We defined the compound Poisson process $S(t) = X_1 + ... + X_{N(t)}$ where $N(t)$ is a Poisson process, i.e. the probability that $N(t+h) - N(t) = k$ is $e^{-\lambda h}\frac{(\lambda h)^k}{k!}$ and the $X_1, ...$ are i.i.d. copies of a severity random variable X with a cdf $P(x)$, pdf $p(x)$, and the literature often writes the raw moments as $p_1, p_2, ...$(note these are **raw** moments even though there is no $\prime$ above them).

A few facts can be noted:

Define $F_{S(t+h|t)}(x) = \Pr[S(t+h) - S(t) \leq x \mid S(u) \text{ for all } u \leq t]$. Then

$$\begin{aligned} F_{S(t+h|t)}(x) &= \sum_{k=0}^{\infty} e^{-\lambda h}\frac{(\lambda h)^k}{k!}P^{*k}(x) \\ f_{S(t+h|t)}(x) &= \sum_{k=0}^{\infty} e^{-\lambda h}\frac{(\lambda h)^k}{k!}p^{*k}(x) \end{aligned}$$

Contingent on $S(u)$ for all $u \leq t$, the probability that the next claim occurs within the interval $[t+h, t+h+dh]$ and that the size of that claim is $\leq x$ is

$$\begin{aligned} & \Pr\{\text{no claims in } [t, t+h]\} \Pr\{1 \text{ claim in } [t+h, t+h+dh]\} \Pr\{\text{claim} \leq x\} \\ = & \left\{e^{-\lambda h}\right\}\left\{e^{-\lambda dh}(\lambda dh)\right\}\{P(x)\} \\ = & e^{-\lambda h}(1 - \lambda dh + ...)(\lambda dh)P(x) \\ = & e^{-\lambda h}\lambda dh P(x), \text{ ignoring } dh^2 \text{ etc.} \\ = & \lambda e^{-\lambda h}P(x)dh \end{aligned} \tag{1}$$

$C_{S(t+h|t)}(z) = \ln\left[M_{S(t+h|t)}(z)\right] = \ln\left[P_{N(t+h|t)}(M_X(z))\right] = \ln\left[e^{\lambda h(M_X(z)-1)}\right]$ $= \lambda h\left(M_X(z) - 1\right)$ so $\kappa_{S(t+h|t)k} = \lambda h p_k$ gives all the cumulants of $S(t+h \mid t)$ which could hardly be more convenient. For example, the first four central moments of $S(t+h \mid t)$ are: $\mu = \kappa_1 = \lambda h p_1$, $\sigma^2 = \kappa_2 = \lambda h p_2$, $\mu_3 = \kappa_3 = \lambda h p_3$, $\mu_4 = \kappa_4 + 3\sigma^4 = \lambda h p_4 + 3(\lambda h)^2 p_2^2$, etc. These formulas are why a lot of the literature uses the p_k = raw moment of X notation. We will use them repeatedly in what follows, so don't forget them.

Now, define the capitalization (or surplus) process $U(t) = u + ct - S(t)$. We're thinking of a process that starts out with an amount of capital (or surplus) $U(0) = u \geq 0$ and proceeds through time with a continuous addition of funds (or a premium inflow) at a rate of c, so that it has received ct in any length of time t, while absorbing risk costs (or losses) according to the compound Poisson process $S(t)$ defined above. Since u, c, and t are not random variables we know that $\mathbb{E}[U(t)] = u + ct - \lambda t p_1$ and the variance $\mathbb{V}[U(t)] = \mathbb{V}[S(t)] = \lambda t p_2$.

The question we want to study is whether and when $U(t) < 0$ ever occurs, and how much is it affected by the amount of starting capital (or surplus) u and the funding rate (or premium rate) c? If $U(t) < 0$ ever occurs it is called "ruin" (the process has run out of money, the enterprise is bankrupt) and the first time T that $U(T) < 0$ occurs is called the "time of ruin" T (in stochastic process terms, T satisfies the definition of a "stopping time"). The event $U(T) < 0$ (i.e. the first time it occurs) is called the "event of ruin." If the event of ruin has not yet occured, we are still in the "state of survival." We have some notational definitions:

