Hoeffding's Inequality

Ezgi Korkmaz

October 2024

Markov's Inequality

Proposition (*Markov's Inequality*)

Let $X \geq 0$ be a non-negative random variable. Then for all $t \geq 0$

$$\mathbb{P}(X \geq t) \leq \frac{\mathbb{E}[X]}{t}$$

Proof.

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} Xf(X)dX \qquad \textbf{Since } X \geq 0$$

$$\mathbb{E}[X] = \int_{0}^{\infty} Xf(X)dX \qquad \textbf{Since } t \geq 0$$

$$\mathbb{E}[X] \geq \int_{t}^{\infty} Xf(X)dX \qquad \textbf{Since } X \geq t\textbf{: X is in the integrated region}$$

$$\mathbb{E}[X] \geq \int_{t}^{\infty} tf(X)dX = t\int_{t}^{\infty} f(X)dX = t\mathbb{P}(X \geq t)$$

□

$$\mathbb{P}(X \geq t) \leq \frac{\mathbb{E}[X]}{t}$$

Hoeffding's Lemma [Part I]

Lemma (*Hoeffding's Lemma*)

Let X be a random variable with $X \in [a, b]$

$$\mathbb{E}[\exp(\lambda(X - \mathbb{E}[X]))] \leq \exp\left(\frac{\lambda^2(b-a)^2}{8}\right) \quad \textit{for all}\, \lambda \in \mathbb{R}$$

Proof.
If $f : \mathbb{R} \to \mathbb{R}$ is a convex function then

$$f(\mathbb{E}[X]) \leq \mathbb{E}[f(X)] \qquad \to \text{Jensen's Inequality}$$

Let $\hat{X}$ be an independent copy of X where $\mathbb{E}[X] = \mathbb{E}[\hat{X}]$.

$$\begin{aligned}\mathbb{E}_X[\exp(\lambda(X - \mathbb{E}[X]))] &= \mathbb{E}_X[\exp(\lambda(X - \mathbb{E}_{\hat{X}}[\hat{X}]))] \\ &\leq \mathbb{E}_X[\mathbb{E}_{\hat{X}}[(\exp(\lambda(X - \hat{X})))]]\end{aligned}$$

□

Hoeffding's Lemma [Part I]

Proof.

$$\mathbb{E}[\exp(\lambda(X - \mathbb{E}[X]))] \leq \mathbb{E}[(\exp(\lambda(X - \hat{X})))]$$

Let Z be a random sign variable $\{-1, 1\}$, and note that $X - \hat{X}$ symmetric around 0. Then $X - \hat{X}$ has the exact same distribution with $Z(X - \hat{X})$

$$\begin{aligned}\mathbb{E}_{X,\hat{X}}[\exp(\lambda(X - \hat{X}))] &= \mathbb{E}_{X,\hat{X},Z}[\exp(\lambda Z(X - \hat{X}))] \\ &= \mathbb{E}_{X,\hat{X}}[\mathbb{E}_Z[\exp(\lambda Z(X - \hat{X})) \mid X, \hat{X}]]\end{aligned}$$

Now we are going to look at the moment generating function of the random sign. □

Moment Generating Functions

$$M_X(\lambda) \coloneqq \mathbb{E}[e^{\lambda X}]$$

Rademacher Random Variable:

$$\mathbb{E}[e^{\lambda X}] \leq \exp\left(\frac{\lambda^2}{2}\right) \quad \text{for all } \lambda \in \mathbb{R}$$

$$e^X = \sum_{k=0}^{\infty} \frac{X^k}{k!} \quad \text{Taylor expansion of exponential function}$$

$$\mathbb{E}[e^{\lambda X}] = \sum_{k=0}^{\infty} \frac{\lambda^k \mathbb{E}[X^k]}{k!}$$

$$= \sum_{k=0,2,4\ldots}^{\infty} \frac{\lambda^k}{k!} = \sum_{k=0}^{\infty} \frac{\lambda^{2k}}{2k!}$$

Moment Generating Functions

$$\mathbb{E}[e^{\lambda X}] = \sum_{k=0}^{\infty} \frac{\lambda^k \mathbb{E}[X^k]}{k!}$$
$$= \sum_{k=0,2,4\ldots}^{\infty} \frac{\lambda^k}{k!} = \sum_{k=0}^{\infty} \frac{\lambda^{2k}}{2k!}$$

Note that $(2k)! \geq 2^k k!$

$$\mathbb{E}[e^{\lambda X}] \leq \sum_{k=0}^{\infty} \frac{\lambda^{2k}}{2^k \cdot k!} = \sum_{k=0}^{\infty} \left(\frac{\lambda^2}{2}\right)^k \frac{1}{k!} = \exp\left(\frac{\lambda^2}{2}\right)$$

$$\mathbb{E}[e^{\lambda X}] \leq \exp\left(\frac{\lambda^2}{2}\right)$$

Hoeffding's Lemma [Part II]

Since we proved the bounds for the Rademacher random variable now let us continue to Hoeffding's Lemma

Proof.

$$\mathbb{E}_{X,\hat{X}}[\exp(\lambda(X - \hat{X}))] = \mathbb{E}_{X,\hat{X}}[\mathbb{E}_Z[\exp(\lambda Z(X - \hat{X})) \mid X, \hat{X}]]$$

$$\mathbb{E}_Z[\exp(\lambda Z(X - \hat{X})) \mid X, \hat{X}] \leq \exp\left(\frac{(\lambda(X - \hat{X}))^2}{2}\right)$$

$$\mathbb{E}_{X,\hat{X}}[\exp(\lambda(X - \hat{X}))] \leq \exp\left(\frac{(\lambda(b - a))^2}{2}\right)$$

