Contents

1. Preliminaries

1.1. Stochastic processes

We start with some definitions.

Definition 1.1. *Let $\mathcal{I}$ be either $\mathbb{N}$ or $[0, \infty)$. A* **stochastic process** *on $\mathcal{I}$ with* **state space** *E is a family of random variables $\{X_t : t \in \mathcal{I}\}$ on E. Let E be a topological space. A stochastic process is called* **cadlag** *if it has a.s. right continuous paths and the limits from the left exist. It is called* **continuous** *if its paths are a.s. continuous. We will often identify a cadlag (continuous, respectively) stochastic process with a random variable on the space of cadlag (continuous) functions. For the rest of these notes we will always assume that all stochastic processes on $[0, \infty)$ are cadlag.*

Definition 1.2. *A cadlag stochastic process $\{N_t\}$ on $[0, \infty)$ is called* **point process** *if a.s.*

i) $N_0 = 0$,

ii) $N_t \geq N_s$ *for all* $t \geq s$ *and*

iii) $N_t \in \mathbb{N}$ *for all* $t \in (0, \infty)$.

We denote the jump times by $T_1, T_2, \ldots$, i.e. $T_k = \inf\{t \geq 0 : N_t \geq k\}$ for all $k \in \mathbb{N}$. In particular $T_0 = 0$. A point process is called **simple** *if $T_0 < T_1 < T_2 < \cdots$.*

Definition 1.3. *A stochastic process $\{X_t\}$ is said to have* **independent increments** *if for all $n \in \mathbb{N}$ and all real numbers $0 = t_0 < t_1 < t_2 < \cdots < t_n$ the random variables $X_{t_1} - X_{t_0}, X_{t_2} - X_{t_1}, \ldots, X_{t_n} - X_{t_{n-1}}$ are independent. A stochastic process is said to have* **stationary increments** *if for all $n \in \mathbb{N}$ and all real numbers $0 = t_0 < t_1 < t_2 < \cdots < t_n$ and all $h > 0$ the random vectors $(X_{t_1} - X_{t_0}, X_{t_2} - X_{t_1}, \ldots, X_{t_n} - X_{t_{n-1}})$ and $(X_{t_1+h} - X_{t_0+h}, X_{t_2+h} - X_{t_1+h}, \ldots, X_{t_n+h} - X_{t_{n-1}+h})$ have the same distribution.*

Definition 1.4. *An increasing family $\{\mathcal{F}_t\}$ of σ-algebras is called* **filtration** *if $\mathcal{F}_t \subset \mathcal{F}$ for all t. A filtration is called* **right continuous** *if $\mathcal{F}_{t+} := \bigcap_{s>t} \mathcal{F}_s = \mathcal{F}_t$ for all $t \geq 0$.A stochastic process $\{X_t\}$ is called* **$\{\mathcal{F}_t\}$-adapted** *if X_t is $\mathcal{F}_t$-measurable* *for all $t \geq 0$. The* **natural filtration** $\{\mathcal{F}_t^X\}$ *of a stochastic process* $\{X_t\}$ *is the smallest right continuous filtration such that the process is adapted. If a stochastic process is considered then we use its natural filtration if nothing else is mentioned.*

Definition 1.5. *An* $\{\mathcal{F}_t\}$**-stopping time** *is a random variable T on $[0, \infty]$ or $\mathbb{N}$ such that $\{T \leq t\} \in \mathcal{F}_t$ for all $t \geq 0$. If we stop a stochastic process then we usually assume that T is a stopping time with respect to the natural filtration of the stochastic process. The σ-algebra*

$$\mathcal{F}_T := \{A \in \mathcal{F} : A \cap \{T \leq t\} \in \mathcal{F}_t \text{ for all } t \geq 0\}$$

is called the information up to the stopping time T.

1.2. Martingales

Definition 1.6. *A stochastic process $\{M_t\}$ with state space $\mathbb{R}$ is called an* $\{\mathcal{F}_t\}$**-martingale** *(***-submartingale,-supermartingale***, respectively) if*

i) *$\{M_t\}$ is adapted,*

ii) *$\mathbb{E}[M_t]$ exists for all $t \geq 0$,*

iii) *$\mathbb{E}[M_t \mid \mathcal{F}_s] = (\geq, \leq)\, M_s$ a.s. for all $t \geq s \geq 0$.*

We simply say $\{M_t\}$ is a martingale (submartingale,supermartingale) if it is a martingale (submartingale,supermartingale) with respect to its natural filtration.

The following two propositions are very important for dealing with martingales.

Proposition 1.7. (Martingale stopping theorem) *Let $\{M_t\}$ be an $\{\mathcal{F}_t\}$-martingale (-submartingale, -supermartingale, respectively) and T be an $\{\mathcal{F}_t\}$-stopping time. Assume that $\{\mathcal{F}_t\}$ is right continuous. Then also the stochastic process $\{M_{T\wedge t} : t \geq 0\}$ is an $\{\mathcal{F}_t\}$-martingale (-submartingale, -supermartingale). Moreover, $\mathbb{E}[M_t \mid \mathcal{F}_T] = (\geq, \leq) M_{T\wedge t}$.*

Proof. See [19, Thm. 12], [25, Thm. II.2.13] or [45, Thm. 10.2.5]. □

Proposition 1.8. (Martingale convergence theorem) *Let $\{M_t\}$ be an $\{\mathcal{F}_t\}$-martingale such that $\lim_{t\to\infty} \mathbb{E}[M_t^-] < \infty$ (or equivalently $\sup_{t\geq 0} \mathbb{E}[|M_t|] < \infty$). If $\{\mathcal{F}_t\}$ is right continuous then the random variable $M_\infty := \lim_{t\to\infty} M_t$ exists a.s. and is integrable.*

Proof. See for instance [19, Thm. 6] or [45, Thm. 10.2.2]. □

Note that in general

$$\mathbb{E}[M_\infty] = \mathbb{E}[\lim_{t\to\infty} M_t] \neq \lim_{t\to\infty} \mathbb{E}[M_t] = \mathbb{E}[M_0]\,.$$

1.3. Poisson processes

Definition 1.9. *A point process $\{N_t\}$ is called (homogeneous)* **Poisson process** *with rate λ if*

i) *$\{N_t\}$ has stationary and independent increments,*

ii) $$\mathbb{P}[N_h = 0] = 1 - \lambda h + o(h) \qquad \text{as } h \to 0,$$

iii) $$\mathbb{P}[N_h = 1] = \lambda h + o(h) \qquad \text{as } h \to 0.$$

Remark. It follows readily that $\mathbb{P}[N_h \geq 2] = o(h)$ and that the point process is simple. ■

We give now some alternative definitions of the Poisson process.

Proposition 1.10. *Let $\{N_t\}$ be a point process. Then the following are equivalent:*

i) *$\{N_t\}$ is a Poisson process with rate λ.*

ii) *$\{N_t\}$ has independent increments and $N_t \sim$ Pois(λt) for each fixed $t \geq 0$.*

iii) *The interarrival times $\{T_k - T_{k-1} : k \geq 1\}$ are independent and Exp(λ) distributed.*

iv) *For each fixed $t \geq 0$, $N_t \sim$ Pois(λt) and given $\{N_t = n\}$ the occurrence points have the same distribution as the order statistics of n independent uniformly on $[0, t]$ distributed points.*

v) $\{N_t\}$ *has independent increments such that* $\mathbb{E}[N_1] = \lambda$ *and given* $\{N_t = n\}$ *the occurrence points have the same distribution as the order statistics of* n *independent uniformly on* $[0, t]$ *distributed points.*

vi) $\{N_t\}$ *has independent and stationary increments such that* $\mathbb{P}[N_h \ge 2] = o(h)$ *and* $\mathbb{E}[N_1] = \lambda$.

Proof. "*i*) $\Longrightarrow$ *ii*)" Let $p_n(t) = \mathbb{P}[N_t = n]$. Then show that $p_n(t)$ is continuous and differentiable. Finding the differential equations and solving them shows *ii*). The details are left as an exercise.
"*ii*) $\Longrightarrow$ *iii*)" We first show that $\{N_t\}$ has stationary increments. It is enough to show that $N_{t+h} - N_h$ is Pois(λt) distributed. For the moment generating functions we obtain

$$\mathrm{e}^{\lambda(t+h)(\mathrm{e}^r - 1)} = \mathbb{E}[\mathrm{e}^{r(N_{t+h} - N_h)}\mathrm{e}^{rN_h}] = \mathbb{E}[\mathrm{e}^{r(N_{t+h} - N_h)}]\mathrm{e}^{\lambda h(\mathrm{e}^r - 1)}\,.$$

It follows that

$$\mathbb{E}[\mathrm{e}^{r(N_{t+h} - N_h)}] = \mathrm{e}^{\lambda t(\mathrm{e}^r - 1)}\,.$$

This proves that $N_{t+h} - N_h$ is Pois(λt) distributed.

Let $t_0 = 0 \le s_1 < t_1 \le s_2 < t_2 \le \cdots \le s_n < t_n$. Then

$$\begin{aligned}
&\mathbb{P}[s_k < T_k \le t_k, 1 \le k \le n] \\
&\quad = \mathbb{P}[N_{s_k} - N_{t_{k-1}} = 0, N_{t_k} - N_{s_k} = 1,\ 1 \le k \le n-1, \\
&\qquad\qquad\qquad\qquad N_{s_n} - N_{t_{n-1}} = 0, N_{t_n} - N_{s_n} \ge 1] \\
&\quad = \mathrm{e}^{-\lambda(s_n - t_{n-1})}(1 - \mathrm{e}^{-\lambda(t_n - s_n)}) \prod_{k=1}^{n-1} \mathrm{e}^{-\lambda(s_k - t_{k-1})}\lambda(t_k - s_k)\mathrm{e}^{-\lambda(t_k - s_k)} \\
&\quad = (\mathrm{e}^{-\lambda s_n} - \mathrm{e}^{-\lambda t_n})\lambda^{n-1}\prod_{k=1}^{n-1}(t_k - s_k) \\
&\quad = \int_{s_1}^{t_1} \cdots \int_{s_n}^{t_n} \lambda^n \mathrm{e}^{-\lambda y_n}\,\mathrm{d}y_n \cdots \mathrm{d}y_1 \\
&\quad = \int_{s_1}^{t_1} \int_{s_2 - z_1}^{t_2 - z_1} \cdots \int_{s_n - z_1 - \cdots - z_{n-1}}^{t_n - z_1 - \cdots - z_{n-1}} \lambda^n \mathrm{e}^{-\lambda(z_1 + \cdots + z_n)}\,\mathrm{d}z_n \cdots \mathrm{d}z_1\,.
\end{aligned}$$

It follows that the joint density of $T_1, T_2 - T_1, \ldots, T_n - T_{n-1}$ is

$$\lambda^n \mathrm{e}^{-\lambda(z_1 + \cdots + z_n)}$$

and therefore they are independent and $\mathrm{Exp}(\lambda)$ distributed.
"*iii*) $\Longrightarrow$ *iv*)" Note that T_n is $\Gamma(n, \lambda)$ distributed. Thus $\mathbb{P}[N_t = 0] = \mathbb{P}[T_1 > t] = \mathrm{e}^{-\lambda t}$ and for $n \ge 1$

$$\begin{aligned} \mathbb{P}[N_t = n] &= \mathbb{P}[N_t \ge n] - \mathbb{P}[N_t \ge n+1] \\ &= \mathbb{P}[T_n \le t] - \mathbb{P}[T_{n+1} \le t] \\ &= \int_0^t \frac{\lambda^n s^{n-1}}{(n-1)!} \mathrm{e}^{-\lambda s} \,\mathrm{d}s - \int_0^t \frac{\lambda^{n+1} s^n}{n!} \mathrm{e}^{-\lambda s} \,\mathrm{d}s \\ &= \int_0^t \frac{\mathrm{d}}{\mathrm{d}s} \frac{(\lambda s)^n}{n!} \mathrm{e}^{-\lambda s} \,\mathrm{d}s = \frac{(\lambda t)^n}{n!} \mathrm{e}^{-\lambda t} \,. \end{aligned}$$

The joint density of $T_1, \dots, T_{n+1}$ is for $t_0 = 0$

$$\prod_{k=1}^{n+1} \lambda \mathrm{e}^{-\lambda(t_k - t_{k-1})} = \lambda^{n+1} \mathrm{e}^{-\lambda t_{n+1}} \,.$$

Thus the joint conditional density of $T_1, \dots, T_n$ given $N_t = n$ is

$$\frac{\int_t^\infty \lambda^{n+1} \mathrm{e}^{-\lambda t_{n+1}} \,\mathrm{d}t_{n+1}}{\int_0^t \int_{t_1}^t \cdots \int_{t_{n-1}}^t \int_t^\infty \lambda^{n+1} \mathrm{e}^{-\lambda t_{n+1}} \,\mathrm{d}t_{n+1} \cdots \mathrm{d}t_1} = \frac{n!}{t^n} \,.$$

"*iv*) $\Longrightarrow$ *v*)" It is clear that $\mathbb{E}[N_1] = \lambda$. Let $x_k \in \mathbb{N}$ and $t_0 = 0 < t_1 < \cdots < t_n$. Then for $x = x_1 + \cdots + x_n$

$$\begin{aligned} &\mathbb{P}[N_{t_k} - N_{t_{k-1}} = x_k, 1 \le k \le n] \\ &\quad = \mathbb{P}[N_{t_k} - N_{t_{k-1}} = x_k, 1 \le k \le n \mid N_{t_n} = x] \mathbb{P}[N_{t_n} = x] \\ &\quad = \frac{(\lambda t_n)^x}{x!} \mathrm{e}^{-\lambda t_n} \frac{x!}{x_1! \cdots x_n!} \prod_{k=1}^n \Bigl(\frac{t_k - t_{k-1}}{t_n}\Bigr)^{x_k} = \prod_{k=1}^n \frac{(\lambda(t_k - t_{k-1}))^{x_k}}{x_k!} \mathrm{e}^{-\lambda(t_k - t_{k-1})} \end{aligned}$$

and therefore $\{N_t\}$ has independent increments.
"*v*) $\Longrightarrow$ *vi*)" Note that for $x_k \in \mathbb{N}$, $t_0 = 0 < t_1 < \cdots < t_n$ and $h > 0$

$$\begin{aligned} &\mathbb{P}[N_{t_k+h} - N_{t_{k-1}+h} = x_k, 1 \le k \le n \mid N_{t_n+h}] \\ &\qquad = \mathbb{P}[N_{t_k} - N_{t_{k-1}} = x_k, 1 \le k \le n \mid N_{t_n+h}] \,. \end{aligned}$$

Integrating with respect to N_{t_n+h} yields that $\{N_t\}$ has stationary increments. Let now $0 < h < 1$ and recall that $(1-h)^k \ge 1 - kh$. Then

$$\begin{aligned}\frac{1}{h}(1 - \mathbb{P}[N_h = 0]) &= \frac{1}{h}\Big(1 - \sum_{k=0}^{\infty} \mathbb{P}[N_1 = k](1-h)^k\Big) \\ &= \sum_{k=1}^{\infty} \mathbb{P}[N_1 = k]\Big(\frac{1-(1-h)^k}{h}\Big) \\ &\le \sum_{k=1}^{\infty} k\mathbb{P}[N_1 = k] = \mathbb{E}[N_1] = \lambda\,.\end{aligned}$$

We can therefore interchange limit and summation to obtain

$$\lim_{h\to 0} \frac{1}{h}(1 - \mathbb{P}[N_h = 0]) = \lambda$$

from which

$$\mathbb{P}[N_h = 0] = 1 - \lambda h + o(h)$$

follows. Furthermore,

$$\lim_{h\to 0} \frac{1}{h}\mathbb{P}[N_h = 1] = \lim_{h\to 0} \sum_{k=1}^{\infty} \mathbb{P}[N_1 = k]k(1-h)^{k-1} = \lambda$$

and thus

$$\mathbb{P}[N_h = 1] = \lambda h + o(h)\,.$$

It follows that $\mathbb{P}[N_h \ge 2] = o(h)$.
"$vi) \implies i)$" We have

$$\mathbb{P}[N_{t+s} = 0] = \mathbb{P}[N_{t+s} - N_s = 0, N_s = 0] = \mathbb{P}[N_t = 0]\mathbb{P}[N_s = 0]\,.$$

It is left as an exercise to show that $\mathbb{P}[N_t = 0] = (\mathbb{P}[N_1 = 0])^t$. Let now $\lambda_0 := -\log \mathbb{P}[N_1 = 0]$. Then $\mathbb{P}[N_h = 0] = 1 - \lambda_0 h + o(h)$ and consequently $\mathbb{P}[N_h = 1] = \lambda_0 h + o(h)$, i.e. $i)$ holds with λ replaced by λ_0. From what we already proved $v)$ holds with λ replaced by λ_0. Therefore $\lambda_0 = \lambda$. $\square$

Remark. The condition $\mathbb{E}[N_1] = \lambda$ in $v)$ and $vi)$ is only used for identifying the parameter λ. From the theory of Lévy processes it follows that the jumps larger than $\frac{1}{2}$, say, form a compound Poisson process. Because the process considered has jumps of size 1 only, $\{N_t\}$ must be a Poisson process. Indeed, for $vi)$ this follows directly from the proof above. Assume $v)$ without requiring a finite mean. As in the proof above, $\{N_t\}$ has stationary increments. As in the proof $vi) \implies i)$ we have $\mathbb{P}[N_t = 0] = \mathrm{e}^{-\lambda_0 t}$ for some λ_0. In particular, $\mathbb{P}[N_h \ge 1] = o(1)$. Moreover, $\mathbb{P}[N_t = 1]$ is a continuous function of t because

$$\mathbb{P}[N_{t+h} = 1] = \mathrm{e}^{-\lambda_0 t}\mathbb{P}[N_h = 1] + \mathrm{e}^{-\lambda_0 h}\mathbb{P}[N_t = 1]\,,$$

(for left continuity replace t by $t - h$). We have

$$\mathrm{e}^{-\lambda_0 t} = \mathbb{P}[N_t = 0] = \mathrm{e}^{-\lambda_0(t+h)} + \frac{h}{t+h}\mathbb{P}[N_{t+h} = 1] + \sum_{k=2}^{\infty}\Bigl(\frac{h}{t+h}\Bigr)^k \mathbb{P}[N_t = k]\,.$$

Reordering the terms yields

$$\frac{\mathrm{e}^{-\lambda_0 t} - \mathrm{e}^{-\lambda_0(t+h)}}{h} = \frac{1}{t+h}\mathbb{P}[N_{t+h} = 1] + \frac{h}{(t+h)^2}\sum_{k=2}^{\infty}\Bigl(\frac{h}{t+h}\Bigr)^{k-2} \mathbb{P}[N_t = k]\,.$$

Because the sum is bounded by 1, letting $h \to 0$ gives $\mathbb{P}[N_t = 1] = \lambda_0 t \mathrm{e}^{-\lambda_0 t}$. This implies i). ■

The Poisson process has the following properties.

Proposition 1.11. *Let $\{N_t\}$ and $\{\tilde{N}_t\}$ be two independent Poisson processes with rates λ and $\tilde{\lambda}$ respectively. Let $\{I_i : i \in \mathbb{N}\}$ be an iid sequence of random variables independent of $\{N_t\}$ with $\mathbb{P}[I_i = 1] = 1 - \mathbb{P}[I_i = 0] = q$ for some $q \in (0, 1)$. Furthermore let $a > 0$ be a real number. Then*

i) *$\{N_t + \tilde{N}_t\}$ is a Poisson process with rate $\lambda + \tilde{\lambda}$.*

ii) *$\Bigl(\sum_{i=1}^{N_t} I_i\Bigr)$ is a Poisson process with rate λq.*

iii) *$\{N_{at}\}$ is a Poisson process with rate λa.*

Proof. Exercise. □

Definition 1.12. *Let $\Lambda(t)$ be an increasing right continuous function on $[0, \infty)$ with $\Lambda(0) = 0$. A point process $\{N_t\}$ on $[0, \infty)$ is called* **inhomogeneous Poisson process** *with* **intensity measure** *$\Lambda(t)$ if*

i) *$\{N_t\}$ has independent increments,*

ii) *$N_t - N_s \sim$ Pois$(\Lambda(t) - \Lambda(s))$.*

If there exists a function $\lambda(t)$ such that $\Lambda(t) = \int_0^t \lambda(s)\,\mathrm{d}s$ then $\lambda(t)$ is called **intensity** *or* **rate** *of the inhomogeneous Poisson process.*

Note that a homogeneous Poisson process is a special case with $\Lambda(t) = \lambda t$. Define $\Lambda^{-1}(x) = \sup\{t \ge 0 : \Lambda(t) \le x\}$ the inverse function of $\Lambda(t)$.

Proposition 1.13. *Let $\{\tilde{N}_t\}$ be a homogeneous Poisson process with rate 1. Define $N_t = \tilde{N}_{\Lambda(t)}$. Then $\{N_t\}$ is an inhomogeneous Poisson process with intensity measure $\Lambda(t)$. Conversely, let $\{N_t\}$ be an inhomogeneous Poisson process with intensity measure $\Lambda(t)$. Let $\tilde{N}_t = N_{\Lambda^{-1}(t)}$ at all points where $\Lambda(\Lambda^{-1}(t)) = t$. On intervals $(\Lambda(\Lambda^{-1}(t)-), \Lambda(\Lambda^{-1}(t)))$ where $\Lambda(\Lambda^{-1}(t)) \neq t$ let there be $N_{\Lambda^{-1}(t)} - N_{(\Lambda^{-1}(t)-)}$ occurrence points uniformly distributed on the interval $(\Lambda(\Lambda^{-1}(t)-), \Lambda(\Lambda^{-1}(t)))$ independent of (N_t). Then $\{\tilde{N}_t\}$ is a homogeneous Poisson process with rate 1.*

Proof. Exercise. □

For an inhomogeneous Poisson process we can construct the following martingales.

Lemma 1.14. *Let $r \in \mathbb{R}$. The following processes are martingales.*

i) $\{N_t - \Lambda(t)\}$,

ii) $\{(N_t - \Lambda(t))^2 - \Lambda(t)\}$,

iii) $\{\exp[rN_t - \Lambda(t)(\mathrm{e}^r - 1)]\}$.

Proof. i) Because $\{N_t\}$ has independent increments

$$\mathbb{E}[N_t - \Lambda(t) \mid \mathcal{F}_s] = \mathbb{E}[N_t - N_s] + N_s - \Lambda(t) = N_s - \Lambda(s)\,.$$

ii) Analogously,

$$\begin{aligned}
&\mathbb{E}[(N_t - \Lambda(t))^2 - \Lambda(t) \mid \mathcal{F}_s] \\
&\qquad = \mathbb{E}[(N_t - N_s - \{\Lambda(t) - \Lambda(s)\} + N_s - \Lambda(s))^2 \mid \mathcal{F}_s] - \Lambda(t) \\
&\qquad = \mathbb{E}[(N_t - N_s - \{\Lambda(t) - \Lambda(s)\})^2] \\
&\qquad\qquad + 2(N_s - \Lambda(s))\mathbb{E}[N_t - N_s - \{\Lambda(t) - \Lambda(s)\}] + (N_s - \Lambda(s))^2 - \Lambda(t) \\
&\qquad = \Lambda(t) - \Lambda(s) + (N_s - \Lambda(s))^2 - \Lambda(t) = (N_s - \Lambda(s))^2 - \Lambda(s)\,.
\end{aligned}$$

iii) Analogously

$$\mathbb{E}[\mathrm{e}^{rN_t - \Lambda(t)(\mathrm{e}^r - 1)} \mid \mathcal{F}_s] = \mathbb{E}[\mathrm{e}^{r(N_t - N_s)}]\mathrm{e}^{rN_s - \Lambda(t)(\mathrm{e}^r - 1)} = \mathrm{e}^{rN_s - \Lambda(s)(\mathrm{e}^r - 1)}\,.$$

□

1.4. Renewal processes

Definition 1.15. *A simple point process $\{N_t\}$ is called* **ordinary renewal process** *if the interarrival times $\{T_k - T_{k-1} : k \ge 1\}$ are iid. If T_1 has a different distribution then $\{N_t\}$ is called* **delayed renewal process**. *If $\lambda^{-1} = \mathbb{E}[T_2 - T_1]$ exists and*

$$\mathbb{P}[T_1 \le x] = \lambda \int_0^x \mathbb{P}[T_2 - T_1 > y] \, \mathrm{d}y \tag{1.1}$$

then $\{N_t\}$ is called **stationary renewal process**. *If $\{N_t\}$ is an ordinary renewal process then the function $U(t) = \mathbb{I}_{\{t \ge 0\}} + \mathbb{E}[N_t]$ is called the* **renewal measure**.

In the rest of this section we denote by F the distribution function of $T_2 - T_1$. For simplicity we let T be a random variable with distribution F. Note that because the point process is simple we implicitly assume that $F(0) = 0$. If nothing else is said we consider in the sequel only ordinary renewal processes.

Recall that F^{*0} is the indicator function of the interval $[0, \infty)$.

Lemma 1.16. *The renewal measure can be written as*

$$U(t) = \sum_{n=0}^{\infty} F^{*n}(t) \,.$$

Moreover, $U(t) < \infty$ for all $t \ge 0$ and $U(t) \to \infty$ as $t \to \infty$.

Proof. Exercise. □

In the renewal theory one often has to solve equations of the form

$$Z(x) = z(x) + \int_0^x Z(x-y) \, \mathrm{d}F(y) \qquad (x \ge 0) \tag{1.2}$$

where $z(x)$ is a known function and $Z(x)$ is unknown. This equation is called the **renewal equation**. The equation can be solved explicitly.

Proposition 1.17. *If $z(x)$ is bounded on bounded intervals then*

$$Z(x) = \int_{0-}^x z(x-y) \, \mathrm{d}U(y) = z * U(x)$$

is the unique solution to (1.2) that is bounded on bounded intervals.

Proof. We first show that $Z(x)$ is a solution. This follows from

$$\begin{aligned} Z(x) &= z * U(x) = z * \sum_{n=0}^{\infty} F^{*n}(x) = \sum_{n=0}^{\infty} z * F^{*n}(x) \\ &= z(x) + \sum_{n=1}^{\infty} z * F^{*(n-1)} * F(x) = z(x) + z * U * F(x) = z(x) + Z * F(x)\,. \end{aligned}$$

Let $Z_1(x)$ be a solution that is bounded on bounded intervals. Then

$$|Z_1(x)-Z(x)| = \Big|\int_0^x (Z_1(x-y)-Z(x-y))\,\mathrm{d}F(y)\Big| \le \int_0^x |Z_1(x-y)-Z(x-y)|\,\mathrm{d}F(y)$$

and by induction

$$|Z_1(x) - Z(x)| \le \int_0^x |Z_1(x-y) - Z(x-y)|\,\mathrm{d}F^{*n}(y) \le \sup_{0\le y\le x} |Z_1(y) - Z(y)| F^{*n}(x)\,.$$

The latter tends to 0 as $n \to \infty$. Thus $Z_1(x) = Z(x)$. □

Let $z(x)$ be a bounded function and $h > 0$ be a real number. Define

$$\overline{m}_k(h) = \sup\{z(t) : (k-1)h \le t < kh\}, \qquad \underline{m}_k(h) = \inf\{z(t) : (k-1)h \le t < kh\}$$

and the Riemann sums

$$\bar{\sigma}(h) = h\sum_{k=1}^{\infty} \overline{m}_k(h), \qquad \underline{\sigma}(h) = h\sum_{k=1}^{\infty} \underline{m}_k(h).$$

Definition 1.18. *A function $z(x)$ is called* **directly Riemann integrable** *if*

$$-\infty < \underline{\lim}_{h\downarrow 0}\, \underline{\sigma}(h) = \overline{\lim}_{h\downarrow 0}\, \bar{\sigma}(h) < \infty\,.$$

The following lemma gives some criteria for a function to be directly Riemann integrable.

Lemma 1.19.

- *The space of directly Riemann integrable functions is a linear space.*
- *If $z(t)$ is monotone and $\int_0^\infty z(t)\,\mathrm{d}t < \infty$ then $z(t)$ is directly Riemann integrable.*
- *If $a(t)$ and $b(t)$ are directly Riemann integrable and $z(t)$ is continuous Lebesgue a.e. such that $a(t) \le z(t) \le b(t)$ then $z(t)$ is directly Riemann integrable.*
- *If $z(t) \ge 0$, $z(t)$ is continuous Lebesgue a.e. and $\bar{\sigma}(h) < \infty$ for some $h > 0$ then is $z(t)$ directly Riemann integrable.*

Proof. See for instance [1, p.69]. □

Definition 1.20. *A distribution function F of a random variable X is called* **arithmetic** *if for some γ one has $\mathbb{P}[X \in \{\gamma, 2\gamma, \dots\}] = 1$. The* **span** *$\gamma$ is the largest number such that the above relation is fulfilled.*

The most important result in renewal theory is a result on the asymptotic behaviour of the solution to (1.2).

Proposition 1.21. (Renewal theorem) *If $z(x)$ is a directly Riemann integrable function then the solution $Z(x)$ to the renewal equation satisfies*

$$\lim_{t\to\infty} Z(t) = \lambda \int_0^\infty z(y)\,\mathrm{d}y$$

if F is non-arithmetic and

$$\lim_{n\to\infty} Z(t+n\gamma) = \gamma\lambda \sum_{j=0}^\infty z(t+j\gamma)$$

for $0 \le t < \gamma$ if F is arithmetic with span γ.

Proof. See for instance [26, p.364] or [45, p.218]. □

Example 1.22. Assume that F is non-arithmetic and that $\mathbb{E}[T_1^2] < \infty$. We consider the function $Z(t) = \mathbb{E}[T_{N_t+1} - t]$, the expected time till the next occurrence of an event. Consider first $\mathbb{E}[T_{N_t+1} - t \mid T_1 = s]$. If $s \le t$ then there is a renewal process starting at s and thus $\mathbb{E}[T_{N_t+1} - t \mid T_1 = s] = Z(t-s)$. If $s > t$ then $T_{N_t+1} = T_1 = s$ and thus $\mathbb{E}[T_{N_t+1} - t \mid T_1 = s] = s - t$. Hence we get the renewal equation

$$Z(t) = \mathbb{E}[\mathbb{E}[T_{N_t+1} - t \mid T_1]] = \int_t^\infty (s-t)\,\mathrm{d}F(s) + \int_0^t Z(t-s)\,\mathrm{d}F(s)\,.$$

Now

$$z(t) = \int_t^\infty (s-t)\,\mathrm{d}F(s) = \int_t^\infty \int_t^s \mathrm{d}y\,\mathrm{d}F(s) = \int_t^\infty (1-F(y))\,\mathrm{d}y\,.$$

The function is monotone and

$$\int_0^\infty \int_t^\infty (s-t)\,\mathrm{d}F(s)\,\mathrm{d}t = \int_0^\infty \int_0^s (s-t)\,\mathrm{d}t\,\mathrm{d}F(s) = \frac{1}{2}\mathbb{E}[T_1^2]\,.$$

Thus $z(t)$ is directly Riemann integrable and by the renewal theorem

$$\lim_{t\to\infty} Z(t) = \frac{\lambda}{2} E\left[T_1^2\right]\,.$$

■

1.5. Brownian motion

Definition 1.23. *A (cadlag) stochastic process $\{W_t\}$ is called (m, η^2)-***Brownian motion** *if a.s.*

- $W_0 = 0$,
- $\{W_t\}$ *has independent increments and*
- $W_t \sim N(mt, \eta^2 t)$.

A $(0, 1)$-Brownian motion is called **standard Brownian motion**.

It can be shown that Brownian motion exists. Moreover, one can proof that a Brownian motion has continuous paths.

From the definition it follows that $\{W_t\}$ has also stationary increments.

Lemma 1.24. *A (m, η^2)-Brownian motion has stationary increments.*

Proof. Exercise. □

We construct the following martingales.

Lemma 1.25. *Let $\{W_t\}$ be a (m, η^2)-Brownian motion and $r \in \mathbb{R}$. The following processes are martingales.*

i) $\{W_t - mt\}$.

ii) $\{(W_t - mt)^2 - \eta^2 t\}$.

iii) $\{\exp[r(W_t - mt) - \frac{\eta^2}{2} r^2 t]\}$.

Proof. i) Using the stationary and independent increments property

$$\mathbb{E}[W_t - mt \mid \mathcal{F}_s] = \mathbb{E}[W_t - W_s] + W_s - mt = W_s - ms\,.$$

ii) It suffices to consider the case $m = 0$.

$$\mathbb{E}[W_t^2 - \eta^2 t \mid \mathcal{F}_s] = \mathbb{E}[(W_t - W_s)^2 + 2(W_t - W_s)W_s + W_s^2 \mid \mathcal{F}_s] - \eta^2 t = W_s^2 - \eta^2 s\,.$$

iii) It suffices to consider the case $m = 0$.

$$\mathbb{E}[\mathrm{e}^{rW_t - \frac{\eta^2}{2} r^2 t} \mid \mathcal{F}_s] = \mathbb{E}[\mathrm{e}^{r(W_t - W_s)}]\mathrm{e}^{rW_s - \frac{\eta^2}{2} r^2 t} = \mathrm{e}^{rW_s - \frac{\eta^2}{2} r^2 s}\,.$$

□

1.6. Random walks and the Wiener-Hopf factorisation

Let $\{X_i\}$ be a sequence of iid. random variables. Let $U(x)$ be the distribution function of X_i. Define

$$S_n = \sum_{i=1}^{n} X_i$$

and $S_0 = 0$.

