集中不等式总结

邓文平

2025 年 2 月 22 日

1 基础不等式

定理 1 (马尔可夫不等式)**.** 设 X 为非负随机变量，则对任意 $t > 0$:

$$\mathbb{P}(X \geq t) \leq \frac{\mathbb{E}[X]}{t}.$$

证明. 通过期望的定义：

$$\mathbb{E}[X] = \int_0^\infty x d\mathbb{P}_X(x) \geq \int_t^\infty x d\mathbb{P}_X(x) \geq t \int_t^\infty d\mathbb{P}_X(x) = t\mathbb{P}(X \geq t).$$

两边除以 t 即得结论。 □

定理 2 (切比雪夫不等式)**.** 设 X 为随机变量，均值为 μ，方差为 σ^2，则:

$$\mathbb{P}(|X - \mu| \geq t) \leq \frac{\sigma^2}{t^2}.$$

证明. 对随机变量 $Y = (X - \mu)^2$ 应用马尔可夫不等式：

$$\mathbb{P}(|X - \mu| \geq t) = \mathbb{P}(Y \geq t^2) \leq \frac{\mathbb{E}[Y]}{t^2} = \frac{\mathrm{Var}(X)}{t^2}.$$

□

定理 3 (切尔诺夫界)**.** 设 $X = \sum_{i=1}^n X_i$，其中 X_i 独立。对任意 $t > 0$:

$$\mathbb{P}(X \geq t) \leq \inf_{\lambda > 0} e^{-\lambda t} \mathbb{E}\left[e^{\lambda X}\right].$$

证明. 对任意 $\lambda > 0$，应用马尔可夫不等式于 $e^{\lambda X}$：

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

取右端关于 λ 的下确界即得结果。 □

2 高斯分布的尾部概率估计

定理 4 (高斯尾部概率估计). 设 $X \sim \mathcal{N}(0, \sigma^2)$ 为标准高斯随机变量，则对任意 $t > 0$，其尾部概率满足：

$$\mathbb{P}(X \geq t) \leq \frac{1}{2} e^{-t^2/(2\sigma^2)}.$$

更精确的双侧估计为：

$$\mathbb{P}(|X| \geq t) \leq \sqrt{\frac{2}{\pi}} \frac{\sigma}{t} e^{-t^2/(2\sigma^2)} \quad (t > \sigma).$$

证明. (单侧估计) 对 $X \sim \mathcal{N}(0, \sigma^2)$，计算尾部概率：

$$\mathbb{P}(X \geq t) = \int_t^\infty \frac{1}{\sqrt{2\pi\sigma^2}} e^{-x^2/(2\sigma^2)} dx.$$

作变量替换 $y = x/\sigma$，得：

$$= \int_{t/\sigma}^\infty \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy.$$

利用不等式 $\int_u^\infty e^{-y^2/2} dy \leq \frac{1}{u} e^{-u^2/2}$ （当 $u > 0$），取 $u = t/\sigma$ 即得：

$$\mathbb{P}(X \geq t) \leq \frac{\sigma}{t\sqrt{2\pi}} e^{-t^2/(2\sigma^2)} \leq \frac{1}{2} e^{-t^2/(2\sigma^2)}.$$

(精确双侧估计) 通过分部积分可得递推式：

$$\int_t^\infty e^{-y^2/2} dy = \frac{e^{-t^2/2}}{t} - \int_t^\infty \frac{e^{-y^2/2}}{y^2} dy \leq \frac{e^{-t^2/2}}{t}.$$

结合对称性 $\mathbb{P}(|X| \geq t) = 2\mathbb{P}(X \geq t)$ 即得结果。 □

3 次高斯分布与次指数分布

定义 5 (次高斯分布). 随机变量 X 称为参数为 σ 的**次高斯分布**（记作 $X \sim \mathrm{Subg}(\sigma^2)$），如果满足:

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

尾概率证明. 对任意 $t > 0$，取 $\lambda = t/\sigma^2$，应用切尔诺夫界：

$$\mathbb{P}(X - \mathbb{E}X \geq t) \leq e^{-\lambda t} e^{\lambda^2\sigma^2/2} = e^{-t^2/(2\sigma^2)}.$$

