Proof of the Hartman-Wintner Law of the Iterated Logarithm

Brian Myungjin Choi

A THESIS

in

Mathematics

Presented to the Faculties of the University of Pennsylvania in Partial

Fulfillment of the Requirements for the Degree of Master of Arts

2012

Supervisor of Thesis

Graduate Group Chairman

In this paper, I seek to present a proof for the Hartman-Wintner law of iterated logarithm. The law states that for any random walk, $\{S_n\}$ with the increment of zero mean and finite variance, σ^2, the following holds almost surely:

$$\limsup_{n\to\infty} \frac{S_n}{\sqrt{2\sigma^2 n \log\log(n)}} = 1$$

The proof presented in this paper requires the use of Skorokhod embedding theorem, which is different from the original proof for Hartman-Wintner law of iterated logarithm, and along the way, I will also prove the law of iterated logarithm for Brownian motion. The proofs for both Skorokhod embedding theorem and the law of iterated logarithm make use of properties of Brownian motion, so in this paper, I also included how to construct a standard Brownian motion as well as proofs for Skorokhod embedding theorem and law of iterated logarithm. For the construction of Brownian motion, I used Haar wavelet approach and closely followed the corresponding chapter from "Stochastic Calculus and Financial Applications" by M.Steele, and for proofs of Skorokhod embedding theorem and law of iterated logarithm, I closely followed the corresponding chapters from "Brownian Motion" by P. Morters and Y. Peres.

1 Introduction

Let's assume you are playing rock, paper, scissors with your friend. In each game, you have $\frac{1}{3}$ chance of winning, $\frac{1}{3}$ chance of losing, and $\frac{1}{3}$ chance of a draw. If you win, you receive a dollar from your friend, and if you lose, you give a dollar to your friend, and nothing happens in the case of a draw game. Let X_n be your payoff at nth game, then $\{X_N\}$ is a sequence of i.i.d. random variables that can take a value of 1 or 0 or -1, each with equal probability, $\frac{1}{3}$, then if we define S_n as

$$S_n = X_1 + X_2 + ... + X_n = \sum_{j=1}^{n} X_j$$

then S_n represents your total payoff at n. We are interested in the behavior of S_n in the long run.

It is a well-known consequence from Hewitt-Savage 0-1 law that if $X_1, X_2, ..., X_n$ are i.i.d. non-degenerate real random variables of with a distribution that is symmetric about zero, and $S_n = \sum_{j=1}^{n} X_j$, then

$$-\infty = \liminf_{n\to\infty} S_n < \limsup_{n\to\infty} S_n = \infty$$

Now, the question is can we do better than that? In particular, is it possible to find some tractable, closed-form function, $\phi(n)$ such that

$$\limsup_{n\to\infty} \frac{S_n}{\phi(n)} = \liminf_{n\to\infty} \frac{S_n}{-\phi(n)} = 1 \qquad \text{almost surely}$$

The answer is yes, and such $\phi(n)$ can be found by Hartman-Wintner law of iterated logarithm. In the case of repeated rock, paper, scissors games shown above, the answer would be $\phi(n) = \sqrt{\frac{4}{3} n \log\log(n)}$, so we get

$$\limsup_{n\to\infty} \frac{S_n}{\sqrt{\frac{4}{3} n \log\log(n)}} = \liminf_{n\to\infty} \frac{S_n}{-\sqrt{\frac{4}{3} n \log\log(n)}} = 1 \qquad \text{almost surely}$$

The real power of the theorem is that it can be generalized to any nondegenerate symmetric random walk in which each increment has zero mean and finite variance. Hartman-Wintner law of iterated logarithm states that for any random walk S_n with increments, $S_n - S_{n-1}$ of zero mean and finite variance σ^2, we get

$$\limsup_{n\to\infty} \frac{S_n}{\sqrt{2\sigma^2 n \log\log(n)}} = 1 \qquad \text{almost surely}$$

The primary aim of this paper is to present a proof Hartman-Winter law of iterated logarithm, and to do so, we will prove another powerful and related theorem called Skorokhod embedding theorem. The proofs for both theorems make use of a stochastic process called Brownian Motion, so we will start by defining what Brownian Motion is and proving its existence by explicitly constructing one. (Also, all random variables mentioned will be assumed to be real-valued).

2 Construction of a Brownian Motion

A standard Brownian motion, $\{B_t : 0 \leq t\}$ is a real-valued continuous-time stochastic process satisfying the following four properties:

- $B_0 = 0$
- It has independent increments. In other words, for any finite set, $0 \leq t_1 < t_2 < t_3 ... < t_n$, $B_{t_2} - B_{t_1}$, $B_{t_3} - B_{t_2}$, ..., $B_{t_n} - B_{t_{n-1}}$ are independent.
- For any $0 \leq s \leq t$, the increment $B_t - B_s$ has the Gaussian distribution with mean 0 and variance $t - s$.
- B_t is a continuous function of t with probability one.

We will show that such a process exists by explicitly constructing one. While there are other methods of constructing a standard Brownian motion, we will make use of Haar Wavelets to construct one. This part will follow chapter 3 from "Stochastic Calculus and Financial Applications" by M.Steele very closely[1]. The idea is to construct a standard Brownian motion on $[0, 1]$, so that for each $1 \leq n < \infty$, we can have an independent standard Brownian motion, $B_s^{(n)}$ for $s \in [0, 1]$, and for any $0 \leq t < \infty$, we can get B_t by

$$B_t = B_{t-n}^{(n+1)} + \sum_{k=1}^{n} B_1^{(k)} \quad \text{for} \quad t \in [n, n+1)$$

The Haar wavelet is a sequence of functions proposed by Alfred Haar in 1909. To construct a Haar wavelet, we first need to define a "mother wavelet," $H(t)$ as

$$H(t) = \begin{cases} 1 & \text{for } 0 \leq t < \frac{1}{2}, \\ -1 & \text{for } \frac{1}{2} \leq t \leq 1, \\ 0 & \text{otherwise} \end{cases}$$

then from this $H(t)$, we define a sequence of functions

$$\psi_{j,k}(t) = H(2^j t - k) \qquad \text{for } \ 0 \leq j, \ \ 0 \leq k < 2^j$$

this sequence of functions, $\{\psi_{j,k}\}$, is called Haar wavelet. Now, from this Haar wavelet, we define $H_0, H_1, H_2, ...$ as

$$H_0(t) = 1$$

$$H_n(t) = 2^{\frac{j}{2}} \psi_{j,k}(t) \qquad \text{for } \ n = 2^j + k \ \text{ where } \ j \geq 0 \ \text{ and } \ 0 \leq k < 2^j$$

