Topics in
Probability Theory and Stochastic Processes
Steven R. Dunbar

The Weak Law of Large Numbers

Rating
Mathematicians Only: prolonged scenes of intense rigor.
Question of the Day

Consider a fair \((p = 1/2 = q)\) coin tossing game carried out for 1000 tosses. Explain in a sentence what the “law of averages” says about the outcomes of this game. Be as precise as possible.

---

Key Concepts

1. Markov's Inequality: Let \(X\) be a random variable taking only non-negative values. Then for each \(a > 0\)
   \[
   \mathbb{P} [X > a] \leq \mathbb{E} [X] / a;
   \]

2. Chebyshev's Inequality: Let \(X\) be a random variable. Then for \(a > 0\)
   \[
   \mathbb{P} [|X - \mathbb{E} [X]| \geq a] \leq \frac{\text{Var} [X]}{a^2}.
   \]

3. Weak Law of Large Numbers: For \(\epsilon > 0\)
   \[
   \mathbb{P}_n \left[ \left| \frac{S_n}{n} - p \right| > \epsilon \right] \to 0 \text{ as } n \to \infty
   \]
   and the convergence is uniform in \(p\).

4. Let \(f\) be a real function that is defined and continuous on the interval \([0, 1]\). Then
   \[
   \sup_{0 \leq x \leq 1} \left| f(x) - \sum_{k=0}^{n} f \left( \frac{k}{n} \right) \binom{n}{k} x^k (1 - x)^{n-k} \right| \to 0
   \]
   as \(n \to \infty\).
Vocabulary

1. The **Weak Law of Large Numbers** says that for $\epsilon > 0$

$$\mathbb{P}_n\left[\left| \frac{S_n}{n} - p \right| > \epsilon \right] \to 0 \text{ as } n \to \infty$$

and the convergence is uniform in $p$.

2. The polynomials

$$B_{n,k}(t) = \binom{n}{k} x^k (1 - x)^{n-k}$$

are called the *Bernstein polynomials*.

---

Mathematical Ideas

**Proof of the Weak Law Using Chebyshev’s Inequality**

**Proposition 1** (Markov’s Inequality). *Let $X$ be a random variable taking only non-negative values. Then for each $a > 0$*

$$\mathbb{P}[X \geq a] \leq \mathbb{E}[X]/a.$$  

*Proof.*

$$\mathbb{P}[X \geq a] = \mathbb{E}[I_{X \geq a}]$$

$$= \int_{x \geq a} d\mathbb{P}[\cdot]$$

$$\leq \int \frac{1}{a} d\mathbb{P}[\cdot]$$

$$\leq \frac{1}{a} \mathbb{E}[X]$$

$\blacksquare$
Proposition 2 (Chebyshev’s Inequality). Let \( X \) be a random variable. Then for \( a > 0 \)
\[
P \left[ |X - \mathbb{E}[X]| \geq a \right] \leq \frac{\text{Var}[X]}{a^2}.
\]
Proof. This immediately follows from Markov’s inequality applied to the non-negative random variable \( (X - \mathbb{E}[X])^2 \).

Theorem 3 (Weak Law of Large Numbers). For \( \epsilon > 0 \)
\[
\mathbb{P}_n \left[ \left| \frac{S_n}{n} - p \right| > \epsilon \right] \to 0 \text{ as } n \to \infty
\]
and the convergence is uniform in \( p \).

Remark. The notation \( \mathbb{P}_n [\cdot] \) indicates that we are considering a family of probability measures on the sample space \( \Omega \). The Weak Law establishes the convergence of the sequence of measures in a particular way.

Proof. The variance of the random variable \( S_n \) is \( np(1 - p) \). Rewrite the probability as the equivalent event:
\[
\mathbb{P}_n \left[ \left| \frac{S_n}{n} - p \right| > \epsilon \right] = \mathbb{P}_n \left[ |S_n - np| > n\epsilon \right].
\]

By Chebyshev’s inequality
\[
\mathbb{P}_n \left[ |S_n - np| > n\epsilon \right] \leq \frac{\text{Var}[S_n]}{(n\epsilon)^2} = \frac{p(1 - p)}{\epsilon^2} \frac{1}{n}.
\]
Since \( p(1 - p) \leq 1/4 \), the proof is complete.

