SLOVENSKÁ ŠTATISTIKA a DEMOGRAFIA

SLOVAK STATISTICS and DEMOGRAPHY

1/2018
ročník/volume 28

Recenzovaný vedecký časopis so zameraním na prezentáciu moderných štatistických a demografických metód a postupov.

Scientific reviewed journal focusing on the presentation of modern statistical and demographic methods and procedures.

Článok/Article: 1

Typ článku/Type of article: vedecký článok/scientific article

Strany/Pages: 3 – 17

Dátum vydania/Publication date: 15. január 2018/January 15, 2018

Michal PÁLEŠ
Katedra matematiky a aktuárstva, Fakulta hospodárskej informatiky Ekonomickej univerzity v Bratislave

VYUŽITIE JAZYKA R NA ODHAD MIER RIZIKA S VYUŽITÍM SIMULÁCIÍ

THE USE OF LANGUAGE R FOR ESTIMATING RISK MEASURES USING SIMULATIONS

ABSTRAKT
Jedna z úloh analýzy rizika môže predstavovať určenie najvhodnejšieho stochastického modelu v podobe náhodnej premennej opisujúcej dostupné údaje. Tieto modely sú základnými jednotkami sofistikovanejších modelov určených na kvantifikáciu rizika. V rámci kolektívneho modelu rizika potom predikujeme rozdelenie pravdepodobnosti počtu škôd, individuálnej výšky škody a agregovanej škody. Následne je dôležité zvoliť miery rizika, ktorými budeme skúmané riziko merať. Jednou z užitočných metód, ktorá slúži na meranie rizík je metóda hodnoty v riziku (VaR) spočívajúca v určení kvantilov príslušných rozdelení, ktorú možno rozšíriť prostredníctvom koherentnej podmienenej hodnoty v riziku (CVaR). Na základe týchto hodnôt vie aktuár modelovať ekonomický kapitál, ktorý zabezpečí solventnosť poisťovne s vysokou pravdepodobnosťou. Miery rizika VaR a najmä CVaR možno určiť viacerými metódami, z ktorých väčšina nie je jednoduchá. V tomto príspevku predstavíme simulačné nástroje jazyka R, ktoré aktuárovi umožnia sofistikovane odhadovať tieto hodnoty.

ABSTRACT
One of the tasks of risk analysis is to determine the most appropriate stochastic model using a random variable describing the data available. These models are the basic units of the more sophisticated models intended for risk quantification. In the collective risk model, the probability distribution of the number of claims, individual claim amount and the aggregate claim distribution is thereafter predicted. Consequently, it is important to choose the risk measures by which the examined risk will be measured. One of the useful risk measurement methods is the value at risk (VaR) method for the determination of quantiles of the respective distributions and which may be extended by a coherent conditional value at risk (CVaR). On the basis of these values, the business capital assuring the insurance company solvency with high probability can be modelled by an actuary. The VaR risk measures and in particular the CVaR can be determined by a number of methods, most of which are not simple. This paper introduces the simulation tools of the R language, enabling the actuary the sophisticated estimation of values.

KĽÚČOVÉ SLOVÁ
miery rizika, hodnota v riziku, *CVaR*, Solventnosť II, analýza rizík, jazyk R

KEY WORDS
risk measures, Value at Risk, *CVaR*, Solvency II, risk analysis, R language

1. ÚVOD
Riziká (finančné alebo poistné) možno rozdeliť do rôznych kategórií. Finančné riziká sa najčastejšie spájajú s trhovými a kreditnými rizikami, rizikom likvidity a operačným rizikom. V poisťovníctve je ale popri týchto rizikách významné aj poistno-technické

riziko. Riziká možno kvantifikovať pomocou vhodných mier rizika. Metodika merania rizík súvisí najmä s direktívou Solventnosť II, kde sa určité vhodné miery rizika inštitucionalizujú tak, aby čo najlepšie vyhovovali požiadavkám regulátora a súčasne rovnako aj vnútornému controllingu poisťovne (v bankách metodika Basel III).

Meranie rizika sa dá potom chápať ako komplex činností (prístup, nástroje, metodiky, postupy) pri analýze (riadení) rizika (rizikových faktorov) v určitej oblasti. Prístupy k meraniu rizika sa delia na stochastické a deterministické. Významným je v súčasnosti stochastický prístup, ktorý pracuje s *mierami rizika* založenými na rozdeleniach pravdepodobnosti náhodnej premennej (ďalej tiež NP) opisujúcej napr. stratu portfólia v oblasti finančných rizík a škodu v oblasti poistných rizík. Tieto môžeme najčastejšie rozdeliť na kvantilové (pozri nižšie) a rozptylové (rozptyl, štandardná odchýlka).

Kvantilové miery rizika predstavujú minimálnu kapitálovú požiadavku, aby príslušná pozícia bola takmer bezriziková, resp. aby kapitál vo výške tejto kapitálovej požiadavky nepokryl možnú stratu len s veľmi malou pravdepodobnosťou známej výšky. Najčastejším predstaviteľom kvantilovej miery rizika je ***hodnota v riziku*** (Value at Risk, *VaR*, pozri ďalej). Miery rizika sa obvykle využívajú pri meraní trhového rizika (napr. v bankovom portfóliu v krátkych časových intervaloch) a pri meraní poistného rizika (v intervale napr. jedného roka). Prostredníctvom mier rizika je možné určiť ekonomický kapitál pre dané riziko.

Ekonomický kapitál (tiež rizikový, regulatórny) je kapitál, ktorý vlastníci (napr. akcionári) musia investovať do spoločnosti, aby udržali jej solventnosť, ktorá s určitou pravdepodobnosťou zaručuje nepretržitý chod spoločnosti v danom období. Tento kapitál by mal garantovať, že možné riziká nespôsobia úpadok spoločnosti, teda že nesolventnou by sa mohla stať len pri katastrofických a veľmi nepravdepodobných udalostiach (pričom ich výskyt sa podľa regulatórnych metodík odhaduje najčastejšie s 5 %, resp. 1 % pravdepodobnosťou). Ekonomický kapitál sa určuje na základe rizík, ktorým je spoločnosť vystavená, pričom podľa druhu rizika sa určí ekonomický kapitál potrebný na ich krytie. Súvisí so stanovením ***kapitálovej požiadavky na solventnosť*** (Solvency Capital Requirement).

