

Generalized integration model for Improved Statistical Inference by Leveraging External Summary Data

Kai Yu

*National Cancer Institute
E-mail: yuka@mail.nih.gov*

Abstract: Meta-analysis has become a powerful tool for enhanced inference by gathering evidence from multiple sources. It pools summary-level data from different studies to improve estimating efficiency with the assumption that all participating studies are analyzed under the same statistical model. It is challenging to integrate external summary data calculated from different models with a newly conducted internal study in which individual-level data is collected. We develop a novel statistical inference framework based on a novel generalized integration model, which effectively synthesizes internal and external information for integrative analysis. The new framework is versatile enough to incorporate various types of summary data from multiple sources. We establish asymptotic properties for the proposed procedure and prove that the new estimate is theoretically more efficient than the internal data based maximum likelihood estimate, as well as a recently developed constrained maximum likelihood approach that incorporates the outside information. We illustrate an application of our method by evaluating cervical cancer risk using data from a large cervical screening program.