

Quantification of risks of natural catastrophes

Viera Pacáková, Pavla Jindrová, Tomáš Musil¹

Abstract

Catastrophic events affect various regions of the world with increasing frequency and intensity. The aim of this paper is to describe parametric curve-fitting methods for modelling extreme historical losses of natural catastrophes in the world. Article summarizes relevant theoretical results of Extreme value theory (EVT) and Excess over Threshold Method (EOT) and presents results of their application to the data about amounts of damages in world catastrophe events in time period 2010-2014, published by Swiss Re studies Sigma. Article presents the models for extreme catastrophic losses by selecting particular probability distributions through statistical analysis of empirical data with the best possible estimate of the upper tail area. These models allowed us to quantify the risks of natural and man-made catastrophes by risk measures.

Key words

Catastrophe losses, Block maxima models, Excess over Threshold Method, Extreme value distribution, Generalized Pareto Distribution, Risk measures.

JEL Classification: C4, C6, C8

1. Úvod

V posledních letech roste počet katastrofických událostí a tím rostou i výdaje pojišťoven či zajišťoven na vyplacení pojistného plnění. Proto je důležité tyto události správně modelovat a prognózovat jejich důsledky. Následně budou analyzovány katastrofické události od roku 2010 do roku 2014 z hlediska počtu katastrof, výše ekonomických ztrát a také výše pojistného plnění, které musely pojišťovny a zajišťovny v souvislosti s pojištěnými škodami vyplatit.

Katastrofou můžeme označit událost, která zasáhne velké území přírody nebo společnosti a je důsledkem přírodní či lidské činnosti. Tato událost negativním způsobem mění původní prostředí.

V článku jsou údaje převzata od švýcarské zajišťovací společnosti Swiss Re (SIGMA: insurance research) [6], která každoročně publikuje periodikum s názvem Sigma, ve kterém jsou poskytnuty veřejnosti informace týkající se katastrofických událostí ve světě.

Aby byla událost definována jako katastrofa, musí být podle švýcarské zajišťovny Swiss Re v roce 2014 spojena se škodou minimálně 97 milionů USD nebo pojištěnou škodou 48 mil. USD (v leteckém pojištění 39 mil. USD, v dopravním 19 mil. USD), nebo minimálně 20 osob, které zemřely či se ztratily, nebo 2000 osob, které přišly kvůli této události o střechnu nad hlavou.

¹ prof. RNDr. Viera Pacáková, PhD., Faculty of Economics and Administration, University of Pardubice, e-mail: Viera.Pacakova@upce.cz

Mgr. Pavla Jindrová, Ph.D., Faculty of Economics and Administration, University of Pardubice, Pardubice, e-mail: Pavla.Jindrova@upce.cz

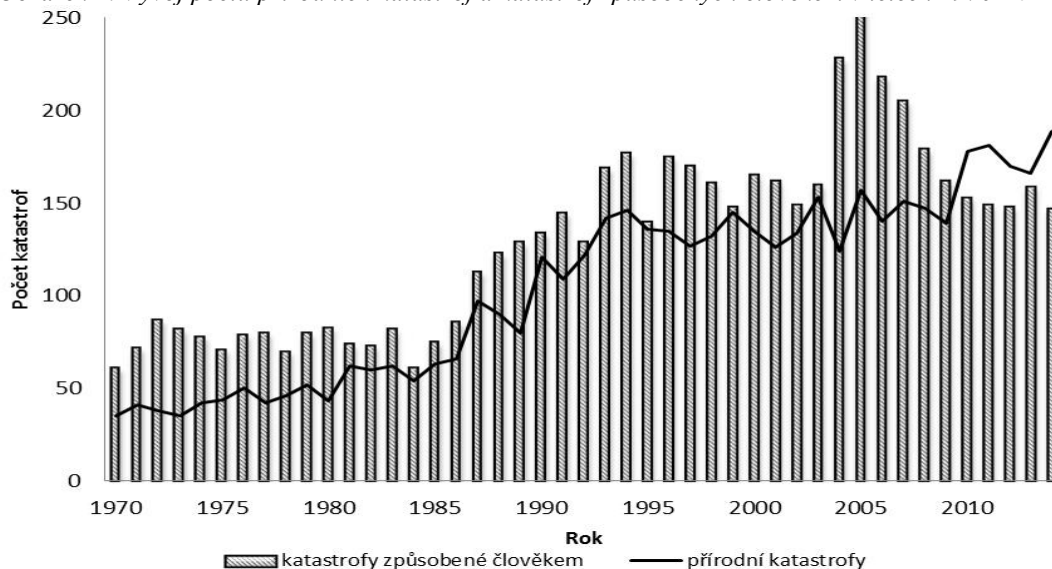
Bc. Tomáš Musil, Faculty of Economics and Administration, University of Pardubice, Pardubice, e-mail: tomas.musil1@student.upce.cz

