The distribution function as a tool for judging the extent of risk

Distribučná funkcia ako nástroj posudzovania miery rizika

Barbora Simanová [1]

Abstract

The aim of the paper is to present some functions for determining the extent of the risk described by an appropriate random variable, whose incidence directly relates to the distribution function. We can specify the distribution function, as a tool for measuring risk, on the basis of date relating to individual claim amounts, available within an insurer. For typical distributions describing individual claim amounts we can derive using them three functions suitable for measuring risk, with whose help we can analyse the positively-skewed joint distribution. The application part concentrates on the use of the derived functions for distributions with light and heavy tails to analyse the level of risk and to contrast the results with the Value-at-Risk values.

Key words

Individual loss, light-tailed distribution, heavy-tailed ditribution, survival function, hazard rate, mean excess loss.

JEL Classification: G22

1. Úvod

V posledných rokoch sa aktuári sústreďujú nielen na identifikáciu rizík, ale aj na ich presnejšie merania, ktoré vedú k hlbším analýzam a spoľahlivejšiemu riadeniu prevzatých rizík. Poisťovne majú k dispozícii dáta o počte a výške škôd, a z nich je potrebné určiť akým rozdelením, akou distribučnou funkciou sa tieto dáta záujmu riadia.

V tomto príspevku sa zameriame na náhodnú premennú popisujúcu výšku individuálnej škody v neživotnom poistení a naším cieľom bude vyjadriť miery rizika prostredníctvom príslušnej distribučnej funkcie analyzovanej náhodnej premennej.

Prvým krokom je špecifikovať z dát, ktoré riziko určujú, rozdelenie. Všeobecný postup pri výbere najvhodnejšieho rozdelenia výšky poistných plnení môžeme rozdeliť do troch krokov, ktoré sú grafická analýza, na základe ktorej určíme predpokladaný typ rozdelenia, odhad parametrov rozdelenia (metódou maximálnej vierohodnosti, metódou momentov, metódou kvantilov), ktoré sme si zvolili pri grafickej analýze a overenie vhodnosti zvoleného rozdelenia na základe testov, napr. Pearsonovho $\chi 2$ testu dobrej zhody, Kolmogorovov-Smirnovovho testu, Cramérov Von Misesovho testu, Andersonov-Darlingovho testu alebo Kuiperovho testu.

[1] Ing. Barbora Simanová, Katedra matematiky, Fakulta hospodárskej informatiky, EU Dolnozemská 1/b, 852 35 Bratislava, e-mail: barbora.simanova@gmail.com
Príspevok vznikol v rámci riešenia grantu VEGA č. : 1/0931/11, Analýza a modelovanie rizík v zmysle kvantitatívnych štúdií QIS projektu SOLVENCY II.

Platí, že ak pravdepodobnosť výskytu hodnôt individuálnych škôd vyšších ako je priemerná škoda je nezanedbateľný, tak rozdelenia, ktoré tieto škody opisujú majú veľký rozptyl a pozitívny koeficient asymetrie. Medzi vhodné modely individuálnej výšky škody patria napr. rozdelenie exponenciálne, gama, Paretovo, Weibullovo, lognormálne, atď.

Niektoré z uvedených rozdelení majú ľahký a niektoré ťažký chvost. Informácia o charaktere pravej strany rozdelenia je veľmi dôležitá pri kvantifikácii rizika. V aktuárskych, ale aj vo finančných aplikáciách štúdium rozdelení s ťažkými koncami poskytuje informácie o potenciáli rizika, a preto distribučná funkcia je základným nástrojom pre modelovanie veľkej straty.

Pre splnenie stanoveného cieľa v druhej časti príspevku uvedieme matematický aparát, ktorý umožní riziko individuálnej škody kvantifikovať a v tretej časti aplikujeme prezentovanú teóriu.

2. Rozdelenia individuálnej výšky škody

Pre výšky poistných plnení je typické, že väčšinou riziko ovplyvňujú hodnoty vyššie ako sú priemerné hodnoty, resp. hodnoty konkrétnych kvantilov s nezanedbateľnými pravdepodobnosťami.

2.1 Rozdelenia s ťažkými a ľahkými chvostmi

Rozdelenie s *ľahkým chvostom* je rozdelenie, pre ktoré existujú také konštanty $\boldsymbol{a} > 0$, $\boldsymbol{b} > 0$, že platí

$$\boldsymbol{S_X(x) \leq ae^{-bx}} \tag{2.1}$$

alebo existuje také $z > 0$, pre ktoré existuje momentová vytvárajúca funkcia, pričom $S_X(x) = 1 - F_X(x)$.

Rozdelenie s *ťažkým chvostom* je rozdelenie, pre ktoré platí

$$S_X(x) > ae^{-bx} \tag{2.2}$$

pričom $a > 0$, $b > 0$, alebo pre $\forall z > 0$, platí $M_X(z) = \infty$.
Teda môžeme povedať, že rozdelenie má ťažký chvost, ak konverguje k nule pomalšie ako exponenciálne rozdelenie.

