Concentration bounds for CVaR estimation: The cases of light-tailed and heavy-tailed distributions

Prashanth L.A. [1] **Krishna Jagannathan** [2] **Ravi Kumar Kolla** [3]

Abstract

Conditional Value-at-Risk (CVaR) is a widely used risk metric in applications such as finance. We derive concentration bounds for CVaR estimates, considering separately the cases of sub-Gaussian, light-tailed and heavy-tailed distributions. For the sub-Gaussian and light-tailed cases, we use a classical CVaR estimator based on the empirical distribution constructed from the samples. For heavy-tailed random variables, we assume a mild 'bounded moment' condition, and derive a concentration bound for a truncation-based estimator. Our concentration bounds exhibit exponential decay in the sample size, and are tighter than those available in the literature for the above distribution classes. To demonstrate the applicability of our concentration results, we consider the CVaR optimization problem in a multi-armed bandit setting. Specifically, we address the best CVaR-arm identification problem under a fixed budget. Using our CVaR concentration results, we derive an upper-bound on the probability of incorrect arm identification.

1. Introduction

In applications such as portfolio optimization in finance, the quality of a portfolio is not satisfactorily captured by the expected value of return. Indeed, in such applications, a more risk-sensitive metric is desirable, so as to capture typical losses in the case of adverse events. Value-at-Risk (VaR) and Conditional-Value-at-Risk (CVaR) are two risk-aware metrics, which are widely used in applications such as portfolio optimization and insurance. VaR at level $\alpha \in (0,1)$ conveys the maximum loss incurred by the portfolio with a confidence of α. In other words, the portfolio incurs a loss greater than VaR at level α with probability $1-\alpha$. In turn, CVaR at level $\alpha \in (0,1)$ captures the expected loss incurred by the portfolio, *given* that the losses exceed VaR at level α. CVaR has an advantage over VaR, in that the former is a *coherent*[1] risk measure (Artzner et al., 1999).

1.1. Our Contributions

In this paper, we derive concentration bounds for CVaR estimators, for the cases of sub-Gaussian, light-tailed and heavy-tailed random variables. For sub-Gaussian and light-tailed distributions, our concentration bound uses a classical CVaR estimator based on the empirical distribution. For the heavy-tailed case, we employ a truncation-based CVaR estimator, and derive a concentration result under a mild assumption: the pth moment of the distribution is assumed to exist, for some $p > 1$. Notably, our concentration bounds enjoy an exponential decay in the sample size, for heavy-tailed as well as light-tailed distributions. Our results also subsume or strengthen several existing CVaR concentration results, as we discuss later. We believe our bounds are order optimal, and the dependence on the number of samples as well as the accuracy cannot be improved.

In order to highlight an important application for our CVaR concentration results, we consider a stochastic bandit setup with a risk-sensitive metric for measuring the quality of an arm. In particular, we consider a K-armed stochastic bandit setting, and study the problem of finding the arm with the *lowest CVaR value* in a fixed budget setting. We propose an algorithm for the best CVaR arm identification that is inspired by successive-rejects (SR) (Audibert et al., 2010). Using our CVaR concentration results, we establish an upper bound on the probability of incorrect arm identification for the SR-based algorithm for CVaR.

[1]Department of Computer Science and Engineering, Indian Institute of Technology Madras. [2]Department of Electrical Engineering, Indian Institute of Technology Madras. [3]ABInBev, Bangalore. Correspondence to: Prashanth L.A. <prashla@cse.iitm.ac.in>, Krishna Jagannathan <krishnaj@ee.iitm.ac.in>, Ravi Kumar Kolla <kolla.422@gmail.com>.

Proceedings of the 37^{th} International Conference on Machine Learning, Online, PMLR 119, 2020.

[1]A risk measure is said to be coherent, if it is monotonic, translation invariant, sub-additive, and positive homogeneous.

1.2. Related Work

For the case of bounded distributions, a popular CVaR estimate has been shown to exponentially concentrate around the true CVaR — see (Brown, 2007; Wang & Gao, 2010). In comparison to CVaR, obtaining a concentration result for VaR is easier, and does not require assumptions on the tail of the distribution – see (Kolla et al., 2019), a paper which also derives a one-sided CVaR concentration bound. More recent work (Thomas & Learned-Miller, 2019) considers CVaR concentration for distributions with bounded support on one side, but the form of their result (which uses order statistics) is not directly useful for important applications such as multi-armed bandits. In another recent paper (Bhat & Prashanth, 2019), the authors derive an exponentially decaying concentration bound for the case of sub-Gaussian distributions, using a concentration result (Fournier & Guillin, 2015) for the Wasserstein distance between the empirical and the true distributions. However, the above approach leads to poor concentration bounds (with power law decay in the sample size) for other relevant distribution classes, such as light-tailed and bounded-moment distributions.

While bandit learning has a long history, dating back to (Thompson, 1933), risk-based criteria have been considered only recently. (Sani et al., 2012) consider mean-variance optimization in a regret minimization framework. In the best arm identification setting, VaR-based criteria has been studied by (David & Shimkin, 2016) and (David et al., 2018). CVaR-based criteria has been explored in a bandit context by (Galichet et al., 2013), albeit with an assumption of bounded arms' distributions. More recently, (Kagrecha et al., 2019) incorporate a CVaR-based risk criterion in a best arm identification problem under a distribution oblivious setting.

The rest of this paper is organized as follows: Section 2 presents the preliminaries. Section 3 presents the key concentration bounds for sub-Gaussian and light-tailed distributions. Section 3.3 presents algorithms and their analyses for CVaR-based multi-armed stochastic bandits. Section 4 deals with CVaR estimation and concentration for heavy-tailed distributions.The proofs are contained in Section 5, and Section 6 concludes the paper.

2. Preliminaries

Given a r.v. X with cumulative distribution function (CDF) $F(\cdot)$, the VaR $v_\alpha(X)$ and CVaR $c_\alpha(X)$ at level $\alpha \in (0,1)$ are defined as follows [2]:

$$v_\alpha(X) = \inf\{\xi : \mathbb{P}\left[X \leq \xi\right] \geq \alpha\}, \text{ and} \tag{1}$$

$$c_\alpha(X) = v_\alpha(X) + \frac{1}{1-\alpha}\mathbb{E}\left[X - v_\alpha(X)\right]^+, \tag{2}$$

where we have used the notation $[X]^+ = \max(0, X)$. Typical values of α chosen in practice are 0.95 and 0.99. We make the following assumption on the r.v. X for the concentration bounds derived later.

(A1) The r.v. X is continuous with a density f that satisfies the following condition: There exist universal constants $\eta, \delta > 0$ such that $f(x) > \eta$ for all $x \in \left[v_\alpha - \frac{\delta}{2}, v_\alpha + \frac{\delta}{2}\right]$.

Under (A1), we have

$$v_\alpha(X) = F^{-1}(\alpha), \text{ and } c_\alpha(X) = \mathbb{E}\left[X | X \geq v_\alpha(X)\right].$$

Next, we recall standard definitions of sub-Gaussian and light-tailed (or sub-exponential) distributions — see Chapter 2 of (Wainwright, 2019).

