Topics in
Probability Theory and Stochastic Processes
Steven R. Dunbar

Strong Law of Large Numbers

Study Tip

Rating
Mathematicians Only: prolonged scenes of intense rigor.
Section Starter Question

Explain what is meant by the “law of averages” and how it applies to an infinite sequence of coin flips.

---

Key Concepts

1. Almost surely

\[ \lim_{n \to \infty} \frac{S_n(\omega)}{n} = p. \]

2. Almost surely, the asymptotic proportion of any outcome in an infinite sequence of trials is the probability of that outcome for a single trial.

---

Vocabulary

1. Borel’s strong law of large numbers is that almost surely

\[ \lim_{n \to \infty} \frac{S_n(\omega)}{n} = p. \]
2. If \( b = (b_1, b_2, \ldots, b_j) \) is a word constructed from the alphabet \( \{0, 1\} \) and 
\( s = \sum_{i=1}^{j} b_j \) we say that \( p^s q^{1-s} \) is the probability of the word \( b \).

3. The event "\( X_n \) is in \( A_n \) infinitely often" denoted \( \{ X_n \in A_n \ i.o. \} \) is defined as 
\[ \{ X_n \in A_n \ i.o. \} = \bigcap_{n=1}^{\infty} \bigcup_{j=n}^{\infty} A_j \]

---

**Mathematical Ideas**

**Borel’s Strong Law**

This section presents three different proofs of Borel’s Strong Law of Large Numbers. The first uses the set theory of negligible sets and the Large Deviations estimate to show that the pointwise convergence fails on a negligible set. The second proof uses a carefully chosen subsequence and elementary estimates. The third proof uses the Borel-Cantelli Lemma and a fourth-moment condition.

**Proof with the Large Deviations Estimate**

**Theorem 1** (Borel’s Strong Law of Large Numbers). *Almost surely*

\[
\lim_{n \to \infty} \frac{S_n(\omega)}{n} = p.
\]

*Proof.* Let \( R_n(\omega) = \frac{S_n(\omega)}{n} - p \). The sequence \( \{ R_n(\omega) \}_{n=1}^{\infty} \) fails to approach 0 if and only if \( \liminf_{n \to \infty} |R_n(\omega)| > 0 \). That is, \( \{ R_n(\omega) \}_{n=1}^{\infty} \) fails to approach 0 if and only if there is an \( m \geq 1 \) depending on \( \omega \) such that for that for each \( n \geq 1 \) there exists a \( k \geq n \) satisfying \( |R_k(\omega)| > \frac{1}{m} \). Therefore, the set of points \( \omega \) in the sample space *not* satisfying \( \lim_{n \to \infty} \frac{S_n(\omega)}{n} = p \) is contained in 
\[
\bigcup_{m \geq 1} \bigcap_{n \geq 1} \bigcup_{k \geq n} \left\{ \omega \in \Omega : |R_k(\omega)| > \frac{1}{m} \right\}.
\]
The goal is to show that this set is a negligible event. Since a countable union of negligible sets is negligible, it suffices to show that

\[ N_m = \bigcap_{n \geq 1} \bigcup_{k \geq n} \left\{ \omega \in \Omega : |R_k(\omega)| > \frac{1}{m} \right\} \]

is negligible for each \( m \geq 1 \).

For each \( k \geq 1 \), let \( A_{m,k} = \left\{ \omega \in \Omega : |R_k(\omega)| > \frac{1}{m} \right\} \). By the Large Deviations Estimate, there is a constant \( c = c(p, m) > 0 \) such that \( \mathbb{P}[A_{m,k}] \leq e^{-ck} \). Since the series \( \sum_{k \geq 1} e^{-ck} \) converges, for every \( \epsilon > 0 \) there is an \( n \geq 1 \) such that \( \sum_{k \geq n} e^{-ck} < \epsilon \). Because \( N_m \subset \bigcup_{k \geq n} A_{m,k} \),

\[ \mathbb{P}[N_m] \leq \mathbb{P}[\bigcup_{k \geq n} A_{m,k}] \leq \sum_{k \geq n} \mathbb{P}[A_{m,k}] \leq \epsilon. \]

This proves that each \( N_m \) is a negligible set.

\[ \square \]

**Proof with subsequences**

Here is another proof of Borel’s Strong Law of Large Numbers, adapted from Theorem 1.21, in [1, page 11].

