EE5111: Estimation Theory

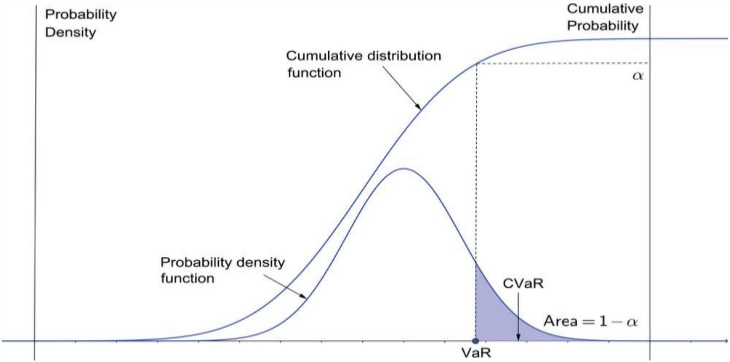
Final Project Presentation

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

Jay Shah (EE18B158) Aravint Annamalai (EE18B125) Sayan Mitra (EE18B156)
Rohith D (EE18B148) Siddharth D P (EE18B072)

Setting the background...

Need for VaR (Value at Risk) and CVaR (Conditional Value at Risk) Metrics:



- When portfolios are described by their mean and variance, the assumption of normal distribution is natural. However most financial and real world variables have fatter tails than normal- implying extreme events are more likely to occur.

- **VaR and CVaR** are two risk measures focusing on adverse events.

Introduction

VaR at level $\alpha \in (0, 1)$

- Conveys the maximum loss incurred by the portfolio with a confidence of α.

$\rightarrow$ Loss will exceed the VaR at level α with probability $1 - \alpha$

CVaR at level $\alpha \in (0, 1)$

- Captures the expected loss incurred by the portfolio, given that the losses exceed VaR at level α, i.e. after the VaR threshold has been breached

Why CVaR?

Coherent risk measure:

- Monotonic
- Translation invariant
- Sub-additive
- Positive homogeneous.

On the other hand VaR isn't coherent as it doesn't satisfy the sub-additive property

Preliminaries

VaR $v_\alpha(X)$ is formally defined as

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

CVaR $c_\alpha(X)$ is formally defined as

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

When we assume the following condition:

(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 $x \in \left[v_\alpha - \frac{\delta}{2}, v_\alpha + \frac{\delta}{2}\right]$ which can be simplified as:

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

Abstract

What exactly does the paper do?

Deriving concentration bounds for CVaR estimates.

Novelty of Result:

- Concentration bounds exhibit **exponential** decay.
- Tighter than those available in the literature for the below distribution classes.

Distribution	Estimator Type
Sub-Gaussian	Classical CVaR estimator (empirical distribution constructed from the samples)
Light tailed distribution	Classical CVaR estimator
Heavy tailed distribution	Truncation based estimator assuming mild bounded conditions i.e p^{th} moment of the distribution is assumed to exist, for some $p > 1$.)

Author’s Contributions

- Exponential decay in the sample size, for heavy tailed as well as light-tailed distributions.
- Bounds are **order optimal**, and the dependence on the number of samples as well as the **accuracy cannot be improved.**

Application of CVaR optimization problem in a multi-armed bandit setting.

- Finding the arm with the lowest CVaR value in a fixed budget setting in a **K-armed stochastic bandit** problem.
- Proposal of an algorithm for the best CVaR arm identification that is inspired by successive-rejects (SR) (Audibert et al., 2010)
- Establishing an upper bound on the probability of incorrect arm identification for the SR-based algorithm for CVaR.

Literature Review

Description	Contribution
For bounded distributions, estimate has been shown to exponentially concentrate around the true CVaR	Brown, 2007; Wang & Gao, 2010
• Obtaining a concentration result for VaR is does not require assumptions on the tail of the distribution. • One-sided CVaR concentration bound	Kolla et al, 2019
Exponentially decaying concentration bound for the case of sub-Gaussian distributions	Bhat and Prashanth
Power law decay using the Wasserstein distance between the empirical and the true distribution	Fournier and Guillin
Mean variance optimization in a regret minimization framework	Hani et al
In the best arm identification setting, VaR-based criteria studied by	David et al
CVaR-based risk criterion in a best arm identification problem under a distribution oblivious setting	Kagreecha et all
CVaR-based criteria explored in a bandit context	Galichet et al

Example: Concentration bounds for Gaussian

Gaussian RV with mean $\boldsymbol{\mu}$ and variance $\boldsymbol{\sigma^2}$.

