DEVELOPMENT OF VARIANCE FORMULAE FOR META-ANALYSES

DEVELOPMENT OF VARIANCE FORMULAE FOR OPTIMALLY WEIGHTED STUDIES IN META-ANALYSES WITH CONTINUOUS OUTCOMES

By HYUNA SEO, BHSC

A Thesis Submitted to the School of Graduate Studies
in Partial Fulfillment of the Requirements for
the Degree Master of Science

McMaster University

MASTER OF SCIENCE (2024)

Hamilton, Ontario, Canada (Mathematics and Statistics)

TITLE:	Development of Variance Formulae for Optimally Weighted Studies in Meta-Analyses with Continuous Outcomes
AUTHOR:	Hyuna Seo BHSc
SUPERVISOR:	Narayanaswamy Balakrishnan, PhD Stephen Walter, PhD
NUMBER OF PAGES:	vii, 36

Abstract

A meta-analysis provides a convenient way to integrate findings from multiple studies. The conventional methods of conducting a meta-analysis use inverse sample variance as weights, which are biased. However, this bias can easily be remedied using a multiplicative correction factor under a fixed-effects model, when the outcome is continuous and the treatment groups share a common variance. To investigate the effects of the bias-correction, Taylor series approximation is used to derive new estimators for the variance of the summary treatment effect. Results obtained from a simulation study show that the Taylor-approximated estimators return superior coverage with near-maximum precision. The bias-correction leads to increased coverage in some cases, although the results are inconclusive. The conventional inverse sum-of-weights estimator for the summary effect variance always underestimates the variance, decreasing the coverage. The work here demonstrates how the bias-correction impacts the precision of the overall treatment effect estimate and provides improved estimators for the variance, with which confidence intervals can be constructed, for example.

사랑하는 가족들을 위해

Acknowledgements

First and foremost I thank my supervisors, Drs. Narayanaswamy Balakrishnan and Stephen Walter, for their guidance and encouragement. It has been nothing but the utmost privilege to have worked with and learned from you both.

I thank my family for their unwavering love and support. All of my achievements are yours just as much, and I am forever indebted to you for all the things you made possible for me. I know I could never have done this without you. I love you, more than you think. Also, infinite kisses for Strawberry.

I am beyond grateful for my friends, thank you all so much for your company through this journey. You are my community and my chosen family. Special thanks to Fei, Lauren, and Daniel– you are the pillars of my daily life. There is no limit to the distance my heart will stretch to reach you guys. Friends I see around the beloved Hamilton Hall: I hope you know that you make my good days better and my soul-crushing days bearable. It means a lot to me that I have such amazing people I can talk to about the joys and struggles of being a graduate student (among other things!).

I would also like to thank my supervisors at Women's College Hospital for their support and accommodation. You inspire me daily to continue working for the things I believe in. You are where I learn what school can't teach me.

Table of Contents

List of Tables

Chapter 1

Introduction

A meta-analysis is a statistical method by which results from multiple studies with similar protocols can be condensed, usually into one summary statistic. Compared to a single study, meta-analyses can provide more accurate results with a greater statistical power. These results are often used to help answer important questions informing policy, evidence-based research, and practice. The celebrated statistician Karl Pearson was among the first to implement meta-analyses in the early 20$^{\text{th}}$ century, and it was later popularized by the American statistician Gene Glass in the 70s who coined the term [7, 19]. Ever since, meta-analyses have been widely used in many disciplines such as education, psychology, business, criminology, ecology, evidence-based medicine, and healthcare [2].

Meta-analyses with normally distributed continuous outcomes often use raw mean-difference (MD) for its interpretability and convenience [2]. Say a researcher has identified k studies from the same population with the same mean that satisfy her eligibility criteria and has decided to use a fixed-effects model. The studies report the effect of a new drug A on blood pressure on the standard scale, in mmHg. In other words, all studies report the outcome measures on the same scale (or different measures that can be easily converted into a common scale). Each study reports the mean blood pressure of the placebo/control group ($\overline{Y}_2$) and of the drug A group ($\overline{Y}_1$). The raw mean difference is then $\overline{Y}_1 - \overline{Y}_2$, and the researcher would have k mean differences. To summarize these k measures into one statistic, weights (w_i) are given to the studies to calculate the weighted mean $\Sigma\, w_i \mathrm{MD}_i / \Sigma\, w_i$. Historically, guidelines, handbooks, and textbooks have recommended the inverse variance weights $w_i = 1/\mathrm{Var}\,(\mathrm{MD}_i)$ [1–3, 11, 12] as it is shown to maximize the precision of the summary statistic [20]. In this thesis, it is assumed that the treatment groups share a common variance, thus the natural estimator for $\mathrm{Var}(Y)$ is the pooled sample variance

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}.$$

As such, the conventional study weights are given by

$$\hat{w}_i = \frac{1}{\widehat{\mathrm{Var}\,(\mathrm{MD}_i)}} = \frac{1}{s_{p,i}^2 \left(\dfrac{1}{n_{1i}} + \dfrac{1}{n_{2i}} \right)}.$$

Walter and Balakrishnan [20] show, however, that this $\hat{w}_i$ is a biased estimator of the intended true weight. The reciprocal operation is nonlinear, which results in

$$\mathbb{E}\left[\frac{1}{s_{p,i}^2 \left(\dfrac{1}{n_{1i}} + \dfrac{1}{n_{2i}} \right)} \right] \neq \frac{1}{\mathrm{Var}\,(\mathrm{MD}_i)}$$

despite $\mathbb{E}\left[s_{p,i}^2 \cdot (1/n_{1i} + 1/n_{2i}) \right] = \mathrm{Var}\,(\mathrm{MD}_i)$. It can be shown that $\hat{w}_i$ has a multiplicative bias with factor $1{+}2/(n_1 + n_2 - 4)$, where n_1 and n_2 are the sample sizes of the two groups [20]. Without correction, study weights are always inflated as this bias factor is always greater than 1. Since the bias factor increases as $n_1 + n_2$ decreases, this overestimation is worse with smaller studies. In other words, without the bias-correction, the relative contribution of smaller studies to the summary statistic is too large. Such disproportionate overestimation potentially introduces noise from the smaller studies into the overall treatment effect estimate, while inhibiting bigger studies from contributing their due amount.

An easy solution to obtaining an unbiased estimator for w_i is to divide the estimator by the multiplicative bias factor [20].

$$\begin{aligned}
\text{Let } \hat{w}_i* &= \frac{1}{1 + \dfrac{2}{n_{1i} + n_{2i} - 4}} \cdot \frac{1}{\text{Var}\,\widehat{(\text{MD}_i)}} \\
\Rightarrow \mathbb{E}\left[\hat{w}_i*\right] &= \frac{1}{1 + \dfrac{2}{n_{1i} + n_{2i} - 4}} \cdot \mathbb{E}\left[\frac{1}{\text{Var}\,\widehat{(\text{MD}_i)}}\right] \\
&= \frac{1}{1 + \dfrac{2}{n_{1i} + n_{2i} - 4}} \cdot \left(1 + \frac{2}{n_{1i} + n_{2i} - 4}\right) \cdot \frac{1}{\text{Var}\,(\text{MD}_i)} \\
&= \frac{1}{\text{Var}\,(\text{MD}_i)}.
\end{aligned}$$

Walter and Balakrishnan [20] showed that this correction usually results in a small change in the overall treatment effect estimates. The biggest changes were observed when the meta-analysis included a study with a small sample size– particularly when there were few other studies, or when the small study reported treatment effect estimates far from the center [20]. In addition to the changes in the point estimate, the variance is also affected by the bias-correction. With the new corrected weights, the variance of the summary treatment effect estimate is

$$\frac{\Sigma\, w_i^2 \cdot \mathrm{Var}\,(\mathrm{MD}_i)}{(\Sigma\, w_i)^2}\ [20].$$

The authors stated that the conventional estimator of the variance $1/\Sigma\ \hat{w}_i$ will inevitably be negatively biased (i.e., expected to underestimate) due to the inflation of the individual study weights [20]. The underestimation of variance can lead to unreasonably narrow confidence intervals and thus poor coverage probabilities.

As an extension of Walter and Balakrishnan's work [20], this thesis aims to obtain an approximate expression for the variance of the summary treatment effect and investigate how it is affected by the bias-correction. The works below also study the extent to which the conventional estimator deflates the variance and adversely affects coverage. Comparisons between the coverage probabilities using biased and bias-corrected weights will aid in improving meta-analysis protocols for optimal inference.

Chapter 2

Methods

There are two principal criteria in optimal inference: bias and variance. Sections 2.2 and 2.3 below explore how the bias-correction affects them.

