Steven R. Dunbar
Department of Mathematics
203 Avery Hall
University of Nebraska-Lincoln
Lincoln, NE 68588-0130
http://www.math.unl.edu
Voice: 402-472-3731
Fax: 402-472-8466

Topics in Probability Theory and Stochastic Processes Steven R. Dunbar

---

Law of the Iterated Logarithm

Rating

Mathematicians Only: prolonged scenes of intense rigor.

Section Starter Question

Key Concepts

1. The *Law of the Iterated Logarithm* tells very precisely how far the number of successes in a coin-tossing game will make excursions from the average value.

2. The Law of the Iterated Logarithm is a high-point among increasingly precise limit theorems characterizing how far the number of successes in a coin-tossing game will make excursions from the average value. The theorems start with the Strong Law of Large Numbers and the Central Limit Theorem, to Hausdorff's Estimate, and the Hardy-Littlewood Estimate leading to the Law of the Iterated Logarithm.

3. Khinchin's Law of the Iterated Logarithm says: Almost surely, for all $\epsilon > 0$, there exist infinitely many n such that

$$S_n - np > (1-\epsilon)\sqrt{n}\sqrt{2p(1-p)\ln(\ln(n))}$$

and furthermore, almost surely, for all $\epsilon > 0$, for every n larger than a threshold value N

$$S_n - np < (1+\epsilon)\sqrt{n}\sqrt{2p(1-p)\ln(\ln(n))}.$$

Vocabulary

1. The **limsup**, abbreviation for *limit superior* is a refined and generalized notion of limit, being the largest dependent-variable subsequence limit. That is, among all subsequences of independent-variable values tending to some independent-variable value, usually infinity, there will be a corresponding dependent-variable subsequence. Some of these dependent-variable sequences will have limits, and among all these, the largest is the limsup.

2. The **liminf**, abbreviation for *limit inferior* is analogous, it is the least of all dependent-variable subsequence limits.

3. Khinchin's **Law of the Iterated Logarithm** says: Almost surely, for all $\epsilon > 0$ there exist infinitely many n such that

$$S_n - np > (1-\epsilon)\sqrt{n}\sqrt{2p(1-p)\ln(\ln(n))}$$

and furthermore, almost surely, for all $\epsilon > 0$, for every n larger than a threshold value N

$$S_n - np < (1+\epsilon)\sqrt{n}\sqrt{2p(1-p)\ln(\ln(n))}.$$

Mathematical Ideas

Overview

We again consider the number of successes in a coin-tossing game. That is, we consider the sum S_n where the independent, identically distributed random variables in the sum $S_n = X_1 + \cdots + X_n$ are the Bernoulli random variables $X_i = +1$ with probability p and $X_i = 0$ with probability $q = 1 - p$. Note that the mean $\mu = p$ is and the variance is $\sigma^2 = p(1-p)$ for each of the summands X_i.

The *Strong Law of Large Numbers* says that

$$\lim_{n\to\infty} \frac{S_n - np}{n} = 0$$

with probability 1 in the sample space of all possible coin flips. This says the denominator n is "too strong", it "condenses out" all variation in the sum S_n.

The Central Limit Theorem applied to this sequence of coin flips says

$$\lim_{n\to\infty} \frac{S_n - np}{\sqrt{np(1-p)}} = Z$$

where $Z \equiv N(0,1)$ is a normal random variable and the limit is interpreted as convergence in distribution. In fact, this implies that for large n about 68% of the points in the sample space of all possible coin flips satisfy

$$\left| \frac{S_n - np}{\sqrt{np(1-p)}} \right| \le 1$$

and about 95% of the points in the sample space of all possible coin flips satisfy

$$\left| \frac{S_n - np}{\sqrt{np(1-p)}} \right| \le 2.$$

This says the denominator $\sqrt{n}$ is "too weak", it doesn't condense out enough information. In fact, using the Kolmogorov zero-one law and the Central Limit Theorem, almost surely

$$\liminf_{n\to\infty} \frac{S_n - np}{\sqrt{np(1-p)}} = -\infty$$

and almost surely

$$\limsup_{n\to\infty} \frac{S_n - np}{\sqrt{np(1-p)}} = +\infty.$$

The Strong Law and the Central Limit Theorem together suggest that "somewhere in between n and $\sqrt{n}$" we might be able to make stronger statements about convergence and the variation in the sequence S_n.

In fact, Hausdorff's estimate tells us:

$$\lim_{n\to\infty} \left| \frac{S_n - np}{n^{1/2+\epsilon}} \right| = 0$$

with probability 1 in the sample space of all possible coin flips for all values of $\epsilon > 0$. This says the denominator $n^{1/2+\epsilon}$ is "still too strong", it condenses out too much information.

Even better, Hardy and Littlewood's estimate tells us:

$$\lim_{n\to\infty} \left| \frac{S_n - np}{\sqrt{n \ln n}} \right| \le \text{constant}$$

with probability 1 in the sample space of all possible coin flips for all values of $\epsilon > 0$. In a way, this says $\sqrt{n \ln n)}$ is "still a little too strong", it condenses out most information.

Khinchin's Law of the Iterated Logarithm has a denominator that is "just right." It tells us very precisely how the deviations of the sums from the mean vary with n. Using a method due to Erdös, it is possible to refine the law even more, but for these notes a refinement is probably past the point of diminishing returns.

Like the Central Limit Theorem, the Law of the Iterated Logarithm illustrates in an astonishingly precise way that even completely random sequences obey precise mathematical laws.

Khinchin's Law of the Iterated Logarithm says that:

> Almost surely, for all $\epsilon > 0$, there exist infinitely many n such that
>
> $$S_n - np > (1-\epsilon)\sqrt{np(1-p)}\sqrt{2\ln(\ln(n))}$$
>
> and furthermore, almost surely, for all $\epsilon > 0$, for every n larger than a threshold value N depending on ϵ
>
> $$S_n - np < (1+\epsilon)\sqrt{np(1-p)}\sqrt{2\ln(\ln(n))}.$$

These appear in a slightly non-standard way, with the additional factor $\sqrt{2 \ln \ln n}$ times the standard deviation from the Central Limit Theorem to emphasize the similarity to and the difference from the Central Limit Theorem.

Theorem 1 (Law of the Iterated Logarithm)**.** *With probability* 1*:*

$$\limsup_{n\to\infty} \frac{S_n - np}{\sqrt{2np(1-p)\ln(\ln(n))}} = 1.$$

This means that with probability 1 *for any* $\epsilon > 0$*, only finitely many of the events:*

$$S_n - np > (1+\epsilon)\sqrt{n}\sqrt{2p(1-p)\ln(\ln(n))}$$

occur; on the other hand, with probability 1*,*

$$S_n - np > (1-\epsilon)\sqrt{n}\sqrt{2p(1-p)\ln(\ln(n))}$$

occurs for infinitely many n*.*

For reasons of symmetry, for $\epsilon > 0$, the inequality

$$S_n - np < -(1+\epsilon)\sqrt{n}\sqrt{2p(1-p)\ln(\ln(n))}$$

can only occur for finitely many n; while

$$S_n - np < -(1-\epsilon)\sqrt{n}\sqrt{2p(1-p)\ln(\ln(n))}$$

must occur for infinitely many n. That is,

$$\liminf_{n\to\infty} \frac{S_n - np}{\sqrt{2np(1-p)\ln(\ln(n))}} = -1$$

with probability 1.

