SIAM J. OPTIM.
Vol. 12, No. 2, pp. 479–502

THE SAMPLE AVERAGE APPROXIMATION METHOD FOR STOCHASTIC DISCRETE OPTIMIZATION*

ANTON J. KLEYWEGT[†], ALEXANDER SHAPIRO[†], AND TITO HOMEM-DE-MELLO[‡]

Abstract. In this paper we study a Monte Carlo simulation–based approach to stochastic discrete optimization problems. The basic idea of such methods is that a random sample is generated and the expected value function is approximated by the corresponding sample average function. The obtained sample average optimization problem is solved, and the procedure is repeated several times until a stopping criterion is satisfied. We discuss convergence rates, stopping rules, and computational complexity of this procedure and present a numerical example for the stochastic knapsack problem.

Key words. stochastic programming, discrete optimization, Monte Carlo sampling, law of large numbers, large deviations theory, sample average approximation, stopping rules, stochastic knapsack problem

AMS subject classifications. 90C10, 90C15

PII. S1052623499363220

1. Introduction.

In this paper we consider optimization problems of the form

$$\min_{x\in\mathcal{S}} \{g(x) := \mathbb{E}_P G(x, W)\}. \tag{1.1}$$

Here W is a random vector having probability distribution P, $\mathcal{S}$ is a *finite* set (e.g., $\mathcal{S}$ can be a finite subset of $\mathbb{R}^n$ with integer coordinates), $G(x, w)$ is a real valued function of two (vector) variables x and w, and $\mathbb{E}_P G(x, W) = \int G(x, w)P(dw)$ is the corresponding expected value. We assume that the expected value function $g(x)$ is well defined, i.e., for every $x \in \mathcal{S}$ the function $G(x, \cdot)$ is measurable and $\mathbb{E}_P\{|G(x, W)|\} < \infty$.

We are particularly interested in problems with the following characteristics:

1. The expected value function $g(x) := \mathbb{E}_P G(x, W)$ cannot be written in a closed form, and/or its values cannot be easily calculated.
2. The function $G(x, w)$ is easily computable for given x and w.
3. The set $\mathcal{S}$ of feasible solutions, although finite, is very large, so that enumeration approaches are not feasible. For instance, in the example presented in section 4, $\mathcal{S} = \{0, 1\}^k$ and hence $|\mathcal{S}| = 2^k$; i.e., the size of the feasible set grows exponentially with the number of variables.

It is well known that many discrete optimization problems are hard to solve. Another difficulty here is that the objective function $g(x)$ can be complicated and/or difficult to compute even approximately. Therefore stochastic discrete optimization problems are difficult indeed and little progress in solving such problems numerically has been reported so far. There is an extensive literature addressing stochastic discrete optimization problems in which the number of feasible solutions is sufficiently small to

*Received by the editors November 1, 1999; accepted for publication (in revised form) May 14, 2001; published electronically December 14, 2001.

http://www.siam.org/journals/siopt/12-2/36322.html

[†]School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA 30332-0205 (Anton.Kleywegt@isye.gatech.edu, Alexander.Shapiro@isye.gatech.edu). The first author's work was supported by the National Science Foundation under grant DMI-9875400. The second author's work was supported by the National Science Foundation under grant DMS-0073770.

[‡]Department of Industrial, Welding and Systems Engineering, The Ohio State University, Columbus, OH 43210-1271 (homem-de-mello.1@osu.edu).