Lemma 1.26.

i) *If* $\mathbb{E}[X_1] < 0$ *then* S_n *converges to* $-\infty$ *a.s..*

ii) *If* $\mathbb{E}[X_1] = 0$ *then a.s.*

$$\overline{\lim_{n\to\infty}} S_n = -\underline{\lim_{n\to\infty}} S_n = \infty \,.$$

iii) *If* $\mathbb{E}[X_1] > 0$ *then* S_n *converges to* ∞ *a.s. and there is a strictly positive probability that* $S_n \geq 0$ *for all* $n \in \mathbb{N}$.

Proof. See for instance [26, p.396] or [45, p.233]. □

For the rest of this section we assume $-\infty < \mathbb{E}[X_i] < 0$. Let $\tau_+ = \inf\{n > 0 : S_n > 0\}$ and $\tau_- = \inf\{n > 0 : S_n \leq 0\}$, $H(x) = \mathbb{P}[\tau_+ < \infty, S_{\tau_+} \leq x]$ and $\rho(x) = \mathbb{P}[S_{\tau_-} \leq x]$. Note that τ_- is defined a.s. because $S_n \to -\infty$ as $n \to \infty$, see Lemma 1.26. Define $\psi_0(x) = \mathbb{1}_{\{x\geq 0\}}$ and

$$\psi_n(x) = \mathbb{P}[S_1 > 0, S_2 > 0, \ldots, S_n > 0, S_n \leq x]$$

for $n \geq 1$. Let

$$\psi(x) = \sum_{n=0}^{\infty} \psi_n(x) \,.$$

Lemma 1.27. *We have for* $n \geq 1$

$$\psi_n(x) = \mathbb{P}[S_n > S_j, \, (0 \leq j \leq n-1), S_n \leq x]$$

and therefore

$$\psi(x) = \sum_{n=0}^{\infty} H^{*n}(x) \,.$$

Moreover, $\psi(\infty) = (1 - H(\infty))^{-1}$, $\mathbb{E}[\tau_-] = \psi(\infty)$ *and* $\mathbb{E}[S_{\tau_-}] = \mathbb{E}[\tau_-]\mathbb{E}[X_i]$.

Proof. Let for n fixed

$$S_k^* = S_n - S_{n-k} = \sum_{i=n-k+1}^{n} X_i \,.$$

Then $\{S_k^*, k \le n\}$ follows the same law as $\{S_k, k \le n\}$. Thus

$$\begin{aligned}\mathbb{P}[S_n > S_j,\, (0 \le j \le n-1), S_n \le x] &= \mathbb{P}[S_n^* > S_j^*,\, (0 \le j \le n-1), S_n^* \le x] \\ &= \mathbb{P}[S_n > S_n - S_{n-j},\, (0 \le j \le n-1), S_n \le x] \\ &= \mathbb{P}[S_j > 0,\, (1 \le j \le n), S_n \le x] = \psi_n(x)\,.\end{aligned}$$

Denote by τ_n the n-th ascending ladder time. Note that S_{τ_n} has distribution $H^{*n}(x)$. We found that $\psi_n(x)$ is the probability that S_n is a maximum of the random walk and lies in the interval $(0, x]$. Thus $\psi_n(x)$ is the probability that there is a ladder height at n and $S_n \le x$. We obtain

$$\begin{aligned}\sum_{n=1}^{\infty} \psi_n(x) &= \sum_{n=1}^{\infty} \mathbb{P}[\exists k : \tau_k = n, 0 < S_n \le x] = \sum_{n=1}^{\infty}\sum_{k=1}^{\infty} \mathbb{P}[\tau_k = n, 0 < S_{\tau_k} \le x] \\ &= \sum_{k=1}^{\infty} \mathbb{P}[0 < S_{\tau_k} \le x, \tau_k < \infty] = \sum_{k=1}^{\infty} H^{*k}(x)\,.\end{aligned}$$

Because $\mathbb{E}[X_i] < 0$ we have $H(\infty) < 1$ (Lemma 1.26) and $H^{*n}(\infty) = (H(\infty))^n$ from which $\psi(\infty) = (1 - H(\infty))^{-1}$ follows.

For the expected value of τ_-

$$\mathbb{E}[\tau_-] = \sum_{n=0}^{\infty} \mathbb{P}[\tau_- > n] = \sum_{n=0}^{\infty} \psi_n(\infty) = \psi(\infty)\,.$$

Finally let $\tau_-(n)$ denote the n-th descending ladder time and let L_n be the n-th descending ladder height $S_{\tau_-(n-1)} - S_{\tau_-(n)}$. Then

$$\frac{L_1 + \cdots + L_n}{n} = \frac{-S_{\tau_-(n)}}{\tau_-(n)} \frac{\tau_-(n)}{n}$$

from which the last assertion follows by the strong law of large numbers. □

Let now

$$\rho_n(x) = \mathbb{P}[S_1 > 0, S_2 > 0, \ldots, S_{n-1} > 0, S_n \le x]\,.$$

Note that for $x \le 0$

$$\rho(x) = \sum_{n=1}^{\infty} \rho_n(x)\,,$$

that $\psi_1(x) = U(x) - U(0)$ if $x > 0$ and $\rho_1(x) = U(x)$ if $x \le 0$. For $n \ge 1$ one obtains for $x > 0$

$$\psi_{n+1}(x) = \int_0^\infty (U(x-y) - U(-y))\psi_n(\mathrm{d}y)$$

and for $x \le 0$

$$\rho_{n+1}(x) = \int_0^\infty U(x-y)\psi_n(\mathrm{d}y)\,.$$

If we sum over all n we obtain for $x > 0$

$$\psi(x) - 1 = \int_{0-}^\infty (U(x-y) - U(-y))\psi(\mathrm{d}y)$$

and for $x \le 0$

$$\rho(x) = \int_{0-}^\infty U(x-y)\psi(\mathrm{d}y)\,.$$

Note that $\rho(0) = \psi(0) = 1$ and thus for $x \ge 0$

$$\psi(x) = \int_{0-}^\infty U(x-y)\psi(\mathrm{d}y)\,.$$

The last two equations can be written as a single equation

$$\psi + \rho = \psi_0 + \psi * U\,.$$

Then

$$\begin{aligned}\psi - \psi_0 + \rho * H &= \psi * H + \rho * H = H + \psi * H * U = H + \psi * U - U \\ &= H - U + \psi + \rho - \psi_0\end{aligned}$$

from which it follows that

$$U = H + \rho - H * \rho\,. \tag{1.3}$$

The latter is called the **Wiener-Hopf factorisation**.

1.7. Subexponential distributions

Definition 1.28. *A distribution function F with $F(x) = 0$ for $x < 0$ is called* **subexponential** *if*

$$\lim_{t\to\infty} \frac{1 - F^{*2}(t)}{1 - F(t)} = 2\,.$$

We want to show that the moment generating function of such a distribution does not exist for strictly positive values. We first need the following

Lemma 1.29. *If F is subexponential then for all $t \in \mathbb{R}$*

$$\lim_{x\to\infty} \frac{1-F(x-t)}{1-F(x)} = 1\,.$$

Proof. Let $t \geq 0$. We have

$$\begin{aligned}
\frac{1-F^{*2}(x)}{1-F(x)} - 1 &= \frac{F(x)-F^{*2}(x)}{1-F(x)} \\
&= \int_{0-}^{t} \frac{1-F(x-y)}{1-F(x)}\,\mathrm{d}F(y) + \int_{t}^{x} \frac{1-F(x-y)}{1-F(x)}\,\mathrm{d}F(y) \\
&\geq F(t) + \frac{1-F(x-t)}{1-F(x)}(F(x)-F(t))\,.
\end{aligned}$$

Thus

$$1 \leq \frac{1-F(x-t)}{1-F(x)} \leq (F(x)-F(t))^{-1}\Big(\frac{1-F^{*2}(x)}{1-F(x)} - 1 - F(t)\Big)\,.$$

The assertion for $t \geq 0$ follows by letting $x \to \infty$. If $t < 0$ then

$$\lim_{x\to\infty} \frac{1-F(x-t)}{1-F(x)} = \lim_{x\to\infty} \frac{1}{\dfrac{1-F((x-t)-(-t))}{1-F(x-t)}} = \lim_{y\to\infty} \frac{1}{\dfrac{1-F(y-(-t))}{1-F(y)}} = 1$$

where $y = x - t$. □

Lemma 1.30. *Let F be subexponential and $r > 0$. Then*

$$\lim_{t\to\infty} \mathrm{e}^{rt}(1-F(t)) = \infty\,,$$

in particular

$$\int_{0-}^{\infty} \mathrm{e}^{rx}\,\mathrm{d}F(x) = \infty\,.$$

Proof. We first observe that for any $0 < x < t$

$$\mathrm{e}^{rt}(1-F(t)) = \frac{1-F(t)}{1-F(t-x)}(1-F(t-x))\mathrm{e}^{r(t-x)}\mathrm{e}^{rx} \tag{1.4}$$

and by Lemma 1.29 the function $\mathrm{e}^{rn}(1-F(n))$ ($n \in \mathbb{N}$) is increasing for n large enough. Thus there exist a limit in $(0,\infty]$. Letting $t \to \infty$ in (1.4) shows that this limit can only be infinite.

Let now $\{t_n\}$ be an arbitrary sequence such that $t_n \to \infty$. Then

$$\begin{aligned}
\mathrm{e}^{rt_n}(1-F(t_n)) &= \frac{1-F(t_n)}{1-F(\lfloor t_n \rfloor)}(1-F(\lfloor t_n \rfloor))\mathrm{e}^{r\lfloor t_n \rfloor}\mathrm{e}^{r(t_n - \lfloor t_n \rfloor)} \\
&\geq \frac{1-F(t_n)}{1-F(t_n-1)}(1-F(\lfloor t_n \rfloor))\mathrm{e}^{r\lfloor t_n \rfloor}
\end{aligned}$$

which tends to infinity as $n \to \infty$.

The moment generating function can be written as

$$\begin{aligned}
\int_{0-}^{\infty} \mathrm{e}^{rx} \, \mathrm{d}F(x) &= 1 + \int_0^\infty \int_0^x r\mathrm{e}^{ry} \, \mathrm{d}y \, \mathrm{d}F(x) \\
&= 1 + r\int_0^\infty \int_y^\infty \mathrm{d}F(x)\mathrm{e}^{ry} \, \mathrm{d}y = 1 + r\int_0^\infty \mathrm{e}^{ry}(1-F(y)) \, \mathrm{d}y = \infty
\end{aligned}$$

because the integrand diverges to infinity. □

The following lemma gives a condition for subexponentiality.

Lemma 1.31. *Assume that for all $z \in (0,1]$ the limit*

$$\gamma(z) = \lim_{x\to\infty} \frac{1-F(zx)}{1-F(x)}$$

exists and that $\gamma(z)$ is left-continuous at 1. Then F is a subexponential distribution function.

Proof. Note first that $F^{*2}(x) = \mathbb{P}[X_1 + X_2 \leq x] \leq \mathbb{P}[X_1 \leq x, X_2 \leq x] = F^2(x)$. For simplicity assume $F(0) = 0$. Hence

$$\varliminf_{x\to\infty} \frac{1-F^{*2}(x)}{1-F(x)} \geq \varliminf_{x\to\infty} \frac{1-F^2(x)}{1-F(x)} = \varliminf_{x\to\infty} 1 + F(x) = 2\,.$$

Let $n \geq 1$ be fixed. Then

$$\begin{aligned}
\varlimsup_{x\to\infty} \frac{1-F^{*2}(x)}{1-F(x)} &= 1 + \varlimsup_{x\to\infty} \int_0^x \frac{1-F(x-y)}{1-F(x)} \, \mathrm{d}F(y) \\
&\leq 1 + \varlimsup_{x\to\infty} \sum_{k=1}^n \frac{1-F(x-kx/n)}{1-F(x)}(F(kx/n) - F((k-1)x/n)) \\
&= 1 + \gamma(1-1/n)\,.
\end{aligned}$$

Because n is arbitrary and $\gamma(z)$ is left continuous at 1 the assertion follows. □

Example 1.32. Consider the $\mathrm{Pa}(\alpha,\beta)$ distribution.

$$\frac{1-F(zx)}{1-F(x)} = \frac{(\frac{\beta}{\beta+zx})^\alpha}{(\frac{\beta}{\beta+x})^\alpha} = \Big(\frac{\beta+x}{\beta+zx}\Big)^\alpha \to z^{-\alpha}$$

as $x\to\infty$. It follows that the Pareto distribution is subexponential. ■

Next we give an upper bound for the tails of the convolutions.

Lemma 1.33. *Let F be subexponential. Then for any $\varepsilon>0$ there exist a $D\in\mathbb{R}$ such that*

$$\frac{1-F^{*n}(t)}{1-F(t)} \le D(1+\varepsilon)^n$$

for all $t>0$ and $n\in\mathbb{N}$.

Proof. Let

$$\alpha_n = \sup_{t\ge 0}\frac{1-F^{*n}(t)}{1-F(t)}$$

and note that $1-F^{*(n+1)}(t) = 1-F(t)+F*(1-F^{*n})(t)$. Choose $T\ge 0$ such that

$$\sup_{t\ge T}\frac{F(t)-F^{*2}(t)}{1-F(t)} < 1+\frac{\varepsilon}{2}\,.$$

Let $A_T = (1-F(T))^{-1}$. For $n\ge 1$

$$\begin{aligned}
\alpha_{n+1} &\le 1+\sup_{0\le t\le T}\int_{0-}^t \frac{1-F^{*n}(t-y)}{1-F(t)}\,\mathrm{d}F(y) + \sup_{t\ge T}\int_{0-}^t \frac{1-F^{*n}(t-y)}{1-F(t)}\,\mathrm{d}F(y)\\
&\le 1+A_T+\sup_{t\ge T}\int_{0-}^t \frac{1-F^{*n}(t-y)}{1-F(t-y)}\,\frac{1-F(t-y)}{1-F(t)}\,\mathrm{d}F(y)\\
&\le 1+A_T+\alpha_n\sup_{t\ge T}\frac{F(t)-F^{*2}(t)}{1-F(t)} \le 1+A_T+\alpha_n\Big(1+\frac{\varepsilon}{2}\Big)\,.
\end{aligned}$$

Choose

$$D=\max\Big\{\frac{2(1+A_T)}{\varepsilon},1\Big\}$$

and note that $\alpha_1 = 1 < D(1+\varepsilon)$. The assertion follows by induction. □

Lemma 1.34. *Let $F(x)=0$ for $x<0$. The following are equivalent:*

i) *F is subexponential.*

ii) *For all $n\in\mathbb{N}$*

$$\lim_{x\to\infty}\frac{1-F^{*n}(x)}{1-F(x)} = n\,. \tag{1.5}$$

iii) *There exists $n \geq 2$ such that (1.5) holds.*

Proof. "*i)* $\Longrightarrow$ *ii)*" We proof the assertion by induction. The assertion is trivial for $n = 2$. Assume that it is proved for some $n \in \mathbb{N}$ with $n \geq 2$. Let $\varepsilon > 0$. Choose t such that

$$\Big|\frac{1 - F^{*n}(x)}{1 - F(x)} - n\Big| < \varepsilon$$

for $x \geq t$.

$$\begin{aligned}\frac{1 - F^{*(n+1)}(x)}{1 - F(x)} &= 1 + \int_{0-}^{x-t} \frac{1 - F^{*n}(x-y)}{1 - F(x-y)} \frac{1 - F(x-y)}{1 - F(x)} \,\mathrm{d}F(y) \\ &\qquad + \int_{x-t}^{x} \frac{1 - F^{*n}(x-y)}{1 - F(x)} \,\mathrm{d}F(y)\,.\end{aligned}$$

The second integral is bounded by

$$\int_{x-t}^{x} \frac{1 - F^{*n}(x-y)}{1 - F(x)} \,\mathrm{d}F(y) \leq \frac{F(x) - F(x-t)}{1 - F(x)} = \frac{1 - F(x-t)}{1 - F(x)} - 1$$

which tends to 0 as $x \to \infty$ by Lemma 1.29. The expression

$$\begin{aligned}&\int_{0-}^{x-t} n \frac{1 - F(x-y)}{1 - F(x)} \,\mathrm{d}F(y) \\ &\qquad = n\Big(\frac{F(x) - F^{*2}(x)}{1 - F(x)} - \int_{x-t}^{x} \frac{1 - F(x-y)}{1 - F(x)} \,\mathrm{d}F(y)\Big) \to n\end{aligned}$$

as $x \to \infty$ by the same arguments as before. Then

$$\begin{aligned}&\Big|\int_{0-}^{x-t} \Big(\frac{1 - F^{*n}(x-y)}{1 - F(x-y)} - n\Big) \frac{1 - F(x-y)}{1 - F(x)} \,\mathrm{d}F(y)\Big| \\ &\qquad \leq \varepsilon\Big(\frac{F(x) - F^{*2}(x)}{1 - F(x)} - \int_{x-t}^{x} \frac{1 - F(x-y)}{1 - F(x)} \,\mathrm{d}F(y)\Big) \to \varepsilon\end{aligned}$$

by the arguments used before. Thus

$$\varlimsup_{x\to\infty} \Big|\frac{1 - F^{*(n+1)}(x)}{1 - F(x)} - (n+1)\Big| \leq \varepsilon\,.$$

This proves the assertion because ε was arbitrary.
"*ii)* $\Longrightarrow$ *iii)*" Trivial.
"*iii)* $\Longrightarrow$ *i)*" Assume now (1.5) for some $n > 2$. We show that the assertion also holds for n replaced by $n - 1$. Since

$$\frac{1 - F^{*n}(x)}{1 - F(x)} - 1 = \int_{0-}^{x} \frac{1 - F^{*(n-1)}(x-y)}{1 - F(x)} \,\mathrm{d}F(y) \geq F(x) \frac{1 - F^{*(n-1)}(x)}{1 - F(x)}$$

it follows that

$$\overline{\lim_{x\to\infty}} \frac{1-F^{*(n-1)}(x)}{1-F(x)} \le n-1\,.$$

Because $F^{*(n-1)}(x) \le F(x)^{n-1}$ it is trivial that

$$\underline{\lim}_{x\to\infty} \frac{1-F^{*(n-1)}(x)}{1-F(x)} \ge n-1\,.$$

It follows that (1.5) holds for $n=2$ and F is a subexponential distribution. □

Lemma 1.35. *Let U and V be two distribution functions with $U(x) = V(x) = 0$ for all $x < 0$. Assume that*

$$1-V(x) \sim a(1-U(x)) \qquad \text{as } x\to\infty$$

for some $a > 0$. If U is subexponential then also V is subexponential.

Proof. We have to show that

$$\overline{\lim_{x\to\infty}} \int_{0-}^{x} \frac{1-V(x-y)}{1-V(x)}\,\mathrm{d}V(y) \le 1\,.$$

Let $0 < \varepsilon < 1$. There exists y_0 such that for all $y \ge y_0$

$$(1-\varepsilon)a \le \frac{1-V(y)}{1-U(y)} \le (1+\varepsilon)a\,.$$

Let $x > y_0$. Note that

$$\int_{x-y_0}^{x} \frac{1-V(x-y)}{1-V(x)}\,\mathrm{d}V(y) \le \frac{V(x)-V(x-y_0)}{1-V(x)} = \frac{\frac{1-V(x-y_0)}{1-U(x-y_0)}}{\frac{1-V(x)}{1-U(x)}}\,\frac{1-U(x-y_0)}{1-U(x)} - 1 \to 0$$

by Lemma 1.29. Moreover,

$$\begin{aligned}\int_{0-}^{x-y_0} \frac{1-V(x-y)}{1-V(x)}\,\mathrm{d}V(y) &\le \frac{1+\varepsilon}{1-\varepsilon}\int_{0-}^{x-y_0} \frac{1-U(x-y)}{1-U(x)}\,\mathrm{d}V(y)\\ &\le \frac{1+\varepsilon}{1-\varepsilon}\int_{0-}^{x} \frac{1-U(x-y)}{1-U(x)}\,\mathrm{d}V(y)\,.\end{aligned}$$

The last integral is

$$\begin{aligned}\int_{0-}^{x} \frac{1-U(x-y)}{1-U(x)}\,\mathrm{d}V(y) &= \frac{V(x)-U*V(x)}{1-U(x)}\\ &= 1-\frac{1-V(x)}{1-U(x)} + \frac{U(x)-V*U(x)}{1-U(x)}\\ &= 1-\frac{1-V(x)}{1-U(x)} + \int_{0-}^{x} \frac{1-V(x-y)}{1-U(x)}\,\mathrm{d}U(y)\,.\end{aligned}$$

$1-(1-U(x))^{-1}(1-V(x))$ tends to $1-a$ as $x \to \infty$. And

$$\int_{x-y_0}^{x} \frac{1-V(x-y)}{1-U(x)} \,\mathrm{d}U(y) \le \frac{U(x)-U(x-y_0)}{1-U(x)}$$

tends to 0. Finally

$$\begin{aligned}\int_{0-}^{x-y_0} \frac{1-V(x-y)}{1-U(x)} \,\mathrm{d}U(y) &\le (1+\varepsilon)a \int_{0-}^{x-y_0} \frac{1-U(x-y)}{1-U(x)} \,\mathrm{d}U(y) \\ &\le (1+\varepsilon)a \int_{0-}^{x} \frac{1-U(x-y)}{1-U(x)} \,\mathrm{d}U(y)\end{aligned}$$

which tends to $(1+\varepsilon)a$. Putting all these limits together we obtain

$$\varlimsup_{x\to\infty} \int_{0-}^{x} \frac{1-V(x-y)}{1-V(x)} \,\mathrm{d}V(y) \le \frac{(1+\varepsilon)(1+a\varepsilon)}{1-\varepsilon}\,.$$

Because ε was arbitrary the assertion follows. □

1.8. Concave and convex functions

In this chapter we let I be an interval, finite or infinite, but not a singleton.

Definition 1.36. *A function $u : I \to \mathbb{R}$ is called* **(strictly) concave** *if for all $x, z \in I$, $x \neq z$, and all $\alpha \in (0,1)$ one has*

$$u((1-\alpha)x + \alpha z) \ge (>)\, (1-\alpha)u(x) + \alpha u(z)\,.$$

u is called **(strictly) convex** *if $-u$ is (strictly) concave.*

Because results on concave functions can easily translated for convex functions we will only consider concave functions in the sequel.

Concave functions have nice properties.

Lemma 1.37. *A concave function $u(y)$ is continuous, differentiable from the left and from the right. The derivative is decreasing, i.e. for $x < y$ we have $u'(x-) \ge u'(x+) \ge u'(y-) \ge u'(y+)$. If $u(y)$ is strictly concave then $u'(x+) > u'(y-)$.*

Remark. The theorem implies that $u(y)$ is differentiable almost everywhere. ∎

Proof. Let $x < y < z$. Then

$$u(y) = u\Big(\frac{z-y}{z-x}x + \frac{y-x}{z-x}z\Big) \geq \frac{z-y}{z-x}u(x) + \frac{y-x}{z-x}u(z)$$

or equivalently

$$(z-x)u(y) \geq (z-y)u(x) + (y-x)u(z)\,. \tag{1.6}$$

This implies immediately

$$\frac{u(y)-u(x)}{y-x} \geq \frac{u(z)-u(x)}{z-x} \geq \frac{u(z)-u(y)}{z-y}\,. \tag{1.7}$$

Thus the function $h \to h^{-1}(u(y) - u(y-h))$ is increasing in h and bounded from below by $(z-y)^{-1}(u(z)-u(y))$. Thus the derivative $u'(y-)$ from the left exists. Analogously, the derivative from the right $u'(y+)$ exists. The assertion in the concave case follows now from (1.7). The strict inequality in the strictly concave case follows analogously. □

Concave functions have also the following property.

Lemma 1.38. *Let $u(y)$ be a concave function. There exists a function $k : I \to \mathbb{R}$ such that for any $y, x \in I$*

$$u(x) \leq u(y) + k(y)(x-y)\,. \tag{1.8}$$

Moreover, the function $k(y)$ is decreasing. If $u(y)$ is strictly concave then the above inequality is strict for $x \neq y$ and $k(y)$ is strictly decreasing. Conversely, if a function $k(y)$ exists such that (1.8) is fulfilled, then $u(y)$ is concave, strictly concave if the strict inequality holds for $x \neq y$.

Proof. Left as an exercise. □

Corollary 1.39. *Let $u(y)$ be a twice differentiable function. Then $u(y)$ is concave if and only if its second derivative is negative. It is strictly concave if and only if its second derivative is strictly negative almost everywhere.*

Proof. This follows readily from Theorem 1.37 and Lemma 1.38. □

The following result is very useful.

Theorem 1.40. (Jensen's inequality) *The function $u(y)$ is (strictly) concave if and only if*

$$\mathbb{E}[u(Y)] \le (<)\, u(\mathbb{E}[Y]) \tag{1.9}$$

for all I-valued integrable random variables Y with $\mathbb{P}[Y \ne \mathbb{E}[Y]] > 0$.

Proof. Assume (1.9) for all random variables Y. Let $\alpha \in (0,1)$. Let $\mathbb{P}[Y = z] = 1 - \mathbb{P}[Y = x] = \alpha$. Then the (strict) concavity follows. Assume $u(y)$ is strictly concave. Then it follows from Lemma 1.38 that

$$u(Y) \le u(\mathbb{E}[Y]) + k(\mathbb{E}[Y])(Y - \mathbb{E}[Y])\,.$$

The strict inequality holds if $u(y)$ is strictly concave and $Y \ne \mathbb{E}[Y]$. Taking expected values gives (1.9). □

Also the following result is often useful.

Theorem 1.41. (Ohlin's lemma) *Let $F_i(y)$, $i = 1, 2$ be two distribution functions defined on I. Assume*

$$\int_I y \,\mathrm{d}F_1(y) = \int_I y \,\mathrm{d}F_2(y) < \infty$$

and that there exists $y_0 \in I$ such that

$$F_1(y) \le F_2(y), \quad y < y_0, \qquad\qquad F_1(y) \ge F_2(y), \quad y > y_0\,.$$

Then for any concave function $u(y)$

$$\int_I u(y) \,\mathrm{d}F_1(y) \ge \int_I u(y) \,\mathrm{d}F_2(y)$$

provided the integrals are well defined. If $u(y)$ is strictly concave and $F_1 \ne F_2$ then the inequality holds strictly.

Proof. Recall that Fubini's theorem yields the well-known formulae $\int_0^\infty y \,\mathrm{d}F_i(y) = \int_0^\infty (1 - F_i(y)) \,\mathrm{d}y$ and $\int_{-\infty}^0 y \,\mathrm{d}F_i(y) = -\int_{-\infty}^0 F_i(y) \,\mathrm{d}y$. Thus it follows that $\int_I (F_2(y) - F_1(y)) \,\mathrm{d}y = 0$. We know that $u(y)$ is differentiable almost everywhere and continuous. Thus $u(y) = u(y_0) + \int_{y_0}^y u'(z) \,\mathrm{d}z$, where we can define $u'(y)$ as the right derivative. This yields

$$\begin{aligned}\int_{-\infty}^{y_0} u(y) \,\mathrm{d}F_i(y) &= u(y_0)F_i(y_0) - \int_{-\infty}^{y_0}\int_y^{y_0} u'(z) \,\mathrm{d}z \,\mathrm{d}F_i(y) \\ &= u(y_0)F_i(y_0) - \int_{-\infty}^{y_0} F_i(z)u'(z) \,\mathrm{d}z\,.\end{aligned}$$

Analogously

$$\int_{y_0}^{\infty} u(y)\,\mathrm{d}F_i(y) = u(y_0)(1 - F_i(y_0)) + \int_{y_0}^{\infty} (1 - F_i(z))u'(z)\,\mathrm{d}z\,.$$

Putting the results together we find

$$\int_{-\infty}^{\infty} u(y)\,\mathrm{d}F_1(y) - \int_{-\infty}^{\infty} u(y)\,\mathrm{d}F_2(y) = \int_{-\infty}^{\infty} (F_2(y) - F_1(y))u'(y)\,\mathrm{d}y\,.$$

If $y < y_0$ then $F_2(y) - F_1(y) \geq 0$ and $u'(y) \geq u'(y_0)$. If $y > y_0$ then $F_2(y) - F_1(y) \leq 0$ and $u'(y) \leq u'(y_0)$. Thus

$$\int_{-\infty}^{\infty} u(y)\,\mathrm{d}F_1(y) - \int_{-\infty}^{\infty} u(y)\,\mathrm{d}F_2(y) \geq \int_{-\infty}^{\infty} (F_2(y) - F_1(y))u'(y_0)\,\mathrm{d}y = 0\,.$$

The strictly concave case follows analogously. □

Corollary 1.42. *Let X be a real random variable taking values in some interval I_1, and let $g_i : I_1 \to I_2$, $i = 1, 2$ be increasing functions with values on some interval I_2. Suppose*

$$\mathbb{E}[g_1(X)] = \mathbb{E}[g_2(X)] < \infty\,.$$

Let $u : I_2 \to \mathbb{R}$ be a concave function such that $\mathbb{E}[u(g_i(X))]$ is well-defined. If there exists x_0 such that

$$g_1(x) \geq g_2(x), \quad x < x_0, \qquad\qquad g_1(x) \leq g_2(x), \quad x > x_0,$$

then

$$\mathbb{E}[u(g_1(X))] \geq \mathbb{E}[u(g_2(X))]\,.$$

Moreover, if $u(y)$ is strictly concave and $\mathbb{P}[g_1(X) \neq g_2(X)] > 0$ then the inequality is strict.

Proof. Choose $F_i(y) = \mathbb{P}[g_i(X) \leq y]$. Let $y_0 = g_1(x_0)$. If $y < y_0$ then

$$F_1(y) = \mathbb{P}[g_1(X) \leq y] = \mathbb{P}[g_1(X) \leq y, X < x_0] \leq \mathbb{P}[g_2(X) \leq y, X < x_0] \leq F_2(y)\,.$$

If $y > y_0$ then

$$1 - F_1(y) = \mathbb{P}[g_1(X) > y, X > x_0] \leq \mathbb{P}[g_2(X) > y, X > x_0] \leq 1 - F_2(y)\,.$$

The result follows now from Theorem 1.41. □

1.9. Hilbert spaces

1.9.1. Metric spaces

Let $\mathcal{X}$ be a set and $d : \mathcal{X} \times \mathcal{X} \to \mathbb{R}_+$ be some function.

Definition 1.43. *The function $d(x, y)$ is called a* **metric** *if*

i)
$$d(x, y) = 0 \iff x = y\,,$$

ii)
$$d(x, y) = d(y, x)\,,$$

iii)
$$d(x, y) \leq d(x, z) + d(z, y)\,.$$

If $d(x, y)$ is a metric then $(\mathcal{X}, d)$ is called a **metric space**.

Let $x \in \mathcal{X}$. A subset $\mathcal{Y} \subset \mathcal{X}$ is called an environment of x if there is an $\varepsilon > 0$ such that $\{y \in \mathcal{X} : d(x, y) < \varepsilon\} \subset \mathcal{Y}$. We call a subset $\mathcal{Y} \subset \mathcal{X}$ **open** if $\mathcal{Y}$ is an environment of each point $x \in \mathcal{Y}$. $\mathcal{Y}$ is called **closed** if $\mathcal{X} \setminus \mathcal{Y}$ is open.

Let $\{x_n\} \subset \mathcal{X}$ be a sequence. We say $\{x_n\}$ is **convergent** if there exists $x \in \mathcal{X}$ such that $\lim_{n\to\infty} d(x_n, x) = 0$. If $y \in \mathcal{X}$ also has the property $\lim_{n\to\infty} d(x_n, y) = 0$ then $d(x, y) \leq d(x, x_n) + d(x_n, y)$ and it follows that $d(x, y) = 0$, i.e. $x = y$. Thus a limit is unique. We write $\lim_{n\to\infty} x_n = x$.

We have the following characterisation of closed sets.

Lemma 1.44. *A subset $\mathcal{Y} \subset \mathcal{X}$ is closed if and only if for any sequence $\{x_n\} \subset \mathcal{Y}$ that is convergent in $\mathcal{X}$ we have $\lim_{n\to\infty} x_n \in \mathcal{Y}$.*

Proof. Suppose $\mathcal{Y}$ is closed. Let $\{x_n\} \subset \mathcal{Y}$ be convergent and denote the limit by x. Suppose $x \notin \mathcal{Y}$. Then there exists $\varepsilon > 0$ such that $\{y : d(x, y) < \varepsilon\} \subset \mathcal{X} \setminus \mathcal{Y}$. Thus $d(x, y_n) \geq \varepsilon$ for all n, which is a contradiction. Suppose $\mathcal{Y}$ is not closed. Then there is $x \in \mathcal{X} \setminus \mathcal{Y}$ such that $\{y : d(x, y) < \varepsilon\} \cap \mathcal{Y} \neq \emptyset$. Choose $y_n \in \mathcal{Y}$ such that $d(x, y_n) < 1/n$. Then $\lim_{n\to\infty} y_n = x \notin \mathcal{Y}$. $\square$

A sequence $\{x_n\} \subset \mathcal{X}$ is called a **Cauchy sequence** if for each ε there exists a number n_ε such that $d(x_n, x_m) < \varepsilon$ for all $n, m \geq n_\varepsilon$. From $d(x_n, x_m) \leq d(x_n, x) + d(x_m, x)$ it follows that every convergent sequence is a Cauchy sequence. If every Cauchy sequence is convergent, then we say $(\mathcal{X}, d)$ is **complete**.