Riziká, ktoré poisťovne plánujú podľa Solvency II obsiahnuť *čiastočnými internými modelmi*, sú napríklad v oblasti neživotného poistenia najmä upisovacie riziko pri poistení zodpovednosti za škodu spôsobenú prevádzkou motorových vozidiel a v havarijnom poistení. Práve na základe znalostí pravdepodobnostných zákonitostí a aj s využitím matematického modelovania majú poisťovne možnosť rozlišovať a posudzovať riziko tak, že straty spojené s poistnými udalosťami sú relatívne nižšie a takisto ich možno využiť na vývoj a oceňovanie nových produktov. Aktuár preto potrebuje pri svojej činnosti poznať pravdepodobnostné rozdelenia, ktoré sú vhodné na modelovanie počtu, individuálnej výšky škôd aj celkovej škody pri rôznych typoch poistných produktov. Posúdenie vystavenia sa riziku, ktoré by malo obsahovať detailnú ukážku procesu poisťovne na hodnotenie rizík (kvalitatívne aj kvantitatívne hodnotenie) pre štandardné aj extrémne podmienky simulované buď stresovými testami, alebo stochastickými modelmi.

Pre ďalšie informácie o metodike Solventnosť II, kapitálových požiadavkách, aktuárskych modeloch a prístupu k rizikám pozri napr. [1], [6].

2. MIERY RIZIKA

Kvantilové miery rizika súvisia s konkrétnym kvantilom rozdelenia pravdepodobnosti náhodnej premennej a pomocou nich je následne možné kalkulovať kapitál potrebný na krytie neočakávaných škôd. *VaR* je nezanedbateľný nástroj charakterizácie rizika ako porovnávajúca hodnota rizík, alebo tiež ako miera rizika. Definujeme ju ako **najhoršiu možnú stratu (škodu), ku ktorej môže dôjsť s vopred stanovenou (požadovanou) pravdepodobnosťou za určité časové obdobie**.

$VaR_p(X)$ náhodnej premennej X opisujúcej škodu s pravdepodobnosťou p je 100 p % kvantil, označovaný ako x_p , $0 < p < 1$, pre ktorý platí

$$VaR_p(X) = \inf\{x \in R : F_X(x) \geq p\} = x_p.$$

Nedostatkom *VaR* je, že nespĺňa vlastnosti ***koherentnej (relevantnej) miery rizika***, konkrétne podmienku *subaditivity* ($\alpha\ (X+Y) \leq \alpha\ (X) + \alpha\ (Y)$). Tiež nič nehovorí o rozdelení extrémnych strát (škôd), ktoré sú väčšie ako táto hodnota. Dôležitou koherentnou mierou rizika, ktorá odstraňuje tieto nedostatky je ***podmienená hodnota v riziku*** (Conditional Tail Expectation, *CTE*, tiež *AVaR*, *TVaR*, *ES*, ďalej ako Conditional *VaR*, *CVaR*).

$CVaR_p(X)$ je **očakávaná strata (škoda) zo všetkých škôd prekračujúcich hodnotu kvantilu** x_p príslušného rozdelenia náhodnej premennej X. Pre spojitú náhodnú premennú $CVaR_p(X)$ vyjadríme vzťahom

$$CVaR_p(X) = E\left(X \middle| X > x_p\right) = \frac{\int_{x_p}^{\infty} x \cdot f_X(x)\,\mathrm{d}x}{P(X > x_p)}.$$

Okrem vyššie uvedených dvoch základných mier rizika sa v praxi objavujú aj iné (modifikované) miery rizika, resp. miery rizika založené na iných prístupoch, (pre viac informácií pozri [1]).

Pre ekonomický kapitál *EC* potom platí

$$EC = VaR_p(X) - E(X),$$

kde $E(X)$ je stredná hodnota náhodnej premennej *X*. Analogicky platí pre $CVaR_p(X)$.

V aktuárskych analýzach štúdium rozdelení s ťažkými koncami *CVaR* poskytuje dôležité informácie o potenciálnom vysokom riziku. Tieto informácie súvisia s distribučnou funkciou, ktorá je základným nástrojom modelovania škôd. Pre každé uvedené spojité rozdelenie, je hodnota *CVaR* vyjadriteľná. Vyjadrenie príslušnej hodnoty *CVaR* je však vo väčšine prípadov náročné, čo komplikuje jej aproximáciu do rizikových modelov.

Tweedie rozdelenia sú špeciálnym prípadom exponenciálnych disperzných modelov používaných na opísanie rozdelenia chýb zovšeobecnených lineárnych modelov (GLM). Názov sa vzťahuje k exponenciálnemu tvaru, ktorý sa využije na opísanie zákona rozdelenia. Numerický postup, ktorý vedie k vyjadreniu hodnoty *CVaR* pre rozdelenia z triedy Tweedie, ktoré sú vhodné na aktuárske analýzy vrátanie stabilných a extrémne stabilných rozdelení, uvádza [3]. Tieto nadobúdajú nezáporné hodnoty, pričom každé z uvedených rozdelení je špecifikované definičným oborom parametra, pomocou ktorého je vyjadrený vzťah disperzie a strednej hodnoty, a ich momentové vytvárajúce funkcie sú vyjadrené predpismi, ktoré umožňujú stanovenie tejto hodnoty. Tweedie modely majú jednoduché vyjadrenie momentových vytvárajúcich funkcií, na základe ktorých je možné poskytnúť rýchle a bezproblémové stanovenie koherentnej miery rizika.

3. MODELOVANIE POČTU ŠKÔD, INDIVIDUÁLNEJ A CELKOVEJ ŠKODY

Viaceré rozdelenia pravdepodobnosti môžu slúžiť ako rozdelenie počtu alebo výšky škody, resp. poistných plnení. Všetky dané rozdelenia závisia od jedného, príp. viacerých parametrov. V konkrétnych praktických situáciách je určenie týchto parametrov prvým závažným problémom pri hľadaní vhodného pravdepodobnostného modelu počtu, resp. výšky škody. Pretože neznáme parametre určujeme na základe neúplných informácií o konkrétnych poistných prípadoch, táto úloha patrí do okruhu problémov štatistickej indukcie.

