Proceedings of the 2020 Winter Simulation Conference
K.-H. Bae, B. Feng, S. Kim, S. Lazarova-Molnar, Z. Zheng, T. Roeder, and R. Thiesing, eds.

SAMPLE AVERAGE APPROXIMATION FOR FUNCTIONAL DECISIONS UNDER SHAPE CONSTRAINTS

Dashi I. Singham

Operations Research Department
Naval Postgraduate School
1411 Cunningham Road
Monterey, CA 93940, USA

Henry Lam

Dept. of Industrial Engineering & Operations Research
Columbia University
500 W. 120th Street
New York, NY 10027, USA

ABSTRACT

Sample average approximation methods are most often applied when the set of decision variables is finite. This research develops a method of finding optimal solutions to infinite-dimensional simulation optimization problems when the decision variable is a monotone function on a random variable used to model the uncertainty itself. This problem is motivated from approximately solving the principal-agent problem in economics, but also has close connections to nonparametric statistical estimation. We demonstrate how to approximate the infinite-dimensional problem with a discrete formulation that allows the use of standard sample average approximation methods. We also demonstrate how to utilize related bounding techniques on the optimal value, and show convergence results for the estimated optimal values and solutions.

1 INTRODUCTION

We present an approach for using sample average approximation (SAA) to solve problems with decision variables that are defined as functions. We focus on situations where the decision variable (function) to be optimized is monotone, defined over a random variable that itself models the uncertainty in the objective. When the random variable is continuous, this problem is in a sense infinite-dimensional. SAA is used to approximately solve a variety of optimization problems with expected value objectives, by replacing the expectation with a sample average. It is most useful when this expectation is difficult to evaluate in closed form or when only data or Monte Carlo samples are available. In our considered context, since the decision variable relates to the underlying random variable through a functional relation, our sampling equivalent to the expectation, in its implementable form, would involve an increasing number of real-valued decision variables as the sample size increases, which renders our setting different from the conventional one in SAA.

This paper was originally motivated by attempts to solve the principal-agent problem, which is a classical model in economics whereby a seller (principal) attempts to optimize her expected profit by selling a product to a heterogeneous population of buyers (agents). One form of the problem involves the principal incentivizing agents with differing valuations of a product to purchase it when there are infinite types of agents following a continuous distribution for their hidden true valuation of the product. There exists *information asymmetry* because only the agent knows his true valuation, though the principal knows the distribution of the valuations across the population of agents. The principal offers a menu of options to an agent listing different quantity and price combinations, and the agent picks the option that maximizes his value function. It is optimal for the principal to design an option for each type of agent, so if the types of agents follow a continuous distribution, the menu of options will be continuous functions. Thus, the problem is to derive the optimal price and quantity functions to offer the agents. For background on continuous principal-agent models, see Chapter 3 of Laffont and Martimort (2002).

978-1-7281-9499-8/20/$31.00 ©2020 IEEE

Traditionally, these continuous principal-agent problems were solved using simplified functions such that calculus could be used to find analytical solutions. More commonly, the agent types are assumed discrete, often taking only two possible values. Singham and Cai (2017) was the first attempt to use an SAA-type method to numerically approximate continuous principal-agent problems when these continuous settings failed to be analytically tractable. That method was further developed in Singham (2019), which used a bootstrap method to obtain approximate bounds on the optimal objective value. This paper generalizes the approach from the specifics of the principal-agent problem, and provides stronger bounds and new convergence results.

We describe more related literature. Shapiro et al. (2014) provide a comprehensive survey on SAA. Most of the results apply to formulations with finite-dimensional decision variables which, intuitively speaking, can be separately modeled from the randomness in the objective. In addition to convergence results on the solution or the objective value, including central limit theorems and concentration inequalities (e.g., Kleywegt et al. (2002)), data-driven bounding techniques have been derived to quantify the optimality gap or the quality of solutions. Mak et al. (1999) use a batching approach to estimate optimality gaps under minimal assumptions, a technique that we will later use. Bayraksan and Morton (2006) provide a single-replication approach that reduces the computational load of Mak et al. (1999). In the simulation literature, recent work along these lines include Jaiswal et al. (2018) who derive an optimal linear sample allocation rule when evaluating a sample average function over a grid of finite decision points, Lam and Qian (2018) who employ a bagging approach to obtain optimality gap bounds, and Phelps et al. (2016) who use SAA to obtain consistent solutions for an uncertain optimal control problem where the objective functional depends on stochastic parameters.

