Data-Driven Sequential Search

David B. Brown
Fuqua School of Business, Duke University
dbbrown@duke.edu

Cagin Uru
Fuqua School of Business, Duke University
cagin.uru@duke.edu

Abstract

We consider a sequential search problem where the distribution of alternative values is unknown. In our data-driven setting, feasible policies are based solely on the history of explored alternative values. We seek to identify a policy that maximizes the worst-case ratio of expected reward compared to an oracle (referred to as *Pandora*) with full knowledge of the value distribution. We design static policies that commit to a prespecified number of explorations. We show that these policies guarantee a competitive ratio of at least $1/e \approx 37\%$ of the Pandora benchmark for any arbitrary value distribution. Our approach involves studying nature's problem to select a distribution to counter a policy and identifying worst-case distributions. Moreover, we study how the structure of the unknown value distribution influences achievable performance guarantees by considering a setting where feasible distributions belong to the class of monotone hazard rate distributions, where we improve our guarantee to $(e/(e+1))^2 \approx 53\%$ of the Pandora benchmark. We show that static policies are especially effective against smaller classes of unknown distributions, guaranteeing at least $e/(e+1) \approx 73\%$ against exponential distributions with an unknown rate and $9/(8(4-\sqrt{7})) \approx 83\%$ against uniform distributions with an unknown maximum. In the latter case, we show that static policies achieve the best possible performance among all feasible policies, including the dynamic ones. Finally, we derive performance limits for all feasible policies to further highlight the efficiency and robustness of our static policies for data-driven search problems.

1 Introduction

We consider a data-driven variant of the classical sequential search problem of Weitzman (1979), where alternative value distributions are unknown and the decision maker (DM) must rely on the history of alternative value realizations to design feasible search policies. Such a setting, where the DM searches in the dark, applies to many search problems where acquiring informative and reliable data is challenging (e.g., dynamic environments where market conditions change rapidly) or modeling accurate alternative distributions is error-prone (e.g., due to multiattribute valuations of alternatives).

We evaluate the performance of a feasible search policy by using the worst-case ratio over the set of feasible distributions of the expected reward of the policy with respect to the expected reward of the optimal policy with full access to the value distribution, referred to as *Pandora's rule*. We refer to the performance of Pandora's rule as the *Pandora benchmark* and the worst-case performance ratio over all feasible distributions as the *competitive ratio*. In this setting, the DM first chooses a feasible policy, then nature chooses a feasible distribution to counter the chosen policy. The problem is to identify a feasible policy that maximizes the competitive ratio, given the set of feasible distributions.

We consider two main settings. First, we assume that alternative values may follow any arbitrary distribution. Second, to capture the influence of the structure of the unknown value distribution on achievable performance guarantees and limits, we restrict feasible distributions to the *monotone*

39th Conference on Neural Information Processing Systems (NeurIPS 2025) Workshop: ML×OR Workshop: Mathematical Foundations and Operational Integration of Machine Learning for Uncertainty-Aware Decision-Making.

hazard rate (MHR) class. In both settings, we design static policies that commit to a predetermined number of explorations. Our central insight is that such policies are not only simple and easy to implement in practice, but they also achieve strong performance guarantees in data-driven settings with unknown value distributions. Because of their efficiency and robustness, our policies effectively mitigate errors and uncertainty, addressing the challenges of converting data into reliable decisions in sequential search contexts. Proofs and numerical examples are presented in Brown and Uru (2025).

2 Model

We consider a DM sequentially exploring alternatives in an attempt to maximize acquisition value. We assume that there are infinitely many alternatives, each with an a priori unknown value $\tilde{v}$, assumed nonnegative and representing the DM's valuation for the alternative. Alternative values are independent and follow an unknown distribution F. While the exact value distribution F is unknown, the DM knows that it belongs to a set of feasible distributions $\mathcal{F}$, which ranges from arbitrary distributions to distributions with unknown parameters in different settings we consider.

In each period, the DM can explore a new alternative and learn its value. We assume that explorations take time; therefore, the values of alternatives are discounted at a rate $\delta \in (0,1)$ during each exploration attempt. In our setting, the search cost is multiplicative and is represented by the discount factor δ. After an exploration attempt, the DM can explore an alternative with an unknown value or select any previously explored alternative. The search ends when the DM selects an alternative. The problem is to find a feasible search policy that maximizes the worst-case ratio of expected reward compared with the performance of Pandora's rule that fully knows the alternative value distribution.

