Limit theorems

Samy Tindel

Purdue University

Probability - MA 416

Mostly taken from *A first course in probability* by S. Ross

Outline

1. Weak law of large numbers
2. Central limit theorem

Outline

1. Weak law of large numbers

2. Central limit theorem

Weak law of large numbers

Theorem 1.

Consider

independent and identically distributed

- A sequence $\{X_i;\ i \geq 1\}$ of i.i.d random variables
- Write $\mathbf{E}[X_i] = \mu$
- Set

empirical mean

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i$$

Then for any $\varepsilon > 0$ we have

$$\lim_{n \to \infty} \mathbf{P}\left(|\bar{X}_n - \mu| > \varepsilon\right) = 0$$

As $n \to \infty$, empirical mean is close to the theoretical mean

Application We flip a fair coin $n = 400$ times. We set

$$X_i = \mathbb{1}(i\text{-th flip is } H)$$

Hyp: X_i's are iid with

$$X_i \sim \mathcal{B}(p = \tfrac{1}{2}), \quad \mathbb{E}[X_i] = p = \frac{1}{2}$$

Then LLN says

$$\mathbb{P}\left(\left| \frac{1}{400} \sum_{i=1}^{400} X_i - \frac{1}{2} \right| > 0.05 \right)$$

is "small" → Pb: Not a very precise statement

With CLT: We will get statements like

$$\mathbb{P}\left(\sum_{i=1}^{400} X_i > 210 \right) \simeq \text{a probability}$$

that can be computed

Outline

1. Weak law of large numbers

2. Central limit theorem

↘ = important

DeMoivre-Laplace theorem (repeated)

Theorem 2.

Let

- $n \geq 1$, $p \in (0, 1)$
- $X_n \sim \text{Bin}(n, p)$
- $a < b$

Rmk : $X_n = \sum_{i=1}^{n} Y_i$, with Y_i i.i.d with $Y_i \sim \mathcal{B}(p)$

Then

$$\lim_{n\to\infty} \mathbf{P}\left(a < \frac{X_n - \overbrace{np}^{=\mathbb{E}[X_n]}}{(\underbrace{np(1-p)}_{\mathrm{Var}(X_n)})^{1/2}} < b \right) = \Phi(b) - \Phi(a)$$

Another way to write De Moivre's theorem

Theorem 3.

Consider

- A sequence $\{Y_i;\ i \geq 1\}$ of indep. $\mathcal{B}(p)$ random variables
- We have $\mathbf{E}[Y_i] = p$ and $\mathbf{Var}(Y_i) = p(1-p) \equiv \sigma^2$
- Set

$$\bar{Y}_n = \frac{1}{n} \sum_{i=1}^{n} Y_i = \frac{X_n}{n}$$

Then for any $\varepsilon > 0$ we have

$$\lim_{n\to\infty} \mathbf{P}\left(a < \sqrt{n}\left(\frac{\bar{Y}_n - p}{\sigma}\right) < b\right) = \Phi(b) - \Phi(a)$$

Central limit theorem

Theorem 4.

Consider

- A sequence $\{X_i;\ i \geq 1\}$ of i.i.d random variables
- Write $\mathbf{E}[X_i] = \mu$ and $\mathbf{Var}(X_i) = \sigma^2$
- Set

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i$$

Then for any $\varepsilon > 0$ we have

$$\lim_{n \to \infty} \mathbf{P}\left(a < \sqrt{n}\left(\frac{\bar{X}_n - \mu}{\sigma}\right) < b\right) = \Phi(b) - \Phi(a)$$

Interpretation of CLT

As $n \to \infty$

$$\sqrt{n}\left(\frac{\bar{X}_n - \mu}{\sigma}\right) \approx Z, \quad Z \sim \mathcal{N}(0,1)$$

$$\Rightarrow \quad \bar{X}_n - \mu \simeq \frac{\sigma Z}{\sqrt{n}}$$

$$\Rightarrow \quad \bar{X}_n \simeq \mu + \frac{\sigma Z}{\sqrt{n}}$$

$$\boxed{\bar{X}_n = E[X_i] + \text{Gauss. fluctuations of order } \frac{1}{\sqrt{n}}}$$

Problem 8.5 (1)

Situation: We have

- Fifty numbers rounded off to the nearest integer and then summed
- The individual round-off errors are uniformly distributed over $(-0.5, 0.5)$

Question:
Approximate the probability that the resultant sum differs from the exact sum by more than 3.

