Tail Inequalities (2)

SubExponential, Bernstein, McDiarmid, Levy

Sun Woo Lim

Dec 2023

Outline

Subexponential random variable

A random variable X is SE(ν, α) if $\mathbb{E}e^{\lambda(X-\mu)} \leq e^{\frac{\nu^2\lambda^2}{2}}, \quad |\lambda| < 1/\alpha.$
Note that SG(σ^2) random variable is subexponential with $\nu = \sigma$ and $\alpha = 0$.

Ex) Chi-square distribution
For $X \sim \chi^2(1)$, $X \in$ SE$(\nu = 2, \alpha = 4)$.
pf) Note that $\mathbb{E}X = 1$. Let $Z \sim$ N$(0,1)$.

$$\mathbb{E}e^{X-1} = \mathbb{E}e^{Z^2-1} = \int_{-\infty}^{\infty} e^{\lambda(z^2-1)} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz = e^{-\lambda} \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} e^{-(\frac{1}{2}-\lambda)z^2} dz$$
$$= e^{-\lambda}\sigma \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{z^2}{2\sigma^2}} dz = e^{-\lambda\sigma} = e^{-\lambda} \frac{1}{\sqrt{1-2\lambda}},$$

by choosing λ such that $1/2 - \lambda = 1/(2\sigma^2)$. $\sigma > 0 \iff \lambda < 1/2$.

Properties of subexponential random variable

Square of subGaussian is subexponential
Let X be zero-mean $\mathsf{SG}(\sigma^2)$ random variable. Then, $X^2 \in \mathsf{SE}(\nu = 16\sigma^2, \alpha = 16\sigma^2)$.

Subexponential tail bound
For $X \in \mathsf{SE}(\nu, \alpha)$,

$$\mathbb{P}(|X-\mu| \geq t) \leq \begin{cases} 2e^{-t^2/(2\nu^2)} & \text{if } 0 \leq t \leq \nu^2/\alpha \\ 2e^{-t/(2\alpha)}, & \text{if } t > \nu^2/\alpha. \end{cases} \Longleftrightarrow \mathbb{P}(|X-\mu| \geq t) \leq 2e^{-\frac{1}{2}\min(\frac{t}{\alpha}, \frac{t}{\nu^2})}.$$

Sum of independent subexponential random variables
Let $X_1, \ldots, X_n$ be independent $\mathsf{SE}(\nu_i, \alpha_i)$ random variables. Then, for $\nu_* = \sqrt{\sum_{i=1}^n \nu_i^2}$ and $\alpha_* = \max_{i=1,\ldots,n} \alpha_i$,

$$\sum_{i=1}^n (X_i - \mu_i) \in \mathsf{SE}(\nu_*, \alpha_*)$$

$$\mathbb{P}\left(\frac{1}{n}\sum_{i=1}^n |X_i - \mu_i| \geq t\right) \leq \begin{cases} 2e^{-nt^2/(2\nu_*^2)} & \text{if } 0 \leq t \leq \nu_*^2/(n\alpha_*) \\ 2e^{-nt/(2\alpha_*)}, & \text{if } t > \nu_*^2/\alpha_*. \end{cases}$$

Outline

Bernstein Condition

Suppose $\mid \mathbb{E}(X-\mu)^k \mid \leq \frac{1}{2}k!\sigma^2 b^{k-2}, \quad k=2,3,\ldots$. Then, X satisfies Bernstein condition with parameter b.
Note that all bounded random variables $\mid X-\mu \mid \leq b$ satisfy this.