Rozdelenia s ľahkými chvostmi je vhodné použiť pri modelovaní malých poistných plnení a rozdelenia s ťažkými chvostmi sa používajú v prípade nadmerných škôd, ako sú napr. zemetrasenia a iné živelné pohromy, ktoré výrazne ovplyvňujú výšku celkovej škody.

2.2 Identifikácia rozdelení s ťažkými a ľahkými pravými koncami

Existujú štyri základné indície, pomocou ktorých sa dá rozlíšiť, či skúmaná náhodná premenná nadobúda vyššie hodnoty s nezanedbateľnými pravdepodobnosťami.

A. ***Existencia momentov***.
Pre danú náhodnú veličinu, existencia všetkých začiatočných momentov naznačuje, že je to rozdelenie s ľahkým pravým koncom, chvostom rozdelenia. Naopak existencia začiatočných momentov *k*-tého rádu až od určitého *k*, kde *k* je kladné celé číslo, je známkou toho, že rozdelenie má ťažký pravý chvost. Teda napr. exponenciálne aj gama rozdelenie sú rozdelenia s ľahkým chvostom, pretože všetky momenty existujú a Paretovo rozdelenie je rozdelenie s ťažkým koncom, pretože prvé dva začiatočné momenty neexistujú. Začiatočné momenty *k*-tého rádu určíme na základe vzťahu

$$E(X^k) = \int_{-\infty}^{\infty} x^k f_X(x)\mathrm{d}x \tag{2.3}$$

za predpokladu, že daný integrál je konvergentný.

B. ***Rýchlosť klesania funkcie prežitia*** $S_X(x)$ ***k nule.***

Funkcia prežitia $S_X(x)$ náhodnej premennej X je definovaná ako pravdepodobnosť, že náhodná premenná X nadobúda hodnoty vyššie ako daná hodnota x. Označujeme ju aj $\bar{F}_X(x)$ a platí

$$S_X(x) = P(X > x) \tag{2.4}$$

Funkcia prežitia v neživotnom poistení zachytáva pravdepodobnosť pravých koncov rozdelení a je dôležitá pri ich analýzach. Fakt, že rozdelenie, ktorého funkcia prežitia sa približuje pomaly k nule, (ekvivalentne distribučná funkcia ide pomaly k jednotke), je ďalším znakom toho, že ide o rozdelenie s ťažkým chvostom.

C. Miera rizika prostredníctvom ***funkcie intenzity rizika***.

Ďalšou dôležitou funkciu využívanou pri kvantifikácii rizika je funkcia intenzity rizika (*hazard rate*), ktorá vyjadruje stupeň rizika a je definovaná podielom funkcie hustoty $f_X(x)$ náhodnej premennej X a príslušnej funkcie prežitia $\bar{F}_X(x)$ vo všetkých bodoch, v ktorých je funkcia hustoty definovaná, teda platí

$$h_X(x) = \frac{f_X(x)}{\bar{F}_X(x)} \tag{2.5}$$

Dôležitou vlastnosťou funkcie $h_X(x)$ je, že môže generovať funkciu prežitia. Platí

$$S_X(x) = e^{-\int_0^x h_X(t)\mathrm{d}t} \tag{2.6}$$

D. ***Funkcia strednej nadbytočnej škody***

Pre danú hodnotu d s $P(X > d) > 0$ náhodná premenná $Y = X - d$ určuje nadbytočnú škodu, za predpokladu, že $X > d$. Funkcia

$$e_X(d) = E(X - d/X > d) \tag{2.7}$$

sa nazýva *funkcia strednej nadbytočnej škody.*

Náhodná premenná Y má podmienené rozdelenie, zľava useknuté, a pre konkrétnu hodnot d jej stredná hodnota predstavuje očakávanú škodu zo všetkých hodnôt, ktoré prekračujú prahovú hodnotu d, napr. nejaký kvantil.

3. Aplikácia mier rizika

Význam a využitie definovaných mier rizika podrobne priblížime na Paretovom rozdelení a následne ich vyjadríme aj pre niektoré rozdelenia individuálnej výšky škody.

3.1 Paretovo rozdelenie

Funkcie definované vzťahmi (2.4), (2.5) a (2.7), prezentované ako miery rizika súvisiace s distribučnou funkciou, priblížime na Paretovom rozdelení. Prečo je to rozdelenie s ťažkým pravým koncom vysvetľuje aj fakt, že je to rozdelenie zmiešané, teda exponenciálne rozdelenie $E(\Theta)$, ktorého parameter je náhodná premenná s gama rozdelením $\Gamma(\alpha, \beta)$.

