1 Weak law of large numbers

In class we saw a simple proof of the weak law of large numbers for random variables with finite variance:

**Theorem 1** (WLLN for finite variance). If $X_i, i \geq 1$ are pairwise independent random variables with mean $\mu$ and $\sup_{i \geq 1} \mathbb{E} \{X_i^2\} = K < \infty$, then writing $S_n = \sum_{i=1}^{n} X_i$, for all $\varepsilon > 0$,

$$
\mathbb{P} \left\{ \left| \frac{S_n}{n} - \mu \right| \geq \varepsilon \right\} \to 0,
$$
as $n \to \infty$.

**Lemma 1** (Chebyshev’s inequality for sums.). If $X_i, 1 \leq i \leq n$ are pairwise independent and $\mathbb{E} \{X_i^2\} < \infty$ for each $1 \leq i \leq n$ then

$$
\mathbb{P} \{|S_n - \mathbb{E} S_n| \geq t\} \leq \frac{\sum_{i=1}^{n} \mathbf{V} \{X_i\}}{t^2}
$$

**Proof.** Since the random variables are pairwise independent, $\mathbf{V}(S_n) = \sum_{i=1}^{n} \mathbf{V}(X_i)$, and the result follows by Chebyshev’s inequality. \(\square\)

In the case that the $X_i$ all have the same variance $\sigma^2$, this yields the bound

$$
\mathbb{P} \{|S_n - \mathbb{E} S_n| \geq t\} \leq \frac{n \sigma^2}{t^2}.
$$

**Proof of WLLN for finite variance.** For fixed $\varepsilon > 0$ we have

$$
\mathbb{P} \left\{ \left| \frac{S_n}{n} - \mu \right| \geq \varepsilon \right\} = \mathbb{P} \{|S_n - \mathbb{E} S_n| \geq \varepsilon n\} \\
\leq \frac{\sum_{i=1}^{n} \mathbf{V}\{X_i\}}{\varepsilon^2 n^2} \\
\leq \frac{K n}{\varepsilon^2 n^2} = \frac{K}{\varepsilon^2 n},
$$

and the last term tends to zero as $n \to \infty$. \(\square\)
We now use the same idea for random variables with possibly infinite variance (but additionally assuming the random variables are identically distributed). We obviously can’t directly use the same proof in this case; we will instead argue by truncation. Before stating the theorem, notice that if a random variable $X$ almost surely satisfies $a \leq X \leq b$ then we always have $|X - \mathbb{E}X| \leq b - a$, and so

$$\mathbb{V}\{X\} = \mathbb{E}\{(X - \mathbb{E}X)^2\} \leq |b - a|^2.$$  

**Exercise.** Strengthen the above bound to $|b - a|^2/4$.

**Theorem 2.** If $X_i, i \geq 1$ are identically distributed, pairwise independent random variables with $\mathbb{E}|X| < \infty$ and $\mathbb{E}|X| = \mu$, then writing $S_n = \sum_{i=1}^n X_i$, for all $\varepsilon > 0$,

$$\mathbb{P}\left\{ \left| \frac{S_n}{n} - \mu \right| \geq \varepsilon \right\} \rightarrow 0,$$

as $n \rightarrow \infty$.

**Proof.** For fixed $N > 0$, we define $X_{k,N}^\leq$ and $X_{k,N}^>$ as follows: $X_{k,N}^\leq = X_k 1_{|X_k| \leq N}$ and $X_{k,N}^> = X_k - X_{k,N}^\leq$.

