

Univerzita Karlova v Praze
Matematicko-fyzikální fakulta

DIPLOMOVÁ PRÁCE



Tereza Zachařová

Statistické metody analýzy složených bodových procesů

Katedra pravděpodobnosti a matematické statistiky

Vedoucí diplomové práce: Doc. Petr Wolf, CSc.

Studijní obor: pravděpodobnost, matematická statistika a ekonometrie

2008

Děkuji vedoucímu práce Doc. Petru Volfovi, CSc. za pokytnutou literaturu a hlavně za cenné rady náměty a připomínky, které mi pomohli při tvorbě této práce. Děkuji také Prof. RNDr. Tomáši Ciprovi, DrSc. za věnovanou knihu Teorie rizika v pojistné matematice, která mi velmi pomohla při prvním seznámení s tematikou a s českou terminologií.

Prohlašuji, že jsem svou diplomovou práci napsala samostatně a výhradně s použitím citovaných pramenů. Souhlasím se zapůjčováním práce.

V Praze dne 2. 8. 2008

Tereza Zachařová

Obsah

Použité značení a konvence	7
1 Úvod	9
2 Definice a základní pojmy	11
2.1 Náhodný proces, bodový proces a složený bodový proces	11
2.2 Konvoluce	14
2.3 Transformace náhodných veličin a jiné operace s rozděleními	15
2.4 Rozdělení s lehkými \times těžkými chvosty	17
2.4.1 Rozdělení s lehkými chvosty	17
2.4.2 Rozdělení s těžkými chvosty	18
2.4.3 Diskrétní rozdělení	22
2.5 Modely rizika v pojišťovnictví	24
3 Složená rozdělení a jejich vlastnosti	26
3.1 Složené rozdělení	26
3.1.1 Složené Poissonovo rozdělení	27
3.1.2 Složené geometrické rozdělení	28
3.2 Panjerův algoritmus a rekurentní metody	29
3.2.1 Panjerův algoritmus	29
3.2.2 Individuální model	31
3.3 Lundebergovy meze	32
3.3.1 Složené geometrické rozdělení	33
3.3.2 Obecnější složená rozdělení	34
3.4 Simulační studie I	36
3.4.1 Geometrické rozdělení	36
3.4.2 Poissonovo rozdělení	38
4 Proces rizika a ruinování	42
4.1 Proces rizika a s ním spjaté procesy	42

4.2	Ruinování a příbuzné pojmy	44
4.3	Pravděpodobnost ruinování	46
5	Složený Poissonův model	48
5.1	Homogenní Poissonův proces	48
5.2	Složený Poissonův proces	49
5.3	Pravděpodobnosti ruinování a pravděpodobnosti přežití	50
5.3.1	Pollaczek-Khinchinova formule	51
5.3.2	Laplaceova transformace	53
5.4	Asymptotické chování a approximace pravděpodobnosti ruinování	54
5.4.1	Lundebergův koeficient	54
5.4.2	Lundebergovy meze	57
5.4.3	Rozdělení s těžkými chvosty	58
5.4.4	Aproximace na základě momentů	62
5.4.5	Další approximační metody	65
5.5	Ruinování v konečném čase	67
5.5.1	Deterministické přírůstky	67
5.5.2	Exponenciální přírůstky	68
5.5.3	Sealova formule	69
5.6	Kdy a jak nastává ruinování	70
5.6.1	Lehké chvosty	72
5.6.2	Těžké chvosty	74
5.7	Přesné řešení	76
6	Simulace	79
6.1	Přímá Monte Carlo metoda	79
6.1.1	Metody redukce rozptylu	80
6.2	Simulace pomocí Pollaczek-Kinchinovy formule	82
6.3	Importance sampling pomocí Lundebergovy approximace	83
6.4	Generování náhodných veličin	84
6.5	Simulační studie II	85
7	Statistické metody odhadu parametrů a testy hypotéz pro složené bodové procesy	91
7.1	Odhad parametrů v homogenním Poissonově procesu	91
7.2	Testování hypotéz o homogenním Poissonově procesu	92
7.2.1	Testy nezávislosti na čase	93
7.2.2	Testy exponenciality	95
7.3	Odhady rozdělení přírůstků X_i	97
7.4	Testování hypotéz o přírůstcích X_i	98

7.4.1	Testy nezávislosti a náhodnosti přírůstků X_i	98
7.4.2	Nezávislost na $N(t)$	99
8	Závěr	100
Literatura		102
Rejstřík		104
Přílohy		106

Název práce: Statistické metody analýzy složených bodových procesů

Autor: Tereza Zachařová

Katedra: Katedra pravděpodobnosti a matematické statistiky

Vedoucí diplomové práce: Doc. Petr Wolf, CSc.

e-mail vedoucího: volf@utia.cas.cz

Abstrakt: Teorie rizika je část matematické statistiky zabývající se hlavně složenými bodovými procesy a procesy od nich odvozenými, tedy dvojně stochastickými procesy, kdy v náhodném čase nastává náhodný počet událostí, či je pozorován přírůstek náhodné velikosti. Tato práce si kade za hlavní cíl shrnout, přehledně usporádat a simulačně ověřit teoretické i approximativní výsledky z teorie rizika.

Podstatná část práce je věnována procesům rizika a tzv. pravděpodobnosti ruinování. Při podrobnějším studiu se omezujeme na složený Poissonův model, kde se příchody událostí řídí Poissonovým procesem a přírůstky jsou nezávislé stejně rozdělené. Výsledky se liší podle toho, zda přírůstky pochází z rozdělení s lehkými nebo těžkými chvosty, a proto uvádíme i základní přehled a dělení pravděpodobnostních rozdělení.

Klíčová slova: Složený bodový proces, proces rizika, Poissonův proces, ruinování, složená náhodná veličina, Monte Carlo metody.

Title: Statistical Methods of the Analysis of Compound Point Processes

Author: Tereza Zachařová

Department: Department of Probability and Mathematical Statistics

Supervisor: Doc. Petr Wolf, CSc.

Supervisor's e-mail address: volf@utia.cas.cz

Abstract: The risk theory studies mainly the behaviour of compound point processes and processes derived, where in random times increments of random size occur. The main objective of the present thesis is to collect in a systematic way the results on compound point processes and verify them by simulations.

The essential parts of this work deal with risk processes and so called ruin event. We concentrate mostly to the case of compound Poisson process, with independent and identically distributed increments. The results concerning both light-tailed and heavy-tailed distributions are presented. To this end, the classification of probability distributions along their tails is recalled, too.

Keywords: Compound point process, risk process, Poisson process, ruin probability, compound distribution, Monte-Carlo methods.

Použité značení a konvence

\mathbb{C}	množina komplexních čísel
\mathcal{D}	
$\mathbb{E}X$, $\text{var } X$	střední hodnota, rozptyl náhodné veličiny X , používáno také jako označení společné střední hodnoty či rozptylu iid náhodných veličin $X_i, i \in \mathbb{N}$
F_X, G_X	distribuční funkce náhodné veličiny X
$f_X(x)$	husota náhodné veličiny X
\bar{F}_X	chvost $\bar{F}_X(x) = 1 - F_X(x) = P(X > x)$
$F_1 * F_2, F^{*k}$	konvoluce distribučních funkcí, k -tá konvoluce distribuční funkce F (též k -tá konvoluční mocnina)
F_X^s	integrovaný chvost rozdělení, $F_X^s = \frac{1}{\mathbb{E}X} \int_0^x F(y)dy$
$\hat{g}(s)$	pravděpodobnostní vytvořující funkce
$I\{A\}$	indikátor množiny, jevu A
iid	nezávislé stejně rozdělené náhodné veličiny (independent identically distributed)
\mathcal{K}	rozdělení s těžkými chvosty
\mathcal{L}	množina distribučních funkcí, kde $\lim_{x \rightarrow \infty} \frac{\bar{F}(x-y)}{\bar{F}} = 1 \forall y > 0$
$\hat{l}(s)$	Laplace-Stieltjesova transformace
$\hat{m}(s)$	momentová vytvořující funkce
\mathbb{N}, \mathbb{N}_0	množina přirozených čísel, $\mathbb{N} \cup \{0\}$
P	pravděpodobnost
p	intenzita pojistného
$\mathbb{R}, \overline{\mathbb{R}}, \mathbb{R}_+$	množina reálných čísel, $\mathbb{R} \cup \{\infty\}$, interval $[0, \infty)$
$\mathcal{R}_0, \mathcal{R}_\alpha, \mathcal{R}$	pomalu se měnící funkce, regulárně se měnící funkce, množina distribučních funkcí, pro které $\bar{F} \in \mathcal{R}_{-\alpha}, \alpha \geq 0$
\mathcal{S}	subexponenciální rozdělení
u	počáteční rezerva
$u(\alpha)$	kritická hodnota normálního rozdělení $N(0, 1)$ na hladině α
\mathbb{Z}	množina celých čísel
γ	Lundebergův koeficient

λ	intenzita Poissonova procesu
μ	střední hodnota přírůstků ve složeném Poissonovu modelu
$\tau(u)$	čas ruinování
$\psi(s)$	charakteristická funkce
$\psi(u), \psi(u, x)$	pravděpodobnost ruinování, pravděpodobnost ruinování do času x
$ \cdot $	absolutní hodnota
\wedge	min
\vee	max
$[x]$	dolní celá část x $[x] = \max\{n \in \mathbb{N} : n \leq x\}$
\leq_{st}	stochasticky menší (viz definice 3.21)
\sim	používáno pouze ve smyslu, že náhodná veličina má určité rozdělení
\approx	$a(x) \approx b(x) \iff \lim_{x \rightarrow \infty} \frac{a(x)}{b(x)} = 1$
\cdot	používáno ve smyslu approximace, ale nikoli tak přesná, poměr veličiny vlevo a vpravo už není v limitě jedna, např. approximace nekonečného součtu několika prvními členy

Dále používáme konvence:

$$\inf \emptyset = \infty$$

$$\sum_0^0 a_i = 0$$

$\mathbb{E}X, \text{var } X$ používáme jako označení společné střední hodnoty iid náhodných veličin $X_i, i \in \mathbb{N}$

Kapitola 1

Úvod

V této práci se budeme zabývat složenými bodovými procesy, tj. procesy $\{C(t), t \in T\}$, kde $T \subseteq [0, \infty)$ je nějaká indexová množina a $C(t)$ je pro každé $t \in T$ složená náhodná veličina. Složené náhodné veličiny vznikají kumulací náhodného počtu náhodných veličin, tedy $C(t) = \sum_{i=1}^{N(t)} X_i$, kde $N(t)$ a $X_i, i \in \mathbb{N}$ jsou náhodné veličiny. My se omezíme pouze na případ, kdy X_i jsou nezávislé stejně rozdělené a navíc nezávislé na $N(t)$. $N(t)$ musí být zřejmě diskrétní, ale X_i mohou být libovolné, ovšem obvykle se uvažují nezáporné.

Pro jednoduchou představu uvedeme příklad z pojišťovnictví, kdy do pojišťovny přicházejí během dne (měsíce, roku) hlášení o pojistných událostech a z nich plyne pro pojišťovnu nějaké pojistné plnění (X_i) a my se zajímáme jaká celková suma bude potřeba k uhrazení všech závazků za celý den (měsíc, rok). Je zřejmé, že předem nevíme kolik pojistných událostí ($N(t)$) nastane a jaká bude jejich výše (X_i).

Spíše než samotný složený bodový proces budeme analyzovat tzv. proces rizikových rezerv $\{R(t), t \in T\}$, pro $R(t) = u + p(t) - C(t)$, kde $C(t)$ je složený bodový proces, u je nějaká počáteční rezerva, $p(t)$ je příjem do času t . V našem příkladu s pojišťovnou je to pojistné inkasované do času t . Pokud zkoumáme pravděpodobnost, že $\{R(t)\}$ klesne do záporných hodnot je to ekvivalentní situaci, kdy $\{C(t)\}$ překročí určitou mez. Tento jev se obvykle nazývá ruinování.

Takovýto model lze použít k popisu mnoha situací z reálného života. Kromě zmíněné pojišťovací praxe, například také v ekonomii, biologii, medicíně, ale i v oblastech jako je životní prostředí.

Pro zjednodušení vyjadřování budeme v práci v některých případech používat terminologii z pojišťovnictví a náhodné veličiny X_i budeme nazývat nejčastěji přírůstky, ale také pojistné události a množinu T budeme interpretovat jako čas (ať už spojity nebo diskrétní), i když její faktický význam

může být různý. Dále používáme $\mathbb{E}X$ jako označení společné střední hodnoty iid náhodných veličin X_i a analogicky i pro dalsí iid náhodné veličiny a všechny momenty.

Cílem práce je především shrnout a přehledně uspořádat známé výsledky pro složené náhodné veličiny a složené bodové procesy, které jsou sice popsané v rozsáhlé literatuře, ale tato literatura není příliš systematická a některé výsledky jsou odvozené pro jiné partie matematiky, jako třeba náhodné procházky či teorie front, a jejich aplikace v teorii rizika nemusí být hned zřejmá. Některé z teoretických a approximativních výsledků ověřujeme simulačně a zkoumáme jejich kvalitu. Tato práce má ambice být dostačujícím textem pro první seznámení s teorií rizika, pro každého kdo má základní znalosti teorie pravděpodobnosti a matematické statistiky. Snahou bylo vytvořit text přehledný, ve kterém se bude snadno vyhledávat, proto je součástí rejstřík pro diplomové práce poněkud netypický. U některých pojmu uvádíme v závorce jejich anglické ekvivalenty pro snadnější návaznost na anglicky psanou literaturu.

Nejprve v práci připomeneme některé definice související s náhodnými procesy, různé transformace náhodných veličin a také definice některých pravděpodobnostních rozdělení, které jsou dále v textu používány (kapitola 2). Zmiňujeme i základní modely, které se používají k popisu pojistných událostí v pojistné matematice. Ve třetí kapitole se budeme zabývat složenými náhodnými veličinami. Největší pozornost je věnována složenému geometrickému a složenému Poissonovu rozdělení, neboť právě pro tato dvě rozdělení existují nejzajímavější teoretické výsledky. Uvedeme několik způsobů jak spočítat rozdělení složené náhodné veličiny nebo ho alespoň omezit. Kvalitu uvedených omezení zkoumáme simulačně. Pomocí simulací se také snažíme udělat si obrázek o závislosti chování složené náhodné veličiny na rozdělení přírušků a dalších parametrech.

Čtvrtá kapitola se již zabývá procesy rizika. Definujeme v ní proces rizika a další příbuzné procesy. Dále definujeme pojem ruinování a odvodíme pravděpodobnost ruinování v nejobecnějším případě. V páté kapitole zavedeme složený Poissonův model a podrobně se budeme zabývat pravděpodobností ruinování v tomto modelu. Odvodíme různé formule a omezení pro pravděpodobnost ruinování v závislosti na typu přírušků. V této kapitole také uvádíme situace, v jakých umíme určit pravděpodobnost ruinování přesně.

V šesté kapitole uvedeme několik postupů pro simulaci pravděpodobnosti ruinování v nekonečném časovém horizontu a prozkoumáme chování pravděpodobnosti ruinování a kvalitu omezení z předchozí kapitoly simulačně. Závěrečná sedmá kapitola je věnována některým statistickým metodám pro analýzu složených bodových procesů.

Kapitola 2

Definice a základní pojmy

Na začátek si připomeneme některé základní definice a věty, týkající se procesů a součtů náhodných veličin. Tato kapitola má také sloužit jako upřesnění požívané varianty definice v případech, kdy se vyskytuje definice různě modifikované. Všechny pojmy najdeme v základních učebnicích teorie pravděpodobnosti, statistiky a náhodných procesů, my vycházíme hlavně z [Prášková and Lachout, 2005], [Jacobsen, 2006], [Rolski et al., 1998], [Anděl, 2005].

2.1 Náhodný proces, bodový proces a složený bodový proces

V této části si definujeme náhodný proces a některé jeho charakteristiky, bodový a čítací proces a nakonec složený bodový proces.

Definice 2.1. Nechť (Ω, \mathcal{A}, P) je pravděpodobnostní prostor a nechť $T \subseteq \mathbb{R}$. Potom soubor náhodných veličin $\{X(t), t \in T\}$ definovaných na (Ω, \mathcal{A}, P) nazýváme **náhodný proces**.

Nebude-li hrozit nedorozumění, budeme používat zkrácený zápis $\{X(t)\}$, kde množinu T budeme vynechávat. V případě, že je $T \subseteq \mathbb{Z}$ používáme častěji zápis pomocí indexů $\{X_t, t \in T\}$.

Definice 2.2. Nechť $\{X(t), t \in T\}$ je náhodný proces takový, že pro každé $t \in T$ existuje střední hodnota $\mathbb{E}X(t)$. Potom funkce $\mu_t = \mathbb{E}X(t)$, definovaná na T se nazývá **střední hodnota** náhodného procesu $\{X(t)\}$. Jestliže $\mathbb{E}|X(t)|^2 < \infty$ pro všechna $t \in T$, potom funkce dvou proměnných definovaná na $T \times T$ předpisem $R(s, t) = \mathbb{E}(X(s) - \mu_s)(X(t) - \mu_t)$ se nazývá **autokovarianční funkce** procesu $\{X(t)\}$. Hodnota $R(t, t)$ se nazývá **rozptyl** procesu $\{X(t)\}$ v čase t .

Definice 2.3. Nechť $T = \{T_n, n \in \mathbb{N}\}$ je posloupnost náhodných veličin definovaných na (Ω, \mathcal{A}, P) , které nabývají hodnot z intervalu $[0, \infty]$. Pak T nazýváme **jednoduchý bodový proces** pokud splňuje:

1. $P(0 < T_1 \leq T_2 \leq \dots) = 1$,
2. pro $n \geq 1$ je $P(T_n < T_{n+1}, T_n < \infty) = P(T_n < \infty)$,
3. $P(\lim_{n \rightarrow \infty} T_n = \infty) = 1$.

Jednoduchý bodový proces tedy definujeme jako skoro jistě rostoucí posloupnost kladných, ne nutně konečných náhodných veličin (bod 1. definice 2.3). Tato posloupnost je rostoucí dokud jsou T_n konečné (bod 2. definice 2.3) přičemž jejich limita je skoro jistě nekonečná (bod 3. definice 2.3). Pokud je náhodná veličina T_n konečná, interpretujeme ji jako čas, kdy nastává n -tá událost. Pokud je $T_n = \infty$ znamená to, že nastává méně než n událostí. Podle této definice žádná událost nemůže nastat v čase 0 (i když jako časovou osu používáme interval $[0, \infty)$).

V teorii bodových procesů se obvykle používá obecnější definice, ale pro naše účely je tato dostačující a názornější.

Chceme-li počítat různé události a rozlišovat je, dáme každé nějakou značku. Dostaváme tak značkovaný bodový proces. Celou teorii složených bodových procesů můžeme řešit v obecnějším kontextu značkovaných bodových procesů. To sice není záměrem této práce, ale přesto nám dovolte uvést definici značkovaného bodového procesu.

Nejprve si definujeme prostor značek jako měřitelný prostor (E, \mathcal{E}) . Definujeme si značku ∇ , která bude popisovat událost, která má v konečném čase nulovou pravděpodobnost a jejíž význam je vidět v následující definici. A dále označme $\bar{E} = E \cup \{\nabla\}$ a $\bar{\mathcal{E}} = \sigma(\mathcal{E}, \{\nabla\})$ označme σ -algebru generovanou \bar{E} .

Definice 2.4. Značkovaný bodový proces s prostorem značek E definujeme jako posloupnost $(T, Y) = (\{T_n, n \in \mathbb{N}\}, \{Y_n, n \in \mathbb{N}\})$ náhodných veličin T_n s hodnotami v $[0, \infty]$ a náhodných veličin Y_n s hodnotami v \bar{E} takovou, že $T = \{T_n\}$ je jednoduchý bodový proces a platí:

1. $P(Y_n \in E, T_n < \infty) = P(T_n < \infty)$,
2. pro $n \geq 1$ je $P(Y_n = \nabla, T_n = \infty) = P(T_n = \infty)$.

Tato definice nám dává posloupnost časových okamžiků, kdy nastává nějaká událost a díky značkám můžeme rozlišovat různé typy událostí. Díky značce ∇ můžeme definovat Y_n pro všechna n dokonce i když n -tá událost v konečném čase nenastává ($T_n = \infty$).

Chceme-li složený bodový proces (definováno dále, viz definice 2.6) převést na značkovaný bodový proces, je to velmi jednoduché. T_n budou časové okamžiky, kdy nastává událost a značky nechť označují velikost přírůstku X_n v čase T_n .

Definice 2.5. Nechť $T = \{T_n, n \in \mathbb{N}\}$ je jednoduchý bodový proces. Pak definujeme **čítací proces** $N = \{N(t), t \geq 0\}$ příslušný bodovému procesu T jako:

$$N(t) = \sum_{n=1}^{\infty} \mathbb{I}\{T_n \leq t\}.$$

Platí $N(0) = 0$ a $N(t)$ udává počet událostí v intervalu $[0, t]$. Zřejmě $N(t)$ nabývá pouze hodnot z \mathbb{N}_0 . Předpoklad, že T_n jsou (striktně) rostoucí dokud jsou konečné zaručuje, že funkce $t \mapsto N(t)$ je skokovitá se skokem velikosti 1.

Při definici můžeme postupovat také opačně a jednoduchý bodový proces identifikovat pomocí příslušného čítacího procesu, pak platí:

$$T_n = \inf\{t \geq 0 : N(t) = n\},$$

kde tradičně definujeme $\inf \emptyset = \infty$. Dále platí rovnost jevů

$$(T_n \leq t) = (N(t) \geq n).$$

Složený bodový proces můžeme formálně definovat různě, ale interpretace jako kumulace náhodných přírůstků v náhodných časech $0 < T_1 < T_2 < \dots$ zůstává vždy v platnosti.

Definice 2.6. (Složený bodový proces)

Položme $C(0) = 0$ a pro dané $t \in (0, \infty)$ definujme

$$C(t) = \sum_{i=1}^{\infty} X_i \cdot \mathbb{I}\{T_i \leq t\},$$

kde X_i jsou náhodné veličiny a $T = \{T_i, i \in \mathbb{N}\}$ je jednoduchý bodový proces, potom $\{C(t), t \in [0, \infty]\}$ nazýváme složený bodový proces s přírůstky X_i .

Nejčastější bývá časová interpretace, proto $T_i, i \in \mathbb{N}$ nabývají hodnot z intervalu $[0, \infty)$ nebo nějaké jeho podmnožiny, např. hodnot $0, 1, 2, \dots$, v takovém případě mluvíme o složeném bodovém procesu s diskrétním časem. Ovšem t nemusí mít jen časovou interpretaci. Může jít např. o pořadí nebo vzdálenost sledovaných jednotek. Přírůstky X_i se často uvažují kladné, ze

smyslu věci (např. označují vzniklou škodu nebo naopak vyprodukovaný zisk).

Ekvivalentní definici složeného bodového procesu dostaneme pomocí čítacího procesu $\{N(t), t \geq 0\}$ příslušného bodovému procesu $\{T_n, n \in \mathbb{N}\}$. $C(t)$ pro $t > 0$ pak mají tvar:

$$C(t) = \sum_{i=1}^{N(t)} X_i.$$

Tento zápis je dále častěji používaný.

2.2 Konvoluce

Věty o konvoluci řeší rozdělení součtu dvou či více náhodných veličin.

Věta 2.7. (*O konvoluci*)

Nechť náhodné veličiny X a Y jsou nezávislé a nechť mají distribuční funkce F_X a F_Y , pak náhodná veličina $Z = X + Y$ má distribuční funkci

$$F_Z(z) = \int_{-\infty}^{\infty} F_Y(z - x) dF_X(x).$$

Důkaz. [Anděl, 2005] □

Funkci F_Z z předchozí věty nazýváme **konvoluce** distribučních funkcí F_X a F_Y a značíme $F_Z = F_X * F_Y$. Platí $F_X * F_Y = F_Y * F_X$.

Věta 2.8. Nechť X a Y jsou nezávislé diskrétní náhodné veličiny. Označme $p_k = P(X = k)$, $q_k = P(Y = k)$, $k \in \mathbb{Z}$. Pak náhodná veličina $Z = X + Y$ nabývá rovněž jen celočíselných hodnot a platí

$$P(Z = k) = \sum_{i=-\infty}^{\infty} p_i q_{k-i}.$$

Důkaz. [Anděl, 2005] □

Nám budou stačit tyto dva případy, ale existují věty pro různé speciální situace. Zájemce odkazujeme na [Anděl, 2005] paragraf 3.4.

2.3 Transformace náhodných veličin a jiné operace s rozděleními

V této části připomeneme definice různých transformací náhodných veličin, které se využívají při výpočtech rozdělení složených náhodných veličin nebo při řešení problému ruinování, o kterém bude řeč později.

Definice 2.9. Bud' X náhodná veličina pak definujeme:

1. **momentovou vytvořující funkci** $\hat{m} : \mathbb{R} \rightarrow \overline{\mathbb{R}}$ náhodné veličiny X danou vztahem $\hat{m}(s) = \mathbb{E}e^{sX}$,
2. **Laplace-Stieltjesovu transformaci** $\hat{l} : \mathbb{R} \rightarrow \overline{\mathbb{R}}$ náhodné veličiny X danou vztahem $\hat{l}(s) = \mathbb{E}e^{-sX}$,
3. **charakteristickou funkci** $\psi : \mathbb{R} \rightarrow \mathbb{C}$ reálné náhodné veličiny X danou vztahem $\psi(s) = \mathbb{E}e^{isX}$.

Pokud má X diskrétní rozdělení pak definujeme

4. **pravděpodobnostní vytvořující funkci** $\hat{g} : [-1, 1] \rightarrow \overline{\mathbb{R}}$ náhodné veličiny X vztahem $\hat{g}(s) = \sum_{k=0}^{\infty} P(X = k)s^k = \mathbb{E}s^X$.

Mimo Laplace-Stieljesovy transformace náhodné veličiny uvažujeme někdy také Laplaceovu transformaci funkce.

Definice 2.10. Bud' $c : \mathbb{R} \rightarrow \mathbb{R}_+$ funkce. Pak její **Laplaceova transformace** je $\hat{L}(s) = \int_{-\infty}^{\infty} e^{-sx}c(x)dx$.

Pokud náhodná veličina X má absolutně spojité rozdělení s hustotou f pak její Laplace-Stieljesova transformace je rovna Laplaceově transformaci funkce f .

Pro funkce z definice 2.9 zřejmě platí $\hat{m}(s) = \hat{l}(-s) = \hat{g}(e^s)$ a $\psi(s) = \hat{m}(is) = \hat{l}(-is) = \hat{g}(e^{is})$ jsou-li všechny funkce definované.

Věta 2.11. Pokud n -té momenty náhodné veličiny $|X|$ jsou konečné, pak existují n -té derivace $\hat{m}^{(n)}(0), \hat{l}^{(n)}(0)$ a platí $\mathbb{E}X^n = \hat{m}^{(n)}(0) = (-1)^n \hat{l}^{(n)}(0) = (-i)^n \psi^{(n)}(0)$.

Důkaz. V [Lachout, 2004] můžeme nalézt důkaz pro charakteristickou funkci, zbytek dostaneme snadno ze vztahu mezi jednotlivými funkcemi. \square

Věta 2.12. *Mezi rozdelením náhodné veličiny a její charakteristickou funkcí je vzájemně jednoznačný vztah.*

Důkaz. [Lachout, 2004] □

Pro momentovou vytvořující funkci, Laplace-Stieltjesovu transformaci a pravděpodobnostní vytvořující funkci náhodné veličiny je potřeba zavést další předpoklady, aby vztah mezi rozdelením náhodné veličiny a příslušnou transformací byl vzájemně jednoznačný.

Definice 2.13. Buděte X_1, X_2, \dots reálné náhodné veličiny a F_1, F_2, \dots jim příslušné distribuční funkce a $\{p_n, n \in \mathbb{N}\}$ nechť je pravděpodobnostní funkce. Potom rozdelení dané distribuční funkcí $F = \sum_{n=1}^{\infty} p_n F_n$ nazýváme **směs** distribucí F_1, F_2, \dots daná pravděpodobnostmi $\{p_n, n \in \mathbb{N}\}$.

Definice 2.14. Buď X náhodná veličina s distribuční funkcí F , pak definujeme následující operace s rozdelením X .

1. **Oříznutí** (truncation). Bud' C nějaká borelovská podmnožina \mathbb{R} , pak definujeme oříznuté rozdelení

$$F^C(B) = P(X \in B | X \notin C), \quad B \in \mathcal{B}(\mathbb{R}).$$

2. **Modifikace.** Je-li X diskrétní náhodná veličina s pravděpodobnostmi $\{p_k = P(X = k), k \in \mathbb{N}_0\}$ a $0 < \theta < 1$, pak θ -modifikací F rozumíme $G = (1 - \theta)\delta_0 + \theta F$ a pro Y s distribuční funkcí G platí

$$P(Y = k) = \begin{cases} 1 - \theta + \theta p_0 & k = 0 \\ \theta p_k & k \geq 1. \end{cases}$$

Oříznuté rozdelení F^C je podmíněné rozdelení dané restrikcí hodnot X na doplněk množiny C . Modifikace je v některých případech inverzní operací k nulovému oříznutí.

Příklad. Pokud X má geometrické rozdelení $X \sim Ge(p)$ a uvažujeme jeho nulové oříznutí pak $P(X = k | X \geq 1) = (1 - p)p^{k-1}$, $k = 1, 2, \dots$ a v dalším textu ho označujeme jako $TG(p)$. Naopak p -modifikace $TG(p)$ je rovna $Ge(p)$.

2.4 Rozdělení s lehkými × těžkými chvosty

Některé výsledky týkající se složených bodových procesů platí jen pro rozdělení s lehkými chvosty. Uvedeme proto v jakém smyslu je tento pojem používán a také která rozdělení to splňují. Obecně můžeme říci, že rozdělení s těžkými chvosty bývají problematické, a to nejen při analýze složených bodových procesů, ale už existence momentů není zaručena, problematická je i platnost limitních vět. K podrobnější klasifikaci rozdělení s těžkými chvosty se dostaneme v kapitole 5.

2.4.1 Rozdělení s lehkými chvosty

Definice 2.15. O rozdělení náhodné veličiny X říkáme, že má **lehké chvosty** (light-tailed distribution) pokud $\bar{F}_X(x) = P(X > x) \leq ae^{-bx}$ pro každé $x > 0$, kde $a, b > 0$ jsou konstanty.

Ekvivalentní definice rozdělení s lehkými chvosty je pomocí momentové vytvářející funkce. Rozdělení s lehkými chvosty jsou taková, pro která existuje $s > 0$ takové, že $\hat{m}_X(s) < \infty$.

Mezi rozdělení s lehkými chvosty používané jako rozdělení přírůstků v pojišťovací praxi patří např.

- **Exponenciální rozdělení** $Exp(\lambda)$ s hustotou

$$f(x) = \lambda e^{-\lambda x}$$

pro $x > 0, \lambda > 0$. Exponenciálně rozdělené přírůstky mají obvykle velmi příjemné vlastnosti a v teorii rizika, jako i v jiných oblastech matematiky, bývá nejsnazší pro ně najít nějaké řešení. Často jsou jediným typem přírůstků, pro které explicitní řešení vůbec existuje.