Regarding the problem of interest, our formulation can be viewed as a functional optimization. An example of related classical contexts include variational problems where the goal is to find an extremal function with certain smoothness conditions. For these types of problems, calculus of variations can be used to determine the optimal functionals (see Gelfand et al. (2000) for a reference). These problems, however, typically do not involve randomness as in stochastic optimization. Our problem is closer to statistical problems where the goal is to minimize the expected loss of a statistical procedure. These problems have a similar structure as ours, in that they typically use an empirical optimization as an approximation to obtain estimated parameters (such as in M-estimation). More relatedly, in the nonparametric setting, the decision variable, or the estimation or training target, can be an entire function. Examples include density estimation and nonparametric regression (Tsybakov 2008). Furthermore, the line of literature on shape-restricted inference considers estimating functions that satisfy certain monotonicity or convexity properties (Royset and Wets (2015), Groeneboom and Jongbloed (2014) and Guntuboyina and Sen (2018)). Compared to using kernels, these approaches have the advantages of not relying on bandwidth selection that could be challenging to well-tune in high dimensions. Moreover, these approaches can also provably speed up estimation convergence, as exemplified by results in adaptive risk bounds (Guntuboyina and Sen 2018). For example, Royset and Wets (2017) derive a unified framework for constrained M-estimation that considers many types of soft information to improve convergence. Lastly, our work is also related to Blanchet et al. (2017) who, motivated from applications in optimal transport and robust estimation, sought to find the worst-case joint distribution between two random variables. They apply an SAA approach and similarly face a growing dimension of decision variables as sample size increases.

In the following, we will describe our procedure to obtain nearly optimal solutions to optimization problems over the space of monotone functions. We also provide procedures to estimate the optimal objective value, and present consistency and convergence rate properties. These results add rigor to past work in Singham (2019) which was specific to the principal-agent problem, and relied on bootstrapping to calculate approximate bounds on the true objective. While Singham (2019) suggested consistency might be possible through numerical results, this paper formally derives consistency. Section 2 describes the continuous problem. Section 3 presents the corresponding SAA formulation. Section 4 describes the details of the computational algorithm to deliver solution estimates to the continuous problem. Section 5 derives bounds on the optimal values and convergence results, and Section 6 concludes.

2 PROBLEM FORMULATION

This section describes our main formulation. The decision variable is a continuous univariate function $x:\Theta \to \mathbb{R}$ defined over a single-dimensional random variable θ, the latter itself describing the uncertainty in the objective function. We assume θ lies in a bounded range $\Theta = [\underline{\theta}, \overline{\theta}]$, with density $f(\theta)$.

Our objective function is the expectation of the cost $c(x(\theta))$, which is a function of the function $x(\theta)$ evaluated at θ. We want to solve:

$$\Phi = \min_{x(\cdot)} \quad E[c(x(\theta))] = \int_{\underline{\theta}}^{\overline{\theta}} c(x(\theta)) f(\theta) d\theta \tag{1}$$

$$\text{subject to} \quad x(\cdot) \in \Omega. \tag{2}$$

We refer to this formulation as the "continuous formulation". Here, the feasible region Ω contains requirements on the geometry of the considered functions. In particular, we consider:

Assumption 1 Assume that Ω is the space of non-decreasing functions $x:\Theta \to \mathbb{R}$, and $c(\cdot)$ is Lipschitz continuous.

