Concentration Inequalities

Peng Lingwei

August 5, 2019

Contents

1 Hoeffding's Inequality

Consider the sum $S_n = \sum_{i=1}^{n} X_i$ of independent random variable $X_1, \ldots, X_n$. Then, $\forall s > 0$,

$$\begin{aligned}
\mathbb{P}\left\{S_n - \mathbb{E}S_n \geq t\right\} =& \mathbb{P}\left\{\exp\left\{s(S_n - \mathbb{E}S_n)\right\} \geq e^{st}\right\} \\
\leq& e^{-st}\mathbb{E}e^{s(S_n - \mathbb{E}S_n)} \\
=& e^{-st}\prod_{i=1}^{n}\mathbb{E}e^{s(X_i - \mathbb{E}X_i)}
\end{aligned}$$

Lemma 1. *For a random varialble X with $\mathbb{E}X = 0$ and $a \leq X \leq b$ then for $s \geq 0$,*

$$\mathbb{E}e^{sX} \leq e^{s^2(b-a)^2/8}$$

Proof. Because $\mathbb{E}X = 0, a \leq X \leq b$, therefore $a = 0 \wedge b = 0$ or $a < 0 \wedge b > 0$.
If $a = 0 \wedge b = 0$, $\mathbb{E}e^{sX} \leq e^{s^2(b-a)^2/8} = 1$
Else if $a < 0 \wedge b > 0$:
From convexity,

$$e^{sx} \leq \frac{x-a}{b-a}e^{sb} + \frac{b-x}{b-a}e^{sa} \Rightarrow \mathbb{E}e^{sX} \leq \frac{be^{sa} - ae^{sb}}{b-a}$$

Let $f(s) = \ln\frac{be^{sa} - ae^{sb}}{b-a}$, we can analyse $f(s)$.

1. $f(0) = 0$;
2. $f'(s) = \frac{abe^{sa} - abe^{sb}}{be^{sa} - ae^{sb}} = \frac{a^2/b - a}{e^{s(a-b)} - a/b} + 1$, $f'(0) = 0$;
3. $f''(s) = -\frac{a}{b}(a-b)^2\frac{e^{s(a-b)}}{(e^{s(a-b)} - a/b)^2} = -\frac{a}{b}(a-b)^2 / \left\{e^{s(a-b)} + \frac{a^2/b^2}{e^{s(a-b)}} - 2a/b\right\}$ $\leq -\frac{a}{b}(a-b)^2 / \{-4a/b\} = \frac{1}{4}(a-b)^2$
4. $f(s) = f(0) + f'(0)s + \frac{f''(\theta s)}{2}s^2 \leq \frac{1}{8}(a-b)^2 s^2$

□

$$\begin{aligned}
\mathbb{P}(S_n - \mathbb{E}S_n \geq t) \leq& \inf_{s>0}\left(e^{-st}\prod_{i=1}^{n}e^{s^2(b_i-a_i)^2/8}\right) \\
=& \inf_{s>0}\exp\left(-st + \frac{s^2}{8}\sum_{i=1}^{n}(b_i - a_i)^2\right) \\
=& \exp\left(\frac{-2t^2}{\sum_{i=1}^{n}(b_i - a_i)^2}\right)
\end{aligned}$$

Theorem 1. ***(Hoeffding's Inequality).*** *For bounded random variables $X_i \in [a_i, b_i]$, where $X_1, \ldots, X_n$ are independent, then*

$$\mathbb{P}\left\{S_n - \mathbb{E}S_n \geq t\right\} \leq \exp\left(\frac{-2t^2}{\sum_{i=1}^{n}(b_i - a_i)^2}\right) \tag{1}$$

$$\mathbb{P}\left\{\mathbb{E}S_n - S_n \geq t\right\} \leq \exp\left(\frac{-2t^2}{\sum_{i=1}^{n}\left(b_i - a_i\right)^2}\right) \tag{2}$$

In PAC learning task, if $l(h, (\vec{x}, y)) \in [0, 1]$,

$$\mathbb{P}\left\{|L_S(h) - L_{\mathcal{D}}(h)| \geq \epsilon\right\} \leq 2\exp\left(\frac{-2\epsilon^2 m}{\frac{1}{m}\sum_{i=1}^{m}\left(b_i - a_i\right)^2}\right) = 2e^{-2\epsilon^2 m}$$

2 McDiarmid's Inequality

Theorem 2. ***(McDiarmid's Inequality).*** *Consider independent random variables $X_1, \ldots, X_n \in \mathcal{X}$, the function ϕ satisfies*

$$\phi(x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_n) - \phi(x_1, \ldots, x_{i-1}, {x'}_i, x_{i+1}, \ldots, x_n) \in [a_i, b_i]$$

then,

$$\mathbb{P}\left\{\phi(X_1, \ldots, X_n) - \mathbb{E}\phi \geq t\right\} \leq \exp\left\{\frac{-2t^2}{\sum_{i=1}^{n}\left(a_i - b_i\right)^2}\right\} \tag{3}$$

$$\mathbb{P}\left\{\mathbb{E}\phi - \phi(X_1, \ldots, X_n) \geq t\right\} \leq \exp\left\{\frac{-2t^2}{\sum_{i=1}^{n}\left(a_i - b_i\right)^2}\right\} \tag{4}$$

Proof. We denote $V_i = \mathbb{E}_{X_{i+1},\ldots,X_n}[\phi|X_1, \ldots, X_i] - \mathbb{E}_{X_i,\ldots,X_n}[\phi|X_1, \ldots, X_{i-1}]$, then $\sum_{i=1}^{n} V_i = \phi(X_1, \ldots, X_n) - \mathbb{E}_{X_1,\ldots,X_n}\phi$.

$$L_i = \inf_x \mathbb{E}_{X_{i+1},\ldots,X_n}[\phi|X_1, \ldots, X_{i-1}, x] - \mathbb{E}_{x_i,\ldots,X_n}[\phi|X_1, \ldots, X_{i-1}]$$
$$U_i = \sup_x \mathbb{E}_{X_{i+1},\ldots,X_n}[\phi|X_1, \ldots, X_{i-1}, x] - \mathbb{E}_{x_i,\ldots,X_n}[\phi|X_1, \ldots, X_{i-1}]$$

Then, $L_i \leq V_i \leq U_i$.Because $L_i \geq L_i - U_i \geq a_i$ and $U_i \leq U_i - L_i \leq b_i$, therefore $a_i \leq V_i \leq b_i$. By Hoeffding, we have

$$\mathbb{P}(\phi - \mathbb{E}\phi \geq t) \leq \inf_{s>0} e^{-st}\mathbb{E}_{X_1,\ldots,X_n}\left(\prod_{i=1}^{n} e^{sV_i}\right)$$

Now, we look into $\mathbb{E}_{X_1,\ldots,X_n}\left(\prod_{i=1}^{n} e^{sV_i}\right)$:

$$\begin{aligned}
\mathbb{E}_{X_1,\ldots,X_n}\left(\prod_{i=1}^{n} e^{sV_i}\right) =& \mathbb{E}_{X_1,\ldots,X_{n-1}}\left(\prod_{i=1}^{n-1} e^{sV_i} \cdot \mathbb{E}_{X_n} e^{sV_n}\right) \\
\leq& \mathbb{E}_{X_1,\ldots,X_{n-1}}\left(\prod_{i=1}^{n-1} e^{sV_i} \cdot\right) e^{s^2(a_n - b_n)^2/8} \\
\ldots \leq& \exp\left(s^2 \sum_{i=1}^{n}\left(a_i - b_i\right)^2/8\right)
\end{aligned}$$

$\square$