Risk-Sensitive Online Algorithms

Nicolas Christianson
Caltech
nchristi@caltech.edu

Bo Sun
University of Waterloo
bo.sun@uwaterloo.ca

Steven Low
Caltech
slow@caltech.edu

Adam Wierman
Caltech
adamw@caltech.edu

ABSTRACT

We study the design of *risk-sensitive online algorithms*, in which risk measures are used in the competitive analysis of randomized online algorithms. We introduce the CVaR_δ-competitive ratio (δ-CR) using the conditional value-at-risk of an algorithm's cost, which measures the expectation of the $(1-\delta)$-fraction of worst outcomes against the offline optimal cost, and use this measure to study three online optimization problems: continuous-time ski rental, discrete-time ski rental, and one-max search. The structure of the optimal δ-CR and algorithm varies significantly between problems: we prove that the optimal δ-CR for continuous-time ski rental is $2 - 2^{-\Theta(\frac{1}{1-\delta})}$, obtained by an algorithm described by a delay differential equation. In contrast, in discrete-time ski rental with buying cost B, there is an abrupt phase transition at $\delta = 1 - \Theta(\frac{1}{\log B})$, after which the classic deterministic strategy is optimal. Similarly, one-max search exhibits a phase transition at $\delta = \frac{1}{2}$, after which the classic deterministic strategy is optimal; we also obtain an algorithm that is asymptotically optimal as $\delta \downarrow 0$ that arises as the solution to a delay differential equation.

1 Introduction

Randomness can improve decision-making performance in many online problems; for instance, randomization improves the competitive ratio of online ski rental from 2 to $\frac{e}{e-1}$ [10] and of online search from polynomial to logarithmic in the fluctuation ratio [6]. However, this improved performance can only be obtained on average over multiple problem instances, as a randomized algorithm can vary wildly in its performance on any particular run, which may be undesirable if an agent is sensitive to risks of a particular size or likelihood. Fields such as economics and finance have fielded research on risk aversion and alternative *risk measures* that enable modifying decision-making objectives to accommodate these risk preferences (e.g., [11, 2]). One of the most well-studied risk measures in recent years, due to its nice properties (as a *coherent* risk measure) and computational tractability, is the *conditional value-at-risk* (CVaR_δ), which measures the expectation of a random loss/reward on its $(1-\delta)$-fraction of worst outcomes [12]. CVaR_δ and other risk measures have seen wide application across domains and have been studied as an objective in place of the expectation in various online learning problems (e.g., [7, 3, 13]).

Despite the significant extent of literature on risk-sensitive algorithms for online learning, there has been no work on the design and analysis of *competitive* algorithms for online optimization problems like ski rental, online search, knapsack, function chasing, or metrical task systems with risk-sensitive objectives; the closest related work is a recent paper that studies ski rental with risk *constraints* [5]. Thus we ask: how can we design competitive online algorithms when we care about the CVaR_δ of the cost/reward, and what are the optimal competitive ratios for different problems?

In this work, we begin to work toward answering this question, studying risk sensitivity in competitive online algorithms for online optimization. In particular, we focus on two of the prototypical problems in online optimization: *ski rental*, which encapsulates the fundamental "rent vs. buy" tradeoff inherent in online optimization with switching costs, and *one-max search*, which exhibits a complementary "accept vs. wait" tradeoff fundamental to constrained online optimization. While these problems are both simple to pose, they have applications to more complex problems including TCP acknowledgement and buffering problems [9, 8], and they reflect crucial components of the difficulty of more complicated online optimization problems. They thus serve as ideal analytic testbeds for investigating the design of risk-sensitive algorithms in online optimization.

In this extended abstract, we give an overview of our main results on risk-sensitive algorithms for continuous- and discrete-time ski rental and one-max search, discussing algorithms, optimality, and phase transitions exhibited by the problems. We refer the reader to the full version of the paper for further details and proofs [4].

2 Preliminaries

In this section, we introduce the conditional value-at-risk and the three online problems studied in this work.

