}_{e_i}^2 \quad . \tag{1.3}$$

To carry out tests for the common mean μ we use the test statistic

$$Z := \frac{\hat{\mu}}{\sqrt{\hat{\sigma}_{\hat{\mu}}^2}} \quad , \tag{1.4}$$

which is approximately normal with mean μ and variance 1.

Due to the estimated weights we get too much significant results because the estimator (1.2) systematically underestimates the parameter $\sigma_{\hat{\mu}}^2 := 1/\sum_{i=1}^{k} \frac{1}{\sigma_i^2}$, see Li et al. (1994).

Table 1 and table 2 show the simulated actual significance levels (10000 runs) at nominal level of 5 % for $H_0 : \mu = 0$ with the test statistic (1.4) for some constellations of two or three samples.

Table 1

Simulated actual significance levels (10000 runs) at nominal level of $\alpha = 5\%$ for $k = 2$ and $H_0 : \mu = 0$ with the test statistic (1.4)

(n_1, n_2)	$(\sigma^2_{e_1}, \sigma^2_{e_2})$	$\hat{\alpha}(\%)$
(15,5)	(1,4)	9.2
	(4,1)	11.5
	(2,3)	9.1
	(1,1)	10.5
(12,8)	(1,4)	9.4
	(4,1)	10.1
	(2,3)	9.2
	(1,1)	9.1
(10,10)	(1,4)	8.1
	(2,4)	9.1
	(2,3)	10.2
	(1,1)	8.9

Table 2

Simulated actual significance levels (10000 runs) at nominal level of $\alpha = 5\%$ for $k = 3$ and $H_0 : \mu = 0$ with the test statistic (1.4)

(n_1, n_2, n_3)	$(\sigma^2_{e_1}, \sigma^2_{e_2}, \sigma^2_{e_3})$	$\hat{\alpha}(\%)$
(5,5,5)	(2,3,4)	19.2
(5,10,15)	(1,3,5)	13.8
	(5,3,1)	10.0
(10,10,10)	(1,3,5)	10.3
	(4,4,4)	10.7
(10,10,15)	(1,3,5)	10.5
	(5,3,1)	8.9

We assess some extremely high actual significance levels particularly in cases of small sample sizes, i.e. 19.2 % in the balanced case $n_1 = n_2 = n_3 = 5$ with variances $\sigma^2_{e_1} = 2$, $\sigma^2_{e_2} = 3$ and $\sigma^2_{e_3} = 4$.

Our intention in this article is to derive some suitable corrections of the test statistic and to receive consequently better approximations of the nominal significance level.

2. Theoretical methods

First we need some results of the examination for convexity in general cases, see HARTUNG (1976).

Definition 2.1.

(i) A function $f : \mathbb{R}^k \to \mathbb{R}^l$ is called convex iff
($\mathbf{x}, \mathbf{y} \in \mathbb{R}^k$, $\lambda \in [0,1] \implies f(\lambda\mathbf{x} + (1-\lambda)\mathbf{y}) \leq \lambda f(\mathbf{x}) + (1-\lambda)f(\mathbf{y})$)
with the natural semi-ordering, i.e. ordering by components.

(ii) A function $f : \mathbb{R}^k \to \mathbb{R}^l$ is called quasi-convex iff
($\mathbf{y} \in \mathbb{R}^l \implies \{\mathbf{x} \in \mathbb{R}^k \mid f(\mathbf{x}) \leq \mathbf{y}\}$ is convex).

(iii) A function f is called (quasi-)concave if $(-f)$ is (quasi-)convex.

For composed functions we obtain the following results:

Lemma 2.2. Let $f : \mathbb{R}^k \to \mathbb{R}^l$ be convex [concave] and $T : \mathbb{R}^l \to \mathbb{R}^m$ be (quasi-)convex [(quasi-)concave] and increasing by the natural semi-ordering in $\mathbb{R}^m$. Then $T \circ f$ is (quasi-)convex [(quasi-)concave].

Lemma 2.3. Let $f : \mathbb{R}^k \to \mathbb{R}^l$ be convex [concave] and $T : \mathbb{R}^l \to \mathbb{R}^m$ be (quasi-)concave [(quasi-)convex] and decreasing by the natural semi-ordering. Then $T \circ f$ is (quasi-)concave [(quasi-)convex].