Definition 2.1. *A r.v. X is said to be σ-sub-Gaussian for some $\sigma > 0$ if*

$$\mathbb{E}\left[\exp\left(\lambda X\right)\right] \leq \exp\left(\frac{\lambda^2\sigma^2}{2}\right), \textit{ for any } \lambda \in \mathbb{R}.$$

Definition 2.2. *A r.v. X is said to be light-tailed if there exist non-negative parameters σ and b such that*

$$\mathbb{E}\left[\exp\left(\lambda X\right)\right] \leq \exp\left(\frac{\lambda^2\sigma^2}{2}\right), \textit{ for any } |\lambda| < \frac{1}{b}. \tag{3}$$

3. CVaR estimation for sub-Gaussian and light-tailed distributions

In this section, we define the empirical CVaR, and present CVaR concentration results for sub-Gaussian and light-tailed distributions. We also discuss a multi-armed bandit application using the derived concentration bounds.

3.1. VaR and CVaR estimation

Let $\{X_i\}_{i=1}^n$ be n i.i.d. samples drawn from the distribution of X. Let $\{X_{[i]}\}_{i=1}^n$ be the order statistics of $\{X_i\}_{i=1}^n$, i.e., $X_{[1]} \leq X_{[2]} \cdots \leq X_{[n]}$. Let $\hat{F}_n(\cdot)$ be the empirical distribution function calculated using $\{X_i\}_{i=1}^n$, defined as $\hat{F}_n(x) = \frac{1}{n}\sum_{i=1}^n \mathbb{I}\left\{X_i \leq x\right\}, \forall x \in \mathbb{R}$. Notice that CVaR is a conditional expectation, where the conditioning event requires VaR. Thus, CVaR estimation requires VaR to be estimated as well. Let $\hat{v}_{n,\alpha}$ and $\hat{c}_{n,\alpha}$ denote the estimates of VaR and CVaR at level α using the n samples above. These quantities are defined as follows (Serfling, 2009):

$$\hat{v}_{n,\alpha} = X_{[\lceil n\alpha \rceil]}, \text{ and} \tag{4}$$

$$\hat{c}_{n,\alpha} = \frac{1}{n(1-\alpha)}\sum_{i=1}^n X_i \mathbb{I}\left\{X_i \geq \hat{v}_{n,\alpha}\right\}. \tag{5}$$

[2] For notational brevity, we omit X from the notations $v_\alpha(X)$ and $c_\alpha(X)$ whenever the underlying the r.v. can be understood from the context.

3.2. Concentration bounds

In the case of distributions with bounded support, a concentration result for CVaR exists in the literature (Gao et al., 2010). For the case of unbounded distributions, deriving a CVaR concentration result becomes considerably easier when the form of distributions are known, *i.e.,* when the closed-form expressions of VaR and CVaR can be derived. To illustrate, consider the case of a Gaussian r.v. X with mean μ and variance σ^2. Let $Q(\xi) = \frac{1}{\sqrt{2\pi}}\int_{\xi}^{\infty}\exp\left(-x^2/2\right)dx$. Notice that $Q(-x) = 1 - Q(x)$ and also that $F_X(\xi) = Q\left(\frac{\mu-\xi}{\sigma}\right)$. Hence, $v_\alpha(X)$ is the solution to $Q\left(\frac{\mu-\xi}{\sigma}\right) = \alpha$, which implies that

$$v_\alpha(X) = \mu - \sigma Q^{-1}(\alpha). \tag{6}$$

The CVaR $c_\alpha(X)$ for Gaussian X can be shown, using Acerbi's formula (see p. 329 in (Chatterjee, 2014)), to be equal to $\mu\left(\frac{\alpha}{1-\alpha}\right) + \sigma c_\alpha(Z)$, where Z is the standard Gaussian random variable *i.e.,* $Z \sim \mathcal{N}(0,1)$.

It is clear from the above argument that estimates of μ and σ are sufficient to estimate $c_\alpha(X)$ for the Gaussian case. Sample mean $\hat{\mu}_n$ and sample variance $\hat{\sigma}_n^2$ (computed using n samples from the distribution of X) would serve this purpose and we obtain $\hat{c}_n = \hat{\mu}\left(\frac{\alpha}{1-\alpha}\right) + \hat{\sigma}c_\alpha(Z)$ as a proxy for $c_\alpha(X)$. Given standard concentration bounds for these quantities through Hoeffding and Bernstein's inequalities, it is straightforward to establish that $\hat{c}_{n,\alpha}$ concentrates exponentially around $c_\alpha(X)$. Similarly, for the case of exponential random variables, we can exploit the memoryless property to derive an explicit expression for CVaR, in terms of the mean μ and the level α.

We therefore focus on distributions that do not have closed-form expressions for VaR and CVaR. In such a setting, the CVaR has to be estimated directly from the available samples. However, for establishing concentration bounds for the CVaR, which involves conditioning on a tail event, it is common to make some assumptions on the tail distribution. The following result presents concentration bounds for the cases when the underlying distribution is either sub-Gaussian or light-tailed.

Theorem 3.1 (***CVaR concentration***)**.** *Assume (A1). Let $\{X_i\}_{i=1}^n$ be a sequence of i.i.d. r.v.s. Let $\hat{c}_{n,\alpha}$ be the CVaR estimate given in (5) formed using the above set of samples.*

(i) *Suppose that X_i, $i = 1,\dots,n$ are σ-sub-Gaussian. Then, for any $\epsilon > 0$, we have*

$$\mathbb{P}\left[|\hat{c}_{n,\alpha} - c_\alpha| > \epsilon\right] \le 2\exp\left[-\frac{n\epsilon^2(1-\alpha)^2}{8\sigma^2}\right] + 4\exp\left[-\frac{n(1-\alpha)^2\eta^2\min\left(\epsilon^2, 4\delta^2\right)}{64}\right], \tag{7}$$

where η and δ are the constants specified in (A1).

(ii) *Suppose that X_i, $i = 1,\dots,n$ are light-tailed with parameters σ, b. Then, for any $\epsilon > 0$, we have*

$$\begin{aligned}&\mathbb{P}\left[|\hat{c}_{n,\alpha} - c_\alpha| > \epsilon\right]\\ &\le 2\exp\left[-\frac{n}{4}\min\left(\frac{\epsilon^2(1-\alpha)^2}{2\sigma^2}, \frac{\epsilon(1-\alpha)}{b}\right)\right]\\ &\quad + 4\exp\left[-\frac{n(1-\alpha)^2\eta^2\min\left(\epsilon^2, 4\delta^2\right)}{64}\right].\end{aligned} \tag{8}$$

Remark 3.2. *For a σ-sub-Gaussian r.v. X, with mean μ, the classic concentration result for sample mean, say $\hat{\mu}_n$, takes the form*

$$\mathbb{P}\left[|\hat{\mu}_n - \mu| > \epsilon\right] \le 2\exp\left[-\frac{n\epsilon^2}{2\sigma^2}\right].$$

In comparison, the tail bound in (7) involves additional factors inside the exponential term, and this because of the VaR estimate $\hat{v}_{n,\alpha}$ that features in the CVaR estimate $\hat{c}_{n,\alpha}$.