($\{H_n\}$ actually forms a complete orthonormal sequence for $L^2[0,1]$, and we will use this fact later in this chapter. For the detailed proof of how this forms a complete orthonormal sequence, you can look at page 135 of "A Basic Course in Probability" by Rabindra Nath Bhattacharya, Edward C. Waymire). To list the first few H_n for $n \geq 1$, we have

$$H_1(t) = \begin{cases} 1 & \text{for } 0 \leq t < \frac{1}{2}, \\ -1 & \text{for } \frac{1}{2} \leq t \leq 1, \\ 0 & \text{otherwise} \end{cases}$$

$$H_2(t) = \begin{cases} \sqrt{2} & \text{for } 0 \leq t < \frac{1}{4}, \\ -\sqrt{2} & \text{for } \frac{1}{4} \leq t \leq \frac{1}{2}, \\ 0 & \text{otherwise} \end{cases} \qquad H_3(t) = \begin{cases} \sqrt{2} & \text{for } \frac{1}{2} \leq t < \frac{3}{4}, \\ -\sqrt{2} & \text{for } \frac{3}{4} \leq t \leq 1, \\ 0 & \text{otherwise} \end{cases}$$

$$H_4(t) = \begin{cases} 2 & \text{for } 0 \leq t < \frac{1}{8}, \\ -2 & \text{for } \frac{1}{8} \leq t \leq \frac{1}{4}, \\ 0 & \text{otherwise} \end{cases} \qquad H_5(t) = \begin{cases} 2 & \text{for } \frac{1}{4} \leq t < \frac{3}{8}, \\ -2 & \text{for } \frac{3}{8} \leq t \leq \frac{1}{2}, \\ 0 & \text{otherwise} \end{cases}$$

$$H_6(t) = \begin{cases} 2 & \text{for } \frac{1}{2} \leq t < \frac{5}{8}, \\ -2 & \text{for } \frac{5}{8} \leq t \leq \frac{3}{4}, \\ 0 & \text{otherwise} \end{cases} \qquad H_7(t) = \begin{cases} 2 & \text{for } \frac{3}{4} \leq t < \frac{7}{8}, \\ -2 & \text{for } \frac{7}{8} \leq t \leq 1, \\ 0 & \text{otherwise} \end{cases}$$

Now, we define another sequence of functions, $\{\Psi_{j,k}(t)\}$ by

$$\Psi_{j,k}(t) = \int_0^t \psi_{j,k}(u)du$$

Just like the way we constructed $\{\psi_{j,k}(t)\}$, this sequence can also be more conveniently represented as the wavelet by defining a mother wavelet first and constructing other functions from that mother wavelet. We will denote the mother wavelet by $\Psi(t)$ and define it as following:

$$\Psi(t) = \begin{cases} 2t & \text{for } 0 \leq t < \frac{1}{2}, \\ 2(1-t) & \text{for } \frac{1}{2} \leq t \leq 1, \\ 0 & \text{otherwise} \end{cases}$$

and from this $\Psi(t)$, we define a sequence of functions

$$\Psi_{j,k}(t) = \Psi(2^j t - k) \qquad \text{for } \ 0 \leq j, \ \ 0 \leq k < 2^j$$

Now, let's define $\{\Delta_n(t)\}$ as $\Delta_{2^j+k} = \Psi_{j,k}(t)$ (note that $|\Delta_n(t)| \leq 1$ from construction), then by defining $\{\lambda_n\}$ as $\lambda_0 = 1$, and for $n \geq 1$

$$\lambda_n = \frac{2^{-j/2}}{2} \qquad \text{where } \ n \geq 1 \ \text{ and } \ n = 2^j + k \ \text{ with } \ 0 \leq k < 2^j$$

we get the following expression

$$\int_0^t H_n(u)du = \lambda_n \Delta_n(t)$$

Now, we have all the necessary raw materials to construct a standard Brownian motion for $t \in [0,1]$. We will do so by proving the following proposition.

Proposition 2.1. *Let $\{Z_n : 0 \leq n < \infty\}$ be a sequence of independent Gaussian random variables with mean 0 and variance 1, then the series defined by*

$$X_t = \sum_{n=0}^{\infty} \lambda_n Z_n \Delta_n(t)$$

is a standard Brownian motion for $0 \leq t \leq 1$.

We will prove the above proposition by showing that X_t satisfies the required properties of a standard Brownian motion. First, we will start by proving that X_t converges uniformly on $[0,1]$ with probability 1. To do so, we will first prove the following lemma.

Lemma 2.2. *Let $\{Z_n : 0 \leq n < \infty\}$ be a sequence of independent Gaussian random variables with mean 0 and variance 1, then there is a random variable C that is finite with probability one and*

$$|Z_n| \leq C\sqrt{\log n} \qquad \textit{for all} \;\; n \geq 2$$

Proof. This is actually a natural consequence from the density of a standard Gaussian random variable and the Borel-Cantelli lemma. Note that for $x \geq 1$

$$P(|Z_n| \geq x) = \frac{2}{\sqrt{2\pi}} \int_0^\infty \exp(-u^2/2)du \leq \sqrt{\frac{2}{\pi}} \int_x^\infty u\exp(-u^2/2)du = \exp(-x^2/2)\sqrt{\frac{2}{\pi}}$$

Thus, for any $\alpha > 1$, we have

$$P(|Z_n| \geq \sqrt{2\alpha \log n}) \leq \exp(-\alpha \log n)\sqrt{\frac{2}{\pi}} = n^{-\alpha}\sqrt{\frac{2}{\pi}}$$

Note that for $\alpha > 1$, we have

$$\sum_{n=1}^{\infty} n^{-\alpha} < \infty$$

thus, using the Borel-Cantelli lemma, we get

$$P(|Z_n| \geq \sqrt{2\alpha \log n} \quad \text{for infinitely many } n) = 0$$

and, therefore, the random variable defined by

$$\sup_{2 \leq n < \infty} \frac{|Z_n|}{\sqrt{\log n}} = C$$

is finite with probability one. □

Using the above lemma, we will now prove that X_t converges uniformly on $[0, 1]$ with probability one. From construction of $\{\Delta_n\}$, for any $0 \leq x \leq 1$, we have