We define feasible search policies as mappings from the observed data (i.e., the history of explored alternative values) to stopping decisions. That is, a feasible policy π is defined by a sequence of functions that map the sets of explored alternative values to stopping probabilities. Other than the realizations of alternative values, the only information accessible to feasible search policies are the discount factor δ and the set of feasible value distributions $\mathcal{F}$. We let Π denote the set of all feasible data-driven search policies and $V^{\pi}_{F,\delta}$ denote the expected reward of a policy $\pi \in \Pi$, given F and δ.

When the alternative value distribution is known, Pandora's rule maximizes the expected discounted value of the acquired alternative. We let $V^{*}_{F,\delta}$ denote the performance of Pandora's rule and refer to it as the Pandora benchmark. We can find $V^{*}_{F,\delta}$ by solving $V = \delta\mathbb{E}[\max\{V, \tilde{v}\}]$. We evaluate the performance of a feasible search policy $\pi \in \Pi$ by using the worst-case ratio over the set of feasible distributions $\mathcal{F}$ of the expected reward of the policy $V^{\pi}_{F,\delta}$ with respect to the Pandora benchmark $V^{*}_{F,\delta}$. Under this performance metric, our goal is to choose a feasible search policy $\pi \in \Pi$ to maximize the competitive ratio $\min_{F\in\mathcal{F}} \{V^{\pi}_{F,\delta}/V^{*}_{F,\delta}\}$. Hence, our problem is given by the *maximin ratio*

$$\max_{\pi\in\Pi} \min_{F\in\mathcal{F}} \frac{V^{\pi}_{F,\delta}}{V^{*}_{F,\delta}} .$$

3 Maximin Ratio for Arbitrary Distributions

We first consider arbitrary distributions, denoted by $\mathcal{F}_{\text{arb}}$, as the set of feasible value distributions. We design static policies that commit to a predetermined number of explorations. These policies sequentially explore a fixed, δ-dependent number of alternatives and select the highest-valued alternative after the explorations. We let π_n denote the static policy that explores $n \geq 1$ alternatives and Π_{STATIC} denote the collection of all static policies. We study the maximin ratio for arbitrary value distributions by focusing on static policies. For policy $\pi_n \in \Pi_{\text{STATIC}}$, we study nature's problem to choose an arbitrary distribution to minimize the performance with respect to the Pandora benchmark. We solve this problem and identify the worst-case distributions for static policies.

Proposition 1 *For $\delta \in (0,1)$ and $n \geq 1$, we have*

$$\min_{F\in\mathcal{F}_{arb}} \frac{V^{\pi_n}_{F,\delta}}{V^{*}_{F,\delta}} = \min_{F\in\mathcal{F}_{two\text{-}pt}} \frac{V^{\pi_n}_{F,\delta}}{V^{*}_{F,\delta}} = \min_{\sigma\in(0,1]} \frac{\delta^{n-1}(1-(1-\sigma)^n)(1-\delta(1-\sigma))}{\sigma} ,$$

where $\mathcal{F}_{two\text{-}pt}$ denotes the class of two-point distributions with support $\{0, x\}$ and respective probabilities $\{1-\sigma, \sigma\}$ for some $x > 0$ and $\sigma \in (0,1]$.

In the proof of Proposition 1, we show that we can rearrange the probability mass for a given arbitrary value distribution to obtain a perturbed distribution that achieves a lower performance ratio for static policies. In Proposition 2 in Brown and Uru (2025), we solve the minimization problem in Proposition 1 over $\sigma \in (0,1]$ and show that either $\sigma = 1$ or $\sigma \to 0$ minimizes the performance ratio. This highlights two key subclasses within the class $\mathcal{F}_{\text{two-pt}}$: the subclass where $\sigma = 1$, denoted by $\mathcal{F}_{\text{pt}}$, and the subclass where $\sigma \to 0$, denoted by $\mathcal{F}^{\text{zero}}_{\text{two-pt}}$. After solving nature's problem and identifying the worst-case distributions, we show that the static policy π_{arb} that explores $n_{\text{arb}} := \lfloor 1/(1-\delta) \rfloor$ many alternatives achieves strong performance guarantees for arbitrary value distributions.

