Uniformly Randomized Markov Inequality

Christian Covington, Dominic DiSanto

BST 235 Final Presentation

Markov Inequality Refresher

Theorem (Markov Inequality (MI))

Let $X : \Omega \mapsto \mathbb{R}^{\geq 0}$ *and* $a \geq 0$*. Then*

$$\Pr(X \geq a) \leq a^{-1}\mathbb{E}X.$$

Markov Inequality Refresher

Theorem (Markov Inequality (MI))

Let $X : \Omega \mapsto \mathbb{R}^{\geq 0}$ *and* $a \geq 0$. *Then*

$$\Pr\left(X \geq a\right) \leq a^{-1}\mathbb{E}X.$$

MI is the core idea powering many concentration inequalities (Chebyshev, Chernoff, etc.)

Markov Inequality Refresher

Theorem (Markov Inequality (MI))

Let $X : \Omega \mapsto \mathbb{R}^{\geq 0}$ *and* $a \geq 0$. *Then*

$$\Pr(X \geq a) \leq a^{-1}\mathbb{E}X.$$

MI is the core idea powering many concentration inequalities (Chebyshev, Chernoff, etc.)

In general: tighter characterization of bounds $\implies$ tighter uncertainty quantification.

Is the Markov Inequality tight?

We know MI is tight in the sense that for any a there exists X where $\Pr(X \geq a) = a^{-1}\mathbb{E}X$.

Is the Markov Inequality tight?

We know MI is tight in the sense that for any a there exists X where $\Pr(X \geq a) = a^{-1}\mathbb{E}X$.

$$\begin{aligned} X \sim a \cdot \mathrm{Bern}(p) \implies \mathbb{E}X &= ap \quad \text{(for any } p \in (0,1)) \\ \Pr(X \geq a) &= \Pr(X = a) \\ &= p \\ &= a^{-1}\mathbb{E}X. \end{aligned}$$

Is the Markov Inequality tight?

We know MI is tight in the sense that for any a there exists X where $\Pr(X \geq a) = a^{-1}\mathbb{E}X$.

$$
\begin{aligned}
X \sim a \cdot \mathrm{Bern}(p) \implies \mathbb{E}X &= ap \quad \text{(for any } p \in (0,1)\text{)} \\
\Pr(X \geq a) &= \Pr(X = a) \\
&= p \\
&= a^{-1}\mathbb{E}X.
\end{aligned}
$$

What about in other senses? Is there any way to retain the generality (i.e. for all non-negative X, a) but also give almost surely tighter bounds in downstream applications?

Uniformly Randomized Markov Inequality

Theorem (Uniformly Randomized Markov Inequality (URMI) [RM23])

Let $X : \Omega \to \mathbb{R}^{\geq 0}$, $a \geq 0$, and $U \sim Unif(0,1)$ such that $U \perp X$. Then

$$\Pr\left(XU^{-1} \geq a\right) \leq a^{-1}\mathbb{E}X.$$

Proof.

$$\Pr\left(XU^{-1} \geq a\right) = \mathbb{E}\Pr\left(XU^{-1} \geq a|X\right) = \mathbb{E}\Pr\left(U \leq Xa^{-1}|X\right) = \mathbb{E}\min(1, Xa^{-1}) \leq a^{-1}\mathbb{E}X$$

□

This trivially holds more generally by replacing U with any $B \perp X$ such that B is stochastically larger than U, which we denote by $U \preceq B$.

That is, $\forall y \geq 0 : \Pr(B \leq y) \leq \Pr(U \leq y)$.

Confidence Intervals

Confidence intervals are a nice application of Markov's (via Chebyshev) with intuitive "tightness" interpretation

Confidence Intervals

Confidence intervals are a nice application of Markov's (via Chebyshev) with intuitive "tightness" interpretation

$$\text{Naïve } \mathrm{CI}_{\text{Naïve}} = \overline{X}_n \pm \frac{\sigma}{\sqrt{\alpha n}}$$

$$\text{“Typical” Random CI} = \overline{X}_n \pm \frac{\sigma\sqrt{U}}{\sqrt{\alpha n}}$$

$$\text{Randomized Hoeffding CI} = \overline{X}_n \pm \left(\sigma\sqrt{\frac{2\log(2/\alpha)}{n}} + \frac{\log(U)}{\sqrt{2n\log(2/\alpha)}}\right)$$

U-Hacking and Improving URMI?

[RM23] cites possibility of "U-hacking" (akin to p-hacking).

U-Hacking and Improving URMI?

[RM23] cites possibility of “U-hacking” (akin to p-hacking).

