

MODELOVÁNÍ KATASTROFICKÝCH ŠKOD

MODELLING OF CATASTROPHIC LOSSES

Viera Pacáková, Lukáš Kubec

***Abstract:** Catastrophe modelling is a risk management tool that uses computer technology to help insurers, reinsurers and risk managers better assess the potential losses caused by natural and man-made catastrophes. The models use historical disaster information to simulate the characteristics of potential catastrophes and to determine the potential losses cost. The aim of this paper is to describe parametric curve-fitting methods for modelling extreme historical losses. Article summarizes relevant theoretical results above Extreme value theory (EVT) and Excess over Threshold Method (EOT) and provide example of their application to Danish data on large fire insurance losses. Application of these methods is not possible without appropriate software packages. Article refers to these options too.*

***Keywords:** Extreme Value Theory, Generalized Pareto Distribution, Excess over Threshold Method.*

***JEL Classification:** C14, C15, C46.*

Úvod

Pro katastrofické (extrémní, závažné, mimořádné) události jsou typické tři základní rysy: zřídkačnost a náhodnost výskytu a závažné důsledky.

Katastrofické události postihují různé oblasti světa se stále rostoucí frekvencí a intenzitou. Mnohé regiony jsou ohrožovány katastrofickými riziky velkého rozsahu, kde se rozsáhlé ničivé události často opakují, někdy i vícekrát za rok. Velké katastrofické události mohou být způsobené přírodními jevy nebo jsou zapříčiněny člověkem. Je třeba si uvědomit, že i mnohé události přírodního charakteru jsou do značné míry ovlivněny činností člověka. Jde hlavně o klimatické změny, ale například i o vliv těžebního průmyslu. Závažné události v posledních letech jsou často důsledkem teroristických činů.

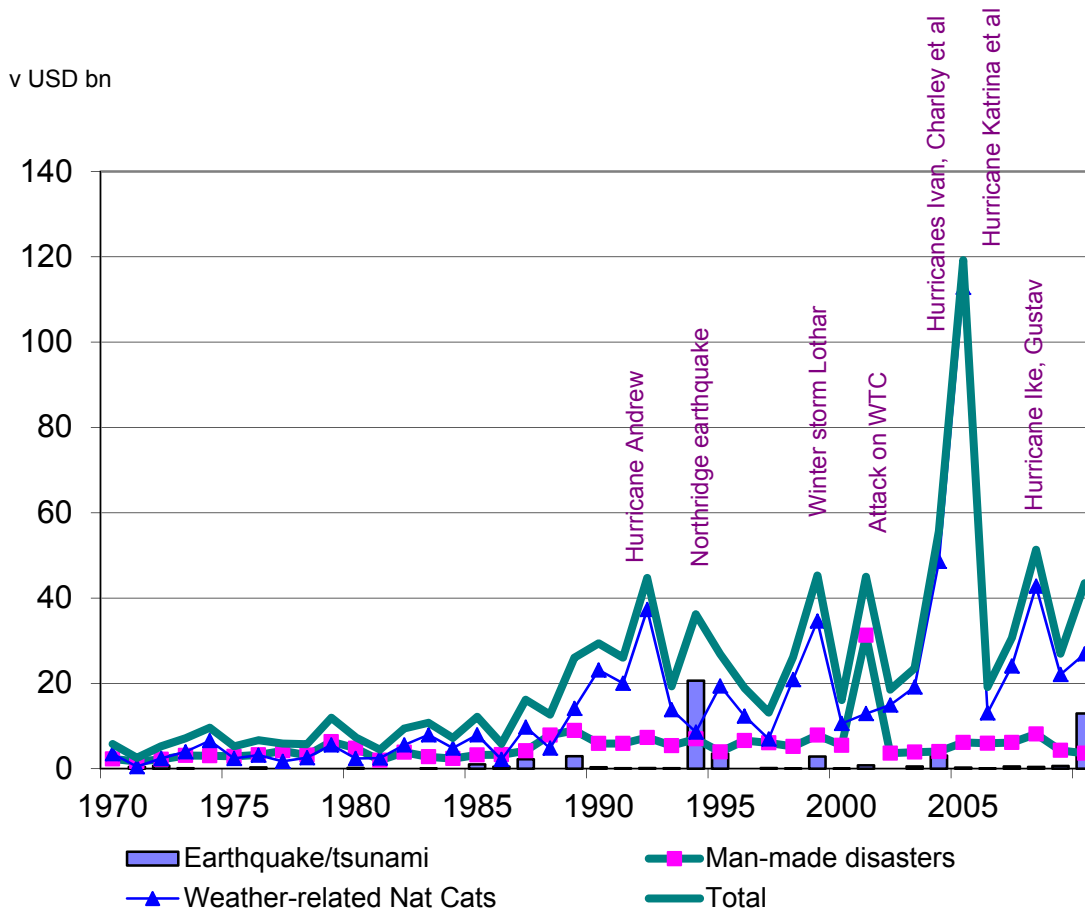
Výskyt katastrofických událostí je čím dál četnější a dále také rostou pojistné plnění pojišťoven a zajišťoven při těchto událostech (obr. 1). Statistiky informují, že katastrofické události tvoří až 80 % celkového pojistného plnění, ale jenom 20 % celkového počtu pojištěných škod [2].

