Skorokhod's embedding and concentration inequality*

Jialin Yi †

Abstract: We analyze the uniqueness of Skorokhod's embedding theorem. For general L2 random variable, we demonstrate that two embedding stopping times mush have the same negative moments. And we show that for Rademacher distribution there are at least countably many embedding stopping times. Besides, we use Skorokhod's embedding theorem to analyze the lower bound of Hoeffding's inequality and give a sharper concentration inequality for martingale difference sequence.

1. Uniqueness analysis of Skorokhod's embedding theorem

Skorokhod's embedding theorem [1] says every $L2$ random variable could be embedded into a Brownian motion but does not ensure the uniqueness of such embedding. We first give a general result about how similar the two embedding stopping times could be in terms of moments and then show an example where there exist countably many embedding stopping times.

1.1. General analysis

Theorem 1.1. *Suppose X is a L2 random variable with mean 0 and $\mathbf{P}(X \neq 0) > 0$. τ_1 and τ_2 are the two Skorokhod's embedding stopping times for X such B_{τ_1} and B_{τ_2} are continuous random variables then*

$$\mathbf{E}\left[\tau_1^{-\frac{1}{2}} e^{-\frac{x^2}{2\tau_1}}\right] = \mathbf{E}\left[\tau_2^{-\frac{1}{2}} e^{-\frac{x^2}{2\tau_2}}\right] \tag{1.1}$$

for all x. Specially, $\mathbf{E}[\tau_1^{-\frac{1}{2}}] = \mathbf{E}[\tau_2^{-\frac{1}{2}}]$.

Proof. Note that $B_{\tau_1} \overset{d}{=} X \overset{d}{=} B_{\tau_2}$, then $\mathbf{P}(B_{\tau_1} \leq x) = \mathbf{P}(B_{\tau_2} \leq x)$ for any x, and since $\mathbf{P}(X \neq 0) > 0$ then $\tau_1, \tau_2 > 0$ applying Fubini's theorem

$$\mathbf{P}(B_{\tau_1} \leq x) = \mathbf{E}\left[\int_{-\infty}^{x} (2\pi\tau_1)^{-\frac{1}{2}} e^{-\frac{y^2}{2\tau_1}} \, dy\right] = \int_{-\infty}^{x} \mathbf{E}\left[(2\pi\tau_1)^{-\frac{1}{2}} e^{-\frac{y^2}{2\tau_1}}\right] dy \tag{1.2}$$

Since B_{τ_1} and B_{τ_2} have the same distribution, they have the same probability density function then

$$\mathbf{E}\left[\tau_1^{-\frac{1}{2}} e^{-\frac{x^2}{2\tau_1}}\right] = \mathbf{E}\left[\tau_2^{-\frac{1}{2}} e^{-\frac{x^2}{2\tau_2}}\right] \tag{1.3}$$

holds for all x. Let $x = 0$ yields the special case □

*This is the 4-page paper for STAT 931: *Stochastic Process* final project.

†jialinyi@sas.upenn.edu

1.2. An example: Rademacher distribution

Consider the Rademacher random variable

$$X = \begin{cases} 1 & \text{with probability } \frac{1}{2} \\ -1 & \text{with probability } \frac{1}{2} \end{cases} \tag{1.4}$$

Define $\tau_1 = \min\{t \geq 0 : |B_t| = 1\}$ and $\tau_n = \min\{t > \tau_{n-1} : |B_t| = 1\}$ for $n \geq 2$. We show that $\{\tau_n\}_{n=1}^{\infty}$ are all the embedding stopping times for X.

Lemma 1.1. *Suppose $\{\tau_n\}_{n=1}^{\infty}$ is defined as above, then $P(\tau_n < \infty) = 1$. Further, we have $\mathbf{E}[\tau_1] \leq 1$.*

Proof. We first prove when $n = 1$. Since $B_t^2 - t$ is a martingale and τ_1 is stopping time, then by Doob's optional stopping theorem

$$\mathbf{E}[B_{\tau_1 \wedge n}^2 - \tau_1 \wedge n] = 0 \tag{1.5}$$

Note that $|B_{\tau_1 \wedge n}| \leq 1$ since B_t is continuous with t. Otherwise, by intermediate value theorem, there exists some $t' < \tau_1$ such that $|B_{t'}| = 1$.