Exponenciální rozdělení bývá nazýváno rozdělení bez paměti, neboť pro exponenciálně rozdělené X platí $P(X > u+v|X > v) = P(X > u)$ pro každé $u, v \geq 0$.

$$\mathbb{E}X = \frac{1}{\lambda}, \quad \text{var } X = \frac{1}{\lambda^2}, \quad \hat{m}(s) = \frac{\lambda}{\lambda - s}, \quad s < \lambda.$$

- **Normální rozdělení** $N(\mu, \sigma^2)$ s hustotou

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp \left\{ -\frac{(x-\mu)^2}{2\sigma^2} \right\}$$

pro $x \in \mathbb{R}, \mu \in \mathbb{R}$ a $\sigma > 0$. Normální rozdělení se jako rozdělení příruček nepoužívá, protože nesplňuje podmínu nezápornosti. Ale díky centrální limitní větě se používá jako rozdělení průměrů.

$$\mathbb{E}X = \mu, \quad \text{var } X = \sigma^2 \quad \hat{m}(s) = \exp \left\{ \mu s + \frac{1}{2} \sigma^2 s^2 \right\}.$$

- **Gamma rozdělení** $\Gamma(a, b)$ s hustotou

$$f(x) = \frac{a^b}{\Gamma(b)} x^{b-1} e^{-ax}$$

pro $x \geq 0, a > 0, b > 0$.

$$\mathbb{E}X = \frac{b}{a}, \quad \text{var } X = \frac{b}{a^2} \quad \hat{m}(s) = \left(\frac{a}{a-s} \right)^b, \quad s < a.$$

Rozdělení $\Gamma(\lambda, 1)$ je totožné s rozdělením $Exp(\lambda)$.

- **Erlangovo rozdělení** $Erl(a, n)$ je speciální případ gamma rozdělení $\Gamma(a, b)$, kdy parametr b je přirozené číslo.
- **Beta rozdělení** $B(a, b, c)$ s hustotou

$$f(x) = \frac{1}{\mathbf{B}(a, b)} \frac{x^{a-1}(c-x)^{b-1}}{c^{a+b-1}}$$

pro $0 < x < c$ a $a > 0, b > 0, c > 0$, kde $\mathbf{B}(a, b)$ označuje beta funkci.

$$\mathbb{E}X = \frac{a}{a+b}, \quad \text{var } X = \frac{ab}{(a+b)^2(a+b+1)}.$$

2.4.2 Rozdělení s těžkými chvosty

Definice 2.16. O rozdělení náhodné veličiny X říkáme, že má **těžké chvosty** (heavy-tailed distribution) pokud její momentová vytvořující funkce $\hat{m}_X(s) = \infty$ pro každé $s > 0$.

Označme si $\Lambda(x) = -\log \bar{F}(x)$ v analýze přežití známé jako kumulovaná riziková funkce a definujme

$$\alpha_F = \limsup_{x \rightarrow \infty} \frac{\Lambda(x)}{x}.$$

Věta 2.17. Pokud $\alpha_F = 0$, pak F má těžké chvosty.

Důkaz. [Rolski et al., 1998] □

Pro rozdělení s těžkými chvosty platí

$$\lim_{x \rightarrow \infty} e^{sx} \bar{F}(x) = \infty. \quad (2.1)$$

Mezi rozdělení s těžkými chvosty používané jako rozdělení přírůstků v pojišťovací praxi patří např.

- **Logaritmicko-normální rozdělení $LN(\mu, \sigma^2)$** s hustotou

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma x} \exp \left\{ -\frac{(\log x - \mu)^2}{2\sigma^2} \right\}$$

pro $x > 0, \mu \in \mathbb{R}, \sigma > 0$. Logaritmicko-normální rozdělení má náhodná veličina e^X , kde X má normální rozdělení.

$$\mathbb{E}X = \exp \left\{ \mu + \frac{\sigma^2}{2} \right\}, \quad \text{var } X = e^{2\mu+\sigma^2} (e^{\sigma^2} - 1).$$

- **Weibullovo rozdělení $W(a, b)$** s hustotou

$$f(x) = abx^{b-1}e^{-ax^b}$$

pro $x > 0, a > 0, b > 0$. Weibullovo rozdělení $W(\lambda, 1)$ je totožné s exponenciálním $Exp(\lambda)$. Weibullovo rozdělení má těžké chvosty pouze pro $b < 1$, jinak je to rozdělení s lehkými chvosty.

$$\mathbb{E}X = a^{-\frac{1}{b}} \Gamma \left(\frac{1}{b} + 1 \right), \quad \text{var } X = a^{-\frac{2}{b}} \left(\Gamma \left(\frac{2}{b} + 1 \right) - \left(\Gamma \left(\frac{1}{b} + 1 \right) \right)^2 \right).$$

- **Paretovo rozdělení $Par(a, b)$** s hustotou

$$f(x) = \frac{a}{b} \left(\frac{b}{x} \right)^{a+1}$$

pro $x \geq b, a > 0, b > 0$. Chvost Paretova rozdělení je $\bar{F}(x) = (\frac{b}{x})^a$, $x \geq b$.

$$\mathbb{E}X = \frac{ab}{a-1}, \quad a > 1, \quad \text{var } X = \frac{ab^2}{(a-1)^2(a-2)}, \quad a > 2.$$

- **Cauchyovo rozdělení** $C(a, b)$ s hustotou

$$f(x) = \frac{1}{\pi} \frac{b}{b^2 + (x - a)^2}$$

pro $x \in \mathbb{R}, a \in \mathbb{R}, b > 0$. Pro Cauchyovo rozdělení neexistují momenty (dokonce ani střední hodnota). Nejznámější reprezentace Cauchyova rozdělení je podél dvou normálních rozdělení ($X \sim N(0, 1), Y \sim N(0, 1)$, pak $\frac{X}{Y} \sim C(0, 1)$). Jako rozdělení přírůstků se používá jednostranná varianta Cauchyova rozdělení s hustotou $f(x) = \frac{2}{\pi} \frac{b}{b^2 + (x-a)^2}, x \in [a, \infty), a \in \mathbb{R}, b > 0$.

Definice rozdělení s těžkými chvosty je příliš obecná na to, aby umožňovala nějaké obecně platné netriviální výsledky. Proto se používá jemnější dělení a z rozdělení z těžkými chvosty se ještě vydělují tzv. subexponenciální rozdělení.

Definice 2.18. Říkáme, že rozdělení nezáporné náhodné veličiny X s distribuční funkcí F je **subexponenciální** pokud

$$\lim_{x \rightarrow \infty} \frac{1 - F^{*2}(x)}{1 - F(x)} = 2 \quad (2.2)$$

a množinu distribučních funkcí, které přísluší nějakému subexponenciálnímu rozdělení označujeme \mathcal{S} .

Jako triviální důsledek (2.2) dostáváme $\bar{F}(x) > 0$ pro každé $x \geq 0$. To samozřejmě platí i pro některá rozdělení s lehkými chvosty (např. exponenciální rozdělení), ale dává nám to nutnou podmínu subexponentiality.

Dále platí

$$\begin{aligned} \frac{\bar{F}^{*2}(x)}{\bar{F}(x)} &= \frac{1}{\bar{F}(x)} \int_0^\infty \bar{F}(x-y) d\bar{F}(y) \\ &= \frac{1}{\bar{F}(x)} \left(\int_0^x \bar{F}(x-y) d\bar{F}(y) + \int_x^\infty \bar{F}(x-y) d\bar{F}(y) \right) \end{aligned}$$

a protože pro $y > x$ máme $\bar{F}(x-y) = 1$ snadno dostáváme rovnost

$$\frac{\bar{F}^{*2}(x)}{\bar{F}(x)} = 1 + \int_0^x \frac{\bar{F}(x-y)}{\bar{F}(x)} d\bar{F}(y),$$

z které dostaneme

$$\liminf_{x \rightarrow \infty} \frac{\overline{F}^{*2}(x)}{\overline{F}(x)} \geq 2.$$

Tedy subexponenciální rozdělení jsou taková, která dosahují minimální možné hodnoty podílu $\frac{\overline{F}^{*2}(x)}{\overline{F}(x)}$.

Věta 2.19. *Pro každé $F \in \mathcal{S}$ a $y > 0$ platí*

$$\lim_{x \rightarrow \infty} \frac{\overline{F}(x-y)}{\overline{F}(x)} = 1.$$

Důkaz. [Rolski et al., 1998] □

Věta 2.20. *Je-li rozdělení náhodné veličiny X subexponenciální, pak X má těžké chvosty.*

Důkaz. [Rolski et al., 1998] □

Věta 2.21. *Nechť X je nezáporná náhodná veličina, pak X je subexponenciální právě tehdy, když pro každé $n = 2, 3, \dots$ platí*

$$\lim_{x \rightarrow \infty} \frac{\overline{F}^{*n}(x)}{\overline{F}(x)} = n.$$

Důkaz. [Rolski et al., 1998] □

Věta 2.22. *Logaritmicko-normální rozdělení $LN(\mu, \sigma^2)$, Weibullovo rozdělení $W(a, b)$ s parametrem $b < 1$ a Paretovo rozdělení $Par(a, b)$ jsou subexponenciální.*

Důkaz. [Rolski et al., 1998], k důkazu jsou potřeba některé pokročilejší partie o rozděleních s těžkými chvosty, které nechceme v tomto textu prezentovat a čtenáře odkazujeme na kapitolu 2.5 citované literatury. □

Další zajímavé výsledky se dají odvodit pro tzv. pomalu se měnící (slowly varying) distribuční funkce.

Definice 2.23. Říkáme, že kladná lebesgueovsky měřitelná funkce L na $(0, \infty)$ je pomalu se měnící v ∞ a značíme $L \in \mathcal{R}_0$, pokud

$$\lim_{x \rightarrow \infty} \frac{L(tx)}{L(x)} = 1, \quad \forall t > 0.$$

Počítáme-li konvoluce distribučních funkcí s pomalu se měnícími chvosty dostáváme zajímavý výsledek.

Věta 2.24. *Nechť F_1, F_2 jsou dvě distribuční funkce takové, že $\bar{F}_i = x^{-\alpha} L_i(x)$ pro $i = 1, 2$ $\alpha \geq 0$ a $L_i \in \mathcal{R}_0$. Pak konvoluce $G = F_1 * F_2$ má pomalu se měnící chrost a platí*

$$\bar{G}(x) \approx x^{-\alpha} (L_1(x) + L_2(x)), \quad x \rightarrow \infty.$$

Důkaz. [Embrechts et al., 1997]. □

Důsledek 2.25. *Nechť $\bar{F}(x) = x^{-\alpha} L(x)$ pro $\alpha \geq 0$ a $L \in \mathcal{R}_0$, pak pro každé $n \in \mathbb{N}$ platí:*

$$\bar{F}^{*n}(x) \approx n\bar{F}(x), \quad x \rightarrow \infty.$$

Důvod, proč se zajímáme o n -té konvoluční mocninu, bude zřejmý později (viz následující kapitola a také věta 5.9). Pokud uvažujeme X_1, \dots, X_n iid s distribuční funkcí F a označíme $S_n = X_1 + \dots + X_n$ jejich součet a $M_n = \max\{X_1, \dots, X_n\}$ jejich maximum, pak pro každé $n \geq 2$ máme:

$$P(S_n > x) = \bar{F}^{*n}(x)$$

a

$$P(M_n > x) = \bar{F}^n(x) = \bar{F}(x) \sum_{k=0}^{n-1} F^k(x) \approx n\bar{F}(x), \quad \text{pro } x \rightarrow \infty.$$

Pokud budeme navíc předpokládat, že F má pomalu se měnící chrost, tj. $F(x) = x^{-\alpha} L(x)$, $\alpha \geq 0$, $L \in \mathcal{R}_0$ můžeme aplikovat důsledek 2.25 a dostáváme

$$P(S_n > x) \approx P(M_n > x), \quad x \rightarrow \infty.$$

Neboli rozdelení součtu náhodných veličin s pomalu se měnícím chrostem je z velké části určeno chrostem rozdelení maxima.

2.4.3 Diskrétní rozdělení

I diskrétní rozdělení se dají rozdělit podle jejich vlastností.

Definice 2.26. Říkáme, že pravděpodobnostní funkce $\{p_k, k \in \mathbb{N}_0\}$ je

1. **logaritmicko-konvexní**, pokud $p_{k+1}^2 \leq p_{k+2}p_k$ pro každé $k \in \mathbb{N}_0$;
2. **logaritnicko-konkávní**, pokud $p_{k+1}^2 \geq p_{k+2}p_k$ pro každé $k \in \mathbb{N}_0$.

Na závěr ještě zmíníme v dalším textu používaná diskrétní rozdělení, protože jejich definice se mohou mírně lišit a také jako ujasnění použitého značení.

- **Degenerované rozdělení** δ_a , $a \in \mathbb{R}$, kde

$$P(N = x) = \begin{cases} 1 & x = a \\ 0 & x \neq a. \end{cases}$$

$$\hat{m}_N(s) = e^{as}$$

- **Binomické rozdělení** $B(n, p)$, $p \in (0, 1)$, $n \in \mathbb{N}$ kde

$$P(N = k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, 2, \dots,$$

$$\hat{m}_N(s) = (pe^s + 1 - p)^n.$$

- **Poissonovo rozdělení** $Po(\lambda)$, $\lambda > 0$, kde

$$P(N = k) = e^{-\lambda} \frac{\lambda^k}{k!}, \quad k = 0, 1, 2, \dots,$$

$$\hat{m}_N(s) = \exp\{\lambda(e^s - 1)\}.$$

- **Negativně binomické rozdělení** $NBi(\alpha, p)$, pro $p \in (0, 1)$, $\alpha \in \mathbb{N}$

$$P(N = k) = \binom{\alpha + k + 1}{k} (1-p)^\alpha p^k, \quad k = 0, 1, 2, \dots$$

Někdy se s výhodou používá ekvivalentní definice

$$P(N = k) = \binom{-\alpha}{k} (1-p)^\alpha (-p)^k.$$

$$\hat{m}_N(s) = \left(\frac{1-p}{1-pe^s} \right)^\alpha.$$

- **Geometrické rozdělení** $Ge(p)$ je speciální varianta negativně binomického pro $\alpha = 1$.

$$P(N = k) = (1-p)p^k, \quad k = 0, 1, 2, \dots,$$

$$\hat{m}_N(s) = \frac{1-p}{1-pe^s}.$$

Geometrické rozdělení je diskrétní rozdělení bez paměti, platí $P(N \geq i+j | N \geq j) = P(N \geq i)$ pro každé $i, j \in \mathbb{N}$.

Bud' $\alpha \in \mathbb{N}$, pak součet α iid náhodných veličin s $Ge(p)$ rozdělením má $NB(\alpha, p)$ rozdělení.

Nulové oříznutí geometrického rozdělení $Ge(p)$ označujeme $TG(p)$ z anglického truncated geometric distribution. Pro X s 0-oříznutým geometrickým rozdělením platí $P(X = k) = (1-p)p^{k-1}$, $k = 1, 2, \dots$

- **Logaritmické rozdělení** s parametrem $\alpha \in (0, 1)$, kde

$$P(N = k) = \frac{1}{-\log(1-\alpha)} \frac{\alpha^k}{k}, \quad k = 0, 1, 2, \dots,$$

$$\hat{m}_N(s) = \frac{\log(1 - \alpha e^s)}{\log(1 - \alpha)}.$$

2.5 Modely rizika v pojišťovnictví

V pojistné matematice se používají dva základní modely příchodu pojistných událostí.

Individuální model

Uvažujeme portfolio, které se skládá z n jednotek. Každá jednotka i má své individuální riziko X_i , $i = 1, \dots, n$ (během daného časového úseku např. jeden rok). Předpokládáme, že náhodné veličiny X_1, \dots, X_n jsou nezávislé, ale nemusí být nutně stejně rozdělené. O rozdělení F_{X_i} náhodné veličiny X_i předpokládáme, že je směsí $F_{X_i} = (1 - \theta_i)\delta_0 + \theta_i F_{V_i}$, kde $0 < \theta_i \leq 1$ a F_{V_i} je rozdělení kladné náhodné veličiny V_i , $i = 1, \dots, n$. V praxi jsou obvykle θ_i malé a mohou být interpretovány jako pravděpodobnost, že i -tá jednotka produkuje kladný přírůstek V_i . Celkové pojistné plnění potom je $C^{ind} = \sum_{i=1}^n X_i$ a jeho rozdělení je $F_{X_1} * \dots * F_{X_n}$. Portfolio se nazývá homogenní pokud platí $F_{V_1} = \dots = F_{V_n}$.

C^{ind} je součet pevného rozsahu, kde sčítáme náhodné veličiny, které v mnoha případech nabývají nulové hodnoty. Rozdělení C^{ind} je dáno větou o konvoluci 2.7, ovšem konkrétní výpočet je pro velká n obtížný a obvykle se používají nějaké approximace.

Kolektivní model:

Předpokládáme, že se portfolio skládá z velkého počtu anonymních jednotek, které nepozorujeme jednotlivě. Celkový počet událostí N , které nastanou za daný časový úsek (např. jeden rok) je náhodný (nejčastěji předpokládáme, že má Poissonovo, binomické nebo negativně binomické rozdělení, ale i další rozdělení jsou možná). Dále předpokládáme, že přírůstky X_i , $i \in \mathbb{N}$ jsou kladné (tedy nenulové!), nezávislé a stejně rozdělené náhodné veličiny. Dále předpokládáme, že velikosti přírůstků X_i jsou nezávislé na jejich celkovém počtu N . Celkové pojistné plnění z takového portfolia je součet náhodného rozsahu $C^{kol} = \sum_{i=1}^N X_i$.

Jak bylo řečeno, je pro velká n výpočet C^{ind} v individuálním modelu obtížný, a proto se snažíme approximovat individuální model vhodným kolektivním modelem. Kolektivní model je ve většině případů z matematického hlediska jednodušší.

Celkové pojistné plnění C^{kol} je složená náhodná veličina, o kterých bude řeč v následující kapitole. Sledujeme-li portfolio založené na kolektivním modelu po celou dobu, nejen na konci období, lze ho považovat za složený bodový proces.

Kapitola 3

Složená rozdělení a jejich vlastnosti

V této kapitole se budeme věnovat vlastnostem rozdělení, která vznikají součtem náhodného počtu náhodných veličin.

3.1 Složené rozdělení

Definice 3.1. Nechť N je nezáporná celočíselná náhodná veličina a nechť X_1, X_2, \dots je posloupnost nezáporných náhodných veličin. Potom náhodná veličina

$$C = \sum_{i=1}^N X_i \quad (3.1)$$

(používáme konvenci $\sum_1^0 = 0$) se nazývá **složená náhodná veličina** (compound).

V této kapitole budeme často uvažovat následující předpoklad.

Předpoklad 1. Náhodné veličiny N, X_1, X_2, \dots jsou nezávislé a náhodné veličiny X_1, X_2, \dots jsou navíc stejně rozdělené s distribuční funkcí F_X .

Označme $p_k = P(N = k)$, $k \in \mathbb{N}$ a F_X distribuční funkci iid náhodných veličin X_i , $i \in \mathbb{N}$. O náhodné veličině $C = \sum_{i=1}^N X_i$ z předchozí definice za předpokladu 1 říkáme, že má **složené rozdělení** dané $\{p_k, k \in \mathbb{N}\}$ a F_X .

Věta 3.2. Bud' C náhodná veličina se složeným rozdělením daným $\{p_k, k \in \mathbb{N}\}$ a F_X , pak pro její distribuční funkci platí

$$F_C = \sum_{k=0}^{\infty} p_k F_X^{*k},$$

kde F_X^{*k} značí k -tou konvoluční mocninu $F_X^{*k} = F_X * \dots * F_X$ k -krát.

Důkaz. [Cipra, 1991] □

Věta 3.3. Bud' $C = \sum_{i=1}^N X_i$ složená náhodná veličina a nechť platí předpoklad 1, pak

1. pro Laplace-Stieljesovu transformaci C platí:

$$\hat{l}_C(s) = \hat{g}_N(\hat{l}_X(s)) \quad \forall s \geq 0,$$

2. pro momentovou vytvářející funkci platí:

$$\hat{m}_C(s) = \hat{g}_N(\hat{m}_X(s)) \quad \forall s \geq 0,$$

kde $\hat{g}_N(s)$ je pravděpodobnostní vytvářející funkce náhodné veličiny N .

Důkaz. Pro Laplace-Stieltesovu transformaci je věta dokázána např. v [Rolski et al., 1998] a pro momentovou vytvářející funkci v [Cipra, 1991]. □

Důsledek 3.4. Předpokládejme, že příslušné momenty existují, pak

$$\mathbb{E}C = \mathbb{E}N\mathbb{E}X, \quad \text{var } C = \text{var } N(\mathbb{E}X)^2 + \mathbb{E}N\text{var } X.$$

Důkaz. $\hat{m}'_C(s) = \hat{g}'_N(\hat{m}_X(s))\hat{m}'_X(s)$ a jak známo $\mathbb{E}C = \hat{m}'_C(0)$ a dostáváme tedy $\mathbb{E}C = \hat{g}'_N(1)\hat{m}'_X(0) = \mathbb{E}N\mathbb{E}X$.

Vzorec pro rozptyl dostaneme analogicky s využitím $\text{var } C = \mathbb{E}C^2 - (\mathbb{E}C)^2$ a $\mathbb{E}C^2 = \hat{m}''_C(0)$. □

Poznámka 3.5. Na základě vlastností momentové vytvářející funkce můžeme snadno odvodit vzorce i pro další momenty.

3.1.1 Složené Poissonovo rozdělení

V literatuře nejvíce studovanou variantou je tzv. složené Poissonovo rozdělení, které doznameme, když N má Poissonovo rozdělení.

Definice 3.6. Nechť $N \sim Po(\lambda)$ a nechť platí předpoklad 1. Potom rozdělení náhodné veličiny $C = \sum_{i=1}^N X_i$ se nazývá **složené Poissonovo rozdělení** (compound Poisson distribution) s charakteristikou (λ, F_X) a značí se $CP(\lambda, F_X)$.

Věta 3.7. Nechť $C \sim CP(\lambda, F_X)$, pak platí:

1. $F_C(x) = \sum_{k=0}^{\infty} e^{-\lambda} \frac{\lambda^k}{k!} F_X^{*k}(x),$
2. $\hat{m}_C(s) = \exp\{\lambda(\hat{m}_X(s) - 1)\},$
3. $\mathbb{E}C = \lambda \mathbb{E}X,$
4. $\text{var } C = \lambda \mathbb{E}X^2.$

Důkaz. Věta je přímým důsledkem věty 3.2 a důsledku 3.4. \square

Věta 3.8. Bud' $n \in \mathbb{N}$ a $C = C_1 + \dots + C_n$, kde $C_i \sim CP(\lambda_i, F_i)$, $i = 1, \dots, n$. Pak C má složené Poissonovo rozdělení s charakteristikou (λ, F) , kde

$$\lambda = \sum_{i=1}^n \lambda_i \quad a \quad F = \sum_{i=1}^n \frac{\lambda_i}{\lambda} F_i.$$

Důkaz. [Rolski et al., 1998] \square

3.1.2 Složené geometrické rozdělení

Dalším ze složených rozdělení, kterými se budeme v následujícím textu zabývat podrobněji, je složené geometrické rozdělení.

Definice 3.9. Nechť $N \sim Ge(p)$ a nechť platí předpoklad 1. Potom rozdělení náhodné veličiny $C = \sum_{i=1}^N X_i$ nazýváme **složené geometrické rozdělení** (geometric compound) s charakteristikou (p, F_X) .

Věta 3.10. Nechť C má složené geometrické rozdělení s charakteristikou (p, F_X) . Pak platí:

1. $F_C(x) = \sum_{k=0}^{\infty} (1-p)p^k F_X^{*k}(x),$
2. $\hat{m}_C(s) = \frac{1-p}{1-p+\hat{m}_X(s)},$
3. $\mathbb{E}C = \frac{p}{1-p} \mathbb{E}X,$
4. $\text{var } C = \frac{2p-p^2}{(1-p)^2} \mathbb{E}X^2 + \frac{p}{1-p} (\mathbb{E}X)^2.$

Důkaz. Věta je přímým důsledkem věty 3.2 a důsledku věty 3.3. \square

3.2 Panjerův algoritmus a rekurentní metody

V této kapitole si uvedeme několik rekurentních metod pro výpočet rozdělení složené náhodné veličiny C . Ovšem jejich přímé použití je limitované, protože jde většinou o poměrně složité formule.

V pojistné matematice je největší pozornost věnována třem speciálním případům, kdy N má Poissonovo, binomické nebo negativně binomické rozdělení. C má pak složené Poissonovo, resp. složené binomické, resp. složené negativně binomické rozdělení zvané častěji složené Pascalovo rozdělení (jeho speciálním případem je složené geometrické rozdělení). V této kapitole ukážeme, že právě tato tři rozdělení mají některé příjemné vlastnosti.

Pravděpodobnostní rozdělení $\{p_k, k \in \mathbb{N}\}$ nezáporné diskrétní náhodné veličiny splňuje **Panjerův rekurentní vztah**, když platí

$$p_k = \left(a + \frac{b}{k}\right)p_{k-1}, \quad k = 1, 2, \dots \quad (3.2)$$

kde $a < 1$ a $b \in \mathbb{R}$ jsou konstanty.

Bylo dokázáno, že pouze Poissonovo, binomické a negativně binomické rozdělení splňují Panjerův rekurentní vztah (3.2) přesně. Stačí však i malé oslabení a dostáváme mnoho dalších kandidátů (např. logaritmické rozdělení splňuje Panjerův rekurentní vztah pro $k = 2, 3, \dots$).

Věta 3.11. *Nechť N je diskrétní náhodná veličina a nechť její rozdělení $\{p_k = P(N = k), k \in \mathbb{N}\}$ splňuje Panjerův rekurentní vztah (3.2), pak N má Poissonovo, binomické nebo negativně binomické rozdělení. Navíc platí*

1. když $a = 0$, pak $b = \lambda > 0$ a $N \sim Po(\lambda)$;
2. když $a < 0$, pak existuje $n \in \mathbb{N}$ takové, že $b = -a(n+1)$ a $N \sim Bi(n, p)$, kde $p = \frac{a}{a-1}$ a $n = -\frac{b}{a} - 1$;
3. když $0 < a < 1$, pak $a + b > 0$ a $N \sim NB(\alpha, p)$, kde $p = a$ a $\alpha = 1 + \frac{b}{p}$.

Důkaz. [Rolski et al., 1998] □

3.2.1 Panjerův algoritmus

Panjerův algoritmus nám umožňuje za platnosti Panjerova rekurentního vztahu (3.2) počítat rekurentně pravděpodobnosti $P(C = \sum_{i=1}^N X_i = k)$.

Předpokládejme, že platí kolektivní model neboli $C = C^{kol} = \sum_{i=1}^N X_i$ a platí předpoklad 1 a navíc X_i jsou diskrétní. Označme $\{p_k = P(N = k), k \in \mathbb{N}\}$, $\{q_k = P(X_i = k), k \in \mathbb{N}, i = 1, 2, \dots\}$ a $p_k^C = P(C^{kol} = k)$.

Věta 3.12. (*Panjerův algoritmus*)

Nechť platí Panjerův rekurentní vztah (3.2). Pak platí

$$p_j^C = \begin{cases} \hat{g}_N(q_0) & \text{pro } j = 0, \\ (1 - aq_0)^{-1} \sum_{k=1}^j (a + \frac{bk}{j}) q_k p_{j-k}^C & \text{pro } j = 1, 2, \dots \end{cases} \quad (3.3)$$

Důkaz. [Rolski et al., 1998] \square

Nevýhodou této verze Panjerova algoritmu je, že jednotlivé přírůstky musí být diskrétní (nebo alespoň nabývat pouze hodnot na nějaké mřížce). Tento nedostatek odstraňuje spojitá verze Panjerova algoritmu, která ovšem vede na integrální rovnici.

Věta 3.13. (*Spojitá verze Panjerova algoritmu*)

Nechť rozdelení $\{p_k\}$ náhodné veličiny N splňuje Panjerův rekurentní vztah (3.2) s parametry a , b a nechť $F_X(0) = 0$. Pak složené rozdelení $F_C = \sum_{k=0}^{\infty} p_k F_X^{*k}$ splňuje integrální rovnici

$$F_C(x) = p_0 + aF_X * F_C(x) + b \int_0^x \int_0^{x-v} \frac{dF_C(y)}{v+y} dF_X(v) \quad x > 0.$$

Pokud $F_X(0) = \alpha > 0$, pak

$$F_C(0) = \begin{cases} e^{(\alpha-1)b} & a = 0 \\ \left(\frac{1-a}{1-a\alpha}\right)^{\frac{a+b}{a}} & a \neq 0. \end{cases} \quad (3.4)$$

Důkaz. [Rolski et al., 1998] \square

Předpokládejme, že distribuční funkce přírůstků F_X je absolutně spojitá a že jí příslušná hustota $f_X(x)$ je navíc omezená. Potom $F_C(x)$ můžeme rozložit na absolutně spojitou část a diskrétní část, kterou tvoří atom v počátku. Pak $F_C(B) = p_0\delta_0(B) + \int_B \tilde{f}_C(x)dx$ pro každou borelovskou množinu $B \in \mathcal{B}(\mathbb{R})$, kde $\tilde{f}_C(x) = \sum_{k=1}^{\infty} p_k f_X^{*k}(x)$ a f_X^{*k} je hustota F_X^{*k} .

Věta 3.14. Nechť $\{p_k\}$ splňuje Panjerův rekurentní vztah (3.2) s parametry a, b a p_0 . Nechť F_X je absolutně spojitá s omezenou hustotou $f_X(x)$, potom hustota $\tilde{f}_C(x)$ absolutně spojité části distribuční funkce $F_C = \sum_{k=0}^{\infty} p_k F_X^{*k}$

splňuje integrální rovnici

$$\tilde{f}_C(x) - \frac{1}{x} \int_0^x (ax + by) f_X(y) \tilde{f}_C(x-y) dy = p_1 f_X(x) \quad x > 0. \quad (3.5)$$

Navíc funkce $\tilde{f}_C(x)$ je jediné řešení rovnice (3.5) na množině všech integrovatelných funkcí na $(0, \infty)$.

Důkaz. [Rolski et al., 1998] □

3.2.2 Individuální model

Předpokládáme-li speciální individuální model, pak pro výpočet pravděpodobností $p_k^C = P(C^{ind} = k)$ nepotřebuje žádné další předpoklady.

Předpokládejme portfolio n jednotek (pojistných smluv), o kterém dále předpokládáme, že se dá rozdělit do několika tříd obsahujících pojistky se stejnou pravděpodobností pojistné události a se stejným podmíněným rozdělením výše nároku za podmínky, že pojistná událost nastala ($F_X|X > 0$). Nechť tedy n_{ij} je počet pojistek s pravděpodobností škodné události $\theta_j < 1$ a s podmíněným rozdělením výše nároku daným $\{p_1^{(i)}, \dots, p_{m_i}^{(i)}\}$, tzn. individuální distribuční funkce pro každou pojistnou smlouvou je směs $F_{ij} = (1 - \theta_j)\delta_0 + \theta_j \sum_{k=1}^{m_i} p_k^{(i)}\delta_k$. Pravděpodobnostní vytvářející funkce celkových pojistných nároků na pojišťovnu C^{ind} v tomto modelu je

$$\hat{g}(s) = \prod_{i=1}^a \prod_{j=1}^b \left(1 - \theta_j + \theta_j \sum_{k=1}^{m_i} p_k^{(i)} s^k\right)^{n_{ij}}, \quad (3.6)$$

kde a je počet možných podmíněných rozdělení výše nároku a b je počet různých pravděpodobností pojistné události. S využitím (3.6) můžeme rekurentně počítat pravděpodobnosti $p_k = P(C^{ind} = k)$ pro $k = 0, 1, \dots, m$, kde $m = \sum_{i=1}^a \sum_{j=1}^b n_{ij} m_i$ představuje maximální možné celkové pojistné plnění.

Věta 3.15. Pro pravděpodobnosti $\{p_k = P(C^{ind} = k)\}$ platí rekurentní vztah

$$p_0 = \prod_{i=1}^a \prod_{j=1}^b (1 - \theta_j)^{n_{ij}} \quad p_k = \frac{1}{k} \sum_{i=1}^a \sum_{j=1}^b n_{ij} v_{ij}(k), \quad k = 1, \dots, m, \quad (3.7)$$

kde

$$v_{ij}(k) = \begin{cases} \frac{\theta_j}{1-\theta_j} \sum_{l=1}^{m_i} p_l^{(i)} (l p_{k-l} - v_{ij}(k-l)) & k = 1, \dots, m \\ 0 & jinak. \end{cases}$$

Důkaz. [Rolski et al., 1998] □

Ve speciálním případě, kde individuální pravděpodobnosti $p_i^{(i)} = 1$ (např. životní pojištění) lze vztah (3.7) zjednodušit a zefektivnit výpočet. Postup se pak obvykle nazývá De Prilův algoritmus.

