Contribution ID: **188** Type: **Oral Presentation**

Prediction Intervals for the Target Sample Size during Information-Based Monitoring

Wednesday 23 July 2025 17:30 (15 minutes)

Oral presentation

Prediction Intervals for the Target Sample Size during Information-Based Monitoring

Author

Ole Schacht, Tom Loeys, Beatrijs Moerkerke

Affiliation

Department of Data-Analysis, Ghent University

Abstract

A key feature of reproducible research in psychology is having a sufficiently large set of observations to support that reliable conclusions can be drawn from the data, which includes having enough statistical power. The current practice for sample size calculation in psychology typically entails specifying the a priori power probability with which a presumed effect is to be detected at a prespecified significance level. This also requires information on the nuisance parameter(s) of the test statistic at hand (such as the variance). Unfortunately, due to often limited information on these nuisance parameters at the study design stage, the sample size may easily be misjudged.

The last two decades, more flexible frequentist alternatives have been developed that enable re-estimation of the target sample size based on interim inspections of the variability (i.e., during ongoing data collection) without inflating the type I error. More precisely, nuisance parameters are monitored as new data comes in, and data collection is terminated only when the effect of interest can be sufficiently precisely estimated. A standard statistical test is then performed with the desired power at the prespecified significance level. Moreover, effect estimators remain unbiased, and no loss in efficiency is observed. Clearly, using such flexible methods may be attractive to applied researchers as accurate information on the variability in the data is no longer needed when planning a study or experiment.

While such approaches offer an adjustment for incorrect assessment of necessary study resources, it is also often criticized because researchers do not know the final sample size at the start of their study. It would therefore be desirable to have an estimator for the target sample size along with a measure of its uncertainty that can updated while data is accumulating. The target sample size that is estimated during the course of the study is a random variable that has not been studied yet. This unknown dispersion of the target sample size is often regarded as the main impediment of information-based monitoring.

In this talk we present a new, intuitively easy, and general approach to make interval predictions for the target sample size of a study. These prediction intervals are based on distributional properties of the Fisher information and can be made at arbitrary points throughout the sampling process. We demonstrate the approach in a typical setting where interest lies in the effect of a focal predictor in a regression model while adjusting for other covariates. We provide a user-friendly Shiny app to facilitate the usability of the prediction intervals.

Keywords

power, information-based monitoring, prediction interval

Primary authors: SCHACHT, Ole (Department of Data-Analysis, Ghent University); Prof. LOEYS, Tom (Department of Data-Analysis, Ghent University); Prof. MOERKERKE, Beatrijs (Department of Data-Analysis, Ghent University)

Presenters: SCHACHT, Ole (Department of Data-Analysis, Ghent University); Prof. LOEYS, Tom (Department of Data-Analysis, Ghent University); Prof. MOERKERKE, Beatrijs (Department of Data-Analysis, Ghent University)

Session Classification: Session 17 : "Sampling and Responses in experience sampling studies"

Track Classification: Design/Research methods: Design/Research methods