



**MATEMATICKO-FYZIKÁLNÍ
FAKULTA**
Univerzita Karlova

BAKALÁRSKA PRÁCA

Viktória Antalicová

Aproximácia rozdelenia úhrnov škôd

Katedra pravděpodobnosti a matematické statistiky

Vedúci bakalárskej práce: RNDr. Lucie Mazurová, Ph.D.

Študijný program: Finanční matematika

Študijný obor: MFMP

Praha 2023

Prehlasujem, že som túto bakalársku prácu vypracoval(a) samostatne a výhradne s použitím citovaných prameňov, literatúry a ďalších odborných zdrojov. Tato práca nebola využitá k získaniu iného alebo rovnakého titulu.

Beriem na vedomie, že sa na moju prácu vzťahujú práva a povinnosti vyplývajúce zo zákona č. 121/2000 Sb., autorského zákona v platnom znení, najmä skutočnosť, že Univerzita Karlova má právo na uzavrenie licenčnej zmluvy o užití tejto práce ako školského diela podľa §60 odst. 1 autorského zákona.

V dne

Podpis autora

Touto formou by som sa rada podakovala vedúcej mojej bakalárskej práce RNDr. Lucii Mazurovej, Ph.D. za jej užitočné pripomienky a rady pri písaní tejto práce.

Názov práce: Aproximácia rozdelenia úhrnov škôd

Autor: Viktória Antalicová

Katedra: Katedra pravdepodobnosti a matematické statistiky

Vedúci bakalárskej práce: RNDr. Lucie Mazurová, Ph.D., Katedra pravdepodobnosti a matematické statistiky

Abstrakt: Táto práca sa zaoberá aproximáciou rozdelenia úhrnov škôd. Najprv predstavíme postup modelovania úhrnov škôd, ktorý zahŕňa výber vhodného rozdelenia počtov a výšok škôd. Ako ďalšie je vysvetlený výpočet úhrnov škôd ako súčtu príslušného počtu jednotlivých škôd. V druhej kapitole rozoberieme aproximáciu rozdelenia simulovaných úhrnov škôd. Uvedieme rozdelenia zvolené pre aproximáciu, metódu odhadu parametrov týchto rozdelení a následné testovanie zhody týchto rozdelení so skutočným rozdelením simulovaných úhrnov škôd. V tretej kapitole zobrazíme výsledky tejto aproximácie a uvedieme vhodnosť použitia jednotlivých uvažovaných rozdelení na modelovanie úhrnov škôd. V poslednej časti predstavíme Edgeworthovu aproximáciu ako metódu pre aproximáciu rozdelenia úhrnov škôd.

Kľúčové slová: úhrn škôd, rozdelenie úhrnov škôd, Edgeworthova aproximácia

Title: Approximations of the Aggregate Loss Distribution

Author: Viktória Antalicová

Department: Department of Probability and Mathematical Statistics

Supervisor: RNDr. Lucie Mazurová, Ph.D., Department of Probability and Mathematical Statistics

Abstract: This thesis is focused on the approximation of the distribution of aggregate losses. We first present a method for modelling aggregate losses, which involves selecting an appropriate frequency and severity distributions. Next, the computation of aggregate losses as the sum of the respective number of individual losses is explained. In the second section, we discuss the approximation of the distribution of the simulated aggregate losses. We present the distributions chosen for the approximation, the method for estimating the parameters of these distributions, and the subsequent testing of fit of these distributions with the actual distribution of the simulated aggregate losses. In the third chapter we show the results of this approximation and indicate the suitability of using each of the considered distributions for modelling aggregate losses. In the last section, we introduce the Edgeworth approximation as a method for approximating the distribution of aggregate losses.

Keywords: aggregate loss, distribution of aggregate losses, Edgeworth approximation

Obsah

Úvod	2
1 Modelovanie úhrnov škôd	3
1.1 Volba rozdelenia počtov škôd	3
1.2 Volba rozdelenia výšok škôd	4
1.3 Konštrukcia úhrnov škôd	7
2 Aproximácia rozdelenia úhrnov škôd	9
2.1 Rozdelenia zvolené pre aproximáciu rozdelenia agregátnych škôd .	9
2.2 Odhad parametrov	11
2.3 Testovanie dobrej zhody	13
3 Výsledky použitia uvažovaných aproximácií na simulované dáta	15
3.1 Výsledky scenáru 1 pre bežné škody	16
3.2 Výsledky scenáru 2 pre bežné škody	17
3.3 Výsledky scenáru 3 pre bežné škody	18
3.4 Výsledky scenáru 1 pre škody podliehajúce zaisteniu	19
3.5 Výsledky scenáru 2 pre škody podliehajúce zaisteniu	20
4 Edgeworthova aproximácia	21
4.1 Odvodenie Edgeworthovej aproximácie	22
4.2 Výsledky použitia Edgeworthovej aproxi- mácie na simulované dáta	24
Záver	28
Zoznam použitej literatúry	29
Zoznam obrázkov	30
Zoznam tabuliek	31

Úvod

Rozdelenia agregátnych škôd majú široké využitie v aktuárskej praxi, a to ako pri stanovovaní sadzieb, tak aj pri tvorbe rezerv. V priebehu posledných 40 rokov bola vyvinutá rada možností pre výpočet rozdelení agregátnych škôd, ako napríklad Panjerova metóda alebo Rýchla Fourierova Transformácia (FFT), avšak, tieto metódy sú založené na predpoklade znalosti separátnych dát ako pre počty, tak pre výšky jednotlivých škôd. V praxi sa však môžeme stretnúť s tým, že pre potrebnú analýzu máme k dispozícii iba úhrny škôd. V takejto situácii sa stáva voľba rozdelenia agregátnych škôd a jeho tvar veľmi dôležitou.

V tejto práci sa preto budeme snažiť nájsť vhodné odporúčania pre voľbu pravdepodobnostného rozdelenia pre agregátne škody. Budeme sa zaoberať aproximáciou vhodného rozdelenia pre úhrny menších, bežných škôd ako aj pre škody vyššie, ktoré podliehajú individuálnemu XL-zaisteniu.

Náš postup pozostáva z niekoľkých krokov. Ako prvé volíme vhodné rozdelenia pre počty a výšky jednotlivých škôd. Počty škôd simulujeme z Poissonovho rozdelenia. Pre výšky škôd volíme dve rôzne rozdelenia, a to gama rozdelenie na ohraničenom intervale pre bežné škody a Paretovo rozdelenie pre škody podliehajúce individuálnemu XL-zaisteniu. Následne po simulácii dostatočného množstva pozorovaní spočítame škodné úhrny ako súčet príslušného počtu jednotlivých škôd. Ďalším krokom bude odhad parametrov pre zvolené teoretické rozdelenia pomocou nasimulovanej vzorky úhrnov škôd a ako posledné budeme analyzovať zhodu odhadnutých rozdelení so simulovanými škodnými úhrnmi.

Rozdelenia, ktorých zhodu s nasimulovanými dátami budeme testovať podľa predchádzajúceho postupu boli zvolené z rozdelení, ktoré sa bežne vyskytujú v poisťno-matematickej praxi. V poslednej časti tejto práce preto porovnáme rozdelenie agregátnych škôd s rozdelením získaným pomocou Edgeworthovej aproximácie. Túto aproximáciu najprv všeobecne odvodíme a následne budeme znovu testovať zhodu s empirickým rozdelením škodných úhrnov.

Hlavným zdrojom práce boli publikácie Papush a kol. (2001) a Papush a kol. (2021). Pre teoretickú časť boli použité najmä Anděl (2011) a Wuthrich (2022). Všetky simulácie a výpočty boli prevedené pomocou softwaru *Wolfram Mathematica 13*.

1. Modelovanie úhrnov škôd

V prvej kapitole sa budeme zaoberať modelovaním úhrnov škôd, ktorých rozdelenie budeme v ďalších kapitolách aproximovať. Definujeme najprv rozdelenia používané pre simulácie príslušných počtov a výšok škôd a následne vysvetlíme, ako sme jednotlivé úhrny zkonštruovali.

Budeme predpokladať, že skutočný, ročný úhrn škôd na určitej skupine poistných zmlúv je náhodná veličina

$$S = \sum_{i=1}^N X_i,$$

kde N je počet škôd za dané obdobie a X_1, \dots, X_N sú príslušné výšky škôd. To znamená, že S má zložené rozdelenie, kde jednotlivé škody sú vzájomne nezávislé a rovnako rozdelené náhodné veličiny, nezávislé na počte škôd N .

Pri definíciách rozdelení použitých v podkapitolách 1.1 a 1.2 sme vychádzali z knihy Anděl (2011).

1.1 Voľba rozdelenia počtov škôd

Definícia 1. *Nech náhodná veličina X nadobúda hodnoty $0, 1, \dots$, a to s pravdepodobnosťami*

$$P[X = x] = \frac{\lambda^x}{x!} e^{-\lambda}, \quad x=0, 1, \dots$$

kde $\lambda > 0$ je dané číslo. Potom povieme, že X má Poissonovo rozdelenie s parametrom λ . Značíme $X \sim Po(\lambda)$.