The continuous principal-agent problem is an example of a common problem meeting these assumptions. The random variable θ represents the demand of the agent for a product sold by the principal. The demand of a specific agent is unknown to the principal, but she knows the distribution $f(\theta)$ across different agent types. It has been shown that the optimal quantity and price functions that the principal offers the agent must be increasing in θ, with the intuition being that agents with higher demands will buy more of the product at an overall higher price (though a lower per unit price). Thus, the continuous principal-agent problem seeks to jointly optimize two univariate functions (quantity and price) over the space of the same random variable (agent demand). The cost function here takes the form $c(\theta, x_1(\theta), x_2(\theta))$, and in general the cost function $c(\cdot)$ could take many forms for different problems. But for our first attempt to generalize this type of SAA problem beyond the principal-agent context we focus on $c(x(\theta))$ since that is the simplest extension.

We note that our continuous formulation does not fall immediately into the typical SAA framework, as our decision variable is infinite-dimensional and, as sample size grows, the number of real-valued decision variables would correspondingly grow when we approximate the objective function of (1). Next, we present an algorithm for estimating Φ and then provide some bounding techniques and convergence guarantees in our SAA estimates.

3 DISCRETE SAA FORMULATION

We suppose that $f(\cdot)$ is unknown but samples of θ are available, or that the expectation cannot be obtained in closed form easily. We now formulate the corresponding SAA for Φ:

$$\hat{\Phi}_N = \min_{x(\cdot)} \quad \frac{1}{N}\sum_{n=1}^{N} c(x(\theta_n)) \tag{3}$$

$$\text{subject to} \quad x(\cdot) \in \Omega.$$

Here, $\theta_1, \ldots, \theta_N$ are N samples of θ drawn from density $f(\theta)$. The SAA formulation replaces the expectation with a sample average. Note that this is still an optimization problem with an infinite-dimensional decision variable $x(\cdot)$.

Next, we construct an equivalent formulation that has only a finite set of real-valued decision variables. Note that the objective function in (3) only depends on x through $x(\theta_n), n = 1,\ldots,N$. Thus, let us consider a new set of decision variables $x_1,\ldots,x_N$, corresponding to each sample of $\theta_n, n = 1,\ldots N$ in that $x_n = x(\theta_n)$ is the function value at θ_n. We consider

$$\Phi_N = \min_{\{x_n\}_{n=1,\ldots,N}} \quad \frac{1}{N}\sum_{n=1}^{N} c(x_n) \tag{4}$$

$$\text{subject to} \quad \{(\theta_n, x_n)\}_{n=1,\ldots,N} \in \Omega_N. \tag{5}$$

Here Ω_N is the constraint space for the pairs $\{(\theta_n,x_n)\}$. The constraint on $x(\cdot)$ imposed by Ω must be carried over to constraining the relations among $\{(\theta_n,x_n)\}$, such that an appropriate interpolation of $\{(\theta_n,x_n)\}$ falls in Ω. Intuitively, Ω_N is a relaxation of the space of Ω, whereby the constraints defining Ω_N need only hold at points $\{(\theta_n,x_n)\}$ as opposed to all $\{(\theta,x(\theta))\}$ values. Note that while the sampled points θ_n do not appear in the objective function in the relaxed problem in (4), they play an important role in the constraint in (5). Below we show some examples of constraint types that can be handled using this framework:

Nonnegativity: If Ω is the set of all nonnegative functions, then $\Omega_N = \{x_n \geq 0, n = 1,\ldots N\}$.

Monotonicity: If Ω is the set of all non-decreasing functions, then for θ_n ordered so that $\theta_n \leq \theta_{n+1}$ for $n = 1,\ldots,N-1$, we have $\Omega_N = \{x_n \leq x_{n+1}, n = 1,\ldots,N-1\}$.

Boundedness: If Ω is the set of functions bounded between $[l,u]$, then $\Omega_N = \{l \leq x_n \leq u, n = 1,\ldots,N\}$.

When Ω is any intersection of the above constraints, Ω_N would consist of the intersection of the corresponding discretized versions. Moreover, there are other constraints that can be handled, such as unimodality and convexity, but we only consider the above (especially monotonicity) as they are relevant to our problem.

