38. Limit Theorems

Chebyshev's inequality and Weak Law of Large Numbers [Ross S8.2]

Proposition 38.1 (Markov inequality) *If X is a non-negative random variable, then for any $a > 0$:*

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

Why? [textbook explanation]

Let $\quad I = \begin{cases} 1 & \text{if } X \geq a \\ 0 & \text{else} \end{cases}$

Then $\quad I \leq \dfrac{X}{a}$

Hence: $\quad E[I] \leq \dfrac{E[X]}{a}$

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

[Second approach for continuous rvs]

$$\begin{aligned} P[X \geq a] &= \int_a^\infty f_X(x)dx \\ &\leq \int_a^\infty \frac{x}{a} f_X(x)dx \qquad \text{since } x/a \geq 1 \text{ and } f_X(x) \geq 0 \\ &\leq \int_0^\infty \frac{x}{a} f_X(x)dx \end{aligned}$$

$$= E[X]/a \qquad \text{since } X \text{ is non-negative}$$

Proposition 38.2 (Chebyshev's inequality) *If X is a random variable with mean μ and variance σ^2, then for any $b > 0$:*

$$P\left[|X - \mu| \geq b\right] = P\left[(X - \mu)^2 \geq b^2\right] \leq \frac{\sigma^2}{b^2}$$

Why?

$(X - \mu)^2$ is a non-negative random variable. With $b^2 > 0$, apply Markov's inequality to it:

$$\begin{aligned} P\left[(X - \mu)^2 \geq b^2\right] &\leq \frac{E\left[(X - \mu)^2\right]}{b^2} \\ &= \frac{\sigma^2}{b^2} \end{aligned}$$

Note: Markov (or Chebyshev) let us derive bounds on probabilities when all we know is the mean (or both the mean and variance) of a random variable.

Example 38.1: The mean number of items per week that a factory produces is 50.

a) What can you say about the probability that it produces at least 75 items in a week?

b) If the variance of the weekly production is 25, what can you say about the probability that it produces more than 40 but fewer than 60 items?

Solution:

Example 38.2: Let $X \sim U(0, 10)$. Use Chebyshev to approximate $P[|X - 5| \geq 4]$ and compare to the exact value.

Solution:

Chebyshev can be used to prove theoretical results:

Proposition 38.3 *Weak Law of Large Numbers [WLLN]*

Let $X_1, X_2, \ldots,$ be a sequence of iid random variables with $E[X_i] = \mu$. Then, for any $\epsilon > 0$:

$$P\left[\left|\underbrace{\frac{X_1 + X_2 + \cdots + X_n}{n}}_{\textit{sample mean}} - \mu\right| \geq \epsilon\right] \to 0 \qquad \textit{as } n \to \infty$$

Why? [Under assumption that $Var[X_i] = \sigma^2$ is finite.]

$$E\left[\frac{X_1 + X_2 + \cdots + X_n}{n}\right] = \mu$$

$$Var\left[\frac{X_1 + X_2 + \cdots + X_n}{n}\right] = \frac{\sigma^2}{n}$$

By Chebyshev

$$P\left[\left|\frac{X_1 + X_2 + \cdots + X_n}{n} - \mu\right| \geq \epsilon\right] \leq \frac{\sigma^2/n}{\epsilon^2}$$

and

$$\frac{\sigma^2}{n\epsilon^2} \to 0 \qquad \text{as } n \to \infty$$

Example 38.3: A fair coin has a 0 on one side and a 1 on the other.

You conduct a sequence of independent trials that consists of repeatedly flipping the coin.

Let Z_n be the fraction of flips that result in the number 1 after n flips.

What can you say about the probability that Z_n is between 0.499 and 0.501 as $n \to \infty$?

Solution: