CHAPTER 12

Random-Effects Model

INTRODUCTION

In this chapter we introduce the random-effects model. We discuss the assumptions of this model, and show how these are reflected in the formulas used to compute a summary effect, and in the meaning of the summary effect.

THE TRUE EFFECT SIZES

The fixed-effect model, discussed above, starts with the assumption that the true effect size is the same in all studies. However, in many systematic reviews this assumption is implausible. When we decide to incorporate a group of studies in a meta-analysis, we assume that the studies have enough in common that it makes sense to synthesize the information, but there is generally no reason to assume that they are *identical* in the sense that the true effect size is *exactly the same* in all the studies.

For example, suppose that we are working with studies that compare the proportion of patients developing a disease in two groups (vaccinated versus a placebo). If the treatment works we would expect the effect size (say, the risk ratio) to be *similar but not identical* across studies. The effect size might be higher (or lower) when the participants are older, or more educated, or healthier than others, or when a more intensive variant of an intervention is used, and so on. Because studies will differ in the mixes of participants and in the implementations of interventions, among other reasons, there may be *different effect sizes* underlying different studies.

Or, suppose that we are working with studies that assess the impact of an educational intervention. The magnitude of the impact might vary depending on the other resources

Introduction to Meta-Analysis, Second Edition.
Michael Borenstein, Larry V. Hedges, Julian P.T. Higgins, and Hannah R. Rothstein.

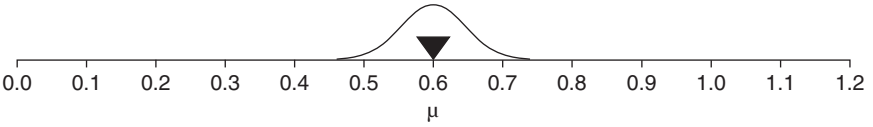

Figure 12.1 Random-effects model – distribution of the true effects.

available to the children, the class size, the age, and other factors, which are likely to vary from study to study.

We might not have assessed these covariates in each study. Indeed, we might not even know what covariates actually are related to the size of the effect. Nevertheless, logic dictates that such factors do exist and will lead to variations in the magnitude of the effect.

One way to address this variation across studies is to perform a *random-effects* meta-analysis. In a random-effects meta-analysis we usually assume that the true effects are normally distributed. For example, in Figure 12.1 the mean of all true effect sizes is 0.60 but the individual effect sizes are distributed about this mean, as indicated by the normal curve. The width of the curve suggests that most of the true effects fall in the range of 0.50 to 0.70.

IMPACT OF SAMPLING ERROR

Suppose that our meta-analysis includes three studies drawn from the distribution of studies depicted by the normal curve, and that the true effects (denoted θ_1, θ_2, and θ_3) in these studies happen to be 0.50, 0.55 and 0.65 (see Figure 12.2).

If each study had an infinite sample size the sampling error would be zero and the observed effect for each study would be the same as the true effect for that study. If we were to plot the observed effects rather than the true effects, the observed effects would exactly coincide with the true effects.

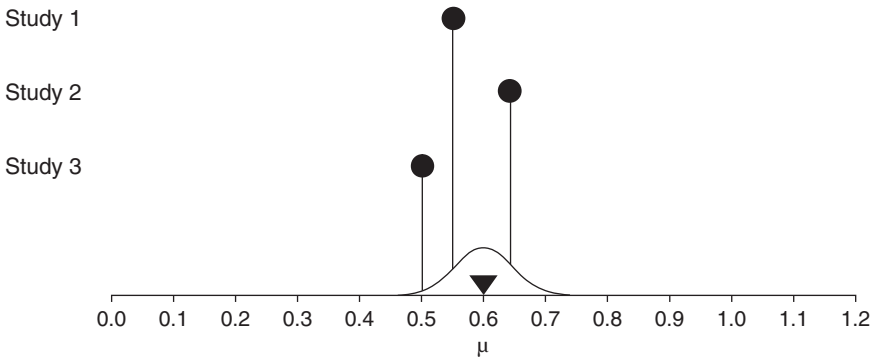

Figure 12.2 Random-effects model – true effects.