$\Delta_n(x) = 0$ for all but one value of n in the interval $[2^j, 2^{j+1})$, and for any $n \in [2^j, 2^{j+1})$, we have $\log n < j+1$. Then from Lemma 2.2, we can find a random variable C, which is finite with probability one and satisfies $|Z_n| \leq C\sqrt{\log n}$ for all $n \geq 2$, so for any $J \geq 1$, if we let $M \geq 2^J$, we get

$$\sum_{n=M}^{\infty} \lambda_n |Z_n| \Delta_n(t) \leq C \sum_{n=M}^{\infty} \lambda_n \sqrt{\log n} \Delta_n(t)$$

$$\leq \sum_{j=J}^{\infty} \sum_{k=0}^{2^j-1} \frac{2^{-j/2}}{2} \sqrt{j+1} \Delta_{2^j+k}(t) \leq C \sum_{j=J}^{\infty} \frac{2^{-j/2}}{2} \sqrt{j+1}$$

and note that $\sum_{j=1}^{\infty} \frac{2^{-j/2}}{2} \sqrt{j+1} < \infty$. Therefore, we have

$$\lim_{J \to \infty} C \sum_{j=J}^{\infty} \frac{2^{-j/2}}{2} \sqrt{j+1} = 0$$

Now, let's recall the definition of X_t, which was the following

$$X_t = \sum_{n=0}^{\infty} \lambda_n Z_n \Delta_n(t)$$

where λ_n is just a real constant, and $\Delta_n(t)$ is a bounded continuous function on $[0, 1]$. Lemma 2.2 shows that X_t converges uniformly on $[0, 1]$ with probability one, and if an infinite sum of continuous functions on $[0, 1]$ uniformly converges on $[0, 1]$, then that infinite sum must also be a continuous function on $[0, 1]$. Thus, the fact that X_t converges uniformly on $[0, 1]$ with probability one implies that the paths of the process $\{X_t : 0 \leq t \leq 1\}$ are continuous with probability one.

Next, we will show that X_t has independent increments by proving that X_t satisfies an equivalent condition assuming X_t is Gaussian: $Cov(X_s, X_t) = \min(s, t)$ for all $0 \leq s, t \leq T$.

Lemma 2.3. *If a process* $\{X_t : 0 \leq t \leq T\}$ *is Gaussian and has* $E(X_t) = 0$ *for all* $0 \leq t \leq T$ *and if*

$$Cov(X_s, X_t) = \min(s,t) \quad \textit{for all} \;\; 0 \leq s,t \leq T$$

then $\{X_t\}$ *has independent increments, and if this process has continuous paths and* $X_0 = 0$*, then it is a standard Brownian motion on* $[0, T]$*.*

Proof. Assuming the first part is true, the second part of this lemma is a natural consequence from the definition of a standard Brownian motion on $[0, T]$, so it suffices to prove only the first part. Recall that the coordinates of a multivariate Gaussian vector are independent if and only if the covariance matrix is a diagonal matrix. Thus, to prove the first part, it suffices to show that for any finite set, $0 \leq t_1 < t_2 < t_3 ... \leq t_n \leq T$, the vector of the process increments

$$(X_{t_2} - X_{t_1}, X_{t_3} - X_{t_2}, X_{t_4} - X_{t_3}, ..., X_{t_n} - X_{t_{n-1}})$$

has a diagonal covariance matrix. Then it follows that, $X_{t_2} - X_{t_1}, X_{t_3} - X_{t_2}, X_{t_4} - X_{t_3}, ..., X_{t_n} - X_{t_{n-1}}$ are independent random variables. To show that the covariance matrix is a diagonal matrix, note that for $i < j$, we have

$$E[(X_{t_i} - X_{t_{i-1}})(X_{t_j} - X_{t_{j-1}})] = E[X_{t_i} X_{t_j}] - E[X_{t_i} X_{t_{j-1}}] - E[X_{t_{i-1}} X_{t_j}] + E[X_{t_{i-1}} X_{t_{j-1}}]$$

$$= t_i - t_i - t_{i-1} + t_{i-1} = 0$$

and the diagonal entries would be $t_1, t_2, ..., t_n$, so the resulting matrix would be a diagonal matrix. $\square$

Note that from our construction of X_t, it's obvious that $X_0 = 0$ as $\Delta_n(0) = 0$ for all n. Furthermore, we already proved in Lemma 2.2 that X_t has continuous paths. So using Lemma 2.3, we only have to show that $Cov(X_s, X_t) = \min(s,t)$ for all $0 \leq s, t \leq 1$ and for any finite set $0 \leq t_1 < t_2 < t_3 ... \leq t_n \leq 1$, $(X_{t_1}, X_{t_2}, ..., X_{t_n})$ has the multivariate Gaussian distribution to prove that X_t is a standard Brownian motion for $t \in [0,1]$. The first part can be done by simple calculations:

$$E(X_s X_t) = E\left[\sum_{n=0}^{\infty} \lambda_n Z_n \Delta_n(s) \sum_{m=0}^{\infty} \lambda_m Z_m \Delta_m(t)\right] = \sum_{n=0}^{\infty} \lambda_n^2 \Delta_n(s) \Delta_n(t)$$

$$= \sum_{n=0}^{\infty} \int_0^s H_n(u)du \int_0^t H_n(u)du = \min(s,t)$$

in which the last part of equality follows from Parseval's identity, and the fact that $\{H_n\}$ is a complete orthonormal sequence for $L^2[0,1]$. (as illustrated below)

$$\min(s,t) = \int_0^1 1_{[0,s]}(x) 1_{[0,t]}(x) dx = \sum_{n=0}^{\infty} \langle 1_{[0,s]}, H_n \rangle \ \langle 1_{[0,t]}, H_n \rangle$$

$$= \sum_{n=0}^{\infty} \int_0^s H_n(u)du \int_0^t H_n(u)du$$

For the second part, it suffices to calculate directly the multivariate characteristic function of $(X_{t_1}, X_{t_2}, ..., X_{t_n})$ and check whether the resulting characteristic function matches that of a multivariate Gaussian with mean zero and covariance matrix, $\Sigma = (\min(t_i, t_j))$

$$E\left[\exp(i \sum_{j=1}^{n} \theta_j X_{t_j})\right] = E\left[\exp(i \sum_{j=1}^{n} \theta_j \sum_{k=0}^{\infty} \lambda_k Z_k \Delta_k(t_j))\right]$$

$$= \prod_{k=0}^{\infty} E\left[\exp(i \lambda_k Z_k \sum_{j=1}^{n} \theta_j \Delta_k(t_j))\right] = \prod_{k=0}^{\infty} \exp\left(-\frac{1}{2} \lambda_k^2 (\sum_{j=1}^{n} \theta_j \Delta_k(t_j))^2\right)$$

$$= \exp\left(-\frac{1}{2}\sum_{k=0}^{\infty}\lambda_k^2(\sum_{j=1}^{n}\theta_j\Delta_k(t_j))^2\right) = \exp\left(-\frac{1}{2}\sum_{j=1}^{n}\sum_{k=1}^{n}\theta_j\theta_k\min(t_j,t_k)\right)$$

