Exercise 21. *Let $\mathscr{A}_1, \dots, \mathscr{A}_n$ be π-systems of $\mathcal{F}$-sets such that*

$$P\Big(\bigcap_{k=1}^n A_k\Big) = \prod_{k=1}^n P(A_k) \tag{28}$$

for each choice of $A_k \in \mathscr{A}_k$ for $k = 1, \dots, n$. (a) Show by simple example that the $\mathscr{A}_k$'s need not be independent. (b) Show that the $\mathscr{A}_k$'s will be independent if for each k, Ω is the countable union of $\mathscr{A}_k$-sets. Hint: for fixed $A_2, \dots, A_n$ use the following general inclusion-exclusion formula to show that (28) with A_1 replaced by any finite union of $\mathscr{A}_1$-sets. Here is the general inclusion-exclusion formula:

$$P\Big(\bigcup_{i=1}^n B_i\Big) = \sum_{m=1}^n (-1)^{m-1} \sum_{i_1 < i_2 < \cdots < i_m} P(B_{i_1} \cap B_{i_2} \cap \cdots \cap B_{i_m})$$

for any $\mathcal{F}$-sets $B_1, B_2, \dots, B_n$.

Exercise 22. *(a) For each $k = 1, \dots, n$ let $\mathcal{P}_k$ be a partition of Ω into countably many $\mathcal{F}$-sets. Show that the σ-fields $\sigma\langle \mathcal{P}_1 \rangle, \dots, \sigma\langle \mathcal{P}_n \rangle$ are independent if and only if (28) holds for each choice of A_k from $\mathcal{P}_k$ for $k = 1, \dots, n$. (b) Use part (a) to show that $\mathcal{F}$-sets $A_1, \dots, A_n$ are independent if and only if*

$$P\Big(\bigcap_{k=1}^n B_k\Big) = \prod_{k=1}^n P(B_k)$$

for each choice of B_k as A_k or A_k^c for $k = 1, \dots, n$. (c) Use part (b) to show that the events $H_1, \dots, H_n$ in Theorem 44 are independent.

Exercise 23. *Let $A_1, A_2, \dots$ be $\mathcal{F}$-sets. Show that $P(A_n \text{ i.o.}_n) = 1$ if and only if $\sum_{n=1}^\infty P(A_n|A)$ diverges for every $\mathcal{F}$-set A of nonzero probability. Hint: show $P(A_n \text{ i.o.}_n) < 1 \iff \sum_{n=1}^\infty P(A_n|A) < \infty$ for some $\mathcal{F}$-set A with $P(A) > 0$.*

Exercise 24. *Let P and Q be probability measures on a σ-field $\mathcal{F}$ of subsets of a sample space Ω.*

- *P and Q are said to be* **singular**, *denoted $P \perp Q$, if and only if there exists a set $F \in \mathcal{F}$ such that*

$$P(F^c) = 0 = Q(F).$$

- *P is said to be* **absolutely continuous with respect to** *Q, denoted $P \ll Q$, if and only if*

$$P(F) = 0 \text{ for every } \mathcal{F}\text{-set } F \text{ for which } Q(F) = 0.$$

Show that

$$P \perp Q \iff \left[\begin{array}{c} \text{there exists } \mathcal{F}\text{-sets } F_1, F_2, \dots \text{ such that} \\ P(F_n^c) \to 0 \text{ and } Q(F_n) \to 0 \text{ as } n \to \infty \end{array}\right]$$

and

$$P \ll Q \iff \lim_{\delta \downarrow 0}\Big(\sup\{P(F) \colon F \in \mathcal{F} \text{ with } Q(F) \le \delta\}\Big) = 0.$$

4.2 Application: law of the iterated logarithm for coin flips

Section Assumption. *Throughout this section let $P : \mathcal{B}^{(0,1]} \to [0,1]$ be the probability model developed in Section 1 (and extended from the Carathéodory) for a uniform random number in $w \in (0,1]$. Also let s_n be defined as in Section 1.*

To motivate the law of the iterated logarithm lets start by discussing the difference between the weak law and strong law of large numbers.

$$\textbf{Weak law}: \quad \lim_{n\to\infty} P\Big(\big|\frac{s_n}{n}\big| < \epsilon\Big) = 1, \text{ for all } \epsilon > 0;$$
$$\textbf{Strong law}: \quad P\Big(\lim_{n\to\infty} \frac{s_n}{n} = 0\Big) = 1.$$

In some sense the difference can be explained as follows. The weak law fixes each n then analyzes the ensemble of $s_n(\omega)/n$ over $\omega \in (0,1]$. In particular, the weak law says that for large n it becomes increasingly rare to find ω's which satisfy $|s_n(\omega)/n| \ge \epsilon$. Conversely the strong law fixes each $\omega \in (0,1]$ and analyzes the ensemble of $s_n(\omega)/n$ over n. In particular for almost all ω, $s_n(\omega)/n \to 0$ as $n \to \infty$.

