

Chapter 2

The Wald Martingale and the Maximum

In this chapter, we shall introduce the first of our two key martingales and consider two immediate applications. In the first application, we will use the martingale to construct a change of measure with respect to \mathbb{P} and thereby consider the dynamics of X under the new law. In the second application, we shall use the martingale to study the law of the process $\bar{X} = \{\bar{X}_t : t \geq 0\}$, where we recall that

$$\bar{X}_t = \sup_{s \leq t} X_s, \quad t \geq 0. \quad (2.1)$$

In particular, we shall discover that the position of the trajectory of \bar{X} , when sampled at an independent and exponentially distributed time, is again exponentially distributed.

2.1 Laplace Exponent

A key quantity in the forthcoming analysis is the *Laplace exponent* of the Cramér–Lundberg process, whose definition is contained in the following lemma.

Lemma 2.1 *For all $\theta \geq 0$ and $t \geq 0$, we have*

$$\mathbb{E}(e^{\theta X_t}) = \exp\{\psi(\theta)t\},$$

where the Laplace exponent ψ satisfies

$$\psi(\theta) := c\theta - \lambda \int_{(0,\infty)} (1 - e^{-\theta x}) F(dx). \quad (2.2)$$

Proof Given the definition (1.1), one easily sees that it suffices to prove that

$$\mathbb{E}(e^{-\theta \sum_{i=1}^{N_t} \xi_i}) = \exp\left\{-\lambda t \int_{(0,\infty)} (1 - e^{-\theta x}) F(dx)\right\}, \quad (2.3)$$

for $\theta, t \geq 0$. To establish (2.3), we make use of the fact that N_t is independent of $\{\xi_i : i \geq 1\}$ and Poisson distributed with rate λt . More precisely,

$$\begin{aligned} \mathbb{E}(e^{-\theta \sum_{i=1}^{N_t} \xi_i}) &= \sum_{n=0}^{\infty} \mathbb{E}(e^{-\theta \sum_{i=1}^n \xi_i}) e^{-\lambda t} \frac{(\lambda t)^n}{n!} \\ &= \sum_{n=0}^{\infty} [\mathbb{E}(e^{-\theta \xi_1})]^n e^{-\lambda t} \frac{(\lambda t)^n}{n!} \\ &= \exp\{-\lambda t (1 - \mathbb{E}(e^{-\theta \xi_1}))\} \\ &= \exp\left\{-\lambda t \int_{(0, \infty)} (1 - e^{-\theta x}) F(dx)\right\}, \end{aligned}$$

for all $\theta, t \geq 0$. □

As we shall see, the Laplace exponent (2.2) is used as a way of identifying certain characteristics of Cramér–Lundberg processes. To this end, let us start by looking at the shape of (2.2). Straightforward differentiation, with the help of the Dominated Convergence Theorem, tells us that, for all $\theta > 0$,

$$\psi''(\theta) = \lambda \int_{(0, \infty)} x^2 e^{-\theta x} F(dx) > 0,$$

which in turn implies that ψ is strictly convex on $(0, \infty)$. Integration by parts allows us to write

$$\psi(\theta) = c\theta - \lambda\theta \int_{(0, \infty)} e^{-\theta x} \overline{F}(x) dx, \quad \theta \geq 0, \quad (2.4)$$

where $\overline{F}(x) := 1 - F(x)$, $x \geq 0$. Moreover, this representation allows us to deduce that

$$\lim_{\theta \rightarrow \infty} \frac{\psi(\theta)}{\theta} = c$$

and

$$\psi'(0+) = \lim_{\theta \rightarrow 0} \frac{\psi(\theta)}{\theta} = c - \lambda\mu = \mathbb{E}(X_1),$$

where the left-hand side is the right derivative of ψ at the origin, the final equality follows from (1.4) and we recall that $\mu := \int_{(0, \infty)} x F(dx) \in (0, \infty]$. The security loading condition (1.3) can thus be alternatively expressed simply as $\psi'(0+) > 0$.

