1. Measures induced by random variables

A probability space is generally defined as a triple \((\Omega, \mathcal{F}, P)\), where \(\Omega\) is a set, \(\mathcal{F}\) is a Borel algebra of subsets of \(\Omega\), and \(P : \mathcal{F} \to \mathbb{R}\) a probability measure. Hence \(P\) is a positive measure on \(\mathcal{F}\) which satisfies \(P(\Omega) = 1\). A real valued random variable \(X\) on \(\Omega\) is then a Borel measurable function \(X : \Omega \to \mathbb{R}\), which means

\[
\text{an open subset of } \mathbb{R} \text{ implies } \{\omega \in \Omega : X(\omega) \in O\} \in \mathcal{F}.
\]

Rather than begin in such an abstract way, we can start with the cumulative distribution function (cdf) \(F(x) = P(X \leq x)\) of \(X\). The function \(F : \mathbb{R} \to [0, 1]\) is monotone increasing on \(\mathbb{R}\), and right continuous i.e.

\[
\lim_{x' \downarrow x} F(x') = F(x), \quad \lim_{x \to \infty} F(x) - \lim_{x \to -\infty} F(x) = 1.
\]

Standard measure theory tells us that there is a Borel measure \(P\) on \(\mathbb{R}\) such that

\[
P(\{x : a < x \leq b\}) = F(b) - F(a), \quad -\infty < a < b < \infty.
\]

Thus given the cdf \(F(.)\), we can construct \((\Omega, \mathcal{F}, P)\) and \(X : \Omega \to \mathbb{R}\) as above, in which \(\Omega = \mathbb{R}, \mathcal{F} = \text{Borel algebra generated by open sets of } \mathbb{R}\), and \(P\) is given by (1.3). The variable \(X\) is just the identity \(X(\omega) = \omega, \omega \in \mathbb{R}\). The probability space \((\Omega, \mathcal{F}, P)\) so constructed is the simplest probability space associated with the cdf \(F(.)\).

We can generalize the previous construction to many random variables \(X_1, ..., X_N\), once we know the joint cdf \(F(x_1, ..., x_N) = P(X_1 \leq x_1, ..., X_N \leq x_N)\). In that case \(\Omega = \mathbb{R}^N, \mathcal{F} = \text{Borel algebra generated by open sets of } \mathbb{R}^N\), and \(P\) is defined similarly to (1.3). This can be further generalized to a countably infinite set of variables \(X_1, X_2, ...\), provided the joint cdfs of the variables satisfy an obvious consistency condition. Thus for \(1 \leq j_1 < j_2 < \cdots < j_N\), let \(F_{j_1,j_2,...,j_N}(x_1, x_2, ... x_N)\) be the cdf of \(X_{j_1}, X_{j_2}, ..., X_{j_N}\). Consistency then just means:

\[
\lim_{x_k \to \infty} F_{j_1,j_2,...,j_N}(x_1, x_2, ... x_{k-1}, x_k, x_{k+1}, ..., x_N) = F_{j_1,j_2,...,j_{k-1},j_{k+1},...,j_N}(x_1, x_2, ... x_{k-1}, x_{k+1}, ..., x_N),
\]

for all possible choices of the \(j_i \geq 1, x_i \in \mathbb{R}\). This construction of \((\Omega, \mathcal{F}, P)\) for a countably infinite set of variables is known as the Kolmogorov construction. Here \(\Omega = \mathbb{R}^\infty\) and \(\mathcal{F}\) is the \(\sigma\) field generated by finite dimensional rectangles.

For a particularly simple set of variables the Kolmogorov construction is already done for us by the construction of the Lebesgue measure. Thus consider a Bernoulli variable \(X = 1\) with probability \(1/2\), and \(X = 0\) with probability \(1/2\). We wish to carry out the Kolmogorov construction for an infinite set \(X_1, X_2, ...\) of independent
Bernoulli variables. This can be done by setting \( \Omega = [0, 1] \), \( P = \text{Lebesgue measure} \), and \( X_j(\omega) \) the \( j \)th entry in the binary expansion of \( \omega \in [0, 1] \). Thus we have

\[
\omega = \sum_{j=1}^{\infty} \frac{X_j(\omega)}{2^j}, \quad \omega \in [0, 1].
\]

Note that this construction is actually simpler than the Kolmogorov construction since for that we have \( \Omega = \mathbb{R}^\infty \). Here we are using the fact that the space \( \{0, 1\}^\infty \) with probability measure induced by the counting measure on finite subsets, is identical to the interval \([0, 1]\) with Lebesgue measure. This illustrates an important point in probability theory. What matters are the values a random variable can take and the associated cdf. Generally we try to construct the simplest probability space \((\Omega, \mathcal{F}, P)\) consistent with this information.

2. Law of Large Numbers

If we take the Bernoulli variables \( X_1, X_2, \ldots \), generated above, then we have a model for the problem of independent tosses of a fair coin. Thus \( X_j = 1 \) if the \( j \)th toss is H and \( X_j = 0 \) if it is T. Hence \( S_N = X_1 + \cdots + X_N \) is the number of heads in \( N \) tosses of the coin. The law of averages tells us that

\[
\lim_{N \to \infty} \frac{S_N}{N} = \frac{1}{2} \quad \text{with probability 1.}
\]

The abstract measure theory we just presented gives us a precise way of formulating this law of averages.

