

A new method for combining data from heterogeneous sources

Pär Villner*

June 2022

Abstract

In this thesis, we introduce a new method for summarizing data from heterogeneous sources. Assuming that there is a true data-generating model for a given phenomenon, we construct a statistical model that aims to include the true data-generating model as a special case. Inference on the parameters in the postulated model is made by performing simulations that replicate the data-generating process behind observed data. If the results from the simulations are sufficiently close to the observed data, the model is deemed plausible. The plausible model can then be used to calculate quantities of interest, such as risk measures.

The new method allows us to make inferences based on data from widely different sources. Examples of these could be summary statistics or raw data; results from research studies with heterogeneous experimental designs and study populations; general scientific facts or results of lab experiments.

In this thesis, we limit the applications to data where traditionally one would use meta-analysis. This is a set of methods that summarize the results of several studies by calculating the weighted average of reported intervention effects. We argue that meta-analysis faces disadvantages that the new method can avoid.

The performance of the new method is explored in two extensive simulation studies, and we also apply the method to data that was previously used in a traditional meta-analysis.

*Postal address: Mathematical Statistics, Stockholm University, SE-106 91, Sweden.
E-mail: par.villner@ki.se. Supervisor: Matteo Bottai.