1.) The Central Limit Theorem

**Theorem 3.1 (The Central Limit Theorem):** Suppose that $X_1, X_2, \cdots$ is a sequence of IID RVs with finite mean $\mu$ and variance $\sigma^2$. Then the distribution of the normalized partial sum

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

tends to the standard normal $\mathcal{N}(0,1)$ as $n \to \infty$: for all $z \in \mathbb{R}$,

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

The proof depends on the following result about moment generating functions which we state without proof.

**Lemma 3.1:** Let $Z_1, Z_2, \cdots$ be a sequence of random variables having cumulative distribution functions $F_{Z_n}$ and moment generating functions $\psi_{Z_n}(t)$, and let $Z$ be a random variable with cumulative distribution function $F_Z$ and moment generating function $\psi_Z$. Then pointwise convergence of the moment generating functions implies pointwise convergence of the cumulative distribution functions at every point of continuity of $F_Z$, i.e., if

$$\lim_{n \to \infty} \psi_{Z_n}(t) = \psi_Z(t)$$

for all $t \in \mathbb{R}$, then

$$\lim_{n \to \infty} F_{Z_n}(x) = F_Z(x)$$

at every $x$ at which $F_Z(x)$ is continuous.

**Proof of the C.L.T.** We will prove the C.L.T. under the assumption that the moment generating function of the random variables $X_i$, denoted $\psi(t)$ exists and is finite for all $t \in \mathbb{R}$. Let us first assume that $\mu = 0$ and $\sigma^2 = 1$. Notice that the moment generating function of the scaled random variable $X_i/\sqrt{n}$ is

$$\mathbb{E}\left[ \exp\left\{ \frac{tX_i}{\sqrt{n}} \right\} \right] = \psi\left( \frac{t}{\sqrt{n}} \right)$$

and that the moment generating function of the sum $\sum_{i=1}^{n} X_i/\sqrt{n}$ is equal to

$$\left[ \psi\left( \frac{t}{\sqrt{n}} \right) \right]^n.$$

Let

$$L(t) = \log \psi(t)$$
and observe that
\[ L(0) = 0 \]
\[ L'(0) = 0 \]
\[ L''(0) = 1. \]

By lemma 3.1, it suffices to show that
\[ \lim_{n \to \infty} \left[ \psi\left( \frac{t}{\sqrt{n}} \right) \right]^n = e^{t^2/2}, \]
which is equivalent to
\[ \lim_{n \to \infty} nL\left( \frac{t}{\sqrt{n}} \right) = \frac{t^2}{2}. \]

However, this last identity can be verified using L'Hôpital's rule
\[
\begin{align*}
\lim_{n \to \infty} \frac{L(t/\sqrt{n})}{n^{-1}} &= \lim_{n \to \infty} \frac{-L'(t/\sqrt{n})n^{-3/2}t}{-2n^{-2}} \\
&= \lim_{n \to \infty} \left[ \frac{L'(t/\sqrt{n})t}{2n^{-1/2}} \right] \\
&= \lim_{n \to \infty} \left[ \frac{-L''(t/\sqrt{n})n^{-3/2}t^2}{-2n^{-3/2}} \right] \\
&= \lim_{n \to \infty} \left[ L''\left( \frac{t}{\sqrt{n}} \right) \frac{t^2}{2} \right] \\
&= \frac{t^2}{2}.
\end{align*}
\]

The general case can then be handled by applying this result to the standardized variables
\[ X_i^* = \frac{X_i - \mu}{\sigma}, \]
which have mean 0 and variance 1.

\[ \square \]

Remark: What is amazing about the central limit theorem is that it says that the distribution of a sum of sufficiently many IID random variables is approximately normal, irrespective of the distribution of the individual variables (provided that these have finite variance). This may explain why so many natural phenomena are approximately normally distributed.

