习题

7.1 设随机变量 X 的期望 $E[X]=\mu>0$, 方差为 σ^2, 证明对任意 $\epsilon>0$ 有

$$P(X-\mu\leqslant-\epsilon)\leqslant\frac{\sigma^2}{\sigma^2+\epsilon^2}.$$

7.2 设随机变量X和Y满足$E(X)=-2$, $E(Y)=2$, $\mathrm{Var}(X)=1$, $Var(Y)=4$, $\rho_{XY}=-1/2$. 利用Chebyshev不等式估计$\mathrm{Pr}(|X+Y|\geqslant 6)$的上界.

7.3 独立同分布随机变量 $X_1,X_2,\cdots,X_n$ 满足 $E[X_i]=\mu$ 和 $\mathrm{Var}(X_i)\leqslant v$. 证明对任意 $\epsilon>0$ 有

$$\left[\left|\frac{1}{n}\sum_{i=1}^{n}X_i-\mu\right|\geqslant\epsilon\right]\leqslant\frac{v}{n\epsilon^2}\ .$$

7.4 阐述什么是chernoff方法。

7.5 随机变量 $X_1,X_2,\cdots,X_n$ 相互独立且满足 $X_i\sim\mathrm{Ber}(p_i)$ $(p_i>0)$. 利用chernoff方法给出下列概率的上界

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

7.6 若独立同分布随机变量 $X_1,X_2,\cdots,X_n$ 满足 $X_i\in\{a,b\}$ $(b>a)$ 且 $P(X_i=a)=P(X_i=b)=1/2$. 求下列概率的上界

$$P\left[\frac{1}{n}\sum_{i=1}^{n}\left(X_i-\frac{a+b}{2}\right)\geqslant\epsilon\right]\quad\text{和}\quad \mathrm{Pr}\left[\frac{1}{n}\sum_{i=1}^{n}\left(X_i-\frac{a+b}{2}\right)\leqslant-\epsilon\right]\ .$$

7.7 随机变量 $X_1,X_2,\cdots,X_n$ 相互独立且满足 $X_i\sim\mathrm{Ber}(p_i)$ $(p_i>0)$. 证明对任意 $0<\epsilon<1$ 有不等式

$$P\left[\sum_{i=1}^{n}X_i\geqslant(1+\epsilon)\sum_{i=1}^{n}p_i\right]\leqslant e^{-\mu\epsilon^2/3}\ .$$

7.8 随机变量 $X\in[a,b]$ 且期望 $\mu=\mathbb{E}[x]$, 证明对任意 $t>0$ 有

$$\mathbb{E}\left[e^{tx}\right]\leqslant\exp\left(\mu t+t^2(b-a)^2/8\right)$$

7.9 利用chernoff方法证明: 设$X_1,X_2,\cdots,X_k$ 是k 个独立的随机变量, 且$X_i\sim N(0,1)$, 则有

$$\mathrm{Pr}\left(\sum_{i=1}^{k}X_i^2\geqslant(1+\epsilon)k\right)\leqslant\exp\left(-k\left(\epsilon^2-\epsilon^3\right)/4\right)$$

7.10 证明 Bennet 不等式.

7.11 证明 Berstein 不等式.

7.12 已知 Berstein 不等式

$$P\left[\frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)\geqslant\epsilon\right]\leqslant\exp\left(\frac{-n\epsilon^2}{2\sigma^2+2b\epsilon}\right),$$

给出其等价 $1-\delta$ 描述。

7.13 已知独立同分布随机变量 $X_1,X_2,\cdots,X_n$ 满足 $X_i\sim N\left(\mu,\sigma^2\right)$, 给出 $E\left[\max_{i\in[n]}\{X_i\}\right]$ 的上界, 并给出严格证明。

7.14 假设训练数据集 $S_n=\{(\boldsymbol{x}_1,y_1),(\boldsymbol{x}_2,y_2),\ldots,(\boldsymbol{x}_n,y_n)\}$ 根据分布 $\mathcal{D}$ 独立采样所得, 分类器 f 在训练集 S_n 的错误率为零 (全部预测正确), 求分类器 f 在分布$\mathcal{D}$上的错误率介于 0 和 ϵ 之间的概率 ($\epsilon>0$).

7.15 假设训练数据集 $S_n=\{(\boldsymbol{x}_1,y_1),(\boldsymbol{x}_2,y_2),\ldots,(\boldsymbol{x}_n,y_n)\}$ 根据分布 $\mathcal{D}$ 独立采样所得, 分类器 f 在训练集 S_n 的错误率为 $\hat{p}$, 求分类器 f 在分布$\mathcal{D}$上的错误率介于 $\hat{p}-\epsilon$ 和 $\hat{p}+\epsilon$ 之间的概率 ($\epsilon>0$).