Důsledek 3.16. (*De Prilův algoritmus*)

Když $p_i^{(i)} = 1$ pro každé $i = 1, \dots, a$, pak pro $\{p_k\}$ platí vztah (3.7), kde

$$v_{ij}(k) = \frac{\theta_j}{1 - \theta_j} (ip_{k-i} - v_{ij}(k-i)).$$

Uved'me ještě jedno zjednodušení, které umožňuje i approximaci pravděpodobností p_k .

Důsledek 3.17. *Když $p_i^{(i)} = 1$ pro každé $i = 1, \dots, a$, pak pro $\{p_k\}$ platí*

$$p_k = \frac{1}{k} \sum_{i=1}^{a \wedge k} \sum_{l=1}^{\lfloor \frac{k}{i} \rfloor} c_{il} p_{k-l}, \quad (3.8)$$

kde

$$c_{il} = (-1)^{l+1} i \sum_{j=1}^b n_{ij} \left(\frac{\theta_j}{1 - \theta_j} \right)^l.$$

Důkaz. [Rolski et al., 1998] □

V praktických aplikacích jsou koeficienty c_{il} blízké nule pro dostatečně velké l , protože pravděpodobnosti θ_j jsou malé. Rekurentní výraz (3.8) může být použit approximativně.

3.3 Lundebergovy meze

Přesný výpočet rozdelení složené náhodné veličiny $C = \sum_{i=1}^N X_i$ je možný pouze ve speciálních případech, a i v těch je obtížný. V této kapitole se budeme proto zabývat asymptotickým chováním chvostu $\bar{F}_C(x) = P(C > x)$ pro velká x . V celé kapitole budeme předpokládat předpoklad 1 (tj. X_i jsou iid a nezávislé na N) a navíc také, že F_X má lehké chvosty.

3.3.1 Složené geometrické rozdělení

Nechť N má geometrické rozdělení s parametrem $p \in (0, 1)$ a předpokládejme, že rozdělení F_X je takové, že rovnice

$$\hat{m}_X(\gamma) = \frac{1}{p} \quad (3.9)$$

má řešení $\gamma > 0$. Takové řešení obvykle nazýváme **vyrovnávací koeficient** (adjustment coefficient) nebo též **Lundebergův koeficient**. Pak umíme odvodit horní a dolní mez pro chvost $\bar{F}_C(x)$.

Pokud neumíme (3.9) řešit explicitně, nebývá těžké vyřešit rovnici numericky. Existují také statistické metody odhadu γ z dat. Postup je analogický jako v kapitole 5.4.1, proto ho zde neuvádíme a čtenáře odkazujeme a zmíněnou kapitolu. Pokud řekneme, že rovnice (3.9) má řešení $\gamma > 0$, pak automaticky předpokládáme, že $\hat{m}_X(s) < \infty$ pro každé $s \leq \gamma$.

Věta 3.18. (Lundebergova nerovnost)

Nechť C má složené geometrické rozdělení a charakteristikou (p, F_X) takové, že rovnice (3.9) má řešení $\gamma > 0$. Označme $x_0 = \sup\{x : F_X(x) < 1\}$, pak platí

$$a_- e^{-\gamma x} \leq \bar{F}_C(x) \leq a_+ e^{-\gamma x}, \quad x \geq 0, \quad (3.10)$$

kde

$$a_- = \inf_{x \in [0, x_0)} \frac{e^{\gamma x} \bar{F}_X(x)}{\int_x^\infty e^{\gamma y} dF_X(y)}, \quad a_+ = \sup_{x \in [0, x_0)} \frac{e^{\gamma x} \bar{F}_X(x)}{\int_x^\infty e^{\gamma y} dF_X(y)}. \quad (3.11)$$

Důkaz. [Rolski et al., 1998] □

Mají-li přírůstky exponenciální rozdělení, dává nám Lundebergova nerovnost (3.10) přesné rozdělení chvostu \bar{F}_C .

Důsledek 3.19. Nechť $F_X \sim Exp(\delta)$, potom $\gamma = (1-p)\delta$, $a_- = a_+ = p$ a tedy

$$\bar{F}_C(x) = p e^{-(1-p)\delta x}, \quad x \geq 0.$$

Důkaz. Pokud $X \sim Exp(\delta)$ pak $\hat{m}_X(\gamma) = \frac{\delta}{\delta - \gamma}$ a vyrovnávací koeficient má splňovat rovnici

$$\frac{\delta}{\delta - \gamma} = \frac{1}{p}.$$

Je tedy $\gamma = \delta(1-p) > 0$, neboť $\delta > 0$ a $p \in (0, 1)$. Dále je

$$\begin{aligned} a_- &= \inf_{x \in [0, x_0)} \frac{e^{\gamma x} \bar{F}_X(x)}{\int_x^\infty e^{\gamma y} dF_X(y)} = \inf_{x \in [0, x_0)} \frac{e^{\delta(1-p)x} \cdot e^{\delta x}}{\int_x^\infty e^{\delta(1-p)y} \cdot \delta e^{\delta y}} \\ &= \inf_{x \in [0, x_0)} \frac{e^{-\delta px}}{\delta \int_x^\infty e^{-\delta py}} = \inf_{x \in [0, x_0)} \frac{p e^{-\delta px}}{e^{-\delta px}} = p \end{aligned}$$

a protože všechny výrazy s x se zkrátí je zároveň $a_+ = p$ a v Lundebergově nerovnosti musí platit rovnost. Tedy $\bar{F}_C(x) = pe^{-(1-p)\delta x}$. \square

Poznámka 3.20. Pro konstanty a_-, a_+ z Lundebergovy věty zřejmě platí $a_- \leq a_+$. Navíc vezmeme-li v úvahu, že

$$\int_x^\infty e^{\gamma y} dF_X(y) \geq e^{\gamma x} \bar{F}_X(x)$$

dostaneme, že $a_+ \leq 1$.

3.3.2 Obecnější složená rozdělení

V této části se pokusíme rozšířit předchozí výsledky i na jiná, nejen složené geometrické rozdělení. Nechť rozdělení N je dáno pravděpodobnostmi $\{p_k, k \in \mathbb{N}\}$ a předpokládejme, že existuje $\theta \in (0, 1)$ tak, že N je stochasticky menší než $(1 - p_0)$ -modifikace oříznutého geometrického rozdělení $TG(\theta)$.

Připomeňme, co rozumíme pojmem stochasticky menší.

Definice 3.21. Nechť jsou X, Y reálné náhodné veličiny, pak říkáme, že X je stochasticky menší než Y a značíme $X \leq_{st} Y$, když pro každou rostoucí funkci $f : \mathbb{R} \rightarrow \mathbb{R}$ je

$$\mathbb{E}f(X) \leq \mathbb{E}f(Y),$$

za předpokladu, že obě střední hodnoty $\mathbb{E}f(X), \mathbb{E}f(Y)$ existují a jsou konečné.

Užitečnou klasifikaci stochasticky menších náhodných veličin dává následující věta.

Věta 3.22. $X \leq_{st} Y$ právě tehdy, když pro každé $x \in \mathbb{R}$ je $\bar{F}_X(x) \leq \bar{F}_Y(x)$.

Důkaz. [Rolski et al., 1998] \square

Vraťme se k rozširování Lundebergových mezí na obecnější rozdělení. Předpokládejme, že pro pravděpodobnostní rozdělení $\{p_k = P(N = k), k \in \mathbb{N}\}$ platí

$$\{p_k\} \leq_{st} \{p'_k\} = p_0\delta_0 + (1 - p_0)TG(\theta), \quad (3.12)$$

pak pro rozdělení složené náhodné veličiny $C = \sum_{i=1}^N X_i$ lze odvodit jednostrannou obdobu Lundebergovy nerovnosti.

Věta 3.23. Nechť (3.12) platí pro nějaké $\theta \in (0, 1)$ a nechť γ splňuje rovnici $\theta \hat{m}_X(\gamma) = 1$. Pak

$$\bar{F}_C(x) \leq \frac{1 - p_0}{\theta} a_+ e^{-\gamma x}, \quad x \geq 0,$$

kde a_+ je definováno v (3.11).

Důkaz. [Rolski et al., 1998] \square

Důsledek 3.24. Nechť $\{p_k\}$ je logaritmicko-konkávní a nechť $p_0 + p_1 < 1$, pak

$$\overline{F}_C(x) \leq \frac{(1-p_0)^2}{1-p_0-p_1} a_+ e^{-\gamma x}, \quad x \geq 0, \quad (3.13)$$

kde a_+ je definováno v (3.11) a $\gamma > 0$ je řešení rovnice $\hat{m}_X(\gamma) = \frac{1-p_0}{1-p_0-p_1}$.

Důkaz. [Rolski et al., 1998] \square

Příklad. Uvažujme složenou náhodnou veličinu $C = \sum_{i=1}^N X_i$, kde N má Poissonovo rozdělení s parametrem λ a přírůstky X_i jsou iid s $Exp(\delta)$. Pak na C můžeme aplikovat důsledek 3.24, neboť rozdělení N je logaritmicko-konkávní. Zřejmě také platí $p_0 + p_1 < 1$ a dále má platit $p_{k+1}^2 \geq p_{k+2}p_k$, tj.

$$\left(e^{-\lambda} \frac{\lambda^{k+1}}{(k+1)!} \right)^2 \geq e^{-\lambda} \frac{\lambda^{k+2}}{(k+2)!} \cdot e^{-\lambda} \frac{\lambda^k}{k!}$$

neboli

$$\frac{1}{k+1} \geq \frac{1}{k+2},$$

což platí pro všechna $k \in \mathbb{N}_0$.

Momentová vytvářející funkce exponenciálního rozdělení je $\hat{m}_X(s) = \frac{\delta}{\delta-s}$ a koeficient $\gamma > 0$ dostaváme řešením rovnice

$$\frac{\delta}{\delta-\gamma} = \frac{1-p_0}{1-p_0-p_1} = \frac{1-e^{-\lambda}}{1-e^{-\lambda}-\lambda e^{-\lambda}}.$$

$$\gamma = \frac{\delta \lambda e^{-\lambda}}{1-e^{-\lambda}}$$

a analogickým výpočtem jako v důkazu důsledku 3.19 dostaneme

$$a_+ = \frac{e^{-\lambda}(\delta+1)-1}{1-e^{-\lambda}}.$$

Po dosazení do (3.13) máme omezení pro chvost složeného Poissonova rozdělení

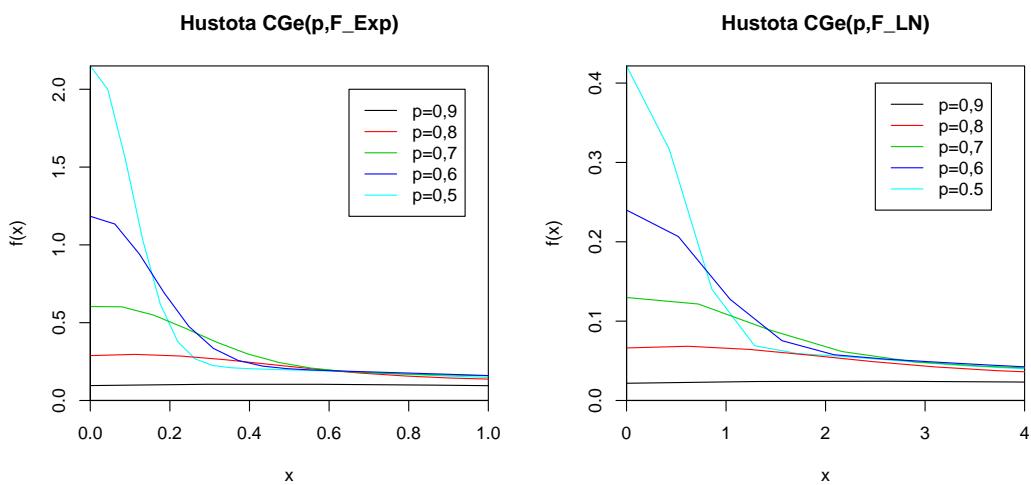
$$\overline{F}_C(x) \leq \frac{(\delta+1)(e^{-\lambda}-e^{-2\lambda})+e^{-\lambda}-1}{1-e^{-\lambda}-\lambda e^{-\lambda}} \cdot e^{\gamma x}.$$

3.4 Simulační studie I

V předchozích částech jsem zkoumali chování složených náhodných veličin teoreticky, v této části prozkoumáme jejich chování na základě simulací. Veškeré výsledky jsou získané na základě 100 000 výběrů. Zaměřili jsme se jen na dvě nejdůležitější rozdělení čítací náhodné veličiny N a to geometrické a Poissonovo. Pro přírůstky jsme vybrali rozdělení exponeční $Exp(1)$, protože pro něj platí mnoho speciálních výsledků, a logaritmicko-normální $LN(1, 1)$ jako zástupce rozdělení s těžkými chvosty. Celá studie byla provedena v software R2.4.0, zdrojové kódy uvádíme v příloze.

3.4.1 Geometrické rozdělení

Zajímali jsme se o chování složeného geometrického rozdělení $CGe(p, F_X)$ v závislosti na parametru p .



Obrázek 3.1: Hustota složeného geometrického rozdělení v závislosti na parametru p a na rozdělení přírůstku (vlevo exponenciální, vpravo logaritmicko-normální).

Zdůrazněme, že složené geometrické rozdělení má atom v počátku $P(C = 0) = 1 - p$. Hustoty byly získány jako jádrové odhady, s pomocí standardních kódů z R.

Z obrázku 3.1 můžeme vidět, že s klesajícím parametrem geometrického rozdělení klesá střední hodnota C a mnohem výrazněji rozptyl. Průběh v závislosti na p se pro exponenciální a logaritmicko-normální přírůstky v zásadě neliší, pouze pro logaritmicko-normální přírůstky je střední hodnota

i rozptyl výrazně vyšší (srovnej tab. 3.1). S klesajícím parametrem p dále klesá střední hodnota a rozptyl, jinak se průběh hustoty nemění.

p	$\mathbb{E}CGe(p, F_{Exp})$	$\text{var } CGe(p, F_{Exp})$	$\mathbb{E}CGe(p, F_{LN})$	$\text{var } CGe(p, F_{LN})$
0,9	9,03	99,42	40,58	2129,50
0,8	3,98	23,78	17,88	539,20
0,7	2,34	10,19	10,49	240,94
0,6	1,49	5,15	6,78	129,88
0,5	1,00	3,01	4,51	75,55
0,4	0,67	1,77	3,00	45,80
0,3	0,43	1,04	1,93	26,86
0,2	0,25	0,56	1,12	15,00
0,1	0,11	0,23	0,50	6,81

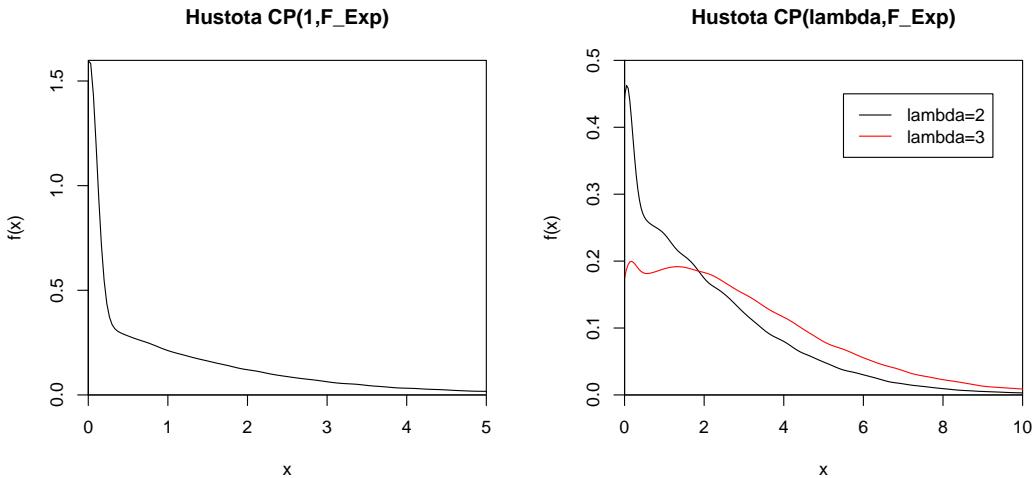
Tabulka 3.1: Odhady střední hodnoty a rozptylu pro složené geometrické rozdělení v závislosti na parametru p .

Lundebergovy meze

Pro exponenciální přírůstky platí Lundebergova nerovnost s rovností a důsledek 3.24 dává přesné rozdělení chvostu $P(C > x)$. To potvrzuje i simulační studie. Odhad kumulované distribuční funkce se překrývá s distribuční funkcí získanou na základě důsledku 3.24. Obrázek neuvádíme.

3.4.2 Poissonovo rozdělení

V této části jsme se zajímali o chování složeného Poissonova rozdělení $CP(\lambda, F_X)$ v závislosti na parametru λ .



Obrázek 3.2: Hustota složeného Poissonova rozdělení v závislosti na parametru λ pro exponenciální přírůstky, $\lambda = 1$ vlevo a pro $\lambda = 2, 3$ vpravo.

Průběh hustoty se mění v závislosti na parametru λ nejen co se týče střední hodnoty a rozptylu, ale také se mění tvar. Od $\lambda = 4$ se tvar ustaluje a pouze se posouvá střední hodnota a zvyšuje rozptyl viz obrázky 3.2 a 3.3.

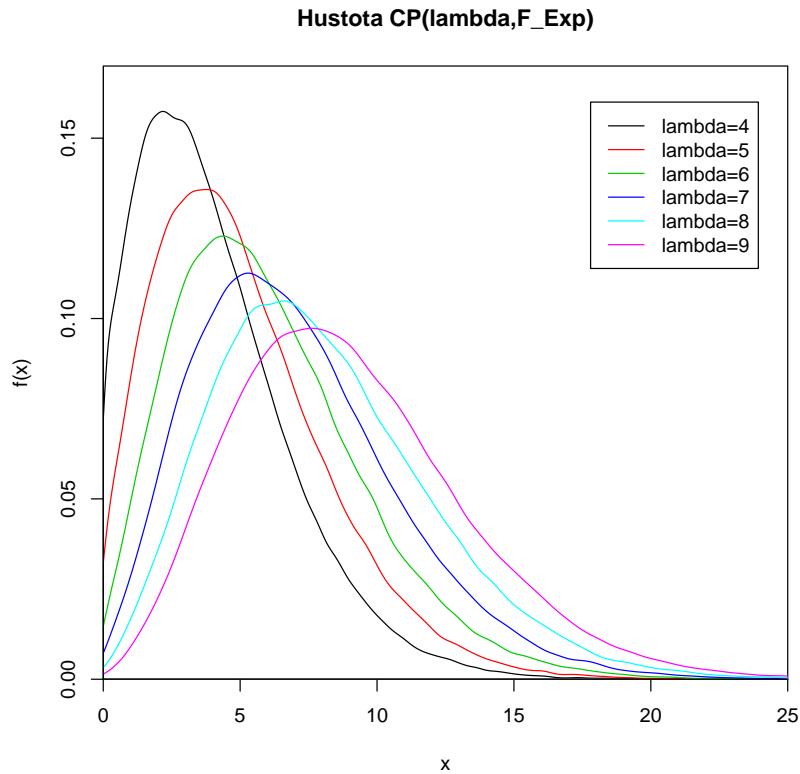
Pro logaritmicko-normální přírůstky se tvar hustoty ustaluje už od $\lambda = 3$ viz obrázek 3.4. Opět platí, že se tvar hustoty složeného Poissonova rozdělení příliš neliší pro exponenciální a logaritmicko-normální přírůstky, jen pro logaritmicko-normální přírůstky má výrazně vyšší rozptyl i střední hodnotu (srovnej tab. 3.2).

Ještě připomeňme, že složené Poissonovo rozdělení má atom v počátku $P(C = 0) = e^{-\lambda}$. Hustoty byly opět získány jako jádrové odhady.

Lundebergovy meze

Pro Poissonovo rozdělení čítací náhodné veličiny N platí důsledek 3.24, který nám dává pouze horní mez pro rozdělení chvostu. Jak přesné je toto omezení jsme studovali pomocí simulací.

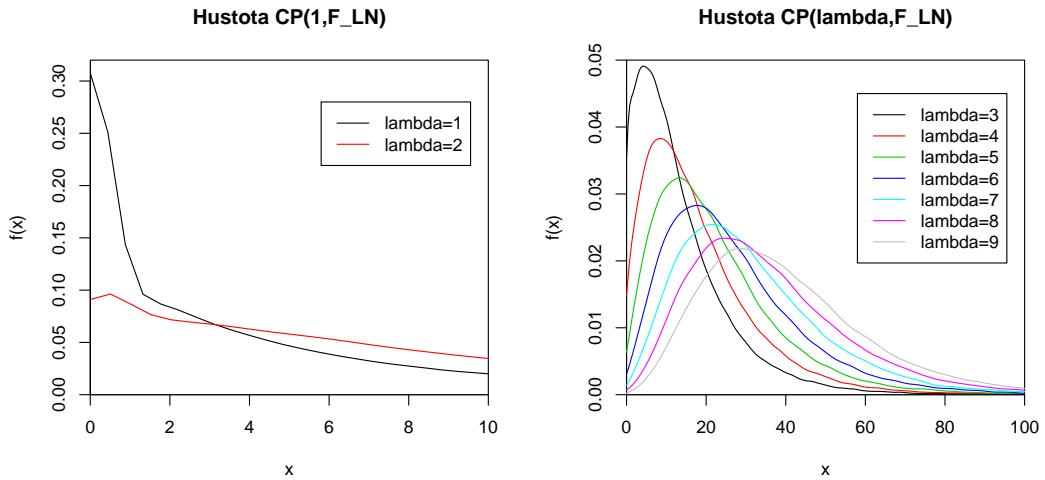
Na obrázku 3.5 uvádíme porovnání empirické distribuční funkce a vypočtené meze na základě důsledku 3.24 (samozřejmě hodnoty 1-mez, protože jde



Obrázek 3.3: Hustota složeného Poissonova rozdělení v závislosti na parametru λ pro exponenciální přírůstky.

o nerovnost pro chvost $1 - F(x)$ a také z horního omezení se stává dolní omezení).

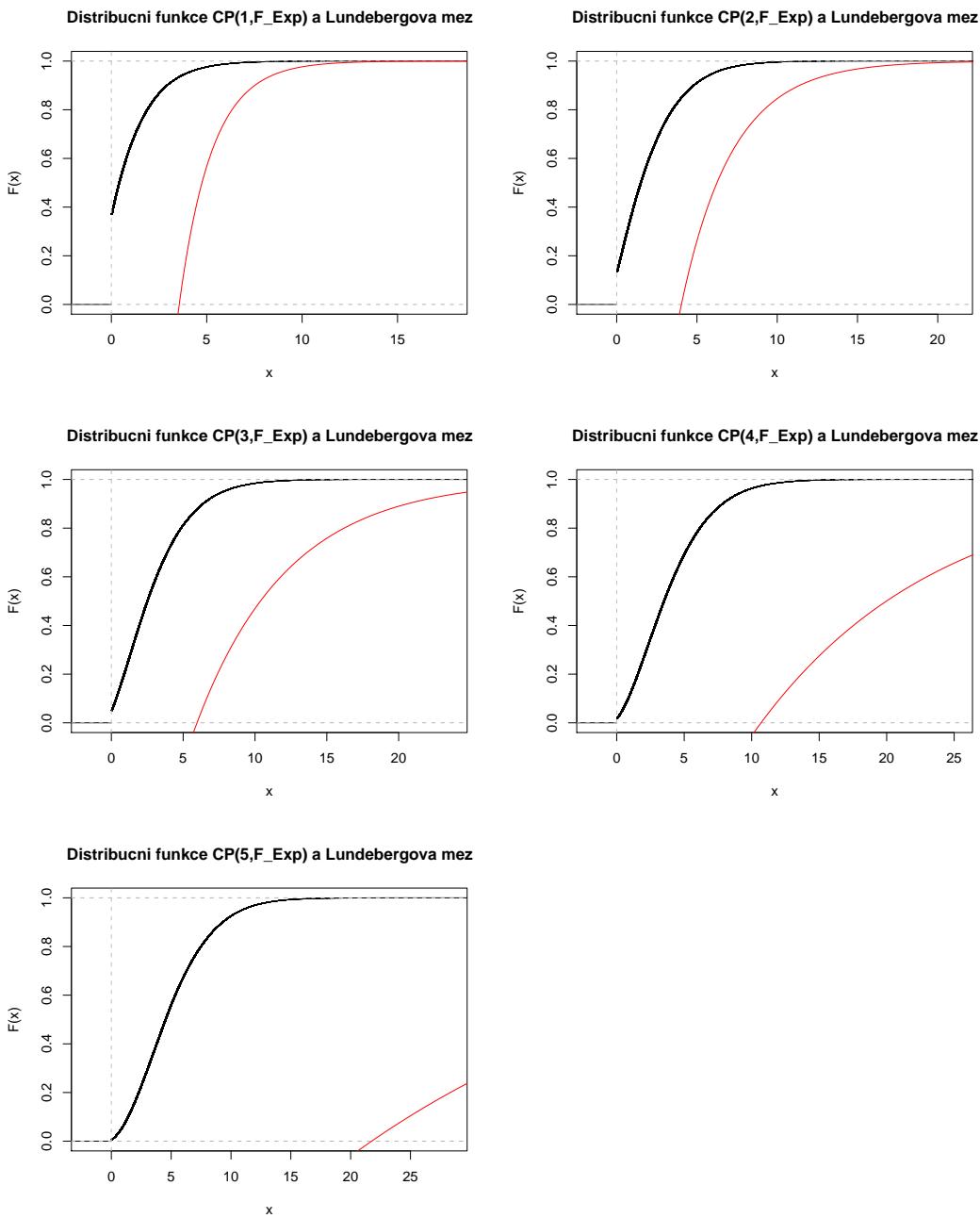
S rostoucí hodnotou parametru λ kvalita omezení klesá. Pro $\lambda = 6$ se na zvoleném úseku omezení ani nedostává do kladných hodnot. Pro velké hodnoty x se omezení vždy přiblíží distribuční funkci. Pro složené Poissonovo rozdělení dává Lundebergova mez dobré omezení až pro velké hodnoty x , ale to není až tak překvapující, protože mez byla odvozena pro chvost.



Obrázek 3.4: Hustota složeného Poissonova rozdělení v závislosti na parametru λ pro logaritmicko-normální přírůstky.

p	$\mathbb{E}CP(\lambda, F_{Exp})$	$\text{var } CP(\lambda, F_{Exp})$	$\mathbb{E}CP\lambda(p, F_{LN})$	$\text{var } CP(\lambda, F_{LN})$
1	1,00	2,01	4,50	54,81
2	2,01	4,01	8,97	107,15
2	3,00	6,07	13,43	161,66
4	3,98	7,94	17,95	218,06
5	5,00	9,94	22,43	274,54
6	6,01	11,99	26,90	331,08
7	7,00	14,00	31,46	383,80
8	8,01	15,97	36,00	445,12
9	9,00	17,94	40,38	485,84

Tabulka 3.2: Odhadování střední hodnoty a rozptylu pro složené Poissonovo rozdělení v závislosti na parametru λ .



Obrázek 3.5: Distribuční funkce (černá) a Lundebergova mez (červená) pro složené Poissonovo rozdělení s exponenciálními přírůstky $Exp(1)$.

Kapitola 4

Proces rizika a ruinování

4.1 Proces rizika a s ním spjaté procesy

Pojem riziko může být v různých oblastech používán různě, v teorii rizika se používá velmi obecná definice rizika.

Definice 4.1. Posloupnost $R = \{X_i, i \in \mathbb{N}\}$, kde X_i jsou nezáporné náhodné veličiny, se nazývá **riziko**. Množina rizik $K = \{R_1, R_2, \dots\}$ se nazývá **portfolio** (rizik). Posloupnost $\{T_1, X_1, T_2, X_2, \dots\}$, kde T_i a X_i jsou nezáporné náhodné veličiny, se nazývá **proces rizika**.

V praxi pak realizace X_i představuje konkrétní požadavek na pojišťovnu, v naší práci je nejčastěji nazýváme přírůstky. Interpretace procesu rizika je taková, že přírůstek X_n nastává v čase $\sigma_n = T_1 + \dots + T_n$, neboli T_i je mezičas mezi událostmi X_{i-1} a X_i .

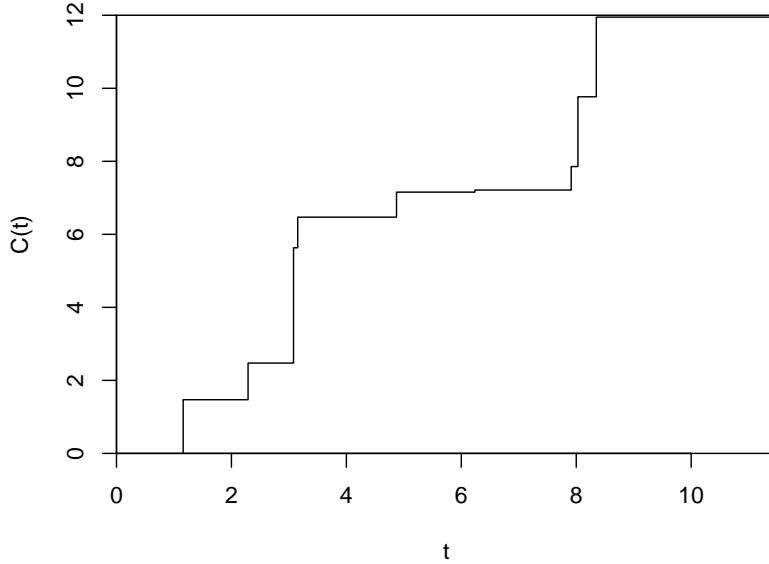
Spíše než samotný vývoj procesu rizika nás zajímá, jaký bude mít celkový dopad na pojišťovnu.

Definice 4.2. Nechť $\{\sigma_n, n \in \mathbb{N}\}$ je jednoduchý bodový proces a $\{N(t), t \geq 0\}$, kde $N(t) = \sum_{n=1}^{\infty} I\{\sigma_n \leq t\}$ příslušný čítací proces a nechť posloupnost $\{X_n, n \in \mathbb{N}\}$ je riziko, kde jsou všechna X_n nenulová. Pak

$$C(t) = \sum_{n=1}^{\infty} X_n I\{\sigma_n \leq t\} = \sum_{n=1}^{N(t)} X_n$$

nazýváme proces **celkových pojistných nároků** (cumulative arrival process, total claim process).

Proces celkových pojistných nároků je zřejmě složený bodový proces.



Obrázek 4.1: Proces celkových pojistných nároků $C(t)$.

Další dva procesy, které si definujeme, jsou úzce spjaty s pojišťovací praxí. Předpokládejme, že pojišťovna má na počátku (své činnosti, nebo sledovaného období) rezervu $R_0 = u \geq 0$. Dále předpokládejme, že celkové pojistné inkasované do času t je lineární funkcí času tvaru $p(t) = pt$, $t \geq 0$, kde p je konstanta (v některých materiálech se dokonce předpokládá $p = 1$). Za těchto předpokladů definujeme proces rezerv.

Definice 4.3. Nechť $\{C(t), t \geq 0\}$ je proces celkových pojistných nároků a nechť inkasované pojistné do času t je $p(t) = pt$. Definujme

$$R(t) = u + pt - C(t),$$

pak $\{R(t), t \geq 0\}$ nazýváme **proces pojistných rezerv** (risk reserve process).

