Sample Average Approximation and Model Predictive Control for Multistage Stochastic Optimization

Dominic S. T. Keehan*
University of Auckland

Andy B. Philpott†
University of Auckland

Edward J. Anderson‡
Imperial College London

Sample average approximation-based stochastic dynamic programming and model predictive control are two different methods of approaching multistage stochastic optimization. Model predictive control—despite a lack of theoretical backing—is often used instead of stochastic dynamic programming due to computational necessity. For settings where the stage reward is a convex function of the random terms, the stage dynamics are deterministic, and the random variables are stage-wise independent, we show that model predictive control is equivalent to a distributional robustification of stochastic dynamic programming with an ambiguity set that consists of distributions with matched means. This motivates tools to compare the out-of-sample performance of each method. We study a simple inventory control problem which illustrates their differences, and find that model predictive control can outperform stochastic dynamic programming when the distribution of the underlying random variable is skewed or has weight in its tails. The results are supported by analytic and numeric examples.

Keywords: multistage stochastic optimization, sample average approximation, model predictive control, distributionally robust optimization

*dkee331@aucklanduni.ac.nz
†a.philpott@auckland.ac.nz
‡e.anderson@imperial.ac.uk

1 Introduction

In practice multistage stochastic optimization problems often have to be solved without explicit knowledge of the distributions involved. Although one can create scenario-tree approximations of such problems based on samples of the random variables in each stage (called *sample average approximation* or SAA), the number of samples required to solve the true problem to a specified accuracy grows exponentially with the number of stages [14, 11] and the resulting optimization problems are computationally intractable for a large number of samples [4, 12]. It follows that for problems with a large number of stages, the SAA method may only be able to be practically applied when there is an insufficient number of samples to ensure good performance; we are interested in the performance of SAA in this regime.

Multistage stochastic optimization problems become easier when the random variables are stage-wise independent or follow a Markov process and the problem can be formulated as a stochastic optimal control problem. In principle, such problems are amenable to solution by stochastic dynamic programming methods as long as the dimension of the state variable is not too large. But this requires knowledge of the distribution of the random variables, whereas we are interested in the case where this must be deduced from samples of past values.

There are alternatives to the use of dynamic programming for stochastic optimal control problems. In many practical settings (e.g., where state dimension is high and controls and states are subject to complicated constraints) *model predictive control* (MPC) can be used. MPC algorithms are suboptimal control methods, which replace the future value function with a cheap-to-evaluate but not necessarily consistent approximation. This enables MPC algorithms to incorporate complicated constraints while being efficiently implementable in situations with a large number of state dimensions and stages [2]. For a survey of applications, see [7, 8].

In this paper we consider a relatively simple setting in which the state variables and constraints are deterministic, and the random variables appearing in the stage rewards are stage-wise independent. We study a simple form of MPC that fixes random variables at their expectation and solves a deterministic optimal control problem. (One can either assume that the expectations are known exactly, or estimate them from a random sample. We focus on the second case in this work.) This determines a control that is applied in the first period. The state then evolves and the process is repeated.

Out-of-sample comparison of SAA and MPC by simulation shows that MPC does well in certain circumstances [6]. However, the reasons for this have not been fully explored. Although the SAA and MPC solutions coincide when the certainty equivalence property holds [16, 17], this does not explain the success of MPC under more general conditions. Our aim in this paper is to advance our understanding of SAA and MPC applied to stochastic control problems when the number of samples is insufficient to ensure SAA's good performance. To do this we first show that MPC can be viewed through the lens of *distributionally robust optimization*, which is often applied in stochastic optimization problems when the number of samples is limited. By taking a conservative approach and optimizing over a set of distributions that are close to the empirical sample distribution, distributionally robust solutions can achieve better performance on the out-of-sample problem [5, 1].

To gain a deeper understanding, we then restrict attention to a specific class of stochastic inventory control problems with a one-dimensional state variable. This inventory problem involves decisions on how much of the existing inventory to sell at the current price, given uncertainty on the future prices that will occur. The problem is simple enough to admit a closed-form optimal policy for any bounded distribution for the uncertain random variable, but at the same time is able to capture critical aspects of the relative performance of SAA and MPC.

Given the inventory problem and some ground-truth distribution for the random (price) variable we can compute an optimal SAA policy for any given sample of the random variable and hence compute its expected value under the true distribution. Similarly, we can compute an optimal MPC policy based on the average of the same sample, and compute its expected value under the true distribution. These two out-of-sample values provide statistics that can be used to understand the sensitivities of each method to the specific samples from the random variable and the parameterization of the underlying problem. Furthermore, the expectation of these two statistics over the sampling distribution gives a measure of the average performance of each approach. Our study is motivated by the question:

> Under what conditions does model predictive control outperform stochastic dynamic programming based on sample average approximation?

We observe that the performance of SAA can suffer when the random variables have ground-truth probability distributions with significant weight in the tails. Small samples drawn from such distributions can have empirical distributions that look very different from the ground-truth, and yield SAA policies that can linger in regions of state-space with a low expected future reward which turns out to be very expensive over an infinite horizon.

In cases where the constraints are deterministic, the stage-reward is a convex function of the random variable, and the random variable is stage-wise independent, MPC attenuates this effect due to its distributionally robust properties—its conservative nature drives it towards regions of state-space with a future reward less dependent on randomness, and its performance under the ground-truth distribution is therefore less sensitive to the effects of sampling.

The paper is arranged as follows. We begin in Section 2 by formulating a general stochastic optimal control problem and showing that its solution via MPC is equivalent to a distributional robustification of its solution via stochastic dynamic programming. In Section 3 we study the out-of-sample value of a predetermined policy applied to the true problem and prove a lemma which allows us to compare the out-of-sample value of two different policies. Then in Section 4 we introduce our one-dimensional stochastic inventory control problem and derive a formula for its optimal solution as a function of the price probability distribution. This formula can be used to determine an optimal SAA policy based on the empirical distribution of price samples, as well as an optimal MPC policy based on the sample-average price. In Section 5 we compare the out-of-sample performance of these two policies under some simple assumptions on the ground-truth price distribution, and provide conditions on the price samples which ensure that the MPC policy performs at least as well as the SAA policy. In the final two sections we report on some examples. Section 6 assumes an exponential distribution for price and shows that the expected out-of-sample improvement from using MPC instead of SAA becomes arbitrarily large as the discount factor approaches 1. Section 7 includes some numerical experiments that support the theoretical results of previous sections. The paper then concludes with a discussion in Section 8.

2 Stochastic optimal control

To study the performance of SAA and MPC, we will look at the following stochastic optimal control problem

$$\text{SOC:} \quad \sup_{\{u^0,u^1,\ldots\}} \quad \mathbb{E}\left[\sum_{t=0}^{\infty} \beta^t C(x^t, u^t, \tilde{\xi}^t)\right]$$

where x^t is a state and u^t is a control satisfying

$$\begin{aligned} x^{t+1} &= f(x^t, u^t) \quad t = 0, 1, \ldots \\ u^t &\in \mathcal{U}(x^t), \quad t = 0, 1, \ldots, \end{aligned}$$

and u^t depends only on the previous states $x^1, x^2, \ldots, x^t$ and realizations of the random variables $\tilde{\xi}^1, \tilde{\xi}^2, \ldots, \tilde{\xi}^t$ (i.e. the standard non-anticipativity constraints). The value of x^0, the initial state at $t = 0$, is given. Here

$\beta \in (0,1)$ is a discount factor and C is a stage-reward function. Throughout this paper we use superscripts to denote indexing with time, though β^t is the t^{th} power of the stage discount factor β. We also use tildes to denote random variables.

Since the dynamics are deterministic, this is equivalent to a problem in which the control y^t is simply the state in the next period, defined by $y^t = f(x^t, u^t)$. Letting $\mathcal{Y}(x^t) = f(x^t, \mathcal{U})$ we can rewrite the dynamics as

$$\begin{aligned} x^{t+1} &= y^t, \quad t = 0, 1, \ldots \\ y^t &\in \mathcal{Y}(x^t), \quad t = 0, 1, \ldots, \end{aligned}$$

and we rewrite the stage rewards so that

$$C(x^t, y^t, \tilde{\xi}^t) = \sup_{\{u^t : f(x^t, u^t) = y^t\}} C(x^t, u^t, \tilde{\xi}^t),$$

where we assume that f is such that this supremum is achieved. Throughout our discussion we will use this form of the problem, in which the control y^t is simply the state at the next period.

To keep our analysis simple, we make the following assumption:

Assumption 1.

(i) The states x^t are contained in a compact and convex subset $\mathcal{X}$ of Euclidean space and the set valued function $\mathcal{Y} : \mathcal{X} \mapsto \mathcal{X}$ is non-empty, compact-valued, and continuous.

(ii) The multivariate random variables $\tilde{\xi}^t$ take values in an open subset Ξ of Euclidean space, and are stage-wise independent and identically

distributed with probability distribution P, and the expectation $\mathbb{E}[\tilde{\xi}^t]$ is well-defined and finite.[1]

(iii) The stage-reward function $C : \mathcal{X} \times \mathcal{X} \times \Xi \mapsto \mathbb{R}$ is continuous.

At times we will further require P to have compact support so that the random variables are bounded.

Under Assumption 1 (ii) we drop dependence of the random variable $\tilde{\xi}^t$ on the index t and denote each $\tilde{\xi}^t$ by $\tilde{\xi}$. For $(x, \xi) \in \mathcal{X} \times \Xi$ let the value function $V : \mathcal{X} \times \Xi \mapsto \mathbb{R}$ be defined through the dynamic programming functional equation

$$V(x, \xi) = \sup_{y \in \mathcal{Y}(x)} \left\{ C(x, y, \xi) + \beta \mathbb{E}\big[V(y, \tilde{\xi})\big] \right\}. \tag{1}$$

The connection between SOC and (1) is well known. Using the results of [15, Chapter 9], under Assumption 1 and if P has compact support, the identity (1) has a unique and continuous solution $V : \mathcal{X} \times \Xi \mapsto \mathbb{R}$ and SOC has a finite optimal value that is attained and equal to $\mathbb{E}[V(x^0, \tilde{\xi})]$. For a given $(x, \xi) \in \mathcal{X} \times \Xi$, the optimal control under dynamic programming is then any $y : \mathcal{X} \times \Xi \mapsto \mathcal{X}$ which attains the supremum in (1).

To determine the MPC policy first define the functional equation

$$V_{\mathrm{M}}(x, \xi) = \sup_{y \in \mathcal{Y}(x)} \left\{ C(x, y, \xi) + \beta V_{\mathrm{M}}(y, \mathbb{E}[\tilde{\xi}]) \right\}. \tag{2}$$

Notice that if $\mu = \mathbb{E}[\tilde{\xi}]$ is well-defined and finite, (2) shares the same existence and concavity properties as (1). Observe that a solution to (2) can be obtained through first solving the deterministic recursion $V_{\mathrm{M}}(x) =$

[1]Throughout this paper, in all probability spaces we assume that the corresponding σ-algebra is the Borel σ-algebra.

