

Concentration bounds for empirical conditional value-at-risk: The unbounded case

Ravi Kumar Kolla^{*1}, Prashanth L.A.^{†2}, Sanjay P. Bhat^{‡3} and Krishna Jagannathan^{§1}

¹Department of Electrical Engineering, Indian Institute of Technology Madras, Chennai, Tamilnadu 600036, India

²Department of Computer Science and Engineering, Indian Institute of Technology Madras, Chennai, Tamilnadu 600036, India

³TCS Research, Hyderabad, Telangana 500081, India

Abstract

In several real-world applications involving decision making under uncertainty, the traditional expected value objective may not be suitable, as it may be necessary to control losses in the case of a rare but extreme event. Conditional Value-at-Risk (CVaR) is a popular risk measure for modelling the aforementioned objective. We consider the problem of estimating CVaR from i.i.d. samples of an unbounded random variable, which is either sub-Gaussian or sub-exponential. We derive a novel one-sided concentration bound for a natural sample-based CVaR estimator in this setting. Our bound relies on a concentration result for a quantile-based estimator for Value-at-Risk (VaR), which may be of independent interest.

1 Introduction

In several practical decision problems, the presence of uncertainty complicates the decision making process as decisions typically are required to be taken before the uncertainty is resolved. Traditionally, this difficulty is overcome by averaging the costs (or rewards) over all possible realizations of the uncertainty, and then optimizing the averaged cost thus obtained. However, it has been argued that considering averaged outcomes is not appropriate in situations where low-probability events such as financial crashes and category 4 hurricanes can cause huge costs. The possibility of occurrence of such tail events has led to the introduction of risk measures such as Value-at-Risk (VaR) and Conditional Value-at-Risk (CVaR) for quantification of risk. In financial risk management, the VaR of a risky portfolio at a confidence level α is a loss threshold such that the probability of the loss exceeding the threshold is no greater than $1 - \alpha$. The CVaR of a portfolio at a confidence level α is the expected loss on the portfolio conditioned on the event that the loss exceeds the VaR. Loosely speaking, VaR quantifies the maximum loss that can occur in the absence of a catastrophic tail event, while the CVaR gives the expected loss given the occurrence of such a tail event. CVaR has several desirable properties as a risk measure. In particular, it is a convex, coherent risk measure (see the survey paper [4] and references therein). As a result, CVaR continues to receive increasing attention in operations research, mathematical finance, and decision science for problems involving risk quantification or risk minimization.

*ee12d024@ee.iitm.ac.in

†prashla@cse.iitm.ac.in

‡sanjay.bhat@tcs.com

§krishnaj@ee.iitm.ac.in

In most applications involving uncertainty, the distributions characterizing the underlying uncertain factors are not known, and risk measures such as CVaR have to be estimated from sampled values of the random variable of interest. This is true, for instance, in a multi-armed bandit problem [5, 3] in which pulling an arm leads to a random loss, and one seeks to identify the arm whose loss random variable has the lowest CVaR by observing a sample of outcomes that result from multiple arm pulls. An obvious estimator for the CVaR of a distribution is the sample CVaR of an i.i.d. sample drawn from the distribution. Naturally, one seeks error bounds for the estimator that help to understand the trade-off between accuracy and sample size. Previous results on CVaR estimation either provide asymptotic error bounds for a general r.v. [7], or provide non-asymptotic error bounds that hold with high probability, but under the stringent assumption that the underlying r.v. is bounded [2, 8].

In this paper, we consider the problem of estimating the CVaR of an unbounded, albeit sub-Gaussian or sub-exponential random variable. Sub-Gaussian r.v.s include bounded, Gaussian and any other r.v. whose tail decays as fast as a Gaussian. On the other hand, sub-exponential r.v.s include exponential, Poisson and squared-Gaussian r.v. and are characterized by a tail heavier than Gaussian, resembling that of an exponential distribution. To the best of our knowledge, there are no concentration bounds for CVaR estimator for these two popular classes of unbounded distributions. We believe, imposing a tail decay assumption (sub-Gaussian or sub-exponential) is not restrictive, as such an imposition is common in concentration results for sample mean. Also, the task of CVaR estimation is more challenging in comparison, as it relates to a tail event. We derive a one-sided concentration bound for the empirical CVaR of an i.i.d. sample. Our bound relies on one of the two concentration results (of possibly independent interest) that we provide for a quantile-based estimator for VaR.

The rest of the paper is organized as follows: Section 2 introduces VaR, CVaR and their estimators from i.i.d. samples, Section 3 presents the concentration bounds for VaR and CVaR estimators. Section 4 provides detailed proofs of the concentration results, and Section 5 concludes the paper.

2 Background

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

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

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

where we have used the notation that $[x]^+ = \max(0, x)$ for a real number x . Typical values of α chosen in practice are 0.95 and 0.99. Note that, if X has a continuous and strictly increasing CDF, then $v_\alpha(X)$ is a solution to the following $\mathbb{P}[X \leq \xi] = \alpha$ *i.e.*, $v_\alpha(X) = F^{-1}(\alpha)$. CVaR also admits another form under the following assumption:

(A1) The r.v. X is continuous and has strictly increasing CDF.