$$\text{Let } Q(\xi) = \frac{1}{\sqrt{2\pi}} \int_{\xi}^{\infty} e^{-\frac{x^2}{2}} dx$$

Clearly, $Q(-x) = 1 - Q(x)$ and $F_X(\xi) = Q\left(\frac{\mu-\xi}{\sigma}\right)$. Hence, $v_\alpha(X)$ is the solution to $Q\left(\frac{\mu-\xi}{\sigma}\right) = \alpha$

$$\Rightarrow v_\alpha(X) = \mu - \sigma Q^{-1}(\alpha) \;\; \ldots\ldots (6)$$

Acerbi's formula states that the CVaR of a random variable X, which represents loss, at the confidence level α can be expressed as

$$CVaR_{\alpha(X)} = \frac{1}{1-\alpha} \int_{\alpha}^{1} VaR_\beta(X)\, d\beta$$

Using **Acerbi's formula**, CVaR $\boldsymbol{c_\alpha(X) = \mu\left(\frac{\alpha}{1-\alpha}\right) + \sigma c_\alpha(Z)}$, where $Z \sim N(0,1)$.

Estimates of μ and σ are **sufficient** to estimate $c_\alpha(X)$ for the Gaussian case.

$\Rightarrow \hat{c}_n = \hat{\mu}\left(\frac{\alpha}{1-\alpha}\right) + \hat{\sigma} c_\alpha(Z)$ serves as a proxy for $c_\alpha(X)$.

Inequality Bounds

Hoeffding's Inequality:

Let x_i be **independent bounded random variables** such that the random variable x_i falls in the interval $[p_i, q_i]$. Then for any a > 0 we have:

$$P\left(\sum_{i=1}^{n} x_i - E\left(\sum_{i=1}^{n} x_i\right) \geq a\right) \leq e^{-\frac{2a^2}{\sum_{i=1}^{n}(q_i - p_i)^2}}$$

Bernstein's Inequality:

Let $x_1, x_2, \dots, x_n$ be **independent bounded random variables** such that $\boldsymbol{E(x_i) = 0}$ and $|x_i| \leq \varsigma$ with probability 1 and let $\sigma^2 = \frac{1}{n}\sum_{i=1}^{n} var(x_i)$. Then for any a > 0 we have:

$$P\left(\frac{1}{n}\sum_{i=1}^{n} x_i \geq \epsilon\right) \leq e^{-\frac{n\epsilon^2}{2\sigma^2 + 2\varsigma\epsilon/3}}$$

- Using the above inequalities, $\hat{\boldsymbol{c}}_{\boldsymbol{n},\boldsymbol{\alpha}}$ concentrates **exponentially around $\mathbf{c}_{\boldsymbol{\alpha}}(\boldsymbol{X})$.**

Basic Definitions

In the majority of the work, we have only considered the sub-gaussian and light-tailed cases, and only towards the end have we considered the heavy-tailed distribution.

Sub-Gaussian - A random variable X is said to be **σ-sub-Gaussian** for some σ > 0 if

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

$$\mathbb{P}[|X-\mu| \geq t] \leq 2\, e^{-\frac{t^2}{2\sigma^2}} \qquad \text{for all } t \in \mathbb{R}.$$

Light-tailed - A random variable X is said to be **light-tailed** if there exists non-negative parameters σ and b such that

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

Empirical Estimation of VaR and CVaR

Let $\{X\}_{i=1}^{n}$ be n samples drawn out of an i.i.d process.
Now, let $\{X\}_{i=1}^{n}$ be the order statistic of the above samples, i.e,
$X_{[1]} \leq X_{[2]} \ldots \leq X_{[n]}$ holds.

Let $\widehat{F}_n(x)$be the empirical cumulative distribution function calculated as -

$$\hat{F}_n(x) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{I}\{X_i \leq x\}$$

Let estimates of VaR and CVaR at level α be calculated using the n samples above. Then the estimates for the VaR and CVaR are as follows -

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

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

CVaR estimation for sub-Gaussian and light-tailed distributions

Theorem - Assuming (**A1**),

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

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

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

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

Proof (i)

Proof: (i) Sub-Gaussian case

Split $\hat{c}_{n,\alpha} - c_\alpha = I_n + e_n$ .....(22)

Letting $\mu_\alpha^+ = \mathbb{E}\left[(X - v_\alpha)^+\right]$, we have

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

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

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 -

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

The claim for the sub-Gaussian bound given in the previous slide follows by using -

