Title:

Discrepancy Theory for Randomized Controlled Trials

Abstract:

Discrepancy theory tells us that it is possible to partition vectors into sets that look surprisingly similar to each other. By "surprisingly similar," we mean much more similar than the sets produced by a random partition.

Randomized Controlled Trials are used to test the effectiveness of interventions, like medical treatments and educational innovations. Randomization is used to ensure that the test and control groups are probably similar. When we know nothing about the experimental subjects, a random partition into test and control groups is the best choice.

When we do have prior information about the experimental subjects, we can combine the strengths of randomization with the guarantees of discrepancy theory. This allows us to obtain more accurate estimates of the effectiveness of treatments, or to conduct trials with fewer experimental subjects.

We show how to use a recent advance in Algorithmic Discrepancy Theory, the Gram-Schmidt Walk algorithm of Bansal, Dadush, Garg, and Lovett, to construct balanced RCTs that produce more accurate estimates of treatment effects when treatment outcomes are correlated with linear functions of the covariates, and which can not be much worse than uniformly random RCTs in the worst case.

This is joint work with Chris Harshaw, Fredrik Sävje, and Peng Zhang.