Beyond IID: data-driven decision-making in heterogeneous environments

Anonymous Author(s)
Affiliation
Address
email

Abstract

In this work, we study data-driven decision-making and depart from the classical identically and independently distributed (i.i.d.) assumption. We present a new framework in which historical samples are generated from unknown and different distributions, which we dub *heterogeneous environments*. These distributions are assumed to lie in a heterogeneity ball with known radius and centered around the (also) unknown future (out-of-sample) distribution on which the performance of a decision will be evaluated. We quantify the asymptotic worst-case regret that is achievable by central data-driven policies such as Sample Average Approximation, but also by rate-optimal ones, as a function of the radius of the heterogeneity ball. Our work shows that the type of achievable performance varies considerably across different combinations of problem classes and notions of heterogeneity. We demonstrate the versatility of our framework by comparing achievable guarantees for the heterogeneous version of widely studied data-driven problems such as pricing, ski-rental, and newsvendor. En route, we establish a new connection between data-driven decision-making and distributionally robust optimization.

1 Introduction

In optimization under uncertainty, the desirability of a decision (e.g., inventory) depends on an unknown future outcome (e.g., demand). Typically, past data is collected to be indicative of the future, and hence inform our decision. However, ideal data-driven decision-making requires postulating beliefs about the reliability of the past data, and importantly, whether the future may *deviate* from it. In practice, past data may depend on contexts, some of which can be controlled for, and some of which cannot (unobserved confounders). This may introduce data heterogeneity that is not "correctable."

In this paper, a new framework is introduced for modeling how the future may deviate from past data. At a high level, it accomplishes three goals: $i.)$ capture different forms of future deviation, including no deviation, under the same umbrella; $ii.)$ understand the performance of a central policy, which makes the decision that optimizes the *average* objective value over past data, under different forms of unexpected deviation; $iii.)$ if the central policy performs poorly, then derive modifications that are robust to different forms of anticipated deviation, and quantify the achievable performance.

Our framework is general, capturing different problems but also leading to problem-specific conclusions. The only assumption is that the policy defined in $ii.)$ above can be computed, which requires perfect counterfactual evaluation of any decision on all data points, as opposed to settings where past data are affected by previous actions [4, 12, 24, 32, 33, 34]. This policy is widely studied across different fields, and hereafter we refer to it as *Sample Average Approximation (SAA)* [28, 27, 30].[1]

[1] It is also known under different names, e.g., Empirical Optimization or Empirical Risk Minimization.

Submitted to 36th Conference on Neural Information Processing Systems (NeurIPS 2022). Do not distribute.

1.1 Framework description and general results

Framework (Section 2.1). Our first contribution lies in introducing a framework which models the future outcome as being drawn from an unknown distribution, and past data as being drawn independently from (also unknown) "nearby" distributions. In this way, past data is indicative of the future, and we emphasize that the samples of data can be drawn from different nearby distributions. In accord, we call this setting a *heterogeneous environment*. Using varying definitions of "nearby distribution" (e.g. based on Kolmogorov/Wasserstein distances; see Section 1.2), we analyze different forms of deviation possible between the past and future, and let a radius ϵ bound the allowed deviation. When $\epsilon = 0$, our framework captures the classical independent and identically distributed (i.i.d.) setting, in which past data are drawn from the same distribution as the future outcome [28, 35, 3].

Performance measure (Section 2.2). We consider the performance of different data-driven policies, which map past data into a decision for a given problem; SAA defines a feasible data-driven policy for any problem. Regret is measured as the difference in objective between the policy's decision and the optimal decision knowing the future distribution, taking an expectation over the draws of past data (which affect the policy's decision), any intrinsic randomness in the policy, and the outcome realization (which affects the objective of both the policy's decision and the optimal decision). We then take a worst case regret over all possible distributions that could have been chosen for the future outcome and the data points, to evaluate the performance of a fixed data-driven policy. We note that this performance depends on the problem (see Section 1.2 for examples), the number of data points n, the definition of "nearby distribution," and the radius ϵ.

General reduction in asymptotic regime (Section 3). Our second contribution lies in developing a general reduction in Theorem 1 in which we establish the following: asymptotically, as the number of samples goes to ∞, the worst-case regret of a policy (in a broad subclass we call sample-size-agnostic) is upper-bounded by that of an alternative problem with two major simplifications: $i.$) the worst-case is now taken over only a single distribution for historical environments as opposed to a sequence of heterogeneous distributions; and $ii.$) the DM has access to the exact historical distribution. This upper bound can be seen as a uniform distributionally robust optimization problem under a regret performance metric, and offers a systematic way to derive upper bounds for policies of interest.

General result for Lipschitz-continuous functions (Section 4). We develop problem-independent results that hold whenever the objective function is Lipschitz as a function of the future outcome that materializes. For such environments, we establish in Theorem 2 that SAA scales linearly in ϵ for both the Kolmogorov and Wasserstein forms of deviation.

1.2 Problem-specific, deviation-specific results

Finally, we consider three classical problems in the context of our framework—newsvendor, pricing, and ski-rental—that have received significant attention in the literature across Operations Research, Economics, and Computer Science in the known-prior, data-driven, distributionally robust, adversarial, and advice-augmented settings. Our third contribution lies in deriving problem-specific conclusions in our general heterogeneous environments. These are obtained by using: $i.$) our general reduction and results to upper-bound their asymptotic worst-case regret, and $ii.$) a combination of algorithm design and the development of impossibility results (lower bounds) on the worst-case regret. We analyze the regret for both the SAA policy as well as our robustified policies.

Problem descriptions. First, we consider the newsvendor problem (Section 4.1), in which a decision-maker (DM) must decide how much capacity or inventory to plan in the face of unknown demand. The objective is to minimize the sum of *underage cost*, paid for each unit by which capacity falls short of realized demand, and *overage cost*, paid for each unit of capacity in excess of demand that needs to sit idle. This is a foundational problem in Operations Research, with early focus on the setting where the demand distribution is known [see 36], and since then studied in various distributionally robust [41, 18, 38] settings but also in data-driven ones [31, 11, 5]. Newsvendor captures the fundamental tradeoff in guessing an unknown value when there are asymmetric consequences for going under/over.