2. Vývoj počtu a důsledků katastrofických událostí

Počet katastrofických událostí se za posledních 40 let výrazně zvýšil, přibližně o 350 %. Jednou z teorií, kterým lze tento výrazný nárůst výskytu katastrof vysvětlit, je nárůst lidské populace, kdy lidí na planetě je stále více a tak vznikají hustěji osídlená území, díky čemuž dochází v případě katastrofy k daleko větším následkům, než v neobydleném území. Klimatické změny jsou dalším negativním jevem. Především skleníkové plyny, které mají vliv na globální oteplování.

Na obrázku 1 je zobrazen vývoj celkového počtu přírodních i člověkem způsobených katastrof od roku 1970.

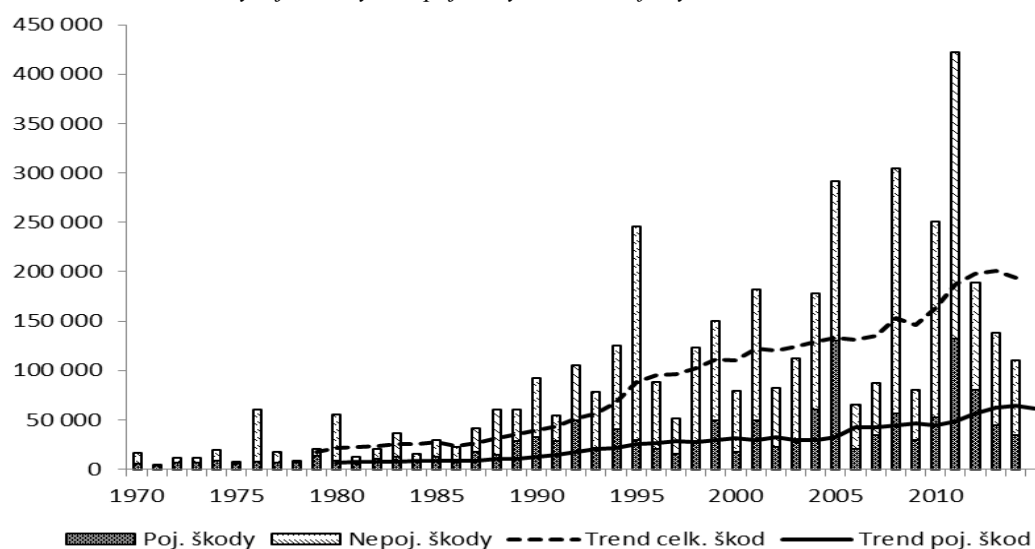
Obrázek 1: Vývoj počtu přírodních katastrof a katastrof způsobených člověkem v letech 1970-2014



Zdroj: vlastní zpracování dle [6]

Obrázek 2 znázorňuje vývoj celkových a pojištěných katastrofických škod a jejich trendy vývoje. Rozdíl mezi těmito trendy se neustále zvyšuje, trend celkových škod roste rychleji než trend pojištěných škod.

Obrázek 2: Vývoj celkových a pojištěných katastrofických škod v letech 1970-2014



Zdroj: vlastní zpracování dle [6]

3. Metody modelování extrémních škod

Výskyt a důsledky katastrofických událostí se v posledních letech neustále zvyšují, a tím se zvyšují i vyplácená pojistná plnění pojišťoven. Pro pojišťovny je důležité tyto události správně modelovat a tím i vhodně zvolit možnost krytí takovýchto rizik.