2.1 The Conditional Value-at-Risk

A *risk measure* is a mapping from the set of $\mathbb{R}$-valued random variables to $\mathbb{R}$ that gives a deterministic valuation of the *risk* associated with a particular random loss. One of the most well-studied risk measures in recent years is the *conditional value-at-risk* (CVaR), defined as the expectation of a random variable X on its $(1-\delta)$-fraction of worst outcomes. It is defined formally as follows.

Definition 2.1 (Conditional Value-at-Risk). *Let X be a real-valued random loss variable with inverse CDF F_X^{-1}.*

Its conditional value-at-risk $\mathrm{CVaR}_\delta[X]$ *is defined as [1]:*

$$\mathrm{CVaR}_\delta[X] = \frac{1}{1-\delta}\int_\delta^1 F_X^{-1}(p)\,\mathrm{d}p.$$

This definition, which is one of several ways of expressing CVaR_δ, highlights the intuition that $\mathrm{CVaR}_\delta[X]$ computes the expected loss of X on the largest $(1-\delta)$-fraction of outcomes in its distribution. For random rewards, CVaR_δ is defined similarly as the expectation on the smallest, rather than largest, $(1-\delta)$-fraction of outcomes. Note that $\mathrm{CVaR}_0[X] = \mathbb{E}[X]$ and $\mathrm{CVaR}_1[X] := \lim_{\delta\uparrow 1}\mathrm{CVaR}_\delta[X] = \operatorname{ess\,sup} X$.

2.2 Competitive Analysis

In the study of online algorithms, algorithm performance is typically measured via the *competitive ratio*, or the worst case ratio in (expected) cost between an algorithm and the offline optimal strategy that knows all uncertainty in advance. In this work, we introduce a modified version of the competitive ratio for randomized algorithms that goes beyond expected performance: instead, we penalize a randomized algorithm via the ratio between the conditional value-at-risk of its cost and the offline optimal algorithm's cost, terming this metric the CVaR_δ*-competitive ratio* (δ-CR).

Definition 2.2 ($\mathbf{CVaR}_\delta$-Competitive Ratio). *Consider an online problem with uncertainty drawn adversarially from a set of instances $\mathcal{I}$. Let* ALG *be a randomized algorithm, and let* OPT *be the offline optimal algorithm. The* **$\mathbf{CVaR}_\delta$-*Competitive Ratio* (δ-CR)** *is defined as the worst-case ratio between the* CVaR_δ *of* ALG*'s cost and the offline optimal cost:*

$$\delta\text{-CR}(\textsc{Alg}) := \sup_{I\in\mathcal{I}} \frac{\mathrm{CVaR}_\delta[\textsc{Alg}(I)]}{\textsc{Opt}(I)},$$

where the CVaR_δ *is taken over* ALG*'s randomness.*

Note that the δ-CR can be viewed as an interpolation between the classic randomized case where the adversary has no power over ALG's randomness and ALG pays its expected cost ($\delta = 0$), and the case where the adversary has full control over ALG's randomness, so ALG pays the worst cost in its support and determinism is optimal ($\delta = 1$).

2.3 Online Problems Studied

We now provide brief descriptions of the three problems studied in this work.

Continuous-Time Ski Rental. In the *continuous-time ski rental (CSR)* problem, a player faces a ski season of unknown and adversarially-chosen duration $s \in \mathbb{R}_{++}$, and must choose how long to rent skis before purchasing them. In particular, the player pays cost equal to the duration of renting, and cost 1 (without loss of generality) for purchasing the skis. Deterministic algorithms for ski rental are wholly determined by the day $x \in \mathbb{R}_{++}$ on which the player stops renting and purchases the skis: an algorithm that rents until day x and then purchases pays cost $s\cdot\mathbb{1}_{x>s} + (x+1)\cdot\mathbb{1}_{x\le s}$. Randomized algorithms can be described by a random variable X over purchase days, in which case the algorithm pays (random) cost $s\cdot\mathbb{1}_{X>s} + (X+1)\cdot\mathbb{1}_{X\le s}$. Given knowledge of the total number of skiing days s, the offline optimal strategy is to rent for the entire season if $s < 1$, incurring cost s, and to buy immediately otherwise, yielding cost 1. Defining $\alpha_\delta^{\mathrm{CSR},\mu}$ as the δ-CR of a strategy $X \sim \mu$, we have

$$\alpha_\delta^{\mathrm{CSR},\mu} := \sup_{s\in\mathbb{R}_{++}} \frac{\mathrm{CVaR}_\delta[s\cdot\mathbb{1}_{X>s} + (X+1)\cdot\mathbb{1}_{X\le s}]}{\min\{s,1\}}.$$

