Chapter 5

Limit Theorems

Peng-Hua Wang

Graduate Institute of Communication Engineering

National Taipei University

Chapter Contents

5.1 Some Useful Inequalities

Markov Inequality

If a random variable $X \geq 0$, then

$$\mathrm{P}(X \geq a) \leq \frac{E[X]}{a}, \quad \text{for all } a > 0$$

Proof.

$$\begin{aligned} E[X] &= \int_0^\infty x f_X(x)dx \geq \int_a^\infty x f_X(x)dx \\ &\geq \int_a^\infty a f_X(x)dx = a\int_a^\infty f_X(x)dx \\ &= a\mathrm{P}(X \geq a) \qquad \blacksquare \end{aligned}$$

Markov Inequality

- Relationship between probability and mean. Use mean to estimate probability.
- Example 5.1. X is uniformly distributed in $(0, 4)$. $E[X] = 2$.
  - $\mathrm{P}(X > 2) \leq 2/2 = 1, \mathrm{P}(X > 2) = 1/2$
  - $\mathrm{P}(X > 3) \leq 2/3, \mathrm{P}(X > 3) = 1/4$
  - The bounds can be very loose.

Chebyshev Inequality

If a random variable X has mean μ and variance σ^2, then

$$P(|X - \mu| \geq c) \leq \frac{\sigma^2}{c^2}, \quad \text{for all } c > 0$$

Proof.

$$\begin{aligned} P(|X - \mu| \geq c) &= P((X - \mu)^2 \geq c^2) \\ &\leq \frac{E[(X - \mu)^2]}{c^2} = \frac{\sigma^2}{c^2} \end{aligned} \qquad \blacksquare$$

Remark.

$$P(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2}$$

Chebyshev Inequality

- Relationship between probability and mean and variance. Use mean and variance to estimate probability.
- Example 5.2. X is uniformly distributed in $(0, 4)$. $E[X] = 2$ and $\mathrm{Var}(X) = 16/12$
  - $P(|X - 2| > 1) \leq 16/12$,
  - The bounds can be very loose.

5.2 The Weak Law Of Large Numbers

Inequality for Sample Mean

- Let $X_1, X_2, \ldots$ be a sequence of independent identically distributed random variables with mean μ and variance σ^2. Let M_n be the sample mean:

$$M_n = \frac{X_1 + X_2 + \cdots + X_n}{n}$$

- We have $E[M_n] = \mu$ and $\mathrm{Var}(M_n) = \sigma^2/n$.
- Apply Chebyshev inequality. For any $\epsilon > 0$ we have

$$\mathrm{P}(|M_n - \mu| \geq \epsilon) \leq \frac{\sigma^2}{n\epsilon^2}$$

- The right-hand side of this inequality goes to zero as n increases.

The Weak Law of Large Numbers

- Therefore, sample mean (a random variable) approaches the mean in probability as n approaches infinity.
- The Weak Law of Large Numbers

$$P\left(\left|\frac{X_1 + X_2 + \cdots + X_n}{n} - \mu\right| \geq \epsilon\right) \to 0$$

5.4 The Central Limit Theorem

The Central Limit Theorem

- Let $X_1, X_2, \ldots$ be a sequence of independent identically distributed random variables with mean μ and variance σ^2. Define Z_n as:

$$Z_n = \frac{X_1 + X_2 + \cdots + X_n - n\mu}{\sigma\sqrt{n}}$$

- It can be shown that the CDF of Y_n converges to the CDF of standard normal distribution.

$$\lim_{n\to\infty} \mathrm{P}(Z_n \le z) = \Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx$$

Example 5.12

Toss a fair coin for 36 times. Estimate the probability that the head occurs for less than ro equal to 21 times.

Let X_i be iid Bernoulli rv's with parameter 0.5. We have $\mu = 0.5$ and $\sigma^2 = 0.5(1-0.5)$. The probability of interest is

$$
\begin{aligned}
& \mathrm{P}\left(X_1 + X_2 + \cdots + X_{36} \leq 21\right) \\
=& \mathrm{P}\left(\frac{X_1 + X_2 + \cdots + X_{36} - 36\mu}{\sqrt{36}\sigma} \leq \frac{21 - 36\mu}{\sqrt{36}\sigma}\right) \approx \Phi(1) \approx 0.8413
\end{aligned}
$$

The exact value is

$$\sum_{n=0}^{21} \binom{36}{n} (0.5)^{36} = 0.8785$$

Binomial Distribution and CLT

- Let X_i be iid Bernoulli rv's with parameter p. Let $S_n = X_1 + X_2 + \cdots X_n$. We know that S_n be a binomial rv with parameter n and p.
- The probability of $\mathrm{P}(k \le S_n \le \ell)$ is

$$\mathrm{P}(k \le S_n \le \ell) = \sum_{i=k}^{\ell} \binom{n}{i} p^i (1-p)^{n-i}$$

- By CLT, we have

$$\begin{aligned}\mathrm{P}(k \le S_n \le \ell) &= \mathrm{P}\left(\frac{k-np}{\sqrt{np(1-p)}} \le \frac{S_n - np}{\sqrt{np(1-p)}} \le \frac{\ell - np}{\sqrt{np(1-p)}}\right) \\ &\approx \Phi\left(\frac{\ell - np}{\sqrt{np(1-p)}}\right) - \Phi\left(\frac{k-np}{\sqrt{np(1-p)}}\right)\end{aligned}$$

Binomial Distribution and CLT

- A more accurate approximation can be obtained by replacing k and ℓ by $k - 1/2$ and $\ell + 1/2$. The is called the De Moivre-Laplace approximation.

$$P(k \le S_n \le \ell) \approx \Phi\left(\frac{\ell + 1/2 - np}{\sqrt{np(1-p)}}\right) - \Phi\left(\frac{k - 1/2 - np}{\sqrt{np(1-p)}}\right)$$

- In Example 5.12,

$$\begin{aligned} & P\left(X_1 + X_2 + \cdots + X_{36} \le 21\right) \\ =& P\left(\frac{X_1 + X_2 + \cdots + X_{36} - 36\mu}{\sqrt{36}\sigma} \le \frac{21 + 1/2 - 36\mu}{\sqrt{36}\sigma}\right) \\ \approx& \Phi((21.5 - 6)/3) \approx 0.879 \end{aligned}$$