Remark 2.4. Especially from lemma 2.3 we see that if y is a concave function, y^{-1} is convex because $(.)^{-1}$ is decreasing and convex.
Further if y is convex the function y^{-1} is quasi-concave because $(.)^{-1}$ is decreasing and quasi-concave.

Lemma 2.5. If $f : \mathbb{R}^l_+ \to \mathbb{R}_+$ is quasi-convex [quasi-concave] and $f(\lambda\mathbf{x}) = \lambda f(\mathbf{x})$, $\lambda > 0$, $\mathbf{x} \neq \mathbf{0}$, (positively homogeneous), then f is also convex [concave].

For further considerations we need the well known Jensen Inequality.

Jensen Inequality 2.6. For a random variable $\hat{\theta}$ we have $\mathrm{E}f(\hat{\theta}) \geq f(\mathrm{E}\hat{\theta})$ if f is convex and $\mathrm{E}g(\hat{\theta}) \leq g(\mathrm{E}\hat{\theta})$ if g is concave.

Now we use the above considerations to receive results for combined functions.

Lemma 2.7.

(i) $h_1 := \sum_{i=1}^{k} \frac{1}{\sqrt{\sigma_i^4}} = \sum_{i=1}^{k} \frac{1}{\sigma_i^2}$ is convex in σ_i, σ_i^2 and σ_i^4.

(ii) $h_2 := \left(\sum_{i=1}^{k} \frac{1}{\sqrt{\sigma_i^4}}\right)^2 = \left(\sum_{i=1}^{k} \frac{1}{\sigma_i^2}\right)^2$ is convex in σ_i, σ_i^2 and σ_i^4.

(iii) $h_3 := \sigma_{\hat{\mu}}^2 = \left(\sum_{i=1}^{k} \frac{1}{\sqrt{\sigma_i^4}}\right)^{-1} = \left(\sum_{i=1}^{k} \frac{1}{\sigma_i^2}\right)^{-1}$ is quasi-concave in σ_i, σ_i^2 and σ_i^4.

(iv) $h_3 := \sigma_{\hat{\mu}}^2 = \left(\sum_{i=1}^{k} \frac{1}{\sigma_i^2}\right)^{-1}$ is concave in σ_i^2.

(v) $h_4 := \sigma_{\hat{\mu}} = \left(\sum_{i=1}^{k} \frac{1}{\sigma_i^2}\right)^{-1/2}$ is concave in σ_i.

(vi) $h_4 := \sigma_{\hat{\mu}}$ is concave in σ_i^2.

(vii) $h_5 := (\sigma_{\hat{\mu}}^2)^2 = \left(\sum_{i=1}^{k} \frac{1}{\sqrt{\sigma_i^4}}\right)^{-2}$ is concave in σ_i^4.

(viii) $h_4 := \sigma_{\hat{\mu}} = \left(\sum_{i=1}^{k} \frac{1}{\sigma_i^2}\right)^{-1/2}$ is convex in $\frac{1}{\sigma_i^2}$.

(ix) $h_3 := \sigma_{\hat{\mu}}^2 = \left(\sum_{i=1}^{k} \frac{1}{\sigma_i^2}\right)^{-1}$ is convex in $\frac{1}{\sigma_i^2}$.

Proof.

(i): Because $\frac{1}{x}$ is convex, $(.)^2$ is increasing and convex, we get with lemma 2.2 that $\frac{1}{x^2}$ is convex and therefore we have the convexity of h_1 in σ_i^2 and σ_i.
Now $\sqrt{x}$ is concave and with remark 2.4 one sees that $\frac{1}{\sqrt{x}}$ is convex. Therefore we have the convexity of h_1 in σ_i^4.

(ii): The function $(.)^2$ is increasing and convex. We conclude the statement 2.7 (ii) by lemma 2.2 and 2.7 (i).

(iii): With $h_3 = \frac{1}{h_1}$ we have the quasi-concavity with 2.7 (i) and remark 2.4.

(iv): The function h_3 is concave in σ_i^2 because of 2.7 (iii) and the positive homogeneity of h_3 in σ_i^2.

(v): On account of 2.7 (iii) we conclude that $h_4 = \sqrt{h_3}$ is quasi-concave in σ_i. With lemma 2.5 and the positive homogeneity of h_4 in σ_i we have the concavity of h_4 in σ_i.

(vi): With 2.7 (iv) h_3 is concave in σ_i^2. Because $\sqrt{(.)}$ is increasing and concave we see with lemma 2.2 that $h_4 = \sqrt{h_3}$ is concave in σ_i^2.