Reálne podmienky, ktoré pri rôznych druhoch poistenia vedú k vzniku poistnej udalosti spôsobujú, že **počet škôd** opísaných náhodnou premennou N, má najčastejšie niektoré z týchto diskrétnych rozdelení pravdepodobnosti – *alternatívne, geometrické, binomické, Poissonovo a negatívne binomické rozdelenie*. Základným orientačným kritériom na výber vhodného diskrétneho rozdelenia je vzťah medzi strednou hodnotou $E(N)$ a rozptylom $D(N)$ náhodnej premennej, pričom tieto charakteristiky sa odhadujú na základe známych údajov o počte škôd (viac v [3]).

Modely **výšky škôd** vychádzajú z pravdepodobnostného rozdelenia náhodnej veličiny X opisujúcej výšku škody alebo aj výšku poistného nároku vzťahujúcu sa na jednu poistnú udalosť. Pri ich modelovaní treba brať do úvahy, že sa tu vyskytujú aj poistné udalosti s výrazne vyššou hodnotou výšky škody. Náhodné premenné preto majú veľký rozptyl a bývajú zošikmené. Skúsenosti z modelovania výšky škody poukazujú na skutočnosť, že rozdelenie výšky škody výrazne závisí aj od najväčšieho plnenia, napríklad veľkého požiaru, záplavy a pod. Ako najvhodnejšie rozdelenia na modelovanie výšky škody sa javia teda rozdelenia s pomerne rýchlou konvergenciou k nule a so zošikmením z pravej strany. Na modelovanie individuálnej výšky škody sa často využívajú – *exponenciálne, gama, Paretovo, Weibullovo a lognormálne rozdelenie*. Z hľadiska akturáskej terminológie je potrebné rozlišovať výšku poistnej škody a poistné plnenie.

Základom riešenia rôznych rozhodujúcich otázok pre aktuára (súvisiacich so stanovením poistného, výškou rezerv, zaistením, pravdepodobnosťou krachu, agregáciou rizík a pod.) je znalosť základných charakteristík a rozdelenia pravdepodobnosti celkovej škody, tzv. ***kolektívny model rizika***. Publikácie [3] a [6] opisujú stochastické aktuárske modely pri riadení poistného rizika vychádzajúce prevažne z „americkej“ teórie rizika, ktorá sa zaoberá analyzovaním počtu, individuálnej výšky a celkovej škody. Definuje sa náhodná premenná S, ktorá predstavuje **celkovú** (súhrnnú, agregátnu) **škodu**, resp. poistné plnenie priameho poisťovateľa za všetky poistné udalosti počas roka. Cieľom je vo všeobecnosti nájsť rozdelenie pravdepodobnosti a základné charakteristiky celkovej škody S.

Potom nech náhodná premenná N opisuje počet škôd, resp. počet poistných plnení, ktoré v sledovanom období nastanú, a náhodná premenná X_i opisuje výšku i-tej škody, resp. celkového poistného plnenia. Potom celkovú škodu vyjadríme ako súčet všetkých individuálnych škôd, čo zapíšeme vzťahom

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

pričom je zrejmé, že ak $N = 0$, tak $S = 0$. Ďalej musia byť splnené tieto predpoklady, a to $X_1, X_2, \dots, X_N$ sú nezávislé a identicky rozdelené náhodné premenné a náhodné premenné *N* a $X_1, X_2, \dots, X_N$ sú vzájomne nezávislé. Celková škoda je však závislá od počtu poistných udalostí, pričom rozdelenie individuálnej škody sa počas celého sledovaného obdobia, tak ako aj počet poistných zmlúv, nemení. Z toho vyplýva, že uvažujeme riziko celkové, a nie riziko z jednotlivých poistných zmlúv, a fakt, že na každú poistnú zmluvu môže nastať viacero poistných plnení. Podmienky vo vzťahu vyššie zodpovedajú len skutočným škodám ($X > 0$). Potom náhodná premenná, ktorá spĺňa dané predpoklady, má ***zložené rozdelenie pravdepodobnosti***

$$S \sim Co(p_N(n); F_X(x)),$$

kde $p_N(n)$ je pravdepodobnostná funkcia a $F_X(x)$ je distribučná funkcia bez špecifikácie rozdelení. Matematicky možno ďalej vyjadriť distribučnú funkciu $F_S(x)$, pravdepodobnostnú funkciu $p_S(x)$ a momenty rozdelenia pravdepodobnosti celkovej škody pre ľubovoľné rozdelenie počtu škôd *N* a individuálnej výšky škody *X*. V praxi sa využívajú viaceré metódy na určenie $F_S(x)$ (konvolúcie, aproximácie, Panjerove rekurentné vzťahy, simulácie Monte Carlo), ktoré často súvisia s využitím výpočtovej techniky.

4. VYUŽITIE SIMULÁCIÍ NA ODHAD MIER RIZIKA V JAZYKU R

Jazyk R poskytuje možnosti, ktorým najmä v oblasti metodík výpočtu technických rezerv, simulácií a širokom využití zovšeobecnených lineárnych modelov konkuruje máloktorý komerčný softvér. Súčasne tvorba doplnkových knižníc z oblasti aktuárskych vied s pokročilými funkciami rastie závratnou rýchlosťou. Použitie open-source softvéru R v aktuárskych analýzach široko opisuje napr. publikácia [7].

R samozrejme umožňuje prácu so zabudovanými rozdeleniami diskrétnej a spojitej náhodnej premennej. Ďalšie špecifické rozdelenia môže používateľ nájsť v doplňujúcich balíčkoch (knižniciach, napr. *actuar*). Je dôležité si uvedomiť, že názov aj parametre rozdelení sa v jednotlivých knižniciach môžu líšiť. Takisto treba poznať tvar hustoty pravdepodobnosti, resp. distribučnej funkcie daného rozdelenia, preto je dôležité sledovať dostupnú dokumentáciu k danej knižnici.

Náhodné čísla generované v R sú generované na základe zabudovaných algoritmov (pozri napr. `?rnorm/RNG`). Podľa pravdepodobnosti výskytu jednotlivých hodnôt môžeme napr. generovať čísla z rôznych typov rozdelení funkciou `r***` (napr. `rnorm(30,0,1)` na generovanie 30 hodnôt z normovaného normálneho rozdelenia) alebo môžeme efektívne využiť funkciu `sample`. Špecifickým otázkam simulácií v jazyku R sa venujú rôzni autori (Charpentier, Jones, Kaas a kol., Robinson, a i.).