Second, we consider the single-item pricing problem (Section 5.1), in which a DM must find the best price to offer to a customer with unknown willingness to pay (wtp). The objective is to maximize revenue. This is a foundational problem in Economics, with early focus on the setting with known prior [see 37, 40], and more recently in data-driven [17, 22, 2, 1, 13] and distributionally robust [8, 19] settings. Pricing captures the fundamental tradeoff between margin and volume.

Finally, we consider the ski-rental problem (Section 5.2), in which a DM must decide each day whether to buy skis or keep renting, without knowing the length of their trip. The objective is to minimize total cost, where buying eliminates the rental cost going forward. Although typically presented as an online problem that requires a decision to be made every day, any (randomized) buy-or-rent policy can be described by a (randomized) duration for which to rent skis, before committing to buying (if the trip has not ended by then). This is a foundational problem in Computer Science, critical to the development of competitive analysis [see 7] which considers the *adversarial* setting where nothing is known, and recently also important to the development of *advice-augmented algorithms* which receive a prediction on the length of the ski trip [39, 15]. Ski rental captures the fundamental tradeoff between proceeding with a suboptimal process vs. paying the cost of refactoring.

Definitions of "nearby distribution." Our framework generally models past data to be drawn from distributions that lie within an ϵ-distance of the future distribution, for *any* notion of distance between probability distributions. For our problems of interest, we focus on two specific distance metrics: Kolmogorov and Wasserstein (see Section 2.1 for formal definitions). This dichotomy suffices to demonstrate the importance of distance metric in determining performance in our problems of interest, and moreover, represents contrasting inductive philosophies about how the future might deviate from the past: Wasserstein allows *every* data point to be erroneous by ϵ, whereas Kolmogorov allows an ϵ fraction of data points to be *arbitrarily* erroneous.

SAA vs. robustified policies. Finally, performance is demarcated by the policy used. We first provide tight asymptotic performance bounds for the SAA policy, which we note is agnostic to the heterogeneity that might arise. We then analyze the best-possible asymptotic performance bounds that can be achieved by any policy when the notion of distance and the radius ϵ are known.

In Table 1, we present a high level summary of our results across problem classes. We assume that the support of the unknown environment is in $[0, M]$, and we track the dependence on both the heterogeneity radius ϵ, as well as M. These results are for the asymptotic setting in which $n \to \infty$.

			Kolmogorov	Wasserstein
Lipschitz	Newsvendor (§4.1)	SAA	$\Theta(M\epsilon)$	$\Theta(\epsilon)$
		best policy	$\Theta(M\epsilon)$	$\Theta(\epsilon)$
Beyond Lipschitz	Pricing (§5.1)	SAA	$\Theta(M\epsilon)$	$\Omega(M)$
		best policy	$\Theta(M\epsilon)$	$\Theta(\sqrt{M\epsilon})$
	Ski-rental (§5.2)	SAA	$\Theta(M\epsilon)$	$\Theta(1)$
		best policy	$\Theta(\epsilon)$	$\Theta(\sqrt{\epsilon})$

Table 1: **SAA and achievable performance for different problem classes and heterogeneity balls.**

For newsvendor problems, through a combination of the general result for Lipschitz-continuous problems and lower bounds, we show that SAA actually achieves the optimal rate for asymptotic worst-case regret both as a function of ϵ and M and that this rate is linear.

We then turn to problems whose objective is not Lipschitz. For the pricing problem, we establish that while under Kolmogorov distance, SAA is near-optimal and the optimal performance is again of order $M\epsilon$, the picture is starkly different under the Wasserstein distance. In this case, SAA has arbitrarily poor performance. We then show that a robustification procedure that deflates SAA by an appropriate factor can yield the best achievable asymptotic worst-case regret that shrinks to zero at rate $\sqrt{M\epsilon}$. The comparison of the pricing problem to newsvendor highlights the important fact that the quality of SAA in heterogeneous environments is highly dependent on the problem class. Furthermore, the dependence of the achievable performance in the support scale M and the heterogeneity ϵ varies considerably across problem classes and notions of heterogeneity.

A second prototypical problem that is not Lipschitz is the ski-rental problem. While SAA was near-optimal for the Kolmogorov distance for pricing and newsvendor, we establish the surprising fact that SAA, even under the Kolmogorov distance, has arbitrarily poor performance. Indeed, its asymptotic worst-case regret scales as $\Theta\left(M\epsilon\right)$, while the optimal policy achieves a performance of $\Theta\left(\epsilon\right)$. As a consequence, the performance of SAA can be arbitrarily worse than the optimal one as the support grows to infinity. Meanwhile, for the Wasserstein distance, we show that the performance of SAA is also suboptimal. Indeed, SAA incurs a regret of $\Theta(1)$; but by inflating SAA appropriately, one can construct a policy that achieves order $\sqrt{\epsilon}$ performance, which is the best dependence on ϵ.

These examples demonstrates that SAA fails for both heterogeneity types for ski-rental, but also more broadly the intricate interplay between problem class and type of heterogeneity.

Connection to offline data corruption models. Different notions of offline data corruptions and robustness have been previously studied in statistics [23, 20], learning theory [42, 21, 26, 25, 42, 29, 14, 6, 16] and, more recently in revenue management [9, 8, 10, 19]. Most approaches consider ϵ-contamination settings [23] in which the adversary can shift a small proportion of data arbitrarily far from the true distribution. They consider either oblivious adversaries who fix a *single* corrupted distribution from which samples are generated or adaptive one who observe samples before corrupting them. [6] establishes that adaptive adversaries yield equivalent performance to oblivious ones in many settings. Closest to us are [8, 19] who study optimal auctions. They assume the distribution is common across all past observations; the former focuses on regret and the the latter on a ratio metric when data is generated from MHR distributions close in Kolmogorov distance. The framework we develop is anchored around a setting that allows for *heterogeneous* "nearby" distributions for each past observation, and is general in that it allows us to unify a variety of problems/metrics and highlight how these affect the levels of achievable performance.

2 Problem formulation: data-driven decisions in heterogeneous environments

We consider a general stochastic optimization problem in which a DM wants to optimize an objective in a stochastic environment. Formally, we denote by $\mathcal{X}$ the decision space and by Ξ the space of environment realizations. The objective of the DM is represented by a function $g : \mathcal{X} \times \Xi \to \mathbb{R}$ which maps each pair of decision $x \in \mathcal{X}$ and environment realization $\xi \in \Xi$ to an objective value.