Strednú hodnotu a rozptyl Poissonovho rozdelenia odvodíme nasledujúcim spôsobom:

$$EX = \sum_{x=0}^{\infty} x e^{-\lambda} \frac{\lambda^x}{x!} = \sum_{x=1}^{\infty} x e^{-\lambda} \frac{\lambda^x}{x!} = \lambda e^{-\lambda} \underbrace{\sum_{x=1}^{\infty} \frac{\lambda^{x-1}}{(x-1)!}}_{= e^{\lambda}} = \lambda e^{-\lambda} e^{\lambda} = \lambda.$$

Rozptyl spočítame ako $\text{Var } X = EX^2 - (EX)^2$ a k výpočtu využijeme vzťah $EX^2 = E[X(X-1)] + EX$, kde

$$E[X(X-1)] = \sum_{x=0}^{\infty} x(x-1) e^{-\lambda} \frac{\lambda^x}{x!} = \lambda^2 e^{-\lambda} \sum_{x=2}^{\infty} \frac{\lambda^{x-2}}{(x-2)!} = \lambda^2 e^{-\lambda} e^{\lambda} = \lambda^2.$$

Pre rozptyl Poissonovho rozdelenia teda dostávame :

$$\text{Var } X = E[X(X-1)] + EX - (EX)^2 = \lambda^2 + \lambda - \lambda^2 = \lambda.$$

V aktuárskej praxi je Poissonovo rozdelenie základným rozdelením používaným na modelovanie frekvencie poistných udalostí, čo je hlavným dôvodom voľby tohto rozdelenia pre simulácie v tejto práci.

Voľby parametrov pre jednotlivé simulácie počtov škôd, ktoré sme následne využili pri konštrukcii úhrnov škôd, sú uvedené v tabuľkách 1.1 a 1.2.

1.2 Voľba rozdelenia výšok škôd

Definícia 2. Pre $\alpha > 0$ a $\beta > 0$ definujeme gama rozdelenie, $\Gamma[\alpha, \beta]$, ako spojité rozdelenie s hustotou

$$f(x) = \frac{1}{\Gamma(\alpha)} e^{-\frac{x}{\beta}} x^{-1+\alpha} \beta^{-\alpha},$$

$$\text{kde } x > 0 \text{ a } \Gamma(\alpha) = \int_0^{\infty} t^{\alpha-1} e^{-t} dt.$$

Strednú hodnotu a rozptyl gama rozdelenia taktiež odvodíme.

$$\begin{aligned} EX &= \int_0^{\infty} x \frac{1}{\Gamma(\alpha) \beta^{\alpha}} x^{\alpha-1} e^{-\frac{x}{\beta}} dx = \frac{1}{\Gamma(\alpha) \beta^{\alpha}} \int_0^{\infty} x^{\alpha} e^{-\frac{x}{\beta}} dx = \\ &= \frac{1}{\Gamma(\alpha) \beta^{\alpha}} \int_0^{\infty} (t\beta)^{\alpha} e^{-t} \beta dt = \frac{1}{\Gamma(\alpha) \beta^{\alpha}} \beta^{\alpha} \beta \int_0^{\infty} t^{\alpha} e^{-t} dt = \\ &= \frac{\beta}{\Gamma(\alpha)} \int_0^{\infty} t^{\alpha} e^{-t} dt = \frac{\beta}{\Gamma(\alpha)} \Gamma(\alpha + 1) = \alpha\beta \end{aligned}$$

Pri odvodení strednej hodnoty sme v tretej rovnosti použili substitúciu $t = x/\beta$. Po úprave sme využili vzťah $\Gamma(\alpha + 1) = \alpha \Gamma(\alpha)$ pre gama funkciu.

Pri odvodzovaní rozptylu opäť vyjdeme zo vzťahu $\text{Var } X = EX^2 - (EX)^2$. Spočítame najprv EX^2 , kde znovu použijeme substitúciu $t = x/\beta$.

$$\begin{aligned} EX^2 &= \int_0^{\infty} x^2 \frac{1}{\Gamma(\alpha) \beta^{\alpha}} x^{\alpha-1} e^{-\frac{x}{\beta}} dx = \frac{1}{\Gamma(\alpha) \beta^{\alpha}} \int_0^{\infty} x^{\alpha+1} e^{-\frac{x}{\beta}} dx = \\ &= \frac{1}{\Gamma(\alpha) \beta^{\alpha}} \int_0^{\infty} (t\beta)^{\alpha+1} e^{-t} \beta dt = \frac{\beta^2}{\Gamma(\alpha)} \int_0^{\infty} t^{\alpha+2-1} e^{-t} dt = \frac{\beta^2 \Gamma(\alpha + 2)}{\Gamma(\alpha)} = \\ &= \alpha(\alpha + 1) \beta^2 \end{aligned}$$

K finálnej úprave sme opäť použili vzťah pre gama funkciu. Konkrétne :

$$\Gamma(\alpha + 2) = (\alpha + 1) \Gamma(\alpha + 1) = (\alpha + 1) \alpha \Gamma(\alpha).$$

Výsledný rozptyl má teda tvar :

$$\text{Var } X = \text{EX}^2 - (\text{EX})^2 = \alpha(\alpha + 1)\beta^2 - \alpha^2\beta^2 = \alpha\beta^2(\alpha + 1 - \alpha) = \alpha\beta^2.$$

Definícia 3. *Nech $\theta > 0$, $\alpha > 0$. Paretovo rozdelenie definujeme ako spojité rozdelenie s distribučnou funkciou*

$$F(x) = 1 - \left(\frac{\theta}{x}\right)^\alpha, \quad x \geq \theta.$$

Strednú hodnotu a rozptyl Paretovho rozdelenia odvodíme tak, že najprv vyjadříme EX^k a následne EX a $\text{Var } X$ dopočítame správnou voľbou za k .

Hustota Paretovho rozdelenia má tvar $f(x) = \alpha \theta^\alpha x^{-\alpha-1}$ pre $x \geq \theta$ a teda EX^k dostaneme ako:

$$\begin{aligned} \text{EX}^k &= \int_{\theta}^{\infty} x^k \alpha \theta^\alpha x^{-\alpha-1} dx = \alpha \theta^\alpha \int_{\theta}^{\infty} x^{k-\alpha-1} dx = \\ &= \alpha \theta^\alpha \left[\frac{x^{k-\alpha}}{k-\alpha} \right]_{\theta}^{\infty} = \frac{\alpha \theta^k}{\alpha - k}, \quad k < \alpha \end{aligned} \tag{1.1}$$

Z čoho vidíme, že Paretovo rozdelenie má konečný k -ty moment len pre $k < \alpha$.

Voľbou $k = 1$ dostávame strednú hodnotu, ktorá má tvar:

$$\text{EX} = \frac{\alpha \theta}{\alpha - 1}, \quad \alpha > 1.$$

Rozptyl Paretovho rozdelenia dopočítame ako :

$$\begin{aligned} \text{Var } X &= \text{EX}^2 - (\text{EX})^2 = \frac{\alpha \theta^2}{\alpha - 2} - \frac{\alpha^2 \theta^2}{(\alpha - 1)^2} = \\ &= \frac{\alpha \theta^2}{(\alpha - 1)^2(\alpha - 2)}, \quad \alpha > 2. \end{aligned}$$

Je dôležité poznamenať, že pre $\alpha \leq 1$, resp. $\alpha \leq 2$, je EX , resp. $\text{Var } X$ rovné nekonečnu, čo vyplýva z 1.1.

Cieľom pri výbere rozdelení pre simuláciu výšok jednotlivých škôd bolo vybrať jedno rozdelenie, ktoré je vhodné použiť na modelovanie výšok menších, bežných škôd a jedno rozdelenie pre modelovanie škôd, ktoré podliehajú XL-zaisteniu a teda chceme, aby boli zahrnuté aj škody vysokých hodnôt.

Poznámka. Dôležitou vlastnosťou rozdelení výšok škôd je chovanie "chvostov", t.j. limitné chovanie $P[X > x] = 1 - F(x)$ pre $x \rightarrow \infty$.

Rozdeleniam, pre ktoré $P[X > x]$ klesá pri $x \rightarrow \infty$ k nule rýchlo, hovoríme rozdelenia s *lahkým chvostom*. Takýmito rozdeleniami sú napríklad exponenciálne alebo gama rozdelenie. V opačnom prípade hovoríme o rozdeleniach s *ťažkým chvostom*.

Pre prístup uvažujúci zaistenie, sme zvolili Paretovo rozdelenie, keďže má ťažké chvosty a teda poskytuje vhodný popis pre rozdelenie škodných nárokov, kde predpokladáme, že môžu nastať značne vysoké škody.

Voľby parametrov pre všetky simulácie a jednotlivé rozdelenia v nich sú zhrnuté v tabuľkách 1.1 a 1.2.

Pre simulácie menších výšok škôd sme zvolili gama rozdelenie, ktoré označujeme ako rozdelenie s lahkým chvostom, čo zabezpečí, že pravdepodobnosť výskytu veľmi vysokých škôd bude malá.

Aby sme čo najviac predišli výskytu vysokých škodných nárokov, uvažovali sme navyše gama rozdelenie na ohraničenom intervale $(0, m]$, $m < \infty$. Hustotu tohto "useknutého" gama rozdelenia sme odvodili pomocou nasledujúceho tvrdenia.

Tvrdenie 1. *Nech X je náhodná veličina so spojitým rozdelením definovaným na intervale $(-\infty, \infty)$. Ďalej nech $f(x)$ je hustota a $F(x)$ distribučná funkcia veličiny X . Hustota podmieneného rozdelenia náhodnej veličiny X za podmienky $X \leq m$, kde $m < \infty$, má za predpokladu, že $F(m) \neq 0$ tvar*

$$f_m(x) = \begin{cases} \frac{f(x)}{F(m)}, & x \leq m, \\ 0, & x > m. \end{cases}$$

Dôkaz. Distribučná funkcia useknutého rozdelenia má tvar $P[X \leq x | X \leq m]$, z čoho dostávame :

$$P[X \leq x | X \leq m] = \frac{P[X \leq x, X \leq m]}{P[X \leq m]} = \frac{P[X \leq \min(x, m)]}{P[X \leq m]}$$

Pričom prvá rovnosť plynie z definície podmienenej pravdepodobnosti.