1.9.2. Vector spaces

Definition 1.45. *A set $\mathcal{X}$ is called* **vector space** *if there are a vector addition + and a multiplication by real numbers such that*

i) $$\boldsymbol{x} + \boldsymbol{y} = \boldsymbol{y} + \boldsymbol{x}\,,$$

ii) $$\boldsymbol{x} + (\boldsymbol{y} + \boldsymbol{z}) = (\boldsymbol{x} + \boldsymbol{y}) + \boldsymbol{z}\,,$$

iii) $$c(\boldsymbol{x} + \boldsymbol{y}) = c\boldsymbol{x} + c\boldsymbol{y}\,,$$

iv) $$(c + d)\boldsymbol{x} = c\boldsymbol{x} + d\boldsymbol{x}\,,$$

v) $$(cd)\boldsymbol{x} = c(d\boldsymbol{x})\,,$$

vi) $$1\boldsymbol{x} = \boldsymbol{x}$$

vii) *There is a neutral element $\mathbf{0}$ such that $\boldsymbol{x} + \mathbf{0} = \boldsymbol{x}$ for all $\boldsymbol{x} \in \mathcal{X}$.*

viii) $$0\boldsymbol{x} = \mathbf{0}\,.$$

Note that $\mathbf{0}$ is unique. Let $\boldsymbol{y}$ be an element such that $\boldsymbol{x} + \boldsymbol{y} = \boldsymbol{x}$ for all $\boldsymbol{x}$. Then $\boldsymbol{y} = \boldsymbol{y} + \mathbf{0} = \mathbf{0}$. Moreover,

$$\boldsymbol{x} + (-1)\boldsymbol{x} = (1 + (-1))\boldsymbol{x} = 0\boldsymbol{x} = \mathbf{0}\,.$$

Thus $(-1)\boldsymbol{x}$ is the inverse element of $\boldsymbol{x}$. We write $-\boldsymbol{x} = (-1)\boldsymbol{x}$ and $\boldsymbol{x} - \boldsymbol{y} = \boldsymbol{x} + (-\boldsymbol{y})$.

Definition 1.46. *A norm $\|\cdot\| : \mathcal{X} \to \mathbb{R}_+$ on a vector space $\mathcal{X}$ is a mapping such that*

i) $$\|\boldsymbol{x} + \boldsymbol{y}\| \le \|\boldsymbol{x}\| + \|\boldsymbol{y}\|\,,$$

ii) $$\|c\boldsymbol{x}\| = |c|\|\boldsymbol{x}\|\,,$$

iii) $$\|\boldsymbol{x}\| = 0 \quad \Longrightarrow \quad \boldsymbol{x} = \mathbf{0}\,.$$

It follows readily that the mapping $d(\boldsymbol{x}, \boldsymbol{y}) = \|\boldsymbol{x} - \boldsymbol{y}\|$ is a metric. We now consider any normed space as a metric space endowed with the above metric. A complete normed space is called a **Banach space**.

1.9.3. Hilbert spaces

Let $\mathcal{X}$ be a vector space.

Definition 1.47. *An* **inner product** *is a mapping* $\langle \cdot, \cdot \rangle : \mathcal{X} \times \mathcal{X} \to \mathrm{I\!R}$ *with the following properties*

i)
$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \langle \boldsymbol{y}, \boldsymbol{x} \rangle \,,$$

ii)
$$\langle c\boldsymbol{x} + d\boldsymbol{y}, \boldsymbol{z} \rangle = c\langle \boldsymbol{x}, \boldsymbol{z} \rangle + d\langle \boldsymbol{y}, \boldsymbol{z} \rangle \,,$$

iii)
$$\langle \boldsymbol{x}, \boldsymbol{x} \rangle \geq 0 \,,$$

iv)
$$\langle \boldsymbol{x}, \boldsymbol{x} \rangle = 0 \quad \Longrightarrow \quad \boldsymbol{x} = \boldsymbol{0} \,.$$

A space $(\mathcal{X}, \langle \cdot, \cdot \rangle)$ *is called an* **inner product space**.

Consider the mapping $\| \cdot \| : \mathcal{X} \to \mathrm{I\!R}_+$, with $\|\boldsymbol{x}\| = \sqrt{\langle \boldsymbol{x}, \boldsymbol{x} \rangle}$. We have that

$$0 \leq \|\boldsymbol{x} - (\langle \boldsymbol{x}, \boldsymbol{y} \rangle / \|\boldsymbol{y}\|^2)\boldsymbol{y}\|^2 = \|\boldsymbol{x}\|^2 - \langle \boldsymbol{x}, \boldsymbol{y} \rangle^2 / \|\boldsymbol{y}\|^2$$

and it follows that $|\langle \boldsymbol{x}, \boldsymbol{y} \rangle| \leq \|\boldsymbol{x}\| \|\boldsymbol{y}\|$. This implies

$$\|\boldsymbol{x} + \boldsymbol{y}\|^2 = \langle \boldsymbol{x} + \boldsymbol{y}, \boldsymbol{x} + \boldsymbol{y} \rangle = \|\boldsymbol{x}\|^2 + 2\langle \boldsymbol{x}, \boldsymbol{y} \rangle + \|\boldsymbol{y}\|^2 \leq (\|\boldsymbol{x}\| + \|\boldsymbol{y}\|)^2 \,.$$

Hence $\| \cdot \|$ is a norm. We now identify each inner product space with the corresponding normed space. If the space $(\mathcal{X}, \langle \cdot, \cdot \rangle)$ is complete, we call it a **Hilbert space**.

Let $(\mathcal{X}, \langle \cdot, \cdot \rangle)$ be an inner product space. Then

$$\|\boldsymbol{x}+\boldsymbol{y}\|^2 + \|\boldsymbol{x}-\boldsymbol{y}\|^2 = \|\boldsymbol{x}\|^2 + 2\langle \boldsymbol{x}, \boldsymbol{y} \rangle + \|\boldsymbol{y}\|^2 + \|\boldsymbol{x}\|^2 - 2\langle \boldsymbol{x}, \boldsymbol{y} \rangle + \|\boldsymbol{y}\|^2 = 2\|\boldsymbol{x}\| + 2\|\boldsymbol{y}\|^2 \,. \tag{1.10}$$

Let $\mathcal{X}$ be a vector space. A subset $\mathcal{Y}$ is called **convex** if for any $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{Y}$ we have $\alpha\boldsymbol{x} + (1-\alpha)\boldsymbol{y} \in \mathcal{Y}$ for all $\alpha \in (0,1)$. A linear space is necessarily convex.

Let now $(\mathcal{X}, \langle \cdot, \cdot \rangle)$ be a Hilbert space and $\mathcal{Y} \subset \mathcal{X}$ be a subset. Let $\boldsymbol{x} \in \mathcal{X}$. The distance from $\boldsymbol{x}$ to $\mathcal{Y}$ is defined as $d(\boldsymbol{x}; \mathcal{Y}) = \inf_{\boldsymbol{y} \in \mathcal{Y}} \|\boldsymbol{x} - \boldsymbol{y}\|$.

Lemma 1.48. *Suppose* $\mathcal{Y}$ *is closed and convex. There exists a unique* $\boldsymbol{y}_0 \in \mathcal{Y}$ *such that* $d(\boldsymbol{x}; \mathcal{Y}) = \|\boldsymbol{x} - \boldsymbol{y}_0\|$.

Proof. Let $\{\boldsymbol{y}_n\} \subset \mathcal{Y}$ be a sequence such that $\|\boldsymbol{x} - \boldsymbol{y}_n\|$ tends to $d(\boldsymbol{x};\mathcal{Y})$. By (1.10) we have

$$\|(\boldsymbol{y}_n - \boldsymbol{x}) - (\boldsymbol{y}_m - \boldsymbol{x})\|^2 + \|(\boldsymbol{y}_n - \boldsymbol{x}) + (\boldsymbol{y}_m - \boldsymbol{x})\|^2 = 2\|\boldsymbol{y}_n - \boldsymbol{x}\|^2 + 2\|\boldsymbol{y}_m - \boldsymbol{x}\|^2$$

or equivalently

$$\|\boldsymbol{y}_n - \boldsymbol{y}_m\|^2 = 2\|\boldsymbol{y}_n - \boldsymbol{x}\|^2 + 2\|\boldsymbol{y}_m - \boldsymbol{x}\|^2 - 4\|\tfrac{1}{2}(\boldsymbol{y}_n + \boldsymbol{y}_m) - \boldsymbol{x}\|^2 .$$

Since $\mathcal{Y}$ is convex be have $\frac{1}{2}(\boldsymbol{y}_n + \boldsymbol{y}_m) \in \mathcal{Y}$, implying

$$\|\boldsymbol{y}_n - \boldsymbol{y}_m\|^2 \le 2\|\boldsymbol{y}_n - \boldsymbol{x}\|^2 + 2\|\boldsymbol{y}_m - \boldsymbol{x}\|^2 - 4d(\boldsymbol{x};\mathcal{Y})^2 .$$

This shows that $\{\boldsymbol{y}_n\}$ is a Cauchy sequence. Thus there is $\boldsymbol{y}_0 \in \mathcal{X}$ such that $\lim_{n\to\infty} \boldsymbol{y}_n = \boldsymbol{y}_0$. Because $\mathcal{Y}$ is closed we have $\boldsymbol{y}_0 \in \mathcal{Y}$. From

$$d(\boldsymbol{x};\mathcal{Y}) \le \|\boldsymbol{x} - \boldsymbol{y}_0\| \le \|\boldsymbol{x} - \boldsymbol{y}_n\| + \|\boldsymbol{y}_n - \boldsymbol{y}_0\| \to d(\boldsymbol{x};\mathcal{Y})$$

it follows that $\|\boldsymbol{x} - \boldsymbol{y}_0\| = d(\boldsymbol{x};\mathcal{Y})$. We now have to show that $\boldsymbol{y}_0$ is unique. Let $\boldsymbol{z} \in \mathcal{Y}$ such that $\|\boldsymbol{x} - \boldsymbol{z}\| = d(\boldsymbol{x};\mathcal{Y})$. Then $\alpha\boldsymbol{y}_0 + (1-\alpha)\boldsymbol{z} \in \mathcal{Y}$ for all $\alpha \in [0,1]$. Thus $\|\boldsymbol{x} - \alpha\boldsymbol{y}_0 - (1-\alpha)\boldsymbol{z}\| \ge d(\boldsymbol{x};\mathcal{Y})$. The triangle inequality gives

$$\|\boldsymbol{x} - \alpha\boldsymbol{y}_0 - (1-\alpha)\boldsymbol{z}\| \le \alpha\|\boldsymbol{x} - \boldsymbol{y}_0\| + (1-\alpha)\|\boldsymbol{x} - \boldsymbol{z}\| = d(\boldsymbol{x};\mathcal{Y})$$

and $\|\boldsymbol{x} - \alpha\boldsymbol{y}_0 - (1-\alpha)\boldsymbol{z}\| = d(\boldsymbol{x};\mathcal{Y})$. On the other hand

$$\|\boldsymbol{x} - \alpha\boldsymbol{y}_0 - (1-\alpha)\boldsymbol{z}\|^2 = \|\boldsymbol{x} - \boldsymbol{z}\|^2 + 2\alpha\langle \boldsymbol{x} - \boldsymbol{z}, \boldsymbol{z} - \boldsymbol{y}_0\rangle + \alpha^2\|\boldsymbol{z} - \boldsymbol{y}_0\|^2 .$$

The right hand side can only be constant if $\|\boldsymbol{z} - \boldsymbol{y}_0\| = 0$. □

We say two vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ are **orthogonal** if $\langle \boldsymbol{x}, \boldsymbol{y}\rangle = 0$. We write $\boldsymbol{x}\perp\boldsymbol{y}$. Orthogonal is equivalent to $\|\boldsymbol{x} + \boldsymbol{y}\|^2 = \|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2$, known as **Pythagoras equality**.

Let $\mathcal{Y} \subset \mathcal{X}$ be a subset. A vector $\boldsymbol{x} \in \mathcal{X}$ is said to be **orthogonal** to $\mathcal{Y}$, $\boldsymbol{x}\perp\mathcal{Y}$, if $\boldsymbol{x}\perp\boldsymbol{y}$ for all $\boldsymbol{y} \in \mathcal{Y}$. The set of vectors $\{\boldsymbol{x} : \boldsymbol{x}\perp\mathcal{Y}\}$ is called the **annihilator** of $\mathcal{Y}$ and is denoted by $\mathcal{Y}^\perp$. Obviously, $\mathcal{Y}^\perp$ is a linear subspace of $\mathcal{X}$. Two subsets $\mathcal{Y}$ and $\mathcal{Z}$ are **orthogonal** if $\mathcal{Z} \subset \mathcal{Y}^\perp$, i.e. $\boldsymbol{x}\perp\boldsymbol{y}$ for all $\boldsymbol{x} \in \mathcal{Y}$ and $\boldsymbol{y} \in \mathcal{Z}$.

Let $\mathcal{Y} \subset \mathcal{X}$ be a linear subspace and $\boldsymbol{x} \in \mathcal{X}$. Suppose there is $\boldsymbol{y} \in \mathcal{Y}$ such that $\|\boldsymbol{x} - \boldsymbol{y}\| = d(\boldsymbol{x};\mathcal{Y})$. Then $\boldsymbol{y}$ is unique. The latter follows as in the proof of Lemma 1.48. We call $\boldsymbol{y}$ the **orthogonal projection** of $\boldsymbol{x}$ onto $\mathcal{Y}$ and denote it by $\mathrm{pro}(\boldsymbol{x} \mid \mathcal{Y})$. Let $\boldsymbol{z} \in \mathcal{Y}$. Then $\boldsymbol{y} + \alpha\boldsymbol{z} \in \mathcal{Y}$ for all $\alpha \in \mathbb{R}$. The function

$$\alpha \mapsto \|\boldsymbol{x} - \boldsymbol{y} - \alpha\boldsymbol{z}\|^2 = \|\boldsymbol{x} - \boldsymbol{y}\|^2 - 2\alpha\langle \boldsymbol{x} - \boldsymbol{y}, \boldsymbol{z}\rangle + \alpha^2\|\boldsymbol{z}\|^2$$

has its minimum at $\alpha = 0$. Thus $\langle \boldsymbol{x} - \boldsymbol{y}, \boldsymbol{z} \rangle = 0$. It follows that $\boldsymbol{x} - \boldsymbol{y} \perp \mathcal{Y}$. Thus we can write $\boldsymbol{x} = \boldsymbol{y} + (\boldsymbol{x} - \boldsymbol{y})$ as the sum of a vector from $\mathcal{Y}$ and a vector from $\mathcal{Y}^\perp$.

Observe that $\{\boldsymbol{0}\} \subset \mathcal{Y} \cap \mathcal{Y}^\perp$. Let $\boldsymbol{x} \in \mathcal{Y} \cap \mathcal{Y}^\perp$. Then $\boldsymbol{x} \perp \boldsymbol{x}$, i.e. $\langle \boldsymbol{x}, \boldsymbol{x} \rangle = 0$ implying that $\boldsymbol{x} = \boldsymbol{0}$. Thus $\mathcal{Y} \cap \mathcal{Y}^\perp = \{\boldsymbol{0}\}$. Let now $\boldsymbol{x} \in \mathcal{X}$. Suppose $\operatorname{pro}(\boldsymbol{x} \mid \mathcal{Y})$ exists and that there are vectors $\boldsymbol{y} \in \mathcal{Y}$ and $\boldsymbol{z} \in \mathcal{Y}^\perp$ such that $\boldsymbol{x} = \boldsymbol{y} + \boldsymbol{z}$. Then $\operatorname{pro}(\boldsymbol{x} \mid \mathcal{Y}) + (\boldsymbol{x} - \operatorname{pro}(\boldsymbol{x} \mid \mathcal{Y})) = \boldsymbol{y} + \boldsymbol{z}$ or equivalently $\operatorname{pro}(\boldsymbol{x} \mid \mathcal{Y}) - \boldsymbol{y} = \boldsymbol{z} - (\boldsymbol{x} - \operatorname{pro}(\boldsymbol{x} \mid \mathcal{Y}))$. Thus $\operatorname{pro}(\boldsymbol{x} \mid \mathcal{Y}) - \boldsymbol{y} \in \mathcal{Y} \cap \mathcal{Y}^\perp$. This implies $\boldsymbol{y} = \operatorname{pro}(\boldsymbol{x} \mid \mathcal{Y})$ and the decomposition is unique. We therefore write $\boldsymbol{x}_{\mathcal{Y}} = \operatorname{pro}(\boldsymbol{x} \mid \mathcal{Y})$ and $\boldsymbol{x}_{\mathcal{Y}^\perp} = \boldsymbol{x} - \operatorname{pro}(\boldsymbol{x} \mid \mathcal{Y})$. Because $\boldsymbol{x}_{\mathcal{Y}} \perp \boldsymbol{x}_{\mathcal{Y}^\perp}$ we get from $\|\boldsymbol{x}\|^2 = \|\boldsymbol{x}_{\mathcal{Y}}\|^2 + \|\boldsymbol{x}_{\mathcal{Y}^\perp}\|^2$ the useful formula $\|\boldsymbol{x}_{\mathcal{Y}}\|^2 = \|\boldsymbol{x}\|^2 - \|\boldsymbol{x}_{\mathcal{Y}^\perp}\|^2$.

If $\mathcal{Y}$ is complete then $\operatorname{pro}(\boldsymbol{x} \mid \mathcal{Y})$ exists for all $\boldsymbol{x} \in \mathcal{X}$, see the proof of Lemma 1.48. One easily checks that

$$\operatorname{pro}\Big(\sum_{i=1}^k \alpha_i \boldsymbol{x}_i \;\Big|\; \mathcal{Y}\Big) = \sum_{i=1}^k \alpha_i \operatorname{pro}(\boldsymbol{x}_i \mid \mathcal{Y})\,.$$

Thus the projection is linear. One also checks easily that $\operatorname{pro}(\operatorname{pro}(\boldsymbol{x} \mid \mathcal{Y}) \mid \mathcal{Y}) = \operatorname{pro}(\boldsymbol{x} \mid \mathcal{Y})$.

Let $\mathcal{Y}$ and $\mathcal{Z}$ be closed linear subspaces of $\mathcal{X}$ such that $\mathcal{Y} \subset \mathcal{Z}$. Let us find $\boldsymbol{y} = \operatorname{pro}(\operatorname{pro}(\boldsymbol{x} \mid \mathcal{Z}) \mid \mathcal{Y})$. For $\boldsymbol{z} = \operatorname{pro}(\boldsymbol{x} \mid \mathcal{Z})$ we can write

$$\boldsymbol{x} = (\boldsymbol{x} - \boldsymbol{z}) + (\boldsymbol{z} - \boldsymbol{y}) + \boldsymbol{y}\,.$$

We have $\boldsymbol{x} - \boldsymbol{z} \in \mathcal{Z}^\perp \subset \mathcal{Y}^\perp$ and $\boldsymbol{z} - \boldsymbol{y} \in \mathcal{Y}^\perp$. By the uniqueness of the decomposition we find $\operatorname{pro}(\boldsymbol{x} \mid \mathcal{Y}) = \boldsymbol{y}$. Hence we found the iterated projections

$$\operatorname{pro}(\operatorname{pro}(\boldsymbol{x} \mid \mathcal{Z}) \mid \mathcal{Y}) = \operatorname{pro}(\boldsymbol{x} \mid \mathcal{Y})\,.$$

We can nest iterated projections. Let $\{\boldsymbol{0}\} = \mathcal{X}_0 \subset \mathcal{X}_1 \subset \cdots \subset \mathcal{X}_n = \mathcal{X}$ be nested closed subspaces and let $\boldsymbol{x}_j = \operatorname{pro}(\boldsymbol{x} \mid \mathcal{X}_j) - \operatorname{pro}(\boldsymbol{x} \mid \mathcal{X}_{j-1}) \in \mathcal{X}_j \cap \mathcal{X}_{j-1}^\perp$. Then $\boldsymbol{x} = \boldsymbol{x}_1 + \cdots + \boldsymbol{x}_n$. The vectors $\{\boldsymbol{x}_i\}$ are orthogonal, yielding $\|\boldsymbol{x}\|^2 = \|\boldsymbol{x}_1\|^2 + \cdots + \|\boldsymbol{x}_n\|^2$. Alternatively, we could write $\operatorname{pro}(\boldsymbol{x} \mid \mathcal{X}_j) = \boldsymbol{x}_1 + \cdots + \boldsymbol{x}_j$ and $\|\operatorname{pro}(\boldsymbol{x} \mid \mathcal{X}_j)\|^2 = \|\boldsymbol{x}_1\|^2 + \cdots + \|\boldsymbol{x}_j\|^2$.

1.9.4. Special Hilbert spaces

Example 1.49. Let $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ be a symmetric positive definite matrix. Consider the space $\mathcal{L}^2$ of all square integrable random vectors $\boldsymbol{X} \in \mathbb{R}^n$. We identify random variables which are a.s. the same. Endowed with the usual addition and multiplication by scalars $\mathcal{L}^2$ is a vector space. Let $\langle \boldsymbol{X}, \boldsymbol{Y} \rangle = \mathbb{E}[\boldsymbol{X}^\top \boldsymbol{A} \boldsymbol{Y}]$. By the properties of $\boldsymbol{A}$ $\langle \cdot, \cdot \rangle$ is an inner product. We want to show that $(\mathcal{L}^2, \langle \cdot, \cdot \rangle)$ is a Hilbert space. One can show that it is no loss of generality to assume $\boldsymbol{A} = \boldsymbol{I}$. For the two spaces $\boldsymbol{X}_n \to \boldsymbol{X}$ is equivalent, and both spaces have the same Cauchy sequences. Let $\{\boldsymbol{X}_k\} \subset \mathcal{L}_2$ be a Cauchy sequence. Note that now $\|\boldsymbol{X}\| = \||\boldsymbol{X}|\|$, where the absolute value is taken entry-wise. Choose n_k such that $\|\boldsymbol{X}_i - \boldsymbol{X}_j\| < k^{-2}$ for $i, j \ge n_k$. Without loss of generality we can assume that n_k is strictly increasing in k. Let $\boldsymbol{Y}_k = \boldsymbol{X}_{n_{k+1}} - \boldsymbol{X}_{n_k}$. Then

$$\Big\| \sum_{i=1}^k |\boldsymbol{Y}_i| \Big\| \le \sum_{i=1}^k \|\boldsymbol{Y}_i\| \le \sum_{i=1}^\infty \|\boldsymbol{Y}_i\| \le \sum_{i=1}^\infty i^{-2} < \infty\,.$$

Thus by monotone convergence the Euclidean norm of $\sum_{i=1}^\infty |\boldsymbol{Y}_i|$ has finite expectation and is therefore finite almost surely. Hence $\sum_{i=1}^\infty \boldsymbol{Y}_i$ exists almost surely. Let $\boldsymbol{X} = \boldsymbol{X}_{n_1} + \sum_{i=1}^\infty \boldsymbol{Y}_i$. We show that $\boldsymbol{X}_{n_k}$ converges to $\boldsymbol{X}$. We have

$$\|\boldsymbol{X}_{n_k} - \boldsymbol{X}\| = \Big\| \sum_{i=k}^\infty \boldsymbol{Y}_i \Big\| \le \sum_{i=k}^\infty \|\boldsymbol{Y}_i\| \le \sum_{i=k}^\infty i^{-2}\,.$$

The right hand side converges to zero as $k \to \infty$. Let now n be arbitrary. Let $k(n) = \sup\{m \in \mathbb{N} : n_m \le n\}$ and note that $k(n) \to \infty$ as $n \to \infty$. Then

$$\|\boldsymbol{X}_n - \boldsymbol{X}\| \le \|\boldsymbol{X}_n - \boldsymbol{X}_{k(n)}\| + \|\boldsymbol{X}_{k(n)} - \boldsymbol{X}\| \le k(n)^{-2} + \sum_{i=k(n)}^\infty i^{-2}\,.$$

The right hand side tends to zero as $n \to \infty$. Hence $(\mathcal{L}^2, \langle \cdot, \cdot \rangle)$ is a Hilbert space. ■

1.10. Matrix algebra

Definition 1.50. *A $m \times n$* **matrix** $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ *is a rectangular scheme of numbers*

$$\boldsymbol{A} = (a_{ij}) = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}.$$

The **transpose** $\boldsymbol{A}^\top$ *of $\boldsymbol{A}$ is the $n \times m$ matrix*

$$\boldsymbol{A}^\top = \begin{pmatrix} a_{11} & \cdots & a_{m1} \\ \vdots & \ddots & \vdots \\ a_{1n} & \cdots & a_{mn} \end{pmatrix},$$

where the rôle of rows and columns are interchanged. If $n = m$ we say $\boldsymbol{A}$ is a **square matrix**. *A square matrix $\boldsymbol{A}$ is called symmetric, if $\boldsymbol{A}^\top = \boldsymbol{A}$. A* **vector** $\boldsymbol{x} \in \mathbb{R}^n$ *is a $n \times 1$ matrix.*

A **diagonal matrix** is a square matrix with off-diagonal element zero. More specifically, we let

$$\operatorname{diag}(a_1, \ldots, a_n) = \begin{pmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_n \end{pmatrix}.$$

We define the sum of two $m \times n$ matrices as the entry-wise sum, $\boldsymbol{A}+\boldsymbol{B} = (a_{ij} + b_{ij})$. The multiplication with a scalar $\alpha \in \mathbb{R}$ is defined as multiplication of each entry, $\alpha\boldsymbol{A} = (\alpha a_{ij})$. We now also define a matrix multiplication. Let $\boldsymbol{A}$ be a $m \times n$ matrix and $\boldsymbol{B}$ be a $n \times k$ matrix. Then the product $\boldsymbol{C} = \boldsymbol{AB}$ is the $m \times k$ matrix $\boldsymbol{C} = (c_{ij})$ with entries

$$c_{ij} = \sum_{h=1}^{n} a_{ih} b_{hj}\,.$$

From the definition it follows that $(\boldsymbol{AB})^\top = \boldsymbol{B}^\top \boldsymbol{A}^\top$. Matrix multiplication has the neutral element $\boldsymbol{I} = \operatorname{diag}(1, \ldots, 1)$, i.e. $\boldsymbol{AI} = \boldsymbol{IA} = \boldsymbol{A}$ for all $\boldsymbol{A} \in \mathbb{R}^{n\times n}$.

A set of vectors $\boldsymbol{a}_1, \ldots \boldsymbol{a}_r$ in $\mathbb{R}^n$ is said to be **linearly dependent** if the only numbers $c_1, \ldots, c_r$ such that $\sum_{i=1}^{r} c_r \boldsymbol{a}_r = \boldsymbol{0}$ are $c_1 = \cdots = c_r = 0$. Otherwise, we say the vectors are **linearly dependent**. Considering the matrix $\boldsymbol{A} = (\boldsymbol{a}_1, \ldots, \boldsymbol{a}_r)$ then we can formulate linear independence also in the following way. If $\boldsymbol{Ac} = \boldsymbol{0}$ then $\boldsymbol{c} = \boldsymbol{0}$. Note that if $r > n$ then necessarily the vectors are linearly dependent.

Let $\mathcal{L}$ be a **linear subspace** of $\mathbb{R}^n$, i.e. if $\boldsymbol{a}, \boldsymbol{b} \in \mathcal{L}$ then $\alpha\boldsymbol{a} + \beta\boldsymbol{b} \in \mathcal{L}$ for all $\alpha, \beta \in \mathbb{R}$. A **basis** for $\mathcal{L}$ is a set of linearly independent vectors $\boldsymbol{b}_1, \ldots \boldsymbol{b}_q$ such that for each vector $\boldsymbol{a} \in \mathcal{L}$ there are numbers $c_1, \ldots c_q$ such that $\boldsymbol{a} = c_1\boldsymbol{b}_1 + \cdots + c_q\boldsymbol{b}_q$. One can show that a basis always exists, and the number q is unique. The number q is called the **dimension** of $\mathcal{L}$.

We denote by $\mathbb{R}(\boldsymbol{A})$ the space $\{\boldsymbol{b} : \boldsymbol{b} = \boldsymbol{Ac} \text{ for some } \boldsymbol{c} \in \mathbb{R}^r\}$, the space spanned by the column vectors of $\boldsymbol{A}$. The dimension of $\mathbb{R}(\boldsymbol{A})$ is call the **rank** of $\boldsymbol{A}$ and denoted by $\operatorname{rank}(\boldsymbol{A})$. One can proof that $\operatorname{rank}(\boldsymbol{A}) = \operatorname{rank}(\boldsymbol{A}^\top)$. If $\boldsymbol{A}$ is an $m \times n$ matrix, then $\operatorname{rank}(\boldsymbol{A}) \leq \min(m, n)$. If $\operatorname{rank}(\boldsymbol{A}) = \min(m, n)$ then we say $\boldsymbol{A}$ is of **full rank**. If $\boldsymbol{AB}$ is well-defined, then $\operatorname{rank}(\boldsymbol{AB}) \leq \min(\operatorname{rank}(\boldsymbol{A}), \operatorname{rank}(\boldsymbol{B}))$.

If $\boldsymbol{A}$ is a square matrix then $\boldsymbol{A}$ is of full rank if and only if there exists an **inverse** matrix $\boldsymbol{A}^{-1}$, that is

$$\boldsymbol{A}^{-1}\boldsymbol{A} = \boldsymbol{A}\boldsymbol{A}^{-1} = \boldsymbol{I}\,.$$

The matrix $\boldsymbol{A}^{-1}$ is unique. We say that $\boldsymbol{A}$ is **invertible** or **non-singular**. Obviously, $\boldsymbol{A}$ is singular if the equation $\boldsymbol{Ax} = \boldsymbol{0}$ has a solution $\boldsymbol{x} \neq \boldsymbol{0}$. It follows readily, that if $\boldsymbol{AB}$ is invertible, then $(\boldsymbol{AB})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$.

The following identity is sometimes useful.

Lemma 1.51. *Suppose $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ is invertible, $\boldsymbol{B} \in \mathbb{R}^{n\times m}$ and $\boldsymbol{C} \in \mathbb{R}^{m\times m}$ is invertible such that the inverses below exist. Then*

$$(\boldsymbol{A} + \boldsymbol{BCB}^\top)^{-1} = \boldsymbol{A}^{-1} - \boldsymbol{A}^{-1}\boldsymbol{B}(\boldsymbol{C}^{-1} + \boldsymbol{B}^\top\boldsymbol{A}^{-1}\boldsymbol{B})^{-1}\boldsymbol{B}^\top\boldsymbol{A}^{-1}\,.$$

Proof. Let $\boldsymbol{D} = (\boldsymbol{A} + \boldsymbol{BCB}^\top)^{-1}$. Then by definition

$$\boldsymbol{DA} + \boldsymbol{DBCB}^\top = \boldsymbol{I}\,.$$

Multiplying with $\boldsymbol{A}^{-1}\boldsymbol{B}$ from the right gives

$$\boldsymbol{DB} + \boldsymbol{DBCB}^\top\boldsymbol{A}^{-1}\boldsymbol{B} = \boldsymbol{A}^{-1}\boldsymbol{B}\,.$$

Solving for $\boldsymbol{DB}$ yields

$$\boldsymbol{DB} = \boldsymbol{A}^{-1}\boldsymbol{B}(\boldsymbol{I} + \boldsymbol{CB}^\top\boldsymbol{A}^{-1}\boldsymbol{B})^{-1}\,.$$

If we plug the expression into the first equation we get

$$\boldsymbol{DA} + \boldsymbol{A}^{-1}\boldsymbol{B}(\boldsymbol{I} + \boldsymbol{CB}^\top\boldsymbol{A}^{-1}\boldsymbol{B})^{-1}\boldsymbol{CB}^\top = \boldsymbol{I}$$

Solving for $\boldsymbol{D}$ gives the desired expression noting that

$$(\boldsymbol{I} + \boldsymbol{CB}^\top\boldsymbol{A}^{-1}\boldsymbol{B})^{-1} = (\boldsymbol{C}^{-1} + \boldsymbol{B}^\top\boldsymbol{A}^{-1}\boldsymbol{B})^{-1}\boldsymbol{C}^{-1}\,.$$

□

The **trace** of a square matrix $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ is the sum of its diagonal entries,

$$\operatorname{tr}(\boldsymbol{A}) = \sum_{i=1}^{n} a_{ii}\,.$$

The trace is a linear operator $\operatorname{tr}(\alpha\boldsymbol{A} + \beta\boldsymbol{B}) = \alpha\operatorname{tr}(\boldsymbol{A}) + \beta\operatorname{tr}(\boldsymbol{B})$. The trace has also the property that $\operatorname{tr}(\boldsymbol{AB}) = \operatorname{tr}(\boldsymbol{BA})$ if both $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ and $\boldsymbol{B} \in \mathbb{R}^{n\times m}$.