$$\begin{aligned}
\varphi(u,t) &= \Pr\{U(\tau) \geq 0 \text{ for all } 0 \leq \tau \leq t \mid U(0) = u\} \\
&\quad \text{is the finite time survival probability function} \\
\varphi(u) &= \Pr\{U(\tau) \geq 0 \text{ for all } 0 \leq \tau < \infty \mid U(0) = u\} \\
&= \lim_{t\to\infty} \varphi(u,t) \text{ is the ultimate survival probability function} \\
\psi(u,t) &= \Pr\{U(\tau) < 0 \text{ for some } 0 \leq \tau \leq t \mid U(0) = u\} \\
&= \Pr\{T \leq t \mid U(0) = u\} \\
&= 1 - \varphi(u,t) \text{ is the finite time ruin probability function} \\
\psi(u) &= \Pr\{U(\tau) < 0 \text{ for some } 0 \leq \tau < \infty \mid U(0) = u\} \\
&= \lim_{t\to\infty} \psi(u,t) \\
&= \Pr\{T < \infty \mid U(0) = u\} \\
&= 1 - \varphi(u) \\
&\quad \text{is the ultimate ruin probability function}
\end{aligned}$$

Knowledge of these functions is of interest to anyone trying to manage a risky situation through a combination of starting capital u and on-going funding ct.

The ruin-survival question is interesting only when the inflow of funding equals or exceeds the expected value of the outflow of losses, $ct \geq \mathbb{E}[S(t)] = \lambda t p_1$, so that the question is whether random fluctuations cause ruin somewhere along the way even though the expected value of the outcome is survival. It is conventional to write $c = (1+\theta)\lambda p_1$ where $\theta \geq 0$ is called the "safety margin" in the funding (or the "safety loading" in the premium). θ represents the extent to which ongoing funding (or premiums) exceed the expected size of the on-going risk costs (or losses). In terms of θ, $\mathbb{E}[U(t)] = u + \theta\lambda t p_1$ follows easily from the definition of $U(t)$ and the expression for c in terms of θ.

The main **theorem** about this set-up is:

$$\psi(u) = \frac{e^{-Ru}}{\mathbb{E}[e^{-RU(T)} \mid T < \infty]} \tag{2}$$

where for $\theta > 0$, R is defined to be the smallest positive solution to the equation

$$R(1+\theta)p_1 = M_X(R) - 1$$

and for $\theta = 0$, R is defined to be 0. R depends only on the safety margin θ and the distribution $P(x)$ of X.

Proof: Pick any $t > 0$ and let $\bar{T} = T \wedge t$. (Later we'll let $t \to \infty$). We will look at $\mathbb{E}[e^{-RU(\bar{T})}]$ in two different ways. First,

$$\begin{aligned}
\mathbb{E}[e^{-RU(\bar{T})}] &= \mathbb{E}[e^{-R\left(u+c\bar{T}-S(\bar{T})\right)}] \text{ from def. of } U(\bar{T}) \\
&= e^{-Ru}\mathbb{E}[e^{-R\left(c\bar{T}-S(\bar{T})\right)}] \\
&= e^{-Ru}\mathbb{E}[e^{-Rc\bar{T}}\mathbb{E}[e^{RS(\bar{T})}]] \\
&= e^{-Ru}\mathbb{E}[e^{-Rc\bar{T}}M_{S(\bar{T})}(R)] \text{ from def. of } M_{S(\bar{T})}(R) \\
&= e^{-Ru}\mathbb{E}[e^{-Rc\bar{T}}P_{N(\bar{T})}(M_X(R))] \\
&= e^{-Ru}\mathbb{E}[e^{-Rc\bar{T}}e^{\lambda\bar{T}(M_X(R)-1)}] \text{ from } N(\bar{T}) \text{ Poisson} \\
&= e^{-Ru}\mathbb{E}[e^{-Rc\bar{T}}e^{\lambda\bar{T}(R(1+\theta)p_1)}] \text{ by hypothesis} \\
&= e^{-Ru}\mathbb{E}[1] \text{ because } c = (1+\theta)\lambda p_1 \\
&= e^{-Ru} \begin{pmatrix} \text{notice that both } \bar{T} \text{ and } \lambda \text{ have} \\ \text{disappeared from the equation;} \\ R \text{ depends only on } \theta \text{ and } P(x) \end{pmatrix} \\
\text{So } \lim_{t\to\infty} \mathbb{E}[e^{-RU(\bar{T})}] &= e^{-Ru} \text{ since that's what it equals for all value of } t.
\end{aligned}$$

