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

Prashanth L A*, Krishna Jagannathan*, Ravi Kolla†

Paper #2156

∗ IIT Madras, Chennai, India
† AB InBev

One-Slide Summary

Objective: Estimate the Conditional Value-at-Risk (CVaR) $c_\alpha(X)$ of a r.v. X from n i.i.d. samples

Our Contributions: Concentration bounds $\mathbb{P}\left[|c_{n,\alpha} - c_\alpha(X)| > \epsilon\right]$

One-Slide Summary

Objective: Estimate the Conditional Value-at-Risk (CVaR) $c_\alpha(X)$ of a r.v. X from n i.i.d. samples

Our Contributions: Concentration bounds $\mathbb{P}\left[|c_{n,\alpha} - c_\alpha(X)| > \epsilon\right]$

1) Sub-Gaussian distributions: Our bounds match an existing result, but with better constants

One-Slide Summary

Objective: Estimate the Conditional Value-at-Risk (CVaR) $c_\alpha(X)$ of a r.v. X from n i.i.d. samples

Our Contributions: Concentration bounds $\mathbb{P}\left[|c_{n,\alpha} - c_\alpha(X)| > \epsilon\right]$

1) Sub-Gaussian distributions: Our bounds match an existing result, but with better constants

2) Light-tailed distributions: We derive an $O(\exp(-cn \min(\epsilon, \epsilon^2)))$ tail bound for an empirical CVaR estimator

One-Slide Summary

Objective: Estimate the Conditional Value-at-Risk (CVaR) $c_\alpha(X)$ of a r.v. X from n i.i.d. samples

Our Contributions: Concentration bounds $\mathbb{P}\left[|c_{n,\alpha} - c_\alpha(X)| > \epsilon\right]$

1) Sub-Gaussian distributions: Our bounds match an existing result, but with better constants

2) Light-tailed distributions: We derive an $O(\exp(-cn\min(\epsilon, \epsilon^2)))$ tail bound for an empirical CVaR estimator

3) Heavy-tailed distributions with bounded variance: We derive an $O(\exp(-cn\epsilon^2))$ tail bound for a truncated CVaR estimator.

One-Slide Summary

Objective: Estimate the Conditional Value-at-Risk (CVaR) $c_\alpha(X)$ of a r.v. X from n i.i.d. samples

Our Contributions: Concentration bounds $\mathbb{P}\left[|c_{n,\alpha} - c_\alpha(X)| > \epsilon\right]$

1) Sub-Gaussian distributions: Our bounds match an existing result, but with better constants

2) Light-tailed distributions: We derive an $O(\exp(-cn\min(\epsilon, \epsilon^2)))$ tail bound for an empirical CVaR estimator

3) Heavy-tailed distributions with bounded variance: We derive an $O(\exp(-cn\epsilon^2))$ tail bound for a truncated CVaR estimator.

4) Bandit application: Best CVaR arm identification and error bounds

What is Conditional Value-at-Risk (CVaR)?

VaR and CVaR are Risk Metrics

- Widely used in financial portfolio optimization, credit risk assessment and insurance
- Let X be a continuous random variable
- Fix a `risk level' $\alpha \in (0, 1)$ (say $\alpha = 0.95$)

VaR and CVaR are Risk Metrics

- Widely used in financial portfolio optimization, credit risk assessment and insurance
- Let X be a continuous random variable
- Fix a `risk level' $\alpha \in (0, 1)$ (say $\alpha = 0.95$)

Value at Risk:

$$v_\alpha(X) = F_X^{-1}(\alpha)$$

VaR and CVaR are Risk Metrics

- Widely used in financial portfolio optimization, credit risk assessment and insurance
- Let X be a continuous random variable
- Fix a `risk level' $\alpha \in (0,1)$ (say $\alpha = 0.95$)

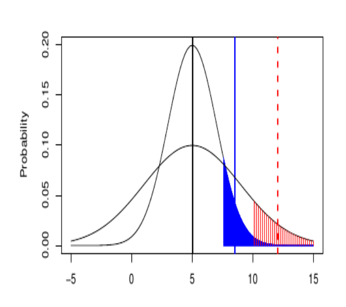
Value at Risk:

$$v_\alpha(X) = F_X^{-1}(\alpha)$$

Conditional Value at Risk:

$$c_\alpha(X) = \mathbb{E}\left[X | X > v_\alpha(X)\right]$$

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



CVaR Estimation and Concentration bounds

CVaR estimation

Problem: Given i.i.d. samples $X_1, \ldots, X_n$ from the distribution F of r.v. X, estimate

$$c_\alpha(X) = \mathbb{E}\left[X | X > v_\alpha(X)\right]$$

Nice to have: Sample complexity $O\left(1/\epsilon^2\right)$ for accuracy ϵ

Empirical VaR and CVaR Estimates

Empirical distribution function (EDF): Given samples $X_1, \ldots, X_n$ from distribution F,

$$\hat{F}_n(x) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{I}\left\{X_i \leq x\right\}, \; x \in \mathbb{R}$$