Proof. (Sketch) We split the CVaR estimation error as $\hat{c}_{n,\alpha} - c_\alpha = I_n + e_n$, where

$$I_n = \frac{1}{1-\alpha}\left[\frac{1}{n}\sum_{i=1}^{n}(X_i - v_\alpha)^+ - \mathbb{E}\left[(X - v_\alpha)^+\right]\right], \tag{9}$$

$$\begin{aligned}e_n &= \frac{\hat{v}_{n,\alpha} - v_\alpha}{1-\alpha}\left[\hat{F}_n(\hat{v}_{n,\alpha}) - \alpha\right]\\ &+ \frac{1}{n}\sum_{i=1}^{n}\frac{X_i - v_\alpha}{1-\alpha}\left[\mathbb{I}\left\{X_i \ge \hat{v}_{n,\alpha}\right\} - \mathbb{I}\left\{X_i \ge v_\alpha\right\}\right].\end{aligned} \tag{10}$$

The I_n term can be bounded using a standard sub-Gaussian concentration result (cf. Theorem 2.1 in (Wainwright, 2019)), after observing that $(X - v_\alpha)^+$ is σ-sub-Gaussian (see Section 5.1 for a rigorous justification), and this leads to a bound of the form:

$$\mathbb{P}\left[I_n > \epsilon\right] \le 2\exp\left[-\frac{n\epsilon^2(1-\alpha)^2}{2\sigma^2}\right]. \tag{11}$$

On the other hand, the bound on the e_n term requires VaR concentration, and takes the following form:

$$\mathbb{P}\left[e_n > \epsilon\right] \le 4\exp\left(-\frac{n(1-\alpha)^2\eta^2\min\left(\epsilon^2, \delta^2\right)}{16}\right). \tag{12}$$

The final bound in (7) can be obtained by combining the bounds in (11) and (12). A similar observation holds for the light-tailed concentration result in (8). Note that the bound on e_n in (12) holds for any distribution satisfying (A1), while the high-probability bound on I_n relates to sum of i.i.d. r.v.s, and would require tail assumptions. The detailed proof is provided in Section 5.1. □

Remark 3.3. *The bound in the theorem above is significantly better than the two-sided bound obtained in (Bhat & Prashanth, 2019) for the light-tailed case. In particular, the bound in the theorem above has an exponential tail decay irrespective of whether ϵ is large or small, while the bound in (Bhat & Prashanth, 2019) has an exponential decay for small ϵ, and a power law for large ϵ. For a light-tailed r.v., one expects a tail behavior similar to that of Gaussian with constant variance for small ϵ, and an exponential decay for large ϵ, and our bound is consistent with this expected behavior. For small ϵ, the bound that we have in (8) is tighter, in terms of the constants, than the corresponding bound in (Bhat & Prashanth, 2019). This is because the Wasserstein distance-based approach shows concentration of the empirical CDF around the true CDF everywhere, while our approach involves concentration of the empirical CDF only around the VaR.*

Remark 3.4. *In comparison to the one-sided bound for light-tailed r.v.s, in (Kolla et al., 2019), our bound exhibits better dependence w.r.t. the number of samples n as well as the accuracy ϵ. More importantly, since our bound is two-sided, it opens avenues for a bandit application, while a one-sided bound is insufficient for this purpose.*

Remark 3.5. *Note that the constant η, which has the interpretation of a 'minimum slope' in a neighbourhood of v_α, appears in the concentration bounds of Theorem 3.1. This η dependence comes from the VaR concentration result in Lemma 5.1. We wish to point out that this is inevitable, and is not an artifact of our proof technique. Indeed, if the CDF is allowed to be arbitrarily 'flat' over a wide interval, it is possible to construct examples to show that the empirical VaR does not concentrate well around the true VaR.*

In the following section, we present multi-armed bandit algorithms that incorporate a CVaR-based objective, and analyze the finite-time performance of these algorithms using the bound derived in Theorem 3.1. Unlike previous CVaR-based bandit algorithms (cf. (Galichet et al., 2013)) that impose a bounded support requirement on the arms' distribution, our bounds in the bandit setting hold for more general sub-Gaussian and light-tailed distributions.

3.3. Application: Multi-armed bandits

We consider a K-armed stochastic bandit problem, with arms' distributions $\mathcal{P}_1, \ldots, \mathcal{P}_K$. We consider a variant of the bandit problem, where the goal is to find the arm with the lowest CVaR. Let c_α^i and v_α^i denote the CVaR and VaR of the arm i at level α. Let $c^* = \min_{i=1,\ldots,K} c_\alpha^i$, and i^* be the arm that achieves this minimum. Let $\Delta_i = c_\alpha^i - c_\alpha^{i^*}$ denote the gap between the CVaR values of arm-i and the optimal arm.

We consider a lowest CVaR arm identification problem in the fixed budget setting, and devise an algorithm based on successive rejects (SR) (Audibert et al., 2010) to incorporate the CVaR criterion. A bandit algorithm interacts with the environment over a given budget of n rounds. In each round $t = 1, \ldots, n$, the algorithm pulls an arm $I_t \in \{1, \ldots, K\}$ and observes a sample cost from the distribution $\mathcal{P}_{I_t}$. At the end of the budget n rounds, the bandit algorithm recommends an arm J_n and is judged based on the probability of incorrect identification, i.e., $\mathbb{P}[J_n \neq i^*]$, where i^* denotes the best arm. Earlier works use the expected value to define the best arm, while we use CVaR.

The goal is to minimize the probability of incorrect identification, i.e., $\mathbb{P}[J_n \neq i^*]$. Let arm-$[i]$ denotes the i^{th} lowest CVaR valued arm.

Algorithm 1 CVaR-SR algorithm

Initialization: Set $A_1 = \{1, \ldots, K\}$, $\overline{\log} K = \frac{1}{2} + \sum_{i=2}^{K} \frac{1}{i}$, $n_0 = 0$, $n_k = \left\lceil \frac{1}{\overline{\log} K} \frac{n-K}{K+1-k} \right\rceil$, $k = 1, \ldots, K-1$.

for $k = 1, 2, \ldots, K-1$ **do**

Play each arm in A_k for $(n_k - n_{k-1})$ times.

Set $A_{k+1} = A_k \setminus \arg\max_{i \in A_k} \hat{c}^i_{\alpha, n_k}$.

end for

Output: Return the solitary element in A_K.

Algorithm 1 presents the pseudo code of our CVaR-SR algorithm, designed to find the CVaR-optimal arm under a fixed budget. The algorithm is a variation of the regular SR algorithm, with the following key difference: regular SR uses sample mean to estimate the expected value of each arm, while CVaR-SR used empirical CVaR, as defined in (5), to estimate CVaR for each arm. The elimination logic, i.e., having $K-1$ phases, and removing the worst arm (according to sample estimates of CVaR) at the end of each phase, is borrowed from regular SR.

In the following result, we analyze the performance of CVaR-SR algorithm for light-tailed distributions.