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

The bound on the e_n term requires concentration, and takes the following form:

$$\mathrm{P}[e_n > \epsilon] \leq 4e^{-\frac{n(1-\alpha^2)\eta^2 \min(\epsilon^2, \delta^2)}{16}} \text{ ... ... } (12)$$

Proof (ii)

Proof: (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 $\mu_\alpha^+ = \mathbb{E}[(X - v_\alpha)^+]$, we have -

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

c_1, c_2 and c_4 are distribution-dependent constants

Comparing

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

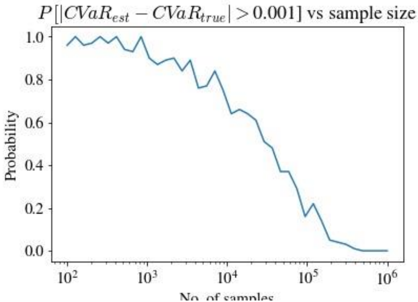
and

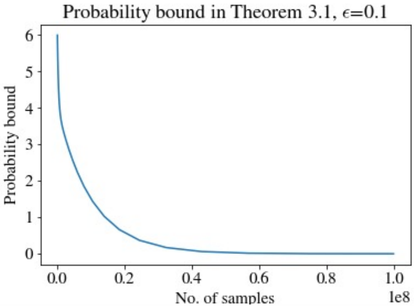
$$\exp\left(\frac{\lambda^2\sigma^2}{2}\right) = 1 + \frac{\lambda^2\sigma^2}{2} + \frac{\lambda^2 v_\alpha^2}{2} + o\left(\lambda^2\right)$$

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

The bound on the e_n term requires concentration, and takes the following form:

$$\mathrm{P}[e_n > \epsilon] \leq 4e^{-\frac{n(1-\alpha^2)\eta^2 \min(\epsilon^2, \delta^2)}{16}} \quad \text{… …} \ (12)$$

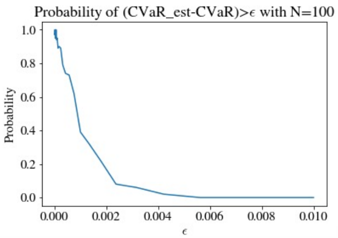


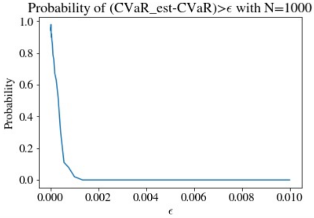


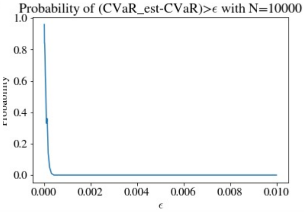
Inference –

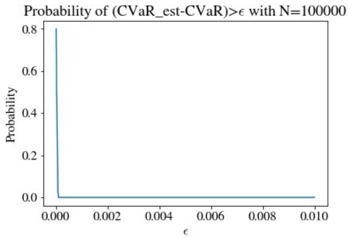
- As the sample size increases, the estimated CVaR gets closer to the true value of CVaR.
- The probability is decaying in an exponential manner.
- The bounds specified in Theorem 3.1 is satisfied.

Simulation Results – Theorem 3.1





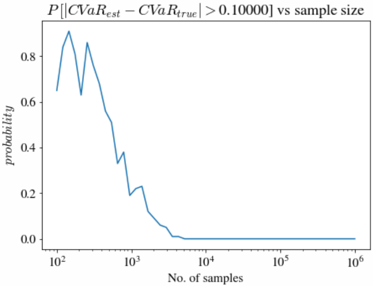


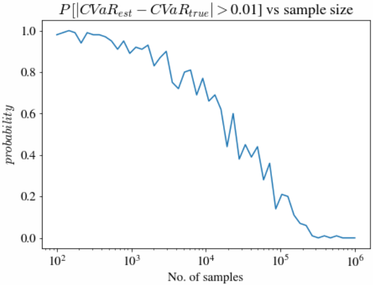


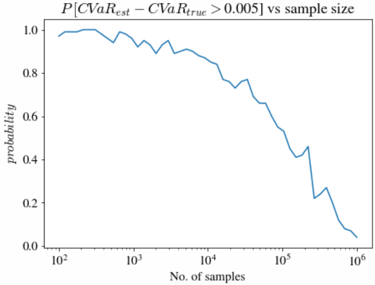
Inference –

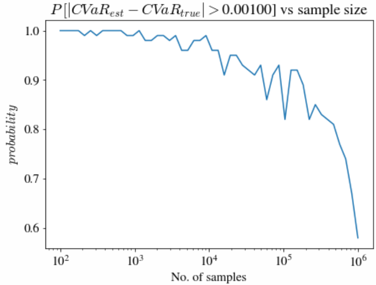
- With the same sample size, as we decrease ϵ, after a certain point, the probability goes to 1.
- However, with increasing sample size, the "ϵ threshold" decreases.
- This implies that the estimator is converging to the true value and is consistent.