Podľa [3] niektoré metódy generovania pseudonáhodných čísiel generujú hodnoty z rovnomerného normovaného rozdelenia, teda výskyt ľubovoľného čísla z intervalu (0;1) je rovnako pravdepodobný a mimo tohto intervalu je pravdepodobnosť výskytu nulová. Čísla s rovnomerným normovaným rozdelením možno transformovať na čísla s požadovaným rozdelením. Existujú rôzne metódy transformácie, jednoduché aj zložitejšie. Často používanou je ***metóda inverznej transformácie***.

Ak napr. $F_X(x)$ je spojitá a rastúca distribučná funkcia náhodnej premennej *X*. Jediné riešenie rovnice $F_X(x) = r$ je hodnota *x* a jediná hodnota *r* prislúcha hodnote *x*. Potom, podľa [3], platí, že $\ddot{x}_i = F_X^{-1}(r_i)$, kde $R \sim Unif(0;1)$ a r_i je realizáciou náhodnej premennej R. A ak napríklad uvažujeme situáciu, že $X \sim Exp(\delta)$ a $F_X(x) = 1 - e^{-\delta \cdot x}$, potom dostávame

$$r_1 = 1 - e^{-\delta \cdot \ddot{x}_1} \Rightarrow \ddot{x}_1 = -\frac{ln(1-r_1)}{\delta}.$$

V jazyku R by bola syntax pre $\delta = 0{,}1$ nasledovná

```
n<-100000; r<-runif(n)
delta<-0.1; x<--(log(1-r))/delta
```

pričom riešenie (napr. $E(X)$) zodpovedá syntaxe funkcie `rexp(n,delta)`.

Ak by sme teda metodicky chceli prepojiť informácie uvedené v predchádzajúcich častiach a nadviazať na využitie simulácií prostredníctvom jazyka R, boli by podstatné tieto informácie:

- mieru rizika *VaR* je možné odhadnúť jednoducho ako kvantil príslušného rozdelenia náhodnej premennej (najčastejšie ako $x_{0,95}$, resp. $x_{0,99}$),
- ak poznáme rozdelenie náhodnej premennej *X* (individuálna výška škody) s príslušnými parametrami, môžeme vyjadriť *VaR* pomocou simulácií a funkcie `r***` pre akékoľvek v R dostupné spojité rozdelenie pravdepodobnosti a funkcie `quantile`,
- ak poznáme rozdelenie počtu škôd *N* a rozdelenie individuálnej výšky škody náhodnej premennej *X* s príslušnými parametrami, môžeme vyjadriť *VaR* pre náhodnú premennú celková škoda *S* pomocou simulácií (`r***`) a funkcie `replicate` (replikovanie, opakovanie) pre v R dostupné kombinácie primárneho a sekundárneho rozdelenia (ďalej využívame funkciu `quantile`),
- mieru rizika *CVaR* je často problematické vyjadriť pre „náročné“ tvary distribučných funkcií,
- odhad *CVaR* mierne uľahčuje využitie vlastností exponenciálnych disperzných tried rozdelení, resp. môžeme využiť techniku simulácií Monte Carlo napr. prostredníctvom jazyka VBA MS Excel, ktorá však si vyžaduje pokročilé programátorské zručnosti a má rôzne obmedzenia, príp. využiť komerčný softvér,
- ak poznáme rozdelenie náhodnej premennej *X* (individuálna výška škody) s príslušnými parametrami, môžeme vyjadriť *CVaR* pomocou simulácií a funkcie `r***` pre akékoľvek v R dostupné spojité rozdelenie pravdepodobnosti a funkcií `quantile`, `subset` a `mean`.
- ak poznáme rozdelenie počtu škôd *N* a rozdelenie individuálnej výšky škody náhodnej premennej *X* s príslušnými parametrami, môžeme vyjadriť *CVaR* pre náhodnú premennú celková škoda *S* pomocou simulácií (`r***`) a funkcie

`replicate` pre v R dostupné kombinácie primárneho a sekundárneho rozdelenia (ďalej využívame funkciu `quantile`, `subset` a `mean`).

Ak realizujeme dostatočne veľa simulácií, ich využitie poskytuje mimoriadne presné výsledky odhadu mier rizika pre rôzne kombinácie primárneho a sekundárneho rozdelenia pre rôzne zložené rozdelenia. Simulácie v R, ktoré sú na prvý pohľad jednoduché, sú značne efektívne a eliminujú problémy pri výpočte *CVaR*. Ako je z dostupných zdrojov autorovi príspevku známe, využitie funkcie `replicate` prezentovali v oblasti kolektívneho modelu rizika len autori Driscoll – Murphy odvolávajúc sa aj na prácu Meyersa (viac na: https://www.casact.org). Teórii mier rizika sa venuje viacero autorov – v ČR a SR napr. Cipra (finančné a poistné riziká), Horáková (poistné riziká, exaktné výpočty mier rizika a komparácia exaktných výpočtov a simulácií), Mucha (simulácie Monte Carlo v jazyku VBA a ich využitie v analýzach v súvislosti s kolektívnym modelom rizika, [5]), Tichý (finančný sektor, [12] a i.).