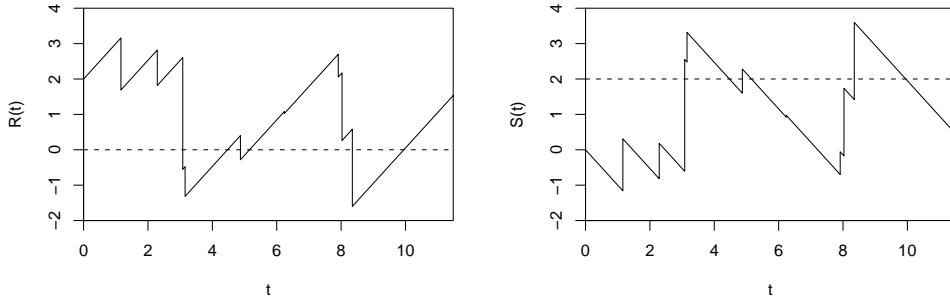
Pro matematické účely bývá někdy vhodnější uvažovat proces vyjadřující přebytek nároků kladených na pojišťovnu nad výší inkasovaného pojistného. Pro tento proces není, narozdíl od anglické terminologie, ustálený český název. My ho budeme označovat proces přebytku nároků.

Definice 4.4. Za stejných předpokladů jako v definici 4.3 definujeme **proces přebytku nároků** (claim surplus process) $\{S(t), t \geq 0\}$, kde

$$S(t) = C(t) - pt.$$

Za určitých předpokladů lze proces $\{S(t)\}$ považovat za náhodnou procházku a na řešení problému ruinování aplikovat teorii o náhodných procházkách.

Na obrázcích 4.1 a 4.2 uvádíme pro představu jednu realizaci právě definovaných procesů pro $\{N(t)\}$ Poissonův proces, X_i exponenciálně rozdělené $Exp(1)$, počáteční rezervu $u = 2$ a pojistné $p = 1$.



Obrázek 4.2: Proces rizikových rezerv $R(t)$ vlevo a proces přebytku nároků $S(t)$ vpravo.

4.2 Ruinování a příbuzné pojmy

O (technickém) ruinování (portfolio) hovoříme, když se proces pojistných rezerv $\{R(t)\}$ dostane do záporných hodnot nebo ekvivalentně, když proces $\{S(t)\}$ překročí počáteční rezervu u . Základní problém, který se v teorii rizika řeší, je nalézt pravděpodobnost ruinování, je-li počáteční rezerva u . Dále se také často vyšetřuje čas prvního ruinování.

Definice 4.5. Při počáteční rezervě $u \geq 0$ definujeme:

1. čas prvního ruinování (time of ruin) $\tau(u) = \inf\{t > 0 : R(t) < 0\} = \inf\{t > 0 : S(t) > u\},$
2. pravděpodobnost ruinování do času $x < \infty$ (finite-horizon ruin probability) $\psi(u, x) = P(\tau(u) \leq x),$

3. pravděpodobnost ruinování v nekonečném čase (infinite-horizon ruin probability, probability of ultimate ruin) $\psi(u) = \lim_{x \rightarrow \infty} \psi(u, x)$,
4. pravděpodobnost přežití (survival probability) $\bar{\psi}(u) = 1 - \psi(u)$.

Přívlásek "v nekonečném čase" se obvykle vynechává a $\psi(u)$ nazýváme krátce pravděpodobnost ruinování. Uvažujeme-li $\psi(u)$ jako funkci u , pak hovoríme o tzv. funkci ruinování (ruin function). Zřejmě platí $\psi(u) = P(\tau(u) < \infty)$.

Kromě pravděpodobnosti a času prvního ruinování existují i další charakteristiky ruinování.

Překročení hodnoty u (overshoot) procesem $\{S(t)\}$ nebo též závažnost ruinování v čase $\tau(u)$

$$Y^+(u) = \begin{cases} S(\tau(u)) - u & \text{když } \tau(u) < \infty \\ \infty & \text{když } \tau(u) = \infty, \end{cases}$$

nebo totéž můžeme vyjádřit pomocí procesu rezerv $\{R(t)\}$

$$Y^+(u) = \begin{cases} -R(\tau(u)) & \text{když } \tau(u) < \infty \\ \infty & \text{když } \tau(u) = \infty. \end{cases}$$

Přebytek do doby ruinování (surplus prior to ruin)

$$X^+(u) = \begin{cases} u - S(\tau(u)) & \text{když } \tau(u) < \infty \\ \infty & \text{když } \tau(u) = \infty. \end{cases}$$

Zřejmě $X^+(u) + Y^+(u)$ je velikost přírůstku, který způsobí ruinování v čase $\tau(u)$.

Čas v červených číslech nebo u nás již také používaný pojem time in the red je další charakteristikou popisující jak vážné bylo ruinování.

$$T'(u) = \begin{cases} \tau'(u) - \tau(u) & \text{když } \tau(u) < \infty \\ 0 & \text{když } \tau(u) = \infty, \end{cases}$$

kde $\tau'(u) = \inf\{t > \tau(u), R(t) > 0\}$ je čas kdy $\{R(t)\}$ poprvé překročí nulu poté co nastalo ruinování. $T'(u)$ udává jak dlouho zůstal proces v záporných hodnotách, ale nepřináší žádnou informaci o jeho chování.

Jako poslední charakteristika se používá **maximální celkový deficit**

$$Z^+(u) = \max\{-R(t) : \tau(u) \leq t\}$$

po čase $\tau(u)$. Vezmeme-li v úvahu všechny periody v červených číslech, můžeme uvažovat maximální celkový deficit pro každou periodu zvlášť.

$$Z_1^+(u) = \max\{-R(t) : \tau(u) \leq t \leq \tau'(u)\}$$

je pak maximální deficit během první periody v červených číslech.

I když se pravděpodobnost ruinování (a další právě definované procesy a náhodné veličiny) uvažují v závislosti na počáteční rezervě u , závisí samozřejmě i na p a na distribucích tvořících proces pojistných nároků $C(t)$.

4.3 Pravděpodobnost ruinování

Jen v několika speciálních případech lze odvodit pravděpodobnost ruinování přesně, ale bylo odvozeno několik asymptotických a approximačních metod.

Definujme si náhodnou procházku $\{S_n, n \in \mathbb{N}_0\}$, kde

$$S_n = \sum_{i=1}^n Y_i, \quad Y_i = X_i - pT_i. \quad (4.1)$$

Ve značení z definic 2.6 a 4.3 T_i označují mezičasy mezi jednotlivými událostmi a p je pojistné inkasované za jednotku času. $\{S_n\}$ je pak diskrétní náhodná procházka, která nese většinu informace o procesu $\{S(t), t \geq 0\}$ a nazývá se diskrétní kostra, nebo též skeleton procesu $\{S(t)\}$. Přírůstky $\{S_n, n \in \mathbb{N}_0\}$ odpovídají přírůstkům procesu $\{C(t), t \geq 0\}$ poníženým o pojistné inkasované za období mezi příslušnými dvěma přírůstky. Jak už značení naznačuje, je $S(t) = S_n$ pro $t = T_n$.

Ze silného zákona velkých čísel plyne, že $\frac{S_n}{n} \xrightarrow{\text{sj}} \mathbb{E}Y$ a $\limsup_{n \rightarrow \infty} S_n = \infty$ pokud $\mathbb{E}Y > 0$. Toto platí dokonce i pro $\mathbb{E}Y = 0$, ale to už není tak snadné dokázat. Čtenáře odkazujeme např. na [Rolski et al., 1998], paragraf 6.3. Tedy pro $\mathbb{E}Y \geq 0$ náhodná procházka překročí jakoukoli mez a ruinování nastává skoro jistě, tj. $\psi(u) \equiv 1$. Pokud $\mathbb{E}Y < 0$, lze ukázat, že $\psi(u) < 1$ pro všechna u (např. [Asmussen, 2000] kapitola 3).

K stejnemu závěru dojdeme, i když uvažujeme spojitý proces $\{S(t)\}$, ovšem odvození je složitější a vychází z diskretizace procesu (viz [Asmussen, 2000]).

Na $\{S(t)\}$ lze také aplikovat centrální limitní větu a ukázat, že limitní rozdělení $\frac{S(t) - \mathbb{E}S(t)}{\sqrt{t}}$ je normální.

Věta 4.6. Nechť $\{S(t), t \geq 0\}$ je proces přebytku nároků příslušný složenému Poissonovu procesu s charakteristikou (λ, F_X) a intenzita pojistného bud' p . Pak

$$\frac{S(t) - t(\lambda \mathbb{E}X - p)}{\sqrt{t}} \xrightarrow{\text{as.}} N(0, \lambda \mathbb{E}X^2)$$

pro $t \rightarrow \infty$.

Důkaz. [Asmussen, 2000]

□

V takto obecném případě více výsledků neodvodíme. Nejvíce výsledků bylo odvozeno pro proces, kde se příchody událostí řídí Poissonovým procesem, ale tomu se věnuje další kapitola.

Kapitola 5

Složený Poissonův model

V této kapitole budeme uvažovat klasický složený Poissonův proces jako model příchodu událostí. Tento model je sice vhodný spíše jako approximace praktických situací. Jeho přímé použití je limitované (obzvláště v pojistné praxi), ale má velmi příjemné vlastnosti a proto je nejvíce studovaným procesem v literatuře. Je také základem pro různá zobecnění.

5.1 Homogenní Poissonův proces

Nejprve si definujeme homogenní Poissonův proces.

Definice 5.1. Buď $\{\sigma_n, n \in \mathbb{N}\}$ jednoduchý bodový proces, kde $T_1 = \sigma_1$ a $T_n = \sigma_n - \sigma_{n-1}$, $n = 2, 3, \dots$ jsou iid náhodné veličiny s exponenciálním rozdělením $Exp(\lambda)$, $\lambda > 0$. Potom čítací proces $\{N(t), t \geq 0\}$, $N(t) = \sum_{n=1}^{\infty} I\{\sigma_n \leq t\}$ příslušný $\{\sigma_n\}$ nazýváme homogenní Poissonův proces s intenzitou λ .

Jednou ze základních vlastností homogenního Poissonova procesu je, že má nezávislé a stacionární přírůstky. Připomeňme pro jistotu definici těchto pojmu.

Definice 5.2. O reálném náhodném procesu $\{X(t), t \geq 0\}$ říkáme, že má

1. **nezávislé přírůstky**, když $\forall n \in \mathbb{N}$ a $0 \leq t_0 < t_1 < \dots < t_n$ jsou náhodné veličiny $X(t_0), X(t_1) - X(t_0), X(t_2) - X(t_1), \dots, X(t_n) - X(t_{n-1})$ nezávislé;
2. **stacionární přírůstky**, když $\forall n \in \mathbb{N}$ a $0 \leq t_0 < t_1 < \dots < t_n$ a $h \geq 0$ rozdělení náhodného vektoru $(X(t_1 + h) - X(t_0 + h), \dots, X(t_n + h) - X(t_{n-1} + h))$ nezávisí na h .

Homogenní Poissonův proces můžeme definovat různě a každá z definic zdůrazňuje jeho jinou vlastnost. Každý bod následující věty je ekvivalentní definicí Poissonova procesu.

Věta 5.3. *Bud' $\{N(t), t \geq 0\}$ čítací proces. Potom jsou následující tvrzení ekvivalentní.*

1. *$\{N(t)\}$ je homogenní Poissonův proces s intenzitou λ .*
2. *$\forall t \geq 0$ a $n \in \mathbb{N}$ má náhodná veličina $N(t)$ Poissonovo rozdělení $Po(\lambda t)$ a je-li $N(t) = n$, pak body $\sigma_1, \dots, \sigma_n$ jsou rovnoraměrně rozdělené na $[0, t]$.*
3. *$\{N(t)\}$ má nezávislé přírůstky a platí $\mathbb{E}N(1) = \lambda$ a pro každé $t \geq 0$, $n \in \mathbb{N}$ a dané $N(t) = n$ jsou body $\sigma_1, \dots, \sigma_n$ rovnoraměrně rozdělené na $[0, t]$.*
4. *$\{N(t)\}$ má stacionární nezávislé přírůstky a pro $h \searrow 0$ $P(N(h) = 0) = 1 - \lambda h + o(h)$, $P(N(h) = 1) = \lambda h + o(h)$.*
5. *$\{N(t)\}$ má stacionární nezávislé přírůstky a pro každé pevné $t \geq 0$ má náhodná veličina $N(t)$ Poissonovo rozdělení $Po(\lambda t)$.*

Důkaz. [Rolski et al., 1998] □

5.2 Složený Poissonův proces

Definice 5.4. Nechť $\{N(t), t \geq 0\}$ je homogenní Poissonův proces s intenzitou λ a nechť X_i jsou iid náhodné veličiny s distribuční funkcí F_X a nezávislé na $\{N(t), t \geq 0\}$. Potom složený bodový proces $\{C(t), t \geq 0\}$, kde

$$C(t) = \sum_{i=1}^{N(t)} X_i$$

se nazývá složený Poissonův proces s charakteristikou (λ, F_X) , tj. s intenzitou λ a rozdělením velikosti přírůstků daným F_X .

Terminologie je motivována faktem, že $C(t) \sim CP(\lambda t, F_X)$, neboť podle věty 5.3 má $N(t)$ Poissonovo rozdělení s parametrem λt .

Uveděme si nějaké základní vlastnosti složeného Poissonova procesu.

Věta 5.5. *Nechť $\{C(t), t \geq 0\}$ je složený Poissonův proces s charakteristikou (λ, F_X) , pak*

1. proces $\{C(t)\}$ má stacionární a nezávislé přírůstky,
2. momentová vytvářející funkce náhodné veličiny $C(t)$ je daná vzorcem

$$\hat{m}_{C(t)}(s) = \exp\{\lambda t(\hat{m}_X(s) - 1)\},$$

a střední hodnota a rozptyl jsou

$$\mathbb{E}C(t) = \lambda t \mathbb{E}X, \quad \text{var } C(t) = \lambda t \mathbb{E}X^2.$$

Důkaz. Bod 1. můžeme nalézt např. v [Rolski et al., 1998], bod 2. je pouhým přeformulováním věty 3.7. \square

Stacionarita a nezávislost přírůstků složeného Poissonova procesu nám dává vlastnost, pro kterou bychom ho mohli nazvat procesem bez paměti. Uvažujeme-li proces pojistných rezerv $\{R(t), t \geq 0\}$ (z definice 4.3) od času $t = T$ dále do budoucnosti, je to jako kdybychom jej restartovali se stejnými parametry, jen počáteční rezerva by byla $u = R(T)$. Navíc, pokud do času T nenastalo ruinování, pak pravděpodobnost (podmíněná) ruinování v budoucnu je $\psi(R(T))$.

Ve zbytku této kapitoly budeme předpokládat, že příchod událostí se řídí složeným Poissonovým modelem.

Předpoklad 2. (*Složený Poissonův model*)

Nechť $\{C(t), t \geq 0\}$ je složený Poissonův proces s charakteristikou (λ, F_X) a pojistné je inkasováno spojité s intenzitou p za jednotku času.

5.3 Pravděpodobnosti ruinování a pravděpodobnosti přežití

Vraťme se k náhodné procházce $\{S_n, n \in \mathbb{N}_0\}$ ze (4.1). Jak víme, pokud $\mathbb{E}Y \geq 0$, pak ruinování nastává skoro jistě. Předpokládejme tedy $\mathbb{E}Y < 0$, po dosazení dostáváme tzv. **podmínu čistého zisku** (net profit condition)

$$p > \lambda \mathbb{E}X. \tag{5.1}$$

kterou budeme předpokládat ve zbytku této kapitoly.

Dále budeme studovat pravděpodobnost přežití $\bar{\psi}(u) = 1 - \psi(u)$ a ukážeme, že jako funkce u je diferencovatelná všude na \mathbb{R}_+ až na spočetně mnoho bodů.

Věta 5.6. Funkce přežití $\bar{\psi}(u)$ je spojitá na \mathbb{R}_+ s jednostrannými derivacemi $\bar{\psi}'_+(u)$ a $\bar{\psi}'_-(u)$ a platí

$$p\bar{\psi}'_+(u) = \lambda \left(\bar{\psi}(u) - \int_0^u \bar{\psi}(u-y) dF_X(y) \right) \quad (5.2)$$

$$p\bar{\psi}'_-(u) = \lambda \left(\bar{\psi}(u) - \int_0^{u-} \bar{\psi}(u-y) dF_X(y) \right) \quad (5.3)$$

Důkaz. [Rolski et al., 1998] \square

Poznámka 5.7. Přímý důsledek věty 5.6 je, že spojitá funkce $\bar{\psi}(u)$ je diferecovatelná až na spočetnou množinu bodů, kde $F_X(x)$ je nespojitá, což implikuje $\int_u^\infty \bar{\psi}'(v) dv = \bar{\psi}(u)$, $u \geq 0$, neboli $\bar{\psi}(u)$ je absolutně spojitá vzhledem k Lebesguově míře.

Bohužel ve většině případů rovnice (5.2) a (5.3) nelze řešit analyticky, ale pouze numericky. Obtížnost řešení spočívá v tom, že rovnice obsahují jak derivaci tak integrál. Jejich integrováním můžeme dostat integrální rovnici.

Věta 5.8. (*Defektivní rovnice obnovy - defective renewal equation*)
Funkce ruinování $\psi(u)$ splňuje integrální rovnici

$$p\psi(u) = \lambda \left(\int_u^\infty \bar{F}_X(x) dx + \int_0^u \psi(u-x) \bar{F}_X(x) dx \right). \quad (5.4)$$

Důkaz. [Rolski et al., 1998] \square

5.3.1 Pollaczek-Khinchinova formule

Přestože rovnice (5.4) je jednoduší než rovnice (5.2) a (5.3), stále je její řešení problematické. Označme pro přehlednost $\mathbb{E}X = \mu$ a dále označme

$$F_X^s(x) = \frac{1}{\mu} \int_0^x \bar{F}_X(y) dy, \quad x \geq 0 \quad (5.5)$$

integrovaný chvost, pak lze pravděpodobnost ruinování reprezentovat jako součet konvolučních mocnin.

Věta 5.9. (*Pollaczek-Kinchinova formule*)
Pro každé $u \geq 0$ platí

$$\psi(u) = \left(1 - \frac{\lambda\mu}{p}\right) \sum_{n=0}^{\infty} \left(\frac{\lambda\mu}{p}\right)^n \overline{(F_X^s)^{*n}}(u). \quad (5.6)$$

Důkaz. [Rolski et al., 1998] \square

Poznámka 5.10. Poznamenejme, že věta 5.9 mimo jiné také říká, že $1 - \psi(u)$ je distribuční funkce náhodné veličiny se složeným geometrickým rozdělením s charakteristikou $(\frac{\lambda \mathbb{E}X}{p}, F_X^s)$ (tedy rozdělení $1 - \psi(u)$ má atom v počátku $1 - \frac{\lambda \mathbb{E}X}{p}$).

Existuje několik speciálních případů, kdy se Pollaczek-Kinchinova formule zjednoduší, ale v ostatních případech není pro výpočet $\psi(u)$ dostačující, protože obsahuje nekonečný součet konvolučních mocnin. Reprezentace jako nekonečný součet je vhodná spíše pro teoretické úvahy nebo pro approximace, kdy sčítáme jen několik prvních členů. Chyba takové approximace je menší než ρ^{k+1} , kde $\rho = \frac{\lambda \mu}{p} < 1$ z podmínky čistého zisku a většinou půjde rychle k nule.

V případě $u = 0$ dostáváme z Pollaczek-Kinchinovy formule (5.6) výraz, který závisí pouze na střední hodnotě přírůstku $\mathbb{E}X$, nikoli na jejich rozdělení.

Věta 5.11. $\psi(0) = \frac{\lambda \mathbb{E}X}{p}$.

Důkaz. [Asmussen, 2000] \square

Pokud jsou přírůstky exponenciálně rozdělené, má Pollaczek-Kinchinova formule (5.6) analytické řešení.

Věta 5.12. Nechť $X_i \sim \text{Exp}(\delta)$, $i \in \mathbb{N}$, pak

$$\psi(u) = \begin{cases} \frac{\lambda}{p\delta} e^{-u(\delta - \frac{\lambda}{p})} & \text{pro } \lambda \leq \delta p, \\ 1 & \text{pro } \lambda > \delta p. \end{cases}$$

Je-li $\lambda < \delta p$, pak pro čas prvního ruinování $\tau(u)$ platí:

$$\mathbb{E}[\tau(u)|\tau(u) < \infty] = \frac{1 + \frac{\lambda u}{p}}{\delta p - \lambda}.$$

Důkaz. [Cipra, 1991] \square

Stejný výsledek dostaneme také řešením rovnice (5.2), kterou lze v tomto případě řešit analyticky, nebo také pomocí Laplaceovy transformace.

Zjednodušení lze dosáhnout také v případě deterministických přírůstků.

Věta 5.13. Nechť $P(X = \mu) = 1$ a nechť $p = 1$ pak

$$\psi(u) = 1 - (1 - \lambda\mu) \sum_{k=0}^{\lfloor u/\mu \rfloor} e^{-\lambda(k\mu-u)} \frac{(\lambda(k\mu-u))^k}{k!}.$$

Důkaz. [Asmussen, 2000] \square

5.3.2 Laplaceova transformace

Podívejme se, jaké informace nám může přinést Laplaceova transformace funkce ruinování, nebo funkce přežití. Snadným výpočtem dostáváme

$$\hat{L}_\psi(s) = \int_0^\infty (1 - \bar{\psi}(u)) e^{-su} du = \frac{1}{s} - \hat{L}_{\bar{\psi}}(s) \quad s > 0$$

a předpokládáme-li složený Poissonův model (předpoklad 2), pak dostáváme následující vyjádření Laplaceových transformací.

Věta 5.14. *Nechť platí předpoklad 2 a $\mathbb{E}X = \mu$. Pak pro Laplaceovy transformace $\hat{L}_\psi(s)$ a $\hat{L}_{\bar{\psi}}(s)$ platí vztahy*

$$\begin{aligned}\hat{L}_\psi(s) &= \frac{1}{s} - \frac{p - \lambda\mu}{ps - \lambda(1 - \hat{l}_X(s))}, & s > 0 \\ \hat{L}_{\bar{\psi}}(s) &= \frac{p - \lambda\mu}{ps - \lambda(1 - \hat{l}_X(s))}, & s > 0,\end{aligned}$$

kde $\hat{l}_X(s)$ značí Laplaceovu-Stieltjesovu transformaci iid náhodných veličin X_i , $i \in \mathbb{N}$.

Důkaz. [Rolski et al., 1998] □

Označme $M = \sup_{0 \leq t < \infty} S(t)$, pak $\psi(u) = P(M > u)$ a důsledkem předchozí věty jsou formule pro momenty M .

Důsledek 5.15. *Předpokládáme-li složený Poissonův model (předpoklad 2), pak první dva momenty M jsou*

$$\mathbb{E}M = \frac{\lambda\mathbb{E}X^2}{2(p - \lambda\mathbb{E}X)} \quad \mathbb{E}M^2 = \frac{\lambda\mathbb{E}X^3}{3(p - \lambda\mathbb{E}X)} + \frac{\lambda^2(\mathbb{E}X^2)^2}{2(p - \lambda\mathbb{E}X)^2}.$$

Důkaz. Víme, že $\bar{\psi}(u) = P(M \leq u)$ je distribuční funkce náhodné veličiny se složeným geometrickým rozdělením s parametrem $\frac{\lambda\mathbb{E}X}{p}$ a rozdělením přírůstku daným F_X^s . Pro momentovou vytvořující funkci M dostáváme výraz

$$\hat{m}_M(s) = \frac{(p - \lambda\mu)s}{ps - \lambda(\hat{m}_X(s) - 1)}.$$

Podrobněji najdeme odvození v [Asmussen, 2000].

Dále dostáváme

$$\hat{m}'_M(s) = \frac{\lambda(p - \lambda\mu)(s\hat{m}'_X(s) - \hat{m}_X(s) + 1)}{(ps - \lambda(\hat{m}_X(s) - 1))^2}.$$

A $\mathbb{E}M = \lim_{s \rightarrow 0} \hat{m}'_M(s)$. Jde o výraz typu $\frac{0}{0}$, a proto použijeme l'Hospitalovo pravidlo a dostáváme

$$\mathbb{E}M = \lim_{s \rightarrow 0} \frac{\lambda(p - \lambda\mu)s\hat{m}''_X(s)}{2(ps - \lambda(\hat{m}_X(s) - 1))(p - \lambda\hat{m}'_X(s))}.$$

Stále máme výraz typu $\frac{0}{0}$ a druhým použitím l'Hospitalova pravidla dostáváme

$$\lim_{s \rightarrow 0} \frac{\lambda(p - \lambda\mu)(\hat{m}''_X(s) + s\hat{m}'''_X(s))}{2(p - \lambda\hat{m}'_X(s))^2 + 2(ps - \lambda(\hat{m}_X(s) - 1))(-\lambda\hat{m}''_X(s))} = \frac{\lambda(p - \lambda\mu)\mathbb{E}X^2}{2(p - \lambda\mathbb{E}X)^2}.$$

Stačí si uvědomit, že $\mu = \mathbb{E}X$ a máme požadovaný výsledek.

Vzorec pro $\mathbb{E}M^2$ dostaneme analogicky.

□

5.4 Asymptotické chování a approximace pravděpodobnosti ruinování

Viděli jsme, že v obecném případě máme problémy odvodit $\psi(u)$ přesně. To nás vede ke snaze odvodit nějaké omezení nebo approximace pro pravděpodobnost ruinování. V této části odvodíme analogii Lundebergovy nerovnosti (3.18) pro pravděpodobnost ruinování $\psi(u)$. Nejprve se zabýveme podrobněji Lundebergovým koeficientem.

5.4.1 Lundebergův koeficient

Uvažujme proces přebytku nároků $\{S(t), t \geq 0\}$, pak má náhodná veličina $S(t)$ v čase $t \geq 0$ složené Poissonovo rozdělení s charakteristikou $(\lambda t, F_X)$ posunuté o $-pt$. Její momentová vytvořující funkce je

$$\hat{m}_{S(t)}(s) = \mathbb{E}e^{sS(t)} = \exp\{t(\lambda(\hat{m}_X(s) - 1) - ps)\} = e^{t\theta(s)},$$

kde

$$\theta(s) = \lambda(\hat{m}_X(s) - 1) - ps.$$

Pokud $\hat{m}_X(s_0) < \infty$ pro nějaké s_0 , potom funkce $\theta(s)$ je diferencovatelná nekonečně-krát na intervalu $(-\infty, s_0)$. Její druhá derivace je

$$\theta''(s) = \lambda\hat{m}''_X(s) = \lambda\mathbb{E}(X^2e^{Xs}) > 0$$

tedy $\theta(s)$ je konvexní funkce. Pro první derivaci platí

$$\theta'(s) = \lambda \hat{m}'_X(s) - p$$

a $\theta'(0) = \lambda \mathbb{E}X - p < 0$, protože uvažuje podmínu čistého zisku, aby pravděpodobnost ruinování nebyla skoro jistě jedna. Funkce $\theta(s)$ je tedy klesající konvexní v nule a $\theta(0) = \lambda(\hat{m}_X(0) - 1) = 0$. Navíc může existovat další kořen $\theta(s) = 0$. Pokud takový kořen existuje pak je jednoznačně určený a kladný a nazýváme ho Lundebergův koeficient a značíme jej γ . γ tedy splňuje rovnici

$$\lambda(\hat{m}_X(\gamma) - 1) - p\gamma = 0.$$

Příklad. Uvažujme například exponenciální rozdělení $X \sim \text{Exp}(\delta)$, jeho momentová vytvářející funkce je $\hat{m}_X(s) = \frac{\delta}{\delta-s}$, $s < \delta$. Řešíme tedy rovnici

$$\theta(s) = \lambda\left(\frac{\delta}{\delta-s} - 1\right) - ps = 0,$$

tj.

$$s(\lambda - p(\delta - s)) = 0.$$

Protože hledáme kladné řešení, je γ řešením rovnice

$$\lambda - p(\delta - s) = 0$$

a

$$\gamma = \delta - \frac{\lambda}{p}.$$

Pro většinu ostatních rozdělení není řešení tak snadné, ale poměrně snadno zjistíme, kdy Lundebergův koeficient existuje. Základní podmínkou existence Lundebergova koeficientu je $\hat{m}_X(s) < \infty$ pro nějaké $s > 0$. To vylučuje rozdělení s těžkými chvosty (v praxi často používaná), protože pro ty platí $\hat{m}_X(s) = \infty$ pro všechna $s > 0$. Další podmínky existence Lundebergova koeficientu udává následující věta.

Věta 5.16. Nechť existuje $s_\infty \in \overline{\mathbb{R}}$ takové, že $\hat{m}_X(s) < \infty$ pokud $s < s_\infty$ a $\lim_{\substack{s \nearrow s_\infty}} \hat{m}_X(s) = \infty$. Pak existuje jednoznačně určené řešení $\gamma > 0$ rovnice $\theta(s) = 0$.

Důkaz. [Rolski et al., 1998] □

I když Lundebergův koeficient existuje, nemusí být vždy snadné ho určit přesně, a proto bylo odvozeno jeho omezení.

Věta 5.17. Platí

$$\gamma < \frac{2(1 - \lambda \mathbb{E}X)}{\lambda \mathbb{E}X^2}.$$

Důkaz. [Asmussen, 2000] □

V praxi nastává však ještě jeden problém. V okamžiku, kdy vycházíme z reálných dat, neznáme ani λ ani rozdělení přírušků F_X , a přesto bychom potřebovali alespoň nějaký odhad γ . Označme \hat{m}_T odhad \hat{m}_X a λ_T odhad λ na základě X_1, \dots, X_{N_T} :

$$\hat{m}_T(s) = \frac{1}{N_T} \sum_{i=1}^{N_T} e^{sX_i}, \quad \lambda_T = \frac{N_T}{T}.$$

Definujme funkci

$$\theta_T(s) = \lambda_T(\hat{m}_T(s) - 1) - ps$$

a γ_T definujme rovnicí $\theta_T(\gamma_T) = 0$. Pokud $N_T = 0$, pak \hat{m}_T a tedy i γ_T nejsou definované. A také pokud $\frac{\lambda_T}{N_T}(U_1 + \dots + U_{N_T}) > 1$ pak $\gamma_T < 0$. Nicméně pravděpodobnosti obou těchto jevů konvergují k nule pro $T \rightarrow \infty$.

Věta 5.18. $\gamma_T \xrightarrow{as} \gamma$ pro $T \rightarrow \infty$. Pokud navíc $\hat{m}_X(2\gamma) < \infty$ pak

$$\gamma_T - \gamma \sim N\left(0, \frac{\sigma_\gamma^2}{T}\right),$$

$$kde \sigma_\gamma^2 = \frac{\lambda\theta(2\gamma)}{(\theta'(\gamma))^2}.$$

Důkaz. [Asmussen, 2000] □

Asymptotická normalita platí i pro odhady momentové vytvořující funkce a $\theta(s)$.

Věta 5.19. Pro $T \rightarrow \infty$ platí

$$\hat{m}_T(\gamma) \sim N\left(\hat{m}_X(\gamma), \frac{\hat{m}_X(2\gamma) - \hat{m}_X(\gamma)^2}{\lambda T}\right),$$

$$\theta_T(\gamma) \sim N\left(0, \frac{\theta(2\gamma)}{T}\right).$$

Důkaz. [Asmussen, 2000] □

Díky větě 5.18 platí i asymptotická normalita veličiny $e^{-\gamma_T u}$ používané v Lundebergových nerovnostech a mezích (viz další kapitola).

$$e^{-\gamma_T u} \sim N(e^{-\gamma u}, u^2 e^{-2\gamma u} \frac{\sigma_\gamma^2}{T})$$

a snadno dostáváme $1 - \alpha$ konfidenční interval pro $e^{-\gamma u}$

$$\left(e^{-\gamma_T u} - u\left(\frac{\alpha}{2}\right) \cdot ue^{-\gamma_T u} \frac{\sigma_{\gamma,T}^2}{\sqrt{T}}, e^{-\gamma_T u} + u\left(\frac{\alpha}{2}\right) \cdot ue^{-\gamma_T u} \frac{\sigma_{\gamma,T}^2}{\sqrt{T}} \right),$$

kde $\sigma_{\gamma,T}^2 = \frac{\lambda_T \theta(2\gamma_T)}{(\theta'_T(\gamma_T))^2}$ je výběrová varianta σ_γ^2 a $u\left(\frac{\alpha}{2}\right)$ je kritická hodnota (tj. $(1 - \frac{\alpha}{2})$ -kvantil) normálního rozdělení.