If (A1) holds and X has a positive density at v_α , then $c_\alpha(X)$ admits the following equivalent form (cf. [7]):

$$c_\alpha(X) = \mathbb{E}[X | X \geq v_\alpha(X)].$$

Let $\{X_i\}_{i=1}^n$ denote n i.i.d. samples from the distribution of X . Then, the estimates of VaR and CVaR at level α , denoted by $\hat{v}_{n,\alpha}$ and $\hat{c}_{n,\alpha}$, are formed as follows [6]:

$$\hat{v}_{n,\alpha} = \hat{F}_n^{-1}(\alpha) := \inf\{x : \hat{F}_n(x) \geq \alpha\} \quad (3)$$

¹For notational brevity, we omit X from $v_\alpha(X)$ and $c_\alpha(X)$ whenever the r.v. can be understood from the context.

$$\hat{c}_{n,\alpha} = \hat{v}_{n,\alpha} + \frac{1}{n(1-\alpha)} \sum_{i=1}^n (X_i - \hat{v}_{n,\alpha})^+, \quad (4)$$

where $\hat{F}_n(x) = \frac{1}{n} \sum_{i=1}^n \mathbb{I}\{X_i \leq x\}$ is the empirical distribution function of X . Note that, from the order statistics $X_{[1]}, \dots, X_{[n]}$, the empirical VaR can be computed as follows: $\hat{v}_{n,\alpha} = X_{[[n\alpha]]}$.

3 Concentration bounds

In this section, we present four concentration bounds. The first two bounds are for the VaR estimator given in (3) and these bounds do not impose any restrictions on the underlying distribution. The next two concentration results are for the CVaR estimator given in (4), and for these results, we assume that the underlying distribution is either sub-Gaussian or sub-exponential (see Definitions 1–2 below).

In each of the result presented below, the estimates are calculated using n i.i.d. samples $\{X_i\}_{i=1}^n$ drawn from the r.v. X with CDF $F(\cdot)$, and for a given $\alpha \in (0, 1)$.

Proposition 1. (VaR concentration bound) *Let $\alpha \in (0, 1)$, $n \in \mathbb{N}$ and $s \in (0, \frac{1}{2})$. Define $\alpha^- = \alpha - \frac{1}{2n^s}$ and $\alpha^+ = \alpha + \frac{1}{2n^s}$. Further, let $a_n = \hat{F}_n^{-1}(\alpha^-)$ and $b_n = \hat{F}_n^{-1}(\alpha^+)$, where $\hat{F}_n^{-1}(\cdot)$ is defined by (3). Then,*

$$\mathbb{P}[v_\alpha(X) \in [a_n, b_n]] \geq \left(1 - 2 \exp\left(-\frac{n^{1-2s}}{8}\right)\right).$$

Note that, the above concentration bound is free of any distribution dependent parameters.

Proof. See Section 4.1. □

The following result presents another concentration bound for the VaR estimator which will have distribution parameters in the bound. However, unlike Proposition 1, the result presented below is symmetric, and more importantly, bounds the estimation error $|\hat{v}_{n,\alpha} - v_\alpha|$ directly.

Proposition 2 (VaR concentration bound). *Suppose that (A1) holds. For any $\epsilon > 0$, we have*

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

where c is a constant that depends on the value of the density f of the r.v. X in a neighbourhood of VaR.

Proof. See Section 4.2. □

The bound above implies that to estimate the VaR to an accuracy of ϵ , one would require an order $O(1/\epsilon^2)$ number of samples. Notice that no restrictive assumptions on the tail of the underlying distribution are made in arriving at the concentration bounds for VaR in Propositions 1 and 2. However, for establishing concentration bounds for the CVaR, which involves conditioning on a tail event, it is necessary to assume that the distribution is not heavy-tailed. In fact, even for the case of estimating the expected value of a r.v., exponential concentration bounds are available under an assumption that restricts the tail to be light (cf. Chapter 2 of [1]).

In this paper, we present concentration bounds under two popular assumptions on the tail of a r.v. The first restricts the r.v. to be sub-Gaussian, while the second requires the same to be sub-exponential. These two classes of r.v.s include bounded r.v.s and more importantly, several unbounded r.v.s as well. Sub-Gaussian r.v.s include the Gaussian r.v.s as well as several other r.v.s whose moment generating functions do not exceed that of a Gaussian, while sub-exponential r.v.s include heavier tailed r.v.s. These two notions are made precise in the following definitions.

Definition 1. A r.v. X with $\mathbb{E}[X] = \mu < \infty$ is said to be σ -sub-Gaussian if

$$\mathbb{E}[\exp(\lambda X)] \leq \exp\left(\lambda\mu + \frac{\lambda^2\sigma^2}{2}\right), \quad \forall \lambda \in \mathbb{R}.$$

Definition 2. A r.v. X with mean $\mu < \infty$ is said to be (σ, b) -sub-exponential if

$$\mathbb{E}[\exp(\lambda X)] \leq \exp\left(\lambda\mu + \frac{\lambda^2\sigma^2}{2}\right), \quad \forall |\lambda| < \frac{1}{b}.$$

It is worth noting that all sub-Gaussian r.v.s are sub-exponential, but the converse is not true. The following result presents a one-sided concentration bound for the CVaR estimator in (4), for the case when the underlying r.v. is sub-Gaussian.

