

Univerzita Pardubice
Fakulta ekonomicko-správní

Pravděpodobnostní modely kolektivního rizika
Bc. Lucie Pešková

Diplomová práce
2020

Univerzita Pardubice
Fakulta ekonomicko-správní
Akademický rok: 2018/2019

ZADÁNÍ DIPLOMOVÉ PRÁCE

(PROJEKTU, UMĚLECKÉHO DÍLA, UMĚLECKÉHO VÝKONU)

Jméno a příjmení: **Bc. Lucie Pešková**
Osobní číslo: **E17858**
Studijní program: **N6209 Systémové inženýrství a informatika**
Studijní obor: **Pojistné inženýrství: Management finančních rizik**
Název tématu: **Pravděpodobnostní modely kolektivního rizika**
Zadávající katedra: **Ústav matematiky a kvantitativních metod**

Z á s a d y p r o v y p r a c o v á n í :

Cíl práce: Teoreticky popsat a aplikovat na reálných datech základní přesné, přibližné a simulační metody pravděpodobnostního modelování kolektivního rizika.

Osnova:

- Kolektivní model rizika a jeho základní charakteristiky.
- Rekurentní vyjádření rozdělení kolektivního rizika.
- Přibližné pravděpodobnostní modely kolektivního rizika.
- Simulační modely kolektivního rizika.
- Prezentace všech modelů na reálných datech.

Rozsah grafických prací: -
Rozsah pracovní zprávy: cca 50 stran
Forma zpracování diplomové práce: tištěná/elektronická

Seznam odborné literatury:

BOLAND, P. J., 2007, *Statistical and Probabilistic Methods in Actuarial Science*, Chapman&Hall/CRC, London, ISBN 1-58488-695-1.
HORÁKOVÁ, G., PÁLEŠ, M., SLANINKA, F., 2015, *Teória rizika v poistení*, Wolters Kluwer, Bratislava, ISBN 978-80-8168-273-5.
PACÁKOVÁ, V., 2004, *Aplikovaná poistná statistika*, IURA EDITION, Bratislava, ISBN 80-8078-004-8.
PACÁKOVÁ, V., 2007, *Modelovanie a simulácia rizík v neživotnom poistení*. E+M Ekonomie a Management, Technická univerzita v Liberci, 3/2007, ročník X, s. 109-121, ISSN 1212-3609.
TSE Y.-K., 2009, *Nonlife Actuarial Models*, University Press, Cambridge, ISBN 978-0-521-76465-0.

Vedoucí diplomové práce: prof. RNDr. Viera Pacáková, Ph.D.
Ústav matematiky a kvantitatívnych metod

Datum zadání diplomové práce: 3. září 2018
Termín odevzdání diplomové práce: 30. dubna 2019

doc. Ing. Romana Provozíková, Ph.D.
děkanka

L.S.

doc. RNDr. Bohdan Linda, CSc.
vedoucí ústavu

V Pardubicích dne 3. září 2018

Prohlašuji:

Tuto práci jsem vypracovala samostatně. Veškeré literární prameny a informace, které jsem v práci využila, jsou uvedeny v seznamu použité literatury.

Byla jsem seznámena s tím, že se na moji práci vztahují práva a povinnosti vyplývající ze zákona č. 121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon), ve znění pozdějších předpisů, zejména se skutečností, že Univerzita Pardubice má právo na uzavření licenční smlouvy o užití této práce jako školního díla podle § 60 odst. 1 autorského zákona, a s tím, že pokud dojde k užití této práce mnou nebo bude poskytnuta licence o užití jinému subjektu, je Univerzita Pardubice oprávněna ode mne požadovat přiměřený příspěvek na úhradu nákladů, které na vytvoření díla vynaložila, a to podle okolností až do jejich skutečné výše.

Beru na vědomí, že v souladu s § 47b zákona č. 111/1998 Sb., o vysokých školách a o změně a doplnění dalších zákonů (zákon o vysokých školách), ve znění pozdějších předpisů, a směrnicí Univerzity Pardubice č. 7/2019 Pravidla pro odevzdávání, zveřejňování a formální úpravu závěrečných prací, ve znění pozdějších dodatků, bude práce zveřejněna prostřednictvím Digitální knihovny Univerzity Pardubice.

V Pardubicích dne 31.05.2020

Lucie Pešková

PODĚKOVÁNÍ

Ráda bych poděkovala vedoucí mé diplomové práce, prof. RNDr. Vieri Pacákové, Ph.D., za cenné rady, trpělivost a čas strávený při konzultacích. Dále bych chtěla poděkovat své rodině a přátelům za podporu při studiu.

ANOTACE

Cílem diplomové práce je teoreticky popsat a aplikovat na reálných datech základní, přesné, přibližné a simulační metody pravděpodobnostního modelování kolektivního rizika. V teoretické části jsou uvedeny důležité definice modelů a základní charakteristiky kolektivního rizika a simulace pomocí metody Monte Carlo. V praktické části jsou prezentovány teoreticky popsané modely na reálných datech. Vstupní data jsou zpracována ve statistickém programu STATGRAPHICS Centurion XVII a následné výpočty jsou provedeny v tabulkovém procesoru MS Excel.

KLÍČOVÁ SLOVA

Kolektivní model rizika, složená rozdělení, smíšená rozdělení, aproximační modely, riziková přirážka, simulace Monte Carlo.

TITLE

Probability models of collective risk

ANNOTATION

The aim of the diploma thesis is to theoretically describe and apply basic, accurate, approximate and simulation methods of probabilistic modeling of collective risk to real data. The theoretical part contains important definitions of models and basic characteristics of collective risk and simulation using the Monte Carlo method. The practical part presents theoretically described models on real data. The input data are processed in the statistical program STATGRAPHICS Centurion XVII and subsequent calculations are performed in the MS Excel spreadsheet.

KEYWORDS

Collective risk models, mixed distributions, mixture distributions, approximately models, risk premium, Monte Carlo simulation.

OBSAH

Seznam obrázků	9
Seznam tabulek	10
Úvod	11
1 Kolektivní model rizika a jeho základní charakteristiky	13
1.1 Kolektivní riziko a jeho rozdělení bez specifikace N	13
1.2 Základní charakteristiky kolektivního rizika S	14
1.3 Rozdělení počtu pojistných plnění	16
1.4 Rozdělení výšky pojistných plnění	19
1.5 Základní typy složených rozdělení	23
1.6 Smíšená rozdělení v heterogenních portfoliích	27
2 Rekurentní vyjádření rozdělení kolektivního rizika	30
2.1 Rekurentní vztah pro diskrétní rozdělení pojistných plnění	30
2.2 Panjerův rekurentní vzorec pro výpočet $G(x)$	31
3 Přibližné pravděpodobnostní modely kolektivního rizika	34
3.1 Aproximace normálním rozdělením	34
3.2 Aproximace posunutým gama rozdělením	35
3.3 Určení rizikové přírážky k čistému pojistnému	36
4 Simulační modely kolektivního rizika	38
4.1 Metoda Monte Carlo pro stanovení rozdělení celkové škody	39
5 Prezentace modelů na reálných datech	41
5.1 Analýza vstupních dat	41
5.2 Hledání vhodného rozdělení pro vstupní data	45
5.2.1 Vhodné rozdělení počtu škod N a odhad parametrů	45
5.2.2 Vhodné rozdělení pro výšky pojistných plnění X a odhad parametrů	49
5.2.3 Hledání vhodného rozdělení a odhad parametrů v MS Excel	55
5.3 Základní charakteristika kolektivního modelu rizika S	59
5.4 Aproximace kolektivního modelu rizika S	60

5.5	Simulace metodou Monte Carlo	61
Závěr	63
Použitá literatura	65
Přílohy	67

SEZNAM OBRÁZKŮ

Obrázek 1: Pravděpodobnostní funkce binomického rozdělení	17
Obrázek 2: Pravděpodobnostní funkce Poissonova rozdělení	18
Obrázek 3: Pravděpodobnostní funkce negativně binomického rozdělení.....	19
Obrázek 4: Hustota pravděpodobnosti exponenciálního rozdělení	20
Obrázek 5: Hustota pravděpodobnosti Weibullova rozdělení	21
Obrázek 6: Hustota pravděpodobnosti lognormálního rozdělení	22
Obrázek 7: Hustota pravděpodobnosti gama rozdělení	23
Obrázek 8: Histogram počtu individuálních škod	42
Obrázek 9: Krabicový graf výšky pojistných plnění	43
Obrázek 10: Histogram výšky pojistných plnění.....	44
Obrázek 11: Histogram X s hustotou testovaných rozdělení	50
Obrázek 12: Histogram výšky pojistného plnění s lognormálním rozdělením	53

SEZNAM TABULEK

Tabulka 1: Panjerovy konstanty	33
Tabulka 2: Rozdělení počtu škod do tříd	42
Tabulka 3: Analýza počtu škod	42
Tabulka 4: Rozdělení pojistných plnění dle intervalů	43
Tabulka 5: Základní charakteristiky výšky pojistných plnění	44
Tabulka 6: χ^2 test – Poissonovo rozdělení.....	46
Tabulka 7: χ^2 test – Negativně binomické rozdělení.....	46
Tabulka 8: χ^2 test počtu škod pro vybraná rozdělení	47
Tabulka 9: Odhadnuté parametry negativně binomického rozdělení	48
Tabulka 10: Základní charakteristiky negativně binomického rozdělení.....	48
Tabulka 11: χ^2 test – exponenciální rozdělení.....	50
Tabulka 12: χ^2 test – gama rozdělení	51
Tabulka 13: χ^2 test – lognormální rozdělení	52
Tabulka 14: χ^2 test – Weibullovo rozdělení	52
Tabulka 15: K-S test	53
Tabulka 16: Odhadnuté parametry lognormálního rozdělení	54
Tabulka 17: Základní charakteristiky lognormálního rozdělení.....	54
Tabulka 18: Odhad parametru Poissonova rozdělení MM	55
Tabulka 19: Výpočtová tabulka χ^2 testu – Poissonovo rozdělení.....	56
Tabulka 20: Výpočtová tabulka χ^2 testu – Poissonovo rozdělení	56
Tabulka 21: Odhadnuté parametry negativně binomického rozdělení MM.....	57
Tabulka 22: Odhad parametrů lognormálního rozdělení MM.....	58
Tabulka 23: Výpočet ověření $X \sim LN(\mu; \sigma)$ K-S test	58
Tabulka 24: Počáteční momenty $E(X^k)$	60
Tabulka 25: Základní charakteristiky S	60
Tabulka 26: Aproximace normálním rozdělením a riziková přírážka	60
Tabulka 27: Parametry α, β , a k	61
Tabulka 28: Aproximace posunutým gama rozdělením a rizikové přírážky.....	61
Tabulka 29: Porovnání $S_{p(N)}$ a $S_{p(G)}$	61

ÚVOD

Lidská společnost je během svého vývoje ovlivňována působením různých nahodilostí, které mohou mít celkově kladné důsledky, ale také důsledky negativní. Vzhledem k existenci negativních důsledků, jsou ekonomické subjekty stále v ohrožení. Jsou vystaveny riziku vzniku události s odchylným výsledkem od jejich cíle s určitou objektivní pravděpodobností. Realizace tohoto rizika vede ke vzniku škody pro daný ekonomický subjekt. Subjekty se snaží různými opatřeními riziku vyhnout nebo ho alespoň zmírnit. Mohou také využít finanční krytí prostřednictvím státu (živelné události, sociální podpora aj.), individuální krytí rizika (různé formy individuálních rezerv) nebo pojištění (přenos rizika na pojistitele). Pojistitel, v posledním případě pomocí příspěvků na pojištění od zúčastněných, tvoří kolektivní rezervy, rozdělení rizika mezi více zúčastněných, a tak není krytí rizika omezeno naspořenými prostředky pouze jednoho účastníka. [4]

Pro pojistitele je důležité znát základní charakteristiky a rozdělení pravděpodobnosti celkového pojistného plnění S , tj. kolektivní model rizika. Díky těmto znalostem může pojistitel dále řešit rozhodující otázky související se zajištěním, spoluúčastí, tvorbou rezerv, pravděpodobností krachu aj. [6] [10]

Diplomová práce na téma Pravděpodobnostní modely kolektivního rizika má za cíl teoreticky popsat a aplikovat na reálných datech základní, přesné, přibližné a simulační metody pravděpodobnostního modelování kolektivního rizika. Práce je členěna na pět kapitol. V kapitolách 1 až 4 jsou teoreticky popsány jednotlivé modely a základní charakteristiky kolektivního rizika. Kapitola 5 se dělí na pět podkapitol, které aplikují modely a simulaci na reálných datech.

První kapitola definuje kolektivní model rizika a jeho základní charakteristiky. Jsou zde uvedeny základní typy rozdělení pravděpodobnosti počtu a výšky pojistných plnění, složených rozdělení a smíšená rozdělení v heterogenních portfoliích.

V druhé kapitole se práce zabývá rekurentním vyjádřením rozdělení kolektivního rizika, kde je popsán Panjerův rekurentní vzorec pro výpočet $G(x)$ a Panjerovy konstanty pro Poissonovo, binomické a negativně binomické rozdělení.

Ve třetí teoretické kapitole se práce zaměřuje na přibližné pravděpodobnostní modely kolektivního rizika, a to na aproximaci normálním rozdělením a posunutým gama rozdělením. Závěrem této části je uvedeno využití aproximačních modelů pro určení rizikové přírážky θ k čistému pojistnému.

Náplní čtvrté, poslední teoretické části je představení simulačních modelů kolektivního rizika. Tato kapitola obsahuje vysvětlení postupu simulačního modelu pomocí metody Monte Carlo pro stanovení rozdělení celkové škody S .

V poslední kapitole této práce jsou na reálných datech o počtu a výšce pojistných plnění aplikovány pravděpodobnostní modely kolektivního rizika, teoreticky popsány v předchozích kapitolách. Výpočty byly provedeny pomocí statistického programu STATGRAPHICS Centurion XVII a tabulkového procesoru MS Excel. V první části jsou vstupní data podrobena statistické analýze, kde jsou vypočteny základní charakteristiky počtu a výšky pojistných plnění. Dalším krokem je hledání vhodného pravděpodobnostního modelu počtu a výšky pojistných plnění pomocí testů dobré shody a odhad parametrů navrhovaných rozdělení. V této podkapitole je znázorněn výpočet pouze pomocí tabulkovém procesoru MS Excel. Pomocí získaných odhadů jsou v následujících dvou částech provedeny výpočty základních charakteristik a aproximace kolektivního rizika S . Poslední podkapitola této práce je věnována simulaci pomocí metody Monte Carlo. Na základě získaných hodnot z předchozího šetření je vysvětlen postup pro získání simulovaného výběrového souboru celkových pojistných škod S .

1 KOLEKTIVNÍ MODEL RIZIKA A JEHO ZÁKLADNÍ CHARAKTERISTIKY

1.1 Kolektivní riziko a jeho rozdělení bez specifikace N

Mějme náhodnou proměnnou S , která představuje celkové (souhrnné, agregátní) pojistné plnění, tj. kolektivní riziko pojistitele za celé portfolio v daném období, nejčastěji během jednoho roku. Písmenem N označíme počet pojistných plnění během roku a písmenem X_i výšku i -tého pojistného plnění pro $i = 1, 2, \dots, N$. Pak S je definováno vztahem:

$$S = \sum_{i=1}^N X_i = X_1 + X_2 + \dots + X_N. \quad 1.1$$

Přitom platí, že X_1, X_2, \dots, X_N jsou nezávislé a identicky rozdělené náhodné proměnné, N, X_1, X_2, \dots, X_N jsou vzájemně nezávislé náhodné proměnné. Dále platí, že pokud $N = 0$, pak i $S = 0$. [10]

Distribuční funkci náhodné proměnné S definujeme vztahem:

$$G(x) = P(S \leq x) = F_S(x) \text{ pro } x \in R$$

s těmito vlastnostmi:

1. $\lim_{x \rightarrow -\infty} G(x) = 0$,
2. $\lim_{x \rightarrow \infty} G(x) = 1$,
3. $G(x)$ je neklesající, tj. je-li $x < y \Rightarrow G(x) \leq G(y)$,
4. $G(x)$ je spojitá zprava, tj. $\lim_{h \rightarrow 0^+} G(x) = G(x + h)$.

Společnou distribuční funkci náhodné proměnné X_i , pro $i = 1, 2, \dots, N$, definujeme vztahem:

$$F(x) = \begin{cases} P(X_i \leq x) & \text{pro } x \geq 0, \\ 0 & \text{pro } x < 0. \end{cases}$$

Distribuční funkce v bodě x se rovná pravděpodobnosti, že celková škoda nepřesáhne hodnotu x . K tomu dochází, když nenastane žádná pojistná událost, ale i když nastane právě jedna, anebo právě dvě, anebo všeobecně právě n pojistných událostí, při kterých celková škoda nepřekročí peněžní sumu x . Na základě disjunktnosti jednotlivých událostí a využitím věty o násobení podmíněných pravděpodobností můžeme tuto úvahu zapsat takto:

$$\begin{aligned}
G(x) = F_S(x) &= P(S \leq x) = P\left(\bigcup_{n=0}^{\infty} \{S \leq x \wedge N = n\}\right) = \sum_{n=0}^{\infty} P(S \leq x \wedge N = n) = \\
&= \sum_{n=0}^{\infty} P(S \leq x | N = n) \cdot P(N = n).
\end{aligned}$$

Protože n značí fixní počet pojistných plnění, platí:

$$P(S \leq x | N = n) = P(X_1 + X_2 + \dots + X_n \leq x) = F^{*n}(x),$$

a můžeme určit vztah všeobecného vyjádření distribuční funkce $G(x)$ celkového plnění S pro libovolné rozdělení počtu pojistných plnění N a výšky pojistných plnění X_i takto:

$$G(x) = F_S(x) = \begin{cases} \sum_{n=0}^{\infty} P(N = n) \cdot F^{*n}(x) & \text{pro } x \geq 0, \\ 0 & \text{pro } x < 0, \end{cases} \quad 1.2$$

kde $F^{*n}(x)$ je definované vztahem $F^{*n}(x) = P(X_1 + X_2 + \dots + X_n \leq x)$ a vyjadřuje n -násobnou konvoluci distribuční funkce $F(x)$. Pro nultou konvoluci platí:

$$F^{*0}(x) = 1 \text{ pro } x = 0. \text{ [6] [10]}$$

1.2 Základní charakteristiky kolektivního rizika S

Pro určení střední hodnoty $E(S)$, rozptylu $D(S)$ a momentové vytvořující funkce $M_S(z)$ využijeme definice podmíněných středních hodnot pro libovolné náhodné proměnné X a Y :

$$E(X) = E(E(X|Y)),$$

$$D(X) = E(D(X|Y)) + D(E(X|Y)).$$

Využijeme také počáteční moment k -tého řádu identicky rozdělených individuálních pojistných plnění X_i , značíme $m_k = E(X_i^k)$ a definujeme vztahy:

- $m_k = \sum_i x_i^k \cdot P(x_i)$ – jestliže X je diskrétní proměnná,
- $m_k = \int_{-\infty}^{+\infty} x^k \cdot f(x) dx$ – jestliže X je spojitá náhodná proměnná.