Z těchto skutečností plyne nutnost takové události co nejvěrohodněji modelovat a tím zvolit nejvhodnější možnosti krytí přejímaných rizik a správné nastavení pojistného, respektive zajištěného.

Vývoj finančních důsledků katastrof má zásadní dopad na celosvětový pojistný trh a nutí pojišťovny a zajišťovny hledat nové cesty a způsoby krytí těchto rizik. Vzniká

obava, že kapacity světových pojistných a zajistných trhů nebudou v budoucnu postačující pro krytí těchto rizik a hledají se alternativní možnosti jejich transferu.

Obr. 1: Pojištěné katastrofické škody v letech 1970-2010



Zdroj: [8]

1 Modelování extrémních hodnot

V závislosti na konkrétních okolnostech se na modelování katastrofických (extrémních) škod používají různé metody modelování extrémních hodnot. Tyto metody se využívají také v jiných oblastech výzkumu či praxe, např. v hydrologii, klimatologii, meteorologii, strojírenství či finančnictví.

Klasická teorie extrémních hodnot (EVT, Extreme Value Theory) řeší problém pravděpodobnostního rozdělení maxima náhodných veličin. Nejdůležitějším výsledkem této teorie je Fisher-Tippettova věta z roku 1928 [4], která specifikuje typ limitního rozdělení pro maxima náhodných veličin. Tato věta je svým významem srovnatelná s významem centrální limitní věty v teorii pravděpodobnosti.

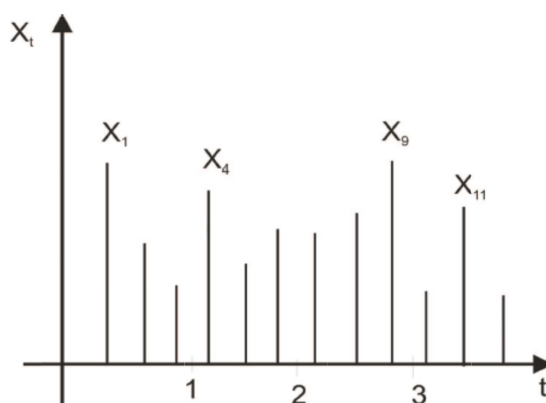
Ať náhodné veličiny X_1, X_2, \dots, X_n vyjadřují pojistné škody během určitého časového období. Používají se dvě základní metody, pomocí kterých identifikujeme extrémní hodnoty: metoda blokového maxima a metoda excedentů přes vysoký práh. V každé z těchto metod se zkoumají možné limitní rozdělení pro příslušné náhodné veličiny.

1.1 Metoda blokového maxima

Metoda blokového maxima považuje za extrémní hodnoty maxima, která pozorujeme v několika za sebou následujících obdobích (blocích) stejné délky. Délka n jednoho bloku by měla být taková, aby bylo možné zanedbat sezónní vlivy. Obvykle představuje délky přirozených časových období, jako jsou měsíce, týdny nebo roky.

Zřejmým nedostatkem této metody je to, že používá jen jednu hodnotu z každého bloku, avšak v jednom bloku může být více hodnot, které lze považovat za extrémní.