Theorem 3.6 (***Probability of incorrect identification***). *Consider a K-armed stochastic bandit, where the arms' distributions satisfy (A1) and are either sub-Gaussian or light-tailed. For a given budget n, the arm, say J_n, returned by the CVaR-SR algorithm satisfies:*

$$\mathbb{P}[J_n \neq i^*] \leq 3K(K-1) \exp\left[-\frac{(n-K)(1-\alpha)^2 G_{\max}}{8H \overline{\log} K} \right],$$

where $G_{\max}$ is a problem dependent constant that does not depend on the underlying CVaR gaps and n, and the hard-

ness measure H is defined by

$$H = \max_{i\in\{1,\ldots,K\}} \frac{i}{\min\left(\Delta_{[i]}^2, 4\delta_{[i]}^2\right)} \textit{ for } \sigma\textit{-sub-Gaussian arms},$$

$$H = \max_{i\in\{1,\ldots,K\}} \frac{i}{\min\{\Delta_{[i]}, \Delta_{[i]}^2, 4\delta_{[i]}^2\}}$$

for (σ, b)-light-tailed arms.

In the above, δ_i denotes the constant from (A1) corresponding to the distribution of arm i, for $i = 1, \ldots, K$.

Proof. The proof follows arguments similar to that used in deriving the corresponding bound for SR, and is provided in Section 5.3. □

4. CVaR estimation: the bounded moment case

In this section, we derive CVaR concentration results under a much milder assumption than in the previous section. In particular, we assume that a bounded pth moment exists, enabling us to handle heavy-tailed distributions as well as light-tailed ones.

As mentioned before, an alternative proof approach using Wasserstein distance (Bhat & Prashanth, 2019) provides weak concentration rates for distributions with bounded higher moments - a gap that we address in this work. In particular, we employ a truncation-based estimator for CVaR to handle this bounded moment case. Throughout this section, we assume that the underlying distribution satisfies:

(A2) $\exists p \in (1, 2], u$ such that $\mathbb{E}[|X|^p] < u < \infty$.

Note that if a pth moment is bounded for some $p > 2$, Lyapunov's inequality (see pp. 143 in (Grimmett & Stirzaker, 2001)) implies that the second moment is bounded, and we simply take $p = 2$.

4.1. CVaR estimation

Recall that $\{X_{[i]}\}_{i=1}^n$ denote the order statistics of n i.i.d. samples drawn from the distribution of X. Using the VaR estimate $\hat{v}_{n,\alpha}$, as defined earlier in Section 3.1, we propose a truncation-based estimator $\hat{c}_{n,\alpha}$ for CVaR at level α, defined as follows:

$$\hat{c}_{n,\alpha} = \frac{1}{n(1-\alpha)} \sum_{i=1}^{n} X_i \mathbb{I}\left\{\hat{v}_{n,\alpha} \le X_i \le B_i\right\}, \qquad (13)$$

$$\text{where } B_i = \left(\frac{ui}{\log\left(\frac{3}{\xi}\right)}\right)^{\frac{1}{p}}.$$

In (5), B_i represents a truncation level of X_i, and the choice for B_i given above is under the assumption that $\mathbb{E}[|X|^p] < u < \infty$ for some $p \in (1, 2]$. Such a truncation based estimator has been employed in the context of expected regret minimization with heavy-tailed random variables in (Bubeck et al., 2013). Intuitively, the truncation level serves to discard very large samples values early on, as B_i is set to grow slowly with i.

4.2. Concentration bounds

In particular, the following result is more general, as it can handle heavy-tailed distributions that satisfy (A2).

Theorem 4.1 (***CVaR concentration: Bounded moment case***)**.** *Let $\{X_i\}_{i=1}^n$ be a sequence of i.i.d. r.v.s satisfying (A1) and (A2). Let $\hat{c}_{n,\alpha}$ be the CVaR estimate given in (13) formed using the above set of samples. Then, for any $p \in (1, 2]$, and $\xi \in (0, 1)$, we have with probability (w.p.) at least $1 - \xi$,*

$$|\hat{c}_{n,\alpha} - c_\alpha| \le \frac{(5u^{\frac{1}{p}} + v_\alpha)}{(1-\alpha)} \left(\frac{1}{n}\right)^{1-\frac{1}{p}} \sqrt{\log\left(\frac{3}{\xi}\right)} + \max\left(\frac{4}{\eta(1-\alpha)}\sqrt{\frac{\log\left(\frac{4}{\xi}\right)}{n}}, \delta\right), \qquad (14)$$

where η and δ are as defined in (A1).

We have presented a high-confidence bound in the theorem above, and not a tail bound, as in Theorem 3.1. The rationale behind this choice is to make the dependence on the probability ξ apparent, esp. considering the truncation level B_i in (13) is a function of ξ.

Proof. We split the estimation error $c_\alpha - \hat{c}_{n,\alpha}$ into three components as follows:

$$c_\alpha - \hat{c}_{n,\alpha} = I_{n,1} - I_{n,2} + e_n, \text{ where}$$

$$I_{n,1} = \frac{1}{1-\alpha}\mathbb{E}\left[X \,\mathbb{I}\left\{v_\alpha \le X\right\}\right] - \frac{1}{n(1-\alpha)} \sum_{i=1}^{n} X_i \,\mathbb{I}\left\{v_\alpha \le X_i \le B_i\right\},$$

$$I_{n,2} = \frac{1}{1-\alpha}\mathbb{E}\left[v_\alpha \,\mathbb{I}\left\{v_\alpha \le X\right\}\right] - \frac{1}{n(1-\alpha)} \sum_{i=1}^{n} v_\alpha \,\mathbb{I}\left\{v_\alpha \le X_i \le B_i\right\}, \text{ and}$$

$$e_n = (\hat{v}_{n,\alpha} - v_\alpha) + \frac{1}{n(1-\alpha)} \sum_{i=1}^{n} (X_i - \hat{v}_{n,\alpha}) \times \left[\mathbb{I}\left\{\hat{v}_{n,\alpha} \le X_i \le B_i\right\} - \mathbb{I}\left\{v_\alpha \le X_i \le B_i\right\}\right].$$

Using a technique from (Bubeck et al., 2013), we obtain the following bound for the first two terms, i.e., $I_{n,1}$ and

$I_{n,2}$:

$$|I_{n,1} - I_{n,2}| \leq \frac{(5u^{\frac{1}{p}} + v_\alpha)}{(1-\alpha)} \left(\frac{1}{n}\right)^{1-\frac{1}{p}} \sqrt{\log\left(\frac{3}{\xi}\right)}. \quad (15)$$

The above bound does not require VaR concentration, as $I_{n,1}$ and $I_{n,2}$ do not have the empirical VaR $\hat{v}_{n,\alpha}$ in their definitions. On the other hand, VaR concentration is required to arrive at the following bound on the e_n term:

$$|e_n| \leq \max\left(\frac{4}{\eta(1-\alpha)}\sqrt{\frac{\log\left(\frac{4}{\xi}\right)}{n}}, \delta\right), \text{ w.p. } (1-\xi). \quad (16)$$

The main claim follows by combining the bounds in (15) and (16).

The detailed proof is provided in Section 5.2. □

Substituting $p = 2$ in the bound derived above leads to the following CVaR concentration bound for distributions with a bounded second moment:

Corollary 4.2. *Assume conditions of Theorem 4.1, for the case when the distribution of X has a bounded second moment, i.e., $p = 2$. Then, for any $\xi \in (0,1)$, we have w.p. at least $1-\xi$,*

$$|\hat{c}_{n,\alpha} - c_\alpha| \leq \frac{(5\sqrt{u} + v_\alpha)}{(1-\alpha)\sqrt{n}}\sqrt{\log\left(\frac{3}{\xi}\right)} + \max\left(\frac{4}{\eta(1-\alpha)}\sqrt{\frac{\log\left(\frac{4}{\xi}\right)}{n}}, \delta\right),$$

where η and δ are as defined in (A1).