The DM faces a subset of possible probability measures $\mathcal{P} \subset \Delta(\Xi)$, where $\Delta(\Xi)$ denotes the space of probability measures supported on Ξ. For every probability measure $\mu \in \mathcal{P}$, the goal of the DM is to optimize the objective defined as

$$\mathcal{G}(\cdot, \mu) : \begin{cases} \mathcal{X} \to \mathbb{R} \\ x \mapsto \mathbb{E}_{\xi \sim \mu}\left[g(x, \xi)\right]. \end{cases} \tag{1}$$

We extend the definition of $\mathcal{G}(\cdot, \mu)$ to allow for randomized decisions $\rho \in \Delta(\mathcal{X})$, in which case we define $\mathcal{G}(\rho, \mu) = \mathbb{E}_{x \sim \rho}\left[\mathcal{G}(x, \mu)\right]$ to also take an expectation over the randomized decision.

Remark that depending on the nature of the problem, the objective could represent for example a function to maximize (profit) or to minimize (cost). In turn, we formally define for every $\mu \in \mathcal{P}$, the optimal objective $\text{opt}(\mu)$ as the objective value achieved by an *optimal action*, given by an *oracle* operator `ORACLE` that is a mapping from $\Delta(\Xi)$ to $\mathcal{X}$.

2.1 Heterogeneous environments

In the absence of distributional knowledge on the underlying environment, a common approach consists in assuming that historical samples are independent and identically drawn from a fixed probability measure μ. We relax the assumption that the samples are identically distributed and introduce a generalization of this framework that we refer to as *ϵ-heterogeneity*.

Let $d : \mathcal{P} \times \mathcal{P} \to \mathbb{R}$ be a metric on $\mathcal{P}$. Given a target probability measure $\mu \in \mathcal{P}$ and a level of heterogeneity ϵ, we define the heterogeneity ball anchored around μ and with radius ϵ as

$$\mathcal{U}_\epsilon(\mu) := \{\nu \in \mathcal{P},\ \text{s.t}\ d(\mu, \nu) \leq \epsilon\}.$$

The heterogeneity ball $\mathcal{U}_\epsilon(\mu)$ includes all distributions that are within a distance of ϵ from μ. Next, we formally define a data-driven decision-making problem in a heterogeneous environment.

Definition 1 (Data-driven decision-making problem in heterogeneous environment) *A data-driven decision-making problem in a heterogeneous environment is defined by the tuple: $(\mathcal{X}, \Xi, \mathcal{P}, d, g, \texttt{ORACLE})$, where $\mathcal{X}$ is the set of possible actions, Ξ is the environment space, $\mathcal{P}$ is the space of measures nature can select from, d is a metric on the space of measures, g is an objective function, and* `ORACLE` *is a mapping from environment measures to target actions.*

Focal metrics. The definition above is general, and the framework can be applied to a variety of metrics. In this work, we aim at developing an understanding of the effect of the extent and the type

of heterogeneity in various problem classes. We focus on two prototypical and central distances: *Kolmogorov* and *Wasserstein*. We formally define them in one dimension as follows.

Definition 2 (Kolmogorov and Wasserstein distances) *Given an environment space $\Xi \subset \mathbb{R}$ and a subset of probability measures $\mathcal{P}$ supported on Ξ, the Kolmogorov (resp. Wasserstein) distance d_K (resp. d_W) is defined for every $\mu_1, \mu_2 \in \mathcal{P}$ as*

$$d_K(\mu_1, \mu_2) = \sup_{\xi \in \Xi} |F_1(\xi) - F_2(\xi)| \qquad \textit{and} \qquad d_W(\mu_1, \mu_2) = \int_{\xi \in \Xi} |F_1(\xi) - F_2(\xi)| \, d\xi,$$

where F_1 (resp. F_2) is the cumulative distribution associated with μ_1 (resp. μ_2).

Note that when Ξ is a compact interval of $\mathbb{R}$, both distances are related as for every μ_1, $\mu_2 \in \mathcal{P}$,

$$d_W(\mu_1, \mu_2) \leq \operatorname{diam} \Xi \cdot d_K(\mu_1, \mu_2). \tag{2}$$

2.2 Data-driven policies and performance

Data-driven policies. In most settings, the probability measure μ is unknown to the DM and the optimal objective $\operatorname{opt}(\mu)$ is not achievable. In turn, the DM only observes a sequence of historical samples $\xi_1, \xi_2, \ldots, \xi_n$ representing the previous environment realizations they observed. A data-driven policy is then a mapping from the past samples to a (randomized) decision. Define the empirical probability measure $\hat{\mu}_{\boldsymbol{\xi^n}}$ as

$$\hat{\mu}_{\boldsymbol{\xi^n}}(\xi) := \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}\{\xi = \xi_i\}, \qquad \text{for every } \xi \in \Xi.$$

In a setting in which the order of the samples does not matter (as the one we will consider), to represent a policy taking as input $\boldsymbol{\xi^n} := (\xi_1, \ldots, \xi_n)$, one only needs to consider mappings that take as an input the number of samples and the empirical measure. We define a data-driven policy π as a mapping from $\mathbb{N}^* \times \mathcal{P}$ to $\Delta(\mathcal{X})$, the space of measures on the action space $\mathcal{X}$, that associates $(n, \hat{\mu})$ to a distribution of actions $\pi(n, \hat{\mu})$. We let Π denote the set of all such mappings. We assume that such a mapping can be made with knowledge of the metric d and radius ϵ. Moreover, we say that π is *sample-size-agnostic* if its decision does not depend on n, in which case we write $\pi(\hat{\mu})$ for brevity.

A widely-studied policy that achieves excellent performance across a vast range of data-driven decision-making problems is Sample Average Approximation (SAA). Given an empirical distribution $\hat{\mu}$, SAA (denoted by π^{SAA}) selects the target action $\texttt{ORACLE}(\hat{\mu})$, the best action associated with $\hat{\mu}$.