We claim that $\hat{\Phi}_N = \Phi_N$ by showing that the optimal solution for each formulation can be converted into a feasible solution for the other formulation with the same objective value. It is easy to see that any solution for $\hat{\Phi}_N$ can be converted into one for Φ_N with the same objective value via a projection $x_n = x(\theta_n)$ on the points $\theta_n, n = 1,\ldots,N$. To see the other direction, we would need to show that, for any set of pairs $\{(\theta_n,x_n)\}$ in Ω_N, there exists an interpolation such that $x(\theta)$ lies in Ω. To facilitate discussion, define a function I as an interpolation rule such that $\tilde{x}(\cdot) = I(\{(\theta_n,x_n)\})(\cdot)$. That is, I applied to $\{(\theta_n,x_n)\}$ results in a function on a continuous range. We need $I(\{(\theta_n,x_n)\}) \in \Omega$ whenever $\{(\theta_n,x_n)\} \in \Omega_N$.

To this end, consider two examples of interpolation rules I: 1) linear interpolation and 2) piecewise constant, right-continuous. We can see which interpolations will maintain certain types of generic function constraints in Table 1. For example, if $\{(\theta_n,x_n)\}$ meets the nonnegativity constraint (i.e., all x_n are nonnegative), then both types of interpolations will also be nonnegative. For any specific constraint, we need to do a specific but straightforward check to see that the interpolation is feasible for the continuous problem.

From Table 1, we see that by using either linear or piecewise constant interpolation, we can convert a feasible solution in Ω_N into a corresponding feasible solution in Ω with the same objective value. Thus $\hat{\Phi}_N = \Phi_N$. We call Φ_N the discrete SAA formulation.

4 ALGORITHM

The discussion in Section 3 suggests to solve formulation Φ_N to obtain a set of pairs $\{(\theta_n,x_n)\}$, on which we interpolate to get a continuous function $x(\cdot)$ from $I(\{(\theta_n,x_n)\})$. This is summarized by the simple algorithm as follows:

1. Sample N values of θ from density $f(\theta)$ to obtain $\theta_1,\ldots,\theta_N$.

Table 1: Feasibility of constraints maintained after interpolation from Ω_N to Ω.

Constraint	Linear Interp	Piecewise Constant Interp
Nonnegativity	✓	✓
Unimodal	✓	
Monotonicity	✓	✓
Continuity	✓	
Bounded	✓	✓
Convexity	✓	

2. Solve the discrete formulation in (4)-(5) using these samples to obtain optimal $x_n^*, n = 1,\dots,N$, and the corresponding optimal objective Φ_N.
3. Interpolate the solution $\{(\theta_n, x_n^*)\}$ to obtain a feasible solution $\tilde{x}(\cdot) = I(\{(\theta_n, x_n^*)\})$, i.e., $\tilde{x}(\theta) \in \Omega$ for (1)-(2). I here can represent either linear or piecewise constant interpolation.
4. Evaluate the continuous objective function using $\tilde{x}(\theta)$, by way of numerical integration if the objective can be directly calculated:

$$\tilde{\Phi}_N = \int_{\underline{\theta}}^{\overline{\theta}} c(\tilde{x}(\theta)) f(\theta) d\theta \tag{6}$$

or using J new samples $\theta_1,\dots,\theta_J$ drawn from density $f(\theta)$ applied to $\tilde{x}(\theta)$ to estimate the continuous objective

$$\tilde{\Phi}_N^J = \frac{1}{J}\sum_{j=1}^{J} c(\tilde{x}(\theta_j)). \tag{7}$$

The first two steps involve sampling θ and solving the corresponding discrete problem. The third step interpolates the discrete solution to obtain a feasible solution to the continuous problem. The fourth step is needed only if we wish to evaluate the objective value at the attained feasible solution. We note that the solution of the discrete problem and the estimation of the objective function in (7) are standard calculations as part of SAA problems. Thus, this algorithm is as easy to implement as any SAA problem, though it becomes harder to solve as N increases due to the increasing number of decision variables and constraints.