**Proposition 2.1.** Let \( X_1, X_2, \ldots, \) be independent identically distributed (i.i.d.) variables on \((\Omega, \mathcal{F}, P)\) which have the property that \( \langle X_1^2 \rangle < \infty \). Then if \( S_N = X_1 + \cdots + X_N \), there is for any \( \varepsilon > 0 \) the limit

\[
\lim_{N \to \infty} \mathbb{P}\left( \left| \frac{S_N}{N} - \langle X_1 \rangle \right| > \varepsilon \right) = 0.
\]

**Proof.** We use the Chebyshev inequality that for any random variable \( Y \) and \( p, \delta > 0 \), one has

\[
\mathbb{P}(|Y| > \delta) \leq \langle |Y|^p \rangle / \delta^p.
\]

Hence taking \( p = 2 \) in (2.3), we have that

\[
P\left( \left| \frac{S_N}{N} - \langle X_1 \rangle \right| > \varepsilon \right) \leq \frac{\text{var}[S_N]}{\varepsilon^2 N^2} = \frac{N \text{var}[X_1]}{\varepsilon^2 N^2} \leq \frac{\langle X_1^2 \rangle}{\varepsilon^2 N}.
\]

Proposition 2.1 is known as the weak law of large numbers. It tells us that \( S_N/N \) converges in probability i.e. in measure to the average \( \langle X_1 \rangle \). However (2.1) is really saying something stronger than this.

**Proposition 2.2** (Strong Law of Large Numbers). Let \( X_1, X_2, \ldots, \) be i.i.d. variables on \((\Omega, \mathcal{F}, P)\) which have the property that \( \langle X_1^4 \rangle < \infty \). If \( S_N = X_1 + \cdots + X_N \), then

\[
\lim_{N \to \infty} \frac{S_N}{N} = \langle X_1 \rangle \quad \text{with probability 1.}
\]
Proof: Just as the weak law followed from estimating the variance of $S_N$, so the strong law (2.5) follows by estimating the fourth moment of $S_N - \langle S_N \rangle = S_N - N \langle X_1 \rangle$. It is easy to see that

$$\langle \{ S_N - \langle S_N \rangle \}^4 \rangle \leq CN^2$$

(2.6) for some constant $C$. Applying (2.3) with $p = 4$ yields

$$P\left( \left| \frac{S_N}{N} - \langle X_1 \rangle \right| > \varepsilon \right) \leq \frac{CN^2}{\varepsilon^4 N^4} = \frac{C}{\varepsilon^4 N^2}.$$  

(2.7) Thus for any $N_0 \geq 1$,

$$P\left( \limsup_{N \to \infty} \left| \frac{S_N}{N} - \langle X_1 \rangle \right| > \varepsilon \right) \leq \sum_{N=N_0}^{\infty} P\left( \left| \frac{S_N}{N} - \langle X_1 \rangle \right| > \varepsilon \right) \leq \sum_{N=N_0}^{\infty} \frac{C}{\varepsilon^4 N^2} \leq \frac{C'}{\varepsilon^4 N_0},$$

(2.8) for some constant $C'$. We conclude that for any $\varepsilon > 0$,

$$P\left( \limsup_{N \to \infty} \left| \frac{S_N}{N} - \langle X_1 \rangle \right| > \varepsilon \right) = 0.$$  

(2.9) Since $\varepsilon > 0$ is arbitrary we have then that

$$P\left( \limsup_{N \to \infty} \left| \frac{S_N}{N} - \langle X_1 \rangle \right| > 0 \right) = 0,$$

(2.10) whence (2.5) holds.  

\[\square\]

Remark 1. Proposition 2.2 applies to the coin tossing problem since in this case $\langle X_1^4 \rangle = 1$. However there are many cases where the fourth moment of $X_1$ is not finite. Since $\langle X_1 \rangle$ enters the RHS of (2.5), we might reasonably expect that (2.5) holds provided $\langle |X_1| \rangle < \infty$. This stronger result is not so easy to prove, and requires the development of new ideas which we shall discuss in future lectures.

3. The Central Limit Theorem

The SLLN tells us that the variable $S_N/N$ converges to $\langle X_1 \rangle$ as $N \to \infty$ for sums of i.i.d. variables. The central limit theorem allows us to estimate the fluctuation of $S_N/N$ from its average. Recall that if we define the variable $Z_N$ by

$$S_N = N \langle X_1 \rangle + \sqrt{N \text{var}[X_1]} Z_N,$$

(3.1) then $\langle Z_N \rangle = 0$ and $\text{var}[Z_N] = 1$. The CLT tells us that $Z_N$ is for large $N$ approximately a standard normal variable. Recall that a variable $Z$ is standard normal if it is defined by

$$P(Z \in O) = \int_O \frac{1}{\sqrt{2\pi}} e^{-z^2/2} \, dz,$$

(3.2) for all open sets $O$. Now the characteristic function $\chi_Y : \mathbb{R} \to \mathbb{C}$ of a variable $Y$ is defined by

$$\chi_Y(\sigma) = \langle e^{i\sigma Y} \rangle, \quad \sigma \in \mathbb{R},$$

(3.3)
Evidently \( \chi_Y \) satisfies \( \| \chi_Y(\cdot) \|_\infty \leq 1, \chi_Y(0) = 1 \). For the standard normal variable \( Z \) we easily see that \( \chi_Z(\sigma), \sigma \in \mathbb{R}, \) extends to an entire function \( \chi_Z : \mathbb{C} \to \mathbb{C} \). We can also evaluate \( \chi_Z(\sigma) \) for pure imaginary \( \sigma = it \) with \( t \in \mathbb{R} \) by change of variable,

\[
(3.4) \quad \chi_Z(it) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp(-tz - z^2/2) \, dz = e^{t^2/2} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp(-(z + i)^2/2) \, dz = e^{t^2/2} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp(-z'^2/2) \, dz' = e^{t^2/2} \chi_Z(0) = e^{t^2/2}.
\]

Using the fact now that \( \chi_Z : \mathbb{C} \to \mathbb{C} \) is entire, we conclude that \( \chi_Z(\sigma) = e^{-\sigma^2/2}, \sigma \in \mathbb{C} \).