Before the proof, first give a quick definition.

**Definition.** Suppose \( A_1 \subseteq A_2 \subseteq \ldots \). Then \( \lim_{n \to \infty} A_n = \bigcup_{n=1}^{\infty} A_n \). Also, if \( A_1 \supseteq A_2 \supseteq \ldots \), then \( \lim_{n \to \infty} A_n = \bigcap_{n=1}^{\infty} A_n \).

**Theorem 2** (Strong Law of Large Numbers).

\[ \mathbb{P} \left[ \lim_{n \to \infty} \frac{S_n}{n} \neq p \right] = 0. \]

**Proof.** Break the proof into a sequence of simple steps.

1. The first step is the claim that \( \lim_{n \to \infty} \frac{S_n}{n} = p \) if and only if \( \lim_{m \to \infty} \frac{S_{m^2}}{m^2} = p \).

To see this, first note that the implication from the limit of the full sequence to the subsequential limit along the squares is immediate. For the other direction, for any \( n \), find \( m \) so that \( m^2 \leq n < (m+1)^2 \),
or \(0 \leq n - m^2 \leq 2m\). Then
\[
\left| \frac{S_n - S_{m^2}}{n - m^2} \right| = \frac{S_n - S_{m^2}}{m^2} + \left( \frac{1}{n} - \frac{1}{m^2} \right) S_n
\leq \frac{|n - m^2|}{m^2} + \left| \frac{1}{n} - \frac{1}{m^2} \right| n
\leq \frac{2m}{m^2} + \frac{2}{m} = \frac{4}{m}.
\]
The term \(\frac{|n - m^2|}{m^2}\) at the second line comes from estimating \(\frac{|S_n - S_{m^2}|}{m^2}\). Since \(n \geq m\), the largest the difference could be would be having each \(\omega_i = 1\) for \(i > m\), thus the difference could be at most \(n - m^2\). Then the claim that \(\lim_{m \to \infty} \frac{S_{m^2}}{m^2} = p\) implies \(\lim_{n \to \infty} \frac{S_n}{n} = p\) follows by the triangle inequality.

2. For \(m_0 < m_1\) define
\[
E_{m_0, m_1} = \bigcup_{m = m_0}^{m_1} \{ \omega : \left| \frac{S_m}{m^2} - p \right| > \epsilon \}.
\]
Note that each of these sets depends only on the coordinates \(\omega_1, \ldots, \omega_{m^2}\), so \(E_{m_0, m_1}\) is of finite type.

3. Use the Chebyshev inequality on each set in the union to obtain
\[
P \left[ \left| \frac{S_m}{m^2} - p \right| \right] < \frac{1}{\epsilon^2} \mathbb{E} \left[ \left| \frac{S_m}{m^2} - p \right|^2 \right] = \frac{1}{\epsilon^2} \frac{p(1 - p)}{m^2}.
\]

4. Thus
\[
P \left[ E_{m_0, m_1} \right] \leq \sum_{m = m_0}^{m_1} P \left[ \left| \frac{S_m}{m^2} - p \right| > \epsilon \right] = \frac{p(1 - p)}{\epsilon^2} \sum_{m = m_0}^{m_1} \frac{1}{m^2}.
\]

5. Define
\[
E_{m_0} = \lim_{m_1 \to \infty} E_{m_0, m_1} = \bigcup_{m = m_0}^{\infty} \left\{ \omega : \left| \frac{S_m}{m^2} - \frac{1}{2} \right| > \epsilon \right\}.
\]
Note that $E_{m_0}$ is the set of $\omega$’s for which $\left| \frac{S_{m^2}}{m^2} - p \right| > \epsilon$ at least once for $m \geq m_0$.