Bernstein implies subexponential
Let X satisfy Bernstein condition with parameter b. Then, Claim) $X \in \mathsf{SE}(\sqrt{2}\sigma, 2b)$.
Pf)

$$\begin{aligned}
\mathbb{E}e^{\lambda(X-\mu)} &= 1 + 0 + \frac{\lambda^2\sigma^2}{2} + \sum_{k=3}^{\infty} \lambda^k \frac{\mathbb{E}(X-\mu)^k}{k!} \\
&\leq 1 + \frac{\lambda^2\sigma^2}{2} + \frac{\lambda^2\sigma^2}{2}\sum_{k=3}^{\infty}(\mid \lambda \mid b)^{k-2} = 1 + \frac{\lambda^2\sigma^2/2}{1-b\mid\lambda\mid} \\
&\leq e^{\frac{\lambda^2\sigma^2/2}{1-b\mid\lambda\mid}} \qquad \because 1+x \leq e^x, \forall x \in \mathbb{R} \\
&\leq e^{\lambda^2(\sqrt{2}\sigma)^2/2}.
\end{aligned} \tag{1}$$

In expanding the geometric series, $\mid \lambda \mid < 1/\lambda$ is assumed and the last inequality is based on the assumption of $\mid \lambda \mid < 1/(2b)$.

Bernstein type inequality

For X following Bernstein condition with parameter b,

$$
\begin{aligned}
\mathbb{E}e^{\lambda(X-\mu)} &\leq e^{\frac{\lambda^2\sigma^2}{1-b|\lambda|}}, \quad \forall \mid \lambda \mid < 1/b, \\
\mathbb{P}(\mid X-\mu \mid \geq t) &\leq 2e^{-\frac{t^2}{2(\sigma^2+bt)}}, \quad t \geq 0.
\end{aligned} \tag{2}
$$

The first inequality is proven.
The second inequality is based on setting $\lambda = \frac{t}{bt+\sigma^2} \in [0, 1/b)$ in Chernoff bound.

Outline

Bounded Difference Condition and McDiarmid's inequality

Bounded Difference Condition

For independent random variables $X_1, \dots, X_n$, the bounded difference condition denotes when $f : \mathbb{R}^n \to \mathbb{R}$ satisfies

$$| f(x_1, \dots, x_k, x_{k+1}, \dots, x_n) - f(x_1, \dots, x_k', x_{k+1}, \dots, x_n) | \leq L_k, \quad \forall x, x' \in \mathbb{R}^n$$

McDiarmid's inequality

For random variables $X_1, \dots, X_n$, independent and satisfying bounded difference condition,

$$\mathbb{P}(| f(x_1, \dots, x_n) - \mathbb{E} f(x_1, \dots, x_n) | \geq t) \leq 2 \exp(-\frac{2t^2}{\sum_{i=1}^n L_k^2}).$$

The proof is quite involved.

Ex) Hoeffding's hound

Take $f(x) = \frac{1}{n} \sum_{i=1}^n x_i$ and each $X_i \in [a, b]$ almost surely. Then, $L_k = (b-a)/n$. Then,

$$\mathbb{P}(\frac{1}{n} | \sum_{i=1}^n (X_i - \mu_i) | \geq t) \leq 2 \exp(-\frac{2nt^2}{(b-a)^2}).$$

Levy's inequality: for Lipschitz smooth function of Gaussian random variable

Assume f satisfy

$$| f(x) - f(y) | \leq L \|x - y\|_2 \iff | f(x_1, \ldots, x_n) - f(y_1, \ldots, y_n) | \leq L \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}, \quad \forall x, y \in \mathbb{R}^n.$$

Further, if $X_1, \ldots, X_n \overset{iid}{\sim} \mathsf{N}(0,1)$,

$$\mathbb{P}(| f(x_1, \ldots, x_n) - \mathbb{E} f(x_1, \ldots, x_n) | \geq t) \leq 2 \exp(-\frac{t^2}{2L^2}), \quad \forall t \geq 0.$$

Ex) Order statistic

Consider f as an order statistic (which is a function): $X_{(1)} \leq \cdots \leq X_{(n)}$.

Then, for the iid sequence $Y_1, \ldots, Y_n \overset{iid}{\sim} \mathsf{N}(0,1)$,

$$| f(x_1, \ldots, x_n) - f(y_1, \ldots, y_n) | = | X_{(k)} - Y_{(k)} | \leq 1 \cdot \|X - Y\|_2, \quad \forall k = 1, \ldots, n.$$

Therefore,

$$\mathbb{P}(| X_{(k)} - \mathbb{E} X_{(k)} | \geq t) \leq 2 \exp(-\frac{t^2}{2}), \quad \forall t \geq 0.$$