We denote by $\alpha_\delta^{\mathrm{CSR},*}$ the smallest δ-CR of any strategy. It is well known that $\alpha_1^{\mathrm{CSR},*} = 2$, which is achieved by purchasing skis deterministically at time 1, and $\alpha_0^{\mathrm{CSR},*} = \frac{e}{e-1}$ [10].

Discrete-Time Ski Rental. In the *discrete-time ski rental (DSR)* problem, a player faces a ski season of unknown and adversarially-chosen duration $s \in \mathbb{N}$ and must choose an integer number of days to rent skis before purchasing them; renting for a day costs 1, and purchasing skis has an integer cost $B \ge 2$. The cost structure is essentially identical to the continuous-time case, except the algorithm's and adversary's decisions are restricted to lie in $\mathbb{N}$: if a player buys skis at the start of day $x \in \mathbb{N}$ and the true season duration is $s \in \mathbb{N}$, their cost will be $s\cdot\mathbb{1}_{x>s} + (B+x-1)\cdot\mathbb{1}_{x\le s}$. Thus for a random strategy $X \sim \mu$ with support on $\mathbb{N}$, the δ-CR is defined as

$$\alpha_\delta^{\mathrm{DSR}(B),\mu} := \sup_{s\in\mathbb{N}} \frac{\mathrm{CVaR}_\delta[s\cdot\mathbb{1}_{X>s} + (B+X-1)\cdot\mathbb{1}_{X\le s}]}{\min\{s,B\}}.$$

As in the continuous-time setting, we denote by $\alpha_\delta^{\mathrm{DSR}(B),*}$ the smallest δ-CR of any strategy. It is well known that $\alpha_1^{\mathrm{DSR}(B),*} = 2 - \frac{1}{B}$, achieved by deterministically purchasing skis at the start of day B, and $\alpha_0^{\mathrm{DSR}(B),*} = \frac{1}{1-\left(1-B^{-1}\right)^B}$, which approaches $\alpha_0^{\mathrm{CSR},*} = \frac{e}{e-1}$ as $B \to \infty$.

One-Max Search. In the *one-max search (OMS)* problem, a player faces a sequence of prices $v_t \in [L,U]$ arriving online, with $U \ge L > 0$ known upper and lower bounds on the price sequence; the *fluctuation ratio* $\theta = \frac{U}{L}$ is defined as the ratio between these bounds. The player seeks to sell an indivisible item for the greatest possible price; after observing a price v_t, they can choose to either accept the price and earn profit v_t, or to wait and observe the next price. The duration $T \in \mathbb{N}$ of the sequence is *a priori* unknown to the player; if T elapses and the player has not yet sold the item, they sell it for the smallest possible price L in a compulsory trade. An algorithm that sells the item at the first price $\ge x$ will earn profit, in the worst case, $L\cdot\mathbb{1}_{x>v} + x\cdot\mathbb{1}_{x\le v}$ when the true maximum price in the sequence is v. Thus the δ-CR of a "random threshold" algorithm that accepts the first price meeting or exceeding some random value $X \sim \mu$ is defined:

$$\alpha_\delta^{\mathrm{OMS}(\theta),\mu} := \sup_{v\in[L,U]} \frac{v}{\mathrm{CVaR}_\delta[L\cdot\mathbb{1}_{X>v} + X\cdot\mathbb{1}_{X\le v}]};$$

Note that in this definition, we take the ratio between the offline optimal cost and the CVaR_δ of the algorithm's cost, since we are maximizing profit rather than minimizing cost. Similarly, the CVaR_δ definition employed here is the reward form; see our full paper [4] for full details. We denote by $\alpha_\delta^{\mathrm{OMS}(\theta),*}$ the optimal δ-CR for the problem; it is known that $\alpha_1^{\mathrm{OMS}(\theta),*} = \sqrt{\theta}$ and $\alpha_0^{\mathrm{OMS}(\theta),*} = 1 + W_0\left(\frac{\theta-1}{e}\right)$, where W_0 is the principal branch of the Lambert W function [6].

