MOŽNOSTI APROXIMACE ROZDĚLENÍ KOLEKTIVNÍHO RIZIKA

[a)]Viera Pacáková, [b)]Veronika Balcárková

[a)]Univerzita Pardubice, Fakulta ekonomicko-správní, Ústav matematiky, b)Univerzita Pardubice, Fakulta ekonomicko-správní, studentka studijního oboru Pojistné **inženýrství**

Abstract: *The probability of a claim and the amount paid out in most short term non-life instance contracts is hard to predict. This article presents some mathematical/statistical techniques that allow finding the compound distribution of aggregate claims by combining the frequency of claims with the distribution of the amounts paid out on individual claims. The algorithm for exact expressions and calculating of the distribution of S is often too complicated in practical use. A quick approximation to the distribution of S can prove very useful and in many situations may be used.*

Keywords: *Frequency of Claims, Claim Size Distribution, Collective Risk Model, Normal Approximation, Translated Gamma Distribution*

1. Úvod

Základním pojmem v pojišťovnictví je pojem riziko, které představuje možnost vzniku pojistné události. Pojistitel, který profesionálně přebírá na sebe velké množství rizik různého charakteru, musí k ohodnocení rizika zaujmout velmi zodpovědný postoj. Riziko pojistitele spočívá v nebezpečí, zda přijaté pojistné bude postačující na vyplacení všech oprávněných pojistných plnění.

Teorie rizika je první oblast v pojistných vědách, ve které se vzhledem na praktický rozvoj zejména neživotního pojištění začali používat pokročilé metody teorie pravděpodobnosti a matematické statistiky.

Modely kolektivního rizika, pokud bychom vůbec nebrali v úvahu jejich mimořádný praktický význam pro krátkodobé pojistné smlouvy, zejména při neživotním pojištění, může sloužit jako ukázka široké aplikace pravděpodobnostně-statistických metod. Jejich základem jsou rozdělení počtu a výše individuálních pojistných plnění

2. Pravděpodobnostní model kolektivního rizika

2.1 Definice kolektivního rizika

Označíme písmenem N počet pojistných plnění za celé portfolio pojišťovny v daném období, písmenem X_i velikost $\boldsymbol{i}$-tého pojistného plnění pro $\boldsymbol{i} = 1,2,......N$. Pak

$$S = X_1 + X_2 + ... + X_N$$

představuje celkovou (agregovanou) sumu pojistných plnění, tzv. *kolektivní riziko* pojišťovny po uvažované období, nejčastěji kalendářní rok. Přitom platí:

- $X_1, X_2, ..., X_N$ jsou nezávislé a identicky rozdělené náhodné veličiny,
- náhodní veličiny $N, X_1, X_2, .., X_N$ jsou vzájemně nezávislé,
- pokud $N = 0$, tak i $S = 0$.

Našim cílem je najít alespoň přibližné rozdělení náhodné proměnné S. Prvním krokem k dosažení tohoto cíle je najít základní charakteristiky kolektivního riziko S, teda střední hodnotu $E(S)$, rozptyl $D(S)$ a koeficient šikmosti γ_1.

2.2 Základní charakteristiky kolektivního rizika S

Poměrně snadno lze dokázat, že charakteristiky kolektivního rizika můžeme vyjádřit pomocí charakteristik náhodných veličin N a X_i [3, s. 107-109].

Společnou distribuční funkce identicky rozdělených náhodných proměnných X_i označme $F(x)$. Dále ať symbol m_k, $k = 1,2,3,….$ značí k-tý počáteční moment identicky rozdělených individuálních pojistných plnění X_i, tj. $m_k = E\left(X_i^k\right)$. Potom platí

$$E(S) = E(N)m_1$$

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

Speciální případy složeného rozdělení S dostáváme v závislosti od rozdělení počtu pojistných plnění N. Pro praktické použití je nejdůležitější případ, kde N má Poissonovo rozdělení.

Ať je S celkové pojistné plnění, přičemž N má Poissonovo rozdělení s parametrem λ, symbolicky $N \sim Po(\lambda)$ a $F(x)$ je společná distribuční funkce identicky rozdělených individuálních pojistných plnění X_i. Pak S má *složené Poissonovo rozdělení* s parametry λ a $F(x)$ a označuje se $CoPo(\lambda, F(x))$ (v souladu s označením v anglické literatuře).

V tomto případě platí [3, s. 110]:

$$E(S) = lm_1, \quad D(S) = lm_2 \quad g_1 = \frac{lm_3}{(lm_2)^{\frac{3}{2}}} > 0$$

Protože $\lambda > 0$ a $X_i \geq 0$, platí i $g_1 > 0$. Složené Poissonovo rozdělení je tedy vždy pozitivně zešikmené, dokonce i tehdy, když je individuální pojistné plnění X negativně zešikmené. Pro velké hodnoty parametru λ se rozdělení $CoPo(\lambda, F(x))$ stává symetričtějším, protože platí, že $g_1 \to 0$, když $l \to \infty$.