A quantity which will also repeatedly appear in our computations is the right inverse of ψ . That is,

$$\Phi(q) := \sup\{\theta \geq 0 : \psi(\theta) = q\}, \quad (2.5)$$

for $q \geq 0$. Thanks to the strict convexity of ψ and that $\lim_{\theta \rightarrow \infty} \psi(\theta) = \infty$, we can say that there is exactly one solution in $[0, \infty)$ to the equation $\psi(\theta) = q$, when $q > 0$, and at most two when $q = 0$. The number of solutions in the latter of these

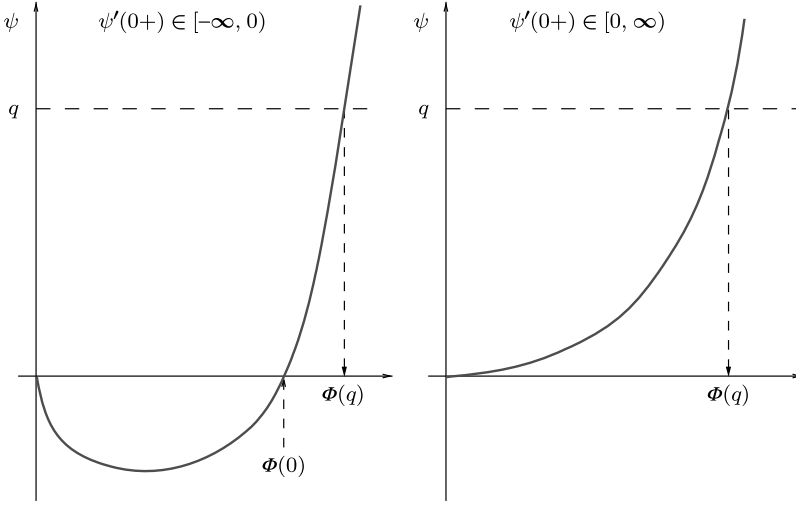


Fig. 2.1 Two examples of ψ , the Laplace exponent of a Cramér–Lundberg process, corresponding to the cases $\psi'(0+) < 0$ and $\psi'(0+) \geq 0$, respectively

two cases depends on the value of $\psi'(0+)$. Indeed, when $\psi'(0+) \geq 0$, then $\theta = 0$ is the only solution to $\psi(\theta) = 0$. When $\psi'(0+) < 0$, there are two solutions, one at $\theta = 0$ and a second solution, in $(0, \infty)$, which, by definition, gives the value of $\Phi(0)$; see Fig. 2.1.

2.2 First Exponential Martingale

For each $\beta > 0$, define the process $\mathcal{E}(\beta) = \{\mathcal{E}_t(\beta) : t \geq 0\}$ by

$$\mathcal{E}_t(\beta) := e^{\beta X_t - \psi(\beta)t}, \quad t \geq 0. \quad (2.6)$$

Theorem 2.2 Fix $\beta > 0$. The process $\mathcal{E}(\beta)$ is a \mathbb{P} -martingale with respect to \mathbb{F} .

Proof Note that the process $\mathcal{E}(\beta)$ is \mathbb{F} -adapted. With this in hand, it suffices to check that, for all $\beta > 0$ and $s, t \geq 0$, $\mathbb{E}[\mathcal{E}_{t+s}(\beta) | \mathcal{F}_t] = \mathcal{E}_t(\beta)$. On account of positivity, this would immediately show that $\mathbb{E}[|\mathcal{E}_t(\beta)|] < \infty$, for all $t \geq 0$, which is also required for $\mathcal{E}(\beta)$ to be a martingale.

Thanks to stationary and independent increments, \mathcal{F} -adaptedness as well as Lemma 2.1, for all $\beta, s, t \geq 0$,

$$\begin{aligned} \mathbb{E}[\mathcal{E}_{t+s}(\beta) | \mathcal{F}_t] &= \mathcal{E}_t(\beta) \mathbb{E}[e^{\beta(X_{t+s} - X_t) - \psi(\beta)s} | \mathcal{F}_t] \\ &= \mathcal{E}_t(\beta) \mathbb{E}[e^{\beta X_s}] e^{-\psi(\beta)s} \\ &= \mathcal{E}_t(\beta) \end{aligned}$$

and the proof is complete. \square

This martingale is known as the *Wald martingale*. See Sect. 2.5 for further historical details.