2.1 Notation

Let the random variables Y_1 and Y_2 denote the continuous outcomes from the two treatment groups. One assumes that the two groups are independent and that the studies are independent. If the treatment groups are homogeneous in their variances, $Y_1 \sim N(\mu_1, \lambda^2)$ and $Y_2 \sim N(\mu_2, \lambda^2)$. For a given study $i \in \{1, 2, \ldots, k\}$, the estimator for mean difference $\hat{\Delta}_i$ is then given by

$$\hat{\Delta}_i = \overline{Y}_{1i} - \overline{Y}_{2i} \dot{\sim} N\left(\Delta = \mu_1 - \mu_2,\ \lambda^2\left(\frac{1}{n_{1i}} + \frac{1}{n_{2i}}\right)\right). \tag{2.1.1}$$

Given the sample sizes for each treatment group n_{1i} and n_{2i}, the degrees of freedom is $v_i = n_{1i} + n_{2i} - 2$. Let $\hat{w}_i$ denote the study weights such that $\hat{w}_i = c_i/\widehat{\mathrm{Var}\left(\hat{\Delta}_i\right)}$, where

$$c_i = \begin{cases} 1 & \text{if } \hat{w}_i\text{s are conventional (i.e., biased)} \\ 1 - \dfrac{2}{\nu_i} & \text{if } \hat{w}_i\text{s are bias-corrected (i.e., unbiased).} \end{cases}$$

As seen in Equation 2.1.1,

$$\mathrm{Var}\left(\hat{\Delta}_i\right) := \varphi_i^2 = \lambda^2\left(\frac{1}{n_{1i}} + \frac{1}{n_{2i}}\right),$$

which can be estimated by

$$\widehat{\mathrm{Var}\left(\hat{\Delta}_i\right)} := \hat{\varphi}_i^2 = \hat{\lambda}_i^2\left(\frac{1}{n_{1i}} + \frac{1}{n_{2i}}\right).$$

Since $\mathrm{Var}\left(Y_i\right) = \mathrm{Var}\left(Y_2\right) = \lambda^2$ under the homogeneity assumption, its estimator is

$$\hat{\lambda}_i^2 = \frac{(n_{1i}-1)s_{1i}^2 + (n_{2i}-1)s_{2i}^2}{n_{1i} + n_{2i} - 2},$$

i.e., the pooled sample variance. The summary treatment effect estimate $\hat{\Delta}$ is the weighted mean of the individual $\hat{\Delta}_i$s, and we let

$$\hat{\theta}_i = \frac{\hat{w}_i}{\Sigma\, \hat{w}_i} = \frac{c_i/\hat{\varphi}_i^2}{\Sigma\ c_i/\hat{\varphi}_i^2},\ \text{such that}$$

$$\hat{\Delta} = \frac{\hat{w}_i\hat{\Delta}_i}{\Sigma\ \hat{w}_i} = \Sigma\, \hat{\theta}_i\hat{\Delta}_i \text{ with } \Sigma\, \hat{\theta}_i = 1.$$

If the studies are heterogeneous, λ^2s are simply replaced by λ_i^2s †.

2.2 Bias

Given the independence of $\hat{\theta}_i$ and $\hat{\Delta}_i$,

$$\mathbb{E}\left[\hat{\Delta}\right] = \mathbb{E}\left[\Sigma\, \hat{\theta}_i\hat{\Delta}_i\right] = \Sigma\, \mathbb{E}\left[\hat{\theta}_i\right]\cdot\mathbb{E}\left[\hat{\Delta}_i\right] = \Delta\cdot\Sigma\, \mathbb{E}\left[\hat{\theta}_i\right] = \Delta\cdot\mathbb{E}\left[\Sigma\, \hat{\theta}_i\right] = \Delta\cdot\mathbb{E}\left[1\right] = \Delta.$$

Therefore, $\hat{\Delta}$ is an unbiased estimator for Δ with or without the bias-correction as long as $\Sigma\, \hat{\theta}_i = 1$. This results in Precision$(\hat{\Delta}) = \mathrm{Var}\left(\hat{\Delta}\right) + \mathrm{Bias}(\hat{\Delta})^2 = \mathrm{Var}\left(\hat{\Delta}\right)$, which justifies the direct comparison of variances.

2.3 Variance

Under the homogeneity assumption, the variance of the summary treatment effect $\hat{\Delta}$ is given by

$$\mathrm{Var}\left(\hat{\Delta}\right) = \mathbb{E}\left[\hat{\Delta}^2\right] - \mathbb{E}\left[\hat{\Delta}\right]^2 = \mathbb{E}\left[(\Sigma\, \hat{\theta}_i\hat{\Delta}_i)^2\right] - \Delta^2, \text{ where}$$

$$\mathbb{E}\left[(\Sigma\, \hat{\theta}_i\hat{\Delta}_i)^2\right] = \sum \mathbb{E}\left[\hat{\theta}_i^2\right]\mathbb{E}\left[\hat{\Delta}_i^2\right] + \Delta^2\sum_{i\neq j}\mathbb{E}\left[\hat{\theta}_i\hat{\theta}_j\right], \tag{2.3.1}$$

$$\text{and } \mathbb{E}\left[\hat{\Delta}_i^2\right] = \mathrm{Var}\left(\hat{\Delta}_i\right) + \mathbb{E}\left[\hat{\Delta}_i\right]^2 = \varphi^2 + \Delta^2 = \lambda^2\left(\frac{1}{n_{1i}} + \frac{1}{n_{2i}}\right) + \Delta^2.$$

However, the exact expression for this variance cannot be obtained. Multivariate Taylor series expansion is thus employed below.

† There are two types of heterogeneity in meta-analyses. One is when the treatment groups, Y_1 and Y_2, have different variances ($\lambda_1^2 \neq \lambda_2^2$). The other is when the studies have different variances (λ_i^2 for $i = 1, 2, \ldots, k$). In this thesis, heterogeneity refers to the latter kind. See Section 2.3.3.

2.3.1 Inverse χ^2 Parametrization

Let $z_i = c_i/\hat{\varphi}_i^2$, then $\hat{\theta}_i$ can be written as a function of the zs: $\hat{\theta}_i(\vec{z}) = z_i/\Sigma_j z_j$. Then,

$$
\begin{aligned}
\mathbb{E}[z_i] &= \mathbb{E}\left[c_i/\hat{\varphi}_i^2\right] \\
&= c_i \cdot \mathbb{E}\left[1/\hat{\varphi}_i^2\right] \qquad \text{since } c_i \text{ is a function of } \nu_i \text{ only} \\
&= c_i \cdot \frac{1}{\dfrac{1}{n_{1i}} + \dfrac{1}{n_{2i}}} \cdot \mathbb{E}\left[1/\hat{\lambda}_i^2\right] \\
&= c_i \cdot \frac{1}{\dfrac{1}{n_{1i}} + \dfrac{1}{n_{2i}}} \cdot \frac{N_i - 2}{N_i - 4} \cdot \frac{1}{\lambda^2}, \qquad \text{where } N_i = n_{1i} + n_{2i}, \\
&\qquad \text{using results from Walter and Balakrishnan [20].} \\
&= \frac{c_i \cdot n_{1i} n_{2i} \cdot \nu_i}{\lambda^2(\nu_i + 2)(\nu_i - 2)}, \qquad \text{where } \nu_i = n_{1i} + n_{2i} - 2 \\
&:= \alpha_i \qquad \text{with common } \lambda^2 \text{ across studies} \\
&(\nu_i > 2, N_i > 4).
\end{aligned}
$$

Similarly,

$$
\begin{aligned}
\mathbb{E}\left[z_i^2\right] &= \mathbb{E}\left[(c_i/\hat{\varphi}_i^2)^2\right] \\
&= c_i^2 \cdot \mathbb{E}\left[\frac{1}{\hat{\lambda}_i^2} \cdot \frac{1}{\left(\dfrac{1}{n_{1i}} + \dfrac{1}{n_{2i}}\right)^2}\right] \\
&= \frac{c_i^2}{\left(\dfrac{1}{n_{1i}} + \dfrac{1}{n_{2i}}\right)^2} \cdot \mathbb{E}\left[1/\hat{\lambda}_i^4\right] \\
&= \frac{c_i^2}{\left(\dfrac{1}{n_{1i}} + \dfrac{1}{n_{2i}}\right)^2} \cdot \frac{\nu_i^2}{\lambda^4} \cdot \mathbb{E}\left[\frac{\lambda^4}{\nu_i^2 \hat{\lambda}_i^4}\right]
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{c_i^2 \nu_i^2}{\lambda^4 \left(\frac{1}{n_{1i}} + \frac{1}{n_{2i}}\right)^2} \cdot \mathbb{E}\left[1/T^2\right] \\
&\text{where } T = \frac{\nu_i \hat{\lambda}_i^2}{\lambda^2} \sim \chi^2_{\nu_i}.
\end{aligned}
$$

$$
\begin{aligned}
\mathbb{E}\left[1/T^2\right] &= \int_0^\infty \frac{1}{t^2} \cdot \frac{1}{2^{\nu_i/2}\Gamma(\nu_i/2)} \cdot e^{-\frac{t}{2}} t^{\frac{\nu_i}{2}-1} dt \\
&= 2^{-2}\left(\frac{2}{\nu_i - 4}\right)\left(\frac{2}{\nu_i - 2}\right) \int_0^\infty \frac{1}{2^{\frac{\nu_i-4}{2}}\Gamma(\frac{\nu_i - 4}{2})} \cdot e^{-\frac{t}{2}} t^{\frac{\nu_i-4}{2}-1} dt \\
&= 2^{-2}\left(\frac{2}{\nu_i - 4}\right)\left(\frac{2}{\nu_i - 2}\right) \\
&= \frac{1}{(\nu_i - 4)(\nu_i - 2)} \\
\Rightarrow \mathbb{E}\left[z_i^2\right] &= \frac{c_i^2 \nu_i^2}{\lambda^4 \left(\frac{1}{n_{1i}} + \frac{1}{n_{2i}}\right)^2} \cdot \frac{1}{(\nu_i - 4)(\nu_i - 2)}.
\end{aligned}
$$

This gives

$$
\begin{aligned}
\mathrm{Var}\left(z_i\right) &= \mathbb{E}\left[z_i^2\right] - \mathbb{E}\left[z_i\right]^2 \\
&= \frac{c_i^2 \nu_i^2}{\lambda^4 \left(\frac{1}{n_{1i}} + \frac{1}{n_{2i}}\right)^2} \cdot \frac{1}{(\nu_i - 4)(\nu_i - 2)} - \left(\frac{c_i \cdot n_{1i} n_{2i} \cdot \nu_i}{\lambda^2 (\nu_i + 2)(\nu_i - 2)}\right)^2 \\
&= \frac{2\nu_i^2 c_i^2}{\lambda^4 \left(\frac{1}{n_{1i}} + \frac{1}{n_{2i}}\right)^2 (\nu_i - 4)(\nu_i - 2)^2} \\
&:= \beta_i^2 \qquad \text{again with common } \lambda^2 \text{ across studies} \\
&(\nu_i > 4, N_i > 6).
\end{aligned}
$$

Note $\mathrm{Cov}(z_i, z_j)$:=$\beta_{i,j}$=0 for all pairs (i, j) since all studies are independent.

