Doob's inequalities

Hongyi Jin 518030910333

July 1, 2020

1 Joseph Leo Doob

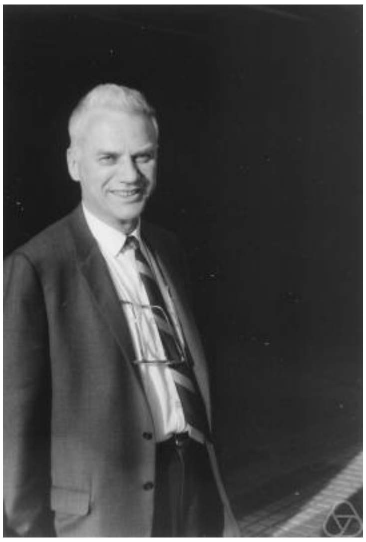

Figure 1: Joseph Leo "Joe" Doob (February 27, 1910 – June 7, 2004) was an American mathematician, specializing in analysis and probability theory. The theory of martingales was developed by Doob.

Joseph Doob was an American mathematician who worked in probability and measure theory. Doob's work was in probability and measure theory, in particular he studied the relations between probability and potential theory. The paper looks at many of the areas of probability to which Doob made major contributions such as separability, stochastic processes, martingales, optimal stopping, potential theory, and classical potential theory and its probabilistic counterpart. Doob built on work by Paul Lévy and, during the 1940's and 1950's, he developed basic martingale theory and many of its applications. Doob's work has become one of the most powerful tools available to study stochastic processes. In 1953 he published a book which gives a comprehensive treatment of stochastic processes, including much of his own development of martingale theory. This book Stochastic Processes has become a classic and was reissued in 1990. In fact he undertook the work of writing the book because he had become intellectually bored while undertaking war work in Washington and so was enthusiastic when, in 1945, Shewhart invited him to publish a volume in the Wiley series in statistics.[2]

The spirit of his book Stochastic Processes is clear by his statement:

... probability theory is simply a branch of measure theory, with its own special emphasis and field of application ...

2 Doob's martingale inequality

In mathematics, Doob's martingale inequality, also known as Kolmogorov's submartingale inequality is a result in the study of stochastic processes. The inequality is due to the American mathematician Joseph L. Doob.

In class, we have discussed the proof of Doob's martingale inequality.

Theorem 1 *Let $\{X\}$ be a submartingale. For all $\lambda > 0$, and $n \in \mathbb{N}$,*

$$\lambda P(\max_{k \le n} X_k \ge \lambda) \le E[X_n^+] \le E[|X_n|] \tag{1}$$

Proof. Let $F_n = \sigma(X_1, X_2, \cdots, X_n)$, $\tau = n \wedge \inf\{k : X_k >= \lambda\}$. Then τ is bounded above by n, from which we can derive that τ is a stopping time. Since X is a submartingale, we get

$$\begin{aligned} E(X_n) &\ge E(X_\tau) && (2) \\ &= E(X_\tau 1_{\max_{k\le n} X_k \ge \lambda}) + E(X_\tau 1_{\max_{k\le n} X_k < \lambda}) && (3) \\ &\ge \lambda P(\max_{k\le n} X_k \ge \lambda) + E(X_n 1_{\max_{k\le n} X_k < \lambda}) && (4) \end{aligned}$$

After transforming the formula, we get

$$\lambda P(\max_{k\le n} X_k \ge \lambda) \le E(X_n 1_{\max_{k\le n} X_k \ge \lambda}) \le E[X_n^+] \le E[|X_n|] \tag{5}$$

■

2.1 Relationship with Kolmogrov's inequality

Kolmogorov's inequality is a so-called "maximal inequality" that gives a bound on the probability that the partial sums of a finite collection of independent random variables exceed some specified bound. The inequality is named after the Russian mathematician Andrey Kolmogorov. Statement of the inequality is as below:

Theorem 2 *Let $X_1, X_2, \cdots, X_n$ be independent random variables defined on (Ω, F, P), with expected values $E[X_k] = 0$ and variance $Var[X_k] < \infty$ for $k = 1, 2, \cdots, n$. Then for each $\lambda > 0$,*

$$P(\max_{1\le k\le n} |S_k| \ge \lambda) \le \frac{1}{\lambda^2} Var[S_n] \tag{6}$$

, where $S_k = X_1 + \cdots + X_k$

The original proof of Kolmogrov's inequality requires some effort. However, we can obtain a quite simple proof with the help of Doob's martingale inequality[5]. First recall a lemma from class.