5 BOUNDS AND CONVERGENCE

Next we establish bounds on Φ using our capability to solve Φ_N by employing the results in Mak, Morton, and Wood (1999). Such bounds are used commonly in the SAA literature and readily apply to our setting, as they do not require any assumptions on the form of decision variables and are very general. Thus, once we have seen that Φ_N is equivalent to $\hat{\Phi}_N$, we can use Mak, Morton, and Wood (1999) directly. This gives the following proposition. For completeness, we provide a proof that pinpoints exactly where we use $\hat{\Phi}_N = \Phi_N$:

Proposition 1 $E[\Phi_N] \le \Phi \le E[\tilde{\Phi}_N^J]$, and $E[\Phi_N]$ monotonically increases as N increases.

Proof: To show the upper bound, note that since $\tilde{x}(\theta)$ is a feasible solution for the continuous problem, plugging it into (6) will yield an upper bound $\Phi \le \tilde{\Phi}_N$. As the estimator used in (7) is unbiased, we have $E[\tilde{\Phi}_N^J] = \tilde{\Phi}_N$.

To show the lower bound, we consider

$$E[\Phi_N] = E[\hat{\Phi}_N] = E\left[\min_{x(\cdot)\in\Omega} \frac{1}{N}\sum_{n=1}^{N} c(x(\theta_n))\right] \le \min_{x(\cdot)\in\Omega} E\left[\frac{1}{N}\sum_{n=1}^{N} c(x(\theta_n))\right] = \min_{x(\cdot)\in\Omega} E[c(x(\theta))].$$

Additionally, to show $E[\Phi_N] \leq E[\Phi_{N+1}]$, we consider

$$E[\Phi_{N+1}] = E[\hat{\Phi}_{N+1}] = E\left[\min_{x(\cdot)\in\Omega} \frac{1}{N+1}\sum_{n=1}^{N+1} c(x(\theta_n))\right] = E\left[\min_{x(\cdot)\in\Omega} \frac{1}{N+1}\sum_{n=1}^{N+1} Z_N^{(n)}\right]$$
$$\geq \quad \frac{1}{N+1}\sum_{n=1}^{N+1} E\left[\min_{\{x(\cdot)\in\Omega} Z_N^{(n)}\right] = E[\hat{\Phi}_N] = E[\Phi_N]$$

where $Z_N^{(n)}$ is the leave-one-out sample average that removes the n-th sample, i.e., $Z_N^{(n)} = (1/N)\sum_{m\neq n} c(x(\theta_m))$.
□

Thus the simple algorithm presented in Section 4 can be used to provide bounds in expectation on the true optimal objective value. Singham (2019) derives bounds using a bootstrap method that develops an approximation to the objective Φ using a fixed number of samples N'. Then an SAA method using a bootstrap with smaller samples $N < N'$ provides a lower bound to the problem with size N'. The method employed in Proposition 1 is not only generalizable beyond the principal-agent problem, but is more direct in that it does not rely on the bootstrap approximation. The numerical examples in Singham and Cai (2017) and Singham (2019) show the bounds shrinking around the true optimal objective value as N increases using the bootstrap method.

Next we show convergence from Φ_N to Φ. Note that the standard SAA convergence results (Shapiro, Dentcheva, and Ruszczyński 2014) do not directly apply due to the dimensionality of the decision variables. However, we can use the underpinning empirical process to handle the problem. We derive two convergence results. First is the convergence of the estimated optimal value in the discrete SAA formulation, Φ_N, to the optimal value of the original problem, Φ. Second is the convergence of the obtained solution $\tilde{x}(\cdot)$ in the algorithm in Section 4, to the optimal in terms of the true objective function evaluation in the continuous formulation.

For these convergences to hold, we need a stronger assumption that $x(\cdot)$ is bounded. Thus, from now on, we strengthen Assumption 1 as:

Assumption 2 Assume that Ω is the space of non-decreasing functions $x : \Theta \to \mathbb{R}$, bounded in say $[l,u]$, and $c(\cdot)$ is Lipschitz continuous.