Using these results for a second-order Taylor approximation with respect to the

z_is, one obtains

$$\begin{aligned}
\hat{\theta}_i^2(\vec{z}) &:= u_i(\vec{z}) \qquad \text{function } u_i \text{ of } \vec{z} \\
&= \frac{z_i^2}{(\Sigma_j z_j)^2} \\
\Rightarrow \hat{\theta}_i^2 &\approx u_i(\vec{\alpha}) + \sum_{p=1}^{k}(z_p - \alpha_p) \cdot u_i'(\vec{\alpha}) \\
&+ \frac{1}{2}\sum_{p=1}^{k}(z_p - \alpha_p)^2 \cdot u_i''(\vec{\alpha}) + \frac{1}{2}\sum_{p \neq q}(z_p - \alpha_p)(z_q - \alpha_q) \cdot \frac{\partial^2}{\partial z_q \partial z_p} u_i(\vec{\alpha}) \\
\Rightarrow \mathbb{E}\left[\hat{\theta}_i^2\right] &\approx u_i(\vec{\alpha}) + \sum_p (\mathbb{E}\left[z_p\right] - \alpha_p) \cdot u_i'(\vec{\alpha}) + \frac{1}{2}\sum_p \mathbb{E}\left[(z_p - \alpha_p)^2\right] \cdot u_i''(\vec{\alpha}) \\
&+ \frac{1}{2}\sum_{p \neq q} \mathbb{E}\left[(z_p - \alpha_p)(z_q - \alpha_q)\right] \cdot \frac{\partial^2}{\partial z_q \partial z_p} u_i(\vec{\alpha}) \\
&= u_i(\vec{\alpha}) + \sum \beta_p^2 \cdot u_i''(\vec{\alpha}), \\
&\text{as } \mathbb{E}\left[z_p\right] - \alpha_p = \mathbb{E}\left[(z_p - \alpha_p)(z_q - \alpha_q)\right] = 0.
\end{aligned}$$

Given

$$u_i'' = \begin{cases} \dfrac{2(\Sigma\, z_j - z_i)(\Sigma\, z_j - 3z_i)}{(\Sigma z_j)^4} & \text{if } p = i \\ \dfrac{6z_i^2}{(\Sigma z_j)^4} & \text{if } p \neq i, \end{cases}$$

$$\mathbb{E}\left[\hat{\theta}_i^2\right] \approx \frac{\alpha_i^2}{(\Sigma_j\, \alpha_j)^2} + \beta_i^2 \cdot \frac{(\Sigma_j\, \alpha_j - \alpha_i)(\Sigma_j\, \alpha_j - 3\alpha_i)}{(\Sigma_j\, \alpha_j)^4} + \sum_{p \neq i} \beta_p^2 \cdot \frac{3\alpha_i^2}{(\Sigma_j\, \alpha_j)^4}.$$

Second-order Taylor approximation is used again to obtain an expression for the last term in Equation 2.3.1:

$$\mathbb{E}\left[\hat{\theta}_i \hat{\theta}_j\right] = \mathbb{E}\left[\frac{z_i}{\Sigma_l\, \alpha_l} \cdot \frac{z_j}{\Sigma_l\, \alpha_l}\right]$$

$$= \mathbb{E}\left[\frac{z_i z_j}{(\Sigma_l\ \alpha_l)^2}\right], \qquad i \neq j.$$

Let $g_{i,j}(\vec{z}) = \hat{\theta}_i \hat{\theta}_j = \frac{z_i z_j}{(\Sigma_l\ \alpha_l)^2}$.

$$\begin{aligned}
\Rightarrow g_{i,j}(\vec{z}) \approx g_{i,j}(\vec{\alpha}) &+ \sum_{p=i}^{k}(z_p - \alpha_p) \cdot \frac{\partial}{\partial z_p} g_{i,j}(\vec{\alpha}) + \frac{1}{2}\sum_{p=1}^{k}(z_p - \alpha_p)^2 \cdot \frac{\partial^2}{\partial z_p^2} g_{i,j}(\vec{\alpha}) \\
&+ \frac{1}{2}\sum_{p \neq q}(z_p - \alpha_p)(z_q - \alpha_q) \cdot \frac{\partial^2}{\partial z_q \partial z_p} g_{i,j}(\vec{\alpha})
\end{aligned}$$

$$\begin{aligned}
\Rightarrow \mathbb{E}\left[\hat{\theta}_i \hat{\theta}_j\right] &= \mathbb{E}\left[g_{i,j}\right] \\
&\approx g_{i,j}(\vec{\alpha}) + \sum_{p=i}^{k}(\mathbb{E}\left[z_p\right] - \alpha_p) \cdot \frac{\partial}{\partial z_p} g_{i,j}(\vec{\alpha}) + \frac{1}{2}\sum_{p=1}^{k}\mathbb{E}\left[(z_p - \alpha_p)^2\right] \cdot \frac{\partial^2}{\partial z_p^2} g_{i,j}(\vec{\alpha}) \\
&\quad + \frac{1}{2}\sum_{p \neq q}\mathbb{E}\left[(z_p - \alpha_p)(z_q - \alpha_q)\right] \cdot \frac{\partial^2}{\partial z_q \partial z_p} g_{i,j}(\vec{\alpha}) \\
&= g_{i,j}(\vec{\alpha}) + \frac{1}{2}\sum_{p} \beta_p^2 \cdot \frac{\partial^2}{\partial z_p^2} g_{i,j}(\vec{\alpha})
\end{aligned}$$

With

$$g''_{i,j} = \begin{cases} \dfrac{2z_j(3z_i - 2\Sigma_l\ \alpha_l)}{(\Sigma_l\ \alpha_l)^4} & \text{if } p = i \\ \dfrac{2z_i(3z_j - 2\Sigma_l\ \alpha_l)}{(\Sigma_l\ \alpha_l)^4} & \text{if } p = j \\ \dfrac{6z_i z_j}{(\Sigma_l\ \alpha_l)^4} & \text{if } p \neq i \wedge p \neq j, \end{cases}$$

one obtains

$$\mathbb{E}\left[\hat{\theta}_i \hat{\theta}_j\right] \approx \frac{\alpha_i \alpha_j}{(\Sigma_l\ \alpha_l)^2} - \frac{2\beta_i^2 \alpha_j + 2\beta_j^2 \alpha_i}{(\Sigma_l\ \alpha_l)^3} + \frac{3\alpha_i \alpha_j \Sigma_l\ \beta_l^2}{(\Sigma_l\ \alpha_l)^4}.$$

Therefore, Equation 2.3.1 yields

$$\mathrm{Var}\left(\hat{\Delta}\right) = \mathbb{E}\left[(\Sigma\ \hat{\theta}_i \hat{\Delta}_i)^2\right] - \Delta^2, \text{ where}$$

$$\begin{aligned}
\mathbb{E}\left[(\Sigma\,\hat{\theta}_i\hat{\Delta}_i)^2\right] &= \sum \mathbb{E}\left[\hat{\theta}_i^2\right]\mathbb{E}\left[\hat{\Delta}_i^2\right] + \Delta^2 \sum_{i\neq j} \mathbb{E}\left[\hat{\theta}_i\hat{\theta}_j\right] \\
&\approx \sum_{i=1}^{k} \left[\frac{\alpha_i^2}{(\Sigma_j\,\alpha_j)^2} + \beta_i^2 \cdot \frac{(\Sigma_j\,\alpha_j - \alpha_i)(\Sigma_j\,\alpha_j - 3\alpha_i)}{(\Sigma_j\,\alpha_j)^4} + \sum_{p\neq i} \beta_p^2 \cdot \frac{3\alpha_i^2}{(\Sigma_j\,\alpha_j)^4}\right] \\
&\quad \times \left[\lambda^2\left(\frac{1}{n_{1i}} + \frac{1}{n_{2i}}\right) + \Delta^2\right] \\
&\quad + \Delta^2 \cdot \sum_{i\neq j} \left[\frac{\alpha_i\alpha_j}{(\Sigma_l\,\alpha_l)^2} - \frac{2\beta_i^2\alpha_j + 2\beta_j^2\alpha_i}{(\Sigma_l\,\alpha_l)^3} + \frac{3\alpha_i\alpha_j\Sigma_l\,\beta_l^2}{(\Sigma_l\,\alpha_l)^4}\right]. \qquad (2.3.2)
\end{aligned}$$

It is important to note that the total sample size of $N_i \geq 7$ is required for all studies $i = 1, \ldots, k$ due to the $\frac{1}{\nu_i - 4}$ term in the second moment of z_i. Hereinafter, this approximation in Equation 2.3.2 will thus be referred to as *N7*.

One may instead opt for a first-order approximation to avoid the sample size restriction:

$$\begin{aligned}
\mathbb{E}\left[\hat{\theta}_i^2\right] &\approx u_i(\vec{\alpha}) + \sum_p (\mathbb{E}\left[z_p\right] - \alpha_p) \cdot u_i'(\vec{\alpha}) \\
&= u_i(\vec{\alpha}) \\
&= \frac{\alpha_i^2}{(\Sigma\,\alpha_j)^2} \\
\mathbb{E}\left[\hat{\theta}_i\hat{\theta}_j\right] &\approx g_{i,j}(\vec{\alpha}) + \sum_{p=i}^{k} (\mathbb{E}\left[z_p\right] - \alpha_p) \cdot g_{i,j}'(\vec{\alpha}) \\
&= g_{i,j}(\vec{\alpha}) \\
&= \frac{\alpha_i\alpha_j}{(\Sigma\,\alpha_l)^2} \\
\Rightarrow \mathrm{Var}\left(\hat{\Delta}\right) &= \mathbb{E}\left[(\Sigma\,\hat{\theta}_i\hat{\Delta}_i)^2\right] - \Delta^2 \\
&\approx \sum_{i=1}^{k} \frac{\alpha_i^2}{(\Sigma\,\alpha_j)^2} \cdot \left[\lambda^2\left(\frac{1}{n_{1i}} + \frac{1}{n_{2i}}\right) + \Delta^2\right] + \Delta^2 \sum_{i\neq j} \frac{\alpha_i\alpha_j}{(\Sigma\,\alpha_l)^2} - \Delta^2 \qquad (2.3.3)
\end{aligned}$$

Without β_i^2, this approximation only requires $N_i \geq 5$. Equation 2.3.3 will thus be referred to as *N5* in the following sections.

