MIKLOS CSORGO, Carleton University

Randomized Empirical Processes with Applications to Infinite Super-Populations and Big Data Sets

A super-population outlook regards a finite population as a large imaginary random sample of N labeled units $\{X_1, \ldots, X_N\}$ of real valued random variables from a hypothetical infinite super-population. One may also view a Big Data set of univariate observations as if it were a large imaginary random sample of N labeled units. In both of these scenarios, instead of trying to process the entire data set, we sample it via its index set $\{1, \ldots, N\}$ and thus reduce the problem to dealing with significantly smaller sub-samples that, in turn, we study with the help of appropriately randomized empirical processes.