**Lemma 3.1.** Suppose \( Z_N \) is defined by (3.1), where \( X_1 \) satisfies \( \langle |X_1|^3 \rangle < \infty \). Then

\[
(3.5) \quad \lim_{N \to \infty} \chi_{Z_N}(\sigma) = e^{-\sigma^2/2}, \quad \sigma \in \mathbb{R}.
\]

**Proof.** By independence of the variables \( X_1, \cdots, X_N \), we have that

\[
(3.6) \quad \log \chi_{Z_N}(\sigma) = N \log \chi_X(\sigma/\sqrt{N}),
\]

where \( X \) has the distribution of \( [X_1 - (X_1)]/\sqrt{\text{var}(X_1)} \), whence \( \langle X \rangle = 0, \langle X^2 \rangle = 1 \) and \( \langle |X|^3 \rangle < \infty \). Observe now that \( \chi_X(\cdot) \) is a \( C^3 \) function and

\[
(3.7) \quad \sup_{\sigma \in \mathbb{R}} \left| \frac{d^3 \chi_X(\sigma)}{d\sigma^3} \right| \leq \langle |X|^3 \rangle.
\]

Hence from Taylor’s theorem and the fact that \( \langle X \rangle = 0, \langle X^2 \rangle = 1 \), we conclude that

\[
(3.8) \quad \left| \chi_X \left( \frac{\sigma}{\sqrt{N}} \right) - 1 + \frac{\sigma^2}{2N} \right| \leq \frac{|\sigma|^3 \langle |X|^3 \rangle}{6N^{3/2}}.
\]

The result follows from (3.8) and the inequality

\[
(3.9) \quad z \leq -\log(1 - z) \leq z + z^2, \quad 0 \leq z < 1/2.
\]

Lemma 3.1 tells us that

\[
(3.10) \quad \lim_{N \to \infty} \langle f(Z_N) \rangle = \langle f(Z) \rangle,
\]

provided \( f \) is a function of the type \( f(z) = \exp[i\sigma z], \sigma \in \mathbb{R} \). We say the variables \( Z_N, N = 1, 2, \cdots \), converge in distribution to the variable \( Z \) if the cdfs of \( Z_N \) converge to the cdf of \( Z \) as \( N \to \infty \) i.e.

\[
(3.11) \quad \lim_{N \to \infty} P(Z_N \leq \xi) = P(Z \leq \xi), \quad \xi \in \mathbb{R}.
\]

Evidently we can rewrite (3.11) as (3.10) with \( f(z) = H(\xi - z), z \in \mathbb{R} \), where \( H(\cdot) \) is the Heaviside function

\[
(3.12) \quad H(z) = 0, \quad z < 0, \quad H(z) = 1, \quad z \geq 0.
\]
We can obtain a proof of (3.11) from Lemma 3.1 and some Fourier analysis. To do this we introduce the Schwartz space $S$ of functions $f : \mathbb{R} \to \mathbb{C}$ defined as follows: $f \in S$ if
\begin{equation}
\sup_{z \in \mathbb{R}} (1 + |z|)^m \left| \frac{d^n f(z)}{dz^n} \right| < \infty
\end{equation}
for all non-negative integers $m, n$. If we define now the Fourier transform of a function $f : \mathbb{R} \to \mathbb{C}$ by
\begin{equation}
\mathcal{F} f(\sigma) = \hat{f}(\sigma) = \int_{-\infty}^{\infty} f(z) e^{i\sigma z} \, dz, \quad \sigma \in \mathbb{R},
\end{equation}
then it is easy to see that $\mathcal{F}$ is an isomorphism of $S$ with inverse $\mathcal{F}^{-1}$ given by
\begin{equation}
f(z) = \mathcal{F}^{-1} \hat{f}(z) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{f}(\sigma) e^{-i\sigma z} \, d\sigma, \quad z \in \mathbb{R}.
\end{equation}

**Lemma 3.2.** Assume the conditions of Lemma 3.1. Then the limit (3.10) holds for all $f \in S$.

**Proof.** From (3.15) we have that
\begin{equation}
\langle f(Z_N) \rangle = \left\langle \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{f}(\sigma) e^{-i\sigma Z_N} \, d\sigma \right\rangle = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{f}(\sigma) \chi_{Z_N}(\sigma) \, d\sigma.
\end{equation}
Observe that in (3.16) we have used Fubini’s theorem. From Lemma 3.1 and the dominated convergence theorem we then see that
\begin{equation}
\lim_{N \to \infty} \langle f(Z_N) \rangle = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{f}(\sigma) e^{-\sigma^2/2} \, d\sigma.
\end{equation}
Now from (3.14) we have that
\begin{equation}
\frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{f}(\sigma) e^{-\sigma^2/2} \, d\sigma = \frac{1}{2\pi} \int_{-\infty}^{\infty} \left[ \int_{-\infty}^{\infty} f(z) e^{i\sigma z} \, dz \right] e^{-\sigma^2/2} \, d\sigma
\end{equation}
\begin{equation*}
= \int_{-\infty}^{\infty} \left[ \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\sigma z} e^{-\sigma^2/2} \, d\sigma \right] f(z) \, dz = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-z^2/2} f(z) \, dz = \langle f(Z) \rangle.
\end{equation*}

**Theorem 3.1** (Central Limit Theorem). Suppose $Z_N$ is defined by (3.1), where $X_1$ satisfies $\langle |X_1|^3 \rangle < \infty$. Then $Z_N$ converges in distribution as $N \to \infty$ to the standard normal variable $Z$.

**Proof.** Consider a finite interval $[a, b]$. Then for any $\varepsilon > 0$ there is a $C^\infty$ function $f_{1,\varepsilon} : \mathbb{R} \to \mathbb{R}$ with support in the open interval $(a, b)$ such that $\|f_{1,\varepsilon}\|_\infty \leq 1$ and $f_{1,\varepsilon}(z) = 1$ for $a + \varepsilon \leq z \leq b - \varepsilon$. Now since $f_{1,\varepsilon} \in S$, Lemma 3.2 implies that
\begin{equation}
\lim_{N \to \infty} \inf P(a < Z_N < b) \geq \langle f_{1,\varepsilon}(Z) \rangle.
\end{equation}
If we let $\varepsilon \to 0$ in (3.19) we conclude that
\begin{equation}
\lim_{N \to \infty} \inf P(a < Z_N < b) \geq P(a \leq Z \leq b).
\end{equation}
Similarly by choosing a $C^\infty$ function $f_{2,\varepsilon} : \mathbb{R} \to \mathbb{R}$ with support in the open interval $(a - \varepsilon, b + \varepsilon)$ such that $\|f_{2,\varepsilon}\|_\infty \leq 1$ and $f_{2,\varepsilon}(z) = 1$ for $a \leq z \leq b$, we conclude that
\begin{equation}
\lim_{N \to \infty} \sup P(a \leq Z_N \leq b) \leq P(a < Z < b).
\end{equation}
From (3.20), (3.21) it follows that for any $a < b$,
\begin{equation}
\lim_{N \to \infty} [P(Z_N \leq b) - P(Z_N \leq a)] = P(Z \leq b) - P(Z \leq a).
\end{equation}