5.4.2 Lundebergovy meze

V této části uvedeme různá omezení pravděpodobnosti ruinování $\psi(u)$. Připomeňme, že stále předpokládáme podmínu čistého zisku (5.1).

Věta 5.20. (*Oboustranné Lundebergovy meze*)

Předpokládejme, že existuje Lundebergův koeficient $\gamma > 0$ a označme $x_0 = \sup\{x : F_X(x) < 1\}$. Potom

$$a_- e^{-\gamma u} \leq \psi(u) \leq a_+ e^{-\gamma u} \quad \forall u \geq 0,$$

kde

$$a_- = \inf_{x \in [0, x_0)} \frac{e^{\gamma x} \int_x^\infty \bar{F}_X(y) dy}{\int_x^\infty e^{\gamma y} \bar{F}_X(y) dy}, \quad a_+ = \sup_{x \in [0, x_0)} \frac{e^{\gamma x} \int_x^\infty \bar{F}_X(y) dy}{\int_x^\infty e^{\gamma y} \bar{F}_X(y) dy}.$$

Důkaz. [Rolski et al., 1998] □

Zřejmě platí $a_- \leq a_+$ a protože $\int_x^\infty e^{\gamma y} \bar{F}_X(y) dy \geq e^{\gamma x} \int_x^\infty \bar{F}_X(y) dy$ máme $a_+ \leq 1$. Navíc lze ukázat, že nerovnosti ve větě 5.20 jsou ostré pokud platí $a_- < \psi(0)$, $a_+ > \psi(0)$ (viz např. [Rolski et al., 1998] paragraf 5.4).

V některých případech si vystačíme s jednodušší variantou Lundebergových mezí.

Věta 5.21. (*Cramér-Lundebergova nerovnost*)

Nechť existuje Lundebergův koeficient $\gamma > 0$, potom

$$\psi(u) \leq e^{-\gamma u}.$$

Důkaz. [Cipra, 1991] □

Dále se zaměříme na asymptotické chování pravděpodobnosti ruinování pro velká u a ukážeme, že existuje limita pro $\psi(u)e^{\gamma u}$.

Věta 5.22. (*Cramér-Lundebergova*)

Předpokládejme, že existuje Lundebergův koeficient $\gamma > 0$. Pokud $\hat{m}'_X(\gamma) < \infty$, pak

$$\lim_{u \rightarrow \infty} \psi(u)e^{\gamma u} = \frac{p - \lambda \mathbb{E}X}{\lambda \hat{m}'_X(\gamma) - p}.$$

Pokud $\hat{m}'_X(\gamma) = \infty$, pak $\lim_{u \rightarrow \infty} \psi(u)e^{\gamma u} = 0$.

Důkaz. [Rolski et al., 1998] □

Za předpokladů předchozí věty se používá Cramér-Lundebergova approximace

$$\psi(u) \approx \frac{p - \lambda \mathbb{E}X}{\lambda \hat{m}'_X(\gamma) - p} e^{-\gamma u}. \quad (5.7)$$

Věta 5.23. *Cramér-Lundebergova approximace (5.7) platí přesně \iff přírůstky $X_i, i \in \mathbb{N}$ mají exponenciální rozdělení.*

Důkaz. [Rolski et al., 1998] □

5.4.3 Rozdělení s těžkými chvosty

V předchozí části jsme předpokládali existenci Lundebergova koeficientu a tím jsme automaticky vyřadili rozdělení s těžkými chvosty jako možné rozdělení přírůstků. Naneštěstí rozdělení s těžkými chvosty jsou v některých oblastech velmi potřebná, protože spíše odpovídají skutečnosti, např. v úrazovém pojistění, nebo pojistění proti živelným katastrofám. Příklady použití rozdělení s těžkými chvosty můžeme najít např. v [Cipra, 1991] nebo [Rolski et al., 1998].

V případě přírůstků s těžkými chvosty je asymptotické chování pravděpodobnosti ruinování $\psi(u)$ velmi odlišné od toho z věty (5.22). Vezmeme-li v úvahu, že $\lim_{x \rightarrow \infty} e^{sx} \bar{F}(x) = \infty$ (viz (2.1)) a Pollaczek-Kinchinovu formuli (5.6), dostáváme pro $u \rightarrow \infty$ a pro každé $s > 0$

$$\psi(u)e^{su} \geq \left(1 - \frac{\lambda\mu}{p}\right) \frac{\lambda\mu}{p} (1 - F_X^s(u)) e^{su} \rightarrow \infty. \quad (5.8)$$

Pro speciální rozdělení s těžkými chvosty máme následující asymptotický výsledek.

Věta 5.24. (*Cramér-Lundebergova pro těžké chvosty I*)

Položme $\rho = \frac{\lambda\mu}{p}$ a předpokládejme, že integrovaný chvost F_X^s (definováno v (5.5)) je subexponenciální. Pak

$$\lim_{u \rightarrow \infty} \frac{\psi(u)}{F_X^s(u)} = \frac{\rho}{1 - \rho}.$$

Důkaz. [Rolski et al., 1998]

□

Na základě předchozí věty se používá approximace

$$\psi(u) \approx \frac{\rho}{1 - \rho}(1 - F_X^s(u)). \quad (5.9)$$

Věta 5.25. (*Cramér-Lundebergova pro těžké chvosty II*)

Položme opět $\rho = \frac{\lambda\mu}{p}$, pak následující tvrzení jsou ekvivalentní.

1. $F_X^s \in \mathcal{S}$.
2. $1 - \psi(u) \in \mathcal{S}$.
3. $\lim_{u \rightarrow \infty} \frac{\psi(u)}{F_X^s(u)}(u) = \frac{\rho}{1 - \rho}$.

Důkaz. [Embrechts et al., 1997].

□

Aproximace Cramér-Lundebergova typu (5.9) je možná pouze pro rozdělení, pro která je integrovaný chvost F_X^s subexponenciální, ale nás by spíše zajímalo, jaké podmínky musí splňovat distribuční funkce přírůstků F_X . Obecně totiž neplatí, že by $F_X \in \mathcal{S}$ implikovalo $F_X^s \in \mathcal{S}$. Obecně neplatí ani opačná implikace.

K tomu potřebujeme ještě hlouběji zkoumat rozdělení s těžkými chvosty. Nejprve si proto rozšíříme definici pomalu se měnících funkcí.

Definice 5.26. Říkáme, že kladná Lebesgueovsky měřitelná funkce h na $(0, \infty)$ je regulárně se měnící (regularly varying) v ∞ indexu $\alpha \in \mathbb{R}$ a značíme $h \in \mathcal{R}_\alpha$, pokud

$$\lim_{x \rightarrow \infty} \frac{h(tx)}{h(x)} = t^\alpha \quad \forall t > 0.$$

Regulárně se měnící funkce indexu α jsou např. funkce x^α , $x^\alpha \log x$, $\alpha \in \mathbb{R}$.

Dále budeme označovat

$$\mathcal{R} = \{F\text{distribuční funkce na } (0, \infty) : \bar{F} \in \mathcal{R}_{-\alpha} \text{ pro nějaké } \alpha \geq 0\}.$$

Uvažujeme-li distribuční funkce s chvostem tvaru $\bar{F}(x) = x^{-\alpha}L(x)$ pro $\alpha \geq 0$ a $L \in \mathcal{R}_0$ jako v kapitole 2.4.2 dostáváme

$$\lim_{x \rightarrow \infty} \frac{\bar{F}(tx)}{\bar{F}(x)} = \lim_{x \rightarrow \infty} \frac{(tx)^{-\alpha}L(tx)}{x^{-\alpha}L(x)} = t^{-\alpha}$$

a tedy pro takovéto distribuční funkce platí $\bar{F} \in \mathcal{R}_{-\alpha}$.

Uvedeme ještě podrobnější dělení distribucí s těžkými chvosty. Označme si všechna rozdelení s těžkými chvosty

$$\mathcal{K} = \{F \text{distribuční funkce na } (0, \infty) : \hat{m}(s) = \infty, \forall s > 0\}.$$

Dále označíme

$$\mathcal{L} = \{F \text{distribuční funkce na } (0, \infty) : \lim_{x \rightarrow \infty} \frac{\bar{F}(x-y)}{\bar{F}(x)} = 1, \forall y > 0\},$$

jak už víme, pokud je $F \in \mathcal{L}$ pak $F \in \mathcal{S}$, ale tato skupina je širší.

Definujme si ještě poslední uvažovaný typ funkcí. Pro tyto funkce se nám nepodařilo nalézt vhodný český ekvivalent, proto nám promiňte použití anglického termínu.

Definice 5.27. Říkáme, že distribuční funkce nezáporné náhodné veličiny je dominatedly varying v nekonečnu pokud

$$\limsup_{x \rightarrow \infty} \frac{\bar{F}(x/2)}{\bar{F}(x)} < \infty$$

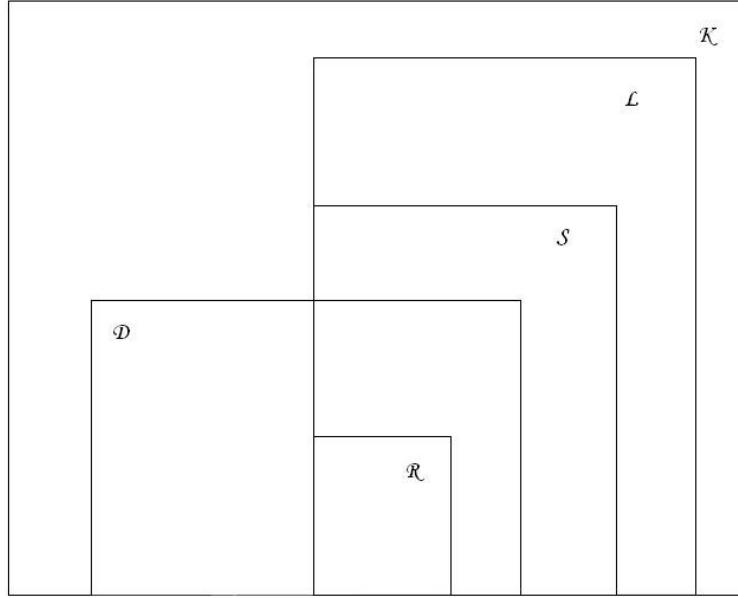
a množinu dominatedly varying distribučních funkcí značíme \mathcal{D} .

Mezi jednotlivými třídami rozdelení s těžkými chvosty platí následující vztahy:

1. $\mathcal{R} \subset \mathcal{S} \subset \mathcal{L} \subset \mathcal{K}$ a $\mathcal{R} \subset \mathcal{D} \subset \mathcal{K}$,
2. $(\mathcal{L} \cap \mathcal{D}) \subset \mathcal{S}$,
3. $\mathcal{D} \not\subseteq \mathcal{S}$ a $\mathcal{S} \not\subseteq \mathcal{D}$.

Veškeré relace shrnuje obrázek 5.1. $\mathcal{R} \subset \mathcal{S}$ a tedy vety 5.24 a 5.25 platí i pro distribuční funkce, pro které je $F_X^s \in \mathcal{R}$.

Přejděme nyní k podmínkám subexponenciality pro integrovaný chvost F_X^s , kvůli kterým jsme prováděli celou analýzu.



Obrázek 5.1: Třídy rozdělení s těžkými chvosty

Věta 5.28. (*Podmínky subexponenciality*)

1. Pokud $F \in \mathcal{L} \cap \mathcal{D}$, pak $F \in \mathcal{S}$.
2. Pokud náhodná veličina X s distribuční funkcí F_X má konečnou střední hodnotu μ a $F_X \in \mathcal{D}$, pak $F_X^s \in \mathcal{L} \cap \mathcal{D}$ a díky bodu 1. také $F_X^s \in \mathcal{S}$.

Důkaz. [Embrechts et al., 1997] □

Následující věta uvádí postačující podmínky pro F_X , aby integrovaný chvost F_X^s byl subexponenciální.

Věta 5.29. (*Podmínky subexponenciality F_X^s*)

Nechť F_X je distribuční funkce s hustotou f_X a kumulovanou rizikovou funkcí $\Lambda = -\log \bar{F}$ a označme $q(x) = \frac{f(x)}{\bar{F}(x)}$. Pokud platí některá z následujících podmínek pak $F_X^s \in \mathcal{S}$.

1. $\limsup_{x \rightarrow \infty} xq(x) < \infty$, nebo
2. $\lim_{x \rightarrow \infty} q(x) = 0$, $\lim_{x \rightarrow \infty} xq(x) = \infty$ a platí některá z následujících podmínek:
 - (a) $\limsup_{x \rightarrow \infty} \frac{xq(x)}{\Lambda(x)} < 1$, nebo

- (b) $q \in \mathcal{R}_\delta$ pro $-1 \leq \delta \leq 1$, nebo
- (c) $\Lambda \in \mathcal{R}_\delta$ pro $0 < \delta < 1$ a $q(x)$ je pro dostatečně velká x nerostoucí nebo
- (d) $q(x)$ jde pro dostatečně velká x monotónně k 0, $q \in \mathcal{R}_0$ a $\Lambda(x) - xq(x) \in \mathcal{R}_1$.

Důkaz. Větu jsme převzali z [Embrechts et al., 1997]. Autoři ve větě shrnují netriviální výsledky z několika článků, zájemce o důkaz odkazujeme na citace uvedené v [Embrechts et al., 1997]. \square

Na základě této věty lze dokázat, že pro logaritmicko-normální a Weibullovo rozdělení s parametrem $b < 1$ je $F_X^s \in \mathcal{S}$.

5.4.4 Aproximace na základě momentů

Pomocí momentové vytvářející funkce snadno spočteme momenty složeného bodového procesu, a proto by nás mohlo napadnout použít Čebyšovu nerovnost jako bodové omezení. Naneštěstí Čebyšova nerovnost je příliš jednoduchá a omezení získané na jejím základě je skutečným hodnotám velmi vzdálené a prakticky nepoužitelné.

V teorii rizika se používají hlavně dvě metody approximace pravděpodobnosti ruinování $\psi(u)$, a to De Vylderova a Beekman-Bowersova. V obou případech se snažíme nahradit proces $\{R(t), t \geq 0\}$ procesem $\{R'(t), t \geq 0\}$ takovým, že se shodují jejich momenty až do řádu $n \geq 1$ a $\{R'(t)\}$ je vybrán tak, že je pro něj snažší odvodit jeho pravděpodobnost ruinování $\psi'(u)$.

1. De Vylderova approximace

Pro exponenciálně rozdělené přírůstky ($X \sim Exp(\delta)$) umíme odvodit pravděpodobnosti ruinování přesně:

$$\psi(u) = \frac{\lambda}{p\delta} e^{-u(\delta - \frac{\lambda}{p})}.$$

Idea De Vylderovy approximace je nahradit $\{R(t)\}$ procesem $\{R'(t)\}$, který má exponenciálně rozdělené přírůstky s parametrem δ' a

$$\mathbb{E}(R(t))^k = \mathbb{E}(R'(t))^k \quad t \geq 0, \quad k = 1, 2, 3.$$

Snadno spočteme, že

$$\mathbb{E}R(t) = u + pt - \lambda t \mathbb{E}X$$

$$\begin{aligned}\text{var } R(t) &= \lambda t \mathbb{E} X^2 \\ \mathbb{E}(R(t) - \mathbb{E} R(t))^3 &= -\lambda t \mathbb{E} X^3.\end{aligned}$$

Rovnost prvních tří momentů tedy platí když

$$p - \lambda \mathbb{E} X = p' - \frac{\lambda'}{\delta'} \quad \lambda \mathbb{E} X^2 = \frac{2\lambda'}{(\delta')^2} \quad \lambda \mathbb{E} X^3 = \frac{6\lambda'}{(\delta')^3}.$$

Řešením této soustavy rovnic dostáváme podmínky na parametry procesu $\{R'(t)\}$

$$\delta' = \frac{3\mathbb{E} X^2}{\mathbb{E} X^3} \quad \lambda' = \frac{\lambda \mathbb{E} X^2 (\delta')^2}{2} \quad p' = p - \lambda \mathbb{E} X + \frac{\lambda'}{\delta'}.$$

De Vylderova approximace pravděpodobnosti ruinování pak je

$$\psi(u) \approx \frac{\lambda'}{p'\delta'} e^{-u(\delta' - \frac{\lambda'}{p'})}.$$

Samozřejmě v případě exponenciálně rozdělených přírůstků platí approximace přesně a numerické příklady ukazují, že platí poměrně přesně i pro ostatní rozdělení (viz [Rolski et al., 1998]).

2. Beekman-Bowersova approximace

Tato metoda je založena na změně disribuční funkce. Uvažujme nějakou náhodnou veličinu Z s distribuční funkcí $F_Z(x) = 1 - \frac{p\psi(x)}{\lambda \mathbb{E} X}$ (pak $\psi(u) = \frac{\lambda \mathbb{E} X}{p}(1 - F_Z(u))$). Víme, že $\psi(0) = \frac{\lambda \mathbb{E} X}{p}$, tedy $F_Z(0) = 0$ a tedy $F_Z(x)$ je distribuční funkce kladné náhodné veličiny Z . Navíc podle diferenciálně-integrální věty 5.6 je funkce F_Z absolutně spojitá a má tedy hustotu $f_Z(x) = \frac{p}{\lambda \mathbb{E} X} \bar{\psi}'_+(x)$. Momentová vytvořující funkce náhodné veličiny Z je

$$\hat{m}_Z(s) = \mathbb{E} e^{sZ} = \int_0^\infty e^{sx} dF_Z(x) = \int_0^\infty e^{sx} \frac{p}{\lambda \mathbb{E} X} \bar{\psi}'_+(x) dx.$$

Integrací per-partes a dalšími standardními úpravami dostaneme

$$\hat{m}_Z(s) = \frac{p}{\lambda \mathbb{E} X} \left(\frac{-s(p - \lambda \mathbb{E} X)}{-ps - \lambda(1 - \hat{m}_X(s))} - \left(1 - \frac{\lambda \mathbb{E} X}{p}\right) \right),$$

což snadno převedeme na tvar

$$\hat{m}_Z(s) = 1 - \frac{p}{\lambda \mathbb{E} X} + \frac{p(p - \lambda \mathbb{E} X)}{\lambda \mathbb{E} X} \frac{s}{ps - \lambda(\hat{m}_X(s) - 1)}.$$

Idea Beekman-Bowersovy approximace je nahradit distribuční funkci $F_Z(x)$ distribuční funkcí $F'(x)$ gamma rozdělení $\Gamma(a, b)$ tak, aby se shodovaly první dva momenty.

Momenty náhodné veličiny Z odvodíme derivací její momentové vytvořující funkce $\hat{m}_Z(s)$.

$$\begin{aligned}\hat{m}'_Z(s) &= \frac{p(p - \lambda \mathbb{E}X)}{\lambda \mathbb{E}X} \frac{ps - \lambda(\hat{m}_X(s) - 1) - s(p - \lambda\hat{m}'_X(s))}{(ps - \lambda(\hat{m}_X(s) - 1))^2} \\ &= \frac{p(p - \lambda \mathbb{E}X)}{\mathbb{E}X} \frac{s\hat{m}'_X(s) - \hat{m}_X(s) + 1}{(ps - \lambda(\hat{m}_X(s) - 1))^2}.\end{aligned}\quad (5.10)$$

L'Hospitalovým pravidlem snadno spočteme $\lim_{s \rightarrow 0} \frac{ps - \lambda(\hat{m}_X(s) - 1)}{s} = p - \lambda \mathbb{E}X$ a pak

$$\begin{aligned}\mathbb{E}Z &= \lim_{s \rightarrow 0} \hat{m}'_Z(s) = \frac{p(p - \lambda \mathbb{E}X)}{\mathbb{E}X} \frac{1}{(p - \lambda \mathbb{E}X)^2} \cdot \lim_{s \rightarrow 0} \frac{s\hat{m}'_X(s) - \hat{m}_X(s) + 1}{s^2} \\ &= \frac{p \mathbb{E}X^2}{2 \mathbb{E}X(p - \lambda \mathbb{E}X)}.\end{aligned}\quad (5.11)$$

Analogicky spočteme

$$\begin{aligned}\hat{m}''_Z(s) &= \frac{p(p - \lambda \mathbb{E}X)}{\mathbb{E}X(ps - \lambda(\hat{m}_X(s) - 1))^3} \cdot \left(s\hat{m}''_X(s)(ps - \lambda(\hat{m}_X(s) - 1)) - \right. \\ &\quad \left. - 2(s\hat{m}'_X(s) - (\hat{m}_X(s) - 1))(p - \lambda\hat{m}'_X(s)) \right)\end{aligned}\quad (5.12)$$

a

$$\begin{aligned}\mathbb{E}Z^2 &= \lim_{s \rightarrow 0} \hat{m}''_Z s = \frac{p(p - \lambda \mathbb{E}X)}{\mathbb{E}X(p - \lambda \mathbb{E}X)^3} \left(\frac{\mathbb{E}X^3(p - \lambda \mathbb{E}X)}{3} + \frac{\lambda(\mathbb{E}X^2)^2}{2} \right) \\ &= \frac{p}{\mathbb{E}X} \left(\frac{\mathbb{E}X^3}{3(p - \lambda \mathbb{E}X)} + \frac{\lambda(\mathbb{E}X^2)^2}{2(p - \lambda \mathbb{E}X)^2} \right).\end{aligned}\quad (5.13)$$

Jak známo, první dva momenty gamma rozdělení $\Gamma(a, b)$ jsou $\mathbb{E}X = \frac{a}{b}$ a $\mathbb{E}X^2 = \frac{b(b+1)}{a^2}$. Parametry a, b jsou tedy dány rovnicemi:

$$\frac{a}{b} = \frac{p \mathbb{E}X^2}{2 \mathbb{E}X(p - \lambda \mathbb{E}X)} \quad \frac{b(b+1)}{a^2} = \frac{p}{\mathbb{E}X} \left(\frac{\mathbb{E}X^3}{3(p - \lambda \mathbb{E}X)} + \frac{\lambda(\mathbb{E}X^2)^2}{2(p - \lambda \mathbb{E}X)^2} \right).$$

Beekman-Bowersova approximace pravděpodobnosti ruinování $\psi(u) = \frac{\lambda \mathbb{E}X}{p} \cdot (1 - F_Z(u))$ pak je

$$\psi(u) \approx \frac{\lambda \mathbb{E}X}{p} \cdot (1 - F'(u)),$$

kde $F'(u)$ je distribuční funkce rozdělení $\Gamma(a, b)$.

Ještě poznamenejme, že pokud pro parametr a platí $2a \in \mathbb{N}$ pak náhodná veličina $2bZ$ má χ^2_{2a} rozdělení a Beekman-Bowersovu aproximaci si můžeme ještě zjednodušit approximací parametru $2a$ přirozenými čísly a výpočtem pomocí χ^2 rozdělení.

Poznámka 5.30. Úplně jiné odvození Beekman-Bowersovy approximace nabízí [Asmussen, 2000] v 7. kapitole.

5.4.5 Další approximační metody

Další dvě approximační metody předpokládají, že intenzita příchodu požadavků λ je blízká některé z krajních hodnot 0 nebo λ_{max} , kde λ_{max} je hodnota, pro kterou je podmínka čistého zisku splněna s rovností, a pravděpodobnost ruinování approximují pomocí těchto mezních případů. Třetí approximační metoda využívá interpolaci mezi těmito krajními případy.

1. Heavy traffic approximace

Pojem heavy traffic pochází z teorie front, ale má interpretaci i v teorii rizika. V pojišťovací praxi se běžně můžeme setkat se situací, kdy pojištěné v průměru jen mírně převyšuje očekávané požadavky. Heavy traffic podmínka známená, že je sice splněna podmínka čistého zisku (5.1), ale p je jen o málo větší než $\lambda \mathbb{E}X$. Formálně toto budeme reprezentovat pomocí limity, kdy λ konverguje k maximální hodnotě $\lambda \nearrow \lambda_{max}$, pro kterou platí $p = \lambda_{max} \mathbb{E}X$. Rozdělení přírůstků zůstává neměnné, dané distribuční funkcí F_X .

Věta 5.31. Označme $M = \sup_{0 \leq t < \infty} S(t)$, pak pro $\lambda \nearrow \lambda_{max}$ náhodná veličina $(\lambda_{max} - \lambda)M$ konverguje v distribuci k exponenciálnímu rozdělení s parametrem $\delta = \frac{2(\mathbb{E}X)^2}{\mathbb{E}X^2}$.

Důkaz. [Asmussen, 2000] □

Důsledek 5.32. Nechť $\lambda \nearrow \lambda_{max}$ a $u \rightarrow \infty$ tak, že $(\lambda_{max} - \lambda)u \rightarrow v$, pak $\psi(u) \rightarrow e^{-\delta v}$

Důkaz. Stačí si uvědomit, že $\psi(u) = P((\lambda_{max} - \lambda)M > (\lambda_{max} - \lambda)u)$. □

Na základě uvedených výsledků se používá approximace

$$\psi(u) \approx e^{-\delta(\lambda_{max} - \lambda)u}.$$

2. Light traffic approximace

Pojem light traffic také pochází z teorie front a označuje situaci, kdy pojistné je v průměru mnohem větší než očekávané přírůstky. Light traffic podmínka znamená, že λ je malé ve srovnání s $\mathbb{E}X$. Matematicky tuto skutečnost opět reprezentujeme limitou $\lambda \searrow 0$ a rozdelení přírůstků zůstává pevné dané distribuční funkcí F_X . Samozřejmě v teorii rizika je mnohem častější situace blížící se stavu popsanému v heavy traffic approximaci, než situace, kdy je $\lambda \ll \mathbb{E}X$. Light traffic situace je v teorii rizika považována spíše za doplněk k heavy traffic stavu a její využití je hlavně pro interpolaci.

Věta 5.33. *Pro $\lambda \searrow 0$ platí*

$$\psi(u) \stackrel{\cdot}{\approx} \lambda \int_u^\infty \bar{F}_X(x) dx = \lambda \mathbb{E}[X - u; X > u] = \lambda \mathbb{E}(X - u)^+.$$

Důkaz. [Asmussen, 2000] □

Heuristicky approximace na základě věty 5.33 znamená, že ruinování může nastat pouze v čase σ_1 prvního přírůstku, tedy $\psi(u) \stackrel{\cdot}{\approx} P(X - \sigma_1 > u)$ a

$$P(X - \sigma_1 > u) = \int_0^\infty \bar{F}_X(x + u) \lambda e^{-\lambda x} dx \stackrel{\cdot}{\approx} \lambda \int_u^\infty \bar{F}_X(x) dx.$$

3. Interpolace mezi light a heavy traffic approximací

Pokud parametr λ není blízký ani jedné z krajiných hodnot, můžeme si představu o pravděpodobnosti ruinování udělat pomocí interpolace mezi light traffic a heavy traffic approximací. První myšlenka by mohla být položit

$$\psi(u) \stackrel{\cdot}{\approx} (1 - \frac{\lambda}{\lambda_{max}}) \lim_{\lambda \searrow 0} \psi(u) + \frac{\lambda}{\lambda_{max}} \lim_{\lambda \nearrow \lambda_{max}} \psi(u) \quad (5.14)$$

$$= (1 - \frac{\lambda}{\lambda_{max}}) \cdot 0 + \frac{\lambda}{\lambda_{max}} \cdot 1 = \frac{\lambda}{\lambda_{max}}, \quad (5.15)$$

což je ovšem naprostě nepoužitelná approximace.

Použitelné výsledky dostaneme, zkombinujeme-li pravděpodobnost ruinování $\psi(u)$ pro exponenciální přírůstky, kterou známe přesně, s výše

uvedenými heavy a light traffic approximacemi. Po úpravách dostáváme odhad ve tvaru:

$$\psi(u) \approx \lambda\mu(1 - \lambda\mu)\lambda_{max} \int_{u(1-\lambda\mu)}^{\infty} \bar{F}_X(x)dx + (\lambda\mu)^2 e^{-\delta u(\lambda_{max} - \lambda)},$$

kde $\delta = \frac{2(\mathbb{E}X)^2}{\mathbb{E}X^2}$ je parametr z heavy traffic approximace. Odvození uvádí [Asmussen, 2000].

5.5 Ruinování v konečném čase

Jak už bylo zmíněno, pro pravděpodobnosti ruinování v konečném čase je ještě mnohem složitější odvodit prakticky použitelné vzorce, ale pro složený Poissonův proces s deterministickými a exponenciálními přírůstky má $\psi(u, x)$ poměrně jednoduchý tvar.

5.5.1 Deterministické přírůstky

Předpokládejme, že $P(X = \mu) = 1$ pro nějaké $\mu > 0$, potom $C(t) = \sum_{i=1}^{N(t)} \mu = \mu N(t)$ a ruinování nastává, když $C(t)$ překročí hranici $u + pt$. Zajímáme se tedy o pravděpodobnost $P(\mu N(t) > u + pt) = P(N(t) > \frac{u+pt}{\mu})$ pro nějaké $t \leq x$ a o $N(t)$ víme že má Poissonovo rozdělení s parametrem λt .

Věta 5.34. *Předpokládejme, že $P(X = \mu) = 1$ pro nějaké $\mu > 0$. Bud' $\phi(u, x, y) = P(\tau(u) > x, R(x) = y)$.*

1. *Pak pro $y > 0$ je $\phi(u, 0, y) = \delta_u(y)$. A dále, když $\frac{u}{\mu}, \frac{xp}{\mu} \in \mathbb{N}$, pak pro y takové, že $\frac{y}{\mu} \in \mathbb{N}$, platí*

$$\phi(u, x, y) = \sum_{i=0}^{u/\mu} e^{-\frac{\lambda\mu}{p}} \frac{1}{i!} \left(\frac{\lambda\mu}{p}\right)^i \phi(u - (i-1)\mu, x - \frac{\mu}{p}, y). \quad (5.16)$$

2. *Nechť $z = \lfloor \frac{u}{\mu} \rfloor + 1 - \frac{u}{\mu}$ a předpokládejme, že $\frac{u+px}{\mu} \in \mathbb{N}$. Pak pro y takové, že $\frac{y}{\mu} \in \mathbb{N}$, platí*

$$\phi(u, x, y) = \sum_{i=0}^{\lfloor u/\mu \rfloor} e^{-\frac{\lambda\mu}{p}} \frac{1}{i!} \left(\frac{\lambda\mu z}{p}\right)^i \phi(u - (i-z)\mu, x - \frac{\mu z}{p}, y). \quad (5.17)$$

3. Nechť jsou $u, x \geq 0$ libovolné a nechť $z = \frac{u+xp}{\mu} - \lfloor \frac{u+xp}{\mu} \rfloor$. Pak pro pravděpodobnost přežití $\bar{\psi}(u, x)$ platí

$$\bar{\psi}(u, x) = \begin{cases} \sum_{k=1}^{\lfloor \frac{u+xp}{\mu} \rfloor} \phi(u, x - \frac{z\mu}{p}, k\mu) P(N(z\mu) \leq pk) & \text{pro } px \geq z\mu, \\ P(N(t)\mu > u) & \text{jinak.} \end{cases} \quad (5.18)$$

Důkaz. [Rolski et al., 1998] \square

Poznámka 5.35. Uvědomme si, že $\bar{\psi}(u, x) = 1 - \psi(u, x) = 1 - P(\tau(u) \leq x) = P(\tau(u) > x)$ a $\bar{\psi}(u, x) = \sum_y \phi(u, x, y)$, kde sčítáme přes všechny možné hodnoty y (je-li jich pouze spočetně mnoho). Pokud je $y \in \mathbb{N}$ pak rovnice (5.16) a (5.17) zůstávají v platnosti i když $\phi(u, x, y)$ nahradíme $\bar{\psi}(u, x)$. V případě kdy $\frac{R(x)}{\mu} \notin \mathbb{Z}$ se bez $\phi(u, x, y)$ neobejdeme. Navíc pokud ruinování nastalo před časem \tilde{x} , kdy $\frac{R(\tilde{x})}{\mu} \in \mathbb{Z}$, ale nikoli před časem $\tilde{x} - \frac{\mu}{p}$, nejsme pro $x \in (\tilde{x} - \frac{\mu}{p}, \tilde{x})$ schopni rozhodnout, zda ruinování nastalo před časem x nebo až po něm.

