# Probability Theory

Textbook: A First Course in Probability, Sheldon Ross, 2019.

Prof. Hicham Elmongui

elmongui@alexu.edu.eg

## Chapter 08: Limit Theorems



## Introduction

### Limit Theorems

- The most important theoretical results in probability theory.
- The most important theorems are those classified either under
  - *Laws of large numbers*
  - *Central limit theorems*

### Laws of large numbers

They are concerned with stating conditions under which the average of a sequence of RVs converges to the expected average.

### Central limit theorems

They are concerned with determining conditions under which the sum of a large number of RVs has a probability distribution that is approximately normal.





## Markov's Inequality

### Proposition

If $X$ is a RV that takes only nonnegative values, then, for any value $a > 0$,

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

### Proof

For $a > 0$, let

$$I = \begin{cases} 1 & X \geq a \\ 0 & \text{otherwise} \end{cases}$$

Since $X \geq 0$,

$$I \leq \frac{X}{a}$$

Taking expectations of the preceding inequality yields

$$E[I] \leq \frac{E[X]}{a}$$

Since $E[I] = P\{X \geq a\}$,

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





## Chebyshev's Inequality

### Proposition

If $X$ is a RV with finite mean $\mu$ and variance $\sigma^2$, then, for any value $k > 0$,

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

### Proof

Since $(X - \mu)^2$ is a nonnegative random variable, we can apply Markov's inequality (with $a = k^2$) to obtain

$$P\left\{(X-\mu)^2 \geq k^2\right\} \leq \frac{E[(X-\mu)^2]}{k^2}$$

But since $(X-\mu)^2 \geq k^2$ if and only if $|X - \mu| \geq k$,

$$P\{|X-\mu| \geq k\} \leq \frac{E[(X-\mu)^2]}{k^2} = \frac{\sigma^2}{k^2} \quad \blacksquare$$





## Examples

Suppose that it is known that the number of items produced in a factory during a week is a random variable with mean 50.

- What's the probability that this week's production will exceed 75?
- If the variance of a week's production is found to be 25, what's the probability that this week's production will be between 40 and 60?

### Solution

Let $X$ be the number of items that will be produced in a week.

- By Markov's inequality,

$$P\{X > 75\} \leq \frac{E[X]}{75} = \frac{50}{75} = \frac{2}{3}$$

- By Chebyshev's inequality,

$$P\{|X - 50| \geq 10\} \leq \frac{\sigma^2}{10^2} = \frac{25}{10^2} = \frac{1}{4}$$

$$P\{|X - 50| < 10\} \geq 1 - \frac{1}{4} = \frac{3}{4}$$





## Examples (cont'd)

As Chebyshev's inequality is valid for all distributions of the random variable $X$, we cannot expect the bound on the probability to be very close to the actual probability in most cases.

### Example

If $X \sim \mathrm{Normal}(\mu, \sigma^2)$, Chebyshev's inequality states that

$$P\{|X - \mu| > 2\sigma\} \leq \frac{1}{4}$$

whereas the actual probability is given by

$$P\{|X-\mu| > 2\sigma\} = P\left\{\left|\frac{X-\mu}{\sigma}\right| > 2\right\} = 2\,(1 - \Phi(2)) \approx .0456$$

Chebyshev's inequality is often used as a theoretical tool in proving results.

## The Weak Law of Large Numbers

### Proposition

Let $X_1, X_2, \cdots, X_n$ be a sequence of i.i.d. RVs, each having finite mean $E[X_i] = \mu$. Then, for any $\varepsilon > 0$,

$$\lim_{n\to\infty} P\left\{\left|\frac{X_1 + X_2 + \cdots + X_n}{n} - \mu\right| \geq \varepsilon\right\} = 0$$

### Part of the proof ($X_i$ has finite variance $\sigma^2$)

Since

$$E\left[\frac{X_1 + X_2 + \cdots + X_n}{n}\right] = \mu \quad \text{and} \quad \mathrm{Var}\left(\frac{X_1 + X_2 + \cdots + X_n}{n}\right) = \frac{\sigma^2}{n}$$

it follows from Chebyshev's inequality that

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

and the result is proven.





## The Central Limit Theorem

### The central limit theorem

Let $X_1, X_2, \cdots, X_n$ be a sequence of i.i.d. RVs, each having mean $\mu$ and variance $\sigma^2$. If $\overline{X}$ is the sample mean, then the distribution of

$$\frac{\overline{X} - \mu}{\sigma/\sqrt{n}}$$

tends to the standard normal as $n \to \infty$.