□

Proposition (*Chernoff's Bounds*)

Let X be a random variable. Then for any $t \geq 0$

$$\mathbb{P}(X \geq \mathbb{E}[X] + t) \leq \min_{\lambda \geq 0} \mathbb{E}[e^{\lambda(X - \mathbb{E}[X])}]e^{-\lambda t} = \min_{\lambda \geq 0} M_{X - \mathbb{E}[X]}(\lambda)e^{-\lambda t}$$

and

$$\mathbb{P}(X \leq \mathbb{E}[X] - t) \leq \min_{\lambda \geq 0} \mathbb{E}[e^{\lambda(\mathbb{E}[X] - X)}]e^{-\lambda t} = \min_{\lambda \geq 0} M_{\mathbb{E}[X] - X}(\lambda)e^{-\lambda t}$$

Chernoff's Bounds

Proposition (*Chernoff's Bounds*)
Let X be a random variable. Then for any $t \geq 0$

$$\mathbb{P}(X \geq \mathbb{E}[X] + t) \leq \min_{\lambda \geq 0} \mathbb{E}[e^{\lambda(X - \mathbb{E}[X])}]e^{-\lambda t} = \min_{\lambda \geq 0} M_{X - \mathbb{E}[X]}(\lambda)e^{-\lambda t}$$

$$\text{For } \lambda > 0, \quad X \geq \mathbb{E}[X] + t \text{ if and only if } e^{\lambda X} \geq e^{\lambda \mathbb{E}[X] + \lambda t}$$

$$\begin{aligned} \mathbb{P}(X \geq \mathbb{E}[X] + t) &= \mathbb{P}(e^{\lambda(X - \mathbb{E}[X])} \geq e^{\lambda t}) \rightarrow \textbf{Markov Inequality} \\ &\leq \frac{\mathbb{E}[e^{\lambda(X - \mathbb{E}[X])}]}{e^{\lambda t}} = \mathbb{E}[e^{\lambda(X - \mathbb{E}[X])}]e^{-\lambda t} \end{aligned}$$

Hoeffding's Inequality

Proposition

Let $X_1, \ldots, X_n$ *be independent bounded random variables with* $X_i \in [a, b]$ *for all* i*, where* $-\infty < a \leq b < \infty$*. Then*

$$\mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{n}(X_i - \mathbb{E}[X_i]) \geq t\right) \leq \exp\left(-\frac{2nt^2}{(b-a)^2}\right)$$

and

$$\mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{n}(X_i - \mathbb{E}[X_i]) \leq t\right) \leq \exp\left(-\frac{2nt^2}{(b-a)^2}\right)$$

Hoeffding's Inequality

Proposition

Let $X_1, \ldots, X_n$ be independent bounded random variables with $X_i \in [a,b]$ for all i, where $-\infty < a \leq b < \infty$. Then

$$\mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{n}(X_i - \mathbb{E}[X_i]) \geq t\right) \leq \exp\left(-\frac{2nt^2}{(b-a)^2}\right)$$

We are going to use Hoeffding's Lemma and Chernoff's Bounds

$$\begin{aligned}\mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{n}(X_i - \mathbb{E}[X_i]) \geq t\right) &= \mathbb{P}\left(\sum_{i=1}^{n}(X_i - \mathbb{E}[X_i]) \geq nt\right) \textbf{[Chernoff's Bounds]} \\ &\leq \mathbb{E}\left[\exp\left(\lambda\sum_{i=1}^{n}(X_i - \mathbb{E}[X_i])\right)\right] e^{-\lambda tn} \\ &\leq \prod_{i=1}^{n}\mathbb{E}\left[\exp\left(\lambda(X_i - \mathbb{E}[X_i])\right)\right] e^{-\lambda tn}\end{aligned}$$

Hoeffding's Inequality

We are going to use Hoeffding's Lemma and Chernoff's Bounds

$$\begin{aligned}
\mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{n}(X_i - \mathbb{E}[X_i]) \geq t\right) &= \mathbb{P}\left(\sum_{i=1}^{n}(X_i - \mathbb{E}[X_i]) \geq nt\right) \\
&\leq \mathbb{E}\left[\exp\left(\lambda\sum_{i=1}^{n}(X_i - \mathbb{E}[X_i])\right)\right] e^{-\lambda tn} \textbf{[Chernoff's Bounds]} \\
&= \prod_{i=1}^{n}\mathbb{E}\left[\exp\left(\lambda(X_i - \mathbb{E}[X_i])\right)\right] e^{-\lambda tn} \\
&\leq \prod_{i=1}^{n}\exp\left(\frac{\lambda^2(b-a)^2}{8}\right) e^{-\lambda tn} \textbf{[Hoeffding's Lemma]}
\end{aligned}$$

Let us rewrite and minimize over λ

$$\mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{n}(X_i - \mathbb{E}[X_i]) \geq t\right) \leq \min_{\lambda \geq 0} \exp\left(\frac{n\lambda^2(b-a)^2}{8} - \lambda nt\right) = \exp\left(-\frac{2nt^2}{(b-a)^2}\right)$$