设 $a=(E[|X|^p])^{1/p}$ 和 $b=(E[|Y|^q])^{1/q}$, 根据上述不等式有

$$\frac{|XY|}{ab}=\frac{|X|}{a}\frac{|Y|}{b}\leqslant\frac{1}{p}\frac{|X|^p}{a^p}+\frac{1}{q}\frac{|Y|^q}{b^q}.$$

对上式两边同时取期望有

$$\frac{E[|XY|]}{ab}\leqslant\frac{1}{p}\frac{E[|X|^p]}{c^p}+\frac{1}{q}\frac{E[|Y|^q]}{d^q}=\frac{1}{p}+\frac{1}{q}=1,$$

从而完成证明.

7.2 Chernoff 不等式

Chernoff 方法 是证明集中不等式一种有效的工具, 首先介绍；其基本原理: 设 X 是一个随机变量, 对给定任意 $t>0$ 和 $\epsilon>0$, 利用 Markov 不等式有

$$\begin{aligned}P\left[X\geqslant E[X]+\epsilon\right] &= P\left[tX\geqslant tE[X]+t\epsilon\right]\\ &=P\left[\exp(tX)\geqslant\exp(tE[X]+t\epsilon)\right]\leqslant\exp(-t\epsilon-tE[X])E[\exp(tX)],\end{aligned}$$

由于上述不等式对任意 $t>0$ 都成立, 于是有

$$P\left[X\geqslant E[X]+\epsilon\right]\leqslant\min_{t>0}\left\{\exp(-t\epsilon-tE[X])E[\exp(tX)]\right\}.$$

类似地, 对任意 $\epsilon>0$ 和 $t<0$ 有

$$\begin{aligned}P\left[X\leqslant E[X]-\epsilon\right] &= P\left[tX\geqslant tE[X]-t\epsilon\right]\\ &= P\left[\exp(tX)\geqslant\exp(tE[X]-t\epsilon)\right]\leqslant\exp(t\epsilon-tE[X])E[\exp(tX)],\end{aligned}$$

同理可得

$$P[X\leqslant E[X]-\epsilon]\leqslant\min_{t<0}\left\{\exp(t\epsilon-tE[X])E[\exp(tX)]\right\}.$$

这种通过引入指数函数和变量 t 方法求解集中不等式的方法称为 Chernoff 方法. 后续将根据随机变量的特征来计算 $E[e^{tX_i}]$.

定理 7.5 若独立同分布的随机变量 $X_1,X_2,\cdots,X_n$ 满足 $P(X_i=+1)=P(X_i=-1)=1/2$, 对任意 $\epsilon>0$ 有

$$P\left(\frac{1}{n}\sum_{i=1}^n X_i\geqslant\epsilon\right)\leqslant\exp(-n\epsilon^2/2)\quad\text{和}\quad P\left(\frac{1}{n}\sum_{i=1}^n X_i\leqslant-\epsilon\right)\leqslant\exp(-n\epsilon^2/2).$$

证明　对任意 $t > 0$, 根据独立性和 Chernoff 方法有

$$P\left(\frac{1}{n}\sum_{i=1}^n X_i \geqslant \epsilon\right) \leqslant \exp(-nt\epsilon)E\left[\exp\left(\sum_{i=1}^n tX_i\right)\right] = \exp(-nt\epsilon)\prod_{i=1}^n E\left[\exp(tX_i)\right].$$

由随机变量 $P(X_i = 1) = P(X_i = -1) = 1/2$ 有

$$E[e^{tX_i}] = \frac{1}{2}\exp(t) + \frac{1}{2}\exp(-t) = \frac{1}{2}\sum_{i\geqslant 0}\frac{t^i}{i!} + \frac{1}{2}\sum_{i\geqslant 0}\frac{(-1)^i t^i}{i!} = \sum_{i\geqslant 0}\frac{t^{2i}}{(2i)!} \leqslant \sum_{i\geqslant 0}\frac{(t^2/2)^i}{i!} = \exp(t^2/2),$$

这里利用了指数函数的 Taylor 展开式和 $(2i)! \geqslant 2^i i!$. 由此可得

$$P\left(\frac{1}{n}\sum_{i=1}^n X_i \geqslant \epsilon\right) \leqslant \inf_{t>0}\left\{\exp(-nt\epsilon + nt^2/2)\right\} = \exp(-n\epsilon^2/2).$$

上式中当 $t = \epsilon$ 时取得最小值. 同理可证明另一个不等式.