Note that by our discussion in converting Ω_N to Ω in the case of boundedness constraints in Section 3, all our previous results do not change by using the Ω in Assumption 2 instead of Assumption 1. We are ready to state our theorem:

Theorem 1 Under Assumption 2, we have

$$\Phi_N - \Phi = O_p\left(\frac{1}{\sqrt{N}}\right)$$

Proof: The class of bounded monotone functions is Donsker (Theorem 2.7.5 of Van Der Vaart and Wellner (1996)). Since x is bounded, we have a finite envelope, i.e., $E[\sup_x |c(x(\theta))|] < \infty$, and since $c(\cdot)$ is Lipschitz, $c(x(\cdot))$ is Donsker (Theorem 2.10.6 of Van Der Vaart and Wellner (1996)). Thus we have

$$\sqrt{N}\left(\frac{1}{N}\sum_{n=1}^{N} c(x(\theta_n)) - E[c(x(\theta))]\right) \Rightarrow v(x) \quad \text{in} \quad \ell^\infty(\Omega) \tag{8}$$

where $v(\cdot)$ denotes a mean-zero Gaussian process parametrized by $x \in \Omega$, with covariance

$$cov(v(x_1), v(x_2)) = cov(c(x_1(\theta)), c(x_2(\theta))),$$

and $\ell^\infty(\Omega)$ denotes the space of uniformly bounded functions from Ω to $\mathbb{R}$.

Now, consider $\Phi_N - \Phi = \hat{\Phi}_N - \Phi$. Suppose there is an optimal solution to the true continuous formulation given by $x^*(\cdot)$. We have

$$\hat{\Phi}_N - \Phi \leq \frac{1}{N}\sum_{n=1}^{N} c(x^*(\theta_n)) - \Phi$$

by the definition of $\hat{\Phi}_N$ as the optimal value for the discrete SAA, and the right hand side above satisfies

$$\sqrt{N}\left(\frac{1}{N}\sum_{n=1}^{N} c(x^*(\theta_n)) - \Phi\right) \Rightarrow N(0, var(c(x^*(\theta))). \tag{9}$$

On the other hand,

$$\hat{\Phi}_N - \Phi \geq \hat{\Phi}_N - E[c(\tilde{x}(\theta))]$$

by the definition of Φ as the optimal value for the continuous formulation. By (8), we have that the right hand side above satisfies

$$\hat{\Phi}_N - E[c(\tilde{x}(\theta))] \geq -\sup_x \left|\frac{1}{N}\sum_{n=1}^{N} c(x(\theta_n)) - E[c(x(\theta))]\right| = O_p\left(\frac{1}{\sqrt{N}}\right). \tag{10}$$

Putting these together, we thus have that

$$\hat{\Phi}_N - \Phi = O_p\left(\frac{1}{\sqrt{N}}\right).$$

Note that this above assumes $x^*(\cdot)$ exists. If this is not the case, we find an ε-near optimal solution and the arguments apply analogously. This concludes the theorem. □

Besides the convergence of the estimated optimal value of the discrete SAA formulation to the continuous optimal, we can also show the convergence of the solution in terms of its continuous objective value. This is presented in the next theorem:

Theorem 2 Under Assumption 2, we have

$$E[c(\tilde{x}(\theta))] - \Phi = O_p\left(\frac{1}{\sqrt{N}}\right).$$

Proof: Suppose there is an optimal solution to the true continuous formulation given by $x^*(\cdot)$. We consider $E[c(\tilde{x}(\theta))] - \Phi$, which is always non-negative by the definition of Φ. We write

$$\begin{aligned} & E[c(\tilde{x}(\theta))] - \Phi \\ = \ & \left(E[c(\tilde{x}(\theta))] - \frac{1}{N}\sum_{n=1}^{N} c(\tilde{x}(\theta_n))\right) + \left(\frac{1}{N}\sum_{n=1}^{N} c(\tilde{x}(\theta_n)) - \frac{1}{N}\sum_{n=1}^{N} c(x^*(\theta_n))\right) + \left(\frac{1}{N}\sum_{n=1}^{N} c(x^*(\theta_n)) - \Phi\right). \end{aligned}$$