6. The claim is that
\[
P [E_{m_0}] = \lim_{m_1 \to \infty} P_{m_1} [E_{m_0, m_1}] \leq \frac{1}{4\epsilon^2} \sum_{m=m_0}^{\infty} \frac{1}{m^2}.
\]

7. Note that
\[
\limsup_m \left| \frac{S_{m^2}}{m^2} - p \right| > \epsilon
\]
if and only if for any $m_0$ there exists $m \geq m_0$ so that $\left| \frac{S_{m^2}}{m^2} - p \right| > \epsilon$.
That is,
\[
\left\{ \omega : \limsup_m \left| \frac{S_{m^2}}{m^2} - p \right| > \epsilon \right\} \subseteq E_{m_0},
\]
for all $m_0$. Notice that $E_{m_0} \supseteq E_{m_0+1}$ and so consider $E_\epsilon = \lim_{m_0 \to \infty} E_{m_0}$. Again the claim is that
\[
P [E_\epsilon] = \lim_{m_0 \to \infty} P [E_{m_0}].
\]

8. Then
\[
P [E_\epsilon] \leq \lim_{m_0 \to \infty} \left[ \frac{p(1 - p)}{\epsilon^2} \sum_{m=m_0}^{\infty} \frac{1}{m^2} \right] = 0.
\]

9. Let $E_{1/k}$ be a special case of $E_\epsilon$ for $\frac{1}{k} = \epsilon$ and $k \in \mathbb{Z}$. Note that
\[
E_{1/k} = \left\{ \omega : \limsup_m \left| \frac{S_{m^2}}{m^2} - p \right| > \frac{1}{k} \right\}.
\]

Let $E = \lim_k E_{1/k}$ and then note that $P [E] = \lim_k P [E_{1/k}] = \lim_k 0 = 0$.

\hfill \Box

Remark. At steps 6, 7, and 9 a limit is pulled through the probability measure. This is valid if given a sequence of finite probability measures $P_n [\cdot]$ defined on a field of finite type events $\mathcal{F}_0$, there exists a well-defined probability measure $P [\cdot]$ defined on $\mathcal{F}$, the smallest $\sigma$-field containing $\mathcal{F}_0$ agreeing with $P_n [\cdot]$ on $\mathcal{F}_0$. That is true from the following theorem, with proof deferred to another section.
Theorem 3 (Kolmogorov Consistency Theorem). Given a consistent family of finite-dimensional distributions $\mathbb{P}_n[\cdot]$ there exists a unique probability $\mathbb{P}[\cdot]$ on $(\Omega, \Sigma)$ such that for every $n$, under the natural projection $\pi_n(\omega) = (x_1, x_2, \ldots, x_n)$ the induced measure $\mathbb{P}_{\pi_n^{-1}}$ is identical to $\mathbb{P}_n$ on $\mathbb{R}^n$.

Proof with the Borel-Cantelli Lemma

Lemma 4 (Borel-Cantelli). The direct half: For a sequence of events $A_n$, if

$$\sum_{n=1}^{\infty} \mathbb{P}[A_n] < \infty$$

then

$$\mathbb{P}\left[\omega : \lim_{n \to \infty} 1_{A_n}(\omega) = 0\right] = 1.$$

The converse half: If the events $A_n$ are mutually independent, the converse is also true.

Definition. Let $X_1, X_2, X_3, \ldots,$ be a sequence of random variables and $A_1, A_2, A_3, \ldots$ a sequence of Borel sets in $\mathbb{R}$. The event “$X_n$ is in $A_n$ infinitely often” denoted $\{X_n \in A_n \text{ i.o.}\}$ is defined as

$$\{X_n \in A_n \text{ i.o.}\} = \bigcap_{n=1}^{\infty} \bigcup_{j=n}^{\infty} A_j$$

Remark. Note that the complementary event to the event in the Borel-Cantelli Lemma is

$$\left[\omega : \limsup_{n \to \infty} 1_{A_n}(\omega) = 1\right].$$

This is exactly the event $\bigcap_{n=1}^{\infty} \bigcup_{j=n}^{\infty} A_j$, that is, the event that infinitely many of the events occur. Thus the direct half of the Borel-Cantelli Lemma can be expressed as

If $\sum_{n=1}^{\infty} \mathbb{P}[A_n] < \infty$ then $\mathbb{P}[A_n \text{ i.o.} ] = 0$.

The converse half of the Borel-Cantelli Lemma can be expressed as

If $\sum_{n=1}^{\infty} \mathbb{P}[A_n] = \infty$ then $\mathbb{P}[A_n \text{ i.o.} ] = 1$. 