We say two vectors $\boldsymbol{a}_1, \boldsymbol{a}_2$ are **orthogonal** if $\boldsymbol{a}_1^\top\boldsymbol{a}_2 = 0$. We sometimes write $\boldsymbol{a}_1 \perp \boldsymbol{a}_2$. Let $\boldsymbol{c}_1, \ldots, \boldsymbol{c}_n$ be a basis in $\mathbb{R}^n$. We say $\boldsymbol{c}_1, \ldots, \boldsymbol{c}_n$ is an **orthonormal basis** if $\boldsymbol{c}_i^\top\boldsymbol{c}_j = \delta_{ij}$, where δ_{ij} are the Kronecker delta (1 if $i = j$, 0 otherwise). The $n \times n$ matrix $\boldsymbol{C} = (\boldsymbol{c}_1, \ldots, \boldsymbol{c}_n)$ is then called **orthogonal**. A matrix $\boldsymbol{C}$ is orthogonal if and only if $\boldsymbol{C}^{-1} = \boldsymbol{C}^\top$. Thus also $\boldsymbol{C}^\top$ is orthogonal. It follows readily that a product of two orthogonal matrices is orthogonal.

The following result is often useful.

Lemma 1.52. *Let $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ be a symmetric matrix. Then there exists a diagonal matrix $\boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$ and an orthogonal matrix $\boldsymbol{C}$ such that $\boldsymbol{A} = \boldsymbol{C}\boldsymbol{\Lambda}\boldsymbol{C}^\top$.*

Let $\boldsymbol{A} \in \mathbb{R}^{n\times n}$. The function $q : \mathbb{R}^n \to \mathbb{R}$, $\boldsymbol{x} \mapsto \boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x}$ is called **quadratic form**. Without loss of generality, $\boldsymbol{A}$ can be chosen to be symmetric ($a_{ij}x_ix_j + a_{ji}x_jx_i = \frac{1}{2}(a_{ij} + a_{ji})x_ix_j + \frac{1}{2}(a_{ij} + a_{ji})x_jx_i$) and we will assume that this is the case for all quadratic forms. From the definition one has $q(\mathbf{0}) = 0$. If $q(\boldsymbol{x}) \geq 0$ for all $\boldsymbol{x}$ then we say $\boldsymbol{A}$ is **positive semi-definite**. If $q(\boldsymbol{x}) > 0$ for all $\boldsymbol{x} \neq \mathbf{0}$ then we say $\boldsymbol{A}$ is **positive definite**. Let $\boldsymbol{\Lambda}$ be a diagonal matrix and $\boldsymbol{C}$ be an orthogonal matrix such that $\boldsymbol{A} = \boldsymbol{C}\boldsymbol{\Lambda}\boldsymbol{C}^\top$ then $\boldsymbol{A}$ is positive (semi-) definite if and only of $\boldsymbol{\Lambda}$ is positive (semi-) definite. Thus positive semi-definite is equivalent to $\lambda_i \geq 0$ and positive definite is equivalent to $\lambda_i > 0$. It follows that a positive definite matrix is invertible.

1.11. Bibliographical remarks

For further literature on stochastic processes see for instance [19], [25], [37], [38], [46] or [63]. The theory of point processes can also be found in [12] or [18]. For further literature on renewal theory see also [1] and [26]. Section 1.6 is taken from [26]. The class of subexponential distributions was introduced by Chistyakov [14]. The results presented here can be found in [14], [8] and [22].

2. The Cramér-Lundberg model

2.1. Definition of the Cramér-Lundberg process

Modelling a risk by a compound Poisson distribution is very popular. That is because the compound Poisson model has nice properties. For instance it can be derived as a limit of individual models. This was the reason for Filip Lundberg to postulate a continuous time risk model where the aggregate claims in any interval have a compound Poisson distribution. Moreover, the premium income should be modelled. In a portfolio of insurance contracts the premium payments will be spread all over the year. Thus he assumed that that the premium income is continuous over time and that the premium income in any time interval is proportional to the interval length. This leads to the following model for the surplus of an insurance portfolio

$$C_t = u + ct - \sum_{i=1}^{N_t} Y_i \,.$$

u is the initial capital, c is the premium rate. The number of claims in $(0, t]$ is a Poisson process N_t with rate λ. The claim sizes $\{Y_i\}$ are a sequence of iid positive random variables independent of $\{N_t\}$. This model is called **Cramér-Lundberg process** or **classical risk process**.

We denote the distribution function of the claims by G, its moments by $\mu_n = \mathbb{E}[Y_1^n]$ and its moment generating function by $M_Y(r) = \mathbb{E}[\exp\{rY_1\}]$. Let $\mu = \mu_1$. We assume that $\mu < \infty$. Otherwise no insurance company would insure such a risk. Note that $G(x) = 0$ for $x < 0$. We will see later that it is no restriction to assume that $G(0) = 0$.

For an insurance company it is important that $\{C_t\}$ stays above a certain level. This level is given by legal restrictions. By adjusting the initial capital it is no loss of generality if we assume this level to be 0. We define the ruin time

$$\tau = \inf\{t > 0 : C_t < 0\}\,, \qquad\qquad (\inf \emptyset = \infty)\,.$$

We will mostly be interested in the probability of ruin in a time interval $(0, t]$

$$\psi(u, t) = \mathbb{P}[\tau \le t \mid C_0 = u] = \mathbb{P}[\inf_{0<s\le t} C_s < 0]$$

and the probability of ultimate ruin

$$\psi(u) = \lim_{t\to\infty} \psi(u, t) = \mathbb{P}[\tau < \infty \mid C_0 = u] = \mathbb{P}[\inf_{t>0} C_t < 0]\,.$$

It is easy to see that $\psi(u,t)$ is decreasing in u and increasing in t.

As in Chapter 1 we denote the claim times by $T_1, T_2, \dots$ and by convention $T_0 = 0$. Let $X_i = c(T_i - T_{i-1}) - Y_i$. If we only consider the process at the claim times we can see that

$$C_{T_n} = u + \sum_{i=1}^n X_i$$

is a random walk. Note that $\psi(u) = \mathbb{P}[\inf_{n\in\mathbb{N}} C_{T_n} < 0]$. From the theory of random walks we can see that ruin occurs a.s. iff $\mathbb{E}[X_i] \le 0$, compare with Lemma 1.26 or [26, p.396]. Hence we will assume in the sequel that

$$\mathbb{E}[X_i] > 0 \iff c\frac{1}{\lambda} - \mu > 0 \iff c > \lambda\mu \iff \mathbb{E}[C_t - u] > 0\,.$$

Recall that

$$\mathbb{E}\Bigl[\sum_{i=1}^{N_t} Y_i\Bigr] = \lambda t \mu\,.$$

The condition can be interpreted that the mean income is strictly larger than the mean outflow. Therefore the condition is also called the **net profit condition**.

If the net profit condition is fulfilled then C_{T_n} tends to infinity as $n \to \infty$. Hence

$$\inf\{C_t - u : t > 0\} = \inf\{C_{T_n} - u : n \ge 1\}$$

is a.s. finite. So we can conclude that

$$\lim_{u\to\infty} \psi(u) = 0\,.$$

2.2. A note on the model and reality

In reality the mean number of claims in an interval will not be the same all the time. There will be a claim rate $\lambda(t)$ which may be periodic in time. Moreover, the number of individual contracts in the portfolio may vary with time. Let $a(t)$ be the volume of the portfolio at time t. Then the claim number process $\{N_t\}$ is an inhomogeneous Poisson process with rate $a(t)\lambda(t)$. Let

$$\Lambda(t) = \int_0^t a(s)\lambda(s)\,\mathrm{d}s$$

and $\Lambda^{-1}(t)$ be its inverse function. Then $\{\tilde N_t\}$ where $\tilde N_t = N_{\Lambda^{-1}(t)}$ is a Poisson process with rate 1. Let now the premium rate vary with t such that $c_t = ca(t)\lambda(t)$ for some constant c. This assumption is natural for changes in the risk volume.

It is artificial for the changes in the intensity. For instance we assume that the company gets more new customers at times with a higher intensity. This effect may arise because customers withdraw from their old insurance contracts and write new contracts with another company after claims occurred because they were not satisfied by the handling of claims by their old companies.

The premium income in the interval $(0, t]$ is $c\Lambda(t)$. Let $\tilde{C}_t = C_{\Lambda^{-1}(t)}$. Then

$$\tilde{C}_t = u + c\Lambda(\Lambda^{-1}(t)) - \sum_{i=1}^{N_{\Lambda^{-1}(t)}} Y_i = u + ct - \sum_{i=1}^{\tilde{N}_t} Y_i$$

is a Cramér-Lundberg process. Thus we should not consider time to be the real time but **operational time**.

The event of ruin does almost never occur in practice. If an insurance company observes that their surplus is decreasing they will immediately increase their premiums. On the other hand an insurance company is built up on different portfolios. Ruin in one portfolio does not mean bankruptcy. Therefore ruin is only a technical term. The probability of ruin is used for decision taking, for instance the premium calculation or the computation of reinsurance retention levels.

The surplus will also be a technical term in practice. If the business is going well then the share holders will decide to get a higher dividend. To model this we would have to assume a premium rate dependent on the surplus. But then it would be hard to obtain any useful results. For some references see for instance [28] and [7].

Data give evidence that a negative binomial model would better fit the number of claims in a certain time interval. This is mainly due that there are two parameters to fit in contrast to the Cramér-Lundberg model where only one parameter is present. However, the present model is the easiest to handle. Therefore it is very popular. For the purpose of this lecture it is the ideal model for a basic introduction to ruin theory. Understanding the Cramér-Lundberg model will help to understand more complicated models, as for instance *Cox risk models*. A further reason for using the model is that it is used to take decisions. As long as the decisions will lead to results close to the results obtained from "optimal" decisions the model will be of practical use. The reader, however, should always have in mind that the model is purely technical and does not describe the real cash-flows of an insurance company.

2.3. A differential equation for the ruin probability

We first prove that $\{C_t\}$ is a strong Markov process.

Lemma 2.1. *Let $\{C_t\}$ be a Cramér-Lundberg process and T be a finite stopping time. Then the stochastic process $\{C_{T+t} - C_T : t \ge 0\}$ is a Cramér-Lundberg process with initial capital 0 and independent of $\mathcal{F}_T$.*

Proof. We can write $\{C_{T+t} - C_T\}$ as

$$ct - \sum_{i=N_T+1}^{N_{T+t}} Y_i \,.$$

Because the claim amounts are iid. and independent of $\{N_t\}$ we only have to prove that $\{N_{T+t} - N_T\}$ is a Poisson process independent of $\mathcal{F}_T$. Because $\{N_t\}$ is a renewal process it is enough to show that $T_{N_T+1} - T$ is $\text{Exp}(\lambda)$ distributed and independent of $\mathcal{F}_T$. Condition on T_{N_T} and T. Then

$$\begin{aligned}&\mathbb{P}[T_{N_T+1} - T > x \mid T_{N_T}, T] \\ &\qquad = \mathbb{P}[T_{N_T+1} - T_{N_T} > x + T - T_{N_T} \mid T_{N_T}, T, T_{N_T+1} - T_{N_T} > T - T_{N_T}] = \mathrm{e}^{-\lambda x}\end{aligned}$$

by the lack of memory property of the exponential distribution. The assertion follows because T_{N_T+1} depends on $\mathcal{F}_T$ via T_{N_T} and T only. The latter, even though intuitively clear, follows readily noting that $\mathcal{F}_T$ is generated by sets of the form $\{N_T = n, A_n\}$ where $n \in \mathbb{N}$ and $A_n \in \sigma(Y_1, \ldots, Y_n, T_1, \ldots, T_n)$. $\square$

Let h be small. If ruin does not occur in the interval $(0, h]$ then a new Cramér-Lundberg process starts at time h with new initial capital C_h. Let $\delta(u) = 1 - \psi(u)$ denote the probability that ruin does not occur. Using the definition of the Poisson process we obtain

$$\begin{aligned}\delta(u) &= (1 - \lambda h + o(h))\delta(u + ch) + (\lambda h + o(h))(\mathbb{E}[\delta((u - Y_1)+)] + o(1)) + o(h) \\ &= (1 - \lambda h)\delta(u + ch) + \lambda h \int_0^u \delta((u - y)+) \,\mathrm{d}G(y) + o(h)\end{aligned}$$

where $\delta((u-y)+)$ exists because $\delta(u)$ is increasing. Letting h tending to 0 shows that $\delta(u)$ is right continuous, i.e. $\int_0^u \delta((u-y)+) \,\mathrm{d}G(y) = \int_0^u \delta(u-y) \,\mathrm{d}G(y)$. Rearranging the terms yields

$$c\,\frac{\delta(u + ch) - \delta(u)}{ch} = \lambda\Big[\delta(u + ch) - \int_0^u \delta(u - y) \,\mathrm{d}G(y)\Big] + o(1)\,.$$

Letting h tend to 0 shows that $\delta(u)$ is differentiable from the right. A similar argument shows the differentiability from the left. But

$$c\,\frac{\delta(u - ch) - \delta(u)}{-ch} = \lambda\Big[\delta(u) - \int_0^{u-} \delta(u - y) \,\mathrm{d}G(y)\Big] + o(1)\,.$$

Thus $\delta(u)$ is continuous everywhere and differentiable at all points where $G(u)$ is continuous, i.e. almost everywhere with respect to the Lebesgue measure. The derivative (in the sense of a density) is given by

$$c\delta'(u) = \lambda\Bigl[\delta(u) - \int_0^u \delta(u-y)\,\mathrm{d}G(y)\Bigr]\,. \tag{2.1}$$

The difficulty with equation (2.1) is that it contains both the derivative of δ and an integral. Let us try to get rid of the derivative.

$$\begin{aligned}
\frac{c}{\lambda}(\delta(u)-\delta(0)) &= \frac{1}{\lambda}\int_0^u c\delta'(x)\,\mathrm{d}x = \int_0^u \delta(x)\,\mathrm{d}x - \int_0^u\int_0^x \delta(x-y)\,\mathrm{d}G(y)\,\mathrm{d}x\\
&= \int_0^u \delta(x)\,\mathrm{d}x - \int_0^u\int_y^u \delta(x-y)\,\mathrm{d}x\,\mathrm{d}G(y)\\
&= \int_0^u \delta(x)\,\mathrm{d}x - \int_0^u\int_0^{u-y} \delta(x)\,\mathrm{d}x\,\mathrm{d}G(y)\\
&= \int_0^u \delta(x)\,\mathrm{d}x - \int_0^u\int_0^{u-x} \mathrm{d}G(y)\,\delta(x)\,\mathrm{d}x\\
&= \int_0^u \delta(x)(1-G(u-x))\,\mathrm{d}x = \int_0^u \delta(u-x)(1-G(x))\,\mathrm{d}x
\end{aligned}$$

Note that $\int_0^\infty (1-G(x))\,\mathrm{d}x = \mu$. Letting $u\to\infty$ we can by the bounded convergence theorem interchange limit and integral and obtain

$$c(1-\delta(0)) = \lambda\int_0^\infty (1-G(x))\,\mathrm{d}x = \lambda\mu$$

where we used that $\delta(u)\to 1$. It follows that

$$\delta(0) = 1-\frac{\lambda\mu}{c}, \qquad \psi(0) = \frac{\lambda\mu}{c}\,.$$

Replacing $\delta(u)$ by $1-\psi(u)$ yields

$$\begin{aligned}
c\psi(u) &= \lambda\mu - \lambda\int_0^u (1-\psi(u-x))(1-G(x))\,\mathrm{d}x\\
&= \lambda\Bigl(\int_u^\infty (1-G(x))\,\mathrm{d}x + \int_0^u \psi(u-x)(1-G(x))\,\mathrm{d}x\Bigr)\,.
\end{aligned} \tag{2.2}$$

In Section 2.8 we shall obtain a natural interpretation of (2.2).

Example 2.2. Let the claims be $\mathrm{Exp}(\alpha)$ distributed. Then equation (2.1) can be written as

$$c\delta'(u) = \lambda\Bigl[\delta(u) - \mathrm{e}^{-\alpha u}\int_0^u \delta(y)\alpha\mathrm{e}^{\alpha y}\,\mathrm{d}y\Bigr]\,.$$

Differentiating yields

$$c\delta''(u) = \lambda\Big[\delta'(u) + \alpha e^{-\alpha u}\int_0^u \delta(y)\alpha e^{\alpha y}\,dy - \alpha\delta(u)\Big] = \lambda\delta'(u) - \alpha c\delta'(u)\,.$$

The solution to this differential equation is

$$\delta(u) = A + Be^{-(\alpha-\lambda/c)u}\,.$$

Because $\delta(u) \to 1$ as $u \to \infty$ we get $A = 1$. Because $\delta(0) = 1 - \lambda/(\alpha c)$ the solution is

$$\delta(u) = 1 - \frac{\lambda}{\alpha c}\,e^{-(\alpha-\lambda/c)u} \qquad \text{or} \qquad \psi(u) = \frac{\lambda}{\alpha c}\,e^{-(\alpha-\lambda/c)u}\,.$$

■

2.4. The adjustment coefficient

Let

$$\theta(r) = \lambda(M_Y(r) - 1) - cr \tag{2.3}$$

provided $M_Y(r)$ exists. Then we find the following martingale.

Lemma 2.3. *Let $r \in \mathbb{R}$ such that $M_Y(r) < \infty$. Then the stochastic process*

$$\exp\{-rC_t - \theta(r)t\} \tag{2.4}$$

is a martingale.

Proof. By the Markov property

$$\begin{aligned}\mathbb{E}[e^{-rC_t-\theta(r)t} \mid \mathcal{F}_s] &= \mathbb{E}[e^{-r(C_t-C_s)}]e^{-rC_s-(\lambda(M_Y(r)-1)-cr)t}\\ &= \mathbb{E}[e^{r\sum_{N_s+1}^{N_t} Y_i}]e^{-rC_s-\lambda(M_Y(r)-1)t+crs}\\ &= e^{-rC_s-\lambda(M_Y(r)-1)s+crs} = e^{-rC_s-\theta(r)s}\,.\end{aligned}$$

□

It would be nice to have a martingale which is only dependent on the surplus and not explicitly on time. Let us therefore consider the equation $\theta(r) = 0$. This equation has obviously the solution $r = 0$. We differentiate $\theta(r)$.

$$\theta'(r) = \lambda M_Y'(r) - c\,, \qquad \theta''(r) = \lambda M_Y''(r) = \lambda\mathbb{E}[Y^2e^{rY}] > 0\,.$$

The function $\theta(r)$ is strictly convex. For $r = 0$

$$\theta'(0) = \lambda M_Y'(0) - c = \lambda\mu - c < 0$$

by the net profit condition. There might be at most one additional solution R to the equation $\theta(R) = 0$ and $R > 0$. If this solution exists we call it the **adjustment coefficient** or the **Lundberg exponent**. The Lundberg exponent will play an important rôle in the estimation of the ruin probabilities.

Example 2.2 (continued). For Exp(α) distributed claims we have to solve

$$\lambda\Bigl(\frac{\alpha}{\alpha - r} - 1\Bigr) - cr = \frac{\lambda r}{\alpha - r} - cr = 0\,.$$

For $r \neq 0$ we find $R = \alpha - \lambda/c$. Note that

$$\psi(u) = \frac{\lambda}{\alpha c}\,\mathrm{e}^{-Ru}\,.$$

■

In general the adjustment coefficient is difficult to calculate. So we try to find some bounds for the adjustment coefficient. Note that the second moment μ_2 exists if the Lundberg exponent exists. Consider the function $\theta(r)$ for $r > 0$.

$$\begin{aligned}
\theta''(r) &= \lambda \mathbb{E}[Y^2 \mathrm{e}^{rY}] > \lambda \mathbb{E}[Y^2] = \lambda\mu_2\,,\\
\theta'(r) &= \theta'(0) + \int_0^r \theta''(s)\,\mathrm{d}s > -(c - \lambda\mu) + \lambda\mu_2 r\,,\\
\theta(r) &= \theta(0) + \int_0^r \theta'(s)\,\mathrm{d}s > \lambda\mu_2\frac{r^2}{2} - (c - \lambda\mu)r\,.
\end{aligned}$$

The last inequality yields for $r = R$

$$0 = \theta(R) > R\Bigl(\lambda\mu_2\frac{R}{2} - (c - \lambda\mu)\Bigr)$$

from which an upper bound of R

$$R < \frac{2(c - \lambda\mu)}{\lambda\mu_2}$$

follows.

We are not able to find a lower bound in general. But in the case of bounded claims we get a lower bound for R. Assume that $Y_1 \leq M$ a.s..

Let us first consider the function

$$f(x) = \frac{x}{M}\left(e^{RM} - 1\right) - \left(e^{Rx} - 1\right) .$$

Its second derivative is

$$f''(x) = -R^2 e^{Rx} < 0 .$$

$f(x)$ is strictly concave with $f(0) = f(M) = 0$. Thus $f(x) > 0$ for $0 < x < M$. The function

$$h(x) = x e^x - e^x + 1$$

has for $x > 0$ a strictly positive derivative. Thus $h(x) > h(0) = 0$ for $x > 0$, in particular

$$\frac{1}{RM}\left(e^{RM} - 1\right) < e^{RM} .$$

We calculate

$$M_Y(R) - 1 = \int_0^M \left(e^{Rx} - 1\right) \, dG(x) \le \int_0^M \frac{x}{M}\left(e^{RM} - 1\right) \, dG(x) = \frac{\mu}{M}\left(e^{RM} - 1\right) .$$

From the equation determining R we get

$$0 = \lambda(M_Y(R) - 1) - cR \le \frac{\lambda\mu}{M}\left(e^{RM} - 1\right) - cR < \lambda\mu R e^{RM} - cR .$$

It follows readily

$$R > \frac{1}{M} \log \frac{c}{\lambda\mu} .$$

2.5. Lundberg's inequality

We will now connect the adjustment coefficient and ruin probabilities.

Theorem 2.4. *Assume that the adjustment coefficient R exists. Then*

$$\psi(u) < e^{-Ru} .$$

Proof. Assume that the theorem does not hold. Let

$$u_0 = \inf\{u \ge 0 : \psi(u) \ge e^{-Ru}\} .$$

Because $\psi(u)$ is continuous we get $\psi(u_0) = e^{-Ru_0}$. Because $\psi(0) < 1$ we conclude that $u_0 > 0$. Consider equation (2.2) for $u = u_0$.

$$\begin{aligned}
c\psi(u_0) &= \lambda\Big[\int_{u_0}^\infty (1 - G(x))\,dx + \int_0^{u_0} \psi(u_0 - x)(1 - G(x))\,dx\Big] \\
&< \lambda\Big[\int_{u_0}^\infty (1 - G(x))\,dx + \int_0^{u_0} e^{-R(u_0-x)}(1 - G(x))\,dx\Big] \\
&\le \lambda \int_0^\infty e^{-R(u_0-x)}(1 - G(x))\,dx = \lambda e^{-Ru_0} \int_0^\infty \int_x^\infty e^{Rx}\,dG(y)\,dx \\
&= \lambda e^{-Ru_0} \int_0^\infty \int_0^y e^{Rx}\,dx\,dG(y) = \lambda e^{-Ru_0} \int_0^\infty \frac{1}{R}(e^{Ry} - 1)\,dG(y) \\
&= \frac{\lambda}{R} e^{-Ru_0}(M_Y(R) - 1) = c e^{-Ru_0}
\end{aligned}$$

which is a contradiction. This proves the theorem. □

We now give an alternative proof of the theorem. We will use the positive martingale (2.4) for $r = R$. By the stopping theorem

$$e^{-Ru} = e^{-RC_0} = \mathbb{E}[e^{-RC_{\tau\wedge t}}] \ge \mathbb{E}[e^{-RC_{\tau\wedge t}}; \tau \le t] = \mathbb{E}[e^{-RC_\tau}; \tau \le t]\,.$$

Letting t tend to infinity yields by monotone convergence

$$e^{-Ru} \ge \mathbb{E}[e^{-RC_\tau}; \tau < \infty] > \mathbb{P}[\tau < \infty] = \psi(u) \tag{2.5}$$

because $C_\tau < 0$. We have got an upper bound for the ruin probability. We would like to know whether R is the best possible exponent in an exponential upper bound. This question will be answered in the next section.

Example 2.5. Let $G(x) = 1 - pe^{-\alpha x} - (1 - p)e^{-\beta x}$ where $0 < \alpha < \beta$ and $0 < p \le \beta(\beta - \alpha)^{-1}$. Note that the mean value is $\frac{p}{\alpha} + \frac{1-p}{\beta}$ and thus

$$c > \frac{\lambda p}{\alpha} + \frac{\lambda(1 - p)}{\beta}\,.$$

For $r < \alpha$

$$M_Y(r) = \int_0^\infty e^{rx} \left(\alpha p e^{-\alpha x} + \beta(1 - p)e^{-\beta x}\right)\,dx = \frac{\alpha p}{\alpha - r} + \frac{\beta(1 - p)}{\beta - r}\,.$$

Thus we have to solve

$$\lambda\Big(\frac{\alpha p}{\alpha - r} + \frac{\beta(1 - p)}{\beta - r} - 1\Big) - cr = \lambda\Big(\frac{pr}{\alpha - r} + \frac{(1 - p)r}{\beta - r}\Big) - cr = 0\,.$$

We find the obvious solution $r = 0$. If $r \neq 0$ then

$$\lambda p(\beta - r) + \lambda(1-p)(\alpha - r) = c(\alpha - r)(\beta - r)$$

or equivalently

$$cr^2 - ((\alpha+\beta)c - \lambda)r + \alpha\beta c - \lambda((1-p)\alpha + p\beta) = 0\,. \tag{2.6}$$

The solution is

$$\begin{aligned} r &= \frac{1}{2}\Bigl(\alpha + \beta - \frac{\lambda}{c} \pm \sqrt{\Bigl(\alpha+\beta-\frac{\lambda}{c}\Bigr)^2 - 4\Bigl(\alpha\beta - \frac{\lambda}{c}p\beta - \frac{\lambda}{c}(1-p)\alpha\Bigr)}\Bigr) \\ &= \frac{1}{2}\Bigl(\alpha + \beta - \frac{\lambda}{c} \pm \sqrt{\Bigl(\beta-\alpha-\frac{\lambda}{c}\Bigr)^2 + 4p\frac{\lambda}{c}(\beta - \alpha)}\Bigr)\,. \end{aligned}$$

Now we got three solutions. But there should only be two. Note that

$$\alpha\beta - \frac{\lambda}{c}p\beta - \frac{\lambda}{c}(1-p)\alpha = \frac{\alpha\beta}{c}\Bigl(c - \lambda\Bigl(\frac{p}{\alpha} + \frac{1-p}{\beta}\Bigr)\Bigr) > 0$$

and thus both solutions are positive. The larger of the solutions can be written as

$$\alpha + \frac{1}{2}\Bigl(\beta - \alpha - \frac{\lambda}{c} + \sqrt{\Bigl(\beta-\alpha-\frac{\lambda}{c}\Bigr)^2 + 4p\frac{\lambda}{c}(\beta - \alpha)}\Bigr)$$

and is thus larger than α. But the moment generating function does not exist for $r \geq \alpha$. Thus

$$R = \frac{1}{2}\Bigl(\alpha + \beta - \frac{\lambda}{c} - \sqrt{\Bigl(\beta-\alpha-\frac{\lambda}{c}\Bigr)^2 + 4p\frac{\lambda}{c}(\beta - \alpha)}\Bigr)\,.$$

From Lundberg's inequality it follows that

$$\psi(u) < \mathrm{e}^{-Ru}\,.$$

■

2.6. The Cramér-Lundberg approximation

Consider the equation (2.2). This equation looks almost like a renewal equation, but

$$\int_0^\infty \frac{\lambda}{c}(1 - G(x))\,\mathrm{d}x = \frac{\lambda\mu}{c} < 1$$

is not a probability distribution. Can we manipulate (2.2) to get a renewal equation? We try for some measurable functions h and $\bar{h}$

$$\psi(u)h(u) = h(u)\int_u^\infty \frac{\lambda}{c}(1-G(x))\,\mathrm{d}x + \int_0^u \psi(u-x)h(u-x)\bar{h}(x)\frac{\lambda}{c}(1-G(x))\,\mathrm{d}x \quad (2.7)$$

where $h(u) = h(u-x)\bar{h}(x)$. We can assume that $h(0) = 1$. Setting $x = u$ shows that $\bar{h}(u) = h(u)$. From a general theorem it follows that $h(u) = \mathrm{e}^{ru}$ where $r = \log h(1)$. In order that (2.7) is a renewal equation we need

$$1 = \int_0^\infty \mathrm{e}^{rx}\frac{\lambda}{c}(1-G(x))\,\mathrm{d}x = \frac{\lambda}{c}\int_0^\infty\int_x^\infty \mathrm{e}^{rx}\,\mathrm{d}G(y)\,\mathrm{d}x = \frac{\lambda(M_Y(r)-1)}{cr}\,.$$

The only solution to the latter equation is $r = R$. Let us assume that R exists. Moreover, we need that

$$\int_0^\infty x\mathrm{e}^{Rx}\frac{\lambda}{c}(1-G(x))\,\mathrm{d}x < \infty \quad (2.8)$$

which is equivalent to $M_Y'(R) < \infty$, see below. The equation is now

$$\psi(u)\,\mathrm{e}^{Ru} = \mathrm{e}^{Ru}\int_u^\infty \frac{\lambda}{c}(1-G(x))\,\mathrm{d}x + \int_0^u \psi(u-x)\,\mathrm{e}^{R(u-x)}\mathrm{e}^{Rx}\frac{\lambda}{c}(1-G(x))\,\mathrm{d}x\,.$$

It can be shown that

$$\mathrm{e}^{Ru}\int_u^\infty \frac{\lambda}{c}(1-G(x))\,\mathrm{d}x$$

is directly Riemann integrable. Thus we get from the renewal theorem

Theorem 2.6. *Assume that the Lundberg exponent exists and that (2.8) is fulfilled. Then*

$$\lim_{u\to\infty}\psi(u)\,\mathrm{e}^{Ru} = \frac{c-\lambda\mu}{\lambda M_Y'(R)-c}\,.$$

Proof. It only remains to compute the limit.

$$\begin{aligned}\int_0^\infty \mathrm{e}^{Ru}\int_u^\infty \frac{\lambda}{c}(1-G(x))\,\mathrm{d}x\,\mathrm{d}u &= \frac{\lambda}{c}\int_0^\infty\int_0^x \mathrm{e}^{Ru}\,\mathrm{d}u\,(1-G(x))\,\mathrm{d}x\\ &= \frac{\lambda}{cR}\int_0^\infty (\mathrm{e}^{Rx}-1)(1-G(x))\,\mathrm{d}x = \frac{1}{R}-\frac{\lambda\mu}{cR} = \frac{1}{cR}(c-\lambda\mu)\,.\end{aligned}$$

The mean value of the distribution is

$$\begin{aligned}\int_0^\infty x\mathrm{e}^{Rx}\frac{\lambda}{c}(1-G(x))\,\mathrm{d}x &= \frac{\lambda}{c}\int_0^\infty\int_x^\infty x\mathrm{e}^{Rx}\,\mathrm{d}G(y)\,\mathrm{d}x\\ &= \frac{\lambda}{c}\int_0^\infty\int_0^y x\mathrm{e}^{Rx}\,\mathrm{d}x\,\mathrm{d}G(y) = \frac{\lambda}{cR^2}\int_0^\infty (Ry\mathrm{e}^{Ry}-\mathrm{e}^{Ry}+1)\,\mathrm{d}G(y)\\ &= \frac{\lambda(RM_Y'(R)-M_Y(R)+1)}{cR^2} = \frac{\lambda RM_Y'(R)-cR}{cR^2} = \frac{\lambda M_Y'(R)-c}{cR}\,.\end{aligned}$$

The limit value follows readily. □

u	0	0.25	0.5	0.75	1
$\psi(u)$	0.6111	0.5246	0.4547	0.3969	0.3479
app(u)	0.5508	0.4879	0.4322	0.3828	0.3391
Er	-9.87	-6.99	-4.97	-3.54	-2.54
u	1.25	1.5	1.75	2	2.25
$\psi(u)$	0.3059	0.2696	0.2379	0.2102	0.1858
app(u)	0.3003	0.2660	0.2357	0.2087	0.1849
Er	-1.82	-1.32	-0.95	-0.69	-0.50

Table 2.1: *Cramér-Lundberg approximation to ruin probabilities*

Remark. The theorem shows that it is not possible to obtain an exponential upper bound for the ruin probability with an exponent strictly larger than R. ■

The theorem can be written in the following form

$$\psi(u) \sim \frac{c - \lambda\mu}{\lambda M_Y'(R) - c}\, \mathrm{e}^{-Ru}\,.$$

Thus for large u we get an approximation to $\psi(u)$. This approximation is called the **Cramér-Lundberg approximation**.