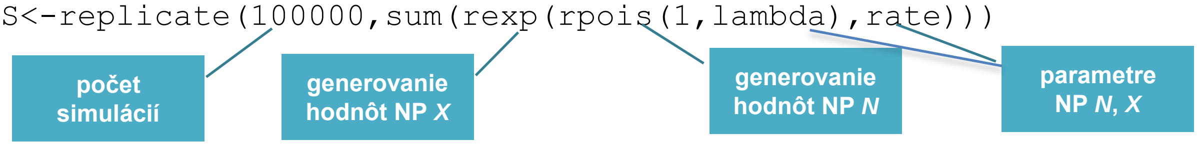
Zavedenie syntaxe funkcie `replicate` na odhad mier rizika *VaR*, *CVaR* náhodnej premennej *S*, ak $N \sim Po(\lambda = 30)$ a $X \sim Exp(\delta = 0{,}1)$, teda ak

$$S \sim CoPo(30;\ X \sim Exp(0{,}1))$$

je nasledovná

```
lambda <-30     # parameter rozdelenia počtu škôd
rate <-0.1      # parameter rozdelenia individuálnej výšky škody

S<-replicate(100000,sum(rexp(rpois(1,lambda),rate)))
```




```
mean(S)    # očakávaná celková škoda (stredná hodnota, E(S))

quantile(S,probs=c(0.95,0.99))          # VaR0.95; VaR0.99

mean(subset(S,S>quantile(S,0.95)))      # CVaR0.95
mean(subset(S,S>quantile(S,0.99)))      # CVaR0.99
```

Vidíme, že nasimulované hodnoty náhodnej premennej *S* sú v jazyku R uložené v rovnomennej premennej `S`. Výstup riešenia uvádza tabuľka č. 1.

Tabuľka č. 1: Výstup riešenia, ak $S \sim CoPo(30;\ X \sim Exp(0{,}1))$; $n = 100\,000$

p	$E(S)$	$VaR_p(S)$	$CVaR_p(S)$	EC^{VaR}	EC^{CVaR}
0,95	299,7721	435,1814	475,7283	135,4093	175,9562
0,99	299,7721	501,7717	536,3991	201,9996	236,6270

Zdroj: vlastné spracovanie

Na komparáciu v tabuľke č. 2 uvádzame aj exaktné riešenie uvedené v [2]. Následne posledný stĺpec v tejto tabuľke je potom porovnaním exaktnej a simulačnej metódy.

Tabuľka č. 2: Výstup riešenia, ak $S \sim CoPo(30;\ X \sim Exp(0{,}1))$

p	$E(S)$	$VaR_p(S)$	$CVaR_p(S)$	EC^{CVaR}	*Chyba aproximácie* EC^{CVaR}
0,95	300	435,4290	476,1157	176,1157	**0,1595**
0,99	300	501,5590	536,6592	236,6592	**0,0322**

Zdroj: [2], upravené

Exaktnú strednú hodnotu náhodnej premennej *S* vypočítame ako

$$E(S) = E(N) \cdot E(X) = 30 \cdot \left(\frac{1}{0{,}1}\right) = 300.$$

V tabuľke č. 3 opíšeme postup na odhad mier rizika na odhad mier rizika *VaR*, *CVaR* pre rôzne kombinácie rozdelení počtu a výšky škody, pričom získané výsledky boli konfrontované aj s exaktným riešením napr. podľa príkladov obsiahnutých v [3], [9].

Tabuľka č. 3: Simulácie hodnôt NP *N*, *X*, *S*, *IN* ***pre rôzne primárne a sekundárne rozdelenia***

Rozdelenie NP N	**Rozdelenie NP** X	**R kód na výpočet (pozri komentár #)**

$N \sim Po(0{,}123)$	-	`lambda<-0.123`
		`N<-rpois(7691,lambda)`
		`mean(N) # E(N)`
		`table(N) # kontingenčná tabuľka`

-	$X \sim N(20;10)$	`mean<-20; sd<-10`
		`X<-rnorm(100000,mean,sd)`
		`quantile(X,0.995) # VaR0.995`
		`# CVaR0.995`
		`mean(subset(X,X>quantile(X,0.995)))`

-	$X \sim \Gamma(10;2)$	`shape<-10`
		`rate<-2`
		`X<-rgamma(100000,shape,rate)`
		`# CVaR0.99`
		`mean(subset(X,X>quantile(X,0.99)))`

-	$X \sim Pa(3;10)$	`library(actuar)`
		`alpha<-3`
		`delta<-10`
		`X<-rpareto(100000,alpha,delta)`
		`# CVaR0.99`
		`mean(subset(X,X>quantile(X,0.99)))`

-	$X \sim IG(20; 80)$	library(statmod) mean<-20 disp<-80 X<-rinvgauss(100000,mean,disp) # $CVaR_{0.995}$ mean(subset(X,X>quantile(X,0.995)))
$N \sim Po(3)$ $M \sim Po(1)$	-	lambda1<-3 lambda2<-1 S<-replicate(10000,sum(rpois(rpois(1, lambda1),lambda2))) mean(S) # E(S) quantile(S,0.945) # $x_{0.945}$
$N \sim Po(10)$	$X \sim Exp(1)$	lambda<-10 rate<-1 S<-replicate(30000,sum(rexp(rpois(1, lambda),rate))) mean(S) # E(S) quantile(S,0.999) # $x_{0.999}$ = VaR
$N \sim Ge(0{,}8)$	$X \sim Exp(1/6)$	prob<-0.8 rate<-1/6 S<-replicate(30000,sum(rexp(rgeom(1, prob),rate))) quantile(S,probs=c(0.8,0.897, 0.947,0.999)) # VaR_p

$N \sim Bi(1000;$
$0{,}15)$
	$X \sim \Gamma(100;$
	$0{,}02)$
		size<-1000; prob<-0.15
		shape<-100; rate<-0.02
		S<-
		replicate(100000,sum(rgamma(rbinom(1,
		size,prob),shape,rate)))
		mean(S) # E(S)
		sd(S) # σ(S)

$N \sim Po(100)$	$X \sim \Gamma(5; 2)$	lambda<-100 shape<-5; rate<-2 S<-replicate(10000,sum(rgamma(rpois(1, lambda),shape,rate))) mean(S) # E(S) quantile(S,0.95) # $VaR_{0.95}$
$N \sim Po(98{,}25)$	$X \sim W(1{,}9; 0{,}615)$	lambda<-98.25 shape<-1.9; scale<-1/0.615 X<-rweibull(10000,shape,scale) mean(X) # E(X) = 1.44 S<- replicate(10000,sum(rweibull(rpois(1, lambda),shape,scale))) mean(S) quantile(S,0.99) # $VaR_{0.99}$

$N \sim Po(73)$	$X \sim LN(4{,}938; 0{,}837)$	lambda<-73 meanlog<-4.938; sdlog<-0.837 S<-replicate(50000,sum(rlnorm(rpois(1, lambda),meanlog,sdlog))) mean(S) # E(S) quantile(S,0.995) # $\text{VaR}_{0.995}$ # $\text{CVaR}_{0.995}$ mean(subset(S,S>quantile(S,0.995)))

