

第四讲

1 Chernoff 技巧

若存在 $b > 0$ 使得对任意 $|\lambda| \leq b$ 函数 $\psi(\lambda) = \mathbb{E}[e^{\lambda(X-\mu)}]$ 存在, 则对任意 $\lambda \in [0, b]$, 由 Markov 不等式, 有

$$\mathbb{P}[X - \mu \geq t] = \mathbb{P}[e^{\lambda(X-\mu)} \geq e^{\lambda t}] \leq \frac{\mathbb{E}[e^{\lambda(X-\mu)}]}{e^{\lambda t}}. \quad (1)$$

上述不等式关于 λ 最小化得到此项的最小上界, 即

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

命题 1.1

设 $X \sim \mathcal{N}(\mu, \sigma^2)$, 则 $\mathbb{P}[X - \mu \geq t] \leq e^{-\frac{t^2}{2\sigma^2}}, \quad \forall t \geq 0$.

证明. 由直接计算可得, 对任意 $\lambda \in \mathbb{R}$,

$$\mathbb{E}[e^{\lambda X}] = e^{\mu\lambda + \frac{\sigma^2\lambda^2}{2}}$$

由 Chernoff 技巧, 对任意 $t \geq 0$

$$\begin{aligned} \mathbb{P}[X - \mu \geq t] &\leq \inf_{\lambda \in [0, +\infty)} \frac{\mathbb{E}[e^{\lambda(X-\mu)}]}{e^{\lambda t}} \\ &= \inf_{\lambda \in [0, +\infty)} e^{-\lambda t + \frac{\sigma^2\lambda^2}{2}} \end{aligned}$$

求 $e^{-\lambda t + \frac{\sigma^2\lambda^2}{2}}$ 的最小值点等价求 $\log e^{-\lambda t + \frac{\sigma^2\lambda^2}{2}}$ 的最小值点, 后者的最小值点为 $\frac{t}{\sigma^2}$.

因此,

$$\begin{aligned} \mathbb{P}[X - \mu \geq t] &\leq \inf_{\lambda \in [0, +\infty)} e^{-\lambda t + \frac{\sigma^2\lambda^2}{2}} \\ &= e^{-\frac{t}{\sigma^2}t + \frac{\sigma^2}{2}(\frac{t}{\sigma^2})^2} = e^{-\frac{t^2}{2\sigma^2}}. \end{aligned}$$

□

2 Sub-Gaussian 变量

定义 2.1: Sub-Gaussian 变量

设有随机变量 X , 其期望 $\mathbb{E}[X] = \mu$, 若存在正数 $\sigma > 0$ 使得

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

则称 X 为 *sub-Gaussian* 随机变量, 称 σ 为 *sub-Gaussian* 参数.

评论. X 为 sub-Gaussian 当且仅当 $-X$ 为 sub-Gaussian.

由 Chernoff 技巧, 若 X 为 sub-Gaussian 随机变量, 对任意 $t \geq 0$

$$\mathbb{P}[X - \mu \geq t] \leq e^{-\frac{t^2}{2\sigma^2}},$$

$-X$ 也为 sub-Gaussian 随机变量, 同理

$$\mathbb{P}[-X + \mu \geq t] = \mathbb{P}[X - \mu \leq -t] \leq e^{-\frac{t^2}{2\sigma^2}}.$$

因此, 参数为 σ 的 sub-Gaussian 随机变量 X 有尾界

$$\mathbb{P}[|X - \mu| \geq t] = \mathbb{P}[X - \mu \geq t] + \mathbb{P}[X - \mu \leq -t] \leq 2e^{-\frac{t^2}{2\sigma^2}}.$$

Rademacher 变量 Rademacher 随机变量 ϵ 等概率取 $+1$ 或 -1 , 为参数 $\sigma = 1$ 的 sub-Gaussian 变量.