Model For $i = 1, \dots, n$, $n = 50$ let

X_i = i-th error

Hyp X_i's are iid, $X_i \sim \mathcal{U}([-0.5, 0.5])$

Recall: for $Y \sim \mathcal{U}([a,b])$ then

$$f_Y(y) = \frac{1}{b-a} \mathbb{1}_{[a,b]}(y)$$

$$\mathbb{E}[Y] = \frac{a+b}{2} \qquad V(Y) = \frac{(b-a)^2}{12}$$

Here

$$\boxed{\begin{aligned} \mathbb{E}[X_i] &= 0 \qquad V(Y) = \frac{1}{12} = \sigma^2 \\ &= \mu \end{aligned}}$$

$n = 50$ $\quad \mu = 0$

$\sigma = \frac{1}{\sqrt{12}}$

We wish to compute

$$\mathbb{P}\left(\left|\sum_{i=1}^{n} X_i\right| > 3\right)$$

in CLT: $\sqrt{n}\,\frac{(\bar{X}_n - \mu)}{\sigma}$

$$= 1 - \mathbb{P}\left(-3 \leq \sum_{i=1}^{n} X_i \leq 3\right)$$

$$= 1 - \mathbb{P}\left(-3/n \leq \bar{X}_n \leq 3/n\right)$$

$$= 1 - \mathbb{P}\left(\frac{(-3/n - \mu)}{\sigma}\sqrt{n} \leq \sqrt{n}\frac{(\bar{X}_n - \mu)}{\sigma} \leq \frac{(3/n - \mu)}{\sigma}\sqrt{n}\right)$$

CLT

$$\simeq 1 - \mathbb{P}\left(\frac{-3}{\sqrt{n}\,\sigma} \leq Z \leq \frac{3}{\sqrt{n}\,\sigma}\right)$$

$$= 1 - \mathbb{P}\left(-\frac{3\sqrt{12}}{\sqrt{50}} \leq Z \leq \frac{3\sqrt{12}}{\sqrt{50}}\right)$$

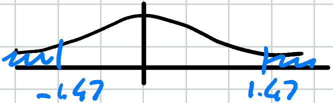
$$= 1 - \mathbb{P}\left(-1.47 \leq Z \leq 1.47\right)$$

Summary : We have found

$$\boxed{\mathbb{P}\left(\left|\sum_{i=1}^{n} X_i\right| > 3\right)}$$

$$\simeq 1-\mathbb{P}(-1.47 \le Z \le 1.47)$$

$$= \mathbb{P}(|Z| > 1.47)$$

$=$ 

$$= 2\,(1-\Phi(1.47))$$

$$\boxed{= \quad 0.14}$$

Problem 8.5 (2)

Model: Set

$$X_i \equiv i\text{-th error}$$

We wish to find

$$\mathbf{P}\left(\left|\sum_{i=1}^{n} X_i\right| > 3\right)$$

Law of X_i: We have

- X_i's i.i.d
- $X_i \sim \mathcal{U}([-0.5, 0.5)$
- $\mathbf{E}[X_i] = 0$, and $\mathbf{Var}(X_i) = \sigma^2 = \frac{1}{12}$

Problem 8.5 (3)

Application of CLT: Write

$$
\begin{aligned}
\mathbf{P}\left(\left|\sum_{i=1}^{n} X_i\right| > 3\right) &= \mathbf{P}\left(\left|\sqrt{n}\,\bar{X}_n\right| > \frac{3}{\sqrt{50}}\right) \\
&= \mathbf{P}\left(\left|\sqrt{n}\,\frac{\bar{X}_n}{\sigma}\right| > \frac{3\sqrt{12}}{\sqrt{50}}\right) \\
&\overset{CLT}{\simeq} \mathbf{P}\left(|Z| > 1.47\right) \\
&= .14
\end{aligned}
$$