In practice, the terms involving the unknown λ^2 in α_i and β_i^2 have to be estimated. Without the bias-correction, one simply replaces λ^2s with $\hat{\lambda}_i^2$, i.e., $1/\hat{\lambda}_i^2$ and $1/\hat{\lambda}_i^4$ as estimators for $1/\lambda^2$ and $1/\lambda^4$, respectively. On the other hand, with the bias-correction ($c_i = 1 - 2/\nu_i$), one uses $c_i/\hat{\lambda}_i^2$ to estimate $1/\lambda^2$ given

$$\mathbb{E}\left[\frac{1}{\hat{\lambda}_i^2}\right] = \frac{1}{c_i} \cdot \frac{1}{\lambda^2}.$$

Similarly, given

$$\mathbb{E}\left[\frac{1}{\hat{\lambda}_i^4}\right] = \frac{\nu_i^2}{(\nu_i - 4)(\nu_i - 2)} \cdot \frac{1}{\lambda^4} = \frac{\nu_i}{c_i(\nu_i - 4)} \cdot \frac{1}{\lambda^4},$$

one uses

$$\frac{c_i(\nu_i - 4)}{\nu_i} \cdot \frac{1}{\hat{\lambda}_i^4}$$

to estimate $1/\lambda^4$. Then, with the bias-correction,

$$\begin{aligned}
\hat{\alpha}_i &= \frac{c_i}{\hat{\lambda}_i^2} \cdot \frac{c_i \cdot n_{1i} n_{2i} \cdot \nu_i}{(\nu_i + 2)(\nu_i - 2)} \\
\hat{\beta}_i^2 &= \frac{c_i(\nu_i - 4)}{\nu_i} \cdot \frac{1}{\hat{\lambda}_i^2} \cdot \frac{2\nu_i^2}{\left(\dfrac{1}{n_{1i}} + \dfrac{1}{n_{2i}}\right)^2 (\nu_i - 4)(\nu_i - 2)^2}.
\end{aligned}$$

2.3.2 χ^2 Parametrization

An alternative parametrization is proposed to address the sample size restriction of N7. Define the reciprocal of z_i as $z_i' = \hat{\varphi}_i^2/c_i$, and $\hat{\theta}_i(\vec{z}') = \dfrac{1/z_i'}{\sum_j 1/z_j'}$. The first moment of z_i' is:

$$\begin{aligned}
\mathbb{E}\left[z_i'\right] &= \mathbb{E}\left[\hat{\varphi}_i^2/c_i\right] \\
&= \frac{1}{c_i} \cdot \mathbb{E}\left[\hat{\varphi}_i^2\right] \\
&= \frac{1}{c_i} \cdot \mathbb{E}\left[\hat{\lambda}_i^2 \cdot \left(\frac{1}{n_{1i}} + \frac{1}{n_{2i}}\right)\right] \\
&= \frac{1}{c_i} \cdot \left(\frac{1}{n_{1i}} + \frac{1}{n_{2i}}\right) \cdot \mathbb{E}\left[\hat{\lambda}_i^2\right] \\
&= \frac{\lambda^2}{c_i} \cdot \left(\frac{1}{n_{1i}} + \frac{1}{n_{2i}}\right) \\
&:= \alpha_i'.
\end{aligned}$$

Similarly, the second moment is given by:

$$\begin{aligned}
\mathbb{E}\left[z_i^{2\prime}\right] &= \mathbb{E}\left[\left(\hat{\varphi}_i^2/c_i\right)^2\right] \\
&= \frac{1}{c_i^2} \cdot \left(\frac{1}{n_{1i}} + \frac{1}{n_{2i}}\right)^2 \cdot \mathbb{E}\left[\hat{\lambda}_i^4\right], \\
&= \frac{1}{c_i^2} \cdot \left(\frac{1}{n_{1i}} + \frac{1}{n_{2i}}\right)^2 \cdot \mathbb{E}\left[T^2\right] \cdot \frac{\lambda^4}{\nu_i^2}, \\
&\text{where } T = \frac{\nu_i \hat{\lambda}_i^2}{\lambda^2} \sim \chi^2_{\nu_i} \text{ again} \\
&= \frac{1}{c_i^2} \cdot \left(\frac{1}{n_{1i}} + \frac{1}{n_{2i}}\right)^2 \cdot \nu_i(\nu_i + 2) \cdot \frac{\lambda^4}{\nu_i^2} \\
&= \frac{\lambda^4(\nu_i + 2)}{c_i^2 \nu_i} \cdot \left(\frac{1}{n_{1i}} + \frac{1}{n_{2i}}\right)^2.
\end{aligned}$$

$$
\begin{aligned}
\Rightarrow \operatorname{Var}\left(z_i'\right) &= \mathbb{E}\left[z_i^{2\prime}\right] - \mathbb{E}\left[z_i'\right]^2 \\
&= \frac{\lambda^4(\nu_i + 2)}{c_i^2 \nu_i} \cdot \left(\frac{1}{n_{1i}} + \frac{1}{n_{2i}}\right)^2 - \frac{\lambda^4}{c_i^2} \cdot \left(\frac{1}{n_{1i}} + \frac{1}{n_{2i}}\right)^2 \\
&= \frac{\lambda^4}{c_i^2} \cdot \left(\frac{1}{n_{1i}} + \frac{1}{n_{2i}}\right)^2 \cdot \frac{2}{\nu_i} \\
&:= \beta_i'^2 \\
&(\nu_i > 0 \text{ or } N_i > 2)
\end{aligned}
$$

It is seen here that this parametrization only requires $N_i \geq 3$ [†]– This approximation will hereinafter be referred to as *N3* (See Equation 2.3.4 below). For all pairs of $i \neq j$, $\operatorname{Cov}(z_i', z_j') = 0$ again given the independence of studies.

The λ^2 term in α_i' can simply be estimated by its unbiased estimator $\hat{\lambda}_i^2$. Without the bias-correction, the λ^4 term in $\beta_i'^2$ is estimated by $\hat{\lambda}_i^4$. However, given

$$
\mathbb{E}\left[\hat{\lambda}_i^4\right] = \frac{\nu_i + 2}{\nu_i} \cdot \lambda^4,
$$

the bias-corrected N3 uses

$$
\frac{\nu_i}{\nu_i + 2} \cdot \hat{\lambda}_i^4
$$

to estimate λ^4 in $\beta_i'^2$. Thus, the bias-correction results in

$$
\hat{\beta}_i'^2 = \frac{\nu_i}{\nu_i + 2} \cdot \hat{\lambda}_i^4 \cdot \frac{1}{c_i^2} \cdot \left(\frac{1}{n_{1i}} + \frac{1}{n_{2i}}\right)^2 \cdot \frac{2}{\nu_i}.
$$

[†]Although the second moment of z_i' only requires $N_i \geq 3$, the effective lower bound for N_i is 5 under the heterogeneity assumption. A sample size of at least 2 is required for both groups to calculate the sample variances ($n_1 \geq 2 \wedge n_2 \geq 2$). This gives $N_i \geq 4$ for all i, but $N_i = 4$ leads to invalid bias-correction factor $1 - \nu_i/2 = 1 - (N_i - 2)/2 = 0$. Therefore, $N_i \geq 5$ is required. Technically, the assumption of homogeneity in treatment group variances permits $n_i = 1$ for the estimation of pooled variance, in which case $N_i = 2$ is possible. However, as mentioned, the conventional estimator for the pooled variance ($\hat{\lambda}_i^2$) requires a sample variance for each treatment group.

Now let

$$
\begin{aligned}
f_i(\vec{z'}) &= \hat{\theta}_i^2(\vec{z'}) \\
&= \frac{1/z_i^{2\prime}}{(\Sigma\, 1/z_j')^2} \\
\Rightarrow f_i(\vec{z'}) &\approx f_i(\vec{\alpha'}) + \sum_{p=1}^{k}(z_p' - \alpha_p') \cdot f_i'(\vec{\alpha'}) \\
&\quad + \frac{1}{2}\sum_{p=1}^{k}(z_p' - \alpha_p')^2 \cdot f_i''(\vec{\alpha'}) + \frac{1}{2}\sum_{p\neq q}(z_p' - \alpha_p')(z_q' - \alpha_q') \cdot \frac{\partial^2}{\partial z_q' \partial z_p'} \cdot f_i(\vec{\alpha'}),
\end{aligned}
$$

using second-order Taylor approximation again.