Možností, jak modelovat extrémní škody je několik. V článku budeme využívat dvě z nich, *metodu blokového maxima* a *metodu excedentů překračujících vysoký práh*.

3.1 Metoda blokového maxima

Metoda blokového maxima (BM) považuje za extrémní hodnoty maximální hodnoty z n bloků po sobě jdoucích nezávislých pozorování $x_1, x_2, x_3, \dots, x_n$ v průběhu zvoleného časového období. Maximum z těchto nezávislých, identicky rozdělených náhodných veličin s distribuční funkcí F dostaneme jako:

$$M_n = \max\{x_1, x_2, x_3, \dots, x_n\} \text{ pro } n \geq 1 \quad (1)$$

Limitní rozdělení blokových maxim je dané Fisher-Tippetovou větou [1], [4]:

Nechť X_1, X_2, \dots, X_n je posloupnost identicky rozdělených náhodných veličin. Jestliže existují normované konstanty $c_n > 0$ a $d_n \in \mathbb{R}$ a existuje nějaká nedegenerovaná distribuční funkce $H(x)$ taková, že:

$$P(c_n^{-1}(M_n - d_n) \leq x) \rightarrow H(x) \quad (2)$$

Poté $H(x)$ pochází z jednoho ze tří typů distribučních funkcí extrémních hodnot, které jsou Gumbelova distribuční funkce, Fréchetova nebo Weibullova distribuční funkce:

- Gumbel: $\Lambda(x) = \exp\{-e^{-x}\}$ pro $x \in \mathbb{R}$
- Fréchet: $\Phi_\alpha(x) = \begin{cases} 0 & \text{pro } x \leq 0, \\ \exp\{-x^{-\alpha}\} & \text{pro } x > 0, \end{cases} \alpha > 0$
- Weibull: $\Psi_\alpha(x) = \begin{cases} 1 & \text{pro } x \leq 0, \\ \exp\{-(-x)^{-\alpha}\} & \text{pro } x > 0, \end{cases} \alpha > 0$

kde α je parametr tvaru funkce.

Gumbelovo, Fréchetovo a Weibullovo rozdělení se nazývají max-stabilní rozdělení. Zachovávají typ rozdělení pro maximum identicky rozdělených náhodných proměnných a jsou speciálním případem zobecněné distribuční funkce extrémních hodnot GEV definované vztahem:

$$H_\xi(x) = \begin{cases} \exp\{-(1 + \xi x)^{-1/\xi}\} & \text{pro } \xi \neq 0, \\ \exp\{e^{-x}\} & \text{pro } \xi = 0, \end{cases} \alpha > 0 \quad (3)$$

V případě praktického využití je vhodnější využít flexibilní rozdělení extrémních hodnot, které je vyjádřeno pomocí tří parametrů μ , σ a ξ [21]:

$$H(x) = \exp\left\{-\left(1 + \xi\left(\frac{x - \mu}{\sigma}\right)\right)^{-1/\xi}\right\} \quad (4)$$

kde μ je parametr polohy (location), σ je parametr rozsahu (scale) a ξ je parametr tvaru (shape).

Nevýhodou metody blokového maxima je to, že uvažuje vždy pouze s jednou hodnotou v bloku, ačkoli by mohla být v bloku i další hodnota velice vysoká. Další z nevýhod metody blokového maxima je to, že vyžaduje poměrně velký počet pozorování, proto se v praxi více používá metoda excedentů přes vysoký práh, který bude popsán v následující kapitole.

3.2 Metoda excedentů přes vysoký práh

Metoda excedentů přes vysoký práh, neboli Peaks over Threshold (POT), je svým modernějším přístupem další metodou, která se využívá pro pravděpodobnostní modelování extrémních hodnot. Tato metoda uvažuje za extrémní všechny hodnoty, které jsou vyšší než zvolený práh u . Zvolení tohoto práhu by mělo být takové, aby excedentů překračující práh bylo poměrně málo. Limitním rozdělením pro excedenty, které překračují práh u , je podle Pickands, Balkema a de Haanovy věty [4] *zobecněné Pareto rozdělení (GDP)* [2], [3]:

$$G_{\xi}(x) = \begin{cases} 1 - e^{-x} & \text{pro } \xi = 0 \\ 1 - (1 + \xi x)^{-1/\xi} & \text{pro } \xi \neq 0 \end{cases} \quad (5)$$