Obr. 2: Ukázka stanovení blokových maxim



Zdroj: [9]

Rozdělení tzv. blokových maxim je dané Fisher-Tippetovou větou, která říká, že pokud existuje nedegenerované limitní rozdělení vhodně normalizovaných maxim stejně rozdělených náhodných veličin (pro počet náhodných veličin $n \rightarrow \infty$), pochází z jednoho ze tří „zobecněných rozdělení extrémních hodnot“ (Fréchetova, Gumbelova či Weibullova rozdělení) [10]:

Tři druhy rozdělení tzv. blokových maxim definuje Fisher-Tippetova věta [4, s. 121], která říká, že pokud existuje nedegenerované limitní rozdělení vhodně normalizovaných maxim stejně rozdělených náhodných veličin (pro počet náhodných veličin $n \rightarrow \infty$), přísluší k jednomu ze tří zobecněných rozdělení extrémních hodnot:

Fréchet

$$\Phi_{\alpha}(x) = \begin{cases} 0 & \text{pro } x \leq 0 \\ \exp\{-x^{-\alpha}\} & \text{pro } x > 0 \end{cases}$$

Gumbel

$$\Lambda_{\alpha}(x) = \exp\{-e^{-x}\} \quad \text{pro } x \in \mathbb{R}$$

Weibull

$$\Psi_{\alpha}(x) = \begin{cases} 1 & \text{pro } x \geq 0 \\ \exp\{-(-x)^{-\alpha}\} & \text{pro } x < 0 \end{cases}$$

Zobecněné Gumbelovo, Fréchetovo a Weibullovo rozdělení může být zkombinováno do tvaru společné distribuční funkce (d. f.):

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

která se nazývá zobecněná d. f. extrémních hodnot (GEV). Pro různý parametr ξ , který se nazývá indexem extrémních hodnot, dostaneme jedno ze třech rozdělení podle Fisher-Tippetovy věty:

- Frechetovo rozdělení pro $\xi = \alpha - 1, x > -\xi - 1$
- Weibullovo rozdělení pro $\xi = \alpha - 1, x < -\xi - 1$
- Gumbelovo rozdělení pro $\xi = 0, x$

Pro praktické aplikace je výhodnější použít flexibilnější rozdělení extrémních hodnot, vyjádřené pomocí trojice parametrů:

$$G(x) = \exp\left\{-\left[1 + \xi\left(\frac{x - \mu}{\sigma}\right)\right]^{-\frac{1}{\xi}}\right\}$$

$$g(x) = \exp\left\{-\left[1 + \xi\left(\frac{x - \mu}{\sigma}\right)\right]^{-\frac{1}{\xi}}\right\} \cdot \frac{1}{\sigma}\left[1 + \xi\left(\frac{x - \mu}{\sigma}\right)\right]^{(-\frac{1}{\xi})-1}$$

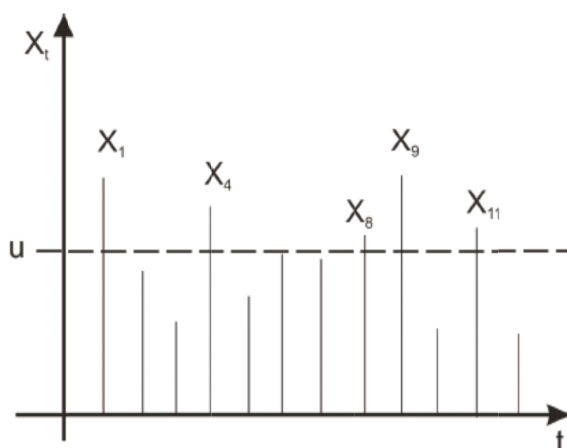
kde $\{x: 1 + \xi(x - \mu)/\sigma > 0\}, -\infty < \mu < \infty$ a $-\infty < \xi < \infty$.

Metoda blokového maxima je klasickým příkladem přístupu EVT, která se snaží obvykle najít nejvhodnější parametry jednoho ze tří výše zmíněných rozdělení podle Fisher-Tippetovy věty.

1.2 Metoda excedentů přes vysoký práh

Metoda excedentů přes vysoký práh (POT, Peaks Over Threshold) považuje za extrémní všechny hodnoty vyšší než je zvolený práh u . Princip této metody znázorňuje obrázek 3, kde X_1, X_4, X_8, X_9 a X_{11} jsou extrémní překračující zvolený práh u .