(vii): The function $h_5 = \dfrac{1}{h_2}$ is quasi-concave in σ_i^4 because of 2.7 (ii) and remark 2.4. h_5 is even concave in σ_i^4 due to the positive homogeneity, see lemma 2.5.

(viii): Due to $\sqrt{\sum_{i=1}^{k} \dfrac{1}{\sigma_i^2}}$ is concave in $\dfrac{1}{\sigma_i^2}$ we conclude with remark 2.4 the convexity of $h_4 = \left(\sqrt{\sum_{i=1}^{k} \dfrac{1}{\sigma_i^2}}\right)^{-1}$ in $\dfrac{1}{\sigma_i^2}$, see remark 2.4.

(ix): The function $(.)^2$ is increasing and convex. Therefore $h_3 = h_4^2$ is convex in $\dfrac{1}{\sigma_i^2}$ with 2.7 (viii) and lemma 2.2.

Further we need the following moments, see Patel et al. (1976), making use of the $\chi^2_{n_i-1}$-distribution of $(n_i-1)\hat{\sigma}_i^2/\sigma_i^2$ from (1.3), $i = 1, \ldots, k$:

$$
\begin{aligned}
\mathrm{E}\left(\sqrt{\hat{\sigma}_i^2}\right) &= \frac{1}{\gamma_{n_i-1}} \cdot \sigma_i \quad \text{with} \quad \gamma_{n_i-1} := \sqrt{\frac{n_i-1}{2}} \cdot \frac{\Gamma(\frac{n_i-1}{2})}{\Gamma(\frac{n_i}{2})} > 1\,, && (2.1)\\
\mathrm{E}(\hat{\sigma}_i^4) &= b_{n_i} \cdot \sigma_i^4 \quad \text{with} \quad b_{n_i} := \frac{n_i+1}{n_i-1} > 1\,, && (2.2)\\
\mathrm{E}\left(\frac{1}{\hat{\sigma}_i^2}\right) &= c_{n_i} \cdot \frac{1}{\sigma_i^2} \quad \text{with} \quad c_{n_i} := \frac{n_i-1}{n_i-3} > 1\,, && (2.3)
\end{aligned}
$$

for $i = 1, \ldots, k$.

3. Upper estimates of the variance and standard deviation

The results of section 2 now serve for the construction of useful upper limits for the standard deviation $\sigma_{\hat{\mu}}$.

Theorem 3.1. For the estimator $\hat{\sigma}^2_{\hat{\mu}}$ in (1.2) we have

$$\operatorname{var}(\hat{\sigma}^2_{\hat{\mu}}) \leq \mathrm{E}\left\{ \left(\sum_{i=1}^{k} \frac{1}{\hat{\sigma}_i^2} \right)^{-2} - \left(\sum_{i=1}^{k} \frac{\sqrt{n_i^2-1}}{n_i-3} \cdot \frac{1}{\hat{\sigma}_i^2} \right)^{-2} \right\} =: \mathrm{E}(\hat{\vartheta}) \quad .$$

Proof. From the definition of variance we have

$$\begin{aligned} \operatorname{var}(\hat{\sigma}^2_{\hat{\mu}}) &= \mathrm{E}(\hat{\sigma}^2_{\hat{\mu}})^2 - \left(\mathrm{E}\left(\sum_{i=1}^{k} \frac{1}{\hat{\sigma}_i^2} \right)^{-1} \right)^2 \\ &\leq \mathrm{E}(\hat{\sigma}^2_{\hat{\mu}})^2 - \left(\sum_{i=1}^{k} \mathrm{E}\left(\frac{1}{\hat{\sigma}_i^2} \right) \right)^{-2} \end{aligned}$$

after application of 2.7 (ix) and the Jensen Inequality 2.6.