2.3 Esscher Transform

Fix $\beta > 0$ and $x \in \mathbb{R}$. Normalising $\mathcal{E}(\beta)$ by its expectation, we may use the resulting mean-one martingale to perform a change of measure on (X, \mathbb{P}_x) via

$$\left. \frac{d\mathbb{P}_x^\beta}{d\mathbb{P}_x} \right|_{\mathcal{F}_t} = \frac{\mathcal{E}_t(\beta)}{\mathcal{E}_0(\beta)} = e^{\beta(X_t - x) - \psi(\beta)t}, \quad t \geq 0. \quad (2.7)$$

In the special case that $x = 0$, we shall write \mathbb{P}^β in place of \mathbb{P}_0^β . Since the process X under \mathbb{P}_x may be written as $x + X$ under \mathbb{P} , it is not difficult to see that the change of measure on (X, \mathbb{P}_x) corresponds to the analogous change of measure on (X, \mathbb{P}) . Also known as the *Esscher transform*, (2.7) alters the law of X . It is related to the Esscher transform for random variables. For example, for the distribution F , its Esscher transform is the distribution

$$F_\beta(dx) := \frac{e^{-\beta x}}{m(\beta)} F(dx), \quad x > 0,$$

for some $\beta > 0$, where $m(\beta) = \int_{(0, \infty)} e^{-\beta x} F(dx)$. For the forthcoming computations, it is important that we understand the dynamics of X under \mathbb{P}^β .

Theorem 2.3 *Fix $\beta > 0$. The process (X, \mathbb{P}^β) is equal in law to a Cramér–Lundberg process with premium rate c and claims that arrive at rate $\lambda m(\beta)$ with common distribution F_β . Said another way, the process (X, \mathbb{P}^β) is equal in law to X^β , where $X^\beta := \{X_t^\beta : t \geq 0\}$ is a Cramér–Lundberg process with Laplace exponent*

$$\psi_\beta(\theta) := \psi(\theta + \beta) - \psi(\beta), \quad \theta \geq 0.$$

Proof For all $0 \leq s \leq t \leq u < \infty$, $\theta \geq 0$ and $A \in \mathcal{F}_s$, with the help of stationary and independent increments of (X, \mathbb{P}) , we have that

$$\begin{aligned} \mathbb{E}^\beta[\mathbf{1}_A e^{\theta(X_u - X_t)}] &= \mathbb{E}[\mathbf{1}_A e^{\beta X_t - \psi(\beta)t} e^{(\theta + \beta)(X_u - X_t)}] e^{-\psi(\beta)(u-t)} \\ &= \mathbb{E}[\mathbf{1}_A e^{\beta X_t - \psi(\beta)t}] \mathbb{E}[e^{(\theta + \beta)X_{u-t}}] e^{-\psi(\beta)(u-t)} \\ &= \mathbb{E}[\mathbf{1}_A e^{\beta X_s - \psi(\beta)s}] \mathbb{E} e^{\psi(\beta + \theta)(u-t)} e^{-\psi(\beta)(u-t)} \\ &= \mathbb{P}^\beta(A) e^{\psi_\beta(\theta)(u-t)}, \end{aligned} \quad (2.8)$$

where in the second equality we have conditioned on \mathcal{F}_t and in the third equality we have conditioned on \mathcal{F}_s and used the martingale property of $\mathcal{E}(\beta)$. It now follows

from (2.8) that, for all $0 \leq v \leq s \leq t \leq u < \infty$ and $\theta_1, \theta_2 \geq 0$,

$$\mathbb{E}^\beta \left[e^{\theta_1(X_s - X_v)} e^{\theta_2(X_u - X_t)} \right] = e^{\psi_\beta(\theta_1)(s-v) + \psi_\beta(\theta_2)(u-t)}.$$