根据上述定理, 很容易得到下列推论:

推论 7.3　若独立同分布随机变量 $X_1, X_2, \cdots, X_n$ 满足 $P(X_1 = 0) = P(X_1 = 1) = 1/2$, 对任意 $\epsilon > 0$ 有

$$P\left(\frac{1}{n}\sum_{i=1}^n X_i - \frac{1}{2} \geqslant \epsilon\right) \leqslant \exp(-2n\epsilon^2) \quad \text{和} \quad P\left(\frac{1}{n}\sum_{i=1}^n X_i - \frac{1}{2} \leqslant -\epsilon\right) \leqslant \exp(-2n\epsilon^2).$$

证明　设随机变量 $Y_i = 2X_i - 1$, 易得 $P(Y_i = -1) = P(Y_i = 1) = 1/2$, 于是有

$$P\left(\frac{1}{n}\sum_{i=1}^n X_i - \frac{1}{2} \geqslant \epsilon\right) = P\left(\frac{1}{n}\sum_{i=1}^n (2X_i - 1) \geqslant 2\epsilon\right) = P\left(\frac{1}{n}\sum_{i=1}^n Y_i \geqslant 2\epsilon\right) \leqslant \exp(-2n\epsilon^2),$$

最后一个不等式根据定理 7.5 所得.

针对标准正态分布的随机变量有

定理 7.6　若独立同分布的随机变量 $X_1, X_2, \cdots, X_n$ 服从标准正态分布, 对任意 $\epsilon > 0$ 有

$$P\left(\frac{1}{n}\sum_{i=1}^n X_i \geqslant \epsilon\right) \leqslant \exp(-n\epsilon^2/2) \quad \text{和} \quad P\left(\frac{1}{n}\sum_{i=1}^n X_i \leqslant -\epsilon\right) \leqslant \exp(-n\epsilon^2/2).$$

证明　对任意 $t > 0$, 根据独立性和 Chernoff 方法有

$$P\left(\frac{1}{n}\sum_{i=1}^n X_i \geqslant \epsilon\right) \leqslant \exp(-nt\epsilon)E\left[\exp\left(t\sum_{i=1}^n X_i\right)\right] = \exp(-nt\epsilon)\prod_{i=1}^n E[\exp(tX_i)].$$

根据标准正态分布的密度函数有

$$E[\exp(tX_i)] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{xt} e^{-x^2/2} dx = e^{t^2/2} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-(x-t)^2/2} = \exp(t^2/2).$$

由此可知

$$P\left(\frac{1}{n}\sum_{i=1}^{n} X_i \geqslant \epsilon\right) \leqslant \inf_{t>0} \left\{\exp(-nt\epsilon + nt^2/2)\right\} = \exp(-n\epsilon^2/2).$$

上式中 $t = \epsilon$ 时取得最小值. 同理可证明另一个不等式.

根据上述定理可知

推论 7.4 若独立同分布随机变量 $X_1, X_2, \cdots, X_n$ 服从正态分布 $N(\mu, \sigma^2)$, 对任意 $\epsilon > 0$ 有

$$P\left(\frac{1}{n}\sum_{i=1}^{n} X_i - \mu \geqslant \epsilon\right) \leqslant \exp(-n\epsilon^2/2\sigma^2) \quad \text{和} \quad P\left(\frac{1}{n}\sum_{i=1}^{n} X_i - \mu \leqslant -\epsilon\right) \leqslant \exp(-n\epsilon^2/2\sigma^2).$$

7.2.1 有界随机变量的 Chernoff 不等式

下面研究有界的随机变量 $X \in [a, b]$ 的 Chernoff 不等式, 首先介绍 Chernoff 引理.