Now lets make a similar analogy with the central limit theorem and the law of the iterated logarithm. From the strong law we know that $s_n(\omega)/n$ converges to 0 for nearly every ω. We can then ask: at what rate? In particular, can we find a smaller denominator than n, call it ℓ_n, so that s_n/ℓ_n doesn't converge to zero. An answer is given by the central limit theorem

$$\textbf{CLT}: \quad \lim_{n\to\infty} P\Big(\frac{s_n}{\sqrt{n}} < x\Big) = \Phi(x)$$

where $\Phi(x) = P(Z \le x)$ and Z is a standard normal random variable. Notice two things. First, this suggests that the $s_n(\omega)/\sqrt{n}$ reaches up to ∞ and down to $-\infty$ for different values of ω and n. In particular for every cut-off M, there exists n large enough so that $P\Big(s_n/\sqrt{n} \ge M\Big) \approx 1 - \Phi(M) > 0$ to arbitrary precision. Also, the central limit theorem is similar to the weak law in that it fixes each n then analyzes the ensemble of $s_n(\omega)/\sqrt{n}$ over $\omega \in (0,1]$. The question then becomes, can one find an analogous form of the strong law such that for each fixed ω one analyzes the ensemble rate of $s_n(\omega)$ as $n \to \infty$. The law of the iterated logarithm gives the right rate

$$\textbf{LIL}: \quad P\Big(\limsup_{n\to\infty} \frac{s_n}{\sqrt{2n\log\log n}} = 1\Big) = 1.$$

Another way to think about the rate $\sqrt{2n\log\log n}$ is the effect due to the correlation of between s_n across different values of n. The expected maximum of n independent standard Gaussian random variables behaves as $\sqrt{2\log n}$. That maximum occurs uniformly on $\{1, 2, \dots, n\}$ and is therefore is expected to occur at index $n/2$. If $s_n/\sqrt{n}$ was not correlated across n, the central limit theorem might suggest that the maximum of $s_k/\sqrt{k}$ over $k \in \{1, 2 \dots, n\}$ behaves on the order of $\sqrt{2\log n}$. This loosely suggests the maximum of s_k over $k \in \{1, 2 \dots, n\}$ behaves on the order $\sqrt{n \log n}$. Now, in some sense, the LIIL says that the correlation across n will dampen the maximum excursions to be at most $\sqrt{n \log \log n}$.

Lemma 1 (Half of large deviation result). *For all $n \in \mathbb{N}$ and $x > 0$ one has*

$$P\big(s_n/\sqrt{n} \geq x\big) \leq \exp\Big(-\frac{x^2}{2}\Big)$$

Proof. This was established in exercise 2. $\square$

The other half of the large deviation result we need is Lemma 2, below. Combined these two lemmas give us good approximations to $P\big(s_n/\sqrt{n} \geq x_n\big)$ for large-ish values of x_n: large compared to 0 but still small compared to $\sqrt{n}$ (if x_n was larger then $\sqrt{n}$ then $P(s_n/\sqrt{n} > x_n) = 0$). This is the key for deriving the Law of the Iterated Logarithm. Also note that Lemma 1 was proved as an exercise but it is typically established using Markov's inequality, the moment generating function and the fact that $\frac{e^x + e^{-x}}{2} \leq \exp\big(x^2/2\big)$.