The second term is non-positive by the optimality of $\tilde{x}$ for $\hat{\Phi}_N$. The third term satisfies (9). Analogous to (10), the first term satisfies

$$\left|E[c(\tilde{x}(\theta))] - \frac{1}{N}\sum_{n=1}^{N} c(\tilde{x}(\theta_n))\right| \leq \sup_x \left|\frac{1}{N}\sum_{n=1}^{N} c(x(\theta_n)) - E[c(x(\theta))]\right| = O_p\left(\frac{1}{\sqrt{N}}\right).$$

Thus, putting together, we have

$$E[c(\tilde{x}(\theta))] - \Phi = O_p\left(\frac{1}{\sqrt{N}}\right).$$

This concludes the theorem. □

Like the proof of Theorem 1, Theorem 2 assumes $x^*(\cdot)$ exists. If this is not the case, we find an ε-near optimal solution and the corresponding arguments apply analogously.

These bounds and convergence results can strengthen the algorithm in Section 4. The algorithm yields computed values of Φ_N in Step 2 and $\tilde{\Phi}_N^J$ in Step 4. Replicating the algorithm M times and averaging the computed values yields estimates of the lower bound $E[\Phi_N]$ and the upper bound $E[\tilde{\Phi}_N^J]$. The convergence results imply that larger values of N will yield values of Φ_N that are closer to the optimal value Φ, so N can be increased until the optimality gap estimate is sufficiently small. For multiple numerical examples using a specific implementation, see Singham (2019).

6 CONCLUSIONS

This work was inspired by attempts to solve the continuous principal-agent model, which determines the optimal quantity and price functions to offer to heterogeneous agents whose demand preferences are drawn from a continuous distribution. Past work presented detailed formulations, algorithms, and numerical results that are specialized to the principal-agent problem, and the observed convergence of upper and lower bounds motivated us to develop a generic algorithm and consider whether similar results hold in a broader infinite-dimensional setting.

We present a straightforward SAA algorithm to obtain approximate solutions to infinite-dimensional stochastic optimization problems when the decision variable is a monotone function over the random variable that defines the uncertainty in the objective. Our algorithm comprises an optimization with a finite number of real-valued decision variables that consistently represent the functional constraint in the original problem. We demonstrate how we can utilize bounding techniques in the SAA literature in this setting. We also show the canonical rate of the estimated optimal value using our algorithm in converging to the true optimal value, and our estimated solution in converging to optimal in terms of the objective value. The former also explains the observed convergence in the optimal values observed in past bootstrap experiments. Past numerical work suggests convergence of the function solution itself to the true optimal, and this will be considered in future work. Additionally, there are many potential applications aside from the principal-agent problem that will be explored, such as acceptance/rejection algorithms, and distributionally robust estimation.

ACKNOWLEDGMENTS

The authors are extremely grateful to Shane Henderson for noticing the underlying parallel between the principal-agent problem in Singham and Cai (2017) and the worst-case expectation problem in Blanchet, He, and Lam (2017), and to Raghu Pasupathy for helpful discussions on this problem. We also gratefully acknowledge support from the National Science Foundation under grants CAREER CMMI-1834710 and IIS-1849280.

REFERENCES

Bayraksan, G., and D. P. Morton. 2006. "Assessing Solution Quality in Stochastic Programs". *Mathematical Programming* 108(2-3):495–514.

Blanchet, J., F. He, and H. Lam. 2017. "Computing Worst-Case Expectations Given Marginals Via Simulation". In *Proceedings of the 2017 Winter Simulation Conference.*, edited by W. K. V. Chan, A. D'Ambrogio, G. Zacharewicz, N. Mustafee, G. Wainer, and E. Page, 2315–2323. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Gelfand, I. M., S. V. Fomin, and R. A. Silverman. 2000. *Calculus of Variations*. Mineola, NY: Dover Publications.