On the other hand,

$$
\begin{aligned}
\mathbb{E}[e^{-RU(\bar{T})}] &= \psi(u,t)\mathbb{E}[e^{-RU(T)} \mid T \le t] + \varphi(u,t)\mathbb{E}[e^{-RU(t)} \mid T > t] \\
&\quad \big(\text{since } \bar{T} = T \wedge t \text{ and using def. of } \psi \text{and } \varphi\big) \\
\text{So } \lim_{t\to\infty} \mathbb{E}[e^{-RU(\bar{T})}] &= \psi(u)\mathbb{E}[e^{-RU(T)} \mid T < \infty] + \varphi(u) \lim_{t\to\infty} \mathbb{E}[e^{-RU(t)} \mid T > t] \\
\text{which means that } e^{-Ru} &= \psi(u)\mathbb{E}[e^{-RU(T)} \mid T < \infty] + \varphi(u) \lim_{t\to\infty} \mathbb{E}[e^{-RU(t)} \mid T > t]
\end{aligned}
$$

So, to prove (2) it only remains to prove either that $\lim_{t\to\infty} \mathbb{E}\left[e^{-RU(t)} \mid T > t\right] = 0$ or that $\varphi(u) = 0$. **When $\boldsymbol{\theta} > \mathbf{0}$**, so that $ct > \mathbb{E}[S(t)] = \lambda t p_1$, we will prove that $\lim_{t\to\infty} \mathbb{E}[e^{-RU(t)} \mid T > t] = 0$, as follows: For any arbitrary positive real number K,

$$
\begin{aligned}
&\lim_{t\to\infty} \mathbb{E}[e^{-RU(t)} \mid T > t] = \\
&= \lim_{t\to\infty} \left\{ \begin{array}{l} \frac{\Pr[U(t)\ge K>0]}{\Pr[U(t)\ge 0]}\mathbb{E}[e^{-RU(t)} \mid T > t, U(t) \ge K \ge 0] \\ +\frac{\Pr[K>U(t)\ge 0]}{\Pr[U(t)\ge 0]}\mathbb{E}[e^{-RU(t)} \mid T > t, K > U(t) \ge 0] \end{array} \right\} \\
&\le \lim_{t\to\infty} \left\{ 1 \cdot e^{-RK} + \frac{\Pr[K > U(t) \ge 0]}{\Pr[U(t) \ge K > 0]} \cdot 1 \right\} \\
&\quad \left(\begin{array}{c} \text{because } e^{-RU(t)} \le e^{-RK} \text{ when } U(t) \ge K \\ \text{and } \quad e^{-RU(t)} \le 1 \text{ when } U(t) \ge 0 \\ \text{and } \Pr[U(t) \ge K > 0] \quad \le \Pr[U(t) \ge 0] \end{array} \right) \\
&= e^{-RK} + \lim_{t\to\infty} \frac{\Pr[K > U(t) \ge 0]}{1 - \Pr[K > U(t) \ge 0]} \\
&= e^{-RK}, \text{ if we can prove that } \lim_{t\to\infty} \frac{\Pr[K > U(t) \ge 0]}{1 - \Pr[K > U(t) \ge 0]} = 0
\end{aligned}
$$

But that can be proved by

$$
\begin{aligned}
\lim_{t\to\infty} \Pr[K &> U(t) \geq 0] = \\
&= \lim_{t\to\infty} \Pr[u + \lambda t\theta p_1 \geq u + \lambda t\theta p_1 - U(t) > u + \lambda t\theta p_1 - K], \\
&\quad \text{since it's just subtracting from the same constant} \\
&\leq \lim_{t\to\infty} \Pr[u + \lambda t\theta p_1 - U(t) > u + \lambda t\theta p_1 - K], \\
&\quad \text{since eliminating one of the inequalities just adds more possibilities} \\
&\leq \lim_{t\to\infty} \Pr[|u + \lambda t\theta p_1 - U(t)| > u + \lambda t\theta p_1 - K], \\
&\quad \text{since the } | \quad | \text{ just adds more possibilites} \\
&= \lim_{t\to\infty} \Pr\left[|u + \lambda t\theta p_1 - U(t)| > \frac{u + \lambda t\theta p_1 - K}{\sqrt{\lambda t p_2}}\sqrt{\lambda t p_2}\right], \\
&\quad \text{since it's multiplying and dividing by same thing} \\
&\leq \lim_{t\to\infty} \frac{1}{\left(\frac{u+\lambda t\theta p_1 - K}{\sqrt{\lambda t p_2}}\right)^2}, \text{ by Chebychev's inequality} \\
&\quad \left(\begin{array}{c} \text{which says that } \Pr[|x-\mu| > \alpha\sigma] \leq \frac{1}{\alpha^2} \text{whenever } \sigma \text{ exists,} \\ \text{which you can prove easily if you don't remember it.} \end{array} \right) \\
&= 0 \text{ when you take the limit.}
\end{aligned}
$$

So we have $\lim_{t\to\infty} \mathbb{E}[e^{-RU(t)} \mid T > t] \leq e^{-RK}$ for any arbitrary positve real number K. But that means $\lim_{t\to\infty} \mathbb{E}[e^{-RU(t)} \mid T > t] = 0$, so that for $\theta > 0$ we now have proven (2).