7
Proof. The direct half:

\[ \mathbb{P} [A \text{ i.o.}] = \mathbb{P} [\cap_{n=1}^{\infty} \cup_{j=n}^{\infty} A_j] \]
\[ = \lim_{n \to \infty} \mathbb{P} [\cup_{j=n}^{\infty} A_j] \]
\[ \leq \lim_{n \to \infty} \sum_{j=m}^{\infty} \mathbb{P} [A_j] \]

But obviously, \( \sum_{n=1}^{\infty} \mathbb{P} [A_n] < \infty \) implies \( \lim_{n \to \infty} \sum_{j=m}^{\infty} \mathbb{P} [A_j] = 0 \).

The converse half: This will be proved in the form

If \( \sum_{n=1}^{\infty} \mathbb{P} [A_n] = \infty \) then \( \mathbb{P} [A \text{ i.o.}] = 1 \).

The complement of the set \( \{A_n \text{ i.o.}\} = \cap_{n=1}^{\infty} \cup_{j=n}^{\infty} A_j \) is \( \cup_{n=1}^{\infty} \cap_{j=n}^{\infty} A_j^C \). Then the goal is to show this is a negligible set. It is enough to show that each set \( \cap_{j=n}^{\infty} A_j^C \) is negligible. Temporarily fix \( n \) and for each \( m \geq n \), let \( B_m \) be the event \( B_m = \cap_{j=n}^{m} A_j^C \), so that \( \cap_{j=n}^{\infty} A_j^C \subset B_m \).

Because the events \( (A_n)_{n=1}^{\infty} \) are independent

\[ \mathbb{P} [B_m] = \mathbb{P} [\cap_{j=n}^{m} A_j^C] = \prod_{j=n}^{m} \mathbb{P} [A_j^C] = \prod_{j=n}^{\infty} (1 - \mathbb{P} [A_n]) \]
\[ \leq \prod_{j=n}^{\infty} \exp \left( -\mathbb{P} [A_n] \right) = \exp \left( -\sum_{j=n}^{\infty} \mathbb{P} [A_j] \right) . \]

This can be made arbitrarily small by choosing sufficiently large \( m \), so by definition, \( \cap_{j=n}^{\infty} A_j^C \) is negligible.

\[ \square \]

Theorem 5 (Strong Law of Large Numbers). If \( X_1, X_2, \ldots, X_n, \ldots \) is a sequence of independent identically distributed random variables with \( \mathbb{E} [|X_i|^4] = C < \infty \), then

\[ \lim_{n \to \infty} \frac{S_n}{n} = \lim_{n \to \infty} \frac{X_1 + X_2 + \cdots + X_n}{n} = \mathbb{E} [X_1] \]

with probability 1.
Proof. We can assume without loss of generality that $E[X_i] = 0$, otherwise just consider $Y_i = X_i - E[X_i]$. Expansion of $(X_1 + \cdots + X_n)^4$, then using the independence and identical distribution assumptions shows

$$E[S_n^4] = nE[X_1^4] + 3n(n-1)E[X_1^2]^2 \leq nC + 3n^2\sigma^4.$$ 

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

$$P\left[\left|\frac{S_n}{n}\right| \geq \delta\right] = P\left[|S_n|^4 \geq n^4\delta^4\right]$$

$$= \int_{|S_n|^4 \geq n^4\delta^4} dP$$

$$\leq \int_{\Omega} \frac{|S_n|^4}{n^4\delta^4} dP$$

$$= \frac{E[|S_n|^4]}{n^4\delta^4}$$

$$\leq nC + 3n^2\sigma^4 \frac{n^4}{n^4\delta^4}.$$

Then from the comparison theorem for series convergence

$$\sum_{n=1}^{\infty} P\left[\left|\frac{S_n}{n}\right| \geq \delta\right] < \infty,$$