Obr. 3: Ukázka stanovení excedentů přes práh u



Zdroj: [9]

Nechť náhodná veličina X má distribuční funkci $F(x)$ s pravým koncovým bodem $x_F = \sup\{x \in \mathbb{R}; F(x) < \infty\}$, potom rozdělení excedentů přes vysoký práh lze vyjádřit podmíněnou distribuční funkcí [10]:

$$F_u(x) = P(X - u \leq x / X > u) = \frac{F(x+u) - F(u)}{\bar{F}(u)} \quad \text{pro } 0 \leq u < x_F$$

Podmíněná střední hodnota jako funkce zvoleného práhu $e(u)$, se nazývá funkce průměrných excedentů překračující práh u a je definována [10]:

$$e(u) = E(X - u / X > u) = \frac{1}{\bar{F}(u)} \int_u^{x_F} \bar{F}(x) dx$$

Vyjadřuje střední hodnotu excedentů náhodné veličiny X přes měnící se práh u .

Tak jako GEV je limitním rozdělením pro bloková maxima nezávislých náhodných veličin, tak pro excedenty přesahující určitý práh u je limitním rozdělením zobecněné Paretovo rozdělení (GPD) [10]:

$$G_\xi(x) = \begin{cases} 1 - e^{-x} & \text{pro } \xi = 0 \\ 1 - (1 + \xi x)^{-1/\xi} & \text{pro } \xi \neq 0 \end{cases} \quad (2)$$

Toto rozdělení je možné dále parametrizovat na tvar [9]:

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

Parametr $\beta = \beta(u)$ závisí na výšce zvoleného práhu u a charakterizuje rozptýlenost. Je tedy velmi důležitá správná volba práhu, který by měl být podle [10] mezi 90. až 95. percentilem a to z toho důvodu, že už běžná rozdělení výše škod špatně aproximují pravé konce (chvosty) empirických dat.

2 Aplikace metod modelování extrémních hodnot

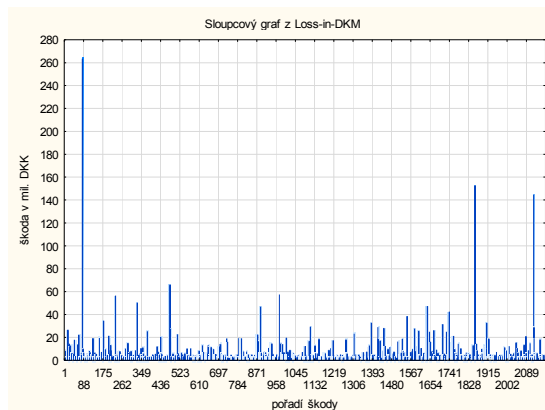
Zde se zaměříme na modelování extrémních škod pomocí metody blokového maxima a metody excedentů přes vysoký práh v programech STATISTICA, STATGRAPHIC *Centurion* a MS EXCEL. Některé výpočty jsou provedeny ve statistickém programu R, který je volně dostupný na <http://www.r-project.org/>, spolu s modulem Extremes Toolkit (volně dostupný na

<http://www.isse.ucar.edu/extremevalues/evtk.html>), který nabízí GEV i GPD modely ve tvaru, vyjádřeném pomocí vztahů (1) a (2).

Pro ukázkou modelování budou použity data o výškách škod přesahující jeden milion dánských korun (DKK), které způsobily požáry v Dánsku v letech 1980-1990. Analyzovaný soubor dat (obr. 4) obsahuje 2167 záznamů, které informují o výši vzniklých škod požárem v jednotlivých dnech, ale pouze pokud celková škoda v jednom dni přesáhla jeden milion DKK. Výškou škody se rozumí celková škoda

způsobená požárem na budovách, zařízení budov i případný ušlý zisk. Data jsou převzata z [5].

Obr. 4: Časová řada výše modelovaných škod



Zdroj: Vlastní zpracování v programu STATISTIKA

2.1 Metoda blokového maxima

Při aplikaci této metody je nejprve potřeba zvolit vhodnou délku jednoho bloku, pro jednotlivé bloky určit maxima a vytvořit nové datové soubory pro další analýzu. Datové soubory byly vytvořeny v MS EXCEL a byly zvoleny délky bloků $n = 5$, $n = 10$, $n = 15$ a $n = 20$. Jejich základní charakteristiky jsou uvedeny v tabulce 1.