Proposition 3 (CVaR concentration bound: sub-Gaussian case). Suppose that (A1) holds. Let $\alpha \in (0, 1)$, and X be a σ -sub-Gaussian r.v. with mean μ . Suppose that α is large enough to ensure $(v_\alpha - \mu) > 0$ and the sub-Gaussian parameter σ satisfies $\sigma < \sqrt{\frac{(v_\alpha - \mu)^2}{2 \ln(1/(1-\alpha))}}$. Then, for any $\epsilon > 0$, we have

$$\begin{aligned} \mathbb{P}[\hat{c}_{n,\alpha} - c_\alpha > \epsilon] \leq & \exp\left(-\frac{n\epsilon(1-\alpha)(v_\alpha - \mu)}{2\sigma^2}\right) + 2 \exp(-2nc_1\epsilon^2) + 2 \exp(-2nc_2\epsilon) \\ & + \exp(-2n\epsilon(1-\alpha)^2), \end{aligned} \quad (5)$$

where c_1 and c_2 are constants that depend on the value of the density f of the r.v. X in a neighbourhood of VaR.

Proof. See Section 4.3. □

Suppose that the accuracy ϵ is greater than 1. Then, it is apparent that the dominant terms on the RHS of (5) are those involving an exponential with ϵ . Further, $\hat{c}_{n,\alpha} \leq c_\alpha + \epsilon$ with probability (w.p.) at least $(1 - \delta)$, when the number of samples n is of the order $O\left(\frac{1}{\epsilon} \ln\left(\frac{1}{\delta}\right)\right)$. On other hand, an order $O\left(\frac{1}{\epsilon^2} \ln\left(\frac{1}{\delta}\right)\right)$ number of samples are enough to ensure that $\hat{v}_{n,\alpha} \leq v_\alpha + \epsilon$ w.p. at least $(1 - \delta)$. Hence, CVaR estimation requires more samples in comparison to VaR, when $\epsilon > 1$. In the complementary case, i.e., when $\epsilon < 1$, both VaR and CVaR estimate can be ϵ -accurate w.p. $(1 - \delta)$, if the number of samples is of the order $O\left(\frac{1}{\epsilon^2} \ln\left(\frac{1}{\delta}\right)\right)$.

Next, we analyse the concentration of the CVaR estimator in (4) for the case when the underlying r.v. is sub-exponential.

Proposition 4 (CVaR concentration bound: sub-exponential case). Suppose that (A1) holds. Let X be a (σ, b) -sub-exponential r.v. with mean μ . Suppose that α is large enough to ensure $v_\alpha - \mu > 0$ and the parameter σ satisfies $\sigma < \sqrt{\frac{2 \ln(1-\alpha) + 2(v_\alpha - \mu)m_b}{m_b^2}}$ where $m_b = \min\left\{\frac{v_\alpha - \mu}{\sigma^2}, b'\right\}$ where $b' < 1/b$. Then for any $\epsilon > 0$, we have

$$\mathbb{P}[\hat{c}_{n,\alpha} - c_\alpha > \epsilon] \leq \exp\left(-\frac{n\epsilon(1-\alpha)m_b}{2}\right) + 2 \exp(-2nc_1\epsilon^2) + 2 \exp(-2nc_2\epsilon) + \exp(-2n\epsilon(1-\alpha)^2), \quad (6)$$

where c_1 and c_2 are as in Proposition 3.

From the result above, it is apparent that the rate of CVaR concentration for sub-exponential r.v.s matches that of sub-Gaussian ones.

Proof. See Section 4.4. □

4 Proofs

In this section, we present the proofs of the results presented in Section 3.

4.1 Proof of Proposition 1

Proof. Recall the Dvoretzky-Kiefer-Wolfowitz (DKW) inequality, which provides a finite-sample bound on the distance between the empirical distribution and the true distribution: For any $\epsilon > 0$,

$$\mathbb{P} \left[\sup_{x \in \mathbb{R}} |\hat{F}_n(x) - F(x)| > \epsilon \right] \leq 2e^{-2n\epsilon^2}.$$

Consider the following event

$$A = \left\{ \sup_{x \in [a_n, b_n]} |\hat{F}_n(x) - F(x)| \leq \frac{1}{4n^s} \right\},$$

with $a_n = \hat{F}_n^{-1}(\alpha^-)$ and $b_n = \hat{F}_n^{-1}(\alpha^+)$ as defined in the theorem statement. By the DKW inequality, we have

$$\mathbb{P}[A] \geq 1 - 2 \exp\left(\frac{-2n}{16n^{2s}}\right) = 1 - 2 \exp\left(\frac{-n^{1-2s}}{8}\right). \quad (7)$$

On the event A , we have

$$\begin{aligned} F(a_n) &\stackrel{(a)}{=} \lim_{a \uparrow a_n} F(a) \stackrel{(b)}{\leq} \lim_{a \uparrow a_n} \left(\hat{F}_n(a) + \frac{1}{4n^s} \right) \stackrel{(c)}{\leq} \alpha^- + \frac{1}{4n^s} = \alpha - \frac{1}{4n^s} < \alpha, \text{ and} \\ F(b_n) &\geq \alpha^+ - \frac{1}{4n^s} = \alpha + \frac{1}{4n^s} > \alpha, \end{aligned}$$

where (a) follows from continuity of F , (b) follows from the definition of the event A , and (c) follows from the definition of a_n . Thus, $v_\alpha \in [a_n, b_n]$ and the main claim follows from the lower bound on $\mathbb{P}[A]$ in (7). \square

4.2 Proof of Proposition 2

Proof.