$\sup_{y\in\mathcal{Y}(x)}\{C(x,y,\mu)+\beta V_{\mathrm{M}}(y)\}$ in which it is assumed that $\tilde{\xi}$ always takes its expected value. Then (2) has value $\sup_{y\in\mathcal{Y}(x)}\{C(x,y,\xi)+\beta V_{\mathrm{M}}(y)\}$. For a given $(x,\xi)\in\mathcal{X}\times\Xi$, the optimal control under MPC is then any $y:\mathcal{X}\times\Xi\mapsto\mathcal{X}$ which attains the supremum in (2).

2.1 A distributionally robust interpretation

We now demonstrate a connection between MPC and distributionally robust optimization. Let $\mathcal{P}$ be an ambiguity set of probability distributions on Ξ. Now define the distributionally robust functional equation[2]

$$V_{\mathrm{R}}(x,\xi)=\sup_{y\in\mathcal{Y}(x)}\left\{C(x,y,\xi)+\inf_{Q\in\mathcal{P}}\beta\mathbb{E}_Q\left[V_{\mathrm{R}}(y,\tilde{\xi})\right]\right\}. \tag{3}$$

For $\xi\in\Xi$ denote by δ_ξ the point-mass distribution at ξ.

Theorem 1. *Under Assumption 1, if the function $C(x,y,\cdot)$ is convex for all $x,y\in\mathcal{X}$, the expectation $\mu=\mathbb{E}_P[\tilde{\xi}]$ is well-defined and finite, and $\mathcal{P}$ is chosen so that $\delta_\mu\in\mathcal{P}$ and $\mathbb{E}_Q[\tilde{\xi}]=\mu$ for all $Q\in\mathcal{P}$, then the DRO recursion (3) has the same solution as the MPC recursion (2).*

[2]In multistage distributionally robust optimization, there are two different approaches to formulating the robust problem: recursively evaluating the worst-case distributions in each stage, or evaluating the worst-case distributions for each stage in advance of any realised states and random variables [9]. In the setting of Theorem 1, it can be seen that both of these approaches are equivalent.

Proof. For any possible V_{R} satisfying (3) and $Q \in \mathcal{P}$ it follows that

$$
\begin{aligned}
&\mathbb{E}_Q\big[V_{\mathrm{R}}(x,\xi)\big] \\
&= \mathbb{E}_Q\left[\sup_{y\in\mathcal{Y}(x)}\left\{C(x,y,\xi) + \inf_{Q'\in\mathcal{P}} \beta\mathbb{E}_{Q'}\big[V_{\mathrm{R}}(y,\tilde{\xi})\big]\right\}\right] \\
&\geq \sup_{y\in\mathcal{Y}(x)}\left\{\mathbb{E}_Q\left[C(x,y,\xi) + \inf_{Q'\in\mathcal{P}} \beta\mathbb{E}_{Q'}\big[V_{\mathrm{R}}(y,\tilde{\xi})\big]\right]\right\} \\
&\geq \sup_{y\in\mathcal{Y}(x)}\left\{C(x,y,\mathbb{E}_P[\xi]) + \inf_{Q'\in\mathcal{P}} \beta\mathbb{E}_{Q'}\big[V_{\mathrm{R}}(y,\tilde{\xi})\big]\right\} \\
&= V_{\mathrm{R}}(x,\mu)
\end{aligned}
$$

where the second inequality follows due to convexity of $C(x, y, \cdot)$. But the probability distribution $\delta_\mu \in \mathcal{P}$, so from this inequality we can deduce that $\inf_{Q\in\mathcal{P}} \beta\mathbb{E}_Q[V_{\mathrm{R}}(x,\tilde{\xi})]$ is attained by δ_μ and has value $\beta V_{\mathrm{R}}(x,\mu)$. Hence $\inf_{Q\in\mathcal{P}} \beta\mathbb{E}_Q[V_{\mathrm{R}}(y,\tilde{\xi})] = \beta V_{\mathrm{R}}(y,\mu)$. It follows that the DRO recursion (3) is

$$
V_{\mathrm{R}}(x,\xi) = \sup_{y\in\mathcal{Y}(x)} \{C(x,y,\xi) + \beta V_{\mathrm{R}}(y,\mu)\}
$$

which is equivalent to the MPC recursion (2). Lastly, under Assumption 1 and with $\mathbb{E}_P[\tilde{\xi}]$ finite, we know that a solution to (2) exists, concluding the proof. □

Theorem 1 shows that the DRO value function and optimal policy is the same as the MPC value function and optimal policy for an ambiguity set chosen to have a matched mean. Use of ambiguity sets in which one or more moments are fixed is a classical approach in distributionally robust optimization (see the seminal work [10]). It is easy to see that Theorem 1 also holds in the finite-horizon setting or when the distributions of $\tilde{\xi}_t$ differ between stages.

Since distributionally robust estimates can be less susceptible to overfitting than sample average estimates, Theorem 1 motivates us to study the conditions under which the MPC policy outperforms the SAA policy. We develop tools for this analysis in the next section.

3 Comparing policies

When studying out-of-sample performance for specific policies, we do not need to use the dynamic programming functional equation and we can drop the requirement for P to have compact support. We make the following Assumption 2 which replaces the compact support requirement with an integrability condition when Ξ is unbounded.

Assumption 2. *There exists a positive valued random variable $L(\tilde{\xi})$ with $\mathbb{E}[L(\tilde{\xi})] < \infty$, and $|C(x, y, \xi)| \leq L(\xi)$ for all $x, y \in \mathcal{X}$ and almost every $\xi \in \Xi$.*

Our approach to compare two different policies is to consider starting with one policy and then switching to the other policy after a certain number of stages. A policy of interest $\pi : \mathcal{X} \times \Xi \mapsto \mathcal{X}$ is said to be *feasible* if it satisfies the constraint $\pi(x, \xi) \in \mathcal{Y}(x)$ on $\mathcal{X} \times \Xi$ and is measurable. Observe that for a feasible policy under Assumption 1 the function $C(x, \pi(x, \cdot), \cdot)$ is measurable for all $x \in \mathcal{X}$ due to continuity. Let $\bar{V}_\pi(x^0)$ be the value of the SOC objective function under the policy π starting from initial state x^0. We have

$$\bar{V}_\pi(x^0) = \mathbb{E}\left[\sum_{t=0}^{\infty} \beta^t C(x^t, \pi(x^t, \tilde{\xi}^t), \tilde{\xi}^t)\right] = \sum_{t=0}^{\infty} \beta^t \mathbb{E}\left[C(x^t, \pi(x^t, \tilde{\xi}^t), \tilde{\xi}^t)\right],$$

where the exchange of expectation and infinite summation follows from Assumption 2 and the Lebesgue Dominated Convergence Theorem using $L(\tilde{\xi})$. Since each term in the sum on the right-hand side is bounded and $\beta < 1$, it follows that $\bar{V}_\pi$ is also bounded on $\mathcal{X}$.

Having defined $\bar{V}_\pi$ as a function of the initial state, it satisfies the functional equation

$$\bar{V}_\pi(x) = \mathbb{E}\left[C(x, \pi(x, \tilde{\xi}), \tilde{\xi}) + \beta \bar{V}_\pi(\pi(x, \tilde{\xi}))\right].$$

Definition 2. *For feasible policies $\pi, \tau : \mathcal{X} \times \Xi \mapsto \mathcal{X}$, let*

$$\bar{V}^1_{(\pi,\tau)}(x) = \mathbb{E}\left[C(x, \pi(x, \tilde{\xi}), \tilde{\xi}) + \beta \bar{V}_\tau(\pi(x, \tilde{\xi}))\right],$$

and for $t > 1$, let

$$\bar{V}^t_{(\pi,\tau)}(x) = \mathbb{E}\left[C(x, \pi(x, \tilde{\xi}), \tilde{\xi}) + \beta \bar{V}^{t-1}_{(\pi,\tau)}(\pi(x, \tilde{\xi}))\right]. \tag{4}$$

The value $\bar{V}^t_{(\pi,\tau)}(x)$ is the out-of-sample value starting from initial state x if the policy π is used for the first t stages and then the policy τ is used forevermore. It is clear that $\bar{V}^t_{(\pi,\tau)}$ is well-defined and bounded in the same way that $\bar{V}_\pi$ and $\bar{V}_\tau$ are.

Proposition 3. *Under Assumptions 1 and 2, if $\pi, \tau : \mathcal{X} \times \Xi \mapsto \mathcal{X}$ are feasible* *policies and $x \in \mathcal{X}$, then*

$$\lim_{t\to\infty} \left|\bar{V}^t_{(\pi,\tau)}(x) - \bar{V}_\pi(x)\right| = 0.$$

Proof. The terms $\bar{V}^t_{(\pi,\tau)}(x^0)$ and $\bar{V}_\pi(x^0)$ are both out-of-sample values when implementing the policy π for the first t periods starting at state x^0. So

$$\left|\bar{V}^t_{(\pi,\tau)}(x^0) - \bar{V}_\pi(x^0)\right| = \left|\beta^t \mathbb{E}\left[\bar{V}_\tau(x^t) - \bar{V}_\pi(x^t)\right]\right| \tag{5}$$

where the expectation is with respect to the state x^t after t applications of the policy π on the previous states and realizations of $\tilde{\xi}$. Since $\bar{V}_\pi(\cdot)$ and $\bar{V}_\tau(\cdot)$ are bounded and $\beta^t \to 0$ as $t \to \infty$, the right-hand side of (5) goes to 0 as $t \to \infty$. Replacing x^0 with x concludes the proof. □

Lemma 4. *Under Assumptions 1 and 2, if $\pi, \tau : \mathcal{X} \times \Xi \mapsto \mathcal{X}$ are feasible policies and $\bar{V}_\tau(x) \geq \bar{V}^1_{(\pi,\tau)}(x)$ for all $x \in \mathcal{X}$, then $\bar{V}_\tau(x) \geq \bar{V}_\pi(x)$ for all $x \in \mathcal{X}$.*

Proof. We make the inductive hypothesis: $\bar{V}^{t-1}_{(\pi,\tau)}(x) \geq \bar{V}^t_{(\pi,\tau)}(x)$ for all $x \in \mathcal{X}$. Now, $\pi(x,\xi) \in \mathcal{X}$ for all $(x,\xi) \in \mathcal{X} \times \Xi$ since π is feasible. By the hypothesis, $\bar{V}^{t-1}_{(\pi,\tau)}(\pi(x,\xi)) \geq \bar{V}^t_{(\pi,\tau)}(\pi(x,\xi))$ and it follows that

$$\begin{aligned} \bar{V}^t_{(\pi,\tau)}(x) &= \mathbb{E}\left[C(x, \pi(x,\tilde{\xi}), \tilde{\xi}) + \beta \bar{V}^{t-1}_{(\pi,\tau)}(\pi(x,\tilde{\xi}))\right] \\ &\geq \mathbb{E}\left[C(x, \pi(x,\tilde{\xi}), \tilde{\xi}) + \beta \bar{V}^{t}_{(\pi,\tau)}(\pi(x,\tilde{\xi}))\right] = \bar{V}^{t+1}_{(\pi,\tau)}(x) \end{aligned} \tag{6}$$

for all $x \in \mathcal{X}$. Equation (6) is the inductive step, and the condition in the statement of the Lemma is the base case of the induction. Hence, $\bar{V}_\tau(x) \geq \bar{V}^t_{(\pi,\tau)}(x)$ for all t and $\bar{V}_\tau(x) \geq \lim_{t\to\infty} \bar{V}^t_{(\pi,\tau)}(x) = \bar{V}_\pi(x)$, where Proposition 3 yields the final equality. □

Rather than having to calculate integrals directly to compare out-of-sample values, Lemma 4 allows us to check a uniform condition involving recursions which have similar forms. This will prove useful in Section 5.