Remark. This inequality demonstrates that
\[
\mathbb{P}_n \left[ \left| \frac{S_n}{n} - p \right| > \epsilon \right] = O(1/n)
\]
uniformly in \( p \).

Remark. Jacob Bernoulli originally proved the Weak Law of Large Numbers in 1713 for the special case when the \( X_i \) are binomial random variables. Bernoulli had to create an ingenious proof to establish the result, since Chebyshev’s inequality was not known at the time. The theorem then
became known as Bernoulli’s Theorem. Simeon Poisson proved a generalization of Bernoulli’s binomial Weak Law and first called it the Law of Large Numbers. In 1929 the Russian mathematician Aleksandr Khinchin proved the general form of the Weak Law of Large Numbers presented here. Many other versions of the Weak Law are known, with hypotheses that do not require such stringent requirements as being identically distributed, and having finite variance.

**Remark.** Another proof of the Weak Law of Large Numbers using moment generating functions is in Mathematical Finance/Central Limit Theorem

**Bernstein’s Proof of the Weierstrass Approximation Theorem**

**Theorem 4.** Let \( f \) be a real function defined and continuous on the interval \([0, 1]\). Then

\[
\sup_{0 \leq x \leq 1} \left| f(x) - \sum_{k=0}^{n} f \left( \frac{k}{n} \right) \cdot \binom{n}{k} x^k (1-x)^{n-k} \right| \to 0
\]

as \( n \to \infty \).

**Proof.**

1. Fix \( \epsilon > 0 \). Since \( f \) continuous on the compact interval \([0, 1]\) it is uniformly continuous on \([0, 1]\). Therefore there is an \( \eta > 0 \) such that \( |f(x) - f(y)| < \epsilon \) if \( |x - y| < \eta \).

2. The expectation \( \mathbb{E} [ f(S_n/n) ] \) can be expressed as a polynomial in \( p \):

\[
\mathbb{E} \left[ f \left( \frac{S_n}{n} \right) \right] = \sum_{k=0}^{n} f \left( \frac{k}{n} \right) \cdot \mathbb{P}_n \left[ S_n = k \right] = \sum_{k=0}^{n} f \left( \frac{k}{n} \right) \cdot \binom{n}{k} p^k (1-p)^{n-k}.
\]

3. By the Weak Law of Large Numbers, for the given \( \epsilon > 0 \), there is an \( n_0 \) such that

\[
\mathbb{P}_n \left[ \left| \frac{S_n}{n} - p \right| > \eta \right] < \epsilon.
\]

4. \[
\left| \mathbb{E} \left[ f \left( \frac{S_n}{n} \right) - f(p) \right] \right| = \left| \sum_{k=0}^{n} \left( f \left( \frac{k}{n} \right) - f(p) \right) \mathbb{P}_n \left[ S_n = k \right] \right|.
\]
5. Apply the triangle inequality to the right hand side and express in terms of two summations:

\[
\leq \sum_{|\frac{k}{n} - p| \leq \eta} \left( f \left( \frac{k}{n} \right) - f(p) \right) \mathbb{P}_n [S_n = k] + \\
\sum_{|\frac{k}{n} - p| > \eta} \left( |f \left( \frac{k}{n} \right)| + |f(p)| \right) \mathbb{P}_n [S_n = k]
\]

Note the second application of the triangle inequality on the second summation.