and the direct half of the Borel-Cantelli Lemma applies to show that

$$P\left[\left|\frac{S_n}{n}\right| \geq \delta \text{ i.o.}\right] = 0$$

or equivalently

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

with probability 1. \qed

The Strong Law for Pairwise Independent Random Variables

Example. Bernstein's example of pairwise independence that are not all independent.
Consider flipping a quarter and a dime. Let $X_1 = 1$ if the quarter is heads, $X_2 = 1$ if the dime is heads, and $X_3 = 1$ if the quarter is the same as the dime. First, $\mathbb{P}[X_i] = \frac{1}{2}$ for $i = 1, 2, 3$. Next note the following probabilities:

\[
\mathbb{P}[X_1 = 1 \cap X_2 = 1] = \mathbb{P}[X_1 = 1] \mathbb{P}[X_2 = 1] = \frac{1}{4},
\]

\[
\mathbb{P}[X_1 = 1 \cap X_3 = 1] = \mathbb{P}[X_1 = 1] \mathbb{P}[X_3 = 1] = \frac{1}{4},
\]

\[
\mathbb{P}[X_2 = 1 \cap X_3 = 1] = \mathbb{P}[X_2 = 1] \mathbb{P}[X_3 = 1] = \frac{1}{4},
\]

but

\[
\frac{1}{4} = \mathbb{P}[X_1 = 1 \cap X_2 = 1 \cap X_3 = 1] \neq \mathbb{P}[X_1 = 1] \mathbb{P}[X_2 = 1] \mathbb{P}[X_3 = 1] = \frac{1}{8}.
\]

**Lemma 6.** Let $(X_n)_{n \geq 1}$ be a sequence of random variables. If $\sum_{n=1}^{\infty} \mathbb{P}[X_n > \epsilon]$ converges for all $\epsilon > 0$, then $\lim_{n \to \infty} X_n = 0$ almost surely.

**Proof.**

1. Set $A_n = [X_n > \epsilon]$.

2. Proposition ?? implies for that for every $\epsilon > 0$ there exists almost surely an $n_0(\omega, \epsilon)$ such that $|X_n(\omega)| \leq \epsilon$ for every $n \geq n_0(\omega, \epsilon)$.

3. By considering a countable union of negligible events, we conclude that for every positive integer $m$ there exists almost surely $n_0(\omega, 1/m)$ such that $|X_n(\omega)| \leq 1/m$ for all $n \geq n_0(\omega, 1/m)$.

4. This implies that the sequence $(X_n)_{n \geq 1}$ almost surely converges to 0.

\[\square\]

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

**Theorem 7** (Pairwise Independent Strong Law). Let $(X_n)_{n \geq 1}$ be a sequence of pairwise independent random variables. Suppose that $\mathbb{E}[X_n] < \infty$, and $\sup_{n \geq 1} \mathbb{E}[X_n^2] < \infty$. Then for $S_n = \sum_{i=1}^{n} X_n$, we have

\[
\lim_{n \to \infty} \mathbb{P} \left[ \frac{|S_n|}{n} \geq \epsilon \right] = 0
\]