Lemma 3 *Suppose X_t is a martingale to F_t, and ϕ is a convex function. Then $\phi(X_t) = Y_t$ is a submartingale relative to F_t if $\phi(X_t)$ is integrable.*

This lemma comes directly from Jensen's lemma.

Now let's start the proof of Kolmogrov's inequality.

Proof. First, it's clear that

$$
\begin{aligned}
E[S_{n+1}|\sigma(S_n)] &= E[X_1 + X_2 + \cdots + X_{n+1}|\sigma(X_1, X_2, \cdots, X_n)] &(7)\\
&= X_1 + X_2 + \cdots + X_n + E[X_{n+1}|\sigma(X_1, X_2, \cdots, X_n)] &(8)\\
&= X_1 + X_2 + \cdots + X_n + E[X_{n+1}] &(9)\\
&= X_1 + X_2 + \cdots + X_n &(10)\\
&= S_n &(11)
\end{aligned}
$$

So S_n is a martingale. Let $\phi(x) = x^2$, ϕ is convex, then S_n^2 is a submartingale by the lemma above. Use Doob's martingale inequality,

$$\lambda^2 P(\max_{k \le n} S_k^2 \ge \lambda^2) \le E[S_n^2] \tag{12}$$

Also, we have

$$P(\max_{k \le n} S_k^2 \ge \lambda^2) = P(\max_{k \le n} |S_k| \ge \lambda) \tag{13}$$

So

$$P(\max_{k \le n} |S_k| \ge \lambda) \le \frac{1}{\lambda^2} E[S_n^2] = \frac{1}{\lambda^2} Var[S_n] \tag{14}$$

■

2.2 Apply Doob's martingale inequality to Brownian motion

Brownian motion, or pedesis , is the random motion of particles suspended in a fluid (a liquid or a gas) resulting from their collision with the fast-moving molecules in the fluid. In mathematics, Brownian motion is described by the Wiener process, a continuous-time stochastic process named in honor of Norbert Wiener. It is one of the best known Lévy processes (càdlàg stochastic processes with stationary independent increments) and occurs frequently in pure and applied mathematics, economics and physics.[4] The Wiener process W_t is characterized by four facts:

- $W_0 = 0$
- W_t is almost surely continuous
- W_t has independent increments
- $W_t - W_s \sim N(0, t - s)$ (for $0 \le s \le t$)

$N(\mu, \sigma^2)$ denotes the normal distribution with expected value μ and variance σ^2. The condition that it has independent increments means that if $0 \le s_1 < t_1 \le s_2 < t_2$ then $W_{t_1} - W_{s_1}$ and $W_{t_2} - W_{s_2}$ are independent random variables. The Wiener process can be constructed as the scaling limit of a random walk, or other discrete-time stochastic processes with stationary independent increments.

The problem can be modeled as below:

Problem 4 *Suppose we have a sequence of i.i.d. zero-centered Gaussian random variables X_t with variance σ^2 . Let $S_n = \sum_{t=1}^{n} X_t$. How to calculate $P(\max_{1 \le t \le n} S_t > \epsilon)$?*

Solution. Recall some properties of Gaussian distribution:
Suppose X has a zero-centered Gaussian distribution, then an easy computation shows that

$$E[e^{\lambda X}] = e^{\frac{\sigma^2\lambda^2}{2}} \tag{15}$$

for any $\lambda \in \mathbb{R}$. By extended version of Markov Inequality, we get

$$P(X > \epsilon) \leq e^{-\lambda\epsilon} E[e^{\lambda X}] = e^{-\lambda\epsilon} e^{\frac{\sigma^2\lambda^2}{2}} = e^{-\lambda\epsilon + \frac{\sigma^2\lambda^2}{2}} \tag{16}$$

Take $\lambda = \frac{\epsilon}{\sigma^2}$ to minimize right hand side. Then we get

$$P(X > \epsilon) \leq e^{-\frac{\epsilon^2}{2\sigma^2}} \tag{17}$$

A naive approach to this problem is to simply apply this property to S_t: It's easy to see that S_t has a gaussian distribution $N(0, t\sigma^2)$. So

$$P(S_t > \epsilon) \leq e^{-\frac{\epsilon^2}{2t\sigma^2}} \tag{18}$$

Then, we get

$$P(\max_{1\leq t\leq n} S_t > \epsilon) \leq \sum_{t=1}^{n} e^{-\frac{\epsilon^2}{2t\sigma^2}} \leq ne^{-\frac{\epsilon^2}{2n\sigma^2}} \tag{19}$$

The upper bound is quite loose, and we'll get a tighter bound with Doob's martingale inequality. From the class, we know random walk S_t is a martingale. Since $\phi(x) = e^{\lambda x}$ is convex for any $\lambda \geq 0$, $e^{\lambda S_t}$ is a submartingale.

