第四讲

1 Chernoff 技巧

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

$$\mathbb{P}[X-\mu \ge t] = \mathbb{P}\left[e^{\lambda(X-\mu)} \ge e^{\lambda t}\right] \le \frac{\mathbb{E}\left[e^{\lambda(X-\mu)}\right]}{e^{\lambda t}}. \tag{1}$$

上述不等式关于 λ 最小化得到此项的最小上界, 即

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

命题 1.1

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

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

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

由 Chernoff 技巧, 对任意 $t \ge 0$

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

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,

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

□

2 Sub-Gaussain 变量

定义 2.1: Sub-Gaussain 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

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

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 $t\ge 0$

$$\mathbb{P}\left[X-\mu\ge t\right]\le e^{-\frac{t^2}{2\sigma^2}},$$

$-X$ 也为 sub-Gaussian 随机变量, 同理

$$\mathbb{P}\left[-X+\mu\ge t\right]=\mathbb{P}\left[X-\mu\le -t\right]\le e^{-\frac{t^2}{2\sigma^2}}.$$

因此, 参数为 σ 的 sub-Gaussian 随机变量 X 有尾界

$$\mathbb{P}\left[|X-\mu|\ge t\right]=\mathbb{P}\left[X-\mu\ge t\right]+\mathbb{P}\left[X-\mu\le -t\right]\le 2e^{-\frac{t^2}{2\sigma^2}}.$$

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 $\lambda\in\mathbb{R}$,

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

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 $\lambda \in \mathbb{R}$,

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

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

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

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. □

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 $t \ge 0$, 有

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

证明. 对任意 $\lambda \ge 0$

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

不等式右端关于 λ 最小化即完成证明. $\square$

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

定理 4.1: Chernorff 不等式

设 X_i 为独立的, 以 p_i 为参数的伯努利随机变量, 其 N 项和 $S_N = \sum_{i=1}^{N} X_i$, 记 $\mathbb{E}S_N = \sum_{i=1}^{N} p_i = \mu$, 则对于任意 $t > \mu$, 有

$$\mathbb{P}\left[S_N \geq t\right] \leq e^{-\mu}\left(\frac{e\mu}{t}\right)^t.$$

证明. 类似 Hoeffding 不等式证明, 这里使用 Chernoff 技巧, 取 $\lambda \geq 0$.

$$\mathbb{P}\left[S_N \geq t\right] = \mathbb{P}\left[e^{\lambda S_N} \geq e^{\lambda t}\right] \leq e^{-\lambda t}\mathbb{E}\left[e^{\lambda S_N}\right] = e^{-\lambda t}\prod_{i=1}^{N}\mathbb{E}e^{\lambda X_i}.$$

由于

$$\mathbb{E}e^{\lambda X_i} = (1-p_i)e^{\lambda \cdot 0} + p_i e^{\lambda} = 1 - p_i + p_i e^{\lambda} \leq \exp((e^{\lambda}-1)p_i),$$

其中不等号源于不等式 $1 + x \leq e^x$.

故

$$\begin{aligned}\mathbb{P}\left[S_N \geq t\right] &\leq e^{-\lambda t}\prod_{i=1}^{N}\exp((e^{\lambda}-1)p_i) \\ &= e^{-\lambda t}\exp\left((e^{\lambda}-1)\sum_{i=1}^{N} p_i\right) \\ &= e^{-\lambda t}\exp\left((e^{\lambda}-1)\mu\right)\end{aligned}$$

不等式右端关于 λ 最小化, 取 $\lambda = \ln\frac{t}{\mu}$, 代入上述不等式, 即有

$$\mathbb{P}\left[S_N \geq t\right] \leq e^{-\mu}\left(\frac{e\mu}{t}\right)^t.$$

$\square$

类似地, 可以得到以下定理

定理 4.2: Chernorff 不等式下界

设 X_i 为独立的, 以 p_i 为参数的伯努利随机变量, 其 N 项和 $S_N = \sum_{i=1}^{N} X_i$, 记 $\mathbb{E}S_N = \sum_{i=1}^{N} p_i = \mu$, 则对于任意 $0 < t < \mu$, 有

$$\mathbb{P}\left[S_N \le t\right] \le e^{-\mu}\left(\frac{e\mu}{t}\right)^t.$$

证明. 对任意 $\lambda \le 0$,

$$\mathbb{P}[S_N \le t] = \mathbb{P}[\lambda S_N \ge \lambda t] = \mathbb{P}\left[e^{\lambda S_N} \ge e^{\lambda t}\right],$$