Lemma 2 (Other half of large deviation result). *If the sequence $\{x_n\}_{n\in\mathbb{N}}$ satisfies $0 \leq x_n \to \infty$ and $x_n/\sqrt{n} \to 0$ as $n \to \infty$ then*

$$P\big(s_n/\sqrt{n} \geq x_n\big) \geq \exp\Big(-\frac{x_n^2}{2}(1 + o(1))\Big).$$

$d_k \equiv X_k$

Proof. The general idea is to use the fact that $s_n = 2\big(\sum_{k=1}^n d_k\big) - n$ where $\sum_{k=1}^n d_k \sim \text{Bin}(n, 1/2)$. Therefore we can write $P(s_n \geq \sqrt{n}x_n)$ as a sum $\sum_{i\in\mathcal{I}_n} P(s_n = i)$ where i is the set of integers greater than $\sqrt{n}x_n$ and less than or equal to n. In fact, since we are trying to construct a lower bound we are free to discard terms in $\mathcal{I}_n$ which will give

$$P(s_n \geq \sqrt{n}x_n) \geq \sum_{i\in\mathcal{I}_n} P(s_n = i).$$

The main problem is how to find $\mathcal{I}_n$ so that the right hand side is $\exp\big(-\frac{x_n^2}{2}(1 + o(1))\big)$.

Lets start by getting some idea of how many integers we should include in $\mathcal{I}_n$ by analysing how $P(s_n = i)$ behaves when $i \approx \sqrt{n}x_n$. To make things a bit more precise let i_n be a sequence of integers depending on n such that $i_n \to \infty$ but $i_n/n \to 0$.

$$\begin{aligned}
P(s_n = i_n) &= P\Big(\sum_{k=1}^n d_k = (i_n + n)/2\Big) \\
&= \binom{n}{(i_n+n)/2}\frac{1}{2^n}, \quad \text{if } (i_n + n)/2 \text{ is an integer} \\
&= \frac{n!}{\frac{i_n+n}{2}!\,\frac{n-i_n}{2}!}\frac{1}{2^n} \\
&= \frac{2}{\sqrt{2\pi n}}[1 + o(1)]\exp\Big(-\frac{(1+o(1))i_n^2}{2n}\Big) \qquad (29)
\end{aligned}$$

To see why (29) is true notice that by Stirling's formula we have $n! = (1 + o(1))\sqrt{2\pi n}n^n e^{-n}$. Therefore

$$\begin{aligned}
\frac{n!}{\frac{i_n+n}{2}!\,\frac{n-i_n}{2}!}\frac{1}{2^n} &= \frac{(1+o(1))}{(1+o(1))^2} \\
&\quad\times \frac{\sqrt{2\pi n}}{\sqrt{2\pi(n+i_n)/2}\sqrt{2\pi(n-i_n)/2}} \\
&\quad\times \frac{n^n}{(n+i_n)^{(n+i_n)/2}(n-i_n)^{(n-i_n)/2}} \\
&\quad\times \frac{e^{-n}}{e^{-(n+i_n)/2}e^{-(n-i_n)/2}} \\
&= \frac{(1+o(1))}{(1+o(1))^2} \\
&\quad\times \underbrace{\frac{1}{\sqrt{2\pi(n+i_n)(n-i_n)/(4n)}}}_{=:I} \\
&\quad\times \underbrace{\Big(\frac{n}{n+i_n}\Big)^{(n+i_n)/2}\Big(\frac{n}{n-i_n}\Big)^{(n-i_n)/2}}_{=:II}
\end{aligned}$$

Notice

$$\begin{aligned}
I &= \frac{1}{\sqrt{2\pi(n+i_n)(n-i_n)/(4n)}} \\
&= \frac{2}{\sqrt{2\pi n}} \times \frac{1}{\sqrt{(n+i_n)(n-i_n)/(n^2)}} \\
&= \frac{2}{\sqrt{2\pi n}} \times \frac{1}{\sqrt{1-(i_n/n)^2}} = \frac{2}{\sqrt{2\pi n}}(1+o(1)).
\end{aligned}$$

Secondly notice that $(1 + x)\log(1 + x) = x + \frac{1}{2}x^2 + O(x^3)$ as $x \to 0$. Therefore

$$\begin{aligned}
\log II &= -\frac{1}{2}\big[(n+i_n)\log\big(1+\frac{i_n}{n}\big) + (n-i_n)\log\big(1-\frac{i_n}{n}\big)\big] \\
&= -\frac{n}{2}\big[\frac{i_n}{n} + \frac{1}{2}\frac{i_n^2}{n^2} - \frac{i_n}{n} + \frac{1}{2}\frac{i_n^2}{n^2} + O(i_n^3/n^3)\big] \\
&= -\frac{n}{2}\big[\frac{i_n^2}{n^2} + O(i_n^3/n^3)\big] = -\frac{i_n^2}{2n}\big[1 + \underbrace{O(i_n/n)}_{o(1)}\big].
\end{aligned}$$

To finish notice that $\frac{(1+o(1))^2}{(1+o(1))^2} = [1 + o(1)]$ which implies (29).

