Ch 6 集中不等式 (Concentration)

回顾前一次课

随机变量X的矩生成函数为$M_X(t)=E[e^{tX}]$、性质

Chernoff方法： $P[X\geq\epsilon]\leq\min_{t>0}\{e^{-t\epsilon}E[e^{tX}]\}.$

0/1-随机变量$P\left[\sum_{i=1}^{n}X_i\geq(1+\epsilon)\mu\right]\leq\left(\frac{e^{\epsilon}}{(1+\epsilon)^{1+\epsilon}}\right)^{\mu}$

Rademacher随机变量： $P\left[\frac{1}{n}\sum_{i=1}^{n}X_i\geq\epsilon\right]\leq e^{-n\epsilon^2/2}$

有界：Chernoff引理

$$P\left[\frac{1}{n}\sum_{i=1}^{n}X_i-\frac{1}{n}\sum_{i=1}^{n}E[X_i]\geq\epsilon\right]\leq e^{-2n\epsilon^2/(b-a)^2}$$

Gaussian随机变量

设随机变量$X_1, \ldots, X_n$相互独立、且服从$X_i \sim \mathrm{N}(\mu, \sigma)$, 对任意$\varepsilon > 0$有

$$P\left[\frac{1}{n}\sum_{i=1}^{n}(X_i - \mu) \geq \epsilon\right] \leq \frac{1}{2}e^{-n\epsilon^2/2\sigma^2}$$

$$P\left[\frac{1}{n}\sum_{i=1}^{n}(X_i - \mu) \leq -\epsilon\right] \leq \frac{1}{2}e^{-n\epsilon^2/2\sigma^2}.$$

Sub-Gaussian随机变量

对任意$t \in (-\infty, +\infty)$, 若随机变量X满足

$$E\left[e^{(X-E[X])t}\right] \le e^{bt^2/2}$$

则称随机变量X是服从参数为b的亚高斯(Sub-Gaussian)随机变量

亚高斯随机变量表示随机变量的尾部分布不会比一个高斯分布更严重

例如

对任意有界的随机变量$X \in [a, b]$, 根据Chernoff引理有

$$E\left[e^{(X-\mu)t}\right] \leq e^{\frac{t^2(b-a)^2}{8}}$$

有界随机变量是参数为$(b-a)^2/4$的亚高斯随机变量

例如

随机变量X服从高斯分布$N\ (\mu, \sigma^2)$, 则有

$$E\left[e^{(X-\mu)t}\right] = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{xt} e^{-x^2/2\sigma^2} dx \ = \ e^{\sigma^2 t^2/2}$$

Gaussian随机变量是参数为σ^2的亚高斯随机变量.

高斯随机变量和有界的随机变量都是亚高斯随机变量

定理

设$X_1, \ldots, X_n$是独立同分布的、且参数为b的亚高斯随机变量, 对任意$\epsilon > 0$有

$$P\left[\frac{1}{n}\sum_{i=1}^{n}(X_i - \mu) \geq \epsilon\right] \leq e^{-n\epsilon^2/2b}$$

$$P\left[\frac{1}{n}\sum_{i=1}^{n}(X_i - \mu) \leq -\epsilon\right] \leq e^{-n\epsilon^2/2b}$$

定理

设$X_1, \ldots, X_n$是n个相互独立的、参数为b的亚高斯随机变量, 且满足$E[X_i] = 0$, 有

$$E\left[\max_{i \in [n]} X_i\right] \leq \sqrt{2b \ln n}$$

另外一种表达形式

假设$X_1, \ldots, X_n$是独立的、且参数为b的亚高斯随机变量，对任意$\epsilon > 0$有

$$P\left[\frac{1}{n}\sum_{i=1}^{n}(X_i - \mu) \geq \epsilon\right] \leq e^{-n\epsilon^2/2b}$$

令$\delta = e^{-n\epsilon^2/2b}$求解出

$$\epsilon = \sqrt{\frac{2b}{n}\ln\frac{1}{\delta}}$$

另外一种表达形式

至少以$1-\delta$的概率有下面的不等式成立

$$\frac{1}{n}\sum_{i=1}^{n} X_i \leq \mu + \sqrt{\frac{2b}{n}\ln\frac{1}{\delta}}$$

前面讲的所有不等式都可以采用$1-\delta$的形式描述.

Bennet不等式

定理：独立同分布随机变量$X_1, \ldots, X_n$满足$X_i - E[X_i] \leq 1$，均值为μ和方差为σ^2，有

$$P\left[\frac{1}{n}\sum_{i=1}^{n}(X_i - \mu) \geq \epsilon\right] \leq \exp\left(-\frac{n\epsilon^2}{2\sigma^2 + 2\epsilon/3}\right)$$

令$\delta = \exp\left(-\frac{n\epsilon^2}{2\sigma^2+2\epsilon/3}\right)$，至少以$1-\delta$的概率有

$$\frac{1}{n}\sum_{i=1}^{n}X_i \leq \mu + \frac{2}{3n}\ln\frac{1}{\delta} + \sqrt{\frac{2\sigma^2}{n}\ln\frac{1}{\delta}}$$

Bernstein不等式

定理：独立同分布随机变量$X_1, \dots, X_n$均值为μ和方差σ^2，若存在常数$b > 0$，使得对任意正整数$m \geq 2$有$E[X_i^m] \leq m!\, b^{m-2}\sigma^2/2$，那么有

$$P\left[\frac{1}{n}\sum_{i=1}^{n} X_i - \mu \geq \epsilon\right] \leq \exp\left(-\frac{n\epsilon^2}{2\sigma^2 + 2b\epsilon}\right)$$