Mitigating the Curse of Dimensionality of the Bayesian Beta-Bernoulli Bandit Problem

Peter Jacko
Lancaster University, UK

Abstract

Statistical testing based on randomized equal allocation is a widespread state-of-the-art approach in the design of experiments today, known as *randomized controlled trial* in biostatistics, *between-group design* in social sciences, and *A/B testing* in Internet marketing. Already in the 1930s, William R. Thompson, a biostatistician from Yale University, considering the case of Bernoulli outcomes, proposed an algorithm which would in expectation yield higher outcomes than (alternating) equal allocation. This was done in a Bayesian framework, using the Beta distribution for maintaining the belief about the success probability, and it has become known as the first formulation of the *bandit problem* in academic literature; the two-armed case in 1933, extended to the multi-armed case in 1935 by the same author. The practical application of the Bayesian approach has however been long hindered by its computational complexity. A variety of approximations (including Thompson's Bayesian posterior sampling and non-Bayesian algorithms based on a stay-with-the-winner property or on confidence bounds) have been developed and studied across several disciplines in order to overcome this issue, but failed to come sufficiently close to Bayes-optimality for finite horizon problems. We will show that, computationally, much larger problems can be solved to Bayes-optimality than what is commonly believed. In particular, using an efficient code in Julia programming language run on a standard laptop, a completely enumerated solution of the two-armed problem with a horizon of around $1,000$ periods can be computed in a few minutes and stored for an offline use, while even longer horizons can be solved by calculating optimal allocations in an online fashion.