Example 2.2 (continued). From $M_Y(r)$ we get that

$$M_Y'(r) = \frac{\alpha}{(\alpha - r)^2}$$

and thus

$$\lim_{u\to\infty} \psi(u)\, \mathrm{e}^{Ru} = \frac{c - \frac{\lambda}{\alpha}}{\frac{\lambda\alpha}{(\alpha-R)^2} - c} = \frac{c - \frac{\lambda}{\alpha}}{\frac{\lambda\alpha}{\left(\frac{\lambda}{c}\right)^2} - c} = \frac{\lambda}{\alpha c}\frac{\alpha c - \lambda}{\alpha c - \lambda} = \frac{\lambda}{\alpha c}\,.$$

Hence the Cramér-Lundberg approximation

$$\psi(u) \sim \frac{\lambda}{\alpha c}\mathrm{e}^{-(\alpha - \lambda/c)u}$$

becomes exact in this case. ■

Example 2.7. Let $c = \lambda = 1$ and $G(x) = 1 - \frac{1}{3}(\mathrm{e}^{-x} + \mathrm{e}^{-2x} + \mathrm{e}^{-3x})$. The mean value of claim sizes is $\mu = 0.611111$, i.e. the net profit condition is fulfilled. One can show that

$$\psi(u) = 0.550790\, \mathrm{e}^{-0.485131u} + 0.0436979\, \mathrm{e}^{-1.72235u} + 0.0166231\, \mathrm{e}^{-2.79252u}\,.$$

From Theorem 2.6 it follows that $\mathrm{app}(u) = 0.550790\,\mathrm{e}^{-0.485131u}$ is the Cramér-Lundberg approximation to $\psi(u)$. Table 2.1 below shows the ruin function $\psi(u)$, its Cramér-Lundberg approximation $\mathrm{app}(u)$ and the relative error $(\mathrm{app}(u)-\psi(u))/\psi(u)$ multiplied by 100 (Er). Note that the the relative error is below 1% for $u \geq 1.71358 = 2.8\mu$. ■

2.7. Reinsurance and ruin

2.7.1. Proportional reinsurance

Recall that for proportional insurance the insurer covers $Y_i^{\mathrm{I}} = \alpha Y_i$ of each claim, the reinsurer covers $Y_i^{\mathrm{R}} = (1-\alpha)Y_i$. Denote by c^{I} the insurer's premium rate. The insurer's adjustment coefficient is obtained from the equation

$$\lambda(M_{\alpha Y}(r) - 1) - c^{\mathrm{I}} r = 0\,.$$

The new moment generating function is

$$M_{\alpha Y}(r) = \mathbb{E}[\mathrm{e}^{r\alpha Y_i}] = M_Y(\alpha r)\,.$$

Assume that both insurer and reinsurer use an expected value premium principle with the same safety loading. Then $c^{\mathrm{I}} = \alpha c$ and we have to solve

$$\lambda(M_Y(\alpha r) - 1) - c\alpha r = 0\,.$$

This is almost the original equation, hence $R = \alpha R^{\mathrm{I}}$ where R^{I} is the adjustment coefficient under reinsurance. The new adjustment coefficient is larger, hence the risk has become smaller.

Example 2.8. Let $Y_i \sim \mathrm{Exp}(\beta)$ and let the gross premium rate be fixed $(1+\xi)\lambda\mathbb{E}[Y] = (1+\xi)\frac{\lambda}{\beta}$. We want to reinsure the risk via proportional reinsurance with retention level α. The reinsurer charges the premium $(1+\vartheta)(1-\alpha)\frac{\lambda}{\beta}$. We assume that $\vartheta \geq \xi$, otherwise the insurer could choose $\alpha = 0$ and he would make a profit without any risk. How shall we choose α in order to maximise the insurer's adjustment coefficient? The case $\vartheta = \xi$ had been considered above. We only have to consider the case $\vartheta > \xi$. Because the net profit condition must be fulfilled there is a lower bound for α.

$$c^{\mathrm{I}} = (1+\xi)\frac{\lambda}{\beta} - (1+\vartheta)(1-\alpha)\frac{\lambda}{\beta} = (\alpha(1+\vartheta) - (\vartheta - \xi))\frac{\lambda}{\beta} > \alpha\frac{\lambda}{\beta}\,.$$

Thus

$$\alpha > \frac{\vartheta - \xi}{\vartheta} = 1 - \frac{\xi}{\vartheta} . \tag{2.9}$$

The moment generating function of αY is

$$M_{\alpha Y}(r) = M_Y(\alpha r) = \frac{\beta}{\beta - \alpha r} = \frac{\frac{\beta}{\alpha}}{\frac{\beta}{\alpha} - r} .$$

The claims after reinsurance are $\operatorname{Exp}(\beta/\alpha)$ distributed and the adjustment coefficient is

$$R^{\mathrm{I}}(\alpha) = \frac{\beta}{\alpha} - \frac{\lambda}{(\alpha(1+\vartheta) - (\vartheta - \xi))\frac{\lambda}{\beta}} = \beta\Big(\frac{1}{\alpha} - \frac{1}{\alpha(1+\vartheta) - (\vartheta - \xi)}\Big) .$$

Setting the derivative to 0 yields

$$\alpha^2 = \frac{(\alpha(1+\vartheta) - (\vartheta - \xi))^2}{1+\vartheta} = (1+\vartheta)\alpha^2 - 2\alpha(\vartheta - \xi) + \frac{(\vartheta - \xi)^2}{1+\vartheta} .$$

The solution to this equation is

$$\begin{aligned}
\alpha &= \frac{1}{\vartheta}\Bigg(\vartheta - \xi \pm \sqrt{(\vartheta-\xi)^2 - \frac{\vartheta}{1+\vartheta}(\vartheta - \xi)^2}\Bigg) \\
&= \Big(1 - \frac{\xi}{\vartheta}\Big)\Bigg(1 \pm \sqrt{1 - \frac{\vartheta}{1+\vartheta}}\Bigg) \\
&= \Big(1 - \frac{\xi}{\vartheta}\Big)\Bigg(1 + \sqrt{\frac{1}{1+\vartheta}}\Bigg)
\end{aligned}$$

where the last equality follows from (2.9). This solution is strictly larger than 1 if $\vartheta > \xi^2 + 2\xi$. Thus the solution is

$$\alpha_{\max} = \min\Bigg\{\Big(1 - \frac{\xi}{\vartheta}\Big)\Bigg(1 + \sqrt{\frac{1}{1+\vartheta}}\Bigg), 1\Bigg\} .$$

By plugging in the values $1 - \frac{\xi}{\vartheta}$ and 1 into the derivative $R^{\mathrm{I}\prime}(\alpha)$ shows that indeed $R^{\mathrm{I}}(\alpha)$ has a maximum at $\alpha_{\max}$.

In [61] it is proved that for any distribution function $G(x)$ such that the adjustment coefficient exists the function $R(\alpha)$ is unimodal. Therefore $R'(\alpha) = 0$ always yields the correct optimal reinsurance level, provided the solution α is smaller than one. ■

2.7.2. Excess of loss reinsurance

Under excess of loss reinsurance with retention level M the insurer has to pay $Y_i^{\mathrm{I}} = \min\{Y_i, M\}$ of each claim. The adjustment coefficient is the strictly positive solution to the equation

$$\lambda\Bigl(\int_0^M \mathrm{e}^{rx}\,\mathrm{d}G(x) + \mathrm{e}^{rM}(1 - G(M)) - 1\Bigr) - c^{\mathrm{I}} r = 0\,.$$

There is no possibility to find the solution from the problem without reinsurance. We have to solve the equation for every M separately. But note that R^{I} exists in any case. Especially for heavy tailed distributions this shows that the risk has become much smaller. By the Cramér-Lundberg approximation the ruin probability decreases exponentially as the initial capital increases.

We will now show that for the insurer the excess of loss reinsurance is optimal. We assume that both insurer and reinsurer use an expected value principle.

Proposition 2.9. *Let all premiums be computed via the expected value principle. Under all reinsurance forms acting on individual claims with premiums rate c^{I} and c^{R} fixed the excess of loss reinsurance maximises the insurer's adjustment coefficient.*

Proof. Let $h(x)$ be an increasing function with $0 \le h(x) \le x$ for $x \ge 0$. We assume that the insurer pays $Y_i^{\mathrm{I}} = h(Y_i)$. Let $h^*(x) = \min\{x, U\}$ be the excess of loss reinsurance. Because c^{I} is fixed U can be determined from

$$\int_0^U y\,\mathrm{d}G(y) + U(1 - G(U)) = \mathbb{E}[h^*(Y_i)] = \mathbb{E}[h(Y_i)] = \int_0^\infty h(y)\,\mathrm{d}G(y)\,.$$

Because $\mathrm{e}^z \ge 1 + z$ we obtain

$$\mathrm{e}^{r(h(y)-h^*(y))} \ge 1 + r(h(y) - h^*(y))$$

and

$$\mathrm{e}^{rh(y)} \ge \mathrm{e}^{rh^*(y)}(1 + r(h(y) - h^*(y)))\,.$$

Thus

$$\begin{aligned} M_{h(Y)}(r) &= \int_0^\infty \mathrm{e}^{rh(y)}\,\mathrm{d}G(y) \ge \int_0^\infty \mathrm{e}^{rh^*(y)}(1 + r(h(y) - h^*(y)))\,\mathrm{d}G(y) \\ &= M_{h^*(Y)}(r) + r\int_0^\infty (h(y) - h^*(y))\,\mathrm{e}^{rh^*(y)}\,\mathrm{d}G(y)\,. \end{aligned}$$

For $y \le U$ we have $h(y) \le y = h^*(y)$. We obtain for $r > 0$

$$\begin{aligned}
&\int_0^\infty (h(y) - h^*(y))\mathrm{e}^{rh^*(y)}\,\mathrm{d}G(y) \\
&\qquad = \int_0^U (h(y) - h^*(y))\mathrm{e}^{rh^*(y)}\,\mathrm{d}G(y) + \int_U^\infty (h(y) - h^*(y))\mathrm{e}^{rh^*(y)}\,\mathrm{d}G(y) \\
&\qquad \ge \int_0^U (h(y) - h^*(y))\mathrm{e}^{rU}\,\mathrm{d}G(y) + \int_U^\infty (h(y) - h^*(y))\mathrm{e}^{rU}\,\mathrm{d}G(y) \\
&\qquad = \mathrm{e}^{rU}\int_0^\infty (h(y) - h^*(y))\,\mathrm{d}G(y) = \mathrm{e}^{rU}(\mathbb{E}[h(Y)] - \mathbb{E}[h^*(Y)]) = 0\,.
\end{aligned}$$

It follows that $M_{h(Y)}(r) \ge M_{h^*(Y)}(r)$ for $r > 0$. Since

$$0 = \theta(R^{\mathrm{I}}) = \lambda(M_{h(Y)}(R^{\mathrm{I}}) - 1) - c^{\mathrm{I}}R^{\mathrm{I}} \ge \lambda(M_{h^*(Y)}(R^{\mathrm{I}}) - 1) - c^{\mathrm{I}}R^{\mathrm{I}} = \theta^*(R^{\mathrm{I}})$$

follows the assertion from the convexity of $\theta^*(r)$. $\square$

We consider now the portfolio of the reinsurer. What is the claim number process of the claims the reinsurer is involved in? Because the claim amounts are independent of the claim arrival process we delete independently points from the Poisson process with probability $G(M)$ and do not delete them with probability $1 - G(M)$. By Proposition 1.11 this process is a Poisson process with rate $\lambda(1 - G(M))$. Because the claim sizes are iid. and independent of the claim arrival process the surplus of the reinsurer is a Cramér-Lundberg process with intensity $\lambda(1 - G(M))$ and claim size distribution

$$\tilde{G}(x) = \mathbb{P}[Y_i - M \le x \mid Y_i > M] = \frac{G(M + x) - G(M)}{1 - G(M)}\,.$$

2.8. The severity of ruin and the distribution of $\inf\{C_t : t \ge 0\}$

For an insurance company ruin is not so dramatic if $-C_\tau$ is small, but it could ruin the whole company if $-C_\tau$ is very large. So we are interested in the distribution of $-C_\tau$ if ruin occurs. Let

$$\psi_x(u) = \mathbb{P}[\tau < \infty, C_\tau < -x]\,.$$

We proceed as in Section 2.3. For h small

$$\psi_x(u) = (1 - \lambda h)\psi_x(u + ch) + \lambda h\Bigl[\int_0^u \psi_x(u - y)\,\mathrm{d}G(y) + \int_{u+x}^\infty \mathrm{d}G(y)\Bigr] + o(h)\,.$$

Note that in a first step $\psi_x(u-y)$ should had been replaced by $\psi_x((u-y)+)$, but it follows that $\psi_x(u)$ is right continuous. We arrive at the integro-differential equation

$$c\psi_x'(u) = \lambda\Big[\psi_x(u) - \int_0^u \psi_x(u-y)\,\mathrm{d}G(y) - (1-G(u+x))\Big]\,.$$

Integration yields

$$\begin{aligned}\frac{c}{\lambda}(\psi_x(u)-\psi_x(0)) &= \int_0^u \psi_x(u-y)(1-G(y))\,\mathrm{d}y - \int_0^u (1-G(y+x))\,\mathrm{d}y \\ &= \int_0^u \psi_x(u-y)(1-G(y))\,\mathrm{d}y - \int_x^{x+u} (1-G(y))\,\mathrm{d}y\,.\end{aligned}$$

Because $\psi_x(u) \le \psi(u)$ we can see that $\psi_x(u) \to 0$ as $u \to \infty$. By the bounded convergence theorem we obtain

$$-\frac{c}{\lambda}\psi_x(0) = -\int_x^\infty (1-G(y))\,\mathrm{d}y$$

and

$$\mathbb{P}[C_\tau < -x \mid \tau < \infty, C_0 = 0] = \frac{\frac{\lambda}{c}\int_x^\infty (1-G(y))\,\mathrm{d}y}{\frac{\lambda\mu}{c}} = \frac{1}{\mu}\int_x^\infty (1-G(y))\,\mathrm{d}y\,. \quad (2.10)$$

The random variable $-C_\tau$ is called **severity of ruin**.

Remark. In order to get an equation for the ruin probability we can consider the first time point τ_1 where the surplus is below the initial capital. At this point, by the strong Markov property, a new Cramér-Lundberg process starts. We get three possibilities:

- The process gets never below the initial capital.
- $\tau_1 < \infty$ but $C_{\tau_1} \ge 0$.
- Ruin occurs at τ_1.

Thus we get

$$\psi(u) = \Big(1-\frac{\lambda\mu}{c}\Big)0 + \frac{\lambda}{c}\int_0^u \psi(u-y)(1-G(y))\,\mathrm{d}y + \frac{\lambda}{c}\int_u^\infty 1\,(1-G(y))\,\mathrm{d}y\,.$$

This is equation (2.2). Hence we have now a natural interpretation of (2.2). ■

Let $\tau_0 = 0$ and $\tau_i = \inf\{t > \tau_{i-1} : C_t < C_{\tau_{i-1}}\}$, called the **ladder times**, and define $L_i = C_{\tau_{i-1}} - C_{\tau_i}$, called the **ladder heights**. Note that L_i only is defined if $\tau_i < \infty$. By Lemma 2.1 we find $\mathbb{P}[\tau_i < \infty \mid \tau_{i-1} < \infty] = \lambda\mu/c$. Let $K = \sup\{i \in \mathbb{N} : \tau_i < \infty\}$ be the number of ladder epochs. We have just seen that $K \sim \text{NB}(1, 1 - \lambda\mu/c)$ and that, given K, the random variables $(L_i : i \le K)$ are iid and absolutely continuous with density $(1 - G(x))/\mu$. We only have to condition on K because L_i is not defined for $i > K$. If we assume that all $\{L_i : i \ge 1\}$ have the same distribution, then we can drop the conditioning on K and $\{L_i\}$ is independent of K. Then

$$\inf\{C_t : t \ge 0\} = u - \sum_{i=1}^{K} L_i$$

and

$$\mathbb{P}[\tau < \infty] = \mathbb{P}[\inf\{C_t : t \ge 0\} < 0] = P\Bigl[\sum_{i=1}^{K} L_i > u\Bigr] .$$

Denote by

$$B(x) = \frac{1}{\mu}\int_0^x (1 - G(y))\,\mathrm{d}y$$

the distribution function of L_i. We can use Panjer recursion to approximate $\psi(u)$ by using an appropriate discretisation.

A formula that is useful for theoretical considerations, but too messy to use for the computation of $\psi(u)$ is the Pollaczek-Khintchine formula.

$$\begin{aligned}\psi(u) = P\Bigl[\sum_{i=1}^{K} L_i > u\Bigr] &= \sum_{n=1}^{\infty} P\Bigl[\sum_{i=1}^{n} L_i > u\Bigr]\mathbb{P}[K = n] \\ &= \Bigl(1 - \frac{\lambda\mu}{c}\Bigr)\sum_{n=1}^{\infty}\Bigl(\frac{\lambda\mu}{c}\Bigr)^n (1 - B^{*n}(u)) . \end{aligned} \tag{2.11}$$

2.9. The Laplace transform of ψ

Definition 2.10. *Let f be a real function on $[0, \infty)$. The transform*

$$\hat{f}(s) := \int_0^\infty \mathrm{e}^{-sx} f(x)\,\mathrm{d}x \qquad\qquad (s \in \mathbb{R})$$

is called the **Laplace transform** *of f.*

If X is an absolutely continuous positive random variable with density f then $\hat{f}(s) = M_X(-s)$. The Laplace transform has the following properties.

Lemma 2.11.

i) *If* $f(x) \geq 0$ *a.e.*
$$\hat{f}(s_1) \leq \hat{f}(s_2) \quad \iff \quad s_1 \geq s_2.$$

ii)
$$\widehat{|f|}(s_1) < \infty \quad \implies \quad \widehat{|f|}(s) < \infty \text{ for all } s \geq s_1.$$

iii)
$$\widehat{f'}(s) = \int_0^\infty \mathrm{e}^{-sx} f'(x) \,\mathrm{d}x = s\hat{f}(s) - f(0)$$
provided $f'(x)$ *exists a.e. and* $\widehat{|f|}(s) < \infty$.

iv)
$$\lim_{s\to\infty} s\hat{f}(s) = \lim_{x\to 0} f(x)$$
provided $f'(x)$ *exists a.e.,* $\lim_{x\to 0} f(x)$ *exists and* $\widehat{|f|}(s) < \infty$ *for an* s *large enough.*

v)
$$\lim_{s\downarrow 0} s\hat{f}(s) = \lim_{x\to\infty} f(x)$$
provided $f'(x)$ *exists a.e.,* $\lim_{x\to\infty} f(x)$ *exists and* $\widehat{|f|}(s) < \infty$ *for all* $s > 0$.

vi)
$$\hat{f}(s) = \hat{g}(s) \text{ on } (s_0, s_1) \quad \implies \quad f(x) = g(x) \text{ Lebesgue a.e. } \quad \forall x \in [0,\infty).$$

□

We want to find the Laplace transform of $\delta(u)$. We multiply (2.1) with e^{-su} and then integrate over u. Let $s > 0$.

$$c\int_0^\infty \delta'(u)\mathrm{e}^{-su} \,\mathrm{d}u = \lambda \int_0^\infty \delta(u)\mathrm{e}^{-su} \,\mathrm{d}u - \lambda \int_0^\infty \int_0^u \delta(u-y) \,\mathrm{d}G(y)\mathrm{e}^{-su} \,\mathrm{d}u.$$

We have to determine the last integral.

$$\begin{aligned}\int_0^\infty \int_0^u \delta(u-y) \,\mathrm{d}G(y)\mathrm{e}^{-su} \,\mathrm{d}u &= \int_0^\infty \int_y^\infty \delta(u-y)\mathrm{e}^{-su} \,\mathrm{d}u \,\mathrm{d}G(y) \\ &= \int_0^\infty \int_0^\infty \delta(u)\mathrm{e}^{-s(u+y)} \,\mathrm{d}u \,\mathrm{d}G(y) = \hat{\delta}(s) M_Y(-s)\,.\end{aligned}$$

Thus we get the equation

$$c(s\hat{\delta}(s) - \delta(0)) = \lambda\hat{\delta}(s)(1 - M_Y(-s))$$

which has the solution

$$\hat{\delta}(s) = \frac{c\delta(0)}{cs - \lambda(1 - M_Y(-s))} = \frac{c - \lambda\mu}{cs - \lambda(1 - M_Y(-s))}\,.$$

The Laplace transform of ψ can easily be found as

$$\hat{\psi}(s) = \int_0^\infty (1 - \delta(u)) \mathrm{e}^{-su} \, \mathrm{d}u = \frac{1}{s} - \hat{\delta}(s) \, .$$

Example 2.2 (continued). For exponentially distributed claims we obtain

$$\hat{\delta}(s) = \frac{c - \lambda/\alpha}{cs - \lambda(1 - \alpha/(\alpha + s))} = \frac{c - \lambda/\alpha}{s(c - \lambda/(\alpha + s))} = \frac{(c - \lambda/\alpha)(\alpha + s)}{s(c(\alpha + s) - \lambda)} \, .$$

and

$$\begin{aligned}\hat{\psi}(s) &= \frac{1}{s} - \frac{(c - \lambda/\alpha)(\alpha + s)}{s(c(\alpha + s) - \lambda)} = \frac{c(\alpha + s) - \lambda - (c - \lambda/\alpha)(\alpha + s)}{s(c(\alpha + s) - \lambda)} \\ &= \frac{\lambda}{\alpha} \frac{1}{c(\alpha + s) - \lambda} = \frac{\lambda}{\alpha c} \frac{1}{\alpha - \lambda/c + s} \, .\end{aligned}$$

By comparison with the moment generating function of the exponential distribution we recognise that

$$\psi(u) = \frac{\lambda}{\alpha c} \, \mathrm{e}^{-(\alpha - \lambda/c)u} \, .$$

■

Example 2.5 (continued). For the Laplace transform of δ we get

$$\begin{aligned}\hat{\delta}(s) &= \frac{c - \lambda p/\alpha - \lambda(1 - p)/\beta}{cs - \lambda(1 - \alpha p/(\alpha + s) - \beta(1 - p)/(\beta + s))} \\ &= \frac{c - \lambda p/\alpha - \lambda(1 - p)/\beta}{s(c - \lambda p/(\alpha + s) - \lambda(1 - p)/(\beta + s))}\end{aligned}$$

and for the Laplace transform of ψ

$$\begin{aligned}\hat{\psi}(s) &= \frac{c - \lambda p/(\alpha + s) - \lambda(1 - p)/(\beta + s) - c + \lambda p/\alpha + \lambda(1 - p)/\beta}{s(c - \lambda p/(\alpha + s) - \lambda(1 - p)/(\beta + s))} \\ &= \lambda \frac{sp/(\alpha(\alpha + s)) + s(1 - p)/(\beta(\beta + s))}{s(c - \lambda p/(\alpha + s) - \lambda(1 - p)/(\beta + s))} \\ &= \lambda \frac{p(\beta + s)/\alpha + (1 - p)(\alpha + s)/\beta}{c(\alpha + s)(\beta + s) - \lambda p(\beta + s) - \lambda(1 - p)(\alpha + s)} \\ &= \lambda \frac{p(\beta + s)/\alpha + (1 - p)(\alpha + s)/\beta}{cs^2 + ((\alpha + \beta)c - \lambda)s + \alpha\beta c - \lambda((1 - p)\alpha + p\beta)} \, .\end{aligned}$$

The denominator can be written as $c(s + R)(s + \bar{R})$ where R and $\bar{R}$ are the two solutions to (2.6) with $R < \bar{R}$. The Laplace transform of ψ can be written in the form

$$\hat{\psi}(s) = \frac{A}{R + s} + \frac{B}{\bar{R} + s}$$

for some constants A and B. Hence

$$\psi(u) = A\mathrm{e}^{-Ru} + B\mathrm{e}^{-\bar{R}u}\,.$$

A must be the constant appearing in the Cramér-Lundberg approximation. We can see that in this case the Cramér-Lundberg approximation is not exact. ■

Recall that τ_1 is the first ladder epoch. On the set $\{\tau_1 < \infty\}$ the random variable $\sup\{u - C_t : t \ge 0\}$ is absolutely continuous with distribution function

$$\mathbb{P}[\sup\{u - C_t : t \ge 0\} \le x \mid \tau_1 < \infty] = 1 - \frac{c}{\lambda\mu}\psi(x)\,.$$

Let Z be a random variable with the above distribution. Its moment generating function is

$$\begin{aligned} M_Z(r) &= \int_0^\infty \mathrm{e}^{ru}\frac{c}{\lambda\mu}\delta'(u)\,\mathrm{d}u = \frac{c}{\lambda\mu}\Bigl(-r\frac{c-\lambda\mu}{-cr-\lambda(1-M_Y(r))} - \Bigl(1-\frac{\lambda\mu}{c}\Bigr)\Bigr) \\ &= 1 - \frac{c}{\lambda\mu} + \frac{c(c-\lambda\mu)}{\lambda\mu}\,\frac{r}{cr-\lambda(M_Y(r)-1)}\,. \end{aligned}$$

We will later need the first two moments of the above distribution function. Assume that $\mu_2 < \infty$. The first derivative of the moment generating function is

$$\begin{aligned} M_Z'(r) &= \frac{c(c-\lambda\mu)}{\lambda\mu}\,\frac{cr-\lambda(M_Y(r)-1)-r(c-\lambda M_Y'(r))}{(cr-\lambda(M_Y(r)-1))^2} \\ &= \frac{c(c-\lambda\mu)}{\mu}\,\frac{rM_Y'(r)-(M_Y(r)-1)}{(cr-\lambda(M_Y(r)-1))^2}\,. \end{aligned}$$

Note that

$$\lim_{r\to 0}\frac{1}{r}(cr-\lambda(M_Y(r)-1)) = c-\lambda\mu\,.$$

We find

$$\lim_{r\to 0}\frac{rM_Y'(r)-(M_Y(r)-1)}{r^2} = \lim_{r\to 0}\frac{M_Y'(r)+rM_Y''(r)-M_Y'(r)}{2r} = \frac{\mu_2}{2}$$

and thus

$$\mathbb{E}[Z] = \frac{c(c-\lambda\mu)}{\mu}\,\frac{\mu_2}{2(c-\lambda\mu)^2} = \frac{c\mu_2}{2\mu(c-\lambda\mu)}\,. \tag{2.12}$$

Assume now that $\mu_3 < \infty$. The second derivative of $M_Z(r)$ is

$$\begin{aligned} M_Z''(r) = \frac{c(c-\lambda\mu)}{\mu}\,\frac{1}{(cr-\lambda(M_Y(r)-1))^3}\Bigl(&rM_Y''(r)(cr-\lambda(M_Y(r)-1)) \\ &- 2(rM_Y'(r)-(M_Y(r)-1))(c-\lambda M_Y'(r))\Bigr)\,. \end{aligned}$$

For the limit to 0 we find

$$\begin{aligned}
&\lim_{r\to 0}\frac{1}{r^3}\Big(rM_Y''(r)(cr-\lambda(M_Y(r)-1)) \\
&\qquad\qquad -2(rM_Y'(r)-(M_Y(r)-1))(c-\lambda M_Y'(r))\Big) \\
&=\lim_{r\to 0}\frac{1}{3r^2}\Big((M_Y''(r)+rM_Y'''(r))(cr-\lambda(M_Y(r)-1)) \\
&\qquad +rM_Y''(r)(c-\lambda M_Y'(r))-2rM_Y''(r)(c-\lambda M_Y'(r)) \\
&\qquad +2(rM_Y'(r)-(M_Y(r)-1))\lambda M_Y''(r)\Big) \\
&=\frac{\mu_3}{3}(c-\lambda\mu)+\lambda\mu_2\lim_{r\to 0}\frac{rM_Y'(r)-(M_Y(r)-1)}{r^2} \\
&=\frac{\mu_3}{3}(c-\lambda\mu)+\frac{\lambda}{2}\mu_2^2\,.
\end{aligned}$$

Thus the second moment of Z becomes

$$\mathbb{E}[Z^2]=\frac{c(c-\lambda\mu)}{\mu}\,\frac{\mu_3(c-\lambda\mu)/3+\lambda\mu_2^2/2}{(c-\lambda\mu)^3}=\frac{c}{\mu}\Big(\frac{\mu_3}{3(c-\lambda\mu)}+\frac{\lambda\mu_2^2}{2(c-\lambda\mu)^2}\Big)\,. \tag{2.13}$$

2.10. Approximations to ψ

2.10.1. Diffusion approximations

Diffusion approximations are based on the following

Proposition 2.12. *Let $\{C_t^{(n)}\}$ be a sequence of Cramér-Lundberg processes with initial capital u, claim arrival intensities $\lambda^{(n)}=\lambda n$, claim size distributions $G^{(n)}(x)=G(x\sqrt{n})$ and premium rates*

$$c^{(n)}=\Big(1+\frac{c-\lambda\mu}{\lambda\mu\sqrt{n}}\Big)\lambda^{(n)}\mu^{(n)}=c+(\sqrt{n}-1)\lambda\mu\,.$$

Let $\mu=\int_0^\infty y\,\mathrm{d}G(y)$ and assume that $\mu_2=\int_0^\infty y^2\,\mathrm{d}G(y)<\infty$. Then

$$\{C_t^{(n)}\}\xrightarrow{\mathrm{d}}\{u+W_t\}$$

in distribution in the topology of uniform convergence on finite intervals where $\{W_t\}$ is a $(c-\lambda\mu,\lambda\mu_2)$-Brownian motion.

Proof. See [36] or [31]. □

Intuitively we let the number of claims in a unit time interval go to infinity and make the claim sizes smaller in such a way that the distribution of $C_1^{(n)} - u$ tends to a normal distribution and $\mathbb{E}[C_1^{(n)} - u] = c - \lambda\mu$. Let $\tau^{(n)}$ denote the ruin time of $(C_t^{(n)})$ and $\tau = \inf\{t \ge 0 : u + W_t < 0\}$ the ruin probability of the Brownian motion. Then

Proposition 2.13. *Let $\{C_t^{(n)}\}$ and $\{W_t\}$ be as above. Then*

$$\lim_{n\to\infty} \mathbb{P}[\tau^{(n)} \le t] = \mathbb{P}[\tau \le t]$$

and

$$\lim_{n\to\infty} \mathbb{P}[\tau^{(n)} < \infty] = \mathbb{P}[\tau < \infty]\,.$$

Proof. The result for a finite time horizon is a special case of [62, Thm.9], see also [36] or [31]. The result for the infinite time horizon can be found in [48]. □

The idea of the diffusion approximation is to approximate $\mathbb{P}[\tau^{(1)} \le t]$ by $\mathbb{P}[\tau \le t]$ and $\mathbb{P}[\tau^{(1)} < \infty]$ by $\mathbb{P}[\tau < \infty]$. Thus we need the ruin probabilities of the Brownian motion.

Lemma 2.14. *Let $\{W_t\}$ be a (m, η^2)-Brownian motion with $m > 0$ and $\tau = \inf\{t \ge 0 : u + W_t < 0\}$. Then*

$$\mathbb{P}[\tau < \infty] = \mathrm{e}^{-2um/\eta^2}$$

and

$$\mathbb{P}[\tau \le t] = 1 - \Phi\Big(\frac{mt + u}{\eta\sqrt{t}}\Big) + \mathrm{e}^{-2um/\eta^2}\Phi\Big(\frac{mt - u}{\eta\sqrt{t}}\Big)\,.$$

Proof. By Lemma 1.25 the process

$$\Big\{\exp\Big\{-\frac{2m(u + W_t)}{\eta^2}\Big\} : t \ge 0\Big\}$$

is a martingale. By the stopping theorem

$$\Big\{\exp\Big\{-\frac{2m(u + W_{\tau\wedge t})}{\eta^2}\Big\} : t \ge 0\Big\}$$

is a positive bounded martingale. Thus, because $\lim_{t\to\infty} W_t = \infty$,

$$\exp\Big\{-\frac{2um}{\eta^2}\Big\} = \mathbb{E}\Big[\exp\Big\{-\frac{2m(u + W_\tau)}{\eta^2}\Big\}\Big] = \mathbb{P}[\tau < \infty]\,.$$

u	0	0.25	0.5	0.75	1
$\psi(u)$	0.6111	0.5246	0.4547	0.3969	0.3479
DA	1.0000	0.8071	0.6514	0.5258	0.4244
Er	63.64	53.87	43.26	32.49	21.98
u	1.25	1.5	1.75	2	2.25
$\psi(u)$	0.3059	0.2696	0.2379	0.2102	0.1858
DA	0.3425	0.2765	0.2231	0.1801	0.1454
Er	11.96	2.54	-6.22	-14.32	-21.78

Table 2.2: *Diffusion approximation to ruin probabilities*

It is easy to see that $sW_{1/s} - m$ is a $(0, \eta^2)$-Brownian motion (see [37, p.351]) and thus $s(u + W_{1/s}) - m$ is a (u, η^2)-Brownian motion. Denote the latter process by $\{\tilde{W}_s\}$. Then

$$\begin{aligned}
\mathbb{P}[\tau \le t] &= \mathbb{P}[\inf\{u + W_s : 0 < s \le t\} < 0] \\
&= \mathbb{P}[\inf\{s(u + W_{1/s}) : s \ge 1/t\} < 0] \\
&= \mathbb{E}[\mathbb{P}[\inf\{m + \tilde{W}_s : s \ge 1/t\} < 0 \mid \tilde{W}_{1/t}]] \\
&= \int_{-\infty}^{-m} \frac{1}{\sqrt{2\pi\eta^2/t}} \mathrm{e}^{-\frac{(y-u/t)^2}{2\eta^2/t}} \,\mathrm{d}y + \int_{-m}^{\infty} \mathrm{e}^{-\frac{2u(y+m)}{\eta^2}} \frac{1}{\sqrt{2\pi\eta^2/t}} \mathrm{e}^{-\frac{(y-u/t)^2}{2\eta^2/t}} \,\mathrm{d}y \\
&= \Phi\Bigl(\frac{-m - u/t}{\eta/\sqrt{t}}\Bigr) + \mathrm{e}^{-\frac{2um}{\eta^2}} \int_{-m}^{\infty} \frac{1}{\sqrt{2\pi\eta^2/t}} \mathrm{e}^{-\frac{(y+u/t)^2}{2\eta^2/t}} \,\mathrm{d}y \\
&= 1 - \Phi\Bigl(\frac{mt + u}{\eta\sqrt{t}}\Bigr) + \mathrm{e}^{-\frac{2um}{\eta^2}} \Phi\Bigl(\frac{mt - u}{\eta\sqrt{t}}\Bigr).
\end{aligned}$$

□

Diffusion approximations only work well if $c/(\lambda\mu)$ is near 1. There also exist corrected diffusion approximations which work much better, see [54] or [6].