1. Střední hodnota $E(S)$

Podle definice kolektivního rizika a vlastností střední hodnoty platí:

$$E(S|N = n) = E(X_1 + X_2 + \dots + X_n) = \sum_{i=1}^n E(X_i) = nm_1.$$

Pak podle definice podmíněné střední hodnoty pro $E(S)$ platí:

$$E(S) = E(E(S|N)) = E(Nm_1) = E(N)m_1. \quad 1.3$$

Střední hodnota celkového pojistného plnění $E(S)$ je součinem střední hodnoty pojistných plnění $E(N)$ a střední hodnoty m_1 výšky individuálního pojistného plnění X_i . [10]

2. Rozptyl $D(S)$

Pro určení rozptylu $D(S)$ opět využijeme podmíněnou střední hodnotu:

$$D(S) = E(D(S|N)) + D(E(S|N)).$$

$D(E(S|N))$ získáme dosazením vztahu $E(S|N) = Nm_1$ a získáváme:

$$D(E(S|N)) = D(Nm_1) = D(N)m_1^2.$$

$D(S|N)$ získáme pomocí skutečnosti, že individuální pojistné plnění jsou nezávislé:

$$D(S|N = n) = D(X_1 + X_2 + \dots + X_n) = \sum_{i=1}^n (m_2 - m_1^2) = n(m_2 - m_1^2).$$

Tedy $D(S|N) = N(m_2 - m_1^2)$ a pro rozptyl $D(S)$ platí:

$$D(S) = E(N(m_2 - m_1^2)) + D(N)m_1^2.$$

Po úpravě dostáváme:

$$D(S) = E(N)(m_2 - m_1^2) + D(N)m_1^2. \quad 1.4$$

Rozptyl kolektivního rizika $D(S)$ je vyjádřen pomocí známých středních hodnot a rozptylů proměnné N a proměnných X_i . [10]

3. Momentová vytvořující funkce $M_S(z)$

Podle definice $M_S(z) = E(e^{zS})$ a podle definice podmíněné střední hodnoty můžeme napsat vztah $M_S(z) = E(E(e^{zS}|N))$. Protože platí $E(e^{zS}|N = n) = E(e^{z(X_1+X_2+\dots+X_n)})$ a náhodné proměnné X_1, X_2, \dots, X_n jsou nezávislé a identicky rozdělené, pak získáme vztah:

$$E(e^{z(X_1+X_2+\dots+X_n)}) = \prod_{i=1}^n E(e^{zX_i}) = \prod_{i=1}^n M_{X_i}(z) = (M_X(z))^n.$$

Odtud dostáváme:

$$E(e^{zS}|N) = (M_X(z))^N$$

a po dosazení:

$$M_S(z) = E\left(\left(M_X(z)\right)^N\right).$$

Upravíme pomocí vztahu $\left(M_X(z)\right)^N = e^{N \cdot \ln M_X(z)}$ a pro definici momentové vytvořující funkce získáme vztah [1][10]:

$$M_S(z) = E\left(e^{N \cdot \ln M_X(z)}\right) = M_N(\ln \cdot M_X(z)). \quad 1.5$$

Momentovou vytvořující funkci celkového pojistného plnění S tedy získáme pomocí momentových vytvořujících funkcí počtu pojistných plnění N a výšky pojistných plnění X_i a jedná se o všeobecné vyjádření $M_S(z)$. [10]

1.3 Rozdělení počtu pojistných plnění

Reálné podmínky, které při různých typech pojištění vedou k nastání pojistné události, způsobují, že počet pojistných plnění N má nejčastěji některé z těchto diskrétních rozdělení pravděpodobnosti [10, 11]:

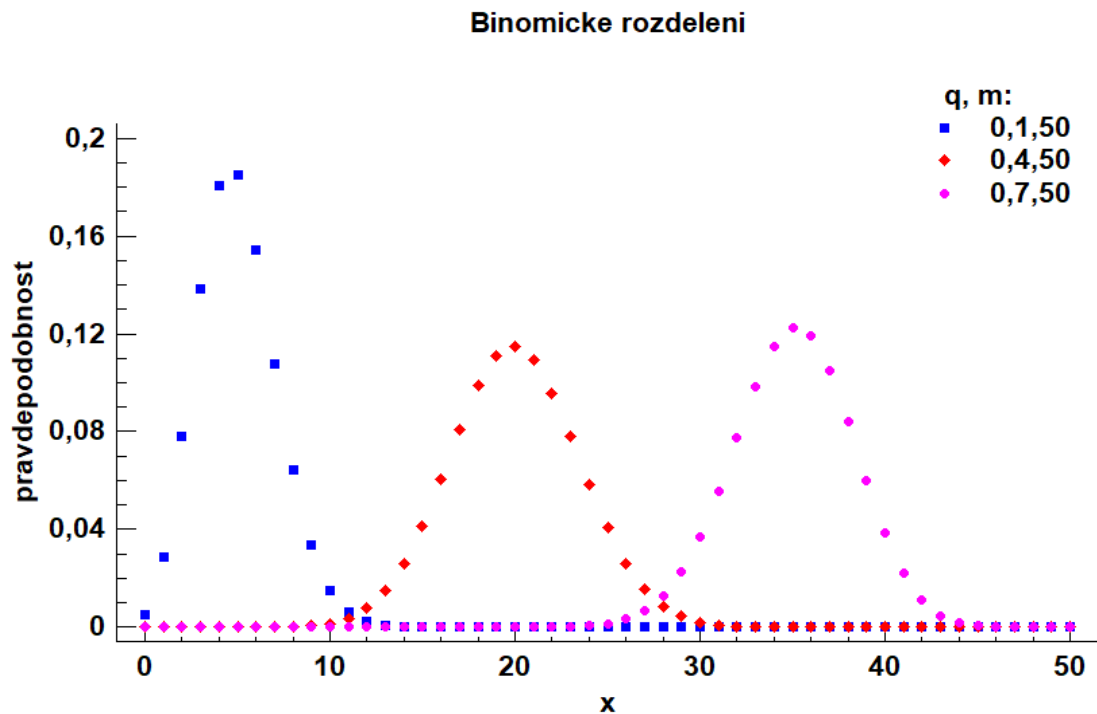
- Alternativní (Bernoulliho),
- binomické,
- Poissonovo,
- negativně binomické.

1. Binomické rozdělení $Bi(m; q)$

Jedná se o n -krát nezávisle opakovaný Bernoulliho pokus. N značí počet výskytů sledované události v n nezávislých pokusech se stejnou pravděpodobností výskytu π při každém realizovaném pokuse. Využívá se při různých typech pojištění, např. u pojištění pro případ úmrtí [11]

Binomické rozdělení $Bi(m; q)$ je definováno pravděpodobnostní funkcí:

$$P_m(x) = \binom{m}{x} q^x (1 - q)^{m-x} \text{ pro } x = 0, 1, 2, \dots, m.$$



Obrázek 1: Pravděpodobnostní funkce binomického rozdělení

Zdroj: STATGRAPHICS Centurion XVII

Základní charakteristiky binomického rozdělení [7][12]:

- Střední hodnota

$$E(X) = mq,$$

- Rozptyl

$$D(X) = mq(1 - q),$$

- Momentová vytvořující funkce

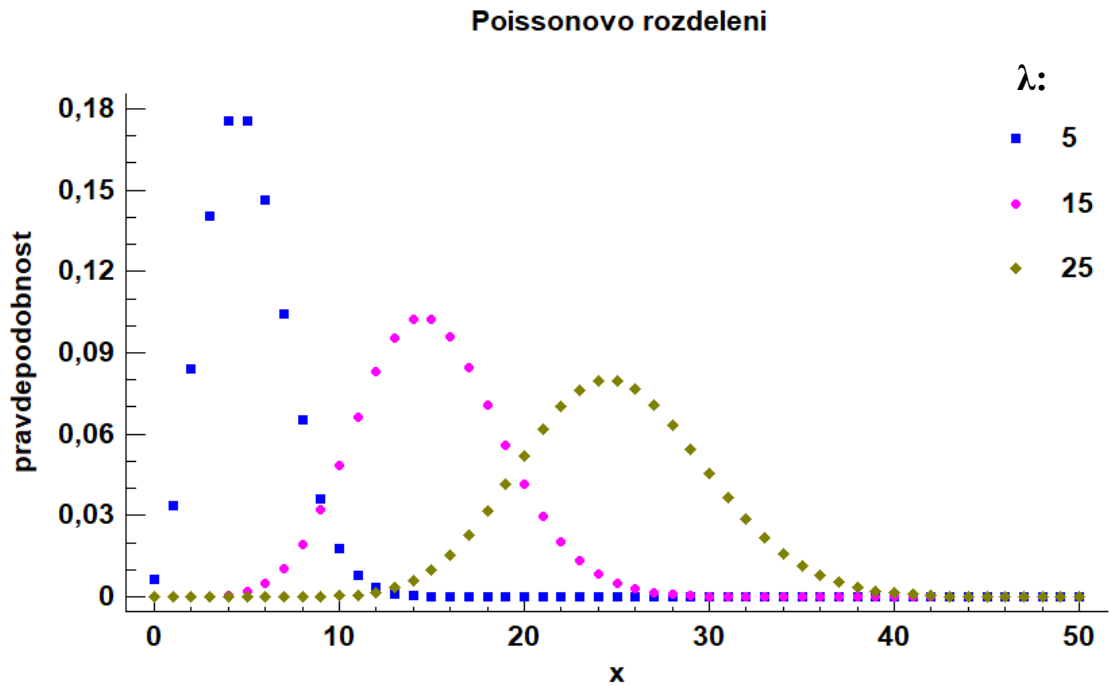
$$M_X(z) = (qe^z - q + 1)^m.$$

2. Poissonovo rozdělení $Po(\lambda)$

Toto rozdělení vzniká jako limitní případ binomického rozdělení $Bi(m; q)$, když současně platí $m \rightarrow \infty$ a $p \rightarrow \infty$, přičemž střední hodnota $mp = \lambda$. [11]

Poissonovo rozdělení $Po(\lambda)$ je definováno pravděpodobnostní funkcí:

$$P(x) = \frac{\lambda^x}{x!} e^{-\lambda} \text{ pro } x = 0, 1, 2, \dots$$



Obrázek 2: Pravděpodobnostní funkce Poissonova rozdělení

Zdroj: STATGRAPHICS Centurion XVII

Základní charakteristiky Poissonova rozdělení [12]:

- Střední hodnota a rozptyl

$$E(X) = D(X) = \lambda,$$

- Koeficient šikmosti

$$\gamma_1 = \frac{1}{\sqrt{\lambda}},$$

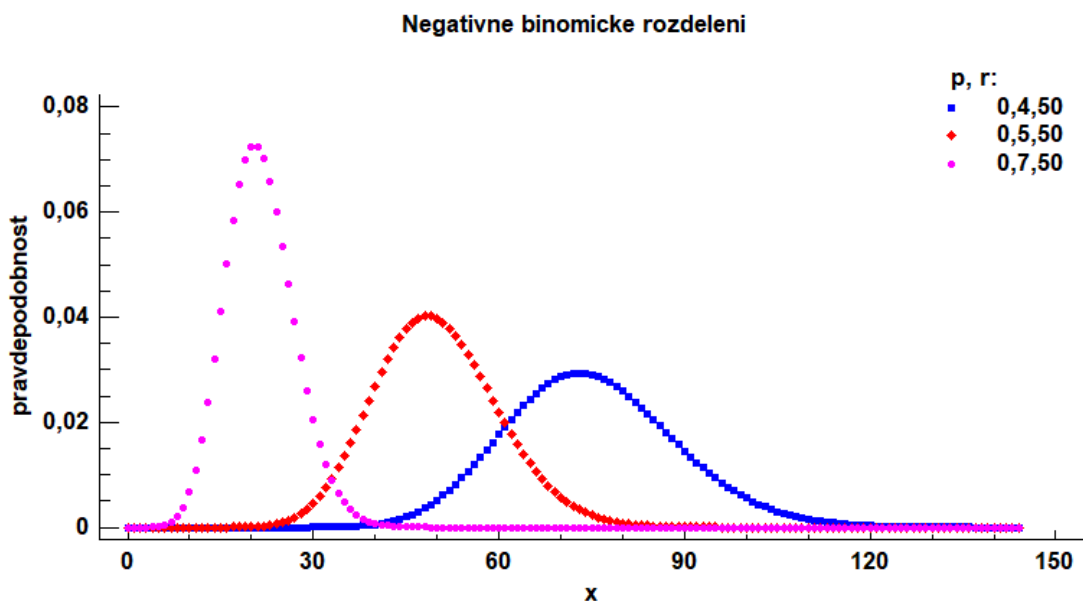
- Momentová vytvořující funkce

$$M_X(z) = e^{\lambda(e^z - 1)}.$$

3. Negativně binomické rozdělení $NBi(r; p)$

Negativně binomické rozdělení $NBi(r; p)$ je definováno pravděpodobnostní funkcí:

$$P(x) = \binom{x-1}{r-1} p^r (1-p)^{x-r} \text{ pro } x = k, k+1, \dots$$



Obrázek 3: Pravděpodobnostní funkce negativně binomického rozdělení

Zdroj: STATGRAPHICS Centurion XVII

Základní charakteristiky negativně binomického rozdělení [10]:

- Střední hodnota

$$E(X) = \frac{r(1-p)}{p},$$

- Rozptyl

$$D(X) = \frac{r(1-p)}{p^2},$$

- Momentová vytvořující funkce

$$M_X(z) = \left(\frac{p}{1-e^z(1-p)} \right)^r.$$

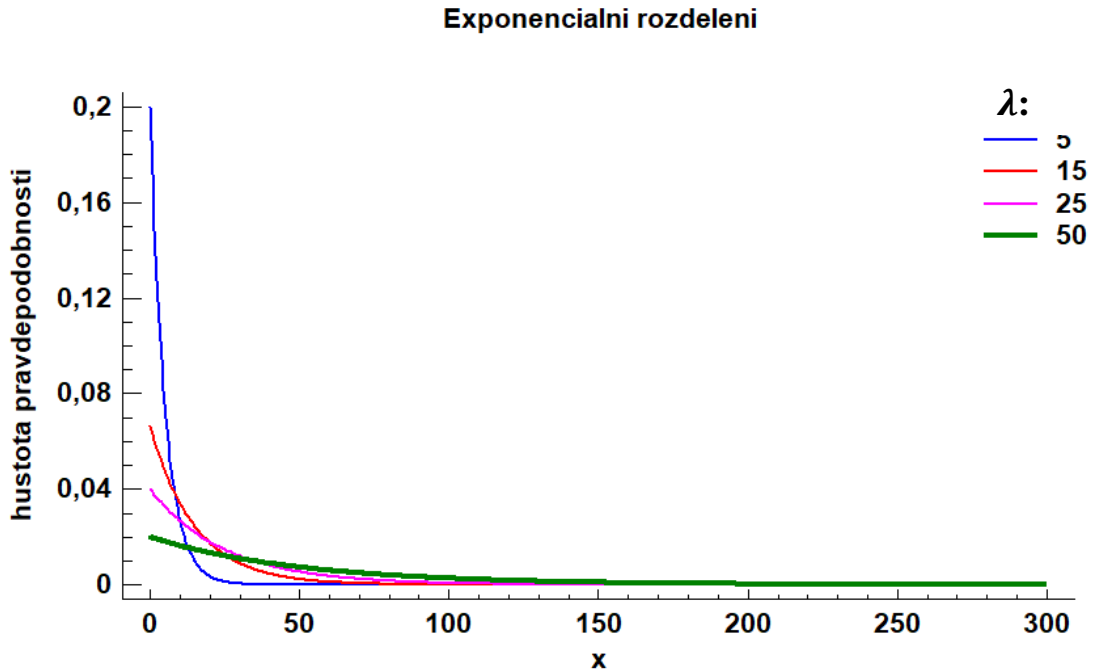
1.4 Rozdělení výšky pojistných plnění

Vhodnými rozděleními pravděpodobností pro výšky škod pojistných událostí jsou zejména pravostranně zešikmená spojitá rozdělení, tj. exponenciální, Weibullovo, lognormální, gama a Paretovo rozdělení. Je to z toho důvodu, že u většiny druhů neživotního pojištění mají výšky pojistných plnění hodnotu nižší, než je průměrná hodnota, ale i přesto jsou velmi pravděpodobné i extrémní škody. [12]

1. Exponenciální rozdělení $Exp(\lambda)$

Je definováno hustotou pravděpodobností [5]:

$$f(x) = \lambda e^{-\lambda x} \text{ pro } x > 0, \lambda > 0.$$



Obrázek 4: Hustota pravděpodobnosti exponenciálního rozdělení

Zdroj: STATGRAPHICS Centurion XVII

Exponenciální rozdělení je definováno distribuční funkcí:

$$F(x) = \begin{cases} 0 & \text{pro } x < 0, \\ 1 - e^{-\lambda x} & \text{pro } x \geq 0. \end{cases}$$

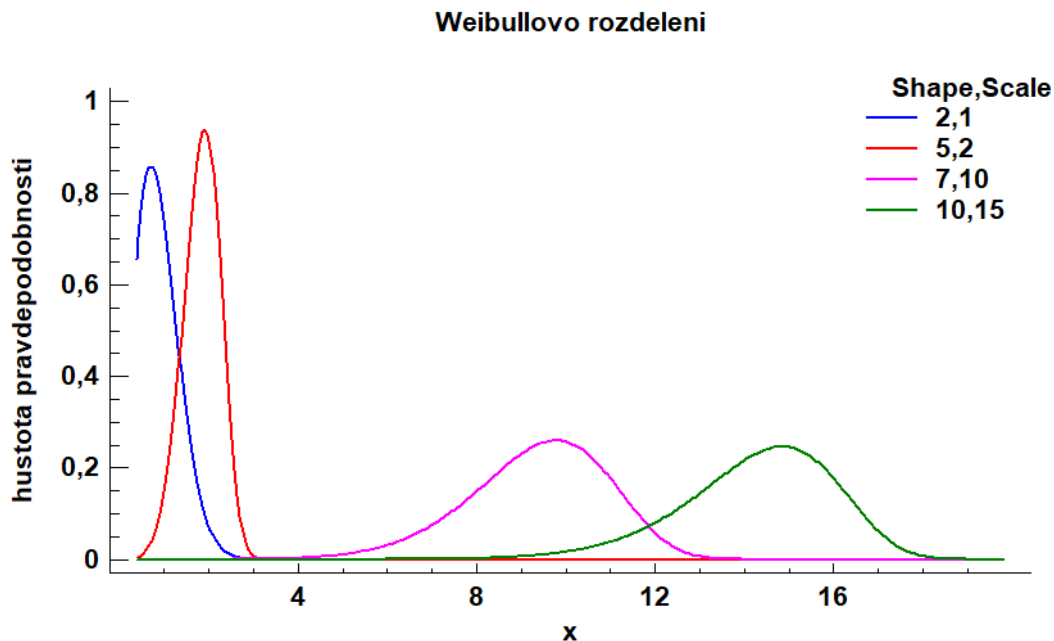
Základní charakteristiky $Exp(\lambda)$:

- Střední hodnota: $E(X) = \frac{1}{\lambda}$,
- Rozptyl: $D(X) = \frac{1}{\lambda^2}$,
- Koeficient šikmosti: $\gamma_1 = 2$
- Momentová vytvořující funkce: $M_X(z) = \frac{\lambda}{\lambda - z}$ pro $z > 0$.