$N \sim Po(200)$	$X \sim Pa(5; 1000)$	library(actuar)
		lambda<-200
		alpha<-5; delta<-1000
		S<-
		replicate(30000,sum(rpareto(rpois(1, lambda),alpha,delta)))
		mean(S) # E(S)
		sd(S) # σ(S)

$N \sim Po(2)$	$X \sim IG(1180{,}6347; 659{,}70608)$	library(statmod) lambda<-2 mean<-1180.6347; disp<-659.70608 S<- replicate(30000,sum(rinvgauss(rpois(1, lambda),mean,disp))) mean(S) # E(S) # VaR_p quantile(S,probs=c(0.2331,0.9903))

Poznámka: Úplný názov jednotlivých pravdepodobnostných rozdelení uvádza [3]. Náhodná premenná IN je náhodná premenná celkového počtu poistných udalostí, t. j. zložené diskrétne rozdelenie, podľa [3] rozdelenie Neyman Typ A.

Zdroj: vlastné spracovanie a vlastné spracovanie podľa [7]

Počet opakovania simulácií je potrebné voliť obozretne v súvislosti s vypočítanou exaktnou strednou hodnotou, pričom $E(S)_{simul} \cong E(S)_{exakt}$ a počet simulácií n s ohľadom na disperziu jednotlivých rozdelení (napr. lognormálneho). Aj napriek vysokej zadanej hodnote n R výsledky poskytuje obratom, najčastejšie od 1 do 60 sekúnd. Ďalej sa odporúča pri tvorbe kódu parametre označovať podľa ich názvu v jazyku R (napr. `shape`, `rate`,..., viac v [7]) a brať do úvahy rozlišovanie veľkých a malých písmen. Taktiež si je potrebné v tabuľke č. 3 všimnúť, že pre niektoré rozdelenia je potrebné nainštalovať príslušnú doplnkovú knižnicu (`library()`), avšak samotný výpočet simulácií prebieha v štandardnom rozhraní jazyka R (verzia ≥ 3.3.0 na výpočty podľa tabuľky č. 3).

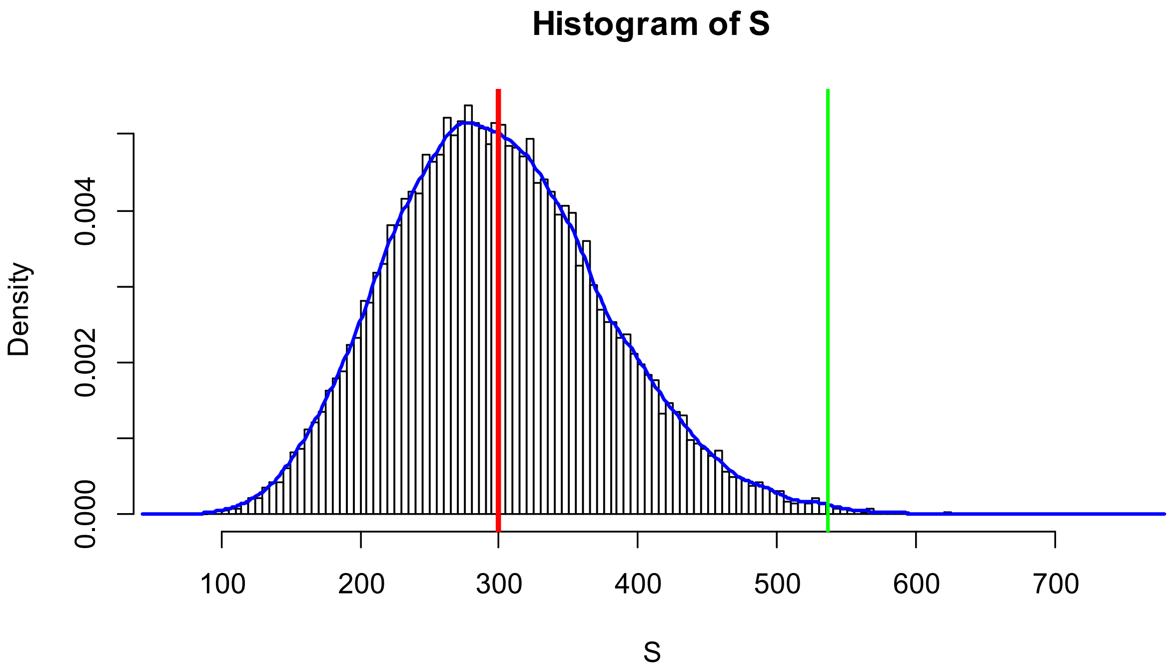
V prípade záujmu môže používateľ generovať histogram nasimulovaných hodnôt náhodnej premennej S, resp. X. Pre prípad uvedený v záhlaví tabuľky č. 1 prostredníctvom kódu

```
hist(S,breaks=200,freq=FALSE)
lines(density(S),col="blue",lwd=2)
abline(v=mean(S),col="red",lwd=3)
abline(v=
```

```
mean(subset(S,S>quantile(S,0.99))),col="green",lwd=2)
```

dostávame výstup, ktorý zobrazuje obrázok č. 1, kde modrá čiara je jadrovou krivkou hustoty náhodnej premennej S, červená priamka zobrazuje na grafe strednú hodnotu $E(S)$ a zelená priamka zobrazuje hodnotu $CVaR_{0,99}(S)$.

Obrázok č. 1: Histogram nasimulovaných hodnôt náhodnej premennej S



Zdroj: vlastné spracovanie

Iný grafický výstup (obrázok č. 2) odhadu jadrovej hustoty so zabudovaným zvýraznením mier rizika môže používateľ získať využitím funkcionality knižnice `ggplot2`.