When $\boldsymbol{\theta} = \mathbf{0}$, (2) will follow if we can prove that $\lim_{\theta\to 0} R = 0$. To do that, remember that for $\theta > 0$, R is the smallest positive solution of

$$R(1+\theta)p_1 = M_X(R) - 1$$

But the first two derivatives of $M_X(R) - 1$ at $R = 0$ are $p_1$and p_2 which are both bigger than 0. We also know that as $R \to \infty$, $M_X(R) = \mathbb{E}[e^{RX}]$ will go to ∞ exponentially fast (viz. faster than $R(1+\theta)p_1$.) So, as a function of R, $M_X(R) - 1$ starts out equal to $R(1+\theta)p_1$ at $R = 0$, has a smaller first derivative than $R(1+\theta)p_1$at $R = 0$ so it begins smaller than

$R(1+\theta)p_1$ for small R, and then must overtake $R(1+\theta)p_1$ on its way to ∞. The first R at which it equals $R(1+\theta)p_1$ is the value in (2). Now, if $\theta \to 0$ then the first derivative of $R(1+\theta)p_1 \to p_1 =$ the first derivative of $M_X(R)-1$. Moreover, the second derivative of $R(1+\theta)p_1 = 0$ while the second derivative of $M_X(R) - 1 = p_2 > 0$, so in the limit as $\theta \to 0$, $M_X(R) - 1$ exceeds $R(1+\theta)p_1$ for small R and therefore the limit as $\theta \to 0$ of the smallest R where they are equal will be 0. By continuity, then, the value for R in (2) when $\theta = 0$ must be $R = 0$. This completes the proof of the main theorem (2). ■

Note that the discussion in the last paragraph of the proof establishes that the solution R in fact exists whenever the moment-generating function $M_X(R)$ exists.

Now we can draw some quick conclusions that bear on risk management.

First, if the safety loading $\theta = 0$ then $\psi(u) = 1$ no matter what the value of the starting capital u might be, because $R = 0$ when $\theta = 0$. This result tells us that no matter how much capital you start with, the probability is 100% to be ruined at some point in time if your inflow is only equal to the expected value of your outflow. Profits aren't just something nice to have. With an expected profit of 0, ruin is certain no matter how much you start with. ("Certain" in the sense that the exceptions to the rule have probability 0, they constitute a set of measure 0.) This is a well known fact about stochastic processes, but it takes careful thinking to prove it. The proof above is probably no more difficult that any other (although this proof only works when the moment-generating function exists; the result is true regardless of the existence of the moment generating function.)

Second, if the safety loading $\theta > 0$ then there is some $R > 0$ with $\psi(u) < e^{-Ru}$ because $U(T) < 0$ (remember, T is the moment when ruin occurs) so $e^{-RU(T)} > 1$ and therefore $\mathbb{E}[e^{-RU(T)} \mid T < \infty] > 1$. This result tells us that as soon as we have a positive expected value of net inflow (after losses) then we can improve our chances of survival (i.e. limit the chance of ruin) by increasing our starting capital, and the improvement proceeds at least as fast as an exponential determined by θ and the distribution of the individual risk event costs (loss amounts) X.

Third, if we could understand the distribution of the strange random variable $U(T)$ well enough to make an exact calculation of the expected value in the denominator of (2), then we could calculate the probability of ruin $\psi(u)$ exactly. $U(T)$ is the random variable representing what the value of capital is at the moment you get ruined (assuming that ruin occurs). It is the measure of "how bad" the ruin is, how far under you go at the moment of ruin. It seems strange, but it is true, that this random variable holds the key to the probability of ruin occuring in the first place.

It turn out that we can in fact understand the random variable $U(T)$ and the understanding will come from our old friend the equilibrium distribution. **First consider only the case when $\mathbf{u = 0}$.** Since $U(T)$ is negative (moment of ruin, remember) make life easier by defining a random variable $L_1 = -U(T)$ that will be positive (the reason for the subscript will come out later). Now consider the amount of surplus $U(T^-)$ the instant before the event of ruin occurred. A more precise definition is $U(T^-) = \lim_{t \to T} U(t) \neq U(T)$ where the limit is from the left and the nonequality reflects the fact that at precisely the moment T the ruinous event occurs, not before. $U(T^-)$ is also a random variable and it's non-negative (remember T is the first time ruin occurs so that before T all the $U(t)$ are non-negative).