Example 2.7 (continued). Let $c = \lambda = 1$ and $G(x) = 1 - \frac{1}{3}(\mathrm{e}^{-x} + \mathrm{e}^{-2x} + \mathrm{e}^{-3x})$. We find $c - \lambda\mu = 7/18$ and $\lambda\mu_2 = 49/54$. This leads to the diffusion approximation $\psi(u) \approx \exp\{-6u/7\}$. Table 2.2 shows exact values ($\psi(u)$), the diffusion approximation (DA) and the relative error multiplied by 100 (Er). Here we have $c/(\lambda\mu) = 18/11 = 1.63636$ is not near one. This is also indicated by the figures. ■

2.10.2. The deVylder approximation

In the case of exponentially distributed claim amounts we know the ruin probabilities explicitly. The idea of the deVylder approximation is to replace $\{C_t\}$ by $\{\tilde{C}_t\}$ where $\{\tilde{C}_t\}$ has exponentially distributed claim amounts and

$$\mathbb{E}[(C_t - u)^k] = \mathbb{E}[(\tilde{C}_t - u)^k] \qquad \text{for } k = 1, 2, 3.$$

The first three (centralised) moments are

$$\mathbb{E}[C_t - u] = (c - \lambda\mu)t = \Big(\tilde{c} - \frac{\tilde{\lambda}}{\tilde{\alpha}}\Big)t\,,$$

$$\operatorname{Var}[C_t] = \operatorname{Var}[u + ct - C_t] = \lambda\mu_2 t = \frac{2\tilde{\lambda}}{\tilde{\alpha}^2}t$$

and

$$\mathbb{E}[(C_t - \mathbb{E}[C_t])^3] = -\mathbb{E}[(u + ct - C_t - \mathbb{E}[u + ct - C_t])^3] = -\lambda\mu_3 t = -\frac{6\tilde{\lambda}}{\tilde{\alpha}^3}t\,.$$

The parameters of the approximation are

$$\tilde{\alpha} = \frac{3\mu_2}{\mu_3}\,, \qquad \tilde{\lambda} = \frac{\lambda\mu_2\tilde{\alpha}^2}{2} = \frac{9\mu_2^3}{2\mu_3^2}\lambda$$

and

$$\tilde{c} = c - \lambda\mu + \frac{\tilde{\lambda}}{\tilde{\alpha}} = c - \lambda\mu + \frac{3\mu_2^2}{2\mu_3}\lambda\,.$$

Thus the approximation to the probability of ultimate ruin is

$$\psi(u) \approx \frac{\tilde{\lambda}}{\tilde{\alpha}\tilde{c}}\mathrm{e}^{-\left(\tilde{\alpha} - \frac{\tilde{\lambda}}{\tilde{c}}\right)u}\,.$$

There is also a formula for the probability of ruin within finite time. Let $\eta = \sqrt{\tilde{\lambda}/(\tilde{\alpha}\tilde{c})}$. Then

$$\psi(u,t) \approx \frac{\tilde{\lambda}}{\tilde{\alpha}\tilde{c}}\mathrm{e}^{-(\tilde{\alpha} - \tilde{\lambda}/\tilde{c})u} - \frac{1}{\pi}\int_0^\pi f(x)\,\mathrm{d}x \tag{2.14}$$

where

$$f(x) = \eta\frac{\exp\{2\eta\tilde{\alpha}\tilde{c}t\,\cos x - (\tilde{\alpha}\tilde{c} + \tilde{\lambda})t + \tilde{\alpha}u(\eta\cos x - 1)\}}{1 + \eta^2 - 2\eta\cos x} \times (\cos(\tilde{\alpha}u\eta\sin x) - \cos(\tilde{\alpha}u\eta\sin x + 2x))\,.$$

Numerical investigations show that the approximation is quite accurate.

Example 2.7 (continued). In addition to the previously calculated values we also need $\mu_3 = 251/108$. The approximation parameters are $\tilde{\alpha} = 1.17131$, $\tilde{\lambda} = 0.622472$ and $\tilde{c} = 0.920319$. This leads to the approximation $\psi(u) \approx 0.577441\mathrm{e}^{-0.494949u}$. It turns out that the approximation works well. ∎

u	0	0.25	0.5	0.75	1
$\psi(u)$	0.6111	0.5246	0.4547	0.3969	0.3479
DV	0.5774	0.5102	0.4509	0.3984	0.3520
Er	-5.51	-2.73	-0.86	0.38	1.18
u	1.25	1.5	1.75	2	2.25
$\psi(u)$	0.3059	0.2696	0.2379	0.2102	0.1858
DV	0.3110	0.2748	0.2429	0.2146	0.1896
Er	1.67	1.95	2.07	2.09	2.03

Table 2.3: *DeVylder approximation to ruin probabilities*

2.10.3. The Beekman-Bowers approximation

Recall from (2.12) and (2.13) that

$$F(u) = 1 - \frac{c}{\lambda\mu}\psi(u)$$

is a distribution function and that

$$\int_0^\infty z \,\mathrm{d}F(z) = \frac{c\mu_2}{2\mu(c-\lambda\mu)}$$

and that

$$\int_0^\infty z^2 \,\mathrm{d}F(z) = \frac{c}{\mu}\Big(\frac{\mu_3}{3(c-\lambda\mu)} + \frac{\lambda\mu_2^2}{2(c-\lambda\mu)^2}\Big)\,.$$

The idea is to approximate the distribution function F by the distribution function $\tilde{F}(u)$ of a $\Gamma(\gamma,\alpha)$ distributed random variable such that the first two moments coincide. Thus the parameters γ and α have to fulfil

$$\begin{aligned}\frac{\gamma}{\alpha} &= \frac{c\mu_2}{2\mu(c-\lambda\mu)}\,,\\ \frac{\gamma(\gamma+1)}{\alpha^2} &= \frac{c}{\mu}\Big(\frac{\mu_3}{3(c-\lambda\mu)} + \frac{\lambda\mu_2^2}{2(c-\lambda\mu)^2}\Big)\,.\end{aligned}$$

The Beekman-Bowers approximation to the ruin probability is

$$\psi(u) = \frac{\lambda\mu}{c}(1-F(u)) \approx \frac{\lambda\mu}{c}(1-\tilde{F}(u))\,.$$

Remark. If $2\gamma \in \mathbb{N}$ then $2\alpha Z \sim \chi^2_{2\gamma}$ is χ^2 distributed. ■

u	0	0.25	0.5	0.75	1
$\psi(u)$	0.6111	0.5246	0.4547	0.3969	0.3479
BB1	0.6111	0.5227	0.4553	0.3985	0.3498
Er	0.00	-0.35	0.12	0.42	0.54
BB2	0.6111	0.5105	0.4456	0.3914	0.3450
Er	0.00	-2.68	-2.02	-1.38	-0.83
u	1.25	1.5	1.75	2	2.25
$\psi(u)$	0.3059	0.2696	0.2379	0.2102	0.1858
BB1	0.3076	0.2709	0.2387	0.2106	0.1859
Er	0.54	0.47	0.34	0.19	0.04
BB2	0.3046	0.2693	0.2383	0.2110	0.1869
Er	-0.42	-0.11	0.18	0.40	0.59

Table 2.4: *Beekman-Bowers approximation to ruin probabilities*

Example 2.7 (continued). For the Beekman-Bowers approximation we have to solve the equations

$$\frac{\gamma}{\alpha} = 1.90909\,, \qquad \frac{\gamma(\gamma+1)}{\alpha^2} = 7.71429\,,$$

which yields the parameters $\gamma = 0.895561$ and $\alpha = 0.469104$. From this the Beekman Bowers approximation can be obtained. Here we have $2\gamma = 1.79112$ which is not close to an integer. Anyway, one can interpolate between the χ_1^2 and the χ_2^2 distribution function to get the approximation

$$0.20888\chi_1^2(2\alpha u) + 0.79112\chi_2^2(2\alpha u)$$

to $1 - c/(\lambda\mu)\,\psi(u)$. Table 2.4 shows the exact values ($\psi(u)$), the Beekman-Bowers approximation (BB1) and the approximation obtained by interpolating the χ^2 distributions (BB2). The relative errors (Er) are given in percent. One can clearly see that both the approximations work well. ∎

2.11. Subexponential claim size distributions

Let us now consider subexponential claim size distributions. In this case the Lundberg exponent does not exist (Lemma 1.30).

Theorem 2.15. *Assume that the ladder height distribution*

$$\frac{1}{\mu}\int_0^x (1-G(y))\,\mathrm{d}y$$

is subexponential. Then

$$\lim_{u\to\infty}\frac{\psi(u)}{\int_u^\infty(1-G(y))\,\mathrm{d}y}=\frac{\lambda}{c-\lambda\mu}\,.$$

Remark. Recall that the probability that ruin occurs at the first ladder time given there is a first ladder epoch is

$$\frac{1}{\mu}\int_u^\infty (1-G(y))\,\mathrm{d}y\,.$$

Hence the ruin probability is asymptotically $(c-\lambda\mu)^{-1}\lambda\mu$ times the probability of ruin at the first ladder time given there is a first ladder epoch. But $(c-\lambda\mu)^{-1}\lambda\mu$ is the expected number of ladder times. Intuitively for u large ruin will occur if one of the ladder heights exceeds u. ■

Proof. Let $B(x)$ denote the distribution function of the first ladder height L_1, i.e.

$$B(x)=\frac{1}{\mu}\int_0^x (1-G(y))\,\mathrm{d}y\,.$$

Choose $\varepsilon>0$ such that $\lambda\mu(1+\varepsilon)<c$. By Lemma 1.33 there exists D such that

$$\frac{1-B^{*n}(x)}{1-B(x)}\le D(1+\varepsilon)^n\,.$$

From the Pollaczek-Khintchine formula (2.11) we obtain

$$\begin{aligned}\frac{\psi(u)}{1-B(u)}&=\Bigl(1-\frac{\lambda\mu}{c}\Bigr)\sum_{n=1}^\infty\Bigl(\frac{\lambda\mu}{c}\Bigr)^n\frac{1-B^{*n}(u)}{1-B(u)}\\&\le D\Bigl(1-\frac{\lambda\mu}{c}\Bigr)\sum_{n=1}^\infty\Bigl(\frac{\lambda\mu}{c}\Bigr)^n(1+\varepsilon)^n<\infty\,.\end{aligned}$$

Thus we can interchange sum and limit. Recall from Lemma 1.34 that

$$\lim_{u\to\infty}\frac{1-B^{*n}(u)}{1-B(u)}=n\,.$$

Thus

$$\begin{aligned}\lim_{u\to\infty}\frac{\psi(u)}{1-B(u)}&=\Bigl(1-\frac{\lambda\mu}{c}\Bigr)\sum_{n=1}^\infty n\Bigl(\frac{\lambda\mu}{c}\Bigr)^n=\Bigl(1-\frac{\lambda\mu}{c}\Bigr)\sum_{n=1}^\infty\sum_{m=1}^n\Bigl(\frac{\lambda\mu}{c}\Bigr)^n\\&=\Bigl(1-\frac{\lambda\mu}{c}\Bigr)\sum_{m=1}^\infty\sum_{n=m}^\infty\Bigl(\frac{\lambda\mu}{c}\Bigr)^n=\sum_{m=1}^\infty\Bigl(\frac{\lambda\mu}{c}\Bigr)^m=\frac{\lambda\mu/c}{1-\lambda\mu/c}=\frac{\lambda\mu}{c-\lambda\mu}\,.\end{aligned}$$

□

u	$\psi(u)$	App	Er
1	0.364	$8.79 \cdot 10^{-3}$	-97.588
2	0.150	$1.52 \cdot 10^{-4}$	-99.898
3	$6.18 \cdot 10^{-2}$	$8.58 \cdot 10^{-6}$	-99.986
4	$2.55 \cdot 10^{-2}$	$9.22 \cdot 10^{-7}$	-99.996
5	$1.05 \cdot 10^{-2}$	$1.49 \cdot 10^{-7}$	-99.999
10	$1.24 \cdot 10^{-4}$	$3.47 \cdot 10^{-10}$	-100
20	$1.75 \cdot 10^{-8}$	$5.40 \cdot 10^{-13}$	-99.997
30	$2.50 \cdot 10^{-12}$	$1.10 \cdot 10^{-14}$	-99.56
40	$1.60 \cdot 10^{-15}$	$6.71 \cdot 10^{-16}$	-58.17
50	$1.21 \cdot 10^{-16}$	$7.56 \cdot 10^{-17}$	-37.69

Table 2.5: *Approximations for subexponential claim sizes*

Example 2.16. Let $G \sim \mathrm{Pa}(\alpha, \beta)$. Then (see Example 1.32)

$$\lim_{x \to \infty} \frac{1 - B(zx)}{1 - B(x)} = z \lim_{x \to \infty} \frac{1 - G(zx)}{1 - G(x)} = z^{-(\alpha - 1)} .$$

By Lemma 1.31 B is a subexponential distribution. Note that we assume that $\mu < \infty$ and thus $\alpha > 1$. Because

$$\int_x^\infty \Bigl(\frac{\beta}{\beta + y}\Bigr)^\alpha \mathrm{d}y = \frac{\beta}{\alpha - 1}\Bigl(\frac{\beta}{\beta + x}\Bigr)^{\alpha - 1}$$

we obtain

$$\psi(u) \approx \frac{\lambda\beta/(\alpha - 1)}{c - \lambda\beta/(\alpha - 1)}\Bigl(\frac{\beta}{\beta + u}\Bigr)^{\alpha - 1} = \frac{\lambda\beta}{c(\alpha - 1) - \lambda\beta}\Bigl(\frac{\beta}{\beta + u}\Bigr)^{\alpha - 1} .$$

Choose now $c = 1$, $\lambda = 9$, $\alpha = 11$ and $\beta = 1$. Table 2.5 gives the exact value ($\psi(u)$), the approximation (App) and the relative error in percent (Er). Consider for instance $u = 20$. The ruin probability is so small, that it is not interesting for practical purposes anymore. But the approximation still underestimates the true value by almost 100%. That means we are still not far out enough in the tail. This is the problem by using the approximation. It should, however, be remarked, that for small values of α the approximation works much better. Values $\alpha \in (1, 2)$ are also more interesting from a practical point of view. ∎

Remark. The conditions of the theorem are also fulfilled for $\mathrm{LN}(\mu, \sigma^2)$, for $\mathrm{LG}(\gamma, \alpha)$ and for $\mathrm{Wei}(\alpha, c)$ ($\alpha < 1$) distributed claims, see [22] and [39]. ∎

2.12. The time of ruin

Consider the function

$$f_\alpha(u) = \mathbb{E}[\mathrm{e}^{-\alpha\tau} 1\!\!1_{\{\tau<\infty\}} \mid C_0 = u]\,.$$

The function is defined at least for $\alpha \ge 0$. We will first find a differential equation for $f_\alpha(u)$.

Lemma 2.17. *The function $f_\alpha(u)$ is absolutely continuous and fulfils the equation*

$$cf_\alpha'(u) + \lambda\Bigl[\int_0^u f_\alpha(u-y)\,\mathrm{d}G(y) + 1 - G(u) - f_\alpha(u)\Bigr] - \alpha f_\alpha(u) = 0\,. \tag{2.15}$$

Proof. For h small we get

$$\begin{aligned} f_\alpha(u) &= (1-\lambda h + o(h))\mathrm{e}^{-\alpha h} f_\alpha(u+ch) \\ &\qquad +(\lambda h + o(h))\Bigl[\int_0^u f_\alpha(u-y)\,\mathrm{d}G(y) + (1-G(u)) + o(1)\Bigr] + o(h)\,. \end{aligned}$$

We see that $f_\alpha(u)$ is right continuous. Reordering of the terms yields

$$\begin{aligned} &c\,\frac{f_\alpha(u+ch)-f_\alpha(u)}{ch} - \frac{1-\mathrm{e}^{-\alpha h}}{h} f_\alpha(u+ch) \\ &\qquad + \lambda\Bigl[\int_0^u f_\alpha(u-y)\,\mathrm{d}G(y) + 1 - G(u) - \mathrm{e}^{-\alpha h} f_\alpha(u+ch)\Bigr] + o(1) = 0\,. \end{aligned}$$

Letting $h \to 0$ shows that $f_\alpha(u)$ is differentiable and (2.15) for the derivative from the right. A similar argument can be used from the left. □

This differential equation is hard to solve. Let us take the Laplace transform with respect to the initial capital. Let $\hat f_\alpha(s) = \int_0^\infty \mathrm{e}^{-su} f_\alpha(u)\,\mathrm{d}u$. For the moment we assume $s > 0$. Note that (see Section 2.9)

$$\int_0^\infty f_\alpha'(u)\mathrm{e}^{-su}\,\mathrm{d}u = s\hat f_\alpha(s) - f_\alpha(0)\ ,$$

$$\int_0^\infty \int_0^u f_\alpha(u-y)\,\mathrm{d}G(y)\mathrm{e}^{-su}\,\mathrm{d}u = \hat f_\alpha(s) M_Y(-s)$$

and

$$\int_0^\infty \int_u^\infty \mathrm{d}G(y)\mathrm{e}^{-su}\,\mathrm{d}u = \int_0^\infty \int_0^y \mathrm{e}^{-su}\,\mathrm{d}u\,\mathrm{d}G(y) = \frac{1-M_Y(-s)}{s}\,.$$

Multiplying (2.15) by e^{-su} and integrating yields

$$c(s\hat{f}_\alpha(s) - f_\alpha(0)) + \lambda\Big[\hat{f}_\alpha(s)M_Y(-s) + \frac{1-M_Y(-s)}{s} - \hat{f}_\alpha(s)\Big] - \alpha\hat{f}_\alpha(s) = 0\,.$$

Solving for $\hat{f}_\alpha(s)$ yields

$$\hat{f}_\alpha(s) = \frac{cf_\alpha(0) - \lambda s^{-1}(1-M_Y(-s))}{cs - \lambda(1-M_Y(-s)) - \alpha}\,. \tag{2.16}$$

We know that $\hat{f}_\alpha(s)$ exists if $\alpha > 0$ and $s > 0$ and is positive. The denominator

$$cs - \lambda(1-M_Y(-s)) - \alpha$$

is convex, has value $-\alpha < 0$ at 0 and converges to ∞ as $s \to \infty$. Thus there exists a strictly positive root $s(\alpha)$ of the denominator. Because $\hat{f}_\alpha(s)$ exists also for $s = s(\alpha)$ the numerator must have a root at $s(\alpha)$ too. Thus

$$cf_\alpha(0) = \lambda s(\alpha)^{-1}(1 - M_Y(-s(\alpha)))\,.$$

The function $s(\alpha)$ is differentiable by the implicit function theorem

$$s'(\alpha)(c - \lambda M_Y'(-s(\alpha))) - 1 = 0\,. \tag{2.17}$$

Because $s(0) = 0$ we obtain that $\lim_{\alpha\to 0} s(\alpha) = 0$.

Example 2.18. Let the claims be $\mathrm{Exp}(\beta)$ distributed. We have to solve

$$cs - \frac{\lambda s}{\beta + s} - \alpha = 0$$

which admits the two solutions

$$s_\pm(\alpha) = \frac{-(\beta c - \lambda - \alpha) \pm \sqrt{(\beta c - \lambda - \alpha)^2 + 4\alpha\beta c}}{2c}$$

where $s_-(\alpha) < 0 \le s_+(\alpha)$. Thus

$$\hat{f}_\alpha(s) = \frac{\lambda/(\beta + s_+(\alpha)) - \lambda/(\beta + s)}{cs - \lambda s/(\beta+s) - \alpha} = \frac{\lambda}{c(\beta + s_+(\alpha))(s - s_-(\alpha))}\,.$$

It follows that

$$f_\alpha(u) = \frac{\lambda}{c(\beta + s_+(\alpha))}\mathrm{e}^{s_-(\alpha)\,u}\,.$$

■

Noting that

$$\mathbb{E}[\tau 1\!\!1_{\{\tau<\infty\}}] = \lim_{\alpha\downarrow 0}\mathbb{E}[\tau \mathrm{e}^{-\alpha\tau} 1\!\!1_{\{\tau<\infty\}}] = \lim_{\alpha\downarrow 0} -\frac{\mathrm{d}}{\mathrm{d}\alpha}\mathbb{E}[\mathrm{e}^{-\alpha\tau} 1\!\!1_{\{\tau<\infty\}}]$$

we see

$$\int_0^\infty \mathbb{E}[\tau 1\!\!1_{\{\tau<\infty\}} \mid C_0 = u]\mathrm{e}^{-su}\,\mathrm{d}u = \lim_{\alpha\downarrow 0} -\frac{\mathrm{d}}{\mathrm{d}\alpha}\int_0^\infty f_\alpha(u)\mathrm{e}^{-su}\,\mathrm{d}u = \lim_{\alpha\downarrow 0} -\frac{\mathrm{d}}{\mathrm{d}\alpha}\hat{f}_\alpha(s)\,.$$

We can find the following explicit formula

Lemma 2.19. *Assume $\mu_2 < \infty$. Then*

$$\mathbb{E}[\tau 1\!\!1_{\{\tau<\infty\}}] = \frac{1}{c-\lambda\mu}\Bigl[\frac{\lambda\mu_2}{2(c-\lambda\mu)}\delta(u) - \int_0^u \psi(u-y)\delta(y)\,\mathrm{d}y\Bigr]. \tag{2.18}$$

Proof. We get

$$\begin{aligned}-\frac{\mathrm{d}}{\mathrm{d}\alpha}\hat{f}_\alpha(s) = {}& \frac{\lambda s'(\alpha)}{cs-\lambda(1-M_Y(-s))-\alpha}\frac{1-M_Y(-s(\alpha))-M_Y'(-s(\alpha))s(\alpha)}{s(\alpha)^2} \\ &- \frac{\lambda((1-M_Y(-s(\alpha)))s(\alpha)^{-1}-(1-M_Y(-s))s^{-1})}{(cs-\lambda(1-M_Y(-s))-\alpha)^2}. \end{aligned}\tag{2.19}$$

It follows from (2.17) that

$$s'(0) = \frac{1}{c-\lambda\mu}$$

and we have already seen that

$$\lim_{\alpha\downarrow 0}\frac{1-M_Y(-s(\alpha))}{s(\alpha)} = \lim_{s\downarrow 0}\frac{1-M_Y(-s)}{s} = \lim_{s\downarrow 0} M_Y'(-s) = \mu\,.$$

Moreover,

$$\begin{aligned}\lim_{\alpha\downarrow 0}\frac{(1-M_Y(-s(\alpha)))-M_Y'(-s(\alpha))s(\alpha)}{s(\alpha)^2} &= \lim_{s\downarrow 0}\frac{(1-M_Y(-s))-M_Y'(-s)s}{s^2} \\ &= \lim_{s\downarrow 0}\frac{sM_Y''(-s)}{2s} = \frac{\mu_2}{2}\,.\end{aligned}$$

Letting α tend to 0 in (2.19) yields

$$\begin{aligned}&\int_0^\infty \mathbb{E}[\tau 1\!\!1_{\{\tau<\infty\}} \mid C_0 = u]\,\mathrm{e}^{-su}\,\mathrm{d}u \\ &\quad= \frac{\lambda\mu_2}{2(c-\lambda\mu)(cs-\lambda(1-M_Y(-s)))} - \frac{\lambda\mu-\lambda s^{-1}(1-M_Y(-s))}{(cs-\lambda(1-M_Y(-s)))^2} \\ &\quad= \frac{\lambda\mu_2}{2(c-\lambda\mu)^2}\frac{c-\lambda\mu}{cs-\lambda(1-M_Y(-s))} \\ &\qquad\qquad - \frac{1}{c-\lambda\mu}\frac{c-\lambda\mu}{cs-\lambda(1-M_Y(-s))}\Bigl(\frac{1}{s}-\frac{c-\lambda\mu}{cs-\lambda(1-M_Y(-s))}\Bigr) \\ &\quad= \frac{1}{c-\lambda\mu}\Bigl[\frac{\lambda\mu_2}{2(c-\lambda\mu)}\hat{\delta}(s) - \hat{\delta}(s)\hat{\psi}(s)\Bigr].\end{aligned}$$

But this is the Laplace transform of the assertion. □

Corollary 2.20. *Let $t > 0$. Then*

$$\mathbb{P}[t < \tau < \infty] < \frac{\lambda\mu_2}{2(c-\lambda\mu)^2 t}\,.$$

Proof. Because $\delta(u)$ and $\psi(u)$ take values in $[0,1]$ it is clear from (2.18) that

$$\mathbb{E}[\tau 1\!\!1_{\{\tau<\infty\}}] < \frac{\lambda\mu_2}{2(c-\lambda\mu)^2}\,.$$

By Markov's inequality

$$\mathbb{P}[t < \tau < \infty] = \mathbb{P}[\tau 1\!\!1_{\{\tau<\infty\}} > t] < \frac{1}{t}\mathbb{E}[\tau 1\!\!1_{\{\tau<\infty\}}] < \frac{\lambda\mu_2}{2(c-\lambda\mu)^2 t}\,.$$

$\square$

From (2.18) it is possible to get an explicit expression for $u = 0$

$$\mathbb{E}[\tau 1\!\!1_{\{\tau<\infty\}} \mid C_0 = 0] = \frac{\lambda\mu_2}{2(c-\lambda\mu)^2}\Bigl(1 - \frac{\lambda\mu}{c}\Bigr) = \frac{\lambda\mu_2}{2c(c-\lambda\mu)}$$

and

$$\mathbb{E}[\tau \mid \tau < \infty, C_0 = 0] = \frac{\mu_2}{2\mu(c-\lambda\mu)}\,.$$

Example 2.18 (continued). For Exp(β) distributed claims we get for $R = \beta - \lambda/c$

$$\begin{aligned}
&\mathbb{E}[\tau 1\!\!1_{\{\tau<\infty\}}] \\
&\quad = \frac{\beta}{c\beta-\lambda}\Bigl[\frac{2\lambda}{2\beta(c\beta-\lambda)}\Bigl(1 - \frac{\lambda}{\beta c}\mathrm{e}^{-Ru}\Bigr) - \int_0^u \frac{\lambda}{\beta c}\mathrm{e}^{-R(u-y)}\Bigl(1 - \frac{\lambda}{\beta c}\mathrm{e}^{-Ry}\Bigr)\,\mathrm{d}y\Bigr] \\
&\quad = \frac{\beta}{c\beta-\lambda}\Bigl[\frac{\lambda}{\beta(c\beta-\lambda)}\Bigl(1 - \frac{\lambda}{\beta c}\mathrm{e}^{-Ru}\Bigr) - \frac{\lambda}{\beta c}\mathrm{e}^{-Ru}\Bigl(\frac{c}{\beta c-\lambda}(\mathrm{e}^{Ru} - 1) - \frac{\lambda}{\beta c}u\Bigr)\Bigr] \\
&\quad = \frac{\lambda}{\beta c^2(\beta c-\lambda)}\mathrm{e}^{-Ru}(\lambda u + c) = \frac{1}{c(\beta c-\lambda)}\psi(u)(\lambda u + c)\,.
\end{aligned}$$

The conditional expectation of the time of ruin is linear in u

$$\mathbb{E}[\tau \mid \tau < \infty] = \frac{\lambda u + c}{c(\beta c-\lambda)}\,.$$

$\blacksquare$

2.13. Seal's formulae

We consider now the probability of ruin within finite time $\psi(u,t)$. But first we want to find the conditional finite ruin probability given C_t for some t fixed.

Lemma 2.21. *Let t be fixed, $u = 0$ and $0 < y \le ct$. Then*

$$\mathbb{P}[C_s \ge 0, 0 \le s \le t \mid C_t = y] = \frac{y}{ct}\,.$$

Proof. Consider

$$C_t - C_{(t-s)-} = cs - \sum_{i=N_{(t-s)-}+1}^{N_t} Y_i\,.$$

This is also a Cramér-Lundberg model. Thus

$$\begin{aligned}\mathbb{P}[C_s \ge 0, 0 \le s \le t \mid C_t = y] &= \mathbb{P}[C_t - C_{t-s} \ge 0, 0 \le s \le t \mid C_t - C_0 = y] \\ &= \mathbb{P}[C_s \le C_t, 0 \le s \le t \mid C_t = y]\,.\end{aligned}$$

Let $S_s = cs - C_s$ denote the aggregate claims up to time s. Denote by σ the permutations of $\{1, 2, \dots, n\}$. Then

$$\begin{aligned}\mathbb{E}[S_s \mid S_t = y, N_t = n, N_s = k] &= \frac{1}{n!}\mathbb{E}\Bigl[\sum_\sigma \sum_{i=1}^k Y_{\sigma(i)} \Bigm| S_t = y, N_t = n, N_s = k\Bigr] \\ &= \frac{k(n-1)!}{n!}\mathbb{E}\Bigl[\sum_{i=1}^n Y_i \Bigm| S_t = y, N_t = n, N_s = k\Bigr] = \frac{ky}{n}\,.\end{aligned}$$

Because, given $N_t = n$, the claim times are uniformly distributed in $[0,t]$ (Proposition 1.10) we obtain

$$\begin{aligned}\mathbb{E}[S_s \mid S_t = y, N_t = n] &= \sum_{k=0}^n \binom{n}{k}\Bigl(\frac{s}{t}\Bigr)^k\Bigl(\frac{t-s}{t}\Bigr)^{n-k}\frac{ky}{n} \\ &= \sum_{k=1}^n \binom{n-1}{k-1}\Bigl(\frac{s}{t}\Bigr)^k\Bigl(\frac{t-s}{t}\Bigr)^{n-k} y = \frac{sy}{t}\,.\end{aligned}$$

This is independent of n. Thus

$$\mathbb{E}[S_s \mid S_t = y] = \frac{sy}{t}$$

and

$$\mathbb{E}[C_s \mid C_t = y] = \mathbb{E}[cs - S_s \mid S_t = ct - y] = cs - \frac{s(ct-y)}{t} = \frac{sy}{t}\,.$$

Then for $v \le s < t$

$$\mathbb{E}\Big[\frac{y - C_s}{t-s} \;\Big|\; C_v, C_t = y\Big] = \frac{y - C_v - \mathbb{E}[(C_s - C_v) \mid C_v, C_t = y]}{t-s}$$
$$= \frac{y - C_v - (s-v)(y-C_v)/(t-v)}{t-s} = \frac{y - C_v}{t-v}$$

and the process

$$M_s = \frac{y - C_s}{t-s} \qquad (0 \le s < t)$$

is a conditional martingale. Note that $\lim_{s\uparrow t} M_s = c$. Let $T = \inf\{s \ge 0 : C_s = y\}$. Because

$$0 \le M_s 1\!\!1_{\{T>s\}} \le c$$

on $\{C_t = y\}$ we have

$$\lim_{s\uparrow t} \mathbb{E}[M_s 1\!\!1_{\{T>s\}} \mid C_t = y] = \mathbb{E}\Big[\lim_{s\uparrow t} M_s 1\!\!1_{\{T>s\}} \;\Big|\; C_t = y\Big] = c\mathbb{P}[T = t \mid C_t = y]\,.$$

Note that by the stopping theorem $(M_{T\wedge t})$ is a bounded martingale. Thus because $M_T = 0$ on $\{T < t\}$

$$\frac{y}{t} = M_0 = \lim_{s\uparrow t} \mathbb{E}[M_{T\wedge s} \mid C_t = y] = \lim_{s\uparrow t} \mathbb{E}[M_s 1\!\!1_{\{T>s\}} \mid C_t = y] = c\mathbb{P}[T = t \mid C_t = y]\,.$$

Note that $\mathbb{P}[T = t \mid C_t = y] = \mathbb{P}[C_s \le C_t, 0 \le s \le t \mid C_t = y]$. □

Conditioning on C_t is in fact a conditioning on the aggregate claim size S_t. Denote by $F(x;t)$ the distribution function of S_t. For integration we use $\mathrm{d}F(\cdot;t)$ for integration with respect to the measure $F(\cdot;t)$. Moreover, let $\delta(u,t) = 1 - \psi(u,t) = \mathbb{P}[\tau > t]$.

Theorem 2.22. *For initial capital 0 we have*

$$\delta(0,t) = \frac{1}{ct}\mathbb{E}[C_t \vee 0] = \frac{1}{ct}\int_0^{ct} F(y;t)\,\mathrm{d}y\,.$$

Proof. By Lemma 2.21 it follows readily that

$$\delta(0,t) = \mathbb{E}[\mathbb{P}[\tau > t \mid C_t]] = \mathbb{E}\Big[\frac{C_t \vee 0}{ct}\Big]\,.$$

Moreover,

$$\mathbb{E}[C_t \vee 0] = \mathbb{E}[(ct - S_t) \vee 0] = \int_0^{ct} (ct - z)\,\mathrm{d}F(z;t)$$
$$= \int_0^{ct}\int_z^{ct} \mathrm{d}y\,\mathrm{d}F(z;t) = \int_0^{ct}\int_0^{y} \mathrm{d}F(z;t)\,\mathrm{d}y = \int_0^{ct} F(y;t)\,\mathrm{d}y\,.$$

□

Let now the initial capital be arbitrary.

