An algorithm for the computation of the power of Monte Carlo tests with guaranteed precision

Patrick Rubin-Delanchy[1], **Axel Gandy**[1,⋆]

1. Deparment of Mathematics, Imperial College London, SW7 2AZ, London, UK
⋆Contact author: a.gandy@imperial.ac.uk

Keywords: Monte Carlo testing, power of a test, sequential testing.

The power of a hypothesis test is the probability of rejecting the null hypothesis if the data follows a given distribution. For Monte Carlo tests, for example bootstrap or permutation tests, this quantity must usually be evaluated by simulation. N datasets are generated from the specified data distribution, a Monte Carlo test is performed on each, and the frequency with which the null hypothesis is rejected serves as an estimate of the power of the test.

The existing research into computing the power of Monte Carlo tests focuses on obtaining an estimate that is as accurate as possible for a fixed computational effort. However, the methods that are proposed have no guarantee of precision, in the sense that they cannot report a (non-trivial) confidence interval for the power with a certain coverage probability.

In this presentation, we describe an algorithm that runs N Monte Carlo tests simultaneously and indefinitely until a user-specified confidence interval length and coverage probability is met. We demonstrate that under some minor regularity conditions the algorithm terminates in finite expected time, and discuss some optimisation issues.

We are currently augmenting the *R* **simctest** package to incorporate this work. The user must input a confidence interval length and coverage probability, and a function that generates N binary streams, where each value is one if the simulated statistic under the null is at least as extreme as the base test-statistic for this stream. The method returns an estimate of the power with a confidence interval that meets the user-specified requirements.