$$P(\max_{1\leq t\leq n} S_t > \epsilon) = P(\max_{1\leq t\leq n} e^{\lambda S_t} > e^{\lambda\epsilon}) \text{ for } \lambda \geq 0 \tag{20}$$

Using Doob's martingale inequality,

$$P(\max_{1\leq t\leq n} e^{\lambda S_t} > e^{\lambda\epsilon}) \leq \frac{E[e^{\lambda S_n}]}{e^{\lambda\epsilon}} \tag{21}$$

$$= e^{\frac{1}{2}\lambda^2\sigma^2 n - \lambda\epsilon} \tag{22}$$

Take $\lambda = \frac{\epsilon}{\sigma^2 n}$ to minimize the right hand side.
Finally we get

$$P(\max_{1\leq t\leq n} S_t > \epsilon) \leq e^{-\frac{\epsilon^2}{2\sigma^2 n}} \tag{23}$$

The bound is far better than the bound calculated by the naive approach. ■

3 Doob's L^p maximal inequality

Doob's L^p maximal inequality is a corollary of the Doob's martingale inequality.

Theorem 5 *Let X_n be a non-negative bounded submartingale in L^p ($p > 1$), then*

$$||X_n^*||_p \leq q||X_n||_p \tag{24}$$

, where $X_n^* = \max_{0\leq k\leq n} X_k$ and $p^{-1} + q^{-1} = 1$.

Proof.

$$
\begin{aligned}
E[(X_n^*)^p] &= \int_\infty^0 t^p dP(X_n^* > t) && (25)\\
&= 0 + \int_0^\infty pt^{p-1} P(X_n^* > t) dt && (26)\\
&\le \int_0^\infty pt^{p-2} E[X_n 1_{X_n^* \ge t}] dt \qquad \text{(use Doob's martingale inequality)} && (27)\\
&= E(\int_0^\infty pt^{p-2} X_n 1_{X_n^* \ge t} dt) && (28)\\
&= E(\int_0^{X_n^*} pt^{p-2} X_n dt) && (29)\\
&= \frac{p}{p-1} E(X_n (X_n^*)^{p-1}) && (30)\\
& && (31)
\end{aligned}
$$

Thus, by Holder's inequality,

$$||X_n^*||_p^p = E[(X_n^*)^p] \le \frac{p}{p-1} E[X_n (X_n^*)^{p-1}] \le q||X_n||_p \cdot ||(X_n^*)^{p-1}||_q = q||X_n||_p \cdot ||(X_n^*)||_p^{p-1} \tag{32}$$

So we can derive that

$$||X_n^*||_p \le q||X_n||_p \tag{33}$$

■

3.1 A tighter bound

Dubins and Gilat [1] showed that the constant q is optimal, that is, cannot be replaced by a strictly smaller constant. It is also natural to ask whether equality can be attained. It turns out that this happens only in the trivial case $X_n = 0$. Otherwise, the inequality is strict. When it comes to continuous martingale, the bound can be tighter.[3]

$$E((X_T^*)^p) + \frac{p-1}{p} X_0^p \le (\frac{p}{p-1})^p E(X_T^p) \tag{34}$$

In contrary, the original Doob L^p inequality tells that

$$E((X_T^*)^p) \le (\frac{p}{p-1})^p E(X_T^p) \tag{35}$$

To start the proof, we'll need an identity:

Lemma 6 *Let X_t be a non-negative continuous martingale. For any $p > 0$, if*

$$\int_0^T (X_t^*)^{2p} d[X]_t < \infty \tag{36}$$

, where $[X]_t$ is X_t's quadratic variation, then,

$$E(X_T (X_T^*)^p) = \frac{p}{p+1} E((X_T^*)^{p+1}) + \frac{1}{p+1} X_0^{p+1} \tag{37}$$

Proof. We find there are 2 ways to represent $\int_0^T d(X_t(X_t^*)^p)$. One way is just to calculate directly: $X_T(X_T^*)^p - X_0^{p+1}$. Another way is to consider the following differential identity:

$$d(X_t(X_t^*)^p) = (X_t^*)^p dX_t + p(X_t^*)^{p-1} X_t dX_t^* \tag{38}$$

, then split the integral into 2 integrals. This can be stated as:

$$X_T(X_T^*)^p - X_0^{p+1} = \int_0^T d(X_t(X_t^*)^p) = \int_0^T (X_t^*)^p dX_t + \int_0^T p(X_t^*)^{p-1} X_t dX_t^* \tag{39}$$