4 A stochastic inventory control problem

To study the performance of SAA and MPC, we will look at a particular stochastic control problem within the general stochastic control setting developed in Section 2. We formulate the one-dimensional *stochastic inventory control* (SIC) problem

$$\text{SIC:} \quad \max_{\{y^1,y^2,\ldots\}} \quad \mathbb{E}\left[\sum_{t=0}^{\infty} \beta^t \left(\tilde{p}^t(x^t - y^t) - C(y^t)\right)\right]$$

where x^t and y^t satisfy

$$\begin{aligned} x^{t+1} &= y^t, \quad t = 0, 1, \ldots \\ y^t &\in \left[0, x^t\right], \quad t = 0, 1, \ldots \end{aligned}$$

and $x^0 \geq 0$. As with the general problem SOC a decision at time t is made depending only on $x^1, x^2, \ldots, x^t$ and realizations of $\tilde{p}^1, \tilde{p}^2, \ldots, \tilde{p}^t$. Here $\tilde{p}^t$ is a random per-unit sales price (henceforth price) that is independent and identically distributed having distribution P with finite expectation, and

$C : \mathbb{R}_+ \mapsto \mathbb{R}_+$ is a storage cost function. The function C is increasing, strictly convex and differentiable with derivative c and $C(0) = 0$. Because c is a strictly increasing continuous function, we may define an inverse function, c^{-1}, on the range of c. We can easily check that Assumptions 1 and 2 hold for this example.

The problem SIC can be interpreted as the problem facing a merchant who maximizes expected discounted reward by at each time t selling at a price realization p^t down to an inventory level y^t less than or equal to their initial inventory x^t, while incurring a storage cost $C(y^t)$ on their remaining inventory. This model may be applied in a number of situations, such as an electricity distributor with a fixed battery needing to decide when to dispatch electricity or an investor who needs to decide when to sell a holding of shares.

For inventory $x \geq 0$ and price p, the optimum expected discounted reward from this point on can be found from the dynamic programming functional equation

$$V(x, p) = \max_{0 \leq y \leq x} \left\{ p(x - y) - C(y) + \beta \mathbb{E} \left[V(y, \tilde{p}) \right] \right\}, \tag{7}$$

where the optimal inventory is given by any maximizing y. When P has compact support, using the results of [15, Chapter 9] equation (7) has a unique and continuous solution $V : \mathbb{R}_+ \times \mathbb{R} \mapsto \mathbb{R}$ and the function $V(\cdot, p)$ is concave. With P having compact support, a simple application of the Dominated Convergence Theorem shows $\mathbb{E}[V(\cdot, \tilde{p})]$ is continuous, and it is also concave. It follows that the superdifferential $\partial_x \mathbb{E}[V(x, \tilde{p})]$ is non-empty and compact.

Denote the projection of $f \in \mathbb{R}$ onto the closed interval $[a, b] \subset \mathbb{R}$ by

$(f)_{[a,b]} = \max\{a, \min\{b, f\}\}$. We let $(f)_{[a,\infty)} = \max\{f, a\}$ and $(f)_+ = \max\{f, 0\}$.

Proposition 5. *If P has compact support, for inventory $x \geq 0$ and price p, SIC has optimal solution*

$$y(x,p) = c^{-1}\left((\beta\mathbb{E}[(\tilde{p}-p)_+] + \beta p - p)_{[c(0),c(x)]}\right).$$

Proof. Given the value function V, for $y \geq 0$ let $\bar{V}(y) = \mathbb{E}\left[V(y,\tilde{p})\right]$ and

$$\varphi(y) = p(x-y) - C(y) + \beta\bar{V}(y).$$

The function $\varphi(\cdot)$ is strictly concave and continuous, so the optimal choice of y on the right hand side of the recursion (7), which is given by $\max_{0\leq y\leq x}\varphi(y)$, is attained by a unique $y_*(x,p) \in [0,x]$. Since the derivative function $-c$ is strictly decreasing and unbounded below, the strictly concave function φ is decreasing for y large enough and there is a unique solution $y(p)$ to $\max_{y\geq 0}\varphi(y)$ which is equal to $y_*(x,p)$ when projected onto $[0,x]$. Observe that the function $y(\cdot)$ is decreasing. For $x \geq 0$, let $p_*(x)$ be any value such that for $p \leq p_*(x)$ the term $y(p) \geq x$ and for $p \geq p_*(x)$ the term $y(p) \leq x$. If x is an upper or lower bound of the function y, such a $p_*(x) \in \mathbb{R}$ does not exist. In this case we set $p_*(x) = -\infty$ or ∞ respectively.

In order to find the solution for the optimizer $y(p)$ we need to use the derivative of $\phi(y)$. Our approach is to take derivatives in the recursion and use this to solve for the derivative of V. In fact we will need to work with the superdifferential since V may not be smooth. Denote by $\partial_x V(x,p)$ the superdifferential of the function $V(\cdot,p)$ at x and p. When $p \geq p_*(x)$, the term $y(p) \leq x$, so $y_*(x,p) = y(p)$ and

$$V(x,p) = p(x - y(p)) - C(y(p)) + \beta\bar{V}(y(p)).$$

In this case it follows that $p \in \partial_x V(x,p)$. On the other hand, when $p \leq p_*(x)$ the term $y(p) \geq x$, so $y_*(x,p) = x$ and

$$V(x,p) = -C(x) + \beta\bar{V}(x). \tag{8}$$

If $x > 0$, (8) implies that

$$-c(x) + \beta\partial_x\bar{V}(x) \subseteq \partial_x V(x,p).$$

Now, $\partial_x\bar{V}(x)$ is non-empty. Any $\bar{g} \in \partial_x\bar{V}(x)$ then defines a supergradient $-c(x) + \beta\bar{g}$ in $\partial_x V(x,p)$. Let

$$g(\bar{g},p) = \begin{cases} p & \text{if } p \geq p_*(x) \\ -c(x) + \beta\bar{g} & \text{if } p < p_*(x). \end{cases}$$

By Theorem 7.46 of [13], $\bar{V}(x) = \mathbb{E}[V(x,\tilde{p})]$ has directional derivatives at every $x \geq 0$, so

$$\mathbb{E}[g(\bar{g},\tilde{p})] \in \partial_x\bar{V}(x).$$

It is easy to see that the mapping $F : \partial_x \bar{V}(x) \mapsto \partial_x \bar{V}(x)$ defined by

$$F(\bar{g}) = \mathbb{P}[\tilde{p} < p_*(x)](\beta\bar{g} - c(x)) + \mathbb{P}[\tilde{p} \geq p_*(x)]\mathbb{E}[\tilde{p}|\tilde{p} \geq p_*(x)]$$

is a contraction mapping with Lipschitz constant strictly less than 1, since for any $\bar{g}, \bar{g}' \in \partial_x \bar{V}(x)$

$$|F(\bar{g}) - F(\bar{g}')| = \beta\mathbb{P}[\tilde{p} < p_*(x)] \left|\bar{g} - \bar{g}'\right| < \left|\bar{g} - \bar{g}'\right|.$$

As $\partial_x \bar{V}(x)$ is a non-empty and closed set, by the Banach Fixed Point Theorem, there is a unique $\bar{g}(x) \in \partial_x \bar{V}(x)$ satisfying $F(\bar{g}(x)) = \bar{g}(x)$. But this implies

$$\bar{g}(x) = \mathbb{P}[\tilde{p} < p_*(x)](\beta\bar{g}(x) - c(x)) + \mathbb{P}[\tilde{p} \geq p_*(x)]\mathbb{E}[\tilde{p}|\tilde{p} \geq p_*(x)]$$

so

$$\bar{g}(x) = \frac{\mathbb{P}[\tilde{p} \geq p_*(x)]\mathbb{E}[\tilde{p}|\tilde{p} \geq p_*(x)] - \mathbb{P}[\tilde{p} < p_*(x)]c(x)}{1 - \beta\mathbb{P}[\tilde{p} < p_*(x)]} \in \partial_x \bar{V}(x). \tag{9}$$

An optimal solution $y(p)$ to $\max_{y \geq 0} \varphi(y)$ can be constructed as follows. Let P have support $[a, b]$. First observe that $\beta\mathbb{E}[(\tilde{p} - p)_+] + \beta p - p$ is a strictly decreasing continuous function of p. If

$$\beta\mathbb{E}[(\tilde{p} - p)_+] + \beta p - p > c(0)$$

for all $p \in [a, b]$ then set $z = b$. Otherwise let z be the unique solution to

$\beta\mathbb{E}[(\tilde{p}-p)_+]+\beta p-p=c(0)$. Define

$$y_*(p)=\begin{cases}c^{-1}(\beta\mathbb{E}[(\tilde{p}-p)_+]+\beta p-p) & \text{if } p<z\\ 0 & \text{if } p\in[z,b].\end{cases}$$

If $p<z$ then $y(p)>0$ and

$$\begin{aligned}y(p)&=c^{-1}(\beta\mathbb{E}[(\tilde{p}-p)_+]+\beta p-p)\\ &=c^{-1}\left(\beta\mathbb{P}[\tilde{p}\geq p]\mathbb{E}[\tilde{p}|\tilde{p}\geq p]+\beta\mathbb{P}[\tilde{p}<p]p-p\right)\end{aligned}$$

which can be rearranged to give

$$c(y(p))-\beta\mathbb{P}[\tilde{p}<p]p+p=\beta\mathbb{P}[\tilde{p}\geq p]\mathbb{E}[\tilde{p}|\tilde{p}\geq p].$$

Thus

$$(1-\beta\mathbb{P}[\tilde{p}<p])(p+c(y(p)))=-\beta\mathbb{P}[\tilde{p}<p]c(y(p))+\beta\mathbb{P}[\tilde{p}\geq p]\mathbb{E}[\tilde{p}|\tilde{p}\geq p],$$

and hence

$$p+c(y(p))-\beta\frac{\mathbb{P}[\tilde{p}\geq p]\mathbb{E}[\tilde{p}|\tilde{p}\geq p]-\mathbb{P}[\tilde{p}<p]c(y(p))}{1-\beta\mathbb{P}[\tilde{p}<p]}=0. \tag{10}$$

Observe that p satisfies the definition of $p_*(y(p))$ since the function $y(\cdot)$ is decreasing: $p'\leq p$ implies that $y(p')\geq y(p)$ and $p'\geq p$ implies that $y(p')\leq$

$y(p)$. Set $p_*(y(p)) = p$ and from (10), if $\bar{g}(y(p))$ is defined by (9), then

$$-p - c(y(p)) + \beta\bar{g}(y(p)) = 0 \in \partial_y \varphi(y(p)),$$

showing that $y(p)$ solves $\max_{y \geq 0} \varphi(y)$.

