

Metóda Bootstrap pre odhad hodnôt LIC v štandarde IFRS17

Jana Tušimová, Tatiana Šoltéssová

Abstrakt

Metóda bootstrap je štatistický nástroj, ktorý získava na popularite pri riešení problémov. V kontexte výpočtu LIC podľa štandardu IFRS17 predstavuje bootstrap efektívnu techniku na simuláciu možných budúcich scenárov na základe historických údajov. Cieľom tohto článku je priblížiť možnosť využitia metódy bootstrap pri výpočte LIC.

Kľúčové slová

metóda bootstrap, jazyk R, LIC, IFRS17

1 ÚVOD

Liability for incurred claims (LIC), teda záväzky z už vzniknutých poistných udalostí, je jedným z kľúčových pojmov v štandarde IFRS17, ktorý nahradil predchádzajúci štandard IFRS4. Výpočet a implementácia LIC v rámci IFRS17 kladie vysoké nároky na poisťovne, najmä v oblasti modelovania a kvality a rozsahu dát.

LIC zahŕňa dva druhy poistných udalostí:

- a) poistné udalosti vzniknuté, nahlásené a k dátumu reportovania nevysporiadané (poistné udalosti, ktoré sú pre poisťovňu známe),
- b) poistné udalosti vzniknuté a k dátumu reportovania ešte nenahlásené (poistné udalosti pre poisťovňu neznáme).

Práve výšku poistných udalostí vzniknutých a nenahlásených potrebuje poisťovňa v pravidelných intervaloch odhadnúť. Podľa povahy a historického rozsahu modelovaných dát sa na výpočet tejto časti LIC môže využívať stochastická bootstrap metóda. Ide o štatistickú metódu, ktorá umožňuje pomocou opakovaného náhodného prevzorkovania odhadovať vlastnosti štatistického rozdelenia dát. Bootstrap bol v roku 1979 predstavený Bradleyom Efronom a stal sa významným nástrojom v oblasti analýzy dát a štatistiky. Hlavným zámerom bootstrapu je vytvoriť z pôvodnej dátovej sady veľké množstvo vzoriek s opakovaným náhodným výberom s vrátením. Následne na základe vypočítaných štatistických parametrov umožňuje odhadovať vlastnosti skúmaných dát. Použitie bootstrapu v IFRS17 je preto neoceniteľným nástrojom pre poisťné spoločnosti, ktoré sa snažia zabezpečiť vysokú presnosť a transparentnosť pri vykazovaní svojej finančnej situácie.

R je rozšírený voľne dostupný softvér využívaný na štatistickú analýzu a vizualizáciu dát. Poskytuje robustnú podporu pre implementáciu bootstrap metódy, s použitím niekoľkých knižníc, ktoré zjednodušujú použitie bootstrapu.

Pre výpočet odhadovanej časti LIC pomocou metódy bootstrap použijeme knižnicu „ChainLadder“ v softvéri R, ktorý obsahuje funkciu `BootChainLadder`. Argumentami tejto funkcie sú:

- trojuholník kumulatívnych poistných plnení,
- počet opakovaní bootstrapu,
- určenie pravdepodobnostného rozdelenia.

Výstupom tejto funkcie sú nasledujúce údaje:

- `Latest` – plnenia vyplatené za sledované obdobia (v nekumulovanej podobe),
- `Mean Ultimate` – celková očakávaná výška poistných plnení,
- `Mean IBNR` – odhad výšky záväzkov na nenahlásené poistné udalosti,
- `SD IBNR` – štandardná odchýlka odhadu výšky záväzkov na nenahlásené poistné udalosti,
- `Total IBNR` – odhad výšky záväzkov na nenahlásené poistné udalosti pri stanovenej spoľahlivosti odhadu.

2 VÝPOČET ODHADOVANÝCH HODNÔT LIC

Ako podklad na výpočet záväzkov na nenahlásené poistné udalosti slúžia údaje o poistných udalostiach, ktoré boli nahlásené a vyplatené a nahlásené a nevysporiadané. Tie sú následne sčítané, usporiadané do trojuholníka na základe roku vzniku poistnej udalosti a roku hlásenia poistnej udalosti. Pre modelový prípad životného poistenia sme si zvolili fiktívne hodnoty v trojuholníku, ktoré sme ručne zadali prostredníctvom kódu. Počet simulácií zvolíme na 10 000.