Using EDF and the order statistics $X_{[1]} \leq X_{[2]} \leq \ldots, X_{[n]}$,

VaR estimate:

$$\hat{v}_{n,\alpha} = \inf\{x : \hat{F}_n(x) \geq \alpha\} = X_{[\lceil n\alpha \rceil]}.$$

Empirical VaR and CVaR Estimates

Empirical distribution function (EDF): Given samples $X_1, \ldots, X_n$ from distribution F,

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

Using EDF and the order statistics $X_{[1]} \leq X_{[2]} \leq \ldots, X_{[n]}$,

VaR estimate:

$$\hat{v}_{n,\alpha} = \inf\{x : \hat{F}_n(x) \geq \alpha\} = X_{[\lceil n\alpha \rceil]}.$$

CVaR estimate:

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

Empirical CVaR concentration: What is known ?

Goal: Bound $\mathbb{P}\left[|\hat{c}_{n,\alpha} - c_\alpha(X)| > \epsilon\right]$

Distribution type	Reference	Salient Feature
Bounded support	[Wang et al. ORL 2010]	$\exp(-cn\epsilon^2)$
Sub-Gaussian/ sub-exponential	[Kolla et al. ORL 2019]	VaR conc. One-sided CVaR
Sub-Gaussian	[S. Bhat & P. L.A. NeurIPS 2019]	Wasserstein
Sub-exponential/ Heavy-tailed	This work	

VaR Concentration [1]

Assumption **(A1):** X is a continuous r.v. with a CDF F that satisfies a condition of *sufficient growth* around the VaR v_α: There exists constants $\delta, \eta > 0$ such that

$$\min\left(F\left(v_\alpha + \delta\right) - F\left(v_\alpha\right), F\left(v_\alpha\right) - F\left(v_\alpha - \delta\right)\right) \geq \eta\delta.$$

[1] Concentration bounds for empirical conditional value-at-risk: The unbounded case; R. Kolla, L.A. Prashanth, S. P. Bhat, K. Jagannathan; *Operations Research Letters*, 2019

VaR Concentration [1]

Assumption **(A1):** X is a continuous r.v. with a CDF F that satisfies a condition of *sufficient growth* around the VaR v_α: There exists constants $\delta, \eta > 0$ such that

$$\min\left(F\left(v_\alpha + \delta\right) - F\left(v_\alpha\right), F\left(v_\alpha\right) - F\left(v_\alpha - \delta\right)\right) \geq \eta\delta.$$

Lemma (*VaR concentration*)

Suppose that (A1) holds. We have for all $\epsilon \in (0, \delta)$,

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

Proof uses DKW inequality; no tail assumptions required.

[1] Concentration bounds for empirical conditional value-at-risk: The unbounded case; R. Kolla, L.A. Prashanth, S. P. Bhat, K. Jagannathan; *Operations Research Letters*, 2019

Concentration for CVaR$_\alpha$ Estimator

- Obtaining concentration for CVaR$_\alpha$ estimator is more involved than for VaR$_\alpha$
- Need to make some assumptions on the tail distribution

Concentration for CVaR$_\alpha$ Estimator

- Obtaining concentration for CVaR$_\alpha$ estimator is more involved than for VaR$_\alpha$
- Need to make some assumptions on the tail distribution
- We work with three progressive broader distribution classes
  (i) X is sub-Gaussian or
  (ii) X is sub-exponential (i.e., light-tailed) or
  (iii) X has a bounded second moment

Concentration for CVaR$_\alpha$ Estimator

- Obtaining concentration for CVaR$_\alpha$ estimator is more involved than for VaR$_\alpha$
- Need to make some assumptions on the tail distribution
- We work with three progressive broader distribution classes
  (i) X is sub-Gaussian or
  (ii) X is sub-exponential (i.e., light-tailed) or
  (iii) X has a bounded second moment
- For (i) and (ii), we use the empirical CVaR estimator; for (iii) we use a truncated CVaR estimator

Sub-Gaussian and Sub-Exponential distributions

A random variable is X is sub-Gaussian if $\exists\, \sigma > 0$ s.t.