With (2.3) and (2.2) it follows

$$\begin{aligned} \mathrm{E}(\hat{\sigma}^2_{\hat{\mu}})^2 - \left(\sum_{i=1}^{k} \mathrm{E}\left(\frac{1}{\hat{\sigma}_i^2} \right) \right)^{-2} &= \mathrm{E}(\hat{\sigma}^2_{\hat{\mu}})^2 - \left(\sum_{i=1}^{k} c_{n_i} \cdot \frac{1}{\sigma_i^2} \right)^{-2} \qquad (3.1.1) \\ &= \mathrm{E}(\hat{\sigma}^2_{\hat{\mu}})^2 - \left(\sum_{i=1}^{k} c_{n_i} \cdot \sqrt{\frac{1}{\frac{1}{b_{n_i}} \cdot \mathrm{E}\hat{\sigma}_i^4}} \right)^{-2} \\ &\leq \mathrm{E}(\hat{\sigma}^2_{\hat{\mu}})^2 - \mathrm{E}\left(\sum_{i=1}^{k} c_{n_i} \cdot \sqrt{b_{n_i}} \cdot \frac{1}{\hat{\sigma}_i^2} \right)^{-2} \\ &= \mathrm{E}\left\{ \left(\sum_{i=1}^{k} \frac{1}{\hat{\sigma}_i^2} \right)^{-2} - \left(\sum_{i=1}^{k} \frac{\sqrt{n_i^2-1}}{n_i-3} \cdot \frac{1}{\hat{\sigma}_i^2} \right)^{-2} \right\} \quad , \end{aligned}$$

where the inequality results from the fact that $\left(\sum_{i=1}^{k} \frac{1}{\hat{\sigma}_i^2} \right)^{-2}$ is concave in $\hat{\sigma}_i^4$ (confer 2.7 (vii)) and the Jensen Inequality 2.6. This completes the proof.

Theorem 3.2. An upper limit for the variance of $\hat{\sigma}^2_{\hat{\mu}}$ is given by

$$\operatorname{var}(\hat{\sigma}^2_{\hat{\mu}}) \leq \left(\sum_{i=1}^{k} \sqrt{\frac{n_i-1}{n_i+1}} \cdot \frac{1}{\sigma_i^2} \right)^{-2} - \left(\sum_{i=1}^{k} \frac{n_i-1}{n_i-3} \cdot \frac{1}{\sigma_i^2} \right)^{-2} \quad .$$

Proof. With (3.1.1) we conclude

$$\begin{aligned}
\mathrm{var}(\hat\sigma^2_{\hat\mu}) &\leq \mathrm{E}(\hat\sigma^2_{\hat\mu})^2 - \left(\sum_{i=1}^k c_{n_i}\cdot\frac{1}{\sigma_i^2}\right)^{-2}\\
&= \mathrm{E}\left(\sum_{i=1}^k \frac{1}{\sqrt{\hat\sigma_i^4}}\right)^{-2} - \left(\sum_{i=1}^k c_{n_i}\cdot\frac{1}{\sigma_i^2}\right)^{-2}\\
&\leq \left(\sum_{i=1}^k \frac{1}{\sqrt{b_{n_i}\cdot\sigma_i^4}}\right)^{-2} - \left(\sum_{i=1}^k c_{n_i}\cdot\frac{1}{\sigma_i^2}\right)^{-2}\\
&= \left(\sum_{i=1}^k \sqrt{\frac{n_i-1}{n_i+1}}\cdot\frac{1}{\sigma_i^2}\right)^{-2} - \left(\sum_{i=1}^k \frac{n_i-1}{n_i-3}\cdot\frac{1}{\sigma_i^2}\right)^{-2}
\end{aligned}$$

because $(\hat\sigma^2_{\hat\mu})^2$ is concave in $\hat\sigma_i^4$, see 2.7 (vii), and application of (2.2).

Note that both limit functions from theorem 3.1 and 3.2 are in no relation to each other.

Theorem 3.3. We have the following interrelations between the estimates and the parameter
$\sigma_{\hat\mu} = \left(\sum_{i=1}^k \frac{1}{\sigma_i^2}\right)^{-1/2}$:

$$\mathrm{E}(\hat\sigma_{\hat\mu}) \leq \sigma_{\hat\mu} \leq \mathrm{E}\left(\sum_{i=1}^k \frac{1}{c_{n_i}}\cdot\frac{1}{\hat\sigma_i^2}\right)^{-1/2} ,$$

where $\hat\sigma_{\hat\mu} = \sqrt{\hat\sigma^2_{\hat\mu}}$, see (1.2), and $c_{n_i} = \frac{n_i-1}{n_i-3}$, $i = 1, \dots, k$.