引理 7.1 若随机变量 $X \in [0, 1]$ 的期望为 $\mu = E[X]$, 对任意 $t > 0$ 有

$$E[\exp(tX)] \leqslant \exp(t\mu + t^2/8).$$

证明 由凸函数的性质可知

$$\exp(tX) = \exp(tX + (1-X)0) \leqslant X\exp(t) + (1-X)\exp(0),$$

对上述不等式两边同时对随机变量 X 取期望有

$$E[\exp(tX)] \leqslant 1 - \mu + \mu\exp(t) = \exp(\ln(1 - \mu + \mu e^t)).$$

设函数 $f(t) = \ln(1 - \mu + \mu e^t)$, 于是有 $f(0) = 0$ 和一阶导数

$$f'(t) = \frac{\mu e^t}{1 - \mu + \mu e^t},$$

由此可得 $f'(0) = \mu$. 进一步可求二阶导数

$$f''(t) = \frac{\mu e^t}{1 - \mu + \mu e^t} - \frac{\mu^2 e^{2t}}{(1 - \mu + \mu e^t)^2} \leqslant \frac{1}{4}.$$

根据泰勒中值定理存在 $\xi \in (0,t)$ 使得

$$f(t) = f(0) + tf'(0) + f''(\xi)t^2/2 \leqslant t\mu + t^2/8,$$

由此完成证明.

根据上面的 Chernoff 引理可以推导出有界的随机变量 X 满足

推论 7.5 若随机变量 $X \in [a,b]$ 的期望为 $\mu = E[X]$, 则对任意 $t > 0$ 有

$$E[\exp(tX)] \leqslant \exp(\mu t + (b-a)^2t^2/8).$$

根据上面的推论可以得到有界随机变量的 Chernoff 不等式.

定理 7.7 若相互独立的随机变量 $X_1, \ldots, X_n$ 满足 $X_i \in [a,b]$, 对任意 $\epsilon > 0$ 有

$$\begin{aligned} P\left[\frac{1}{n}\sum_{i=1}^{n}(X_i - E[X_i]) \geqslant \epsilon\right] &\leqslant \exp\left(-\frac{2n\epsilon^2}{(b-a)^2}\right), \\ P\left[\frac{1}{n}\sum_{i=1}^{n}(X_i - E[X_i]) \leqslant -\epsilon\right] &\leqslant \exp\left(-\frac{2n\epsilon^2}{(b-a)^2}\right). \end{aligned}$$

证明 对任意实数 $t > 0$, 根据独立性和 Chernoff 方法有

$$\begin{aligned} &P\left[\frac{1}{n}\sum_{i=1}^{n}(X_i - E[X_i]) \geqslant \epsilon\right] = P\left[\sum_{i=1}^{n}t(X_i - E[X_i]) \geqslant nt\epsilon\right] \\ &\qquad \leqslant \exp(-nt\epsilon)E\left[\exp\left(\sum_{i=1}^{n}t(X_i - E[X_i])\right)\right] = \exp(-nt\epsilon)\prod_{i=1}^{n}E\left[\exp(t(X_i - E[X_i]))\right]. \end{aligned}$$

根据 Chernoff 引理, 对任意随机变量 $X_i \in [a,b]$ 有

$$E\left[\exp(tX_i - tE[X_i])\right] = \exp(-tE[X_i])E\left[\exp(tX_i)\right] \leqslant \exp((b-a)^2t^2/8).$$

由此得到

$$P\left[\frac{1}{n}\sum_{i=1}^{n}X_i - \frac{1}{n}\sum_{i=1}^{n}E[X_i] \geqslant \epsilon\right] \leqslant \min_{t>0}\left\{\exp(-nt\epsilon + nt^2(b-a)^2/8)\right\} \leqslant \exp(-2n\epsilon^2/(b-a)^2).$$

在上式中 $t = 4\epsilon/(b-a)^2$ 时取得最小值, 同理可证明第二个不等式.