Stein's method and applications in high-dimensional statistics

organized by
Jay Bartroff, Larry Goldstein, Stanislav Minsker, and Gesine Reinert

Workshop Summary

Goals of the workshop and background. The workshop considered applications of Stein's method, in both statistics and probability, with a special focus on its use in high-dimensional statistical procedures. Since its inception nearly five decades ago, Stein's method has proven to be a valuable tool for distributional approximation, allowing for the handling of dependence and yielding finite sample bounds for approximating complicated distributions using their simpler limiting behavior. Such results are valuable in statistics in assessing the sample size required to achieve performance guarantees and for producing robust hypothesis tests and confidence regions. In addition, Stein identities have lead to the introduction of a number of novel statistical procedures, such as shrinkage techniques, ways to unbiasedly estimate the risk of certain estimators, methods for relaxing the Gaussian assumptions that are currently necessary in many commonly used procedures, and for assessing the quality of samples generated to approximate a target distribution by, say, Markov chain Monte Carlo. All of the tools mentioned are of ever increasing value in today's 'big data' environment.

The goals of the workshop were to expand the intersection of Stein's method and Stein-related methods in statistics and probability – to use advances on the probabilistic side to analyze existing, or suggest new, statistical procedures, and also to use modern statistical problems as inspiration for new directions on the probability side. In particular, our main aims were to make progress in the following broad areas:

(1) Empirical measures and dimension reduction;
(2) Sequential analysis and change-point detection;
(3) Concentration of measure inequalities and sparse recovery problems;
(4) Distributed estimation from heavy-tailed data and the rates of convergence to normal approximation.

Workshop structure. The mornings on each of the first three days had two lectures each, with each day's talks sharing a theme in order to stimulate interaction. On Monday, Larry Goldstein spoke about Stein's identity as the key to both Stein's shrinkage procedure and Stein's unbiased risk estimation (SURE), and Persi Diaconis spoke about generator method and exchangeable pair approaches to Stein's method. On Tuesday, Murat Erdogdu spoke about using discretized diffusion processes to minimize non-convex functions, and Jon Wellner spoke about rates of convergence to Chernoff's distribution. On Wednesday, Jackson Gorham talked about using Stein's method to assess sample quality, and Stas Minsker spoke about robust estimation for heavy-tailed distributions.

The afternoons were devoted to brainstorming sessions, group work on specific problems, and general discussions. By Tuesday afternoon we identified and discussed a list of roughly a dozen problems, and then through voting whittled the list down to 6 problems and formed groups to work on these. Through the remainder of the week we made concrete progress on all 6 of these problems, described below, and established new connections to continue working on them.

(1) *The limiting arcsine distribution for a class of statistics for gene expression data:* Here convergence to the arcsine law was proved, which was not known before, using a result of Hoeffding. The rate of convergence is still an open question, but the group put together a roadmap for its proof.
(2) *SURE:* Work centered on the relaxation of the Gaussian assumptions in Stein's Unbiased Risk Estimate. The working group produced a basis for proving extensions of SURE in non-Gaussian settings at the cost of a bias term whose magnitude could be expressed in terms of the behavior of the Stein kernel, or distances between the parent distribution and its zero bias transform.
(3) *The rate of convergence for a randomly stopped sum:* This group addressed the question of how a random stopping time affects the classical rate of convergence of a sum in the Central Limit Theorem. The group derived a correction term to the classical Berry-Esseen bound involving moments of the stopping time random variable, and also proved conversely that the rate can be arbitrarily slow without such restrictions on the moments.
(4) *Rates of convergence to normal approximation for stochastic approximation algorithms:* It is known since the work by B. Polyak and A. Juditsky (1992) that the accelerated stochastic approximation scheme for the problem of minimizing a twice continuously differentiable function yields a solution whose error is approximately normal. The group worked on the problem of quantifying the *rates of convergence* of the distribution of the error to normal approximation; for instance, such results can be useful for constructing non-asymptotic confidence balls for the solution. The group created a roadmap of the proof that is based on extending existing results for the rates of convergence in the central limit theorem for the martingale difference sequences to the higher dimensional case; such a result could be of independent interest.
(5) *Convergence of the Galton-Watson process:* Inspired by the result of Rényi and Yaglom that gives the exponential limiting distribution for the properly scaled population size of a critical Galton-Watson process conditioned on survival, this group proved the new result, potentially applicable in this or like settings, that if a nonnegative random variable X with finite mean satisfies the stochastic dominance relation $UX^s \leq_{\text{st}} X$ a.s., where U and X^s are independent standard uniform and an X-size-biased random variable respectively, then X has an exponential tail bound.
(6) *Does a Poisson tail bounds imply a bounded size biased coupling?* This group answered the question affirmatively, proving a new result that if a random variable X has Poisson-like concentration about its mean, then there exists a size biased coupling X^s such that $X^s \leq X + C$ a.s. This results provides a satisfying converse to a known result for $C = 1$.

Conclusions. The workshop was successful in forming coherent and motivated research groups that produced both new problems as well as progress towards, or in some case, complete proofs of new results. The workshop resulted in the creation of new collaborations and new research directions for the participants. This success is due in large part to the care taken by the AIM staff and their workshop structure which provides sufficient small-group and informal work time, and a valuable and practical method for having the participants share potential research problems, and for dividing them into productive working groups for producing results.