Následne ak $x > m$ máme:

$$\frac{P[X \leq \min(x, m)]}{P[X \leq m]} = \frac{P[X \leq m]}{P[X \leq m]} = 1,$$

ak $x \leq m$:

$$\frac{P[X \leq \min(x, m)]}{P[X \leq m]} = \frac{P[X \leq x]}{P[X \leq m]} = \frac{F(x)}{F(m)}.$$

Po derivácii podľa premennej x dostávame hustotu useknutého rozdelenia

$$f_m(x) = \begin{cases} \frac{f(x)}{F(m)}, & x \leq m, \\ 0, & x > m. \end{cases}$$

□

1.3 Konštrukcia úhrnov škôd

Ako už bolo spomenuté, uvažovali sme rôzne scenáre pre generovanie úhrnov tvorených bežnými škodami s menšou výškou ako aj so škodami podliehajúcimi zaisteniu, pre ktoré očakávame vysoké škodné nároky. V tejto podkapitole uvedieme najprv typ zaistenia, ktoré sme v scenároch uvažovali, vyjdeme pri tom z knihy Mandl a Mazurová (1999), následne predstavíme jednotlivé scenáre a na záver vysvetlíme postup zostavenia jednotlivých škodných úhrnov.

Pre simuláciu škôd podliehajúcich zaisteniu sme zvolili základnu formu neproporcionálneho zaistenia a to zaistenie škodného nadbytku, nazývané tiež XL-zaistenie (z anglického excess of loss), konkrétne **individuálne XL-zaistenie** (zaistenie na riziko).

Pre každé riziko zahrnuté do tohto zaistenia sa zaistovňa zaväzuje hradiť tú časť škody, ktorá presiahne stanovenú hranicu (**prioritu**) r . Plnenie zaistovne je ohraničené hornou hranicou (**veľkosťou vrstvy**) a . Ak teda nastane škoda X , zaistovňa hradí

$$\min(\max(X - r, 0), a).$$

V jednotlivých scenároch budeme používať značenie a a r , ktoré predstavuje XL-zaistenie, v ktorom zaistiteľ hradí z každej škody časť spadajúcu do vrstvy veľkosti a nad prioritou r .

V nasledujúcich tabuľkách sú zhrnuté všetky scenáre, ktoré sme testovali.

Scenár	Rozdelenie počtu škôd	Rozdelenie výšok škôd
#1	Poisson, $\lambda = 100$	Useknuté gama, $m = 30, \alpha = 5, \beta = 3$
#2	Poisson, $\lambda = 500$	Useknuté gama, $m = 30, \alpha = 5, \beta = 2$
#3	Poisson, $\lambda = 1000$	Useknuté gama, $m = 30, \alpha = 3, \beta = 2$

Pozn: Počet prevedených simulácií v týchto scenároch bol 5000

Tabuľka 1.1: Rozdelenia použité pre simuláciu bežných škodných úhrnov

Scenár	Rozdelenie počtu škôd	Rozdelenie výšok škôd	Uvažovaná zaistná vrstva
#1	Poisson, $\lambda = 50$	Pareto, $\theta = 50, \alpha = \frac{7}{6}$	750 x 250
#2	Poisson, $\lambda = 10$	Pareto, $\theta = 300, \alpha = \frac{22}{19}$	4000 x 1000

Pozn 1: Uvažované zaistné vrstvy sú vyjadrené v tisícoch

Pozn 2: Počet prevedených simulácií v týchto scenároch bol 50000

Tabuľka 1.2: Rozdelenia použité pre simuláciu škodných úhrnov uvažujúc zaistenie

Nech S značí celkový úhrn škôd získaných v našej simulácii. Hodnotu náhodnej veličiny S sme skonštruovali v nasledujúcich krokoch :

1. Simulujeme hodnotu príslušného počtu škôd N zo zvoleného rozdelenia.
2. Pre dané N generujeme výšky škôd X_1, \dots, X_N zo zvoleného rozdelenia.
3. Skonštruujeme úhrn škôd

$$S = \sum_{i=1}^N X_i,$$

resp. skonštruujeme úhrn plnení zo zaistenia

$$S_z = \sum_{i=1}^N \min(\max(X_i - r, 0), a).$$

2. Aproximácia rozdelenia úhrnov škôd

2.1 Rozdelenia zvolené pre aproximáciu rozdelenia agregátnych škôd

Definícia 4. *Nech $\mu \in \mathbb{R}$, $\sigma > 0$. Logaritmicke-normálne rozdelenie má hustotu*

$$f(x) = \frac{1}{\sqrt{2\pi x\sigma}} \exp\left[-\frac{(\log(x) - \mu)^2}{2\sigma^2}\right], \quad x > 0.$$

Strednú hodnotu a rozptyl logaritmicke-normálneho rozdelenia môžeme nájsť v knihe (Anděl (2011), str. 27). Platí :

$$EX = \exp\left[\mu + \frac{\sigma^2}{2}\right], \quad \text{Var } X = (e^{\sigma^2} - 1) e^{2\mu + \sigma^2}.$$

Z uvedených vzťahov jednoducho vyjadríme EX^2 ako :

$$EX^2 = \text{Var}X + (EX)^2 = e^{2\mu + 2\sigma^2} - e^{2\mu + \sigma^2} + e^{2\mu + \sigma^2} = e^{2\mu + 2\sigma^2}.$$

V nasledujúcej definícii budeme vychádzať z literatúry Wuthrich (2022).

Definícia 5. *Zvolme $\kappa \in \mathbb{R}$. Posunutú gama rozdelenie definujeme ako rozdelenie náhodnej veličiny*

$$X = \kappa + Z, \quad \text{kde } Z \sim \Gamma[\alpha, \beta].$$

Pre strednú hodnotu a rozptyl posunutého gama rozdelenia platí

$$EX = \kappa + \alpha\beta, \quad \text{Var } X = \alpha\beta^2.$$

Z čoho opäť dokážeme odvodiť vzťah pre EX^2 ako :

$$EX^2 = \text{Var}X + (EX)^2 = \alpha\beta^2 + \kappa^2 + 2\kappa\alpha\beta + \alpha^2\beta^2 = \alpha(\alpha + 1)\beta^2 + 2\kappa\alpha\beta + \kappa^2.$$

Pre využitie v ďalších výpočtoch odvodíme aj EX^3 . Budeme pritom vychádzať zo vzťahu pre koeficient šikmosti :

$$\gamma_X = \frac{E(X - EX)^3}{(\text{Var}X)^{\frac{3}{2}}} = \frac{EX^3 - 3EXEX^2 + 2(EX)^3}{(\text{Var}X)^{\frac{3}{2}}}.$$

Pre koeficient šikmosti posunutého gama rozdelenia platí :

$$\gamma_X = \frac{2}{\sqrt{\alpha}},$$

viz. Wuthrich (2022).

Pre EX^3 potom dostávame :

$$\begin{aligned} EX^3 &= \gamma_X(VarX)^{\frac{3}{2}} + 3EXEX^2 - 2(EX)^3 = \\ &= 2\alpha\beta^3 + 3(\kappa + \alpha\beta)(\alpha(\alpha + 1)\beta^2 + 2\kappa\alpha\beta + \kappa^2) - 2(\kappa + \alpha\beta)^3 = \\ &= \alpha(\alpha + 1)(\alpha + 2)\beta^3 + 3\kappa\alpha(\alpha + 1)\beta^2 + 3\kappa^2\alpha\beta + \kappa^3. \end{aligned}$$

Pre aproximáciu rozdelenia simulovaných agregátnych škôd sme zvolili tri, dvoj a trojparametrické, distribúcie. V nasledujúcej tabuľke uvádzame jednotlivé použité rozdelenia.

Typ rozdelenia	Parametre
Gama	$\alpha > 0, \beta > 0$
Log-Normálne	$\mu \in \mathbb{R}, \sigma > 0$
Posunuté gama	$\alpha > 0, \beta > 0, \kappa \in \mathbb{R}$

Tabuľka 2.1: Rozdelenia zvolené pre aproximáciu agregátnych škôd

Pri voľbe dvojparametrických rozdelení sme vychádzali z článku Papush a kol. (2021), kde voľbu dvojparametrických rozdelení argumentujú tým, že rozsah dát môže byť príliš malý pre spoľahlivý odhad troch parametrov.

Z rozdelení použitých v článku sme zvolili logaritnicko-normálne a gama rozdelenie pre ich dobré vlastnosti a široké využitie v aktuárskej praxi.

Ako tretie sme zvolili posunuté gama rozdelenie, ktoré sa od bežného gama rozdelenia líši parametrom umiestenia κ . Parameter κ teda určuje horizontálnu polohu hustoty gama rozdelenia a posúva ho na interval $[\kappa, \infty)$. Parameter κ zároveň zvyšuje flexibilitu gama rozdelenia a umožňuje zohľadniť situácie, pri ktorých je základné gama rozdelenie posunuté.

V literatúre Wuthrich (2022) sa voľba posunutého gama rozdelenia doporučuje práve preto, že dané rozdelenie má tri parametre a teda dokáže vystihnúť aj šikmost zistenú z dát.

2.2 Odhad parametrov

Ďalším krokom bolo odhadnúť parametre pre zvolené pravdepodobnostné rozdelenia pomocou simulovanej vzorky agregátnych škôd. V nasledujúcom texte objasníme použité metódy. Budeme vychádzať najmä zo skript Kulich (2014).