5.5.2 Exponenciální přírůstky

Exponenciálně rozdelené přírůstky jsou druhou variantou, kdy jsme schopni určit pravděpodobnosti ruinování analyticky.

Nejprve si definujme pomocnou funkci

$$J(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!n!}.$$

Tato funkce je úzce spjata s modifikovanou Besselovou funkcí

$$I_{\nu}(x) = \sum_{k=0}^{\infty} \frac{1}{k!\Gamma(\nu + k + 1)} \left(\frac{x}{2}\right)^{2k+\nu}, \quad x \in \mathbb{R}, \nu \in \mathbb{R}.$$

Platí $J(x) = I_0(2\sqrt{x})$. J má několik zajímavých vlastností, ale ty jsou potřeba jen pro důkaz následující věty a proto je nebudeme uvádět a zájemce odkazujeme na [Rolski et al., 1998].

Věta 5.36. Nechť $X \sim \text{Exp}(\delta)$, pak

$$\psi(u, x) = 1 - \exp \left\{ -\delta u - \left(1 + \frac{\delta p}{\lambda}\right) \lambda x \right\} g(\delta u + \delta p x, \lambda x),$$

kde

$$g(z, \theta) = J(\theta z) + \theta J'(\theta z) + \int_0^z e^{z-v} J(\theta v) dv - \frac{\lambda}{\delta p} \int_0^{\frac{\delta p \theta}{\lambda}} \exp \left\{ \frac{\delta p \theta}{\lambda} - v \right\} J \left(\frac{z \lambda v}{\delta p} \right) dv.$$

Důkaz. [Rolski et al., 1998] \square

Vzorec pro $\psi(u, x)$ je sice jednoduchý, ale výpočet $g(z, \theta)$ může být problematický.

Uvedeme ještě alternativní reprezentaci pravděpodobnosti ruinování $\psi(u, x)$.

Věta 5.37. Nechť $X \sim Exp(\delta)$, pak

$$\psi(u, x) = \frac{1}{c} \exp \left\{ -\frac{(c-1)\delta u}{c} \right\} - \frac{1}{\pi} \exp \{-\delta u - (1+c)x\} \int_0^\pi g(\delta u, \lambda x, y) dy,$$

kde $c = \frac{\delta p}{\lambda}$ a

$$g(\omega, \theta, y) = 2\sqrt{c} \frac{\exp \{(2\theta\sqrt{c} + \frac{\omega}{\sqrt{c}}) \cos y\}}{1 + c - 2\sqrt{c} \cos y} \left(\sin y \sin(y + \frac{\omega}{\sqrt{c}} \sin y) \right).$$

Důkaz. [Rolski et al., 1998] \square

5.5.3 Sealova formule

V této části nepředpokládáme žádné speciální rozdělení příruštků a pravděpodobnost přežití $\psi(u, x)$ vyjádříme pomocí rozdělení složeného bodového procesu $C(t) = \sum_{i=1}^{N(t)} X_i$.

Nechť X_i má hustotu f_X , pak rozdělení $C(t)$ můžeme vyjádřit jako

$$F_{C(t)}(x) = e^{-\lambda x} + \int_0^x \tilde{f}_{C(t)}(y) dy \quad x \geq 0,$$

kde $\tilde{f}_{C(t)}(y) = \sum_{n=1}^{\infty} \frac{(\lambda t)^n}{n!} e^{-\lambda t} f_X^{*n}(y)$. Formule (5.19) a (5.20) jsou dobře známé v pojistné matematice pod názvem Sealova formule.

Věta 5.38. (Sealova formule)

Předpokládejme, že $P(X > 0) = 1$.

1. Pak pro počáteční rezervu $u = 0$ je

$$\bar{\psi}(0, x) = \frac{1}{px} \mathbb{E}(R(x)_+) = \frac{1}{px} \int_0^{px} F_{C(x)}(y) dy. \quad (5.19)$$

2. Pro $u > 0$ a pro X s hustotou $f_X(y)$ je

$$\bar{\psi}(u, x) = F_{C(x)}(u + px) - p \int_0^u \bar{\psi}(0, x - y) \tilde{f}_{C(y)}(u + py) dy. \quad (5.20)$$

Důkaz. [Rolski et al., 1998] □

5.6 Kdy a jak nastává ruinování

V této části nadále předpokládáme složený Poissonův model (předpoklad 2) a zajímáme se o to, jestli je realizace vedoucí k ruinování něčím zvláštní a dá se na základě jejího vývoje usoudit, že přijde ruinování dříve než opravdu nastane. I v této části budeme předpokládat, že platí podmínka čistého zisku (5.1). Pokud by neplatila podmínka čistého zisku, víme, že ruinování nastává dříve nebo později pro skoro každou realizaci.

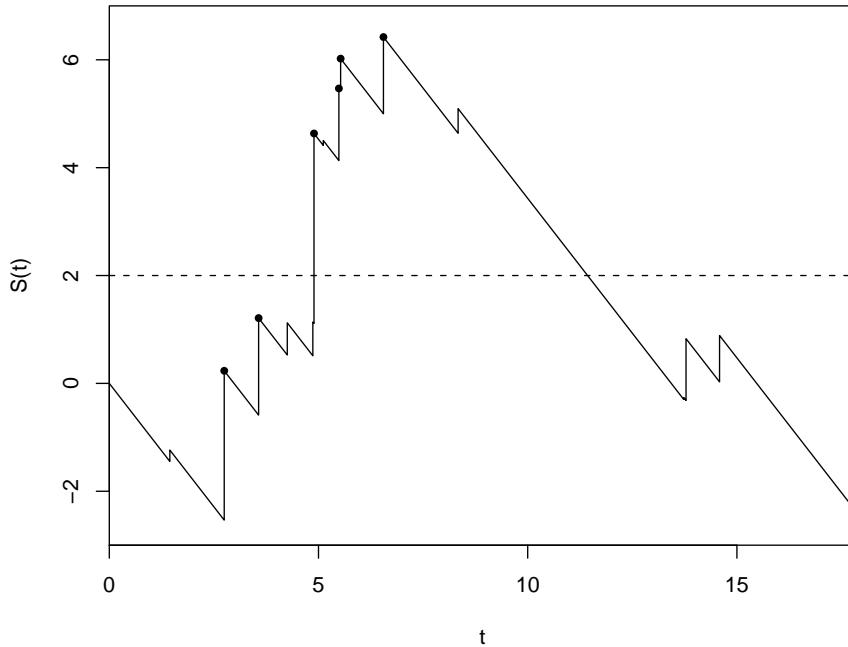
Ruinování zřejmě může nastat pouze v okamžiku, kdy nastává nějaká událost, tj. v čase $\sigma_n = T_1 + \dots + T_n$ pro $n \in \mathbb{N}$, proto v této části budeme s výhodou uvažovat náhodnou procházku $\{S_n, n \in \mathbb{N}\}$, která tvoří diskrétní kostru spojitého procesu $\{S(t), t \geq 0\}$. Připomeňme, že

$$S_n = \sum_{i=1}^n Y_i, \quad Y_i = X_i - pT_i$$

a zaved'me ještě nové označení, které souvisí s tzv. procesy rekordů (ladder heights process). Definujme časy rekordů (ladder indices)

$$\begin{aligned} \tau_+(0) &= 0, \\ \tau_+(1) &= \inf\{t > 0 : S(t) > 0\}, \\ \tau_+(n+1) &= \inf\{t > \tau_+(n) : S(t) > S(\tau_+(n))\}, \quad n \in \mathbb{N}, \end{aligned}$$

tradičně uvažujeme $\inf \emptyset = \infty$. A definujme náhodné veličiny $S(\tau_+(n))$, které se obvykle nazývají rekordy, a mají přesně takový význam. V časech $\tau(n)$, $n = 1, 2, \dots$ realizace procesu $\{S(t)\}$ dosahuje své nové nejvyšší hodnoty (n -tého rekordu) ve svém dosavadním vývoji a hodnota rekordu je



Obrázek 5.2: Realizace náhodného procesu $\{S(t)\}$ pro počáteční rezervu $u = 2$ a pojistné $p = 1$. Dosažené rekordy $S(\tau_+(n))$ jsou označeny body, ruinování způsobí třetí rekord.

$S(\tau_+(n))$. Realizaci procesu $\{S(t)\}$ s vyznačenými rekordy uvádíme na obrázku 5.2.

Díky stacionaritě a nezávislosti přírůstků, se zřejmě v každém rekordu $(\tau_+(k), S(\tau_+(k)))$, $k = 1, 2 \dots$ proces $\{S(t)\}$ restartuje. Standardní výsledek z teorie rekordů je, že úseky procesu $\{S(t)\}$ mezi jednotlivými rekordy představují nezávislé stochastické procesy, které mají stejnou intenzitu příchodu událostí a shodné rozdělení přírůstků.

Pro první rekord definujme

$$V = S(\tau_+(1)) \quad \text{a} \quad \hat{Z} = -S(\tau_+(1)-).$$

Veličiny V a \hat{Z} odpovídají dříve definovaným veličinám $Y^+(u)$, resp. $X^+(u)$ popisujícím překročení hodnoty u resp. přebytek do doby ruinování, pro počáteční rezervu $u = 0$. $A = V + \hat{Z}$ je velikost přírůstku, který způsobí ruinování. Následující výsledek o rozdělení V je klasický. Připomeňme, že $\tau(u)$ označuje čas ruinování při počáteční rezervě u .

Věta 5.39. Pro počáteční rezervu $u = 0$ platí následující tvrzení.

1. $P(V \leq x | \tau(0) < \infty) = P(\hat{Z} \leq x | \tau(0) < \infty) = F_X^s(x).$
2. $P(A \leq x | \tau(0) < \infty) = \frac{1}{\mu} \int_0^x y dF(y).$
3. Nechť U je náhodná veličina rovnoměrně rozdělená na $(0, 1)$ nezávislá na A , pak vektory (V, \hat{Z}) a $(UA, (1-U)A)$ jsou stejně rozdělené vzhledem k podmíněné míře $P(\cdot | \tau(0) < \infty)$.

Důkaz. [Feller, 1971] □

Označme $M = \sup_{t \geq 0} S(t) = \sup_{n \geq 1} S_n$, pak rekordy $S(\tau_+(n))$ určují rozdělení M . Náhodné veličiny $S(\tau_+(n)) - S(\tau_+(n-1))$ jsou nezáporné iid [Embrechts et al., 1997] a podle věty 5.39 pro rozdělení jejich chvostu platí

$$P(S(\tau_+(1)) > x) = P(\tau(0) < \infty)P(V > x | \tau(0) < \infty) = \psi(0)F_X^s(x), \quad x > 0.$$

Pokud $\tau(0) = \infty$ pak $S(t) \leq 0$ pro každé $t > 0$.

5.6.1 Lehké chvosty

V této podkapitole budeme předpokládat, že přírůstky X_i mají lehké chvosty, tj. jejich momentová vytvořující funkce $\hat{m}_X(s) < \infty$ pro nějaké $s > 0$. Budeme také předpokládat, že existuje Lundebergův koeficient γ a navíc předpokládejme, že $\int_0^\infty xe^{\gamma x} \bar{F}(x) dx < \infty$. Za těchto předpokladů platí věta 5.22 a $\lim_{u \rightarrow \infty} \psi(u)e^{\gamma u} = C$, kde C je kladná konstanta závislá na parametrech procesu.

Uvažujme přírůstky $Y_i = X_i - pT_i$ náhodné procházky $\{S_n\}$ a H_Y označme jejich distribuční funkci. Dále označme H_γ jí příslušnou Esscherovu transformaci distribuční funkce,

$$H_\gamma(x) = \int_{-\infty}^x e^{\gamma y} dH_Y(y), \quad x \in \mathbb{R}.$$

H_γ je pak distribuční funkce náhodné veličiny s kladnou konečnou střední hodnotou [Embrechts et al., 1997]. Označme $\{\tilde{Y}_i\}$ posloupnost iid náhodných veličin s distribuční funkcí H_γ . Hlavní myšlenka jak zkoumat trajektorie $\{S_n\}$, ve kterých nastává ruinování, je zkoumat trajektorie přidružené náhodné

procházky $\{\tilde{S}_n = \sum_{i=1}^n \tilde{Y}_i, n = 1, 2, \dots\}$. Zdůrazněme, že přidružená náhodná procházka $\{\tilde{S}_n\}$ má narozdíl od $\{S_n\}$ kladný trend.

Označme \hat{H}_n empirickou distribuční funkci náhodných veličin Y_1, \dots, Y_n . Podle Glivenko-Cantelliho věty (uvedena dále v textu, viz 7.3) platí

$$\sup_{x \in \mathbb{R}} |\hat{H}_n(x) - H_Y(x)| \rightarrow 0$$

s pravděpodobností jedna. Ovšem tato skutečnost se úplně změní, když n nahradíme časem ruinování.

Všechny dále uvedené výsledky jsou podmíněné tím, že nastalo ruinování, označme si proto pro zjednodušení zápisu

$$P^{(u)}(\cdot) = P(\cdot | \tau(u) < \infty), \quad u \geq 0.$$

Dále si označme N pořadí přírůstku, který způsobí ruinování. Uvažujeme-li \hat{H}_N empirickou distribuční funkci Y_1, \dots, Y_N , pak platí následující věta.

Věta 5.40. *Pro $u \rightarrow \infty$ platí*

$$\sup_{x \in \mathbb{R}} |\hat{H}_N(x) - H_\gamma(x)| \rightarrow 0 \text{ v pravděpodobnosti } P^{(u)}.$$

Věta 5.41. *Pro $u \rightarrow \infty$ platí:*

1. $\sup_{t \in [0,1]} \left| \frac{S(t\tau(u))}{\tau(u)} - t\hat{m}'_Y(\gamma) \right| \rightarrow 0$ v pravděpodobnosti $P^{(u)}$.
2. Veličina $\frac{\tau(u)-u/\hat{m}'_Y(\gamma)}{\sqrt{u}s(\gamma)} \rightarrow N$ v distribuci dané pravděpodobnosti $P^{(u)}$, kde $N \sim N(0, 1)$ a $s(\gamma)$ je funkce závislá pouze na Lundebergově koeficientu a momentech X .
3. Veličiny $S(\tau(u)) - u$ a $u - S(\tau(u)-)$ konvergují společně v distribuci dané pravděpodobnosti $P^{(u)}$ k ne degenerovanému sdruženému rozdělení. Navíc $\tau(u)$ a $S(\tau(u)) - u$ jsou asymptoticky nezávislé.

Důkazy uvedených tvrzení vedou na složité partie teorie front. Celá tato partie je shrnutím kapitoly 8.3 z [Embrechts et al., 1997] a zájemce o podrobnější studium odkazujeme na tuto literaturu a citace uvedené v ní.

Pro nás důležitý výsledek uvedených tvrzení je, že těsně než nastane ruinování se rozdělení přírůstků Y_i mění z H_Y na $H\gamma$. Trajektorie vedoucí k ruinování mají lokálně rostoucí trend se směrnicí $\hat{m}'_Y(\gamma) > 0$, narozdíl od globálního klesajícího trendu procesu $\{S(t)\}$.

5.6.2 Těžké chvosty

V případě, že přírůstky X_i mají rozdělení s těžkými chvosty je situace naprosto odlišná. Proces se vyvíjí naprosto standardně až do času ruinování a ruinování nastává jako důsledek jednoho extrémně velkého přírůstku.

Jako i dříve budeme uvažovat jen speciální třídu rozdělení s těžkými chvosty a to rozdělení subexponenciální. Pro pravděpodobnost ruinování pak platí obdoba Cramér-Lundebergovy approximace pro rozdělení s těžkými chvosty (5.9).

Zavedeme si označení:

- $-Z(u) = S(\tau(u)) - u$ nechť označuje rezervu těsně než nastane ruinování,
- $V_n = S(\tau_+(n)) - S(\tau_+(n-1))$ přírůstky mezi rekordy,
- $V(u) = S(\tau(u))$ rezervu v okamžiku ruinování a
- $V(u) - u$ přesah nad počáteční rezervu u v čase ruinování $\tau(u)$.

Klesající trend procesu $\{S(t)\}$ zaručuje, že s pravděpodobností jedna má proces $\{S(t)\}$ jen konečně mnoho rekordů a ruinování zřejmě nastává v okamžiku nějakého rekordu. Označme proto

$$K(u) = \inf\{k \in \mathbb{N}, S(\tau_+(k)) > u\}$$

pořadí rekordu, který způsobí ruinování.

Přírůstky mezi rekordy V_n jsou iid a $P^{(0)}(V_1 < x) = F_X^s(x)$ podle věty 5.39. Dále platí

$$\begin{aligned} P(K(u) = n) &= P(K(u) = n, \tau(u) < \infty) \\ &= P(\tau_+(n) < \infty)P(V_1 + \dots + V_{n-1} \leq u, V_1 + \dots + V_n > u | \tau_+(n) < \infty) \\ &= P(\tau_+(n) < \infty)p_n(u), \end{aligned}$$

kde

$$\begin{aligned} p_n(u) &\approx P(\max(V_1, \dots, V_{n-1}) \leq u, \max(V_1, \dots, V_n) > u | \tau_+(n) < \infty) \\ &\approx P(V_n > u | \tau_+(n) - \tau_+(n-1) < \infty) = P^{(0)}(V_1 > u) = \overline{F_X^s}(u), \quad u \rightarrow \infty \end{aligned}$$

Pro důkaz uvedených tvrzení čtenáře opět odkazujeme na knihu [Embrechts et al., 1997] a v ní uvedené citace. Poznamenejme také, že

$$\begin{aligned} P(\tau_+(u) < \infty) &= P\left(\max_{k=1,\dots,n} (\tau_+(k) - \tau_+(k-1)) < \infty\right) \\ &= P^n(\tau_+(1) < \infty) = (\psi(0))^n. \end{aligned} \tag{5.21}$$

Kombinací s 5.24 dostaneme

$$\begin{aligned}
P^{(u)}(K(u) = n) &= \frac{P(K(u) = n, \tau(u) < \infty)}{\psi(u)} \\
&\approx \frac{(\psi(0))^n \bar{F}_X^s(u)}{\psi(u)} \\
&\rightarrow (1 - \psi(0))(\psi(0))^{n-1},
\end{aligned} \tag{5.22}$$

a vidíme, že počet rekordů než nastane ruinování má asymptoticky geometrické rozdělení s parametrem $1 - \psi(0)$.

Jednotlivé rekordy dělí trajektorie $\{S(t), t \in [0, \tau(u))\}$ do $K(u)$ úseků, kde $K(u)$ má asymptoticky geometrické rozdělení s parametrem $1 - \psi(0)$. Prvních $K(u) - 1$ segmentů se chová ustáleným způsobem podle věty 5.39, ale poslední úsek je odlišný. Ruinování je způsobeno jedním velkým přírůstkem a dalo by se očekávat, že ruinování pro subexponenciální přírůstky bude popisovat teorie extrémních hodnot.

Nechť $(Z(u), V(u))$ je náhodný vektor, který má vzhledem k $P^{(u)}$ stejné rozdělení jako vektor $(-R(\tau(u)-), R(\tau(u)))$, pak chování trajektorie do času ruinování a ruinování popisuje následující věta. $MDA(\Lambda)$ označuje speciální rozdělení, tzv. maximum domain of attraction, kde normované rozdělení maxima konverguje k Gumbelově rozdělení. Více o těchto rozděleních uvádí např. [Embrechts et al., 1997] ve třetí kapitole.

Věta 5.42. (*Ruinování v subexponenciálním případě*)

Předpokládejme, že $\bar{F}_X^s \in \mathcal{R}_{-1/\xi}$ pro $\xi \in (0, \infty)$ nebo $F_X^s \in MDA(\Lambda) \cap \mathcal{S}$ (to odpovídá $\xi = 0$) a položme $a(u) = \int_u^\infty \bar{F}(x)dx/\bar{F}(u)$. Pak

$$\sup_{t \in [0,1]} \left| \frac{S(t\tau(u))}{\tau(u)} + pt(1 - \psi(0)) \right| \rightarrow 0$$

a

$$(a(u))^{-1}(c(1 - \psi(0))\tau(u), Z(u), V(u) - u) \rightarrow (Z_\xi, Z_\xi, V_\xi)$$

pro $u \rightarrow \infty$ v pravděpodobnosti $P^{(u)}$. Náhodné veličiny V_ξ a Z_ξ mají obě zobecněné Paretovo rozdělení kde

$$P(V_\xi > v, Z_\xi > z) = \bar{G}_\xi(v + z) \quad v, z \geq 0,$$

kde

$$\bar{G}_\xi(x) = \begin{cases} (1 + \xi x)^{-1/\xi}, & \xi \in (0, \infty) \\ e^{-x}, & \xi = 0. \end{cases} \tag{5.23}$$

Z věty 5.42 okamžitě dostáváme, že pro každé $\xi \in [0, \infty)$ platí

$$\frac{Z(u) + V(u) - u}{a(u)} \rightarrow Z_\xi + V_\xi$$

v distribuci dané pravděpodobností $P^{(u)}$. Toto nese informaci o velikosti přírůstku $R(\tau(u)) - R(\tau(u)-) = V(u) + Z(u)$, který způsobí ruinování.

Důsledek 5.43. (*Velikost přírůstku, který způsobí ruinování v subexponenciálním případě*)

1. Nechť $\overline{F}_X^s \in \mathcal{R}_{1/\xi}$ pro $\xi \in (0, \infty)$, pak $a(u) \sim \xi u$ a

$$\lim_{u \rightarrow \infty} P^{(u)}\left(\frac{V(u) + Z(u)}{u} > x\right) = \left(1 + \frac{1}{\xi}\left(1 - \frac{1}{x}\right)\right)x^{-\frac{1}{\xi}}, \quad x \geq 0.$$

2. Nechť $F_X^s \in MDA(\Lambda) \cap \mathcal{S}$ a $a(u)$ je jako ve větě 5.42, pak

$$\lim_{u \rightarrow \infty} P^{(u)}\left(\frac{V(u) + Z(u) - u}{a(u)} > x\right) = (1+x)e^{-x}, \quad x \geq 0.$$

5.7 Přesné řešení

Samozřejmě ideální variantou řešení jsou přesné vzorce pro pravděpodobnosti ruinování $\psi(u)$ nebo $\psi(u, x)$, ovšem to je možné jen v několika speciálních případech a vesměs jde o složený Poissonův proces.

- Složený Poissonův model, kde se příchody událostí řídí homogenním Poissonovým procesem s intenzitou λ a přírůstky mají exponenciální rozdělení $Exp(\delta)$, pak je pravděpodobnost ruinování $\psi(u) = \frac{\lambda}{p\delta} e^{-u(\delta - \frac{\lambda}{p})}$. Tento výsledek se dá odvodit různě např. řešením integrálně diferenciální rovnice (5.2), tento postup nabízí [Rolski et al., 1998] v příkladu na str. 163. Jednodušeji se dostaneme k témuž výsledku výpočtem Lundebergových mezí (5.20).

Pro exponenciální rozdělení $Exp(\delta)$ snadno spočteme $\hat{m}_X(s) = \frac{\delta}{\delta-s}$ a Lundebergův koeficient dostaneme řešením rovnice

$$\theta(s) = \lambda(\hat{m}_X(s) - 1) - ps = \lambda\left(\frac{\delta}{\delta-s}\right) - ps = 0.$$

$\gamma = \delta - \frac{\lambda}{p}$ a $\overline{F}_X(x) = e^{-\delta x}$ dosadíme do

$$a_- = \inf_{x \in [0, x_0]} \frac{e^{\gamma x} \int_x^\infty \overline{F}_X(y) dy}{\int_x^\infty e^{\gamma y} \overline{F}_X(y) dy}$$

a dostáváme

$$a_- = \inf_{\substack{x \in [0, x_0) \\ \frac{\lambda}{p}x}} \frac{e^{(\delta - \frac{\lambda}{p})x} \cdot \frac{e^{-\delta x}}{\delta}}{e^{-\frac{\lambda}{p}x}},$$

všechny výrazy s x se zkrátí a zřejmě tedy $a_- = a_+ = \frac{\lambda}{p\delta}$ a v nerovnosti nastává rovnost.

Pro jiná rozdělení tento postup nevede k výsledku, protože nejsme obvykle schopni analityky spočítat ani Lundebergův koeficient.

- Složený Poissonův model, kde se příchody událostí řídí homogenním Poissonovým procesem s intenzitou λ a přírůstky jsou deterministické, řekněme $P(X = \mu) = 1$. Potom pravděpodobnost ruinování

$$\psi(u) = 1 - (1 - \lambda\mu) \sum_{k=0}^{\lfloor u/\mu \rfloor} e^{\lambda\mu(k - \frac{u}{\mu})} \frac{(\lambda\mu(k - \frac{u}{\mu}))^k}{k!}.$$

Odvození můžeme nalézt např. v [Asmussen, 2000].

- Explicitní formule existuje i pro složený Poissonův model s exponenciálními přírůstky ($X \sim Exp(\delta)$), kde intenzita pojistného závisí na aktuální rizikové rezervě $R(t)$. Předpokládejme, že intenzita pojistného v čase t je $p(r)$, pokud $R(t) = r$. Potom za rozumných předpokladů kladených na $p(r)$ je pravděpodobnost ruinování $\psi(u) = \int_u^\infty g(y)dy$, pro

$$g(x) = \frac{\gamma_0 \lambda}{p(x)} e^{\lambda\omega(x) - \delta x},$$

kde

$$\omega(x) = \int_0^x \frac{1}{p(y)} dy$$

a

$$\frac{1}{\gamma_0} = 1 + \int_0^\infty \frac{\lambda}{p(x)} e^{\lambda\omega(x) - \delta x} dx.$$

Důkaz a předpoklady tvrzení lze nalézt v [Asmussen, 2000].

- Analytické řešení lze odvodit také pro složený Poissonův model, kde rozdělení přírůstků je tzv. phase-type, a to jak pro konstantní pojistné tak pro pojistné závislé na aktuální rizikové rezervě. Phase-type distribuce je rozdělení času absorpce v Markovském procesu s konečně mnoha stavami, kde jeden stav je absorpční a ostatní jsou přechodné. Zájemce odkazujeme na [Asmussen, 2000].

- Přesné řešení existuje i pro složený Poissonův model, kde přírůstky mají některé speciální rozdělení s těžkými chvosty. viz [Boxma and Cohen, 1998] a rozšíření jejich práce v [Abate and Whitt, 1999].
- Pro pravděpodobnost ruinování v konečném čase $\psi(u, x)$ existují formule, které by se daly nazvat přesným řešením, pouze pro deterministické a exponenciální přírůstky, viz kapitola 5.5.

Kapitola 6

Simulace

V mnoha případech neexistuje přesné řešení ani numerické metody výpočtu pravděpodobnosti ruinování a představu si nemůžeme udělat ani pomocí nějakých omezení. Potom nám poměrně přesnou představu o chování pravděpodobnosti ruinování mohou dát simulace. Zajímá-li nás pravděpodobnost ruinování v konečném čase, je situace poměrně jednoduchá a řešením jsou běžné Monte Carlo metody. V případě pravděpodobnosti ruinování v nekonečném čase vyvstává problém, neboť sebedelší úsek procesu $\{R(t)\}$ nám neřekne zda nenastane ruinování někdy v budoucnu. Omezení na konečný horizont bylo v minulosti použito, ale podle [Asmussen, 2000] nedává uspokojivé výsledky. Naštěstí i zde existuje řešení a nejjednodušší metody pro složený Poissonův model popíšeme v dalším textu. V praxi nastává ještě jeden problém. A to, že ruinování je obvykle málo častá událost tedy má malou pravděpodobnost a dostáváme se do problémů s časem potřebným na vygenerování dostatečného počtu realizací.

6.1 Přímá Monte Carlo metoda

Tato metoda je označovaná CMC z anglického Crude Monte Carlo. Nechť Z je nějaká náhodná veličina a předpokládejme, že chceme zjistit $z = \mathbb{E}Z$ v situaci, kdy z nemůžeme zjistit analyticky, ale můžeme simulovat Z . CMC metoda spočívá v simulaci iid náhodných veličin Z_1, \dots, Z_n a odhadu z výběrovým průměrem

$$\bar{z} = \frac{1}{n} \sum_{i=1}^n Z_i.$$

Rozptyl Z odhadujeme také jeho výběrovou variantou

$$s^2 = \frac{1}{n} \sum_{i=1}^n (Z_i - \bar{z})^2 = \frac{1}{n} \sum_{i=1}^n Z_i^2 - \bar{z}^2.$$

Aplikací centrální limitní věty dostáváme

$$\sqrt{n}(z - \bar{z}) \xrightarrow{d} N(0, \sigma_Z^2),$$

kde $\sigma_Z^2 = \text{var } Z$. Snadno odvodíme interval spolehlivosti $(\hat{z} \pm \frac{su(\frac{\alpha}{2})}{\sqrt{n}})$, kterým se nejčastěji reprezentují výsledky.

Pokud jde o simulaci pravděpodobnosti ruinování, lze CMC metodu naprostě přímočaře použít k simulaci pravděpodobnosti ruinování v konečném čase $\psi(u, x)$. Simulujeme proces $\{R(t)\}$ do času $t = x$ a Z položíme rovno indikátoru jevu, že nastalo ruinování

$$Z = I(\min_{0 \leq t \leq x} R(t) < 0) = I(\tau(u) \leq x).$$

6.1.1 Metody redukce rozptylu

Ale při simulovalání málo pravděpodobných jevů nastává problém času. Čas potřebný pro simulaci se snažíme zkrátit zmenšením rozptylu simulované veličiny Z . Obvykle modifikujeme Z na Z' tak, že $\text{var } Z' < \text{var } Z$ a $\mathbb{E}Z' = \mathbb{E}Z$. Toto je klasické téma literatury o simulacích, my popíšeme jen dvě nejznámější metody používané pro simulaci pravděpodobnosti ruinování. Typický problém metod redukce rozptylu je, že je potřeba delší čas na vygenerování jedné realizace Z' . Přistoupit na nějakou metodu redukce rozptylu má smysl pokud je $\text{var } Z'$ významně menší než $\text{var } Z$. Pokud například $\text{var } Z' = \frac{\text{var } Z}{2}$ pak $\frac{n}{2}$ opakování přinese stejnou přesnost CMC metody jako n realizací bez redukce rozptylu.

Podmíněná Monte Carlo metoda

Buď Z náhodná veličina z CMC metody a Y buď nějaká další náhodná veličina generovaná ve stejný čas jako Z . Položme $Z' = \mathbb{E}[Z|Y]$, pak $\mathbb{E}Z' = \mathbb{E}Z$ a platí

$$\text{var } Z = \text{var } (\mathbb{E}[Z|Y]) + \mathbb{E}(\text{var } [Z|Y]).$$

Druhý člen je zřejmě nezáporný a tedy $\text{var } Z' \leq \text{var } Z$ a podmíněná Monte Carlo metoda vede vždy k redukci rozptylu.