for the last equality, note that

$$\sum_{k=0}^{\infty}\lambda_k^2(\sum_{j=1}^{n}\theta_j\Delta_k(t_j))^2 = \sum_{k=0}^{\infty}\lambda_k^2\left(\sum_{i=1}^{n}\sum_{j=1}^{n}\theta_i\theta_j\Delta_k(t_i)\Delta_k(t_j)\right)$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n}\theta_i\theta_j\int_0^{t_i}H_n(u)du\int_0^{t_j}H_n(u)du = \sum_{i=1}^{n}\sum_{j=1}^{n}\theta_i\theta_j\min(t_i,t_j)$$

thus, we get the desired equality

$$\exp\left(-\frac{1}{2}\sum_{k=0}^{\infty}\lambda_k^2(\sum_{j=1}^{n}\theta_j\Delta_k(t_j))^2\right) = \exp\left(-\frac{1}{2}\sum_{j=1}^{n}\sum_{k=1}^{n}\theta_j\theta_k\min(t_j,t_k)\right)$$

and the last expression is the characteristic function of a multivariate Gaussian with mean zero and covariance matrix, $\Sigma = (\min(t_i,t_j))$, and, therefore, X_t is indeed a standard Brownian motion.

3 Law of Iterated Logarithm

As complete as the previous chapter looks in constructing a standard Brownian Motion, it still misses one of its key facets: a filtration. A filtration is a family of σ-fields, denoted by $\mathcal{F}_t$, such that $\mathcal{F}_s \subset \mathcal{F}_t$ for $s \leq t$. In our case, we would like to make it right-continuous, meaning $\bigcap_{s:s>t} \mathcal{F}_t = \mathcal{F}_s$. Basically, to complete a construction of a Brownian motion, we will need to define Ω and $\mathcal{F}_t$, so that each B_t is a measurable function from a probability space, $(\Omega, \mathcal{F}_t, P)$ into $[0, 1]$. For Ω, it would be a set of all continuous real functions

$$\Omega = \{\omega(t) : [0, \infty) \to (-\infty, \infty)\}$$

and for $\mathcal{F}_t$, it needs to satisfy that for each Borel set, B, of $\mathbb{R}$, we must have $\{\omega : B_t(\omega) \in B\} \subset \mathcal{F}_t$. The obvious choice for $\mathcal{F}_t$ would be $\sigma(\{B_s : s \leq t\})$, namely, a family of σ-fields generated by Brownian motion up to time t. (Although this filtration is not right-continuous, it is still possible to construct a right-continuous filtration, which is same as $\sigma(\{B_s : s \leq t\})$ up to measure zero, but we will omit the details. For the detailed proof, please refer to the chapter on Brownian motion on "Probability: Theory and Examples" by Richard Durrett).

Note that from this construction, each B_t would be equipped with a different $(\Omega, \mathcal{F}_t)$, as each filtration, $\mathcal{F}_t$, would be different for different values of t, and using the filtration, we can define a new concept: **stopping time**, which is a random variable, T, taking values on $[0, \infty]$ such that for all $t \geq 0$, we have $\{T < t\} \in \mathcal{F}_t$.

Finally, we have finished the construction of a standard Brownian motion. Arguably the most important and interesting stochastic process, the standard Brownian motion has many intriguing properties, one of which is the following reflection principle.

Theorem 3.1. *Let T be a stopping time, and let $\{B_t : t \geq 0\}$ be a standard Brownian motion, then the process $\{B_t^* : t \geq 0\}$ defined by*

$$B_t^* = B_t 1_{\{t \leq T\}} + (2B_T - B_t) 1_{\{t > T\}}$$

is also a standard Brownian motion.

This process, B_t^*, is called Brownian motion reflected at T, and using this reflection principle, we can prove the following result.

Lemma 3.2. *For $a > 0$, $P(\max_{0 \leq s \leq t} B_s > a) = 2P(B_t > a) = P(|B_t| > a)$.*

Proof. The second equality is obvious from the symmetry of Brownian motion, so it suffices to prove the first part to prove the lemma. We define a stopping time, T, as $T = \inf\{t \geq 0 : B_t = a\}$ and let $\{B_t^* : t \geq 0\}$ be Brownian motion reflected at the stopping time T, then

$$P(\max_{0 \leq s \leq t} B_s > a) = P(B_t > a) + P(\max_{0 \leq s \leq t} B_s > a \quad \text{and} \quad B_t \leq a)$$

$$= P(B_t > a) + P(B_t^* > a) = 2P(B_t > a) = P(|B_t| > a)$$

□

Now, we will get to the main theme of this thesis, which is proving the following three theorems: law of Iterated logarithm, Hartman-Wintner law of iterated logarithm, and Skorokhod embedding theorem, and the remaining parts of the thesis will closely follow chapter 5 from "Brownian Motion" by P. Morters and Y. Peres[2]. We will first prove the law of iterated logarithm, which can be stated as following:

Theorem 3.3. *If* $\{B_t : t \geq 0\}$ *is a standard Brownian motion, then*

$$\limsup_{t\to\infty} \frac{B_t}{\sqrt{2t\log\log(t)}} = \liminf_{t\to\infty} \frac{B_t}{-\sqrt{2t\log\log(t)}} = 1 \qquad \textit{almost surely}$$

By symmetry of the standard Brownian motion, it suffices to prove the lim sup case. We will make use of the Borel-Cantelli lemma and Lemma 3.2 to prove Theorem 3.3.

Proof. Let $\psi(t) = \sqrt{2t\log\log(t)}$. We will first prove the upper bound. For some fixed $\epsilon > 0$ and $q > 1$, we define the event A_n as

$$A_n = \max_{0\leq t\leq q^n} B_t \geq (1+\epsilon)\psi(q^n)$$

then by using Lemma 3.2, we get

$$P(A_n) = P(|B_{q^n}| \geq (1+\epsilon)\psi(q^n)) = P\left(\frac{|B_{q^n}|}{\sqrt{q^n}} \geq (1+\epsilon)\frac{\psi(q^n)}{\sqrt{q^n}}\right)$$