Funkci $G_{\xi}(x)$ lze následně dále ještě zevšeobecnit na tříparametrický tvar [3]

$$G_{\xi, \beta, \gamma}(x) = \begin{cases} 1 - e^{-(x-\mu)/\sigma} & \text{pro } \xi = 0 \\ 1 - \left(1 + \xi \frac{x-\mu}{\sigma}\right)^{-1/\xi} & \text{pro } \xi \neq 0 \end{cases} \quad (6)$$

Parametr $\sigma = \sigma(u)$ závisí na práhu u a charakterizuje variabilitu hodnot. Parametr μ charakterizuje polohu. Při aplikaci metody je důležité správně zvolit práh u . Podle [4] velikost práhu u je vhodné zvolit mezi 90. až 95. percentilem empirického souboru dat, neboli z intervalu $(x_{0,9}; x_{0,95})$.

4. Aplikace metod modelování extrémních katastrofických škod

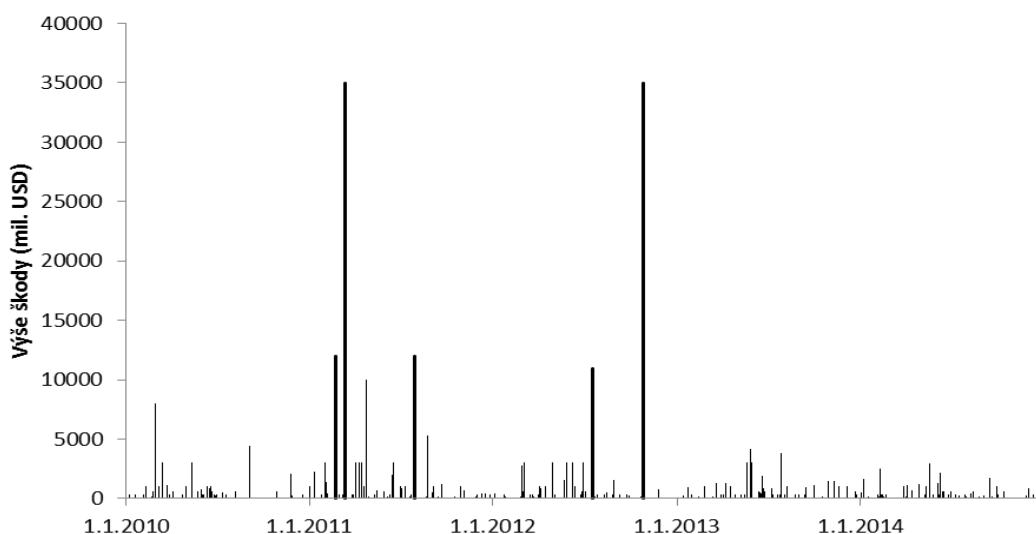
Pro pravděpodobnostní modelování extrémních pojistných plnění z katastrofických událostí podle metod, popsaných v předchozí kapitole, byly použity statistické programové balíky STATISTICA 12, STATGRAPHICS Centurion XVII a statistický program R verze 3.2.2 (volně dostupný na <https://www.r-project.org/>) společně s modulem Extremes Toolkit (volně dostupný na <http://www.isse.ucar.edu/extremevalues/evtk.html>), obsahující možnost modelování rozdělení, zobecněné rozdělení extrémních hodnot (GEV) i zobecněné Pareto rozdělení (GDP).

4.1 Charakteristika analyzovaných dat

Data, která byla použita pro modelování, byla ručně přepsaná z časopisu SIGMA, kterou každoročně vydává švýcarská zajišťovna Swiss Re (viz [6], Table 8) a následně upravena v programu MS EXCEL. Jedná se o data vyjadřující výši vyplaceného pojistného plnění v milionech amerických dolarů (USD) v důsledku přírodních katastrof, které se udály ve světě od roku 2010 do roku 2014. Analyzovaný soubor obsahuje 323 dat.