### The central limit theorem

Let $X_1, X_2, \cdots, X_n$ be a sequence of i.i.d. RVs, each having mean $\mu$ and variance $\sigma^2$. Then the distribution of

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

tends to the standard normal as $n \to \infty$.





## Examples

If 10 fair dice are rolled, find the approximate probability that the sum obtained is between 30 and 40, inclusive.

### Solution

Let $X_i$ denote the value of the $i^{\text{th}}$ die, $i = 1, 2, \cdots, 10$. Since

$$E[X_i] = \frac{7}{2} \qquad \mathrm{Var}(X_i) = \frac{35}{12}$$

the central limit theorem yields

$$\begin{aligned} P\{29.5 \leq X \leq 40.5\} &= P\left\{\frac{29.5 - 35}{\sqrt{\frac{350}{12}}} \leq \frac{X - 35}{\sqrt{\frac{350}{12}}} \leq \frac{40.5 - 35}{\sqrt{\frac{350}{12}}}\right\} \\ &\approx P\{-1.0184 \leq Z \leq 1.0184\} \\ &= \Phi(1.0184) - \Phi(-1.0184) \\ &= 2\Phi(1.0184) - 1 \qquad \approx .692 \end{aligned}$$





## Examples (cont'd)

An instructor has 50 exams that will be graded in sequence. The times required to grade the 50 exams are independent, with a common distribution that has mean 20 minutes and standard deviation 4 minutes. Approximate the probability that the instructor will grade at least 25 of the exams in the first 450 minutes of work.

### Solution

If we let $X_i$ be the time that it takes to grade exam $i$, then the time it takes to grade the first 25 exams is

$$X = \sum_{i=1}^{25} X_i$$

$$E[X] = \sum_{i=1}^{25} E[X_i] = 25(20) = 500$$

$$\mathrm{Var}(X) = \sum_{i=1}^{25} \mathrm{Var}(X_i) = 25(16) = 400$$

$$\begin{aligned} P\{X \leq 450\} &= P\left\{\frac{X - 500}{\sqrt{400}} \leq \frac{450 - 500}{\sqrt{400}}\right\} \\ &\approx P\{Z \leq -2.5\} \\ &= 1 - \Phi(2.5) \qquad \approx .006 \end{aligned}$$





## The Strong Law of Large Numbers

### The strong law of large numbers

Let $X_1, X_2, \cdots, X_n$ be a sequence of i.i.d. RVs, each having a finite mean $\mu = E[X_i]$. Then,

$$P\left\{\lim_{n\to\infty} \frac{X_1 + X_2 + \cdots + X_n}{n} = \mu\right\} = 1$$

### Application

Suppose that a sequence of independent trials of some experiment is performed. Let $E$ be a fixed event of the experiment.

Letting $X_i = \begin{cases} 1 & \text{if } E \text{ occurs on the } i^{\text{th}} \text{ trial} \\ 0 & \text{if } E \text{ does not occur on the } i^{\text{th}} \text{ trial} \end{cases}$

We have, by the strong law of large numbers, that with probability 1,

$$\frac{X_1 + X_2 + \cdots + X_n}{n} \to E[X] = P(E)$$





## The Weak Law vs. Strong Law of Large Numbers

Let $X_1, X_2, \cdots, X_n$ be a sequence of i.i.d. RVs (with finite mean $E[X_i] = \mu$).

### The weak law of large numbers (WLLN)

For any $\varepsilon > 0$, $\quad \lim_{n\to\infty} P\left\{\left|\frac{X_1 + X_2 + \cdots + X_n}{n} - \mu\right| \geq \varepsilon\right\} = 0$

### The strong law of large numbers (SLLN)

$$P\left\{\lim_{n\to\infty} \frac{X_1 + X_2 + \cdots + X_n}{n} = \mu\right\} = 1$$

The WLLN states that for any *specified* large value $n^*$, $(X_1 + X_2 + \cdots + X_{n^*})/n^*$ is likely to be near $\mu$. However, it does not say that $(X_1 + X_2 + \cdots + X_n)/n$ is bound to *stay* near $\mu$ for all values of $n$ larger than $n^*$. Thus, it leaves open the possibility that large values of $|(X_1 + X_2 + \cdots + X_n)/n - \mu|$ can occur infinitely often (though at infrequent intervals).
The SLLN shows that this cannot occur.