$$\mathbb{E}\left[e^{\lambda X}\right] \leq e^{\frac{\sigma^2 \lambda^2}{2}},\ \forall \lambda \in \mathbb{R}.$$

Or equivalently, letting $Z \sim \mathcal{N}(0, \sigma^2)$,

$$\mathbb{P}\left[X > \epsilon\right] \leq c\mathbb{P}\left[Z > \epsilon\right], \forall \epsilon > 0.$$

Tail dominated by a Gaussian

Sub-Gaussian and Sub-Exponential distributions

A random variable is X is sub-Gaussian if $\exists\ \sigma > 0$ s.t.

$$\mathbb{E}\left[e^{\lambda X}\right] \leq e^{\frac{\sigma^2\lambda^2}{2}},\ \forall\lambda \in \mathbb{R}.$$

Or equivalently, letting $Z \sim \mathcal{N}(0, \sigma^2)$,

$\mathbb{P}\left[X > \epsilon\right] \leq c\mathbb{P}\left[Z > \epsilon\right], \forall\epsilon > 0.$ ⟵ Tail dominated by a Gaussian

---

A random variable is X is sub-exponential if $\exists\ c_0 > 0$ s.t.

$$\mathbb{E}\left[e^{\lambda X}\right] < \infty,\ \forall|\lambda| < c_0.$$

Or equivalently, $\exists\sigma, b > 0$ s.t. $\mathbb{E}\left[e^{\lambda X}\right] \leq e^{\frac{\sigma^2\lambda^2}{2}},\ \forall|\lambda| \in \frac{1}{b}.$ Or

$\mathbb{P}\left[X > \epsilon\right] \leq c_1\exp(-c_2\epsilon), \forall\epsilon > 0.$ ⟵ Tail dominated by an exponential r.v

CVaR concentration for Sub-Gaussian case

Recall

$$\hat{v}_{n,\alpha} = \inf\{x : \hat{F}_n(x) \geq \alpha\} = X_{[\lceil n\alpha \rceil]}.$$

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

CVaR concentration for Sub-Gaussian case

Recall

$$\hat{v}_{n,\alpha} = \inf\{x : \hat{F}_n(x) \geq \alpha\} = X_{[\lceil n\alpha \rceil]}.$$

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

Theorem (*CVaR concentration for sub-Gaussian*)

Assume (A1). Suppose that X_i, $i = 1, \ldots, n$ are σ-sub-Gaussian. Then, for any $\epsilon \in (0, \delta)$, we have

$$\mathbb{P}\left[|\hat{c}_{n,\alpha} - c_\alpha| > \epsilon\right] \leq 6 \exp\left[-n\psi_1(\epsilon)\right],$$

where $\psi_1(\epsilon) = \dfrac{\epsilon^2 (1-\alpha)^2 \min\left(\eta^2, 1\right)}{8 \max\left(\sigma^2, 8\right)}$.

CVaR concentration for Sub-Exponential case

Recall

$$\hat{v}_{n,\alpha} = \inf\{x : \hat{F}_n(x) \geq \alpha\} = X_{[\lceil n\alpha \rceil]}.$$

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

Theorem (*CVaR concentration for sub-Exponential*)

Assume (A1). Suppose that X_i, $i = 1, \ldots, n$ are sub-exponential with parameters σ, b. Then, for all $\epsilon \in (0, \delta)$, we have

$$\mathbb{P}\left[|\hat{c}_{n,\alpha} - c_\alpha| > \epsilon\right] \leq 6 \exp\left[-n\psi_2(\epsilon)\right],$$

where $\psi_2(\epsilon) = \min\left(\dfrac{\epsilon^2(1-\alpha)^2 \min\left(\eta^2, 1\right)}{8\max\left(\sigma^2, 8\right)}, \dfrac{\epsilon(1-\alpha)}{8b}\right).$

Handling Heavy-Tailed distributions

- Heavy-tailed distributions occur commonly in finance applications
- Tail of distribution decays slower than any exponential —characterised by atypically large sample values
- Empirical estimates may be `thrown off' due to atypically large values occurring early in the aggregating process
- Raw empirical estimates do not concentrate well around true value

[2] Bubeck et. al., *Bandits with Heavy-Tail*, IEEE Trans. Inf. Thy., 2013.

Handling Heavy-Tailed distributions

- Heavy-tailed distributions occur commonly in finance applications
- Tail of distribution decays slower than any exponential —characterised by atypically large sample values
- Empirical estimates may be `thrown off' due to atypically large values occurring early in the aggregating process
- Raw empirical estimates do not concentrate well around true value
- Key Idea: Truncated estimator!
- Truncate large values, while slowly growing the truncation threshold[2]

[2] Bubeck et. al., *Bandits with Heavy-Tail*, IEEE Trans. Inf. Thy., 2013.