Performance through worst-case regret. Given an out-of-sample distribution $\mu \in \mathcal{P}$ the best achievable objective is $\operatorname{opt}(\mu)$. For a sequence of historical distributions $\nu_1, \ldots, \nu_n$ in $\mathcal{U}_\epsilon(\mu)$, the expected objective of a policy π is

$$\mathbb{E}_{\xi_i \sim \nu_i}\left[\mathcal{G}\left(\pi(n, \hat{\mu}_{\boldsymbol{\xi^n}}), \mu\right)\right].$$

In turn, we define the expected (absolute) regret of a policy $\pi \in \Pi$ against an out-of-sample distribution $\mu \in \mathcal{P}$ and a sequence of n historical distributions $\nu_1, \ldots, \nu_n \in \mathcal{U}_\epsilon(\mu)$ as

$$\mathcal{R}_n(\pi, \mu, \nu_1, \ldots, \nu_n) := \left|\mathbb{E}_{\xi_i \sim \nu_i}\left[\mathcal{G}\left(\pi(n, \hat{\mu}_{\boldsymbol{\xi^n}}), \mu\right)\right] - \operatorname{opt}(\mu)\right|.$$

We note that we use the absolute value in this definition, as our framework applies to minimization and maximization problem. Once the type of problem is specified (i.e., the operator associated with $\texttt{ORACLE}$), this definition of regret coincides with the usual regret for minimization or maximization problems, as we will see in Sections 4 and 5, when describing specific applications.

In this work, we will assess the performance of a policy based on its *worst-case expected regret*. Given an instance $\mathcal{I} = (\mathcal{X}, \Xi, \mathcal{P}, d, g, \texttt{ORACLE})$, we define the worst-case expected regret of a data-driven policy $\pi \in \Pi$ for a degree of heterogeneity $\epsilon \geq 0$ as

$$\mathfrak{R}^{\pi}_{\mathcal{I},n}(\epsilon) := \sup_{\mu \in \mathcal{P}} \sup_{\nu_1, \ldots, \nu_n \in \mathcal{U}_\epsilon(\mu)} \mathcal{R}_n(\pi, \mu, \nu_1, \ldots, \nu_n). \tag{3}$$

We finally define the *worst-case asymptotic regret* of a data-driven policy $\pi \in \Pi$ in a ϵ-heterogeneous environment as the number of samples n grows large as

$$\mathfrak{R}^{\pi}_{\mathcal{I},\infty}(\epsilon) := \limsup_{n \to \infty} \mathfrak{R}^{\pi}_{\mathcal{I},n}(\epsilon). \tag{4}$$

We sometimes write $\mathfrak{R}^{\pi}_{K,\infty}(\epsilon)$ or $\mathfrak{R}^{\pi}_{W,\infty}(\epsilon)$ instead of $\mathfrak{R}^{\pi}_{\mathcal{I},\infty}(\epsilon)$ when the problem class is specified to highlight the dependence on the Kolmogorov or Wasserstein metrics.

3 Reduction to distributionally robust optimization in the asymptotic regime

A significant challenge in the analysis of problem (4) stems from the difficulty to characterize a potentially complex worst-case sequence of historical probability measures $\nu_1, \ldots, \nu_n$ and out-of-sample measure μ. We illustrate in the appendix that analyzing every element in the sequence of historical probability measures as a function of n is a priori challenging as nature will attempt to use different distributions for each of the past environments. We will show that, asymptotically, it suffices to consider cases in which nature selects a common distribution for all of the past environments. Formally, we define for a sample-size-agnostic policy $\pi \in \Pi$ the following optimization problem:

$$\sup_{\mu \in \mathcal{P}} \sup_{\nu \in \mathcal{U}_\epsilon(\mu)} |\text{opt}(\mu) - \mathcal{G}\left(\pi(\nu), \mu\right)|. \tag{5}$$

While Problem (5) resembles Problem (3), it differs on two crucial dimensions: $i.)$ in Problem (5), nature selects the *same* distribution ν in the heterogeneity ball for all past environments; $ii.)$ in Problem (5), the policy π is assumed to know the actual distribution of the past environments ν, whereas in Problem (3), the policy only observed the empirical distribution of past realizations.

Next, we will establish that Problems (5) and (3) are tightly connected under some mild assumptions. We introduce the following definition on the distance d defining the heterogeneity.

Definition 3 (Empirical triangular convergence) *We say that a distance d on the space of probability measures satisfies the empirical triangular convergence (ETC) property if and only if for every triangular array sequence of probability measures $(\mu_{i,n})_{1 \le i \le n, n \in \mathbb{N}^*}$ all belonging to $\mathcal{P}$, we have*

$$\lim_{n \to \infty} d(\hat{\mu}_n, \bar{\mu}_n) = 0 \qquad a.s.,$$

where $\hat{\mu}_n(\xi) := \frac{1}{n} \sum_{i=1}^n \mathbb{1}\{\xi = \xi_{i,n}\}$ for samples $\xi_{i,n} \sim \mu_{i,n}$ and $\bar{\mu}_n = \frac{1}{n} \sum_{i=1}^n \mu_{i,n}$.

The ETC property requires that the sequence of empirical distributions converges to the average of ground truth distributions for arbitrary triangular arrays. It is satisfied by many common distances on probability spaces, including the Kolmogorov and Wasserstein distances on compact sets, as we discuss in the appendix. We now state our first main result which establishes a fundamental reduction.

Theorem 1 (Upper bound reduction) *Let $\mathcal{I} = (\mathcal{X}, \Xi, \mathcal{P}, d, g, \texttt{ORACLE})$ be a data-driven decision problem in a heterogeneous environment and let $\pi \in \Pi$ be a sample-size-agnostic policy. Assume that g is bounded on $\mathcal{X} \times \Xi$ and that d satisfies the empirical triangular convergence property. Then*

$$\mathfrak{R}^{\pi}_{\mathcal{I},\infty}(\epsilon) \le \lim_{\eta \to \epsilon^+} \sup_{\mu \in \mathcal{P}} \sup_{\nu \in \mathcal{U}_\eta(\mu)} |opt(\mu) - \mathcal{G}\left(\pi(\nu), \mu\right)|. \tag{6}$$

Theorem 1 establishes, that under some mild conditions, in the asymptotic regime, one may restrict attention to problem (5). This reduction significantly simplifies the analysis of sample-size-agnostic policies[2] as problem (5) does not involve any form of randomness in the input received by the policy, and only allows the adversary to select a single distribution. We show in the appendix that taking the limit of $\eta \to \epsilon$ on the RHS of (6) is necessary. We also show in the appendix that Theorem 1 holds whenever Ξ is finite, for any distance metric, even without a sample-size-agnostic policy.