Note that for a standard normal random variable, Z, from Lemma 2.2, we have the tail estimate $P(Z > x) \leq \sqrt{\frac{2}{\pi}}\exp(-x^2/2) \leq \exp(-x^2/2)$ for $x \geq 1$, so

$$P(A_n) \leq 2\exp(-(1+\epsilon)^2\log\log(q^n)) = \frac{2}{(n\log(q))^{(1+\epsilon)^2}}$$

and for any $\epsilon > 0$ and $q > 1$,

$$\sum_{n=1}^{\infty} \frac{2}{(n\log(q))^{(1+\epsilon)^2}} < \infty$$

and, therefore, by Borel-Cantelli lemma, we know that only finitely many of these events occur. For any t, we can find n such that $q^{n-1} \leq t < q^n$, giving

$$\frac{B_t}{\psi(t)} = \frac{B_t}{\psi(q^n)} \frac{\psi(q^n)}{q^n} \frac{t}{\psi(t)} \frac{q^n}{t} \leq (1+\epsilon)q$$

because $\frac{\psi(t)}{t}$ is decreasing with respect to t. Since our choice of t was arbitrary, we have

$$\limsup_{t\to\infty} \frac{B_t}{\psi(t)} \leq (1+\epsilon)q \qquad \text{almost surely}$$

and since our choice of $\epsilon > 0$, $q > 1$ was also arbitrary, we have proved that

$$\limsup_{t\to\infty} \frac{B_t}{\psi(t)} \leq 1 \qquad \text{almost surely}$$

Having proved the upper bound, we will now prove the lower bound. Again, we will make use of Borel-Cantelli lemma and geometric sequence. Fix $q > 1$ and define the event D_n as

$$D_n = \{B_{q^n} - B_{q^{n-1}} \geq \psi(q^n - q^{n-1})\}$$

so that $\{D_n\}$ is a sequence of independent events.

Lemma 3.4. *If Z is a standard normal random variable, then for all $x > 0$, we have*

$$\frac{x}{x^2+1} \frac{1}{\sqrt{2\pi}} \exp(-x^2/2) \leq P(Z > x)$$

Proof. Let $f(x) = x\exp(-x^2/2) - (x^2+1)\int_x^\infty \exp(-u^2/2)du$, then it suffices to prove that $f(x) \leq 0$ for all x. It's clear that $f(0) = -1 < 0$ and $\lim_{x\to\infty} f(x) = 0$, and for $x > 0$

$$f'(x) = (1 - x^2 + x^2 + 1)\exp(-x^2/2) - 2x \int_x^\infty \exp(-u^2/2)du$$

$$= -2x\left(\int_x^\infty \exp(-u^2/2)du - \frac{\exp(-x^2/2)}{x}\right) \geq 0$$

so $f(x) \leq 0$ for all $x > 0$. $\square$

Using the above lemma, we know that if Z is a standard normal random variable, then for sufficiently large x, we have $P(Z > x) \geq \frac{\exp(-x^2/2)}{x}$, and from this tail estimate, for sufficiently large n, we get

$$P(D_n) = P\left(Z \geq \frac{\psi(q^n - q^{n-1})}{\sqrt{q^n - q^{n-1}}}\right) \geq \frac{\exp(-\log\log(q^n - q^{n-1}))}{2\log\log(q^n - q^{n-1})}$$

$$\geq \frac{\exp(-\log(n\log(q)))}{\sqrt{2\log(n\log(q))}} > \frac{1}{n\log(n)}$$

and, therefore, $\sum_{n=1}^{\infty} P(D_n)$ diverges, so for infinitely many n

$$B_{q^n} \geq B_{q^{n-1}} + \psi(q^n - q^{n-1})$$

and from the upper bound $B_{q^{n-1}} \leq 2\psi(q^{n-1})$ and symmetry of the standard Brownian motion, we get $B_{q^{n-1}} \geq -2\psi(q^{n-1})$, and, therefore, we can re-write the above inequality as

$$B_{q^n} \geq B_{q^{n-1}} + \psi(q^n - q^{n-1}) \geq -2\psi(q^{n-1}) + \psi(q^n - q^{n-1})$$

Thus, almost surely, for infinitely many n

$$\frac{B_{q^n}}{\psi(q^n)} \geq \frac{-2\psi(q^{n-1}) + \psi(q^n - q^{n-1})}{\psi(q^n)} \geq \frac{-2}{\sqrt{q}} + \frac{q^n - q^{n-1}}{q^n} = 1 - \frac{2}{\sqrt{q}} - \frac{1}{q}$$

as $\frac{\psi(t)}{\sqrt{t}}$ is increasing in t for sufficiently large t, but $\frac{\psi(t)}{t}$ is decreasing in t, and, therefore, we have

$$\limsup_{t\to\infty} \frac{B_t}{\psi(t)} \geq 1 - \frac{2}{\sqrt{q}} - \frac{1}{q}$$

and since our choice of $q > 1$ was arbitrary, we get the "almost sure" lower bound $\limsup_{t\to\infty} \frac{B_t}{\psi(t)} \geq 1$, and combining it with the upper bound, we know that

$$\limsup_{t\to\infty} \frac{B_t}{\psi(t)} = 1 \qquad \text{almost surely.}$$

□

Now, the question is whether we can get a result similar to one in Theorem 3.3 in a discrete case as well, and conveniently, the answer is yes. We will demonstrate the result by proving the following discrete analogue of Theorem 3.3: (from now on, we will denote B_t by $B(t)$)

Theorem 3.5. *If $S_n = X_1 + X_2 + \cdots + X_n$ in which $\{X_n\}$ is an i.i.d. sequence of symmetric random variables such that $P(X_i = 1) = P(X_i = -1) = \frac{1}{2}$ for each i, then almost surely*

$$\limsup_{n\to\infty} \frac{S_n}{\sqrt{2n\log\log(n)}} = 1$$

For the proof of the above theorem, we will need the following lemma:

Lemma 3.6. *If $\{T_n : 1 \leq n\}$ is a sequence of random times such that $T_n \to \infty$ and $\frac{T_{n+1}}{T_n} \to 1$ almost surely, then letting $\psi(t) = \sqrt{2t\log\log(t)}$*

$$\limsup_{n\to\infty} \frac{B(T_n)}{\psi(T_n)} = 1 \qquad \textit{almost surely.}$$

Also, if $\frac{T_n}{n} \to a > 0$ almost surely, then

$$\limsup_{n\to\infty} \frac{B(T_n)}{\psi(an)} = 1 \qquad \textit{almost surely.}$$

Remark 3.7. The upper bound requires the sequence $\{T_n : 1 \leq n\}$ to be unbounded, and the second condition, $\frac{T_{n+1}}{T_n} \to 1$, requires times to be sufficiently dense to make the lemma work. For instance, for the following sequence

$$T_0 = 0 \qquad\qquad T_n = \inf\{t > T_{n-1} + 1 : B(t) < \frac{1}{n}\}$$

the second condition, ($\frac{T_{n+1}}{T_n} \to 1$ almost surely), does not hold and

$$\limsup_{n\to\infty} \frac{B(T_n)}{\psi(T_n)} = 0 \qquad \text{almost surely.}$$