2. Weibullovo rozdělení $W(\gamma; c)$

Je definováno hustotou pravděpodobnosti [9]:

$$f(x) = c\gamma x^{\gamma-1} e^{-cx^\gamma} \text{ pro } x > 0, c > 0, \gamma > 0.$$



Obrázek 5: Hustota pravděpodobnosti Weibullova rozdělení

Zdroj: STATGRAPHICS Centurion XVII

Weibullovo rozdělení je definováno distribuční funkcí:

$$F(x) = 1 - e^{-cx^\lambda} \text{ pro } x > 0.$$

Základní charakteristiky $W(\gamma; c)$:

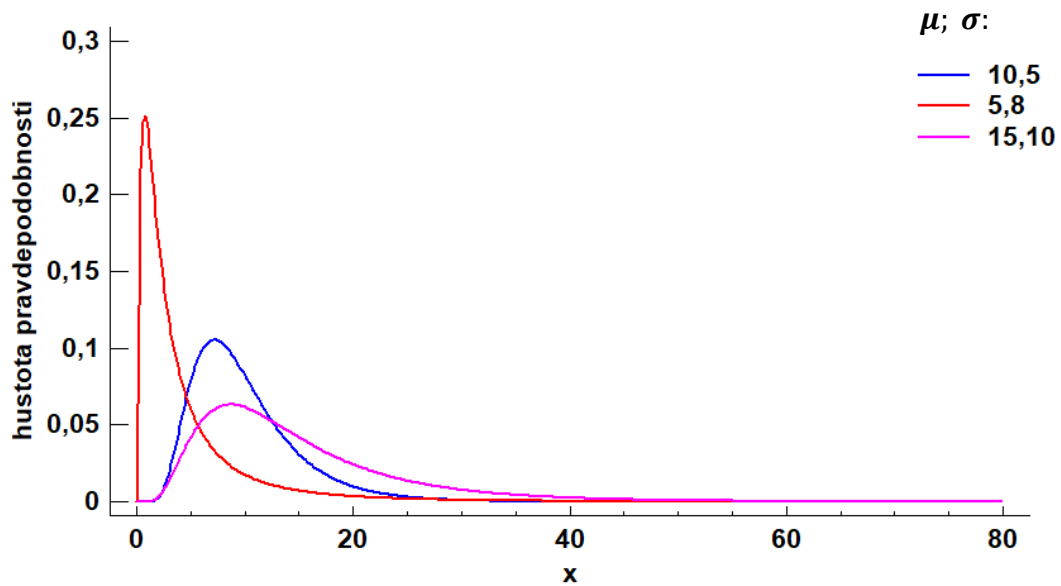
- Střední hodnota: $E(X) = c^{-\frac{1}{\gamma}} \Gamma\left(1 + \frac{1}{\gamma}\right)$,
- Rozptyl: $D(X) = c^{-\frac{2}{\gamma}} \left\{ \Gamma\left(1 + \frac{2}{\gamma}\right) - \left[\Gamma\left(1 + \frac{1}{\gamma}\right) \right]^2 \right\}$.

3. Lognormální rozdělení $LN(\mu; \sigma^2)$

Je definováno hustotou pravděpodobnosti [5]:

$$f(x) = \frac{1}{\sigma x \sqrt{2\pi}} e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}} \text{ pro } x > 0, \text{ jestliže } \ln X \sim N(\mu; \sigma^2).$$

Lognormalní rozdělení



Obrázek 6: Hustota pravděpodobnosti lognormálního rozdělení

Zdroj: STATGRAPHICS Centurion XVII

Lognormální rozdělení je definováno distribuční funkcí:

$$F(x) = \begin{cases} 0 & \text{pro } x \leq 0, \\ \frac{1}{\sigma\sqrt{2\pi}} \int_0^{\ln x} e^{-\frac{(y-\mu)^2}{2\sigma^2}} dy & \text{pro } x > 0. \end{cases}$$

Základní charakteristiky $LN(\mu; \sigma^2)$:

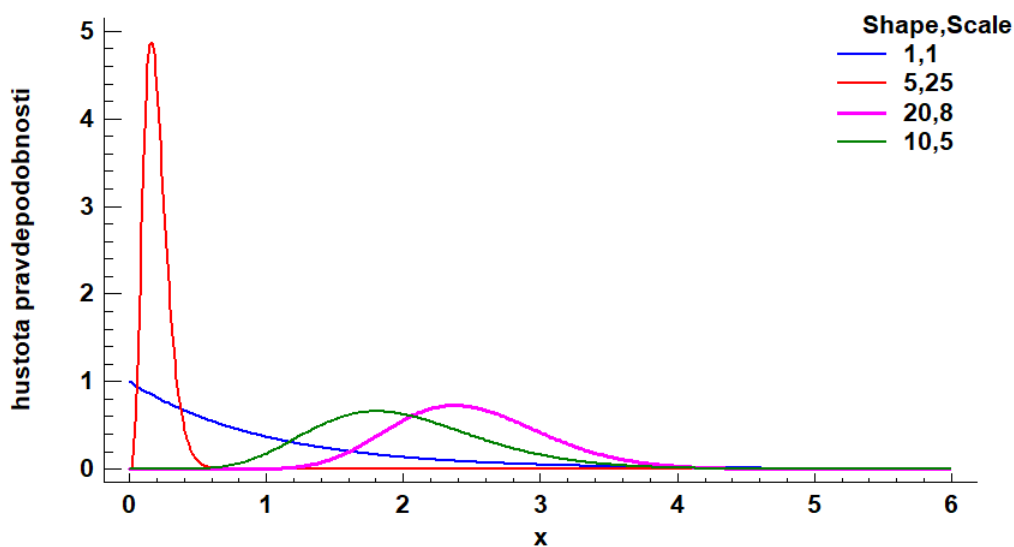
- Střední hodnota: $E(X) = e^{\mu + \frac{\sigma^2}{2}}$,
- Rozptyl: $D(X) = e^{2\mu + \sigma^2}(e^{\sigma^2} - 1)$,
- Koeficient šikmosti: $\gamma_1 = (e^{\sigma^2} + 2)\sqrt{e^{\sigma^2} - 1}$,
- Momentová vytvořující funkce: $M_X(z) = e^{\mu z + \frac{\sigma^2 z^2}{2}}$,
- k-tý počáteční moment: $E(X^k) = m_k = e^{\mu k + \frac{\sigma^2 k^2}{2}}$.

4. Gama rozdělení $G(\alpha; \beta)$

Je definováno hustotou pravděpodobnosti [9]:

$$f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{(\alpha-1)} e^{-\beta x} \text{ pro } x > 0.$$

Gama rozdělení



Obrázek 7: Hustota pravděpodobnosti gama rozdělení

Zdroj: STATGRAPHICS Centurion XVII

Základní charakteristiky $G(\alpha; \beta)$:

- Střední hodnota: $E(X) = \frac{\alpha}{\beta}$,
- Rozptyl: $D(X) = \frac{\alpha}{\beta^2}$,
- Koeficient šikmosti: $\gamma_1 = \frac{2}{\sqrt{\alpha}}$,
- Momentová vytvořující funkce: $M_X(z) = \left(\frac{\beta}{\beta-z}\right)^\alpha$.

1.5 Základní typy složených rozdělení

Doposud jsme se zabývali složeným rozdělením bez specifikace rozdělení počtu škod N a výšky pojistného plnění X_i . V následující kapitole uvedeme tři typy složených rozdělení, kde budeme uvažovat speciální rozdělení pro počet škod N .

1. Složené Poissonovo rozdělení

Prvním typem složeného rozdělení je složené Poissonovo rozdělení. Počet škod v tomto případě se řídí Poissonovým rozdělením $N \sim Po(\lambda)$ a náhodná proměnná S se potom řídí složeným Poissonovým rozdělením $S \sim CoPo(\lambda; F(x))$. [6]

Distribuční funkce podle obecné definice distribuční funkce kolektivního rizika má tvar [6]:

$$G(x) = F_S(x) = \sum_n \frac{e^{-\lambda} \lambda^n}{n!} F^{*n}(x) \text{ pro } x \geq 0.$$

Základní charakteristiky rozdělení $CoPo(\lambda; F(x))$ [10]:

- Střední hodnota

$$E(S) = \lambda m_1, \quad 1.6$$

- Rozptyl

$$D(S) = \lambda(m_2 - m_1^2) + \lambda m_1^2 = \lambda m_2, \quad 1.7$$

- Koeficient šikmosti

$$\gamma_1 = \frac{\lambda m_3}{(\lambda m_2)^2}, \quad 1.8$$

- Momentová vytvořující funkce

$$M_S(z) = e^{\lambda(M_X(z)-1)}.$$

Složené Poissonovo rozdělení je determinováno rozdělením Poissonovým, které opisuje počet škod v daném portfoliu. V případě kolektivního rizika připouštíme, že na jednu pojistnou smlouvu může nastat vícero pojistných plnění. V portfoliu w nezávislých pojistných smluv bude počet škod z první pojistné smlouvy daný náhodnou proměnnou $N_1 \sim Po(\lambda_1)$, počet škod z druhé pojistné smlouvy bude daný náhodnou proměnnou $N_2 \sim Po(\lambda_2)$. Pro w -tou pojistnou smlouvu bude počet škod daný náhodnou proměnnou $N_w \sim Po(\lambda_w)$. Za celé portfolio bude počet škod daný náhodnou proměnnou $N = N_1 + N_2 + \dots + N_w$ s rozdělením $N \sim Po(\lambda)$, $\lambda = \lambda_1 + \lambda_2 + \dots + \lambda_w$, které vystupuje v základním označení složeného Poissonova rozdělení. [6]

Ať S_1, S_2, \dots, S_n jsou nezávislé náhodné proměnné, např. celkové pojistné plnění v n pobočkách pojišťovny. Předpokládejme, že S_i má složené Poissonovo rozdělení s parametry λ_i a $F_i(x)$, tedy $S_i \sim CoPo(\lambda_i; F_i(x))$ pro $i = 1, 2, \dots, n$. Pokud označíme jejich součet $A = S_1 + S_2 + \dots + S_n$, pak A má také složené Poissonovo rozdělení s parametry λ a $F(x)$, kde platí

$$\lambda = \sum_{i=1}^n \lambda_i \text{ a } F(x) = \frac{1}{\lambda} \sum_{i=1}^n \lambda_i F_i(x),$$

λ je součtem Poissonových parametrů λ_i pro $\{S_i\}_{i=1}^n$ a $F(x)$ je váženým aritmetickým průměrem distribučních funkcí $\{F_i(x)\}_{i=1}^n$ s váhami λ_i . [10]

2. Složené binomické rozdělení

Dalším typem složeného rozdělení je složené binomické rozdělení. Počet škod se v tomto případě řídí binomickým rozdělením $N \sim Bi(m; q)$ a náhodná proměnná S se řídí složeným binomickým rozdělením $S \sim CoBi(m; q; F(x))$. [6]

Distribuční funkce podle obecné definice distribuční funkce kolektivního rizika má tvar [6]:

$$G(x) = F_S(x) = \sum_{n=0}^m \binom{m}{n} q^n (1-q)^{m-n} F^{*n}(x) \text{ pro } x \geq 0.$$

Základní charakteristiky rozdělení $CoBi(m; q; F(x))$ [10]:

- Střední hodnota

$$E(S) = mqm_1, \quad 1.9$$

- Rozptyl

$$D(S) = mq(m_2 - m_1^2) + mq(1-q)m_1^2 = mqm_2 - mq^2m_1^2, \quad 1.10$$

- Momentová vytvořující funkce

$$M_S(z) = (qM_X(z) + 1 - q)^m.$$

Konvolucí binomických rozdělení se stejnou pravděpodobností výskytu pojistné události je opět binomické rozdělení. Tedy platí, že v portfoliu w nezávislých pojistných smluv bude počet škod z první pojistné smlouvy daný náhodnou proměnnou $N_1 \sim Bi(m_1; q)$, počet škod z druhé pojistné smlouvy bude daný náhodnou proměnnou $N_2 \sim Bi(m_2; q)$. Pro w -tou pojistnou smlouvu bude počet škod daný náhodnou proměnnou $N_w \sim Bi(m_w; q)$. Za celé portfolio se počet škod popíše náhodnou proměnnou $N = N_1 + N_2 + \dots + N_w$ s rozdělením $N \sim Bi(m; q)$, $m = m_1 + m_2 + \dots + m_w$, které vystupuje v základním označení složeného binomického rozdělení. [6]

Mějme celková pojistná plnění S_1, S_2, \dots, S_n , která jsou nezávislá a mají identické složené binomické rozdělení. Potom jejich součet $A = S_1 + S_2 + \dots + S_n$ má také složené binomické rozdělení a platí $N \sim Bi(nm; q)$. [10]

3. Složené negativně binomické rozdělení

V pojistné praxi je také využíváným složeným rozdělením složené negativně binomické rozdělení. V tomto případě se tedy počet škod řídí negativně binomickým rozdělením $N \sim NBi(r; p)$ a celkové pojistné plnění S se řídí složeným negativně binomickým rozdělením $S \sim CoNBi(r; p; F(x))$. [6]

Distribuční funkce podle obecné definice distribuční funkce kolektivního rizika má tvar [6]:

$$G(x) = F_S(x) = \sum_{n=0}^{\infty} \binom{r+n-1}{r-1} p^r q^n F^{*n}(x) \text{ pro } x \geq 0.$$

Základní charakteristiky rozdělení $CoNBi(r; p; F(x))$ [10]:

- Střední hodnota

$$E(S) = \frac{r(1-p)}{p} m_1, \quad 1.11$$

- Rozptyl

$$D(S) = \frac{r(1-p)}{p} (m_2 - m_1^2) + \frac{r(1-p)}{p^2} m_1^2 = \frac{r(1-p)}{p} m_2 - \frac{r(1-p)^2}{p^2} m_1^2, \quad 1.12$$

- Momentová vytvořující funkce

$$M_S(z) = \frac{p^r}{(1-(1-p)M_X(z))^r}. \quad 1.13$$

Složené binomické rozdělení je determinované rozdělením negativně binomickým a pro rozdělení celkové škody platí stejná úvaha jako pro binomické rozdělení. Tedy pro portfolio w nezávislých pojistných smluv, kde počet škod z první pojistné smlouvy bude daný náhodnou proměnnou $N_1 \sim NBi(r_1; p)$, počet škod z druhé pojistné smlouvy bude daný náhodnou proměnnou $N_2 \sim NBi(r_2; p)$ a pro w -tou pojistnou smlouvu bude počet škod daný náhodnou proměnnou $N_w \sim NBi(r_w; p)$. Za celé portfolio počet škod vyjadřuje náhodná proměnná $N = N_1 + N_2 + \dots + N_w$ s rozdělením $N \sim NBi(r; p)$, $r = r_1 + r_2 + \dots + r_w$, které vystupuje v základním označení složeného binomického rozdělení. [6]

Mějme posloupnost nezávislých náhodných proměnných $\{S_i\}_{i=1}^n$, které mají identické složené negativně binomické rozdělení, kde $N_i \sim NBi(r; p)$. Potom součet $A = S_1 + S_2 + \dots + S_n$ má složené negativně binomické rozdělení, přičemž pro $N = \sum_{i=1}^n N_i$ platí, že $N \sim NBi(nr; p)$. [1][10]

1.6 Smíšená rozdělení v heterogenních portfoliích

Smíšená rozdělení jsou rozdělení, jejichž parametr je náhodná proměnná, která se řídí nějakým dalším typem rozdělení. Jsou to zejména rozdělení pravděpodobnosti počtu škod v heterogenních portfoliích pojistných smluv. [13]

1. Smíšené rozdělení počtu pojistných plnění

Počet pojistných plnění u jednotlivých pojistných smluv daného portfolia se často řídí Poissonovým rozdělením pravděpodobnosti. Označme N_i jako počet škod z i -té pojistné smlouvy během roka a N_i má Poissonovo rozdělení s parametrem λ_i , symbolicky $N_i|\lambda_i \sim Po(\lambda_i)$. Hodnoty λ_i se řídí v daném portfoliu n pojistných smluv podle určitého zákona rozdělení pravděpodobnosti, které označujeme jako smíšené, mixující rozdělení. Vhodným smíšeným rozdělením pro parametr λ_i Poissonova rozdělení je gama rozdělení, to můžeme zapsat takto [10]:

$$N_i|\lambda_i \sim Po(\lambda_i), \text{ kde } \lambda_i \sim G(\alpha; \beta) \text{ pro } i = 1, 2, \dots, n.$$

Pro marginální rozdělení pravděpodobnosti počtu škod v celém portfoliu pojistných smluv dostáváme vztah podle [1][10]:

$$\begin{aligned} P(x) &= P(N = x) = \int_0^{\infty} f_{N,\lambda}(x, \lambda) d\lambda = \int_0^{\infty} f_{\lambda}(\lambda) f_{N|\lambda}(x|\lambda) d\lambda = \\ &= \int_0^{\infty} \frac{\beta^{\alpha}}{\Gamma(\alpha)} \lambda^{\alpha-1} e^{-\beta\lambda} e^{-\lambda} \frac{\lambda^x}{x!} d\lambda = \frac{\beta^{\alpha}}{\Gamma(\alpha)x!} \int_0^{\infty} \lambda^{\alpha+x-1} e^{-\lambda(\beta+1)} d\lambda = \\ &= \frac{\beta^{\alpha}}{\Gamma(\alpha)x!} \cdot \frac{\Gamma(\alpha+x)}{(\beta+1)^{\alpha+x}} = \frac{\Gamma(\alpha+x)}{\Gamma(\alpha)x!} \left(\frac{\beta}{\beta+1}\right)^{\alpha} \left(\frac{1}{\beta+1}\right)^x. \end{aligned}$$

Můžeme je i přepsat ve tvaru [10]:

$$P(x) = \frac{\Gamma(\alpha+x)}{\Gamma(\alpha)x!} \pi^{\alpha} (1-\pi)^x \text{ pro } x = 0, 1, 2, \dots \quad 1.14$$

Tento vztah je odpovídající pravděpodobnostní funkci negativně binomického rozdělení $NBi(\alpha; \pi)$, kde $\pi = \frac{\beta}{\beta+1}$. Dále uvedeme základní charakteristiky. [10]

Pro střední hodnotu $E(N)$ rozdělení $NBi(\alpha; \pi)$ platí:

$$E(N) = E(E(N|\lambda)) = E(\lambda), \text{ protože } N|\lambda \sim Po(\lambda), \quad 1.15$$

$$E(N) = \frac{\alpha}{\beta}, \text{ protože } \lambda \sim G(\alpha; \beta), \quad 1.16$$

$$E(N) = \frac{\alpha(1-\pi)}{\pi}, \quad \text{protože} \quad \pi = \frac{\beta}{\beta+1}. \quad 1.17$$

Podobně dostaneme vztah pro $E(N^2)$:

$$E(N^2) = E(E(N^2|\lambda)) = E(\lambda + \lambda^2) \quad \text{protože} \quad N|\lambda \sim Po(\lambda),$$

$$E(N^2) = \frac{\alpha}{\beta} + \frac{\alpha}{\beta^2} + \left(\frac{\alpha}{\beta}\right)^2, \quad \text{protože} \quad \lambda \sim G(\alpha; \beta),$$

Pro rozptyl $D(N)$ dále platí:

$$D(N) = \frac{\alpha}{\beta} + \frac{\alpha}{\beta^2} = \frac{\alpha(1-\pi)}{\pi^2}. \quad 1.18$$

Pro heterogenní portfolio, které obsahuje n pojistných smluv, kde počty pojistných plnění N_i u jednotlivých pojistných smluv mají Poissonovo rozdělení s možnými parametry λ_i , $N_i|\lambda_i \sim Po(\lambda_i)$, kde $\lambda_i \sim G(\alpha; \beta)$ pro $i = 1, 2, \dots, n$, není Poissonovo rozdělení vhodné jako rozdělení celkového počtu pojistných plnění N za celé portfolio pojistných smluv v daném roce, protože pro Poissonovo rozdělení platí rovnost $E(N) = D(N) = \lambda$. Pro $E(N)$ a $D(N)$ negativně binomického rozdělení je však vztah vyjádřen nerovností:

$$E(N) = \frac{\alpha(1-\pi)}{\pi} < \frac{\alpha(1-\pi)}{\pi^2} = D(N).$$

Proto je toto rozdělení vhodnějším modelem počtu pojistných plnění v heterogenním portfolio pojistných smluv než Poissonovo rozdělení. [10]

2. Smíšené rozdělení výšky pojistných plnění

Předpokládejme, že pojistné plnění u každé individuální pojistné smlouvy v portfolio smluv se řídí exponenciálním rozdělením s parametrem δ . Označme θ_i průměrné pojistné plnění u jednotlivých pojistných smluv pro $i = 1, 2, \dots, n$. Pak $\delta_i = \frac{1}{\theta_i}$ je převrácená hodnota průměrného pojistného plnění i -té pojistné smlouvy. Předpokládejme, že δ_i jsou různé hodnoty náhodné proměnné δ , která má gama rozdělení, tj. $\delta \sim G(\alpha; \beta)$, s hustotou

$$f(\delta) = \frac{\beta^\alpha}{\Gamma(\alpha)} \delta^{\alpha-1} \cdot e^{-\beta\delta} \text{ pro } \delta > 0.$$

Pokud chceme najít rozdělení počtu pojistných plnění za celé portfolio pojistných smluv, použijeme toto rozdělení jako smíšené rozdělení parametru δ exponenciálního rozdělení. Výsledné rozdělení individuálních pojistných plnění v celém portfolio pojistných smluv bude smíšené rozdělení.