4 Lipschitz-continuous problems

In this section, we study a broad class of problems for which the underlying objective function $g(x, \cdot)$ is Lipschitz-continuous as a function of the environment.

Theorem 2 (Upper bounds for Lipschitz-continuous objectives) *Suppose that $\Xi = [0, M]$ for some $M > 0$. Let $\mathcal{I}_K = (\mathcal{X}, \Xi, \mathcal{P}, d_K, g, \texttt{ORACLE})$ and $\mathcal{I}_W = (\mathcal{X}, \Xi, \mathcal{P}, d_W, g, \texttt{ORACLE})$ be instances of the data-driven decision problem under the Kolmogorov and Wasserstein distances, respectively. Assume that for every $x \in \mathcal{X}$, the function $g(x, \cdot)$ is L-Lipschitz-continuous and that* ORACLE *is either a minimization or maximization oracle. Then for every $\epsilon \ge 0$,*

$$\mathfrak{R}^{\pi^{SAA}}_{\mathcal{I}_K,\infty}(\epsilon) \le 2L \cdot M \cdot \epsilon \qquad \text{and} \qquad \mathfrak{R}^{\pi^{SAA}}_{\mathcal{I}_W,\infty}(\epsilon) \le 2L \cdot \epsilon.$$

[2]Note that considering sample-size-agnostic policies is not very restrictive. As we will see in Section 4 and 5 this class is sufficient to achieve optimal rates for the asymptotic worst-case regret as a function of ϵ.

Theorem 2 implies that for this class of problems, the asymptotic regret of SAA vanishes as the radius of the heterogeneity ball goes to 0. Furthermore, this theorem shows that the dependence of the asymptotic worst-case regret in the heterogeneity parameter ϵ is at most linear. We show next that this dependence is the best possible for this class of problems and this notion of heterogeneity.

4.1 Newsvendor problem

Recall the newsvendor problem described in Section 1.2.Using $M > 0$ to denote an upper bound on demand, we let $\Xi = \mathcal{X} = [0, M]$, and our regret bounds may (or may not) depend on M. The objective function is the newsvendor cost which is parameterized by two quantities: c_u the underage cost and c_o the overage cost. The cost of carrying an inventory level $x \in \mathcal{X}$ and then observing demand $\xi \in \Xi$ is,

$$g(x, \xi) = c_u (\xi - x)^+ + c_o (x - \xi)^+ ,$$

where $(\cdot)^+ := \max\{\cdot, 0\}$ is the positive part operator. The measure space $\mathcal{P}$ is the set of probability measures on Ξ without restrictions and for every $\mu \in \mathcal{P}$, we define the cost of the oracle as $\text{opt}(\mu) = \min_{x\in\mathcal{X}} \mathcal{G}(x, \mu)$. Therefore the oracle mapping is defined as $\texttt{ORACLE} : \mu \mapsto \arg\min_{x\in\mathcal{X}} \mathcal{G}(x, \mu)$.

We note that the objective function g satisfies the L-Lipschitz-continuity property on the environment realization variable ξ with $L = \max(c_u, c_o)$. Theorem 2 thus directly yields an upper bound on the achievable regret for this class of problems, for both Kolmogorov and Wasserstein, through the regret of SAA. We next show that these dependencies are the best possible, under any policy.

Proposition 1 (Heterogeneous newsvendor) *For the data-driven newsvendor problem in a heterogenous environment under the Kolmogorov and Wasserstein distances, given any $\epsilon \geq 0$,*

$$\begin{aligned} \frac{c_u + c_o}{2} \cdot \epsilon \cdot M \quad &\leq \inf_{\pi\in\Pi} \mathfrak{R}^{\pi}_{K,\infty}(\epsilon) \leq \mathfrak{R}^{\pi^{SAA}}_{K,\infty}(\epsilon) \leq \quad 2\max(c_u, c_o) \cdot \epsilon \cdot M, \\ \frac{c_u + c_o}{2} \cdot \epsilon \quad &\leq \inf_{\pi\in\Pi} \mathfrak{R}^{\pi}_{W,\infty}(\epsilon) \leq \mathfrak{R}^{\pi^{SAA}}_{W,\infty}(\epsilon) \leq \quad 2\max(c_u, c_o) \cdot \epsilon. \end{aligned}$$

Proposition 1 first shows that SAA is robust to deviation under multiple forms of heterogeneity and that the regret scales linearly with the radius of the heterogeneity ball for both distances. We show furthermore that the linear dependence is, for the newsvendor problem, the best rate achievable for any data-driven policy. We also note that the lower bound obtained for the newsvendor depends linearly on the Lipschitz constant L through the factor $(c_u + c_o)$, since scaling the costs c_u, c_o would clearly scale the regret.

Beyond newsvendor, the asymptotic worst-case regret of SAA over the general class of L-Lipschitz-continuous functions is of the order $\Theta(L \cdot \epsilon)$. This suggests that the performance of SAA may deteriorate for problems in which the objective function g is not smooth.

5 Beyond Lipschitz-continuous problems

5.1 Pricing problem

Recall the pricing problem presented in Section 1.2. The environment space $\Xi = [0, M]$ represents the set of possible wtp. Similarly, the set of possible prices $\mathcal{X}$ is $[0, M]$. The set of considered probability measures $\mathcal{P}$ is the whole set of probability measures on $[0, M]$ without restriction. The revenue generated by the DM with a price decision $x \in \mathcal{X}$ when facing a customer with wtp $\xi \in \Xi$ is,

$$g(x, \xi) = x \cdot \mathbb{1}\{x \leq \xi\} .$$

Therefore for any probability measure $\mu \in \mathcal{P}$ associated with the cumulative distribution F, we have that $\mathcal{G}(x, \mu) = x \cdot (1 - F(x))$. The goal of the DM is to set a price x in order to maximize its revenue. Equivalently, the goal is to minimize the regret against the oracle which posts the optimal price given a wtp distribution. Formally, the oracle operator is defined as, $\texttt{ORACLE} : \mu \mapsto \arg\max_{x\in\mathcal{X}} \mathcal{G}(x, \mu)$.