3 Results

In this section, we describe our algorithmic upper bounds and lower bounds for each of the three problems studied.

Continuous-Time Ski Rental. For the CSR problem, the optimal algorithm arises as the solution to a delay differential equation.

Theorem 3.1. *For any $\delta \in [0,1)$, let $\phi : [0,1] \to [0,1]$ be the solution to the following delay differential equation:*

$$\phi'(t) = \frac{1}{\alpha(1-\delta)}\left[\phi(t) - \phi(t-(1-\delta))\right] \quad \textit{for } t \in [1-\delta, 1],$$

with initial condition $\phi(t) = \log\left(1 + \frac{t}{(\alpha-1)(1-\delta)}\right)$ on $t \in [0, 1-\delta]$. Then when $\alpha = \alpha_\delta^{CSR,*}$, ϕ is the inverse CDF of the unique optimal strategy for continuous-time ski rental with the δ-CR metric. Moreover,

$$\alpha_\delta^{CSR,*} = 2 - 2^{-\Theta(\frac{1}{1-\delta})} \qquad \textit{as } \delta \uparrow 1.$$

Proving this result requires the insight that the inverse CDF is more tractably analyzed than the PDF when considering the δ-CR, and in addition depends on a set of structural results characterizing the optimal algorithm's indifference to the adversary's decision.

Discrete-Time Ski Rental. The DSR problem exhibits a remarkably different structure than the CSR problem: in particular, there is a phase transition at $\delta = 1 - \Theta(\frac{1}{\log B})$ marking a transition from a regime where randomness improves performance to one in which determinism is optimal.

Theorem 3.2. *Let $\alpha_\delta^{DSR(B),*}$ be the optimal δ-CR for discrete-time ski rental with buying cost $B \in \mathbb{N}$. Then $\alpha_\delta^{DSR(B),*}$ exhibits a* ***phase transition*** *at $\delta = 1 - \Theta(\frac{1}{\log B})$, whereby before this transition, $\alpha_\delta^{DSR(B),*}$ strictly improves on the deterministic optimal δ-CR of $2 - \frac{1}{B}$, whereas after this transition, $\alpha_\delta^{DSR(B),*} = 2 - \frac{1}{B}$.*

Moreover, in the discrete setting, the optimal algorithm for $\delta = 0$ remains optimal for sufficiently small δ.

Theorem 3.3. *Suppose $\delta = \mathcal{O}(\frac{1}{B})$. Then the optimal δ-CR $\alpha_\delta^{DSR(B),*}$ and strategy $\mathbf{p}^{B,\delta,*}$ for discrete-time ski rental with buying cost B are*

$$\alpha_\delta^{B,*} = \frac{C-\delta}{1-\delta} \qquad \textit{and} \qquad p_i^{B,\delta,*} = \frac{C}{B}\left(1 - \frac{1}{B}\right)^{B-i}$$

for all $i \in [B]$, where $C = \frac{1}{1-(1-1/B)^B}$ is the optimal competitive ratio for the $\delta = 0$ case. In particular, $\mathbf{p}^{B,\delta,*}$ is identical to the optimal algorithm for the $\delta = 0$ setting.

One-Max Search Finally, we show that OMS admits an algorithm that arises as the solution to a delay differential equation, just like CSR. Like DSR, though, OMS also exhibits a phase transition at $\delta = \frac{1}{2}$, after which randomness does not help and determinism is optimal.