Theorem 2.23. *With the notation used above we have for $u > 0$*

$$\delta(u,t) = F(u+ct;t) - \int_u^{u+ct} \delta\Big(0, t - \frac{v-u}{c}\Big) F\Big(\mathrm{d}v; \frac{v-u}{c}\Big)\,.$$

Proof. Note that

$$\delta(u,t) = \mathbb{P}[C_t > 0] - \mathbb{P}[\exists 0 \le s < t : C_s = 0, C_v > 0 \text{ for } s < v \le t]$$

and $\mathbb{P}[C_t > 0] = F(u+ct;t)$. Let $T = \sup\{0 \le s \le t : C_s = 0\}$ and set $T = \infty$ if $C_s > 0$ for all $s \in [0,t]$. Then, noting that $\mathbb{P}[C_s > 0 : 0 < s \le t \mid C_0 = 0] = \mathbb{P}[C_s \ge 0 : 0 < s \le t \mid C_0 = 0]$,

$$\begin{aligned}\mathbb{P}[T \in [s, s+\mathrm{d}s)] &= \mathbb{P}[C_s \in (-c\,\mathrm{d}s, 0], C_v > 0 \text{ for } v \in [s+\mathrm{d}s, t]]\\ &= (F(u+c(s+\mathrm{d}s);s) - F(u+cs;s))\delta(0, t-s-\mathrm{d}s)\\ &= \delta(0,t-s)c\,\mathrm{d}F(u+cs;s)\,.\end{aligned}$$

Thus

$$\begin{aligned}\delta(u,t) &= F(u+ct;t) - \int_0^t \delta(0,t-s)c\,\mathrm{d}F(u+cs;s)\\ &= F(u+ct;t) - \int_u^{u+ct} \delta\Big(0, t-\frac{v-u}{c}\Big)\,\mathrm{d}F\Big(v, \frac{v-u}{c}\Big)\,.\end{aligned}$$

□

2.14. Finite time Lundberg inequalities

In this section we derive an upper bound for probabilities of the form

$$\mathbb{P}[\underline{y}u < \tau \le \bar{y}u]$$

for $0 \le \underline{y} < \bar{y} < \infty$. Assume that the Lundberg exponent R exists. We will use the martingale (2.4) and the stopping theorem. For $r \ge 0$ such that $M_Y(r) < \infty$,

$$\begin{aligned}\mathrm{e}^{-ru} &= \mathbb{E}[\mathrm{e}^{-rC_{\tau\wedge\bar{y}u} - \theta(r)(\tau\wedge\bar{y}u)}] > \mathbb{E}[\mathrm{e}^{-rC_{\tau\wedge\bar{y}u} - \theta(r)(\tau\wedge\bar{y}u)}; \underline{y}u < \tau \le \bar{y}u]\\ &= \mathbb{E}[\mathrm{e}^{-rC_\tau - \theta(r)\tau} \mid \underline{y}u < \tau \le \bar{y}u]\,\mathbb{P}[\underline{y}u < \tau \le \bar{y}u]\\ &> \mathbb{E}[\mathrm{e}^{-\theta(r)\tau} \mid \underline{y}u < \tau \le \bar{y}u]\,\mathbb{P}[\underline{y}u < \tau \le \bar{y}u]\\ &> \mathrm{e}^{-\max\{\theta(r)\underline{y}u, \theta(r)\bar{y}u\}}\mathbb{P}[\underline{y}u < \tau \le \bar{y}u]\,.\end{aligned}$$

Thus

$$\mathbb{P}[\underline{y}u < \tau \le \bar{y}u \mid C_0 = u] < \mathrm{e}^{-(r - \max\{\theta(r)\underline{y}, \theta(r)\bar{y}\})u} \,.$$

Choosing the exponent as small as possible we obtain

$$\mathbb{P}[\underline{y}u < \tau \le \bar{y}u \mid C_0 = u] < \mathrm{e}^{-R(\underline{y},\bar{y})u} \tag{2.20}$$

where

$$R(\underline{y}, \bar{y}) = \sup_{r \in \mathbb{R}} \min\{r - \theta(r)\underline{y}, r - \theta(r)\bar{y}\} \,.$$

The supremum over $r \in \mathbb{R}$ makes sense because $\theta(r) > 0$ for $r < 0$ and $\theta(r) = \infty$ if $M_Y(r) = \infty$. Since $\theta(R) = 0$ it follows that $R(\underline{y}, \bar{y}) \ge R$. Hence (2.20) could be useful.

For further investigation of $R(\underline{y}, \bar{y})$ we consider the function $f_y(r) = r - y\theta(r)$. Let $r_\infty = \sup\{r \ge 0 : M_Y(r) < \infty\}$. Then $f_y(r)$ exists in the interval $(-\infty, r_\infty)$ and $f_y(r) = -\infty$ for $r \in (r_\infty, \infty)$. If $M_Y(r_\infty) < \infty$ then $|f_y(r_\infty)| < \infty$. Since $f_y''(r) = -yM_Y''(r) < 0$ the function $f_y(r)$ is concave. Thus there exists a unique r_y such that $f_y(r_y) = \sup\{f_y(r) : r \in \mathbb{R}\}$. Because $f_y(0) = 0$ and $f_y'(0) = 1 - y\theta'(0) = 1 + y(c - \lambda\mu) > 0$ we conclude that $r_y > 0$. r_y is either the solution to the equation

$$1 + y(c - \lambda M_Y'(r_y)) = 0$$

or $r_y = r_\infty$. We assume now that $R < r_\infty$ and let

$$y_0 = (\lambda M_Y'(R) - c)^{-1} \,.$$

It follows $r_{y_0} = R$. We call y_0 critical value.

Because $\frac{\mathrm{d}}{\mathrm{d}y} f_y(r) = -\theta(r)$ it follows readily that

$$\frac{\mathrm{d}}{\mathrm{d}y} f_y(r) \gtreqless 0 \iff r \lesseqgtr R \,.$$

We get

$$f_{\underline{y}}(r) \lesseqgtr f_{\bar{y}}(r) \text{ as } r \lesseqgtr R \,.$$

Because $M_Y'(r)$ is an increasing function it follows that

$$y \gtreqless y_0 \iff r_y \lesseqgtr R \,.$$

Thus

$$R(\underline{y}, \bar{y}) = \begin{cases} R & \text{if } \underline{y} \le y_0 \le \bar{y}\,, \\ f_{\bar{y}}(r_{\bar{y}}) & \text{if } \bar{y} < y_0, \\ f_{\underline{y}}(r_{\underline{y}}) & \text{if } \underline{y} > y_0. \end{cases}$$

If $y_0 \in [\underline{y}, \bar{y}]$ then we got again Lundberg's inequality. If $\bar{y} < y_0$ then $R(\underline{y}, \bar{y})$ does not depend on $\underline{y}$. By choosing $\underline{y}$ as small as possible we obtain

$$\mathbb{P}[0 < \tau \le yu \mid C_0 = u] < \mathrm{e}^{-R(0,y)u} \tag{2.21}$$

for $y < y_0$. Note that $R(0, y) > R$. Analogously

$$\mathbb{P}[yu < \tau < \infty \mid C_0 = u] < \mathrm{e}^{-R(y,\infty)u} \tag{2.22}$$

for $y > y_0$. The strict inequality follows from

$$\mathrm{e}^{-ru} \ge \mathbb{E}[\mathrm{e}^{-rC_\tau - \theta(r)\tau}; yu < \tau < \infty] > \mathrm{e}^{-\theta(r)yu}\mathbb{P}[yu < \tau < \infty]$$

if $0 < r < R$, compare also (2.5). Again $R(y, \infty) > R$.

We see that $\mathbb{P}[\tau \notin ((y_0 - \varepsilon)u, (y_0 + \varepsilon)u)]$ goes much faster to 0 than $\mathbb{P}[\tau \in ((y_0 - \varepsilon)u, (y_0 + \varepsilon)u)]$. Intuitively ruin should therefore occur near $y_0 u$.

Theorem 2.24. *Assume that $R < r_\infty$. Then*

$$\frac{\tau}{u} \longrightarrow y_0$$

in probability on the set $\{\tau < \infty\}$ as $u \to \infty$.

Proof. Choose $\varepsilon > 0$. By the Cramér-Lundberg approximation $\mathbb{P}[\tau < \infty \mid C_0 = u] \exp\{Ru\} \to C$ for some $C > 0$. Then

$$\begin{aligned}
P\Big[\Big|\frac{\tau}{u} - y_0\Big| > \varepsilon \;\Big|\; \tau < \infty, C_0 = u\Big]
&= \frac{\mathbb{P}[\tau < (y_0 - \varepsilon)u \mid C_0 = u] + \mathbb{P}[(y_0 + \varepsilon)u < \tau < \infty \mid C_0 = u]}{\mathbb{P}[\tau < \infty \mid C_0 = u]} \\
&\le \frac{\exp\{-R(0, y_0 - \varepsilon)u\} + \exp\{-R(y_0 + \varepsilon, \infty)u\}}{\mathbb{P}[\tau < \infty \mid C_0 = u]} \\
&= \frac{\exp\{-(R(0, y_0 - \varepsilon) - R)u\} + \exp\{-(R(y_0 + \varepsilon, \infty) - R)u\}}{\mathbb{P}[\tau < \infty \mid C_0 = u]\mathrm{e}^{Ru}} \longrightarrow 0
\end{aligned}$$

as $u \to \infty$. □

2.15. Bibliographical remarks

The Cramér-Lundberg model was introduced by Filip Lundberg [40]. The basic results, like Lundberg's inequality and the Cramér-Lundberg approximation go back to Lundberg [41] and Cramér [16], see also [17], [32] or [45]. The differential equation for the ruin probability (2.2) and the Laplace transform of the ruin probability can for instance be found in [17]. The proof of the Cramér-Lundberg approximation using the renewal theorem is due to Feller [26]. The martingale (2.4) was introduced by Gerber [27]. The latter paper contains also the martingale proof of Lundberg's inequality and the proof of (2.21). The proof of (2.22) goes back to Grandell [33]. Section 2.14 can also be found in [23]. Some similar results are obtained in [4] and [5]. Segerdahl [53] proved the asymptotic normality of the ruin time conditioned on ruin occurs.

The results on reinsurance and ruin can be found in [61] and [15]. Optimal reinsurance strategies in a dynamical setup are treated in [50]. Proposition 2.9 can be found in [29, p.130].

The term **severity of ruin** was introduced in [30]. More results on the topic are obtained in [21], [20], [49] and [45]. A diffusion approximation, even though not obtained mathematically correct, was already used by Hadwiger [35]. The modern approach goes back to Iglehart [36], see also [31], [6] and [48]. The deVylder approximation was introduced in [60]. Formula (2.14) was found by Asmussen [6]. It is however stated incorrectly in the paper. For the correct formula see [10] or [45]. The Beekman-Bowers approximation can be found in [11].

Theorem 2.15 was obtain by von Bahr [9] in the special case of Pareto distributed claim sizes and by Thorin and Wikstad [59] in the case of lognormally distributed claims. The theorem in the present form is basically found in [22], see also [24].

Section 2.12 is taken from [47]. Parts can also be found in [48] and [10]. (2.16) was first obtained in [17]. The expected value of the ruin time for exponentially distributed claims can be found in [29, p.138]. Seal's formulae were first obtained by Takács [55, p.56]. Seal [51] and [52] introduced them to risk theory. The present presentation follows [29].

3. The renewal risk model

3.1. Definition of the renewal risk model

The easiest generalisation of the classical risk model is to assume that the claim number process is a renewal process. Then the risk process is not Markovian anymore, because the distribution of the time of the next claim is dependent on the past via the time of the last claim. We therefore cannot find an ordinary differential equation for $\psi(u)$ as in the classical case.

Let

$$C_t = u + ct - \sum_{i=1}^{N_t} Y_i$$

where

- $\{N_t\}$ is a renewal process with event times $0 = T_0, T_1, T_2, \ldots$ with interarrival distribution function F, mean λ^{-1} and moment generating function $M_T(r)$. We denote by T a generic random variable with distribution function F. The first claim time T_1 has distribution function F^1. If nothing else is said we consider the ordinary case where $F^1 = F$.
- The claim amounts $\{Y_i\}$ build an iid sequence with distribution function G, moments $\mu_k = \mathbb{E}[Y_i^k]$ and moment generating function $M_Y(r)$. As in the classical model $\mu = \mu_1$.
- $\{N_t\}$ and $\{Y_i\}$ are independent.

This model is called **renewal risk model** or, after the person who introduced the model, **Sparre Andersen model**.

As in the classical model we define the ruin time $\tau = \inf\{t > 0 : C_t < 0\}$ and the ruin probabilities $\psi(u,t) = \mathbb{P}[\tau \le t]$ and $\psi(u) = \mathbb{P}[\tau < \infty]$. Considering the random walk $\{C_{T_i}\}$ we can conclude that

$$\psi(u) = 1 \qquad \Longleftrightarrow \qquad \mathbb{E}[C_{T_2} - C_{T_1}] = c/\lambda - \mu \le 0\,.$$

In order to avoid ruin a.s. we assume the net profit condition $c > \lambda\mu$. Note that in contrary to the classical case

$$\mathbb{E}[C_t - u] \neq (c - \lambda\mu)t \text{ for most } t \in \mathbb{R}$$

except if T has an exponential distribution. But

$$\lim_{t\to\infty} \frac{1}{t}(C_t - u) = c - \lambda\mu$$

by the law of large numbers. It follows again that $\psi(u) \to 0$ as $u \to \infty$.

3.2. The adjustment coefficient

We were able to construct some martingales of the form $\exp\{-rC_t - \theta(r)t\}$ in the classical case. Because the risk process $\{C_t\}$ is not Markov anymore we cannot hope to get such a martingale again because Lemma 2.1 does not hold any more. But the claim times are renewal times because $\{C_{T_1+t} - C_{T_1}\}$ is a renewal risk model independent of T_1 and Y_1.

Lemma 3.1. *Let $\{C_t\}$ be an (ordinary) renewal risk model. For any $r \ge 0$ such that $M_Y(r) < \infty$ let $\theta(r)$ be the unique solution to*

$$M_Y(r)M_T(-\theta(r) - cr) = 1\,. \tag{3.1}$$

Then the discrete time process

$$\{\exp\{-rC_{T_n} - \theta(r)T_n\} : n \in \mathbb{N}\}$$

is a martingale.

Remark. In the classical case with $F(t) = 1 - \mathrm{e}^{-\lambda t}$ equation (3.1) is

$$M_Y(r)\frac{\lambda}{\lambda + \theta(r) + cr} = 1$$

which is equivalent to the equation (2.3) in the classical case. ■

Proof. If $r \ge 0$ then $M_Y(r) \ge 1$. Because $M_T(r)$ is an increasing and continuous function defined for all $r \le 0$ and because $M_T(r) \to 0$ as $r \to -\infty$ there exists a unique solution $\theta(r)$ to (3.1). Then

$$\begin{aligned}\mathbb{E}[\mathrm{e}^{-rC_{T_{n+1}}-\theta(r)T_{n+1}} \mid \mathcal{F}_{T_n}] &= \mathbb{E}[\mathrm{e}^{-r(c(T_{n+1}-T_n)-Y_{n+1})-\theta(r)(T_{n+1}-T_n)} \mid \mathcal{F}_{T_n}]\mathrm{e}^{-rC_{T_n}-\theta(r)T_n} \\ &= \mathbb{E}[\mathrm{e}^{rY_{n+1}}\mathrm{e}^{-(cr+\theta(r))(T_{n+1}-T_n)}]\mathrm{e}^{-rC_{T_n}-\theta(r)T_n} \\ &= M_Y(r)M_T(-cr-\theta(r))\mathrm{e}^{-rC_{T_n}-\theta(r)T_n} = \mathrm{e}^{-rC_{T_n}-\theta(r)T_n}\,.\end{aligned}$$

□

As in the classical case we are interested in the case $\theta(r) = 0$. In the classical case there existed, besides the trivial solution $r = 0$, at most a second solution to the equation $\theta(r) = 0$. This was the case because $\theta(r)$ was a convex function. Let us therefore compute the second derivative of $\theta(r)$. Define $m_Y(r) = \log M_Y(r)$ and $m_T(r) = \log M_T(r)$. Thus $\theta(r)$ is the solution to the equation

$$m_Y(r) + m_T(-\theta(r) - cr) = 0\,.$$

By the implicit function theorem $\theta(r)$ is differentiable and

$$m'_Y(r) - (\theta'(r) + c)m'_T(-\theta(r) - cr) = 0\,. \tag{3.2}$$

Hence $\theta(r)$ is infinitely often differentiable and

$$m''_Y(r) - \theta''(r)m'_T(-\theta(r) - cr) + (\theta'(r) + c)^2 m''_T(-\theta(r) - cr) = 0\,.$$

Note that $M_T(r)$ is a strictly increasing function and so is $m_T(r)$. Thus $m'_T(-\theta(r) - cr) > 0$. Considering $m''_Y(r)$ as a variance it is not difficult to show that both $m''_Y(r)$ and $m''_T(r)$ are positive. We can assume that not both Y and T are deterministic, otherwise $\psi(u) = 0$. Then at least one of the two functions $m''_Y(r)$ and $m''_T(r)$ is strictly positive. From (3.2) it follows that $\theta'(r) + c > 0$. Thus $\theta''(r) > 0$ and $\theta(r)$ is a strictly convex function.

Let $r = 0$, i.e. $\theta(0) = 0$. Then

$$\theta'(0) = \frac{M'_Y(0)M_T(0)}{M_Y(0)M'_T(0)} - c = \lambda\mu - c < 0 \tag{3.3}$$

by the net profit condition. Thus there exists at most a second solution to $\theta(R) = 0$ with $R > 0$. Again we call this solution **adjustment coefficient** or **Lundberg exponent**. Note that R is the strictly positive solution to the equation

$$M_Y(r)M_T(-cr) = 1\,.$$

Example 3.2. Let $\{C_t\}$ be a renewal risk model with Exp(1) distributed claim amounts, premium rate $c = 5$ and interarrival time distribution

$$F(t) = 1 - \frac{1}{2}(\mathrm{e}^{-3t} + \mathrm{e}^{-7t})\,.$$

It follows that $M_Y(r)$ exists for $r < 1$, $M_T(r)$ exists for $r < 3$ and $\lambda = 4.2$. The net profit condition $5 > 4.2$ is fulfilled. The equation to solve is

$$\frac{1}{1-r}\,\frac{1}{2}\Bigl(\frac{3}{3+5r} + \frac{7}{7+5r}\Bigr) = 1\,.$$

Thus

$$3(7+5r)+7(3+5r)=2(1-r)(3+5r)(7+5r)$$

or equivalently

$$25r^3+25r^2-4r=0\,.$$

We find the obvious solution $r=0$ and the two other solutions

$$r_{1/2}=\frac{-25\pm\sqrt{1025}}{50}=\frac{-5\pm\sqrt{41}}{10}=\begin{cases}0.140312,\\-1.14031.\end{cases}$$

From the theory we conclude that $R=0.140312$. But we proved that there is no negative solution. Why do we get $r_2=-1.14031$? Obviously $M_Y(r_2)<1<\infty$. But $-cr_2=5.70156>3$ and thus $M_T(-cr_2)=\infty$. Thus r_2 is not a solution to the equation $M_Y(r)M_T(-cr)=1$. ■

Example 3.3. Let $T\sim\Gamma(\gamma,\alpha)$. Assume that there exists an r_∞ such that $M_Y(r)<\infty \iff r<r_\infty$. Then

$$\lim_{r\uparrow r_\infty} M_Y(r)=\infty\,.$$

Let $R(\gamma,\alpha)$ denote the Lundberg exponent which exists in this case. Then for $h>0$

$$\Bigl(\frac{\alpha}{\alpha+cR(\gamma,\alpha)}\Bigr)^{\gamma+h}M_Y(R(\gamma,\alpha))=\Bigl(\frac{\alpha}{\alpha+cR(\gamma,\alpha)}\Bigr)^{h}<1$$

and therefore $R(\gamma+h,\alpha)>R(\gamma,\alpha)$. Similarly

$$\Bigl(\frac{\alpha+h}{\alpha+h+cR(\gamma,\alpha)}\Bigr)^{\gamma}M_Y(R(\gamma,\alpha))>\Bigl(\frac{\alpha}{\alpha+cR(\gamma,\alpha)}\Bigr)^{\gamma}M_Y(R(\gamma,\alpha))=1$$

and thus $R(\gamma,\alpha+h)<R(\gamma,\alpha)$. It would be interesting to see what happens if we let $\alpha\to0$ and $\gamma\to0$ in such a way that the mean value remains constant. Let therefore $\alpha=\kappa\gamma$. Recall that $m_Y(r)=\log M_Y(r)$.

Observe that

$$\lim_{\gamma\to0}\gamma\log\Bigl(\frac{\kappa\gamma}{\kappa\gamma+cr}\Bigr)+m_Y(r)=m_Y(r)$$

and therefore $R(\gamma)=R(\gamma,\kappa\gamma)\to0$ as $\gamma\to0$. The latter follows because convergence to a continuous function is uniform on compact intervals. Let now $r(\gamma)=\gamma^{-1}R(\gamma)$. Then $r(\gamma)$ has to solve the equation

$$\log\Bigl(\frac{\kappa}{\kappa+cr}\Bigr)+\frac{m_Y(\gamma r)}{\gamma}=0\,. \tag{3.4}$$

Letting $\gamma \to 0$ yields

$$\log\Bigl(\frac{\kappa}{\kappa + cr}\Bigr) + rm_Y'(0) = \log\Bigl(\frac{\kappa}{\kappa + cr}\Bigr) + r\mu = 0\,.$$

The latter function is convex, has the root $r = 0$, the derivative $-c/\kappa + \mu < 0$ at 0 and converges to infinity as $r \to \infty$. Thus there exists exactly one additional solution r_0 and $r(\gamma) \to r_0$ as $\gamma \to 0$.

Replace r by $r(\gamma)$ in (3.4) and take the derivative, which exists by the implicit function theorem,

$$-\frac{cr'(\gamma)}{\kappa + cr(\gamma)} + \frac{(\gamma r'(\gamma) + r(\gamma))m_Y'(\gamma r(\gamma))}{\gamma} - \frac{m_Y(\gamma r(\gamma))}{\gamma^2} = 0\,.$$

One obtains

$$r'(\gamma) = \frac{\frac{\gamma r(\gamma) m_Y'(\gamma r(\gamma)) - m_Y(\gamma r(\gamma))}{\gamma^2}}{\frac{c}{\kappa + cr(\gamma)} - m_Y'(\gamma r(\gamma))}\,.$$

Note that by Taylor expansion

$$\gamma r(\gamma) m_Y'(\gamma r(\gamma)) - m_Y(\gamma r(\gamma)) = \frac{\sigma^2}{2}\gamma^2 r(\gamma)^2 + O(\gamma^3 r(\gamma)^3) = \frac{\sigma^2}{2}\gamma^2 r_0^2 + O(\gamma^3)$$

uniformly on compact intervals for which $R(\gamma)$ exists. The second equality holds because $\lim_{\gamma \downarrow 0} r'(\gamma) = r_0 > 0$ exists. We denoted by σ^2 the variance of the claims. We found

$$r'(\gamma) = \frac{\sigma^2 r_0^2(\kappa + cr_0)}{2(c - \kappa\mu - c\mu r_0)} + O(\gamma)\,.$$

Because the convergence is uniform it follows that

$$r(\gamma) = r_0 + \frac{\sigma^2 r_0^2(\kappa + cr_0)}{2(c - \kappa\mu - c\mu r_0)}\gamma + O(\gamma^2)$$

or

$$R(\gamma) = \gamma r_0 + \frac{\sigma^2 r_0^2(\kappa + cr_0)}{2(c - \kappa\mu - c\mu r_0)}\gamma^2 + O(\gamma^3)\,.$$

Note that the moment generating function of the interarrival times

$$\Bigl(\frac{\kappa\gamma}{\kappa\gamma - r}\Bigr)^{\gamma} = \Bigl(1 - \frac{r}{\kappa\gamma}\Bigr)^{-\gamma} \to \mathrm{e}^{r/\kappa}$$

converges to the moment generating function of deterministic interarrival times $T = 1/\kappa$. That means that $T(\gamma) \to 1/\kappa$ in distribution as $\gamma \to 0$. But the Lundberg exponents do not converge to the Lundberg exponent of the model with deterministic interarrival times. ■

3.3. Lundberg's inequality

3.3.1. The ordinary case

In order to find an upper bound for the ruin probability we try to proceed as in the classical case. Unfortunately we cannot find an analogue to (2.1). But recall the interpretation of (2.2) as conditioning on the ladder heights. Let $H(x) = \mathbb{P}[\tau < \infty, C_\tau \ge -x \mid C_0 = 0]$. Then

$$\psi(u) = \int_0^u \psi(u-x)\,\mathrm{d}H(x) + (H(\infty) - H(u)) \tag{3.5}$$

where $H(\infty) = \mathbb{P}[\tau < \infty \mid C_0 = 0]$. Unfortunately we cannot find an explicit expression for $H(u)$ except in some special cases. In order to find a result we have to link the Lundberg exponent R to H.

Lemma 3.4. *Let $\{C_t\}$ be an ordinary renewal risk model with Lundberg exponent R. Then*

$$\int_0^\infty \mathrm{e}^{Rx}\,\mathrm{d}H(x) = \mathbb{E}\left[\mathrm{e}^{-RC_\tau} 1\!\!1_{\{\tau<\infty\}} \,\middle|\, C_0 = 0\right] = 1\,.$$

Proof. The discrete time process $\{\mathrm{e}^{-RC_{T_i}}\}$ is a martingale with initial value 1 if $C_0 = 0$. We will drop the conditioning on $C_0 = 0$ in the sequel. By the stopping theorem

$$1 = \mathbb{E}\left[\mathrm{e}^{-RC_{\tau\wedge T_n}}\right] = \mathbb{E}\left[\mathrm{e}^{-RC_\tau} 1\!\!1_{\{\tau\le T_n\}}\right] + \mathbb{E}\left[\mathrm{e}^{-RC_{T_n}} 1\!\!1_{\{\tau> T_n\}}\right]\,.$$

The function $\mathrm{e}^{-RC_\tau} 1\!\!1_{\{\tau\le T_n\}}$ is increasing with n. By the monotone limit theorem it follows that

$$\lim_{n\to\infty} \mathbb{E}[\mathrm{e}^{-RC_\tau} 1\!\!1_{\{\tau\le T_n\}}] = \mathbb{E}\Bigl[\lim_{n\to\infty} \mathrm{e}^{-RC_\tau} 1\!\!1_{\{\tau\le T_n\}}\Bigr] = \mathbb{E}[\mathrm{e}^{-RC_\tau} 1\!\!1_{\{\tau<\infty\}}]\,.$$

For the second term we have $\mathrm{e}^{-RC_{T_n}} 1\!\!1_{\{\tau> T_n\}} \le 1$ and we know that $C_{T_n} \to \infty$ as $n \to \infty$. Thus

$$\lim_{n\to\infty} \mathbb{E}[\mathrm{e}^{-RC_{T_n}} 1\!\!1_{\{\tau> T_n\}}] = \mathbb{E}\Bigl[\lim_{n\to\infty} \mathrm{e}^{-RC_{T_n}} 1\!\!1_{\{\tau> T_n\}}\Bigr] = 0\,.$$

This proves the lemma. □

We are now ready to prove Lundberg's inequality.

Theorem 3.5. *Let $\{C_t\}$ be an ordinary renewal risk model and assume that the Lundberg exponent R exists. Then*

$$\psi(u) < \mathrm{e}^{-Ru} .$$

Proof. Assume that the assertion were wrong and let

$$u_0 = \inf\{u \geq 0 : \psi(u) \geq \mathrm{e}^{-Ru}\} .$$

Let $u_0 \leq u < u_0 + \varepsilon$ such that $\psi(u) \geq \mathrm{e}^{-Ru}$. Then

$$\psi(u_0) \geq \psi(u) \geq \mathrm{e}^{-Ru} > \mathrm{e}^{-R(u_0+\varepsilon)} .$$

Thus $\psi(u_0) \geq \mathrm{e}^{-Ru_0}$. Because $\mathrm{e}^{-RC_\tau} > 1$ it follows from Lemma 3.4 that $\psi(0) = H(\infty) < 1$ and therefore $u_0 \geq -R^{-1} \log H(\infty) > 0$. Assume first that $H(u_0-) = 0$. Then it follows from (3.5) that

$$\begin{aligned}
\mathrm{e}^{-Ru_0} \leq \psi(u_0) &= \psi(0)H(u_0) + \int_{u_0}^\infty \mathrm{d}H(x) < H(u_0) + \int_{u_0}^\infty \mathrm{e}^{R(x-u_0)}\, \mathrm{d}H(x) \\
&= \int_{u_0-}^\infty \mathrm{e}^{R(x-u_0)}\, \mathrm{d}H(x) = \mathrm{e}^{-Ru_0} \int_0^\infty \mathrm{e}^{Rx}\, \mathrm{d}H(x) = \mathrm{e}^{-Ru_0}
\end{aligned}$$

which is a contradiction. Thus $H(u_0-) > 0$.

As in the classical case we obtain from (3.5) and Lemma 3.4

$$\begin{aligned}
\mathrm{e}^{-Ru_0} \leq \psi(u_0) &= \int_0^{u_0} \psi(u_0 - x)\, \mathrm{d}H(x) + \int_{u_0}^\infty \mathrm{d}H(x) \\
&< \int_0^{u_0} \mathrm{e}^{-R(u_0-x)}\, \mathrm{d}H(x) + \int_{u_0}^\infty \mathrm{d}H(x) \\
&\leq \int_0^\infty \mathrm{e}^{-R(u_0-x)}\, \mathrm{d}H(x) = \mathrm{e}^{-Ru_0} .
\end{aligned}$$

This is a contradiction and the theorem is proved. □

As in the classical case we prove the result using martingale techniques. By the stopping theorem

$$\begin{aligned}
\mathrm{e}^{-Ru} = \mathrm{e}^{-RC_{T_0}} &= \mathbb{E}\left[\mathrm{e}^{-RC_{\tau\wedge T_n}}\right] \geq \mathbb{E}\left[\mathrm{e}^{-RC_{\tau\wedge T_n}}; \tau \leq T_n\right] \\
&= \mathbb{E}\left[\mathrm{e}^{-RC_\tau}; \tau \leq T_n\right] .
\end{aligned}$$

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

$$\mathrm{e}^{-Ru} \geq \mathbb{E}\left[\mathrm{e}^{-RC_\tau}; \tau < \infty\right] > \mathbb{P}[\tau < \infty]$$

because $C_\tau < 0$.

Example 3.6. Let Y_i be $\text{Exp}(\alpha)$ distributed and

$$F(t) = 1 - p\mathrm{e}^{-\beta t} - (1-p)\mathrm{e}^{-\gamma t}$$

where $0 < \beta < \gamma$ and $0 < p \le \gamma(\gamma-\beta)^{-1}$. The net profit condition can then be written as

$$\alpha c(\beta + p(\gamma - \beta)) > \beta\gamma\,.$$

The equation for the adjustment coefficient is

$$\frac{\alpha}{\alpha - r}\Bigl(\frac{p\beta}{\beta + cr} + \frac{(1-p)\gamma}{\gamma + cr}\Bigr) = 1\,.$$

This is equivalent to

$$c^2r^3 - c(\alpha c - \beta - \gamma)r^2 - (\alpha c(\beta + p(\gamma-\beta)) - \beta\gamma)r = 0$$

from which the solutions $r_0 = 0$ and

$$r_{1/2} = \frac{\alpha c - \beta - \gamma \pm \sqrt{(\alpha c - \gamma + \beta)^2 + 4\alpha c p(\gamma - \beta)}}{2c}$$

follow. Note that by the net profit condition $r_1 > 0 > r_2$. By the condition on p

$$\begin{aligned} r_2 < r_1 &\le \frac{\alpha c - \beta - \gamma + \sqrt{(\alpha c - \gamma + \beta)^2 + 4\alpha c\gamma}}{2c} \\ &< \frac{\alpha c - \beta - \gamma + \sqrt{(\alpha c + \gamma + \beta)^2}}{2c} = \alpha\,. \end{aligned}$$

Thus $M_Y(r_{1/2}) < \infty$. And

$$\begin{aligned} -cr_2 &= \frac{\beta + \gamma - \alpha c + \sqrt{(\alpha c - \gamma + \beta)^2 + 4\alpha c p(\gamma - \beta)}}{2} \\ &> \frac{\beta + \gamma - \alpha c + \sqrt{(\alpha c - \gamma + \beta)^2}}{2} \ge \beta \end{aligned}$$

and thus $M_T(-cr_2) = \infty$, i.e. r_2 is not a solution. Thus

$$R = \frac{\alpha c - \beta - \gamma + \sqrt{(\alpha c - \gamma + \beta)^2 + 4\alpha c p(\gamma - \beta)}}{2c}\,.$$

It follows from Lundberg's inequality that

$$\psi(u) < \mathrm{e}^{-Ru}\,.$$

$\blacksquare$

3.3.2. The general case

Let now F^1 be arbitrary. Let us denote the ruin probability in the ordinary model by $\psi^{\rm o}(u)$. We know that $\{C_{T_1+t}\}$ is an ordinary model with initial capital C_{T_1}. There are two possibilities: Ruin occurs at T_1 or $C_{T_1} \ge 0$. Thus

$$\begin{aligned}\psi(u) &= \int_0^\infty \Bigl(\int_0^{u+ct} \psi^{\rm o}(u+ct-y)\,{\rm d}G(y) + \int_{u+ct}^\infty {\rm d}G(y)\Bigr)\,{\rm d}F^1(t) \\ &< \int_0^\infty \Bigl(\int_0^{u+ct} {\rm e}^{-R(u+ct-y)}\,{\rm d}G(y) + \int_{u+ct}^\infty {\rm e}^{-R(u+ct-y)}\,{\rm d}G(y)\Bigr)\,{\rm d}F^1(t) \\ &= \int_0^\infty \int_0^\infty {\rm e}^{-R(u+ct-y)}\,{\rm d}G(y)\,{\rm d}F^1(t) \\ &= {\rm e}^{-Ru}\mathbb{E}\left[{\rm e}^{R(Y_1-cT_1)}\right] = M_Y(R)M_{T_1}(-cR){\rm e}^{-Ru}\,.\end{aligned}$$

Thus $\psi(u) < C{\rm e}^{-Ru}$ for $C = M_Y(R)M_{T_1}(-cR)$. In the cases considered so far we always had $C = 1$. But in the general case $C \in (0, M_Y(R))$ can be any value of this interval (let T_1 be deterministic). If $F^1(t) = F(t)$ then $C = 1$ by the equation determining the adjustment coefficient.