Proof. Now, assume that first and second conditions both hold for the sequence of times $\{T_n : 1 \leq n\}$. Since the fact that $\limsup_{n\to\infty} \frac{B(T_n)}{\psi(T_n)} \leq 1$ almost surely follows from the upper bound for continuous time without any conditions on $\{T_n : n \geq 1\}$, it suffices to prove that $\limsup_{n\to\infty} \frac{B(T_n)}{\psi(T_n)} \geq 1$ almost surely. Fix $q > 4$, and define the events D_k, Ω_k, D_k^* as following:

$$D_k = \{B(q^k) - B(q^{k-1}) \geq \psi(q^k - q^{k-1})\}$$

$$\Omega_k = \{\min_{q^k \leq t \leq q^{k+1}} B(t) - B(q^k) \geq -\sqrt{q^k}\}$$

$$D_k^* = D_k \cap \Omega_k$$

Note that D_k, Ω_k are independent events, so we have

$$P(D_k^*) = P(D_k)P(\Omega_k)$$

From the proof of Theorem 3.3, we know that there is a constant, $c > 0$ such that

$$P(D_k) = P\left(B(1) \geq \frac{\psi(q^k - q^{k-1})}{\sqrt{q^k - q^{k-1}}}\right) \geq \frac{c}{k\log(k)}$$

and by scaling (dividing both sides of $\sqrt{q^k}$), we know that for some constant $c_2 > 0$, we have $P(\Omega_k) > c_2$ for all k. Thus, we have

$$c_2 \sum_{j=1}^{\infty} \frac{c}{j \log(j)} \leq c_2 P(D_k) < \sum_{k=1}^{\infty} P(D_k^*)$$

so we know that $\sum_{k=1}^{\infty} P(D_{2k}^*)$ is infinite, and since the events $\{D_{2k}^* : k \geq 1\}$ are independent, by the Borel-Cantelli lemma, for infinitely many even integers, k, we have

$$\min_{q^k \leq t \leq q^{k+1}} B(t) \geq B(q^{k-1}) + \psi(q^k - q^{k-1}) - \sqrt{q^k}$$

and note that if k is sufficiently large, then $B(q^{k-1}) \geq -2\psi(q^{k-1})$ holds almost surely, and $(1 - \frac{1}{q})\psi(q^k) \leq \psi(q^k - q^{k-1})$, so we have, for infinitely many k,

$$\min_{q^k \leq t \leq q^{k+1}} B(t) \geq B(q^{k-1}) + \psi(q^k - q^{k-1}) - \sqrt{q^k} \geq \psi(q^k)(1 - \frac{1}{q} - \frac{2}{\sqrt{q}}) - \sqrt{q^k}$$

Now, let $n(k) = \min\{n : T_n > q^k\}$, then as $\frac{T_{n+1}}{T_n} \to 1$ almost surely, for any fixed $\epsilon > 0$, for all sufficiently large k, we have $q^k \leq T_{n(k)} \leq q^k(1+\epsilon)$, so for infinitely many k, we have

$$\frac{B(T_{n(k)})}{\psi(T_{n(k)})} \geq \frac{\psi(q^k)}{\psi(q^k(1+\epsilon))}(1 - \frac{1}{q} - \frac{2}{\sqrt{q}}) - \frac{\sqrt{q^k}}{\psi(q^k)}$$

but as $\frac{\sqrt{q^k}}{\psi(q^k)} \to 0$ and $\frac{\psi(q^k)}{\psi(q^k(1+\epsilon))} \to \frac{1}{\sqrt{1+\epsilon}}$, we have

$$\limsup_{n\to\infty} \frac{B(T_n)}{\psi(T_n)} \geq \frac{1}{\sqrt{1+\epsilon}}(1 - \frac{1}{q} - \frac{2}{\sqrt{q}})$$

and if we let $\epsilon \to 0$ and $q \to \infty$, as the left side does not depend on either q or ϵ, we get the desired lower bound

$$\limsup_{n\to\infty} \frac{B(T_n)}{\psi(T_n)} \geq 1 \qquad \text{almost surely}$$

which leads to the conclusion

$$\limsup_{n\to\infty} \frac{B(T_n)}{\psi(T_n)} = 1 \qquad \text{almost surely,}$$

and note that if $\frac{T_n}{n} \to a$, then $\frac{\psi(T_n)}{\psi(an)} \to 1$ by scaling. $\square$

Now, using the above lemma, we will prove Theorem 3.5. It turns out to be remarkably simple.

Proof. All we need to do is to construct a correct sequence of times, and we can do so by defining $\{T_n\}$ as following:

$$T_0 = 0$$

$$T_n = \min\{t > T_{n-1} : |B(t) - B(T_{n-1})| = 1\} \quad \text{for } \ 1 \leq n$$

The sequence of times $\{T_n\}$ is just a sequence of stopping times for Brownian motion, and, therefore, the waiting times $T_n - T_{n-1}$ are i.i.d. random variables by the strong Markov property. It's clear that $T_n \to \infty$ from construction, so we just have to check the second condition, $\frac{T_{n+1}}{T_n} \to 1$.

By symmetry, we have

$$P(B(T_n) - B(T_{n-1}) = 1) = P(B(T_n) - B(T_{n-1}) = -1) = \frac{1}{2}$$

so we can think of $\{B(T_n) : n \geq 0\}$ as a simple random walk. To calculate $E(T_n - T_{n-1})$, we will use the following fact:

- For $a < 0 < b$, if we let $T = \min\{t \geq 0 : B(t) \in \{a, b\}\}$, then

$$P(B(T) = a) = \frac{b}{|a| + b} \qquad P(B(T) = b) = \frac{|a|}{|a| + b} \qquad E(T) = |a|b$$

then by letting $a = -1$, $b = 1$, we get $E(T_n - T_{n-1}) = 1$, and, therefore, we have $\frac{T_{n+1}}{T_n} \to 1$, then the statement in Theorem 3.5 is just a natural consequence of Lemma 3.6. $\square$

Although we put some severe restrictions on the distribution of X_i, this is a good beginning for the proof of Hartman-Wintner law of iterated logarithm that we will cover later. (Note that if X_i is any symmetric binomial random variable with mean zero and $\mathrm{Var}(X_i) = \sigma^2$ for any $\sigma > 0$, then we can just normalize it by dividing it by σ and apply Theorem 3.5).