Nejdříve je vhodné vytvořit graf s chronologicky seřazenými pojištěnými škodami, který nám umožní vizuálně identifikovat extrémní hodnoty a časové období jejich výskytu. Údaje jsou uvedeny v milionech USD. Na první pohled jsou na obrázku 3 značně odlehle hodnoty zobrazeny tučnou čarou. Tento graf poskytuje informaci, zda data neporušují předpoklad nezávislosti a identického rozdělení. To by znamenalo, že by se extrémní hodnoty shromažďovaly v nějakém krátkém časovém období, což se u našeho datového souboru nestává. Tím jsou splněny obě podmínky.

Obrázek 3: Vývoj celkových pojištěných katastrofických škod



Zdroj: vlastní zpracování dle [5]

4.2 Analýza extrémních pojistných plnění metodou blokového maxima

První aplikovanou metodou analyzovaných dat, kterými jsou výše pojistných plnění na základě jednotlivých přírodních katastrof ve světě od roku 2010 do roku 2014 v milionech USD, je metoda blokového maxima. Na začátku modelování je důležité si nastavit optimální délku bloku n . Následně určíme pro všechny bloky maximální hodnoty, pomocí kterých bude vytvořen nový soubor, který použijeme pro další analýzy. Metoda je aplikována ve statistickém programovém balíku STATISTICA 12 a pomocí softwarové aplikace R 3.2.2. s modulem *Extremes Toolkit*.

Analyzovaná data budeme modelovat pomocí zobecněné distribuční funkce GEV extrémních hodnot, která je daná vztahem (4). Byly zvoleny délky bloku ($n = 5$, $n = 10$, $n = 15$, $n = 20$) z důvodu, abychom zjistili, která délka bloku dává optimální výsledky.

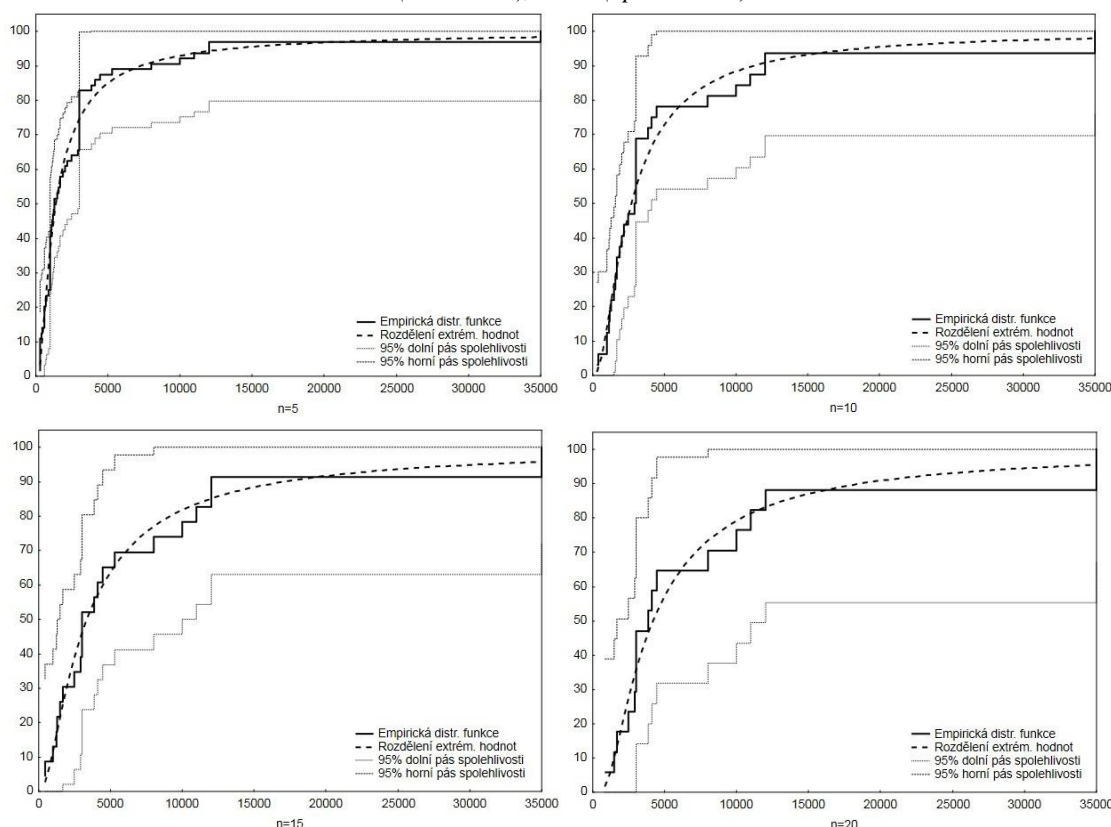
Tabulka 1: Parametry GEV pro různé délky bloků n