Compare the Law of the Iterated Logarithm to the Central Limit Theorem. The Central Limit Theorem, says that $(S_n - np)/\sqrt{np(1-p)}$ is approximately distributed as a $N(0,1)$ random variable for large n. Therefore, for a large but fixed n, there is probability about $1/6$ that the values of $(S_n - np)/\sqrt{np(1-p)}$ can exceed the standard deviation 1, or $S_n - np > \sqrt{np(1-p)}$. For a fixed but large n, with probability about 0.025, $(S_n - np)/\sqrt{np(1-p)}$ can exceed twice the standard deviation 2, or $(S_n - np) > 2\sqrt{np(1-p)}$. The Law of the Iterated Logarithm tells us the more precise information that there are infinitely many n such that

$$S_n - np > (1-\epsilon)\sqrt{2np(1-p)\ln(\ln(n))}$$

for *any* $\epsilon > 0$. The Law of the Iterated Logarithm does *not* tell us how long we will have to wait between such repeated crossings however, and the wait can be very, very long indeed, although it must (with probability 1) eventually occur again. Moreover, the Law of the Iterated Logarithm tells us in addition that

$$S_n - np < -(1-\epsilon)\sqrt{2np(1-p)\ln(\ln(n))}$$

must occur for infinitely many n.

Khinchin's Law of the Iterated Logarithm also applies to the cumulative fortune in a coin-tossing game, or equivalently, the position in a random walk. Consider the independent Bernoulli random variables $Y_i = +1$ with probability p and $Y_i = -1$ with probability $q = 1-p$. The mean is $\mu = 2p-1$ and the variance is $\sigma^2 = 4p(1-p)$ for each of the summands Y_i. Then consider the sum $T_n = Y_1 + \cdots + Y_n$ with mean $(2p-1)n$ and variance $4np(1-p)$. Since $Y_n = 2X_n - 1$, then $T_n = 2S_n - n$ and $S_n = \frac{1}{2}T_n + \frac{n}{2}$. Then applying the Law of the Iterated Logarithm says that with probability 1 for any $\epsilon > 0$, only finitely many of the events:

$$|T_n - n(2p-1)| > (1+\epsilon)2\sqrt{n}\sqrt{2p(1-p)\ln(\ln(n))}$$

occur; on the other hand, with probability 1,

$$|T_n - n(2p-1)| > (1-\epsilon)2\sqrt{n}\sqrt{2p(1-p)\ln(\ln(n))}$$

occurs for infinitely many n. This means that the fortune must (with probability 1) oscillate back and forth across the net zero axis infinitely often, crossing the upper and lower boundaries:

$$\pm(1-\epsilon)2\sqrt{2p(1-p)n\ln(\ln(n))}$$

The statement puts some strength behind an understanding of the long-term swings backs and forth in value of a random process. It also implies a form of *recurrence*, that is, a random walk must visit every integer value.

The Law of the Iterated Logarithm for Bernoulli trials stated here is a special case of an even more general theorem first formulated by Kolmogorov in 1929. It is also possible to formulate even stronger and more general theorems! The proof here uses the Large Deviations and Moderate Deviations results with the Borel-Cantelli Lemmas. In another direction, the Law of the

Iterated Logarithm can be proved using invariance theorems, so it is distantly related to the Central Limit Theorem.

Figure 1 gives an impression of the growth of the function in the Law of the Iterated Logarithm compared to the square root function. Figure 2 gives an impression of the Law of the Iterated Logarithm by showing a piecewise linearly connected graph of 2000 steps of $S_n - np$ with $p = q = 1/2$. In this figure, the random walk must return again to cross the blue curves with $(1 - \epsilon) = 0.9$ infinitely many times, but may only cross the red curve with $1+\epsilon = 1.1$ finitely many times. Of course, this is only a schematic impression since a single random walk (possibly atypical, from the negligible set!) on the finite interval $0 \le n \le 2000$ can only suggest the almost sure infinitely many crossings of $(1 - \epsilon)\alpha(x)$ for any $\epsilon > 0$.

Figure 3 gives a comparison of impressions of four of the limit theorems. The individual figures deliberately are "spaghetti graphs" to give an impression of the ensemble of sample paths. Each figure shows a different scaling of the same 15 sample paths for a sequence of 100,000 fair coin flips, each path a different color. Note that the steps axis has a logarithmic scale, meaning that the shape of the paths is distorted although it still gives an impression of the random sums. The top left figure shows $S_n/n - p$ converging to 0 for all paths in accord with the Strong Law of Large Numbers. The top right figure plots the scaling $(S_n - np)/\sqrt{2p(1-p)n}$. For large values of steps the values over all paths is a distribution ranging from about -2 to 2, consistent with the Central Limit Theorem. The lower left figure plots $(S_n - np)/n^{0.6}$ as an illustration of Hausdorff's Estimate with $\epsilon = 0.1$. It appears that the scaled paths are very slowly converging to 0, the range for $n = 100{,}000$ is within $[-0.5, 0.5]$. The lower right figure shows $(S_n - np)/\sqrt{2p(1-p)n\ln(\ln(x))}$ along with lines ± 1 to suggest the conclusions of the Law of the Iterated Logarithm. It suggests that all paths are usually in the range $[-1, 1]$ but with each path making a few excursions outside the range.

Hausdorff's Estimate

Theorem 2 (Hausdorff's Estimate)**.** *Almost surely, for any $\epsilon > 0$,*

$$S_n - np = o\left(n^{\epsilon + 1/2}\right)$$

as $n \to \infty$.

