1 Moments and Deviations

Theorem 1.1 **(Markov's Inequlity)** Let X be a random variable with only nonnegative values. Then, for all $a > 0$,

$$\mathbf{P}(X \geq a) \leq \frac{\mathbf{E}[X]}{a}$$

Proof. For $a > 0$, let $I = 1$ (if $X \geq a$) or 0 (otherwise), and note that $I \leq X/a$. Taking expectaions on both sides, thus yields $\mathbf{P}(X \geq a) = \mathbf{E}[I] =\leq \mathbf{E}[X/a] = \mathbf{E}[X]/a$. ◀

The **k-th moment** of a random variable X is $\mathbf{E}[X^k]$. The **variance** of random variable X is defined as $\mathbf{Var}[X] = \mathbf{E}[(X - \mathbf{E}[X])^2] = \mathbf{E}[X^2] - \mathbf{E}[X]^2$, and the **standard deviation** of a random variable X is $\sigma[X] = \sqrt{\mathbf{Var}[X]}$. The **convariance** of two random variables X and Y is $\mathbf{Cov}(X, y) = \mathbf{E}[(X - \mathbf{E}[X])(Y - \mathbf{E}[Y])]$, and we have

Lemma 1.2 For any two random variables X and Y, $\mathbf{Var}[X + Y] = \mathbf{Var}[X] + \mathbf{Var}[Y] + 2 \cdot \mathbf{Cov}(X, Y)$.

Lemma 1.3 For any two independent random variables X and Y, $\mathbf{E}[X \cdot Y] = \mathbf{E}[X] \cdot \mathbf{E}[Y]$. (the opposite does not hold)

Corollary 1.4 If X and Y are independent random variables, then $\mathbf{Cov}(X, Y) = 0$.

Theorem 1.5 **(Linearity of variance)** Let $X_1, X_2, \cdots, X_n$ be mutually independent random variables. Then

$$\mathbf{Var}\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} \mathbf{Var}[X_i]$$

For example, a Bernoulli trial with success probability p has variable $p(1 - p)$, therefore the variance of a binomial random variable X with parameters n and p is $np(1 - p)$.

Theorem 1.6 **(Chebyshev's inequality)** Let X be a random variable. Then, for any $a > 0$,

$$\mathbf{P}\left(|X - \mathbf{E}[X]| \geq a\right) \leq \frac{\mathbf{Var}[X]}{a^2}$$

Proof. We can apply Markov's inequality to prove:

$$\mathbf{P}(|X - \mathbf{E}[X]| \geq a) = \mathbf{P}((X - \mathbf{E}[X])^2 \geq a^2) \leq \frac{\mathbf{E}[(X - \mathbf{E}[X])^2]}{a^2} = \frac{\mathbf{Var}[X]}{a^2}$$

A useful variant of Chebyshev's inequality is to substitute a with $t \cdot \sigma[X]$ $(t \geq 1)$. ◀

The **median** of random variable X is defined to be any value m such that $\mathbf{P}(X \leq m) \geq 1/2$ and $\mathbf{P}(X \geq m) \geq 1/2$.

Theorem 1.7 For any random variable X with finite expectaion $\mathbf{E}[X]$ and finite median m,

- the expectaion $\mathbf{E}[X]$ is the value of c that minimizes the expression $\mathbf{E}[(X - c)^2]$.
- the median m is the value of c that minimizes the expression $\mathbf{E}[|X - c|]$.

Corollary 1.8 $|\mu - m| = |\mathbf{E}[X] - m| = |\mathbf{E}[X - m]| \leq \mathbf{E}[|X - m|] \leq \mathbf{E}[|X - \mu|] \leq \sqrt{\mathbf{E}[(X - \mu)^2]} = \sigma$.