3.4. The Cramér-Lundberg approximation

3.4.1. The ordinary case

In order to obtain the Cramér-Lundberg approximation we proceed as in the classical case and multiply (3.5) by ${\rm e}^{Ru}$.

$$\psi(u){\rm e}^{Ru} = \int_0^u \psi(u-x){\rm e}^{R(u-x)}{\rm e}^{Rx}\,{\rm d}H(x) + {\rm e}^{Ru}(H(\infty) - H(u))\,.$$

We obtain the following

Theorem 3.7. *Let $\{C_t\}$ be an ordinary renewal risk model. Assume that R exists and that there exists an $r > R$ such that $M_Y(r) < \infty$. If the distribution of $Y_1 - cT_1$ given $Y_1 - cT_1 > 0$ is not arithmetic then*

$$\lim_{u\to\infty} \psi(u){\rm e}^{Ru} = \frac{1 - H(\infty)}{R\int_0^\infty x{\rm e}^{Rx}\,{\rm d}H(x)} =: C\,.$$

If the distribution of $Y_1 - cT_1$ is arithmetic with span γ then for $x \in [0, \gamma)$

$$\lim_{n\to\infty} \psi(x + n\gamma){\rm e}^{R(x+n\gamma)} = C{\rm e}^{Rx}\frac{1 - {\rm e}^{-R\gamma}}{R}\,.$$

Proof. We prove only the non-arithmetic case. The arithmetic case can be proved similarly. Recall from Lemma 3.4 that $\int_0^x \mathrm{e}^{Ry}\,\mathrm{d}H(y)$ is a proper distribution function. It follows as in the classical case that $\mathrm{e}^{Ru}(H(\infty) - H(u))$ is directly Riemann integrable. From $M_Y(r) < \infty$ for an $r > R$ it follows that $\int_0^\infty x\mathrm{e}^{Rx}\,\mathrm{d}H(x) < \infty$. Thus by the renewal theorem

$$\lim_{u\to\infty} \psi(u)\mathrm{e}^{Ru} = \frac{\int_0^\infty \mathrm{e}^{Ru}(H(\infty) - H(u))\,\mathrm{d}u}{\int_0^\infty x\mathrm{e}^{Rx}\,\mathrm{d}H(x)}\,.$$

Simplifying the numerator we find

$$\int_0^\infty \mathrm{e}^{Ru}\int_u^\infty \mathrm{d}H(x)\,\mathrm{d}u = \int_0^\infty\int_0^x \mathrm{e}^{Ru}\,\mathrm{d}u\,\mathrm{d}H(x) = \frac{1}{R}(1 - H(\infty))$$

where we used Lemma 3.4 in the last equality. □

It should be remarked that in general there is no explicit expression for the ladder height distribution $H(x)$. It is therefore difficult to find an explicit expression for C. Nevertheless we know that $0 < C < 1$.

Assume that we know C. Then

$$\psi(u) \approx C\mathrm{e}^{-Ru}$$

for u large.

Example 3.8. Assume that Y_i is $\mathrm{Exp}(\alpha)$ distributed. Because $M_Y(r) \to \infty$ as $r \to \alpha$ it follows that R exists and that the conditions of Theorem 3.7 are fulfilled. Consider the martingale $\mathrm{e}^{-RC_{T_i}}$. By the stopping theorem

$$\mathrm{e}^{-Ru} = \mathbb{E}\left[\mathrm{e}^{-RC_{\tau\wedge T_i}}\right] = \mathbb{E}\left[\mathrm{e}^{-RC_\tau}\mathbb{1}_{\{\tau\le T_i\}}\right] + \mathbb{E}\left[\mathrm{e}^{-RC_{T_i}}\mathbb{1}_{\{\tau>T_i\}}\right]\,.$$

As in the proof of Lemma 3.4 it follows that

$$\mathrm{e}^{-Ru} = \mathbb{E}\left[\mathrm{e}^{-RC_\tau}\mathbb{1}_{\{\tau<\infty\}}\right] = \mathbb{E}\left[\mathrm{e}^{-RC_\tau}\,\middle|\,\tau<\infty\right]\psi(u)\,.$$

Assume for the moment that we know $C_{\tau-}$. Let $Z = C_{\tau-} - C_\tau$ denote the size of the claim leading to ruin. The only information on Z we have is that $Z > C_{\tau-}$ because ruin occurs at time τ. Then

$$\begin{aligned}\mathbb{P}[-C_\tau > x \mid C_{\tau-} = y, \tau < \infty] &= \mathbb{P}[Z > y + x \mid C_{\tau-} = y, \tau < \infty]\\ &= \mathbb{P}[Y_1 > y + x \mid Y_1 > y] = \mathrm{e}^{-\alpha x}\,.\end{aligned}$$

It follows that

$$\mathbb{E}\left[\mathrm{e}^{-RC_\tau} \mid \tau < \infty\right] = \int_0^\infty \mathrm{e}^{Rx}\alpha\mathrm{e}^{-\alpha x}\,\mathrm{d}x = \frac{\alpha}{\alpha - R}\,.$$

Thus we obtain the explicit solution

$$\psi(u) = \frac{\alpha - R}{\alpha}\mathrm{e}^{-Ru}$$

to the ruin problem. Therefore, as in the classical case, the Cramér-Lundberg approximation is exact and we have an explicit expression for C. ∎

3.4.2. The general case

Consider the equation

$$\psi(u) = \int_0^\infty \Bigl(\int_0^{u+ct} \psi^{\mathrm{o}}(u + ct - y)\,\mathrm{d}G(y) + (1 - G(u + ct))\Bigr)\,\mathrm{d}F^1(t)\,.$$

We have to multiply the above equation by e^{Ru}. We obtain by Lundberg's inequality

$$\begin{aligned}
&\int_0^\infty \int_0^{u+ct} \psi^{\mathrm{o}}(u + ct - y)\mathrm{e}^{Ru}\,\mathrm{d}G(y)\,\mathrm{d}F^1(t) \\
&\quad = \int_0^\infty \int_0^{u+ct} \psi^{\mathrm{o}}(u + ct - y)\mathrm{e}^{R(u+ct-y)}\mathrm{e}^{Ry}\,\mathrm{d}G(y)\mathrm{e}^{-cRt}\,\mathrm{d}F^1(t) \\
&\quad < \int_0^\infty \int_0^{u+ct} \mathrm{e}^{Ry}\,\mathrm{d}G(y)\mathrm{e}^{-cRt}\,\mathrm{d}F^1(t) \qquad (3.6)\\
&\quad \le \int_0^\infty \int_0^\infty \mathrm{e}^{Ry}\,\mathrm{d}G(y)\mathrm{e}^{-cRt}\,\mathrm{d}F^1(t) = M_Y(R)M_{T_1}(-cR) < \infty\,.
\end{aligned}$$

We can therefore interchange limit and integration

$$\begin{aligned}
&\lim_{u\to\infty} \int_0^\infty \int_0^{u+ct} \psi^{\mathrm{o}}(u + ct - y)\mathrm{e}^{R(u+ct-y)}\mathrm{e}^{Ry}\,\mathrm{d}G(y)\mathrm{e}^{-cRt}\,\mathrm{d}F^1(t) \\
&\quad = \int_0^\infty \int_0^\infty C\mathrm{e}^{Ry}\,\mathrm{d}G(y)\mathrm{e}^{-cRt}\,\mathrm{d}F^1(t) \\
&\quad = M_Y(R)M_{T_1}(-cR)C\,.
\end{aligned}$$

The remaining term is converging to 0 because

$$\begin{aligned}
\int_0^\infty \int_{u+ct}^\infty \mathrm{e}^{Ru}\,\mathrm{d}G(y)\,\mathrm{d}F^1(t) &\le \int_0^\infty \int_{u+ct}^\infty \mathrm{e}^{R(y-ct)}\,\mathrm{d}G(y)\,\mathrm{d}F^1(t) \\
&\le \int_0^\infty \int_u^\infty \mathrm{e}^{Ry}\,\mathrm{d}G(y)\,\mathrm{e}^{-cRt}\,\mathrm{d}F^1(t) \\
&= M_{T_1}(-cR) \int_u^\infty \mathrm{e}^{Ry}\,\mathrm{d}G(y)\,.
\end{aligned}$$

Hence

$$\lim_{u\to\infty} \psi(u)e^{Ru} = M_Y(R)M_{T_1}(-cR)C \tag{3.7}$$

where C was the constant obtained in the ordinary case.

Example 3.8 (continued). In the general case we have to change the martingale because in general $\mathbb{E}[e^{-RC_{T_1}}] \neq e^{-Ru}$. Note that

$$\mathbb{E}[e^{-RC_{T_1}}] = e^{-Ru}M_Y(R)M_{T_1}(-cR) = \frac{\alpha}{\alpha - R}e^{-Ru}M_{T_1}(-cR)\,.$$

Let

$$M_n = \begin{cases} e^{-RC_{T_n}} & \text{if } n \geq 1, \\ \frac{\alpha}{\alpha-R}e^{-Ru}M_{T_1}(-cR) & \text{if } n = 0. \end{cases}$$

As in the ordinary case we obtain

$$\frac{\alpha}{\alpha - R}e^{-Ru}M_{T_1}(-cR) = \frac{\alpha}{\alpha - R}\psi(u)$$

or equivalently

$$\psi(u) = M_{T_1}(-cR)e^{-Ru}\,.$$

■

3.5. Diffusion approximations

As in the classical case it is possible to show the following

Proposition 3.9. *Let $\{C_t^{(n)}\}$ be a sequence of renewal risk models with initial capital u, interarrival time distribution $F^{(n)}(t) = F(nt)$, claim size distribution $G^{(n)}(x) = G(x\sqrt{n})$ and premium rate*

$$c^{(n)} = \left(1 + \frac{\frac{c-\lambda\mu}{\lambda\mu}}{\sqrt{n}}\right)\lambda^{(n)}\mu^{(n)} = c + (\sqrt{n} - 1)\lambda\mu$$

where $\lambda = \lambda^{(1)}$ and $\mu = \mu^{(1)}$. Then

$$\{C_t^{(n)}\} \to \{u + W_t\}$$

uniformly on finite intervals where $\{W_t\}$ is a $(c - \lambda\mu, \lambda\mu_2)$-Brownian motion.

Proof. See for instance [34, p.172]. □

Denote the ruin time of the renewal model by $\tau^{(n)}$ and the ruin time of the Brownian motion $\tau = \inf\{t \ge 0 : u + W_t < 0\}$. The diffusion approximation is based on the following

Proposition 3.10. *Let $\{C_t^{(n)}\}$ and $\{W_t\}$ as above. Then*

$$\lim_{n\to\infty} \mathbb{P}[\tau^{(n)} \le t] = \mathbb{P}[\tau \le t]\,.$$

Proof. See [62, Thm.9]. □

Remark. In the classical model we also had

$$\lim_{n\to\infty} \mathbb{P}[\tau^{(n)} < \infty] = \mathbb{P}[\tau < \infty]\,.$$

It is conjectured that the result remains true in the renewal case provided the second moment of the claim size distribution is finite. The author, however, is not aware of a formal proof. ■

As in the classical case numerical investigations show that the approximation only works well if $c/(\lambda\mu)$ is near 1.

3.6. Subexponential claim size distributions

In the classical case we had assumed that the ladder height distribution is subexponential. If we assume that the ladder height distribution H is subexponential then the asymptotic behaviour would follow as in the classical case. Unfortunately, we do hot have an explicit expression for the ladder height distribution in the renewal case. But the following proposition helps us out of this problem. We consider the ordinary case.

Proposition 3.11. *Let U be the distribution of $u - C_{T_1} = Y_1 - cT_1$ and let $m = \int_0^\infty (1-U(x))\,\mathrm{d}x$. Let $G_1(x) = \mu^{-1}\int_0^x (1-G(y))\,\mathrm{d}y$ and $U_1(x) = m^{-1}\int_0^x (1-U(y))\,\mathrm{d}y$. Then*

i) *If G_1 is subexponential, then U_1 is subexponential and*

$$\lim_{x\to\infty} \frac{1-U_1(x)}{1-G_1(x)} = \frac{\mu}{m}\,.$$

ii) *If U_1 is subexponential then $H/H(\infty)$ is subexponential and*

$$\lim_{x\to\infty} \frac{1-U_1(x)}{H(\infty)-H(x)} = \frac{c-\lambda\mu}{\lambda m(1-H(\infty))}\,.$$

Proof.

i) Using Fubini's theorem

$$\begin{aligned} m(1-U_1(x)) &= \int_x^\infty \int_0^\infty (1-G(y+ct))\,\mathrm{d}F(t)\,\mathrm{d}y \\ &= \int_0^\infty \int_x^\infty (1-G(y+ct))\,\mathrm{d}y\,\mathrm{d}F(t) \\ &= \mu \int_0^\infty (1-G_1(x+ct))\,\mathrm{d}F(t)\,. \end{aligned}$$

It follows by the bounded convergence theorem and Lemma 1.29 that

$$\lim_{x\to\infty} \frac{1-U_1(x)}{1-G_1(x)} = \lim_{x\to\infty} \frac{\mu}{m} \int_0^\infty \frac{1-G_1(x+ct)}{1-G_1(x)}\,\mathrm{d}F(t) = \frac{\mu}{m}\,.$$

By Lemma 1.35 U_1 is subexponential.

ii) Consider the random walk $\sum_{i=1}^n (Y_i - cT_i)$. We use the notation of Lemma 1.27. By the Wiener-Hopf factorisation for $y > 0$

$$1-U(y) = \int_{-\infty}^0 (H(y-z)-H(y))\,\mathrm{d}\rho(z)\,.$$

Let $0 < x < b$. Then using Fubini's theorem

$$\begin{aligned} \int_x^b (1-U(y))\,\mathrm{d}y &= \int_{-\infty}^0 \int_x^b (H(y-z)-H(y))\,\mathrm{d}y\,\mathrm{d}\rho(z) \\ &= \int_{-\infty}^0 \Bigl(\int_x^{x-z} (H(\infty)-H(y))\,\mathrm{d}y - \int_b^{b-z} (H(\infty)-H(y))\,\mathrm{d}y\Bigr)\,\mathrm{d}\rho(z)\,. \end{aligned}$$

Because $\int_{-\infty}^0 |z|\,\mathrm{d}\rho(z)$ is a finite upper bound (Lemma 1.27) we can interchange the integral and the limit $b \to \infty$ and obtain

$$\int_x^\infty (1-U(y))\,\mathrm{d}y = \int_{-\infty}^0 \int_x^{x-z} (H(\infty)-H(y))\,\mathrm{d}y\,\mathrm{d}\rho(z)\,.$$

We find

$$\int_{-\infty}^0 |z|(H(\infty)-H(x-z))\,\mathrm{d}\rho(z) \le m(1-U_1(x)) \le (H(\infty)-H(x)) \int_{-\infty}^0 |z|\,\mathrm{d}\rho(z)\,.$$

For $s > 0$ we obtain

$$(H(\infty)-H(x+s)) \int_{-s}^{0} |z| \, \mathrm{d}\rho(z) \le \int_{-s}^{0} |z|(H(\infty)-H(x-z)) \, \mathrm{d}\rho(z) \le m(1-U_1(x))$$

and therefore

$$1 \le \int_{-\infty}^{0} |z| \, \mathrm{d}\rho(z) \frac{H(\infty) - H(x+s)}{m(1 - U_1(x+s))} \le \frac{\int_{-\infty}^{0} |z| \, \mathrm{d}\rho(z)}{\int_{-s}^{0} |z| \, \mathrm{d}\rho(z)} \frac{1 - U_1(x)}{1 - U_1(x+s)} \, .$$

Thus by Lemma 1.29 $(H(\infty) - H(x))^{-1}(1 - U_1(x)) \to m^{-1} \int_{-\infty}^{0} |z| \, \mathrm{d}\rho(z)$ and $H/H(\infty)$ is subexponential by Lemma 1.35. The explicit value of the limit follows from Lemma 1.27 noting that $\int_{-\infty}^{0} |z| \, \mathrm{d}\rho(z)$ is the expected value of the first descending ladder height of the random walk $\sum_{i=1}^{n}(Y_i - cT_i)$.

□

We can now proof the asymptotic behaviour of the ruin probability.

Theorem 3.12. *Let $\{C_t\}$ be a renewal risk process. Assume that the integrated tail of the claim size distribution*

$$G_1(x) = \mu^{-1} \int_0^x (1 - G(y)) \, \mathrm{d}y$$

is subexponential. Then

$$\lim_{u \to \infty} \frac{\psi(u)}{\int_u^\infty (1 - G(y)) \, \mathrm{d}y} = \frac{\lambda}{c - \lambda\mu} \, .$$

Remark. Note that this is exactly Theorem 2.15. The difference is only that $G_1(x)$ is not the ladder height distribution anymore. ■

Proof. Let us first consider an ordinary renewal model. By repeating the proof of Theorem 2.15 we find that

$$\lim_{u \to \infty} \frac{\psi(u)}{1 - H(u)/H(\infty)} = \frac{H(\infty)}{1 - H(\infty)}$$

because $H(x)$ is subexponential by Proposition 3.11. Moreover, by Proposition 3.11

$$\begin{aligned} \lim_{u \to \infty} \frac{\psi(u)}{\int_u^\infty (1 - G(y)) \, \mathrm{d}y} &= \lim_{u \to \infty} \frac{\psi(u)}{1 - H(u)/H(\infty)} \frac{1}{H(\infty)} \frac{H(\infty) - H(u)}{1 - U_1(u)} \frac{1 - U_1(u)}{\mu(1 - G_1(u))} \\ &= \frac{H(\infty)}{1 - H(\infty)} \frac{1}{H(\infty)} \frac{\lambda m(1 - H(\infty))}{c - \lambda\mu} \frac{1}{m} = \frac{\lambda}{c - \lambda\mu} \, . \end{aligned}$$

For an arbitrary renewal risk model the assertion follows similarly as in Section 3.4.2. But two properties of subexponential distributions that are not proved here will be needed. □

3.7. Finite time Lundberg inequalities

Let $0 \le \underline{y} < \bar{y} < \infty$ and define $T = \inf\{T_n : n \in \mathbb{N}, T_n \ge \bar{y}u\}$. Using the stopping theorem

$$
\begin{aligned}
\mathrm{e}^{-ru} &= \mathbb{E}\left[\mathrm{e}^{-rC_{\tau\wedge T\wedge T_n}-\theta(r)(\tau\wedge T\wedge T_n)}\right] \\
&= \mathbb{E}\left[\mathrm{e}^{-rC_{\tau\wedge T}-\theta(r)(\tau\wedge T)}; T_n \ge \tau\wedge T\right] + \mathbb{E}\left[\mathrm{e}^{-rC_{T_n}-\theta(r)T_n}; T_n < \tau\wedge T\right] .
\end{aligned}
$$

By monotone convergence

$$\lim_{n\to\infty} \mathbb{E}\left[\mathrm{e}^{-rC_{\tau\wedge T}-\theta(r)(\tau\wedge T)}; T_n \ge \tau\wedge T\right] = \mathbb{E}\left[\mathrm{e}^{-rC_{\tau\wedge T}-\theta(r)(\tau\wedge T)}\right] .$$

For the second term we obtain

$$\lim_{n\to\infty} \mathbb{E}\left[\mathrm{e}^{-rC_{T_n}-\theta(r)T_n}; T_n < \tau\wedge T\right] \le \lim_{n\to\infty} \max\{\mathrm{e}^{-\theta(r)\bar{y}u}, 1\}\mathbb{P}[T_n < \tau\wedge T] = 0 .$$

Therefore

$$
\begin{aligned}
\mathrm{e}^{-ru} &= \mathbb{E}\left[\mathrm{e}^{-rC_{\tau\wedge T}-\theta(r)(\tau\wedge T)}\right] > \mathbb{E}\left[\mathrm{e}^{-rC_{\tau\wedge T}-\theta(r)(\tau\wedge T)}; \underline{y}u < \tau \le \bar{y}u\right] \\
&= \mathbb{E}\left[\mathrm{e}^{-rC_{\tau}-\theta(r)\tau} \,\middle|\, \underline{y}u < \tau \le \bar{y}u\right] \mathbb{P}[\underline{y}u < \tau \le \bar{y}u] \\
&> \mathbb{E}\left[\mathrm{e}^{-\theta(r)\tau} \,\middle|\, \underline{y}u < \tau \le \bar{y}u\right] \mathbb{P}[\underline{y}u < \tau \le \bar{y}u] \\
&> \mathrm{e}^{-\max\{\theta(r)\underline{y}u, \theta(r)\bar{y}u\}} \mathbb{P}[\underline{y}u < \tau \le \bar{y}u] .
\end{aligned}
$$

We obtain

$$\mathbb{P}[\underline{y}u < \tau \le \bar{y}u] < \mathrm{e}^{-(r-\max\{\theta(r)\underline{y},\theta(r)\bar{y}\})u} = \mathrm{e}^{-\min\{r-\theta(r)\underline{y}, r-\theta(r)\bar{y}\}u}$$

which leads to the finite time Lundberg inequality

$$\mathbb{P}[\underline{y}u < \tau \le \bar{y}u] < \mathrm{e}^{-R(\underline{y},\bar{y})u} .$$

Here, as in the classical case,

$$R(\underline{y}, \bar{y}) := \sup\{\min\{r - \theta(r)\underline{y}, r - \theta(r)\bar{y}\} : r \in \mathbb{R}\} .$$

We have already discussed $R(\underline{y}, \bar{y})$ in the classical case (section 2.14). Assume that $r_\infty = \sup\{r \ge 0 : M_Y(r) < \infty\} > R$. Let

$$y_0 = (\theta'(R))^{-1} = \Bigl(\frac{M_Y'(R)M_T(-cR)}{M_Y(R)M_T'(-cR)} - c\Bigr)^{-1}$$

denote the critical value. Then we obtain as in the classical case

$$\mathbb{P}[0 < \tau \le yu \mid C_0 = u] < \mathrm{e}^{-R(0,y)u}$$

and $R(0, y) > R$ if $y < y_0$, and

$$\mathbb{P}[yu < \tau < \infty \mid C_0 = u] < \mathrm{e}^{-R(y,\infty)u}$$

and $R(y, \infty) > R$ if $y > y_0$. The next result can be proved in the same way as Theorem 2.24 was proved.

Theorem 3.13. *Assume that $R < r_\infty$. Then*

$$\frac{\tau}{u} \longrightarrow y_0$$

in probability on the set $\{\tau < \infty\}$. □

3.8. Bibliographical remarks

The renewal risk model was introduced by E. Sparre Andersen [2]. A good reference are also the papers by Thorin [56], [57] and [58]. Example 3.3 is due to Grandell [32, p.71]. Lundberg's inequality in the ordinary case was first proved by Andersen [2, p.224]. Example 3.6 can be found in [32, p.60]. The Cramér-Lundberg approximation was first proved by Thorin, [56, p.94] in the ordinary case and [57, p.97] in the stationary case. The results on subexponential claim sizes go back to Embrechts and Veraverbeke [24]. The results on finite time Lundberg inequalities can be found in [23]. More results on the renewal risk model is to be found in Chapter 6 of [45].

References

[1] **Alsmeyer, G.** (1991). *Erneuerungstheory.* Teubner, Stuttgart.

[2] **Andersen, E.Sparre** (1957). *On the collective theory of risk in the case of contagion between the claims.* Transactions XVth International Congress of Actuaries, New York, II, 219–229.

[3] **Anderson, T.W.** (1958). *An Introduction to Multivariate Statistical Analysis.* Wiley, New York.

[4] **Arfwedson, G.** (1954). Research in collective risk theory. Part 1. *Skand. Aktuar Tidskr.*, 191–223.

[5] **Arfwedson, G.** (1955). Research in collective risk theory. Part 2. *Skand. Aktuar Tidskr.*, 53–100.

[6] **Asmussen, S.** (1984). Approximations for the probability of ruin within finite time. *Scand. Actuarial J.*, 31–57.

[7] **Asmussen, S. and Nielsen, H.M.** (1995). Ruin probabilities via local adjustment coefficients. *J. Appl. Probab.*, 736–755.

[8] **Athreya, K.B. and Ney, P.** (1972). *Branching processes.* Springer-Verlag, Berlin.

[9] **von Bahr, B.** (1975). Asymptotic ruin probabilities when exponential moments do not exist. *Scand. Actuarial J.*, 6–10.

[10] **Barndorff-Nielsen, O.E. and Schmidli, H.** (1994). Saddlepoint approximations for the probability of ruin in finite time. *Scand. Actuarial J.*, 169–186.

[11] **Beekman, J.** (1969). A ruin function approximation. *Trans. Soc. Actuaries* **21**, ,. 41–48 and 275–279

[12] **Brémaud, P.** (1981). *Point processes and queues.* Springer-Verlag, New York.

[13] **Bühlmann, H. and Straub, E.** (1970). Glaubwürdigkeit für Schadensätze. *Schweiz. Verein. Versicherungsmath. Mitt.* **70**, 111–133.

[14] **Chistyakov, V.P.** (1964). A theorem on sums of independent, positive random variables and its applications to branching processes. *Theory Probab. Appl.* **9**, 640–648.

[15] **Canteno, L.** (1986). Measuring the effects of reinsurance by the adjustment coefficient. *Insurance Math. Econom.* **5**, 169–182.

[16] **Cramér, H.** (1930). *On the mathematical theory of risk.* Skandia Jubilee Volume, Stockholm.

[17] **Cramér, H.** (1955). *Collective risk theory.* Skandia Jubilee Volume, Stockholm.

[18] **Daley, D.J. and Vere-Jones, D.** (1988). *An introduction to the theory of point processes.* Springer-Verlag, New York.

[19] **Dellacherie, C. and Meyer, P.A.** (1980). *Probabilités et potentiel.* Ch. VI, Hermann, Paris.

[20] **Dickson, D.C.M.** (1992). On the distribution of the surplus prior to ruin. *Insurance Math. Econom.* **11**, 191–207.

[21] **Dufresne, F. and Gerber, H.U.** (1988). The surpluses immediately before and at ruin, and the amount of the claim causing ruin. *Insurance Math. Econom.* **7**, 193–199.

[22] **Embrechts, P., Goldie, C.M. and Veraverbeke, N.** (1979). Subexponentiality and infinite divisibility. *Z. Wahrsch. verw. Geb.* **49**, 335–347.

[23] **Embrechts, P., Grandell, J. and Schmidli, H.** (1993). Finite-time Lundberg inequalities in the Cox case. *Scand. Actuarial J.*, 17–41.

[24] **Embrechts, P. and Veraverbeke, N.** (1982). Estimates for the probability of ruin with special emphasis on the possibility of large claims. *Insurance Math. Econom.* **1**, 55–72.

[25] **Ethier, S.N. and Kurtz, T.G.** (1986). *Markov processes.* Wiley, New York.

[26] **Feller, W.** (1971). *An introduction to probability theory and its applications.* Volume II, Wiley, New York.

[27] **Gerber, H.U.** (1973). Martingales in risk theory. *Schweiz. Verein. Versicherungsmath. Mitt.* **73**, 205–216.

[28] **Gerber, H.U.** (1975). The surplus process as a fair game – utilitywise. *ASTIN Bulletin* **8**, 307–322.

[29] **Gerber, H.U.** (1979). *An introduction to mathematical risk theory.* Huebner Foundation Monographs, Philadelphia.

[30] **Gerber, H.U., Goovaerts, M.J. and Kaas, R.** (1987). On the probability and severity of ruin. *ASTIN Bulletin* **17**, 151–163.

[31] **Grandell, J.** (1977). A class of approximations of ruin probabilities. *Scand. Actuarial J.*, 37–52.

[32] **Grandell, J.** (1991). *Aspects of risk theory.* Springer-Verlag, New York.

[33] **Grandell, J.** (1991). Finite time ruin probabilities and martingales. *Informatica* **2**, 3–32.

[34] **Gut, A.** (1988). *Stopped random walks.* Springer-Verlag, New York.

[35] **Hadwiger, H.** (1940). Über die Wahrscheinlichkeit des Ruins bei einer grossen Zahl von Geschäften. *Archiv für mathematische Wirtschafts- und Sozialforschung* **6**, 131–135.

[36] **Iglehart, D.L.** (1969). Diffusion approximations in collective risk theory. *J. Appl. Probab.* **6**, 285–292.

[37] **Karlin, S. and Taylor, H.M.** (1975). *A first course in stochastic processes.* Academic Press, New York.

[38] **Karlin, S. and Taylor, H.M.** (1981). *A second course in stochastic processes.* Academic Press, New York.

[39] **Klüppelberg, C.** (1988). Subexponential distributions and integrated tails. *J. Appl. Probab.* **25**, 132–141.

[40] **Lundberg, F.** (1903). **I**. *Approximerad Framställning av Sannolikhetsfunktionen.* **II**. *Återförsäkering av Kollektivrisker.* Almqvist & Wiksell, Uppsala.

[41] **Lundberg, F.** (1926). *Försäkringsteknisk Riskutjämning.* F. Englunds boktryckeri A.B., Stockholm.

[42] **Norberg, R.** (1986). A contribution to modelling of IBNR claims. *Scand. Actuarial J.*, 155–203.

[43] **Norberg, R.** (1986). Hierarchical credibility: analysis of a random effect linear model with nested classification. *Scand. Actuarial J.*, 204–222.

[44] **Norberg, R.** (1989). Experience rating in group life insurance. *Scand. Actuarial J.*, 194–224.

[45] **Rolski, T., Schmidli, H., Schmidt, V. and Teugels, J.L.** (1999). *Stochastic Processes for Insurance and Finance.* Wiley, Chichester.

[46] **Ross, S.M.** (1983). *Stochastic processes.* Wiley, New York.

[47] **Schmidli, H.** (1992). A general insurance risk model. *Ph.D Thesis, ETH Zürich.*

[48] **Schmidli, H.** (1994). Diffusion approximations for a risk process with the possibility of borrowing and investment. *Comm. Statist. Stochastic Models* **10**, 365–388.

[49] **Schmidli, H.** (1999). On the distribution of the surplus prior and at ruin. *ASTIN Bull.* **29**, 227–244.

[50] **Schmidli, H.** (2001). Optimal proportional reinsurance policies in a dynamic setting. *Scand. Actuarial J.*, 55–68.

[51] **Seal, H.L.** (1972). Numerical calculation of the probability of ruin in the Poisson/exponential case. *Schweiz. Verein. Versicherungsmath. Mitt.* **72**, 77–100.

[52] **Seal, H.L.** (1974). The numerical calculation of $U(w, t)$, the probability of non-ruin in an interval $(0, t)$. *Scand. Actuarial J.*, 121-139.

[53] **Segerdahl, C.-O.** (1955). When does ruin occur in the collective theory of risk?. *Skand. Aktuar Tidskr.*, 22–36.

[54] **Siegmund, D.** (1979). Corrected diffusion approximation in certain random walk problems. *Adv. in Appl. Probab.* **11**, 701–719.

[55] **Takács, L.** (1962). *Introduction to the theory of queues.* Oxford University Press, New York.

[56] **Thorin, O.** (1974). On the asymptotic behavior of the ruin probability for an infinite period when the epochs of claims form a renewal process. *Scand. Actuarial J.*, 81–99.

[57] **Thorin, O.** (1975). Stationarity aspects of the Sparre Andersen risk process and the corresponding ruin probabilities. *Scand. Actuarial J.*, 87–98.

[58] **Thorin, O.** (1982). Probabilities of ruin. *Scand. Actuarial J.*, 65–102.

[59] **Thorin, O. and Wikstad, N.** (1977). Calculation of ruin probabilities when the claim distribution is lognormal. *ASTIN Bulletin* **9**, 231–246.

[60] **deVylder, F.** (1978). A practical solution to the problem of ultimate ruin probability. *Scand. Actuarial J.*, 114–119.

[61] **Waters, H.R.** (1983). Some mathematical aspects of reinsurance insurance. *Insurance Math. Econom.* **2**, 17–26.

[62] **Whitt, W.** (1970). Weak convergence of probability measures on the function space $C[0, \infty)$. *Ann. Statist.* **41**, 939–944.

[63] **Williams, D.** (1979). *Diffusions, Markov processes and martingales.* Volume I, Wiley, New York.