Tab. 1: Základní statistiky datových souborů pro různá n

Summary Statistics for $n=5$		Summary Statistics for $n=10$		Summary Statistics for $n=15$		Summary Statistics for $n=20$	
Count	434	Count	217	Count	145	Count	109
Average	8,91665	Average	13,6107	Average	17,7479	Average	21,1851
Median	4,65707	Median	7,32064	Median	10,8205	Median	13,5
Standard deviation	17,699	Standard deviation	23,956	Standard deviation	28,3361	Standard deviation	31,9034
Coeff. of variation	198,494%	Coeff. of variation	176,008%	Coeff. of variation	159,659%	Coeff. of variation	150,594%
Minimum	1,25655	Minimum	1,5	Minimum	1,85	Minimum	3,24675
Maximum	263,25	Maximum	263,25	Maximum	263,25	Maximum	263,25
Range	261,994	Range	261,75	Range	261,4	Range	260,004

Zdroj: Výstup z programu STATISTIKA

Maximálně věrohodné odhady parametrů pro analyzované datové soubory s různou délkou bloku n jsou uvedeny v tabulce 2. Dobrou shodu lze pozorovat na Q-Q plotu (Quantile plot) na obrázku 5.

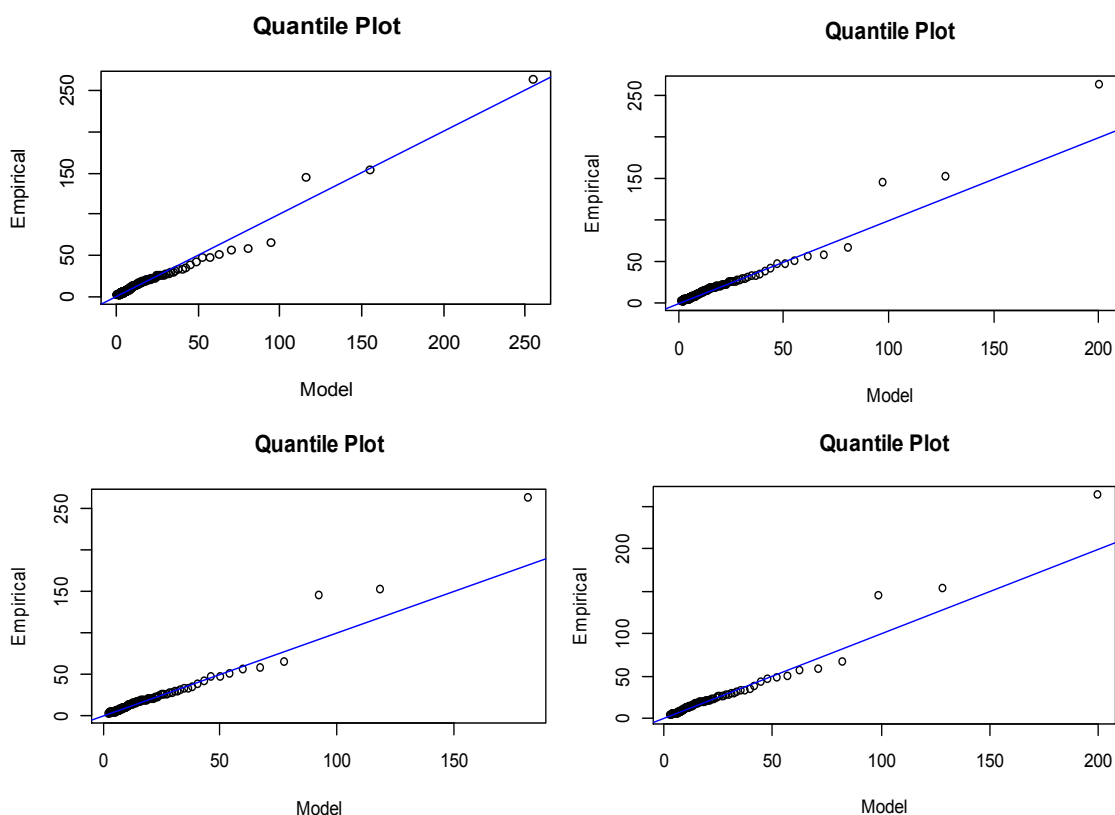
Podle Q-Q plotů na obr. 5 s nižší hodnotou n GEV poskytuje horší aproximaci empirických hodnot, protože se do analyzovaného souboru dostávají i údaje nekatastrofické a s rostoucím n roste standartní chyba odhadu parametrů. Z toho plyne nutnost najít rovnováhu mezi skresleným odhadem a velikostí jejího rozptylu. Podle těchto požadavků bude nejvhodnější zvolit pro případnou další analýzu, nebo další využití, např. pro neproporcionální zajištění, délku bloku $n = 10$ nebo $n = 15$.