Proof. For the standard deviation we have

$$\begin{aligned}
\mathrm{E}(\hat\sigma_{\hat\mu}) &= \mathrm{E}\left(\sum_{i=1}^k \frac{1}{\hat\sigma_i^2}\right)^{-1/2}\\
&\leq \mathrm{E}\left(\sum_{i=1}^k \frac{1}{\gamma^2_{n_i-1}\cdot\hat\sigma_i^2}\right)^{-1/2}\\
&\leq \left(\sum_{i=1}^k \frac{1}{\gamma^2_{n_i-1}\cdot(\mathrm{E}\sqrt{\hat\sigma_i^2})^2}\right)^{-1/2}
\end{aligned}$$

because $1 < \gamma^2_{n_i-1}$, see (2.1), and $\hat\sigma_{\hat\mu}$ is concave in $\hat\sigma_i$, see 2.7 (v).

Now with (2.1) and (2.3) we conclude

$$\begin{aligned}
\left(\sum_{i=1}^{k} \frac{1}{\gamma_{n_i-1}^2 \cdot (\mathrm{E}\sqrt{\hat{\sigma}_i^2})^2}\right)^{-1/2} &= \sigma_{\hat{\mu}} \\
&= \left(\sum_{i=1}^{k} \frac{1}{c_{n_i}} \cdot c_{n_i} \cdot \frac{1}{\sigma_i^2}\right)^{-1/2} \\
&= \left(\sum_{i=1}^{k} \frac{1}{c_{n_i}} \cdot \mathrm{E}\left(\frac{1}{\hat{\sigma}_i^2}\right)\right)^{-1/2} \\
&\leq \mathrm{E}\left(\sum_{i=1}^{k} \frac{1}{c_{n_i}} \cdot \frac{1}{\hat{\sigma}_i^2}\right)^{-1/2}
\end{aligned}$$

because $\hat{\sigma}_{\hat{\mu}}$ is convex in $\frac{1}{\hat{\sigma}_i^2}$, see 2.7 (viii).

Definition 3.4. Let us now define the estimators

$$\begin{aligned}
\hat{\sigma}_{\hat{\mu}}(\mathbf{a}) &:= \left(\sum_{i=1}^{k} \frac{1}{a_i \cdot \hat{\sigma}_i^2}\right)^{-1/2} \qquad , \mathbf{a} := (a_1, \ldots, a_k)^T,\ a_i > 1,\ i = 1, \ldots, k, \\
\hat{\sigma}_{\hat{\mu}}(\hat{\vartheta}, \kappa) &:= \left(\hat{\sigma}_{\hat{\mu}}^2 + \kappa \cdot \sqrt{\hat{\vartheta}}\right)^{1/2} \geq \hat{\sigma}_{\hat{\mu}}
\end{aligned}$$

with $\hat{\vartheta}$ defined in theorem 3.1, $\kappa > 0$.

We state the following corollary to see further interrelations of estimates and parameters which are implied by the theorem above.

Corollary 3.5. With theorem 3.3, the proof of theorem 3.3 and definition 3.4 we have

$$\begin{aligned}
\left(\sum_{i=1}^{k} c_{n_i} \cdot \frac{1}{\sigma_i^2}\right)^{-1/2} &\leq \mathrm{E}(\hat{\sigma}_{\hat{\mu}}) \\
&\leq \mathrm{E}(\hat{\sigma}_{\hat{\mu}}(\boldsymbol{\gamma}^2)) \\
&\leq \sigma_{\hat{\mu}} \\
&\leq \mathrm{E}(\hat{\sigma}_{\hat{\mu}}(\mathbf{c}))
\end{aligned}$$

with $\boldsymbol{\gamma}^2 := (\gamma_{n_1-1}^2, \ldots, \gamma_{n_k-1}^2)^T$ and $\mathbf{c} := (c_{n_1}, \ldots, c_{n_k})^T$.

4. Simulation results

Now we demonstrate by simulations the actual significance levels of the defined estimators in some constellations of three or five samples, see table 3 and table 4.

We have to note that $\sigma_{\hat{\mu}}$ is usually unknown and we have to replace it by estimators to receive test statistics like (1.4). It is remarkable that the levels of $\hat{\sigma}_{\hat{\mu}}(\mathbf{c})$ are in the same order of magnitude like the levels achieved whith the theoretical value $\sigma_{\hat{\mu}}$, while results with the usual estimator $\hat{\sigma}_{\hat{\mu}}$ extremely overestimate the significance level, see table 3, table 4 and also table 1. The improvements achieved with $\hat{\sigma}_{\hat{\mu}}(\mathbf{c})$ are consistent and convincing compared with the usual estimator $\hat{\sigma}_{\hat{\mu}}$.