Proof. The proof resembles the proof of the Strong Law of Large Numbers for independent random variables with mean 0 and uniformly bounded 4th

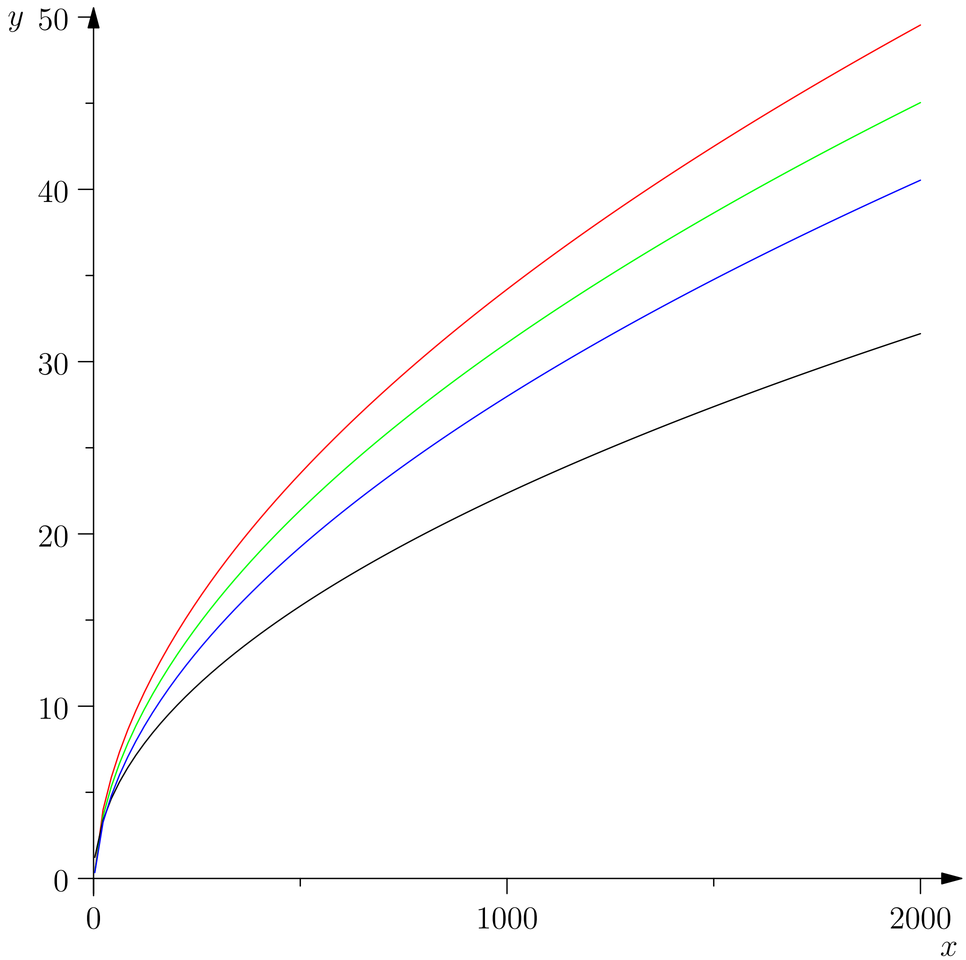


Figure 1: The Iterated Logarithm function $\alpha(x) = \sqrt{2p(1-p)x\ln(\ln(x))}$ in green along with functions $(1+\epsilon)\alpha(x)$ in red and $(1-\epsilon)\alpha(x)$ in blue, with $\epsilon = 0.1$. For comparison, the square root function $\sqrt{2p(1-p)x}$ is in black.

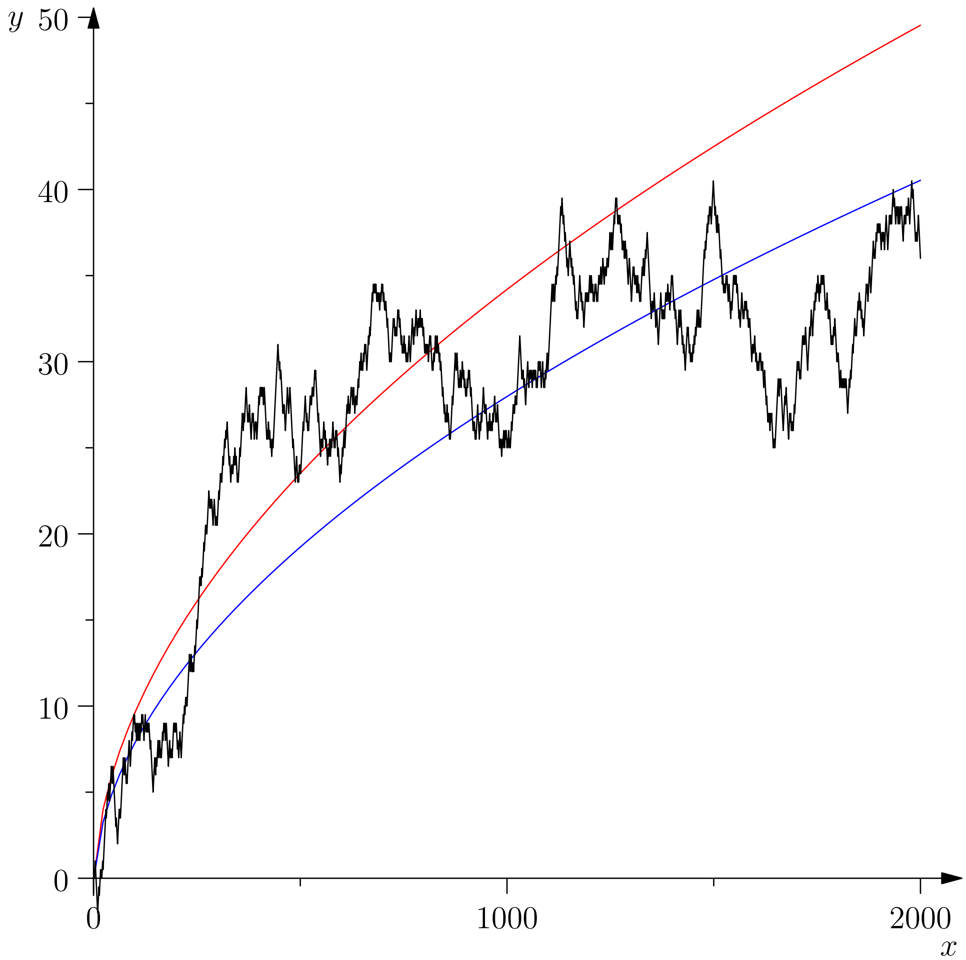


Figure 2: An impression of the Law of the Iterated Logarithm using a piecewise linearly connected graph of 2000 steps of $S_n - np$ with $p = q = 1/2$ with the blue curve with $(1-\epsilon)\alpha(x)$ and the red curve with $(1+\epsilon)\alpha(x)$ for $\epsilon = 0.1$ and $\alpha(x) = \sqrt{2p(1-p)x\ln(\ln(x))}$.