Pro marginální rozdělení individuálních pojistných plnění X v celém heterogenním portfoliu platí:

$$\begin{aligned}
 f_X(x) &= \int_0^{\infty} f_{X,\delta}(x, \delta) d\delta = \int_0^{\infty} f_{\delta}(\delta) \cdot f_{X|\delta}(x|\delta) d\delta = \\
 &= \int_0^{\infty} \frac{\beta^{\alpha}}{\Gamma(\alpha)} \delta^{\alpha-1} \cdot e^{-\beta\delta} \cdot \delta e^{-\delta x} d\delta = \frac{\beta^{\alpha}}{\Gamma(\alpha)} \int_0^{\infty} \delta^{\alpha} e^{-(x+\beta)\delta} d\delta = \\
 &= \frac{\beta^{\alpha}}{\Gamma(\alpha)} \cdot \frac{\Gamma(\alpha+1)}{(x+\beta)^{\alpha+1}} \int_0^{\infty} \frac{(x+\beta)^{\alpha+1}}{\Gamma(\alpha+1)} \delta^{\alpha} e^{-(x+\beta)\delta} d\delta = \\
 &= \frac{\alpha\beta^{\alpha}}{(x+\beta)^{\alpha+1}}
 \end{aligned}$$

což je hustota Paretova rozdělení $Pa(\alpha; \beta)$. Paretovo rozdělení je tedy rozdělení individuálních škod v takovém portfoliu pojistných smluv, kde jsou škody z každé pojistné smlouvy rozdělené exponenciálně a parametr tohoto rozdělení má gama rozdělení. [10]

2 REKURENTNÍ VYJÁDŘENÍ ROZDĚLENÍ KOLEKTIVNÍHO RIZIKA

2.1 Rekurentní vztah pro diskrétní rozdělení pojistných plnění

Předpokládejme, že známe rozdělení počtu pojistných plnění N a identické rozdělení výšky individuálních pojistných plnění X_i , pro $i = 1, 2, \dots, N$, které je diskrétním rozdělením definovaným pro kladná celá čísla. Předpoklad diskrétních rozdělení pro X_i a S může být omezující vzhledem ke spojitosti X_i a S v praxi. Proto je nutné přeskálovat spojitou náhodnou veličinu X_i na diskrétní, můžeme tento proces též nazvat diskretizace.[10]

Necht' $f_k = P(X_i = k)$, pro $k = 1, 2, 3, \dots$, je společná pravděpodobnostní funkce identicky rozdělených individuálních pojistných plnění X_i a $g_k = P(S = k)$ pro $k = 0, 1, 2, \dots$, je funkce pravděpodobnosti celkového pojistného plnění S .

Po diskretizaci náhodné proměnné X_i můžeme pomocí distribuční funkce $G(x)$ vyjádřit vztah g_x , pro hodnoty $x = 1, 2, 3, \dots$, takto:

$$g_x = P(S = x) = G(x) - G(x - 1) = \sum_{n=1}^{\infty} P(N = n)(F^{*n}(x) - F^{*n}(x - 1)).$$

Označíme-li $f_k^{*n} = F^{*n}(k) - F^{*n}(k - 1)$ jako pravděpodobnostní funkci součtu $\sum_{i=1}^n X_i$, dostaneme [10]:

$$g_x = \sum_{n=1}^x P(N = n)f_x^{*n} \text{ pro } x = 1, 2, \dots \quad 2.1$$

V případě, kdy X_i je spojitá náhodná proměnná, platí:

$$g_0 = P(S = 0) = P(N = 0). \quad 2.2$$

Pro znázornění uvedeme několik pravděpodobností g_x [10]:

- Jestliže $x = 1$, pak $S = 1$, $f_1^* = f_1$ a $g_1 = P(N = 1)f_1$.
- Jestliže $x = 2$, pak $S = 2$ a $N \leq 2$,
 $f_2^{*2} = P(X_1 + X_2 = 2) = P(X_1 = 1 \wedge X_2 = 1) = f_1^2$,
 $g_2 = P(S = 2) = P(N = 1)f_2 + P(N = 2)f_1^2$.
- Jestliže $x = 3$, pak $S = 3$ a $N \leq 3$,
 $f_3^{*1} = P(X_1 = 3) = f_3$,
 $f_3^{*2} = P(X_1 + X_2 = 3) = 2f_1 \cdot f_2$,
 $f_3^{*3} = P(X_1 + X_2 + X_3 = 3) = f_1^3$,
 $g_3 = \sum_{n=1}^3 P(N = n)f_3^{*n} = P(N = 1)f_3 + P(N = 2) \cdot 2f_1 \cdot f_2 + P(N = 3)f_1^3$.

- Jestliže $x = 4$, pak $S = 4$ a $N \leq 4$,

$$f_4^{*1} = P(X_1 = 4) = f_4,$$

$$f_4^{*2} = P(X_1 + X_2 = 4) = 2f_1 \cdot f_3 + f_2^2,$$

$$f_4^{*3} = P(X_1 + X_2 + X_3 = 4) = 3f_1^2 f_2,$$

$$f_4^{*4} = P(X_1 + X_2 + X_3 + X_4 = 4) = f_1^4,$$

$$g_4 = \sum_{n=1}^4 P(N = n) f_4^{*n} =$$

$$= P(N = 1)f_4 + P(N = 2)(2f_1 \cdot f_3 + f_2^2) + P(N = 3)3f_1^2 f_2 + P(N = 4)f_1^4.$$

Postup výpočtu pravděpodobností g_x , potřebných pro vyčíslení distribuční funkce $G(x)$ pro velké hodnoty k , je velmi náročný i s využitím výpočetní techniky. Efektivnějším způsobem výpočtu g_k je rekurentní vzorec. [10]

2.2 Panjerův rekurentní vzorec pro výpočet $G(x)$

Pomocí rekurentního vzorce můžeme snadněji vypočítat g_r náhodné veličiny S . Tento vzorec odvodil Harry H. Panjer (r. 1981) a můžeme ho použít, pokud je splněná podmínka, týkající se rozdělení počtu pojistných plnění. Mějme pravděpodobností funkci náhodné veličiny N danou vztahem $p_r = P(N = r)$. Předpokládejme, že existují konstanty a, b , pro které platí [10][14]:

$$p_r = \left(a + \frac{b}{r}\right) p_{r-1} \text{ pro } r = 1, 2, 3 \dots \quad 2.3$$

Upravíme vztah do tvaru:

$$\frac{p_r}{p_{r-1}} = \left(a + \frac{b}{r}\right). \quad 2.4$$

Pokud máme diskrétní individuální plnění X_i , s hodnotami $1, 2, 3 \dots$ a $f_k = P(X_i = k)$, pravděpodobnosti $g_r = P(S = r)$ může vypočítat podle Panjerova rekurentního vzorce [10]:

$$g_0 = p_0, \quad 2.5$$

$$g_r = \sum_{j=1}^r \left(a + \frac{bj}{r}\right) f_j \cdot g_{r-j}. \quad 2.6$$

Nyní nalezneme vyjádření pro konstanty a a b pro Poissonovo, binomické a negativně binomické rozdělení náhodné proměnné N .

- Mějme $N \sim Po(\lambda)$. Pro $r = 0, 1, 2, 3 \dots$ platí [7] [14]:

$$p_r = P(N = r) = \frac{\lambda^r}{r!} e^{-\lambda},$$

$$p_{r-1} = \frac{\lambda^{r-1}}{(r-1)!} e^{-\lambda},$$

$$p_0 = e^{-\lambda}.$$

Vyjádříme $\frac{p_r}{p_{r-1}}$:

$$\frac{p_r}{p_{r-1}} = \frac{\frac{\lambda^r}{r!} e^{-\lambda}}{\frac{\lambda^{r-1}}{(r-1)!} e^{-\lambda}} = \frac{\lambda}{r}.$$

Porovnáme-li toto vyjádření s vyjádřením 2.4, získáme hodnoty pro konstanty a, b :

$$a = 0, b = \lambda. \quad 2.7$$

Pravděpodobnostní funkci můžeme poté přepsat ve zjednodušeném tvaru [14][10]:

$$g_r = \frac{1}{1 - 0f_0} \sum_{j=1}^r \left(0 + \frac{\lambda \cdot j}{r}\right) f_j \cdot g_{r-j} = \frac{\lambda}{r} \sum_{j=1}^r j \cdot f_j \cdot g_{r-j}. \quad 2.8$$

- Mějme $N \sim Bi(n; \pi)$. Pro $r = 0, 1, 2, 3 \dots, n$ platí [7] [14]:

$$p_r = \frac{n!}{r! (n-r)!} \pi^r (1-\pi)^{n-r},$$

$$p_{r-1} = \frac{n!}{(r-1)! (n-r+1)!} \pi^{r-1} (1-\pi)^{n-r+1},$$

$$p_0 = (1-\pi)^n$$

Vyjádříme $\frac{p_r}{p_{r-1}}$:

$$\frac{p_r}{p_{r-1}} = \frac{\pi(n-r+1)}{r(1-\pi)} = \frac{-\pi}{1-\pi} + \frac{\pi(n+1)}{r(1-\pi)}.$$

Opět porovnáme toto vyjádření s vyjádřením 2.4 pro konstanty a, b .

$$a = \frac{-\pi}{1-\pi}, b = \frac{\pi(n+1)}{(1-\pi)}. \quad 2.9$$

Pravděpodobnostní funkci můžeme poté přepsat ve zjednodušeném tvaru [14][10]:

$$g_r = \frac{\pi}{1-\pi} \sum_{j=1}^r \frac{(n+1) \cdot j}{r-1} f_j \cdot g_{r-j}. \quad 2.10$$

- Mějme $N \sim NBi(k; \pi)$. Pro $r = 0, 1, 2, 3 \dots$ platí [7] [14]:

$$p_r = \frac{(k+r-1)!}{r! (k-1)!} \pi^k (1-\pi)^r,$$

$$p_{r-1} = \frac{(k+r-2)!}{(r-1)!(k-1)!} \pi^k (1-\pi)^{r-1},$$

$$p_0 = \pi^k.$$

Vyjádříme $\frac{p_r}{p_{r-1}}$:

$$\frac{p_r}{p_{r-1}} = \frac{(k+r-1)(1-\pi)}{r} = (1-\pi) + \frac{(k-1)(1-\pi)}{r}.$$

Porovnáním dostaneme konstanty [10]:

$$a = 1 - \pi, b = (k-1)(1-\pi). \quad 2.11$$

Pravděpodobnostní funkci můžeme poté přepsat ve zjednodušeném tvaru [6]:

$$g_r = (1-\pi) \sum_{j=1}^r \left(1 + \frac{(k-1) \cdot j}{r}\right) f_j \cdot g_{r-j}. \quad 2.12$$

Panjerovy konstanty pro Poissonovo, binomické a negativně binomické rozdělení jsou shrnuty v tabulce 1.

Tabulka 1: Panjerovy konstanty

Rozdělení	a	b	Parametr
$N \sim Po(\lambda)$	0	λ	$\lambda > 0$
$N \sim Bi(n; \pi)$	$\frac{-\pi}{1-\pi}$	$\frac{\pi(n+1)}{(1-\pi)}$	$\pi \in (0, 1)$
$N \sim NBi(k; \pi)$	$1-\pi$	$(k-1)(1-\pi)$	$\pi \in (0, 1), r > 0$

Zdroj: zpracováno dle [6]

3 PŘIBLIŽNÉ PRAVDĚPODOBNOSTNÍ MODELY KOLEKTIVNÍHO RIZIKA

V předchozí kapitole jsme vysvětlili rekurentní vyjádření distribuční funkce $G(x)$. Toto schéma však není vždy stabilní a hrozí, že i se správným rekurentním vzorcem může model produkovat chybné hodnoty. Aproximační způsoby výpočtu funkčních hodnot distribuční funkce $G(x)$ jsou velmi užitečné, a to zejména v těch případech, kdy je výpočet uskutečněn jednoduše a rychle pomocí dostupného základního softwaru. Tento model vyžaduje znalost základních charakteristik rozdělení náhodných proměnných N a X_i . [6][10]

3.1 Aproximace normálním rozdělením

Náhodnou proměnnou opisující celkovou škodu v kolektivním modelu rizika S můžeme při splnění předpokladů aproximovat pomocí centrální limitní věty Lindeberg-Lévyho.

Pro tuto větu musíme mít nezávislé a identicky rozdělené náhodné proměnné X_1, X_2, \dots, X_N , pro které existuje střední hodnota $\mu = E(S)$ a rozptyl $\sigma^2 = D(S)$, a jejich součet definujeme jako náhodnou proměnnou S , tj. $S = X_1 + X_2 + \dots + X_N$. Kolektivní model rizika S má tedy přibližné normální rozdělení, tj. [11]:

$$S \sim N(E(S); D(S)) = N(\mu; \sigma^2). \quad 3.1$$

Čím větší je počet N pojistných událostí, tím je aproximace $G(s)$ pomocí distribuční funkce normálního rozdělení lepší.

- Aproximace složeného Poissonova rozdělení

Náhodná proměnná S má složené Poissonovo rozdělení $S \sim CoPo(\lambda; F(x))$, potom pro $\lambda \rightarrow \infty$ má náhodná proměnná S normální rozdělení s parametry [6][10]:

$$\mu = \lambda m_1, \quad 3.2$$

$$\sigma^2 = \lambda m_2. \quad 3.3$$

- Aproximace složeného binomického rozdělení

Náhodná proměnná S má složené binomické rozdělení $S \sim CoBi(m; q; F(x))$, pro $m \rightarrow \infty$ a $q \rightarrow 0$ má náhodná proměnná S normální rozdělení s parametry [6][10]:

$$\mu = m q m_1, \quad 3.4$$

$$\sigma^2 = m_1 m_2 - m_1^2. \quad 3.5$$

- Aproximace složeného negativně binomického rozdělení

Náhodná proměnná S má složené negativně binomické rozdělení $S \sim \text{CoNBi}(r; p; F(x))$, pro $r \rightarrow \infty$ má náhodná proměnná S normální rozdělení s parametry [6][10]:

$$\mu = \frac{r(1-p)}{p} m_1, \quad 3.6$$

$$\sigma^2 = \frac{r(1-p)}{p} m_2 - \frac{r(1-p)^2}{p^2} m_1^2. \quad 3.7$$

Hustota pravděpodobnosti normálního rozdělení je symetrická, ale při různých typech pojištění bývají příslušné hustoty pravděpodobnosti výšky škod pomalu klesající, tj. pravděpodobnost extrémně velkých škod není zanedbatelná. Při aproximaci $G(s)$ normálním rozdělením dochází k podhodnocování vysokých hodnot, což pro pojistitele není žádoucí kvůli vysokým pojistným plněním v případě výskytu extrémní škody. Proto uvedeme dále aproximaci posunutým gama rozdělením, která dává přesnější výsledky.[11]

3.2 Aproximace posunutým gama rozdělením

Předpokládejme, že známe první tři začáteční momenty náhodné proměnné S a předpokládejme, že tato náhodná proměnná má přibližně gama rozdělení posunuté o konstantu k . Náhodná proměnná S má podobné rozdělení hodnot jako náhodná proměnná $Y + k$, kde $Y \sim G(\alpha; \beta)$ s distribuční funkcí [6]:

$$F_Y(x) = \int_0^x \frac{\beta^\alpha}{\Gamma(\alpha)} t^{\alpha-1} e^{-\beta t} dt.$$

Pro distribuční funkci náhodné proměnné $Y + k$ potom platí:

$$F_{Y+k}(x) = F_Y(x - k).$$

Jestliže náhodná proměnná S má být aproximována náhodnou proměnnou $Y + k$, musí mít tato náhodná proměnná shodné základní charakteristiky s adekvátními charakteristikami kolektivního rizika S , a to $E(S)$, $D(S)$, $\gamma(S)$, které známe. Neznámé parametry α , β , k vyjádříme podle [1][6][11] soustavou tří rovnic:

$$\mu = k + \frac{\alpha}{\beta}, \quad 3.8$$

$$\sigma^2 = \frac{\alpha}{\beta^2}, \quad 3.9$$

$$\gamma = \frac{2}{\sqrt{\alpha}}. \quad 3.10$$

Pomocí známých hodnot $E(S) = \mu$, $D(S) = \sigma^2$ a $\gamma(S) = \gamma$ celkového plnění S vyjádříme postupně parametry α, β, k [11]:

$$\alpha = \frac{4}{\gamma^2}, \quad 3.11$$

$$\beta = \sqrt{\frac{\alpha}{\sigma^2}}, \quad 3.12$$

$$k = \mu - \frac{\alpha}{\beta}. \quad 3.13$$

3.3 Určení rizikové přírážky k čistému pojistnému

Výše zmíněné aproximace můžeme využít pro výpočet rizikové přírážky θ k čistému neboli netto pojistnému. Netto pojistné NP v neživotním pojištění vyjadřuje střední hodnotu $E(S)$ kolektivního rizika a je určeno k pokrytí průměrných výdajů pojišťovny na pojistná plnění. Při výpočtu netto pojistného se vychází z průměrných hodnot zejména o počtu pojistných událostí a výšky škod z minulých let, proto platí $NP = E(S)$. [4]

Při určování rizikové přírážky, se pojistitel snaží, aby rizikové pojistné bylo sestaveno tak, že celkové pojistné plnění s vysokou pravděpodobností, např. 0,95, bylo nižší než celkové přijaté pojistné. Jestliže průměrné pojistné plnění $E(S)$ je čisté pojistné, potom požadujeme, aby se rizikové pojistné $(1 + \theta)E(S)$ rovnalo 95. percentilu $S_{0,95}$ rozdělení S . To vyjádříme zápisem [11]:

$$RP = E(S)(1 + \theta) = S_{0,95}. \quad 3.14$$

Vyjádřením ze vztahu 3.8, získáváme vztah pro výpočet rizikové přírážky:

$$\theta = \frac{S_{0,95} - E(S)}{E(S)}. \quad 3.15$$

Vztah 3.9 můžeme zapsat ve tvaru [10]:

$$P(S \leq S_{0,95}) = P(S \leq (1 + \theta)E(S)) = 0,95 \quad 3.16$$

Při normální aproximaci celkového pojistného plnění S , $S \sim N(E(S); D(S))$, normováním obou stran nerovnosti v závorce ve vztahu 3.10, dostáváme [10]:

$$P\left(\frac{S - E(S)}{\sqrt{D(S)}} \leq \frac{\theta E(S)}{\sqrt{D(S)}}\right) = 0,95.$$

Protože platí:

$$\frac{S - E(S)}{\sqrt{D(S)}} \approx N(0; 1),$$

dostaneme vztah:

$$\frac{\theta E(S)}{\sqrt{D(S)}} = z_{0,95}, \quad 3.17$$

kde $z_{0,95}$ vyjadřuje 95. percentil normovaného normálního rozdělení. Označme $\sigma(S) = \sqrt{D(S)}$, pak pro obecné z_α můžeme rizikovou přírážku vyjádřit vztahem [10]:

$$\theta = \frac{z_\alpha \sigma(S)}{E(S)}. \quad 3.18$$

4 SIMULAČNÍ MODELY KOLEKTIVNÍHO RIZIKA

V pojistně-matematické praxi mají své významné místo simulace, hlavně pro simulaci extrémních škod. V této kapitole se zaměříme na simulační metodu Monte Carlo. Základní myšlenka vychází z přímého napodobování reálného systému, pomocí poznatků z matematiky, teorie pravděpodobnosti, statistiky a programování. [6]

Při realizaci náhodné veličiny získané skutečnými událostmi, je velmi těžké zaznamenat vysoký počet pozorování. Pomocí simulace Monte Carlo je možné získat uměle vytvořené pozorování, a tím získat tisíce náhodných realizací. Princip simulace vychází ze zákona velkých čísel, kdy s rostoucím počtem realizací náhodné veličiny se budou pozorované charakteristiky i odhadnutá funkce hustoty blížit teoretickému předpokladu.[16]

Metoda Monte Carlo je numerickou metodu a touto metodou je možné řešit libovolné matematické úlohy, a nejen pravděpodobnostního charakteru. [3]

Základní přiblížení simulací [6]:

1. Vytvoříme model S , který závisí na náhodných proměnných X, Y, Z, N, \dots se známým rozdělením a vzájemné závislosti,
2. Pro $j = 1, 2, \dots, n$ se generují pseudonáhodné hodnoty $x_j, y_j, z_j, n_j \dots$ a následně se vypočítá s použitím modelu specifikovaným v kroku 1. hodnota s_j .
3. Rozdělení náhodné proměnné S v podobě distribuční funkce může být aproximované $F(s)$, empirickým rozdělením stanoveným na základě pseudonáhodného vzorku s_1, s_2, \dots, s_n .
4. Vypočítáme hodnoty, jako je průměr, rozptyl, percentily anebo pravděpodobnosti a míry rizika pomocí empirického rozdělení.

