FOUNDATIONAL TECHNIQUES IN MACHINE LEARNING & DATA SCIENCE

CSE 382M

Author
Noah Reef
UT Austin
Spring 2025

Contents

1 Probability

1.1 Concentration Inequalities

Theorem 1.1 (Markov's Inequality)**.** Let x be a non-negative random variable. Then for $a > 0$,

$$\mathbb{P}(x \geq a) \leq \frac{\mathbb{E}[x]}{a}$$

Theorem 1.2 (Chebyshev's Inequality)**.** Let x be a random variable. Then for $c > 0$,

$$\mathbb{P}(|x - \mathbb{E}[x]| \geq c) \leq \frac{\mathrm{Var}[x]}{c^2}$$

Theorem 1.3 (Law of Large Numbers)**.** Let $x_1, x_2, \ldots, x_n$ be n independent samples of a random variable x. Then

$$\mathbb{P}\left[\left|\frac{x_1 + x_2 + \cdots + x_n}{n} - \mathbb{E}[x]\right| \geq \epsilon\right] \leq \frac{\mathrm{Var}[x]}{n\epsilon^2}$$

Theorem 1.4 (Master Tail Bounds Theorem)**.** Let $x = x_1 + x_2 + \cdots + x_n$, where $x_1, x_2, \ldots, x_n$ are mutually independent random variables with zero mean and variance at most σ^2. Let $0 \leq a \leq \sqrt{2}n\sigma^2$. Assume that $|\mathbb{E}[x_i^s]| \leq \sigma^2 s!$ for $s = 3, 4, \ldots, \lfloor(a^2/4n\sigma^2)\rfloor$. Then,

$$\mathbb{P}[|x| \geq a] \leq 3e^{-a^2/(12n\sigma^2)}$$

Theorem 1.5 (General Master Tail Bounds Theorem)**.** Let $x = x_1 + x_2 + \cdots + x_n$, where $x_1, x_2, \ldots, x_n$ are mutually independent random variables with zero mean and variance at most σ^2. Let $0 \leq a \leq \sqrt{2}n\sigma^2$ and $s \leq n\sigma^2/2$ is a positive even integer and $|\mathbb{E}[x_i^r] \leq \sigma^2 r!$ for $r = 3, 4, \ldots, s$.Then,

$$\mathbb{P}[|x_1 + x_2 + \cdots + x_n| \geq a] \leq \left(\frac{2sn\sigma^2}{a^2}\right)^{s/2}$$

If further, $s \geq a^2/(4n\sigma^2)$, then we also have:

$$\mathbb{P}[|x_1 + x_2 + \cdots + x_n| \geq a] \leq 3e^{-a^2/(12n\sigma^2)}$$

Theorem 1.6 (Chernoff Bound)**.** Let x be a Bernoulli random variable, with $\mathbb{E}[x] = p$ and $\mathrm{Var}[x] = p(1-p)$ then we have that

$$\mathbb{P}\left[\left|\frac{y}{n} - p\right| \geq \sqrt{2}cp(1-p)\right] \leq 3e^{-np(1-p)c^2/6}$$

References