Abstract: Parameter estimation in the random effects meta regression model

Meta analysis aims to combine the effect estimates of various related studies, trials, or experiments. It is a highly important statistical tool with various areas of application and has been a field of active research for many years (Hedges and Olkin, 1985), (Whitehead, 2002), (Hartung et al., 2008).

When combining different outcomes into an overall analysis, one cannot expect that each study specific outcome will be centred around the very same value. Such differences in location occur due to individual study specific design features. Usually, such differences get accounted for by including a heterogeneity parameter into the model. Sometimes, however, these differences can be explained by known study specific covariates. Additionally to adding another variance component, one can account for these location differences by including these study specific covariates into the model. This approach also allows to understand the rationale behind the observed differences. Berkey et al. (1995) called the resulting model the *random effects meta regression model*.

Depending on the underlying scientific question, different model parameters of the random effects meta regression model are of interest. Either the interest lies in accessing the amount of heterogeneity itself, as in (Hartung and Knapp, 2005) and (Knapp et al., 2006), or one is interested in the regression coefficients and, thus, aims to study the causes that may explain the occurring heterogeneity. The latter has been studied by Thompson and Sharp (1999) and Knapp and Hartung (2003). One objective of this dissertation is to test whether a construction based on generalised inference principles may improve the performance of statistical inference, especially when dealing with small sample sizes.

Generalised inference principles have already been successfully applied to inference on the overall treatment effect and between study variance in a meta analysis setting with Gaussian distributed responses in (Tian, 2008a) and (Tian, 2008b). The principle was coined under this term by Tsui and Weerahandi (1989). A concise introduction to the generalised inference principle is given in this dissertation. As a contribution to the theory of generalised inference, an extension of the principle to higher dimensions is developed which is original and has not been published yet.

The dissertation develops new point and new interval estimators for the heterogeneity and the regression coefficients based on generalised inference principles which are new and genuine. Different approaches and strategies are discussed including computational issues and generalisations to higher dimensions.

A unifying framework is constructed in which common point and interval estimators are integrated. For multiple point estimators of the heterogeneity parameter

of the random effects meta regression model, this unifying framework allows the formulation of new iterative procedures that have not been previously studied.

The dissertation is accompanied by a general purpose software package called metagen, (Möbius, 2014), written for the statistical software environment R, (R Development Core Team, 2010). The package provides access to the discussed methods and facilitates the set-up and realisation of large scale simulation studies evaluating the performance of these methods. It can be used as a statistical software tool for analysing data in the meta analysis or meta regression framework. Additionally, the package allows to reproduce the simulation studies and all results presented in this dissertation. It also allows to set up custom simulation studies including new methods which have not been coded into the base package of metagen. It, therefore, establishes a tool box for comparing a variety of different methods in a standardised way.

The dissertation includes two simulation studies evaluating the performances of all discussed methods. The specific set-up of the simulation studies has not been applied to the performance estimation of inferential methods of parametric models such as the random effects meta regression model, yet. The set-up and the results of these studies are presented and discussed thoroughly. These studies show that the new developed methods based on generalised inference perform well, and their performances are on comparable scales to the other presented methods. In particular, it is shown that the new generalised interval estimator for the regression coefficients is the second best performing method of all. The results of this dissertation also suggest which method to use as a default. The only other method that shows better performances than the new constructed generalised inference method is an interval estimator by Knapp and Hartung (2003) in conjunction with a heterogeneity estimator by Sidik and Jonkman (2005). The combination of these two constitutes a surprising result that was unknown prior to this dissertation.

The dissertation concludes with some final remarks and a discussion of open problems.

Thomas W. D. Möbius, Dortmund, May 9, 2014

References

Berkey, C. S., Hoaglin, D. C., Mosteller, F. and Colditz, C. A. (1995). A random-effects regression model for meta-analysis. *Statistics in Medicine, 14*, 395-411.

Hartung, J. and Knapp, G. (2005). On confidence intervals for the among-group variance in the one-way random effects model with unequal error variances. *Journal of Statistical Planning and Inference, 127*, 157-177.

Hartung, J., Knapp, G. and Sinha, B. K. (2008). *Statistical Meta-Analysis with Applications*. Hoboken: John Wiley & Sons.

Hedges, L. V. and Olkin, I. (1985). *Statistical Methods for Meta-Analysis*. Orlando: Academic Press.

Knapp, G., Biggerstaff, B. and Hartung, J. (2006). Assessing the amount of heterogeneity in random-effects meta-analysis. *Biometrical Journal, 48*, 271-285.

Knapp, G. and Hartung, J. (2003). Improved tests for a random effects meta-regression with a single covariate. *Statistics in Medicine, 22*, 2693-2710.

Möbius, T. W. D. (2014). *metagen: Inference in Meta Analysis and Meta Regression*. Accessible via http://00tau.github.io/metagen/. R package version 1.0.

R Development Core Team (2010). *R: A Language and Environment for Statistical Computing*. Vienna.: R Foundation for Statistical Computing

Sidik, K. and Jonkman, J. N. (2005). Simple heterogeneity variance estimation for meta-analysis. *Applied Statistical Science, 54*, 367-384.

Thompson, S. G. and Sharp, S. J. (1999). Explaining heterogeneity in meta-analysis: A comparison of methods. *Statistics in Medicine, 18*, 2693-2708.

Tian, L. (2008a). Generalized inferences on the overall treatment effect in meta-analysis with normally distributed outcomes. *Biometrical Journal, 50*, 237-247.

Tian, L. (2008b). Inferences about the between-study variance in meta-analysis with normally distributed outcomes. *Biometrical Journal, 50*, 248-256.

Tsui, K.-W. and Weerahandi, S. (1989). Generalized p-values in significance testing of hypotheses in the presence of nuisance parameters. *Journal of the American Statistical Association, 84*, 602-607.

Whitehead, A. (2002). *Meta-Analysis of Controlled Clinical Trials*. Hoboken: John Wiley & Sons.