4 Skorokhod Embedding Theorem

What if $\{X_i\}$ is any i.i.d. sequence of random variables with finite variance and mean zero? Does the Theorem 3.5 still hold true? The answer is yes, and this remarkably general phenomenon is known as Hartman-Wintner law of the iterated logarithm. To prove this theorem, we need one more theorem, namely Skorokhod embedding theorem, and this chapter will be dedicated to proving that theorem. Before we prove the theorem, we will prove two theorems that will be used to prove a key lemma, which is necessary for the proof of Skorokhod embedding theorem. The proof of Theorem 4.1 is based on the corresponding part from "Probability with Martingales" by David Williams[3].

Theorem 4.1. *If X is an integrable random variable and $X_n = E(X|\mathcal{F}_n)$, then $\{X_n : n \geq 0\}$ is a uniformly integrable martingale and*

$$\lim_{n\to\infty} X_n = E(X|\mathcal{F}_\infty) \qquad \textit{almost surely and in } \mathbf{L}^1$$

in which a filtration, $\mathcal{F}_\infty$ is defined as $\mathcal{F}_\infty = \sigma(\bigcup_n \mathcal{F}_n)$.

Proof. Note that it is obvious from the tower property that $\{X_n : n \geq 0\}$ is martingale. To show that it is uniformly integrable, we will prove the following lemma:

Lemma 4.2. *If $X \in \mathbf{L}^1$, then the class*

$$\{E(X|\mathcal{G}) : \mathcal{G} \textit{ is a sub-}\sigma\textit{-algebra of } \mathcal{F}\}$$

is uniformly integrable.

Proof. For any given $\epsilon > 0$, we can choose $\delta > 0$ such that for $F \in \mathcal{F}$

$$P(F) < \delta \quad \Rightarrow \quad \int_F |X| dP < \epsilon$$

then we let c to be a positive real number such that $c^{-1}E(|X|) < \delta$ and define $\mathcal{G}$ as sub-σ-algebra of $\mathcal{F}$ and Y as any version of $E(X|\mathcal{G})$, then by Jensen's inequality, we get

$$|Y| \leq E(|X|\ |\mathcal{G}) \quad \text{almost surely}$$

and thus, $E(|Y|) \leq E(|X|)$, and

$$cP(|Y| > c) \leq E(|Y|) \leq E(|X|)$$

such that

$$P(|Y| > c) < \delta$$

However, $\{|Y| > c\} \in \mathcal{G}$, so from the definition of conditional expectation and the fact we just derived: $|Y| \leq E(|X|\ |\mathcal{G})$ almost surely, we get the desired result:

$$\int_{|Y| \geq c} |Y| \leq \int_{|X| \geq c} |X| < \epsilon$$

□

Now, we have proved that $\{X_n : n \geq 0\}$ is a uniformly integrable martingale, so all we need to prove is the last part: $\lim_{n\to\infty} X_n = E(X|\mathcal{F}_\infty)$ almost surely. Without a loss of generality, let's assume that $X \geq 0$, and let $\mu = E(X|\mathcal{F}_\infty)$, then consider two measures: Q_1, Q_2 on $(\Omega, \mathcal{F}_\infty)$, which we will define as

$$Q_1 = \int_F \mu\ dP$$

$$Q_2 = \int_F X_\infty \, dP \qquad (X_\infty = \limsup X_n)$$

for any $F \in \mathcal{F}_\infty$. If $F \in \mathcal{F}_n$, then as $E(\mu|\mathcal{F}_n) = E(X|\mathcal{F}_n)$ by tower property, we get

$$\int_F \mu \, dP = \int_F X_n \, dP = \int_F X_\infty \, dP$$

showing that Q_1, Q_2 agree on $\bigcup \mathcal{F}_n$, and from $\pi-\lambda$ theorem, they agree on $\mathcal{F}_\infty$ as well. Note that μ is clearly $\mathcal{F}_\infty$-measurable, and $X_\infty = \limsup X_n$ is also $\mathcal{F}_\infty$-measurable. Thus, we get

$$F = \{\omega : \mu > X_\infty\} \in \mathcal{F}_\infty$$

and as $Q_1(F) = Q_2(F)$, we get

$$\int_{\mu > X_\infty} \mu - X_\infty \, dP = 0 \Rightarrow P(\mu > X_\infty) = P(X_\infty > \mu) = 0$$

which means

$$\lim_{n\to\infty} X_n = E(X|\mathcal{F}_\infty) \qquad \text{almost surely and in } \mathbf{L}^1$$

□

Theorem 4.3. *If the martingale $\{X_n : n \geq 0\}$ is $\mathbf{L}^2$-bounded, then there is a random variable, X, such that*

$$\lim_{n\to\infty} X_n = X \qquad \textit{almost surely and in } \mathbf{L}^2$$

Proof. For $m \geq n$, we have

$$E((X_m - X_n)^2) = \sum_{k=n+1}^{m} E((X_k - X_{k-1})^2) \leq \sum_{k=1}^{\infty} E((X_k - X_{k-1})^2) < \infty$$

Then as $\mathbf{L}^2$-boundedness implies $\mathbf{L}^1$-boundedness, by martingale convergence theorem, X_n converges to an integrable random variable, X almost surely. Then if we let $m \to \infty$, by applying Fatou's lemma, we get the desired result. $\square$

Theorem 4.1 is called Levy's upward theorem, and Theorem 4.3 is called convergence theorem for $\mathbf{L}^2$-bounded martingales. Note that in Theorem 3.5, we constructed a sequence of stopping times, $\{T_n\}$, such that it satisfies the conditions in Lemma 3.6 and $B(T_{n+1}) - B(T_n)$ has the law of a symmetric binomial random variable with mean 0 and variance 1. In that particular case, we put some severe restrictions on the distribution of X_i, and one may wonder for any random variable X with mean zero and finite variance, is it possible to find a stopping time, T, such that $B(T)$ has the law of X. Again, the answer is yes, and this is a result of Skorokhod embedding theorem, which can be formally stated as following:

Theorem 4.4. *Assume that $\{B(t) : t \geq 0\}$ is a standard Brownian motion, and X is a random variable with mean zero and finite variance. Then there is a stopping time, T, with respect to the natural filtration, $(\mathcal{F}(t) : t \geq 0)$ of Brownian motion, such that $B(T)$ has the law of X and $E(T) = E(X^2)$.*

Proof. First, we will define a new concept called **binary splitting**. If $\{X_n : n \in \mathbb{N}\}$ is martingale such that whenever for some $x_0, \cdots, x_n \in \mathbb{R}$, the event

$$A(x_0, \cdots, x_n) = \{X_0 = x_0, X_1 = x_1, \cdots, X_n = x_n\}$$

has positive probability, the random variable X_{n+1} conditioned on $A(x_0, \cdots, x_n)$ is supported on at most two values, then it's called **binary splitting**.