Tab. 2: Parametry GEV rozdělení pro různá n

$n=5$			$n=10$			$n=15$			$n=20$		
	MLE	SE		MLE	SE		MLE	SE		MLE	SE
μ	3.66709	0.13254	μ	5.79621	0.30333	μ	8.00705	0.50890	μ	9.58741	0.75697
σ	2.38980	0.14625	σ	3.95072	0.32429	σ	5.39360	0.52885	σ	6.52925	0.77824
ξ	0.71336	0.05626	ξ	0.64922	0.07033	ξ	0.60786	0.08507	ξ	0.63186	0.12485

Zdroj: Výstup z R modulu Extremes Toolkit

Obr. 5: Q-Q plot GEV rozdělení pro $n=5$ vlevo nahoře, $n=10$ vpravo nahoře, $n=15$ vlevo dole a $n=20$ vpravo dole



Zdroj: Výstup z R modulu Extremes Toolkit

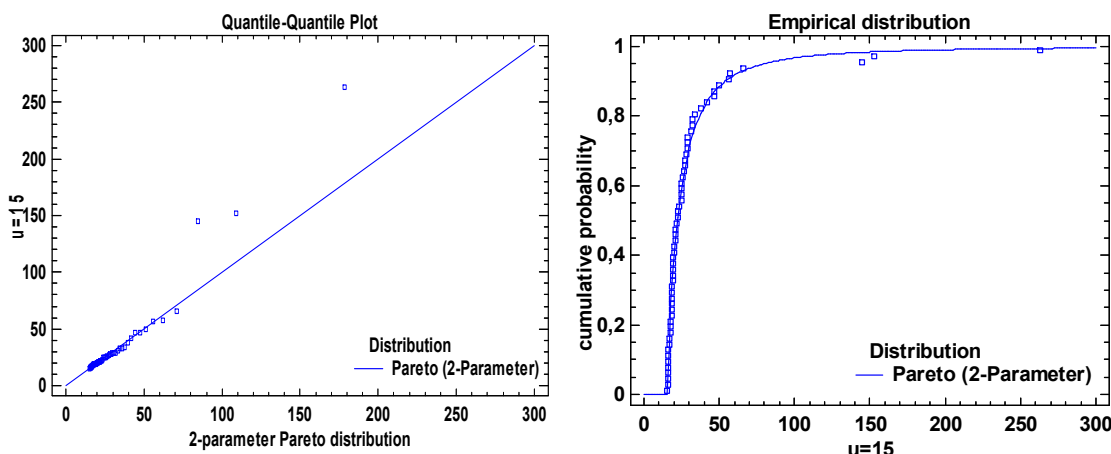
2.2 Metoda excedentů přes vysoký práh

Při aplikaci této metody zvolíme hodnotu prahu u a vytvoříme nový datový soubor, kde všechny hodnoty budou překračovat u . Tento datový soubor využijeme pro určení vhodného pravděpodobnostního modelu, přičemž se místo GPD budeme muset spokojit se dvouparametrickým Pareto rozdělením [6], protože programové systémy STATGRAPHIC *Centurion* ani STATISTICA rozdělení GPD neobsahují.

Pro další zpracování v praktické části byly zvoleny velikosti prahů $u = 5$, $u = 10$, $u = 15$ a $u = 20$. Maximálně věrohodné odhady parametrů dvouparametrického Pareto rozdělení pro různé datové soubory s různým prahem u jsou v tabulce 3. Zde jsou uvedeny také p -hodnoty Kolmogorovova-Smirnovova testu dobré shody pro

různé velikosti u . P -hodnota je nejvyšší pro $u = 15$, čemuž odpovídá 60 excedentů. Proto za nejvhodnější považujeme práh $u = 15$. Dobrou shodu potvrzuje také $Q-Q$ graf a graf shody empirické a teoretické distribuční funkce pro $u = 15$ na obrázku 6.