$$\begin{aligned} \mathbb{P}[|\hat{v}_{n,\alpha} - v_\alpha| \geq \epsilon] &= \mathbb{P}[\hat{v}_{n,\alpha} \geq v_\alpha + \epsilon] + \mathbb{P}[\hat{v}_{n,\alpha} \leq v_\alpha - \epsilon] \\ &\leq \mathbb{P}\left[\hat{F}_n(v_\alpha + \epsilon) \leq \alpha\right] + \mathbb{P}\left[\hat{F}_n(v_\alpha - \epsilon) \geq \alpha\right] \\ &= \mathbb{P}\left[F(v_\alpha + \epsilon) - \hat{F}_n(v_\alpha + \epsilon) \geq F(v_\alpha + \epsilon) - \alpha\right] + \mathbb{P}\left[\hat{F}_n(v_\alpha - \epsilon) - F(v_\alpha - \epsilon) \geq \alpha - F(v_\alpha - \epsilon)\right] \\ &\stackrel{(a)}{\leq} \exp\left(-2n(F(v_\alpha + \epsilon) - F(v_\alpha))^2\right) + \exp\left(-2n(F(v_\alpha) - F(v_\alpha - \epsilon))^2\right) \\ &\leq 2 \exp\left(-2n\delta_\epsilon^2\right), \end{aligned}$$

where (a) is due to the DKW inequality, and $\delta_\epsilon = \min\{F(v_\alpha + \epsilon) - F(v_\alpha), F(v_\alpha) - F(v_\alpha - \epsilon)\}$. Given that the density exists, we have

$$F(v_\alpha + \eta_1) - F(v_\alpha - \eta_2) = f(\bar{v})(\eta_1 + \eta_2),$$

for some $\bar{v} \in [v_\alpha - \eta_2, v_\alpha + \eta_1]$. Using the identity above for the two expressions inside δ_ϵ , we obtain

$$\delta_\epsilon = \min(f(\bar{v}_1), f(\bar{v}_2)) \times \epsilon.$$

for some $\bar{v}_1 \in [v_\alpha, v_\alpha + \epsilon]$ and $\bar{v}_2 \in [v_\alpha - \epsilon, v_\alpha]$. The claim follows. \square

4.3 Proof of Proposition 3

We first prove a more general result without restricting the sub-Gaussian parameter σ and the main claim in Proposition 3 follows in a straightforward fashion.

Proposition 5 (General CVaR concentration bound: sub-Gaussian case). *Assume (A1). Let X be a σ -sub-Gaussian with mean μ . Suppose that $(v_\alpha - \mu) > 0$. Then, for any $\epsilon > 0$, we have*

$$\begin{aligned} \mathbb{P}[\hat{c}_{n,\alpha} - c_\alpha > \epsilon] &\leq \exp\left(-\frac{n\epsilon(1-\alpha)(v_\alpha - \mu)}{2\sigma^2}\right) \left[\alpha + \exp\left(-\frac{(v_\alpha - \mu)^2}{2\sigma^2}\right)\right]^n + 2\exp(-2n\delta_{\epsilon_1}^2) \\ &\quad + 2\exp(-2n\delta_{\epsilon_2}^2) + \exp(-2n\epsilon(1-\alpha)^2), \end{aligned} \quad (8)$$

where $\delta_{\epsilon_1} = \min\{F(v_\alpha + \frac{n(1-\alpha)\epsilon}{8}) - F(v_\alpha), F(v_\alpha) - F(v_\alpha - \frac{n(1-\alpha)\epsilon}{8})\}$ and $\delta_{\epsilon_2} = \min\{F(v_\alpha + \sqrt{\epsilon}/4) - F(v_\alpha), F(v_\alpha) - F(v_\alpha - \sqrt{\epsilon}/4)\}$.

Proof. First, we bound the estimate $\hat{c}_{n,\alpha}$. Notice that

$$\begin{aligned} \hat{c}_{n,\alpha} &= \hat{v}_{n,\alpha} + \frac{1}{n(1-\alpha)} \sum_{i=1}^n (X_i - \hat{v}_{n,\alpha})^+ \\ &= v_\alpha + \frac{1}{n(1-\alpha)} \sum_{i=1}^n (X_i - v_\alpha)^+ + (\hat{v}_{n,\alpha} - v_\alpha) + \frac{1}{n(1-\alpha)} \sum_{i=1}^n [(X_i - \hat{v}_{n,\alpha})^+ - (X_i - v_\alpha)^+] \end{aligned} \quad (9)$$

The last term on the RHS of (9) can be re-written as follows:

$$\begin{aligned} \frac{1}{n(1-\alpha)} \sum_{i=1}^n [(X_i - \hat{v}_{n,\alpha})^+ - (X_i - v_\alpha)^+] &= \frac{1}{n(1-\alpha)} \sum_{i=1}^n [(v_\alpha - \hat{v}_{n,\alpha})\mathbb{I}\{X_i \geq \hat{v}_{n,\alpha}\}] \\ &\quad + \frac{1}{n(1-\alpha)} \sum_{i=1}^n (X_i - v_\alpha) [\mathbb{I}\{X_i \geq \hat{v}_{n,\alpha}\} - \mathbb{I}\{X_i \geq v_\alpha\}], \end{aligned} \quad (10)$$

and

$$\begin{aligned} \frac{1}{n(1-\alpha)} \sum_{i=1}^n [(v_\alpha - \hat{v}_{n,\alpha})\mathbb{I}\{X_i \geq \hat{v}_{n,\alpha}\}] &= \frac{\hat{v}_{n,\alpha} - v_\alpha}{1-\alpha} \left[\hat{F}_n(\hat{v}_{n,\alpha}) - 1 \right] - \frac{1}{n} \sum_{i=1}^n \mathbb{I}\{X_i = \hat{v}_{n,\alpha}\} \\ &= \frac{\hat{v}_{n,\alpha} - v_\alpha}{1-\alpha} \left[\hat{F}_n(\hat{v}_{n,\alpha}) - 1 \right] \text{ w.p. } 1. \end{aligned} \quad (11)$$

The last equality above uses the fact that $\frac{1}{n} \sum_{i=1}^n \mathbb{I}\{X_i = \hat{v}_{n,\alpha}\}$ takes values zero w.p. 1, since X_i is continuous, for each i .