Letting $n \to \infty$ by monotone convergence theorem yields

$$\mathbf{E}[\tau_1] = \lim_{n \to \infty} \mathbf{E}(\tau_1 \wedge n) \leq 1 \tag{1.6}$$

therefore $\mathbf{P}(\tau_1 < \infty) = 1$. To show that $\mathbf{P}(\tau_n < \infty) = 1$ for all n, it suffices to show that $\limsup_{t \to \infty} B_t = +\infty$ and $\liminf_{t \to \infty} B_t = -\infty$ almost surely.

Let $Z_n = B_n - B_{n-1}$, then $Z_1, Z_2, \cdots, Z_n$ are i.i.d. standard normal and $B_n = Z_1 + \cdots + Z_n$, then by Kolmogorov's 0-1 law, $\mathbf{P}(\limsup B_n = +\infty) = 0$ or 1.

Note that applying reverse Fatou's lemma, we have

$$\mathbf{P}(B_n > n^{\frac{1}{2}} \ i.o.) \geq \limsup \mathbf{P}(B_n > n^{\frac{1}{2}}) = \mathbf{P}(B_1 > 1) > 0 \tag{1.7}$$

then $\mathbf{P}(\limsup B_n = +\infty) = 1$ and $\mathbf{P}(\liminf_{t \to \infty} B_t = -\infty) = 1$ follows by the symmetry of Brownian motion. $\square$

By definition $|B_{\tau_n}| = 1$ for all $n \geq 1$ and by the symmetry of standard Brownian motion, the probability of hitting ± 1 are the same, hence we have the following results.

Theorem 1.2. *Suppose $\{\tau_n\}_{n=1}^{\infty}$ are defined as above, then $\{\tau_n\}_{n=1}^{\infty}$ are all Skorokhod's embedding stopping times for X.*

2. Concentration inequality for martingale difference sequence

2.1. A Lower bound for Hoeffding's lemma

We give a lower bound for all the approximation of the expectation of the exponential of a $L2$ random variable, which is also the trick of the Hoeffding's lemma [2]. This lower bound guides us to improve the Hoeffding's inequality

Note that any $L2$ random variable with mean zero could be embedded into the standard Brownian motion by Skorokhod's embedding theorem [1]. And we know the expectation formula for the geometric Brownian motion

$$\mathbf{E}[e^{B_t}] = e^{\frac{1}{2}t} \tag{2.1}$$

hence we can prove the following theorem.

Theorem 2.1. *Suppose* $\mathbf{E}|X|^2 < \infty$*, then*

$$\mathbf{E}[e^{tX}] \geq \exp\left(\frac{1}{2}t^2\sigma^2 + t\mu\right) \tag{2.2}$$

for all $t > 0$ *where* $\mathbf{E}[X] = \mu$ *and* $\mathrm{Var}[X] = \sigma^2$.

Proof. Let $Y = X - \mu$, then $\mathbf{E}[Y] = 0$, $\mathrm{Var}[Y] = \mathrm{Var}[X] = \sigma^2$.

There exists some stopping time τ with respect to the natural filtration $\{\mathcal{F}_t\}_{t\geq 0}$ of standard Brownian motion $\{B_t, t \geq 0\}$ such that

$$B_\tau \overset{d}{=} Y \tag{2.3}$$

with $\mathbf{E}[\tau] = \sigma^2$ and $\mathbf{E}[\tau^2] \leq 4\mathbf{E}[Y^4]$ by Skorokhod's embedding theorem [1]. Hence

$$\mathbf{E}[e^{tY}] = \mathbf{E}\left[e^{tB_\tau}\right] = \mathbf{E}\left[e^{B_{t^2\tau}}\right] = \mathbf{E}\left[\exp\left(\frac{1}{2}t^2\tau\right)\right] \geq \exp\left(\frac{1}{2}t^2\sigma^2\right) \tag{2.4}$$

the second inequality comes from the scaling of the Brownian motion and the inequality follows Jensen's inequality. Multiplying $\exp(t\mu)$ on the both side of the inequality above and we complete the proof. □

Note that if $\mathbf{E}(X) = 0$ and $|X| \leq M$ with probability 1, then X is in L_2 and Theorem 2.1 holds. This shows that the Hoeffding's lemma is sharp with respect to the growth rate of t — exponential squared of t.

However, the Hoeffding's lemma is not sharp since it replace the variance of X with bound M which is larger. Note that Theorem 2.1 tells us that the expectation of the exponential function of the L2 random variable must be at least $O(e^{\frac{1}{2}t^2\sigma^2})$. Based on this approximation, we propose another concentration inequality for martingale difference sequence like Hoeffding's inequality.