If $p = z$ then a similar analysis shows that $\bar{g}(0)$ satisfies

$$-z - c(0) + \beta\bar{g}(0) = 0$$

so for $p \geq z$ the right-hand derivative of $p(x-y) - C(y) + \beta\bar{V}(y)$ at $y = 0$ is less than or equal to 0, implying that $y(p) = 0$ solves $\max_{y \geq 0} \varphi(y)$.

Combining both cases by projecting onto $[c(0), c(x)]$ yields

$$y_*(x,p) = c^{-1}\left((\beta\mathbb{E}[(\tilde{p}-p)_+] + \beta p - p)_{[c(0),c(x)]}\right).$$

□

Proposition 5 shows that for a price p there is a finite target inventory

$$c^{-1}\left((\beta\mathbb{E}[(\tilde{p}-p)_+] + \beta p - p)_{[c(0),\infty)}\right)$$

at which the marginal cost of storage is equal to the discounted expected increase in the random price above p in the next stage. This is the value $\lim_{x\to\infty} y(x,p)$, and for simplicity we supress dependence on the first argument and denote this by $y(p)$. The optimal SIC policy reduces the current inventory level x to $y(p)$ if it is above this level, and does nothing otherwise. Conversely, for any inventory level x, there is a minimal offer price $p(x)$ required for sales to be worthwhile: this is the inverse function of $y(p)$ on the range $[0, \infty)$, i.e. the p that solves $y(p) = \min\{x, c^{-1}(\beta\mathbb{E}[\tilde{p}])\}$. Note that $p(x)$ is well-defined since $\beta\mathbb{E}[(\tilde{p} - p)_+] + \beta p - p$ is continuous, strictly decreasing, and unbounded below in p when $\beta < 1$.

Proposition 5 makes no assumptions about the probability distribution P, except that it has compact support. Thus P could have a density on $[a, b]$, or could be an empirical distribution on N price samples $p_1, p_2, \ldots, p_N$ which assigns probability $\frac{1}{N}$ to each sample, giving the SAA policy

$$y_{\mathrm{S}}(x, p) = c^{-1}\left(\left(\beta \frac{1}{N}\sum_{i=1}^{N}(p_i - p)_+ + \beta p - p\right)_{[c(0), c(x)]}\right). \tag{11}$$

Let $\bar{p} = \frac{1}{N}\sum_{i=1}^{N} p_i$. The MPC policy can be obtained from Proposition 5 by using the probability distribution that assigns probability 1 to the sample average $\bar{p}$, giving

$$y_{\mathrm{M}}(x, p) = c^{-1}\left(\left(\beta\big(\bar{p} - p\big)_+ + \beta p - p\right)_{[c(0), c(x)]}\right). \tag{12}$$

Denote the optimal target inventory function for the SAA policy by

$$y_{\mathrm{S}}(p) = c^{-1}\left(\left(\beta \frac{1}{N}\sum_{i=1}^{N}(p_i - p)_+ + \beta p - p\right)_{[c(0), \infty)}\right) \tag{13}$$

and its inverse function by p_{S}, and similarly for the MPC policy by y_{M} and p_{M} respectively. It is easy to see that the value $p_{\mathrm{S}}(x)$ is increasing in each p_i, and this has a natural explanation: the SAA policy foresees higher prices in the future and therefore requires a higher price before any sales. Clearly

$p_{\mathrm{S}}(x) \leq p_N$ since p_N is the highest price the SAA policy foresees as possible.

Depending on the samples $p_1, p_2, \ldots, p_N$ the policy obtained could either pay too much in storage costs by selling too little inventory, or not be able to take full advantage of future high prices having sold too much inventory. By Jensen's inequality, $\mathbb{E}[(\tilde{p} - p)_+] \geq (\mathbb{E}[\tilde{p}] - p)_+$, and hence $y_{\mathrm{S}}(x) \geq y_{\mathrm{M}}(x)$. Thus, the MPC policy decides that it is worthwhile to sell at a lower price than the SAA policy does. It is then clear that $p_{\mathrm{S}}(x) \geq p_{\mathrm{M}}(x)$ since these are the minimal offer prices required for sales under each policy.

5 Out-of-sample performance

The following assumption allows us to study the out-of-sample performance of the sample-based policies derived using Proposition 5 even when the underlying price distribution has unbounded support.

Assumption 3. *The probability distribution P has non-negative support, a finite mean, and no atoms.*

Note that under Assumption 3, for $x \geq y \geq 0$ the term $|\tilde{p}(x - y) - C(y)| \leq \tilde{p}(x - y) + C(y)$ with $\mathbb{E}[\tilde{p}(x - y) + C(y)]$ finite. This is equivalent to Assumption 2 in SOC. For N price samples $p_1, p_2, \ldots, p_N$ we form the SAA and MPC policies defined by (11) and (12) respectively. It follows that the value of the SIC problem if the possibly sub-optimal SAA policy is used out-of-sample from initial state x^0, which we denote by $\bar{V}_{\mathrm{S}}(x^0)$, is well-defined as in (3) and satisfies the functional equation

$$\bar{V}_{\mathrm{S}}(x) = \mathbb{E}\left[\tilde{p}(x - y_{\mathrm{S}}(x, \tilde{p})) - C(y_{\mathrm{S}}(x, \tilde{p})) + \beta \bar{V}_{\mathrm{S}}(y_{\mathrm{S}}(x, \tilde{p}))\right].$$

Similarly, the value of the SIC problem if the possibly sub-optimal MPC policy is used out-of-sample from initial state x^0, which we denote by $\bar{V}_{\mathrm{M}}(x^0)$, is well-defined and satisfies the functional equation

$$\bar{V}_{\mathrm{M}}(x) = \mathbb{E}\left[\tilde{p}(x - y_{\mathrm{M}}(x,\tilde{p})) - C(y_{\mathrm{M}}(x,\tilde{p})) + \beta \bar{V}_{\mathrm{M}}(y_{\mathrm{M}}(x,\tilde{p}))\right].$$

5.1 Derivative of the expected value function

Before making comparisons between $\bar{V}_{\mathrm{S}}(x)$ and $\bar{V}_{\mathrm{M}}(x)$ we will first calculate their derivatives with respect to the initial inventory. It will be helpful to use the result [13, Theorem 7.44]: suppose $F : \mathbb{R}^M \times \Omega \to \mathbb{R}$ is a random function with expected value $f(x) = \mathbb{E}[F(x, \tilde{\omega})]$.

Lemma 6. *If the following conditions hold:*

(i) the expectation $f(x_0)$ is well defined and finite valued at some point $x_0 \in \mathbb{R}^M$,

(ii) there exists a positive valued random variable $L(\tilde{\omega})$ with $\mathbb{E}[L(\tilde{\omega})] < \infty$, and for all x_1, x_2 in a neighbourhood of x_0 and almost every $\omega \in \Omega$, it holds that $|F(x_1,\omega) - F(x_2,\omega)| \le L(\omega)\|x_1 - x_2\|$,

(iii) for almost every $\omega \in \Omega$ the function $F(x,\omega)$ is differentiable with respect to x at x_0,

then $f(x)$ is differentiable at x_0 and

$$\nabla f(x_0) = \mathbb{E}[\nabla_x F(x_0, \tilde{\omega})].$$

Now the derivative values can be established.

Proposition 7. *Under Assumption 3, for $x \geq 0$ the derivatives $\frac{d}{dx}\bar{V}_{\mathrm{S}}(x)$ and $\frac{d}{dx}\bar{V}_{\mathrm{M}}(x)$ exist.*

Proof. We will first show that $\frac{d}{dx}\bar{V}_{\mathrm{S}}(x)$ exists. Note that $\bar{V}_{\mathrm{S}}(x)$ is well-defined and finite valued, satisfying condition (i) of Lemma 6. For a sequence of prices $\{p^1, p^2, \ldots\} = \omega$, let

$$V_{\mathrm{S}}(x^0, \omega) = \sum_{t=0}^{\infty} \beta^t (p^t(x^t - y^t) - C(y^t)) \tag{14}$$

where $y^t = y_{\mathrm{S}}(x^t, p^t)$ and $x^{t+1} = y^t$ for all t. Note that $\mathbb{E}[V_{\mathrm{S}}(x^0, \tilde{\omega})] = \bar{V}_{\mathrm{S}}(x^0)$. For each p^t, the policy y_{S} has a target inventory $y_{\mathrm{S}}(p^t)$. In (14), the inventory values x^t remain at their initial value x^0, until the first t (say z) with $p^t \geq p_{\mathrm{S}}(x^0)$ and $y_{\mathrm{S}}(p^t) \leq x^0$. Beyond z, all x^t are independent of x^0. Due to P being atomless, almost surely there are no p^t with $y_{\mathrm{S}}(p^t) = x^0$ since $\mathbb{P}[\tilde{p} = p_{\mathrm{S}}(x^0)] = 0$. So there is a neighbourhood of x^0 such that $y_{\mathrm{S}}(p^z) < x^0$, where $x^t = x^0$ for $t \leq z$ and $y^z = y_{\mathrm{S}}(p^z)$, with all other states not depending on the value of x^0. In this neighbourhood $\frac{d}{dx^0}V_{\mathrm{S}}(x_0, \omega) = \beta^z p^z - \sum_{t=0}^{z-1} \beta^t c(x^0)$, meeting condition (iii) of Lemma 6. Since $\mathbb{E}[\sum_{t=0}^{\tilde{z}-1} \beta^t c(x^0)] \leq \frac{1}{1-\beta} c(x^0)$ is bounded on any compact set and $\mathbb{E}[\beta^{\tilde{z}} p^{\tilde{z}}] \leq \mathbb{E}[\tilde{p} | \tilde{p} \geq p_{\mathrm{S}}(x)] \leq \mathbb{E}[\tilde{p} | \tilde{p} \geq \max_i\{p_i\}] < \infty$, it follows that $V_{\mathrm{S}}(x^0, \omega)$ has a Lipschitz constant with finite expectation, meeting condition (ii) of Lemma 6. Thus, $\bar{V}_{\mathrm{S}}(x)$ is differentiable. Similar reasoning shows that $\bar{V}_{\mathrm{M}}(x)$ is differentiable. $\square$