Remark 4.3. *A bandit application for the case of heavy-tailed distributions can be worked out using arguments similar to that in Section 3.3. The main difference is that the SR algorithm in the heavy-tailed case would involve a truncated estimator, and a slightly different hardness measure that is derived using Theorem 4.1. We omit the details due to space constraints.*

5. Proofs

Before providing the main proofs of the CVaR concentration bounds in Theorems 3.1 and 4.1, we note that empirical CVaR, as defined in (5), involves empirical VaR, and it is natural to expect that empirical CVaR concentration would require empirical VaR to concentrate as well. VaR concentration bounds have been derived recently in (Kolla et al., 2019), and we state a straightforward variation to their tail bound below. Note that the result below will be used to establish the bound in Theorems 3.1, and 4.1.

Lemma 5.1 (***VaR concentration***)**.** *Suppose that (A1) holds. For any $\epsilon > 0$, we have*

$$\mathbb{P}\left[|\hat{v}_{n,\alpha} - v_\alpha| \geq \epsilon\right] \leq 2\exp\left(-2n\eta^2 \min\left(\epsilon^2, \delta^2\right)\right), \quad (17)$$

where η is the constant specified in (A1).

Proof. From the initial passage in the proof of Proposition 2 in (Kolla et al., 2019), we have

$$\mathbb{P}\left[|\hat{v}_{n,\alpha} - v_\alpha| \geq \epsilon\right] \leq 2\exp\left(-2n\zeta_\epsilon^2\right), \quad (18)$$

where $\zeta_\epsilon = \min\left(F\left(v_\alpha + \epsilon\right) - F\left(v_\alpha\right), F\left(v_\alpha\right) - F\left(v_\alpha - \epsilon\right)\right)$. From (A1), $\zeta_\epsilon \geq \eta\epsilon$, for $\epsilon < \delta$, where η and δ are constants specified in (A1). Hence, for $\epsilon < \delta$, we have

$$\mathbb{P}\left[|\hat{v}_{n,\alpha} - v_\alpha| \geq \epsilon\right] \leq 2\exp\left(-2n\eta^2\epsilon^2\right).$$

The bound in (17) follows by observing that $\mathbb{P}\left[|\hat{v}_{n,\alpha} - v_\alpha| \geq \epsilon\right] \leq \mathbb{P}\left[|\hat{v}_{n,\alpha} - v_\alpha| \geq \delta\right]$ for $\epsilon > \delta$. □

Notice that

$$\hat{c}_{n,\alpha} = v_\alpha + \frac{1}{n(1-\alpha)}\sum_{i=1}^{n}\left(X_i - v_\alpha\right)\mathbb{I}\left\{v_\alpha \leq X_i\right\} + e_n, \quad (19)$$

where e_n is as defined in (10). We now present a high-probability bound on the error term e_n, which holds for any class of distributions satisfying the growth condition in (A1).

Lemma 5.2. *Assume (A1). For e_n as defined in (10), the following bound holds:*

$$\mathbb{P}\left[|e_n| > \epsilon\right] \leq 4\exp\left(-\frac{n(1-\alpha)^2\eta^2\min\left(\epsilon^2, \delta^2\right)}{16}\right), \quad (20)$$

where η and δ are the constants specified in (A1).

Proof. Notice that

$$\begin{aligned}
|e_n| &\leq \frac{|v_\alpha - \hat{v}_{n,\alpha}|}{1-\alpha}|\alpha - \hat{F}_n(\hat{v}_{n,\alpha})| \\
&\quad + \frac{|v_\alpha - \hat{v}_{n,\alpha}|}{1-\alpha}|\hat{F}_n(v_\alpha) - \hat{F}_n(\hat{v}_{n,\alpha})| \\
&\leq \frac{|v_\alpha - \hat{v}_{n,\alpha}|}{1-\alpha}\Big[2|\hat{F}_n(\hat{v}_{n,\alpha}) - F(v_\alpha)| \\
&\quad + |\hat{F}_n(v_\alpha) - F(v_\alpha)|\Big]. \quad (21)
\end{aligned}$$

Using $|\hat{F}_n(\hat{v}_{n,\alpha}) - F(v_\alpha)| \leq 1/n$ and $|\hat{F}_n(v_\alpha) - F(v_\alpha)| \leq 2$, followed by an application of the VaR concentration result in Lemma 5.1, we obtain

$$\mathbb{P}\left[|e_n| > \epsilon\right] \leq \mathbb{P}\left[\frac{2}{n}\frac{1}{(1-\alpha)}|\hat{v}_{n,\alpha} - v_\alpha| > \frac{\epsilon}{2}\right]$$

$$+\mathbb{P}\left[\frac{1}{1-\alpha}|\hat{v}_{n,\alpha}-v_\alpha||\hat{F}_n(v_\alpha)-F(v_\alpha)|>\frac{\epsilon}{2}\right]$$
$$\leq 2\exp\left[-\frac{n^3(1-\alpha)^2\eta^2\min\left(\epsilon^2,\delta^2\right)}{8}\right]$$
$$+2\exp\left[-\frac{n(1-\alpha)^2\eta^2\min\left(\epsilon^2,\delta^2\right)}{16}\right]$$
$$\leq 4\exp\left(-\frac{n(1-\alpha)^2\eta^2\min\left(\epsilon^2,\delta^2\right)}{16}\right).$$

□

5.1. Proof of Theorem 3.1

Proof. **(i) Sub-Gaussian case**
Using (19), we have

$$\hat{c}_{n,\alpha}-c_\alpha=I_n+e_n, \tag{22}$$

where I_n and e_n are as defined in (9) and (10), respectively. For bounding the I_n term, we use the fact that $(X-v_\alpha)^+$ is a σ-sub-Gaussian r.v. This can be argued as follows: Letting $\mu_\alpha^+=\mathbb{E}\left[(X-v_\alpha)^+\right]$, we have

$$\mathbb{E}\left[\exp\left[\lambda\left((X-v_\alpha)^+-\mu_\alpha^+\right)\right]\right]$$
$$\leq 1+\frac{\lambda^2\mathbb{E}X^2}{2}+o\left(\lambda^2\right).$$

In the above, we have used the following fact that

$$\mathbb{E}\left[[X-v_\alpha]^2\mathbb{I}\left\{X\geq v_\alpha\right\}\right]\leq\mathbb{E}\left[X^2\mathbb{I}\left\{X\geq v_\alpha\right\}\right]$$
$$\leq\mathbb{E}X^2. \tag{23}$$

Comparing with the following identity:

$$\exp\left(\frac{\lambda^2\sigma^2}{2}\right)=1+\frac{\lambda^2\sigma^2}{2}+o\left(\lambda^2\right),$$

it is easy to see that $(X-v_\alpha)^+$ is σ-sub-Gaussian, whenever X is σ-sub-Gaussian.

