σ- Subgaussian random variables

$$M_X(\lambda) = \mathbb{E}[e^{\lambda X}] \leq e^{\lambda^2 \sigma^2 / 2}$$

if X is σ- subgaussian

(i) $\mathbb{E}[X] = 0$

(ii) $\mathrm{Var}[X] \leq \sigma^2$

(iii) cX is $|c|\sigma$ - subgaussian

(iv) if x_1 and x_2 are independent and subgaussian with σ_1 and σ_2, then $x_1 + x_2$ is $\sqrt{\sigma_1^2 + \sigma_2^2}$ - subgaussian

Theorem : $P(|X| \geq \varepsilon) \leq 2\, e^{-\left(\frac{\varepsilon^2}{2\sigma^2}\right)}$ — (1)

|||

with probability greater than $1-\delta$, the random variable takes values in the (confidence) interval $\left(-\sqrt{2\sigma^2 \log(2/\delta)}\,,\; +\sqrt{2\sigma^2\log(2/\delta)}\right)$

(let $\delta = 2e^{-\left(\frac{\varepsilon^2}{2\sigma^2}\right)}$ in ①)

We need $\mathbb{P}(|\text{Sample Mean} - \text{True Mean}| \geq \varepsilon)$

$\hat{\mu} = \frac{S_n}{n}$, $\mathbb{P}(\hat{\mu} - \mu \geq \varepsilon)$?

Sample Mean

$$\mathbb{P}(\hat{\mu} - \mu \geq \varepsilon) = \mathbb{P}\left(\frac{S_n}{n} - \mu \geq \varepsilon\right)$$

$$= \mathbb{P}\left(\frac{S_n - n\mu}{n} \geq \varepsilon\right)$$

$$\frac{S_n - n\mu}{n} = \frac{(X_1 - \mu) + (X_2 - \mu) + \cdots + (X_n - \mu)}{n}$$

$$= \underset{\frac{\sigma}{n}}{\frac{(X_1 - \mu)}{n}} + \underset{\frac{\sigma}{n}}{\frac{(X_2 - \mu)}{n}} + \cdots + \underset{\frac{\sigma}{n}}{\frac{(X_n - \mu)}{n}}$$

if $(X_1 - \mu)$ is σ-subgaussian, then

$\frac{X_1 - \mu}{n}$ is $\frac{\sigma}{n}$-subgaussian

$\frac{S_n - n\mu}{n}$ is $\sqrt{\left(\frac{\sigma}{n}\right)^2 + \cdots + \left(\frac{\sigma}{n}\right)^2}$

n times

$$\sqrt{n \frac{\sigma^2}{n^2}} = \frac{\sigma}{\sqrt{n}}$$

$\frac{S_n - n\mu}{n}$ is $\frac{\sigma}{\sqrt{n}}$ -subgaussian

$$\mathbb{P}(\hat{\mu} - \mu \geq \varepsilon) = \mathbb{P}\left(\frac{S_n - n\mu}{n} \geq \varepsilon\right)$$

$$\leq e^{-\left(\frac{\varepsilon^2}{2(\sigma/\sqrt{n})^2}\right)}$$

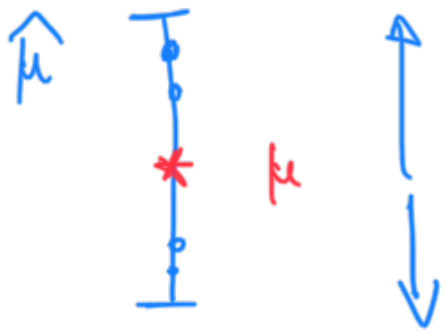
$$= e^{-\left(\frac{n\varepsilon^2}{2\sigma^2}\right)}$$

$$\mathbb{P}(|\hat{\mu} - \mu| \geq \varepsilon) \leq 2e^{-\left(\frac{n\varepsilon^2}{2\sigma^2}\right)}$$

|||

with probability greater than $1-\delta$

$$\hat{\mu} \in \Bigg[\underbrace{\mu - \sqrt{\frac{2\sigma^2 \log(2/\delta)}{n}}, \ \mu + \sqrt{\frac{2\sigma^2 \log(2/\delta)}{n}}}_{\text{Confidence interval}} \Bigg]$$