Proposition 8. *Under Assumption 3, for any $x \geq 0$ the derivatives $\frac{d}{dx}\bar{V}_{\mathrm{S}}(x)$*

and $\frac{d}{dx}\bar{V}_{\mathrm{M}}(x)$ *are given by*

$$\frac{d}{dx}\bar{V}_{\mathrm{S}}(x) = \frac{\mathbb{E}[\tilde{p}|\tilde{p} \geq p_{\mathrm{S}}(x)]\mathbb{P}[\tilde{p} \geq p_{\mathrm{S}}(x)] - c(x)\mathbb{P}[\tilde{p} < p_{\mathrm{S}}(x)]}{1 - \beta\mathbb{P}[\tilde{p} < p_{\mathrm{S}}(x)]}$$

and

$$\frac{d}{dx}\bar{V}_{\mathrm{M}}(x) = \frac{\mathbb{E}[p|p \geq p_{\mathrm{M}}(x)]\mathbb{P}[p \geq p_{\mathrm{M}}(x)] - c(x)\mathbb{P}[p < p_{\mathrm{M}}(x)]}{1 - \beta\mathbb{P}[p < p_{\mathrm{M}}(x)]}.$$

Proof. The proof proceeds by first showing that the derivatives satisfy a recursion. Let

$$V_{\mathrm{S}}(x,p) = \begin{cases} -C(x) + \beta\bar{V}_{\mathrm{S}}(x) & \text{if } p < p_{\mathrm{S}}(x) \\ p(x - y_{\mathrm{S}}(p)) - C(y_{\mathrm{S}}(p)) + \beta\bar{V}_{\mathrm{S}}(y_{\mathrm{S}}(p)) & \text{if } p \geq p_{\mathrm{S}}(x). \end{cases} \tag{15}$$

Note that $\mathbb{E}[V_{\mathrm{S}}(x,\tilde{p})] = \bar{V}_{\mathrm{S}}(x)$. For any $p < p_{\mathrm{S}}(x)$ the first case of (15) holds in a neighbourhood of x since $p_{\mathrm{S}}(x)$ is continuous. It follows that $\frac{d}{dx}V_{\mathrm{S}}(x,p) = -c(x) + \beta\frac{d}{dx}\bar{V}_{\mathrm{S}}(x)$, making use of the fact that $\frac{d}{dx}\bar{V}_{\mathrm{S}}(x)$ exists by Proposition 7. Similarly, for $p > p_{\mathrm{S}}(x)$, following the second case of (15), $\frac{d}{dx}V_{\mathrm{S}}(x,p) = p$. A similar application of Lemma 6 as in the proof of Proposition 7 shows that $\frac{d}{dx}\bar{V}_{\mathrm{S}}(x) = \mathbb{E}\left[\frac{d}{dx}V_{\mathrm{S}}(x,\tilde{p})\right]$. Taking expectations yields

$$\frac{d}{dx}\bar{V}_{\mathrm{S}}(x) = \left(\beta\frac{d}{dx}\bar{V}_{\mathrm{S}}(x) - c(x)\right)\mathbb{P}[\tilde{p} < p_{\mathrm{S}}(x)] + \mathbb{E}[\tilde{p}|\tilde{p} \geq p_{\mathrm{S}}(x)]\mathbb{P}[\tilde{p} \geq p_{\mathrm{S}}(x)],$$

and rearranging gives the required expression:

$$\frac{d}{dx}\bar{V}_{\mathrm{S}}(x) = \frac{\mathbb{E}[\tilde{p}|\tilde{p} \geq p_{\mathrm{S}}(x)]\mathbb{P}[\tilde{p} \geq p_{\mathrm{S}}(x)] - c(x)\mathbb{P}[\tilde{p} < p_{\mathrm{S}}(x)]}{1 - \beta\mathbb{P}\left[\tilde{p} < p_{\mathrm{S}}(x)\right]}.$$

The expression for $\frac{d}{dx}\bar{V}_{\mathrm{M}}(x)$ can be derived via similar reasoning. □

Proposition 9. *Assume P has a probability density function f. Under Assumption 3, if*

$$c(y) \geq \beta \int_{p_{\mathrm{S}}(y)}^{\infty} pf(p)dp$$

for all $y \in [0, x]$, then $\bar{V}_{\mathrm{M}}(x) \geq \bar{V}_{\mathrm{S}}(x)$. That is, the MPC policy performs better than the SAA policy out-of-sample.

Proof. Let $\bar{V}^1_{(\mathrm{S},\mathrm{M})}(y)$ denote the out-of-sample value of starting with inventory y and following the SAA policy for one stage and the MPC policy forevermore. We will first show that $\frac{d}{dy}\bar{V}_{\mathrm{M}}(y) \geq \frac{d}{dy}\bar{V}^1_{(\mathrm{S},\mathrm{M})}(y)$ for all $y \in [0, x]$. Following Proposition 8,

$$\frac{d}{dy}\bar{V}_{\mathrm{M}}(y) = \left(\beta\frac{d}{dy}\bar{V}_{\mathrm{M}}(y) - c(y)\right)\int_{-\infty}^{p_{\mathrm{M}}(y)} f(p)dp + \int_{p_{\mathrm{M}}(y)}^{\infty} pf(p)dp.$$

Inspecting $\bar{V}^1_{(\mathrm{S},\mathrm{M})}(y)$ shows that the similar expression

$$\frac{d}{dy}\bar{V}^1_{(\mathrm{S},\mathrm{M})}(y) = \left(\beta\frac{d}{dy}\bar{V}_{\mathrm{M}}(y) - c(y)\right)\int_{-\infty}^{p_{\mathrm{S}}(y)} f(p)dp + \int_{p_{\mathrm{S}}(y)}^{\infty} pf(p)dp$$

holds. It can be seen that

$$\frac{d}{dy}\bar{V}_{\mathrm{M}}(y) - \frac{d}{dy}\bar{V}^{1}_{(\mathrm{S},\mathrm{M})}(y) = \int_{p_{\mathrm{M}}(y)}^{p_{\mathrm{S}}(y)} pf(p)dp - \left(\beta\frac{d}{dy}\bar{V}_{\mathrm{M}}(y) - c(y)\right)\int_{p_{\mathrm{M}}(y)}^{p_{\mathrm{S}}(y)} f(p)dp. \quad (16)$$

Using Proposition 8, write

$$\begin{aligned}\beta\frac{d}{dy}\bar{V}_{\mathrm{M}}(y) - c(y) &= \beta\left(\frac{\int_{p_{\mathrm{M}}(y)}^{\infty} pf(p)dp - c(y)\int_{-\infty}^{p_{\mathrm{M}}(y)} f(p)dp}{1-\beta\int_{-\infty}^{p_{\mathrm{M}}(y)} f(p)dp}\right) - c(y)\\ &= \frac{\beta\int_{p_{\mathrm{M}}(y)}^{\infty} pf(p)dp - c(y)}{1-\beta\int_{-\infty}^{p_{\mathrm{M}}(y)} f(p)dp}\end{aligned}$$

so applying the condition in the statement of the proposition yields

$$\beta\frac{d}{dy}\bar{V}_{\mathrm{M}}(y) - c(y) \leq \frac{\beta\int_{p_{\mathrm{M}}(y)}^{p_{\mathrm{S}}(y)} pf(p)dp}{1-\beta\int_{-\infty}^{p_{\mathrm{M}}(y)} f(p)dp}. \quad (17)$$

Recall $p_{\mathrm{S}}(y) \geq p_{\mathrm{M}}(y)$. Then

$$\left(\frac{\beta\int_{p_{\mathrm{M}}(y)}^{p_{\mathrm{S}}(y)} pf(p)dp}{1-\beta\int_{-\infty}^{p_{\mathrm{M}}(y)} f(p)dp}\right)\int_{p_{\mathrm{M}}(y)}^{p_{\mathrm{S}}(y)} f(p)dp \leq \int_{p_{\mathrm{M}}(y)}^{p_{\mathrm{S}}(y)} pf(p)dp$$

since $\int_{p_{\mathrm{M}}(y)}^{p_{\mathrm{S}}(y)} pf(p)dp$ can be cancelled from both sides and the identity rearranged to give $\beta\int_{-\infty}^{p_{\mathrm{S}}(y)} f(p)dp \leq 1$. Thus (17) yields

$$\left(\beta\frac{d}{dy}\bar{V}_{\mathrm{M}}(y) - c(y)\right)\int_{p_{\mathrm{M}}(y)}^{p_{\mathrm{S}}(y)} f(p)dp \leq \int_{p_{\mathrm{M}}(y)}^{p_{\mathrm{S}}(y)} pf(p)dp,$$

which shows that (16) is non-negative, whereby $\frac{d}{dy}\bar{V}_{\mathrm{M}}(y) \geq \frac{d}{dy}\bar{V}^{1}_{(\mathrm{S},\mathrm{M})}(y)$ for all

$y \in [0,x]$. This implies that $\bar{V}_{\text{M}}(y) \geq \bar{V}^1_{(\text{S},\text{M})}(y)$ for all $y \in [0,x]$. Lemma 4 can then be used to conclude that $\bar{V}_{\text{M}}(x) \geq \bar{V}_{\text{S}}(x)$, as required. □

Suppose $\tilde{p}_1, \tilde{p}_2, \ldots, \tilde{p}_N$ are independent and identically distributed samples from P which has density f. For any realization of these samples, Proposition 9 requires that $c(y) \geq \beta \int_{p_{\text{S}}(y)}^{\infty} pf(p)dp$ for all $y \in [0,x]$. Without loss of generality, assume $\tilde{p}_1 \leq \tilde{p}_2 \leq \ldots \leq \tilde{p}_N$ and recall $p_{\text{S}}(y) \leq \tilde{p}_N$. Observe that the term $(\tilde{p}_N - p)_+$ in (13) is strictly increasing in $\tilde{p}_N$ at the point $p = p_{\text{S}}(y)$ where $y_{\text{S}}(p) = y$. It follows that the minimal offer price $p_{\text{S}}(y)$ which solves $y_{\text{S}}(p) = y$ is strictly increasing in $\tilde{p}_N$. The term $\int_{p_{\text{S}}(x)}^{\infty} pf(p)dp$ is then decreasing in $\tilde{p}_N$ and eventually vanishes. When f has infinite support we will occasionally sample a $\tilde{p}_N$ that is sufficiently large for the inequality $c(x) \geq \beta \int_{p_{\text{S}}(x)}^{\infty} pf(p)dp$ to hold for all $y \in [0,x]$. So we are more likely to encounter samples where $\bar{V}_{\text{M}}(x_0) \geq \bar{V}_{\text{S}}(x_0)$ when f has more weight in the (high price) tail of the distribution.