证明余下部分与定理 4.2 证明相同. $\square$

综合定理 4.1 和定理 4.2, 可以得到以下不等式尾界

引理 4.1

对任意 $x \in [0,1]$, 有不等式

$$x - (1+x)\log(1+x) \le -\frac{x^2}{6}.$$

定理 4.3: Chernoff 不等式尾界

在定理 4.1 条件下, 对任意 $\delta \in (0,1]$,

$$\mathbb{P}\left[|S_N - \mu| \ge \delta\mu\right] \le 2e^{-\mu\delta^2/6}.$$

证明. 若取 $t = (1+\delta)\mu > \mu$, 由定理 4.1,

$$\begin{aligned}\mathbb{P}\left[S_N \ge (1+\delta)\mu\right] &\le e^{-\mu}\left(\frac{e\mu}{(1+\delta)\mu}\right)^{(1+\delta)\mu} = e^{\delta\mu}\left(\frac{1}{1+\delta}\right)^{(1+\delta)\mu} \\ &= \exp\left\{\mu\left(\delta - (1+\delta)\log(1+\delta)\right)\right\} \\ &\le e^{-\mu\delta^2/6}.\end{aligned}$$

最后一个不等号取自引理 4.1.

若取 $t=(1-\delta)\mu<\mu$, 由定理 4.2,

$$\begin{aligned}\mathbb{P}\left[S_N \le (1-\delta)\mu\right] \le e^{-\mu}\left(\frac{e\mu}{(1-\delta)\mu}\right)^{(1-\delta)\mu} &= e^{-\delta\mu}\left(\frac{1}{1-\delta}\right)^{(1-\delta)\mu} \\ &= \exp\left\{\mu\left(-\delta-(1-\delta)\log(1-\delta)\right)\right\} \\ &\le e^{-\mu\delta^2/6}.\end{aligned}$$

综上, 有

$$\mathbb{P}\left[|S_N-\mu|\ge\delta\mu\right]=\mathbb{P}\left[S_N\ge(1+\delta)\mu\right]+\mathbb{P}\left[S_N\le(1-\delta)\mu\right]\le 2e^{-\mu\delta^2/6}.$$

□

定义 4.1: Eröds-Rényi 模型

给定 n 个节点, 节点间连边相互独立且服从以 p 为参数的伯努利分布, 记此随机图为 $\mathcal{G}(n,p)$.

在 Eröds-Rényi 模型中, 随机图每一个节点度的期望为 $(n-1)p$. 由 Chernoff 尾界, 当随机图的节点度期望足够大时, 每一节点度几乎是相同的.

命题 4.1

考虑随机图 $G\sim\mathcal{G}(n,p)$, 对任意 $\delta\in(0,1)$, 若 $d:=(n-1)p\ge\frac{6}{\delta^2}\log\frac{2n}{\delta}$, 则有

$$\mathbb{P}\left[\forall i, |d_i-d|\le\delta d\right]\ge 1-\delta$$

其中 d_i 为节点 i 的度.

证明. 固定任意节点 i, 由 Chernoff 不等式尾界,

$$\mathbb{P}[|d_i-d|\ge\delta d]\le 2e^{-d\delta^2/6}$$

由事件联合界

$$\mathbb{P}[\exists i, |d_i-d|\ge\delta d]\le\sum_{i=1}^{n}\mathbb{P}[|d_i-d|\ge\delta d]\le 2ne^{-d\delta^2/6}.$$

令 $\delta\ge 2ne^{-d\delta^2/6}$, 解得

$$d\ge\frac{6}{\delta^2}\log\frac{2n}{\delta}.$$

此时

$$\mathbb{P}\left[\forall i, |d_i-d|\le\delta d\right]=1-\mathbb{P}[\exists i, |d_i-d|\ge\delta d]\ge 1-\delta.$$

□