Using a straightforward argument by induction, again using (2.8), we also have that, for all $n \in \mathbb{N}$, $0 \leq s_1 \leq t_1 \leq \dots \leq s_n \leq t_n < \infty$ and $\theta_1, \dots, \theta_n \geq 0$,

$$\mathbb{E}^\beta \left[\prod_{j=1}^n e^{\theta_j(X_{t_j} - X_{s_j})} \right] = \prod_{j=1}^n e^{\psi_\beta(\theta_j)(t_j - s_j)}. \quad (2.9)$$

Moreover, a brief computation shows that

$$\psi_\beta(\theta) = c\theta - \lambda m(\beta) \int_{(0, \infty)} (1 - e^{-\theta x}) \frac{e^{-\beta x}}{m(\beta)} F(dx), \quad \theta \geq 0.$$

Coupled with (2.9), this shows that (X, \mathbb{P}^β) has stationary and independent increments which are equal in law to those of a Cramér–Lundberg process with premium rate c , arrival rate of claims $\lambda m(\beta)$ and distribution of claims $e^{-\beta x} F(dx)/m(\beta)$. Since the measures \mathbb{P}^β and \mathbb{P} are equivalent on \mathcal{F}_t , for all $t \geq 0$, then the property that X has paths that are almost surely right-continuous with left limits and no positive jumps on $[0, t]$ carries over to the measure \mathbb{P}^β . \square

The Esscher transform may also be formulated at stopping times.

Corollary 2.4 *Under the conditions of Theorem 2.3, if τ is an \mathbb{F} -stopping time, then*

$$\left. \frac{d\mathbb{P}^\beta}{d\mathbb{P}} \right|_{\mathcal{F}_\tau} = \mathcal{E}_\tau(\beta) \quad \text{on } \{\tau < \infty\}.$$

Said another way, for all $A \in \mathcal{F}_\tau$, we have

$$\mathbb{P}^\beta(A, \tau < \infty) = \mathbb{E}(\mathbf{1}_{(A, \tau < \infty)} \mathcal{E}_\tau(\beta)).$$

Proof By definition, if $A \in \mathcal{F}_\tau$, then $A \cap \{\tau \leq t\} \in \mathcal{F}_t$. Hence

$$\begin{aligned} \mathbb{P}^\beta(A \cap \{\tau \leq t\}) &= \mathbb{E}(\mathcal{E}_t(\beta) \mathbf{1}_{(A, \tau \leq t)}) \\ &= \mathbb{E}(\mathbf{1}_{(A, \tau \leq t)} \mathbb{E}(\mathcal{E}_t(\beta) | \mathcal{F}_\tau)) \\ &= \mathbb{E}(\mathcal{E}_\tau(\beta) \mathbf{1}_{(A, \tau \leq t)}), \end{aligned}$$

where in the third equality we have used the strong Markov property as well as the martingale property for $\mathcal{E}(\beta)$. Now taking limits as $t \rightarrow \infty$, the result follows with the help of the Monotone Convergence Theorem. \square

2.4 Distribution of the Maximum

We want to use the Esscher transform to characterise the law of the first passage times

$$\tau_x^+ := \inf\{t > 0 : X_t > x\},$$

for $x \geq 0$, and subsequently the law of the running maximum when sampled at an independent and exponentially distributed time. Note that the stopping time τ_x^+ may be infinite in value, depending on the long-term behaviour of the process X . Accordingly, in the theorem below, where τ_x^+ appears in an exponent, we will understand $e^{-\infty} := 0$.

Theorem 2.5 *For $x \geq 0$ and $q > 0$,*

$$\mathbb{E}(e^{-q\tau_x^+}) = e^{-\Phi(q)x},$$

where we recall that $\Phi(q)$ is given by (2.5). By taking limits as $q \rightarrow 0$, it also follows that

$$\mathbb{P}(\tau_x^+ < \infty) = e^{-\Phi(0)x},$$

for $x \geq 0$.

