CHAPTER 6

# Limit laws

Suppose $X_i$ are independent and have the same distribution. In the case of continuous or discrete random variables, this means they all have the same density. We say the $X_i$ are i.i.d., which stands for "independent and identically distributed." Let

$$S_n = \sum_{i=1}^{n} X_i.$$

$S_n$ is called the partial sum process.

THEOREM 1. *Suppose* $\mathbb{E}\,|X_i| < \infty$ *and let* $\mu = \mathbb{E}X_i$. *Then*

$$\frac{S_n}{n} \to \mu.$$

This is known as the strong law of large numbers (SLLN). The convergence here means that $S_n(\omega)/n \to \mu$ for every $\omega \in \Omega$, where $\Omega$ is the probability space, except possibly for a set of $\omega$ of probability $0$.

The proof of Theorem 1 is quite hard, and we prove a weaker version, the weak law of large numbers (WLLN). The WLLN states that for every $a > 0$,

$$\mathbb{P}\left(\left|\frac{S_n}{n} - \mathbb{E}X_1\right| > a\right) \to 0$$

as $n \to \infty$. It is not even that easy to give an example of random variables that satisfy the WLLN but not the SLLN.

Before proving the WLLN, we need an inequality called Chebyshev's inequality.

PROPOSITION 2. *If* $Y \geq 0$, *then for any* $A$,

$$\mathbb{P}\,(Y > A) \leq \frac{\mathbb{E}Y}{A}.$$

PROOF. We do the case for continuous densities, the case for discrete densities being similar. We have

$$\begin{aligned}\mathbb{P}(Y > A) &= \int_A^\infty f_Y(y)dy \le \int_A^\infty \frac{y}{A} f_Y(y)dy \\ &\le \frac{1}{A}\int_{-\infty}^\infty y f_Y(y)dy = \frac{1}{A}\mathbb{E}Y.\end{aligned}$$

□

We now prove the WLLN.

THEOREM 3. *Suppose the $X_i$ are i.i.d. and $\mathbb{E}|X_1|$ and $\mathrm{Var}(X_1)$ are finite. Then for every $a > 0$,*

$$\mathbb{P}\left(\left|\frac{S_n}{n} - \mathbb{E}X_1\right| > a\right) \to 0$$

*as $n \to \infty$.*

PROOF. Recall $\mathbb{E}S_n = n\mathbb{E}X_1$ and by the independence, Var $S_n = n$Var$X_1$, so Var $(S_n/n) = \mathrm{Var}X_1/n$. We have

$$\begin{aligned}\mathbb{P}\left(\left|\frac{S_n}{n} - \mathbb{E}X_1\right| > a\right) &= \mathbb{P}\left(\left|\frac{S_n}{n} - \mathbb{E}\left(\frac{S_n}{n}\right)\right| > a\right) \\ &= \mathbb{P}\left(\left|\frac{S_n}{n} - \mathbb{E}\left(\frac{S_n}{n}\right)\right|^2 > a^2\right) \\ &\le \frac{\mathbb{E}\left|\frac{S_n}{n} - \mathbb{E}\left(\frac{S_n}{n}\right)\right|^2}{a^2} \\ &= \frac{\mathrm{Var}\left(\frac{S_n}{n}\right)}{a^2} \\ &= \frac{\frac{\mathrm{Var}\, X_1}{n}}{a^2} \to 0.\end{aligned}$$

The inequality step follows from Proposition 2 with $A = a^2$ and $Y = \left|\frac{S_n}{n} - \mathbb{E}(\frac{S_n}{n})\right|^2$. □

We now turn to the central limit theorem (CLT).

THEOREM 4. *Suppose the* $X_i$ *are i.i.d. Suppose* $\mathbb{E}X_i^2 < \infty$. *Let* $\mu = \mathbb{E}X_i$ *and* $\sigma^2 = \mathrm{Var}X_i$. *Then*

$$\mathbb{P}\left(a \leq \frac{S_n - n\mu}{\sigma\sqrt{n}} \leq b\right) \to \mathbb{P}\left(a \leq Z \leq b\right)$$

*for every* $a$ *and* $b$, *where* $Z$ *is a* $\mathcal{N}(0,1)$.