Since the terms $\bar{V}_{\text{S}}(x_0)$ and $\bar{V}_{\text{M}}(x_0)$ are themselves random (determined by the samples $\tilde{p}_1, \tilde{p}_2, \ldots, \tilde{p}_N$), Proposition 9 does not say anything explicitly about the *expected* out-of-sample performance of MPC and SAA with respect to the sampling distribution. In fact, in most applications it is likely that there will always be some values of the random variables which result in MPC outperforming SAA, so Proposition 9 is not surprising. However, Proposition 9 demonstrates how the performance of SAA is affected by the parameterization of the SIC problem: the SAA policy can be misled by price samples with large values, which cause it to hold on to inventory for too long. Consequently, heavy-tailed distributions, which occasionally yield samples with particularly large values, will negatively impact SAA's expected out-of-sample performance. MPC—due to its distributionally robust properties—more quickly sells down to lower storage cost levels and is protected.

If $c(0) = 0$, then the premise of Proposition 9 requires $p_{\mathrm{S}}(0)$ to be equal to the upper limit of the support of f and consequently $\tilde{p}_N$ to be at least as large. This occurs with measure 0. But here the failure of Proposition 9 is to be expected; when $c(0) = 0$, for an infinitesimal inventory level selling inventory does not change the storage cost incurred, so waiting for higher prices by using the SAA policy instead of the MPC policy will always perform better out-of-sample. Despite this, for non-negligible initial inventory levels, we present examples below which show that MPC can still outperform SAA out-of-sample under expectation when $c(0) = 0$. These examples all involve densities having a tail at high prices.

6 Exponentially distributed price example

In this section we compare the expected out-of-sample rewards of the sample-based policies on SIC when $x^0 = 1$, $C(x) = \frac{1}{2}x^2$ (so $c(x) = x$), and $\tilde{p}$ is exponentially distributed with rate 1. This has cumulative distribution function

$$F(p) = \begin{cases} 1 - e^{-p} & \text{if } p \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

and probability density function

$$f(p) = \begin{cases} e^{-p} & \text{if } p \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

Note that Assumption 3 applies here since $\mathbb{E}[\tilde{p}] = 1$.

For $N \geq 2$, let $\tilde{p}_1, \tilde{p}_2, \ldots, \tilde{p}_N$ be independent and identically distributed random samples drawn from the exponential distribution with rate 1. First consider the SAA solution to SIC using $\tilde{p}_1, \tilde{p}_2, \ldots, \tilde{p}_N$. The calculation below shows that the resulting SAA policy performs poorly. In fact the expected out-of-sample cost is unbounded as $\beta \to 1$. We will then compare this with the result if the MPC policy is used instead.

For any $p_1, p_2, \ldots, p_N$, the out-of-sample reward of the SAA policy is $\bar{V}_{\mathrm{S}}(1)$. We will evaluate $\mathbb{E}[\bar{V}_{\mathrm{S}}(1)]$, the expectation here being with respect to the sampling distribution of $\tilde{p}_1, \tilde{p}_2, \ldots, \tilde{p}_N$, and show that this is unbounded below as $\beta \to 1$. By Proposition 7, $\frac{d}{dx}\bar{V}_{\mathrm{S}}(x)$ exists. Since $\bar{V}_{\mathrm{S}}(0) = 0$, it follows that $\bar{V}_{\mathrm{S}}(1) = \int_0^1 \frac{d}{dx}\bar{V}_{\mathrm{S}}(x)dx$. Using the expression for $\frac{d}{dx}\bar{V}_{\mathrm{S}}(x)$ from Proposition 8, $\mathbb{E}[\bar{V}_{\mathrm{S}}(1)]$ is equal to

$$\mathbb{E}\left[\int_0^1 \frac{\int_{p_{\mathrm{S}}(x)}^{\infty} pf(p)dp - xF(p_{\mathrm{S}}(x))}{1 - \beta F(p_{\mathrm{S}}(x))} dx\right]. \tag{18}$$

First consider the negative term in (18);

$$\mathbb{E}\left[\int_0^1 \frac{-xF(p_{\mathrm{S}}(x))}{1 - \beta F(p_{\mathrm{S}}(x))} dx\right]. \tag{19}$$

The term inside the expectation in (19) is measurable. Using the sampling densities and applying Tonelli's Theorem, (19) is then equal to

$$\int_0^\infty \int_0^\infty \cdots \int_0^\infty \int_0^1 \frac{-xF(p_{\rm S}(x))}{1-\beta F(p_{\rm S}(x))} dx f(p_N)dp_N \ldots f(p_2)dp_2 f(p_1)dp_1. \quad (20)$$

Let us fix $p_1, p_2, \ldots, p_{N-1}$, and consider the inner-most integral in (20) when p_N is large. The value $p_{\rm S}(x)$ is the p which solves (13) when $c(y_{\rm S}(p)) = x$. Given $\delta \in (0,1)$, if p_N is sufficiently large, then for all $\beta \in [\delta, 1)$ and $x \in [0,1]$ equation (13) becomes $\beta \frac{1}{N}(p_N - p)_+ + \beta p - p = x$ and has solution $p_{\rm S}(x) = \frac{p_N \beta - Nx}{N+\beta-N\beta} \geq 0$. It can then be seen that

$$\frac{-xF(p_{\rm S}(x))}{1-\beta F(p_{\rm S}(x))} \to \frac{-x\left(1-e^{-(p_N-Nx)}\right)}{1-\left(1-e^{-(p_N-Nx)}\right)}$$

uniformly in $x \in [0,1]$ as $\beta \to 1$, whereby

$$\begin{aligned}\lim_{\beta\to 1} \int_0^1 \frac{-xF(p_{\rm S}(x))}{1-\beta F(p_{\rm S}(x))} dx &= \int_0^1 \frac{-x\left(1-e^{-(p_N-Nx)}\right)}{1-\left(1-e^{-(p_N-Nx)}\right)} dx \\ &= \frac{1}{2} + \left(\frac{1}{N} + \frac{1}{N^2}\right) e^{p_N - N} - \frac{1}{N^2} e^{p_N}.\end{aligned}$$

So for each p_N sufficiently large, given $\epsilon > 0$, there exists a $\beta < 1$ beyond which

$$\int_0^1 \frac{-xF(p_{\rm S}(x))}{1-\beta F(p_{\rm S}(x))} dx < \frac{1}{2} + \left(\frac{1}{N} + \frac{1}{N^2}\right) e^{p_N - N} - \frac{1}{N^2} e^{p_N} + \epsilon. \quad (21)$$

But $f(p_N) = e^{-p_N}$ and

$$\lim_{p_N \to \infty} \left(\frac{1}{2} + \left(\frac{1}{N} + \frac{1}{N^2} \right) e^{p_N - N} - \frac{1}{N^2} e^{p_N} + \epsilon \right) e^{-p_N}$$
$$= \left(\frac{1}{N} + \frac{1}{N^2} \right) e^{-N} - \frac{1}{N^2} < 0$$

which means that for any $b > a > 0$ the integral

$$\int_a^b \left(\frac{1}{2} + \left(\frac{1}{N} + \frac{1}{N^2} \right) e^{p_N - N} - \frac{1}{N^2} e^{p_N} + \epsilon \right) e^{-p_N} dp_N$$

can be made arbitrarily negative by increasing b. Using the upper bound (21), it follows that the integral

$$\int_0^\infty \int_0^1 \frac{-xF(p_{\mathrm{S}}(x))}{1 - \beta F(p_{\mathrm{S}}(x))} dx f(p_N) dp_N$$

in (20) can be made arbitrarily negative by increasing β towards 1, regardless of the values of $p_1, p_2, \ldots, p_{N-1}$. Hence, (19) is unbounded below.

Let us now address the positive term in (18);

$$\mathbb{E} \left[\int_0^1 \frac{\int_{p_{\mathrm{S}}(x)}^\infty pf(p)dp}{1 - \beta F(p_{\mathrm{S}}(x))} dx \right]. \tag{22}$$

Again, for each β the integral in (22) is measurable. For any $p_1, p_2, \ldots, p_N$ recall $p_{\mathrm{S}}(x) \leq \max_i\{p_i\}$. Thus, for all β

$$\frac{\int_{p_{\mathrm{S}}(x)}^\infty pf(p)dp}{1 - \beta F(p_{\mathrm{S}}(x))} \leq \frac{\int_{p_{\mathrm{S}}(x)}^\infty pf(p)dp}{1 - F(p_{\mathrm{S}}(x))} = 1 + p_{\mathrm{S}}(x) \leq 1 + \max_i\{p_i\}.$$

The value $\max_i\{p_i\}$ is just the N^{th} order statistic of the N samples from the exponential distribution with rate 1, which has expectation $\sum_{i=1}^{N} \frac{1}{i}$. So (22) is bounded above by the value $1 + \sum_{i=1}^{N} \frac{1}{i}$.

We have shown that (19) is unbounded below and (22) is bounded above. Thus, their sum is defined and the limit (18) is unbounded below, which shows that the expected out-of-sample loss incurred by following the SAA policy is unbounded as $\beta \to 1$, regardless of the choice of N.