Kód jazyka R

```
install.packages("ChainLadder") #1
install.packages("openxlsx") #2
library(ChainLadder) #3
library(openxlsx) #4
data <- (matrix(c(100, 150, 175, 180, 200, #5
                  110, 168, 192, 205, NA,
                  115, 169, 202, NA, NA,
                  125, 185, NA, NA, NA,
                  150, NA, NA, NA, NA), nrow = 5, byrow = TRUE))
rozdelenie <- "gamma" #6
sim <- 10000 #7
data_cum <- incr2cum(data) #8

plot(data_cum/100000, main = "Vývoj škôd po rokoch vzniku") #9 Graf 1

plot(data_cum/100000, lattice = TRUE, main = "Vývoj škôd po rokoch vzniku") #9 Graf 2

set.seed(1) #10
data_bootstrap1 <- BootChainLadder(data_cum, R = sim, process.distr =
as.character(rozdelenie)) #11
data_bootstrap2 <- quantile(data_bootstrap1, c(0.75, 0.8, 0.85, 0.9, 0.95, 0.975)) #12
plot(data_bootstrap1) #13 Graf 3
```

```

stlpec <- c("Latest", "Mean Ultimate", "Mean IBNR", "SD IBNR", "Total IBNR 75 %",
"Total IBNR 95 %") #14
ROK <- c("1. rok", "2. rok", "3. rok", "4. rok", "5. rok", "SPOLU") #15
data_vystup1 <- summary(data_bootstrap1) #16
data_vystup2 <- data_bootstrap2 #17
data_tab1 <- data.frame(stlpec, data_vystup1$Totals) #18
data_tab11 <- data.frame(data_vystup1$ByOrigin) #19
sum1 <- c(sum(data_tab11[1]), sum(data_tab11[2]), sum(data_tab11[3]),
sum(data_tab11[4]), sum(data_tab11[5]), sum(data_tab11[6])) #20
data_tab11 <- rbind(data_tab11, sum1) #21
data_tab11 <- data.frame(ROK, data_tab11) #22
data_tab2 <- data.frame(stlpec, data_vystup2$Totals) #23
data_tab22 <- data.frame(data_vystup2$ByOrigin) #24
sum2 <- c(sum(data_tab22[1]), sum(data_tab22[2]), sum(data_tab22[3]),
sum(data_tab22[4]), sum(data_tab22[5]), sum(data_tab22[6])) #25
data_tab22 <- rbind(data_tab22, sum2) #26
data_tab22 <- data.frame(ROK, data_tab22) #27

```

```

#1 inštalácia balíka "ChainLadder"
#2 inštalácia balíka "openxlsx"
#3 načítanie balíka "ChainLadder"
#4 načítanie balíka "openxlsx"
#5 vytvorenie vzorových dát
#6 nastavenie rozdelenia gamma
#7 nastavenie počtu simulácii
#8 vytvorenie kumulatívneho trojuholníka
#9 grafické znázornenie vývoja škôd po rokoch vzniku
#10 nastavenie stability generovanie výstupov
#11 vykonanie bootstrapu
#12 bootstrap pre rôzne kvantily
#13 grafické znázornenie výsledkov
#14 označenie výsledkov
#15 označenie rokov
#16 uloženie prvého výstupu bootstrapu do novej premennej
#17 uloženie druhého výstupu (kvantily) do novej premennej
#18 uloženie výsledkov bootstrapu do tabuľky
#19 uloženie výsledkov bootstrapu do tabuľky
#20 súčet jednotlivých stĺpcov
#21 pridanie súčtového riadku do tabuľky
#22 pridanie označenia rokov do tabuľky
#23 uloženie výsledkov bootstrapu do tabuľky
#24 uloženie výsledkov bootstrapu do tabuľky
#25 súčet jednotlivých stĺpcov
#26 pridanie súčtového riadku do tabuľky
#27 pridanie označenia rokov do tabuľky

```

Test stability

```

simulacie <- c("5 000", "10 000", "15 000", "20 000", "30 000")
set.seed(1)