The result follows from (3.22) if we can show that
\begin{equation}
\lim_{N \to \infty} P(Z_N \leq a) \leq K_a,
\end{equation}
where the constant $K_a$ satisfies $\lim_{a \to -\infty} K_a = 0$. To prove (3.23) we observe that for a $C^\infty$ function $f_a : \mathbb{R} \to \mathbb{R}$ with compact support in $(a, \infty)$ such that $\|f_a\|_\infty \leq 1$, one has
\begin{equation}
\liminf_{N \to \infty} P(Z_N > a) \geq \langle f_a(Z) \rangle.
\end{equation}

Now for $a << 0$ we choose $f_a$ to have the property that $f_a(z) = 1$ for $a+1 < z < |a|$. If we set the RHS of (3.24) to be $1 - K_a$, then $\lim_{a \to -\infty} K_a = 0$. □

**Remark 2.** We can compare the SLLN with the CLT. Since the SLLN tells us that
\begin{equation}
\lim_{N \to \infty} \frac{Z_N}{\sqrt{N}} = 0 \text{ with probability 1},
\end{equation}
and CLT that $Z_N$ converges to the normal variable $Z$ in distribution, we might expect that for any $\alpha > 0$,
\begin{equation}
\lim_{N \to \infty} \frac{Z_N}{N^{\alpha}} = 0 \text{ with probability 1}.
\end{equation}

We might even expect that $Z_N$ converges with probability 1 to $Z$, but this is always false whereas (3.26) is true for many i.i.d. variables. Thus the function $Z_N(\omega)$, $\omega \in \Omega$, is very oscillatory, but for fixed large $N$ the measure of the set of $\omega$ for which $Z_N(\omega)$ takes values between $a$ and $b$ remains more or less constant.

As an example we might think of $Z_N$ being like the function $Z_N(\omega) = \sin 2\pi N\omega$, $\omega \in [0, 1]$. Thus $Z_N(\omega)$ is very oscillatory but
\begin{equation}
\langle f(Z_N) \rangle = \int_0^1 f(\sin 2\pi N\omega) \, d\omega = \int_0^1 f(\sin 2\pi \omega) \, d\omega.
\end{equation}

Thus all the $Z_N$, $N \geq 1$, have the same distribution but $\lim_{N \to \infty} Z_N$ does not exist with probability 1.

### 4. Recurrence and Tail Events

We shall abstract part of the argument used in the proof of SLLN, in particular (2.8) to (2.10). For a probability space $(\Omega, \mathcal{F}, P)$, suppose that $A_n \in \mathcal{F}$, $n = 1, 2, \ldots$, is an infinite sequence of “events”. We define the lim sup and lim inf of this sequence as follows:
\begin{equation}
\limsup_{n \to \infty} A_n = \bigcap_{N=1}^\infty \bigcup_{n=N}^\infty A_n, \quad \liminf_{n \to \infty} A_n = \bigcup_{N=1}^\infty \bigcap_{n=N}^\infty A_n.
\end{equation}

Hence $\limsup_{n \to \infty} A_n$ is the event that $A_n$ occurs infinitely often, whereas $\liminf_{n \to \infty} A_n$ is the event that $A_n$ eventually always happens.

**Proposition 4.1** (Borel-Cantelli lemma). (a) Suppose that $\sum_{n=1}^\infty P(A_n) < \infty$. Then $P(\limsup_{n \to \infty} A_n) = 0$.
(b) Suppose that $\sum_{n=1}^\infty P(A_n) = \infty$ and the $A_n$, $n = 1, 2, \ldots$, are independent events. Then $P(\limsup_{n \to \infty} A_n) = 1$. 

Proof. (a) From (4.1) we have that

\[ P(\limsup_{n \to \infty} A_n) \leq P(\bigcup_{n=1}^{\infty} A_n) \leq \sum_{n=N}^{\infty} P(A_n) \]

for any \( N \geq 1 \). Since by our assumption the RHS of (4.2) goes to 0 as \( N \to \infty \) we are done.

(b) We observe that the complement of the set \( \limsup_{n \to \infty} A_n \) in \( \Omega \) is \( \liminf_{n \to \infty} \sim A_n \), where \( \sim A_n = \Omega - A_n \), \( n = 1, 2, \ldots \), are the complements of the \( A_n \) in \( \Omega \). Hence we have that

\[ 1 - P(\limsup_{n \to \infty} A_n) = P(\bigcup_{N=1}^{\infty} \cap_{n=N}^{\infty} \sim A_n), \]

which in turn implies that

\[ 1 - P(\limsup_{n \to \infty} A_n) = \lim_{N \to \infty} P(\cap_{n=N}^{\infty} \sim A_n). \]

Using the independence of the events \( A_n, n = 1, 2, \ldots \), we have that

\[ P(\cap_{n=N}^{\infty} \sim A_n) = \prod_{n=N}^{\infty} [1 - P(A_n)]. \]

Now from (3.9) we have that

\[ -\log \left\{ \prod_{n=N}^{\infty} [1 - P(A_n)] \right\} \geq \sum_{n=N}^{\infty} P(A_n) = \infty. \]

□

Remark 3. Note that in (2.8) the set \( A_n \) is the set \( A_n = \{ \omega \in \Omega : |S_n(\omega)/n - \langle X_1 \rangle| > \varepsilon \} \).