Theorem 1 *For $\delta \in (0,1)$, we have*

$$\max_{\pi_n \in \Pi_{\mathit{STATIC}}} \min_{F \in \mathcal{F}_{\mathit{arb}}} \frac{V^{\pi_n}_{F,\delta}}{V^*_{F,\delta}} = \min_{F \in \mathcal{F}_{\mathit{arb}}} \frac{V^{\pi_{\mathit{arb}}}_{F,\delta}}{V^*_{F,\delta}} = \delta^{n_{\mathit{arb}}-1}(1-\delta)n_{\mathit{arb}} \geq \delta^{\frac{\delta}{1-\delta}} \geq \frac{1}{e}\,.$$

Following Theorem 1, we consider feasible search policies that potentially randomize between policies in Π_{STATIC}. In Proposition 3 in Brown and Uru (2025), we show that the best randomized policy explores n_{arb} many alternatives with probability one, coinciding with the best non-randomized policy, and hence offering no improvement in performance guarantees for arbitrary value distributions.

We complement our performance guarantee results by developing performance limits for feasible search policies. We consider a setting in which the unknown alternative value distribution belongs to either $\mathcal{F}_{\text{pt}}$ or $\mathcal{F}^{\text{zero}}_{\text{two-pt}}$. In this setting, there is a misalignment in search incentives. At a high-level, the distributions in $\mathcal{F}_{\text{pt}}$ favor a shorter search process, whereas those in $\mathcal{F}^{\text{zero}}_{\text{two-pt}}$ favor a longer one. This difference in search incentives drives a performance limit for *any* feasible search policy, including the policies outside of Π_{STATIC}. Focusing on this setting, we derive an upper bound for the maximin ratio.

Proposition 2 *For $\delta \in (0,1)$, no feasible policy can achieve more than $1/(1+\delta)$ of the Pandora benchmark when F is arbitrary. In particular, no feasible policy can achieve more than 50% of the Pandora benchmark for all $\delta \in (0,1)$ when F is arbitrary.*

4 Maximin Ratio for Monotone Hazard Rate Distributions

We consider the class of MHR distributions, denoted by $\mathcal{F}_{\text{mhr}}$, which includes many well-known distributions, e.g., uniform and exponential. For MHR distributions, identifying classes of worst-case distributions for static policies that minimize the performance ratio $V^{\pi_n}_{F,\delta}/V^*_{F,\delta}$ remains challenging. Therefore, we consider an approach based on a decomposition of the performance ratio. We show that the class of exponential distributions with an unknown rate, denoted by $\mathcal{F}_{\text{exp}}$, and the class of point distributions with an unknown mass point, denoted by $\mathcal{F}_{\text{pt}}$, separately minimize the two parts of the decomposed performance ratio. At a high-level, these two classes of distributions lie at opposite ends of the MHR class for the search problem, e.g., $\mathcal{F}_{\text{exp}}$ provides the most incentive to explore, whereas $\mathcal{F}_{\text{pt}}$ provides the least. We show that the static policy π_{mhr} that explores $n_{\text{mhr}} := \lceil \log(e/(e+1))/\log(\delta) \rceil$ many alternatives achieves strong performance guarantees for MHR distributions.

Theorem 2 *Let W denote the Lambert function defined as the inverse of $x \mapsto xe^x$. For $\delta \in (0,1)$, we have*

$$\min_{F \in \mathcal{F}_{\mathit{mhr}}} \frac{V^{\pi_{\mathit{mhr}}}_{F,\delta}}{V^*_{F,\delta}} \geq \delta^{2n_{\mathit{mhr}}-1} \left(W\left(\frac{\delta^{n_{\mathit{mhr}}}}{1-\delta^{n_{\mathit{mhr}}}} \right) \right)^{-1} \geq \left(\frac{e}{e+1} \right)^2 \approx 53\%\,.$$

After studying the performance of static policies for several common MHR distributions, we conjecture that the distributions in $\mathcal{F}_{\text{exp}}$ and $\mathcal{F}_{\text{pt}}$ minimize not only the decomposed performance ratio but also the overall performance ratio $V^{\pi_n}_{F,\delta}/V^*_{F,\delta}$ for static policies, thereby achieving the worst-case performance among MHR distributions à la Proposition 1.

