Basic Bounds

1. Markov : $\forall t > 0$; $\mathbb{P}(X \geq t) \leq \dfrac{\mathbb{E}[X]}{t}$

$$\leq \frac{\mathbb{E}[|X|^k]}{t^k}$$

2. Chebyshev : $\forall t > 0$; $\mathbb{P}[|X - \mu| \geq t] \leq \dfrac{\mathbb{V}[X]}{t^2}$

$$\mathbb{P}[X-\mu \geq t]$$

$$MGF_X(s) = \mathbb{E}[e^{sX}]$$

$$= \mathbb{P}[e^{\lambda(X-\mu)} \geq e^{\lambda t}] \leq \frac{\mathbb{E}[e^{\lambda(X-\mu)}]}{\underbrace{\mathbb{E}[e^{\lambda t}]}_{e^{\lambda t}}}$$

for $\lambda \leq b$

Chernoff bound:

$$\log \mathbb{P}[X-\mu \geq t] \leq \inf_{\lambda \in [0,b]} \left\{ \log \mathbb{E}[e^{\lambda(X-\mu)}] - \lambda t \right\}$$

Gaussian MGF: $e^{\mu\lambda + \sigma^2\lambda^2/2}$ $\quad \forall \lambda \in \mathbb{R}$

$$\frac{\sigma^2\lambda^2}{2} - \lambda t$$

$$\sigma^2\lambda - t = 0 \Rightarrow \lambda = \frac{t}{\sigma^2}$$

$$\Rightarrow \quad \boxed{\mathbb{P}[X-\mu \geq t] \leq e^{-t^2/2\sigma^2}}$$

For Gaussian: $\mathbb{E}[e^{\lambda(X-\mu)}] = e^{\frac{\sigma^2\lambda^2}{2}}$

Def. X with $\mathbb{E}[X]=\mu$ is sub-gaussian if $\exists\, \sigma > 0$ s.t.

$$\mathbb{E}[e^{\lambda(X-\mu)}] \leq e^{\sigma^2\lambda^2/2} \qquad \forall \lambda \in \mathbb{R}$$

Chernoff for subG: $\mathbb{P}[|X-\mu| \geq t] \leq 2e^{-\frac{t^2}{2\sigma^2}}$

Example: Rademacher R.V. $\varepsilon \in \{-1, +1\}$ with $\sigma = 1$.

$$\underbrace{\mathbb{E}[e^{\lambda \varepsilon}]}_{e^{\lambda} + e^{-\lambda}} \leq 2e^{\sigma^2 \lambda^2 / 2} \Rightarrow \sigma = 1 \checkmark$$

Bounded R.V. : $X \in [a,b]$ is zero-mean $X' \overset{d}{=} X$

$$\mathbb{E}[e^{\lambda X}] = \mathbb{E}[e^{\lambda (X - \mathbb{E}[X'])}] \leq \mathbb{E}_{X, X'}[e^{\lambda (X - X')}]$$

$$\varepsilon(X - X') \overset{d}{=} X - X'$$

$$= \mathbb{E}_{X, X'}[\mathbb{E}_{\varepsilon}[e^{\lambda \varepsilon (X - X')}]]$$

$$\leq \mathbb{E}[e^{\lambda^2 (X - X')^2 / 2}]$$

$X \in [a,b] \Rightarrow |X - X'| \leq b - a$

$$\Rightarrow \mathbb{E}[e^{\lambda X}] \leq e^{\lambda^2 (b-a)^2/2} \Rightarrow \text{subG with } \sigma = b-a$$

Stronger result: X is subG with $\boxed{\sigma = \frac{b-a}{2}}$

equivalent forms: $(\mathbb{E}[X] = 0)$

1. $\exists \sigma > 0$ s.t. $\mathbb{E}[e^{\lambda X}] \leq e^{\lambda^2 \sigma^2/2} \quad \forall \lambda \in \mathbb{R}$

2. $\exists c > 0$, $Z \sim N(0, \tau^2)$ s.t. $\mathbb{P}[|X| \geq s] \leq c\,\mathbb{P}[|Z| \geq s]$

3. $\exists \theta \geq 0$ s.t. $\mathbb{E}[X^{2k}] \leq \frac{(2k)!}{2^k k!} \theta^{2k} \quad \forall k \in \mathbb{N}$

4. $\exists \sigma \geq 0$ s.t. $\mathbb{E}[e^{\lambda X^2/2\sigma^2}] \leq 1/\sqrt{1-\lambda} \quad \forall \lambda \in [0,1)$

$\ast$ $X_1 \perp\!\!\!\perp X_2$ both subG $\Longrightarrow$ $X_1 + X_2$ is subG with $\sigma = \sqrt{\sigma_1^2 + \sigma_2^2}$

Hoeffding bound : X_i's are independent & $X_i \in [a, b]$ a.s.

