

Theory and Applications of Stochastic Systems

Lecture 10

B60.4312.01

1 Exponential Martingale for Random Walk

Let $(S_n : n \geq 0)$ be a random walk with i.i.d. increments $X = (X_n : n \geq 1)$ and $S_0 = 0$.

We showed earlier that in the mean-zero and finite variance case, the sequences $(S_n : n \geq 0)$, $(S_n^2 - n\sigma^2 : n \geq 0)$ are martingales with respect to X . In this section, we will introduce the exponential martingale which could be viewed as a generalization of those martingales.

We call for $\theta \in \mathbb{R}$, $\phi(\theta) = E \exp(\theta X_1)$, the moment generating function, and $\psi(\theta) = \log \phi(\theta)$, the log-moment generating function known also as the cumulant generating function. Now define the set $\Theta = \{\theta \in \mathbb{R} : \phi(\theta) < \infty\}$. For every $\theta \in \Theta$ define for all $n \geq 0$,

$$M_n(\theta) = \exp(\theta S_n - n\psi(\theta)). \quad (1)$$

Proposition 1.1 *The sequence $(M_n(\theta) : n \geq 0)$ is a martingale (the so-called exponential martingale for random walk).*

Note that for M to be a martingale, S does not need to be a mean-zero random walk. The proof is left as an exercise.

1.1 Properties of ϕ and ψ

One main ingredient of the previous martingale is the moment generating function or its logarithm. Such functions are well behaved. In particular, $\phi(0) = 1$ (or $\psi(0) = 0$) and $\phi'(0) = \psi'(0) = EX_1$. Note also that for all $\theta \in \Theta$, $\phi(\theta) > 0$ is always true and equivalently $\psi(\theta) > -\infty$. However, $\psi(\theta)$ can very well take the value $+\infty$.

Proposition 1.2 *If $\psi(\theta_0) < \infty$, for some $\theta_0 > 0$ ($\theta_0 < 0$), then $\psi(\theta) < \infty$, for all $\theta \in [0, \theta_0]$ ($[\theta_0, 0]$).*

Proof. Assume that $\psi(\theta_0) < \infty$ for $\theta_0 > 0$ and let $\theta \in [0, \theta_0]$. We have,

$$\begin{aligned}\psi(\theta) &= \mathbb{E}[\exp(\theta X_1)I(X_1 > 0)] + \mathbb{E}[\exp(\theta X_1)I(X_1 \leq 0)] \\ &\leq \mathbb{E}(\exp \theta_0 X_1) \\ &< \infty.\end{aligned}$$

A similar proof holds when $\theta_0 < 0$. □

From the previous proposition we conclude that Θ is an interval of either of the following forms: $[a, b]$, (a, b) , $[a, b)$, $(a, b]$ with $a \leq 0 \leq b$. Finally, one can also show that ψ is a convex function of θ and infinitely differentiable (i.e. $\psi \in C^\infty$) on the interior of Θ (i.e. on (a, b)).

We end this section by a theorem that when applying it one can retrieve the Wald's moments martingales $(S_n : n \geq 0)$, $(S_n^2 - n\sigma^2 : n \geq 0)$ etc...

Theorem 1.3 *Suppose $\Theta = (a, b)$ and $a < b$. Under the same conditions as before, $\forall \theta \in \Theta$, $(M_n^{(k)}(\theta) : n \geq 0)$ is a martingale where,*

$$M_n(\theta) = \exp[\theta S_n - n\psi(\theta)] \text{ and } M_n^{(k)}(\theta) = \frac{d^k}{d\theta^k} M_n(\theta).$$

2 Change-of-Measure

Change-of-measure method is a powerful tool of modern probability theory. The idea is basically to define a new probability space in a suitable way, under which one can obtain more tractable expressions or a just a more convenient setting to work with. The so-called likelihood ration will allow one to convert the results into the initial probability measure.