We then have that $|X_{1,N}^\leq|$ increases to $|X_1|$ as $N \rightarrow \infty$, so by monotone convergence

$$\mathbb{E}|X_{1,N}^\leq| \rightarrow \mathbb{E}|X_1|,$$

again as $N \rightarrow \infty$. Since $|X_1| = |X_{1,N}^\leq| + |X_{1,N}^>|$ (check if it isn’t obvious to you), it follows that as $N \rightarrow \infty$ we also have

$$\mathbb{E}|X_{1,N}^>| \rightarrow 0.$$

Now fix $\varepsilon > 0$, and let $N$ be large enough that $\mathbb{E}|X_{1,N}^>| < \varepsilon^2/8$. By Chebyshev’s inequality for sums, we then have

$$\mathbb{P}\left\{ |\bar{S}_{n,N}^\leq - \mathbb{E}\bar{S}_{n,N}^\leq| > \varepsilon/2 \right\} \leq \frac{1}{(\varepsilon/2)^2 n} \mathbb{V}(X_{1,N}^\leq) \leq \frac{4N^2}{\varepsilon^2 n},$$

the last inequality since $-N \leq X_1 \leq N$ so $\mathbb{V}(X_{1,N}^\leq) \leq (2N)^2/4 = N^2$. The last expression is less than $\varepsilon/2$ for $n > 8N^2/\varepsilon^3$. We then have

$$\mathbb{P}\left\{ |\bar{S}_{n,N}^> - \mathbb{E}\bar{S}_{n,N}^>| > \varepsilon/2 \right\} \leq \frac{\mathbb{E}\left\{ |\bar{S}_{n,N}^> - \mathbb{E}\bar{S}_{n,N}^>| \right\}}{(\varepsilon/2)} \leq \frac{\mathbb{E}\left\{ |\bar{S}_{n,N}^>| \right\} + |\mathbb{E}\bar{S}_{n,N}^>|}{(\varepsilon/2)} \leq \frac{4\mathbb{E}\left\{ |\bar{S}_{n,N}^>| \right\}}{\varepsilon} \leq \frac{\varepsilon}{2}$$

(Markov’s inequality)

(Triangle inequality)

(Move absolute value inside expectation)

(Since $\mathbb{E}|\bar{S}_{n,N}^>| \leq \mathbb{E}|X_{1,N}^>| < \varepsilon^2/8$).

It follows that for $n > 8N^2/\varepsilon^3$,

$$\mathbb{P}\left\{ |\bar{S}_n - \mathbb{E}\{X_1\}| > \varepsilon \right\} \leq \mathbb{P}\left\{ |\bar{S}_{n,N}^\leq - \mathbb{E}\bar{S}_{n,N}^\leq| > \varepsilon/2 \right\} + \mathbb{P}\left\{ |\bar{S}_{n,N}^> - \mathbb{E}\bar{S}_{n,N}^>| > \varepsilon/2 \right\} < 2\varepsilon.$$

Since $\varepsilon > 0$ was arbitrary this shows convergence in probability. \qed
2 SLLN

Here is a version of the proof of the Strong Law of Large Numbers that we did in class and that can also be found on Terry Tao’s blog. The first step is reduction to non-negative random variables: by writing $X_n = X_n^+ - X_n^-$, we see that to prove $S_n/n \to \mathbb{E}(X_1)$ almost surely, it suffices to prove that

$$\frac{X_1^+ + \ldots + X_n^+}{n} \to \mathbb{E}(X_1^+),$$

and likewise for $X_n^-$. We thus reduce to proving a “non-negative strong law of large numbers”.

Next, recall that we say a sequence $\{n_k\}_{k \in \mathbb{N}}$ is lacunary if there exists $c > 1$ such that $n_{k+1} > cn_k$ for all $k$ sufficiently large. We then have the following theorem.

**Theorem 3** (Lacunary Strong Law of Large Numbers). Let $\{X_n\}_{n \in \mathbb{N}}$ be iid real random variables defined on a common probability space. If $X_1 \geq 0$ and $\mathbb{E}(X_1) < \infty$ then for any lacunary sequence $\{n_k\}_{k \in \mathbb{N}}$,

$$\mathbb{P}\left(\lim_{j \to \infty} \frac{X_1 + \ldots + X_{n_j}}{n_j} = \mathbb{E}(X_1)\right) = 1.$$  

**Exercise** Let $(s_n, n \geq 0)$ be a non-decreasing sequence with $s_0 = 0$. Fix $\mu > 0, \varepsilon \in (0, 1/3)$, and define a sequence by $n_k = \lceil (1 + \varepsilon)^k \rceil$.