This gives

$$
\begin{aligned}
\mathbb{E}\left[\hat{\theta}_i^2\right] &\approx f_i(\vec{\alpha'}) + \sum_{p=1}^{k}(\mathbb{E}\left[z_p'\right] - \alpha_p') \cdot f_i'(\vec{\alpha'}) \\
&\quad + \frac{1}{2}\sum_{p=1}^{k}\mathbb{E}\left[(z_p' - \alpha_p')^2\right] \cdot f_i''(\vec{\alpha'}) + \frac{1}{2}\sum_{p\neq q}\mathbb{E}\left[(z_p' - \alpha_p')(z_q' - \alpha_q')\right] \cdot \frac{\partial^2}{\partial z_q' \partial z_p'} \cdot f_i(\vec{\alpha'}) \\
&= f_i(\vec{\alpha'}) + \frac{1}{2}\sum_{p=1}^{k}\beta_p^{2\prime} \cdot f_i''(\vec{\alpha'}).
\end{aligned}
$$

Given

$$
f_i'' = \begin{cases} \dfrac{6z_i'^{-4}(z_i'^{-1} - \Sigma\, z_j'^{-1})^2}{(\Sigma\, z_j'^{-1})^4} & \text{if } p = i \\[2ex] \dfrac{2z_i'^{-2}z_p'^{-3}(3z_p'^{-1} - 2\Sigma\, z_j'^{-1})}{(\Sigma\, z_j'^{-1})^4} & \text{if } p \neq i, \end{cases}
$$

$$
\mathbb{E}\left[\hat{\theta}_i^2\right] \approx \frac{\alpha_i'^{-2}}{(\Sigma\, \alpha_j'^{-1})^2} + \beta_i'^2 \cdot \frac{3\alpha_i'^{-4}(\alpha_i'^{-1} - \Sigma\, \alpha_j'^{-1})^2}{(\Sigma\, \alpha_j'^{-1})^4} + \sum_{p\neq i}\beta_p'^2 \cdot \frac{\alpha_i'^{-2}\alpha_p'^{-3}(3\alpha_p'^{-1} - 2\Sigma\, \alpha_j'^{-1})}{(\Sigma\, \alpha_j'^{-1})^4}.
$$

Similarly, let

$$\begin{aligned}
h_{i,j}(\vec{z'}) &= \hat{\theta}_i\hat{\theta}_j \\
&= \frac{z_i'^{-1}z_j'^{-1}}{(\Sigma\, z_l'^{-1})^2} \qquad (i \neq j) \\
\Rightarrow h_{i,j}(\vec{z'}) &\approx h_{i,j}(\vec{\alpha'}) + \sum_{p=1}^{k}(z_p' - \alpha_p') \cdot h_{i,j}'(\vec{\alpha'}) + \frac{1}{2}\sum_{p=1}^{k}(z_p' - \alpha_p')^2 \cdot h_{i,j}''(\vec{\alpha'}) \\
&\quad + \frac{1}{2}\sum_{p\neq q}(z_p' - \alpha_p')(z_q' - \alpha_q') \cdot \frac{\partial^2}{\partial z_q'\partial z_p'} \cdot h_{i,j}(\vec{\alpha'}).
\end{aligned}$$

Taylor approximation gives

$$\begin{aligned}
\mathbb{E}\left[\hat{\theta}_i\hat{\theta}_j\right] &\approx h_{i,j}(\vec{\alpha'}) + \sum_{p=1}^{k}(\mathbb{E}\left[z_p'\right] - \alpha_p') \cdot h_{i,j}'(\vec{\alpha'}) + \frac{1}{2}\sum_{p=1}^{k}\mathbb{E}\left[(z_p' - \alpha_p')^2\right] \cdot h_{i,j}''(\vec{\alpha'}) \\
&\quad + \frac{1}{2}\sum_{p\neq q}\mathbb{E}\left[(z_p' - \alpha_p')(z_q' - \alpha_q')\right] \cdot \frac{\partial^2}{\partial z_q'\partial z_p'} \cdot h_{i,j}(\vec{\alpha'}) \\
&= h_{i,j}(\vec{\alpha'}) + \sum_{p=1}^{k}\beta_p'^2 \cdot h_{i,j}''(\vec{\alpha'}).
\end{aligned}$$

Using

$$h_{i,j}'' = \begin{cases} \dfrac{2z_i'^{-3}z_j'^{-1}(\Sigma_l\, z_l'^{-1} - z_i'^{-1})(\Sigma_l\, z_l'^{-1} - 3z_i'^{-1})}{(\Sigma\, z_l'^{-1})^4} & \text{if } p = i \\ \dfrac{2z_i'^{-1}z_j'^{-3}(\Sigma_l\, z_l'^{-1} - z_j'^{-1})(\Sigma_l\, z_l'^{-1} - 3z_j'^{-1})}{(\Sigma\, z_l'^{-1})^4} & \text{if } p = j \\ \dfrac{2z_i'^{-1}z_j'^{-1}z_p'^{-3}(3z_p'^{-1} - 2\Sigma\, z_l'^{-1})}{(\Sigma\, z_l'^{-1})^4} & \text{if } p \neq i \wedge p \neq j, \end{cases}$$

one has

$$\mathbb{E}\left[\hat{\theta}_i\hat{\theta}_j\right] \approx \frac{\alpha_i'^{-1}\alpha_j'^{-1}}{(\Sigma\, \alpha_l'^{-1})^2} + \beta_i'^2 \cdot \frac{\alpha_i'^{-3}\alpha_j'^{-1}(\Sigma\, \alpha_l'^{-1} - \alpha_i'^{-1})(\Sigma\, \alpha_l'^{-1} - 3\alpha_i'^{-1})}{(\Sigma\, \alpha_l'^{-1})^4}$$

$$+ \beta_j'^2 \cdot \frac{\alpha_i'^{-1}\alpha_j'^{-3}(\Sigma\, \alpha_l'^{-1} - \alpha_j'^{-1})(\Sigma\, \alpha_l'^{-1} - 3\alpha_j'^{-1})}{(\Sigma\, \alpha_l'^{-1})^4}$$

$$+ \sum_{p\neq i \wedge p\neq j} \beta_p'^2 \cdot \frac{\alpha_i'^{-1}\alpha_j'^{-1}\alpha_p'^{-3}(3\alpha_p'^{-1} - 2\Sigma\, \alpha_l'^{-1})}{(\Sigma\, \alpha_l'^{-1})^4}.$$

Therefore, under this parametrization Equation 2.3.1 yields

$$\begin{aligned}
\mathrm{Var}\left(\hat{\Delta}\right) &= \mathbb{E}\left[(\Sigma\, \hat{\theta}_i\hat{\Delta}_i)^2\right] - \Delta^2, \text{ where} \\
\mathbb{E}\left[(\Sigma\, \hat{\theta}_i\hat{\Delta}_i)^2\right] &= \sum \mathbb{E}\left[\hat{\theta}_i^2\right]\mathbb{E}\left[\hat{\Delta}_i^2\right] + \Delta^2 \sum_{i\neq j} \mathbb{E}\left[\hat{\theta}_i\hat{\theta}_j\right] \\
&\approx \sum_{i=1}^{k} \left[\frac{\alpha_i'^{-2}}{(\Sigma\, \alpha_j'^{-1})} + \beta_i'^2 \cdot \frac{3\alpha_i'^{-4}(\alpha_i'^{-1} - \Sigma\, \alpha_j'^{-1})^2}{(\Sigma\, \alpha_j'^{-1})^4} + \sum_{p\neq i} \beta_p'^2 \cdot \frac{\alpha_i'^{-2}\alpha_p'^{-3}(3\alpha_p'^{-1} - 2\Sigma\, \alpha_j'^{-1})}{(\Sigma\, \alpha_j'^{-1})^4}\right] \\
&\quad \times \left[\lambda^2\left(\frac{1}{n_{1i}} + \frac{1}{n_{2i}}\right) + \Delta^2\right] \\
&\quad + \Delta^2 \cdot \sum_{i\neq j} \left[\frac{\alpha_i'^{-1}\alpha_j'^{-1}}{(\Sigma\, \alpha_l'^{-1})^2} + \beta_i'^2 \cdot \frac{\alpha_i'^{-3}\alpha_j'^{-1}(\Sigma\, \alpha_l'^{-1} - \alpha_i'^{-1})(\Sigma\, \alpha_l'^{-1} - 3\alpha_i'^{-1})}{(\Sigma\, \alpha_l'^{-1})^4}\right. \\
&\qquad + \beta_j'^2 \cdot \frac{\alpha_i'^{-1}\alpha_j'^{-3}(\Sigma\, \alpha_l'^{-1} - \alpha_j'^{-1})(\Sigma\, \alpha_l'^{-1} - 3\alpha_j'^{-1})}{(\Sigma\, \alpha_l'^{-1})^4} \\
&\qquad \left. + \sum_{p\neq i \wedge p\neq j} \beta_p'^2 \cdot \frac{\alpha_i'^{-1}\alpha_j'^{-1}\alpha_p'^{-3}(3\alpha_p'^{-1} - 2\Sigma\, \alpha_l'^{-1})}{(\Sigma\, \alpha_l'^{-1})^4}\right]. \qquad (2.3.4)
\end{aligned}$$

Again, unknown λ^2 in α_i' and $\beta_i'^2$ needs to be estimated.

2.3.3 Heterogeneity

All three approximations (N7, N5, and N3) are readily applicable to heterogeneous meta-analyses. Rewrite 2.1.1 to

$$\hat{\Delta}_i = \overline{Y}_{1i} - \overline{Y}_{2i} \mathrel{\dot\sim} N\left(\Delta = \mu_1 - \mu_2,\ \lambda_i^2\left(\frac{1}{n_{1i}} + \frac{1}{n_{2i}}\right)\right),$$

where $\lambda_i^2 \neq \lambda_j^2$ for some $(i, j) \in \{1, 2, \ldots, k\}^2$. The common variance λ^2s are simply replaced by λ_i^2, yielding

$$\text{for N7 and N5} \begin{cases} \alpha_i &= \dfrac{c_i \cdot n_{1i} n_{2i} \cdot \nu_i}{\lambda_i^2 (\nu_i + 2)(\nu_i - 2)} \\ \beta_i^2 &= \dfrac{2\nu_i^2 c_i^2}{\lambda_i^4 \left(\dfrac{1}{n_{1i}} + \dfrac{1}{n_{2i}}\right)^2 (\nu_i - 4)(\nu_i - 2)^2}, \end{cases}$$

$$\text{for N3} \begin{cases} \alpha_i' &= \dfrac{\lambda_i^2}{c_i} \cdot \left(\dfrac{1}{n_{1i}} + \dfrac{1}{n_{2i}}\right) \\ \beta_i'^2 &= \dfrac{\lambda_i^4}{c_i^2} \cdot \left(\dfrac{1}{n_{1i}} + \dfrac{1}{n_{2i}}\right)^2 \cdot \dfrac{2}{\nu_i}, \end{cases}$$

and $\mathbb{E}\left[\hat{\Delta}_i^2\right] = \lambda_i^2 \left(\frac{1}{n_{1i}} + \frac{1}{n_{2i}}\right) + \Delta^2$. All else remains the same as the homogeneous case. In practice, $\alpha_i, \beta_i^2, \alpha_i'$, and $\beta_i'^2$ are all estimated by using the pooled sample variance $\hat{\lambda}_i^2$ in homogeneous settings. One likewise uses $\hat{\lambda}_i^2$ to estimate these parameters under the heterogeneity assumption. The approximation procedure is thus the same for homogeneous and heterogeneous meta-analyses.