We can use Borel-Cantelli (b) to prove recurrence for the standard random walk on the integers \( Z \). Thus let the \( X_j, j = 1, 2, \ldots \), be Bernoulli variables taking the values \( \pm 1 \) with equal probability \( 1/2 \). Then \( S_N = X_1 + \cdots + X_N \) is the position after \( N \) steps of the standard random walk on \( Z \) starting at the origin. We wish to show that

\[ P(S_N = 0 \text{ for infinitely many } N) = 1. \]

Thus (4.7) says that the walk recurs to its starting point 0 infinitely often with probability 1. Following the argument of Borel-Cantelli (b) we need to first show that

\[ \sum_{n=1}^{\infty} P(S_n = 0) = \infty. \]

We can see why this is plausible from CLT. Thus using the fact that \( S_n \) takes only integer values and the fact that \( S_n/\sqrt{n} \) converges in distribution to the standard normal variable \( Z \), we expect that

\[ P(S_n = 0) \simeq \frac{1}{\sqrt{2\pi}} \int_{-1/2\sqrt{n}}^{1/2\sqrt{n}} e^{-z^2/2} \, dz \simeq \frac{1}{\sqrt{2\pi n}}. \]
for large \( n \). The approximation (4.9) is of course not correct because \( P(S_n = 0) = 0 \) if \( n \) is odd. We might however be tempted then to modify (4.9) to be double the RHS of (4.9) in the case of even integer \( n \). In that case we expect

\[
P(S_{2m} = 0) \simeq 2 \frac{1}{\sqrt{2\pi 2m}} = \frac{1}{\sqrt{\pi m}}.
\]

for large integer \( m \). We shall show using Sterling’s formula that (4.10) holds. This illustrates that approximation of \( S_N / \sqrt{N} \) by a standard normal variable can be much finer than what is given by CLT.

**Lemma 4.1.** For \( S_N, N = 1, 2, \ldots \), the standard random walk on \( \mathbb{Z} \) starting at the origin, then \( \lim_{m \to \infty} \sqrt{\pi m} P(S_{2m} = 0) = 1 \).

**Proof.** We have that

\[
P(S_{2m} = 0) = \frac{(2m)!}{(m!)^2} \frac{1}{2^{2m}}.
\]

Recall now Sterling’s formula that

\[
\lim_{n \to \infty} \frac{n!}{n^{n+1/2}e^{-n}} = \sqrt{2\pi}.
\]

Observe now that

\[
\sqrt{m} P(S_{2m} = 0) = \left[ \frac{(2m)!}{(2m)^{2m+1/2}e^{-2m}} \right] \left[ \frac{m^{m+1/2}e^{-m}}{(m!)^2} \right]^2 \sqrt{2}.
\]

From (4.12) we see that the RHS of (4.13) converges to \( 1/\sqrt{\pi} \) as \( m \to \infty \).

Evidently Lemma 4.1 implies (4.8), but (4.8) does not imply (4.7) since the events \( \{S_N = 0\}, N = 1, 2, \ldots \), are not independent. These events are however approximately independent and using that fact we can still prove (4.7). The notion of “approximate independence” is a bit fuzzy, but we can give it some quantitative meaning by considering covariances. Thus if \( X, Y \) are two variables, the covariance of \( X \) and \( Y \) is defined as

\[
\text{cov}[X, Y] = \langle (X - \langle X \rangle)(Y - \langle Y \rangle) \rangle = \langle XY \rangle - \langle X \rangle \langle Y \rangle.
\]

If \( X, Y \) are independent then \( \text{cov}[X, Y] = 0 \), but of course one can have \( \text{cov}[X, Y] = 0 \) and \( X, Y \) not be independent. We define the coefficient of correlation \( \rho \) for the variables \( X, Y \) as

\[
\rho(X, Y) = \frac{\text{cov}[X, Y]}{\sqrt{\text{var}[X]\text{var}[Y]}}.
\]

Evidently we have that \( -1 \leq \rho(X, Y) \leq 1 \). If \( \rho(X, Y) = \pm 1 \) then \( X \) is a multiple of \( Y \), which is the opposite of independence, and \( \rho(X, Y) = 0 \) if \( X, Y \) are independent. Therefore we may use the coefficient of correlation as a measure of independence. In particular if \( |\rho(X, Y)| < 1 \) then we could conclude that \( X, Y \) are “approximately independent”.

Let us apply the above considerations to the events \( \{S_N = 0\}, N = 1, 2, \ldots \). We define \( X_m \) by \( X_m(\omega) = 1 \) if \( S_{2m}(\omega) = 0 \), otherwise \( X_m(\omega) = 0 \). We have now, assuming from Lemma 4.1 that for large \( m \) and \( k \geq 1 \),

\[
\langle X_m \rangle = \frac{1}{\sqrt{\pi m}}, \quad \langle X_m X_{m+k} \rangle = \frac{1}{\pi \sqrt{mk}}.
\]
that \( \rho(X_m, X_{m+k}) \) is given by the formula
\begin{equation}
(4.17) \quad \rho(X_m, X_{m+k}) = \frac{1}{\pi} \left[ \frac{1}{\sqrt{m+k}} - \frac{1}{\sqrt{m}} \right] \frac{1}{\pi} \left[ \frac{1}{\sqrt{m+k}} - \frac{1}{\sqrt{m+k}} \right].
\end{equation}

Thus we have for fixed \( k \),
\begin{equation}
(4.18) \quad \lim_{m \to \infty} \rho(X_m, X_{m+k}) = \frac{1}{\sqrt{\pi k}}.
\end{equation}

We conclude that the sequence of variables \( X_{rk}, r = 1, 2, \ldots \), is approximately independent for large fixed \( k \).

Since Lemma 4.1 implies that
\begin{equation}
(4.19) \quad \sum_{r=1}^{\infty} P(X_{rk} = 1) = \infty,
\end{equation}

we can more or less conclude the recurrence (4.7) from Borel-Cantelli (b). This is of course not a rigorous argument, but still an important indicator of why the result is correct.