Generátor pseudonáhodných čísel je deterministický algoritmus, který generuje posloupnost čísel, která by neměla být klasickými statistickými testy rozlišitelná od náhodné posloupnosti. Některé metody generování pseudonáhodných čísel generují hodnoty z rovnoměrného normovaného rozdělení, tj. výskyt libovolného čísla z intervalu $(0, 1)$ je stejně pravděpodobný a mimo tento interval je pravděpodobnost nulová. Toto rozdělení ale v praxi není dostačující a využívají se různé metody transformace. [6]

Čtyři obecné metody pro transformaci náhodných čísel s rovnoměrným normovaným rozdělením na čísla s požadovaným rozdělením jsou[3]:

- Rozehrávání diskrétní náhodné veličiny,
- Metoda inverzní funkce (transformace),
- Metoda výběru (metoda Neumannova),
- Metoda superpozice.

4.1 Metoda Monte Carlo pro stanovení rozdělení celkové škody

Metoda Monte Carlo numericky řeší pravděpodobnostní úlohy pomocí organizovaných statistických pokusů. Získáváme tak potřebné řešení problémů pomocí uměle realizovaných náhodných procesů. Základní myšlenkou je hledání souvislostí mezi veličinami, které jsou řešením zkoumaného problému, a charakteristikami náhodných procesů modelovaných počítačem. [6]

Simulovaným náhodným procesem budou hodnoty s_1, s_2, \dots, s_n , kterými nabývá náhodná proměnná celkové škody S , realizované prostřednictvím simulací hodnot počtu pojistných plnění N a výšky pojistných plnění X . Předpokládáme znalost rozdělení a parametrů náhodných proměnných N a X . [6]

Prvním krokem pro aplikaci metody pro určení rozdělení pravděpodobnosti celkové škody S , kde $S = X_1, X_2, \dots, X_N$ souvisí s rozdělením počtu škod N . Mějme $r \in (0; 1)$ náhodně vygenerované číslo a \tilde{n}_1 je řešením rovnice $r = F_N(\tilde{n}_1)$, kde $F_N(n)$ je distribuční funkce náhodné proměnné N , která opisuje počet škod v daném portfoliu pojistných smluv. Proto hodnotu \tilde{n}_1 můžeme považovat za simulovaný počet škod. [6]

Následným vygenerováním náhodných čísel $r_{11}, r_{12}, \dots, r_{1\tilde{n}_1}$, kde $r_{1j} \in (0; 1)$ získáme funkční hodnoty distribuční funkce výšky pojistných plnění $F_{X_i}(x)$. Jestliže náhodná proměnná X_i je spojitá, tak k její distribuční funkci v intervale, kde je rostoucí, existuje inverzní funkce F_X^{-1} . Pak ke každému z čísel r_{1j} z definičního oboru F_X^{-1} , kde $j = 1, 2, \dots, \tilde{n}_1$, je jednoznačně určena hodnota \tilde{x}_{1j} předpisem $r_{1j} = F_X(\tilde{x}_{1j})$, resp. $\tilde{x}_{1j} = F_X^{-1}(r_{1j})$, kde \tilde{x}_{1j} představuje hodnotu náhodné proměnné výšky individuální škody X_i , kterou můžeme vygenerovat i jiným algoritmem v závislosti od požadovaného rozdělení. Hodnotu \tilde{x}_{11} můžeme považovat za výšku první škody z počtu \tilde{n}_1 a pak hodnotu $\tilde{s}_1 = \tilde{x}_{11} + \tilde{x}_{12} + \dots + \tilde{x}_{1\tilde{n}_1}$ za první člen vzorku náhodné proměnné S . [6]

Opakováním tohoto algoritmu n -krát dostaneme n realizací hodnot celkové škody S , $\check{s}_1, \check{s}_2, \dots, \check{s}_n$. Jestliže označíme m_x jako počet všech $\check{s}_i, i = 1, 2, \dots, n$, pro které platí $\check{s}_i \leq x$, tak dostaneme odhad distribuční funkce celkové škody S dle vztahu [6]:

$$G(x) = F_S(x) \approx \frac{m_x}{n}. \quad 4.1$$

5 PREZENTACE MODELŮ NA REÁLNÝCH DATECH

V této kapitole bude provedena prezentace dříve vysvětlených modelů kolektivního rizika na reálných datech. Nejprve budou uvedeny základní charakteristiky vstupních dat a hledání vhodného rozdělení pravděpodobnosti. Následně bude provedena aproximace daného modelu. Závěrem kapitoly bude naznačen postup simulace metodou Monte Carlo. Data budou zpracována ve statistickém programu STATGRAPHICS Centurion XVII a výpočty provedeny v tabulkovém procesoru MS Excel.

5.1 Analýza vstupních dat

Prvním krokem při realizaci modelů ze vstupních dat je zjištění základních informací, tj. stanovení základních charakteristik, jako je počet, součet, rozptyl, průměr, medián a koeficient šikmosti a špičatosti. Vhodné je určit i minimální a maximální hodnotu, horní a dolní kvantil datového souboru.

Data můžeme před samotnou analýzou podrobit grafické analýze, která nám usnadní určit následující kroky. Vhodnými porovnávacími nástroji jsou například box plot (krabicový graf, který slouží k zjištění rozdílů v poloze dat a jejich variabilitě), histogram (vhodný pro znázornění distribuce vzorku dat pomocí sloupcových grafů stejné šířky intervalů, kde výška sloupců vyjadřuje četnost sledované veličiny), nebo dot plot (bodový graf, který se používá k zobrazení distribuce dat jednotlivými body na ose). [9]

V této kapitole využijeme reálná data o počtu a výšce pojistných plnění v heterogenním portfoliu pojistných smluv na pojištění odpovědnosti za škodu způsobenou provozem vozidla. K dispozici máme 25 649 pojistných smluv, kdy bylo šetřeno 951 pojistných událostí z 876 pojistných smluv během jednoho sledovaného roku, viz. Příloha A – Vstupní data. Nejprve provedeme analýzu první části dat, tj. počtu škod, a následně výšky pojistných plnění.

V prvním kroku rozdělíme smlouvy do tříd škod podle počtu nahlášených pojistných událostí, v našem případě se jedná o třídy 0 – 3. V tabulce 2 můžeme vidět, že do nulové třídy škod, tj. do třídy, kdy na jedné smlouvě nebyla šetřena žádná pojistná událost, patří 24 773 bezeškodných pojistných smluv, to činí 96,58 % smluv. Do první třídy škod pak patří pojistné smlouvy s jednou pojistnou událostí, tj. 806 smluv, což představuje 3,14 % smluv. V druhé třídě je pouze 65 smluv s relativní četností 0,25 %. Během roka byly nejvíce nahlášeny tři události připadající na jednu pojistnou smlouvu, a to v pěti případech smluv. Celkem bylo pojišťovnou šetřeno

951 událostí z 876 pojistných smluv. V tabulce 3 jsou znázorněny základní charakteristiky datového souboru. Grafické znázornění můžeme vidět na histogramu, obrázek 8.

Tabulka 2: Rozdělení počtu škod do tříd

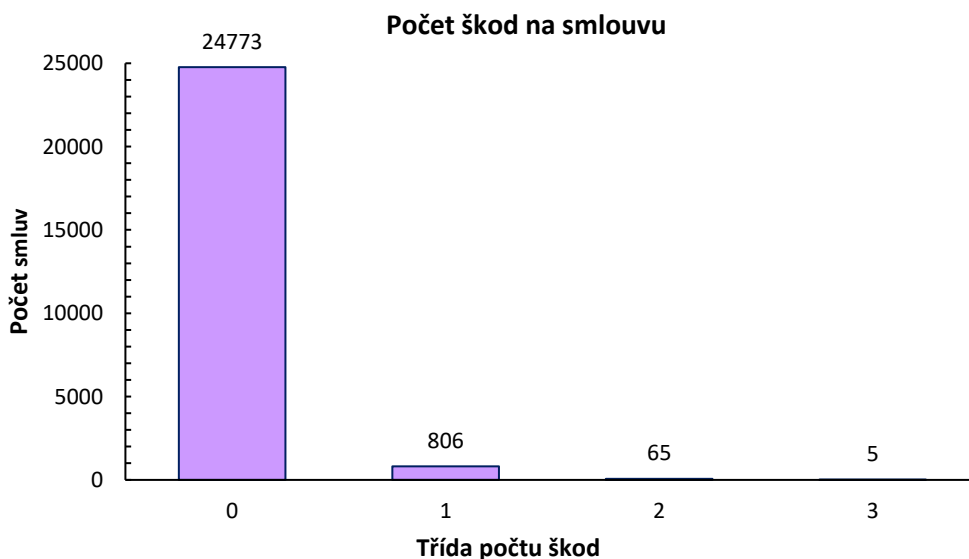
Třída škod	Počet smluv	Relativní četnost
0	24773	0,9658
1	806	0,0314
2	65	0,0025
3	5	0,0002

Zdroj: vlastní zpracování dle STATGRAPHICS Centurion XVII

Tabulka 3: Analýza počtu škod

Počet	25649
Součet	951
Průměr	0,03708
Minimum	0
Maximum	3

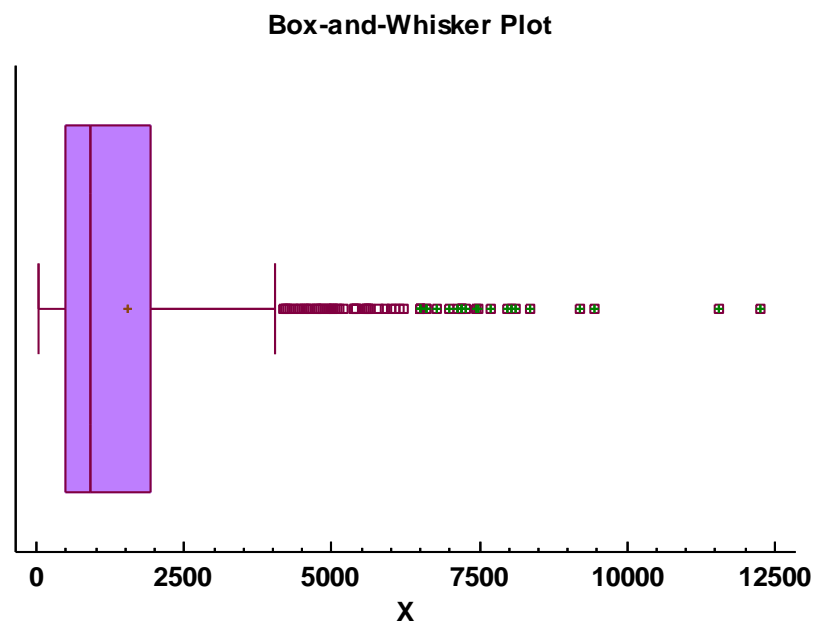
Zdroj: vlastní zpracování dle STATGRAPHICS Centurion XVII



Obrázek 8: Histogram počtu individuálních škod

Zdroj: STATGRAPHICS Centurion XVII

Nyní provedeme analýzu druhé části dat, a to výšky pojistných plnění X_i . Jednotky výšek pojistných plnění jsou 100 CZK, pro usnadnění budeme uvádět pouze označení PJ, peněžní jednotky. Nejprve vykreslíme graf všech pojistných plnění pomocí box plotu zobrazení, obrázek 9. Vidíme, že v datovém souboru je více škod s menším pojistným plněním, než je průměrná hodnota (+). Je zřejmé, že data mají pravostranné zešíkmené rozdělení pravděpodobnosti.



Obrázek 9: Krabicový graf výšky pojistných plnění

Zdroj: STATGRAPHICS Centurion XVII

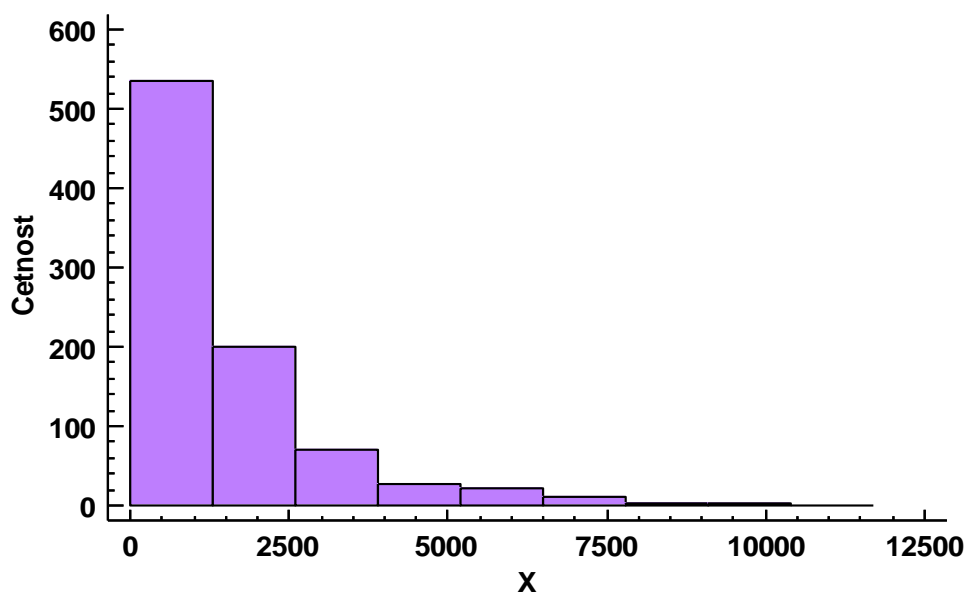
V tabulce 4 je uvedeno rozdělení dat do deseti intervalů s délkou 1 300 PJ. V prvním intervalu (0; 1300) je šetřeno téměř 60,96 % všech plnění, tj. 534 pojistných plnění. V druhém intervalu se nachází 201 škod, což představuje 22,95 %. V sedmém intervalu (7800; 9100) se nachází pouze čtyři pojistná plnění. V devátém a desátém intervalu se nachází vždy pouze jedna škoda s relativní četností 0,11 %. Rozdělení do těchto intervalů je znázorněno v histogramu na obrázku 10.

Tabulka 4: Rozdělení pojistných plnění dle intervalů

Interval: [PJ]	od	do	Četnost	Relativní četnost
1	0	1300	534	0,6096
2	1300	2600	201	0,2295
3	2600	3900	71	0,0811
4	3900	5200	28	0,0320
5	5200	6500	22	0,0251
6	6500	7800	12	0,0137
7	7800	9100	4	0,0046
8	9100	10400	2	0,0023
9	10400	11700	1	0,0011
10	11700	13000	1	0,0011

Zdroj: vlastní zpracování dle STATGRAPHICS Centurion XVII

Histogram X



Obrázek 10: Histogram výšky pojistných plnění

Zdroj: STATGRAPHICS Centurion XVII

Celkem byla vyplacena pojistná plnění ve výšce 1 333 739,22 PJ za škody z 876 pojistných smluv. Minimální výše byla 24,42 PJ a nejvyšší vyplacené pojistné plnění bylo ve výši 12 254,00 PJ. Průměrná výše činí 1 522,53 PJ na jednu škodu. Pravostrannost dat potvrzuje koeficient šikmosti, $\gamma_1 = 2,37 > 0$. Data jsou také výrazně špičatější než normální rozdělení, $\gamma_2 = 7,35 > 0$. Základní charakteristiky těchto dat jsou uvedeny v tabulce 5.

Tabulka 5: Základní charakteristiky výšky pojistných plnění

Počet	876
Součet	1333739,22
Průměr	1522,53
Medián	922,39
Minimum	24,42
Maximum	12254,00
Směrodatná odchylka	1621,48
Variační koeficient [%]	106,50
Koeficient šikmosti	2,36728
Koeficient špičatosti	7,35035

Zdroj: vlastní zpracování dle STATGRAPHICS Centurion XVII

5.2 Hledání vhodného rozdělení pro vstupní data

Dalším postupem bude nalezení vhodného pravděpodobnostního rozdělení, které spolehlivě modeluje vstupní data. Testování vhodnosti daného rozdělení budeme provádět pomocí statistického softwaru STATGRAPHICS Centurion XVII, kde si můžeme vybrat z různých testů dobré shody. Závěrem této části bude ukázka zpracování v tabulkovém procesoru MS Excel.

5.2.1 Vhodné rozdělení počtu škod N a odhad parametrů

Vhodným rozdělením pro počet pojistných plnění N může být alternativní, binomické, Poissonovo nebo negativně binomické rozdělení. [10]. Základní charakteristiky těchto rozdělení jsou uvedeny v kapitole 1.3, kde můžeme vidět grafy jejich pravděpodobnostních funkcí. V našem případě budeme testovat, zda data o počtu pojistných plnění pocházejí z Poissonova nebo negativně binomického rozdělení pravděpodobnosti.

Testování vhodnosti daného rozdělení provedeme pomocí softwaru. Využijeme Pearsonův χ^2 test dobré shody, který slouží na ověření shody rozdělení četností výběrových dat a předpokládaným teoretickým rozdělením s hustotou pravděpodobnosti $f(x; \Theta)$, kde Θ je vektor parametrů, nejčastěji odhadnutých z výběrového souboru. Testujeme nulovou hypotézu H_0 : náhodná proměnná X má rozdělení s hustotou $f(x; \Theta)$. Východiskem pro test jsou empirické údaje x_1, x_2, \dots, x_n náhodné proměnné X rozříděné do k skupin s četnostmi O_1, O_2, \dots, O_n . Pro tento test je dáno testovací kritérium:

$$\chi^2 = \sum_{i=1}^k \frac{(O_i - np_i)^2}{np_i} \sim \chi_{k-1-p}^2, \quad 5.1$$

kde p_i je pravděpodobnost hodnoty x_i diskrétní proměnné s předpokládaným rozdělením, které má p odhadovaných parametrů. H_0 přijmeme na hladině významnosti α , jestliže testovací kritérium χ^2 nepřekročí percentil $\chi_{1-\alpha}^2$ rozdělení χ_{k-1-p}^2 . [10]

Byla zvolena hladina významnosti $\alpha = 0,05$, tj. 95. percentil χ^2 –rozdělení. A postupně testujeme nulovou hypotézu:

- H_0 : Počty pojistných plnění pocházejí z Poissonova rozdělení.
- H_0 : Počty pojistných plnění pocházejí z negativně binomického rozdělení.