Kolmogorov pricing. Remark that the pricing objective is not Lipschitz-continuous and that one cannot use the argument presented in Section 4.1. We show in the appendix that despite the lack of smoothness of the objective, SAA is still near-optimal in the asymptotic regime under the Kolmogorov distance. We derive an upper bound on the asymptotic worst-case regret of SAA and a matching universal lower bound showing that the best achievable performance is order $\Theta(M\epsilon)$.

Failure of SAA in pricing with Wasserstein heterogeneity. In opposition to the strong performance achieved by SAA for pricing against Kolmogorov heterogeneity, our next proposition shows that this result does not hold if one considers instead the Wasserstein heterogeneity ball.

Proposition 2 (SAA and Wasserstein pricing) *For the data-driven pricing problem in a heterogenous environment under the Wasserstein distance, we have for $\epsilon > 0$ that*

$$\mathfrak{R}^{\pi^{SAA}}_{W,\infty}(\epsilon) = M.$$

Proposition 2 shows that, in stark contrast with Lipschitz problems, for the pricing problem, the asymptotic regret of SAA under Wasserstein heterogeneity does not shrink to zero as ϵ goes to zero. As a matter of fact, the regret of M is the worst possible, as the revenue of the oracle is bounded from above by M and the revenue of any data-driven policy is bounded from below by 0. Quite notably, this result holds for any level of heterogeneity ϵ, and an arbitrarily small deviation from the i.i.d. case leads to extremely poor performance, even with infinite data. We next propose a policy which "robustifies" SAA and we analyze it under the Wasserstein distance.

Robustification of SAA for Wasserstein pricing. The key challenge in pricing is that for two distributions μ and ν such that $d_W(\mu, \nu)$ is arbitrarly small, a price x could have a high revenue for μ but a low revenue for ν, and the difference in revenue between distributions may be arbitrarily large. This enabled us to construct an instance for which the optimal price for the in-sample distribution ν generates a low revenue for the out-of-sample distribution μ.

Our main structural result in this section shows that one can control the difference in revenue between two different distributions for two *different* prices. Formally, we show the following.

Proposition 3 *Let μ and $\nu \in \mathcal{P}$ be two different distributions and let x_1 and x_2 be two different prices such that $x_1 < x_2$. We have that,*

$$\mathcal{G}(x_2, \nu) - \mathcal{G}(x_1, \mu) \leq (x_2 - x_1) + M \cdot \frac{d_W(\mu, \nu)}{x_2 - x_1}.$$

Proposition 3 unveils an interesting tradeoff faced by a DM who prices in a heterogeneous environment. By selecting appropriately the difference between two prices x_1 and x_2, one can ensure that the gap in revenues for two different distributions can be controlled by the Wasserstein distance.

We next restrict attention to the class of policies which additively *deviate* from the action selected by SAA and leverage Proposition 3 to define the correct deviation parameter.

Definition 4 (δ-SAA policies) *For any one-dimensional data-driven decision-making problem we say that a policy is a δ-SAA policy if for every empirical distribution $\hat{\mu}$, the policy selects the action $\texttt{ORACLE}(\hat{\mu}) + \delta$, for some $\delta \in \mathbb{R}$ that could be positive or negative. We denote this policy by $\pi^{SAA(\delta)}$.*

We note that when $\delta = 0$, we recover the usual SAA policy. Our next result establishes the best possible regret dependence on ϵ and M for the class of pricing problems, using a δ-SAA policy.

Theorem 3 (Deviation for Wasserstein pricing) *For the data-driven pricing problem in a heterogenous environment under the Wasserstein distance, for any $\epsilon > 0$, let $\tilde{\delta} = -\sqrt{M \cdot \epsilon}$. Then,*

$$C_1 \cdot \sqrt{M \cdot \epsilon} \leq \inf_{\pi \in \Pi} \mathfrak{R}^{\pi}_{W,\infty}(\epsilon) \leq \mathfrak{R}^{\pi^{SAA(\tilde{\delta})}}_{W,\infty}(\epsilon) \leq C_2 \cdot \sqrt{M \cdot \epsilon},$$

where C_1 and C_2 are universal constants.

Theorem 3 shows that by appropriately deflating the price posted by SAA, one is able to be robust against Wasserstein heterogeneity and achieve a worst-case asymptotic regret that vanishes to 0 as the degree of heterogeneity ϵ goes to 0. This proves that one may operate efficiently for pricing under Wasserstein heterogeneity as ϵ goes to 0. Our result establishes that in contrast with Lipschitz problems, it is now impossible to achieve a linear dependence in ϵ (since $\sqrt{\epsilon} > \epsilon$), with the regret growing at a rate of $\Omega(\sqrt{M \cdot \epsilon})$ under any policy. A small radius ϵ has a significantly higher impact in pricing problems than in, e.g., newsvendor problems.

5.2 Ski-rental problem

We now consider the ski-rental problem in which renting skis costs $\$1$ per unit of time while buying them costs $\$b$ up-front, for some real value b. The environment space Ξ represents the set of possible lengths of the ski trip and a decision x represents the duration after which skis should be bought (if the ski trip has not ended by that time). We call x the *rental duration*. Let $\mathcal{X} = \Xi = [0, M]$, where we note that setting $x = M$ indicates that skis should never be bought. The set of probability measures $\mathcal{P}$ is the whole space of probability measures supported on $[0, M]$. For every rental duration $x \in \mathcal{X}$ and any trip length $\xi \in \Xi$, the cost incurred is

$$g(x, \xi) = \xi \cdot \mathbb{1}\{\xi \leq x\} + (b + x) \cdot \mathbb{1}\{\xi > x\}.$$

Finally, as the goal is to minimize cost we have that, $\texttt{ORACLE} : \mu \mapsto \arg\min_{x \in \mathcal{X}} \mathcal{G}(x, \mu)$.

Wasserstein ski-rental. We proved in Section 5.1 that SAA performs arbitrarily poorly in pricing under Wasserstein heterogeneity. We now show that SAA falters similarly for the Wasserstein ski-rental problem. Furthermore, we derive a deviation policy which inflates the SAA action and achieves the best rate possible as a function of the heterogeneity radius. Formally, we show the following.

