Bulletin of the Institute of Mathematics
Academia Sinica (New Series)
Vol. **7** (2012), No. 3, pp. 417-422

AN EXACT WEAK LAW OF LARGE NUMBERS

ANDRÉ ADLER

Department of Applied Mathematics, Illinois Institute of Technology, Chicago, Illinois, 60616.
E-mail: adler@iit.edu

Abstract

This paper explores a Weighted Exact Weak Law, where the classical Weak Law fails and the corresponding Strong Law also fails. This type of result comes from the Fair Games problem and is associated with the St Petersburg Game.

1. Introduction

In this paper we observe the limiting behaviour of a weighted sum from a sequence of independent random variables. The point here is that this is the only way to obtain an Exact Law of Large Numbers for this particular sequence of random variables. We want to establish a nonzero limit between our partial sums and a sequence of constants. We can always divide by a large enough sequence so that the ratio between our partial sums and the norming sequence converges to zero, but that's not fair. We want the ratio to approach a nonzero constant since the sum of the random variables can be considered winning from a gambling game and in order for both parties to partake in this game the ratio should converge to one. That is why this type of result also goes by the name of the 'Fair Games Problem' see [5], pages 248-253. The most famous of these is the St. Petersburg Game itself. Furthermore the only Law of Large Numbers that we can establish is a weighted Weak Law. Both the Classical Weak Law, $(X_1 + \cdots + X_n)/n$ will not converge to a constant, nor will the weighted Strong Law converge, even though $X_n \xrightarrow{P} 0$.

Received July 7, 2009 and in revised form March 15, 2011.

AMS Subject Classification: 60F05, 60F15.

Key words and phrases: Almost sure convergence, weak law of large numbers, strong law of large numbers, exact laws.

In the Exact Strong Law setting we can find a Weighted Strong Law and yet the Classical Strong Law fails. For example if X_n are i.i.d. with common distribution $xP\{X > x\} \sim a(\lg x)^\alpha$, where $\alpha > -1$, then the only type of Strong Law that holds is

$$\lim_{N\to\infty} \frac{\sum_{n=1}^{N} \frac{(\lg n)^{b-\alpha-2}}{n} X_n}{(\lg N)^b} = \frac{a}{(\alpha+1)b} \quad \text{almost surely.}$$

for any $b > 0$, see Example 2 from [2]. In this setting we can establish many Weak Laws including the classical nonweighted Weak Law. Moreover, from [1] we can find sequences $\{b_N\}$ such that

$$\frac{\sum_{n=1}^{N} n^\alpha X_n}{b_N} \xrightarrow{P} 1$$

for $\alpha > -1/2$. This was later extended to $\alpha > -1$, as one would expect.

However in the Exact Weak Law situation we are now forced to look at independent but no longer identically distributed random variables. That comes from [3], since if we were observing i.i.d. random variables then a classical Weak Law implies a Weighted Strong Law. As in the Exact Strong Law setting, these random variables have a similar structure. Just like the St. Petersburg Game, these random variables barely do not have a finite first moment. Similarly, the weights are of the type $a_n = L(n)/n$ where $L(x)$ is a slowly varying function.

The random variables I will explore are such that $P\{X_n \le x\} = 1 - \frac{1}{x+n}$ where $x > 0$. Thus $P\{X_n = 0\} = 1 - \frac{1}{n}$ and $f_{X_n}(x) = (x+n)^{-2} I(x > 0)$. In terms of notation we will let $\lg x = \log\,(\max\{e, x\})$ and $\lg_2 x = \lg(\lg x)$. Also note that the constant C will be used as a generic bound that is not necessarily the same in each appearance.

2. The Weak Law

It's not hard to show that $X_n \xrightarrow{P} 0$ and yet $(X_1 + \cdots + X_n)/n$ does not converge to zero in probability. Thus we must explore weighted partial sums in order to establish a law of large numbers. But once again we need to be exact in our weights.

Theorem 1. *If* $P\{X_n \le x\} = 1 - \frac{1}{x+n}$ *where* $x > 0$*, then*

$$\frac{\sum_{n=1}^{N} \frac{1}{n} X_n}{\lg N \lg_2 N} \xrightarrow{P} 1.$$

Proof. We will establish all three parts of the Degenerate Convergence Criterion, see [4], page 356. Let $a_n = 1/n$ and $b_N = \lg N \lg_2 N$.

Since $EX_n^2 I(X_n \le c) = \int_0^c \frac{x^2 dx}{(x+n)^2} < c$, we have

$$\begin{aligned} b_N^{-2} \sum_{n=1}^{N} a_n^2 E X_n^2 I(a_n X_n \le b_N) &\le C b_N^{-2} \sum_{n=1}^{N} a_n^2 b_N / a_n \\ &= C b_N^{-1} \sum_{n=1}^{N} a_n \\ &= \frac{C \sum_{n=1}^{N} \frac{1}{n}}{\lg N \lg_2 N} \\ &\le \frac{C}{\lg_2 N} \to 0. \end{aligned}$$

Next, we have for all $\epsilon > 0$

$$\begin{aligned} \sum_{n=1}^{N} P\{a_n X_n > \epsilon b_N\} &= \sum_{n=1}^{N} \left[1 - F_{X_n}(\epsilon b_N / a_n) \right] \\ &= \sum_{n=1}^{N} \left[\frac{1}{\frac{\epsilon b_N}{a_n} + n} \right] \\ &= \sum_{n=1}^{N} \left[\frac{1}{\epsilon n \lg N \lg_2 N + n} \right] \\ &= \frac{1}{\epsilon \lg N \lg_2 N + 1} \sum_{n=1}^{N} \frac{1}{n} \\ &\le \frac{C \lg N}{\epsilon \lg N \lg_2 N + 1} \to 0. \end{aligned}$$