(a) Show that for all $n$ sufficiently large (i.e. $n \geq n_0(\varepsilon)$), if $s_n \geq \mu n(1 + 3\varepsilon)$ then letting $k$ be such that $n_{k-1} < n \leq n_k$, we have $s_{n_k} \geq \mu n(1 + \varepsilon)$.

(b) Show that for all $n$ sufficiently large, if $s_n \leq \mu n(1 - 3\varepsilon)$ then letting $k$ be such that $n_{k-1} < n \leq n_k$, we have $s_{n_k} \leq \mu n(1 - \varepsilon)$.

(c) Conclude that $\lim sup_n \{|s_n - \mu n| > 3\varepsilon\} \subset lim sup_k \{|s_{n_k} - \mu n_k| > \varepsilon\}$. Then deduce the non-negative strong law from the lacunary strong law.

In view of the above exercise, to prove the strong law it suffices to prove the lacunary strong law. For this we use the Borel-Cantelli lemma; here is one way it can be stated.

**Lemma 2.** Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $\{E_n\}_{n \in \mathbb{N}}$ be events (elements of $\mathcal{F}$). Let $J : \Omega \to \mathbb{N} \cup \{\infty\}$, $J(\omega) = \sup\{n : \omega \in E_n\}$. If $\sum_{n=1}^{\infty} \mathbb{P}(E_n) < \infty$ then $\mathbb{P}(J < \infty) = 1$.

**Proof of Lacunary SLLN.** In this proof write $\overline{S}_n = (X_1 + \ldots + X_n)/n$. Fix a lacunary sequence $\{n_i\}_{i \in \mathbb{N}}$ with constant $c$. If we show that for all $\varepsilon > 0$,

$$\sum_{i=1}^{\infty} \mathbb{P}\left\{|\overline{S}_{n_i} - \mathbb{E}\{X_1\}| > \varepsilon\right\} < \infty \quad (1)$$

then by Borel–Cantelli, applied with $E_i = \{|\overline{S}_{n_i} - \mathbb{E}\{X_1\}| > \varepsilon\}$, we see that

$$\mathbb{P}\left\{\limsup_{i \to \infty} \overline{S}_{n_i} \leq \mathbb{E}\{X_1\} + \varepsilon \right\} \cap \left\{\liminf_{i \to \infty} \overline{S}_{n_i} \geq \mathbb{E}\{X_1\} - \varepsilon\right\} = 1.$$
As this holds for any \( \varepsilon > 0 \), writing \( \varepsilon_m = 1/m \) we then have

\[
P \left\{ \lim_{i \to \infty} S_{ni} = \mathbb{E} \{ X_1 \} \right\}
\]

\[
= \mathbb{P} \left\{ \bigcap_{m=1}^{\infty} \left( \{ \limsup_{i \to \infty} S_{ni} \leq \mathbb{E} \{ X_1 \} + \varepsilon_m \} \cap \{ \liminf_{i \to \infty} S_{ni} \geq \mathbb{E} \{ X_1 \} - \varepsilon_m \} \right) \right\}
\]

\[
= \lim_{m \to \infty} \mathbb{P} \left\{ \{ \limsup_{i \to \infty} S_{ni} \leq \mathbb{E} \{ X_1 \} + \varepsilon_m \} \cap \{ \liminf_{i \to \infty} S_{ni} \geq \mathbb{E} \{ X_1 \} - \varepsilon_m \} \right\}
\]

\[
= \lim_{m \to \infty} 1 = 1.
\]

Thus, to prove the theorem it remains only to prove (1).

Now fix \( i_0 \) large enough that for all \( i \geq i_0 \), \( \mathbb{E} \{ X_1^{\leq ni} \} > \mathbb{E} \{ X_1 \} - \varepsilon \) and additionally \( n_{i+1} > cn_i \).