The estimator $\hat{\sigma}_{\hat{\mu}}(\hat{\vartheta}, 0.5)$, with $\kappa = 0.5$, results in some further improvement towards the nominal level of 5% in moderate sample sizes for three or five samples, see table 3 and table 4. For further improvement of the significance level κ can be varied.

For large sample sizes like $n_i \geq 20$, $i = 1, \ldots, k$, we recommend the use of $\hat{\sigma}_{\hat{\mu}}(\mathbf{c})$ in the test statistic to avoid the conservatism of $\hat{\sigma}_{\hat{\mu}}(\hat{\vartheta}, 0.5)$ in such sample sizes, regardless of the variance pattern.

When there are very small sample sizes like $n_i = 5$, $i = 1, \ldots, k$, the estimator $\hat{\sigma}_{\hat{\mu}}(\boldsymbol{c\gamma}^3)$ with $\boldsymbol{c\gamma}^3 := (c_{n_1}\gamma^3_{n_1-1}, \ldots, c_{n_k}\gamma^3_{n_k-1})^T$ leads to acceptable actual levels in our simulation study.

Table 3

Simulated actual significance levels (10000 runs) at a nominal level of $\alpha = 5\%$ for $k = 3$ and $H_0 : \mu = 0$ with test statistics like (1.4) using the theoretical value $\sigma_{\hat{\mu}}$ and different estimators of it

test statistic $Z = \hat{\mu}/x$, nominal level $\alpha = 5\%$							
		x					
(n_1, n_2, n_3)	$(\sigma^2_{e_1}, \sigma^2_{e_2}, \sigma^2_{e_3})$	$\sigma_{\hat{\mu}}$	$\hat{\sigma}_{\hat{\mu}}$	$\hat{\sigma}_{\hat{\mu}}(\boldsymbol{\gamma}^2)$	$\hat{\sigma}_{\hat{\mu}}(\mathbf{c})$	$\hat{\sigma}_{\hat{\mu}}(\hat{\vartheta}, 0.5)$	$\hat{\sigma}_{\hat{\mu}}(\boldsymbol{c\gamma}^3)$
(5,5,5)	(2,3,4)	8.6	19.2	16.6	8.4	12.5	6.5
(5,10,15)	(1,3,5)	7.6	13.8	12.5	7.6	8.6	6.2
	(5,3,1)	7.0	10.0	9.4	7.1	6.1	6.3
(10,10,10)	(1,3,5)	6.9	10.3	9.6	7.0	6.3	6.0
	(4,4,4)	7.0	10.7	9.9	7.4	6.9	6.5
(10,10,15)	(1,3,5)	6.7	10.5	9.5	6.9	6.1	6.0
	(5,3,1)	6.6	8.9	8.3	6.5	5.4	5.8
(20,30,40)	(1,3,5)	5.3	6.4	6.2	5.3	4.0	4.9
	(5,3,1)	5.1	5.8	5.6	5.0	3.8	4.7
(40,50,60)	(1,3,5)	5.4	6.2	6.1	5.5	4.2	5.3
	(5,3,1)	5.3	5.8	5.6	5.4	4.2	5.3
(60,70,80)	(1,3,5)	5.2	5.5	5.5	5.2	4.1	5.1
	(5,3,1)	5.0	5.5	5.5	5.1	4.0	5.0

Table 4

Simulated actual significance levels (10000 runs) at a nominal level of $\alpha = 5\%$ for $k = 5$ and $H_0 : \mu = 0$ with test statistics like (1.4) using the theoretical value $\sigma_{\hat{\mu}}$ and different estimators of it