for all $\epsilon > 0$. In other words, $\lim_{n \to \infty} \frac{S_n}{n} = 0$ almost surely.
Proof. Let $M = \sup_{n \geq 1} \mathbb{E}[X_n^2]$ and note that pairwise independence says
\[
\mathbb{E}[X_i X_j] = \mathbb{E}[X_i] \mathbb{E}[X_j] = 0, \quad i \neq j.
\]
Note that we have
\[
\mathbb{E} \left[ \left( \frac{S_n}{n} \right)^2 \right] = \frac{1}{n^2} \sum_{i=1}^{n} \mathbb{E}[X_i^2] \leq \frac{M}{n}.
\]
The Chebyshev inequality gives us that
\[
P \left[ \left| \frac{S_n}{n} \right| \geq \epsilon \right] \leq \frac{\mathbb{E} \left[ \left( \frac{S_n}{n} \right)^2 \right]}{\epsilon^2} \leq \frac{1}{\epsilon^2} M,
\]
and so $\lim_{n \to \infty} P \left[ \left| \frac{S_n}{n} \right| \geq \epsilon \right] = 0$. Since $\sum_{n \geq 1} \mathbb{E} \left[ \left( \frac{S_n}{n} \right)^2 \right] < \infty$, by Corollary 1
\[
\lim_{n \to \infty} \left( \frac{S_n^2}{n^2} \right)^2 = 0 \text{ almost surely, and hence } \lim_{n \to \infty} \frac{S_n^2}{n^2} = 0 \text{ almost surely.}
\]
Now let $m = \lfloor \sqrt{n} \rfloor$; i.e., $m^2 \leq n \leq (m+1)^2$. Then $\lim_{n \to \infty} \frac{S_m^2}{n^2} = 0$ almost surely. This gives
\[
\mathbb{E} \left[ \left( \frac{S_n}{n} - \frac{S_m^2}{n} \right)^2 \right] = \frac{1}{n^2} \mathbb{E} \left[ \left( \sum_{i=m^2+1}^{n} X_i \right)^2 \right]
\leq \frac{1}{n^2} \sum_{i=m^2+1}^{n} \mathbb{E}[X_i^2]
\leq \frac{(n - m^2)M}{n^2}
= \frac{2m + 1}{n^2} M
= \frac{2 \lfloor \sqrt{n} \rfloor + 1}{n^2} M
= O \left( n^{-3/2} \right).
\]
Using Corollary 1 again
\[
\lim_{n \to \infty} \frac{S_n}{n} - \frac{S_m^2}{n} = 0 \text{ a.s.}
\]
\[
\lim_{n \to \infty} \frac{S_n}{n} - \frac{S_m^2}{n} = 0 \text{ a.s.}
\]
\[
\lim_{n \to \infty} \frac{S_n}{n} = 0 \text{ a.s.}
\]
Applications of the Strong Law

The following proposition says that the asymptotic proportion of any outcome in an infinite sequence of trials is “almost surely” the probability of that outcome for a single trial. This is sometimes referred as the “Law of Averages”.

**Proposition 8 (The Law of Averages).** Let \((A_n)_{n=1}^{\infty}\) be a sequence of equiprobable independent random events with probability \(p\). The asymptotic empirical probability that these event will occur is almost surely \(p\); that is,

\[
\lim_{n \to \infty} \frac{1}{n} \left| \{k : 1 \leq k \leq n, \omega \in A_k\} \right| = p
\]

almost surely.

**Proof.** For each \(\omega\) create a sequence \(\rho = (\rho_n)_{n=1}^{\infty}\) of 0’s and 1’s by setting \(\rho_n = 1\) if \(\omega \in A_n\) and \(\rho_n = 0\) otherwise. The proof is to show

\[
\lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} \rho_k = P[\rho_i = 1] = p
\]

almost surely.

For each \(n\) and each \((\epsilon_1, \epsilon_2, \ldots, \epsilon_n) \in \{0,1\}^n\) let \(s = \sum_{k=0}^{n} \epsilon_k\), then

\[
P[\rho_k = \epsilon_k, 1 \leq k \leq n] = p^s(1-p)^{n-s}.
\]

Now the same proof used to show that \(\frac{1}{n} \sum_{k=1}^{n} \omega_k\) almost surely converges to \(p\) can be applied to complete the proof that the convergence of (1) is almost sure. □

**Strong Law for Finite Type Events**

**Corollary 2.** Let \(A\) be a finite type event. For each integer \(n \geq 1\) and each \(\omega \in \Omega\), let \(S(A, n, \omega)\) be the number of integers \(k\) between 1 and \(n\) such that \((\omega_k, \omega_{k+1}, \omega_{k+2}, \ldots) \in A\). Then

\[
\lim_{n \to \infty} \frac{S(A, n, \omega)}{n} = P[A]
\]

almost surely.
Proof. Since the event $A$ is of finite type, $A$ depends only on the coordinates with index less than or equal to $m$ for some positive integer $m$. Equivalently, there exists an $A' \subset \Omega_m$ such that

$$A = \{ \omega : (\omega_1, \omega_2, \ldots, \omega_m) \in A' \}.$$ 

For each integer $j$ from 1 to $m$ consider the sequence $(A_{j,n})_{n \geq 0}$ of events defined by

$$A_{j,n} = \{ \omega : (\omega_{j+nm}, \omega_{j+nm+1}, \omega_{j+nm+2}, \ldots) \in A \} = \{ \omega : (\omega_{j+nm}, \omega_{j+nm+1}, \ldots, \omega_{j+(n+1)m-1}) \in A' \}.$$ 