Now, looking at (29) it is clear that we want $\mathcal{I}_n$ to contain about $\sqrt{2\pi n}$ terms that are near $\sqrt{n}x_n$ (so that we can apply (29)). In particular $\mathcal{I}_n$ denote the set of indices between $\sqrt{n}x_n$ and $\sqrt{n}x_n + 2\sqrt{\pi n}$ which have the same parity at n (i.e. that $(i_n + n)/2$ is an integer). Also let i_n be the maximum integer in $\mathcal{I}_n$, which implies $i_n = \sqrt{n}x_n + 2\sqrt{\pi n} + O(1)$. Now

$$\begin{aligned}
&P(s_n \geq \sqrt{n}x_n) \\
&\quad\geq \sum_{i\in\mathcal{I}_n} P(s_n = i) \\
&\quad\geq [\#\mathcal{I}_n]P(s_n = i_n), \text{ since } \binom{n}{(i+n)/2} \geq \binom{n}{(i_n+n)/2}
\end{aligned}$$

$[\sqrt{\pi n} + O(1)]$

$\times \left[\frac{2}{\sqrt{2\pi n}} (1 + o(1)) \right]$

$= \left[O\left(\frac{1}{\sqrt{n}}\right) + \sqrt{2}\,(1 + o(1)) \right]$

≥ 1 for large n

$$
\begin{aligned}
&= [\sqrt{\pi n} + O(1)]P(s_n = i_n) \\
&\geq \sqrt{2}\exp\left(-\frac{(1+o(1))i_n^2}{2n}\right), \quad \text{by (29)} \\
&\geq \exp\left(-\frac{(1+o(1))i_n^2}{2n}\right) \\
&= \exp\left(-\frac{(1+o(1))\big[\sqrt{n}x_n + 2\sqrt{\pi n} + O(1)\big]^2}{2n}\right) \\
&= \exp\left(-\frac{x_n^2}{2}(1+o(1))\right), \quad \text{since } x_n \to \infty.
\end{aligned}
$$

□

Lemma 3 (Maximal inequality). *For all $n \in \mathbb{N}$ and every nonnegative integer c*

$$P\big(\max_{1\leq j\leq n} s_j \geq c\big) \leq 2P(s_n \geq c).$$

Proof. First write

$$
\begin{aligned}
P\big(\max_{1\leq j\leq n} s_j \geq c\big) &= P\big(\max_{1\leq j\leq n} s_j \geq c,\, s_n \geq c\big) \\
&\quad + P\big(\max_{1\leq j\leq n} s_j \geq c,\, s_n < c\big) \\
&= P\big(s_n \geq c\big) + P\big(\max_{1\leq j\leq n} s_j \geq c,\, s_n < c\big).
\end{aligned}
$$

Therefore all that remains is to show $P\big(\max_{1\leq j\leq n} s_j \geq c,\, s_n < c\big) \leq P\big(s_n \geq c\big)$. Start by segmenting the event $\{\max_{1\leq j\leq n} s_j \geq c\}$ corresponding to the first indicie j for which $s_j = c$ (this must occur when $\max_{1\leq j\leq n} s_j \geq c$ since s_n goes up or down with jumps of size 1 and c is a nonnegative integer). In particular, write

$$\{\max_{1\leq j\leq n} s_j \geq c\} = \bigcup_{j=1}^n \underbrace{\{s_1 < c, \ldots, s_{j-1} < c, s_j = c\}}_{=:F_j}$$

Now

$$
\begin{aligned}
&\{\max_{1\leq j\leq n} s_j \geq c\} \cap \{s_n < c\} \\
&\quad = \bigcup_{j=1}^n F_j \cap \{s_n < c\} \\
&\quad = \bigcup_{j=1}^n \underbrace{F_j \cap \{s_n - s_j < 0\}}_{\text{disjoint since the } F_j\text{'s are}}, \quad \text{since } \omega \in F_j \text{ implies } s_j(\omega) = c.
\end{aligned}
$$