test statistic $Z = \hat{\mu}/x$, nominal level $\alpha = 5\%$								
		x						
$(n_1, n_2, n_3, n_4, n_5)$	$(\sigma^2_{e_1}, \sigma^2_{e_2}, \sigma^2_{e_3}, \sigma^2_{e_4}, \sigma^2_{e_5})$	$\sigma_{\hat{\mu}}$	$\hat{\sigma}_{\hat{\mu}}$	$\hat{\sigma}_{\hat{\mu}}(\boldsymbol{\gamma}^2)$	$\hat{\sigma}_{\hat{\mu}}(\mathbf{c})$	$\hat{\sigma}_{\hat{\mu}}(\hat{\vartheta}, 0.5)$	$\hat{\sigma}_{\hat{\mu}}(\boldsymbol{c\gamma}^3)$
(5,5,5,5,5)	(1,1,1,1,1)	11.3	22.1	19.9	10.6	15.3	8.3
	(9,9,4,4,1)	10.4	20.9	18.6	9.6	14.2	7.5
(10,10,10,10,10)	(1,1,1,1,1)	7.3	11.9	11.0	7.8	7.1	6.6
	(9,9,4,4,1)	6.7	10.0	9.0	6.6	6.0	5.6
(6,8,10,12,14)	(1,1,1,1,1)	8.1	12.6	11.6	8.3	7.7	7.1
	(9,9,4,4,1)	6.7	9.9	9.1	6.7	5.7	5.8
	(1,4,4,9,9)	7.5	13.0	11.8	7.8	8.3	6.5
(15,15,15,15,15)	(1,1,1,1,1)	6.4	8.7	8.1	6.5	5.1	5.7
	(9,9,4,4,1)	6.4	8.7	8.1	6.6	5.3	6.0
(6,10,14,18,22)	(1,1,1,1,1)	7.0	9.7	9.1	7.2	6.0	6.5
	(9,9,4,4,1)	6.1	8.0	7.6	6.2	4.7	5.7
	(1,4,4,9,9)	7.1	11.8	10.6	7.3	7.0	6.2
	(1,4,4,1,1)	6.8	9.5	8.9	6.9	5.6	6.2
(20,25,30,40,50)	(1,1,1,1,1)	5.3	6.0	5.8	5.3	4.1	5.1
	(9,9,4,4,1)	5.1	5.8	5.7	5.3	4.0	5.1
	(1,4,4,9,9)	5.5	6.6	6.3	5.7	4.3	5.3
	(81,81,16,16,1)	5.0	5.4	5.3	5.0	3.9	4.8
	(1,16,16,81,81)	5.2	6.4	6.0	5.2	4.0	4.7
(40,45,50,60,70)	(1,1,1,1,1)	5.1	5.7	5.6	5.1	3.8	5.0
	(9,9,4,4,1)	5.1	5.4	5.2	4.9	4.0	4.8
	(1,4,4,9,9)	5.1	5.8	5.6	5.1	3.8	4.9
(60,65,70,75,80)	(1,1,1,1,1)	5.4	5.8	5.7	5.4	4.2	5.3
	(9,9,4,4,1)	5.3	5.9	5.8	5.5	4.3	5.4
	(1,4,4,9,9)	5.7	6.1	6.1	5.7	4.6	5.6

There is no obvious reason to see why the test statistics are nearly insensitive with respect to the error variances, see the 6th example with higher sample sizes and increased variances in table 4.

It is worth to note that with increasing sample sizes one observes a stabilization of the actual significance levels at all estimators and no growing conservatism.

5. Final remark

In this paper we have shown the consequences of the estimated weights in the test statistic for tests about the common effect in combining estimates from independent studies or experiments. We recommend to use the proposed corrections for the standard deviation to achieve better approximations of the nominal significance level.

The next step in this direction would be to try an extension of the considered methods to the case of random effects models where with the standard procedures we observe the same deficiencies as in the fixed effects models considered above. But more problems will arise in estimating the involved parameters as well as distributional problems with the test statistics.

Acknowledgement

We would like to thank a referee of the *Biometrical Journal* for valuable comments.

References

Cochran, W.G., 1937: Problems arising in the analysis of a series of similar experiments. *Journal of the Royal Statistical Society B* **4**, 102-118.

Hartung, J., 1976: Some inequalities for a random function. *Teorija Verojatnostei i ee Primenenja* **21**, 661-665 (and: *Theory of Probability and its Applications*, SIAM **21**, 1977).

Li, Y., Shi, L. and Roth, H.D., 1994: The bias of the commonly-used estimate of variance in meta-analysis. *Communications in Statistics - Theory and Methods* **23**, 1063-1085.

Patel, J.K., Kapadia, C.P. and Owen, D.B., 1976: *Handbook of Statistical Distributions.* Marcel Dekker, New York.

Whitehead, A. and Whitehead, J., 1991: A general parametric approach to meta-analysis of randomised clinical trials. *Statistics in Medicine* **10**, 1665-1677.

Annette Böckenhoff
Joachim Hartung
Department of Statistics
University of Dortmund
Vogelpothsweg 87
D-44221 Dortmund
Germany