Prvně testujeme, zda data pocházejí z Poissonova rozdělení, zda $N \sim Po(\lambda)$. Výsledky jsou znázorněny v tabulce 6. Vidíme, že v testu došlo ke sloučení druhé a třetí třídy, tedy smluv,

u kterých vznikly dvě a tři pojistné události. Pozorovaná četnost, po sloučení, činí 70 pojistných smluv a odhadovaná očekávaná četnost je 17,20. Testovací kritérium $\chi_1^2 = 175,506$, P-hodnota je rovna nule, kritická hodnota pro zvolenou hladinu významnosti, pro tři třídy a jeden odhadovaný parametr je $KH_1 = 5,99$.

Tabulka 6: χ^2 test – Poissonovo rozdělení

Goodness-of-Fit Tests for N
Chi-Square Test

	Lower Limit	Upper Limit	Observed Frequency	Expected Frequency	Chi-Square
at or below		0	24773	24715,41	0,13
	1	1	806	916,39	13,30
	2		70	17,20	162,08

Chi-Square = 175,506 with 1 d.f. P-Value = **0,0**

Zdroj: výstup z STATGRAPHICS Centurion XVII

Následně budeme testovat druhou námi zvolenou nulovou hypotézu a to, zda data pocházejí z negativně binomického rozdělení, zda $N \sim NBi(r, q)$. Výsledky testu ze softwaru jsou uvedeny v tabulce 7. Zde nedošlo k žádným úpravám tříd a v testu je použito stejné rozložení jako v tabulce 2. Testovací kritérium $\chi_2^2 = 0,530044$, P-hodnota je rovna 0,466587, kritická hodnota pro zvolenou hladinu významnosti, pro čtyři třídy a dva odhadované parametry je $KH_2 = 7,81$.

Tabulka 7: χ^2 test – Negativně binomické rozdělení

Goodness-of-Fit Tests for N
Chi-Square Test

	Lower Limit	Upper Limit	Observed Frequency	Expected Frequency	Chi-Square
at or below		0	24773	24770,81	0,00
	1	1	806	811,91	0,04
	2	2	65	60,39	0,35
	3		5	5,89	0,14

Chi-Square = 0,530044 with 1 d.f. P-Value = **0,466587**

Zdroj: výstup z STATGRAPHICS Centurion XVII

S 95 % spolehlivostí zamítáme hypotézu, že data pocházejí z Poissonova rozdělení pravděpodobnosti, P-hodnota je menší než námi zvolená hladina významnosti, $\chi_1^2 > KH_1$. Přijímáme hypotézu, že data pocházejí z negativně binomického rozdělení, P-hodnota je vyšší než zvolená hladina významnosti, $\chi_2^2 < KH_2$. Dále tedy budeme pracovat s tímto rozdělením pro počty pojistných plnění. Porovnání výsledků obou testů znázorňuje tabulka 8.

Tabulka 8: χ^2 test počtu škod pro vybraná rozdělení

χ^2 test	Poissonovo	Neg. binomické
Chi-Square	175,506	0,530
D.f.	1	1
P-Value	0	0,466587

Zdroj: vlastní zpracování dle STATGRAPHICS Centurion XVII

Nyní, když známe vhodné rozdělení dat, můžeme v softwaru odhadnout parametry rozdělení. Software provádí odhad pomocí metody maximální věrohodnosti (MMV, z anglického maximum likelihood možno i ML). Jedná se o metodu, kterou můžeme aplikovat v mnohých situacích, a odhady, získané pomocí této metody, mají velmi dobré vlastnosti.

Nechť $\mathbf{x} = (x_1, x_2, \dots, x_n)'$ je vektor n nezávislých výběrových pozorování náhodného pozorování X s rozdělením $f(x; \theta_1, \theta_2, \dots, \theta_p)$, kdy $\theta_1, \theta_2, \dots, \theta_p$ jsou neznámé parametry tohoto rozdělení. Pro libovolný parametr θ definujeme funkci maximální věrohodnosti $L(\theta; \mathbf{x})$ vztahem:

$$L(\theta; \mathbf{x}) = \prod_{i=1}^n f(x_i; \theta). \quad 5.2$$

Přirozený logaritmus funkce věrohodnosti je:

$$l(\theta; \mathbf{x}) = \ln L(\theta; \mathbf{x}) = \sum_{i=1}^n \ln f(x_i; \theta). \quad 5.3$$

Maximální věrohodný odhad parametru θ je taková výběrová charakteristika $\hat{\theta} = \hat{\theta}(\mathbf{x})$, která maximalizuje funkci věrohodnosti $L(\theta; \mathbf{x})$, resp. $l(\theta; \mathbf{x})$. [12]

Maximálně věrohodný odhad má tyto výhodné vlastnosti [11]:

- $\hat{\theta}$ má asymptoticky normální rozdělení,
- $\hat{\theta}$ je asymptoticky nezkreslený,
- $\hat{\theta}$ je asymptoticky vydatný,
- $\hat{\theta}$ je konzistentní,
- $\hat{\theta}$ je invariantní, tj. $\hat{\gamma}(\theta) = \gamma(\hat{\theta})$, kde $\gamma(\theta)$ je libovolná funkce odhadnutých parametrů.

Další možností odhadu je metoda momentů (MM), která spočívá v odhadu k -tého začátečního momentu rozdělení náhodné proměnné X příslušným momentem, který jsme určili z výběrového souboru, tj. [12]:

$$\text{est } E(X^k) = m_k = \frac{\sum_{i=1}^n x_i^k}{n}. \quad 5.4$$

Pomocí softwaru odhadneme parametry a určíme základní charakteristiky tohoto rozdělení pro počet pojistných plnění $k = 25649$. Odhadnuté parametry jsou uvedeny v tabulce 9.

Tabulka 9: Odhadnuté parametry negativně binomického rozdělení

Parametry NBi	
r	0,28258
p	0,88401

Zdroj: vlastní zpracování dle STATGRAPHICS Centurion XVII

Dosavadním šetřením jsme zjistili, že počty pojistných plnění N pocházejí s 95 % spolehlivostí z negativně binomického rozdělení pravděpodobnosti s parametry $r = 028258$ a $p = 0,88401$. Pomocí níže uvedených vztahů (5.5, 5.6) dopočítáme střední hodnotu $E(N)$ a rozptyl $D(N)$, tabulka 10.

Negativně binomické rozdělení $NBi(r; p)$ je definováno pravděpodobnostní funkcí:

$$P(x) = \binom{x-1}{r-1} p^r (1-p)^{x-r} \text{ pro } x = k, k+1, \dots$$

Základní charakteristiky negativně binomického rozdělení [10]:

- Střední hodnota

$$E(N) = \frac{r(1-p)}{p}, \quad 5.5$$

- Rozptyl

$$D(N) = \frac{r(1-p)}{p^2}, \quad 5.6$$

- Momentová vytvořující funkce

$$M_X(z) = \left(\frac{p}{1-e^z(1-p)} \right)^r.$$

Tabulka 10: Základní charakteristiky negativně binomického rozdělení

Základní charakteristiky	
$E(N)$	950,99
$D(N)$	1075,78

Zdroj: vlastní výpočet v MS Excel

5.2.2 Vhodné rozdělení pro výšky pojistných plnění X a odhad parametrů

Nalezení vhodného rozdělení pravděpodobnosti výšky pojistných plnění X_i provedeme pomocí χ^2 testu a Kolmogorova – Smirnova testu dobré shody (K-S test) se zvolenou hladinou významnosti $\alpha = 0,05$. Pro vzestupně uspořádaný náhodný výběr (X_1, X_2, \dots, X_n) pocházející ze spojitého rozdělení s distribuční funkcí $F(x)$ a rozsahem n testuje K-S test nulovou hypotézu H_0 : Náhodný výběr pochází ze základního souboru s předpokládaným rozdělením pravděpodobnosti. Test je založen na porovnání rozdílů mezi distribuční funkcí ověřovaného a výběrového rozdělení pravděpodobnosti, tj. $F(x)$ a $F_n(x)$.

Takto uspořádanému náhodnému výběru přísluší výběrová (empirická) distribuční funkce, která je definována vztahem [10]:

$$F_n(x) = \begin{cases} 0 & x \leq x_1 \\ \frac{i}{n} & \\ 1 & \end{cases} \quad 5.7$$

Testovací kritérium má tvar:

$$d_n = \max\{D_1^*, D_2^*, \dots, D_n^*\}, \quad 5.8$$

kde $D_i^* = \max\left\{\left|F(x_i) - \frac{i-1}{n}\right|; \left|\frac{i}{n} - F(x_i)\right|\right\}, i = 1, 2, \dots, n$.

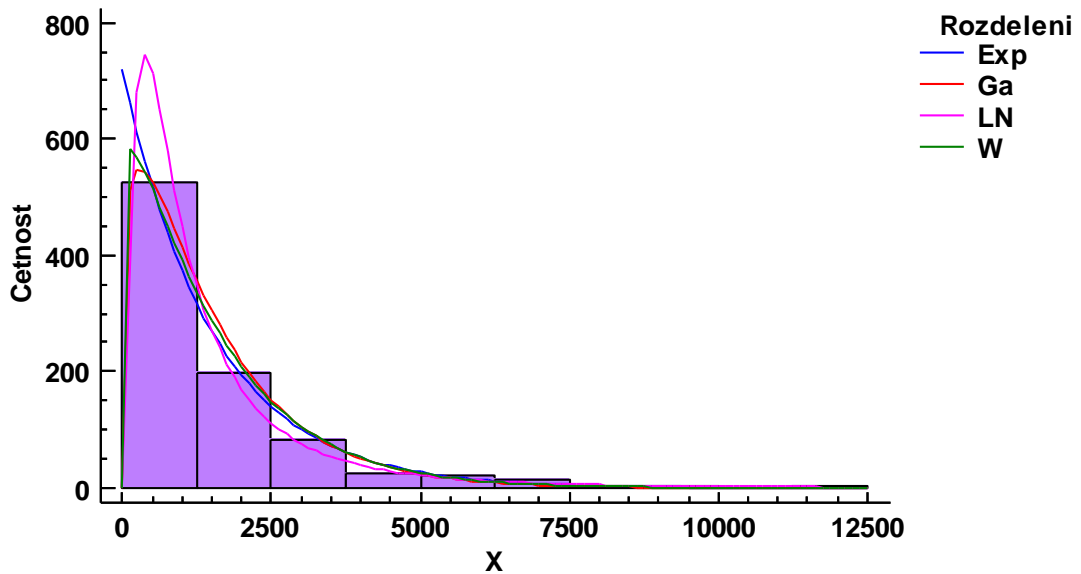
Kritická oblast je definována $d_n \geq d_{n,1-\alpha}$, kde $d_{n,1-\alpha}$ je $1-\alpha$ kvantil rozdělení pravděpodobností testovacího kritéria d_n za předpokladu platnosti hypotézy H_0 . [8]

U vstupních dat výšky pojistných plnění budeme testovat vhodnost exponenciálního, gama, lognormálního a Weibullova rozdělení pravděpodobnosti. Základní charakteristiky těchto rozdělení byly uvedeny v kapitole 1.4.

Všechny testy provedeme na zvolené hladině významnosti $\alpha = 0,05$, graficky znázorněno na obrázku 11. Testujeme postupně nulovou hypotézu:

- H_0 : Výšky individuálních pojistných plnění pocházejí z exponenciálního rozdělení.
- H_0 : Výšky individuálních pojistných plnění pocházejí z gama rozdělení.
- H_0 : Výšky individuálních pojistných plnění pocházejí z lognormálního rozdělení.
- H_0 : Výšky individuálních pojistných plnění pocházejí z Weibullova rozdělení.

Histogram X



Obrázek 11: Histogram X s hustotou testovaných rozdělení

Zdroj: STATGRAPHICS Centurion XVII

Prvně otestujeme pomocí χ^2 testu, zda data pocházejí z exponenciálního rozdělení, tj. $X_i \sim Ex(\lambda)$. Test je znázorněn v tabulce 11. Výšky škod byly rozděleny do deseti intervalů v délce 1000 PJ se získanou četností. Testovací kritérium $\chi_1^2 = 23,8336$, P-hodnota je rovna 0,0024, kritická hodnota pro zvolenou hladinu významnosti, pro deset zvolených intervalů a jeden odhadovaný parametr je $KH_1 = 15,51$. H_0 zamítáme, data nepocházejí z exponenciálního rozdělení, $\chi_1^2 > KH_1$.

Tabulka 11: χ^2 test – exponenciální rozdělení

Goodness-of-Fit Tests for X Chi-Square Test

	Lower Limit	Upper Limit	Observed Frequency	Expected Frequency	Chi-Square
at or below		1000	464	421,79	4,22
	1000	2000	209	218,70	0,43
	2000	3000	89	113,40	5,25
	3000	4000	46	58,80	2,79
	4000	5000	22	30,49	2,36
	5000	6000	18	15,81	0,30
	6000	7000	12	8,20	1,77
	7000	8000	9	4,25	5,31
	8000	9000	3	2,20	0,29
above	9000		4	2,37	1,12

Chi-Square = 23,8336 with 8 d.f. P-Value = **0,00244367**

Zdroj: výstup z STATGRAPHICS Centurion XVII

χ^2 test nulové hypotézy, zda data pocházejí z gama rozdělení, zda $X_i \sim Ga(\alpha; \beta)$ je uvedeno v tabulce 11. Tentokrát byla data rozdělena do devíti intervalů v délce 1000 PJ se získanou četností. Testovací kritérium $\chi_2^2 = 64,6664$, P-hodnota je rovna nule, kritická hodnota pro zvolenou hladinu významnosti, pro devět zvolených intervalů a dva odhadované parametry je $KH_2 = 12,59$. H_0 zamítáme, data nepocházejí z gama rozdělení, $\chi_2^2 > KH_2$.

Tabulka 12: χ^2 test – gama rozdělení

Goodness-of-Fit Tests for X

Chi-Square Test

	Lower Limit	Upper Limit	Observed Frequency	Expected Frequency	Chi-Square
at or below		1000	464	393,54	12,61
	1000	2000	209	246,48	5,70
	2000	3000	89	124,18	9,96
	3000	4000	46	59,69	3,14
	4000	5000	22	28,06	1,31
	5000	6000	18	13,03	1,90
	6000	7000	12	5,99	6,02
	7000	8000	9	2,74	14,28
above	8000		7	2,28	9,74

Chi-Square = 64,6664 with 6 d.f. P-Value = **0,0**

Zdroj: výstup z STATGRAPHICS Centurion XVII

Další testovací nulovou hypotézou bude, zda data pocházejí z lognormálního rozdělení, zda $X_i \sim LN(\mu; \sigma^2)$. Výsledky jsou uvedeny v tabulce 13. Pro tento test byly data rozdělena do jedenácti intervalů v délce 1000 PJ se získanou četností. Testovací kritérium $\chi_3^2 = 12,0857$, P-hodnota je rovna 0,147422, kritická hodnota pro zvolenou hladinu významnosti, pro jedenáct zvolených intervalů a dva odhadované parametry je $KH_3 = 15,51$. H_0 nezamítáme, data s 95 % pravděpodobností pocházejí z lognormálního rozdělení, $\chi_3^2 < KH_3$.

Tabulka 13: χ^2 test – lognormální rozdělení

Goodness-of-Fit Tests for X
Chi-Square Test

	Lower Limit	Upper Limit	Observed Frequency	Expected Frequency	Chi-Square
at or below		1000	464	449,68	0,46
	1000	2000	209	227,21	1,46
	2000	3000	89	92,35	0,12
	3000	4000	46	43,82	0,11
	4000	5000	22	23,25	0,07
	5000	6000	18	13,37	1,61
	6000	7000	12	8,17	1,80
	7000	8000	9	5,23	2,72
	8000	9000	3	3,48	0,07
	9000	10000	2	2,39	0,06
above	10000		2	7,05	3,62

Chi-Square = 12,0857 with 8 d.f. P-Value = **0,147422**

Zdroj: výstup z STATGRAPHICS Centurion XVII

Posledním testem je χ^2 test nulové hypotézy, zda data pocházejí z Weibullova rozdělení, zda $X_i \sim W(\gamma; c)$, tabulka 14. Data byla rozdělena do devíti intervalů v délce 1000 PJ se získanou četností. Testovací kritérium $\chi_4^2 = 46,9112$, P-hodnota je rovna $1,94 \cdot 10^{-8}$, kritická hodnota pro zvolenou hladinu významnosti, pro devět zvolených intervalů a dva odhadované parametry je $KH_4 = 12,59$. H_0 zamítáme, data nepocházejí z Weibullova rozdělení, $\chi_4^2 > KH_4$.

Tabulka 14: χ^2 test – Weibullovo rozdělení

Goodness-of-Fit Tests for X
Chi-Square Test

	Lower Limit	Upper Limit	Observed Frequency	Expected Frequency	Chi-Square
at or below		1000	464	403,19	9,17
	1000	2000	209	233,27	2,53
	2000	3000	89	121,31	8,61
	3000	4000	46	60,89	3,64
	4000	5000	22	29,90	2,09
	5000	6000	18	14,45	0,87
	6000	7000	12	6,90	3,78
	7000	8000	9	3,26	10,12
above	8000		7	2,84	6,11

Chi-Square = 46,9112 with 6 d.f. P-Value = **1,94897E-8**

Zdroj: výstup z STATGRAPHICS Centurion XVII

V tabulce 15 jsou uvedeny souhrnné výsledky K–S testů daných rozdělení. S 95 % spolehlivostí zamítáme nulovou hypotézu, že data pocházejí z exponenciálního, gama a Weibullova rozdělení, stejně jako při χ^2 testu dobré shody, protože p-value je menší než námi zvolená hladina

významnosti $\alpha = 0,05$. Přijímáme pouze nulovou hypotézu H_0 : Vstupní data pocházejí z lognormálního rozdělení s danou hladinou významnosti α . P-hodnota pro lognormální rozdělení je tedy vyšší než zvolená hladina významnosti α .

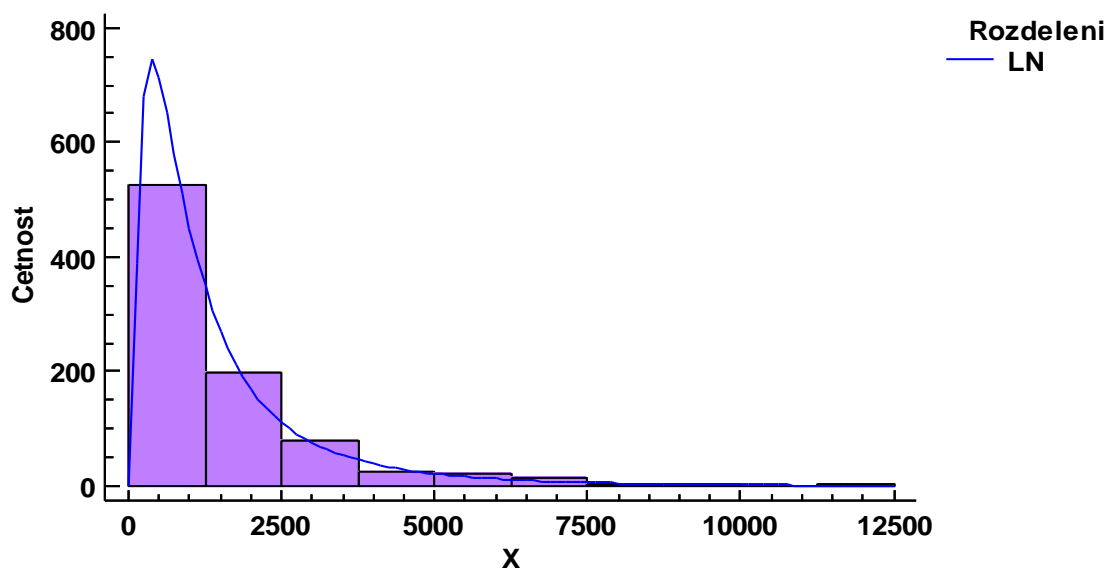
Tabulka 15: K-S test

K-S test	Exponenciální	Gamma	Lognormální	Weibullovo
DPLUS	0,0505154	0,0838946	0,028327	0,0722254
DMINUS	0,0896348	0,0486648	0,0218468	0,0675707
DN	0,0896348	0,0838946	0,028327	0,0722254
P-Value	0,0000015	0,0000088	0,493934	0,0002147

Zdroj: vlastní zpracování dle STATGRAPHICS Centurion XVII

Do histogramu výšky pojistného plnění X_i proložíme hustotu pravděpodobnosti lognormálního rozdělení. Vidíme na obrázku 12, že data odpovídají námi navrhovanému rozdělení a následně s tímto rozdělením budeme dále pracovat.