Theorem 4 *For the data-driven ski-rental problem in a heterogenous environment under the Wasserstein distance, we have for $\epsilon > 0$ that,*

$$\mathfrak{R}^{\pi^{SAA}}_{W,\infty}(\epsilon) = \Theta(1) \qquad \textit{and} \qquad \inf_{\pi \in \Pi} \mathfrak{R}^{\pi}_{W,\infty}(\epsilon) = \Theta\left(\sqrt{\epsilon}\right).$$

Theorem 4 formalizes the failure of SAA for ski-rental under Wasserstein distance. A notable fact about the ski-rental problem under Wasserstein heterogeneity is that both the regret of SAA and of the optimal data-driven decision do not scale with the size of the support. For pricing, SAA scaled linearly with M and the optimal policy scales in $\sqrt{M}$. The fact that $\mathfrak{R}^{\pi^{\text{SAA}}}_{W,\infty}(\epsilon) = O(1)$ in ski-rental requires a separate non-trivial proof in Theorem 4.[3]

Kolmogorov ski-rental. Under the Kolmogorov distance, we had seen that SAA was near-optimal for both the newsvendor and pricing problems. We establish next that the performance of SAA for the ski-rental problem under Kolmogorov distance is, surprisingly, highly suboptimal.

Theorem 5 *For the data-driven ski-rental problem in a heterogenous environment under the Kolmogorov distance, we have for $\epsilon > 0$ that,*

$$\mathfrak{R}^{\pi^{SAA}}_{K,\infty}(\epsilon) = \Theta(M \cdot \epsilon) \qquad \textit{and} \qquad \inf_{\pi \in \Pi} \mathfrak{R}^{\pi}_{K,\infty}(\epsilon) = \Theta(\epsilon).$$

Theorem 5 shows that the asymptotic worst-case regret of SAA scales linearly with the radius of heterogeneity ϵ and with the size of the support M. We also characterize the rate of the optimal asymptotic worst-case regret as a function of ϵ. The upper bound on the asymptotic regret is obtained in [15] through a variant of the SAA policy, which caps the maximum number of days to rent. We complement this result by providing a matching lower bound. Note that, in contrast to the pricing and newsvendor problems, the scaling of SAA is not optimal in M. As the scale of the support grows, the asymptotic worst-case regret of SAA is considerably worse than the optimal achievable rate.

6 Conclusion

All in all, the present results offer a systematic way of understanding and quantifying the implications of operating in heterogeneous environments. Our framework enables to develop a common language to compare achievable performance across central classes of problems. This opens many questions, such as, e.g., quantifying of performance with finite samples in heterogeneous environments and, isolating the properties of the policies and the framework elements that drive different levels of achievable performance.

[3] The fact that $\mathfrak{R}^{\pi^{\text{SAA}}}_{W,\infty}(\epsilon) \leq M$ in pricing was vacuous.

References

[1] Allouah, A., Bahamou, A. and Besbes, O. [2022], 'Pricing with samples', *Operations Research* **70**(2), 1088–1104.

[2] Babaioff, M., Gonczarowski, Y. A., Mansour, Y. and Moran, S. [2018], Are two (samples) really better than one?, *in* 'Proceedings of the 2018 ACM Conference on Economics and Computation', pp. 175–175.

[3] Bertsimas, D., Gupta, V. and Kallus, N. [2018], 'Robust sample average approximation', *Mathematical Programming* **171**(1), 217–282.

[4] Besbes, O., Gur, Y. and Zeevi, A. [2014], 'Stochastic multi-armed-bandit problem with non-stationary rewards', *Advances in neural information processing systems* **27**.

[5] Besbes, O. and Mouchtaki, O. [2021], 'How big should your data really be? data-driven newsvendor: learning one sample at a time', *working paper, Columbia University* .

[6] Blanc, G., Lange, J., Malik, A. and Tan, L.-Y. [2021], 'On the power of adaptivity in statistical adversaries', *arXiv preprint arXiv:2111.10352* .

[7] Borodin, A. and El-Yaniv, R. [2005], *Online computation and competitive analysis*, cambridge university press.

[8] Brustle, J., Cai, Y. and Daskalakis, C. [2020], Multi-item mechanisms without item-independence: Learnability via robustness, *in* 'Proceedings of the 21st ACM Conference on Economics and Computation', pp. 715–761.

[9] Cai, Y. and Daskalakis, C. [2017], Learning multi-item auctions with (or without) samples, *in* '2017 IEEE 58th Annual Symposium on Foundations of Computer Science (FOCS)', IEEE, pp. 516–527.

[10] Chen, X. and Wang, Y. [2020], 'Dynamic pricing with demand learning in the presence of outlier customers', *https://ssrn.com/abstract=3650656* .

[11] Cheung, W. C. and Simchi-Levi, D. [2019], 'Sampling-based approximation schemes for capacitated stochastic inventory control models', *Mathematics of Operations Research* **44**(2), 668–692.

[12] Cheung, W. C., Simchi-Levi, D. and Zhu, R. [2022], 'Hedging the drift: Learning to optimize under nonstationarity', *Management Science* **68**(3), 1696–1713.

[13] Daskalakis, C. and Zampetakis, M. [2020], More revenue from two samples via factor revealing sdps, *in* 'Proceedings of the 21st ACM Conference on Economics and Computation', pp. 257–272.

[14] Diakonikolas, I., Kamath, G., Kane, D., Li, J., Moitra, A. and Stewart, A. [2019], 'Robust estimators in high-dimensions without the computational intractability', *SIAM Journal on Computing* **48**(2), 742–864.

[15] Diakonikolas, I., Kontonis, V., Tzamos, C., Vakilian, A. and Zarifis, N. [2021], Learning online algorithms with distributional advice, *in* 'International Conference on Machine Learning', PMLR, pp. 2687–2696.

[16] Duchi, J. C. and Namkoong, H. [2021], 'Learning models with uniform performance via distributionally robust optimization', *The Annals of Statistics* **49**(3), 1378 – 1406.

[17] Fu, H., Immorlica, N., Lucier, B. and Strack, P. [2015], Randomization beats second price as a prior-independent auction, *in* 'Proceedings of the Sixteenth ACM Conference on Economics and Computation', pp. 323–323.