$$+\sqrt{\frac{2\sigma^2 \log(2/\delta)}{n}}$$

$$-\sqrt{\frac{2\sigma^2 \log(2/\delta)}{n}}$$

Our next goal: $X \in [a,b]$ then X is subgaussian

Lemma II: $\mathrm{Var}[X] = \min_{c} \mathbb{E}[(X-c)^2]$

$$\mathbb{E}[(X-c)^2] = \mathbb{E}[X^2] + \mathbb{E}[c^2] - 2\,\mathbb{E}[X.c]$$

$$= \mathbb{E}[X^2] + c^2 - 2c\,\mathbb{E}[X]$$

differentiate w.r.t c

$$2c_* - 2\mathbb{E}[X] = 0 \Rightarrow c_* = \mathbb{E}[X]$$

$$\mathrm{Var}[X] = \mathbb{E}[(X - \mathbb{E}[X])^2]$$

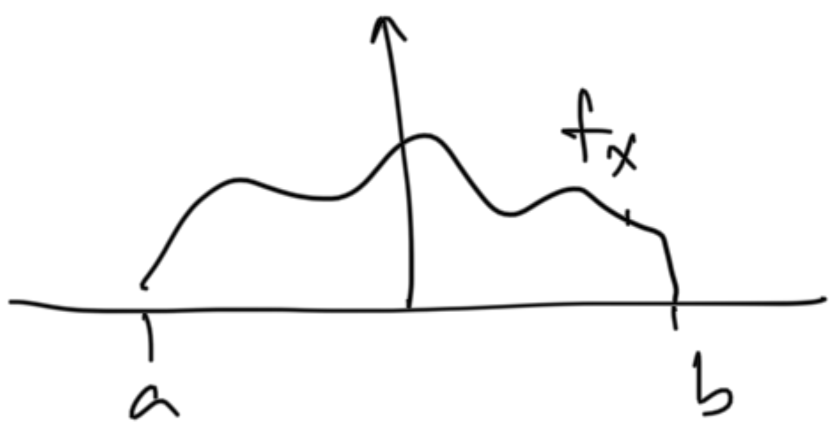
Lemma I: $X \in [a, b]$ we have

$$\mathrm{Var}[X] \leq \frac{(b-a)^2}{4}$$

Proof: $\mathrm{Var}[X] = \min_c \mathbb{E}[(X-c)^2]$

(put $c = \frac{a+b}{2}$) $\leq \mathbb{E}\left[\left(X - \frac{(b+a)}{2}\right)^2\right]$

$\downarrow$



$$\int_a^b \left(x - \frac{(b+a)}{2}\right)^2 f_X(x)\,dx$$

$$\leq \int_a^b \max(e_1, e_2)\, f_X(x)\,dx$$

$$= \max(e_1, e_2) \int_a^b f_X(x)\,dx$$

extreme values of $\left(x - \frac{(b+a)}{2}\right)^2$ is either $e_1 = \left(a - \frac{(b+a)}{2}\right)^2$ $e_2 = \left(b - \frac{(b+a)}{2}\right)^2$

$$\text{Var}[X] \leq \max\left\{ \left(a - \frac{(b+a)}{2}\right)^2, \left(b - \frac{(b+a)}{2}\right)^2 \right\}$$

$$= \frac{(b-a)^2}{4}$$

$\vdots$

our next goal : $X \in [a, b]$ then X is

subgaussian (X is mean 0)

Proof: $M_X(\lambda) = \mathbb{E}[e^{\lambda X}]$

$$\psi(\lambda) = \log M_X(\lambda)$$

$$= \log \mathbb{E}[e^{\lambda X}]$$

$$\psi(0) = \log \mathbb{E}[e^{0X}] = \log \mathbb{E}[1] = \log 1 = 0$$

$$\psi'(\lambda) = \frac{\mathbb{E}[X e^{\lambda X}]}{\mathbb{E}[e^{\lambda X}]}, \quad \psi'(0) = \frac{\mathbb{E}[X e^{0}]}{\mathbb{E}[e^{0}]}$$

$$= \mathbb{E}[X] = 0$$

look at $\psi'(\lambda) = \dfrac{1}{\mathbb{E}[e^{\lambda X}]} \displaystyle\int_{-\infty}^{\infty} x e^{\lambda x} f(x)\, dx$

forget the $\underline{\dfrac{1}{\mathbb{E}[e^{\lambda x}]}}$ factor for a moment

$$\int_{-\infty}^{\infty} x e^{\lambda x} f(x)\, dx = \int_{-\infty}^{\infty} x \cdot g_{\lambda}(a)\, dx$$