```

```

data_bootstrapsim1 <- summary(BootChainLadder(data_cum, R = 5000,
process.distr = as.character(rozdelenie))) #28
IBNRsim1 <- data_bootstrapsim1$Totals[6,] #29
set.seed(1)
data_bootstrapsim2 <- summary(BootChainLadder(data_cum, R = 10000,
process.distr = as.character(rozdelenie))) #30
IBNRsim2 <- data_bootstrapsim2$Totals[6,] #31
set.seed(1)
data_bootstrapsim3 <- summary(BootChainLadder(data_cum, R = 15000,
process.distr = as.character(rozdelenie))) #32
IBNRsim3 <- data_bootstrapsim3$Totals[6,] #33
set.seed(1)
data_bootstrapsim4 <- summary(BootChainLadder(data_cum, R = 20000,
process.distr = as.character(rozdelenie))) #34
IBNRsim4 <- data_bootstrapsim4$Totals[6,] #35
set.seed(1)
data_bootstrapsim5 <- summary(BootChainLadder(data_cum, R = 30000,
process.distr = as.character(rozdelenie))) #36
IBNRsim5 <- data_bootstrapsim5$Totals[6,] #37
vysledkysim <- data.frame(simulacie, IBNR = c(IBNRsim1, IBNRsim2,
IBNRsim3, IBNRsim4, IBNRsim5)) #38

```

Funkcia **set.seed(1)** v jazyku R zabezpečuje, že konkrétna sekvencia náhodných čísel, ktorá bola vygenerovaná, po opakovanom spustení s kódom s rovnakým seed (v tomto prípade so seed 1), vygeneruje rovnaký výsledok ako pri prvom spustení. Táto funkcia je užitočná pre poisťovne, ktoré chcú zabezpečiť reprodukovateľnosť výsledkov (napr. pre audítarov).

```

#28 vykonanie bootstrapu pre sim1
#29 uloženie výsledku 95%-ného kvantilu pre sim1
#30 vykonanie bootstrapu pre sim2
#31 uloženie výsledku 95%-ného kvantilu pre sim2
#32 vykonanie bootstrapu pre sim3
#33 uloženie výsledku 95%-ného kvantilu pre sim3
#34 vykonanie bootstrapu pre sim4
#35 uloženie výsledku 95%-ného kvantilu pre sim4
#36 vykonanie bootstrapu pre sim5
#37 uloženie výsledku 95%-ného kvantilu pre sim5
#38 uloženie výsledkov simulácii do tabuľky

```

Výsledky

```

myfile <- createWorkbook() #39
addWorksheet(myfile, sheetName = "vysledky") #40
writeDataTable(myfile, sheet = 1, x = data_tab11, startCol = 1, startRow = 1,
tableStyle = "TableStyleLight16") #41
writeDataTable(myfile, sheet = 1, x = data_tab1, startCol = 1, startRow = 20,
tableStyle = "TableStyleLight16") #42
writeDataTable(myfile, sheet = 1, x = data_tab22, startCol = 10, startRow = 1,
tableStyle = "TableStyleLight17") #43
writeDataTable(myfile, sheet = 1, x = data_tab2, startCol = 10, startRow = 20,