Using a sub-Gaussian concentration result (cf. Theorem 2.1 in (Wainwright, 2019)), we obtain

$$\mathbb{P}\left[|I_n|>\epsilon\right]\leq 2\exp\left[-\frac{n\epsilon^2(1-\alpha)^2}{2\sigma^2}\right]. \tag{24}$$

The claim for the sub-Gaussian case in (7) follows by using

$$\mathbb{P}\left[|\hat{c}_{n,\alpha}-c_\alpha|>\epsilon\right]\leq\mathbb{P}\left[|I_n|>\frac{\epsilon}{2}\right]+\mathbb{P}\left[e_n>\frac{\epsilon}{2}\right], \tag{25}$$

and substituting the bounds obtained in (20) and (24) in the RHS above.

(ii) Light-tailed case
For bounding the I_n term in (22), we show that $(X-v_\alpha)^+$ is a light-tailed r.v. Denoting $\mu_\alpha^+=\mathbb{E}\left[(X-v_\alpha)^+\right]$, we have

$$\mathbb{P}\left[(X_i-v_\alpha)^+-\mu_\alpha^+>\epsilon\right]$$
$$=\mathbb{P}\left[X>v_\alpha+\mu_\alpha^++\epsilon\right]\leq c_1\exp\left(-c_2\left(v_\alpha+\epsilon\right)\right)$$
$$\leq c_1\exp\left(-c_4\epsilon\right),$$

where c_1, c_2, and c_4 are distribution-dependent constants. Using the fact that X is light-tailed, we have

$$\mathbb{E}\left[\exp\left[\lambda\left((X-v_\alpha)^+-\mu_\alpha^+\right)\right]\right]$$
$$\leq 1+\frac{\lambda^2\mathbb{E}X^2}{2}+o\left(\lambda^2\right).$$

In the above, we have used the following fact from (23). Comparing with the following identity:

$$\exp\left(\frac{\lambda^2\sigma^2}{2}\right)=1+\frac{\lambda^2\sigma^2}{2}+\frac{\lambda^2 v_\alpha^2}{2}+o\left(\lambda^2\right),$$

it is easy to see that $(X-v_\alpha)^+$ is a light-tailed r.v. with parameters (σ^2,b), whenever X is light-tailed with parameters (σ^2,b).

The rest of the proof follows in a similar fashion as part (i), by using a standard light-tailed concentration result (cf. Theorem 2.2. in (Wainwright, 2019)). □

5.2. Proof of Theorem 4.1

Proof. Notice that

$$\hat{c}_{n,\alpha}=v_\alpha+\frac{1}{n}\sum_{i=1}^{n}\frac{(X_i-v_\alpha)}{(1-\alpha)}\mathbb{I}\left\{v_\alpha\leq X_i\leq B_i\right\}+e_n,$$

where

$$e_n=(\hat{v}_{n,\alpha}-v_\alpha)+\frac{1}{n(1-\alpha)}\sum_{i=1}^{n}(X_i-\hat{v}_{n,\alpha})$$
$$\times\left[\mathbb{I}\left\{\hat{v}_{n,\alpha}\leq X_i\leq B_i\right\}-\mathbb{I}\left\{v_\alpha\leq X_i\leq B_i\right\}\right]$$
$$=(\hat{v}_{n,\alpha}-v_\alpha)+\frac{(v_\alpha-\hat{v}_{n,\alpha})}{(1-\alpha)}\left(\hat{F}_n(B_i)-\hat{F}_n(\hat{v}_{n,\alpha})\right)$$
$$+\frac{1}{n(1-\alpha)}\sum_{i=1}^{n}(X_i-v_\alpha)\left[\mathbb{I}\left\{\hat{v}_{n,\alpha}\leq X_i\leq B_i\right\}\right.$$
$$\left.-\mathbb{I}\left\{v_\alpha\leq X_i\leq B_i\right\}\right]$$

Thus,

$$|e_n|\leq\frac{|v_\alpha-\hat{v}_{n,\alpha}|}{1-\alpha}\left[2|\hat{F}_n(\hat{v}_{n,\alpha})-F(v_\alpha)|+|\hat{F}_n(v_\alpha)-F(v_\alpha)|\right].$$

Using Lemma 5.2, and the inequality above, we obtain

$$\mathbb{P}\left[|e_n| > \epsilon\right] \leq 4 \exp\left(-\frac{n(1-\alpha)^2\eta^2 \min(\epsilon^2, \delta^2)}{16}\right),$$

or, equivalently,

$$|e_n| \leq \max\left(\frac{4}{\eta(1-\alpha)}\sqrt{\frac{\log\left(\frac{4}{\xi}\right)}{n}}, \delta\right), \text{ w.p. } (1-\xi). \tag{26}$$

Next, observe that the estimation error can be re-written as

$$c_\alpha - \hat{c}_{n,\alpha} = I_{n,1} - I_{n,2} + e_n, \text{ where}$$

$$\begin{aligned} I_{n,1} =& \frac{1}{1-\alpha}\mathbb{E}\left[X\, \mathbb{I}\left\{v_\alpha \leq X\right\}\right] \\ &- \frac{1}{n(1-\alpha)}\sum_{i=1}^{n} X_i\, \mathbb{I}\left\{v_\alpha \leq X_i \leq B_i\right\}, \text{ and} \\ I_{n,2} =& \frac{1}{1-\alpha}\mathbb{E}\left[v_\alpha\, \mathbb{I}\left\{v_\alpha \leq X\right\}\right] \\ &- \frac{1}{n(1-\alpha)}\sum_{i=1}^{n} v_\alpha\, \mathbb{I}\left\{v_\alpha \leq X_i \leq B_i\right\}. \end{aligned}$$

We bound the $I_{n,1}$ term, using a technique from (Bubeck et al., 2013), as follows:

$$\begin{aligned} &\frac{1}{1-\alpha}\left|\mathbb{E}\left[X\, \mathbb{I}\left\{v_\alpha \leq X\right\}\right] - \frac{1}{n}\sum_{i=1}^{n} X_i\, \mathbb{I}\left\{v_\alpha \leq X_i \leq B_i\right\}\right| \\ &= \frac{1}{n(1-\alpha)}\left|\sum_{i=1}^{n}\mathbb{E}\left[X\, \mathbb{I}\left\{X > B_i\right\}\right]\right. \\ &\left. + \sum_{i=1}^{n}\mathbb{E}\left[X\, \mathbb{I}\left\{v_\alpha \leq X \leq B_i\right\}\right] - X_i\, \mathbb{I}\left\{v_\alpha \leq X_i \leq B_i\right\}\right| \end{aligned} \tag{27}$$

$$\begin{aligned} &\leq \frac{1}{n(1-\alpha)}\sum_{i=1}^{n}\frac{u}{B_i^{p-1}} + \frac{1}{(1-\alpha)}\sqrt{\frac{2B_n^{2-p}u\log\left(\frac{2}{\xi}\right)}{n}} \\ &+ \frac{1}{(1-\alpha)}\frac{2B_n\log\left(\frac{2}{\xi}\right)}{3n}, \text{ holds w.p. } (1-\xi), \end{aligned}$$

where we have used the fact that $\mathbb{E}(X^p) \geq B^{p-1}\mathbb{E}\left[X\, \mathbb{I}\left\{X > B\right\}\right]$ to handle the first term in (27). The bound on the second term is justified as follows: Let $Z_i = X_i\, \mathbb{I}\left\{v_\alpha \leq X_i \leq B_i\right\}$, $i = 1, \ldots, n$. We have $|Z_i| \leq B_n, \forall i$, and $EZ_i^2 \leq uB_n^{2-p}$. Thus, applying Bernstein's inequality leads to the following bound:

$$\mathbb{P}\left[\left|\sum_{i=1}^{n}(Z_i - \mathbb{E}Z_i)\right| > \epsilon\right] \leq 2\exp\left(\frac{\epsilon^2}{2nuB_n^{2-p} + \frac{2}{3}B_n\epsilon}\right).$$