Now to the rigorous argument:

**Proposition 4.2.** For \( S_N, N = 1, 2, \ldots \), the standard random walk on \( \mathbb{Z} \), the recurrence (4.7) holds.

**Proof.** We consider the disjoint events \( A_k, k = 0, 1, 2, \ldots \), defined by
\begin{equation}
(4.20) \quad A_0 = \{ \omega : S_N(\omega) \neq 0 \text{ for all } N = 1, 2, \ldots \},
A_k = \{ \omega : S_k(\omega) = 0, S_{N+k}(\omega) \neq 0 \text{ for all } N = 1, 2, \ldots \}, \quad k \geq 1.
\end{equation}

Thus \( A_k \) is the event that \( S_N \) returns to 0 for the last time at \( N = k \). Since these are disjoint events we have
\begin{equation}
(4.21) \quad \sum_{k=0}^{\infty} P(A_k) \leq 1.
\end{equation}

Observe now that we can rewrite \( A_k \) as
\begin{equation}
(4.22) \quad A_k = \{ \omega : S_k(\omega) = 0, S_{N+k}(\omega) - S_k(\omega) \neq 0 \text{ for all } N = 1, 2, \ldots \}.
\end{equation}

Hence by the independence of the events \( \{S_k = 0\} \) and \( \{S_{N+k} - S_k \neq 0 \text{ for all } N = 1, 2, \ldots\} \), we conclude from (4.22) that
\begin{equation}
(4.23) \quad P(A_k) = P(S_k = 0)P(A_0).
\end{equation}

Substituting (4.23) into (4.21) and using (4.8), we conclude that \( P(A_0) = 0 \). Now (4.23) implies that \( P(A_k) = 0, k = 0, 1, 2, \ldots \). To finish the proof we observe that
\begin{equation}
(4.24) \quad \bigcup_{k=0}^{\infty} A_k = \text{ event that } S_N = 0 \text{ finitely often,}
\end{equation}

and we have shown the probability of the LHS of (4.24) is 0. \( \square \)

We have observed in the Borel-Cantelli Lemma and in the previous proposition that the probability of an event is either 0 or 1. These events are so called “tail events” since they are independent of any finite number of the variables involved. The following systematizes this observation:

**Proposition 4.3** (Kolmogorov zero-one law). Let \( X_1, X_2, \ldots \) be independent variables on a probability space \( (\Omega, \mathcal{F}, P) \). Suppose \( E \in \mathcal{F} \) is in the \( \sigma \) field generated by the \( X_1, X_2, \ldots \), but is independent of any finite number of the variables \( X_1, X_2, \ldots \). Then \( P(E) = 0 \) or \( P(E) = 1 \).
Proof. $E$ can be approximated arbitrarily closely by sets $E_N$ in the $\sigma$ field $\mathcal{F}(X_1, X_2, ..., X_N)$ for large enough $N$. Hence we have that

\begin{equation}
\lim_{N \to \infty} P(E_N) = P(E), \quad \lim_{N \to \infty} P(E_N \cap E) = P(E).
\end{equation}

By assumption we have $P(E_N \cap E) = P(E_N) P(E)$, whence (4.25) implies that $P(E) = P(E)^2$. □

We can apply Proposition 4.3 to the issue raised at the end of §3, in particular (3.26). Observe that for $\alpha > 0$ the set \{ω : lim$_{N \to \infty}$ $Z_N(\omega)/N^\alpha$ exists\} is a tail event, whence it follows that it occurs with probability 1 or 0. We have seen from the SLLN that it occurs with probability 1 if $\alpha = 1/2$ and the limit is 0. Next we show that if $\alpha = 0$ it occurs with probability 0.

**Proposition 4.4.** Let $X_1, X_2, ..., \in i.i.d.$ variables on $(\Omega, \mathcal{F}, P)$ and assume that

\langle|X_1|^3\rangle < \infty. Then $P(\{\omega : \lim_{N \to \infty} Z_N(\omega) \text{ exists}\} ) = 0$.

Proof. By the Kolmogorov law we can assume for contradiction that \{ω : lim$_{N \to \infty}$ $Z_N(\omega)$ exists\} occurs with probability 1. Thus there exists a variable $Z$ and lim$_{N \to \infty} Z_N = Z$ with probability 1. However by the CLT this limiting variable $Z$ must be the standard normal variable, whence

\begin{equation}
P(\{\omega : \lim_{N \to \infty} Z_N(\omega) > 0\} ) = 1/2.
\end{equation}

Since \{ω : lim$_{N \to \infty}$ $Z_N(\omega) > 0\}$ is a tail event we have obtained a contradiction to the Kolmogorov law. Thus \{ω : lim$_{N \to \infty}$ $Z_N(\omega)$ exists\} occurs with probability 0. □

## 5. Law of the Iterated Logarithm

In this section we shall be considering the standard random walk $S_N = X_1 + \cdots + X_N$ on $Z$, so the $X_j, j = 1, 2, ..., \in i.i.d.$ Bernoulli taking values $\pm 1$ with probability 1/2. Suppose now that $b_n, n = 1, 2, ..., \in$ is a positive sequence satisfying

\begin{equation}
\lim_{n \to \infty} b_n = \infty.
\end{equation}

By Kolmogorov the set $A$ given by

\begin{equation}
A = \{ \omega : \limsup_{n \to \infty} |S_n(\omega)|/b_n < \infty \}
\end{equation}

is a tail event and hence $P(A) = 0$ or $P(A) = 1$. Assume now that $P(A) = 1$. Invoking Kolmogorov again, we see that there exists $\gamma \geq 0$ such that

\begin{equation}
\limsup_{n \to \infty} |S_n(\omega)|/b_n = \gamma \quad \text{with probability 1}.
\end{equation}

We have already seen that if $b_n = n$ then $\gamma = 0$. We also know that if $b_n = \sqrt{n}$ then $P(A) = 0$. A natural question to ask therefore is to find a sequence $b_n$ such that (5.3) holds for some $\gamma$ with $0 < \gamma < \infty$. The answer is given by the following:

**Theorem 5.1** (Law of the Iterated Logarithm). If $b_n = [2n \log \log n]^{1/2}$, then (5.3) holds with $\gamma = 1$. 

To prove Theorem 5.1 we use estimates on the fluctuations of the random walk which come from CLT, but we also need an important new idea, the idea of a maximal function. We define the maximal function $M_N$ associated with the random walk $S_N$ to be

\[ M_N(\omega) = \sup_{0 \leq n \leq N} S_n(\omega), \quad \omega \in \Omega. \]

Thus $M_N$ measures the farthest to the right the walk has wandered in $N$ steps. Evidently $M_N$ is a non-negative variable, but there is no obvious relationship between the distribution of $M_N$ and $S_N$. The reflection principle tells us that the distribution of $M_N$ and $|S_N|$ are almost identical.