Importance sampling

Idea této metody spočívá v simulaci z jiného rozdělení, řekněme \tilde{P} , než je původní rozdělení P náhodné veličiny Z . Míru \tilde{P} vybíráme tak, aby existovala náhodná veličina L taková, že

$$\mathbb{I}\{Z(\omega) \neq 0\}P(d\omega) = \mathbb{I}\{Z(\omega) \neq 0\}L(\omega)\tilde{P}(d\omega).$$

Pro simulovaný odhad z pak máme

$$z = \mathbb{E}Z = \tilde{\mathbb{E}}(LZ), \quad (6.1)$$

kde $\tilde{\mathbb{E}}$ je střední hodnota příslušná pravděpodobnostní míře \tilde{P} .

Realizace CMC metody při použití této metody redukce rozptylu známená generování iid dvojic $(Z_1, L_1), \dots, (Z_n, L_n)$ z \tilde{P} a z odhadujeme jako

$$\bar{z}_{IS} = \frac{1}{n} \sum_{i=1}^n L_i Z_i$$

a jeho rozptyl jako

$$s_{IS}^2 = \frac{1}{n} \sum_{i=1}^n (L_i Z_i - \bar{z}_{IS})^2.$$

Redukce rozptylu může být dosažena, ale také nemusí. Záleží to na volbě míry \tilde{P} a problém spočívá v optimální volbě \tilde{P} .

Poznámka 6.1. Existuje optimální volba \tilde{P} . \tilde{P} definujeme vztahem $\frac{d\tilde{P}}{dP} = \frac{Z}{\mathbb{E}Z} = \frac{Z}{z}$, a tedy $L = \frac{z}{Z}$. Pak

$$\text{vár}(LZ) = \tilde{\mathbb{E}}(LZ)^2 - (\tilde{\mathbb{E}}LZ)^2 = \tilde{\mathbb{E}}\left(\frac{z^2}{Z^2}Z^2\right) - \tilde{\mathbb{E}}\left(\frac{z}{Z}Z\right)^2 = z^2 - z^2 = 0.$$

Mohlo by se zdát, že jsme dosáhli náhodné veličiny s nulovým rozptylem. Ovšem veškeré simulace a následné pokusy o redukci rozptylu jsme prováděli, protože z nelze určit analyticky a nemůžeme tedy spočítat $L = \frac{z}{Z}$.

V obecnějším kontextu je existence volby zaručující redukci rozptylu dokázána v [Asmussen and Glynn, 2007].

I když optimální volbu \tilde{P} nemůžeme aplikovat v praxi, dává nám návod, jak dosáhnout velké redukce rozptylu. Snažíme se volit \tilde{P} tak, aby $\frac{d\tilde{P}}{dP}$ co nejvíce odpovídalo Z . To může být někdy také obtížné, pak se snažíme vybrat \tilde{P} tak, aby velké hodnoty Z byly více pravděpodobné.

6.2 Simulace pomocí Pollaczek-Kinchinovy formule

Pro složený Poissonův model máme Pollaczek-Kinchinovu formulaci (5.6). $1 - \psi(u)$ je distribuční funkce složené geometrické náhodné veličiny, a tedy

$$\psi(u) = P(C_N > u),$$

kde $C_N = \sum_{i=1}^N X_i$, kde N má geometrické rozdělení s parametrem $\frac{\lambda\mu}{p}$ a je nezávislé na X_i . X_1, X_2, \dots jsou nezáporné iid náhodné veličiny s distribuční funkcí F_X^s . CMC metodu můžeme aplikovat na $\psi(u) = z = \mathbb{E}Z$, kde $Z = I(C_N > u)$. Popišme ještě postup algoritmicky.

Algoritmus I

1. Generuj $N_i \sim Ge(\frac{\lambda\mu}{p})$.
2. Generuj $X_1^i, \dots, X_{K_i}^i$ z rozdělení daného F_X^s a polož

$$C_{N_i} = X_1^i + \dots + X_{N_i}^i.$$

3. Pokud $C_{N_i} > u$ polož $Z_i = 1$ jinak polož $Z_i = 0$.
4. Opakuj kroky 1 - 3 pro $i = 1, \dots, n$.

5. Odhadni $\psi(u) = z = \mathbb{E}Z$ výběrovým průměrem $\hat{z} = \bar{z} = \frac{1}{n} \sum_{i=1}^n Z_i$.

Tento algoritmus dává řešení pro pravděpodobnosti ruinování v nekonečném čase, ale pro velká u je velmi časově náročný. Zlepšení můžeme dosáhnout redukcí rozptylu.

Podmíněná Monte Carlo metoda vždy vede k redukci rozptylu. Uvažujme

$$\begin{aligned} \psi(u) &= P(X_1 + \dots + X_N > u) \\ &= \mathbb{E}[P(X_1 + \dots + X_N > u) | X_1, \dots, X_{N-1}] \\ &= \mathbb{E}[F_X^s(u - X_1 - \dots - X_{N-1})]. \end{aligned}$$

Generujeme tedy jen X_1, \dots, X_{N-1} a spočteme $Y = u - X_1 - \dots - X_{N-1}$ a položíme $Z = 1 - F_X^s(Y)$, pravděpodobnost, že další přírůstek způsobí ruinování.

Algoritmus II

1. Generuj $N_i \sim Ge(\frac{\lambda\mu}{p})$.
2. Generuj $X_1^i, \dots, X_{K_i-1}^i$ z rozdělení daného F_X^s a polož

$$Y = u - X_1 - \dots - X_{N-1}.$$

3. Polož $Z_i = 1 - F_X^s(Y)$ ($Z_i = 1$ pokud $Y_i < 0$).

4. Opakuj kroky 1 - 3 pro $i = 1, \dots, n$.

5. Odhadni $\psi(u) = z = \mathbb{E}Z$ výběrovým průměrem $\hat{z} = \bar{z} = \frac{1}{n} \sum_{i=1}^n Z_i$.

Rozptyl tohoto odhadu musí být jako podmíněný Monte Carlo menší než rozptyl CMC odhadu. Podrobnější infromace o algoritmech I a II a lze nalézt v [Asmussen and Binswanger, 1997]. Je zde také uvedeno jejich porovnání z hlediska rozptylu. Autoři uvádí ještě jeden komplikovanější, ale efektivnější algoritmus.

6.3 Importance sampling pomocí Lundebergovy approximace

I v tomto případě předpokládáme složený Poissonův model. Dále předpokládejme, že jsou splněny podmínky Cramér-Lundebergovy approximace (5.7) (tj. je splněna podmínka čistého zisku, $\hat{m}_X(\gamma) < \infty$ a existuje Lundebergův koeficient). Položme

$$z(u) = \psi(u) \approx \frac{p - \lambda \mathbb{E}X}{\lambda \hat{m}'_X(\gamma) - p} e^{-\gamma u} = Ce^{-\gamma u}.$$

$\psi(u)$ můžeme reprezentovat pomocí procesu přebytku nároku $\{S(t)\}$, pak $\psi(u) = e^{-\gamma u} \mathbb{E}_\theta e^{S(\tau(u)) - u}$, kde $\tau(u)$ je čas prvního ruinování a \mathbb{E}_θ je střední hodnota příslušná pozměněné míře F_θ definované

$$F_\theta(dx) = \frac{e^{\theta x}}{\hat{m}_X(\theta)} F_X(dx)$$

(podrobněji viz [Asmussen, 2000]).

Přitom $\psi(u) = z(u) = \mathbb{E}Z(u)$, kde $Z(u) = e^{\gamma S(\tau(u))}$. Spojitý proces $\{S(t)\}$ v praxi generujeme jen v časech σ_n , kdy nastává n -tá událost. Pak postup při generování náhodných veličin $Z = Z(u)$ popisuje algoritmus III.

Algoritmus III

1. Vypočítej Lundebergův koeficient $\gamma > 0$ jako řešení rovnice

$$\lambda(\hat{m}_X(\gamma) - 1) - p\gamma = 0$$

a definuj

$$\tilde{\lambda} = \lambda \hat{m}_X(\gamma), \quad \tilde{F}_X(dx) = \frac{e^{\gamma x}}{\hat{m}_X(\gamma)} F(dx).$$

2. Polož $S = 0$.
3. Generuj $T \sim Exp(\tilde{\lambda})$ a U z \tilde{F}_X . Polož $S = S + U - T$.
4. Pokud $S > u$, polož $Z = e^{-\gamma S}$, jinak se vrať do bodu 3.

6.4 Generování náhodných veličin

Ve výše uvedených algoritmech potřebujeme generovat náhodné veličiny, které pochází z poněkud nestandardního rozdělení daného integrovaným chvostem F_X^s . Základní metoda používaná pro generování náhodných veličin, které nepocházejí z rovnoměrného rozdělení, vychází z inverze distribuční funkce F .

Definice 6.2. Bud' $F(x)$ distribuční funkce nějaké náhodné veličiny, pak její inverze $F^{-\leftarrow}$ je

$$F^{-\leftarrow}(u) = \inf\{x : F(x) \geq u\}, \quad 0 < u < 1.$$

Využití inverzní funkce pro generování náhodných veličin vychází z následující věty.

Věta 6.3. Bud' X náhodná veličina s distribuční funkcí F , pak platí následující tvrzení.

1. $u \leq F(x) \iff F^{-\leftarrow}(u) \leq x$.
2. Pokud $U \sim R(0, 1)$, pak $F^{-\leftarrow}(U)$ má distribuční funkci F .
3. Pokud je F spojitá funkce, pak $F(X)$ je rovnoměrně rozdělené na intervalu $(0, 1)$.

Důkaz. [Asmussen and Glynn, 2007]

□

Pokud chceme vygenerovat náhodnou veličinu X z rozdělení s distribuční funkcí F , generujeme náhodnou veličinu U z rovnoměrného rozdělení $R(0, 1)$ a položíme $X = F^{-1}(U)$.

Samozřejmě existuje mnoho dalších sofistikovanějších metod pro generování náhodných veličin, některé další metody můžeme najít v [Asmussen and Glynn, 2007], včetně příkladů jejich použití. My je v textu uvádět nebudeme. Kdykoli je to možné používáme standardní metody naprogramované v použitém software R, v ostatních případech si vystačíme s výše uvedenou metodou založenou na inverzi distribuční funkce.

6.5 Simulační studie II

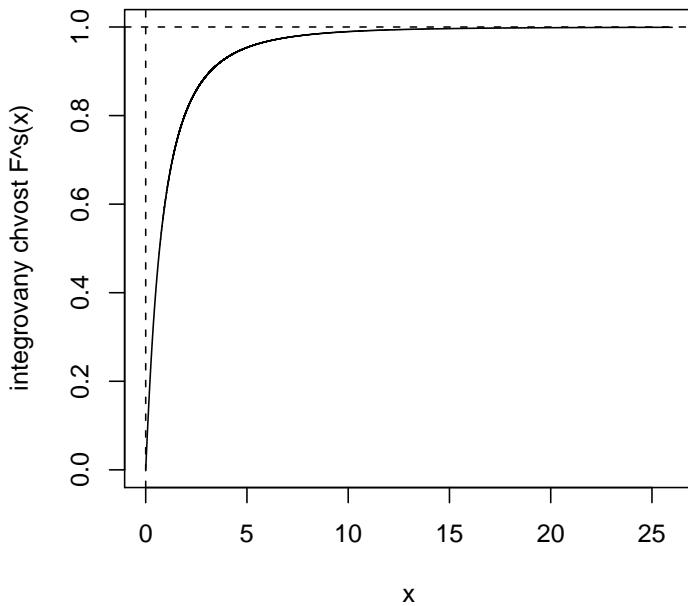
V této části prozkoumáme chování pravděpodobnosti ruinování pro složený Poissonův model s intenzitou $\lambda = 1$ v závislosti na různých parametrech, a to vždy pro degenerované ($P(X = 1) = 1$), exponenciální ($Exp(1)$) a logaritmicko-normální ($LN(\mu, 1)$, kde μ je voleno tak, aby střední hodnota přírůstku byla $\mathbb{E}X = 1$) přírůstky, tedy ve všech případech mají přírůstky stejnou střední hodnotu.

Integrovaný chvost je $F_X^s(x) = \frac{1}{\mu} \int_0^x \bar{F}_X(y) dy$ a snadným výpočtem dostaneme, že pro degenerované přírůstky s $\mathbb{E}X = 1$ je F_X^s distribuční funkce rovnoměrného rozdělení na intervalu $(0, 1)$, pro exponenciální přírůstky $Exp(1)$ dostáváme, že F_X^s je distribuční funkce téhož exponenciálního rozdělení $Exp(1)$. Pro lagaritmicko-normální přírůstky neexistuje analytické vyjádření distribuční funkce a tedy ani jejího integrovaného chvostu. V práci jsme použili nejjednodušší přístup, kdy F_X^s počítáme numericky na mřížce. Náhodnou veličinu s rozdělením daným F_X^s generujeme na základě věty 6.3 opět numericky.

Hodnoty F_X^s jsme generovali na mřížce s krokem 0,0001. Poslední hodnota ma mřížce byla volena tak, aby pravděpodobnost že logaritmicko-normální rozdělení dosáhne hodnoty vyšší, než je poslední bod mřížky, byla menší nebo rovna než požadovaná hodnota α , kde α bylo zvoleno 0,0001. Tvar integrovaného chvostu na zvoleném intervalu můžeme vidět na obrázku 6.1. Pro malé hodnoty x $F_X^s(x)$ velmi rychle roste a proto byla zvolena poměrně hustá síť.

Pro vygenerování náhodné veličiny s rozdělením daným F_X^s jsme generovali U z rovnoměrného rozdělení na intervalu $(0, 1)$. Následně alokovali interval $[i, i + 1)$ takový, že $F_X^s(i) \leq U \leq F_X^s(i + 1)$ a hodnotu X jsme určili lineární interpolací krajních bodů příslušného intervalu.

Pravděpodobnosti ruinování $\psi(u)$ byly získány přímou Monte Carlo metodou (algoritmus I) na základě 50000 opakování pro degenerované a expo-

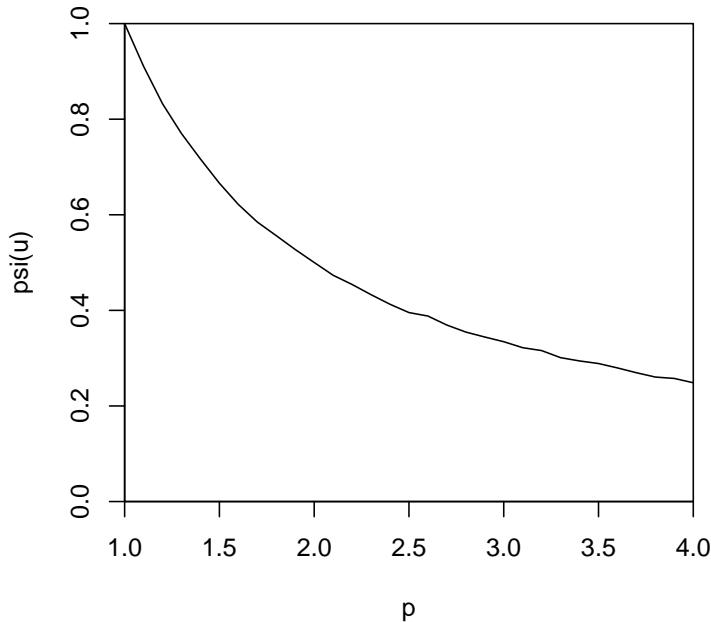


Obrázek 6.1: Integrovaný chvost logaritmicko-normálního rozdělení $LN(\mu, 1)$

nenciální přírůstky. Numerický přístup k logaritmicko-normálním přírůstkům vyžaduje vyšší čas pro vygenerování jednoho výběru, a proto jsme pravděpodobnosti ruinování získaly na základě pouchých 10000 opakování v CMC algoritmu. I pro tento počet opakování trvala simulace na použitém počítačovém vybavení několik hodin. Uvědomujeme si, že zvolený počet opakování může být pro některé kombinace parametrů nedostatečný, ale naším záměrem není určit co nejpřesněji pravděpodobnosti ruinování, ale udělat si představu o jejich chování v závislosti na různých parametrech.

Nejprve jsme zkoumali závislost pravděpodobnosti ruinování na výši pojištěného p pro počáteční rezervu $u = 0$. Hodnoty p byly zvoleny od jedné, kdy je $\psi(0) = 1$ s.j. z podmínky čistého zisku, do hodnoty $p = 4$ s krokem 0,1. Simulační studie potvrdila platnost věty 5.11. Mírná odlišnost pravděpodobnosti ruinování v tabulce 6.1 pro různé rozdělení přírůstků je způsobena nepřesností simulace. V grafickém zobrazení se jednotlivé křivky překrývají, a proto na obrázku 6.2 uvádíme jen tu pro degenerované přírůstky.

Pravděpodobnost ruinování klesá z hodnoty 1 na hodnotu přesahující 20%. To je až překvapivě vysoká hodnota, když si uvědomíme, že (pro $p = 4$)



Obrázek 6.2: Pravděpodobnost ruinování v závislosti na výši pojistného p pro počáteční rezervu $u = 0$.

je ve střední hodnotě příjem za jedno období čtyřnásobný oproti očekávané výši nároků. Problém je pravděpodobně v nulové počáteční rezervě. K ruinování může velmi snadno dojít na začátku vývoje procesu, než rezervy dostatečně narostou.

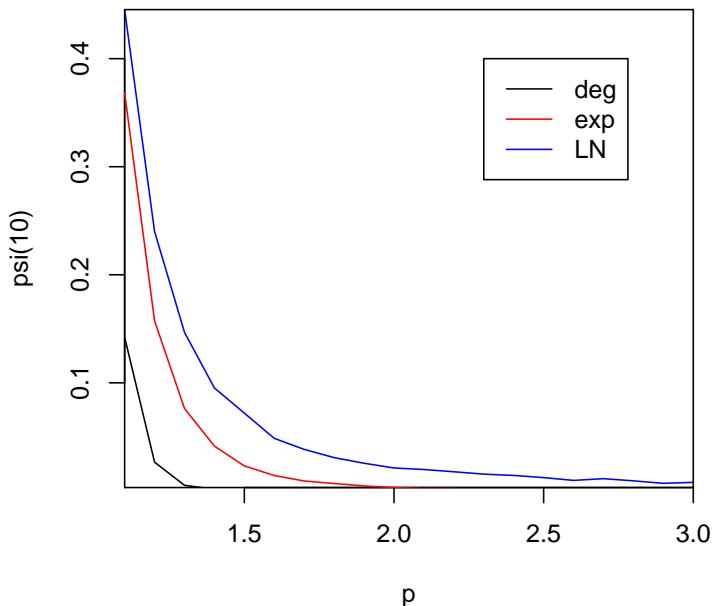
Pro $u \neq 0$ je už pravděpodobnost ruinování pro různé rozdělení přírůstků různá. Na obrázku 6.3 uvádíme pravděpodobnost ruinování pro $u = 10$. Na základě věty 5.22 a vztahu (5.8) bychom očekávali, že pro logaritmicko-normální přírůstky, jako reprezentanta rozdělení s těžkými chvosty, bude pravděpodobnost ruinování nejvyšší. Jde sice o asymptotický výsledek, ale naše očekávání se potvrzuje i pro malé hodnoty u .

Dále můžeme vidět, že se pravděpodobnosti ruinování dostávají rychle do velmi malých hodnot. Pro degenerované přírůstky simulace dokonce dává od $p = 1,8$ (viz tabulka 6.1) nulovou pravděpodobnost ruinování. To je ale způsobeno nepřesností a nedostatečným počtem opakování pro takovouto kombinaci parametrů. Pro exponenciální a logaritmicko-normální přírůstky jsou pravděpodobnosti ruinování nenulové pro všechny hodnoty p . Od $p =$

	$\psi(0)$				$\psi(10)$		
p	deg	exp	LN	p	deg	exp	LN
1	1	1	1	1	1	1	1
1,1	0,91064	0,90988	0,9092	1,1	0,14228	0,36848	0,4454
1,2	0,83230	0,83532	0,8323	1,2	0,02660	0,15732	0,2401
1,3	0,76990	0,76848	0,7700	1,3	0,00516	0,07626	0,1467
1,4	0,71662	0,71440	0,7091	1,4	0,00156	0,04144	0,0952
1,5	0,66602	0,66688	0,6666	1,5	0,00028	0,02312	0,0720
1,6	0,62158	0,62326	0,6238	1,6	0,00010	0,01424	0,0486
1,7	0,58480	0,58818	0,5891	1,7	0,00008	0,00914	0,0384
1,8	0,55630	0,55486	0,5564	1,8	0	0,00684	0,0308
1,9	0,52730	0,52582	0,5274	1,9	0	0,00460	0,0256
2,0	0,50020	0,50162	0,5009	2,0	0	0,00306	0,0212
2,1	0,47346	0,47220	0,4814	2,1	0	0,00218	0,0198
2,2	0,45422	0,45244	0,4542	2,2	0	0,00214	0,0177
2,3	0,43284	0,43450	0,4358	2,3	0	0,00130	0,0155
2,4	0,41274	0,41638	0,4157	2,4	0	0,00126	0,0143
2,5	0,39540	0,40050	0,4029	2,5	0	0,00066	0,0123
2,6	0,38814	0,38530	0,3847	2,6	0	0,00102	0,0097
2,7	0,36926	0,36522	0,3679	2,7	0	0,00060	0,0113
2,8	0,35462	0,36000	0,3513	2,8	0	0,00070	0,0093
2,9	0,34422	0,34364	0,3475	2,9	0	0,00046	0,0070
3,0	0,33442	0,33236	0,3386	3,0	0	0,00042	0,0079
3,1	0,32192	0,32382	0,3243	3,1	0	0,00050	0,0073
3,2	0,31586	0,31420	0,3109	3,2	0	0,00030	0,0058
3,3	0,30104	0,30440	0,3031	3,3	0	0,00018	0,0059
3,4	0,29418	0,29152	0,2934	3,4	0	0,00020	0,0052
3,5	0,28876	0,28488	0,2833	3,5	0	0,00022	0,0058
3,6	0,27960	0,27858	0,2743	3,6	0	0,00030	0,0054
3,7	0,26950	0,27024	0,2632	3,7	0	0,00018	0,0056
3,8	0,26044	0,26190	0,2624	3,8	0	0,00016	0,0058
3,9	0,25754	0,25448	0,2494	3,9	0	0,00020	0,0041
4,0	0,24854	0,24776	0,2501	4,0	0	0,00014	0,0031

Tabulka 6.1: Pravděpodobnost ruinování pro počáteční rezervu $u = 0$ vlevo a $u = 10$ vpravo v závislosti na výši pojistného, pro degenrované $P(X = 1) = 1$, exponenciální $Exp(1)$ a logaritmicko-normální přírůstky $LN(\mu, 1)$.

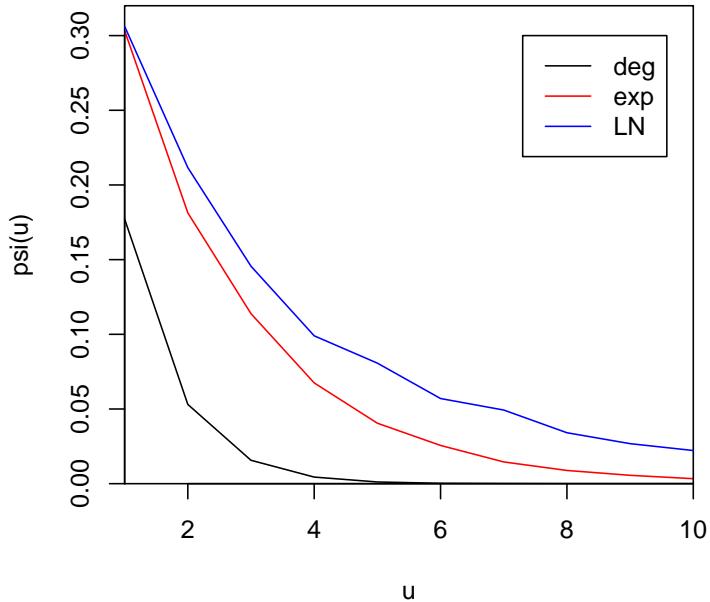
2,7 je rozdíl dokonce celého řádu, ale musíme opět zdůraznit nedostatečnost simulace pro takto malé hodnoty $\psi(u)$. Pro exponenciální přírůstky klesá



Obrázek 6.3: Pravděpodobnost ruinování v závislosti na výši pojistného p pro počáteční rezervu $u = 10$.

$\psi(u)$ až do řádu 10^{-4} , pro logaritmicko-normální do řádu 10^{-3} .

Nakonec jsme se zabývali ještě závislostí pravděpodobnosti ruinování $\psi(u)$ na počáteční rezervě u . Vybrali jsme prostřední hodnotu pojistného $p = 2$. Pro $u = 0$ je pravděpodobnost ruinování pro všechny tři typy přírůstků stejná (viz věta 5.11), s rostoucím u klesá nejrychleji pro degenerované přírůstky a nejpomaleji dle očekávání pro logaritmicko-normální přírůstky.



Obrázek 6.4: Pravděpodobnost ruinování v závislosti na výši počáteční rezervy u pro $p = 2$.

u	$\psi(u)$		
	deg	exp	LN
1	0,17716	0,30338	0,3062
2	0,05306	0,18134	0,2116
3	0,01566	0,11382	0,1456
4	0,00442	0,06748	0,0990
5	0,00114	0,04048	0,0807
6	0,00026	0,02560	0,0570
7	0,00010	0,01456	0,0493
8	0,00002	0,00884	0,0341
9	0,00000	0,00560	0,0268
10	0,00002	0,00334	0,0222

Tabulka 6.2: Pravděpodobnost ruinování pro pojistné $p = 2$ v závislosti na počáteční rezervě u , pro degenrované $P(X = 1) = 1$, exponenciální $Exp(1)$ a logaritmicko-normální příruštky $LN(\mu, 1)$.

Kapitola 7

Statistické metody odhadu parametrů a testy hypotéz pro složené bodové procesy

Předpokládejme, že pozorujeme proces $\{R(t)\}$, který splňuje předpoklad 2, tj. $\{C(t)\}$ je složený Poissonův proces a pojistné je inkasováno spojitě s konstantní intenzitou p . Pak z vývoje procesu $\{R(t)\}$ snadno dostaneme časy σ_n , kdy nastává n -tá událost a stejně tak velikost skoku X_n v čase σ_n . Víme, že přírůstky X_i , $i \in \mathbb{N}$ jsou iid a nezávislé na čítacím procesu $\{N(t) = \sum_{i=1}^{\infty} I\{\sigma_i \leq t\}, t \geq 0\}$. Díky vzájemné nezávislosti můžeme charakteristiky čítacího procesu $\{N(t)\}$ a přírůstků X_i odhadovat odděleně.

7.1 Odhad parametrů v homogenním Poissonově procesu

Pozorujeme časy událostí σ_i , které tvoří homogenní Poissonův proces s intenzitou λ . Proces pozorujeme do nějakého času t , tedy máme vektor časů $\sigma = (\sigma_1, \dots, \sigma_n)$, kde $\sigma_n \leq t$. Odvodíme vektor mezičasů $T = (T_1, \dots, T_n)$, kde $T_1 = \sigma_1$ a $T_i = \sigma_i - \sigma_{i-1}$, $i = 2, 3, \dots$. O T_i víme, že jsou iid s exponenciálním rozdělením $Exp(\lambda)$, tj. s hustotou $f_\lambda(x) = \lambda e^{-\lambda x}$ pro $\lambda > 0, x \geq 0$.

Odhad parametru λ odvodíme metodou maximální věrohodnosti.

$$L(\lambda) = \prod_{i=1}^n f_\lambda(x) = \lambda^n \prod_{i=1}^n e^{-\lambda T_i}.$$

Logaritmická věrohodnost je

$$l(\lambda) = \log L(\lambda) = n \log \lambda - \lambda \sum_{i=1}^n T_i.$$

Derivaci

$$\frac{\partial l}{\partial \lambda} = \frac{n}{\lambda} - \sum_{i=1}^n T_i$$

položíme rovnou nule a dostáváme

$$\hat{\lambda} = \frac{1}{\bar{T}}.$$

7.2 Testování hypotéz o homogenním Poissonově procesu

Na základě našich předpokladů by T_i měli tvořit náhodný výběr z exponenciálního rozdělení. To mimo jiné také znamená, že T_i jsou nezávislé na čase. Tuto hypotézu můžeme testovat různě.

1. Máme-li dostatek pozorování, můžeme parametr λ odhadovat v posuvném okně. Dostáváme tak odhad $\lambda(t)$, který je funkcí času, a pokud je splněn předpoklad nezávislosti, měla by tato funkce být přibližně konstantní. Tato metoda je vhodná jako orientační a může nás vést k nějaké domněnce o tom, jakým způsobem je porušen předpoklad nezávislosti a v následném testování můžeme volit vhodnější testy.
2. Vektor pozorování si rozdělíme na dvě části a testujeme nulovou hypotézu $F_1 = F_2$, že distribuční funkce v obou částech jsou shodné. Existuje mnoho různých testů a každý z nich je citlivý na jiný druh porušení nulové hypotézy. My uvedeme dvouvýběrový Kolmogorov-Smirnovův test a dvouvýběrový Wilcoxonův test.

Než si přesně formulujeme jednotlivé testy definujme si formálně empirickou distribuční funkci a uvedeme některé její vlastnosti.

Definice 7.1. Položme

$$F_n(x) = \frac{1}{n} \sum_{i=1}^n I\{X_i < x\},$$

pak F_n je empirická distribuční funkce náhodné veličiny X získaná na základě pozorování X_1, \dots, X_n .

S rostoucím n se empirická distribuční funkce blíží skutečné distribuční funkci F . Formálně tuto skutečnost uvádí následující věta.

Věta 7.2. Pro každé $x \in \mathbb{R}$ a pro $n \rightarrow \infty$ platí $F_n(x) \xrightarrow{sj} F(x)$.

Důkaz. [Anděl, 2005]. □

Ještě silnější tvrzení nám dává známá Glivenko-Cantelliho věta.

Věta 7.3. (Glivenko-Cantelli)

Označme $D_n = \sup_{x \in \mathbb{R}} |F_n(x) - F(x)|$, pak

$$P\left(\lim_{n \rightarrow \infty} D_n = 0\right) = 1.$$

7.2.1 Testy nezávislosti na čase

Dvouvýběrový Kolmogorov-Smirnovův test

Nechť X_1, \dots, X_m je náhodný výběr z rozdělení se spojitou distribuční funkcí F a nechť Y_1, \dots, Y_n je na něm nezávislý náhodný výběr z rozdělení se spojitou distribuční funkcí G . Kolmogorov-Smirnovovým testem ověřujeme hypotézu $H_0 : F = G$ proti alternativě $H_1 : F \neq G$. Označme F_m empirickou distribuční funkci na základě X_1, \dots, X_m a G_n empirickou distribuční funkci na základě Y_1, \dots, Y_n . Dále označme

$$D_{m,n} = \sup_{x \in \mathbb{R}} |F_m(x) - G_n(x)|.$$

Z věty 7.3 dostáváme $D_{m,n} \xrightarrow{sj} 0$ pro $m, n \rightarrow \infty$. Smirnov odvodil přesnější výsledek, na kterém je založen test hypotézy.