显然 $\mathbb{E}[\epsilon] = 0$, 对任意 $\lambda \in \mathbb{R}$,

$$\begin{aligned} \mathbb{E}[e^{\lambda\epsilon}] &= \frac{1}{2}e^{-\lambda} + \frac{1}{2}e^{\lambda} \\ &= \frac{1}{2} \left\{ \sum_{k=0}^{\infty} \frac{(-\lambda)^k}{k!} + \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} \right\} \\ &= \sum_{k=0}^{\infty} \frac{\lambda^{2k}}{(2k)!} \\ &\leq \sum_{k=0}^{\infty} \frac{\lambda^{2k}}{2^k k!} = e^{\frac{\lambda^2}{2}}. \end{aligned}$$

引理 2.1

设 X 为有界随机变量, 其支撑集为 $[a, b]$, 则随机变量 X 是以 $\sigma = b - a$ 为参数的 *sub-Gaussian* 随机变量.

证明. 不失一般性设 $\mathbb{E}X = 0$.

设 X' 为独立于 X 的来自同一分布的随机变量, ϵ 为 Rademacher 随机变量.

对任意 $\lambda \in \mathbb{R}$,

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

最后的不等号源自 **Jensen** 不等式.

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

$$\mathbb{E}_{X, X'}[e^{\lambda(X - X')}] = \mathbb{E}_{X, X'} \left[\mathbb{E}_\epsilon e^{\lambda(X - X')\epsilon} \right] \leq \mathbb{E}_{X, X'} \left[e^{\frac{\lambda^2(X - X')^2}{2}} \right] \leq e^{\frac{\lambda^2(b-a)^2}{2}}.$$

其中第一个不等号利用 **Rademacher** 随机变量的性质, 第二个不等号源自随机变量的有界性. \square

评论. 上述证明过程通过引入同一分布的独立随机变量 X' 和 **Rademacher** 变量实现, 这种方法称为**对称化技巧**.

引理 2.2

设 X_1, X_2 分别为独立的, 以 σ_1, σ_2 为参数的 *sub-Gaussian* 随机变量, 则 $X_1 + X_2$ 为以 $\sqrt{\sigma_1^2 + \sigma_2^2}$ 为参数的 *sub-Gaussian* 随机变量.

3 Hoeffding 界

综合上述结论, 可以得到以下定理.

定理 3.1: Hoeffding 界

设以 σ_i 为参数的 *sub-Gaussian* 随机变量 $X_i, i = 1, \dots, n$ 相互独立, 且 $\mathbb{E}[X_i] = \mu_i$, 则对任意 $t \geq 0$, 有

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

证明. 对任意 $\lambda \geq 0$

$$\begin{aligned} \mathbb{P} \left[\sum_{i=1}^n (X_i - \mu_i) \geq t \right] &= \mathbb{P} \left[\exp \left(\lambda \sum_{i=1}^n (X_i - \mu_i) \right) \geq e^{\lambda t} \right] \\ &\leq e^{-\lambda t} \mathbb{E} \left[\exp \left(\lambda \sum_{i=1}^n (X_i - \mu_i) \right) \right] \\ &= e^{-\lambda t} \mathbb{E} \left[\prod_{i=1}^n e^{\lambda(X_i - \mu_i)} \right] \\ &= e^{-\lambda t} \prod_{i=1}^n \mathbb{E} e^{\lambda(X_i - \mu_i)} \leq e^{-\lambda t} + \prod_{i=1}^n e^{\frac{\lambda^2 \sigma_i^2}{2}}. \end{aligned}$$

不等式右端关于 λ 最小化即完成证明. □

投硬币回顾 设 X_i 为 Bernoulli 变量, Z_i 为 Rademacher 变量, 则有 $Z_i = 2X_i - 1$, 已知 Z_i 是参数为 1 的 sub-Gaussian 变量, 应用 Hoeffding 不等式, 有

$$\mathbb{P} \left[\sum_{i=1}^n Z_i \geq t \right] \leq \exp \left(-\frac{t^2}{2n} \right), \quad \forall t \geq 0.$$

回顾第三讲中投掷硬币的例子

$$\mathbb{P} \left[\sum_{i=1}^n X_i \geq \frac{3}{4}n \right] = \mathbb{P} \left[\sum_{i=1}^n Z_i \geq \frac{1}{2}n \right] \leq \exp \left(-\frac{n}{8} \right).$$

由此得到比应用中心极限定理及 Markov 不等式更为准确的估计.

4 Chernoff 不等式及其在随机图上的应用

TD