**Lemma 5.1.** Let $r \geq 1$ be an integer. Then there is the inequality,

\[ 2P(S_N \geq r + 1) \leq P(M_N \geq r) \leq 2P(S_N \geq r). \]

**Proof.** Observe that

\[ P(M_N \geq r) - P(S_N \geq r) = P(M_{N-1} \geq r, S_N < r). \]

Reflection symmetry comes in by noting that

\[ P(M_{N-1} \geq r, S_N < r) = P(M_{N-1} \geq r, S_N > r). \]

To see (5.7) we use a random stopping time $n^*$ defined by

\[ n^*(\omega) = \inf\{ n \geq 1 : S_n(\omega) = r \}. \]

Thus we have that

\[ P(M_{N-1} \geq r, S_N < r) = P(\{ \omega : n^*(\omega) < N, S_{N-n^*(\omega)} < 0 \}) = P(\{ \omega : n^*(\omega) < N, S_{N-n^*(\omega)} > 0 \}) = P(M_{N-1} \geq r, S_N > r). \]

Now (5.6), (5.7) imply that

\[ P(M_N \geq r) = \frac{1}{2} P(M_{N-1} \geq r, S_N \neq r) + P(S_N \geq r) \leq \frac{1}{2} P(M_N \geq r) + P(S_N \geq r), \]

which implies the upper bound in (5.5).

We also get the lower bound by using the identity in (5.10). Thus we have from (5.10) that

\[ P(M_N \geq r) \geq \frac{1}{2} P(M_{N-1} \geq r) + \frac{1}{2} P(S_N = r) + P(S_N \geq r + 1) \]

\[ \geq \frac{1}{2} P(M_N \geq r) + P(S_N \geq r + 1). \]

\[ \square \]

Next we obtain some estimates on the asymptotics for the cdf of the normal variable.

**Lemma 5.2.** For $Z$ the standard normal variable and $a > 0$, there are the inequalities

\[ \frac{1}{\sqrt{2\pi}} \left( \frac{1}{a} - \frac{1}{a^3} \right) e^{-a^2/2} \leq P(Z > a) \leq \frac{1}{a\sqrt{2\pi}} e^{-a^2/2}. \]
Proof. We have that
\[ P(Z > a) = \frac{1}{\sqrt{2\pi}} \int_a^\infty e^{-z^2/2} \, dz = \frac{1}{\sqrt{2\pi}} e^{-a^2/2} \int_0^\infty e^{-ax-x^2/2} \, dx , \]
upon making the change of variable \( z = x + a \) in the integration. The upper bound in (5.12) follows from
\[ \int_0^\infty e^{-ax-x^2/2} \, dx \leq \int_0^\infty e^{-ax} \, dx = 1/a . \]
The estimate \( 1/a \) in (5.14) is just the first term in an asymptotic expansion for the integral in powers of \( 1/a \) which should be valid for \( a >> 1 \). The second term is obtained by making the approximation
\[ e^{-x^2/2} \simeq 1 - x^2/2 , \]
in which case we get
\[ \int_0^\infty e^{-ax-x^2/2} \, dx \simeq 1/a - 1/a^3 . \]
We can turn the calculation (5.16) into a rigorous lower bound by noting that
\[ - \frac{d}{dz} \left[ \left( \frac{1}{z} - \frac{1}{z^3} \right) e^{-z^2/2} \right] = \left( 1 - \frac{3}{z^4} \right) e^{-z^2/2} . \]
Hence
\[ \int_0^\infty e^{-ax-x^2/2} \, dx = e^{a^2/2} \int_a^\infty e^{-z^2/2} \, dz \geq e^{a^2/2} \int_a^\infty \left( 1 - \frac{3}{z^4} \right) e^{-z^2/2} \, dz = \frac{1}{a} - \frac{1}{a^3} , \]
whence the lower bound in (5.12) follows. \( \square \)

The fine asymptotics in Lemma 5.2 shows us how the logarithm gets involved in LIL.

**Lemma 5.3.** There is the inequality
\[ \sum_{n=2}^\infty P \left( \frac{S_n}{2n(1 + \varepsilon) \log n}^{1/2} > 1 \right) < \infty \quad \text{or} \quad = \infty \]
according as \( \varepsilon > 0 \) or \( \varepsilon < 0 \).

Proof. From Lemma 5.2 and CLT we have that
\[ P \left( \frac{S_n}{2n(1 + \varepsilon) \log n}^{1/2} > 1 \right) = P \left( \frac{S_n}{\sqrt{n}} > [2(1 + \varepsilon) \log n]^{1/2} \right) \simeq \frac{1}{[4\pi(1 + \varepsilon) \log n]^{1/2} e^{-(1+\varepsilon) \log n}} = \frac{1}{[4\pi(1 + \varepsilon) \log n]^{1/2} n^{1+\varepsilon}} . \]
Since
\[ \sum_{n=1}^\infty \frac{1}{n^{1+\varepsilon}} < \infty \quad \text{or} \quad = \infty \]
according as \( \varepsilon > 0 \) or \( \varepsilon < 0 \), the result follows. \( \square \)
Remark 4. From Lemma 5.3 and Borel-Cantelli (a) we conclude that
\[
\left(5.22\right) \limsup_{n \to \infty} \frac{S_n}{[2n(1+\varepsilon) \log n]^{1/2}} \leq 1 \text{ with probability 1}
\]
if \(\varepsilon > 0\), whence it follows for \(\varepsilon = 0\).