Finally, we investigate the truncated mean. We have

$$E X_n I(X_n \le c) = \int_0^c \frac{x dx}{(x+n)^2} = \lg\left(\frac{c+n}{n}\right) + \frac{n}{c+n} - 1.$$

Thus

$$\begin{aligned} EX_nI(X_n \le b_N/a_n) &= \lg\left(\frac{n\lg N\lg_2 N + n}{n}\right) + \frac{n}{n\lg N\lg_2 N + n} - 1 \\ &= \lg\big(\lg N\lg_2 N + 1\big) + \frac{1}{\lg N\lg_2 N + 1} - 1 \\ &\sim \lg_2 N. \end{aligned}$$

whence

$$\begin{aligned} b_N^{-1}\sum_{n=1}^{N} a_nEX_nI(a_nX_n \le b_N) &\sim \frac{1}{\lg N\lg_2 N}\sum_{n=1}^{N}\frac{\lg_2 N}{n} \\ &= \frac{1}{\lg N}\sum_{n=1}^{N}\frac{1}{n} \to 1 \end{aligned}$$

completing this proof. $\square$

3. Almost sure Results

The Weak Law established in the last section cannot be extended to a Strong Law.

Theorem 2. *If* $P\{X_n \le x\} = 1 - \frac{1}{x+n}$ *where* $x > 0$, *then*

$$\liminf_{N\to\infty} \frac{\sum_{n=1}^{N}\frac{1}{n}X_n}{\lg N\lg_2 N} \le 1 \quad \text{almost surely.}$$

and

$$\limsup_{N\to\infty} \frac{\sum_{n=1}^{N}\frac{1}{n}X_n}{\lg N\lg_2 N} = \infty \quad \text{almost surely.}$$

Proof. As in the last proof, let $a_n = 1/n$ and $b_N = \lg N\lg_2 N$. The lower limit is a direct result of Theorem 1. As for the upper limit, let $M > 0$, then

$$\begin{aligned} \sum_{n=1}^{\infty} P\{a_nX_n/b_n > M\} &= \sum_{n=1}^{\infty} P\{X_n > Mb_n/a_n\} \\ &= \sum_{n=1}^{\infty}\Big[1 - F_{X_n}(Mb_n/a_n)\Big] \\ &= \sum_{n=1}^{\infty}\frac{1}{Mn\lg n\lg_2 n + n} \end{aligned}$$

$$
\begin{aligned}
&> C\sum_{n=1}^{\infty}\frac{1}{n\lg n\lg_2 n}\\
&= \infty.
\end{aligned}
$$

Thus

$$\limsup_{n\to\infty}\frac{a_nX_n}{b_n}=\infty \quad \text{almost surely}$$

hence

$$\limsup_{N\to\infty}\frac{\sum_{n=1}^{N}a_nX_n}{b_N}\geq\limsup_{N\to\infty}\frac{a_NX_N}{b_N}=\infty \quad \text{almost surely}$$

which completes the proof. $\square$

4. Discussion

The distribution used here is a continuous version of the Zipf-Mandelbrot law also known as the Pareto-Zipf law. It's sometimes referred to as the the Yule-Simon distribution. Zipf law states "that the frequency of use of the nth-most-frequently-used word in any natural language is inversely proportional to n." This has been show to apply to music as well as literature. Mandelbrot extended this to include an extra term, n, as in this paper.

These types of distributions that don't have a finite mean are sometimes reserved for such phenomenon as solar flares and wars, see [7]. But the type of distribution examined in this paper is mainly used in linguistics and in ecological field studies. Besides being used to model the ranking of words in various languages it has also been used to establish the number of species according to their abundance, see [6].

The next step is to establish a limiting result for a more general distribution. Instead of observing $P\{X_n > x\} = 1/(x+n)$ we can consider a sequence of independent random variables where

$$P\{X_n\leq x\}=1-\frac{1}{x+c_n}$$

for a general sequence $\{c_n, n\geq 1\}$. But the problem is that the results here are so delicate that we really need to have a specific structure on our sequence $\{c_n, n\geq 1\}$ in order to obtain an Exact Weak Law. The harmonic sequence does play a pivotal role in this type of limit theorem, you can see it in the sequence of coefficients in our Weak Law where we have no choice but to set $a_n = 1/n$.

References

1. A. Adler, Generalized one-sided laws of the iterated logarithm for random variables barely with or without finite mean, *J. Theor. Probab.*, **4** (1990), 587-597.

2. A. Adler, Exact strong laws, *Bull. Inst. Math. Acad. Sinica*, **28** (2000), 141-166.

3. A. Adler and R. Wittmann, Stability of sums of independent random variables, *Stoch. Process. Appl.*, **52** (1994), 179-182.

4. Y. S. Chow and H. Teicher, *Probability Theory: Independence, Interchangeability, Martingales*, 3^{rd} ed., Springer-Verlag, New York, 1997.

5. W. Feller, *An Introduction to Probability Theory and Its Applications*, Vol 1, 3^{rd} ed., John Wiley, New York, 1968.

6. D. Mouillot and A. Lepretre, Introduction of relative abundance distribution (RAD) indices, estimated from the rank-frequency diagrams (RFD), to assess changes in community diversity, *Environmental Monitoring and Assessment*, **63** (2000), 279-295.

7. M. E. J. Newman, Power laws, Pareto distributions and Zipf's law, *Contemporary Physics* **46** (2005), 323-351.