It should also be noted that, under the homogeneity assumption, an alternative choice for the weights exists. Since one assumes that the studies share a common variance, each w_i contains a λ^2 term. Given that all weights are scaled by the common factor λ^2, the term can be dropped entirely, yielding $\hat{w}_i = 1/(1/n_{1i} + 1/n_{2i})$. The treatment variance λ^2 can then be separately estimated by pooling all studies $(\sum n_{1i} + n_{2i})$, in which case $n_{1i} = n_{2i} = 1$ is permitted. However, such a procedure is uncommon in practice, and most meta-analyses (although they may not explicitly state so) assume heterogeneity by choosing to pool the treatment groups in each study separately instead of pooling all the studies.[†]

[†]To reflect the conventional meta-analysis procedure, this alternative choice of weights was not used in the simulation study presented below.

2.4 Simulations

In order to investigate the effects of the bias-correction on $\mathrm{Var}\left(\hat{\Delta}\right)$, a simulation study was conducted. For a chosen set of values for parameters k, n_{1i}, n_{2i}, Δ, and λ^2 (λ_i^2 if heterogeneous), $\hat{\Delta}_i$ and $\hat{\lambda}_i$ were randomly sampled from their distributions– normal and χ^2:

$$\text{if homogeneous} \begin{cases} \hat{\Delta}_i & \dot{\sim} \; N\left(\Delta = \mu_1 - \mu_2,\; \lambda^2\left(\frac{1}{n_{1i}} + \frac{1}{n_{2i}}\right)\right) \\ \frac{\nu_i \hat{\lambda}_i^2}{\lambda^2} & \sim \chi^2_{\nu_i} \;\Rightarrow\; \hat{\lambda}_i^2 \overset{d}{=} \frac{\lambda^2 \cdot x^2_{\nu_i}}{\nu_i} \end{cases}$$

$$\text{if heterogeneous} \begin{cases} \hat{\Delta}_i & \dot{\sim} \; N\left(\Delta = \mu_1 - \mu_2,\; \lambda_i^2\left(\frac{1}{n_{1i}} + \frac{1}{n_{2i}}\right)\right) \\ \frac{\nu_i \hat{\lambda}_i^2}{\lambda_i^2} & \sim \chi^2_{\nu_i} \;\Rightarrow\; \hat{\lambda}_i^2 \overset{d}{=} \frac{\lambda_i^2 \cdot x^2_{\nu_i}}{\nu_i}. \end{cases}$$

$\mathrm{Var}\left(\hat{\Delta}\right)$ was estimated using the methods in Section 2.3 with and without the bias-correction, which was then used to calculate $\hat{w}_i = c_i/\hat{\varphi}_i^2$. The gold standard (GS) variance $1/(\Sigma\, 1/\varphi^2)$ (i.e., variance obtained by assuming that the true λ^2 is known) and the conventional plug-in estimator for the variance $1/\Sigma\, \hat{w}_i$ were calculated as well for comparison. 2,000 simulations were run for each set of parameter values. Scenarios with a small number of studies ($k = 2, 5$) and small to moderate sample sizes were explored. The true effect size Δ was set to zero to explore null cases (where treatment has no effect on the outcomes, i.e., $\mu_1 - \mu_2 = 0$).

Chapter 3

Results

3.1 Homogeneous Meta-Analyses

Tables 3.1 and 3.2 show results from the simulations. As expected, one observes that as the studies get bigger the variance estimates decrease overall while the coverage approaches 95%. Similarly, it is not surprising that the differences in the estimates become negligible as the sample sizes increase. When the number of studies increases, the variance estimates decrease when the sample sizes are comparable (e.g., $N_1 = N_2 = 20$ with $k = 2$ and $N_1 = \cdots = N_5 = 18$ with $k = 5$). However, when N_is are not comparable, having fewer but bigger studies is preferable to having many small studies. This is demonstrated by, for instance, comparing case 6 from Table 3.1 and case 1 from Table 3.2– The total number of observations $\Sigma\, N_i$ is the same (42 for both), yet $k = 2$ reports better coverage for all estimators.

An important result is that the conventional plug-in estimator $1/\Sigma\, \hat{w}_i$ consistently underestimates the variance across all cases. Even when compared to the GS estimates (a lower bound for the variance), the plug-in estimator gives smaller

n_{11}	n_{21}	n_{12}	n_{22}	GS	Plug-in	Empirical; biased	Empirical; corrected	Coverage; plug-in	N7 biased	N7 corrected	Coverage; N7 biased	Coverage; N7 corrected
2	3	5	5	0.270	0.222	0.337	0.322	0.532	NA	NA	NA	NA
3	4	5	6	0.225	0.195	0.255	0.252	0.878	0.290	0.235	0.928	0.902
2	3	20	20	0.089	0.081	0.119	0.100	0.369	NA	NA	NA	NA
7	8	40	40	0.042	0.041	0.042	0.042	0.936	0.043	0.042	0.941	0.943
10	10	11	11	0.095	0.090	0.097	0.097	0.933	0.096	0.095	0.939	0.938
20	20	30	30	0.040	0.039	0.040	0.040	0.938	0.040	0.040	0.943	0.942
50	55	50	55	0.019	0.019	0.019	0.019	0.941	0.019	0.019	0.942	0.942

continued

n_{11}	n_{21}	n_{12}	n_{22}	N5 biased	N5 corrected	Coverage; N5 biased	Coverage; N5 corrected	N3 biased	N3 corrected	Coverage; N3 biased	Coverage; N3 corrected
2	3	5	5	0.255	0.245	0.579	0.578	0.263	0.228	0.592	0.551
3	4	5	6	0.198	0.198	0.880	0.881	0.221	0.193	0.895	0.879
2	3	20	20	0.099	0.087	0.445	0.419	0.088	0.084	0.399	0.409
7	8	40	40	0.041	0.041	0.937	0.941	0.042	0.041	0.939	0.941
10	10	11	11	0.090	0.090	0.933	0.933	0.095	0.092	0.938	0.936
20	20	30	30	0.039	0.039	0.938	0.940	0.040	0.040	0.943	0.940
50	55	50	55	0.019	0.019	0.941	0.941	0.019	0.019	0.942	0.942

Table 3.1: Homogeneous Simulation Results with $k = 2$. $\Delta = 0$ and $\lambda^2 = 1$ for all. Rounded to three decimal places. NAs are presented where sample size requirements are not met for N7. The confidence level is 95% for all coverages reported. The empirical variance and the coverage probabilities are computed with the results from the 2,000 iterations. All estimates of variance, except for the empirical results, are means calculated from the 2,000 simulated results.

n_{11}	n_{21}	n_{12}	n_{22}	n_{13}	n_{23}	n_{14}	n_{24}	n_{15}	n_{25}	GS	Plug-in	Empirical; biased	Empirical; corrected	Coverage; plug-in	N7 biased	N7 corrected	Coverage; N7 biased	Coverage; N7 corrected
3	4	4	4	5	4	5	5	4	4	0.096	0.073	0.130	0.129	0.844	0.130	0.099	0.926	0.893
3	4	4	6	5	5	4	5	6	5	0.086	0.069	0.114	0.112	0.859	0.111	0.089	0.930	0.905
3	4	5	5	5	4	6	6	15	15	0.059	0.050	0.078	0.075	0.879	0.076	0.060	0.937	0.914
8	8	8	7	9	9	10	10	40	40	0.027	0.025	0.031	0.030	0.922	0.028	0.027	0.935	0.932
8	8	10	10	11	11	9	9	10	9	0.042	0.038	0.048	0.048	0.914	0.043	0.042	0.931	0.928
10	10	9	9	11	11	12	11	8	9	0.040	0.037	0.044	0.044	0.915	0.041	0.040	0.935	0.932
30	30	35	35	33	33	37	37	40	40	0.011	0.011	0.012	0.012	0.939	0.011	0.011	0.944	0.944
50	50	51	51	54	54	52	53	55	55	0.008	0.007	0.008	0.008	0.943	0.008	0.008	0.947	0.947

continued

n_{11}	n_{21}	n_{12}	n_{22}	n_{13}	n_{23}	n_{14}	n_{24}	n_{15}	n_{25}	N5 biased	N5 corrected	Coverage; N5 biased	Coverage; N5 corrected	N3 biased	N3 corrected	Coverage; N3 biased	Coverage; N3 corrected
3	4	4	4	5	4	5	5	4	4	0.073	0.073	0.844	0.848	0.090	0.063	0.876	0.814
3	4	4	6	5	5	4	5	6	5	0.069	0.069	0.860	0.862	0.083	0.062	0.893	0.839
3	4	5	5	5	4	6	6	15	15	0.051	0.051	0.884	0.893	0.058	0.048	0.905	0.881
8	8	8	7	9	9	10	10	40	40	0.025	0.025	0.922	0.923	0.027	0.026	0.931	0.924
8	8	10	10	11	11	9	9	10	9	0.038	0.038	0.914	0.914	0.042	0.039	0.927	0.914
10	10	9	9	11	11	12	11	8	9	0.037	0.037	0.915	0.913	0.040	0.037	0.929	0.915
30	30	35	35	33	33	37	37	40	40	0.011	0.011	0.939	0.939	0.011	0.011	0.944	0.943
50	50	51	51	54	54	52	53	55	55	0.007	0.007	0.943	0.944	0.008	0.008	0.947	0.947

Table 3.2: Homogeneous Simulation Results with $k = 5$. $\Delta = 0$ and $\lambda^2 = 1$ for all.

results. The underestimation is more severe when the studies are small. For example, the first two cases in the tables above show that the plug-in estimator underestimates the variance approximately by 25%-40% compared to the empirical variance with biased weights. This is reflected in its coverage as well, with its probability being smaller than those of the approximation methods in almost all cases.