We know that $U(T^-) + L_1 = X_{N(T)}$ since $U(t)$ develops from events that cost a random amount X. Therefore, we can condition on $U(T^-)$ and write

$$\begin{aligned} f_{L_1}(y) &= \int_0^\infty f_{U(T^-)}(v) f_{X|X>v}(y+v) dv \\ &= \int_0^\infty f_{U(T^-)}(v) \frac{f_X(y+v)}{S_X(v)} dv \end{aligned}$$

Now comes the trick. Suppose I know that $f_{L_1}(y) = f_{U(T^-)}(y)$ for all y. Then

$$f_{L_1}(y) = \int_0^\infty f_{L_1}(v) \frac{f_X(y+v)}{S_X(v)} dv$$

This is just an integral equation for an unknown function f_{L_1} and remembering our friend the equilibrium distributition we can guess that a solution might be $f_{L_1}(y) = \frac{S_X(y)}{p_1}$. Try it out:

$$\begin{aligned}
\int_0^\infty f_{L_1}(v) \frac{f_X(y+v)}{S_X(v)} dv &= \int_0^\infty \frac{S_X(v)}{p_1} \frac{f_X(y+v)}{S_X(v)} dv \\
&= \int_0^\infty \frac{f_X(y+v)}{p_1} dv \\
&= \int_y^\infty \frac{f_X(v)}{p_1} dv \\
&= \frac{S_X(y)}{p_1} \\
&= f_{L_1}(y)
\end{aligned}$$

It works. If I can prove that $f_{L_1}(y) = f_{U(T^-)}(y)$ for all y, then I will know that $L_1 = -U(T)$ in the case when $u = 0$ follows the equilibrium distribution of the individual risk cost (or loss) variable X.

To prove that $f_{L_1}(y) = f_{U(T^-)}(y)$ for all y, consider what happens if the process continues after the ruin moment T. The distribution of $U(T+h)$ for any h will have mean $U(T) + h\theta\lambda p_1$ and variance $h\lambda p_2$. Using Chebychev's inequality again (you can do the details for yourself), as $h \to \infty$ the probability that $U(T+h) < 0$ goes to 0. That means that the probability is 1 that there is some time H at which $U(T+H) = 0$.

Now look at the process running backwards from $U(T+H)$ to $U(T)$, and put a minus sign in front of everything (i.e. look at it upside down as well as backwards.) It starts at 0 and has exponential $(1/\lambda)$ interarrival times for downward jumps that themselves have the distribution of X. The only fly in the ointment is that the very first interarrival time might not be exponential. The time from the last event while $U(t)$ was negative until the next event, after it crossed 0 into positive territory, is exponential. But we're interested in the time from that last event while $U(t)$ was negative until the time it crossed 0 into positive territory. But the memoryless property of the exponential makes that time also an exponential with the same parameter.

So, the upside-down backward process from $U(T+H)$ to $U(T)$ is identical to the original process from $U(0)$ to $U(T)$. In the upside-down backward process, $U(T^-)$ plays the same role that L_1 plays in the original process. So $U(T^-)$ and L_1 must have identical distributions, which completes the proof that $f_{L_1}(y) = \frac{S_X(y)}{p_1}$.■

Now (2) tells us that

$$\begin{aligned} \psi(0) &= \frac{e^{R \cdot 0}}{\mathbb{E}\left[e^{RL_1}\right]} \\ &= \frac{1}{M_{L_1}(R)} \end{aligned}$$

and since we know the density of L_1 we can calculate the moment generating function. With an integration by parts it gives

$$\begin{aligned} M_{L_1}(R) &= \frac{1}{p_1 R}\left[M_X(R) - 1\right] \\ &= \frac{1}{p_1 R}\left[R(1+\theta)p_1\right], \text{ by def. of } R \\ &= (1+\theta) \end{aligned}$$

Therefore $\psi(0) = \frac{1}{1+\theta}$ is the probability of ruin if you start with zero capital, $u = 0$.

But now we can get an expression for $\psi(u)$ for any starting capital u by conditioning on the amount that the surplus first drops below the starting point (which it must do if it is ever to reach 0):

$$\psi(u) = \psi(0)\left\{\int_0^u f_{L_1}(y)\psi(u-y)dy + \int_u^\infty f_{L_1}(y)dy\right\}$$