For fixed $j$ and varying $n$, the events $A_{j,n}$ are independent and each has probability $\mathbb{P}[A]$. Let $S(A, j, n, \omega)$ be the number of integers $k$ between 0 and $n - 1$ such that $\omega \in A_{j,k}$. The Proposition implies that

$$\lim_{n \to \infty} \frac{S(A, j, n, \omega)}{n} = \mathbb{P}[A]$$

almost surely. Also,

$$\frac{1}{nm} S(A, nm, \omega) = \frac{1}{m} \sum_{j=1}^{m} \frac{S(A, j, n, \omega)}{n}.$$ 

The corollary then follows from the inequality

$$\frac{1}{(n + 1)m} S(A, nm, \omega) \leq \frac{1}{nm+k} S(A, nm+k, \omega) \leq \frac{1}{nm} S(A, (n+1)m, \omega)$$

which holds for every $k$ from 0 to $m$. \hfill \Box

Definition. If $b = (b_1, b_2, \ldots, b_j)$ is a word constructed from the alphabet $\{0,1\}$ and $s = \sum_{i=1}^{j} b_j$ we say that $p^s q^{1-s}$ is the probability of the word $b$.

Corollary 3. In the sequence $\omega$, every word $b$ almost surely occurs with asymptotic frequency equal to its probability.

Proof. This is a special case of Proposition 2. \hfill \Box

13
Sources

This proof of the Strong Law for sum of Bernoulli random variables is adapted from: *Heads or Tails*, by Emmanuel Lesigne, Student Mathematical Library Volume 28, American Mathematical Society, Providence, 2005, Chapter 11.2, pages 82-85. The proof based on subsequential limits and the Borel Cantelli Lemma and following remarks is adapted from *Probability* by Leo Breiman, Addison-Wesley, Reading MA, 1968, Chapters 1 and 3. The proof using the fourth-moment and the Borel-Cantelli Lemma is adapted from Varadhan.

Algorithms, Scripts, Simulations

Algorithm

‘Comment Post: Observation that any coin flip sequence selected by a pseudo-random-number-generator has the average number of heads converge to the probability of heads, suggesting the Strong Law of Large Numbers.

‘Comment Post: The average number of heads sampled at an exponential sequence of indices, to suggest the Strong Law of Large Numbers.

1 Set probability of success \( p \).
2 Set length of coin-flip sequence \( n \).
3 Initialize and fill length \( n \) array of coin flips.
4 Use vectorization to sum the cumulative sequence \( S_n \) of heads.
5 Use vectorization to compute an exponential sequence of indices for sampling.
6 Use vectorization to compute the average number of heads along the exponential sequence of indices.
7 Print the averages.
Scripts

\textbf{R}  \textit{R script for Strong Law.}

```r
p <- 0.5
n <- 1e+6
coinFlips <- (runif(n) <= p)
S <- cumsum(coinFlips)
tenpows <- 10^(1:6)
cat(S[tenpows]/tenpows , "\n")
```

\textbf{Octave}  \textit{Octave script for Strong Law.}

```octave
p = 0.5;
n = 1e+6;
coinFlips = rand(1, n) <= p;
S = cumsum(coinFlips);
tenpows = 10 .^ (1:6);
S(tenpows) ./ tenpows
```

\textbf{Perl}  \textit{Perl PDL script for Strong Law.}

```perl
$p = 0.5;
$n = 1e+6;
$coinFlips = random($n) <= $p;
$S = cumusumover($coinFlips);
$pows = zeros(6)->xlinvals( 1, 6 );
$tenpows = 10**$pows;
print $S( $tenpows - 1 ) / $tenpows, "\n"
```

\textbf{SciPy}  \textit{Scientific Python script for Strong Law.}

```python

```
import scipy

p = 0.5
n = 1e+6
coinFlips = scipy.random.random(n) <= p
S = scipy.cumsum(coinFlips, axis=0)

pows = scipy.arange(1, 6+1) # 6 entries
tenpows = 10**pows

print(scipy.take(S, tenpows-1)/scipy.array(tenpows, dtype=float))

Problems to Work for Understanding

1.
2.
3.
4.
Reading Suggestion:

References


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 October 20, 2017