Metóda maximalnej vierohodnosti

Nech $\mathbf{X} = (X_1, \dots, X_n)$ je náhodný výber s parametrickým modelom

$$\mathcal{F} = \{\text{rozdelenia s hustotou } f(x|\boldsymbol{\theta}), \boldsymbol{\theta} \in \Theta \subseteq \mathbb{R}^d\},$$

a nech \mathbf{X} má združenú hustotu $\prod_{i=1}^n f(x_i|\boldsymbol{\theta})$. Maximálne vierohodný odhad $\hat{\boldsymbol{\theta}}$ parametru $\boldsymbol{\theta}_{\mathbf{X}}$ je taký bod z Θ , ktorý maximalizuje združenú hustotu spočítanú v pozorovaných hodnotách (X_1, \dots, X_n) cez všetky $\boldsymbol{\theta} \in \Theta$.

Definícia 6.

- *Náhodnú funkciu*

$$L_n(\boldsymbol{\theta}) = \prod_{i=1}^n f(X_i|\boldsymbol{\theta})$$

nazývame vierohodnostná funkcia pre parameter $\boldsymbol{\theta}$ v modeli \mathcal{F} .

- Maximálne vierohodný odhad parametru $\boldsymbol{\theta}_{\mathbf{X}}$ v modeli \mathcal{F} je definovaný ako

$$\hat{\boldsymbol{\theta}}_n = \arg \max_{\boldsymbol{\theta} \in \Theta} L_n(\boldsymbol{\theta}).$$

Odhadnuté parametre získané metódou maximalnej vierohodnosti spomínané v ďalšom texte tejto práce boli získané pomocou softvéru *Wolfram Mathematica*.

Momentová metóda

Cielom je opäť odhadnúť parameter $\boldsymbol{\theta}_{\mathbf{X}}$. Pri momentovej metóde využijeme to, že momenty rozdelenia náhodnej veličiny X_i vieme vyjadriť ako funkcie neznámych parametrov a to, že máme k dispozícii empirické odhady týchto momentov.

Porovnaním momentov daného rozdelenia a výberových momentov získame sústavu rovníc, ktorej riešením získame hľadané odhady. Príslušné sústavy rovníc pre odhad parametrov, pre rozdelenia z tabuľky 2.1 uvedieme v nasledujúcom texte.

Vzťahy pre EX_i a EX_i^2 gama rozdelenia sme odvodili v podkapitole 1.2. Konzistentným odhadom EX_i , resp. EX_i^2 , je výberový priemer \bar{X}_n , resp. výberový priemer $X_i^2, i = 1, \dots, n$, t.j. \bar{X}_n^2 . Na základe týchto informácií dokážeme odvodiť odhady $\hat{\alpha}, \hat{\beta}$ parametrov α, β gama rozdelenia riešením sústavy rovníc:

$$\begin{aligned}\bar{X}_n &= \hat{\alpha}\hat{\beta} \\ \overline{X_n^2} &= \hat{\alpha}(1 + \hat{\alpha})\hat{\beta}^2.\end{aligned}$$

S využitím vzťahov pre EX_i a EX_i^2 log-normálneho rozdelenia formulovaných v podkapitole 2.1 a už spomínaných konzistentných odhadov, dostaneme odhady $\hat{\mu}, \hat{\sigma}$ parametrov μ, σ logaritmickeo-normálneho rozdelenia riešením sústavy rovníc:

$$\begin{aligned}\bar{X}_n &= \exp\left[\hat{\mu} + \frac{\hat{\sigma}^2}{2}\right] \\ \overline{X_n^2} &= \exp\left[2\hat{\mu} + 2\hat{\sigma}^2\right].\end{aligned}$$

Vzťahy pre prvé tri momenty posunutého gama rozdelenia sme taktiež odvodili v podkapitole 2.1. Rovnakou úvahou ako v predchádzajúcich dvoch prípadoch a pomocou konzistentného odhadu $\overline{X_n^3}$ momentu EX_i^3 získame odhady $\hat{\alpha}, \hat{\beta}, \hat{\kappa}$ parametrov α, β, κ posunutého gama rozdelenia riešením sústavy rovníc:

$$\begin{aligned}\bar{X}_n &= \hat{\kappa} + \hat{\alpha}\hat{\beta} \\ \overline{X_n^2} &= \hat{\alpha}(\hat{\alpha} + 1)\hat{\beta}^2 + 2\hat{\kappa}\hat{\alpha}\hat{\beta} + \hat{\kappa}^2 \\ \overline{X_n^3} &= \hat{\alpha}(\hat{\alpha} + 1)(\hat{\alpha} + 2)\hat{\beta}^3 + 3\hat{\kappa}\hat{\alpha}(\hat{\alpha} + 1)\hat{\beta}^2 + 3\hat{\kappa}^2\hat{\alpha}\hat{\beta} + \hat{\kappa}^3.\end{aligned}$$

Pre scenár #1 popísaný v tabulke 1.1 sme pre odhad parametrov použili metódu maximálnej vierohodnosti aj metódu momentov. Odhady parametrov spočítané pomocou oboch metód boli vzájomne veľmi blízko. Avšak, rozdelenia s parametrami získanými momentovou metódou vykazovali po otestovnaí lepšiu zhodu, a preto sme pre daný scenár zobrazili iba výsledky získané touto metódou a pre všetky ostatné scenáre uvedené v tabulkách 1.1 a 1.2 sme hľadali odhady iba s metódou momentov.

Pri simulovaní úhrnov škôd podliehajúcich zaisteniu nastávajú situácie, kde žiaden zo škodných nárokov nezasahuje do uvažovanej zaistnej vsrty, a teda celkový úhrn škôd je rovný nule. Tieto situácie spôsobia nahromadenie sa núl v našom rozdelení. Preto rozdelenia, ktorých parametre sme pre tieto scenáre odhadovali a následne testovali ich zhodu mali tvar :

$$P[S_z \leq x] = p_0 + p_1 F(x), \quad (2.1)$$

kde $p_0 = P[S_z = 0]$, $p_1 = 1 - p_0$ a $F(x)$ je distribučná funkcia niektorého z rozdelení uvedených v tabulke 2.1.

Odhad \hat{p}_0 parametru p_0 dostaneme ako:

$$\hat{p}_0 = \frac{n_0}{n},$$

kde n_0 značí podiel počtu nulových hodnôt v simulovaných úhrnoch S_z a n značí počet všetky simulácii. Pre odhad \hat{p}_1 potom platí: $\hat{p}_1 = 1 - \hat{p}_0$.

Parametre týchto rozdelení opäť odhadneme metódou momentov. K-ty moment rozdelenia s distribučnou funkciou 2.1 má tvar :

$$ES_z^k = p_1 EX^k,$$

kde EX^k je k-ty moment niektorého z rozdelení v tabuľke 2.1. Parametre rozdelení s distribučnou funkciou 2.1 potom odhadujeme porovnaním teoretických vyjadrení EX^k a príslušného priemeru

$$\frac{1}{n_1} \sum_{z=1}^n S_z^k,$$

kde n_1 v menovateli zlomku je počet nenulových hodnôt v simulovaných úhrnoch S_z .

2.3 Testovanie dobrej zhody

Po získaní simulovaných agregátnych škôd a odhadnutí parametrov pre zvolené distribúcie sme testovali zhodu týchto rozdelení s empirickým rozdelením nasimulovaných agregátnych škôd.

Klasické testy odchýlky medzi dvomi rozdeleniami, ako napríklad Kolmogorov-Smirnovov test, poskytujú iba všeobecnú informáciu o tom, ako sa dve rozdelenia od seba navzájom líšia, ale nevedia určiť, či sa rozdelenia systematicky odlišujú pre určité množstvo hodnôt, napríklad v chvostoch. Práve pre agregátne škody sú hodnoty v chvostoch dôležité, keďže reprezentujú vysoké škodné nároky.

Pri testovaní sme vychádzali z postupu popísaného v Papush a kol. (2021) a Papush a kol. (2001), nazývaného autormi "percentile matching test".

Definícia 7. Funkcia prežitia, označovaná $S_X(x)$, je pre náhodnú veličinu X definovaná vzťahom

$$S_X(x) = P[X > x] = 1 - F_X(x).$$

Aby sme čo najlepšie zhodnotili rozdiely medzi jednotlivými distribúciami používanými na aproximáciu a empirickou distribúciou príslušnou simulovaným úhrnom škôd po celom ich rozsahu, zvolili sme **test percentilovej zhody**, ktorý porovnáva ich funkcie prežitia v niekoľkých zvolených hodnotách argumentu x .

Hodnoty x zvolené pre porovnávanie v tomto teste sme vyjadrili ako p percent z výberového priemeru príslušných simulovaných úhrnov škôd, t.j. $x = \frac{p}{100} \bar{X}_n$.

Následne porovnáme hodnoty funkcie prežitia pre jednotlivé rozdelenia s hodnotami tejto funkcie pre rozdelenie simulovaných úhrnov postupne vo všetkých zvolených hodnotách x .

Tento test prehľadne ukazuje, kde a o koľko sa dve rozdelenia líšia.

Jednotlivé voľby hodnôt x budú uvedené v tabuľkách pri príslušných výsledkoch jednotlivých scenárov.

3. Výsledky použitia uvažovaných aproximácií na simulované dáta

V nasledujúcich podkapitolách uvidíme pomocou tabuliek a grafov výsledky aproximácií rozdelenia úhrnov škôd simulovaných podľa zvolených scenárov uvedených v tabuľkách 1.1 a 1.2.

V prvom stĺpci každej tabuľky uvádzame jednotlivé voľby hodnôt p pre každý scenár.

Druhý stĺpec každej tabuľky zobrazuje hodnotu funkcie prežitia vyjadrenú v percentách pre simulované úhrny škôd.