délka bloku	$n = 5$	$n = 10$	$n = 15$	$n = 20$
počet pozorování	64	32	23	17
MLE (μ)	-0,20329	0,36093	0,49964	0,68634
MLE (σ)	1,04353	1,17186	1,02954	0,78838
MLE (ξ)	0,32880	0,33360	0,35764	0,64920
SE (μ)	0,15069	0,22834	0,25550	0,22532
SE (σ)	0,12968	0,16322	0,21891	0,23824
SE (ξ)	0,12032	0,18589	0,22773	0,29399
p-hodnota	0,38079	0,50006	0,94977	0,93914

Zdroj: vlastní zpracování, výstup R 3.2.2 modul *Extremes Toolkit*

V tabulce 1 jsou maximálně věrohodné odhady (MLE) parametrů μ , σ , a ξ rozdělení GEV spolu se standardními chybami (SE) odhadu z jednotlivých bloků délky n . V tabulce jsou také znázorněny p -hodnoty pro modelování blokových maxim. Můžeme vidět, že maxima všech bloků mají zobecněné rozdělení extrémních hodnot. Jako optimální velikost bloku se jeví blok $n = 15$, kde p -hodnota je 0,9497. Dobré výsledky také dává blok s délkou $n = 20$, kde p -hodnota vychází 0,9314.

Obrázek 4: Empirická distribuční funkce GEV rozdělení pro $n=5$ (vlevo nahoře), $n=10$ (vpravo nahoře), $n=15$ (vlevo dole), $n=20$ (vpravo dole)



Zdroj: vlastní zpracování, výstup STATISTICA 12

Na obrázku 4 jsou znázorněny empirické distribuční funkce společně s teoretickými distribučními funkcemi rozdělení GEV pro jednotlivé bloky různé délky n . Jsou zde vyznačeny také 95%-ní intervaly spolehlivosti. Lze vidět, že nejlépe odpovídá GEV rozdělení při délce bloku $n = 15$ nebo délce bloku $n = 20$, kde se teoretické pravděpodobnostní modely nacházejí uvnitř intervalů spolehlivosti [5].

4.3 Analýza extrémních pojistných plnění metodou excedentů přes vysoký práh

Pro analýzu byly zvoleny hodnoty prahů $u = 2000$, $u = 2500$, $u = 3500$ a $u = 4000$. V dalším kroku analýzy byly pro různé hodnoty prahů u vytvořeny nové datové soubory z hodnot, přesahujících příslušný práh u . Nové datové soubory byly použity pro určení vhodného pravděpodobnostního modelu škod, přesahujících zvolený práh, kterým je podle části 3.2 zobecněné Pareto rozdělení (GDP). Přitom byl využit statistický programový balík STATISTICA 12.

Maximálně věrohodný odhad parametrů pomocí tříparametrického zobecněného Pareto rozdělení (6) při různé hodnotě prahů u je znázorněn v tabulce 2. V posledním řádku jsou uvedeny výsledky Kolmogorova-Smirnova (K-S) testu dobré shody s rozdělením GDP pro různé velikosti prahů u . Nejlepší výsledek podle K-S testu dostáváme pro velikost prahu $u = 3500$, čemuž odpovídá 11 excedentů a p -hodnota je rovna 0,922191. Vysoká p -hodnota, rovna 0,883495 je i pro velikost prahu $u = 4000$. Velikosti prahu $u = 4000$ odpovídá 10 excedentů.