Věta 7.4. (Smirnov)

Označme $M = \frac{mn}{m+n}$ a položme

$$K(x) = 1 - s \sum_{k=1}^{\infty} (-1)^{k+1} \exp\{-2k^2 x^2\}.$$

Pak pro každé x platí

$$\lim_{m,n \rightarrow \infty} P(\sqrt{M} D_{m,n} < x) = K(x).$$

Přesné rozdělení a kritické hodnoty pro konečné m, n jsou tabelované. Pro velké hodnoty m, n se používá approximativní přístup. Funkce $K(x)$ se approximuje pomocí prvních členů $K(x) \approx 1 - 2e^{-2x^2}$ (grafické posouzení kvality approximace můžeme najít v [Anděl, 2005]). Použijeme-li tuto approximaci, dostaneme

$$P\left(D_{m,n} < \frac{x}{\sqrt{M}}\right) \doteq 1 - 2e^{-2x^2}.$$

Pro test hypotézy požadujeme, aby se výraz na pravé straně rovnal $1 - \alpha$, což platí pro $x_\alpha = \sqrt{\frac{1}{2} \log \frac{2}{\alpha}}$. Approximativní kritická hodnota pak je

$$D_{m,n}^*(\alpha) = \frac{x_\alpha}{\sqrt{M}} = \sqrt{\frac{1}{2M} \log \frac{2}{\alpha}}.$$

Praktické provedení Kolmogorov-Smirnovova testu spočívá ve výpočtu empirických distribučních funkcí F_m, G_n a následně veličiny $D_{m,n}$. Pro malé rozsahy výběrů m, n se $D_{m,n}$ porovná s přesnými kritickými hodnotami, pro velké m, n se využije approximace na základě Smirnovovy věty 7.4. Položí se $x_0 = \sqrt{M} D_{m,n}$ a spočte se $K(x_0)$. Pokud vyjde $K(x_0) \geq 1 - \alpha$ zamítá se H_0 na hladině přibližně α . Druhou možností je porovnat $D_{m,n}$ s approximativní kritickou hodnotou $D_{m,n}^*(\alpha)$.

Kolmogorov-Smirnovův test je standardně naprogramován snad v každém statistickém software. V R ho najdeme pod příkazem *ks.test*.

Dvouvýběrový Wilcoxonův test

Nechť X_1, \dots, X_m je, stejně jako v případě Kolmogorov-Smirnovova testu, náhodný výběr z rozdělení se spojitou distribuční funkcí F a nechť Y_1, \dots, Y_n je na něm nezávislý náhodný výběr z rozdělení se spojitou distribuční funkcí G . Chceme testovat hypotézu $H_0 : F = G$ proti alternativě $H_1 : F \neq G$.

Sdružený náhodný výběr, tj. všech $m + n$ náhodných veličin $X_1, \dots, X_m, Y_1, \dots, Y_n$ uspořádáme vzestupně podle velikosti. Označíme T_1 součet pořadí hodnot X_1, \dots, X_m a T_2 označíme součet pořadí hodnot Y_1, \dots, Y_n . Zřejmě

$$T_1 + T_2 = \frac{(m+n)(m+n+1)}{2}$$

a za platnosti nulové hypotézy $H_0 : F = G$ je

$$\mathbb{E}T_1 = \frac{m(m+n+1)}{2}, \quad \text{var } T_1 = \frac{mn(m+n+1)}{12}. \quad (7.1)$$

Místo náhodné veličiny T_1 se častěji k testu hypotézy používá veličina $U_1 = mn + \frac{m(m+1)}{2} - T_1$. Testu založenému na U_1 se říká také Mann-Whitneyův

test. Označme dále $U_2 = mn + \frac{n(n+1)}{2} - T_2$, pak $U_1 + U_2 = mn$. Nulovou hypotézu o rovnosti distribučních funkcí zamítáme, pokud $\min\{U_1, U_2\}$ je menší nebo rovno než tabelovaná kritická hodnota.

Pro velké hodnoty m a n se používá asymptotický test. Bylo dokázáno, že pro $m \rightarrow \infty$ a $n \rightarrow \infty$ má náhodná veličina T_1 asymptoticky normální rozdělení. Z definice U_1 a z vzorce (7.1) dostaneme

$$\mathbb{E}U_1 = mn + \frac{m(m+1)}{2} - \frac{m(m+n+1)}{2} = \frac{mn}{2}$$

a

$$\text{var } U_1 = \text{var } T_1 = \frac{mn(m+n+1)}{12}.$$

Jednoduchými úpravami dostaneme $U_2 = mn - U_1$ a tedy $\mathbb{E}U_2 = mn - \mathbb{E}U_1 = \frac{mn}{2} = \mathbb{E}U_1$ a $\text{var } U_2 = \text{var } U_1$, a nakonec

$$U_{MW} = \frac{U_1 - \mathbb{E}U_1}{\sqrt{\text{var } U_1}} = \frac{U_1 - \frac{mn}{2}}{\sqrt{\frac{mn(m+n+1)}{12}}} \stackrel{as}{\sim} N(0, 1)$$

a pokud $|U_{MW}| \geq u(\frac{\alpha}{2})$, pak nulovou hypotézu zamítáme na hladině blížící se α . [Anděl, 2005] uvádí, že asymptotický test založený na U_{MW} se dá použít pro $m > 10, n > 10$.

V R najdeme Wilcoxonův test ve standardní knihovně stats pod příkazem *wilcox.test*.

I když je Wilcoxonův test formulován proti obecné alternativě, je citlivý hlavně na alternativu posunutí, tj. $F(x) = G(x - x_0)$, $x_0 \neq 0$. Pokud se distribuční funkce F a G liší spíše tvarem nebo rozptylem je vhodnější použít výše uvedený Kolmogorov-Smirnovův test.

7.2.2 Testy exponenciality

Dále můžeme testovat zda časy T_i pochází z exponenciálního rozdělení. Ovyklo ovšem neznáme parametry exponenciálního rozdělení a nemůžeme použít běžný jednovýběrový Kolmogorov-Smirnovův test. Ale můžeme použít Lillieforsův test, což je obdoba jednovýběrového Kolmogorov-Smirnovova testu, pro parametry odhadované z dat.

Lillieforsův test

Jednovýběrový Kolmogorov-Smirnovův test se používá pro testování hypotézy, že daná pozorování jsou realizací náhodné veličiny s distribuční funkcí F_0 , přičemž F_0 musí být spojitá a kompletně specifikovaná, tzn. musí být

známé její parametry. Pokud jsou parametry odhadované z dat, nelze Kolmogorov-Smirnovův test použít, alespoň ne s klasickými kritickými hodnotami. Naštěstí pro exponenciální rozdělení existuje řešení viz [Lilliefors, 1969].

Uvažujme náhodný výběr X_1, \dots, X_n a jeho empirickou distribuční funkci F_n . Chceme testovat hypotézu, že X_1, \dots, X_n je náhodný výběr z exponenciálního rozdělení a parametrem λ , kde $\lambda > 0$ je libovolné.

Označme F_λ distribuční funkci exponenciálního rozdělení s parametrem $\lambda = \frac{1}{\bar{X}}$ odhadnutým z X_1, \dots, X_n a definujme

$$D_n = \sup_{x \in \mathbb{R}} |F_n - F_\lambda|.$$

Pokud D_n překročí tabelované kritické hodnoty pak zamítáme nulovou hypotézu, že výběr pochází z exponenciálního rozdělení.

Test exponenciality založený na Kullback-Leiblerově informaci

Zajímavý a silný test exponenciality uvádí [Ebrahimi et al., 1992]. Opět předpokládáme náhodný výběr z rozdělení s distribuční funkcí F a hustotou $f(x)$ pro $x \geq 0$ a zajímáme se o test hypotézy $H_0 : f = f_0 = \lambda e^{-\lambda x}$ proti alternativě $H_1 : f \neq f_0$. Parametr λ může, ale nemusí být specifikován.

Rozhodování mezi hypotézou a alternativou je založeno na Kullback-Leiblerově diskriminační informační funkci

$$I(f, f_0) = \int_0^\infty f(x) \log \left(\frac{f(x)}{f_0(x)} \right) dx. \quad (7.2)$$

$I(f, f_0)$ je nezáporná a $I(f, f_0) = 0$ právě tehdy když $f(x) = f_0(x)$ pro skoro všechna x . Kullback-Leiblerova diskriminační funkce měří míru nepodobnosti mezi f a f_0 . Za platnosti nulové hypotézy by měla být $I(f, f_0) = 0$ a velké hodnoty $I(f, f_0)$ svědčí pro alternativu. Naneštěstí pro přesný výpočet integrálu v (7.2) potřebujeme znát obě hustoty f i f_0 . Ale test můžeme založit na odhadu.

Definujeme entropii f

$$H(f) = - \int_0^\infty f(x) \log f(x) dx$$

pak za platnosti H_0 máme

$$\begin{aligned} I(f, f_0) &= -H(f) - \int_0^\infty f(x) \log f_0(x) dx = \\ &= -H(f) - \log \lambda + \lambda \int_0^\infty x f(x) dx = -H(f) - \log \lambda + 1. \end{aligned} \quad (7.3)$$

Odhadem entropie se zabývalo mnoho studií. Autoři si pro svůj test vybírají Vašíčkův odhad

$$H_{mn} = \frac{1}{n} \sum_{i=1}^n \log \left(\frac{n}{2m} (X_{(i+m)} - X_{(i-m)}) \right),$$

kde se entropie odhaduje v posuvném okně a velikost okna m se volí $0 < m < \frac{n}{2}$. $X_{(1)} \leq \dots \leq X_{(n)}$ jsou pořadové statistiky náhodného výběru a položíme $X(i) = X(1)$ pokud $i < 1$ a $X(i) = X_{(n)}$ pokud $i > n$.

Pokud parametr λ není specifikován v nulové hypotéze, odhadneme ho v rovnici (7.3) na základě výběru X_1, \dots, X_n jako $\hat{\lambda} = \frac{1}{\bar{X}}$ a odhad Kullback-Leiblerovy informace $I(f, f_0)$ mezi daty a nespecifikovaným exponenciálním rozdělením je

$$I_{mn} = -H_{mn} - \log \hat{\lambda} + 1 = -H_{mn} + \log \bar{X} + 1.$$

Velké hodnoty I_{mn} naznačují, že náhodný výběr X_1, \dots, X_n nepochází z exponenciálního rozdělení. Autoři článku uvažují monotonné transformaci $KL_{mn} = e^{-I_{mn}}$, která normuje hodnoty na interval $[0, 1]$. Na základě Monte Carlo simulací tabelovali kritické hodnoty pro tento test a část jich uvádí i v citovaném článku ([Ebrahimi et al., 1992]).

Konkrétní provedení testu je velmi jednoduché. Autoři uvádí i tabulkou, podle které se zvolí pro daný rozsah výběru n nejvhodnější šířka okna m . Pro pevné m se velmi snadno spočte odhad I_{mn} a následně jeho normalizace KL_{mn} . Tato hodnota se porovná s tabelovanými kritickými hodnotami $C_{mn}(\alpha)$. Nulová hypotéza se zamítá na hladině α pokud je $KL_{mn} < C_{mn}(\alpha)$.

Autoři v článku porovnávají svůj test z hlediska síly s několika dalšími testy a ukazují, že mnohdy dosahuje větší síly, nebo je jeho síla alespoň srovnatelná s jinými testy exponenciality.

7.3 Odhad rozdělení příruštek X_i

Do času t pozorujeme konečnou posloupnost událostí X_i a zajímá nás z jakého pocházejí rozdělení. První fází odhadu bude určitě histogram nebo jádrový odhad hustoty. Jádrový odhad hustoty narozdíl od histogramu dává různým bodům různou váhu a jeho výsledkem je hladká funkce, která bývá pro grafickou představu vhodnější.

$$\hat{f}_n(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x - x_i}{h}\right),$$

kde K je nějaká váhová funkce. Bežně se používá několik různých tvarů váhové funkce. Při samotném výpočtu ve statistickém software si musíme dát pozor, aby byly zakázané záporné hodnoty x , protože přírůstky X_i jsou z definice nezáporné a pro jejich hustotu tedy musí platit $f(x) = 0$ pro $X \leq 0$.

7.4 Testování hypotéz o přírůstcích X_i

Odhad hustoty nás může vést k nějaké domněnce o typu rozdělení a tuto domněnku pak můžeme testovat, nejčastěji pomocí empirické distribuční funkce. Opět si musíme dát pozor na to, že obvykle neznáme parametry daného rozdělení.

Než se pustíme do nějakých odhadů, měli bychom otestovat zda jsou jednotlivá pozorování nezávislá a pochází ze stejného rozdělení (a zda má vůbec smysl se zabývat odhadem rozdělení). O X_i předpokládáme jednak, že jsou pozorování nezávislá mezi sebou a jednak, že jsou nezávislé na čítacím procesu $N(t)$.

7.4.1 Testy nezávislosti a náhodnosti přírůstků X_i

Podpůrnou grafickou metodou může být odhad $\mathbb{E}X_i$ a $\text{var } X_i$ v posuvném okně. Pokud popisné statistiky vykazují nějaký trend svědčí to pro závislost pozorování na čase.

Znaménkové testy

Základní a široce použitelné testy nezávislosti jsou tzv. znaménkové testy. O X_i předpokládáme, že jsou iid a že pochází ze spojitého rozdělení. Pak jsou jejich hodnoty skoro jistě různé. V praxi se samozřejmě často stává, že jsou v datech některé hodnoty shodné, to není vylučující pro použití znaménkových testů. Podstatné je, aby tento jev nepřevládal.

- Chceme testovat hypotézu H_0 , že pozorování X_1, \dots, X_n jsou iid, proti obecné aletrenativě $H_1 = \text{non } H_0$. Z pozorovaných hodnot X_1, \dots, X_n spočteme medián \tilde{X} a vytvoříme posloupnost znamének $s_i = \text{sign}(X_i - \tilde{X})$, $i = 1, \dots, n$, pokud jsou v ní nuly, pak je vynecháme. Dostáváme posloupnost délky $N \leq n$, kde je m čísel +1 a n čísel -1. Spočteme P počet skupin stejných znamének (sérií).

Příklad. V posloupnosti $-1, -1, -1, -1, 1, 1, -1, 1, 1, 1, -1$ je $P = 5$ sérií.

Za platnosti H_0 je $\mathbb{E}P = \frac{2mn}{N} + 1$ a $\text{var } P = \frac{2mn(2mn-N)}{N^2(N-1)}$ a samozřejmě $\frac{m}{N}$ a $\frac{n}{N} \rightarrow \frac{1}{2}$ s.j. Naopak pokud H_0 neplatí dá se čekat, že P bude hodně malé či hodně velké. Kritické hodnoty pro $P(m, n)$ jsou tabelovány.

Pro velký rozsah výběru (uvádí se $n > 20$) má za platnosti H_0 náhodná veličina $Z = \frac{P - \mathbb{E}P}{\sqrt{\text{var } P}} \approx \sqrt{N} \left(\frac{NP}{2mn} - 1 \right)$ asymptoticky normální rozdělení $N(0, 1)$. V testu pak porovnáváme Z s kritickými hodnotami pro normální rozdělení na hladině α .

2. Druhá varianta znaménkového testu využívá porovnání dvou sousedních realizací náhodné veličiny X . Spočteme $s_i = \text{sign}(X_{i+1} - X_i)$, $i = 1, \dots, n-1$ a nechť opět P značí počet sérií v posloupnosti s_1, \dots, s_N , kde jsme vynechali všechny nuly ($N \leq n-1$). Při H_0 je $\mathbb{E}P = \frac{2n-1}{3}$ a $\text{var } P = \frac{16n-29}{90}$. Pro velké hodnoty n ($n > 20$) se opět používá normální approximace. $Z = \frac{P - \mathbb{E}P}{\sqrt{\text{var } P}} \stackrel{as}{\sim} N(0, 1)$.

7.4.2 Nezávislost na $N(t)$

Pokud jsou X_i nezávislé na $N(t)$ musí být nezávislé i na čase, neboť $N(t)$ je neklesající transformace času. Proto jakýkoli trend v pozorováních bude svědčit spíše pro zamítnutí nezávislosti. Jako statistický test nám může posloužit Spearmanův test, který testuje nezávislost dvou náhodných veličin, a to z pouhé znalosti pořadí sledovaných hodnot.

Pozorujeme náhodný výběr $(N(t_1), X_1), \dots, (N(t_n), X_n)$ uspořádaný podle velikosti $N(t_i)$ a i odpovídá pořadí jednotlivých $N(t_i)$ v $(N(t_1), \dots, N(t_n))$. Poté uspořádáme podle velikosti přírůstky X_1, \dots, X_n a stanovíme jejich pořadí R_i v uspořádané posloupnosti. Spearmannův korelační koeficient se definuje jako korelační koeficient pořadí, tj.

$$r_S = \frac{\text{cov}(i, R_i)}{\sqrt{\text{var } i} \sqrt{\text{var } R_i}} = 1 - \frac{6}{n(n^2 - 1)} \sum_{i=1}^n (i - R_i)^2.$$

Skutečně má vlastnosti korelačního koeficientu, tedy nabývá hodnot od -1 do +1. Pro test nulové hypotézy $H_0 : N(t)$ a X jsou nezávislé náhodné veličiny jsou tabelované kritické hodnoty. Pro $n > 30$ se používá asymptotická normalita $U = r_S \sqrt{n-1} \sim N(0, 1)$ a nulovou hypotézu zamítáme pokud $|U|$ přesáhne kritické hodnoty normálního rozdělení na požadované hladině α .

Kapitola 8

Závěr

V práci jsme co nejpřehledněji shrnuli množství známých výsledků pro složené náhodné veličiny a složené bodové procesy, které jsou roztríštěné v rozsáhlé literatuře. Záměrem bylo vytvořit text, který bude srozumitelný každému, kdo má alespoň základní znalosti z teorie pravděpodobnosti a matematické statistiky. Text, který bude dostatečný pro první seznámení se složenými bodovými procesy a procesy rizika. Tato oblast matematiky je ale natolik rozsáhlá, že se nám nepodařilo do práce zahrnout mnohé další přístupy k složeným bodovým procesům, jako jsou například martingalové techniky či využití markovských řetězců. Do textu se také nevešlo rozšíření výsledků na jiné modely příchodu událostí, než je složený Poissonův model.

Pomocí simulačních technik jsme si udělali představu o chování složených náhodných veličin pro dvě nejdůležitější a v praxi nejpoužívanější rozdělení čítací náhodné veličiny. Rozdělení přírůstků jsme volili exponenciální, protože pro něj platí nejvíce speciálních výsledků a omezení, a logaritmicko-normální jako zástupce rozdělení s těžkými chvosty. Porovnáváme-li složené náhodné veličiny v závislosti na typu přírůstků, je průběh jejich hustoty podobný, ale ve střední hodnotě je rozdíl přibližně jednoho rádu. Dále jsme zkoumali kvalitu některých approximací či omezení pro složené náhodné veličiny.

Druhá polovina práce je věnována složeným bodovým procesům a procesům rizika. Opět na základě simulací jsme zkoumali závislost pravděpodobnosti ruinování na různých parametrech modelu. Přírůstky jsme volili degenerované, exponenciální a logaritmicko-normální jako zástupce rozdělení s těžkými chvosty. Rozsah simulací rozhodně nebyl dostatečný, aby získaný výsledek odpovídal teoretickým hodnotám pravděpodobnosti ruinování. Ale to ani nebylo záměrem. Simulace dávají představu o tom, jak se mění pravděpodobnost ruinování v závislosti na počáteční rezervě u a v závislosti na pojistném p . V obou případech podáváme také srovnání vývoje pro uvažované rozdělení přírůstků.

Snažili jsme se porovnat i střední hodnotu času ruinování, ale kvůli velké časové náročnosti simulací jsme nedospěli k použitelným výsledkům. Do textu jsme také nezahrnuli studie pravděpodobnosti ruinování v konečném čase.

V závěru práce jsme ukázali jak se dají použít běžné statistické metody a testy při analýze reálných dat pocházejících z nějakého složeného bodového procesu.

Literatura

- [Abate and Whitt, 1999] Abate, J. and Whitt, W. (1999). Explicit M/G/1 waiting-time distributions for a class of long-tail service-time distributions. *Operations Research Letters*, 25:25–31.
- [Anděl, 2002] Anděl, J. (2002). *Základy matematické statistiky*. Preprint MFF UK, Praha.
- [Asmussen, 2000] Asmussen, S. (2000). *Ruin Probabilities*. World Scientific, Singapore.
- [Asmussen and Binswanger, 1997] Asmussen, S. and Binswanger, K. (1997). Simulation of ruin probabilities for subexponential claims. *Astin Bulletin*, 27(2):297–318.
- [Asmussen and Glynn, 2007] Asmussen, S. and Glynn, P. W. (2007). *Stochastic Simulation: Algorithms and Analysis*. Springer Science, New York.
- [Boxma and Cohen, 1998] Boxma, O. J. and Cohen, J. (1998). The M/G/1 queue with heavy-tailed service-time distribution. *IEEE Journal on Selected Areas in Communications*, 16:749–763.
- [Cipra, 1991] Cipra, T. (1991). *Teorie rizika v pojistné matematice*. MFF UK ve spolupráci s Českou pojišťovnou s.p., Praha.
- [Ebrahimi et al., 1992] Ebrahimi, N., Habibullah, M., and Soofi, E. S. (1992). Testing exponentiality based on kullback-leibler information. *Journal of the Royal Statistical Society. Series B (Methodological)*, 54(3):739–748.
- [Embrechts et al., 1997] Embrechts, P., Klüppelberg, C., and Mikosch, T. (1997). *Modelling Extremal Events for Insurance and Finance*. Springer Verlag, Berlin.
- [Feller, 1971] Feller, W. (1971). *An Introduction to Probability Theory and Its Applications I*. Wiley, New York.

- [Jacobsen, 2006] Jacobsen, M. (2006). *Point Process Theory and Applications, Marked Point and Piecewise Deterministic Processes*. Birkhäuser, Boston.
- [Lachout, 2004] Lachout, P. (2004). *Teorie pravděpodobnosti*. Karolinum, Praha.
- [Lilliefors, 1969] Lilliefors, H. W. (1969). On the kolmogorov-smirnov test for the exponential distribution with mean unknown. *Journal of the American Statistical Association*, 64(325):387–389.
- [Prášková and Lachout, 2005] Prášková, Z. and Lachout, P. (2005). *Základy náhodných procesů*. Karolinum, Praha.
- [Rolski et al., 1998] Rolski, T., Schmidli, H., Schmidt, V., and Teugels, J. (1998). *Stochastic Processes for Insurance and Finance*. Wiley, New York.

Index

- čas
 - prvního ruinování, 45
 - v červených číslech, 46
- algoritmus
 - De Prilův, 33
 - Panjerův, 30–32
 - Panjerův spojitý, 31
- aproximace
 - Beekman-Bowersova, 64
 - Cramér-Lundebergova, 59
 - De Vylderova, 63
 - heavy trafic, 66
 - light trafic, 67
- CMC, 80
- deficit
 - maximální celkový, 46
- formule
 - Pollaczek-Kinchinova, 52
 - Sealova, 71
- funkce
 - autokovarianční, 11
 - charakteristická, 15
 - momentová vytvořující, 15
 - pravděpodobnostní vytvořující, 15
- importance sampling, 82
- integrovaný chvost, 52
- koeficient
 - Lundebergův, 34, 55–58
 - vyrovňávací, 34
- konvoluce, 14
- logaritmicko-konkávní, 22
- logaritmicko-konvexní, 22
- metoda
 - Crude Monte Carlo, 80
 - podmíněná Monte Carlo, 81
- model
 - individuální, 25
 - kolektivní, 25
 - složený Poissonův, 49
- modifikace, 16
- nerovnost
 - Cramér-Lundebergova, 58
 - Lundebergova, 34
- oříznutí, 16
- přírůstky
 - nezávislé, 49
 - stacionární, 49
- přebytek do doby ruinování, 46
- podmínka
 - čistého zisku, 51
- portfolio, 43
- pravděpodobnost
 - ruinování, 45, 47–48
- proces
 - čítací, 13
 - bodový jednoduchý, 12
 - bodový značkováný, 12
 - celkových pojistných nároků, 43
 - homogenní Poissonův, 49

- náhodný, 11
- přebytku nároků, 45
- pojistných rezerv, 44
- rizika, 43
- složený bodový, 13
- složený Poissonův, 50
- prostor
 - značek, 12
- riziko, 43
- rovnice
 - defektivní obnovy, 52
- rozdělení
 - beta, 18
 - binomické, 23
 - Cauchyovo, 20
 - degenerované, 23
 - Erlangovo, 18
 - exponenciální, 17
 - gamma, 18
 - geometrické, 23
 - logaritmické, 24
 - logaritmicko-normální, 19
 - negativně binomické, 23
 - normální, 17
 - oříznuté geometrické, 16
 - Paretovo, 19
 - Poissonovo, 23
 - s lehkými chvosty, 17
 - s těžkými chvosty, 18
 - složené, 27
 - složené geometrické, 29
 - složené Poissonovo, 28–29
 - subexponenciální, 20
 - Weibullovo, 19
- ruinování, 45
 - v konečném čase, 68–71
- střední hodnota, 11
- stochasticky menší, 35
- test

Přílohy

Složené geometrické rozdělení

```
#----- slozene geometricke rozdeleni -----
geom<-function(n,p,prirustky=c("deg", "exp","log_norm"),
EX=1,varX=1){ # generuje nahodny vyber ze slozeneho geometrickeho
rozdeleni # p = parametr geom rozdeleni # n = delka vyberu

# EX = stredni hodnota prirustku, resp. parametr LN rozdeleni #
deg = degenerovane prirustky # exp = exponencialni prirustky
Exp(1/EX) # log-norm = logarotmicko normalni prirustky LN(EX,varX)
neni veskuteckosti stredni hodnota a rozptyl

C<-c() for (i in 1:n){
  N<-rgeom(1,1-p) # rozdilna definice geometrickeho rozdeleni v R a v textu
#-----
#PRIRUSTKY
if (prirustky=="deg"){
  #degenerovane prirustky
  X<-seq(EX, by=0, length=N)
}

if (prirustky=="exp"){
  #exponencialni prirustky
  X<-rexp(N,(1/EX))
}

if (prirustky=="log_norm"){
  #logaritmicko-normalni prirustky
  X<-rlnorm(N,EX,varX)
}

#-----
```

```

C<-c(C,sum(X))
}
return(C) }

```

Složené Poissonovo rozdělení

```

----- slozene Poissonovo rozdeleni -----
pois<-function(n,lambda,prirustky=c("deg", "exp","log_norm"),
EX=1,varX=1){ # generuje nahodny vyber ze slozeneho geometrickeho
rozdeleni

# lambda = parametr Poissonova rozdeleni # n velikost vyberu

# EX = stredni hodnota prirustku, resp. parametr LN rozdeleni #
deg = degenerovane prirustky # exp = exponencialni prirustky
Exp(1/EX) # log-norm = logarotmicko normalni prirustky LN(EX,varX)
neni veskuteckosti stredni hodnota a rozptyl

C<-c() for (i in 1:n){
  N<-rpois(1,lambda)
  #-----
  #PRIRUSTKY
  if (prirustky=="deg"){
    #degenerovane prirustky - kontrola
    X<-seq(EX, by=0, length=N)
  }

  if (prirustky=="exp"){
    #exponencialni prirustky
    X<-rexp(N,(1/EX))
  }

  if (prirustky=="log_norm"){
    #logaritmicko-normalni prirustky
    X<-rlnorm(N,EX,varX)
  }
  #-----
  C<-c(C,sum(X))
}

```

```
return(C) }
```

CMC algoritmus pro $\psi(u)$

```
#----- CMC algoritmus pro psi(u)-----
# simulace pravdepodobnosti ruinovani v nekonecнем case
# na zaklade Algoritmu I z textu

CMC <- function(n,u=0,p=1,lambda=1,prirustky=c("deg",
"exp","log_norm"),EX=1,b=1){ parametr<-lambda*EX/p    #parametr
geometrickeho rozdeleni
#pokud >=1 pak ruinovani skoro jiste
if (parametr>=1) return(1) else { # lambda=parametr Poissonova
procesu # EX=stredni hodnota prirustku # b jen pro log-norm
prirustky b druhý param LN(a,b)

Z<-c() for (i in 1:n){
  N<-rgeom(1,(1-parametr))
  s#rozdelna definice geom rozdeleni v R a v textu
#-----
#PRIRUSTKY
  if (prirustky=="deg"){
    #degenerovane prirustky
    X<-seq(EX, by=0, length=N)
  }
  if (prirustky=="exp"){
    #exponencialni prirustky, F^s zustava stejne exp rozdeleni
    X<-rexp(N,(1/EX))
  }
  if (prirustky=="LN"){
    #logaritmicko-normalni prirustky
    a<-log(EX)-(1/2*b^2)
    X<-rlnorm(N,a,b)
  }
#-----

CC<-sum(X)
if(CC>u) x<-1 else x<-0
Z<-c(Z,x)
pravdepodobnost<-mean(Z)
rozptyl<-var(Z)
```

```

    }
return(data.frame(pravdepodobnost,rozptyl)) } }
```

Střední hodnota času ruinování

```

----- stredni hodnota casu ruinovani-----
Etau<-function(n,u=0,p=1,lambda=1,prirustky=c("deg",
"exp","log_norm"),EX=1,b=1){ #simuluje proces R(t) dokud nenastane
ruinovani, vraci vyberovy prumer a rozptyl casu ruinovani #pro
urcite kombinace parametru a poctu opakovani trva neunesne dlouho

#n pocet opakovani realizace procesu #lambda intenzita Poissona
(casy mezi udalostmi exp(lambda)) #prirustky "deg"=degenerovane
velikosti lambdaX #           "exp"=exponencialni Exp(lambdaX)
EX=1/lambdaX

tau<-c()

for (i in 1:n){
  sigma<-0
  R<-u

  while (R>=0){
    t<-rexp(1,lambda) #cas mezi udalostmi
    sigma<-sigma+t
    #cas udalosti (na konci cyklu je cas udalosti, ktera zpusobila ruinovani)

    -----
    #PRIRUSTKY
    if (prirustky=="deg"){
      #degenerovane prirustky
      X<-EX
    }
    if (prirustky=="exp"){
      #exponencialni prirustky
      X<-rexp(1,(1/EX))
    }
    if (prirustky=="LN"){
      #logaritmicko-normalni prirustky
      a<-log(EX)-(1/2*b^2)
      X<-rlnorm(1,a,b)
    }
  }
  tau[i]<-sigma
}
```

```

        }
#-----
R<-R+p*t-X
}
tau<-c(tau,sigma)
}
E<-mean(tau) rozptyl<-var(tau)

return(data.frame(E, rozptyl)) }

```

Pravděpodobnost ruinování v konečném čase

```

----- pravdepodobnost ruinovani do casu T -----
pst<-function(n,T,u=0,p=1,lambda=1,prirustky=c("deg",
"exp","log_norm"),EX=1,b=1){ #simuluje proces do casu T, sleduje
zda nastalo ruinovani a vraci v jakem pomeru nastalo ruinovani

#n pocet opakovani realizace procesu #lambda intenzita Poissona
(casy mezi udalostmi exp(lambda)) #prirustky "deg"=degenerovane
velikosti lambdaX # "exp"=exponencialni Exp(1/EX) #
"log_norm"=logaritmicko-normalni LN(a,b), a voleno tak aby stredni
hodnota =EX

pravdepodobnost<-c()

for (i in 1:n){

-----#
#POISSONUV PROCES N(t)
t<-0
sigma<-c() #casy udalosti

while (t<=T) {
t<-t+rexp(1,lambda)
sigma<-c(sigma,t)
}

# po pruchodu cyklem je posledni cas udalosti v sigma vetsi nez T
# orizneme
sigma<-sigma[-(length(sigma))]
```

```

# pokud sigma prazdny vektor dava varovani ale je to OK

#-----
#PRIRUSTKY
if (prirustky=="deg"){
  #degenerovane prirustky
  X<-seq(EX, by=0, length=length(sigma))
}

if (prirustky=="exp"){
  #exponencialni prirustky
  X<-rexp(length(sigma),(1/EX))
}

if (prirustky=="LN"){
  #logaritmicko-normalni prirustky
  a<-log(EX)-(1/2*b^2)
  X<-rlnorm(length(sigma),a,b)
}

#-----

#SLOZENY POISSONUV PROCES C(t)
C<-cumsum(X)
# PROCES RIZIKA R(t)=u+pt-C(t)
P<-p*sigma      #pojistne
R<-u+P-C      #risk reserve proces
#-----
if(min(R)<0) x<-1 else x<-0
pravdepodobnost<-c(pravdepodobnost,x)
}
return(mean(pravdepodobnost)) }

```

Čebyšova nerovnost

Omezení pravděpodobnosti ruinování na základě Čebyšovy nerovnosti