Conjecture 1 *For $\delta \in (0,1)$ and $n \geq 1$, we have* $\min_{F \in \mathcal{F}_{\mathit{mhr}}} \frac{V^{\pi_n}_{F,\delta}}{V^*_{F,\delta}} = \min_{F \in \mathcal{F}_{\mathit{exp}} \cup \mathcal{F}_{\mathit{pt}}} \frac{V^{\pi_n}_{F,\delta}}{V^*_{F,\delta}}\,.$

Exponential and Point Distributions. In light of Conjecture 1, we study this subcase for the MHR class. We consider the static policy $\pi_{\text{exp-pt}}$ that explores $n \geq 1$ many alternatives for $\delta \in (\delta_{n-1}, \delta_n]$, where $\delta_0 := 0$, $\delta_n := H_n/(H_n + \exp(-H_n))$, and $H_n := \sum_{k=1}^n 1/k$ denotes the n^{th} harmonic number. We let $n_{\text{exp-pt}}$ denote the number of alternatives the policy $\pi_{\text{exp-pt}}$ explores, given $\delta \in (0,1)$.

Proposition 3 *Let $\mathcal{F}_{\textit{exp-pt}} := \mathcal{F}_{\textit{exp}} \cup \mathcal{F}_{\textit{pt}}$. For $\delta \in (0,1)$, we have*

$$\max_{\pi_n \in \Pi_{\textit{STATIC}}} \min_{F \in \mathcal{F}_{\textit{exp-pt}}} \frac{V^{\pi_n}_{F,\delta}}{V^*_{F,\delta}} = \min_{F \in \mathcal{F}_{\textit{exp-pt}}} \frac{V^{\pi_{\textit{exp-pt}}}_{F,\delta}}{V^*_{F,\delta}} = \min \left\{ \underbrace{\frac{H_{n_{\textit{exp-pt}}} \delta^{n_{\textit{exp-pt}}}}{W\left(\frac{\delta}{1-\delta}\right)}}_{\mathcal{F}_{\textit{exp}}}, \underbrace{\delta^{n_{\textit{exp-pt}}-1}}_{\mathcal{F}_{\textit{pt}}} \right\} \geq \frac{e}{e+1} \approx 73\% \,.$$

In particular, the policy $\pi_{\textit{exp-pt}}$ achieves the same performance as Pandora's rule as $\delta \to 1$.

We consider uniform distributions (another subcase of the MHR setting) and derive surprising results.

Uniform Distributions. We assume that the value distribution belongs to the class of uniform distributions with support $(0, u)$, where $u > 0$ is unknown, denoted by $\mathcal{F}_{\text{unif}}$. We analyze the performance of static policies and show that the static policy π_{unif} that explores $n_{\text{unif}} := \lfloor \sqrt{1/(1-\delta)} \rfloor$ many alternatives achieves strong performance guarantees for uniform value distributions in $\mathcal{F}_{\text{unif}}$.

Proposition 4 *Let $\kappa_\delta := \delta/(1 - \sqrt{1-\delta^2})$. For $\delta \in (0,1)$, we have*

$$\max_{\pi_n \in \Pi_{\textit{STATIC}}} \min_{F \in \mathcal{F}_{\textit{unif}}} \frac{V^{\pi_n}_{F,\delta}}{V^*_{F,\delta}} = \min_{F \in \mathcal{F}_{\textit{unif}}} \frac{V^{\pi_{\textit{unif}}}_{F,\delta}}{V^*_{F,\delta}} = \frac{\kappa_\delta \delta^{n_{\textit{unif}}} n_{\textit{unif}}}{n_{\textit{unif}}+1} \geq \frac{9}{8(4-\sqrt{7})} \approx 83.1\% \,.$$

In particular, the policy $\pi_{\textit{unif}}$ achieves the same performance as Pandora's rule as $\delta \to 1$.