2.1 Construction of a New Probability Measure

Under the setting introduced earlier, let $\theta \in \Theta$, such that $\theta > 0$ and F be the distribution of X_1 . Define,

$$F_\theta(dx) = \exp(\theta x - \psi(\theta))F(dx).$$

F_θ is indeed a probability distribution ($\int F_\theta(dx) = 1$) and the sequence $(F_\theta : \theta \in \Theta)$ is the exponential family generated by F . To be more specific, let $\Omega = \mathbb{R}^\infty (= \mathbb{R} \times \mathbb{R} \times \dots)$ be the sample space supporting the random walk. For $\omega = (x_0, x_1, \dots)$, let $X_i(\omega) = x_i$ for $i \geq 0$. For $\theta \in \Theta$, let P_θ be the probability measure defined on Ω such that

$$\mathbb{P}_\theta(dx) = \mathbb{P}_\theta(X_0 \in dx_1, \dots, X_n \in dx_n) = \prod_{i=0}^n F_\theta(dx_i).$$

As usually, one needs to invoke an appropriate “extension theorem” from measure theory to guarantee the existence of a unique probability measure $\mathbb{P}_\theta$ defined as above. In other words, the X_i ’s are i.i.d. with common distribution F_θ under $\mathbb{P}_\theta$. It is helpful to note that under the new measure,

$$\mathbb{E}_\theta X_1 = \psi'(\theta),$$

where ψ is as above the log-moment generating function under the initial distribution.

Now let A be an event generated by the n first random variables $X_1, X_2, \dots, X_n$. We can write $I(A) = f_n(X_1, X_2, \dots, X_n)$ and we obtain

$$\begin{aligned}
\mathbb{P}_\theta(A) &= \int_A \mathbb{P}_\theta(dx) = \int_{\mathbb{R}} \int_{\mathbb{R}} \dots \int_{\mathbb{R}} f_n(x_1, x_2, \dots, x_n) F_\theta(dx_1) F_\theta(dx_2) \dots F_\theta(dx_n) \\
&= \int_{\mathbb{R}} \int_{\mathbb{R}} \dots \int_{\mathbb{R}} \exp(\theta s_n - n\psi(\theta)) f_n(x_1, x_2, \dots, x_n) F(dx_1) F(dx_2) \dots F(dx_n) \\
&= \int_A \exp(\theta s_n - n\psi(\theta)) \mathbb{P}_\theta(dx) \\
&= \mathbb{E}[\exp(\theta S_n - n\psi(\theta)) I(A)].
\end{aligned}$$

We should note here that it is crucial that $M_n(\theta) = \exp[\theta S_n - n\psi(\theta)]$ defines a martingale sequence, to be able to perform a change-of-measure. It is needed to preserve the consistency when enlarging the filtration. We argue this point when discussing later more general change-of-measure.

2.2 Likelihood Ratio Identity (also called Wald's Sequential Identity)

In this part we introduce the most general formulation of the change-of-measure formula. Let τ be a stopping time and W be a nonnegative random variable that is " $\mathcal{F}_\tau$ -measurable", in the sense that W depends on the history up to time τ . It can be written as $W = \sum_{i=1}^{\infty} f_i(X_1, \dots, X_i) I(\tau = i)$, where the f_i 's are deterministic functions. The following is the likelihood ratio identity:

$$\mathbb{E}[W; \tau < \infty] = \mathbb{E}_\theta[\exp(-\theta S_\tau + \tau\psi(\theta)) W; \tau < \infty],$$

and similarly one can obtain that

$$\mathbb{E}_\theta[W; \tau < \infty] = \mathbb{E}[\exp(\theta S_\tau - \tau\psi(\theta)) W; \tau < \infty].$$

Finally, note that by taking $\tau = n$, $W = I(A)$ where A is an event on $\mathcal{F}_n$, we obtain the expression computed at the end of the previous section.

3 Cramér-Lundberg Approximation

We have already seen that the all time maximum of a random walk is a crucial quantity for many applications. Hence, obtaining approximations of the tail distribution of such a quantity is of high interest.

Let $S = (S_n : n \geq 0)$ be a random walk with increments $X = (X_n : n \geq 0)$ such that $\mathbb{E}X_1 < 0$. Form the quantity $M^* = \max_{n \geq 0} S_n < \infty$ a.s. The goal is to obtain an approximation of $\mathbb{P}(M^* > u)$ as u takes large values.

We start by observing that,

$$\mathbb{P}(M^* > u) = \mathbb{P}(T(u) < \infty),$$

where $T(u) = \inf\{n : S_n > u\}$ is a stopping time. Applying the likelihood ratio identity to the event $\{T(u) < \infty\}$, we get

$$\mathbb{P}(T(u) < \infty) = \mathbb{E}I(T(u) < \infty) = \mathbb{E}_\theta[\exp(-\theta S_{T(u)} + T(u)\psi(\theta)); T(u) < \infty].$$

We are assuming here that there exists a non-zero θ for which $\psi(\theta) < \infty$. But what if this θ can be chosen in a particular way so that things get more tractable? Specifically, for the rest we assume the following.