Combining (9), (10) and (11), we obtain

$$\begin{aligned} \hat{c}_{n,\alpha} &= v_\alpha + \frac{1}{n} \sum_{i=1}^n \frac{(X_i - v_\alpha)^+}{1-\alpha} + (\hat{v}_{n,\alpha} - v_\alpha) + \frac{\hat{v}_{n,\alpha} - v_\alpha}{1-\alpha} \left[\hat{F}_n(\hat{v}_{n,\alpha}) - 1 \right] \\ &\quad + \frac{1}{n} \sum_{i=1}^n \frac{X_i - v_\alpha}{1-\alpha} [\mathbb{I}\{X_i \geq \hat{v}_{n,\alpha}\} - \mathbb{I}\{X_i \geq v_\alpha\}] \\ &= v_\alpha + \frac{1}{n(1-\alpha)} \sum_{i=1}^n (X_i - v_\alpha)^+ + B_n, \end{aligned} \quad (12)$$

where

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

From (12), we have

$$\begin{aligned} \hat{c}_{n,\alpha} - c_\alpha &= \left[v_\alpha + \frac{1}{n} \sum_{i=1}^n \frac{(X_i - v_\alpha)^+}{1 - \alpha} - c_\alpha \right] + B_n \\ &= \frac{1}{1 - \alpha} \left[\frac{1}{n} \sum_{i=1}^n (X_i - v_\alpha) \mathbb{I}\{X_i \geq v_\alpha\} - \mathbb{E}[(X - v_\alpha) \mathbb{I}\{X \geq v_\alpha\}] \right] + B_n. \end{aligned} \quad (13)$$

For notational convenience, let

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

It is easy to see that Y_i 's are i.i.d., non-negative, and $\mathbb{E}[Y_i] = (1 - \alpha)(c_\alpha - v_\alpha)$, $\forall i$. We now proceed to bound $\mathbb{P}[\hat{c}_{n,\alpha} - c_\alpha > \epsilon]$, using (13), as follows:

$$\mathbb{P}[\hat{c}_{n,\alpha} - c_\alpha > \epsilon] = \mathbb{P}[A_n + B_n > \epsilon] \leq \mathbb{P}[A_n > \epsilon/2] + \mathbb{P}[B_n > \epsilon/2]. \quad (14)$$

For handling $\mathbb{P}[A_n > \epsilon/2]$, we bound the moment generating function of r.v. Y_i as follows:

$$\begin{aligned} \mathbb{P}[e^{\lambda Y_i/n}] &= \mathbb{E}\left[e^{\frac{\lambda}{n}(X_i - v_\alpha) \mathbb{I}\{X_i \geq v_\alpha\}}\right] = \int_{-\infty}^{v_\alpha} f_X(x) dx + \int_{v_\alpha}^{\infty} e^{\frac{\lambda}{n}(x - v_\alpha)} f_X(x) dx \\ &\leq F_X(v_\alpha) + e^{-\frac{\lambda}{n}v_\alpha} \int_{-\infty}^{\infty} e^{\frac{\lambda}{n}x} f_X(x) dx \\ &\stackrel{(a)}{\leq} \alpha + e^{-\frac{\lambda}{n}v_\alpha + \frac{\lambda}{n}\mu + \frac{\lambda^2\sigma^2}{2n^2}}, \end{aligned} \quad (15)$$

where (a) is due to the sub-Gaussianity of X_i . Thus,

$$\mathbb{P}\left[A_n > \frac{\epsilon}{2}\right] = \mathbb{P}\left[\frac{1}{n} \sum_{i=1}^n Y_i > \frac{(1 - \alpha)\epsilon}{2} + \mathbb{E}[Y]\right] \stackrel{(b)}{\leq} \frac{\prod_{i=1}^n \mathbb{E}[e^{\lambda Y_i/n}]}{e^{\lambda((1 - \alpha)\epsilon/2 + \mathbb{E}[Y])}} \stackrel{(c)}{\leq} \frac{\left[\alpha + e^{-\frac{\lambda}{n}v_\alpha + \frac{\lambda}{n}\mu + \frac{\lambda^2\sigma^2}{2n^2}}\right]^n}{e^{\lambda((1 - \alpha)\epsilon/2 + \mathbb{E}[Y])}}, \quad (16)$$

where (b) uses Markov's inequality and (c) follows from (15). Notice that (16) holds for any $\lambda > 0$. However, for the bound on the RHS above to be meaningful, we require that $\alpha + e^{-\frac{\lambda}{n}v_\alpha + \frac{\lambda}{n}\mu + \frac{\lambda^2\sigma^2}{2n^2}} < 1$. Now, maximizing $\frac{\lambda^2\sigma^2}{2n^2} - \frac{\lambda(v_\alpha - \mu)}{n}$ over λ , we obtain $\lambda_* = \frac{n(v_\alpha - \mu)}{\sigma^2}$. Substituting the value of λ_* in (16), we obtain

$$\mathbb{P}[A_n > \epsilon/2] \leq \exp\left(-\frac{n(v_\alpha - \mu)}{\sigma^2} \left[\frac{(1 - \alpha)\epsilon}{2} + \mathbb{E}[Y]\right]\right) \left[\alpha + \exp\left(-\frac{(v_\alpha - \mu)^2}{2\sigma^2}\right)\right]^n. \quad (17)$$