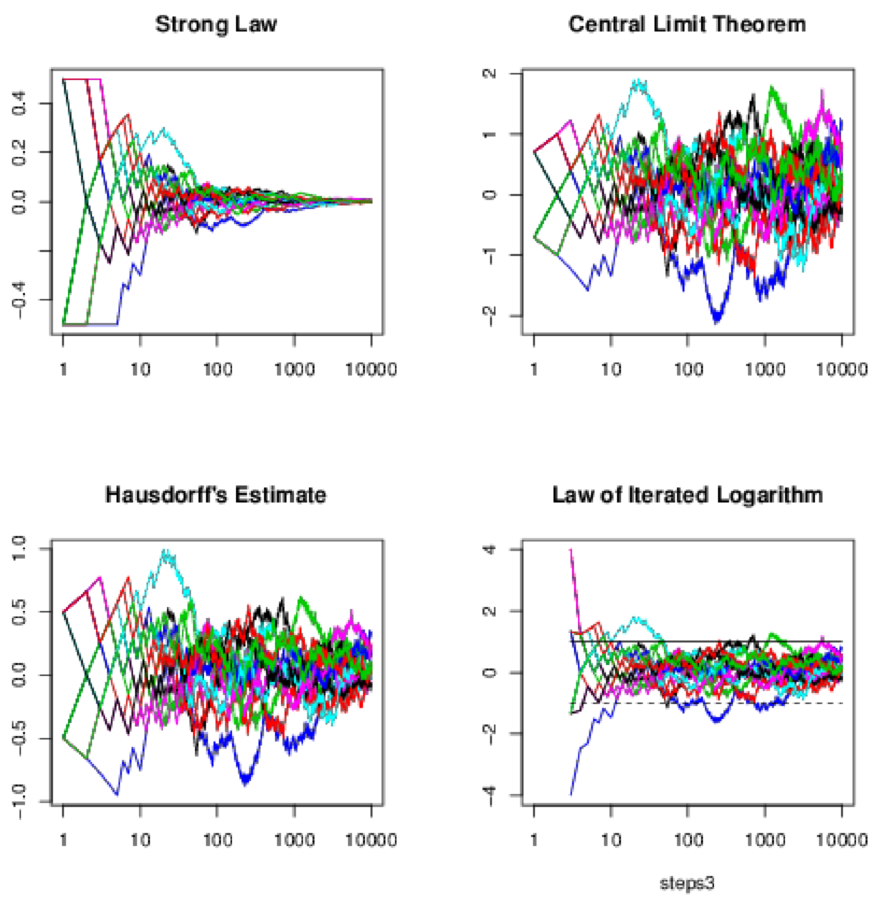


Figure 3: A comparison of four of the limit theorems. The individual figures deliberately are “spaghetti graphs” to give an impression of the ensemble of sample paths. Each figure shows a different scaling of the same 15 sample paths for a sequence of 100,000 fair coin flips, each path a different color. Note that the steps axis has a logarithmic scale, meaning that the shape of the sample paths is distorted.

moments. That proof showed that using the independence and identical distribution assumptions

$$\mathbb{E}\left[S_n^4\right] = n\mathbb{E}\left[X_1^4\right] + 3n(n-1)\mathbb{E}\left[X_1^2\right]^2 \leq nC + 3n^2\sigma^4 \leq C_1 n^2.$$

Then adapting the proof of the Markov and Chebyshev inequalities with fourth moments

$$\frac{\mathbb{E}\left[|S_n|^4\right]}{n^4} \leq \frac{C_1 n^2}{n^4}.$$

Use the Corollary that if $\sum_{n=1}^{\infty} \mathbb{E}\left[|X_n|\right]$ converges, then the sequence $(X_n)_{n\geq 1}$ converges almost surely to 0. By comparison, $\sum_{n=1}^{\infty} \frac{\mathbb{E}\left[|S_n|^4\right]}{n^4}$ converges so that $S_n/n \to 0$ a.s. Using the same set of ideas

$$\frac{\mathbb{E}\left[|S_n|^4\right]}{n^\alpha} \leq \frac{C_1 n^2}{n^{4\alpha}}$$

provided that $\alpha > 3/4$. Then using the same lemma $S_n/n^\alpha \to 0$ for $\alpha > 3/4$ for a simple version of Hausdorff's Estimate.

Now adapt this proof to higher moments. Let k be a fixed positive integer. Recall the definition $R_n(\omega) = S_n(\omega) - np = \sum_{k=1}^{n}(X_k - p) = \sum_{k=1}^{n}(X'_k$ where $X'_k = X_k - p$ and consider $\mathbb{E}\left[R_n^{2k}\right]$. Expanding the product R_n^{2k} results in a sum of products of the individual random variables X'_i of the form $X'_{i_1}X'_{i_2}\cdots X'_{i_{2k}}$. Each product $X'_{i_1}X'_{i_2}\cdots X'_{i_{2k}}$ results from a selection or mapping from indices $\{1, 2, \ldots, 2k\}$ to the set $\{1, \ldots, n\}$. Note that if an index $j \in \{1, 2, \ldots, n\}$ is selected only once so that X'_j appears only once in the product $X'_{i_1}X'_{i_2}\cdots X'_{i_k}$, then $\mathbb{E}\left[X'_{i_1}X'_{i_2}\cdots X'_{i_k}\right] = 0$ by independence. Further notice that for all sets of indices

$$\mathbb{E}\left[X'_{i_1}X'_{i_2}\cdots X'_{i_k}\right] \leq 1.$$

Thus

$$\mathbb{E}\left[R_n^{2k}\right] = \sum_{1\leq i_1,\ldots,i_{2k}\leq n} \mathbb{E}\left[X'_{i_1}X'_{i_2}\cdots X'_{i_{2k}}\right] \leq N(k,n),$$

where $N(k,n)$ is the number of functions from $\{1, \ldots, 2k\}$ to $\{1, \ldots, n\}$ that take each value at least twice. Let $M(k)$ be the number of partitions of $\{1, \ldots, 2k\}$ into subsets each containing at least two elements. If P is such a partition, then P has at most k elements. The number of functions $N(k,n)$

that are constant on each element of P is at most n^k. Thus, $N(k,n) \leq n^k M(k)$.

Now let $\epsilon > 0$ and consider

$$\mathbb{E}\left[\left(n^{-\epsilon-1/2}R_n\right)^{2k}\right] \leq n^{-2k\epsilon-k}N(k,n) \leq n^{-2k\epsilon}M(k).$$

Choose $k > \frac{1}{2\epsilon}$. Then

$$\sum_{n\geq 1}\mathbb{E}\left[\left(n^{-\epsilon-1/2}R_n\right)^{2k}\right] < \infty.$$

Recall the Corollary 2 appearing in the section on the Borel-Cantelli Lemma:

> Let $(X_n)_{n\geq 0}$ be a sequence of random variables. If $\sum_{n=1}^{\infty}\mathbb{E}\left[|X_n|\right]$ converges, then X_n converges to zero, almost surely.

By this corollary, the sequence of random variables

$$\left(n^{-\epsilon-1/2}R_n\right) \to 0$$

almost surely as $n \to \infty$.

This means that for each $\epsilon > 0$, there is a negligible event (depending on ϵ) outside of which $n^{-\epsilon-1/2}R_n$ converges to 0. Now consider a countable set of values of ϵ tending to 0. Since a countable union of negligible events is negligible, then for each $\epsilon > 0$, $n^{-\epsilon-1/2}R_n$ converges to 0 almost surely. □

Hardy-Littlewood Estimate

Theorem 3 (Hardy-Littlewood Estimate)**.**

$$S_n - np = O\left(\sqrt{n\ln n}\right) \;\; \textit{a.s. for } n \to \infty.$$

Remark. The proof shows that $S_n - np \leq \sqrt{n\ln n}$ a.s. for $n \to \infty$.