In order to go further than (5.22) we need to use Lemma 5.1.

Lemma 5.4. With \(b_n\) the sequence in the statement of Theorem 5.1, then
\[
\left(5.23\right) \limsup_{n \to \infty} \frac{S_n}{b_n} \leq 1 \text{ with probability 1}.
\]

Proof. We shall show using Lemma 5.1 that for any \(\alpha > 1\)
\[
\left(5.24\right) \sum_{m=1}^{\infty} P \left( \sup_{\alpha^m \leq n < \alpha^{m+1}} \frac{S_n}{b_n} > \alpha \right) < \infty.
\]

The result follows from (5.24) and Borel-Cantelli (a) on letting \(\alpha \to 1\).

To prove (5.24) consider integers \(r, k\) which satisfy for some integer \(m \geq 1\) the inequality \(\alpha^m \leq r, k \leq \alpha^{m+1}\). Then we have that
\[
\left(5.25\right) \frac{\gamma(m)}{\sqrt{\alpha}} \leq \frac{b_r}{b_k} \leq \frac{\sqrt{\alpha}}{\gamma(m)}.
\]

where \(\lim_{m \to \infty} \gamma(m) = 1\). Thus in intervals \(\alpha^m \leq n \leq \alpha^{m+1}\) the sequence \(b_n\) is essentially constant for \(\alpha\) close to 1. Using this observation we note then that we may estimate
\[
\left(5.26\right) P \left( \sup_{\alpha^m \leq n < \alpha^{m+1}} \frac{S_n}{b_n} > \alpha \right) \leq P \left( \sup_{n \leq \alpha^{m+1}} S_n > \alpha b_{\alpha^m} \right)
\leq 2P \left( S_{\alpha^{m+1}} > \alpha b_{\alpha^m} \right) = 2P \left( \frac{S_{\alpha^{m+1}}}{\sqrt{\alpha^{m+1}}} > \alpha \left[ \frac{2\xi(m) \log m}{\alpha} \right]^{1/2} \right)
\]
where \(\lim_{m \to \infty} \xi(m) = 1\). Using the CLT and Lemma 5.2 we have that
\[
\left(5.27\right) P \left( \frac{S_{\alpha^{m+1}}}{\sqrt{\alpha^{m+1}}} > \alpha \left[ \frac{2\xi(m) \log m}{\alpha} \right]^{1/2} \right) \sim
\frac{1}{\sqrt{4\pi\alpha \xi(m) \log m}} \exp \left[ -\alpha \xi(m) \log m \right] = \frac{1}{\sqrt{4\pi\alpha \xi(m) \log m}} \frac{1}{m^{\alpha \xi(m)}}.
\]
The inequality (5.24) follows from (5.27) since
\[
\left(5.28\right) \sum_{m=1}^{\infty} \frac{1}{m^{\alpha \xi(m)}} < \infty.
\]

\[\square\]

Lemma 5.5. With \(b_n\) the sequence in the statement of Theorem 5.1, then
\[
\left(5.29\right) \limsup_{n \to \infty} \frac{S_n}{b_n} \geq 1 \text{ with probability 1}.
\]
Proof. Taking $\alpha > 1$ again and setting $\beta = \alpha / (\alpha - 1) > 1$ we consider

\begin{equation}
P \left( S_{\alpha m+1} - S_{\alpha m} > \frac{1}{\beta} b_{\alpha m+1} \right) = P \left( \frac{S_{\alpha m} - (\alpha - 1)}{\sqrt{\alpha m (\alpha - 1)}} > \left[ \frac{2\xi(m) \log m}{\beta} \right]^{1/2} \right)
\end{equation}

\[ \simeq \left[ \frac{\beta}{4\pi \xi(m) \log m} \right]^{1/2} \exp \left(-\frac{\xi(m) \log m}{\beta} \right) = \left[ \frac{\beta}{4\pi \xi(m) \log m} \right]^{1/2} \frac{1}{\beta} \frac{1}{m^{\xi(m)/\beta}}. \]

Sine $\beta > 1$ we conclude that

\begin{equation}
\sum_{m=1}^{\infty} P \left( S_{\alpha m+1} - S_{\alpha m} > \frac{1}{\beta} b_{\alpha m+1} \right) = \infty.
\end{equation}

Since the events in (5.31) are independent we conclude from Borel-Cantelli (b) that

\begin{equation}
S_{\alpha m+1} - S_{\alpha m} > \frac{1}{\beta} b_{\alpha m+1} \text{ infinitely often with probability 1.}
\end{equation}

Combining Lemma 5.4 with (5.32) we see that

\begin{equation}
S_{\alpha m+1} > \frac{1}{\beta} b_{\alpha m+1} - \alpha^{1/4} b_{\alpha m} \text{ infinitely often with probability 1,}
\end{equation}

whence we have that

\begin{equation}
\frac{S_{\alpha m+1}}{b_{\alpha m+1}} > \frac{1}{\beta} - \frac{\xi(m)}{\alpha^{1/4}} \text{ infinitely often with probability 1.}
\end{equation}

Now let $\alpha \to \infty$ in (5.34) to get the result, noting that then $\beta \to 1$. $\square$

Proof of Theorem 5.1. Simply observe that

\begin{equation}
\limsup_{n \to \infty} \frac{S_n}{b_n} = \max \left[ \limsup_{n \to \infty} \frac{S_n}{b_n}, \liminf_{n \to \infty} \frac{S_n}{b_n} \right].
\end{equation}

From the previous lemmas we have by symmetry that

\begin{equation}
\liminf_{n \to \infty} \frac{S_n}{b_n} = -\limsup_{n \to \infty} \frac{S_n}{b_n} = -1,
\end{equation}

whence the RHS of (5.35) is 1. $\square$