The Bounded (Second) Moment case: Truncated CVaR estimator

Assume **(A2)** $\exists\, u$ such that $\mathbb{E}[X^2] < u < \infty$.

The Bounded (Second) Moment case: Truncated CVaR estimator

Assume **(A2)** $\exists\, u$ such that $\mathbb{E}[X^2] < u < \infty$.

Truncated CVaR Estimator

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

where $B_i \propto \sqrt{ui}$.

CVaR Concentration for the Bounded Moment case

Theorem (*CVaR concentration: Bounded second 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 truncated CVaR estimate formed using the above set of samples. For all $\epsilon > 0$,

$$\mathbb{P}\left[|\hat{c}_{n,\alpha} - c_\alpha| > \epsilon\right] \leq 2\exp\left(-\frac{n(1-\alpha)^2\epsilon^2}{144\left(\sqrt{u} + v_\alpha\right)^2}\right) + 4\exp\left(-\frac{n\eta^2(1-\alpha)^2\min\left(\epsilon^2, \delta^2\right)}{144}\right),$$

where η and δ are as defined in (A1).

Bandit application

CVaR-aware bandits: Model

Known # of arms K and horizon n

Unknown Distributions $F_k, k = 1, \ldots, K$,

CVaR-values (at fixed risk level α) : $c_1, c_2, \ldots, c_K$

Interaction In each round $t = 1, \ldots, n$

- pull arm $I_t \in \{1, \ldots, K\}$
- observe a sample loss from F_{I_t}

Recommendation Arm J_n

Benchmark: $k^* = \arg\min_{k=1,\ldots,K} c_k$.

CVaR-aware bandits: Model

Known # of arms K and horizon n

Unknown Distributions $F_k, k = 1, \ldots, K$,

CVaR-values (at fixed risk level α) : $c_1, c_2, \ldots, c_K$

Interaction In each round $t = 1, \ldots, n$

- pull arm $I_t \in \{1, \ldots, K\}$
- observe a sample loss from F_{I_t}

Recommendation Arm J_n

Benchmark: $k^* = \arg\min_{k=1,\ldots,K} c_k$.

Goal: Minimize probability of erroneous recommendation

$$\mathbb{P}\left[J_n \neq k^*\right]$$

The Successive Rejects Algorithm[3]

Initialization

$$A_1 = \{1, \ldots, K\},$$
$$n_k = \left\lceil \frac{1}{\overline{\log}(K)} \frac{n-K}{K+1-k} \right\rceil$$

Phase 1

Play each arm $j \in A_1$, n_1 times;
Set $A_2 = A_1 \setminus \arg\max_{j \in A_1} \hat{c}_{j,n_1}$

Phase 2

Play each arm $j \in A_2$, $n_2 - n_1$ times; Eliminate ...

$\vdots$

$\vdots$

Phase $K-1$

Play the remaining two arms $n_{K-1} - n_{K-2}$ times

- One arm played n_1 times, ..., another played n_{K-2} times
- Two arms played n_{K-1} times
- $n_1 + \ldots + n_{K-1} + n_{K-1} \leq n$
- n_k increases with k
- Adaptive exploration: better than uniform (i.e., play each arm n/K times)

[3] Audibert et al., *Best Arm Identification in Multi-armed Bandits*, COLT 2010

Probability of error for Successive Rejects

- Suppose the arm distributions are all 1-sub-exponential.
- Given a simulation budget n, the probability that the SR algorithm identifies a suboptimal arm as being optimal can be bounded as

$$\mathbb{P}\left[J_n \neq k^*\right] \leq 3K(K-1)\exp\left(-\frac{(n-K)(1-\alpha)^2\beta}{H_2\ \overline{\log}(K)}\right),$$

where β is a problem dependent constant (indep. of the gaps), and

$$H_2 = \max_{k=1,2,\ldots,K} \frac{k}{\min(\Delta_k, \Delta_k^2, \delta_k^2)},$$

where δ_k is the constant from (A1) for arm k's distribution

Concluding Remarks

- Derived a concentration bound for empirical CVaR_α estimator for sub-Gaussian and sub-exponential r.v.s
- A truncated CVaR estimator to handle heavy-tailed distributions
- Showed a bandit application for best CVaR_α arm identification, and derived probability of error for SR algorithm

Thank you!