Proof. The proof uses the Large Deviations Estimate as well as the Borel-Cantelli Lemma. Recall the Large Deviations Estimate says

$$\mathbb{P}\left[\frac{S_n}{n} \geq p + \epsilon\right] \leq \mathrm{e}^{-nh_+(\epsilon)},$$

where

$$h_+(\epsilon) = (p+\epsilon)\ln\left(\frac{p+\epsilon}{p}\right) + (1-p-\epsilon)\ln\left(\frac{1-p-\epsilon}{1-p}\right).$$

Note that as $\epsilon \to 0$, $h_+(\epsilon) = \frac{\epsilon^2}{2p(1-p)} + O(\epsilon^3)$.

Note that

$$\mathbb{P}\left[\frac{S_n}{n} \geq p+\epsilon\right] = \mathbb{P}\left[S_n - np \geq n\epsilon\right],$$

and take $\epsilon = \sqrt{\frac{\ln n}{n}}$ and note that

$$\mathbb{P}\left[S_n - np \geq \sqrt{n \ln n}\right] \leq e^{-nh_+\left(\sqrt{\frac{\ln n}{n}}\right)}.$$

Then

$$h_+\left(\sqrt{\frac{\ln n}{n}}\right) = \frac{\ln n}{2p(1-p)n} + o\left(\frac{1}{n}\right),$$

since $O\left(\left(\frac{\ln n}{n}\right)^{3/2}\right) = o\left(\frac{1}{n}\right)$. Thus, the probability is less than or equal to the following:

$$\begin{aligned}
\exp\left(-nh_+\left(\sqrt{\frac{\ln n}{n}}\right)\right) &= \exp\left(-\frac{1}{2p(1-p)}\ln n + o\left(1\right)\right) \\
&= \exp\left(\frac{-\ln n}{2p(1-p)}\right)\exp\left(o\left(1\right)\right) \\
&= n^{\frac{-1}{2p(1-p)}} \cdot \exp\left(o\left(1\right)\right).
\end{aligned}$$

Hence $\exp\left(-nh_+\left(\sqrt{\frac{\ln n}{n}}\right)\right) \sim n^{\frac{-1}{2p(1-p)}}$, and $\sum_{n\geq 1} n^{\frac{-1}{2p(1-p)}}$ is convergent because of the following inequalities

$$\begin{aligned}
p(1-p) &\leq \frac{1}{4} \\
2p(1-p) &\leq \frac{1}{2} \\
\frac{1}{2p(1-p)} &\geq 2 \\
\frac{-1}{2p(1-p)} &\leq -2 \\
n^{\frac{-1}{2p(1-p)}} &\leq n^{-2}.
\end{aligned}$$

Thus,

$$\sum_{n\geq 1} \mathbb{P}\left[S_n - np > \sqrt{n\ln n}\right] < \infty,$$

and so

$$\mathbb{P}\left[S_n - np \leq \sqrt{n\ln n} \text{ i.o.}\right] = 1.$$

□

Proof of Khinchin's Law of Iterated Logarithm

Theorem 4 (Khinchin's Law of Iterated Logarithm)**.** *Almost surely,*

$$\limsup_{n\to\infty} \frac{S_n - np}{\sqrt{2p(1-p)n\ln(\ln n)}} = 1,$$

and

$$\liminf_{n\to\infty} \frac{S_n - np}{\sqrt{2p(1-p)n\ln(\ln n)}} = -1.$$

First establish a two lemmas, and for convenience, let

$$\alpha(n) = \sqrt{2p(1-p)n\ln(\ln n)}.$$

Lemma 5. *For all positive a and δ and large enough n,*

$$(\ln n)^{-a^2(1+\delta)} < \mathbb{P}\left[S_n - np > a\alpha(n)\right] < (\ln n)^{-a^2(1-\delta)}.$$

Proof of Lemma 5. Recall that the Large Deviations Estimate gives

$$\begin{aligned}\mathbb{P}\left[R_n \geq a\alpha(n)\right] &= \mathbb{P}\left[S_n - np \geq a\alpha(n)\right] \\ &= \mathbb{P}\left[\frac{S_n}{n} - p \geq \frac{a\alpha(n)}{n}\right] \\ &\leq \exp\left(-nh_+\left(\frac{a\alpha(n)a}{n}\right)\right).\end{aligned}$$

Note that $\frac{\alpha(n)}{n} \to 0$ as $n \to \infty$. Then

$$h_+\left(\frac{a\alpha(n)}{n}\right) = \frac{a^2}{2p(1-p)}\left(\frac{\alpha(n)}{n}\right)^2 + O\left(\left(\frac{\alpha(n)}{n}\right)^3\right),$$

and so

$$nh_+\left(\frac{a\alpha(n)}{n}\right) = a^2\ln(\ln n) + O\left(\frac{\alpha(n)^3}{n^2}\right) \geq a^2(1-\delta)\ln(\ln n)$$

for large enough n. This means that

$$\mathbb{P}\left[\frac{S_n}{n} - p \geq a\alpha(n)\right] \leq \exp\left(-a^2(1-\delta)\ln(\ln n)\right) = (\ln n)^{-a^2(1-\delta)}.$$

Since $\sqrt{\ln(\ln n)} = o\left(n^{1/6}\right)$, the results of the Moderate Deviations Theorem apply to give

$$\begin{aligned}\mathbb{P}\left[\frac{S_n}{n} - p \geq \frac{a\alpha(n)}{n}\right] &= \mathbb{P}\left[\frac{S_n}{n} - p \geq \sqrt{\frac{p(1-p)}{n}}a\sqrt{2\ln(\ln n)}\right] \\ &\sim \frac{1}{\sqrt{2\pi}a\sqrt{2\ln(\ln n)}}\exp\left(-a^2\ln(\ln n)\right) \\ &= \frac{1}{2a\sqrt{\pi\ln(\ln n)}}(\ln n)^{-a^2}.\end{aligned}$$

Since $\sqrt{\ln(\ln n)} = o\left((\ln n)^{a^2\delta}\right)$,

$$\mathbb{P}\left[\frac{S_n}{n} - p \geq \frac{a\alpha(n)}{n}\right] \geq (\ln n)^{-a^2(1+\delta)}$$

for large enough n. □

Lemma 6 (12.5, Kolmogorov Maximal Inequality)**.** *Suppose* $(Y_n)_{n\geq 1}$ *are independent random variables. Suppose further that* $\mathbb{E}[Y_n] = 0$ *and* $Var(Y_n) = \sigma^2$*. Define* $T_n := Y_1 + \cdots + Y_n$*. Then*

$$\mathbb{P}\left[\max_{1\leq k\leq n} T_k \geq b\right] \leq \frac{4}{3}\mathbb{P}\left[T_n \geq b - 2\sigma\sqrt{n}\right].$$

Remark. Lemma 6 is an example of a class of lemmas called maximal inequalities. Here are two more examples of maximal inequalities.

