8. Limit theorems

Proposition 8.1 (Markov inequality)**.** *Let X be a nonnegative random variable. Then, for $a > 0$,*

$$P(X > a) \leq \frac{E(X)}{a}.$$

Proof. Let $A = \{X > a\}$ and set

$$I(s) = \begin{cases} 1 & \text{if } s \in A \\ 0 & \text{if } x \notin A \end{cases}.$$

Clearly, $X/a \geq I$. This implies

$$P(X > a) = P(A) = E(I) \leq E(X/a) = E(X)/a.$$

□

Proposition 8.2 (Chebyshev inequality)**.** *Let X be a random variable with mean $\mu \in \mathbb{R}$ and variance $\sigma^2 \in (0, \infty)$. Then, for any $a > 0$,*

$$P(|X - \mu| > a) \leq \frac{\sigma^2}{a^2}.$$

Proof. Since $(X - \mu)^2$ is nonnegative, one may apply the Markov inequality to derive

$$P(|X - \mu| > a) = P((X - \mu)^2 > a) \leq \frac{E(X - \mu)^2}{a^2} = \frac{\mathrm{Var}(X)}{a^2}.$$

□

Proposition 8.3 (One-sided Chebyshev inequality)**.** *Let X be a random variable with mean μ and variance σ^2. Then, for any $a > 0$,*

$$P(X - \mu > a) \leq \frac{\sigma^2}{\sigma^2 + a^2}, \quad P(X - \mu < -a\sigma) \leq \frac{\sigma^2}{\sigma^2 + a^2}.$$

Proof. Note that the second inequality is exactly the first one when X is replaced by $-X$. To prove the first inequality, it suffices to consider the specific case of $\mu = 0$ and $\sigma^2 = 1$. Let $a > 0$ and $b > 0$. By the Markov inequality, one has

$$P(X > a) = P(X + b > a + b) \leq P((X + b)^2 > (a + b)^2) \leq \frac{E(X + b)^2}{(a + b)^2} = \frac{1 + b^2}{(a + b)^2}.$$

Observe that the last term, as a function of b, has its minimum at $b = 1/a$. Consequently, this implies

$$P(X > a) \leq \frac{1 + (1/a)^2}{(a + 1/a)^2} = \frac{1}{1 + a^2}.$$

□

Example 8.1. Let X be a binomial random variable with parameters (n, p). By the Markov inequality, one has

$$P(X > a) \leq \frac{E(X)}{a} = \frac{np}{a},$$

while the Chebyshev inequality yields

$$P(|X - np| > a) \leq \frac{np(1 - p)}{a^2}.$$

Theorem 8.4 (The weak law of large numbers)**.** *Let $X_1, X_2, ...$ be i.i.d. random variables with finite mean μ. Then, for any $\epsilon > 0$,*

$$\lim_{n\to\infty} P\left(\left|\frac{X_1 + \cdots + X_n}{n} - \mu\right| \le \epsilon\right) = 1.$$

Proof. We consider the case $\mathrm{Var}(X) = \sigma^2 < \infty$ here. By the Chebyshev inequality, one has

$$P\left(\left|\frac{X_1 + \cdots + X_n}{n} - \mu\right| > \epsilon\right) \le \frac{\mathrm{Var}((X_1 + \cdots + X_n)/n)}{\epsilon^2} = \frac{\sigma^2}{n\epsilon}.$$

Letting n tend to infinity gives the desired limit. □

Theorem 8.5 (The strong law of large numbers)**.** *Let $X_1, X_2, ...$ be i.i.d. random variables with finite mean μ. Then,*

$$P\left(\lim_{n\to\infty} \frac{X_1 + \cdots + X_n}{n} = \mu\right) = 1.$$

Theorem 8.6 (The central limit theorem)**.** *Let $X_1, X_2, ...$ be i.i.d. random variables with mean μ and variance $\sigma^2 \in (0, \infty)$. Then, for any $c \in \mathbb{R}$,*

$$\lim_{n\to\infty} P\left(\frac{X_1 + \cdots + X_n - n\mu}{\sqrt{n}\sigma} \le c\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{c} e^{-t^2/2} dt.$$

Remark 8.1. From the viewpoint of measure theory, the weak law of large number says that $(X_1 + \cdots + X_n)/n$ converges to μ in probability. The strong law of large numbers says that $(X_1 + \cdots + X_n)/n$ converge to μ with probability 1 or almost surely (a.s.). The central limit theorem says that $(X_1 + \cdots + X_n - n\mu)/(\sqrt{n}\sigma)$ converges in distribution to the standard normal random variable.