2.2. Improving through Taylor's formula

Lemma 2.1. *Suppose* X *is random variable in* $(\Omega, \mathcal{F}_0, \mathbf{P})$ *with* $\mathcal{F} \subset \mathcal{F}_0$ *and* $\mathbf{E}[X|\mathcal{F}] = 0$*,* $|X| \leq M$ *almost surely, then*

$$\mathbf{E}[e^{tX}|\mathcal{F}] \leq \exp\left\{\sigma^2\left[\frac{e^{tM} - 1 - tM}{M^2}\right]\right\} \tag{2.5}$$

where $\sigma^2 = \mathbf{E}[X^2|\mathcal{F}]$*. Specially, we have*

$$\mathbf{E}[e^{tX}|\mathcal{F}] \leq \exp\left[e^{tM} - 1 - tM\right] \tag{2.6}$$

Proof. By Taylor's formula, we have

$$e^{tX} = 1 + tX + \sum_{t=2}^{\infty} \frac{(tX)^n}{n!} \tag{2.7}$$

taking the conditional expectation of the both sides and conditional dominated convergence theorem yields

$$\mathbf{E}[e^{tX}|\mathcal{F}] = 1+\sum_{t=2}^{\infty} \frac{(t\mathbf{E}\left[X|\mathcal{F}\right])^n}{n!} \leq \exp\left[\sum_{t=2}^{\infty} \frac{t^n M^{n-2}}{n!}\sigma^2\right] = \exp\{\sigma^2 \left[\frac{e^{tM}-1-tM}{M^2}\right]\} \tag{2.8}$$

the inequality comes from the basic inequality $1 + x \leq e^x$. □

It is worth pointing out that this approximation works pretty well when t is small, since $(e^{tM} - 1 - tM)/M^2 = O(\frac{1}{2}t^2)$ as $t \to 0$. Compare to the lower bound we got in Theorem 2.1, this approximation is sharpe when $t \to 0$.

Theorem 2.2. *Suppose $\{X_i\}_{i=1}^n$ is a martingale difference sequence with respect to $\{\mathcal{F}_i\}_{i=1}^n$ and $|X_i| \leq M$ almost surely for all n, then*

$$\mathbf{P}(S_n/n > \epsilon) \leq \exp\left[-n\psi(\epsilon/M)\right] \tag{2.9}$$

where $S_n = X_1 + X_2 + \cdots + X_n$ and $\psi(x) = (1 + x)\log(1 + x) - x$.

Proof. By Markov's inequality

$$\mathbf{P}(S_n/n > \epsilon) = \mathbf{P}(e^{tS_n} > e^{tn\epsilon}) \leq e^{-tn\epsilon}\mathbf{E}[e^{tS_n}] \tag{2.10}$$

for all $t > 0$ and since $X_i \in \mathcal{F}_{i-1}$,

$$\mathbf{E}[e^{tS_n}] = \mathbf{E}\left[\prod_{i=1}^{n} \mathbf{E}[e^{tX_i}|\mathcal{F}_i]\right] \tag{2.11}$$

by Lemma 2.1 we get

$$\mathbf{P}(S_n/n > \epsilon) \leq \exp\left[-tn\epsilon + n(e^{tM} - 1 - tM)\right] \tag{2.12}$$

which holds for all $t > 0$. Note that

$$e^{tM} - 1 - tM - t\epsilon \geq \frac{\epsilon}{M} - (1 + \frac{\epsilon}{M})\log(1 + \frac{\epsilon}{M}) \tag{2.13}$$

then

$$\mathbf{P}(S_n/n > \epsilon) \leq \exp\left[-n\left((1 + \frac{\epsilon}{M})\log(1 + \frac{\epsilon}{M}) - \frac{\epsilon}{M}\right)\right] \tag{2.14}$$

□

Acknowledgements

Helpful discussion with Prof. Michael Steele, Darrick Lee and Jacob Seidman are greatly acknowledged.

References

[1] Skhorokhod, AV (1965). *Studies in the theory of random processes*, Translated from the Russian by Scripta Technica, Inc. Addison-Wesley Publishing Co., Inc., Reading, Mass.

[2] Hoeffding, Wassily (1963). Probability inequalities for sums of bounded random variables *Journal of the American statistical association* **58** 13–30.

[3] Durrett, Rick (2010). *Probability: theory and examples.* Cambridge university press