Obr. 6: Grafické ověření dobré shody s Pareto rozdělením pro práh $u = 15$



Zdroj: Výstup z programu STATISTIKA

Modelování škod vyšších jako zvolený práh u pomocí zobecněného Paretova rozdělení (GPD) umožňuje statistický program R v modulu Extremes Toolkit. To umožňuje porovnat rozdíly v modelování extrémních škod dvouparametrickým Pareto rozdělením a zobecněným Pareto rozdělením (GPD), které je definováno vztahem 2. Program R s modulem Extremes Toolkit umožňuje analyzovat data pro různé úrovně prahu u a nevyžaduje ruční vytvoření datových souborů s požadovanými daty přesahujícími práh u .

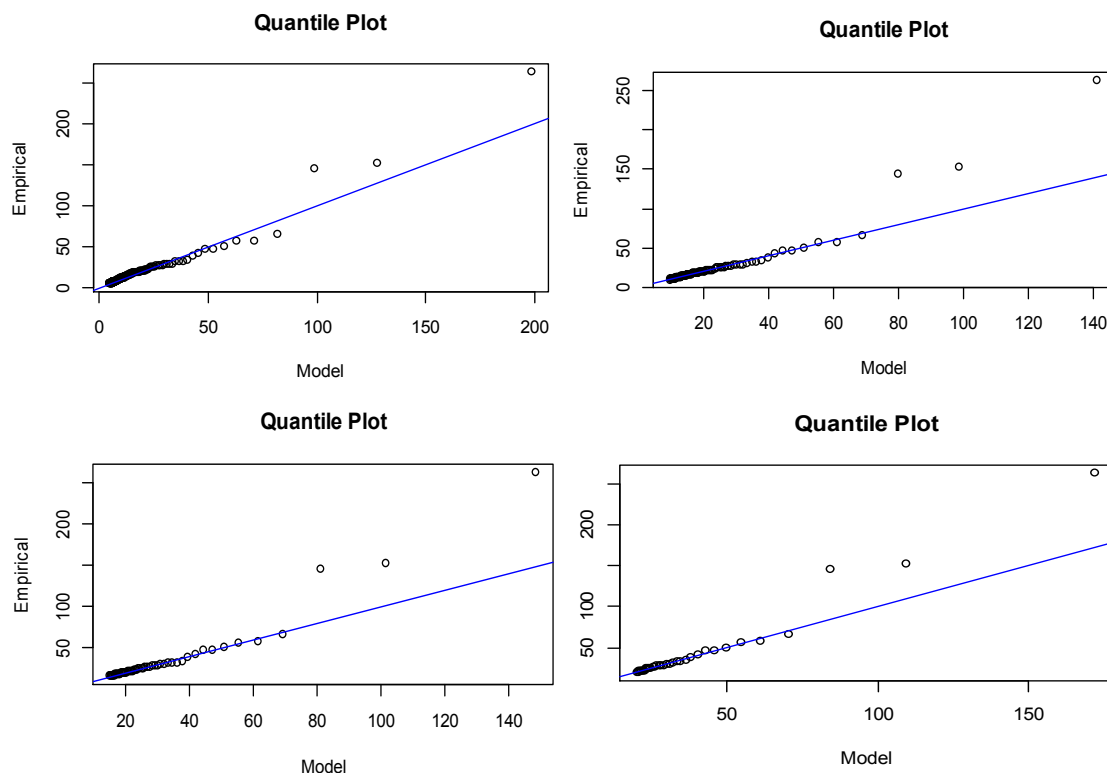
Tab. 3: Parametry GPD rozdělení pro různé hodnoty prahů u

$u=5$			$u=10$			$u=15$			$u=20$		
	MLE	SE		MLE	SE		MLE	SE		MLE	SE
β	3.80786	0.4636	β	6.97385	1.11324	β	8.71846	1.84173	β	9.63143	2.89499
ξ	0.6315	0.1116	ξ	0.4971	0.1363	ξ	0.5428	0.1812	ξ	0.683648	0.2747358
p	0,44425		p	0,76717		p	0,96280		p	0,913898	

Zdroj: Výstup R modulu Extremes Toolkit

Dobrou shodu empirických škod přesahujících zvolené prahy u s rozdělením GPD můžeme posoudit také pomocí $Q-Q$ grafů na obr. 6, $Q-Q$ plot GPD rozdělení pro $u = 5$ je vlevo nahoře, pro $u = 10$ vpravo nahoře, pro $u = 15$ vlevo dole a pro $u = 20$ vpravo dole. Nejlepší shoda empirických kvantilů a kvantilů GPD rozdělení je patrná aj vizuálním posouzením pro hodnotu prahu $u = 15$.