For handling the $\mathbb{P}[B_n > \frac{\epsilon}{2}]$ term in (14), we bound $|B_n|$ as follows. Using the inequality above, we bound $|B_n|$ as follows:

$$\begin{aligned}
|B_n| &\leq \frac{|v_\alpha - \hat{v}_{n,\alpha}|}{1-\alpha} |\alpha - \hat{F}_n(\hat{v}_{n,\alpha})| + \frac{|v_\alpha - \hat{v}_{n,\alpha}|}{1-\alpha} |\hat{F}_n(v_\alpha) - \hat{F}_n(\hat{v}_{n,\alpha})| \\
&= \frac{|v_\alpha - \hat{v}_{n,\alpha}|}{1-\alpha} \left[|F(v_\alpha) - \hat{F}_n(\hat{v}_{n,\alpha})| + |\hat{F}_n(v_\alpha) - F(v_\alpha) - \hat{F}_n(\hat{v}_{n,\alpha}) + F(v_\alpha)| \right] \\
&\leq \frac{|v_\alpha - \hat{v}_{n,\alpha}|}{1-\alpha} \left[2|\hat{F}_n(\hat{v}_{n,\alpha}) - F(v_\alpha)| + |\hat{F}_n(v_\alpha) - F(v_\alpha)| \right]. \tag{18}
\end{aligned}$$

The first inequality above uses the following fact:

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

Using (18), we have

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

It is easy to see that $|\hat{F}_n(\hat{v}_{n,\alpha}) - F(v_\alpha)| \leq \frac{1}{n}$. Hence,

$$\begin{aligned}
\mathbb{P}\left[B_n > \frac{\epsilon}{2}\right] &\leq \mathbb{P}\left[|B_n| > \frac{\epsilon}{2}\right] \leq \mathbb{P}\left[\frac{|\hat{v}_{n,\alpha} - v_\alpha|}{1-\alpha} \left(\frac{2}{n} + |\hat{F}_n(v_\alpha) - F(v_\alpha)|\right) > \frac{\epsilon}{2}\right] \\
&\leq \mathbb{P}\left[\frac{2}{n(1-\alpha)} |\hat{v}_{n,\alpha} - v_\alpha| > \frac{\epsilon}{4}\right] + \mathbb{P}\left[\frac{1}{1-\alpha} |\hat{v}_{n,\alpha} - v_\alpha| |\hat{F}_n(v_\alpha) - F(v_\alpha)| > \frac{\epsilon}{4}\right].
\end{aligned}$$

Let $D_n = \{|\hat{v}_{n,\alpha} - v_\alpha| \leq \frac{\sqrt{\epsilon}}{4}\}$. Then, we have

$$\begin{aligned}
\mathbb{P}\left[\frac{|\hat{v}_{n,\alpha} - v_\alpha| |\hat{F}_n(v_\alpha) - F(v_\alpha)|}{1-\alpha} > \frac{\epsilon}{4}\right] &\leq \mathbb{P}\left[\frac{|\hat{v}_{n,\alpha} - v_\alpha| |\hat{F}_n(v_\alpha) - F(v_\alpha)|}{1-\alpha} > \frac{\epsilon}{4}, D_n\right] + \mathbb{P}[D_n^c] \\
&\leq \mathbb{P}\left[|\hat{F}_n(v_\alpha) - F(v_\alpha)| > \sqrt{\epsilon}(1-\alpha)\right] + \mathbb{P}[D_n^c] \\
&\stackrel{(d)}{\leq} \exp(-2n\epsilon(1-\alpha)^2) + \mathbb{P}[D_n^c],
\end{aligned}$$

where (d) is due to the DKW inequality. Therefore,

$$\mathbb{P}[B_n > \epsilon/2] \leq \mathbb{P}[|\hat{v}_{n,\alpha} - v_\alpha| > n(1-\alpha)\epsilon/8] + \mathbb{P}[|\hat{v}_{n,\alpha} - v_\alpha| > \sqrt{\epsilon}/4] + \exp(-2n\epsilon(1-\alpha)^2). \tag{19}$$