Now notice two things. First $P(s_n - s_j < 0) = P(s_n - s_j > 0)$ by symmetry. Secondly, since $\{s_n - s_j < 0\} \in \sigma\langle z_{j+1}, \ldots, z_n\rangle$ and $F_j \in \sigma\langle z_1, \ldots, z_j\rangle$, the event $\{s_n - s_j < 0\}$ is independent of F_j (by ANOVA). Therefore

$$
\begin{aligned}
&P\big(\max_{1\leq j\leq n} s_j \geq c,\, s_n < c\big) \\
&\quad = \sum_{j=1}^n P\big(F_j \cap \{s_n - s_j < 0\}\big) \\
&\quad = \sum_{j=1}^n P\big(F_j\big)P\big(s_n - s_j < 0\big) \text{ by independence} \\
&\quad = \sum_{j=1}^n P\big(F_j\big)P\big(s_n - s_j > 0\big) \text{ by symmetry} \\
&\quad = \sum_{j=1}^n P\big(F_j \cap \{s_n - s_j > 0\}\big) \text{ by independence again} \\
&\quad = \sum_{j=1}^n P\big(F_j \cap \{s_n > c\}\big) \text{ since } \omega \in F_j \text{ implies } s_j(\omega) = c \\
&\quad = P\big(\max_{1\leq j\leq n} s_j \geq c,\, s_n > c\big) \\
&\quad \leq P\big(s_n > c\big) \\
&\quad \leq P\big(s_n \geq c\big).
\end{aligned}
$$

Therefore $P\big(\max_{1\leq j\leq n} s_j \geq c\big) \leq 2P\big(s_n \geq c\big)$. □

Theorem 49 (Law of the iterated logarithm for coin flips).

1. $P\Big[\limsup_{n\to\infty} \dfrac{s_n}{\sqrt{2n\log\log n}} = 1\Big] = 1$;
2. $P\Big[\liminf_{n\to\infty} \dfrac{s_n}{\sqrt{2n\log\log n}} = -1\Big] = 1$.

Proof. To make the formulas more readable let $\ell_n := \sqrt{2n\log\log n}$. Notice first that

$$
\begin{aligned}
&\{\limsup_n \frac{s_n}{\ell_n} = 1\} \\
&\quad = \bigcap_{\epsilon\in(0,1)\cap\mathbb{Q}} \{s_n/\ell_n > (1-\epsilon) \text{ i.o.}_n\} \cap \{s_n/\ell_n < (1+\epsilon) \text{ a.a.}_n\}.
\end{aligned}
$$

This implies that $\{\limsup_n (s_n/\ell_n) = 1\}$ and $\{\liminf_n (s_n/\ell_n) = -1\}$ are Borel measurable. Secondly notice that by symmetry we have

$$
\begin{aligned}
P\big[\limsup_{n\to\infty}(s_n/\ell_n) = 1\big] &= P\big[\liminf_{n\to\infty}(-s_n/\ell_n) = -1\big] \\
&= P\big[\liminf_{n\to\infty}(s_n/\ell_n) = -1\big].
\end{aligned}
$$

We can also simply our proof by using countable sub-additivity

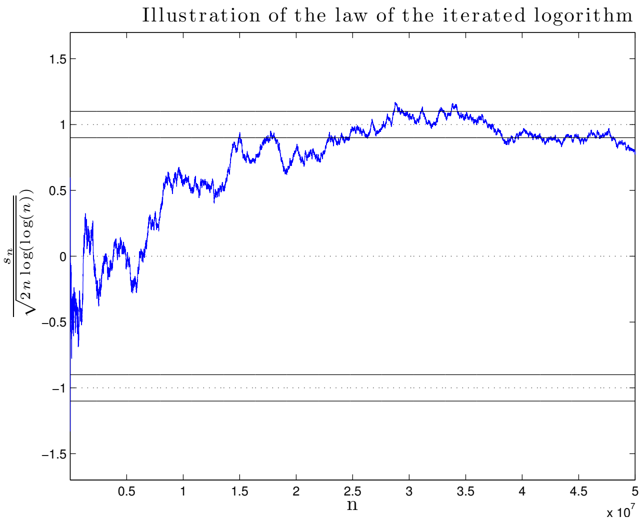
$$
\begin{aligned}
P\Big[\big\{\limsup_{n\to\infty}(s_n/\ell_n) = 1\big\}^c\Big] &\leq \sum_{\epsilon\in(0,1)\cap\mathbb{Q}} P[\{s_n/\ell_n \leq (1-\epsilon) \text{ a.a.}_n\}] \\
&\quad + \sum_{\epsilon\in(0,1)\cap\mathbb{Q}} P[\{s_n/\ell_n \geq (1+\epsilon) \text{ i.o.}_n\}].
\end{aligned}
$$