Lemma 4.5. *If X is a random variable with finite variance, then there is a binary splitting martingale, $\{X_n : n \in \mathbb{N}\}$ such that $X_n \to X$ almost surely in $\mathbf{L}^2$.*

Proof. Let $\mathcal{G}_0$ be the trivial σ-algebra, which only consists of the empty set and the underlying probability space and let $X_0 = E(X)$. We will define the martingale $\{X_n : n \in \mathbb{N}\}$ and the associated filtration $(\mathcal{G}_n : n \in \mathbb{N}\}$ recursively from those. Next, we define the random variable, ξ_0 by

$$\xi_0 = \begin{cases} 1 & \text{if } X \geq X_0, \\ \\ -1 & \text{if } X < X_0. \end{cases}$$

and for $n > 0$, we let $\mathcal{G}_n = \sigma(\xi_0, \xi_1, \cdots, \xi_{n-1})$ and $X_n = E(X|\mathcal{G}_n)$ and define ξ_n as

$$\xi_n = \begin{cases} 1 & \text{if } X \geq X_n, \\ \\ -1 & \text{if } X < X_n. \end{cases}$$

To show that $\{X_n : n \in \mathbb{N}\}$ is binary splitting martingale, note that $\mathcal{G}_n$ is generated by a partition, $\mathcal{P}_n$ of the underlying probability space into 2^n sets, each of which has the form $A(x_0, \cdots, x_n)$, and since each element of $\mathcal{P}_n$ is a union of two elements of $\mathcal{P}_{n+1}$, the martingale $\{X_n : n \in \mathbb{N}\}$ is indeed binary splitting. Note that from construction of X_n, we get

$$E(X^2) = E(((X - X_n) + X_n)^2) = E((X - X_n)^2) + 2E((X - X_n)X_n) + E(X_n^2)$$

$$= E((X - X_n)^2) + E(X_n^2) \geq E(X_n^2)$$

Thus, $\{X_n : n \in \mathbb{N}\}$ is bounded in $\mathbf{L}^2$, and from Levy's upward theorem and the convergence theorem for $\mathbf{L}^2$, Theorem 4.1 and 4.2 respectively, we get

$$X_n \to X_\infty = E(X|\mathcal{G}_\infty) \qquad \text{almost surely in } \mathbf{L}^2$$

in which $\mathcal{G}_\infty = \sigma(\bigcup_{i=0}^{\infty} G_i)$. Now, all we need to show is $X = X_\infty$ almost surely. To do so, we will show that

$$\lim_{n\to\infty} \xi_n(X - X_{n+1}) = |X - X_\infty|$$

Note that if $X(\omega) < X_\infty(\omega)$ then there is an integer N such that $X(\omega) < X_n(\omega)$ for $n < N$, so $\xi_n = -1$, and the above equality holds. Similarly, if $X(\omega) > X_\infty(\omega)$ then there is an integer N such that $X(\omega) > X_n(\omega)$ for $n < N$, so $\xi_n = 1$, and the above equality holds as well, and if $X(\omega) = X_\infty(\omega)$, the equality trivially holds.

Using the fact that ξ_n is $\mathcal{G}_{n+1}$-measurable, we get

$$E[\xi_n(X - X_{n+1})] = E[\xi_n E[X - X_{n+1}|\mathcal{G}_{n+1}]] = 0$$

and since if $Y_n \to Y$ almost surely, and $\{Y_n : n = 0, 1, \cdots\}$ is $\mathbf{L}^2$-bounded, then $E(Y_n) \to E(Y)$, we know that $\lim_{n\to\infty} \xi_n(X - X_{n+1})$ must also be $\mathbf{L}^2$-bounded, so $E|X - X_\infty| = 0$. $\square$

Now, having completed the proof of the lemma, we will get to the proof of Skorokhod embedding theorem. Using the above lemma, we take a binary splitting martingale $\{X_n : n \in \mathbb{N}\}$ such that $X_n \to X$ almost surely in $\mathbf{L}^2$. Then it is well-known fact that if X is supported on a set of two elements $\{-a, b\}$ for some $a, b > 0$

with mean zero, then if we let $T = \inf\{t : B(t) \in \{-a, b\}\}$ to be our stopping time, then $B(T)$ has the same distribution as that of X with $E(T) = -ab < \infty$, so we will use that as our stopping time. Thus, as X_n conditioned on $A(x_0, x_1, \cdots, x_{n-1})$ is supported on at most two values, it is clear that we can find a sequence of monotone increasing stopping times, $T_0 \leq T_1 \leq T_2 \leq \cdots$ such that $B(T_n)$ is distributed as X_n for each n, and $E(T_n) = E(X_n^2)$. Since T_n is a monotone increasing sequence, we have $T_n \to T$ almost surely for some stopping time, T, and by monotone convergence theorem

$$E(T) = \lim_{n\to\infty} E(T_n) = \lim_{n\to\infty} E(X_n^2) = E(X^2)$$

then as $B(T_n)$ converges in distribution to X from construction and converges almost surely to $B(T)$ from the continuity of Brownian sample paths, we get the desired result: $B(T)$ is distributed as X. $\square$

5 Hartman-Wintner Law of Iterated Logarithm

Now we have all the necessary tools to prove Hartman-Wintner law of iterated logarithm, which can be stated formally as following:

Theorem 5.1. *If $\{S_n : n \in \mathbb{N}\}$ is a random walk such that increments, $S_n - S_{n-1}$ has zero mean and finite variance, σ^2, then the following holds almost surely:*

$$\limsup_{n\to\infty} \frac{S_n}{\sqrt{2\sigma^2 n \log\log(n)}} = 1$$

Proof. Having proved Skorokhod embedding theorem, the proof for this law is delightfully simple. Given the distribution of $S_{n+1} - S_n$, we simply replace the sequence of times $\{T_n\}$ in Theorem 3.5 by the sequence of times $\{T_n^*\}$ such that $B(T_{n+1}^*) - B(T_n^*)$ each has the law of S_n, which exists by Skorokhod embedding theorem, then the result naturally follows. □

References

[1] Michael Steele, Stochastic Calculus and Financial Applications, p29-40, Springer, (2000)

[2] Peter Morters and Yuval Peres, Brownian Motion, p118-123, p127-129, Cambridge University Press, (2010)

[3] David Williams, Probability with Martingales, p134-135, Cambridge University Press, (1991)