$S_n = \sum X_i$

$$\Longrightarrow \mathbb{P}[S_n - \mathbb{E}[S_n] \geq t] \leq \exp\left[\frac{-2t^2}{\sum (b_i - a_i)^2}\right] \quad \forall t \geq 0$$

$\mathbb{P}[\sum X_i - \mu n \geq t]$

X_i is subG with $\sigma_i = \dfrac{b_i - a_i}{2}$

or we can simply write subG with σ_i

Def. X is sub-exponential if $\exists \nu, \alpha \geq 0$ s.t.

$$E\left[e^{\lambda(X-\mu)}\right] \leq e^{\nu^2\lambda^2/2} \qquad \forall \lambda < \frac{1}{\alpha}$$

subG is subE with $\nu = \sigma$ and $\alpha \to 0^+$.

Chernoff:

$$\mathbb{P}[X-\mu \geq t] \leq \begin{cases} e^{-t^2/2\nu^2} & \text{if } 0 \leq t \leq \frac{\nu^2}{\alpha} \\ e^{-t/2\alpha} & \text{for } t > \nu^2/\alpha \end{cases}$$

Bernstein Condition: Given X B.C. with param b holds

if $\left|\mathbb{E}[(X-\mu)^k]\right| \leq \frac{1}{2} k! \, \sigma^2 b^{k-2}$ for $k \geq 2$

X bounded $\Rightarrow$ B.C. holds

If X satisfies B.C. then X is sub$E$ with $(\sqrt{2}\sigma, 2b)$.

$$\mathbb{E}\left[e^{\lambda(X-\mu)}\right] \leq e^{\lambda^2(\sqrt{2}\sigma)^2/2} \qquad \forall \, |\lambda| \leq \frac{1}{2b}$$

Bernstein-type bounds.

X satisfies B.C. then: $\mathbb{E}[e^{\lambda(X-\mu)}] \leq e^{\frac{\lambda^2\sigma^2/2}{1-b|\lambda|}} \quad \forall |\lambda| \leq \frac{1}{b}$

setting $\lambda = \frac{t}{bt+\sigma^2} \Rightarrow \mathbb{P}[|X-\mu| \geq t] \leq 2\exp\left(\frac{-t^2}{2(\sigma^2+bt)}\right) \quad \forall t \geq 0$

$|X-\mu| \leq b$ a.s. $\Rightarrow \sigma^2 \leq b^2$

if $\sigma^2 \ll b^2$ then Bernstein is much tighter than Hoeffding.

Martingale Based Bounds.

$\{Y_k\}_{k=1}^{\infty}$ $\quad$ Y_k is $\mathcal{F}_k$-measurable

Filteration: $\mathcal{F}_{k-1} \subseteq \mathcal{F}_k$ $\quad$ $\{\mathcal{F}_k\}$ is a filteration.

$\{(Y_k, \mathcal{F}_k)\}$ is a martingale given $\{Y_k\}$ is adapted to $\{\mathcal{F}_k\}$ and:

forall $k \geq 1$ $\quad$ $\mathbb{E}[|Y_k|] < \infty$, $\mathbb{E}[Y_{k+1} | \mathcal{F}_k] = Y_k$

$\{X_k\}_{k=1}^{\infty}$ s.t. $\mathcal{F}_k = \sigma(X_1, \dots, X_k)$

e.g.

Partial Sums: $S_k = \sum X_i$ iid

$\mathbb{E}[S_{k+1} \mid \mathcal{F}_k] = \mathbb{E}[X_{k+1}] + S_k$

$\Rightarrow Y_k = S_k - k\mu \Rightarrow \{Y_k\}$ defines a martingale.

Consider $f(X_1, \ldots, X_n) : \mathbb{R}^n \longrightarrow \mathbb{R}$

$\{X_k\}_{k=1}^n$ independent.

$$Y_0 = \mathbb{E}[f(X)] \qquad Y_k = \mathbb{E}[f(X) \mid X_1, \ldots, X_k] \qquad Y_n = f(X)$$

Doob Martingale.

$$f(X) - \mathbb{E}[f(X)] = \sum_{k=1}^{n} \underbrace{Y_k - Y_{k-1}}_{D_k}$$

$\{(D_k, \mathcal{F}_k)\}_{k=1}^{\infty}$ is a martingale difference sequence.

Azuma-Hoeffding: If $D_k \in [a_k, b_k]$ a.s.

then $$\mathbb{P}\left[\left|\sum D_k\right| \geq t\right] \leq 2\exp\left[\frac{-2t^2}{\sum (b_k - a_k)^2}\right]$$

McDiarmid's Ineq.

X_i's are independent $\quad f(X)$ has bounded difference property.

$$|f(x_1, \ldots, x_i, \ldots, x_n) - f(x_1, \ldots, x_i', \ldots, x_n)| \leq c_i$$

It can be proved that $Y_k - Y_{k-1}$ is in a range of length c_i.

Plug into Az-Hoef

$$\Longrightarrow \quad \mathbb{P}\left[f(X) - \mathbb{E}[f(X)] \geq t\right] \leq \exp\left[\frac{-2t^2}{\sum_{k=1}^{n} c_k^2}\right]$$