Turning the above inequality into high-confidence form, we obtain the following bound w.p. $(1-\xi)$,

$$\left|\sum_{i=1}^{n}(Z_i - \mathbb{E}Z_i)\right| \leq \sqrt{2nuB_n^{2-p}\log\left(\frac{2}{\xi}\right)} + \frac{2B_n\log(\frac{2}{\xi})}{3}.$$

Along similar lines, the term $I_{n,2}$ is bounded as follows:

$$\begin{aligned} &\frac{1}{1-\alpha}\left|\mathbb{E}\left[v_\alpha\, \mathbb{I}\left\{v_\alpha \leq X\right\}\right] - \frac{1}{n}\sum_{i=1}^{n} v_\alpha\, \mathbb{I}\left\{v_\alpha \leq X_i \leq B_i\right\}\right| \\ &= \frac{v_\alpha}{n(1-\alpha)}\left|\sum_{i=1}^{n}\mathbb{E}\left[\mathbb{I}\left\{X > B_i\right\}\right]\right. \\ &\quad + \frac{v_\alpha}{n(1-\alpha)}\sum_{i=1}^{n}\Big(\mathbb{E}\left[\mathbb{I}\left\{v_\alpha \leq X \leq B_i\right\}\right] \\ &\qquad \left. - \mathbb{I}\left\{v_\alpha \leq X_i \leq B_i\right\}\Big)\right| \end{aligned} \tag{28}$$

$$\begin{aligned} &\leq \frac{1}{n(1-\alpha)}\sum_{i=1}^{n}\frac{u}{B_i^{p-1}} + \frac{v_\alpha}{(1-\alpha)}\sqrt{\frac{\log\left(\frac{2}{\xi}\right)}{2n}}, \\ &\quad \text{holds w.p. } (1-\xi), \end{aligned}$$

where we have used Hoeffding's inequality, and $B_i^p \geq B_i^{p-1}$ for bounding the second term[3] in (28), while the first term is bounded using an argument similar to that used in bounding $I_{n,1}$ term above.

Using $B_i = \left(\frac{ui}{\log(\frac{3}{\xi})}\right)^{\frac{1}{p}}$, we have, w.p. $(1-\xi)$,

$$|I_{n,1}| \leq \frac{4u^{\frac{1}{p}}}{(1-\alpha)}\left(\frac{\log\left(\frac{3}{\xi}\right)}{n}\right)^{1-\frac{1}{p}}, \text{ and}$$

$$|I_{n,2}| \leq \frac{u^{\frac{1}{p}}}{(1-\alpha)}\left(\frac{\log\left(\frac{3}{\xi}\right)}{n}\right)^{1-\frac{1}{p}} + \frac{v_\alpha}{(1-\alpha)}\sqrt{\frac{\log\left(\frac{3}{\xi}\right)}{n}}.$$

Combining the bounds on $I_{n,1}$ and $I_{n,2}$ above, we have w.p. $(1-\xi)$,

$$|I_{n,1} - I_{n,2}| \leq \frac{(5u^{\frac{1}{p}} + v_\alpha)}{(1-\alpha)}\left(\frac{1}{n}\right)^{1-\frac{1}{p}}\sqrt{\log\left(\frac{3}{\xi}\right)}, \tag{29}$$

where we used the fact that $(\log(3/\xi))^{1-\frac{1}{p}} \leq \sqrt{\log\left(\frac{3}{\xi}\right)}$, since $\log\left(\frac{3}{\xi}\right) > 1$. Combining the bounds in (26) and (29), we have w.p. $(1-\xi)$,

$$|\hat{c}_{n,\alpha} - c_\alpha| \leq \frac{(5u^{\frac{1}{p}} + v_\alpha)}{(1-\alpha)}\left(\frac{1}{n}\right)^{1-\frac{1}{p}}\sqrt{\log\left(\frac{3}{\xi}\right)}$$

[3] Note that for a fixed ξ, we can assume $B_i > 1$ for all i by taking a u large enough.

$$+ \max\left(\frac{4}{\eta(1-\alpha)}\sqrt{\frac{\log\left(\frac{4}{\xi}\right)}{n}}, \delta\right).$$

□

5.3. Proof of Theorem 3.6

Proof. **(i) Sub-Gaussian case**

We begin the proof by rewriting the CVaR concentration bound in (7) for the sub-Gaussian case in a simplified manner as follows:

$$\mathbb{P}\left[|\hat{c}_{n,\alpha} - c_\alpha| > \epsilon\right] \leq 6\exp\left[-\frac{nG(1-\alpha)^2\min\left(\epsilon^2,\delta^2\right)}{2}\right], \tag{30}$$

where $G = \frac{\min(\eta^2,1)}{\max(\sigma^2,8)}$.

Note that, if the CVaR-SR algorithm has eliminated the optimal arm in phase i then it implies that at least one of the last i worst arms *i.e.,* one of the arms in $\{[K],[K-1],\cdots,[K-i+1]\}$ must not have been eliminated in phase i. Hence, we obtain

$$\begin{aligned}
\mathbb{P}\left[J_n \neq i^*\right] &\leq \sum_{k=1}^{K-1}\sum_{i=K+1-k}^{K}\mathbb{P}\left[\hat{c}^{i^*}_{n_k,\alpha} \geq \hat{c}^{[i]}_{n_k,\alpha}\right]\\
&= \sum_{k=1}^{K-1}\sum_{i=K+1-k}^{K}\mathbb{P}\left[\hat{c}^{i^*}_{n_k\alpha} - c^{i^*}_\alpha - \hat{c}^{[i]}_{n_k,\alpha} + c^{[i]}_\alpha \geq c^{[i]}_\alpha - c^{i^*}_\alpha\right]\\
&\leq \sum_{k=1}^{K-1}\sum_{i=K+1-k}^{K}\mathbb{P}\left[\hat{c}^{i^*}_{n_k,\alpha} - c^{i^*}_\alpha \geq \frac{\Delta_{[i]}}{2}\right]\\
&+ \sum_{k=1}^{K-1}\sum_{i=K+1-k}^{K}\mathbb{P}\left[c^{[i]}_\alpha - \hat{c}^{[i]}_{n_k,\alpha} \geq \frac{\Delta_{[i]}}{2}\right]
\end{aligned} \tag{31}$$

We now bound the above terms individually as follows.