```

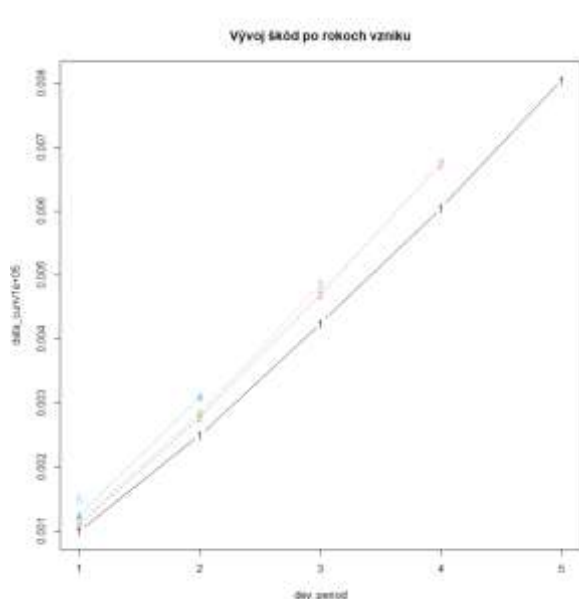
```

tableStyle = "TableStyleLight17") #44
addWorksheet(myfile, sheetName = "stabilita") #45
writeDataTable(myfile, sheet = 2, x = vysledkysim, startCol = 1, startRow = 1,
tableStyle = "TableStyleLight18") #46
saveWorkbook(myfile, file = paste(getwd(), '/vysledky_bootstrapu.xlsx', sep = ""),
overwrite = TRUE) #47
#39 vytvorenie excelovského súboru s výsledkami
#40 vytvorenie hárka "vysledky"
#41 uloženie výsledkov na Hárak1
#42 uloženie výsledkov na Hárak1
#43 uloženie výsledkov na Hárak1
#44 uloženie výsledkov na Hárak1
#45 vytvorenie hárka "stabilita"
#46 uloženie výsledkov na Hárak2
#47 export výsledkov do xlsx

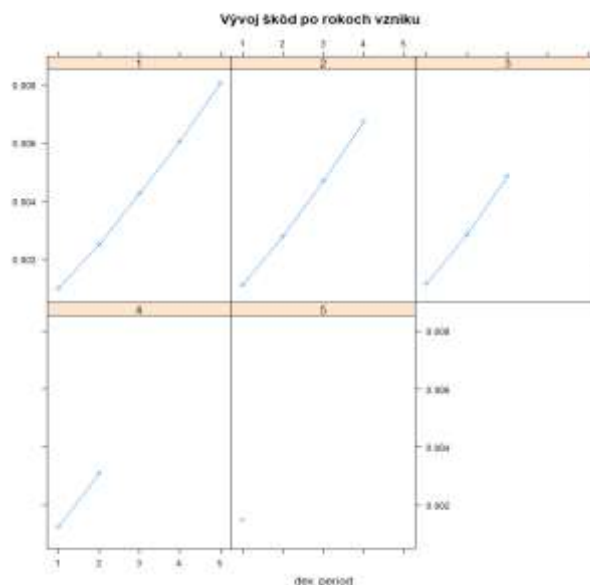
```

3 GRAFICKÉ ZOBRAZENIE

Príkazom #9 *Graf 1* graficky zobrazíme vývoj výšky kumulatívnych poistných plnení po rokoch vzniku. Netreba zabúdať, že v článku pracujeme s fiktívnymi dátami. Každá poisťovňa by mala k dispozícii obširnejšie dáta, s väčšou históriou a iným vývojom v priebehu rokov, čím by dáta zobrazené na grafe 1 nemali takmer totožný priebeh.