A1. There exists $\theta^* > 0$, s.t. $\psi(\theta^*) = 0$

Note that if it exists, θ^* must be unique due to the convexity of ψ . It turns out that assumption A1 is not very restrictive and holds in most interesting cases. Under A1, the quantity of interest becomes,

$$P(T(u) < \infty) = E_{\theta^*}[\exp(-\theta^* S_{T(u)}; T(u) < \infty].$$

We note that $\psi(0) = \psi(\theta^*) = 0$. Hence, by convexity of ψ , $\psi'(0) = \mathbb{E}X_1$ being negative, $\psi'(\theta^*) = \mathbb{E}_{\theta^*}X_1$ must be positive. So that the random walk under the probability

measure induced by θ^* has a positive drift. Hence, $T(u) < \infty$ $\mathbb{P}_{\theta^*}$ a.s. and the tail distribution of M^* is given by

$$\mathbb{P}(M^* > u) = \mathbb{E}_{\theta^*}[\exp(-\theta^* S_{T(u)})]. \quad (2)$$

From the last expression and the definition of $T(u)$, we can first obtain a bound (known as the Kingman's bound).

$$\mathbb{P}(M^* > u) \leq \exp(-\theta^* u).$$

But back to equation (2), by introducing the term $\exp(-\theta^* u)$ on the right hand side, we obtain

$$\mathbb{P}(M^* > u) = \exp(-\theta^* u) \mathbb{E}_{\theta^*}[\exp(-\theta^* (S_{T(u)} - u))].$$

We call $B(u) = S_{T(u)} - u$ and recognize it as the forward recurrence time associated with the random walk S . From renewal theory, we know that under mild assumptions, $B(u) \Rightarrow B(\infty)$ as u goes to infinity. We then conclude under such assumptions, and using the Bounded Convergence Theorem, that there exists $C > 0$ such that $\mathbb{E}_{\theta^*}[\exp(-\theta^* (S_{T(u)} - u)) \rightarrow C$ as $u \rightarrow \infty$. Equivalently, we obtain the so-called Cramér-Lundberg approximation

$$\mathbb{P}(M^* > u) \sim C \exp(-\theta^* u) \quad (3)$$

as $u \rightarrow \infty$.

4 More General Change-of-Measure

Suppose that $(L_n : n \geq 0)$ is a nonnegative martingale adapted to $Z = (Z_n : n \geq 0)$ under which $EL_n = 1$. We can always define for $A \in \mathcal{F}_n$ the following probability

$$\mathbb{P}_n(A) = \mathbb{E}[L_n I(A)],$$

since $P_n(\Omega) = \mathbb{E}[L_n I(\Omega)] = \mathbb{E}[L_n] = 1$. It is desirable though that the probabilities $(P_n : n \geq 0)$ were compatible with one another, in the sense that if an event $A \in \mathcal{F}_n$ (i.e. defined by the X_i 's up to n), then $P_n(A) = P_{n+1}(A)$. This is indeed the case due to the martingale property of the L_n 's.

$$\begin{aligned}
P_{n+1}(A) &= \mathbb{E}L_{n+1}I(A) \\
&= \mathbb{E}[\mathbb{E}[L_{n+1}I(A)|\mathcal{F}_n]] \\
&= \mathbb{E}[I(A)\mathbb{E}[L_{n+1}|\mathcal{F}_n]] \\
&= \mathbb{E}[I(A)L_n] \\
&= P_n(A).
\end{aligned}$$

Since the P_n 's are compatible one can again invoke an extension theorem to guarantee the existence of a unique probability measure $\tilde{\mathbb{P}}$ such that,

$$\tilde{\mathbb{E}}[W; \tau < \infty] = \mathbb{E}[WL_\tau; \tau < \infty],$$

where τ is a stopping time and W is a non-negative random variable adapted to the Z_i 's up to time τ (i.e. $W \in \mathcal{F}_\tau$).