Histogram X



Obrázek 12: Histogram výšky pojistného plnění s lognormálním rozdělením

Zdroj: STATGRAPHICS Centurion XVII

Pro následující výpočty je nutné odhadnout parametry μ a σ^2 lognormálního rozdělení a určit jeho základní charakteristiky. Odhad parametrů opět provedeme pomocí softwaru, a to metodou maximální věrohodnosti. Výsledky odhadů parametrů jsou uvedeny v tabulce 16. Následně podle definic (5.9 a 5.10) určíme základní charakteristiky střední hodnotu $E(X)$, rozptyl $D(X)$ a koeficient γ_1 výšky pojistných plnění X . Výsledky jsou uvedeny v tabulce 17.

Tabulka 16: Odhadnuté parametry lognormálního rozdělení

Parametry LN	
μ	6,8753
σ^2	0,9415
σ	0,9703

Zdroj: vlastní zpracování dle STATGRAPHICS Centurion XVII

Lognormální rozdělení $LN(\mu; \sigma^2)$, je dáno hustotou pravděpodobnosti:

$$f(x) = \frac{1}{\sigma x \sqrt{2\pi}} e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}} \text{ pro } x > 0, \text{ jestliže } \ln X \sim N(\mu; \sigma^2).$$

Mějme náhodnou veličinu $Y = \ln X$, kde právě $X \sim LN(\mu, \sigma^2)$, která má normální rozdělení $Y \sim N(\mu, \sigma^2)$. A naopak platí, pokud $Y \sim N(\mu, \sigma^2)$, má náhodná veličina $X = e^Y$ lognormální rozdělení $X \sim LN(\mu, \sigma^2)$. [5]

Základní charakteristiky rozdělení $LN(\mu, \sigma^2)$ [10]:

- Střední hodnota:

$$E(X) = e^{\mu + \frac{\sigma^2}{2}}, \quad 5.9$$

- Rozptyl:

$$D(X) = e^{2\mu + \sigma^2} (e^{\sigma^2} - 1), \quad 5.10$$

- Koeficient šikmosti:

$$\gamma_1 = (e^{\sigma^2} + 2) \sqrt{e^{\sigma^2} - 1}, \quad 5.11$$

- Momentová vytvořující funkce:

$$M_X(z) = e^{\mu z + \frac{\sigma^2 z^2}{2}},$$

- k -tý počáteční moment:

$$E(X^k) = m_k = e^{\mu k + \frac{\sigma^2 k^2}{2}}. \quad 5.12$$

Tabulka 17: Základní charakteristiky lognormálního rozdělení

Základní charakteristiky	
$E(X)$	1550,08
$D(X)$	3757340,24
γ_1	5,71

Zdroj: vlastní výpočet v MS Excel

Dosavadním šetřením jsme zjistili, že s 95 % pravděpodobností mají výšky pojistných plnění X_i lognormální rozdělení s parametry $\mu = 6,87532$ a $\sigma = 0,9703$.

5.2.3 Hledání vhodného rozdělení a odhad parametrů v MS Excel

V této části provedeme odhad parametrů a hledání vhodného rozdělení vstupních dat bez použití softwaru STATGRAPHICS Centurion XVII. Výpočty budou provedeny v tabulkovém procesoru MS Excel. Pro porovnání provedeme odhad parametrů daných rozdělení pomocí metody momentů MM. Ověření vhodnosti zvoleného rozdělení realizujeme u analýzy počtu škod N Pearsonovým χ^2 testem dobré shody a u výšky pojistných plnění pomocí Kolmogorova-Smirnova testu dobré shody.

Prvním krokem bude odhadnout parametr λ Poissonova rozdělení a parametry π, α negativně binomického rozdělení pro počet škod N a následně ověřit vhodnost takto zvoleného rozdělení.

Pro odhad parametru λ Poissonova rozdělení využijeme rovnici 5.4 a vztah $E(X) = \lambda$. Z čehož plyne, že odhad je roven aritmetickému průměru. Výsledek je uveden v tabulce 18.

$$m_1 = \frac{\sum_{i=1}^n x_i}{n} = \tilde{\lambda}$$

Tabulka 18: Odhad parametru Poissonova rozdělení MM

Parametry Po odhadnuté MM	
$\tilde{\lambda}$	0,03708

Zdroj: vlastní výpočet v MS Excel

Ověření hypotézy, zda data pocházejí z Poissonova rozdělení pro počty škod N , provedeme Pearsonovým χ^2 testem dobré shody s hladinou významnosti $\alpha = 0,05$, jako při šetření v kapitole 5.2.1 pomocí softwaru.

Vytvoříme si výpočtovou tabulku, Tabulka 19, kde data rozdělíme do tříd jako v tabulce 2. Dopočítáme jednotlivé pravděpodobnosti p_i pomocí pravděpodobnostní funkce $P(x) = \frac{\lambda^x}{x!} e^{-\lambda}$ nebo využijeme excelovskou funkci POISSON.DIST($x; \tilde{\lambda}; 1$), která vrátí kumulativní Poissonovo rozdělení pravděpodobnosti. Následně dopočítáme očekávané početnosti $E_i = n \cdot p_i$. Poslední sloupec tabulky obsahuje výpočet testovacího kritéria podle vztahu 5.1.

Tabulka 19: Výpočtová tabulka χ^2 testu – Poissonovo rozdělení

x_i	O_i	p_i	E_i	$\frac{(O_i - E_i)^2}{E_i}$
0	24773	0,963601	24715,41	0,134171
1	806	0,035728	916,39	13,29665
2	65	0,000662	16,99	135,6845
3	5	0,000008	0,21	109,2774
Součet	25649	0,999999	25649	$\chi^2 = 258,3928$

Zdroj: vlastní výpočet v MS Excel

Hodnota $E_3 = 0,21$, což nespĺňuje podmínku, že jednotlivé očekávané četnosti jsou větší než hodnota 2. Předpoklad souvisí s asymptotickými vlastnostmi χ^2 rozdělení. [11] Z tohoto důvodu sloučíme třídy dvě a tři a test provedeme znovu. Výstup ze systému v kapitole 5.2.1 značí stejné seskupení, tabulka 6. Po přepočítání podle nových tříd, jsme vypočetli, že testovací kritérium je $\chi^2 = 178,8478$. Kritická hodnota pro zvolenou hladinu významnosti α , tři třídy a jeden odhadovaný parametr je $KH = 3,8415$. Hypotézu, že data o počtu pojistných událostí pocházejí z Poissonova rozdělení s odhadovaným parametrem $\tilde{\lambda}$ zamítáme, testovací kritérium překračuje kritickou hodnotu, $\chi^2 > KH$. Stejný výsledek jsme získali i při šetření dané hypotézy pomocí systému.

Tabulka 20: Výpočtová tabulka χ^2 testu – Poissonovo rozdělení

x_i	O_i	p_i	E_i	$\frac{(O_i - E_i)^2}{E_i}$
0	24773	0,963601	24715,41	0,134171
1	806	0,035728	916,39	13,2967
2	70	0,000662	16,99	165,4170
Součet	25649	0,999999	25649	$\chi^2 = 178,8478$

Zdroj: vlastní výpočet v MS Excel

Pro odhad parametrů p, r negativně binomického rozdělení využijeme vztahy [18]:

$$p = \frac{\mu_1}{\mu_2} \quad 5.13$$

$$r = \frac{(\mu_1)^2}{\mu_2 - \mu_1} \quad 5.14$$

Úpravou rovnice 5.13 získáme vztah pro výpočet odhadu parametru p z datového souboru:

$$\tilde{p} = \frac{\frac{1}{N} \sum_{i=1}^N X_i}{\frac{1}{N} \sum_{j=1}^N \left(X_j - \frac{1}{N} \sum_{i=1}^N X_i \right)^2} \quad 5.15$$

Úpravou rovnice 5.14 získáme vztah pro výpočet odhadu parametru r z datového souboru:

$$\tilde{r} = \frac{\left(\frac{1}{N} \sum_{i=1}^N X_i \right)^2}{\frac{1}{N} \sum_{j=1}^N \left(X_j - \frac{1}{N} \sum_{i=1}^N X_i \right)^2 - \frac{1}{N} \sum_{i=1}^N X_i} \quad 5.16$$

Z dat o počtu pojistných plnění určíme pomocí excelovských funkcí výběrový průměr (PRŮMĚR = vrátí průměrnou hodnotu výběru) a výběrový rozptyl (VARA = odhadne rozptyl výběru). Následně dopočítáme odhady podle uvedených rovnic 5.17 a 5.18. Výsledky odhadů pomocí výpočtu v MS Excel jsou uvedeny v tabulce 21.

Tabulka 21: Odhadnuté parametry negativně binomického rozdělení MM

Parametry NBi odhadnuté MM	
\tilde{r}	0,28258
\tilde{p}	0,88401

Zdroj: vlastní výpočet v MS Excel

Vidíme, že odhadnuté parametry pomocí metody momentů jsou shodné s odhady softwaru. Z tohoto důvodu nebudeme v ověřování pokračovat a využijeme pouze výstupy z kapitoly 5.2.1.

Další částí bude výpočet odhadu parametrů μ a σ^2 lognormálního rozdělení pravděpodobnosti výšky pojistných plnění X_i . Využijeme k tomu vztahy 5.4 (definice metody momentů) a 5.12 (výpočet k -tého počátečního momentu lognormálního rozdělení). Určíme tedy bodový odhad prvních dvou momentů pomocí výběrových údajů x_1, x_2, \dots, x_n ve tvaru [11]:

$$m_1 = \frac{\sum_{i=1}^n x_i}{n} \quad a \quad m_2 = \frac{\sum_{i=1}^n x_i^2}{n} \quad 5.17$$

a vytvoříme systém rovnic:

$$m_1 = e^{\mu + \frac{\sigma^2}{2}} \quad a \quad m_2 = e^{2(\mu + \sigma^2)}.$$

Řešením těchto rovnic získáme momentové odhady $\tilde{\mu}$ a $\tilde{\sigma}^2$ parametrů μ, σ^2 ve tvaru:

$$\tilde{\mu} = 2 \ln(m_1) - \frac{\ln(m_2)}{2} \quad 5.18$$

$$\tilde{\sigma}^2 = \ln(m_2) - 2\ln(m_1) \quad 5.19$$

Pomocí datového souboru určíme hodnoty m_1, m_2 a dopočítáme podle 5.13 a 5.14 odhady parametrů lognormálního rozdělení, výsledky znázorňuje tabulka 22.

Tabulka 22: Odhad parametrů lognormálního rozdělení MM

Odhad parametrů lognormálního rozdělení MM	
m_1	1540,8236
m_2	6278866,1775
$\tilde{\mu}$	6,8436
$\tilde{\sigma}^2$	0,9718
$\tilde{\sigma}$	0,9858

Zdroj: vlastní výpočet v MS Excel

Ověření, že jsme u výšky pojistných plnění vhodně zvolili lognormální rozdělení, provedeme pomocí Kolmogorova-Smirnova testu dobré shody s hladinou významnosti $\alpha = 0,05$.

Vytvoříme si tabulku, v které provedeme výpočty pro ověření nulové hypotézy H_0 : Výšky pojistných plnění X pocházejí z lognormálního rozdělení $LN(\mu; \sigma)$. Výpočetní postup je uveden v tabulce 23.

Tabulka 23: Výpočet ověření $X \sim LN(\mu; \sigma)$ K-S test

i	x	$Ln(x_i)$	$F_n(x)$	$F_n(x+1)$	$F(x)$	$ F_n(x)-F(x) $	$ F(x)-F_n(x+1) $
1	24,42	3,19540	0	0,00114	0,00007	0,00007	0,00107
2	31,99	3,46542	0,00114	0,00228	0,00022	0,00092	0,00206
3	33,29	3,50526	0,00228	0,00342	0,00026	0,00203	0,00317
4	52,52	3,96119	0,00342	0,00457	0,00134	0,00209	0,00323
5	58,02	4,06079	0,00457	0,00571	0,00186	0,00270	0,00385
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
873	9200,00	9,12696	0,99543	0,99658	0,98984	0,00559	0,00673
874	9445,55	9,15330	0,99658	0,99772	0,99056	0,00602	0,00716
875	11563,00	9,35557	0,99772	0,99886	0,99471	0,00301	0,00415
876	12254,00	9,41361	0,99886	1	0,99555	0,00331	0,00445
E(X)	1522,53	6,87532			MAX	0,02718	0,02833
D(X)	2629200,07	0,94148					

Zdroj: vlastní výpočet v MS Excel

První sloupec značí pořadové číslo i výšky pojistného plnění x , které je uvedeno v druhém sloupci. Výběrový soubor je uspořádán vzestupně. Ve třetím sloupci je vypočten přirozený logaritmus hodnot x_i . Sloupce čtyři a pět obsahují hodnoty empirické distribuční funkce $F_n(x)$

a $F_n(x + 1)$. V šestém sloupci jsou uvedeny hodnoty předpokládané distribuční funkce vypočítané v hodnotách x_i s odhadnutými parametry $\tilde{\mu}, \tilde{\sigma}$, které jsou uvedeny v tabulce 22. V MS Excel využijeme funkci LOGNORM.DIST($x; \tilde{\mu}; \tilde{\sigma}; 1$), která vrátí kumulativní distribuční funkci lognormálního rozdělení. V posledních dvou sloupcích jsou uvedeny absolutní rozdíly hodnot empirické a hypotetické distribuční funkce v bodech nespojitosti. Maximální absolutní odchylky jsou uvedeny v tabulce 23. Z těchto hodnot vezmeme větší hodnotu, která představuje testovací kritérium $d_{876} = 0,02833$. Z tabulek pro kvantily Kolmogorova-Smirnovova testu určíme podle počtu n a zvolené hladiny významnosti α kritickou hodnotu, v našem případě pro $n = 876$ a $\alpha = 0,05$ je kritická hodnota $d_{876;0,95} = 0,04589$. Porovnáním testovacího kritéria a kritické hodnoty, tj. $d_{979} < d_{979;0,95}$, vidíme, že nulovou hypotézu H_0 přijímáme. Datový soubor s výšky pojistných plnění 95 % pravděpodobností pochází z lognormálního rozdělení pravděpodobnosti s parametry $\tilde{\mu}, \tilde{\sigma}$.

Porovnáme-li odhady parametrů z obou metod, tj MMV a MM, zjistíme, že se liší. Odhady získané metodou maximální věrohodnosti jsou vydatnější než odhady metodou momentů, protože při momentové nebo kvantilové metodě mají odhady z různých výběrových souborů velmi vysokou variabilitu. Jsou tedy málo vydatné, což způsobuje dost pravděpodobný výskyt extrémních hodnot. Proto při dalších výpočtech budeme pracovat s výpočty ze softwaru. [12]

5.3 Základní charakteristika kolektivního modelu rizika S

V předchozí části jsme zjistili, že počet pojistných plnění N pocházejí z negativně binomického rozdělení pravděpodobnosti a že výšky pojistných plnění X_i pocházejí z lognormálního rozdělení pravděpodobnosti.

V této části určíme základní charakteristiky kolektivního modelu rizika S podle kapitoly 1.3.1, tj. určíme střední hodnotu $E(S)$, rozptyl $D(S)$ a koeficient $\gamma(S)$. K výpočtu je nutné nejprve určit počáteční momenty náhodné proměnné X_i . Pro k -tý počáteční moment lognormálního rozdělení, platí:

$$E(X^k) = m_k = e^{\mu k + \frac{\sigma^2 k^2}{2}}.$$

Nejprve určíme pomocí parametrů lognormálního rozdělení tři počáteční momenty $E(X^k)$, uvedeno v tabulce 24.

Tabulka 24: Počáteční momenty $E(X^k)$

k	$E(X^k)$
1	1550,08
2	6160079,85
3	62762316357

Zdroj: vlastní výpočet v MS Excel

Pomocí těchto počátečních momentů a parametrů negativně binomického rozdělení podle definice vypočteme základní charakteristiky kolektivního rizika, tedy $E(S)$, $D(S)$, $\sigma(S)$. Výsledky jsou uvedeny v tabulce 25.

Tabulka 25: Základní charakteristiky S

Základní charakteristiky S	
$E(S)$	1474119,93
$D(S)$	6158037035,87
$\sigma(S)$	78473,16
$\gamma(S)$	0,13116

Zdroj: vlastní výpočet v MS Excel

5.4 Aproximace kolektivního modelu rizika S

Podle kapitoly 3 vypočteme aproximace normálním rozdělením a následně posunutým gama rozdělením pro 99., 98., 95. a 90. percentil a náležící rizikovou přírážku θ . Nejdříve určíme $S_{p(N)}$ pomocí excelovské funkce NORM.INV, která vrátí inverzní funkci k distribuční funkci normálního kumulativního rozdělení, a pomocí $E(S)$ a $D(S)$ z tabulky 25. Výsledky jsou uvedeny v tabulce 26.

$$G(x) = P(Z \leq x) \approx \Phi\left(\frac{x - E(S)}{\sqrt{D(S)}}\right) = p$$

Tabulka 26: Aproximace normálním rozdělením a riziková přírážka

p	$S_{p(N)}$	θ^N
0,99	1656675,80	0,1238
0,98	1635284,10	0,1093
0,95	1603196,80	0,0876
0,90	1574687,34	0,0682

Zdroj: vlastní výpočet v MS Excel

Pro výpočet aproximace kolektivního rizika posunutým gama rozdělením je nutné nejdříve určit parametry α, β a k , které určíme opět pomocí základních charakteristik z tabulky 26, kde $E(S) = \mu, D(S) = \sigma^2$. Parametry jsou uvedeny v tabulce 27.

Tabulka 27: Parametry α, β , a k

α	232,5082
β	0,000194311
k	277543,8266

Zdroj: vlastní výpočet v MS Excel

Pomocí těchto parametrů můžeme vypočítat aproximace posunutým gama rozdělením o k . Pomůžeme si opět excelovskou funkcí GAMMA.INV pro hodnoty α, β . Tato funkce vrátí inverzní funkci k distribuční funkci kumulativního gama rozdělení. Výsledné hodnoty jsou znázorněny v tabulce 28.

Tabulka 28: Aproximace posunutým gama rozdělením a rizikové přírážky

p	Z_p	$S_{p(G)}$	θ^G
0,99	25096929,83	63754151,14	0,2707
0,98	23031924,12	61689145,42	0,2296
0,95	20140687,62	58797908,92	0,1719
0,90	17775177,36	56432398,66	0,1248

Zdroj: vlastní výpočet v MS Excel

V tabulce 29 můžeme vidět porovnání hodnot $S_{p(N)}$ a $S_{p(G)}$. Aproximace posunutým gama rozdělením je lepší, neboť gama rozdělení je pravostranně zešikmené a je tedy vhodnější pro naše data, která jsou také pravostranně zešikmená.

Tabulka 29: Porovnání $S_{p(N)}$ a $S_{p(G)}$

p	$S_{p(N)}$	$S_{p(G)}$
0,99	61067097,67	63754151,14
0,98	59790342,96	61689145,42
0,95	57875225,69	58797908,92
0,90	56173650,55	56432398,66

Zdroj: vlastní výpočet v MS Excel

5.5 Simulace metodou Monte Carlo

Pomocí simulace metodou Monte Carlo, podle postupu uvedeném v kapitole 4.1 ve vhodném programu vygenerujeme náhodné číslo n_1 z námi odhadnutého rozdělení počtu pojistných škod

$N \sim NBi(0,28258; 0,88401)$. Pokud daný software neumožňuje generovat náhodná čísla z různých rozdělení, vygenerujeme náhodné číslo r_1 z rovnoměrného rozdělení v intervalu $(0; 1)$ a náhodné číslo n_1 dostaneme výpočtem dle vztahu $n_1 = Q(r_1)$, kde $Q(r)$ je kvantilová funkce rozdělení N .