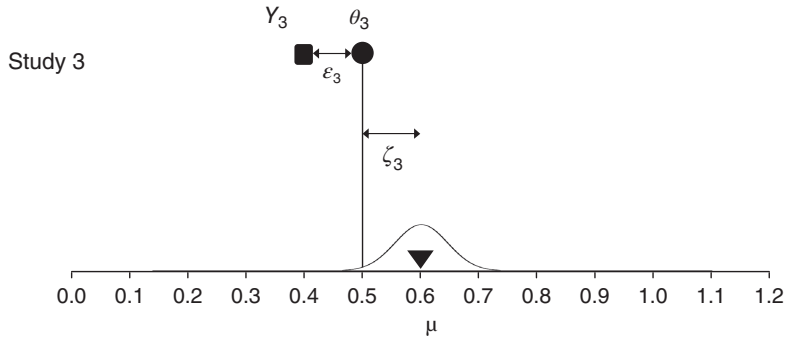

Figure 12.3 Random-effects model – true and observed effect in one study.

Of course, the sample size in any study is not infinite and therefore the sampling error is not zero. If the true effect size for a study is θ_i, then the observed effect for that study will be less than or greater than θ_i because of sampling error. For example, consider Study 3 in Figure 12.2. This study is the subject of Figure 12.3, where we consider the factors that control the observed effect. The true effect for Study 3 is 0.50 but the sampling error for this study is –0.10, and the observed effect for this study is 0.40.

This figure also highlights the fact that the distance between the overall mean and the observed effect in any given study consists of two distinct parts: true variation in effect sizes (ζ_i) and sampling error (ε_i). In Study 3 the total distance from μ to θ_3 is – 0.20. The distance from μ to θ_3 (0.60 to 0.50) reflects the fact that the true effect size actually varies from one study to the next, while the distance from θ_3 to Y_3 (0.5 to 0.4) is sampling error.

More generally, the observed effect Y_i for any study is given by the grand mean, the deviation of the study's true effect from the grand mean, and the deviation of the study's observed effect from the study's true effect. That is,

$$Y_i = \mu + \zeta_i + \varepsilon_i \tag{12.1}$$

Therefore, to predict how far the observed effect Y_i, is likely to fall from μ in any given study we need to consider both the variance of ζ_i, and the variance of ε_i

The distance from μ (the triangle) to each θ_i (the circles) depends on the standard deviation of the distribution of the true effects across studies, called τ (tau) (or τ^2 for its variance). The same value of τ^2 applies to all studies in the meta-analysis, and in Figure 12.4 is represented by the normal curve at the bottom, which extends roughly from 0.50 to 0.70.

The distance from θ_i to Y_i, depends on the sampling distribution of the sample effects about θ_i. This depends on the variance of the observed effect size from each study, V_{Y_i}, and so will vary from one study to the next. In Figure 12.4 the curve for Study 1 is relatively wide while the curve for Study 2 is relatively narrow.

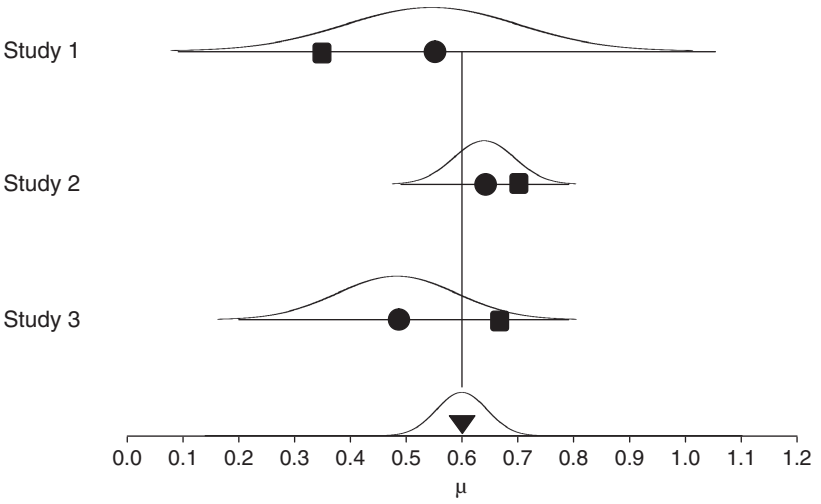

Figure 12.4 Random-effects model – between-study and within-study variance.

PERFORMING A RANDOM-EFFECTS META-ANALYSIS

In an actual meta-analysis, of course, rather than start with the population effect and make projections about the observed effects, we start with the observed effects and try to estimate the population effect. In other words our goal is to use the collection of Y_i to estimate the overall mean, μ. In order to obtain the most precise estimate of the overall mean (to minimize the variance) we compute a weighted mean, where the weight assigned to each study is the inverse of that study's variance.