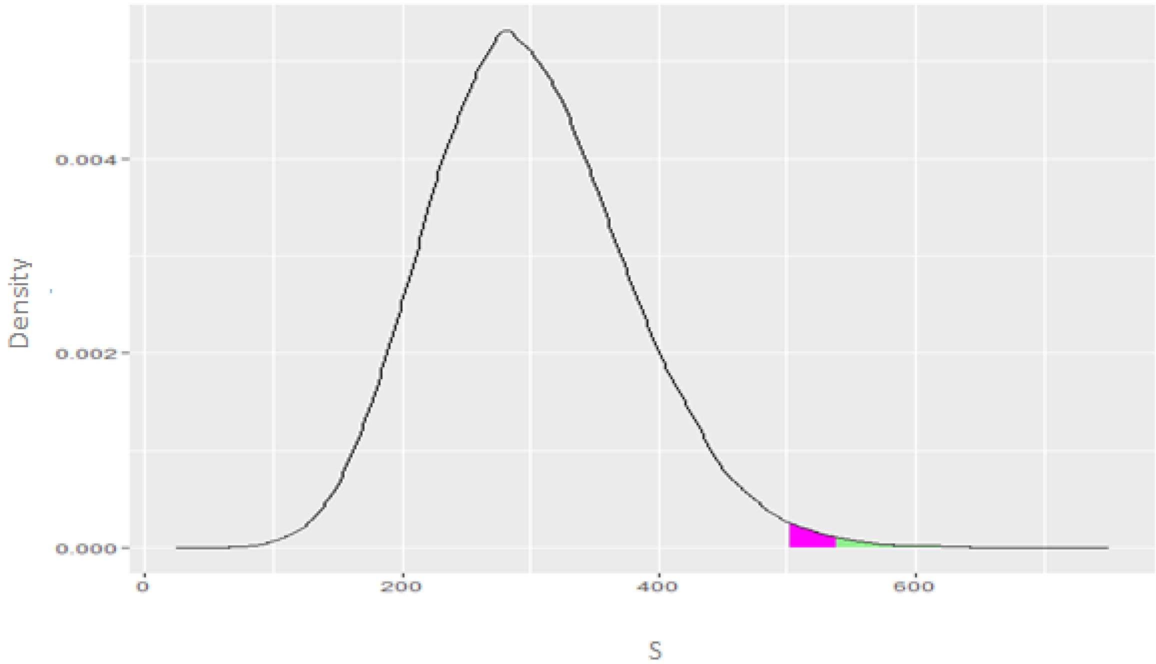
```
library(ggplot2)
VaR<-quantile(S,0.99)
CVaR<-mean(subset(S,S>quantile(S,0.99)))
myd = data.frame(xvar=S,yvar=S)
    xd <- data.frame(density(myd$xvar)[c("x", "y")])
    p <- ggplot(xd, aes(x, y)) +

      geom_area(data = subset(xd, x > VaR), fill = "magenta")
+
      geom_area(data   =   subset(xd,   x   >   CVaR),   fill   =
"lightgreen") +

      geom_line()

    p
```

Obrázok č. 2: Krivka hustoty náhodnej premennej S s mierami rizika



Zdroj: vlastné spracovanie

Možno sa domnievať, že replikovanie bolo v R využité aj v rámci funkcie `aggregateDist` knižnice `actuar` na výpočet mier rizika. Jej konštrukcia (mimo manuálu) je však pre koncového používateľa uzavretá a neposkytuje výstup nasimulovaných údajov. Ukážme teda, že podľa manuálu (pozri napr. prílohu v [7]) ku tejto knižnici a k funkcii *CTE*, *Conditional Tail Expectation* (v spojitom prípade *CTE* ≅ *TVaR* ≅ *CVaR*) môžeme písať

```
library(actuar)
model.freq <- expression(data = rpois(7))
model.sev <- expression(data = rnorm(9, 2))
Fs <- aggregateDist("simulation", model.freq, model.sev,
nb.simul = 100000)
CTE(Fs)
```

pričom rovnaký odhad pre $CTE_{0,99}$ získame aj pomocou nami uvedenej funkcie

```
S<-replicate(100000,sum(rnorm(rpois(1,lambda=7),mean=9,sd=2)))
mean(subset(S,S>quantile(S,0.99)))
```

A teda výhodou tejto metódy tiež je, že používateľ vie takisto ďalej pracovať s nasimulovanými údajmi náhodnej premennej *S* a tieto využiť pre ďalšie analýzy, napr. pre analýzu extrémnych hodnôt (Extreme Value Theory).

5. ZÁVER

Solvency II poskytuje metodiky na výpočet ekonomického kapitálu potrebného na krytie neočakávaných škôd, pri ktorých odporúča zohľadniť aj stochastický prístup. V príspevku sme využili kvantifikáciu poistného rizika pomocou pravdepodobnostných rozdelení. Rozdelenia opisujúce riziká môže poisťovňa získať na základe interných údajov, z ktorých sa vytvorí príslušné zložené rozdelenie agregovanej škody. Jeho znalosť potom umožňuje analyzovať riziko a prispieva k vyhnutiu sa situáciám, že príjmy poisťovne budú nedostatočné na krytie prevzatých záväzkov. Väčšina najčastejšie využívaných mier rizika je založená na využití náhodných veličín opisujúcich pravdepodobnostné rozdelenie strát (škôd) portfólia za určitý čas, teda vedie aj k vyjadreniu mier rizika *VaR* a *CVaR* v poistnej praxi. Stanovenie nielen maximálnej možnej straty (*VaR*), ale aj na základe nej odvodenej miery rizika *CVaR*, teda podmienenej strednej hodnoty pravých koncov rozdelení, je v aktuárskej praxi vysoko aktuálna. Táto hodnota predstavuje očakávanú stratu, ktorá môže byť spôsobená v danom časovom období zo všetkých škôd presahujúcich príslušnú hodnotu *VaR* s konkrétnou pravdepodobnosťou. Aj podľa predchádzajúcich kvantitatívnych dopadových štúdií Solvency II (QIS 5) ide o dôležitú mieru pri kvantifikácii rizika. Preto by poisťovňa mala byť schopná jednak s vysokou presnosťou správne určiť príslušné rozdelenie pravdepodobnosti a jeho distribučnú funkciu a jednak s požadovanou úrovňou spoľahlivosti stanoviť miery rizika a príslušný ekonomický kapitál na krytie prevzatého rizika.

Miery rizika by mali spĺňať štyri vlastnosti: *subaditivitu*, *monotónnosť*, *homogenitu* a *invariantnosť*. Hodnota *CVaR* je miera rizika, ktorá uvedené vlastnosti má, no jej stanovenie nie je v niektorých prípadoch jednoduché. H. H. Panjerom boli vyvinuté postupy na jej pomerne jednoduché určenie pri niektorých rozdeleniach. Rozšírené výsledky týchto postupov sú založené na využití širokej triedy exponenciálnych disperzných modelov.

