How to Generate Uniform Samples on Discrete Sets Using the Splitting Method

P. W. Glynn, A. Dolgin, R. Y. Rubinstein, and R. Vaisman

Probability in the Engineering and Information Sciences, Vol. 24, 405-422 (2010)

- GDolginRubinsteinVaisman10.pdf

The goal of this work is two fold. We show the following:

1. In spite of the common consensus on the classic Markov chain Monte Carlo (MCMC) as a universal tool for generating samples on complex sets, it fails to generate points uniformly distributed on discrete ones, such as that defined by the constraints of integer programming. In fact, we will demonstrate empirically that not only does it fail to generate uniform points on the desired set, but typically it misses some of the points of the set.

2. The *splitting*, also called the *cloning* method – originally designed for combinatorial optimization and for counting on discrete sets and presenting a combination of MCMC, like the Gibbs sampler, with a specially designed splitting mechanism—can also be efficiently used for generating uniform samples on these sets. Without introducing the appropriate splitting mechanism, MCMC fails. Although we do not have a formal proof, we guess (conjecture) that the main reason that the classic MCMC is notworking is that its resulting chain is not irreducible.We provide valid statistical tests supporting the uniformity of generated samples by the splitting method and present supportive numerical results.