Since we can easily observe that $dX_t^* \neq 0 \Rightarrow X_t = X_t^*$, we can calculate that

$$X_T(X_T^*)^p - X_0^{p+1} = \int_0^T (X_t^*)^p dX_t + \int_0^T p(X_t^*)^p dX_t^* \tag{40}$$

From the assumption that $\int_0^T (X_t^*)^{2p} d[X]_t < \infty$ and that X_t is a martingale, we can derive that $\int_0^T (X_t^*)^p dX_t$ is a martingale because the Itô integral preserves the local martingale property. So we have

$$E(\int_0^T (X_t^*)^p dX_t) = 0 \tag{41}$$

After taking expectations to both sides of (40), we get

$$E(X_T(X_T^*)^p) = \frac{p}{p+1} E((X_T^*)^{p+1}) + \frac{1}{p+1} X_0^{p+1} \tag{42}$$

, which finishes the proof. ■

Lemma 7 *Let X_t be a non-negative continuous submartingale. For any $p > 0$, if*

$$\int_0^T (X_t^*)^{2p} d[X]_t < \infty \tag{43}$$

, then,

$$E(X_T(X_T^*)^p) \geq \frac{p}{p+1} E((X_T^*)^{p+1}) + \frac{1}{p+1} X_0^{p+1} \tag{44}$$

Proof. The proof is almost the same as Lemma 6, except that $\int_0^T (X_t^*)^p dX_t$ is a submartingale. So

$$E(\int_0^T (X_t^*)^p dX_t) \geq 0 \tag{45}$$

■

Having these lemmas, it's easy to prove the bound.

Theorem 8 *Let X_t be a non-negative continuous submartingale. For any $p > 0$, if*

$$\int_0^T (X_t^*)^{2p} d[X]_t < \infty \tag{46}$$

, then we have :

$$E((X_T^*)^p) + \frac{p-1}{p} X_0^p \leq (\frac{p}{p-1})^p E(X_T^p) \tag{47}$$

Proof. First, we use Holder inequality:

$$E((X_T^*)^p X_T) \leq E(X_T^{p+1})^{\frac{1}{p+1}} E((X_T^*)^{p+1})^{\frac{p}{p+1}} \tag{48}$$
$$= (\epsilon^{-p} E(X_T^{p+1}))^{\frac{1}{p+1}} (\epsilon E((X_T^*)^{p+1}))^{\frac{p}{p+1}} \tag{49}$$

Then, use Young's inequality for products, which states that

$$ab \leq \frac{a^p}{p} + \frac{b^p}{p} \tag{50}$$

Here we use as

$$a^{\frac{1}{p+1}} b^{\frac{p}{p+1}} \leq \frac{1}{p+1} a + \frac{p}{p+1} b \tag{51}$$

We get

$$E((X_T^*)^p X_T) \leq (\epsilon^{-p} E(X_T^{p+1}))^{\frac{1}{p+1}} (\epsilon E((X_T^*)^{p+1}))^{\frac{p}{p+1}} \leq \frac{\epsilon^{-p}}{p+1} E(X_T^{p+1}) + \frac{p\epsilon}{p+1} E((X_T^*)^{p+1}) \tag{52}$$

Now use Lemma 7,

$$\frac{p}{p+1} E((X_T^*)^{p+1}) + \frac{1}{p+1} X_0^{p+1} \leq \frac{\epsilon^{-p}}{p+1} E(X_T^{p+1}) + \frac{p\epsilon}{p+1} E((X_T^*)^{p+1}) \tag{53}$$

Rearranging the formula, we have

$$E((X_T^*)^{p+1}) + \frac{1}{p(1-\epsilon)} X_0^{p+1} \leq \frac{p+1}{p} \frac{1}{\epsilon^p (1-\epsilon)(p+1)} E(X_T^{p+1}) \tag{54}$$

To minimize the right hand side, we let $\epsilon = \frac{p}{p+1}$, then

$$E((X_T^*)^{p+1}) + \frac{p+1}{p} X_0^{p+1} \leq (\frac{p+1}{p})^{p+1} E(X_T^{p+1}) \tag{55}$$

This finishes the proof. ■

References

[1] Lester Dubins and David Gilat. On the distribution of maxima of martingale. *Proceedings of The American Mathematical Society - PROC AMER MATH SOC*, 68:337–337, 03 1978.

[2] MacTutor History of Mathematics Archive. Joseph leo doob.

[3] Jian Sun. Generalization of doob's inequality and a tighter estimate on look-back option price, 2018.

[4] Wikipedia. Brownian motion.

[5] Wikipedia. Doob's martingale inequality.