Hodnoty *VaR* a *CVaR* v tomto prípade vyjadrujeme pomocou parametrov konkrétnych rozdelení, ktoré zapadajú do koncepcie kolektívneho modelu rizika. Prístup simulácií, v rámci tohto procesu, môže byť efektívne využitý v oblasti eliminácie náročných postupov, ktoré súvisia s exaktnými metódami stanovenia zloženého rozdelenia celkovej škody. Tieto sa potom dajú využiť v aktuárstve nielen na opísanie rizika individuálnej škody, počtu škôd ale aj na zložené rozdelenia celkovej škody, čo je v prípade spojitej náhodnej premennej s existujúcou strednou hodnotou kľúčové pri odhade ekonomického kapitálu. Simulácie sú potom výraznou možnosťou získania hodnôt kvantilov zložených rozdelení. Hodnoty konkrétnych kvantilov získané pomocou simulácií sú nielen porovnateľné, ale sú zastupiteľnou alternatívou pre hodnoty kvantilov získaných exaktnou metódou a to aj v prípade založenom na vlastnostiach exponenciálnych disperzných modelov.

Ako výpočtové prostredie sme v našom prípade zvolili open-source softvér R ako efektívny nástroj (mnohokrát prevyšujúci možnosti komerčných softvérov) pre generovanie pseudonáhodných čísiel, pričom v kolektívnom modeli rizika sme využili funkciu `replicate`. Ide o veľmi účinný a sofistikovaný prístup k simulovaniu hodnôt celkovej škody bez zložitejšieho programovania, čo môže zvýšiť záujem používateľov o tento softvér. Autor zavádza použitie tejto funkcie jazyka R v našom prostredí v súvislosti s publikáciou [7]. Čitateľ sa v tomto príspevku môže podrobne, bez vynechania medzikrokov, zoznámiť s celou problematikou odhadu mier rizika prostredníctvom simulácií s vysvetlením ad hoc postupu v jazyku R.

LITERATÚRA

[1] CIPRA, T.: Riziko ve financích a pojišťovnictví: Basel III a Solvency II. Praha: Ekopress, 2015. ISBN 978-80-87865-24-8.

[2] HORÁKOVÁ, G.: Odhad hodnoty CVaR a jej využitie pri riadení poistných rizík. In: Řízení a modelování finančních rizik – 6. mezinárodní vědecká konference. Ostrava: VŠB – TU Ostrava, 2012.

[3] HORÁKOVÁ, G. – PÁLEŠ, M. – SLANINKA, F.: Teória rizika v poistení. Bratislava: Wolters Kluwer, 2015. ISBN 978-80-8168-273-5.

[4] KLUGMAN, S. A. – PANJER, H. H. – WILLMOT, G. E.: Loss Models (From Data to Decision). New York: John Wiley & Sons, 2012. ISBN 978-1-118-31532-3.

[5] MUCHA, V.: Simulácie ako nástroj riadenia rizika v neživotnom poistení. In: Řízení a modelování finančních rizik – 4. mezinárodní vědecká konference. Ostrava: VŠB – TU Ostrava, 2008.

[6] PÁLEŠ, M.: Aktuárstvo v režime Solventnosť II (S riešenými príkladmi v jazyku R). Bratislava: Vydavateľstvo EKONÓM, 2016. ISBN 978-80-225-4288-3.

[7] PÁLEŠ, M.: Jazyk R v aktuárskych analýzach. Bratislava: Vydavateľstvo EKONÓM, 2017. ISBN 978-80-225-4331-6.

[8] PÁLEŠ, M.: Využitie kopula funkcií pri agregácií rizík. In: Slovenská štatistika a demografia, 2017, č. 1. s. 13 – 22.

[9] PÁLEŠ, M. – POLÁČEK, Š.: Softvérová podpora pri modelovaní rozdelenia celkovej škody v havarijnom poistení. In: Slovenská štatistika a demografia, 2012. č. 4, s. 50 – 58.

[10] R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2017.

[11] SLANINKA, F. – KADEROVÁ, A. – SIMONKA, ZS.: Risk Modeling and Analysis in ModelRisk Software. In: Řízení a modelování finančních rizik – 8. mezinárodní vědecká konference. Ostrava: VŠB – TU Ostrava, 2016.

[12] ZMEŠKAL, Z. – DLUHOŠOVÁ, D. – TICHÝ, T.: Finanční modely. Praha: Ekopress, 2013. ISBN 978-80-86929-91-0.

RESUME

Economic capital modeling is one of the fundamental components of risk management in the insurance company and is a tool for the actuary to protect the insuran-ce company against unexpected risks and also losses. The company must analyse the risks that are relevant to it and by means of models for the determination of the economic capital, to identify the risks which can threaten its profitability. In 2016, the Solvency II methodology entered into force – a project for the regulation of insurance companies, representing a systematic approach to risk management that results in a better risk evaluation and thus guaranteeing a greater protection of the insured persons. Solvency II provides several methods for calculating the capital requirement that can be selected by the insurance companies, having regard to their scale and complexity. The risk management function carried out by the actuary of the insurance company is closely connected with this activity. By using the R language and simulations, the actuary can successfully estimate the risk measures and thus appropriately determine the economic capital.

PROFESIJNÝ ŽIVOTOPIS

Ing. Michal Páleš, PhD., *od roku 2012 pôsobí ako odborný asistent a tajomník Katedry matematiky a aktuárstva Fakulty hospodárskej informatiky Ekonomickej univerzity v Bratislave. V rámci pedagogickej činnosti vyučuje predmety matematika, teória pravdepodobnosti, softvérové aplikácie pre aktuárov, teória rizika v poistení, úvod do aktuárstva a vybrané* *kapitoly z matematiky. Vo svojej vedeckej práci sa orientuje na aktuársku vedu, využitie kvantitatívnych metód v ekonómii a softvérovú podporu riadenia rizík. Je autorom viacerých vysokoškolských učebníc a vedeckých článkov z oblasti aktuárstva.*

KONTAKT
pales.euba@gmail.com