5 Exponential Martingale for Markov random walk

In this section we construct an exponential martingale for a more general class of random walks than the one studied earlier. In particular we consider an irreducible Markov chain $X = (X_n : n \geq 0)$ with finite space $\mathcal{S}$, and a non-negative function f . We form the stochastic process $S = (S_n : n \geq 0)$, where $S_n = \sum_{i=0}^n f(X_i)$. Observe that the increments of the random walk S are not independent as they are governed by the Markov chain X . (e.g. Markov modulated processes)

Recall that in the case where the increments form an i.i.d. sequence, the moment generating function of the random walk was easy to compute as a function of the

moment generating function of the increments. This is not the case here where we need to define the function ψ in a more general way as follows:

$$\exp \psi(\theta) = \lim_{n \rightarrow \infty} (\mathbb{E} \exp \theta S_n)^{1/n}.$$

$$\begin{aligned} \mathbb{E}_x \exp \theta S_n &= \sum_{x_1, \dots, x_n} \exp(\theta f(x_1)) \dots \exp(\theta f(x_n)) P(x, x_1) \dots P(x_{n-1}, x_n) \\ &= \sum_{x_1, \dots, x_n} \prod_{i=1}^{n-1} \exp(\theta f(x_i)) P(x_{i-1}, x_i) \\ &= (K_\theta^n \cdot \mathbf{e})(x), \end{aligned}$$

where K is the $|\mathcal{S}| \times |\mathcal{S}|$ matrix such that

$$K_\theta(x, y) = e^{\theta f(y)} P(x, y),$$

and $\mathbf{e}$ is a vector with ones as coordinates.

One expects K^n to grow as a function of n at the rate of its largest eigenvalue (in modulus). A linear algebra result will be here of great importance.

Theorem 5.1 (*Perron-Frobenius*) *Suppose B is a nonnegative irreducible matrix, then there exists a positive eigenvalue and a strictly positive column and row vector h and ϕ such that*

$$Bh = \lambda h, \text{ and } \phi B = \lambda \phi.$$

Furthermore, λ is the spectral radius of B .

We apply Perron-Frobenius Theorem to K_θ . For a fix θ , let $\lambda(\theta) > 0$ the corresponding eigenvalue given by this Theorem. We can always write $\lambda(\theta) = \exp \psi(\theta)$. By definition of h we have,

$$\sum_y K_\theta(x, y) h(\theta, y) = \lambda(\theta) h(\theta, x), \forall x \in S.$$

We obtain that

$$\begin{aligned}
\sum_y \exp(\theta f(x)) P(x, y) h(\theta, y) \frac{1}{h(\theta, x)} \exp(-\psi(\theta)) &= 1, \\
\exp(\theta f(x) - \psi(\theta)) \frac{1}{h(\theta, x)} \sum_y P(x, y) h(\theta, y) &= 1, \\
\Rightarrow e^{\theta f(x) - \psi(\theta)} \frac{1}{h(\theta, x)} E_x h(\theta, X_1) &= 1.
\end{aligned}$$

The main result of this section is given by the following proposition

Proposition 5.2 *Let X be an irreducible Markov chain living on a finite state space $\mathcal{S}$. Then $\forall \theta \in \mathbb{R}, \exists \psi(\theta), h(\theta)$, s.t., $(M_n(\theta) : n \geq 0)$ is a martingale, where*

$$M_n(\theta) = \exp\left[\theta \sum_{i=1}^{n-1} f(X_i) - n\psi(\theta)\right] \frac{h(\theta, X_n)}{h(\theta, X_0)}.$$

Proof.

$$\begin{aligned}
\mathbb{E}[M_n(\theta) | X_0, \dots, X_{n-1}] &= \exp\left[\theta \sum_{i=1}^{n-1} f(X_i) - n\psi(\theta)\right] \frac{E_{X_{n-1}} h(\theta, X_n)}{h(\theta, X_0)} \\
&= \exp\left[\theta \sum_{i=1}^{n-2} f(X_i) - (n-1)\psi(\theta)\right] \exp[\theta f(X_{n-1}) - \psi(\theta)] \\
&\quad \cdot \frac{E_{X_{n-1}} h(\theta, X_n)}{h(\theta, X_0)}.
\end{aligned}$$

From the above derivation we have,

$$\mathbb{E}_{X_{n-1}} h(\theta, X_n) = \exp[-\theta f(X_{n-1}) + \psi(\theta)] h(\theta, X_{n-1}).$$

By putting together the last two equalities we complete our proof.