Zvyšné tri stĺpce tabuľky zobrazujú rozdiel hodnôt funkcie prežitia v bode x pre empirické rozdelenie a testované rozdelenie, vyjadrený v percentách.

Grafy následne zobrazujú výsledky popísané v tabuľke, pričom červená čiara reprezentuje perfektnú zhodu aproximovaného rozdelenia s empirickým rozdelením simulovaných úhrnov škôd, teda vychýlenie o 0 % pre všetky hodnoty x .

Z výsledkov v podkapitolách 3.1 - 3.5 pozorujeme 2 závery. Pre bežné škody sa javí gama rozdelenie ako najlepšia aproximácia rozdelenia úhrnov škôd z troch uvažovaných rozdelení. Treba ale poznamenať, že posunuté gama rozdelenie taktiež nepreukazuje príliš veľké rozdiely v porovnaní s gama rozdelením.

V scenároch uvažujúcich zaistenie jednoznačne najlepšie figuruje posunuté gama rozdelenie. Druhú najlepšiu zhodu ukazuje gama rozdelenie, ktoré sa ku skutočnému rozdeleniu simulovaných úhrnov škôd približuje najmä v chvostoch, čo je dôležitý predpoklad pre škody, pri ktorých predpokladáme vysoké škodné nároky.

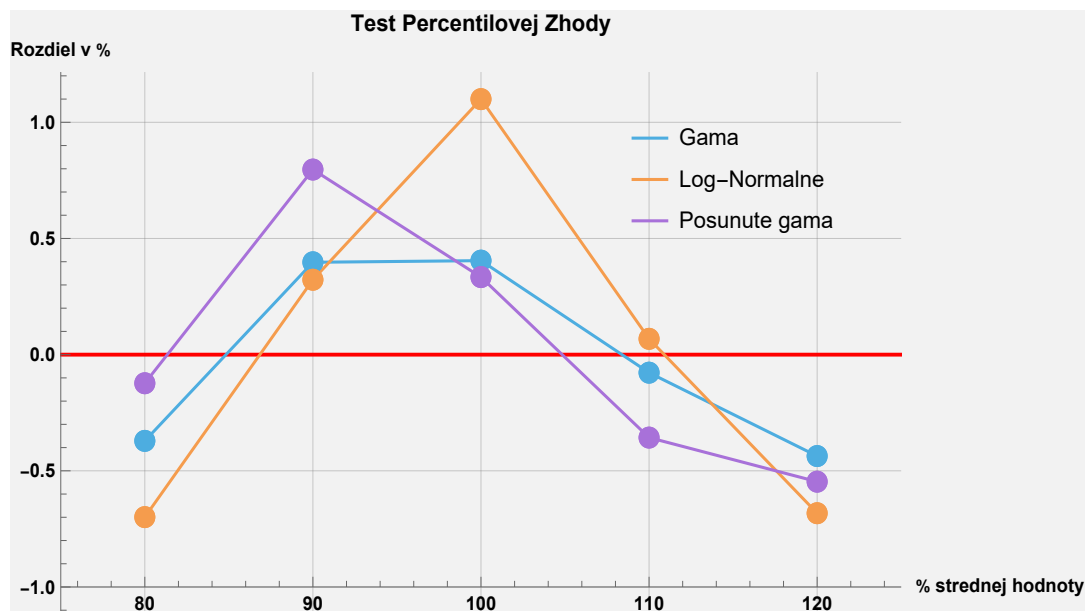
Ak by sme chceli voľbu rozdelenia úhrnov škôd zjednotiť pre všetky scenáre, môžeme pre odporúčanie výjsť z rozsahu dostupných dát. Ak máme k dispozícii dostatočný počet dát, môžeme zvoliť posunuté gama rozdelenie, keďže poskytuje väčšiu flexibilitu a vďaka tretiemu parametru dokáže vystihnúť aj šikmost uvažovaných dát. V prípade, že rozsah dát bude menší, môžeme zvoliť gama rozdelenie, pri ktorom stačí odhadovať 2 parametre a taktiež, ako môžeme vidieť z výsledkov, poskytuje dobrú zhodu s rozdelením simulovaných úhrnov škôd.

3.1 Výsledky scenáru 1 pre bežné škody

Scenár	Rozdelenie počtu škôd	Rozdelenie výšok škôd	Výberový priemer
#1	Poisson, $\lambda = 100$	Gama, $\alpha = 5, \beta = 3$	1 439

% Výberového priemeru	Empiricke	Gama	Log-Normal	Posunute gama
80%	97.36 %	-0.371 %	-0.699 %	-0.123 %
90%	83.16 %	0.398 %	0.322 %	0.797 %
100%	49. %	0.405 %	1.100 %	0.334 %
110%	16.98 %	-0.078 %	0.068 %	-0.357 %
120%	3.04 %	-0.436 %	-0.682 %	-0.547 %

Tabulka 3.1: Výsledky testu percentilovej zhody scenáru 1 pre bežné škody



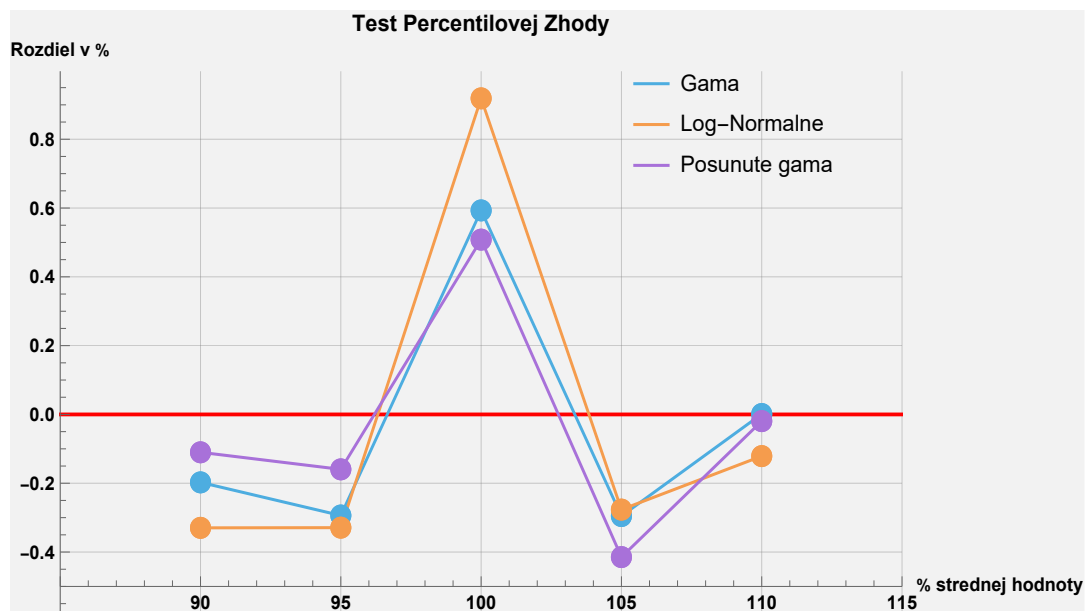
Obr. 3.1: Graf výsledkov testu percentilovej zhody scenáru 1 pre bežné škody

3.2 Výsledky scenáru 2 pre bežné škody

Scenár	Rozdelenie počtu škôd	Rozdelenie výšok škôd	Výberový priemer
#2	Poisson, $\lambda = 500$	Gama, $\alpha = 5, \beta = 2$	4 986

% Výberového priemeru	Empiricke	Gama	Log-Normal	Posunute gama
90%	97.98 %	-0.198 %	-0.330 %	-0.110 %
95%	84.3 %	-0.294 %	-0.329 %	-0.160 %
100%	49.94 %	0.593 %	0.919 %	0.508 %
105%	15.14 %	-0.295 %	-0.277 %	-0.415 %
110%	2.34 %	0.001 %	-0.121 %	-0.019 %

Tabulka 3.2: Výsledky testu percentilovej zhody scenáru 2 pre bežné škody



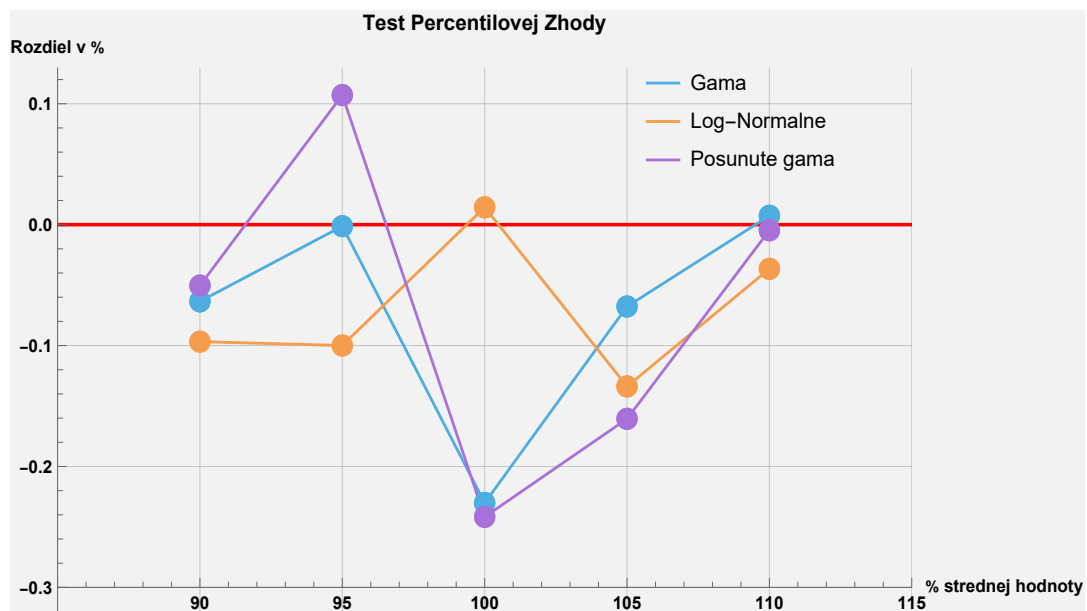
Obr. 3.2: Graf výsledkov testu percentilovej zhody scenáru 3 pre bežné škody