Therefore the proof will follow by establishing Lemmas 4 and 5 which state that for all $\epsilon > 0$

$$
\begin{aligned}
P[s_n/\ell_n \geq (1+\epsilon) \text{ i.o.}_n] &= 0 \\
P[s_n/\ell_n > (1-\epsilon) \text{ i.o.}_n] &= 1
\end{aligned}
$$

□

Theorem 49 is interesting for a number of reasons. First, it gives a very detailed analysis of the Central Limit Theorem. Second, it shows the power of the Borel-Cantelli lemmas. Third, it is one of those theorems in probability which is extremely hard to see in simulations: $\sqrt{2n\log\log n}$ grows to slow for modern computers to probe. The following simulation is an attempt at illustrating Theorem 186 but we still don't see the required fluctuation from $+1$ to -1.



Lemma 4. *For all $\epsilon > 0$*

$$P[s_n/\ell_n \geq (1+\epsilon) \text{ i.o.}_n] = 0$$

Proof. The obvious strategy is to use the first Borel-Cantelli lemma. In particular, it would be nice if we could show

$$\sum_{n=1}^{\infty} P[s_n/\ell_n \geq (1+\epsilon)] < \infty$$

which would give us the lemma. By the first half of the large deviation result we know $P[s_n/\ell_n \geq (1+\epsilon)]$ converge to zero as $n \to \infty$. Unfortunately, they do not converge fast enough for the first Borel-Cantelli. Taking a sub-sequence will get something summable but we would need to show that the events are well behaved between the sub-sequence. Fortunatly we can do this since the sets $\{s_n/\ell_n \geq (1+\epsilon)\}$ overlap a lot. The strategy is to group the events $\{s_n/\ell_n \geq (1+\epsilon)\}$ by unioning them into blocks, then apply Borel-Cantelli on the blocks. This will be sufficient since the blocks occur infinity often if and only if the events $s_n/\ell_n \geq (1+\epsilon)$ occur infinity often. Controlling the probability of the blocks is done with the maximal inequality.

Let the k^{th} block be defined

$$B_k := \bigcup_{j=n_{k-1}}^{n_k} \{s_j/\ell_j \geq (1+\epsilon)\}$$

where n_k is a (yet to be determined) subsequence. We use maximal inequality to bound $P[B_k]$ as follows

$$\begin{aligned}
P[B_k] &\leq P[\max_{n_{k-1}\leq j\leq n_k} s_j \geq (1+\epsilon) \min_{n_{k-1}\leq j \leq n_k} \ell_j] \\
&\leq P[\max_{n_{k-1}\leq j\leq n_k} s_j \geq (1+\epsilon)\ell_{n_{k-1}}] \\
&\leq P[\max_{j\leq n_k} s_j \geq (1+\epsilon)\ell_{n_{k-1}}] \\
&\leq P[\max_{j\leq n_k} s_j \geq \lceil (1+\epsilon)\ell_{n_{k-1}}\rceil], \quad \text{since } s_j \in \mathbb{Z} \\
&\leq 2P[s_{n_k} \geq \lceil (1+\epsilon)\ell_{n_{k-1}}\rceil], \quad \text{maximal ineq} \\
&\leq 2P[s_{n_k} \geq (1+\epsilon)\ell_{n_{k-1}}] \\
&\leq \exp\big(-\tfrac{1}{2}(1+\epsilon)^2\ell^2_{n_{k-1}}/n_k\big), \quad \text{half of large deviation} \\
&\leq \exp\big(-(1+\epsilon)^2 n_{k-1}\log\log n_{k-1}/n_k\big) \\
&= \left(\frac{1}{\log n_{k-1}}\right)^{(1+\epsilon)^2\frac{n_{k-1}}{n_k}}.
\end{aligned}$$