Graf 1: Výstup z R: Vývoj škôd po rokoch vzniku

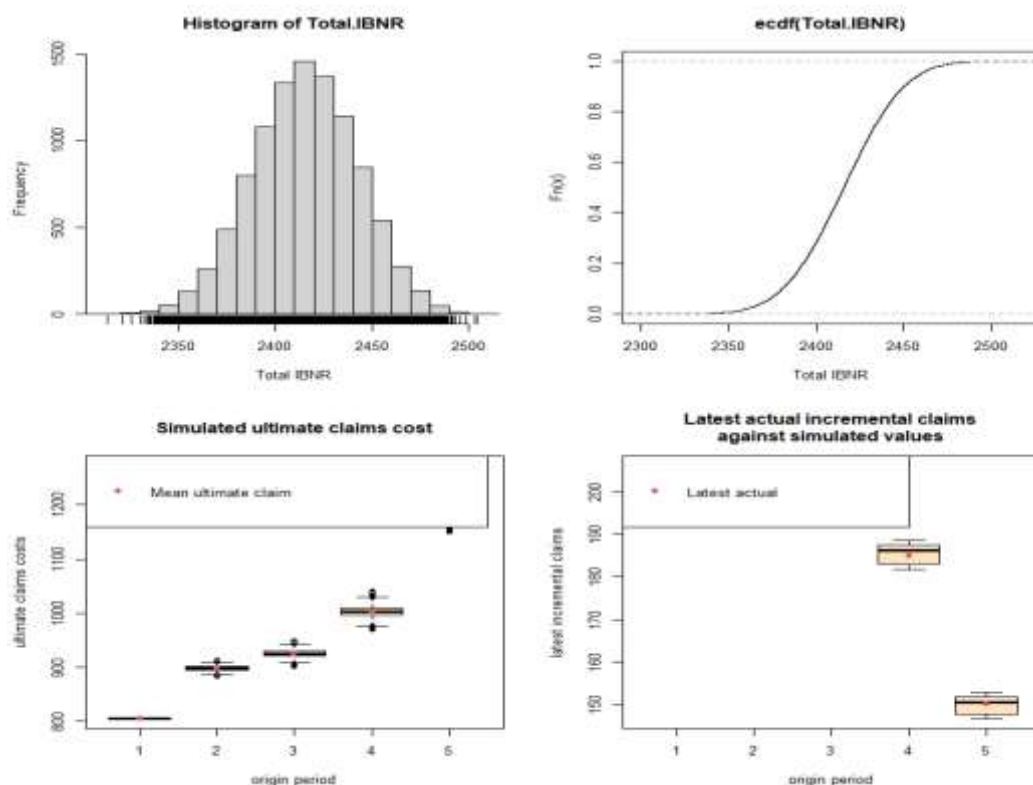


Graf 2: Výstup z R: Vývoj škôd samostatne po rokoch vzniku

Graf 2 zobrazíme príkazom #9 *Graf 2*. Tento graf nám vykresľuje vývoj výšky kumulatívnych poistných plnení podľa roku vzniku poistnej udalosti.

Príkazom #13 *Graf 3* zobrazíme získané výsledky. Na grafe 3 sa nachádza histogram rozdelenia celkových odhadov LIC (poistných udalostí vzniknutých, ale ešte nenahlásených). Ďalej na grafe 3 nájdeme graf distribučnej funkcie a dve sústavy boxplotov. Ľavý boxplot

znázorňuje predpoklad vývoja výšky celkových poistných plnení, pričom pravý boxplot prezentuje hodnoty najaktuálnejších inkrementálnych plnení oproti bootstrapovým variantom.



Graf 3: Výstup z R: Grafické zobrazenie výsledkov

Odhadnuté hodnoty LIC je potrebné následne rozalokovať pre jednotlivé Unit of accounts, ako aj určiť ich cashflowy vývoja. Tie sú určené na základe podrobných analýz vývoja poistných udalostí pre konkrétne produktové portfóliá.

4 ZÁVER

IFRS17, ako globálny štandard pre vykazovanie poistných zmlúv, vyžaduje od poistných spoločností presné modelovanie budúcich peňažných tokov, ktoré sa pri použití bootstrapu stávajú flexibilným a robustným nástrojom. Z toho dôvodu je bootstrap metóda široko využívaná nie len v štatistickej, ale práve aj v aktuárskej praxi. Softvér R poskytuje bohaté možnosti na jej implementáciu, vďaka čomu sa stáva táto metóda obľúbenou pre veľké spektrum odborníkov. Netreba zabúdať, že nie je dôležité iba samotné spracovanie dát pomocou tejto metódy, ale aj správna interpretácia výsledkov.

POUŽITÉ ZDROJE

- [1] R Core Team (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- [2] Schauberger, P. & Walker, A. (2022). openxlsx: Read, Write and Edit xlsx Files. R package version 4.2.5.1. <https://CRAN.R-project.org/package=openxlsx>.

- [3] Markus Gesmann, Daniel Murphy, Yanwei (Wayne) Zhang, Alessandro Carrato, Mario Wuthrich, Fabio Concina & Eric Dal Moro (2024). ChainLadder: Statistical Methods and Models for Claims Reserving in General Insurance. R package version 0.2.19. <https://CRAN.R-project.org/package=ChainLadder>.
- [4] Efron, B., & Tibshirani, R. J. (1994). An Introduction to the Bootstrap. Chapman & Hall/CRC.
- [5] Davison, A. C., & Hinkley, D. V. (1997). Bootstrap Methods and Their Application. Cambridge University Press.
- [6] R Documentation: <https://cran.r-project.org/web/packages/boot/index.html>
- [7] Wickham, H. (2021). Advanced R. Chapman & Hall/CRC.

KONTAKTNÉ ÚDAJE

Tušimová, Jana, Ing., PhD., e-mail: tusimova.janka@gmail.com

Šoltésová, Tatiana, doc., Mgr., PhD., Katedra matematiky a aktuárstva, Fakulta hospodárskej informatiky, Ekonomická univerzita v Bratislave, Dolnozemska cesta 1, 852 35 Bratislava, e-mail: tatiana.soltesova@euba.sk