We wondered if it was possible to use this idea to actually tighten URMI further.

U-Hacking and Improving URMI?

[RM23] cites possibility of "U-hacking" (akin to p-hacking).

We wondered if it was possible to use this idea to actually tighten URMI further.

Instead of using

$$\Pr\left(XU^{-1} \geq a\right) \leq a^{-1}\mathbb{E}X,$$

could we beat it by choosing a ϕ such that

$$\Pr\left(X\phi\left(\min_{i\in[k]} U_i^{-1}\right) \geq a\right) \leq a^{-1}\mathbb{E}X$$

and $\phi\left(\min_{i\in[k]} U_i^{-1}\right) \leq U$ (in some reasonable sense)?

Confidence Intervals

$$\text{Naïve CI}_{\text{Naïve}} = \overline{X}_n \pm \frac{\sigma}{\sqrt{\alpha n}}$$

$$\text{“Typical” Random CI} = \overline{X}_n \pm \frac{\sigma\sqrt{U}}{\sqrt{\alpha n}}$$

$$\text{Randomized Hoeffding CI} = \overline{X}_n \pm \left(\sigma\sqrt{\frac{2\log(2/\alpha)}{n}} + \frac{\log(U)}{\sqrt{2n\log(2/\alpha)}}\right)$$

$$\text{“Adjusted” Random CI} = \overline{X}_n \pm \frac{\sigma\sqrt{kU_{(1)}}}{\sqrt{\alpha n}}$$

Over B=2000 iterations, sampled n=1000 $X_i \overset{iid}{\sim} N(0,1)$

Width

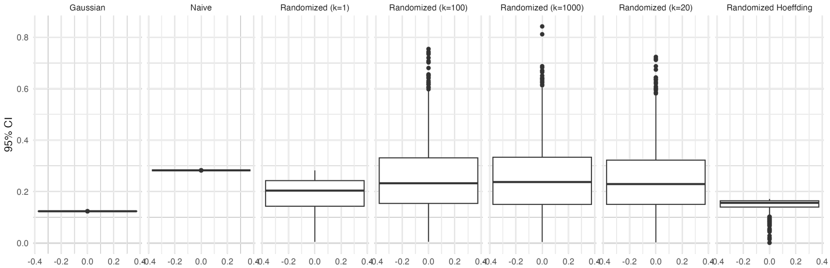

Figure: Comparison of Interval Width by Tail-Bound Randomization for Gaussian Mean (n=1000)

Theoretical Contribution: URMI is Tight!

Theorem (Stochastic upper bounds on U are valid MI randomizers [RM23])

Let $U \sim \mathit{Unif}(0,1)$*. If* $R \succeq U$*, then* R *is a valid MI randomizer.*

Theoretical Contribution: URMI is Tight!

Theorem (Stochastic upper bounds on U are valid MI randomizers [RM23])

Let $U \sim \mathit{Unif}(0,1)$. If $R \succeq U$, then R is a valid MI randomizer.

Theorem (U as stochastic lower bound on valid MI randomizers [CD23])

Let $U \sim \mathit{Unif}(0,1)$. If $R \not\succeq U$, R cannot be a valid MI randomizer.

Theoretical Contribution: URMI is Tight!

Theorem (Stochastic upper bounds on U are valid MI randomizers [RM23])

Let $U \sim Unif(0,1)$. If $R \succeq U$, then R is a valid MI randomizer.

Theorem (U as stochastic lower bound on valid MI randomizers [CD23])

Let $U \sim Unif(0,1)$. If $R \not\succeq U$, R cannot be a valid MI randomizer.

Corollary ("Optimality" of U among valid MI randomizers)

For any valid MI randomizer V such that $V \overset{D}{\neq} U$:

$$\forall y \in (0,1) : \Pr(V \leq y) \leq \Pr(U \leq y)$$
$$\exists y \in (0,1) : \Pr(V \leq y) < \Pr(U \leq y)$$

References I

[CD23] Christian Covington and Dominic DiSanto. Bst235 project, 2023.

[RM23] Aaditya Ramdas and Tudor Manole. Randomized and exchangeable improvements of markov's, chebyshev's and chernoff's inequalities. *arXiv preprint arXiv:2304.02611*, 2023.

Supp: Coverage Table

Type	Coverage Probability
Gaussian	0.95
Naïve	1.00
Randomized Markov's (k=1)	0.95
Randomized Markov's (k=100)	0.95
Randomized Markov's (k=1000)	0.96
Randomized Markov's (k=20)	0.95
Randomized Hoeffding	0.97

Table: Table of Empirical Confidence Interval Coverage Probability over B=2000 Iterations