Now we just find n_k which makes the last term summable over k. If $n_k \approx \theta^k$ one gets

$$\left(\frac{1}{\log n_{k-1}}\right)^{(1+\epsilon)^2\frac{n_{k-1}}{n_k}} \approx \left(\frac{1}{(k-1)\log\theta}\right)^{(1+\epsilon)^2\frac{1}{\theta}}$$

which is summable if $(1+\epsilon)^2\frac{1}{\theta} > 1$. We also need that $\theta > 1$ since we need $n_k \to \infty$ as $k \to \infty$. Luckily there does exist such a θ for which $(1+\epsilon)^2 > \theta > 1$. $\square$

Lemma 5. *For all $\epsilon > 0$*

$$P[s_n/\ell_n > (1-\epsilon) \text{ i.o.}_n] = 1$$

Proof. In the previous lemma we presented a technique to adjust the first Borel-Cantelli lemma in the case the summability condition doesn't hold. For this lemma we want to use the second Borel-Cantelli lemma but, again, it doesn't directly apply since the condition that the events $s_n/\ell_n > (1-\epsilon)$ are independent does not hold. Here is a generic technique to get around this obstacle. Find subsequence n_k and subsets

$$I_k \subset \{s_{n_k}/\ell_{n_k} > (1-\epsilon)\}$$

such that I_k's are independent and $\sum_k P[I_k] = \infty$ so that $P[I_k \text{ i.o.}_k] = 1$ (which would then give the lemma). Unfortunately, even this doesn't work. What ends up working is to find two sets A_k I_k such that

$$A_k \cap I_k \subset \{s_{n_k}/\ell_{n_k} > (1-\epsilon)\} \tag{30}$$
$$I_k \text{ are independent and } \textstyle\sum_k P[I_k] = \infty \tag{31}$$
$$A_k \text{ are not independent but } P[A_k \text{ a.a.}_k] = 1. \tag{32}$$

To see why this is sufficient notice that (31) implies $P[I_k \text{ i.o.}_k] = 1$ by the second Borel-Cantelli lemma. Then

$$\begin{aligned}
&P[A_k \text{ a.a.}_k] = 1 \text{ and } P[I_k \text{ i.o.}_k] = 1 \\
&\implies P[A_k \cap I_k \text{ i.o.}_k] = 1 \\
&\overset{(30)}{\implies} P[s_{n_k}/\ell_{n_k} > (1-\epsilon) \text{ i.o.}_k] = 1
\end{aligned}$$

Therefore (30), (31) and (32) are sufficient to establish the lemma.

since $\{A_k \text{ a.e}\} \cap \{I_k \text{ i.o.}\} \subset \{A_k \cap I_k \text{ i.o}\}$

$\therefore P(\{A_k \cap I_k \text{ i.o.}\}^c) \leq \underbrace{P(\{A_k \text{ a.a.}\}^c) + P(\{I_k \text{ i.o.}\}^c)}_{=0}$

Figuring out how to define I_k and A_k are the tricky parts. The intuition is that if your going to get independent events you need to look at increments of s_n. Define

$$I_k := \{s_{n_k} - s_{n_{k-1}} \geq (1-\epsilon/2)\ell_{n_k}\}$$
$$A_k := \{s_{n_{k-1}} > -(\epsilon/2)\ell_{n_k}\}.$$

Clearly $I_k \cap A_k \subset \{s_{n_k}/\ell_{n_k} > (1-\epsilon)\}$ so that (30) holds. Moreover, the I_k's are independent. To show (32)

$$\begin{aligned} P[A_k \text{ a.a.}_k] &= P\Big[s_{n_{k-1}} > -(\epsilon/2)\ell_{n_k} \text{ a.a.}_k\Big] \\ &= P\Big[s_{n_{k-1}} < (\epsilon/2)\ell_{n_k} \text{ a.a.}_k\Big], \quad \text{by symmetry} \\ &= P\Big[\frac{s_{n_{k-1}}}{\ell_{n_{k-1}}} < (\epsilon/2)\frac{\ell_{n_k}}{\ell_{n_{k-1}}} \text{ a.a.}_k\Big] \\ &= 1 - P\Big[\frac{s_{n_{k-1}}}{\ell_{n_{k-1}}} \geq (\epsilon/2)\frac{\ell_{n_k}}{\ell_{n_{k-1}}} \text{ i.o.}_k\Big] \\ &= 1, \text{ by Lemma 4 if } \frac{\ell_{n_k}}{\ell_{n_{k-1}}} \to \infty. \end{aligned}$$