We next introduce a general approach for deriving performance limits for feasible policies and use this approach to derive upper bounds on the maximin ratio for MHR distributions. At a high-level, our approach draws inspiration from Lagrangian duality and is centered around switching from worst-case performance objective to a measure of "average-case" performance objective. Specifically, we rewrite our maximin ratio problem as a linear program (LP) as

$$\max_{\pi \in \Pi, z} \quad z \qquad \text{subject to} \quad z \leq \frac{V^{\pi}_{F,\delta}}{V^*_{F,\delta}} \quad \forall F \in \mathcal{F} \,.$$

Since it is difficult to solve this infinite-dimensional LP, we relax the constraints and introduce the Lagrange multiplier function μ. In the relaxed problem, μ assigns a weight $\mu(F)$ to each $F \in \mathcal{F}$ and the objective becomes the average performance over the set of feasible distributions $\mathcal{F}$. We refer to any measure μ that satisfies $\mu(F) \geq 0$ for $F \in \mathcal{F}$ and $\int_{F\in\mathcal{F}} \mu(F) d\mu = 1$ as *dual feasible*.

Proposition 5 *For any dual feasible μ, we have*

$$V^\mu := \max_{\pi \in \Pi} \int_{F \in \mathcal{F}} \mu(F) \frac{V^{\pi}_{F,\delta}}{V^*_{F,\delta}} d\mu \geq \max_{\pi \in \Pi} \min_{F \in \mathcal{F}} \frac{V^{\pi}_{F,\delta}}{V^*_{F,\delta}} \,.$$

The relaxed problem has a Bayesian interpretation. As μ assigns a weight $\mu(F) \geq 0$ for each $F \in \mathcal{F}$, it provides a prior over the set of feasible value distributions, which leads to a Bayesian search problem. We let $\mathcal{F} = \mathcal{F}_{\text{unif}}$ and consider a Bayesian search problem where we let μ assign Pareto probabilities to potential values of the unknown maximum u in $\mathcal{F}_{\text{unif}}$. Formulating it as a stochastic dynamic program, we solve the Bayesian search problem and show that static policies are optimal. Combined with Proposition 4, it follows that static policies solve the maximin ratio when $F \in \mathcal{F}_{\text{unif}}$.

Theorem 3 *For $\delta \in (0,1)$, we have*

$$\max_{\pi \in \Pi} \min_{F \in \mathcal{F}_{\textit{unif}}} \frac{V^{\pi}_{F,\delta}}{V^*_{F,\delta}} = \max_{\pi_n \in \Pi_{\textit{STATIC}}} \min_{F \in \mathcal{F}_{\textit{unif}}} \frac{V^{\pi_n}_{F,\delta}}{V^*_{F,\delta}} = \min_{F \in \mathcal{F}_{\textit{unif}}} \frac{V^{\pi_{\textit{unif}}}_{F,\delta}}{V^*_{F,\delta}} = \frac{\kappa_\delta \delta^{n_{\textit{unif}}} n_{\textit{unif}}}{n_{\textit{unif}}+1} \,.$$

That is, the static policy $\pi_{\textit{unif}}$ is maximin optimal for uniform distributions with an unknown maximum.

Theorem 3 reveals that static policies in Π_{STATIC} achieve the best possible performance among *all* feasible data-driven search policies, including the dynamic ones, when $F \in \mathcal{F}_{\text{unif}}$. Since $\mathcal{F}_{\text{unif}} \subset \mathcal{F}_{\text{mhr}}$, Theorem 3 also provides an upper bound on the maximin ratio for MHR distributions.

Corollary 1 *For $\delta \in (0,1)$, no feasible policy can achieve more than $\kappa_\delta n_{\textit{unif}} \delta^{n_{\textit{unif}}}/(n_{\textit{unif}}+1)$ of the Pandora benchmark when $F \in \mathcal{F}_{\textit{mhr}}$. In particular, no feasible policy can achieve more than $9/(8(4-\sqrt{7})) \approx 83.1\%$ of the Pandora benchmark for all $\delta \in (0,1)$ when $F \in \mathcal{F}_{\textit{mhr}}$.*

References

Brown, D. B. and Uru, C. (2025), 'Data-driven sequential search', *Under Review, Available at SSRN 5348531.*

Weitzman, M. (1979), 'Optimal search for the best alternative', *Econometrica* **47**(3), 641–54.