Comparing the estimates with and without the bias-correction, empirical results show that bias-corrected weights give more precise $\hat{\Delta}$s almost always, although by a small margin. The margin is the greatest when the studies vary in their sample sizes, although the gap closes when at least one study has a sample size $\geq$80 (see cases 3-5 in Table 3.1 and cases 4-5 in Table 3.2). On the other hand, Taylor series-approximated estimates show more fluctuation. N5 exhibits superior coverage probabilities with the bias-correction almost always. Meanwhile, the bias-correction is often observed to decrease coverage in N7 and N3, particularly when the studies are small. However, the bias-corrected estimators, even when they perform worse than those without the correction, often show better coverage compared to the plug-in estimator.

Now comparing the three approximation methods, N7 outperforms N5 and N3 (as well as the plug-in estimator) with or without the bias-correction. N7 consistently returns the best coverage probability with $\approx$ 10% improvement from the plug-in estimator in some cases. N7's outperforming N5 is expected, since N7 has the advantage of the second-order terms in addition to the first-order approximation N5 uses under the same parametrization. However, N5 remains advantageous when the sample size dips below seven– N7 cannot be used even when there is only one study in the meta-analysis that does not meet the requirement $N_i \geq 7 \; \forall \; i$. N3's poor performance compared to N7 can be explained by its parametrization. Unlike N7 and N5, N3 has a χ^2 random variable in its numerator. χ^2 distribution is right-skewed, particularly when the degrees of freedom (ν_i) is small. The heavy tail near zero indicates non-negligible probabilities associated with values close to zero with a small ν_i. This pushes its reciprocal towards large values, yet the Taylor approximation around α_i's only allows finite results. Limit issues near zero thus explain the bigger margin of error in N3 than in N7. Following N7, N3 without the bias-correction performs better than N5 according to the coverage probabilities. With the correction, N3 returns worse coverages than N5 when the sample sizes are small, since the correction tends to reduce coverage in N3 as mentioned above. However, when the studies get bigger ($N_i \approx 20$) the trend reverses, and the bias-corrected N3 reports slightly better coverages than N5's.

3.2 Heterogeneous Meta-Analyses

Tables 3.3 and 3.4 show results from the simulations under the heterogeneity assumption. With control variance $\lambda_1^2 = 1$, heterogeneous studies with 25%, 50%, 75%, and 100% increase in the variance ($\lambda_{i \neq 1}^2 = 1.25, 1.50, 1.75, 2.00$) were explored.

n_{11}	n_{21}	n_{12}	n_{22}	λ_2^2	GS	Plug-in	Empirical; biased	Empirical; corrected	Coverage; plug-in	N7 biased	N7 corrected	Coverage; N7 biased	Coverage; N7 corrected
2	3	5	5	1.25	0.313	0.255	0.385	0.377	0.554	NA	NA	NA	NA
2	3	5	5	1.5	0.349	0.283	0.426	0.426	0.570	NA	NA	NA	NA
3	4	5	6	1.75	0.306	0.265	0.345	0.348	0.875	0.380	0.330	0.925	0.904
3	4	5	6	2	0.325	0.283	0.367	0.371	0.876	0.401	0.353	0.922	0.906
20	20	2	3	1.25	0.091	0.084	0.122	0.102	0.383	NA	NA	NA	NA
20	20	2	3	1.5	0.093	0.086	0.124	0.103	0.389	NA	NA	NA	NA
7	8	40	40	1.75	0.066	0.063	0.067	0.066	0.930	0.067	0.066	0.934	0.937
7	8	40	40	2	0.073	0.070	0.074	0.073	0.930	0.074	0.073	0.936	0.936
10	10	11	11	1.25	0.106	0.101	0.109	0.109	0.933	0.107	0.106	0.940	0.939
10	10	11	11	1.5	0.115	0.109	0.118	0.118	0.932	0.116	0.115	0.940	0.940
30	30	20	20	1.75	0.048	0.047	0.047	0.047	0.937	0.048	0.048	0.941	0.941
30	30	20	20	2	0.050	0.049	0.049	0.049	0.940	0.050	0.050	0.942	0.942

n_{11}	n_{21}	n_{12}	n_{22}	λ_2^2	N5 biased	N5 corrected	Coverage; N5 biased	Coverage; N5 corrected	N3 biased	N3 corrected	Coverage; N3 biased	Coverage; N3 corrected
2	3	5	5	1.25	0.293	0.284	0.601	0.597	0.304	0.263	0.615	0.571
2	3	5	5	1.5	0.324	0.318	0.619	0.613	0.339	0.294	0.635	0.586
3	4	5	6	1.75	0.269	0.269	0.877	0.878	0.302	0.263	0.892	0.874
3	4	5	6	2	0.287	0.287	0.878	0.876	0.322	0.280	0.893	0.874
20	20	2	3	1.25	0.102	0.089	0.448	0.426	0.090	0.087	0.405	0.417
20	20	2	3	1.5	0.103	0.090	0.452	0.434	0.092	0.088	0.410	0.422
7	8	40	40	1.75	0.064	0.064	0.931	0.933	0.065	0.064	0.933	0.934
7	8	40	40	2	0.070	0.070	0.931	0.931	0.072	0.071	0.932	0.932
10	10	11	11	1.25	0.101	0.101	0.933	0.934	0.106	0.103	0.938	0.936
10	10	11	11	1.5	0.109	0.109	0.932	0.933	0.115	0.112	0.939	0.934
30	30	20	20	1.75	0.047	0.047	0.937	0.937	0.048	0.048	0.941	0.940
30	30	20	20	2	0.049	0.049	0.940	0.940	0.050	0.049	0.942	0.941

continued

Table 3.3: Heterogeneous Simulation Results with $k = 2$. $\Delta = 0$ and $\lambda_1^2 = 1$ for all.

n_{11}	n_{21}	n_{12}	n_{22}	n_{13}	n_{23}	n_{14}	n_{24}	n_{15}	n_{25}	λ_2^2	λ_3^2	λ_4^2	λ_5^2	GS	Plug-in	Empirical; biased	Empirical; corrected	Coverage; plug-in
3	4	4	4	5	4	5	5	4	4	1.5	1.25	1.5	1.25	0.124	0.094	0.167	0.166	0.842
3	4	4	6	5	5	4	5	6	5	2	2	1	1	0.110	0.087	0.144	0.143	0.856
3	4	5	5	5	4	6	6	15	15	1.25	1.25	1.5	1.5	0.080	0.067	0.106	0.103	0.876
8	8	8	7	9	9	10	10	40	40	1.25	1	1.75	2	0.041	0.038	0.047	0.047	0.925
7	7	8	8	9	9	10	10	10	9	1.25	1	1.25	1	0.050	0.045	0.061	0.061	0.901
8	8	7	7	9	9	10	10	10	9	2	1	1	2	0.057	0.051	0.068	0.068	0.906
10	10	9	9	11	11	12	11	8	9	1	2	1	1	0.045	0.041	0.050	0.050	0.918
30	30	35	35	33	33	37	37	40	40	1.25	1.5	1	1.75	0.014	0.014	0.015	0.015	0.937
50	50	51	51	54	54	52	53	55	55	2	1.5	1.75	1.25	0.011	0.011	0.011	0.011	0.944

N7 biased	N7 corrected	Coverage; N7 biased	Coverage; N7 corrected	N5 biased	N5 corrected	Coverage; N5 biased	Coverage; N5 corrected	N3 biased	N3 corrected	Coverage; N3 biased	Coverage; N3 corrected
0.168	0.128	0.929	0.894	0.094	0.094	0.842	0.843	0.116	0.081	0.876	0.814
0.143	0.113	0.925	0.878	0.088	0.088	0.858	0.859	0.106	0.078	0.888	0.839
0.104	0.082	0.937	0.912	0.069	0.069	0.880	0.888	0.078	0.064	0.899	0.873
0.042	0.041	0.936	0.933	0.038	0.038	0.925	0.923	0.041	0.039	0.933	0.923
0.052	0.050	0.925	0.922	0.045	0.045	0.901	0.903	0.050	0.045	0.922	0.904
0.058	0.057	0.921	0.921	0.051	0.051	0.906	0.905	0.056	0.051	0.918	0.906
0.046	0.045	0.935	0.931	0.041	0.041	0.918	0.917	0.045	0.042	0.930	0.918
0.014	0.014	0.941	0.941	0.014	0.014	0.937	0.937	0.014	0.014	0.941	0.940
0.011	0.011	0.946	0.946	0.011	0.011	0.944	0.944	0.011	0.011	0.946	0.946

continued

Table 3.4: Heterogeneous Simulation Results with $k = 5$. $\Delta = 0$ and $\lambda_1^2 = 1$ for all.

Overall, similar trends are observed as seen in Section 3.1. Coverage probabilities approach 95%, and variance estimates decrease as the sample sizes increase. Similarly, with bigger sample sizes, the results from the estimators exhibit negligible differences. Meta-analyses with fewer, bigger studies again show better results than those with several smaller studies when the degree of heterogeneity is similar– See for example cases 10 and 1 from Tables 3.3 and 3.4, respectively. Otherwise, bigger values of k report better coverage and precision when N_is are similar, as expected.

As in the homogeneous case, the plug-in estimator is seen consistently underestimating the variance and reporting poor coverage. The degree to which it underestimates the empirical variance is similar to that in homogeneous meta-analyses. However, when the studies vary in their sample sizes, the underestimation is slightly more severe under the heterogeneity assumption (e.g., case 3 from Table 3.1 and 5-6 from Table 3.3).

Contrary to what one might expect, coverages often increase with the degree of heterogeneity, particularly when the sample sizes are small. One can potentially attribute this phenomenon to the fact that the variance estimates also increase with heterogeneity. If the heterogeneity results in only a slight change in $\hat{\Delta}$, increased variance simply leads to a wider confidence interval, hence greater coverage. The variable that has the biggest influence on coverage is therefore not the degree of heterogeneity, but the sample size.