In contrast to the SAA policy, the expected out-of-sample cost incurred by following the MPC policy is bounded as $\beta \to 1$. For simplicity we demonstrate this in the case $N = 2$, although it can be shown to hold in general. Following similar reasoning as in the SAA case, $\lim_{\beta \to 1} \mathbb{E}[\bar{V}_{\mathrm{M}}(1)]$ is equal to

$$\lim_{\beta \to 1} \mathbb{E}\left[\int_0^1 \frac{\int_{p_{\mathrm{M}}(x)}^{\infty} p f(p) dp - x F(p_{\mathrm{M}}(x))}{1 - \beta F(p_{\mathrm{M}}(x))} dx\right] \tag{23}$$

which has negative term

$$\lim_{\beta \to 1} \int_0^{\infty} \int_0^{\infty} \int_0^1 \frac{-x F(p_{\mathrm{M}}(x))}{1 - \beta F(p_{\mathrm{M}}(x))} dx f(p_2) dp_2 f(p_1) dp_1. \tag{24}$$

Let $\overline{p} = \frac{1}{2}(p_1 + p_2)$. The iterated integral in (24) can be divided into ranges based on the value of $\overline{p}$. Observe that, depending on the value of $\overline{p}$, either $p_{\mathrm{M}}(x) = \beta\overline{p} - x$ or $p_{\mathrm{M}}(x) = \frac{-x}{1-\beta}$. Since $\overline{p} - x \geq \beta\overline{p} - x \geq \frac{-x}{1-\beta}$, for any value of $\overline{p}$ and β, in (24) the term

$$\int_0^1 \frac{-x F(p_{\mathrm{M}}(x))}{1 - \beta F(p_{\mathrm{M}}(x))} dx \geq \int_0^1 \frac{-x F(\overline{p} - x)}{1 - F(\overline{p} - x)} dx.$$

But

$$
\begin{aligned}
\int_0^1 \frac{-xF(\overline{p}-x)}{1-F(\overline{p}-x)}dx &= \int_0^{\min\{\overline{p},1\}} \frac{-x\left(1-e^{-(\overline{p}-x)}\right)}{1-(1-e^{-(\overline{p}-x)})}dx \\
&= \frac{1}{2}(\min\{\overline{p},1\})^2 + (1+\min\{\overline{p},1\})e^{\overline{p}-\min\{\overline{p},1\}} - e^{\overline{p}} \\
&\geq -e^{\overline{p}}
\end{aligned}
$$

and consequently the negative value (24) is bounded below by

$$
\int_0^\infty \int_0^\infty -e^{\frac{1}{2}(p_1+p_2)}e^{-p_2}dp_2e^{-p_1}dp_1 = -4.
$$

Moreover, similar reasoning to the SAA case shows that the positive term in (23) is bounded above. Thus, the expected out-of-sample loss incurred under the MPC policy is bounded as $\beta \to 1$.

The previous calculations show that, for any $N \geq 2$, the expected out-of-sample performance of the SAA policy can be made arbitrarily worse than that of the MPC policy by choosing β sufficiently close to 1. For any given β the expected out-of-sample performance of the SAA policy may be improved by increasing N, but our example shows that SAA problems with low discount rates can require very large values of N to give good solutions.

It is tempting to conclude for this example that the SAA policy is not convergent almost surely to the optimal policy, but it is easy to see from the explicit form of the policies that this is not true. The SAA policy and its *in-sample* value converge to their true optimal counterparts. However the derivation above shows that the *out-of-sample* performance of the SAA policy for this example might be very bad on average even if N is large. In real applications, even if a sufficient number of samples are available, the SAA policy must be computed numerically, and algorithms for doing this cannot handle very large values of N. The MPC policy does not have this problem.

7 Numerical studies

In this section we use numerical simulation to study the expected out-of-sample performance of the two sample-based policies (SAA and MPC) on different price distributions. In Section 6 we showed that MPC performs far better than SAA when the underlying price distribution is exponential. But this is an exception—we do not always find the extreme behaviour where the two expected out-of-sample values differ by an amount that is unbounded as $\beta \to 1$. However, this does suggest that the amount of skew and the size of the tail in the underlying distribution is important, and we will investigate this in the current section.

To compute the expected out-of-sample performance of the sample-based policies under the sampling distribution of $\{\tilde{p}_1, \tilde{p}_2, \ldots, \tilde{p}_N\}$, we use a simulation coded in the `Julia` programming language [3]. Although the true problem has an infinite number of stages, simulation with a finite number of stages (say T) will give a realistic estimate as long as it is sufficiently large. We set $T = 1000$ and efficiently simulate the repeated sales process by terminating any instances as soon as the inventory level reaches 0. Setting

$\beta = 0.95$,[3] $x_0 = 1$ and $C(x) = \frac{1}{2}x^2$, for each policy we:

1. Sample N random prices from P to construct $\tilde{p}_1, \tilde{p}_2, \ldots, \tilde{p}_N$, which then determines the sample-based policy y (either by SAA or MPC).

2. Sample a random price $\tilde{p}^t$ from P, accrue the stage reward $\beta^t(\tilde{p}^t(x^t - y(x^t, \tilde{p}^t)) - C(y(x^t, \tilde{p}^t)))$, and set $x^{t+1} = y(x^t, \tilde{p}^t)$.

3. Repeat Step 2 from stage $t = 0$ to $T-1$ and sell any remaining inventory at stage T to generate $\sum_{t=0}^{T} \beta^t(\tilde{p}^t(x^t - y(x^t, \tilde{p}^t)) - C(y(x^t, \tilde{p}^t)))$.

We repeat Steps 1 through 3 to generate realizations for use as an estimate of the expected value of the SIC problem when a sample-based policy is used out-of-sample. In our experiments we used 50000 realizations to generate the estimate of the expected out-of-sample value and found that this was sufficient to achieve accurate values. In Figures 1-4 and 5 the standard error ranges are smaller than the markers and so are not shown. Also note that for $N = 1$ the two sample-based policies coincide.

7.1 Skewed price distributions

Suppose $\tilde{p} \sim \text{Triangular}(a, m, b)$, with lower limit a, mode m, and upper limit b. This distribution serves to illustrate the effect of skew on the performance of SAA and MPC on SIC. In what follows we select a, m, and b such that $\mathbb{E}[\tilde{p}] = 1$ and $\text{Var}[\tilde{p}] = \frac{1}{8}$; the intention being to confine differences between SAA and MPC to the sampling effects of skew only and compare them on different distributions as fairly as possible.

[3] The resulting truncation error is on the order of $0.95^{1000} \approx 10^{-23}$.

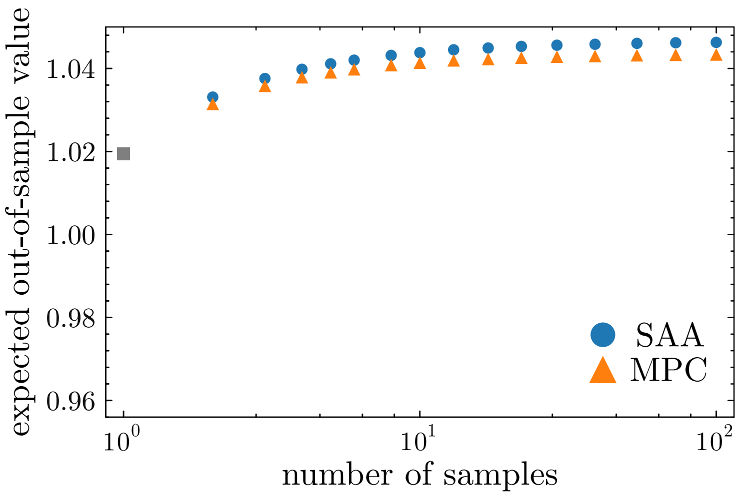

Figure 1: Expected out-of-sample value of SAA and MPC for $p \sim \text{Triangular}\,(0, {}^3\!/\!{}_2, {}^3\!/\!{}_2)$, a left-skewed distribution.

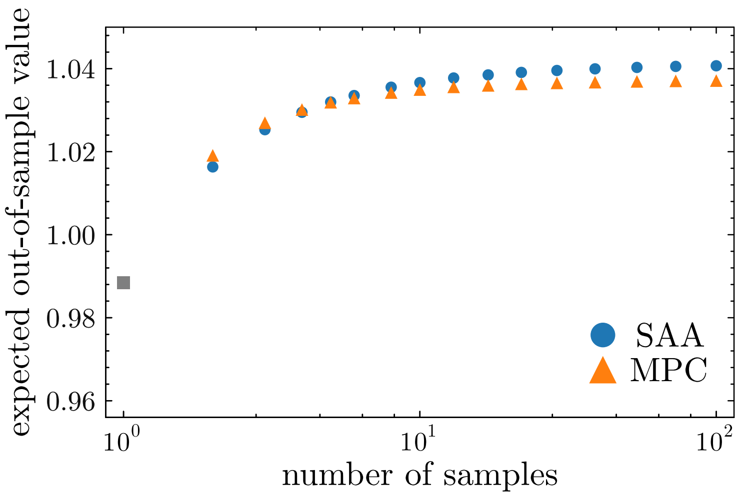

Figure 2: Expected out-of-sample value of SAA and MPC for $\tilde{p} \sim \text{Triangular}\left(1 - {}^1\!/\!{}_2\sqrt{3}, 1, 1 + {}^1\!/\!{}_2\sqrt{3}\right)$, a symmetric distribution.

Figure 1 shows SAA outperforming MPC for all N on a price distribution that is triangular and left-skewed. This is in contrast to Figure 2, which shows MPC outperforming SAA for $N \leq 5$ on a price distribution that is triangular and symmetric. Replacing the left-skewed price distribution that yields Figure 1 with a symmetric distribution increases the value of the support's upper limit. Samples with high prices then cause the SAA policy to under-sell and pay too much in storage costs. The MPC policy attenuates this effect since $y_{\mathrm{M}}(x) \leq y_{\mathrm{S}}(x)$.

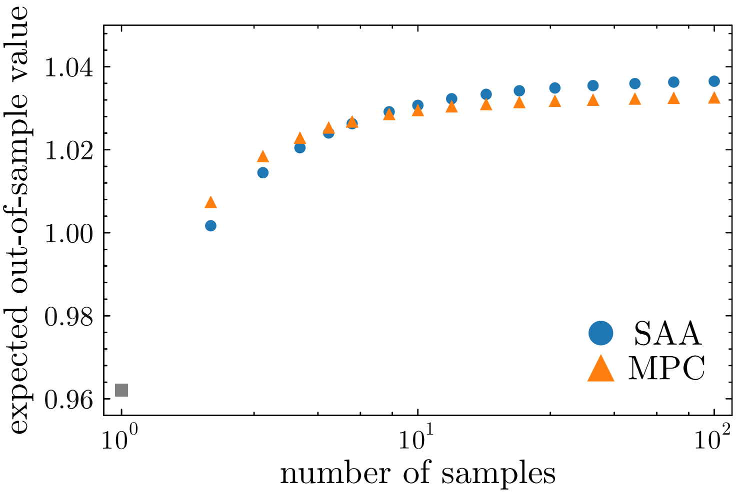

Figure 3: Expected out-of-sample value of SAA and MPC for $\tilde{p} \sim \text{Triangular}\,(1/2, 1/2, 2)$, a right-skewed distribution.

Further increasing the mode to 2 increases the range where MPC outperforms SAA, as can be seen in Figure 3, which shows MPC outperforming SAA for $N \leq 6$ on a price distribution that is triangular and right-skewed.