$$= \mathbb{E}_{G}[X]$$

g_λ is a different density

$$g_\lambda(x) = M \cdot e^{\lambda x} f(x)$$

$$\int_{-\infty}^{\infty} g_\lambda(x)\,dx = \int_{-\infty}^{\infty} M \cdot e^{\lambda x} f(x)\,dx$$

$$= M \int_{-\infty}^{\infty} e^{\lambda x} f(x)\,dx$$

$$= M\, \mathbb{E}_f[e^{\lambda x}]$$

$$M\, \mathbb{E}_f[e^{\lambda x}] = 1, \qquad M = \frac{1}{\mathbb{E}[e^{\lambda x}]}$$

$$\psi'(\lambda) = \int_{-\infty} x \cdot \underbrace{\frac{e \quad f(x)}{\mathbb{E}[e^{\lambda x}]}}_{g_\lambda} \, dx$$

$$\psi'(\lambda) = \mathbb{E}_{g_\lambda}[X]$$

$$\psi'(\lambda) = \frac{\mathbb{E}[X e^{\lambda X}]}{\mathbb{E}[e^{\lambda X}]}$$

$$\psi''(\lambda) = \frac{\mathbb{E}[X^2 e^{\lambda X}]}{\mathbb{E}[e^{\lambda X}]} - \left(\frac{\mathbb{E}[X e^{\lambda X}]}{\mathbb{E}[e^{\lambda X}]}\right)^2$$

$$\psi''(\lambda) = \mathrm{Var}_{g_\lambda}[X]$$

Invoke Taylor theorm

$$\psi(\lambda) = \psi(0) + \psi'(0)\lambda + \psi''(\tilde{\lambda})\frac{\lambda^2}{2}$$

$$\tilde{\lambda} \in [0, \lambda]$$

$$= 0 + 0 + \frac{\lambda^2}{2} \, \mathrm{Var}_{g_{\tilde{\lambda}}}[X]$$

$$\leq \frac{\lambda^2}{2} \, \frac{(b-a)^2}{4}$$

$$= \frac{\lambda^2 (b-a)^2}{8}$$

$$M_X(\lambda) = e^{\psi(\lambda)}$$

$$\leq e^{\frac{\lambda^2 (b-a)^2}{8}}$$

$$\mathbb{E}[e^{\lambda X}] \leq e^{\frac{\lambda^2 (b-a)^2}{2 \times 4}}$$

$X \in [a, b]$ is subgaussian with

$$\sigma = \frac{(b-a)^2}{4}$$

Summary of concentration inequalities

Chebyshev $\quad \mathbb{P}(|\hat{\mu} - \mu| \geq \varepsilon) \leq \frac{\sigma^2}{n\varepsilon^2}$ (rate $\frac{1}{n}$)

Cramer chernoff $\quad \mathbb{P}(|\hat{\mu} - \mu| \geq \varepsilon) \leq 2e^{-\left(\frac{n\varepsilon^2}{2\sigma^2}\right)}$ rate is (exponential)

Chebyshev $\quad \hat{\mu} \in \left[\mu - \sqrt{\frac{\sigma^2}{n\delta}}, \mu + \sqrt{\frac{\sigma^2}{n\delta}}\right]$

as $\delta \to 0$, $\to \infty$

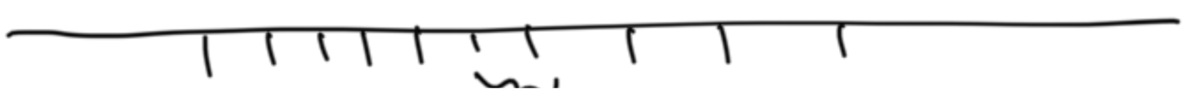
Cramer chernoff $\quad \hat{\mu} \in \left[\mu - \sqrt{\frac{2\sigma^2 \log(2/\delta)}{n}}, \mu + \sqrt{\frac{2\sigma^2 \log(2/\delta)}{n}}\right]$

we are not doing any better than statistical rate, if I take samples from a random variable the standard deviation is of order $\left(\frac{1}{\sqrt{n}}\right)$

$\delta = 0.1$ to $\delta = 0.01$

$\log(2/\delta)$ becomes only 2 times more

$$n = \frac{2\sigma^2 \log(2/\delta)}{\varepsilon^2}$$ (fixing ε and δ)



ϵ