The ratio on the left is $(S_n - \mathbb{E}S_n)/\sqrt{\mathrm{Var}S_n}$. We do not claim that this ratio converges for any $\omega$ (in fact, it doesn't), but that the probabilities converge.

EXAMPLE 5. Suppose we roll a die 3600 times. Let $X_i$ be the number showing on the $i^{th}$ roll. We know $S_n/n$ will be close to 3.5. What's the probability it differs from 3.5 by more than 0.05?

SOLUTION. We want

$$\mathbb{P}\left(\left|\frac{S_n}{n} - 3.5\right| > 0.05\right).$$

We rewrite this as

$$\mathbb{P}\left(|S_n - n\mathbb{E}X_1| > (0.05)(3600)\right) = \mathbb{P}\left(\left|\frac{S_n - n\mathbb{E}X_1}{\sqrt{n}\sqrt{\mathrm{Var}X_1}}\right| > \frac{180}{(60)\sqrt{\frac{35}{12}}}\right)$$

$$r \approx \mathbb{P}(|Z| > 1.756) \approx 0.08.$$

EXAMPLE 6. Suppose the lifetime of a human has expectation 72 and variance 36. What is the probability that the average of the lifetimes of 100 people exceeds 73?

SOLUTION. We want

$$\begin{aligned}\mathbb{P}\left(\frac{S_n}{n} > 73\right) &= \mathbb{P}(S_n > 7300) \\ &= \mathbb{P}\left(\frac{S_n - n\mathbb{E}X_1}{\sqrt{n}\sqrt{\mathrm{Var}X_1}} > \frac{7300 - (100)(72)}{\sqrt{100}\sqrt{36}}\right) \\ &\approx \mathbb{P}(Z > 1.667) \approx 0.047.\end{aligned}$$

The idea behind proving the central limit theorem is the following. It turns out that if $m_{Y_n}(t) \to m(t)$ for every $t$, then $\mathbb{P}(a \leq Y_n \leq b) \to \mathbb{P}(a \leq Z \leq b)$ . (We won't prove this.)

We are going to let $Y_n = (S_n - n\mu)/\sigma\sqrt{n}$. Let $W_i = (X_i - \mu)/\sigma$. Then $\mathbb{E}W_i = 0$, Var $W_i = \dfrac{\mathrm{Var} X_i}{\sigma^2} = 1$, the $W_i$ are independent, and

$$\frac{S_n - n\mu}{\sigma\sqrt{n}} = \frac{\sum_{i=1}^n W_i}{\sqrt{n}}.$$

So there is no loss of generality in assuming that $\mu = 0$ and $\sigma = 1$. Then

$$m_{Y_n}(t) = \mathbb{E}e^{tY_n} = \mathbb{E}e^{(t/\sqrt{n})(S_n)} = m_{S_n}(t/\sqrt{n}) \ .$$

Since the $X_i$ are i.i.d., all the $X_i$ have the same moment generating function. Since $S_n = X_1 + \cdot + X_n$, then

$$m_{S_n}(t) = m_{X_1}(t) \cdots m_{X_n}(t) = [m_{X_1}(t)]^n$$

If we expand $e^{tX_1}$ as a power series,

$$m_{X_1}(t) = \mathbb{E}e^{tX_1} = 1 + t\mathbb{E}X_1 + \frac{t^2}{2!}\mathbb{E}(X_1)^2 + \frac{t^3}{3!}\mathbb{E}(X_1)^3 + \cdots$$

We put the above together and obtain

$$\begin{aligned} m_{Y_n}(t) &= m_{S_n}\left(t/\sqrt{n}\right) \\ &= \left[m_{X_1}\left(t/\sqrt{n}\right)\right]^n \\ &= \left[1 + t \cdot 0 + \frac{\left(t/\sqrt{n}\right)^2}{2!} + R_n\right]^n \\ &= \left[1 + \frac{t^2}{2n} + R_n\right]^n, \end{aligned}$$

where $|R_n|/n \to 0$ as $n \to \infty$. This converges to $e^{t^2/2} = m(t)$ as $n \to \infty$.