7.2 Tailed price distributions

We have already seen that the perfomance of SAA on exponentially distributed prices can be arbitraryly bad. Following Section 6, suppose that $\tilde{p} \sim \text{Exponential}\,(\lambda)$ with rate $\lambda = 1$.

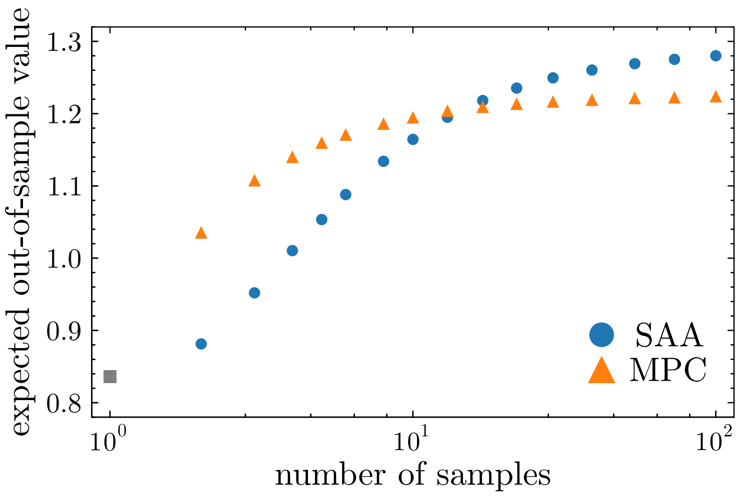

Figure 4: Expected out-of-sample value of SAA and MPC for $\tilde{p} \sim \text{Exponential}\,(1)$. Note $\mathbb{E}[\tilde{p}] = 1$.

Figure 5 shows MPC outperforming SAA for all N less than about 10, a larger range than that in Figure 3. The right-tail of the Exponential(1) distribution increases the propensity for a single very large price sample to be included in $\tilde{p}_1, \tilde{p}_2, \ldots, \tilde{p}_N$ which degrades the approximate price distribution informing the SAA policy.

Now suppose $\tilde{p} \sim \text{LogNormal}\,(\mu, \sigma^2)$, with mean μ and variance σ^2. Log-Normal distributions are often used to model prices in financial applications and have a right-tail which decays slower than that of exponential distributions.

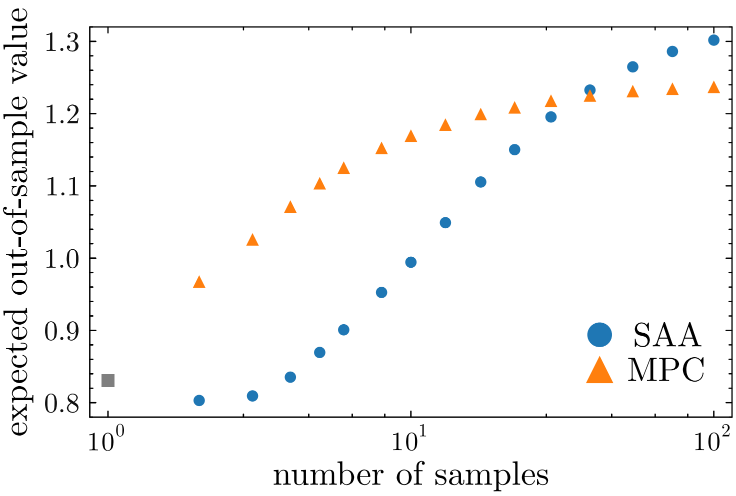

Figure 5: Expected out-of-sample value of SAA and MPC for $\tilde{p} \sim \text{LogNormal}\,(-1/2, 1)$. Note $\mathbb{E}[\tilde{p}] = 1$.

Figure 5 shows MPC outperforming SAA for all N less than about 50, a larger range than that in Figures 3 and 4. Increasing the signifigance of the tail beyond the exponential(1) distribution again results in an increase in the propensity for a single very large price sample to be included in $\tilde{p}_1, \tilde{p}_2, \ldots, \tilde{p}_N$.

Figure 6 explicitly demonstrates that price samples with large values degrade the performance of SAA in the case where $N = 2$; typical samples result in the SAA policy outperforming the MPC policy, but for more extreme events where one sample is very large the reverse occurs and the MPC policy outperforms the SAA policy.

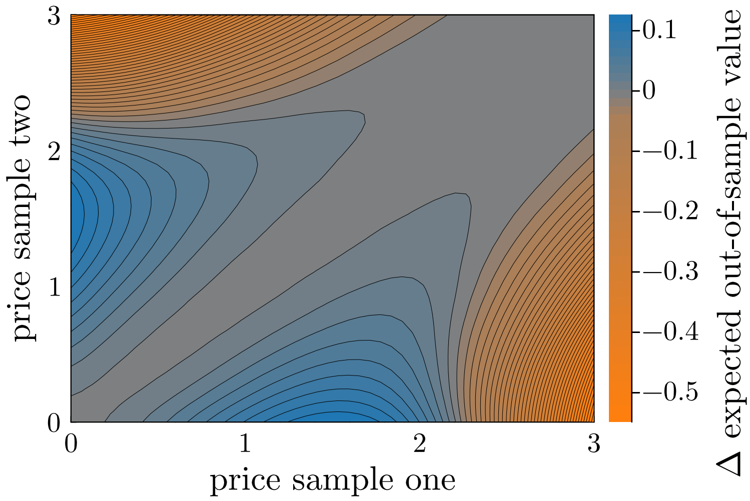

Figure 6: Expected out-of-sample value of SAA minus that of MPC as a function of two price samples p_1 and p_2 over $[0, 3] \times [0, 3]$ for $\tilde{p} \sim \text{LogNormal}\,(-1/2, 1)$. Orange contours indicate regions where the MPC policy outperforms the SAA policy and blue contours indicate the opposite. The contour that the right diagonal lies in is at elevation 0 since the SAA and MPC policies are identical when $p_1 = p_2$.

8 Discussion

We studied the performance of SAA and MPC on a multistage stochastic inventory control problem, finding that MPC can outperform SAA when the underlying price distribution is right-skewed and N is not too large. In the case where the underlying price distribution is exponential and $\beta \to 1$, MPC can outperform SAA regardless of the size of N.

The performance issues which occur when applying sample-based stochastic dynamic programming to this stochastic inventory problem may be alleviated by appending newly observed prices to the sample history and updating the policy before applying it again. However, this is not practical in general as the time complexities of algorithms used to solve for the optimal (in-sample) stochastic dynamic programming policy grow quickly in N. Regardless of this limitation, a first policy must still be constructed, and, in the context of SIC, if this policy is based on a sample of prices which includes a very large price, the same performance issues will occur (albeit to a lesser extent as the influence of the large price diminishes as N grows). For these reasons we have only studied the setting where N is fixed.

The inventory problem we have considered is quite restricted, for example having deterministic dynamics and additional constraints that allow inventory only to decrease. This gives a transient problem of selling inventory down rather than looking at a problem with a steady state component. We have chosen this problem because it enables an analytical solution and a more detailed analysis, but we expect that similar results would be obtained for problems in which there are occasional additional amounts of inventory arriving. For example, if inventory to replenish stocks to level x^0 arrives in each time interval with probability κ then a renewal theory argument shows that the problem of maximizing average reward per unit time is equivalent to SIC with $\beta = 1 - \kappa$.

We have provided an explanation of the better out-of-sample performance of optimizing using sample average prices by viewing it through the lens of distributional robustness. This is not entirely the whole story, since the extent of the improvement depends also on the skew of the underlying ground-truth distribution. Our use of a single average price is also a simplification of model predictive control in practice, which updates the sample averages in a rolling horizon fashion. Nevertheless our results show that model predictive control has some merit beyond computational convenience.

Acknowledgments

We would like to acknowledge helpful discussions with Andrew J. Mason.

References

[1] Edward Anderson and Andy Philpott. Improving sample average approximation using distributional robustness. *INFORMS Journal on Optimization*, 4(1):90–124, 2022.

[2] Dimitri P. Bertsekas. Dynamic programming and suboptimal control: A survey from ADP to MPC. *European Journal of Control*, 11(4):310–334, 2005.

[3] Jeff Bezanson, Alan Edelman, Stefan Karpinski, and Viral B. Shah. Julia: A fresh approach to numerical computing. *SIAM Review*, 59(1):65–98, 2017.

[4] Martin Dyer and Leen Stougie. Computational complexity of stochastic programming problems. *Mathematical Programming*, 106:423–432, 2006.

[5] Peyman M. Esfahani and Daniel Kuhn. Data-driven distributionally robust optimization using the Wasserstein metric: Performance guarantees and tractable reformulations. *Mathematical Programming*, 171:115—-166, 2018.

[6] Thomas Martin. *Stochastic optimization for the procurement of crude oil in refineries.* PhD thesis, École des Ponts ParisTech, 2021.

[7] David Q. Mayne. Model predictive control: Recent developments and future promise. *Automatica*, 50(12):2967–2986, 2014.

[8] Ali Mesbah. Stochastic model predictive control: An overview and perspectives for future research. *IEEE Control Systems Magazine*, 36(6):30–44, 2016.

[9] Alois Pichler and Alexander Shapiro. Mathematical foundations of distributionally robust multistage optimization. *SIAM Journal on Optimization*, 31(4):3044–3067, 2021.

[10] Herbert E. Scarf. A min-max solution of an inventory problem. In Kenneth J. Arrow, Samuel Karlin, and Herbert E. Scarf, editors, *Studies in the Mathematical Theory of Inventory and Production*, pages 201–209. Stanford University Press, CA, 1958.

[11] Alexander Shapiro. On complexity of multistage stochastic programs. *Operations Research Letters*, 34(1):1–8, 2006.

[12] Alexander Shapiro. Analysis of stochastic dual dynamic programming method. *European Journal of Operational Research*, 209(1):63–72, 2011.

[13] Alexander Shapiro, Darinka Dentcheva, and Andrzej Ruszczyński. *Lectures on Stochastic Programming: Modeling and Theory.* Society for Industrial and Applied Mathematics, Philadelphia, PA, 3rd edition, 2021.

[14] Alexander Shapiro and Arkadi Nemirovski. On complexity of stochastic programming problems. In Vaithilingam Jeyakumar and Alexander Rubinov, editors, *Continuous Optimization: Current Trends and Modern Applications*, pages 111–146. Springer US, Boston, MA, 2005.

[15] Nancy L. Stokey, Robert E. Lucas, Jr, and Edward C. Prescott. *Recursive Methods in Economic Dynamics*. Harvard University Press, Cambridge, MA, 1989.

[16] Henri Theil. A note on certainty equivalence in dynamic planning. *Econometrica*, 25(2):346–349, 1957.

[17] William T. Ziemba. Transforming stochastic dynamic programming problems into nonlinear programs. *Management Science*, 17(7):450–462, 1971.