Example (Ross, 3d): Let \( X_1, \cdots, X_{10} \) be independent random variables, each uniformly distributed on \([0, 1]\), and let \( X = X_1 + \cdots + X_{10} \). The central limit theorem can be used to approximate the distribution of \( X \). For example,
\[
\mathbb{P}\{X > 6\} = \mathbb{P}\left\{ \frac{X - 5}{\sqrt{10/12}} > \frac{6 - 5}{\sqrt{10/12}} \right\}
\approx 1 - \Phi(\sqrt{1.2})
\approx 0.137.
\]
A version of the central limit theorem also holds when the random variables \(X_i\) are independent, but not necessarily identically-distributed.

**Theorem 3.2:** Let \(X_1, X_2, \ldots\) be a sequence of independent random variables with \(\mu_i = \mathbb{E}[X_i]\) and \(\sigma^2_i = \text{Var}(X_i)\). If the \(X_i\) are uniformly bounded, i.e., there exists \(M < \infty\) such that \(\mathbb{P}\{X_i < M\} = 1\) for all \(i \geq 1\), and if \(\sum_{i \geq 1} \sigma^2_i < \infty\), then

\[
\lim_{n \to \infty} \mathbb{P}\left\{ \frac{\sum_{i=1}^{n} (X_i - \mu_i)}{\sqrt{\sum_{i=1}^{n} \sigma^2_i}} \leq x \right\} = \Phi(x)
\]

for each \(x \in \mathbb{R}\).

2.) **The Strong Law of Large Numbers**

**Theorem 4.1:** The strong law of large numbers. Let \(X_1, X_2, \ldots\) be a sequence of IID random variables, each with finite mean \(\mu = \mathbb{E}[X]\). Then,

\[
\mathbb{P}\left\{ \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} X_i = \mu \right\} = 1.
\]

**Proof:** We will prove this result under the additional assumption that the \(X_i\) have a finite fourth moment \(K = \mathbb{E}[X_i^4] < \infty\). Without loss of generality, we may also assume that \(\mu = 0\). Let \(S_n = X_1 + \cdots + X_n\) and notice that

\[
\mathbb{E}[S_n^4] = n\mathbb{E}[X_i^4] + 6\binom{n}{2}\mathbb{E}[X_i^2X_j^2] = nK + 3n(n-1)\sigma^4.
\]

Dividing both sides by \(n^4\) gives

\[
\mathbb{E}\left[ \frac{S_n^4}{n^4} \right] = \frac{K}{n^3} + \frac{3\sigma^4}{n^2},
\]

and so

\[
\mathbb{E}\left[ \sum_{n=1}^{\infty} \frac{S_n^4}{n^4} \right] < \infty.
\]

In particular, this last inequality implies that

\[
\mathbb{P}\left\{ \sum_{n=1}^{\infty} \frac{S_n^4}{n^4} < \infty \right\} = 1,
\]

since otherwise the expectation would be infinite. Furthermore, because a series converges only if its terms tend to 0, we can conclude that

\[
\mathbb{P}\left\{ \lim_{n \to \infty} \frac{1}{n} S_n = 0 \right\} = \mathbb{P}\left\{ \lim_{n \to \infty} \frac{1}{n^4} S_n^4 = 0 \right\} = 1,
\]
which completes the proof.

\[ \square \]

**Remark:** Part of the significance of the strong law of large numbers is that it shows that the frequentist definition of probabilities can be deduced using measure-theoretic machinery. In particular, suppose that \((\Omega, \mathcal{F}, \mathbb{P})\) is a probability space and that \(E \in \mathcal{F}\) is an event, and let \(X_1, X_2, \cdots\) be independent indicator variables for \(E\):

\[
X_i = \begin{cases} 
1 & \text{if } E \text{ occurs on the } i\text{'th trial} \\
0 & \text{otherwise.}
\end{cases}
\]

Then \(\mathbb{E}[X_i] = \mathbb{P}(E)\) and so the strong law tells us that

\[
\mathbb{P} \left\{ \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} X_i = \mathbb{P}(E) \right\} = 1.
\]