Obr. 6: Grafické ověření dobré shody s GPD rozdělením pro různé prahy



Zdroj: výstup R modulu Extremes Toolkit

Závěr

Znalost pravděpodobnostních modelů extrémních (mimořádných, katastrofických) škod je pro pojišťovny i zajišťovny mimořádně důležitá, zvláště při kalkulaci cen pojistného a zajistného a stanovení optimálních parametrů při neproporcionálním zajištění.

V článku jsme se pokusili o teoretický popis metod, vyvinutých v aktuárských vědách pro tyto účely v rozsahu, potřebném pro modelování extrémních škod pomocí vyvinutých programových balíčků. Právě tyto programové balíčky, či už statistické, nebo zaměřené speciálně pro potřeby pojistné praxe, umožňují masové využití těchto metod. Jejich aplikace jednak pro teoretickou náročnost a jednak pro výpočetní pracnost by bez informačních technologií a vhodných softwarových produktů nebyla možná.

Metody, teoreticky popsané a aplikované v článku, poskytují užitečné informace pro pojišťovny a zajišťovny. Vyžadují však specifická data ve tvaru dlouhodobých časových řad výše pojistných škod, které by každá pojišťovna měla průběžně kompletizovat. Bez jejich znalosti analýza a prognóza katastrofických rizik není možná.

Tento příspěvek byl zpracován s podporou projektu SGFES01/2012 vědecko-výzkumné aktivity v oblasti „Systémové inženýrství a informatika“.

Reference

- [1] CIPRA, T. *Zajištění a přenos rizik v pojišťovnictví*. 2. Vydání. Praha: Grada, 2004. ISBN 80-247-0838-8.
- [2] Ducháčková, E.: Katastrofy v roce 2003 a komerční pojištění. In: *Pojistný obzor* č. 10, 2004.
- [3] EMBRECHTS, P., KLUPPELBERG, C., MIKOSCH, P. *Modelling extremal events for insurance and finance*. New York: Springer, 1997, 645 s. ISBN 35-406-0931-8.
- [4] FISHER, R. A. – TIPPET, L. H. C.: Limiting forms of the frequency distribution of the largest or smallest number of the sample. In *Proceeding. Cambridge Philos. Soc.* č.24, 1928, s. 180-190.
- [5] McNEIL, A. Estimating The Tails Of Loss Severity Distributions Using Extreme Value Theory. In: *ASTIN BULLETIN* [online]. Zurich, 1997 [cit. 2012-04-08]. Dostupné na WWW: <http://www.actuaries.org/LIBRARY/ASTIN/vol27no1/117.pdf>
- [6] PACÁKOVÁ, V. *Aplikovaná poistná štatistika*. 3. vydání. Bratislava: Iura Edition, 2004, 248 s. ISBN 80-807-8004-8.
- [7] SANDERS, D. E. A. *The Modelling of Extreme Events*. Institute of Actuaries, 2005. Dostupné na WWW: www.actuaries.org.uk/system/files/documents/pdf/sm20050404.pdf
- [8] Sigma, SWISS RE [online]. Zurich, 2010 [cit. 2012-04-08]. Dostupné na WWW: http://media.swissre.com/documents/sigma1_2010_en.pdf
- [9] SKŘIVÁNKOVÁ, V.; TARTALOVÁ, A. Catastrophic Risk Management in Non-life Insurance. In *E + M – Economie a Management*, 2008, č. 2, s. 65-72.
- [10] ZHONGXIAN, H.: *Actuarial modelling of extremal events using transformed generalized extreme value distributions and generalized Pareto distributions*. 2003. [cit. 2012-04-08]. Dostupné na WWW: <http://www.math.ohio-state.edu/history/phds/abstracts/pdf/Han.Zhongxian.pdf>.

Kontaktní adresa

prof. RNDr. Viera Pacáková, PhD.

Bc. Lukáš Kubec

Univerzita Pardubice, Fakulta ekonomicko-správní,

Ústav matematiky a kvantitativních metod

Studentská 84, 532 10 Pardubice

E-mail: vpacakova@upce.cz, kubecl@seznam.cz

Received: 09.06.2012

Reviewed: 02.01.2013

Approved for publication: 23.01.2013