Empirical results again show that bias-corrected $\hat{\Delta}$s are almost always more precise than those calculated with biased weights, though the differences are scant. As in homogeneous meta-analyses, the bias-correction makes the biggest difference in the empirical variance when the studies vary in their sample sizes. However, this difference diminishes quickly when there are one or more larger studies ($N_i \approx 80$). The approximation methods on the other hand show mixed results under the heterogeneity assumption as well. N7 and N3 report worse coverage with bias-corrected $\hat{\Delta}$s in both Tables 3.3 and 3.4. The only exceptions to this trend are cases 5-7 in Table 3.3, where the studies vary in their sizes– In these cases, the bias-correction improves coverage. N5 often gives worse coverage probabilities with the correction particularly when N_is are small with $k = 2$. These differences seem to disappear with an increase in k, as Table 3.4 shows that the correction makes little change in N5's coverages.

Comparisons among the three approximation methods remain mostly the same as those in homogeneous meta-analyses. All three approximations generally report better coverage than the conventional plug-in estimator does, regardless of the sample size and the degree of heterogeneity. N7 provides the best coverage among all the estimators, reporting improvements greater than 5% compared to the plug-in estimator in some cases. Behind N7, N3 without the correction reports better coverage than N5 when the studies are small. The exception to this trend is when $(N_1, N_2) = (40, 5)$, where N5 without the bias-correction returns better coverage. N3 with the correction performs worse than N5 in general, however this trend reverses as the sample sizes increase ($N_i \approx 20$).

Chapter 4

Discussion

Senn et al. [15] (independently from Walter and Balakrishnan [20]) also raised concerns about the bias in study weights, suggesting that it can be a source of variability in treatment effect estimates. Their results are congruent with those presented here and in Walter and Balakrishnan [20]– That is, the conventional inverse variance method can cause serious digressions from the true weights when small studies are being integrated. The works in this thesis extend the propositions by Walter and Balakrishnan [20], exploring the effects of bias-correction on the precision of the summary treatment effect estimator $\hat{\Delta}$. Meta-analyses with normally distributed continuous outcomes sharing a common variance are considered, under a fixed effects model with the mean difference as the measure. Three new estimators for $\mathrm{Var}\left(\hat{\Delta}\right)$ are developed using the Taylor series expansion, and simulation studies were conducted in both homogenous and heterogenous settings.

The simulation results show that the conventional methods of using biased weights and $\widehat{\mathrm{Var}\left(\hat{\Delta}\right)} = 1/\Sigma\, \hat{w}_i$ lead to underestimation of the variance, particularly when the number of studies or the sample size is small. The N7 approximation reliably increases coverage probabilities compared to the conventional estimator in both homogenous and heterogenous outcomes. N5 and N3 are also available when the sample size requirement $N_i \geq 7$ is not met for all studies, although their performances vary depending on parameter values. On the other hand, the results are inconclusive with respect to the bias-correction. The estimator $\hat{\Delta}$ is empirically more precise with the correction. However, the bias-correction does not always lead to a better coverage probability when the variance is estimated with Taylor approximation. The changes are slight especially when the sample sizes are not too small ($N_i \approx 20$). Overall, one learns that sample sizes drive the coverage probabilities, more than the degree of heterogeneity or the bias-correction.

The methods presented here have some limitations, and there are many potential avenues for extensions of this work. For example, one can evaluate the methods presented here using real meta-analyses instead of simulated data. Walter and Balakrishnan's original paper evaluates the effects of the bias-correction using meta-analyses from the biomedical literature available in the Cochrane Database of Systematic Reviews [20]. The paper presents changes in the individual weights w_i, summary effect estimates $\hat{\Delta}$, and its variance with the correction. In addition to these initial results, the new Taylor-approximated estimators can give variance estimates with biased and corrected weights, and comparisons among the precision estimators can be made using published meta-analytic data.

The raw MD may not be an appropriate effect measure when studies report outcomes on different scales that cannot be easily converted. Cohen's d and Hedges' g, both forms of standardized MD, are often chosen as a measure in such circumstances [2, 4, 12]. The identification of optimal weights in meta-analyses using these two statistics is thus another potential path to supplementing the propositions here.

This thesis focuses on meta-analyses reporting continuous outcomes with a common variance, which does not cover the vast literature with different types of data. For instance, Song [18] investigates meta-analyses with binary outcomes and different association measures (e.g., log odds ratio, risk difference). Similar Taylor approximation methods are used to estimate the expected value of the individual study weights, according to which different bias-correction methods are suggested for each measure. Walter and colleagues are also studying heterogeneous ($\text{Var}(Y_1) \neq \text{Var}(Y_2)$) continuous outcomes in meta-analyses and associated biases in weights (unpublished) [21].

A fixed-effects model adopted in this work assumes that the studies in a meta-analysis all come from the same population and have identical true effect sizes. This assumption is implausible in most cases due to the diversity in study designs, demographics of the participants, and locations of the trial sites, among many factors that affect treatment effect estimates. Therefore, a random-effects model is generally more appropriate as it allows different effect sizes for each study [2, 5, 6]. Many have discussed optimal inference in meta-analyses under the random effects model, though they focus on improving the hypothesis tests associated with the conventional DerSimonian-Laird methods [5, 8–10, 13, 14, 16, 17]. Identification of optimal weights in random-effects meta-analyses could also contribute to these efforts, and even improve the estimation of the between-studies variance.

Study weights should not be treated as constants when they are in fact random variables whose estimations should be approached with rigour. In that, the works in this thesis make two contributions: The first involves further investigation into the bias-correction proposed by Walter [20]. The results show that the treatment effect estimator is empirically more precise with the correction than without. Situations where the correction has the biggest impact on the precision are also identified. The second contribution is the provision of three new estimators for the summary effect variance. The simulation study has demonstrated that the conventional inverse sum-of-weights estimator for the variance consistently underestimates the standard error, especially when the sample sizes are small. Thus, they return a confidence interval too narrow and a coverage too low. The Taylor-approximated estimators N7, N5, and N3 are shown to improve coverage probabilities in many cases.

Bibliography

[1] Peter Armitage, Geoffrey Berry, and J. N. S. Matthews. *Statistical methods in medical research*. Blackwell Scientific Publications, Oxford, 4th ed. edition, 2002.

[2] Michael Borenstein, Larry V. Hedges, Julian P. T. Higgins, and Hannah R. Rothstein. *Introduction to meta-analysis*. John Wiley Sons, Chichester, West Sussex, U.K., 2009.

[3] Cochrane. Analysing data and undertaking meta-analyses. In J. J. Deeks, J. P. T. Higgins, and D. G. Altman, editors, *Cochrane Handbook for Systematic Reviews of Interventions*, chapter 10. Cochrane, 6.4 edition, 2023.

[4] Jacob Cohen. *Statistical power analysis for the behavioral sciences*. L. Erlbaum Associates, Hillsdale, N.J, 2nd ed. edition, 1988.

[5] Rebecca DerSimonian and Nan Laird. Meta-analysis in clinical trials. *Controlled Clinical Trials*, 7(3):177–188, 1986.

[6] Joseph R. Dettori, Daniel C. Norvell, and Jens R. Chapman. Fixed-effect vs random-effects models for meta-analysis: 3 points to consider. *Global Spine Journal*, 12(7):1624–1626, 2022. PMID: 35723546.

[7] Gene V. Glass. Primary, secondary, and meta-analysis of research. *Educational Researcher*, 5(10):3–8, 1976.

[8] Joachim Hartung. An alternative method for meta-analysis. *Biometrical Journal*, 41(8):901–916, 1999.

[9] Joachim Hartung and Guido Knapp. On tests of the overall treatment effect in meta-analysis with normally distributed responses. *Statistics in Medicine*, 20(12):1771–1782, 2001.

[10] Joachim Hartung and Guido Knapp. A refined method for the meta-analysis of controlled clinical trials with binary outcome. *Statistics in Medicine*, 20(24):3875–3889, 2001.

[11] Larry V. Hedges. Meta-analysis. *Journal of Educational Statistics*, 17(4):279–296, 1992.

[12] Larry V. Hedges and Ingram Olkin. *Statistical methods for meta-analysis*. Academic Press, Orlando, 1985.

[13] Joanna IntHout, John P. A. Ioannidis, and George F Borm. The Hartung-Knapp-Sidik-Jonkman method for random effects meta-analysis is straightforward and considerably outperforms the standard DerSimonian-Laird method. *BMC Medical Research Methodology*, 14(1):25–25, 2014.

[14] Christian Röver, Guido Knapp, and Tim Friede. Hartung-Knapp-Sidik-Jonkman approach and its modification for random-effects meta-analysis with few studies. *BMC Medical Research Methodology*, 15(91):99–99, 2015.

[15] Stephen Senn, Susanne Schmitz, Anna Schritz, and Artur Araujo. A note regarding alternative explanations for heterogeneity in meta-analysis. *Statistics in Medicine*, 41(22):4501–4509, 2022.

[16] Kurex Sidik and Jeffrey N. Jonkman. A simple confidence interval for meta-analysis. *Statistics in Medicine*, 21(21):3153–3159, 2002.

[17] Kurex Sidik and Jeffrey N. Jonkman. On constructing confidence intervals for a standardized mean difference in meta-analysis. *Communications in Statistics. Simulation and Computation*, 32(4):1191–1203, 2003.

[18] Ge Song. *Identifiation of optimal study weights in meta-analyses with a binary outcome*. MSc thesis, McMaster University, Hamilton, Ontario, Canada, August 2024. Available at http://hdl.handle.net/11375/30216.

[19] Donna F. Stroup and Stephen B. Thacker. Meta-analysis. *Encyclopedia Britannica*, June 2024.

[20] Stephen D. Walter and Narayanaswamy Balakrishnan. A method was developed for correcting the bias in the usual study weights in meta-analyses. *Journal of Clinical Epidemiology*, 152:23–29, 2022.

[21] Stephen D. Walter, Narayanaswamy Balakrishnan, and Ge Song. Correcting the bias in the study weights for meta-analyses: the case of continuous outcome data with heterogeneous variances. [Unpublished Manuscript], 2024.