Theorem 3.4. *Let $\delta \in [0,1]$, and let $\phi : [0,1] \to [L,U]$ be the solution to the following delay differential equation:*

$$\phi'(t) = \frac{\alpha}{1-\delta}\left[\phi(t-\delta) - L\right] \qquad \textit{for } t \in [\delta, 1],$$

with initial condition $\phi(t) = \alpha L$ on $t \in [0,\delta]$, where α is chosen such that $\phi(1) = U$ when $\delta < 1$, and $\alpha := \sqrt{\theta}$ when $\delta = 1$. Then ϕ is the inverse CDF of a random threshold algorithm for one-max search with δ-CR α. In particular, α is upper bounded as:

$$\alpha \le \begin{cases} 1 + W_0\left(\frac{\theta-1}{e}\right) + \mathcal{O}(\delta) & \textit{as } \delta \downarrow 0 \\ \sqrt{\theta} & \textit{when } \delta > \frac{1}{5}. \end{cases}$$

Moreover, there is an identical lower bound of $\alpha_\delta^{OMS(\theta),*} \ge \sqrt{\theta} = \alpha_1^{OMS(\theta),*}$ when $\delta \ge \frac{1}{2}$ and an asymptotically identical lower bound of $\alpha_\delta^{OMS(\theta),*} = 1 + W_0\left(\frac{\theta-1}{e}\right) + \Omega(\delta)$ as $\delta \downarrow 0$.

4 Acknowledgments

The authors acknowledge support from an NSF Graduate Research Fellowship (DGE-2139433), NSF Grants CNS-2146814, CPS-2136197, CNS-2106403, and NGSDI-2105648, the Resnick Sustainability Institute, and a Caltech S2I Grant.

References

[1] C. Acerbi and D. Tasche. On the coherence of expected shortfall. *Journal of Banking & Finance*, 26(7):1487–1503, July 2002.

[2] P. Artzner, F. Delbaen, J.-M. Eber, and D. Heath. Coherent Measures of Risk. *Mathematical Finance*, 9(3):203–228, July 1999.

[3] Y. Chow and M. Ghavamzadeh. Algorithms for CVaR Optimization in MDPs. In *Advances in Neural Information Processing Systems*, volume 27. Curran Associates, Inc., 2014.

[4] N. Christianson, B. Sun, S. Low, and A. Wierman. Risk-Sensitive Online Algorithms, 2024. arXiv:2405.09859.

[5] M. Dinitz, S. Im, T. Lavastida, B. Moseley, and S. Vassilvitskii. Controlling Tail Risk in Online Ski-Rental. In *Proceedings of the 2024 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 4247–4263, 2024.

[6] R. El-Yaniv, A. Fiat, R. M. Karp, and G. Turpin. Optimal Search and One-Way Trading Online Algorithms. *Algorithmica*, 30(1):101–139, May 2001.

[7] E. Even-Dar, M. Kearns, and J. Wortman. Risk-Sensitive Online Learning. In J. L. Balcázar, P. M. Long, and F. Stephan, editors, *Algorithmic Learning Theory*, pages 199–213, Berlin, Heidelberg, 2006. Springer.

[8] S. Geulen, B. Vocking, and M. Winkler. Regret Minimization for Online Buffering Problems Using the Weighted Majority Algorithm. In *23rd Annual Conference on Learning Theory*, 2010.

[9] A. R. Karlin, C. Kenyon, and D. Randall. Dynamic TCP acknowledgement and other stories about e/(e-1). In *Proceedings of the Thirty-Third Annual ACM Symposium on Theory of Computing*, STOC '01, pages 502–509, New York, NY, USA, July 2001. Association for Computing Machinery.

[10] A. R. Karlin, M. S. Manasse, L. A. McGeoch, and S. Owicki. Competitive randomized algorithms for nonuniform problems. *Algorithmica*, 11(6):542–571, June 1994.

[11] H. M. Markowitz. *Portfolio Selection: Efficient Diversification of Investments.* Yale University Press, 1959.

[12] R. T. Rockafellar and S. Uryasev. Conditional value-at-risk for general loss distributions. *Journal of Banking & Finance*, 26(7):1443–1471, July 2002.

[13] A. Tamkin, C. Dann, R. Keramati, and E. Brunskill. Distributionally-Aware Exploration for CVaR Bandits. In *NeurIPS 2019 Workshop on Safety and Robustness on Decision Making*, 2019.