3.3 Výsledky scenáru 3 pre bežné škody

Scenár	Rozdelenie počtu škôd	Rozdelenie výšok škôd	Výberový priemer
#3	Poisson, $\lambda = 1000$	Gama, $\alpha = 3, \beta = 2$	5 998

% Výberového priemeru	Empiricke	Gama	Log-Normal	Posunute gama
90%	99.68 %	-0.063 %	-0.097 %	-0.050 %
95%	91.44 %	-0.001 %	-0.100 %	0.107 %
100%	49.28 %	-0.230 %	0.014 %	-0.242 %
105%	8.82 %	-0.068 %	-0.134 %	-0.161 %
110%	0.42 %	0.007 %	-0.037 %	-0.005 %

Tabulka 3.3: Výsledky testu percentilovej zhody scenáru 3 pre bežné škody



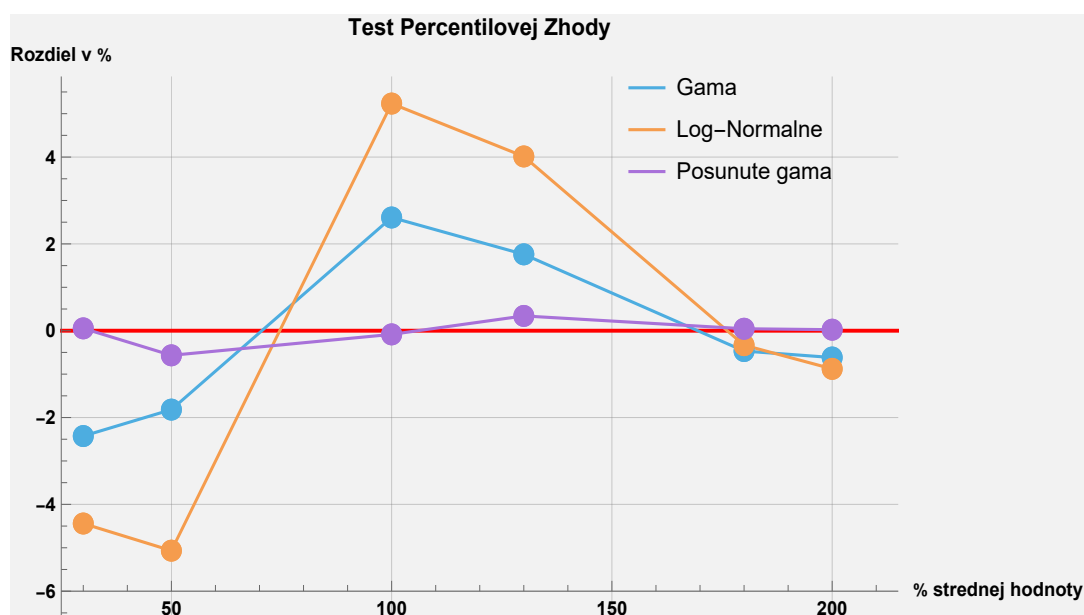
Obr. 3.3: Graf výsledkov testu percentilovej zhody scenáru 2 pre bežné škody

3.4 Výsledky scenáru 1 pre škody podliehajúce zaisteniu

Scenár	Rozdelenie počtu škôd	Rozdelenie výšok škôd	Zaistná vrstva	Výberový priemer
#1	Poisson $\lambda = 50$	Pareto $\theta = 50, \alpha = \frac{7}{6}$	750 x 250	2 365

% Výberového priemeru	Empiricke	Gama	Log-Normal	Posunute gama
30%	94.7 %	-2.427 %	-4.441 %	0.054 %
50%	84.976 %	-1.817 %	-5.067 %	-0.567 %
100%	46.2 %	2.608 %	5.233 %	-0.084 %
130%	25.16 %	1.759 %	4.014 %	0.341 %
180%	6.192 %	-0.468 %	-0.333 %	0.045 %
210%	3.194 %	-0.617 %	-0.879 %	0.024 %

Tabuľka 3.4: Výsledky testu percentilovej zhody scenáru 1 pre škody podliehajúce zaisteniu



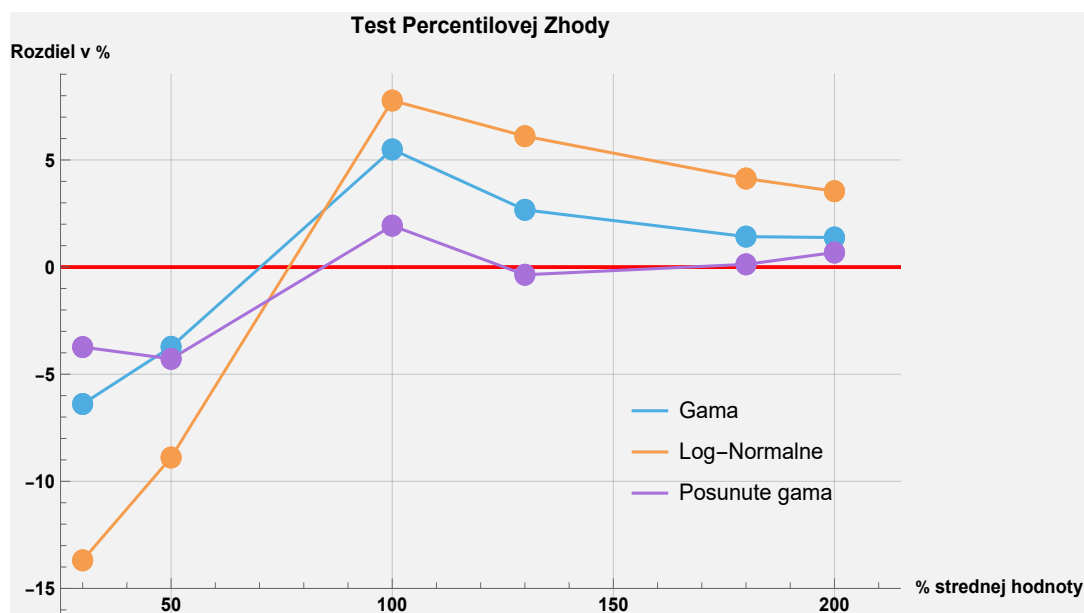
Obr. 3.4: Graf výsledkov testu percentilovej zhody scenáru 1 pre škody podliehajúce zaisteniu

3.5 Výsledky scenáru 2 pre škody podliehajúce zaisteniu

Scenár	Rozdelenie počtu škôd	Rozdelenie výšok škôd	Zaistná vrstva	Výberový priemer
#2	Poisson $\lambda = 10$	Pareto $\theta = 300, \alpha = \frac{22}{19}$	4000 x 1000	3 522

% Výberového priemeru	Empiricke	Gama	Log-Normal	Posunute gama
30%	71.598 %	-6.391 %	-13.690 %	-3.734 %
50%	62.206 %	-3.723 %	-8.888 %	-4.285 %
100%	45.374 %	5.491 %	7.774 %	1.933 %
130%	31.246 %	2.671 %	6.110 %	-0.359 %
180%	17.324 %	1.422 %	4.134 %	0.128 %
200%	13.864 %	1.380 %	3.543 %	0.681 %

Tabuľka 3.5: Výsledky testu percentilovej zhody scenáru 2 pre škody podliehajúce zaisteniu



Obr. 3.5: Graf výsledkov testu percentilovej zhody scenáru 2 pre škody podliehajúce zaisteniu

4. Edgeworthova aproximácia

Na začiatok uvedieme potrebnú teóriu, ktorú budeme následne používať pri odvodzovaní Edgeworthovej aproximácie.

V celej kapitole sme vychádzali najmä z literatúry Wuthrich (2022) a Mandl a Mazurová (1999).

Definícia 8. *Nech X je reálna náhodná veličina. Funkcie $M_x(t) = E[e^{tX}]$, $t \in \mathbb{R}$, sa nazývajú momentová vytvárajúca funkcia náhodnej veličiny X .*

Tvrdenie 2. *Nech existuje $b > 0$ také, že $M_x(t) < \infty$, $t \in (-b, b)$. Potom platí, že*

1. M_X má spojité derivácie všetkých rádov na $(-b, b)$;
2. $M^{(k)}(0) = EX^k = \mu'_k < \infty$ pre $k = 1, 2, \dots$

Dôkaz. Dôkaz vety je uvedený v knihe (Štěpán (1987) , str.165)

□

Lemma 3. *Nech Φ značí distribučnú funkciu normovaného normálneho rozdelenia a $\Phi^{(k)}$ je jej k -ta derivácia. Pre $k \in \mathbb{N}$ a $r \in \mathbb{R}$ platí*

$$r^k e^{\frac{r^2}{2}} = (-1)^k \int_{-\infty}^{\infty} e^{rx} \Phi^{(k+1)}(x) dx. \quad (4.1)$$

Dôkaz. Pri dôkaze budeme vychádzať z literatúry (Wuthrich (2022) , str.109). Dôkaz prevedieme indukciou podľa k . Zvolíme $k = 0$, potom

$$\int_{-\infty}^{\infty} e^{rx} \Phi'(x) dx = \int_{-\infty}^{\infty} e^{rx} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = M_X(r),$$

čo je momentová vytvárajúca funkcia náhodnej veličiny X s rozdelením $N(0,1)$. Pre túto momentovú vytvárajúcu funkciu platí:

$$\begin{aligned} M_X(r) &= \int_{-\infty}^{\infty} e^{rx} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \underbrace{e^{-\frac{x^2}{2} + rx}}_{= e^{-\frac{(x-r)^2}{2}} e^{\frac{r^2}{2}}} dx = \\ &= e^{\frac{r^2}{2}} \underbrace{\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-r)^2}{2}} dx}_{= 1} = e^{\frac{r^2}{2}}, \end{aligned}$$

pričom posledná rovnosť vyplýva z toho, že daný integrál je rovný 1, keďže sa jedná o integrál z hustoty náhodnej veličiny s rozdelením $N(r,1)$ cez celý jej nosič.