Lemma 7 (Karlin and Taylor, page 280)*. Let* $(Y_n)_{n\geq 1}$ *be identical independently distributed random variables with* $\mathbb{E}[Y_n] = 0$ *and* $Var(Y_n) = \sigma^2 < \infty$*. Define* $T_n = \sum_{k=1}^n Y_k$*. Then*

$$\epsilon^2\mathbb{P}\left[\max_{0\leq k\leq n} |T_k| > \epsilon\right] \leq n\sigma^2.$$

Lemma 8 (Karlin and Taylor, page 280). *Let $(X_n)_{n\geq 0}$ be a submartingale for which $X_n \geq 0$. For $\lambda > 0$,*

$$\lambda \mathbb{P}\left[\max_{0\leq k\leq n} X_k > \lambda\right] \leq \mathbb{E}\left[X_n\right].$$

Proof of Lemma 6. Since the Y_k's are independent, then $\mathrm{Var}(T_n - T_k) = (n-k)\sigma^2$ for $1 \leq k \leq n$. Chebyshev's Inequality tells us that

$$\mathbb{P}\left[|T_n - T_k| \leq 2\sigma\sqrt{n}\right] \geq 1 - \frac{\mathrm{Var}(T_n - T_k)}{4\sigma^2 n} = 1 - \frac{n-k}{4n} \geq \frac{3}{4}.$$

Note that

$$\begin{aligned}
\mathbb{P}\left[\max_{0\leq k\leq n} T_k \geq b\right] &= \sum_{k=1}^{n} \mathbb{P}\left[T_1 < b, \dots, T_{k-1} < b, \text{ and } T_k \geq b\right] \\
&\leq \sum_{k=1}^{n} \mathbb{P}\left[T_1 < b, \dots, T_{k-1} < b, \text{ and } T_k \geq b\right] \cdot \frac{4}{3}\mathbb{P}\left[|T_n - T_k| \leq 2\sigma n\right] \\
&= \frac{4}{3}\sum_{k=1}^{n} \mathbb{P}\left[T_1 < b, \dots, T_{k-1} < b, \text{and } T_k \geq b \text{ and } |T_n - T_k| \leq 2\sigma\sqrt{n}\right] \\
&\leq \frac{4}{3}\sum_{k=1}^{n} \mathbb{P}\left[T_1 < b, \dots, T_{k-1} < b, \text{ and } T_k \geq b \text{ and } T_n \geq b - 2\sigma\sqrt{n}\right] \\
&\leq \frac{4}{3}\mathbb{P}\left[T_n \geq b - 2\sigma\sqrt{n}\right].
\end{aligned}$$

□

Remark. Note that the second part of the Law of the Iterated Logarithm, $\liminf_{n\to\infty} \frac{S_n - np}{\alpha(n)} = -1$ follows by symmetry from the first part by replacing p with $(1-p)$ and S_n with $n - S_n$.

Remark. The proof of the Law of the Iterated Logarithm proceeds in two parts. First it suffices to show that

$$\limsup_{n\to\infty} \frac{S_n - np}{\alpha(n)} < 1 + \eta \text{ for } \eta > 0, \text{ a.s.}. \tag{1}$$

The second part of the proof is to establish that

$$\limsup_{n\to\infty} \frac{S_n - np}{\alpha(n)} > 1 - \eta \text{ for } \eta > 0, \text{ a.s.} \tag{2}$$

It will only take a subsequence to prove (2). However this will not be easy because it will use the second Borel-Cantelli Lemma, which requires independence.

Remark. The following is a simplified proof giving a partial result for integer sequences with exponential growth. This simplified proof illustrates the basic ideas of the proof. Fix $\gamma > 1$ and let $n_k := \lfloor \gamma^k \rfloor$. Then

$$\begin{aligned}\mathbb{P}\left[S_{n_k} - pn_k \geq (1+\eta)\alpha(n_k)\right] &< (\ln n_k)^{-(1+\eta)^2(1-\delta)} \\ &= O\left(k^{-(1+\eta)^2(1-\delta)}\right).\end{aligned}$$

Choose δ so that $(1+\eta)^2(1-\delta) < 1$. Then

$$\sum_{k\geq 1} \mathbb{P}\left[S_{n_k} - n_k p \geq (1+\eta)\alpha(n_k)\right] < \infty.$$

By the first Borel-Cantelli lemma,

$$\mathbb{P}\left[S_{n_k} - n_k p \geq (1+\eta)\alpha(n_k) \text{ i.o.}\right] = 0,$$

and so

$$\mathbb{P}\left[\limsup_{k\to\infty} \frac{S_{n_k} - n_k p}{\alpha(n_k)} \leq (1+\eta)\right] = 1,$$

or

$$\frac{S_{n_k} - n_k p}{\alpha(n_k)} \leq (1+\eta) \text{ a.s.}$$

The full proof of the Law of the Iterated Logarithm takes more work to complete.

Proof of (1) *in the Law of the Iterated Logarithm.* Fix $\eta > 0$ and let $\gamma > 1$ be a constant chosen later. For $k \in \mathbb{Z}$, let $n_k = \lfloor \gamma^k \rfloor$. The proof consists of showing that

$$\sum_{k\geq 1} \mathbb{P}\left[\max_{n\leq n_{k+1}} (S_n - np) \geq (1+\eta)\alpha(n_k)\right] < \infty.$$

From Lemma 6

$$\begin{aligned}&\sum_{k\geq 1} \mathbb{P}\left[\max_{n\leq n_{k+1}} (S_n - np) \geq (1+\eta)\alpha(n_k)\right] \\ &\qquad\leq \frac{4}{3}\mathbb{P}\left[R_{n_{k+1}} \geq (1+\eta)\alpha(n_k) - 2\sqrt{n_{k+1}p(1-p)}\right],\end{aligned}$$

where $R_n = S_n - np$. We do know that

$$\mathbb{P}\left[\max_{n\le n_{k+1}}(S_n - np) \ge (1+\eta)\alpha(n_k)\right] \le \frac{4}{3}\mathbb{P}\left[R_{n_{k+1}} \ge (1+\eta)\alpha(n_k) - 2\sqrt{n_{k+1}p(1-p)}\right]. \quad (3)$$

Note that $\sqrt{n_{k+1}} = o\left(\alpha(n_k)\right)$ because this is approximately

$$\sqrt{\gamma^{k+1}} \text{ compared to } c_1\sqrt{\gamma^k \ln(\ln\gamma^\gamma)} = c_1\gamma^{k/2}\sqrt{\ln(\ln\gamma)+\ln(k)},$$

which is the same as

$$\gamma^{1/2} \text{ compared to } c_1\sqrt{c_2+\ln(k)}.$$

Then $2\sqrt{n_{k+1}p(1-p)} < \frac{1}{2}\eta\alpha(n_k)$ for large enough k. Using this inequality in the right hand side of Equation (3), we get

$$\mathbb{P}\left[\max_{n\le n_{k+1}} S_n - np \ge (1+\eta)\alpha(n_k)\right] \le \frac{4}{3}\mathbb{P}\left[S_{n_{k+1}} - n_{k+1}p \ge (1+\eta/2)\alpha(n_k)\right].$$