$$\begin{aligned}
&\sum_{k=1}^{K-1}\sum_{i=K+1-k}^{K}\mathbb{P}\left[c^{[i]}_\alpha - \hat{c}^{[i]}_{n_k,\alpha} \geq \frac{\Delta_{[i]}}{2}\right]\\
&\leq \sum_{k=1}^{K-1}\sum_{i=K+1-k}^{K}\mathbb{P}\left[|\hat{c}^{[i]}_{n_k,\alpha} - c^{[i]}_\alpha| \geq \frac{\Delta_{[i]}}{2}\right]\\
&\overset{(a)}{\leq} \sum_{k=1}^{K-1}\sum_{i=K+1-k}^{K} 6\exp\left[-\frac{n(1-\alpha)^2\min\left[\Delta^2_{[i]},\delta^2_{[i]}\right]G_{[i]}}{8}\right]\\
&\leq \sum_{k=1}^{K-1}\sum_{i=K+1-k}^{K} 6\exp\left([-\frac{n(1-\alpha)^2 G_{\max}}{8}\right.\\
&\qquad \left.\times\min\left[\Delta^2_{[i]},4\delta^2_{[i]}\right]\right)\\
&\leq \sum_{k=1}^{K-1} 6k\exp\left(-\frac{nG_{\max}(1-\alpha)^2}{8}\right.\\
&\qquad \left.\times\min\left(\Delta^2_{[K+1-k]},4\delta^2_{[K+1-k]}\right)\right),
\end{aligned} \tag{32}$$

where we used (30) in arriving at the inequality in (a), and $G_{\max} = \max_i G_i$.

Notice that

$$n\min\left(\Delta^2_{[K+1-k]},4\delta^2_{[K+1-k]}\right) \geq \frac{n-K}{H\overline{\log}K},$$

where H is as defined in the theorem statement. By substituting the above in (32), we obtain

$$\begin{aligned}
&\sum_{k=1}^{K-1}\sum_{i=K+1-k}^{K}\mathbb{P}\left[c^{[i]}_\alpha - \hat{c}^{[i]}_{n_k,\alpha} \geq \frac{\Delta_{[i]}}{2}\right]\\
&\leq \sum_{k=1}^{K-1} 6k\exp\left(-\frac{(n-K)(1-\alpha)^2 G_{\max}}{8H\overline{\log}K}\right).
\end{aligned} \tag{33}$$

Similarly, we can show that

$$\begin{aligned}
&\sum_{k=1}^{K-1}\sum_{i=K+1-k}^{K}\mathbb{P}\left[\hat{c}^{i^*}_{n_k,\alpha} - c^{i^*}_\alpha - \geq \frac{\Delta_{[i]}}{2}\right]\\
&\leq \sum_{k=1}^{K-1} 6k\exp\left(-\frac{(n-K)(1-\alpha)^2 G_{\max}}{8H\overline{\log}K}\right).
\end{aligned} \tag{34}$$

The main claim follows by substituting (33) and (34) in (31).

(ii) Light-tailed case

The proof follows by a completely parallel argument to the proof for the sub-Gaussian case, while using the following bound in place of (30):

$$\begin{aligned}
&\mathbb{P}\left[|\hat{c}_{n,\alpha} - c_\alpha| > \epsilon\right] \leq\\
&\quad 6\exp\left[-\frac{nG(1-\alpha)^2\min\{\epsilon,\epsilon^2,\delta^2\}}{2}\right],
\end{aligned}$$

where $G = \frac{\min(\eta^2,1)}{\max(2\sigma^2,16,b)}$. □

6. Concluding Remarks

We derived concentration bounds for CVaR estimation, separately considering light-tailed and heavy-tailed distributions. For light-tailed distributions, our concentration bound uses a classical CVaR estimator based on the empirical distribution. For the heavy-tailed case, we employ a truncation based CVaR estimator, and derive a concentration result under a mild bounded-moment assumption. Our concentration bound enjoys exponential decay in the sample size even for heavy-tailed random variables.

We highlighted the applicability of the CVaR concentration result by considering a risk-aware bandit problem. We proposed an adaptation of the successive rejects algorithm to the setting where the goal is to find an arm with the lowest CVaR. Using the CVaR concentration bound, we established error bounds for the proposed algorithm.

References

Artzner, P., Delbaen, F., Eber, J., and Heath, D. Coherent measures of risk. *Mathematical finance*, 9(3):203–228, 1999.

Audibert, J. Y., Bubeck, S., and Munos, R. Best arm identification in multi-armed bandits. In *Conference on Learning Theory*, pp. 41–53, 2010.

Bhat, S. P. and Prashanth, L. A. Concentration of risk measures: A Wasserstein distance approach. In *Advances in Neural Information Processing Systems*, pp. 11739–11748, 2019.

Brown, D. B. Large deviations bounds for estimating conditional value-at-risk. *Operations Research Letters*, 35 (6):722–730, 2007.

Bubeck, S., Cesa-Bianchi, N., and Lugosi, G. Bandits with heavy tail. *IEEE Transactions on Information Theory*, 59(11):7711–7717, 2013.

Chatterjee, R. *Practical methods of financial engineering and risk management: tools for modern financial professionals*. Apress, 2014.

David, Y. and Shimkin, N. Pure exploration for max-quantile bandits. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pp. 556–571. Springer, 2016.

David, Y., Szörényi, B., Ghavamzadeh, M., Mannor, S., and Shimkin, N. PAC Bandits with Risk Constraints. In *International Symposium on Artificial Intelligence and Mathematics*, 2018.

Fournier, N. and Guillin, A. On the rate of convergence in Wasserstein distance of the empirical measure. *Probability Theory and Related Fields*, 162(3-4):707–738, 2015.

Galichet, N., Sebag, M., and Teytaud, O. Exploration vs exploitation vs safety: Risk-aware multi-armed bandits. In *Asian Conference on Machine Learning*, pp. 245–260, 2013.

Gao, S., Frejinger, E., and Ben-Akiva, M. Adaptive route choices in risky traffic networks: A prospect theory approach. *Transportation Research Part C: Emerging Technologies*, 18(5):727–740, 2010.

Grimmett, G. and Stirzaker, D. *Probability and random processes*. Oxford university press, 2001.

Kagrecha, A., Nair, J., and Jagannathan, K. Distribution oblivious, risk-aware algorithms for multi-armed bandits with unbounded rewards. In *Advances in Neural Information Processing Systems*, pp. 11269–11278, 2019.

Kolla, R. K., Prashanth, L., Bhat, S. P., and Jagannathan, K. Concentration bounds for empirical conditional value-at-risk: The unbounded case. *Operations Research Letters*, 47(1):16 – 20, 2019.

Sani, A., Lazaric, A., and Munos, R. Risk-aversion in multi-armed bandits. In *Advances in Neural Information Processing Systems*, pp. 3275–3283, 2012.

Serfling, R. J. *Approximation theorems of mathematical statistics*, volume 162. John Wiley & Sons, 2009.

Thomas, P. and Learned-Miller, E. Concentration inequalities for conditional value at risk. In *International Conference on Machine Learning*, pp. 6225–6233, 2019.

Thompson, W. R. On the likelihood that one unknown probability exceeds another in view of the evidence of two samples. *Biometrika*, 25(3/4):285–294, 1933.

Wainwright, M. J. *High-dimensional statistics: A non-asymptotic viewpoint*, volume 48. Cambridge University Press, 2019.

Wang, Y. and Gao, F. Deviation inequalities for an estimator of the conditional value-at-risk. *Operations Research Letters*, 38(3):236–239, 2010.