[18] Gallego, G. and Moon, I. [1993], 'The distribution free newsboy problem: review and extensions', *Journal of the Operational Research Society* **44**(8), 825–834.

[19] Guo, W., Jordan, M. and Zampetakis, E. [2021], 'Robust learning of optimal auctions', *Advances in Neural Information Processing Systems* **34**.

[20] Hampel, F. R. [1971], 'A general qualitative definition of robustness', *The annals of mathematical statistics* **42**(6), 1887–1896.

[21] Haussler, D. [1992], 'Decision theoretic generalizations of the pac model for neural net and other learning applications', *Information and computation* **100**(1), 78–150.

[22] Huang, Z., Mansour, Y. and Roughgarden, T. [2018], 'Making the most of your samples', *SIAM Journal on Computing* **47**(3), 651–674.

[23] Huber, P. J. [1992], Robust estimation of a location parameter, *in* 'Breakthroughs in statistics', Springer, pp. 492–518.

[24] Karnin, Z. S. and Anava, O. [2016], 'Multi-armed bandits: Competing with optimal sequences', *Advances in Neural Information Processing Systems* **29**.

[25] Kearns, M. J., Schapire, R. E. and Sellie, L. M. [1994], 'Toward efficient agnostic learning', *Machine Learning* **17**(2), 115–141.

[26] Kearns, M. and Li, M. [1993], 'Learning in the presence of malicious errors', *SIAM Journal on Computing* **22**(4), 807–837.

[27] Kim, S., Pasupathy, R. and Henderson, S. G. [2015], 'A guide to sample average approximation', *Handbook of simulation optimization* pp. 207–243.

[28] Kleywegt, A. J., Shapiro, A. and Homem-de Mello, T. [2002], 'The sample average approximation method for stochastic discrete optimization', *SIAM Journal on Optimization* **12**(2), 479–502.

[29] Klivans, A. R., Long, P. M. and Servedio, R. A. [2009], 'Learning halfspaces with malicious noise.', *Journal of Machine Learning Research* **10**(12).

[30] Lam, H. [2021], 'On the impossibility of statistically improving empirical optimization: A second-order stochastic dominance perspective', *arXiv preprint arXiv:2105.13419* .

[31] Levi, R., Perakis, G. and Uichanco, J. [2015], 'The data-driven newsvendor problem: new bounds and insights', *Operations Research* **63**(6), 1294–1306.

[32] Luo, H., Wei, C.-Y., Agarwal, A. and Langford, J. [2018], Efficient contextual bandits in non-stationary worlds, *in* 'Conference On Learning Theory', PMLR, pp. 1739–1776.

[33] Lykouris, T., Mirrokni, V. and Paes Leme, R. [2018], Stochastic bandits robust to adversarial corruptions, *in* 'Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing', pp. 114–122.

[34] Lykouris, T., Simchowitz, M., Slivkins, A. and Sun, W. [2021], Corruption-robust exploration in episodic reinforcement learning, *in* 'Conference on Learning Theory', PMLR, pp. 3242–3245.

[35] Mohajerin Esfahani, P. and Kuhn, D. [2018], 'Data-driven distributionally robust optimization using the wasserstein metric: Performance guarantees and tractable reformulations', *Mathematical Programming* **171**(1), 115–166.

[36] Morse, P. M. and Kimball, G. E. [1946], Methods of operations research, Technical report, Center for Naval Analyses, Alexandria VA, Operations Evaluation Group.

[37] Myerson, R. B. [1981], 'Optimal auction design', *Mathematics of operations research* **6**(1), 58–73.

[38] Perakis, G. and Roels, G. [2008], 'Regret in the newsvendor model with partial information', *Operations Research* **56**(1), 188–203.

[39] Purohit, M., Svitkina, Z. and Kumar, R. [2018], 'Improving online algorithms via ml predictions', *Advances in Neural Information Processing Systems* **31**.

[40] Riley, J. and Zeckhauser, R. [1983], 'Optimal selling strategies: When to haggle, when to hold firm', *The Quarterly Journal of Economics* **98**(2), 267–289.

[41] Scarf, H. [1958], 'A min-max solution of an inventory problem', *Studies in the mathematical theory of inventory and production* .

[42] Servedio, R. A. [2003], 'Smooth boosting and learning with malicious noise', *The Journal of Machine Learning Research* **4**, 633–648.

Checklist

1. For all authors...
   (a) Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope? [Yes] See Section 1 for pointers to each contribution.
   (b) Did you describe the limitations of your work? [Yes] See Section 6
   (c) Did you discuss any potential negative societal impacts of your work? [N/A]
   (d) Have you read the ethics review guidelines and ensured that your paper conforms to them? [Yes]
2. If you are including theoretical results...
   (a) Did you state the full set of assumptions of all theoretical results? [Yes]
   (b) Did you include complete proofs of all theoretical results? [Yes] All proofs are in the appendix.
3. If you ran experiments...
   (a) Did you include the code, data, and instructions needed to reproduce the main experimental results (either in the supplemental material or as a URL)? [N/A]
   (b) Did you specify all the training details (e.g., data splits, hyperparameters, how they were chosen)? [N/A]
   (c) Did you report error bars (e.g., with respect to the random seed after running experiments multiple times)? [N/A]
   (d) Did you include the total amount of compute and the type of resources used (e.g., type of GPUs, internal cluster, or cloud provider)? [N/A]
4. If you are using existing assets (e.g., code, data, models) or curating/releasing new assets...
   (a) If your work uses existing assets, did you cite the creators? [N/A]
   (b) Did you mention the license of the assets? [N/A]
   (c) Did you include any new assets either in the supplemental material or as a URL? [N/A]
   (d) Did you discuss whether and how consent was obtained from people whose data you're using/curating? [N/A]
   (e) Did you discuss whether the data you are using/curating contains personally identifiable information or offensive content? [N/A]
5. If you used crowdsourcing or conducted research with human subjects...
   (a) Did you include the full text of instructions given to participants and screenshots, if applicable? [N/A]
   (b) Did you describe any potential participant risks, with links to Institutional Review Board (IRB) approvals, if applicable? [N/A]
   (c) Did you include the estimated hourly wage paid to participants and the total amount spent on participant compensation? [N/A]