Proof Using the fact that X has no positive jumps, it must follow that $X_{\tau_x^+} = x$ on $\{\tau_x^+ < \infty\}$. With the help of the strong Markov property we have that

$$\begin{aligned} & \mathbb{E}(e^{\Phi(q)X_t - qt} | \mathcal{F}_{\tau_x^+}^+) \\ &= \mathbf{1}_{(\tau_x^+ \geq t)} e^{\Phi(q)X_t - qt} + \mathbf{1}_{(\tau_x^+ < t)} e^{\Phi(q)x - q\tau_x^+} \mathbb{E}(e^{\Phi(q)(X_t - X_{\tau_x^+}) - q(t - \tau_x^+)} | \mathcal{F}_{\tau_x^+}^+) \\ &= e^{\Phi(q)X_{t \wedge \tau_x^+} - q(t \wedge \tau_x^+)}, \end{aligned} \tag{2.10}$$

where, in the final equality, we have used the fact that $\mathbb{E}(\mathcal{E}_t(\Phi(q))) = 1$ for all $t \geq 0$. Using this fact again together with the law of total probability, we get, by taking expectations again in (2.10),

$$\mathbb{E}(e^{\Phi(q)X_{t \wedge \tau_x^+} - q(t \wedge \tau_x^+)}) = 1.$$

Noting that the expression in the latter expectation is bounded above by $e^{\Phi(q)x}$, an application of dominated convergence yields

$$\mathbb{E}(e^{\Phi(q)x - q\tau_x^+}) = 1,$$

which is equivalent to the statement of the theorem. \square

We recover the promised distributional information about the maximum process (2.1) in the next corollary. In its statement, we understand an exponential random variable with rate 0 to be infinite in value with probability one.

Corollary 2.6 *Fix $q \geq 0$ and let \mathbf{e}_q be an exponentially distributed random variable with rate q , which is independent of X . Then $\overline{X}_{\mathbf{e}_q}$ is exponentially distributed with parameter $\Phi(q)$.*

Proof First suppose that $q > 0$. The result is an easy consequence of the fact that

$$\mathbb{P}(\overline{X}_{\mathbf{e}_q} > x) = \mathbb{P}(\tau_x^+ < \mathbf{e}_q) = \mathbb{E}\left(\int_0^\infty q e^{-qt} \mathbf{1}_{(\tau_x^+ < t)}\right) = \mathbb{E}(e^{-q\tau_x^+}),$$

together with the conclusion of Theorem 2.5. For the remaining case that $q = 0$, note with the help of the last part of Theorem 2.5 that

$$\mathbb{P}(\overline{X}_\infty > x) = \mathbb{P}(\tau_x^+ < \infty) = e^{-\Phi(0)x},$$

and the proof is complete. \square

2.5 Comments

The idea of *tilting* a distribution by exponentially weighting its probability distribution function was introduced by Esscher (1932). This idea lends itself well to changes of measure in the theory of stochastic processes, in particular for Lévy processes. The Wald martingale can be traced back to Wald (1944, 1945). The associated Esscher transform is analogous to the exponential martingale for Brownian motion and the role that it plays in the classical Cameron–Martin–Girsanov change of measure. Indeed, the theory presented here may be extended to the general class of spectrally negative Lévy processes, which includes Cramér–Lundberg processes and Brownian motion. See for example Chap. 3 of Kyprianou (2013). The Esscher transform plays a prominent role in mathematical finance as well as insurance mathematics; see for example the discussion in the paper of Gerber and Shiu (1994) and references therein. The style of reasoning in the proof of Theorem 2.5 is inspired by the classical computations of Wald (1944) for random walks, see also Bingham (1975) and Gerber (1990).

Gerber-Shiu Risk Theory

Kyprianou, A.E.

2013, VIII, 93 p. 7 illus., 3 illus. in color., Softcover

ISBN: 978-3-319-02302-1