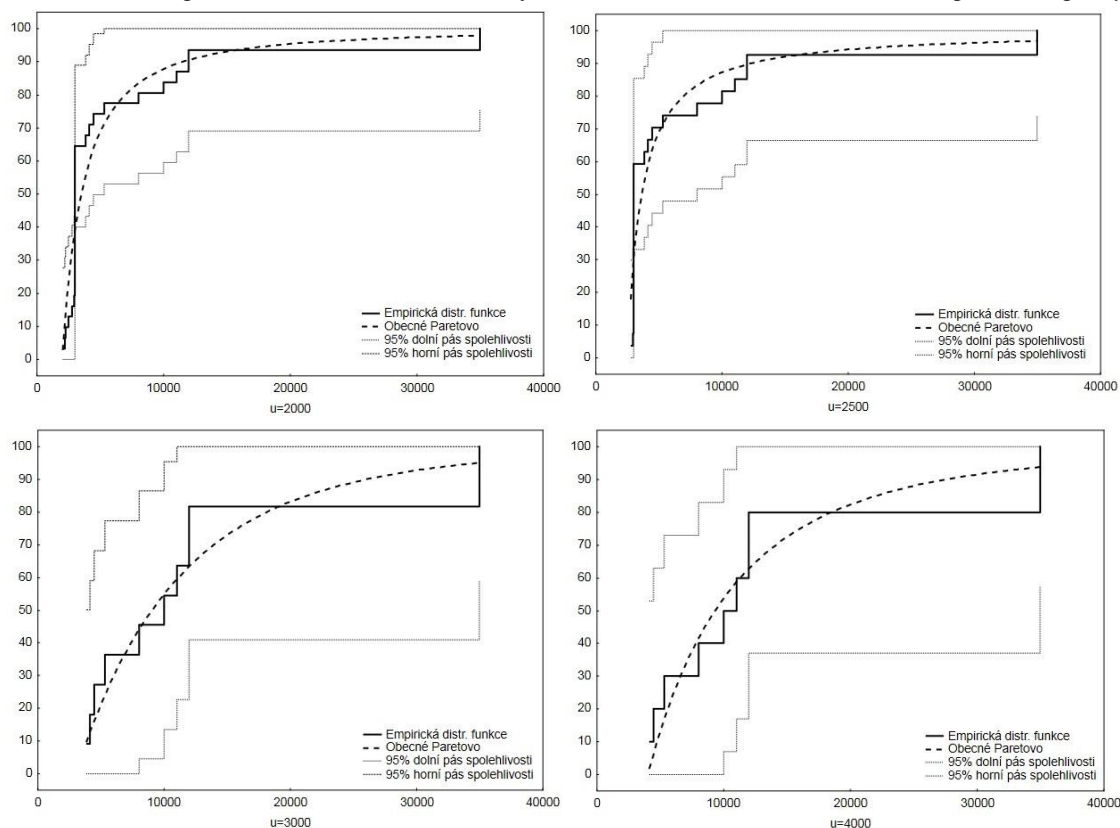
Tabulka 2: Odhad parametrů zobecněného Paretova rozdělení pro různé prahy u a výsledky K-S testů

$u = 2000$ zobecněné Pareto	$u = 2500$ zobecněné Pareto	$u = 3500$ zobecněné Pareto	$u = 4000$ zobecněné Pareto
$\xi = 1681,103$	$\xi = 1155,657$	$\xi = 6447,036$	$\xi = 6849,886$
$\mu = 2000$	$\mu = 2500$	$\mu = 3500$	$\mu = 4000$
$\sigma = 0,690287$	$\sigma = 0,962237$	$\sigma = 0,336181$	$\sigma = 0,325339$
$p = 0,031268$	$p = 0,017163$	$p = 0,922191$	$p = 0,883495$

Zdroj: vlastní zpracování, výstup STATISTICA 12

Na obrázku 5 jsou znázorněny empirické distribuční funkce v porovnání s teoretickou distribuční funkcí pro jednotlivé prahy u při předpokládaném GDP rozdělení pravděpodobnosti. Černá schodovitá křivka představuje empirickou distribuční funkci s porovnáním s teoretickou distribuční funkcí zobecněného Paretova rozdělení, kterou představuje čárkovaná křivka. Tečkovanými křivkami jsou zde vyznačeny 95 %-ní intervaly spolehlivosti. Můžeme vidět, že pro prahy $u = 3500$ a $u = 4000$ empirická distribuční funkce nepřekračuje hranice spolehlivosti [5].