3. Aproximace kolektivního modelu rizika

Aproximativní metody pro stanovení rozdělení pravděpodobnosti kolektivního rizika S se v praxi používají místo přesných analytických vzorců pro výpočet distribuční funkce $F_S(x)$, protože ty jsou často příliš složité a pracné. Pozornost je nejčastěji zaměřena na aproximaci rozdělení S pomocí normálního rozdělení a posunutého gama rozdělení.

3.1 Aproximace normálním rozdělením

Často pro kolektivní riziko S známe základní charakteristiky: střední hodnotu $E(S) = m$

a rozptyl $D(S) = s^2$. Protože $S = \sum_{i=1}^{N} X_i$ je součtem nezávislých a identicky rozdělených náhodných proměnných, nabízí se podle centrální limitní věty aproximace rozdělení S normálním rozdělením $N(\mu = E(S), \quad \sigma^2 = D(S))$.

Hustota normálního rozdělení je symetrická, což znamená, že pravý konec konverguje velmi rychle k nule. Při mnohých typech pojištění je však rozdělení S pravostranné, s pomalým klesáním pravé strany hustoty k ose x, tedy s dost vysokou pravděpodobností i extrémních pojistných plnění. Pro tyto typy pojištění má normální aproximace tendenci podhodnocovat hodnoty $P(S > x)$ pro velké hodnoty x. Z pohledu pojišťovatele je to velmi nežádoucí skutečnost, protože vysoké pojistné plnění má pro něho závažný finanční důsledek.

3.2 Aproximace posunutým gama rozdělením

Nechť jsou známé či spolehlivě odhadnutelné tři základní charakteristiky celkového pojistného plnění S, teda střední hodnotu $E(S) = m$, rozptyl $D(S) = s^2$ a koeficient šikmosti γ. V takovém případě můžeme použít posunuté gama rozdělení pro aproximaci rozdělení S. Posunuté gama rozdělení je rozdělení náhodné veličiny $k + Y$, kde náhodná veličina Y má gama rozdělení s parametry α, β a charakteristiky $E(Y) = \frac{\alpha}{\beta}$, $D(Y) = \frac{\alpha}{\beta^2}$ a koeficientem šikmosti $\gamma = \frac{2}{\sqrt{\alpha}}$. [3, s 63]

Parametry k, α a β se určí za předpokladu, že náhodná proměnná $k + Y$ má první tři momenty shodné s adekvátními momenty S, které jsou známé. Získává se tím systém tří rovnic s neznámými k, α a β [1, s. 93]:

$$g = \frac{2}{\sqrt{a}} \qquad s^2 = \frac{a}{b^2} \qquad m = k + \frac{a}{b}$$

Protože hodnoty μ, σ^2 a γ jsou podle předpokladu známé, postupně se vypočítají hodnoty neznámých parametrů:

$$a = \frac{4}{g^2} \qquad b = \sqrt{\frac{a}{s^2}} \qquad k = E(S) - \frac{a}{b}$$

4. Aplikace metod aproximace kolektivního rizika S

Pro ukázku aproximace rozdělení kolektivního rizika *S* jsme předpokládali, že počet pojistných plnění má Poissonovo rozdělení s parametrem $\lambda = 1000$. Rozdělení výše škod *X* jsme našli pomocí procedury *Distribution Fitting* statistického programového systému STATGRAPHICS *Centurion* XV. Vycházeli jsme ze znalosti 120 hodnot individuálních pojistných plnění, uvedených v [1, s. 36].

Data variable: x

120 values ranging from 3,0 to 32043,0

Fitted Distributions

Lognormal
mean = 2382,01
standard deviation = 7147,91
Log scale: mean = 6,62417
Log scale: std. dev. = 1,51758

Kolmogorov-Smirnov Test

	Lognormal
DPLUS	0,0520834
DMINUS	0,0867263
DN	0,0867263
P-Value	0,329632

Obr. 1: Výstup procedury Distribution Fitting, STATGRAPHICS Centurion XV

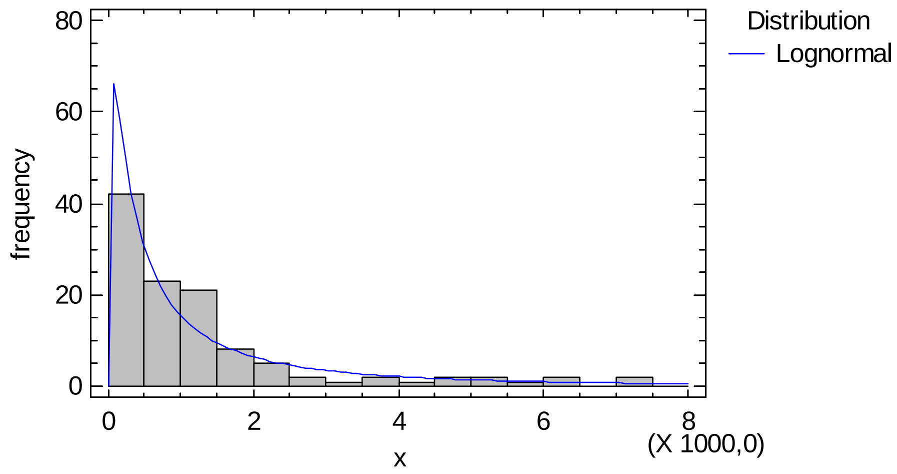
Zdroj: vlastní

Na základě výstupu za systému STATGRAPHICS *Centurion* XV na Obr. 1 můžeme lognormální rozdělení s parametry