Stejným způsobem vygenerujeme z rozdělení výšky škod $X \sim LN(6,8753; 0,9415)$ právě n_1 hodnot individuálních škod $x_1, x_2, x_3, \dots, x_{n_1}$. Podle definice kolektivního rizika, vztah 1.1, určíme součet $s_1 = \sum_{i=1}^{n_1} x_i$, který představuje první vygenerovanou hodnotu celkových škod S .

Takto zopakujeme m -krát tak, aby m bylo dostatečně velké číslo. Získáme tím vygenerované náhodné hodnoty $s_1, s_2, s_3, \dots, s_m$ z neznámého rozdělení celkových škod S , které představují jeden náhodný výběr z tohoto rozdělení.

V předchozích částech jsme zjistili, že v případě, jestliže platí $N \sim NBi(0,28258; 0,88401)$ a $X \sim LN(6,8753; 0,9415)$, celkové škody mají přibližné normální rozdělení pravděpodobnosti $N(E(S) = 1474119,93; D(S) = 6158037035,87)$ nebo posunuté gama rozdělení s parametry $\alpha = 232,5082, \beta = 0,000194311$ a $k = 277543,82$.

Můžeme například vygenerovat $m = 10000$ celkových pojistných škod S , které představují jeden výběrový soubor, a na základě těchto hodnot řešit dvě hlavní úlohy. První úlohou je ověření vhodnosti nalezených aproximačních modelů a druhou úlohou je nalezení teoretického pravděpodobnostního modelu s dobrou shodou s výběrovým souborem, který jsme získaly simulací Monte Carlo.

Z dostupných statistických programových balíčků v žádném nebylo možné realizovat simulaci hodnot z negativního binomického rozdělení s odhadnutými parametry na základě reálných dat, které nejsou celočíselná. Ve všech je tato simulace možná jenom pro celočíselné hodnoty parametrů.

ZÁVĚR

Hlavním cílem této práce bylo teoreticky popsat a následně aplikovat na reálných datech základní, přesné, přibližné a simulační metody pravděpodobnostního modelování kolektivního rizika S . V první kapitole byl čtenář seznámen s definicí a se základními charakteristikami celkového pojistného plnění S , tedy kolektivního rizika za celé portfolio pojistných smluv v daném období. Kapitola se dále věnovala definicím diskrétních a spojitých pravděpodobnostních rozdělení, které je možné využít při modelování počtu a výšky pojistných plnění. Kapitola se také zabývala složenými a smíšenými rozděleními pravděpodobnosti.

V druhé kapitole práce bylo představeno rekurentní vyjádření rozdělení kolektivního rizika. V této části byl popsán Panjerův rekurentní vzorec pro usnadnění výpočtu distribuční funkce $G(x)$. Závěrem byly uvedeny Panjerovy konstanty pro Poissonovo, binomické a negativně binomické rozdělení.

Třetí kapitola se zabývala přibližnými pravděpodobnostními modely kolektivního rizika, za předpokladu znalosti základních charakteristik rozdělení náhodných proměnných N a X_i . Kapitola definuje postup při aproximaci normálním rozdělením a posunutým gama rozdělením a popisuje důležité využití pro výpočet rizikové přírážky θ k čistému netto pojistnému.

V poslední teoretické části se práce věnovala simulačním metodám, kde byl uveden postup simulace pomocí metody Monte Carlo. Tato metoda nám pomáhá nasimulovat náhodné hodnoty s_1, s_2, \dots, s_n , kterými nabývá náhodná proměnná celkové škody S , realizované prostřednictvím simulací hodnot počtu a výšky pojistných plnění.

V kapitole 5 byly na vstupních datech vysvětleny postupy uvedených modelů kolektivního rizika S . Nejprve byly vstupní hodnoty podrobeny grafické analýze, abychom si udělali základní představu o datech a jejich empirickém rozdělení. K dispozici jsme měli údaje o 25 649 pojistných smluv s pojištěním odpovědnosti za škodu způsobenou provozem vozidla. Během sledovaného období 1 rok vzniklo na tomto portfoliu 951 pojistných událostí z 876 smluv. Celkem byla vyplacena pojistná plnění ve výšce 1 333 739,22 PJ, kde minimální výše činila 24,42 PJ a maximální hodnota byla 12 254,00 PJ. V průměru na jednu pojistnou škodu připadá pojistné plnění ve výši 1 522,53 PJ. V další části jsme zjistili pomocí výpočtu v softwaru STATGRAPHICS, že data o počtu pojistných plnění N s pravděpodobnostní 0,95 pocházejí z negativně binomického rozdělení s parametry $r = 0,28258$ a $p = 0,88401$, výšky pojistného plnění pak s pravděpodobnostní 0,95 pocházejí z lognormálního rozdělení s parametry $\mu = 6,8753$

a $\sigma^2 = 0,9415$. Pro porovnání byl uveden postup výpočtů testů dobré shody a odhady parametrů pouze pomocí tabulkového procesoru MS Excel. Na základě předchozích výsledků byly dopočítány základní charakteristiky kolektivního rizika $E(S) = 1474119,93$, $D(S) = 6158037035,87$ a $\gamma(S) = 0,13116$. Průměrné celkové pojistné plnění činí 1 474 119,93 PJ. Další část práce se zabývala výpočtem aproximace kolektivního rizika S normálním rozdělením a posunutým gama rozdělením. Celkové škody mají přibližné normální rozdělení pravděpodobnosti $N(E(S) = 1474119,93; D(S) = 6158037035,87)$ nebo posunuté gama rozdělení s parametry $\alpha = 232,5082$, $\beta = 0,000194311$ a $k = 277543,82$. Závěrem praktické kapitoly je uveden postup simulace pomocí metody Monte Carlo aplikovaný na vypočítané hodnoty.

POUŽITÁ LITERATURA

- [1] BOLAND, P. J., *Statistical and Probabilistic Methods in Actuarial Science*. London: Chapman&Hall/CRC, 2007. ISBN 1-58488-695-1.
- [2] CIPRA, T. *Finanční a pojistné vzorce*. Praha: Grada, 2006. ISBN 80-247-1633.
- [3] DŘÍMAL, J., TRUNEC, D., BRABLEC, A., *Úvod do metody Monte Carlo*. [online] Brno: Masarykova Univerzita, Přírodovědecká fakulta, 2006. [cit. 16.8.2019]. Dostupné z: <http://www.physics.muni.cz/~trunec/mc.pdf>.
- [4] DUCHÁČKOVÁ, E., *Principy pojištění a pojišťovnictví*. Praha: Ekopress, 2009. ISBN 978-80-86929-51-4.
- [5] HINDLS, R. a kol., *Statistika v ekonomii*. Průhonice: Professional Publishing, 2018. ISBN 978-80-88260-09-7.
- [6] HORÁKOVÁ, G., PÁLEŠ, M., SLANINKA, F., *Teória rizika v poistení*. Bratislava: Wolters Kluwer, 2015. ISBN 978-80-8168-273-5.
- [7] KLUGMAN, S. A., PANJER, H. H., WILLMOT, G. E. *Loss Models: from data to decisions*. [online]. John Wiley & Sons, Inc., Hoboken. 2004. [cit. 16.8.2019]. ISBN 0-471-21577-5. Dostupné z: <https://epdf.pub/loss-models-from-data-to-decisions.html>
- [8] KUBANOVÁ, J., *Statistické metody pro ekonomickou a technickou praxi*. Bratislava: STATIS. 2004. ISBN 80-85659-37-9.
- [9] Lean Six Sigma: *Grafická analýza*. [online]. © 2020 Lean Six Sigma. [cit.19.5.2020]. Dostupné z: <https://lean6sigma.cz/graficka-analyza/>
- [10] PACÁKOVÁ, V., *Aplikovaná poistná štatistika*. Bratislava: IURA EDITION, 2004. ISBN 80-8078-004-8.
- [11] PACÁKOVÁ, V. a kol., *Aplikovaná pojistná štatistika*. Pardubice: Univerzita Pardubice, 2019. ISBN 978-80-7560-259-6
- [12] PACÁKOVÁ, V. a kol., *Štatistické metódy pre ekonómov*. Bratislava: IURA EDITION, 2009. ISBN 978-80-8078-284-9.

- [13] PÁLEŠ, M., Aplikácia špecifických rozdelení pravdepodobnosti na analýzu rizika v portfóliu poisťných zmlúv. *Slovenská štatistika a demografia* [online]. Štatistický úrad Slovenskej republiky. 2014, 24(1), str. 3-12. [cit. 09.10.2019]. ISSN 1210-1095. Dostupné z: https://www7.statistics.sk/wps/wcm/connect/30204047-7474-4dfc-ab23-962d985bd051/Slovenska_statistika_a+_demografia_+1_2014.pdf?MOD=AJPERES&CVID=klw1U6u.
- [14] PANJER, Harry H., Recursive Evaluation of a Family of Compound Distributions. *ASTIN Bulletin* [online]. 1981, 12(1), str. 22-26 [cit. 17.02.2020]. DOI: 10.1017/S0515036100006796. ISSN 0515-0361. Dostupné z: https://www.cambridge.org/core/product/identifier/S0515036100006796/type/journal_article
- [15] *Procedure documentation: Probability Distributions*. STATGRAPHICS Centurion XVII, version 17.2.07 (64-bit). Copyright 1982-2016 StatPoint Technologies, Inc. www.STATGRAPHICS.com
- [16] TICHÝ, T., *Simulace Monte Carlo ve financích: aplikace při ocenění jednoduchých opcí*. Ostrava: VŠB-TU Ostrava, 2010. ISBN 978-80-248-2352-2.
- [17] TSE Y.-K., *Nonlife Actuarial Models*. Cambridge: University Press, 2009. ISBN 978-0-521-76465-0.
- [18] WIMMER, G., *Diskrétné jednorozmerné rozdelenia pravdepodobnosti*. Praha: Matfyzpress, 2000, ISBN 80-858-6360-X.

PŘÍLOHY

PŘÍLOHA A – VSTUPNÍ DATA

Individuální celkové škody:

24,42	212,09	290,27	346,46	399,49	463,70	500,00	587,01	659,90	739,84
31,99	216,00	291,36	348,10	400,00	463,71	500,53	588,56	660,57	747,54
33,29	216,14	292,00	348,11	400,00	467,39	503,97	591,60	662,10	748,08
52,52	218,12	299,19	350,41	400,00	469,63	504,78	592,19	664,07	749,49
58,02	218,13	299,88	350,63	400,88	469,65	509,24	594,27	665,11	749,58
74,70	219,50	300,00	352,17	402,64	471,83	512,73	595,15	665,64	751,91
75,89	219,92	302,43	353,67	403,73	472,00	513,23	596,91	668,54	752,13
82,33	225,07	306,81	354,42	404,82	472,75	513,42	599,58	673,57	753,22
85,34	226,10	306,90	355,16	409,09	473,22	519,01	599,90	674,76	756,00
86,22	227,50	307,03	356,65	412,77	474,40	519,28	600,00	675,23	757,83
90,00	232,92	307,07	361,41	412,78	474,43	519,37	602,30	676,49	759,06
102,00	234,45	308,56	361,78	414,72	474,66	519,72	602,84	678,67	760,47
116,51	236,40	312,23	362,30	415,04	480,05	521,78	605,12	679,36	761,12
121,44	239,50	313,34	362,59	416,00	480,50	523,50	610,10	680,00	763,72
123,71	243,09	314,79	364,73	417,42	481,00	525,79	611,62	682,21	774,94
148,00	243,19	314,91	369,91	418,49	481,08	528,40	612,56	683,90	778,09
149,37	243,78	315,68	370,00	419,47	481,24	529,33	612,56	685,29	779,02
150,00	252,00	316,40	374,08	421,06	483,60	529,84	612,64	687,38	779,62
160,00	255,17	318,44	374,22	421,66	484,28	530,00	614,03	690,66	780,89
163,50	257,49	318,61	374,85	422,26	484,90	534,00	616,68	697,39	781,84
165,97	260,60	320,00	375,05	424,61	486,53	534,37	617,30	700,59	783,90
167,36	261,00	324,01	376,00	424,71	487,71	537,49	626,04	702,08	784,17
168,19	265,71	327,06	376,43	425,00	487,80	542,20	626,28	703,20	784,26
174,09	267,28	327,82	376,88	425,87	487,89	547,20	627,85	704,08	787,43
174,67	267,94	327,85	378,17	426,22	487,98	547,91	627,93	705,08	787,48
180,00	270,05	328,50	379,20	427,09	490,13	550,00	630,30	708,71	790,40
186,42	272,99	330,00	379,58	430,00	490,52	550,77	630,95	708,78	797,05
188,50	276,59	331,10	380,00	431,77	490,70	556,39	633,31	710,85	797,52
189,40	278,36	333,09	381,63	432,95	491,15	558,86	634,51	714,19	798,04
191,00	278,46	335,21	382,63	441,85	494,03	560,19	635,09	714,82	801,00
194,83	278,90	335,67	383,73	442,29	495,40	561,42	637,47	719,44	808,95
196,21	281,51	335,84	383,76	442,56	495,50	564,00	639,41	719,81	810,21
199,80	281,56	336,40	389,43	447,47	496,78	564,75	643,05	720,09	813,81
200,00	282,25	337,65	393,73	450,34	498,00	565,54	644,61	721,19	813,88
201,03	282,83	337,99	396,12	450,89	498,00	565,59	647,86	722,15	819,11
203,83	283,57	338,66	396,47	454,03	498,00	566,95	647,98	723,37	820,02
204,48	284,00	339,86	396,48	454,26	498,33	572,28	655,98	734,37	821,76
206,55	285,84	341,43	397,43	454,80	498,59	574,40	656,23	737,00	822,08
206,99	287,00	344,80	398,69	455,20	498,68	576,15	656,30	737,37	824,75
209,76	289,08	345,30	399,41	458,71	499,37	577,00	659,73	737,41	825,62

828,40	926,64	1077,79	1217,27	1436,95	1590,02	1809,35	2040,00	2482,99	2992,66
832,98	929,07	1079,25	1225,87	1443,78	1590,81	1810,30	2047,00	2486,30	3000,00
833,16	934,83	1079,30	1231,30	1445,33	1597,45	1815,00	2066,67	2488,15	3012,77
833,40	937,97	1080,83	1237,00	1447,17	1611,33	1824,40	2072,68	2523,38	3034,02
838,60	938,00	1081,87	1244,72	1451,43	1619,95	1828,37	2082,40	2527,04	3048,10
850,23	941,60	1084,82	1254,29	1457,14	1623,20	1830,00	2112,00	2550,11	3057,14
857,18	943,84	1086,12	1257,13	1459,38	1632,24	1837,07	2118,34	2556,87	3073,81
859,60	948,01	1099,66	1258,54	1460,00	1636,00	1839,84	2146,13	2560,92	3096,90
863,86	960,26	1100,00	1266,67	1462,10	1651,88	1845,86	2159,34	2560,96	3120,84
864,55	960,80	1101,19	1272,68	1462,45	1653,75	1847,58	2192,00	2570,38	3120,88
866,36	961,59	1108,60	1283,78	1462,96	1665,22	1849,00	2201,92	2576,28	3126,90
868,94	962,70	1111,09	1294,05	1470,15	1668,78	1849,28	2220,61	2591,15	3156,38
873,97	965,53	1111,25	1294,40	1470,33	1679,47	1865,70	2235,11	2593,61	3170,08
878,82	967,37	1118,08	1296,00	1478,03	1685,88	1888,96	2242,59	2595,87	3177,58
885,36	967,94	1127,41	1308,66	1483,20	1686,48	1892,55	2251,63	2596,29	3200,00
886,00	968,40	1127,71	1309,63	1489,73	1686,67	1897,17	2263,02	2606,92	3200,07
886,57	970,00	1129,52	1311,00	1490,79	1688,50	1920,24	2267,00	2614,75	3201,85
887,93	976,87	1138,46	1313,47	1494,00	1693,60	1934,53	2274,62	2621,38	3228,64
888,26	981,33	1146,00	1319,56	1494,00	1694,42	1934,71	2296,00	2624,98	3231,36
891,56	983,90	1147,17	1324,24	1496,18	1695,32	1936,93	2300,00	2713,22	3231,72
895,80	992,25	1154,71	1326,91	1499,42	1699,69	1937,00	2300,00	2730,00	3300,00
896,50	994,00	1161,82	1329,10	1500,00	1704,80	1941,94	2300,00	2734,38	3304,39
896,76	996,00	1164,86	1330,98	1504,91	1708,20	1962,90	2300,00	2780,39	3329,74
897,57	997,00	1166,58	1335,36	1520,38	1712,62	1963,06	2300,00	2784,23	3366,98
901,25	1001,00	1167,84	1341,40	1521,31	1715,97	1986,50	2300,00	2793,16	3375,00
903,68	1004,98	1175,47	1345,19	1528,46	1730,93	1992,00	2300,00	2845,15	3391,00
905,45	1013,94	1181,25	1347,78	1528,54	1733,81	1992,00	2300,00	2875,00	3399,59
907,36	1022,71	1181,86	1360,98	1542,03	1740,94	1992,00	2300,00	2875,00	3411,27
907,46	1027,82	1182,97	1363,46	1547,96	1745,67	1992,00	2303,42	2875,00	3439,54
907,93	1029,91	1189,23	1364,09	1554,53	1747,63	1992,00	2324,93	2875,00	3440,57
909,16	1031,08	1193,98	1370,18	1555,45	1749,50	1992,00	2329,52	2875,00	3450,00
912,94	1040,54	1194,17	1376,86	1561,22	1752,45	1992,77	2329,88	2875,00	3450,00
914,16	1045,00	1196,00	1387,24	1562,62	1756,31	1994,52	2348,44	2875,00	3450,00
915,42	1048,98	1196,00	1388,84	1569,18	1760,24	2013,86	2375,00	2875,16	3450,00
918,53	1049,64	1196,02	1390,00	1569,49	1761,97	2014,22	2414,32	2925,57	3450,00
920,00	1050,00	1197,52	1408,73	1570,96	1762,20	2020,00	2433,60	2932,79	3450,00
920,82	1056,70	1198,93	1409,59	1574,84	1779,99	2020,97	2457,50	2965,37	3453,09
922,30	1065,47	1200,00	1414,93	1578,20	1781,55	2021,47	2472,58	2971,60	3500,85
922,48	1067,61	1209,73	1420,42	1578,37	1790,16	2030,90	2476,91	2978,58	3527,76
926,33	1076,51	1210,62	1430,43	1578,47	1798,16	2033,77	2476,92	2988,00	3579,29

3618,61	4604,83	5540,00	7117,87
3625,00	4616,68	5573,08	7196,00
3650,93	4741,47	5590,00	7196,00
3695,38	4744,55	5600,00	7196,00
3756,35	4782,67	5610,00	7281,00
3814,39	4800,00	5662,63	7460,00
3904,50	4850,00	5781,31	7465,98
3911,76	4900,00	5902,51	7685,90
4014,00	4950,00	6001,34	7960,00
4027,59	5000,00	6074,52	8040,00
4171,27	5020,00	6130,90	8107,80
4223,17	5036,00	6224,49	8351,00
4227,41	5100,00	6494,58	9200,00
4235,32	5120,00	6500,00	9445,55
4287,51	5211,00	6500,00	11563,00
4348,00	5374,54	6500,00	12254,00
4404,66	5365,77	6591,90	
4472,50	5397,10	6600,00	
4543,40	5418,98	6772,01	
4577,49	5421,78	6988,21	