Simulation Results – Sub-Gaussian (Gaussian Distribution)









Simulation Results – Sub-Gaussian (Gaussian Distribution)

Theoretical Expression for VaR and CVaR

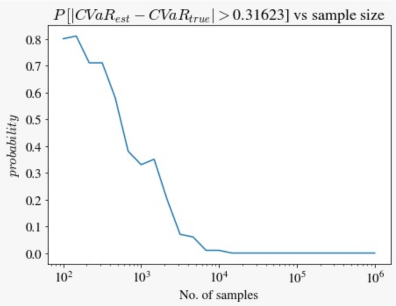
$$VaR = \Phi^{-1}(\alpha)\,\sigma + \mu$$

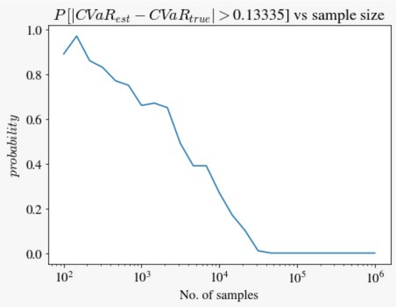
$$CVaR = \mu - \sigma \cdot \frac{\phi(\Phi^{-1}(\alpha))}{1-\alpha}$$

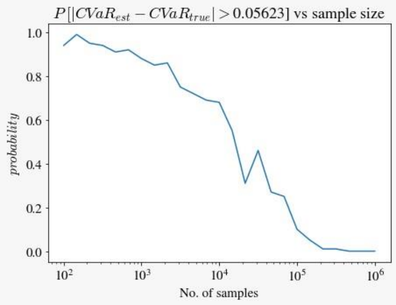
Inference –

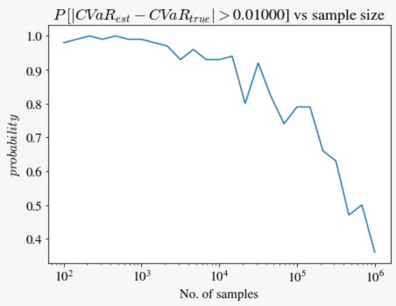
- As the sample size increases, the estimated CVaR gets closer to the true value of CVaR.
- With decreasing ϵ, the probability falls at a lower rate, yet continues to fall.
- This implies that if samples taken were high enough, the estimator would converge to the true value.

Simulation Results – Sub-Exponential (χ_1^2 Distribution)









Simulation Results – Sub-Exponential (χ_1^2 Distribution)

χ_1^2 is sub-exponential with parameters (2,4)

Theoretical Expression for VaR and CVaR

$$VaR = \Phi^{-1}(\alpha)$$

$$CVaR = \int_{\Phi^{-1}(\alpha)}^{\infty} x.\chi_1^2 dx$$

Inference –

- As the sample size increases, the estimated CVaR gets closer to the true value of CVaR.
- With decreasing ϵ, the probability falls at a lower rate, yet continues to fall.
- This implies that if samples taken were high enough, the estimator would converge to the true value.

Notable results

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 small ϵ, the bound 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.

Notable results contd...

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

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**.

Application: Multi-armed bandits

Goal: Finding the arm with the lowest CVaR in a fixed budget setting.

- K-armed stochastic bandit problem with arms' distributions $\mathrm{P}_1 \ldots \mathrm{P}_k$.
- Let $\mathrm{c}_\mathrm{i}^{\alpha}$ and v_i^{α} denote the CVaR and VaR of the arm i at level α.
- Let $\mathrm{c}^{\alpha} = \min_{\mathrm{i}=1,\ldots,\mathrm{K}} \mathrm{c}_\mathrm{i}^{\alpha}$, and i^* be the arm that achieves this minimum.

A brief explanation of the Bandit problem

A bandit algorithm interacts with the environment over a given budget of n rounds.

In each round t = 1, . . . , n, the algorithm pulls an arm $I_t \in \{1, \ldots, K\}$ and observes a sample cost from the distribution 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., $\mathrm{P}\,[\,J_n \neq \mathrm{i}^*]$, where i^* denotes the best arm.

