

Estimating Percentiles in Computer Experiments

Soma Roy

Ohio State University

Abstract:

Computer code is often used to study complex physical processes. The running of such code at a few chosen input settings comprises a computer experiment. The goal is to always use as few runs as possible to keep costs low. Most of the work done in this area focuses on either estimating the unknown complex input-output relationship or finding the maxima/minima of the output. My talk is about estimating percentiles in a computer experiments setting. I consider the case in which there are multiple inputs and a single output. I assume that the inputs have some known distribution and wish to estimate, say, the p^{th} percentile, Q_p , of the induced distribution of the output. The proposed algorithm uses a sequential design methodology. The code is observed at an initial design of a few carefully chosen points. These data are used to estimate the input-output relationship and Q_p . Next, I use specially developed criteria to sequentially add points to our design to improve our estimate of the relationship and refine the estimate of Q_p . I stop when the budget has been expended, and use all available data to obtain the final estimate Q_p^* and the input settings that would produce Q_p^* . I compare results from using a sequential design to those obtained from a fixed design approach. I also look at the case when two types of input variables may be present: control and environmental.