Using (14), (17) and (19), we obtain

$$\begin{aligned}
\mathbb{P}[\hat{c}_{n,\alpha} - c_\alpha > \epsilon] &\leq \exp\left(-\frac{n(v_\alpha - \mu)}{\sigma^2} \left[\frac{(1-\alpha)\epsilon}{2} + \mathbb{E}[Y]\right]\right) \left[\alpha + \exp\left(-\frac{(v_\alpha - \mu)^2}{2\sigma^2}\right)\right]^n \\
&\quad + \mathbb{P}\left[|\hat{v}_{n,\alpha} - v_\alpha| > \frac{n(1-\alpha)\epsilon}{8}\right] + \mathbb{P}\left[|\hat{v}_{n,\alpha} - v_\alpha| \geq \frac{\sqrt{\epsilon}}{4}\right] + \exp(-2n\epsilon(1-\alpha)^2) \\
&\stackrel{(a)}{\leq} \exp\left(-\frac{n\epsilon(1-\alpha)(v_\alpha - \mu)}{2\sigma^2}\right) \left[\alpha + \exp\left(-\frac{(v_\alpha - \mu)^2}{2\sigma^2}\right)\right]^n + 2\exp(-2n\delta_{\epsilon_1}^2) \\
&\quad + 2\exp(-2n\delta_{\epsilon_2}^2) + \exp(-2n\epsilon(1-\alpha)^2),
\end{aligned}$$

where $\delta_{\epsilon_1}, \delta_{\epsilon_2}$ are as defined in the statement of the proposition, and (a) is due to Proposition 2 and the fact that $\mathbb{E}[Y] \geq 0$. \square

Proof of Proposition 3

Proof. For $\sigma < \sqrt{\frac{(v_\alpha - \mu)^2}{2 \ln(1/(1-\alpha))}}$, we note that $\alpha + \exp\left(-\frac{(v_\alpha - \mu)^2}{2\sigma^2}\right) < 1$. Given that the density exists, we have

$$F(v_\alpha + \eta_1) - F(v_\alpha - \eta_2) = f(\bar{v})(\eta_1 + \eta_2),$$

for some $\bar{v} \in [v_\alpha - \eta_2, v_\alpha + \eta_1]$. Using the identity above for the two expressions inside δ_{ϵ_1} , we obtain

$$\delta_{\epsilon_1} = \min(f(\bar{v}_1), f(\bar{v}_2)) \times \frac{n(1-\alpha)\epsilon}{8}.$$

for some $\bar{v}_1 \in [v_\alpha, v_\alpha + \frac{n(1-\alpha)\epsilon}{8}]$ and $\bar{v}_2 \in [v_\alpha - \frac{n(1-\alpha)\epsilon}{8}, v_\alpha]$. Along similar lines, it is easy to infer that

$$\delta_{\epsilon_2} = \min(f(\bar{v}_3), f(\bar{v}_4)) \times \frac{\sqrt{\epsilon}}{4},$$

for some $\bar{v}_1 \in [v_\alpha, v_\alpha + \frac{\sqrt{\epsilon}}{4}]$ and $\bar{v}_2 \in [v_\alpha - \frac{\sqrt{\epsilon}}{4}, v_\alpha]$. The claim follows. \square

4.4 Proof of Proposition 4

Proof. Observe that, in the proof of Proposition 5, sub-Gaussianity is used in bounding $\mathbb{P}[A_n \geq \epsilon/2]$ following (14) there. Here, we bound the same using sub-exponential assumption.

Recall that $A_n = \frac{1}{1-\alpha} \left(\frac{1}{n} \sum_{i=1}^n Y_i - \mathbb{E}[Y] \right)$, where $Y_i = (X_i - v_\alpha) \mathbb{I}\{X_i \geq v_\alpha\}$ and $Y = (X - v_\alpha) \mathbb{I}\{X \geq v_\alpha\}$. Starting as in the derivation of (15),

$$\begin{aligned} \mathbb{E} \left[e^{\lambda Y_i} \right] &= \mathbb{E} \left[e^{\lambda(X_i - v_\alpha) \mathbb{I}\{X_i \geq v_\alpha\}} \right] \leq F_X(v_\alpha) + e^{-\lambda v_\alpha} \int_{-\infty}^{\infty} e^{\lambda x} f_X(x) dx \\ &\leq \alpha + e^{-\lambda v_\alpha + \lambda \mu + \frac{\lambda^2 \sigma^2}{2}}, \quad \forall |\lambda| < \frac{1}{b}, \end{aligned} \quad (20)$$

where the last inequality uses the fact that X_i is sub-exponential, for each i . Thus,

$$\begin{aligned} \mathbb{P} \left[A_n > \frac{\epsilon}{2} \right] &= \mathbb{P} \left[\sum_{i=1}^n Y_i > n \left(\frac{(1-\alpha)\epsilon}{2} + \mathbb{E}[Y] \right) \right] \stackrel{(b)}{\leq} \frac{\prod_{i=1}^n \mathbb{E} \left[e^{\lambda Y_i} \right]}{e^{\lambda(n(1-\alpha)\epsilon/2 + \mathbb{E}[Y])}} \quad \forall \lambda > 0 \\ &\stackrel{(c)}{\leq} \frac{\left[\alpha + e^{-\lambda v_\alpha + \lambda \mu + \frac{\lambda^2 \sigma^2}{2}} \right]^n}{e^{n\lambda((1-\alpha)\epsilon/2 + \mathbb{E}[Y])}}, \quad \forall 0 < \lambda < \frac{1}{b}, \end{aligned} \quad (21)$$

where (b) follows from Markov's inequality and (c) uses (20). As argued in the proof of Proposition 5, choosing $\lambda = \frac{v_\alpha - \mu}{\sigma^2}$ maximizes the term in the numerator on the RHS above, while ensuring the same is less than 1. However, sub-exponential assumption requires that $0 < \lambda < 1/b$. Hence, we choose $\lambda = \{m_b = \min(\frac{v_\alpha - \mu}{\sigma^2}, b')\}$, where $b' < 1/b$. Using the above and the fact that $\mathbb{E}[Y] \geq 0$, we obtain that

$$\mathbb{P}[A_n > \epsilon/2] \leq \exp\left(-\frac{n\epsilon(1-\alpha)m_b}{2}\right) \left[\alpha + \exp\left(m_b(\mu - v_\alpha) + \frac{m_b^2 \sigma^2}{2}\right) \right]^n. \quad (22)$$