6. Now estimate the terms:

\[
\leq \sum_{|\frac{k}{n} - p| \leq \eta} \epsilon \mathbb{P}_n [S_n = k] + \sum_{|\frac{k}{n} - p| > \eta} 2 \sup_{0 \leq x \leq 1} |f(x)| \mathbb{P}_n [S_n = k]
\]

7. Finally, do the addition over the individual values of the probabilities over single values to re-write them as probabilities over events:

\[
= \epsilon \mathbb{P}_n \left[ \left| \frac{S_n}{n} - p \right| \leq \eta \right] + 2 \sup_{0 \leq x \leq 1} |f(x)| \mathbb{P}_n \left[ \left| \frac{S_n}{n} - p \right| > \eta \right]
\]

8. Now apply the Weak Law to the second term to see that:

\[
\left| \mathbb{E} \left[ f \left( \frac{S_n}{n} \right) - f(p) \right] \right| < \epsilon + 2\epsilon \sup_{0 \leq x \leq 1} |f(x)|.
\]

This shows that \( \left| \mathbb{E} \left[ f \left( \frac{S_n}{n} \right) - f(p) \right] \right| \) can be made arbitrarily small, uniformly with respect to \( p \), by picking \( n \) sufficiently large.

\(\Box\)

Remark. The polynomials

\[B_{n,k}(t) = \binom{n}{k} x^k (1 - x)^{n-k}\]

are called the Bernstein polynomials. The Bernstein polynomials have several useful properties:
1. \( B_{i,n}(t) = B_{n-i,n}(1 - t) \)

2. \( B_{i,n}(t) \geq 0 \)

3. \( \sum_{i=0}^{n} B_{i,n}(t) = 1 \) for \( 0 \leq t \leq 1 \).

**Corollary 1.** A polynomial of degree \( n \) uniformly approximating the continuous function \( f(x) \) on the interval \( [a,b] \) is

\[
\sum_{k=0}^{n} f \left( a + (b - a) \frac{k}{n} \right) \left( \frac{n}{k} \right) \left( \frac{x - a}{b - a} \right)^k \left( \frac{b - x}{b - a} \right)^{n-k}
\]

**Sources**

This section is adapted from: *Heads or Tails*, by Emmanuel Lesigne, Student Mathematical Library Volume 28, American Mathematical Society, Providence, 2005, Chapter 5, [3].

---

**Problems to Work for Understanding**

1. Suppose \( X \) is a continuous random variable with mean and variance both equal to 20. What can be said about \( \mathbb{P}[0 \leq X \leq 40] \)?

2. Suppose \( X \) is an exponentially distributed random variable with mean \( \mathbb{E}[X] = 1 \). For \( x = 0.5, 1 \), and 2, compare \( \mathbb{P}[X \geq x] \) with the Markov inequality bound.

3. Suppose \( X \) is a Bernoulli random variable with \( \mathbb{P}[X = 1] = p \) and \( \mathbb{P}[X = 0] = 1 - p = q \). Compare \( \mathbb{P}[X \geq 1] \) with the Markov inequality bound.

4. Let \( X_1, X_2, \ldots, X_{10} \) be independent Poisson random variables with mean 1. First use the Markov Inequality to get a bound on \( \mathbb{P}[X_1 + \cdots + X_{10} > 15] \). Next find the exact probability that \( \mathbb{P}[X_1 + \cdots + X_{10} > 15] \) using that the fact that the sum of independent Poisson random variables with parameters \( \lambda_1, \lambda_2 \) is again Poisson with parameter \( \lambda_1 + \lambda_2 \).
5. Consider a fair \((p = 1/2 = q)\) coin tossing game carried out for \(n = 100\) tosses. Calculate the exact value
\[
P_n \left[ \left| \frac{S_n}{n} - p \right| > 1/10 \right]
\]
and compare it to the estimates in the proof of the Weak Law of Large Numbers.

6. Calculate the Bernstein polynomial approximation of \(\sin(\pi x)\) of degree 1, 2, and 3 and plot the graphs of \(\sin(\pi x)\) and the approximations.

7. Calculate the Bernstein polynomial approximation of \(\cos(\pi x)\) of degree 1, 2, and 3 and plot the graphs of \(\cos(\pi x)\) and the approximations.

8. Calculate the Bernstein polynomial approximation of \(\exp(\pi x)\) of degree 1, 2, and 3 and plot the graphs of \(\exp(\pi x)\) and the approximations.

---

**Reading Suggestion:**

**References**


Outside Readings and Links:

1. Virtual Laboratories in Probability and Statistics / Binomial


3. Wikipedia, Weak Law of Large Numbers

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 May 25, 2011