Obrázek 5: Empirická a teoretická distribuční funkce zobecněného Paretova rozdělení pro různé prahy u



Zdroj: vlastní zpracování, výstup STATISTICA

5. Pravděpodobnostní modely extrémních pojistných plnění

Základem analýzy je 40 nejnákladnějších pojistných událostí na světě (v mil. USD), které se udály od roku 1970 do roku 2014 ([5], 2015, Table 10, str. 41).

V tabulce 3 je výsledek K-S testů dobré shody se třemi tříparametrickými rozděleními pravděpodobnosti a dvouparametrickým Pareto rozdělením [7], s parametry, odhadnutými metodou maximální věrohodnosti při využití systému Statgraphics Centurion XVII. Na základě p -hodnot hypotézu, že empirický soubor pochází z předpokládaných rozdělení,

nezamítáme ani v jednom z případů. Nejlepší rozdělení pro 40 nejnákladnějších katastrofických pojistných plnění je lognormální rozdělení, p -hodnota = 0,99797.

Tabulka 3: Výsledky Kolmogorovova-Smirnovova testu dobré shody

	Gamma	Lognormální	Paretovo	Weibullovo
DPLUS	0,09628	0,04733	0,05764	0,07231
DMINUS	0,06399	0,06186	0,09552	0,07193
DN	0,09628	0,06186	0,09552	0,07231
p-hodnota	0,85225	0,99797	0,85878	0,98494

Zdroj: vlastní zpracování, výstup Statgraphics

System Statgraphics Centurion poskytuje také základní charakteristiky a kvantily rozdělení s dobrou shodou. Pro lognormální rozdělení 40 nejnákladnějších katastrofických pojistných plnění proto můžeme stanovit střední hodnotu $E(X) = 13\,794,1$ mil. USD a míru rizika

$$VaR_{0,95}(X) = x_{0,95} = 42\,832 \text{ mil. USD.}$$

S pravděpodobností 0,05 tedy mohou některé ze 40 nejnákladnějších pojistných plnění přesáhnout hodnotu 42 832 mil. USD. Tuto hodnotu v období roků 1970-2014 přesáhly pojištěné škody při hurikánu Katrina v roce 2005, protože dosáhly výše 79 663 mil. USD.

Acknowledgements

Tento příspěvek vznikl za podpory SGS projektu Univerzity Pardubice, Fakulty ekonomicko-správní, č. SGS_2016_023.

References

- [1] Fisher, R.A., Tippett, L.H. (1928). Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Math. Proc. Camb. Phil. Soc. Mathematical Proceedings of the Cambridge Philosophical Society*, 24(02), p.180.
- [2] Jindrová, P., Pacáková, V. (2016). Modelling of extreme losses in natural disasters. *International Journal of Mathematical Methods in Applied Sciences*, Volume 10, 2016, pp. 171-178.
- [3] Jindrová, P., Papoušková, M. (2016). Modelling Insured Catastrophe Losses. In Proceedings of *The 10th Professor Aleksander Zelias International Conference on Modelling and Forecasting of Socio-Economic Phenomena*, Zakopane, pp. 53-61.
- [4] McNeil, A. (1997). Estimating the Tails of Loss Severity Distributions Using Extreme Value Theory. *ASTIN Bulletin*, 27(1), pp. 117-138.
- [5] Musil, T. (2016). *Možnosti modelování extrémních pojistných plnění*, diplomová práce, Fakulta ekonomicko-správní, Univerzita Pardubice.
- [6] Natural catastrophes and man-made disasters in 2010-2014: SIGMA Swiss Re [online]. [cit. 2016-06-11]. Dostupné z: http://media.swissre.com/documents/sigma_en.pdf
- [7] Pacáková, V., Gogola, J. (2013). Pareto Distribution in Insurance and Reinsurance. Conference proceedings from 9th international scientific conference *Financial Management of Firms and Financial Institutions*, VŠB Ostrava, 2013. pp. 298-306