For $\sigma < \sqrt{\frac{2 \ln(1-\alpha) + 2(v_\alpha - \mu)m_b}{m_b^2}}$, we obtain

$$\mathbb{P}[A_n > \epsilon/2] \leq \exp\left(-\frac{n\epsilon(1-\alpha)m_b}{2}\right)$$

The rest of the proof follows in a similar manner as that of Proposition 3. \square

Remark 1. *The proof technique used to establish Proposition 3 cannot be employed to establish a lower deviations bound for the CVaR estimator in (4). In particular, using the notation from the proof of Proposition 3, we have*

$$\begin{aligned}\mathbb{P}[\hat{c}_{n,\alpha} - c_\alpha \leq -\epsilon] &= \mathbb{P}[A_n + B_n \leq -\epsilon] \leq \mathbb{P}[A_n \leq -\epsilon/2] + \mathbb{P}[B_n \leq -\epsilon/2] \\ &\leq \mathbb{P}[A_n \leq -\epsilon/2] + \mathbb{P}[|B_n| \geq \epsilon/2]\end{aligned}$$

While $\mathbb{P}[|B_n| \leq -\epsilon/2]$ can be bounded as before (see (19)), handling $\mathbb{P}[A_n \leq -\epsilon/2]$ is challenging. For instance, mimicking the steps leading to (14), we have

$$\begin{aligned}\mathbb{P}[A_n \leq -\epsilon/2] &= \mathbb{P}\left[\frac{1}{n(1-\alpha)} \sum_{i=1}^n (Y_i - \mathbb{E}[Y]) \leq -\epsilon/2\right] \\ &= \mathbb{P}\left[\frac{1}{n} \sum_{i=1}^n Y_i \leq \mathbb{E}[Y] - (1-\alpha)\epsilon/2\right] \\ &\stackrel{(a)}{=} \mathbb{P}\left[\exp\left(-\frac{\lambda}{n} \sum_{i=1}^n Y_i\right) \geq \exp\left(-\lambda\left(\mathbb{E}[Y] - \frac{(1-\alpha)\epsilon}{2}\right)\right)\right] \\ &\stackrel{(b)}{\leq} \frac{\prod_{i=1}^n \mathbb{E}\left[e^{\frac{-\lambda Y_i}{n}}\right]}{e^{-\lambda(\mathbb{E}[Y] - (1-\alpha)\epsilon/2)}} \\ &\stackrel{(c)}{\leq} \frac{\left[\alpha + e^{\frac{\lambda}{n}v_\alpha - \frac{\lambda}{n}\mu + \frac{\lambda^2\sigma^2}{2n^2}}\right]^n}{\exp(-\lambda(-(1-\alpha)\epsilon/2 + \mathbb{E}[Y]))},\end{aligned}\tag{23}$$

where (a) holds for any $\lambda > 0$, (b) uses Markov's inequality and (c) is due to (15). Deriving a meaningful upper bound on $\mathbb{P}[A_n \leq -\epsilon/2]$ using (23) is difficult, since the numerator there cannot be controlled. This is because $(v_\alpha - \mu) > 0$ and λ is constrained to be positive.

5 Conclusions

We derived a one-sided concentration bound for a natural sample-based CVaR estimator, when the underlying distribution is unbounded, albeit sub-Gaussian or sub-exponential. We also derived concentration bounds for a quantile-based estimator for VaR, and this may be of independent interest. We believe our concentration bounds for natural estimates VaR and CVaR are interesting not only from a statistical viewpoint, but also for solving sequential decision making problems under uncertainty, for e.g., in the multi-armed bandit framework [3].

As future work, it would be interesting to derive a concentration result that bounds the lower deviations of the CVaR estimator. An orthogonal direction is to derive concentration result for CVaR estimators that incorporate importance sampling.

References

- [1] Stéphane Boucheron, Gábor Lugosi, and Pascal Massart. *Concentration inequalities: A nonasymptotic theory of independence*. Oxford university press, 2013.
- [2] D. B. Brown. Large deviations bounds for estimating conditional value-at-risk. *Operations Research Letters*, 35(6):722–730, 2007.

- [3] S. Bubeck and N. Cesa-Bianchi. Regret analysis of stochastic and nonstochastic multi-armed bandit problems. *Foundations and Trends® in Machine Learning*, 5(1):1–122, 2012.
- [4] P. Krokmal, M. Zabaranin, and S. Uryasev. Modeling and optimization of risk. *Surveys in Operations Research and Management Science*, 16:49–66, 2011.
- [5] Herbert Robbins. Some aspects of the sequential design of experiments. In *Bulletin of the American Mathematical Society*, volume 55. 1952.
- [6] Robert Serfling. Approximation theorems of mathematical statistics. 1980.
- [7] L. Sun and L. J. Hong. Asymptotic representations for importance-sampling estimators of value-at-risk and conditional value-at-risk. *Operations Research Letters*, 38(4):246–251, 2010.
- [8] Y. Wang and F. Gao. Deviation inequalities for an estimator of the conditional value-at-risk. *Operations Research Letters*, 38(3):236–239, 2010.