This is chapter 6: Summary of statistical part (pp. 87-89) from

Schulze, R. (2004). *Meta-analysis: A comparison of approaches*. Hogrefe & Huber.

6
Summary of Statistical Part

In this part of the book, the statistical foundations of several approaches to meta-analysis of correlations have been outlined. The effect size database of interest in the present context was restricted to two families, the correlation coefficient and standardized mean effect sizes with a strong focus on the former. They still represent the most often used effect sizes in the social sciences and properties of estimators for both were therefore examined.

For the correlation coefficient as an effect size, the sample correlation coefficient and its properties were examined. The approximation introduced by Fisher (1921) was presented as a transformation of the correlation coefficient that shows a much more rapid convergence to a normal distribution in comparison to the correlation coefficient. Both estimators are biased and approximate formulae suggest a larger bias in absolute value for Fisher-z. However, the approximate variance of Fisher-z is independent of the population parameter to be estimated whereas the approximate variance of the correlation coefficient is not. For the latter, illustrations of this dependency were given. In addition to the common estimators r and Fisher-z, the unique minimum variance unbiased estimator introduced by Olkin and Pratt (1958) and its variance were presented. The variance was also shown to be dependent on the population parameter to be estimated.

From the d family, three estimators were presented of which d is considered the most important in the present context. It is, however, not the one which best attains desirable statistical properties like unbiasedness. The variance of d was also shown to depend on the the population parameter it is supposed to estimate, but the relationship is very different from the one presented for the correlation coefficients.

A brief examination of the conversion of the effect sizes presented along with revised formulae was given. It was concluded that available formulae may not hold for the nonnull case. The Monte Carlo study in Part III will provide evidence on this subject.

The methods of aggregating effect sizes were first presented in a general framework by specifying the statistical models of fixed effects and random effects. One important difference between the models lies in their assumptions about the distribution of effect sizes in the universe of studies. In the fixed effects model, homogeneity of all effect sizes or subgroups of effect sizes is often assumed. The fixed effects case represents a common assumption made in most applications of meta-analysis in practice but was criticized on various grounds (e.g., Erez et al., 1996; Hunter & Schmidt, 2000; National Research Council, 1992). It was pointed out in this context, that in the presence of heterogeneity application of the fixed effects model demands careful interpretation of the mean effect size. It has to be interpreted like a grand mean in ANOVA and may in some cases be ambiguous. This does not necessarily invalidate statements made on the basis of results from applying fixed effects models in heterogeneous situations. Whether ambiguity is indeed a problem, is a question to be answered by the researcher applying the models in a specific research situation.

In the random effects model, in contrast, heterogeneity of effect sizes is always an integral part of estimation as well as inference (see Hedges & Vevea, 1998). For both models, desired inference is different. In the fixed effects model, interpretation is restricted to studies like those available. In the random effects model, generalization of estimated characteristics of the effect size distribution leads to generalizations of effects to studies different from those examined but from the same research domain (Hedges, 1994b; Hedges & Vevea, 1998). As was pointed out, one important task for a researcher who wants to apply meta-analysis, is to carefully consider the model of the situation of interest and the desired inferences.

Additionally, the principles and concepts of applying mixture models to meta-analysis were outlined. It was pointed out that they provide a very flexible framework for the research situation of meta-analysis and were used to describe the research situations $\mathfrak{S}_1$ to $\mathfrak{S}_3$ that cover many important situations and will be used in Part III in the Monte Carlo study to systematize the design and presentation of results.

As another model class, hierarchical linear models were briefly introduced. These models are often used to assess the effect of observed explanatory variables on the effect size variability. It was shown that these models are very general and most other models can be regarded as special cases of hierarchical linear models.

The specific procedures of the various approaches to meta-analysis were outlined in detail in the subsequent chapter. As the major meta-analytical approaches for correlations as effect sizes, the approaches proposed by Hedges and Olkin (1985), Rosenthal and Rubin (1979), as well as Hunter and Schmidt (1990) were identified. In addition to these approaches, refinements were also presented that draw on the works of Hotelling (1953), Olkin and Pratt (1958), as well as DerSimonian and Laird (1986). All approaches are presented for application to correlations as effect sizes. In addition to these approaches, the procedures for d as an effect size presented by Hedges and Olkin (1985) were also outlined. Distinctions were drawn between the approaches with respect to the effect size to be aggregated (r, Fisher-z or d) and they were categorized according to the general framework introduced. The major approaches as well as OP, HOT, and OP-FE were identified as fixed effects approaches whereas DSL and OP-RE are random effects approaches. HS seemed to be of a hybrid type. Two major approaches, HOr and RR, are indistinguishable and may therefore not count as different approaches at all.

In the penultimate section of the last chapter, it was shown that the choice of an approach is at least consequential in situations $\mathfrak{S}_2$ and $\mathfrak{S}_3$, where heterogeneous situations are given. Fisher-z-based (HOr, RR, HOT, and DSL) and d-based (HOd) approaches were shown to estimate different parameters in comparison to r-based approaches. Since the expected value μ_ρ of the effect size distribution is considered to be the parameter of main interest in meta-analysis of correlations, cautions were raised about the application of Fisher-z based approaches in heterogeneous situations. Furthermore, the use of variances of the estimates in computing weights when the variances are confounded with the population parameters was pointed out to be a potential problem for the pooled estimators of the approaches. For HOd, the effect of applying such a weighting scheme is that the estimates are expected to be closer to μ_ρ than to the theoretically derived parameter without employing weights. In the case of OP-FE and OP-RE problems in estimation may arise. However, the r-based approaches retain the interpretation of the mean effect size estimate for μ_ρ in all situations because n is used in weighting the effect sizes and are therefore preferable in these situations from a theoretical point of view.

Finally, a comparison of approaches was presented that highlighted the major statistical attributes to classify the approaches as presented beforehand. These were the distinction between random versus fixed effects models, the use of transformed correlations, and the weighting scheme. Additionally, a brief overview of previous comparisons of meta-analytic approaches for correlations as effect sizes was presented.

In the following chapters, the design and results for a Monte Carlo study conducted for evaluation of the approaches will be presented. The situations introduced in the present chapter will be incorporated in the design and performance of the approaches with respect to the various estimates they propose will be evaluated.