Note that it suffices to prove that

\[
\sum_{i=i_0}^{\infty} \mathbb{P} \left\{ |S_{ni} - \mathbb{E} \{ X_1 \} | > \varepsilon \right\} < \infty,
\]

(2)

since the difference between the sum in (1) and the latter sum is at most \( i_0 < \infty \). We will use truncation, truncating \( S_{ni} \) at \( n_i \). Writing \( S_{ni} = S_{ni}^{\leq ni} + S_{ni}^{> ni} \) in the obvious notation, for \( i \geq i_0 \) we then have

\[
P \left\{ |S_{ni} - \mathbb{E} \{ X_1 \} | > 2\varepsilon \right\} \leq P \left\{ S_{ni}^{\leq ni} - \mathbb{E} \{ X_1^{\leq ni} \} | > \varepsilon \right\} + P \left\{ S_{ni}^{> ni} \neq 0 \right\}
\]

\[
= P \left\{ S_{ni}^{\leq ni} - \mathbb{E} \{ X_1^{\leq ni} \} | > \varepsilon \right\} + P \left\{ S_{ni}^{> ni} \neq 0 \right\}.
\]

The inequality comes from the fact that for \( i \geq i_0 \), \( |\mathbb{E} \{ X_1^{\leq ni} \} - \mathbb{E} \{ X_1 \} | < \varepsilon \). By Chebyshev’s inequality for sums,

\[
P \left\{ S_{ni}^{\leq ni} - \mathbb{E} \{ X_1^{\leq ni} \} | > \varepsilon \right\} \leq \frac{1}{n_i \varepsilon^2} \mathbb{E} \left\{ (X_1^{\leq ni})^2 \right\}.
\]

By subadditivity of probabilities,

\[
P \left\{ S_{ni}^{> ni} \neq 0 \right\} \leq n_i P \{ X_1 > n_i \},
\]

so we will be done (we will have proved (2) if we can show that both of the sums

\[
\sum_{i=i_0}^{\infty} \frac{\mathbb{E} \left\{ (X_1^{\leq ni})^2 \right\}}{n_i} \quad \text{and} \quad \sum_{i=i_0}^{\infty} n_i P \{ X_1 > n_i \}
\]
are finite. For the first, let \( J = J(\omega) = \min \{ i : n_i \geq X_1(\omega) \} \), and write

\[
\sum_{i=i_0}^{\infty} \mathbb{E} \left\{ (X_1^{n_i})^{2} \right\} = \mathbb{E} \left\{ \sum_{i=i_0}^{\infty} \frac{X_i^{2}1_{X_i \leq n_i}}{n_i} \right\} \\
= \mathbb{E} \left\{ \sum_{i=\max(i_0,J)}^{\infty} \frac{X_i^{2}}{n_i} \right\} \quad \text{(all terms zero below } J) \\
\leq \mathbb{E} \left\{ \sum_{i=\max(i_0,J)}^{\infty} \frac{n_iX_1}{n_i} \right\} \quad \text{(} n_j \leq X_i \text{ for } i \geq J) \\
\leq \mathbb{E} \left\{ \sum_{i=\max(i_0,J)}^{\infty} e^{-\left(j-\max(i_0,J)\right)}X_1 \right\} \quad \text{(lacunarity)} \\
= \frac{c}{c-1} \mathbb{E} \left\{ X_1 \right\} < \infty.
\]

For the second, write

\[
\sum_{i=i_0}^{\infty} n_i 1^p \left\{ X_1 > n_i \right\} = \mathbb{E} \left\{ \sum_{i=i_0}^{\infty} n_i 1_{X_1 > n_i} \right\} \\
= \mathbb{E} \left\{ \sum_{i=i_0}^{J-1} n_i \right\} \quad \text{(} 1_{X_1 > n_i} = 0 \text{ for } i \geq J) \\
\leq \mathbb{E} \left\{ \sum_{i=i_0}^{J-1} e^{-\left(j-1-i\right)}X_1 \right\} \quad \text{(lacunarity)} \\
\leq \mathbb{E} \left\{ \sum_{j=0}^{\infty} c^{-j}X_1 \right\} \\
= \frac{c}{c-1} \mathbb{E} \left\{ X_1 \right\} < \infty.
\]

This completes the proof.