To show (31) notice

$$\begin{aligned} P[I_k] &= P\Big[s_{n_k} - s_{n_{k-1}} \geq (1-\epsilon/2)\ell_{n_k}\Big] \\ &= P\Big[s_{n_k - n_{k-1}} \geq (1-\epsilon/2)\ell_{n_k}\Big] \\ &\geq P\Big[\frac{s_{n_k - n_{k-1}}}{\sqrt{n_k - n_{k-1}}} \geq \frac{(1-\epsilon/2)\ell_{n_k}}{\sqrt{n_k - n_{k-1}}}\Big] \\ &\geq \exp\Big(-\frac{1}{2}\Big[\frac{(1-\epsilon/2)\ell_{n_k}}{\sqrt{n_k - n_{k-1}}}\Big]^2(1+o(1))\Big), \qquad (33) \end{aligned}$$

by Lemma 2 if:

$\frac{\ell_{n_k}}{\sqrt{n_k - n_{k-1}}} \to \infty$ and

$\frac{1}{\sqrt{n_k - n_{k-1}}}\frac{\ell_{n_k}}{\sqrt{n_k - n_{k-1}}} \to 0$ and

$n_k - n_{k-1} \to \infty$.

To finish the proof of (32) and (31) we need to find $n_k \to \infty$ such that ℓ_{n_k} satisfies the above conditions and the sum of (33) diverges.

A subsequence of the form $n_k := \lfloor \exp(k^\theta) \rfloor$ will work. To check the conditions notice

$$\begin{aligned} \frac{\ell_{n_k}}{\ell_{n_{k-1}}} &\sim \frac{\sqrt{2\exp(k^\theta)\log k^\theta}}{\sqrt{2\exp((k-1)^\theta)\log(k-1)^\theta}} \\ &= \exp\Big(\frac{k^\theta - (k-1)^\theta}{2}\Big)\frac{\log k}{\log(k-1)} \\ &= \exp\Big(\frac{\theta(k^*)^{\theta-1}}{2}\Big)(1+o(1)), \quad \text{where } k-1 \leq k^* \leq k \\ &\longrightarrow \infty, \quad \text{if } \theta > 1. \end{aligned}$$

Also

$$\begin{aligned} n_k - n_{k-1} &\sim \exp(k^\theta) - \exp((k-1)^\theta) \\ &= \theta(k^*)^{(\theta-1)}\exp((k^*)^\theta), \quad \text{where } k-1 \leq k^* \leq k \\ &\longrightarrow \infty, \quad \text{if } \theta > 0. \end{aligned}$$

And therefore

$$\begin{aligned} \frac{\ell_{n_k}}{\sqrt{n_k - n_{k-1}}} &\sim \frac{\sqrt{2\exp(k^\theta)\log k^\theta}}{\sqrt{\exp(k^\theta) - \exp((k-1)^\theta)}} \\ &= \frac{\sqrt{2\theta\log k}}{\sqrt{1 - \exp((k-1)^\theta - k^\theta)}} \\ &= \frac{\sqrt{2\theta\log k}}{\sqrt{1 - o(1)}} \\ &\longrightarrow \infty. \end{aligned}$$

Clearly we now have that $\frac{\ell_{n_k}}{n_k - n_{k-1}} \longrightarrow 0$. Finally we need to show that the sum of (33) over k diverges. The individual terms are

$$\begin{aligned} \exp\Big(-\frac{1}{2}\Big[\frac{(1-\epsilon/2)\ell_{n_k}}{\sqrt{n_k - n_{k-1}}}\Big]^2(1+o(1))\Big) \\ &\sim \exp\Big(-\frac{(1-\epsilon/2)^2}{2}2\theta\log k\Big) \\ &= \exp\Big(-(1-\epsilon/2)^2\theta\log k\Big) \\ &= k^{-(1-\epsilon/2)^2\theta} \end{aligned}$$

the sum of the above terms over k diverges if $(1-\epsilon/2)^\theta < 1$, i.e. if $\theta < \frac{1}{(1-\epsilon)^2}$. Now putting all the conditions on θ together says that we simply need to choose θ such that

$$1 < \theta < \frac{1}{(1-\epsilon/2)^2}.$$

$\square$