Groeneboom, P., and G. Jongbloed. 2014. *Nonparametric Estimation Under Shape Constraints*, Volume 38. Cambridge, UK: Cambridge University Press.

Guntuboyina, A., and B. Sen. 2018. "Nonparametric Shape-Restricted Regression". *Statistical Science* 33(4):568–594.

Jaiswal, P., H. Honnappa, and R. Pasupathy. 2018. "Optimal Allocations for Sample Average Approximation". In *Proceedings of the 2018 Winter Simulation Conference.*, edited by M. Rabe, A. Juan, N. Mustafee, A. Skoogh, S. Jain, and B. Johansson, 1874–1885. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Kleywegt, A. J., A. Shapiro, and T. Homem-de Mello. 2002. "The Sample Average Approximation Method for Stochastic Discrete Optimization". *SIAM Journal on Optimization* 12(2):479–502.

Laffont, J.-J., and D. Martimort. 2002. *The Theory of Incentives: The Principal-Agent Model*. 2nd ed. Princeton, NJ: Princeton University Press.

Lam, Henry and Qian, Huajie 2018. "Bounding Optimality Gap in Stochastic Optimization via Bagging: Statistical Efficiency and Stability". arXiv preprint arXiv:1810.02905, accessed September 9, 2020.

Mak, W.-K., D. P. Morton, and R. K. Wood. 1999. "Monte Carlo Bounding Techniques for Determining Solution Quality in Stochastic Programs". *Operations Research Letters* 24(1-2):47–56.

Phelps, C., J. O. Royset, and Q. Gong. 2016. "Optimal Control of Uncertain Systems Using Sample Average Approximations". *SIAM Journal on Control and Optimization* 54(1):1–29.

Royset, J. O., and R. J. Wets. 2015. "Fusion of Hard and Soft Information in Nonparametric Density Estimation". *European Journal of Operational Research* 247(2):532–547.

Royset, J. O., and R. J. Wets. 2017. "Variational Analysis of Constrained M-estimators". arXiv preprint arXiv:1702.08109.

Shapiro, A., D. Dentcheva, and A. Ruszczyński. 2014. *Lectures on Stochastic Programming: Modeling and Theory*. 2nd ed. Philadelphia, PA: SIAM.

Singham, D. 2019. "Sample Average Approximation for the Continuous Type Principal-Agent Problem". *European Journal of Operational Research* 275(3):1050–1057.

Singham, D., and W. Cai. 2017. "Sample Average Approximations for the Continuous Type Principal-Agent Problem: An Example". In *Proceedings of the 2017 Winter Simulation Conference.*, edited by W. K. V. Chan, A. D'Ambrogio, G. Zacharewicz, N. Mustafee, G. Wainer, and E. Page. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Tsybakov, A. B. 2008. *Introduction to Nonparametric Estimation*. New York: Springer Science & Business Media.

Van Der Vaart, A. W., and J. A. Wellner. 1996. *Weak Convergence and Empirical Processes*. New York, NY: Springer-Verlag.

AUTHOR BIOGRAPHIES

DASHI I. SINGHAM is a Research Associate Professor in the Operations Research Department at the Naval Postgraduate School where she researches, teaches, and advises student theses. Her main research interests lie in simulation analysis, methodology and applied statistics, and she has worked in the areas of sequential stopping rules for simulation output analysis, simulation optimization, and standardized time series. She obtained her Ph.D. in Industrial Engineering & Operations Research from the University of California at Berkeley and her B.S.E. in Operations Research and Financial Engineering from Princeton University. Her email address is dsingham@nps.edu, and her website is https://faculty.nps.edu/dsingham/.

HENRY LAM is an Associate Professor in the Department of Industrial Engineering and Operations Research at Columbia University. He received his Ph.D. degree in statistics from Harvard University in 2011, and was on the faculty of Boston University and the University of Michigan before joining Columbia in 2017. His research focuses on Monte Carlo simulation, uncertainty quantification, risk analysis, and stochastic and robust optimization. His email address is henry.lam@columbia.edu.