Based on the Successive Reject (SR) algorithm

Main Idea: In each round t, we play each arm in the active set S_{t-1} for a certain number of times, and then use the empirical means to update the active set S_t by eliminating some bad arms in S_{t-1}. In the end, only one arm is returned by the final round.

What's novel?
Here, **CVaR is used** instead of expected value.

Algorithm: CVaR-SR algorithm

Initialization: Set A_1 = {1, . . . , 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, 2, …, K-1

for k = 1, 2, …, K-1 **do**

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

Set $A_{k+1} = A_k \setminus \underset{i \in A_k}{\operatorname{argmax}}\, \hat{c}_{\alpha,n_k}^{i}$

end for

Output: Return the solitary element in A_K.

Theorem: Probability of incorrect identification

Problem: K-armed stochastic bandit problem where the arms' distributions satisfy **(A1)** and are either sub-Gaussian or light-tailed.

Claim: For a given budget *n,* the arm, say J_n, returned by the CVaR-SR algorithm satisfies:

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

$\boldsymbol{G_{max}}$: Problem dependent constant which does not depend on the underlying CVaR gaps and n.

H: Hardness measure defined by –

$$H = \max_{i\in\{1,\dots,K\}} \frac{i}{\min(\Delta_{[i]}^2, 4\delta_{[i]}^2)} \text{ for σ-sub-Gaussian arms,}$$

$$H = \max_{i\in\{1,\dots,K\}} \frac{i}{\min\left\{\Delta_{[i]}, \Delta_{[i]}^2, 4\delta_{[i]}^2\right\}} \text{ for (σ, b)-light-tailed arms.}$$

$\boldsymbol{\delta_i}$: constant from (A1) corresponding to the distribution of arm i, for i = 1, …, K

Proof: Probability of incorrect identification

(i) Sub-Gaussian case : Rewriting the CVaR concentration bound in (7) from Theorem 3.1,

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

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

(ii) Light-tailed case : Using the following concentration bound instead of the one presented in the previous slide, we get -

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

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

Then following a similar parallel argument as before, we get the desired inequality for the light-tailed case.

Assumption: A milder assumption is adopted - a bounded p^{th} moment exists, enabling us to handle heavy-tailed distribution.

- A truncation-based estimator for CVaR to handle this bounded moment case.

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

Lyapunov's Inequality

For a random variable *X* and numbers $0 \le r \le s < \infty$, we have –

$$E[|X|^r]^{\frac{1}{r}} \le E[|X|^s]^{\frac{1}{s}}$$

Result: If a p^{th} moment is bounded for some $p\ > 2$, from Lyapunov's inequality,

⇒ The second moment is bounded and hence, we simply take p = 2.

CVaR estimation

Proposition: A truncation-based estimator $\hat{c}_{n,\alpha}$ for CVaR at level α defined as –

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

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

- B_i represents a truncation level of X_i.
- Choice for B_i is under the assumption that $E[|X|\,p\,] < u < \infty$ for some $\mathrm{p} \in (1, 2]$.
- Intuitively, the truncation level serves to discard very large samples values early on, as B_i is set to grow slowly with *i*.

Concentration bounds

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 the previous slide formed using the above set of samples. Then, for any $p \in (1,2]$, and $\xi \in (0,1)$, we have with probability at least $1-\xi$,

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

where, η and δ are as defined in **(A1)**

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

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

Proof

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

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

Using a technique from (Bubeck et al., 2013), we obtain the following bound for the first two terms $I_{n,1}$ and $I_{n,2}$:

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

VaR concentration is required to arrive at the following bound on the e_n term:

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

By combining the bounds in (15) and (16), we get the desired inequality.

Corollary of the previous theorem

Assumption: 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 $1 - \xi$,

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

Note - A bandit application for the case of heavy-tailed distribution would involve a truncated estimator in SR algorithm, and a slightly different hardness measure that is derived using Theorem 4.1.

And finally

Conclusion

- Concentration bounds for CVaR estimation were derived
- Light-tailed (using empirical based estimators) and heavy-tailed distributions (truncation based CVaR estimator).

- For the heavy-tailed case, a concentration bound is derived under a mild bounded-moment.
- Concentration bound enjoys exponential decay in the sample size even for heavy-tailed random variables.

- Adaptation of the successive rejects algorithm was proposed to to the multi-armed bandit problem
- Goal of finding the arm with the lowest CVaR
- Established error bounds for the proposed algorithm

Future Work and Scope

- A tighter optimal bound greater than the current one based on CVaR.
- More simulation results and code repositories in the public domain.
- Greater diversity of the concentration bounds’ real-life applications and utility.

THANK YOU!