结合 $-X$ 的情况即得双侧界。 □

4 霍夫丁不等式

定理 6 (对称伯努利变量的霍夫丁不等式)**.** 设 $X_1, \ldots, X_n$ 为独立对称伯努利随机变量（即 $\mathbb{P}(X_i = 1) = \mathbb{P}(X_i = -1) = \frac{1}{2}$），令 $S_n = \sum_{i=1}^n X_i$。则对任意 $t > 0$：

$$\mathbb{P}(S_n \geq t) \leq \exp\left(-\frac{t^2}{2n}\right).$$

双侧形式为：

$$\mathbb{P}(|S_n| \geq t) \leq 2\exp\left(-\frac{t^2}{2n}\right).$$

证明. **方法一**（从一般霍夫丁不等式导出）：由于 $X_i \in [-1, 1]$ 且 $\mathbb{E}X_i = 0$，应用标准霍夫丁不等式：

$$\sum_{i=1}^n (b_i - a_i)^2 = \sum_{i=1}^n (1 - (-1))^2 = 4n.$$

代入得：

$$\mathbb{P}(S_n \geq t) \leq \exp\left(-\frac{2t^2}{4n}\right) = \exp\left(-\frac{t^2}{2n}\right).$$

方法二（直接矩生成函数计算）：计算单个变量的矩生成函数：

$$\mathbb{E}[e^{\lambda X_i}] = \cosh(\lambda) \leq e^{\lambda^2/2}.$$

利用独立性：

$$\mathbb{E}[e^{\lambda S_n}] = \prod_{i=1}^n \mathbb{E}[e^{\lambda X_i}] \leq e^{n\lambda^2/2}.$$

应用切尔诺夫界：

$$\mathbb{P}(S_n \geq t) \leq \inf_{\lambda>0} e^{-\lambda t} e^{n\lambda^2/2} = e^{-t^2/(2n)}.$$

其中最优 $\lambda = t/n$。 □

推论 7 (次高斯性验证)**.** 对称伯努利变量 X_i 是 $\mathrm{Subg}(1)$ 的，其和 $S_n \sim \mathrm{Subg}(n)$。

证明. 由矩生成函数估计：

$$\mathbb{E}[e^{\lambda X_i}] = \frac{e^{\lambda} + e^{-\lambda}}{2} = \cosh(\lambda) \leq e^{\lambda^2/2}.$$

和变量满足 $\mathbb{E}[e^{\lambda S_n}] \leq e^{n\lambda^2/2}$，符合次高斯性定义。 □

定理 8 (霍夫丁不等式). 设 $X_1, \ldots, X_n$ 为独立随机变量，满足 $X_i \in [a_i, b_i]$ 几乎必然成立。令 $S_n = \sum_{i=1}^n (X_i - \mathbb{E}X_i)$，则对任意 $t > 0$:

$$\mathbb{P}(S_n \geq t) \leq \exp\left(-\frac{2t^2}{\sum_{i=1}^n (b_i - a_i)^2}\right).$$

证明. 关键步骤： 1. 对每个 $X_i - \mathbb{E}X_i$ 建立次高斯性，其参数 $\sigma_i = (b_i - a_i)/2$ 2. 利用独立性得 S_n 的矩生成函数：

$$\mathbb{E}[e^{\lambda S_n}] = \prod_{i=1}^n \mathbb{E}[e^{\lambda(X_i - \mathbb{E}X_i)}] \leq \prod_{i=1}^n e^{\lambda^2 (b_i - a_i)^2/8}$$

3. 应用切尔诺夫界并优化 λ □

5 基于鞅的不等式

定理 9 (Azuma-Hoeffding不等式). 设 $\{X_i\}_{i=0}^n$ 为鞅序列，满足 $|X_i - X_{i-1}| \leq c_i$ 几乎必然成立。则对任意 $t > 0$:

$$\mathbb{P}(|X_n - X_0| \geq t) \leq 2\exp\left(-\frac{t^2}{2\sum_{i=1}^n c_i^2}\right).$$

证明. 构造鞅差序列 $D_i = X_i - X_{i-1}$，利用霍夫丁引理证明每个 D_i 的次高斯性，再通过矩生成函数的乘积性质完成证明。 □

6 应用说明

- 次高斯性是推导高斯向量Lipschitz函数尖锐界限的关键
- McDiarmid不等式在机器学习中广泛用于算法稳定性分析
- Azuma不等式适用于依赖过程（例如具有依赖性的随机游走）