To compute a study's variance under the random-effects model, we need to know both the within-study variance and τ^2, since the study's total variance is the sum of these two values. Formulas for computing the within-study variance were presented in Part 3. A method for estimating the between-studies variance is given here so that we can proceed with the worked example, but a full discussion of this method is deferred to Part 4, where we shall pursue the issue of heterogeneity in some detail.

Estimating tau-squared

The parameter τ^2 (tau-squared) is the between-studies variance (the variance of the effect size parameters across the population of studies). In other words, if we somehow knew the *true* effect size for each study, and computed the variance of these effect sizes (across an infinite number of studies), this variance would be τ^2. One method for estimating τ^2 is the method of moments (or the DerSimonian and Laird) method, as follows. We compute

$$T^2 = \frac{Q - df}{C}, \tag{12.2}$$

where

$$Q = \sum_{i=1}^{k} W_i Y_i^2 - \frac{\left(\sum_{i=1}^{k} W_i Y_i\right)^2}{\sum_{i=1}^{k} W_i}, \tag{12.3}$$

$$df = k - 1, \tag{12.4}$$

where k is the number of studies, and

$$C = \sum W_i - \frac{\sum W_i^2}{\sum W_i}. \tag{12.5}$$

Estimating the mean effect size

In the fixed-effect analysis each study was weighted by the inverse of its variance. In the random-effects analysis, too, each study will be weighted by the inverse of its variance. The difference is that the variance now includes the original (within-studies) variance plus the estimate of the between-studies variance, τ^2. In keeping with the book's convention, we use τ^2 to refer to the parameter and T^2 to refer to the sample estimate of that parameter.

To highlight the parallel between the formulas here (random effects) and those in the previous chapter (fixed effect) we use the same notations but add an asterisk (*) to represent the random-effects version. Under the random-effects model the weight assigned to each study is

$$W_i^* = \frac{1}{V_{Y_i}^*}, \tag{12.6}$$

where $V_{Y_i}^*$ is the within-study variance for study i plus the between-studies variance, T^2. That is,

$$V_{Y_i}^* = V_{Y_i} + T^2.$$

The weighted mean, M^*, is then computed as

$$M^* = \frac{\sum_{i=1}^{k} W_i^* Y_i}{\sum_{i=1}^{k} W_i^*}. \tag{12.7}$$

that is, the sum of the products (effect size multiplied by weight) divided by the sum of the weights.

The variance of the summary effect is estimated as the reciprocal of the sum of the weights, or

$$V_{M^*} = \frac{1}{\sum_{i=1}^{k} W_i^*}, \tag{12.8}$$

and the estimated standard error of the summary effect is then the square root of the variance,

$$SE_{M^*} = \sqrt{V_{M^*}}. \tag{12.9}$$

The 95% lower and upper limits for the summary effect would be computed as

$$LL_{M^*} = M^* - 1.96 \times SE_{M^*}, \tag{12.10}$$

and

$$UL_{M^*} = M^* + 1.96 \times SE_{M^*}. \tag{12.11}$$

Finally, a Z-value to test the null hypothesis that the mean effect μ is zero could be computed using

$$Z^* = \frac{M^*}{SE_{M^*}}. \tag{12.12}$$

For a one-tailed test the p-value is given by

$$p^* = 1 - \Phi(\pm|Z^*|), \tag{12.13}$$

where we choose '+' if the difference is in the expected direction or '–' otherwise, and for a two-tailed test by

$$p^* = 2[1 - (\Phi(|Z^*|))], \tag{12.14}$$

where $\Phi(Z^*)$ is the standard normal cumulative distribution. This function is tabled in many introductory statistics books, and is implemented in Excel as the function = NORMSDIST(Z^*).

Illustrative example

As before, we suggest that you turn to one of the worked examples in the next chapter before proceeding with this discussion.

SUMMARY POINTS

- Under the random-effects model, the true effects in the studies are assumed to have been sampled from a distribution of true effects.
- The summary effect is our estimate of the mean of all relevant true effects, and the null hypothesis is that the mean of these effects is 0.0 (equivalent to a ratio of 1.0 for ratio measures).
- Since our goal is to estimate the mean of the distribution, we need to take account of two sources of variance. First, there is within-study error in estimating the effect in each study. Second (even if we knew the true mean for each of our studies), there is variation in the true effects across studies. Study weights are assigned with the goal of minimizing both sources of variance.