$$\mu = 6{,}62417 \text{ a } \sigma = 1{,}51758$$

odhadnutými metodou maximální věrohodnosti považovat na hladině významnosti $\alpha = 0{,}05$ za vhodný model výše individuálních pojistných plnění *X*, protože hodnota P-Value = 0,329632 Kolmogorovova-Smirnovova testu je větší než $\alpha = 0{,}05$. [2, s. 86]

Vizuálně můžeme vhodnost tohoto rozdělení posoudit na Obr. 2, kde je histogram empirických dat proložen hustotou lognormálního rozdělení s odhadnutými parametry.



Obr. 2: Vizuální posouzení výsledku testů dobré shody, STATGRAPHIC Centurion XV

Zdroj: vlastní

Pro počáteční momenty lognormálního rozdělení platí vztah [3, s. 67] $E(X^k) = e^{k \cdot m + k^2 \cdot \frac{s^2}{2}}$. Na základě tohoto vztahu byly vypočítány první tři počáteční momenty m_k náhodné veličiny X (Tab. 1):

Tab. 1: Počáteční momenty pojistných plnění X, MS Excel

k	$k \cdot m + k^2 \cdot \frac{s^2}{2}$	m_k
1	7,611	2 020,292
2	16,789	19 552 940,142
3	27,533	906 554 660 446,952

Zdroj: [vlastní]

Dále byly vypočítány základní charakteristiky celkového pojistného plnění S postupem, uvedeným v části 2.2. Jsou uvedeny v Tab. 2.

Na základě části 3.1 můžeme rozdělení S aproximovat normálním rozdělením s parametry $\mu = 2\,020\,292$ a $\sigma^2 = 19\,682\,951\,585$.

Tab. 2: Základní charakteristiky S

E(S)	2 020 291,66667
D(S)	19 682 951 584,71930
γ (S)	0,12449

Zdroj: vlastní

Aproximaci kolektivního rizika *S* jsme provedli postupem, uvedeným v části 3.2. Hodnoty odhadnutých parametrů obsahuje Tab. 3.

Tab. 3: Hodnoty parametrů posunutého gama rozdělení

α	258,11589
β	0,00011
k	-233 700,91177

Kolektivní riziko *S* má tedy přibližně stejné rozdělení jako náhodná veličina -233 700,9+*Y*, kde *Y* má gama rozdělení s parametry $\alpha = 258{,}116$ a $\beta = 0{,}00011$.

5. Závěr

Znalost rozdělení pravděpodobnosti kolektivního rizika S, resp. celkového pojistného plnění pojišťovny za rok má pro pojišťovnu mimořádný význam. Je základem řešení mnoha klíčových problémů, jako je stanovení výše pojistného, rozhodování o optimální formě zajištění, posouzení rizika pomocí hodnoty v riziku VaR apod. Střední hodnota $E(S)$ přímo definuje tzv. čisté (netto) pojistné. Znalost rozdělení S, alespoň aproximativního, však umožní najít jeho kvantily a stanovit tzv. rizikové pojistné RP. Například jestli rizikové pojistné stanovíme jako 95. percentil rozdělení S, teda $RP = S_{0,95}$, pojišťovna připouští riziko 0,05, že vyinkasované pojistné nebude postačující na výplatu pojistných plnění.

Příspěvek je součástí řešení grantového projektu GAČR 402/09/1866 s názvem *Modelování, simulace a řízení pojistných rizik.*

Použité zdroje:

[1] BOLAND,P. J. *Statistical and Probabilistic Methods in Actuarial Science*. London: Chapman&Hall/CRC, 2007. 351 s.ISBN 978-1-58488-695-2.

[2] KUBANOVÁ, J. *Statistické metody pro ekonomickou a technickou praxi.* Bratislava: Statis, 2004. 249 s. ISBN 80-85659-37-9.

[3] PACÁKOVÁ, V. *Aplikovaná poistná štatistika*. Bratislava: vydavatelství IURA EDITION, 2004. 261 s. ISBN 80-8078-004-8.

[4] PACÁKOVÁ, V. Modelovanie a simulácia rizík v neživotnom pojistěni. *Ekonomika a management*. roč. 10, Technická univerzita v Liberci, 3/2007, ISSN 1212-3609.

Kontaktní adresa:

prof. RNDr. Viera Pacáková, PhD.
Bc. Veronika Balcárková
Univerzita Pardubice
Fakulta ekonomicko-správní
Ústav matematiky
Studentská 84, 532 10 Pardubice
e-mail: vpacakova@gmail.com
e-mail: verabalcarkova@seznam.cz