Now, $\alpha(n_{k+1}) \sim \sqrt{\gamma}\alpha(n_k)$. Choose γ so that $1+\eta/2 > (1+\eta/4)\sqrt{\gamma}$. Then for large enough k, we have

$$(1+\eta/2)\alpha(n_k) > (1+\eta/4)\alpha(n_{k+1}).$$

Now

$$\mathbb{P}\left[\max_{n\le n_{k+1}} S_n - np \ge (1+\eta)\alpha(n_k)\right] \le \frac{4}{3}\mathbb{P}\left[S_{n_{k+1}} - n_{k+1}p \ge (1+\eta/4)\alpha(n_{k+1})\right].$$

Use Lemma 5 with $a = (1-\delta)^{-1} = (1+\eta/4)$. Then we get

$$\mathbb{P}\left[\max_{n\le n_{k+1}} S_n - np \ge (1+\eta)\alpha(n_k)\right] \le \frac{4}{3}\left(\ln n_{k+1}\right)^{-(1+\eta/4)}$$

for k large. Note that

$$\left(\ln n_{k+1}\right)^{-(1+\eta/4)} \sim \left(\ln\gamma\right)^{-(1+\eta/4)} k^{-(1+\eta/4)},$$

which is the general term of a convergent series so

$$\sum_{k\ge1}\mathbb{P}\left[\max_{n\le n_{k+1}} R_n \ge (1+\eta)\alpha(n_k)\right] < \infty.$$

Then the first Borel-Cantelli Lemma implies that

$$\max_{n \le n_{k+1}} R_n \ge (1+\eta)\alpha(n_k) \text{ i.o. with probability 0.}$$

or equivalently

$$\max_{n \le n_{k+1}} S_n - np < (1+\eta)\alpha(n_k) \text{ a.s. for large enough } k.$$

Then in particular

$$\max_{n_k \le n < n_{k+1}} S_n - np < (1+\eta)\alpha(n_k) \text{ a.s. for large enough } k.$$

This in turn implies that almost surely

$$S_n - np < (1+\eta)\alpha(n).$$

for $n > n_k$ and large enough k which establishes (1). □

Proof of (2) *in the Law of the Iterated Logarithm.* Continue with the proof of Equation (2). To prove the second part, it suffices to show that there exists n_k so that $R_{n_k} \ge (1-\eta)\alpha(n_k)$ i.o. almost surely. Let $n_k = \gamma^k$ for γ chosen later with $\gamma \in \mathbb{Z}$ sufficiently large. The proof will show

$$\sum_{n \ge 1} \mathbb{P}\left[R_{\gamma^n} - R_{\gamma^{n-1}} \ge \left(1 - \frac{\eta}{2}\right)\alpha(\gamma^n)\right] = \infty, \tag{4}$$

and also

$$R_{\gamma^{n-1}} \ge \frac{-\eta}{2}\alpha(\gamma^n) \text{ a.s. for large enough } n. \tag{5}$$

Note that $R_{\gamma^n} - R_{\gamma^{n-1}} \overset{\text{dist.}}{=} R_{\gamma^n - \gamma^{n-1}}$. It suffices to consider

$$\mathbb{P}\left[R_{\gamma^n - \gamma^{n-1}} \ge \left(1 - \frac{\eta}{2}\right)\alpha(\gamma^n)\right].$$

Note that

$$\begin{aligned}
\frac{\alpha(\gamma^n - \gamma^{n-1})}{\alpha(\gamma^n)} &= \frac{\sqrt{c\left(\gamma^n - \gamma^{n-1}\right)\ln\left(\ln\left(\gamma^n - \gamma^{n-1}\right)\right)}}{\sqrt{c\gamma^n \ln\left(\ln \gamma^n\right)}} \\
&= \sqrt{\left(1 - \frac{1}{\gamma}\right)\frac{\ln\left(n \ln \gamma + \ln\left(1 - \frac{1}{\gamma}\right)\right)}{\ln\left(n \ln \gamma\right)}} \\
&\to \sqrt{1 - \frac{1}{\gamma}}.
\end{aligned}$$

Choose $\gamma \in \mathbb{Z}$ so that

$$\frac{1-\frac{\eta}{2}}{1-\frac{\eta}{4}} < \sqrt{1-\frac{1}{\gamma}}.$$

Then note that we can choose n large enough so that

$$\frac{\left(1-\frac{\eta}{2}\right)}{\left(1-\frac{\eta}{4}\right)} < \frac{\alpha(\gamma^n - \gamma^{n-1})}{\alpha(\gamma^n)}$$

or

$$\left(1-\frac{\eta}{2}\right)\alpha(\gamma^n) < \left(1-\frac{\eta}{4}\right)\alpha(\gamma^n - \gamma^{n-1}).$$

Now considering equation (4), and the inequality above

$$\mathbb{P}\left[R_{\gamma^n} - R_{\gamma^{n-1}} \geq \left(1-\frac{\eta}{2}\right)\alpha(\gamma^n)\right] \geq \mathbb{P}\left[R_{\gamma^n-\gamma^{n-1}} \geq \left(1-\frac{\eta}{4}\right)\alpha(\gamma^n - \gamma^{n-1})\right].$$

Now using Lemma 5, with $a = (1+\delta)^{-1} = \left(1-\frac{\eta}{4}\right)$, we get

$$\mathbb{P}\left[R_{\gamma^n} - R_{\gamma^{n-1}} \geq \left(1-\frac{\eta}{2}\right)\alpha(\gamma^n)\right] \geq \ln\left(\gamma^n - \gamma^{n-1}\right)^{-\left(1-\frac{\eta}{4}\right)} = \left(n\ln\gamma + \ln\left(1-\frac{1}{\gamma}\right)\right)^{-\left(1-\frac{\eta}{4}\right)}.$$

The series with this as its terms diverges. Thus, we see that Equation (4) has been proven.

Now notice that

$$\begin{aligned}\alpha(\gamma^n) &= \sqrt{c\gamma^n \ln\left(\ln\gamma^n\right)} \\ &= \sqrt{c\gamma^n\left(\ln n + \ln\ln\gamma\right)}\end{aligned}$$

and so

$$\alpha(\gamma^{n-1}) = \sqrt{c\gamma^{n-1}\left(\ln(n-1) + \ln\ln\gamma\right)},$$

which means that

$$\sqrt{\gamma}\alpha(\gamma^{n-1}) = \sqrt{c\gamma^n\left(\ln(n-1) + \ln\ln\gamma\right)}.$$

Thus, $\alpha(\gamma^n) \sim \sqrt{\gamma}\alpha(\gamma^{n-1})$. Now choose γ so that $\eta\sqrt{\gamma} > 4$. Then $\eta\alpha(\gamma^n) \sim \eta\sqrt{\gamma}\alpha(\gamma^{n-1}) > 4\alpha(\gamma^{n-1})$ for large enough n.