Indukčný krok $k \rightarrow k + 1$: Pomocou integrácie per partes dostávame

$$\begin{aligned}
& (-1)^{(k+1)} \int_{-\infty}^{\infty} e^{rx} \Phi^{(k+2)}(x) dx = \\
& = \left[(-1)^{k+1} e^{rx} \Phi^{(k+1)}(x) \right]_{-\infty}^{\infty} - (-1)^{k+1} \int_{-\infty}^{\infty} r e^{rx} \Phi^{(k+1)}(x) dx. \quad (4.2)
\end{aligned}$$

Prvý výraz na pravej strane v rovnosti 4.2 je rovný nule. Keďže $\Phi^{(k+1)}(x) = \phi^k(x)$, kde $\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ je hustota náhodnej veličiny s rozdelením $N(0,1)$, môžeme ukázať, že $\phi^k(x) = \pi_k(x)\phi(x)$, kde $\pi_k(x)$ je polynom stupňa k (pre odvodenie viz Mandl a Mazurová (1999), str.30). Z toho vyplýva, že $\Phi^{(k+1)}(x)$ je vždy polynom v x násobený funkciou $e^{-\frac{x^2}{2}}$, ktorá je pre konvergenciu v $\pm\infty$ rozhodujúca, a to aj v prípade, keď je $\Phi^{(k+1)}(x)$ násobená funkciou e^{rx} .

Predchádzajúci fakt a indukčný predpoklad pre k nám dáva:

$$(-1)^{(k+1)} \int_{-\infty}^{\infty} e^{rx} \Phi^{(k+2)}(x) dx = r(-1)^k \int_{-\infty}^{\infty} e^{rx} \Phi^{(k+1)}(x) dx = r r^k e^{r^2/2},$$

čo dokazuje dané lemma. □

4.1 Odvodenie Edgeworthovej aproximácie

Rozdelenia, ktorých zhodu s nasimulovanými dátami sme testovali metódou objasnenou v kapitolách 1 a 2 boli zvolené z rozdelení, ktoré sa bežne používajú v poisťno-matematickej praxi kvôli ich dobrým vlastnostiam. Pri Edgeworthovej aproximácii budeme vychádzať z normálnej aproximácie. Normálna aproximácia spočíva v priblížení rozdelenia štandardizovanej náhodnej veličiny rozdelením $N(0,1)$.

Edgeworthova aproximácia sa snaží pridať členy k normálnej aproximácii tým, že skúma Taylorov rozvoj momentovej vytvárajúcej funkcie.

Nech $f(x)$ značí hustotu štandardizovanej veličiny $Z = (S - \mu)/\sigma$, kde S značí náhodnú veličinu agregátnych škôd, μ , σ sú stredná hodnota a smerodatná odchýlka veličiny S , v tomto poradí. Pre štandardizovanú veličinu Z platí

$$E[Z] = 0, \text{ Var}[Z] = 1.$$

Ďalej nech

$$M_Z(r) = \int_{-\infty}^{\infty} e^{rx} f(x) dx \quad (4.3)$$

je príslušná momentová vytvárajúca funkcia. Uvažujme Taylorov rozvoj logaritmu $M_Z(r)$ v okolí nuly:

$$\begin{aligned}
b_Z(r) &= \log M_Z(r) = \sum_{k=0}^n r^k \frac{1}{k!} \frac{d^k}{dr^k} \log M_Z(r) \Big|_{r=0} = \\
&= \underbrace{\log M_Z(0)}_{= a_0} + \underbrace{\frac{d}{dr} \log M_Z(r) \Big|_{r=0}}_{= a_1} r + \underbrace{\frac{1}{2} \frac{d^2}{dr^2} \log M_Z(r) \Big|_{r=0}}_{= a_2} r^2 + \dots = \\
&= a_0 + a_1 r + a_2 r^2 + \dots
\end{aligned} \tag{4.4}$$

Hodnoty a_0, a_1, a_2, a_3, a_4 odvodíme pomocou tvrdenia 2 a toho, že pre náhodnú veličinu Z máme $EZ = 0$ a $EZ^2 = 1$. Dostávame :

$$a_0 = \log \underbrace{M_Z(0)}_{= 1} = \log 1 = 0,$$

$$a_1 = \frac{d}{dr} \log M_Z(r) \Big|_{r=0} = \frac{1}{M_Z(0)} M'_Z(0) = EZ = 0,$$

$$a_2 = \frac{1}{2} \frac{d^2}{dr^2} \log M_Z(r) \Big|_{r=0} = \frac{1}{2} \frac{M_Z^{(2)}(0) M_Z(0) - [M'_Z(0)]^2}{M_Z^2(0)} = \frac{1}{2} EZ^2 = \frac{1}{2},$$

$$\begin{aligned}
a_3 &= \frac{1}{6} \frac{d^3}{dr^3} \log M_Z(r) \Big|_{r=0} = \frac{1}{6} [EZ^3 - 3EZ^2EZ + 2[EZ]^3] = \\
&= \frac{1}{6} EZ^3.
\end{aligned}$$

$$\begin{aligned}
a_4 &= \frac{1}{24} \frac{d^4}{dr^4} \log M_Z(r) \Big|_{r=0} = \\
&= \frac{1}{24} [EZ^4 - 4EZ^3EZ - 3[EZ^2]^2 + 12(EZ)^2EZ^2 - 6(EZ)^4] = \\
&= \frac{EZ^4 - 3}{24}.
\end{aligned}$$

Pre získanie $M_Z(r)$ aplikujeme na výraz 4.4 exponenciálnu funkciu. Dostávame :

$$M_Z(r) = \exp\left[r^2/2 + \sum_{k=3}^{\infty} a_k r^k\right] = \exp[r^2/2] \exp\left[\sum_{k=3}^{\infty} a_k r^k\right].$$

Použitím Taylorovho rozvoja $e^x = 1 + x + x^2/2 + \dots$ na $\exp\left[\sum_{k=3}^{\infty} a_k r^k\right]$ dostávame:

$$\begin{aligned}
M_Z(r) &= \exp[r^2/2] \exp[a_3 r^3 + a_4 r^4 + \dots] = \\
&= e^{r^2/2} [1 + a_3 r^3 + a_4 r^4 + a_3^2 r^6/2 + \dots].
\end{aligned} \tag{4.5}$$

S využitím vzťahu (4.1) môžeme 4.5 prepísať:

$$\begin{aligned}
M_Z(r) &= \int_{-\infty}^{\infty} e^{rx} [\Phi^{(1)}(x) - a_3 \Phi^{(4)}(x) + a_4 \Phi^{(5)}(x) + \frac{a_3^2}{2} \Phi^{(7)}(x) + \dots] dx = \\
&= \int_{-\infty}^{\infty} e^{rx} [\phi(x) - a_3 \phi^{(3)}(x) + a_4 \phi^{(4)}(x) + \frac{a_3^2}{2} \phi^{(6)}(x) + \dots] dx,
\end{aligned} \tag{4.6}$$

pričom tretia rovnosť plynie zo vzťahu $\Phi^{(1)}(x) = \phi(x)$, kde $\phi(x)$ značí hustotu normovaného normálneho rozdelenia.

Z porovnania 4.3 a 4.6 plynie aproximácia hustoty $f(x)$ náhodnej veličiny Z pomocou Edgeworthovej aproximácie. Výpočet potom spočíva v použití určitého množstva sčítancov z radu

$$\phi(x) - a_3 \phi^{(3)}(x) + a_4 \phi^{(4)}(x) + \frac{a_3^2}{2} \phi^{(6)}(x) + \dots \tag{4.7}$$

4.2 Výsledky použitia Edgeworthovej aproximácie na simulované dáta

Postup, ktorým sme odvodili aproximovanú hustotu $f(x)$, ktorej zhodu s empirickým rozdelením štandardizovaných nasimulovaných dát sme následne testovali pomocou testu percentilovej zhody bol nasledovný :

1. Štandardizujeme nasimulované úhrny škod odčítaním výberového priemeru a vydelením výberovou smerodatnou odchýlkou.
2. Odhadneme a_3 a a_4 zo štandardizovaných dát.
3. Pomocou radu 4.7, z ktorého sme použili členy :

$$\phi(x) - a_3 \phi^{(3)}(x) + a_4 \phi^{(4)}(x) + \frac{a_3^2}{2} \phi^{(6)}(x),$$

odvodíme aproximovanú hustotu $f(x)$.

Tento postup sme opakovali pre každý scenár z tabuliek 1.1 a 1.2.

V kapitole 2 sme pre porovnanie rozdelení v teste percentilovej zhody použili násobky strednej hodnoty. Keďže pri Edgeworthovej aproximácii uvažujeme štandardizované dáta, ktorých stredná hodnota je 0, zvolili sme pre porovnanie hodnoty z intervalu $[-2, 2]$.

Výsledky testu percentilovej zhody pre rozdelenia získané pomocou Edgeworthovej aproximácie v jednotlivých scenároch uvedieme v tabuľkách na nasledujúcich stranách. Vysvetlíme najprv hodnoty uvádzané v jednotlivých stĺpcoch všetkých tabuliek.

V prvom stĺpci každej tabuľky uvádzame jednotlivé voľby hodnôt x , v ktorých sme porovnávali funkcie prežitia aproximovaného rozdelenia a empirického rozdelenia štandardizovaných simulovaných dát.

Druhý stĺpec zobrazuje hodnotu funkcie prežitia vyjadrenú v percentách pre simulované, štandardizované úhrny škôd.

Ako posledné uvádzame rozdiel hodnôt funkcie prežitia v bode x pre empirické rozdelenie štandardizovaných dát a rozdelenie získané Edgeworthovou aproximáciou, vyjadrený v percentách.

Porovnávané hodnoty	Empiricke	Edgeworthova aproximácia
$x=-2$	97.96 %	-0.044 %
$x=-1$	84.66 %	0.404 %
$x=-1/2$	68.8 %	0.026 %
$x=1/2$	30.24 %	-0.024 %
$x=1$	15.6 %	-0.144 %
$x=2$	2.64 %	0.054 %

Tabuľka 4.1: Výsledky Edgeworthovej aproximácie pre scenár 1

Porovnávané hodnoty	Empiricke	Edgeworthova aproximácia
$x=-2$	97.68 %	-0.164 %
$x=-1$	83.8 %	-0.363 %
$x=-1/2$	69.4 %	0.429 %
$x=1/2$	30.8 %	0.175 %
$x=1$	15.44 %	-0.397 %
$x=2$	2.54 %	0.136 %

Tabuľka 4.2: Výsledky Edgeworthovej aproximácie pre scenár 2

Porovnávané hodnoty	Empiricke	Edgeworthova aproximácia
x=-2	97.82 %	-0.103 %
x=-1	84.3 %	0.031 %
x=-1/2	68.52 %	-0.391 %
x=1/2	30.12 %	-0.245 %
x=1	15.7 %	-0.031 %
x=2	2.5 %	-0.021 %

Tabuľka 4.3: Výsledky Edgeworthovej aproximácie pre scenár 3

Porovnávané hodnoty	Empiricke	Edgeworthova aproximácia
x=-2	99.774 %	0.256 %
x=-1	83.882 %	-0.230 %
x=-1/2	66.324 %	-0.014 %
x=1/2	28.708 %	-0.039 %
x=1	16. %	0.112 %
x=2	3.612 %	0.115 %

Tabuľka 4.4: Výsledky Edgeworthovej aproximácie pre scenár 1 pre škody podliehajúce zaisteniu

Porovnávané hodnoty	Empiricke	Edgeworthova aproximácia
x=-2	100 %	0.000 %
x=-1	83.218 %	-0.716 %
x=-1/2	59.92 %	-3.306 %
x=1/2	26.144 %	-1.473 %
x=1	15.726 %	-0.340 %
x=2	4.474 %	0.409 %

Tabuľka 4.5: Výsledky Edgeworthovej aproximácie pre scenár 2 pre škody podliehajúce zaisteniu

Cielom využitia Edgeworthovej aproximácie v tejto práci bolo ukázať inú metódu pre aproximáciu rozdelenia úhrnov škôd ako fitovanie teoretických distribúcií.

Edgeworthova aproximácia vylepšuje normálnu aproximáciu tým, že okrem strednej hodnoty a rozptylu zachytených normálnou aproximáciou zohľadňuje aj vyššie momenty. Zahnutím vyšších momentov, ako sú šikmosť a špicatosť, umožňuje Edgeworthova aproximácia lepšie zachytenie vlastností skutočného rozdelenia simulovaných úhrnov škôd.

Ako môžeme vidieť z výsledkov uvedených v tabuľkách 4.1.-4.5. rozdelenie získané Edgeworthovou aproximáciou veľmi dobre popisuje skutočné rozdelenie simulovaných úhrnov škôd pre všetky scenáre.

Hoci Edgeworthova aproximácia ponúka dobrú presnosť pri opise rozdelenia úhrnov škôd, je dôležité poznamenať, že rozdelenie získané touto metódou má svoje nevýhody. Ako prvé poznamenajme, že aproximovaná funkcia $f(x)$ nie je hustota. Ďalšou nevýhodou môže byť to, že na Edgeworthovu aproximáciu potrebujeme odhad už spomínaných vyšších momentov, čo môže byť pri menšom počte dostupných dát limitujúce.

Vo všeobecnosti môžeme povedať, že Edgeworthova aproximácia je dobrým nástrojom na vylepšenie normálnej aproximácie, avšak je dôležité brať do úvahy konkrétnu vzorku úhrnov škôd kvôli limitáciám opísaným vyššie.

Záver

V prvej kapitole tejto práce sme objasnili metódu, ktorou sme modelovali úhrny škôd. Predstavili sme najprv voľby rozdelení počtov a výšok škôd, odvodili momenty týchto rozdelení a následne formulovali detaily konštrukcie úhrnov škôd.

Druhá kapitola rozoberá postup aproximácie rozdelenia simulovaných úhrnov škôd. Ako prvé sme uviedli rozdelenia, ktorých zhodu s rozdelením simulovaných úhrnov sme testovali. Nasledoval odhad parametrov týchto rozdelení a to najmä metódou momentov, ktorú sme detailne popísali. Ako posledné sme uviedli spôsob testovania zhody zvolených rozdelení s empirickým rozdelením simulovaných dát.

V tretej kapitole následne uvádzame výsledky použitia uvažovaných aproximácií na simulované dáta. Výsledky pre jednotlivé rozdelenia zobrazujeme v tabuľke, kde sa zameriavame najmä na rozdiel hodnôt funkcie prežitia vo zvolených bodoch pre empirické rozdelenie a testované rozdelenie. Tieto rozdiely zobrazujeme aj pomocou grafov. Z výsledkov usudzujeme, že najlepšou voľbou pre aproximáciu rozdelenia úhrnov škôd z uvažovaných rozdelení je gama alebo posunutá gama rozdelenie, odvíjajúc sa od toho, aký rozsah dát máme k dispozícii.

V poslednej kapitole volíme pre aproximáciu rozdelenia simulovaných úhrnov škôd Edgeworthovu aproximáciu. Najskôr odvodíme všeobecný tvar aproximácie hustoty $f(x)$ pomocou Edgeworthovej aproximácie a následne využijeme určitý počet sčítancov z odvodeného radu pre odvodenie aproximovanej hustoty $f(x)$ pre naše dáta. Výsledky použitia Edgeworthovej aproximácie na simulované dáta opäť uvádzame v tabuľkách.

Zoznam použitej literatúry

- ANDĚL, J. (2011). *Základy matematické statistiky*. Vydání třetí. Matfyzpress, Praha. ISBN 978-80-7378-162-0.
- KULICH, M. (2014). Přehledový větník. http://www.karlin.mff.cuni.cz/~pesta/NMFM301/statistika_fm.pdf. cit. : 2023-04-05.
- MANDL, P. a MAZUROVÁ, L. (1999). *Matematické základy neživotního pojištění*. Vydání první. Matfyzpress, Praha. ISBN 80-85863-42-1.
- PAPUSH, D. E., PATRIK, G. S. a PODGAITS, F. (2001). Approximations of the aggregate loss distribution. *Casualty Actuarial Society Forum*.
- PAPUSH, D. E., POPELYUKHIN, A. S. a ZHANG, J. S. (2021). Approximating the aggregate loss distribution. *Variance 14 (2)*.
- WUTHRICH, M. V. (2022). Non-life insurance: Mathematics and statistics. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2319328. cit. : 2023-04-05.
- ŠTĚPÁN, J. (1987). *Teorie pravdepodobnosti*. Academia, Praha.

Zoznam obrázkov

3.1	Graf výsledkov testu percentilovej zhody scenáru 1 pre bežné škody	16
3.2	Graf výsledkov testu percentilovej zhody scenáru 3 pre bežné škody	17
3.3	Graf výsledkov testu percentilovej zhody scenáru 2 pre bežné škody	18
3.4	Graf výsledkov testu percentilovej zhody scenáru 1 pre škody podliehajúce zaisteniu	19
3.5	Graf výsledkov testu percentilovej zhody scenáru 2 pre škody podliehajúce zaisteniu	20

Zoznam tabuliek

1.1	Rozdelenia použité pre simuláciu bežných škodných úhrnov	7
1.2	Rozdelenia použité pre simuláciu škodných úhrnov uvažujúc zaistenie	8
2.1	Rozdelenia zvolené pre aproximáciu agregátnych škôd	10
3.1	Výsledky testu percentilovej zhody scenáru 1 pre bežné škody . .	16
3.2	Výsledky testu percentilovej zhody scenáru 2 pre bežné škody . .	17
3.3	Výsledky testu percentilovej zhody scenáru 3 pre bežné škody . .	18
3.4	Výsledky testu percentilovej zhody scenáru 1 pre škody podliehajúce zaisteniu	19
3.5	Výsledky testu percentilovej zhody scenáru 2 pre škody podliehajúce zaisteniu	20
4.1	Výsledky Edgeworthovej aproximácie pre scenár 1	25
4.2	Výsledky Edgeworthovej aproximácie pre scenár 2	25
4.3	Výsledky Edgeworthovej aproximácie pre scenár 3	26
4.4	Výsledky Edgeworthovej aproximácie pre scenár 1 pre škody podliehajúce zaisteniu	26
4.5	Výsledky Edgeworthovej aproximácie pre scenár 2 pre škody podliehajúce zaisteniu	27