Thus, we have

$$\left[R_{\gamma^{n-1}} \leq \frac{-\eta}{2}\alpha(\gamma^n)\right] \subseteq \left[-R_{\gamma^{n-1}} \geq 2\alpha(\gamma^{n-1})\right]$$

since

$$\left[R_{\gamma^{n-1}} \leq \frac{-\eta}{2}\alpha(\gamma^n)\right] \subseteq \left[R_{\gamma^{n-1}} \leq -2\alpha(\gamma^{n-1})\right].$$

Now use (4) and we see that $-R_{\gamma^{n-1}} < 2\alpha(\gamma^{n-1})$ happens almost surely for large enough n.

Now $R_{\gamma^n} - R_{\gamma^{n-1}}$ is a sequence of independent random variables, and so the second Borel-Cantelli Lemma says that almost surely

$$R_{\gamma^n} - R_{\gamma^{n-1}} > \left(1 - \frac{\eta}{2}\right)\alpha(\gamma^n) \text{ i.o.}$$

Adding this with Equation (5), we get that

$$R_{\gamma^n} > (1-\eta)\,\alpha(\gamma^n) \text{ i.o.}$$

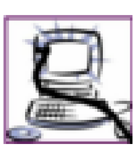
This is enough to show that

$$\liminf_{n\to\infty} \frac{S_n - np}{\alpha(n)} > 1 - \eta \text{ a.s.,}$$

which is enough to show the only remaining part of Khinchin's Law of the Iterated Logarithm. □

Sources

This section is adapted from: W. Feller, in *Introduction to Probability Theory and Volume I*, Chapter III, and Chapter VIII, and also E. Lesigne, *Heads or Tails: An Introduction to Limit Theorems in Probability*, Chapter 12, American Mathematical Society, Student Mathematical Library, Volume 28, 2005. Some of the ideas in the proof of Hausdorff's Estimate are adapted from J. Lamperti, *Probability: A Survey of the Mathematical Theory*, Second Edition, Chapter 8. Figure 3 is a recreation of a figure in the Wikipedia article on the Law of the Iterated Law.



Algorithms, Scripts, Simulations

Algorithm

Scripts

R script for comparison figures.

```
p <- 0.5
k <- 15
n <- 100000
coinFlips <- array(runif(n*k) <= p, dim=c(n,k))
S <- apply(coinFlips, 2, cumsum)
steps <- c(1:n)

steps2 <- steps[2:n]
S2 <- S[2:n, ]
steps3 <- steps[3:n]
S3 <- S[3:n, ]

ones <- cbind( matrix(1,n-2,1), matrix(-1,n-2,1))

par( mfrow = c(2,2))

matplot((S-steps*p)/steps,
        log="x", type="l", lty = 1, ylab="", main="Strong Law")
matplot((S-steps*p)/sqrt(2*p*(1-p)*steps),
        log="x", type="l", lty = 1, ylab="", main="Central Limit Theorem")
matplot((S-steps*p)/(steps^(0.6)),
        log="x", type="l", lty = 1, ylab="", main="Hausdorff's Estimate")
## matplot((S2-steps2*p)/sqrt(steps2*log(steps2)), log="x", xlim=c(1,n), type="l", lty = 1)
matplot(steps3, (S3-steps3*p)/sqrt(2*p*(1-p)*steps3*log(log(steps3))),
        log="x", xlim=c(1,n), type="l", lty = 1, ylab="", main="Law of Iterated Logarithm")
matlines(steps3, ones, type="l", col="black")
```

Problems to Work for Understanding

1. The "multiplier of the variance $\sqrt{2p(1-p)n}$" function $\sqrt{\ln(\ln(n))}$ grows very slowly. To understand how very slowly, calculate a table with $n = 10^j$ and $\sqrt{2\ln(\ln(n))}$ for $j = 10, 20, 30, \dots, 100$. (Remember, in mathematical work above calculus, $\ln(x)$ is the natural logarithm, base e, often written $\ln(x)$ in calculus and below to distinguish it from the "common" or base-10 logarithm. Be careful, some software and technology cannot directly calculate with magnitudes this large.)

2. Consider the sequence
$$a_n = (-1)^{\lfloor n/2 \rfloor} + \frac{(-1)^n}{n}$$
for $n = 1, 2, 3 \dots$. Here $\lfloor x \rfloor$ is the "floor function", the greatest integer less than or equal to x, so $\lfloor 1 \rfloor = 1$, $\lfloor 3/2 \rfloor = 1$, $\lfloor 8/3 \rfloor = 2$, $\lfloor -3/2 \rfloor = -2$, etc. Find
$$\limsup_{n\to\infty} a_n$$
and
$$\liminf_{n\to\infty} a_n.$$
Does the sequence a_n have a limit?

3. Show that the second part of the Law of the Iterated Logarithm, $\liminf_{n\to\infty} \frac{S_n - np}{\alpha(n)} = -1$ follows by symmetry from the first part by replacing p with $(1-p)$ and S_n with $n - S_n$.

Solutions to Problems

Reading Suggestion:

References

[1] William Feller. *An Introduction to Probability Theory and Its Applications, Volume I*, volume I. John Wiley and Sons, third edition, 1973. QA 273 F3712.

[2] John W. Lamperti. *Probability: A Survey of the Mathematical Theory.* Wiley Series in Probability and Statistics. Wiley, second edition edition, 1996.

[3] Emmanuel Lesigne. *Heads or Tails: An Introduction to Limit Theorems in Probability*, volume 28 of *Student Mathematical Library.* American Mathematical Society, 2005.

Outside Readings and Links:

1.
2.
3.
4.

I check all the information on each page for correctness and typographical errors. Nevertheless, some errors may occur and I would be grateful if you would alert me to such errors. I make every reasonable effort to present current and accurate information for public use, however I do not guarantee the accuracy or timeliness of information on this website. Your use of the information from this website is strictly voluntary and at your risk.

I have checked the links to external sites for usefulness. Links to external websites are provided as a convenience. I do not endorse, control, monitor, or guarantee the information contained in any external website. I don't guarantee that the links are active at all times. Use the links here with the same caution as you would all information on the Internet. This website reflects the thoughts, interests and opinions of its author. They do not explicitly represent official positions or policies of my employer.

Information on this website is subject to change without notice.

Steve Dunbar's Home Page, `http://www.math.unl.edu/~sdunbar1`

Email to Steve Dunbar, `sdunbar1 at unl dot edu`

Last modified: Processed from LaTeX source on March 22, 2018