Predpokladajme, že sledujeme exponenciálne rozdelenie. Potom podmienené rozdelenie náhodnej premennej X má funkciu hustoty

$$f_{X/\Theta}(x/\theta) = \theta \cdot e^{-\theta x}, x > 0$$

pričom parameter Θ môže nadobúdať rôzne hodnoty a môže byť považovaný za náhodnú premennú. Predpokladajme, že sa riadi gama rozdelením s parametrom α, β. Potom

$$X/\Theta \sim E(\Theta) \wedge \Theta \sim \Gamma(\alpha,\beta), \Theta > 0, \alpha > 0, \beta > 0$$

$$f_\Theta(x) = \begin{cases} \frac{\beta^\alpha}{\Gamma(\alpha)} \cdot x^{\alpha-1} \cdot e^{-x\beta}, & x > 0 \\ 0, & x \leq 0 \end{cases}$$

a teda funkcia hustoty Paretovho rozdenia má tvar

$$f_X(x) = \int_0^\infty \theta \cdot e^{-\theta x} \cdot \frac{\beta^\alpha}{\Gamma(\alpha)} \cdot \theta^{\alpha-1} \cdot e^{-\beta\theta} \mathrm{d}\theta = \int_0^\infty \frac{\beta^\alpha}{\Gamma(\alpha)} \cdot \theta^\alpha \cdot e^{-\theta(x+\beta)} \mathrm{d}\theta$$

$$= \frac{\beta^\alpha}{\Gamma(\alpha)} \cdot \frac{\Gamma(\alpha+1)}{(x+\alpha)^{\beta+1}} \int_0^\infty \frac{(x+\beta)^{\alpha+1}}{\Gamma(\alpha+1)} \cdot \theta^\alpha \cdot e^{-\theta(x+\beta)} \mathrm{d}\theta = \frac{\alpha \cdot \beta^\alpha}{(x+\beta)^{\alpha+1}}$$

z ktorej odvodíme distribučnú funkciu v tvare

$$F_X(x) = 1 - \left(\frac{\beta}{\beta + x}\right)^\alpha$$

a charakteristiky

$$E(X) = \frac{\beta}{\alpha - 1}, \alpha > 1$$

$$D(X) = \frac{\alpha\beta^2}{(\alpha-1)^2(\alpha-2)}, \alpha > 2$$

Existencia momentov Paretovho rozdelenia je teda závislá od parametra α. Stredná hodnota neexistuje pre $\alpha \leq 1$. Ak toto rozdelenie modeluje straty, v prípade $\alpha = 1$ stredná hodnota neexistuje a riziko je nepoistiteľné. Na druhej strane, keď $\alpha = 2$ rozptyl Paretovho rozdelenia neexistuje, z čoho vyplýva že smerodajnú odchýlku neuvažujeme ako mieru rizika, ale zároveň podľa článku 2.2 bodu A možno usúdiť, že ide o rozdelenie s ťažkým koncom.

Rýchlosť klesania funkcie prežitia špecifikujeme na základe porovnania funkcie prežitia Paretovho a exponenciálneho rozdelenia. Uvažujme náhodnú premennú X s Paretovým rozdelením s parametrami $\alpha = 3$ a $\beta = 10$. V tabuľke 1 v prvom stĺpci sú niektoré vybrané hodnoty, ktoré táto náhodná premenná nadobúda. V druhom stĺpci sú funkčné hodnoty funkcie prežitia, z ktorých napr. vyplýva, že $x_{0,875} = 10$. V treťom stĺpci sú uvedené hodnoty funkcie prežitia premennej Y s exponenciálnym rozdelením s parametrom $\delta = 0{,}2079441584$, pre ktorú je tiež hodnota 10 kvantilom s tou istou pravdepodobnosťou.

Tabuľka 1: Porovnanie hodnôt funkcií prežitia Paretovho a exponenciálneho rozdelenia

x	$S_X(x)$	$S_Y(x)$	$\frac{S_X(x)}{S_Y(x)}$	$\frac{S_Y(x)}{S_X(x)}$
10	0,125	0,125	1	1
20	0,03703703703	0,01562499867	2,370370572	0,4218749640
30	0,015625	0,001953124752	8,000001015	0,1249999841
40	0,008	0,0002441405836	32,76800555	0,03051757294
100	0,0007513148009	$9{,}313221804 \cdot 10^{-10}$	$8{,}067184661 \cdot 10^{5}$	$1{,}239589822 \cdot 10^{-6}$

Z uvedených hodnôt vyplýva, že náhodná premenná X nadobúda vysoké hodnoty s oveľa vyššími pravdepodobnosťami ako náhodná premenná Y, čo je tiež potvrdené pomerom príslušných dvoch funkcií prežitia uvedených v štvrtom stĺpci tabuľky 1, ktorý sa blíži nekonečnu, alebo v opačnom prípade k nule. Na tomto príklade vidíme, že exponenciálne rozdelenie podceňuje výskyt hodnôt vyšších ako 10, aj keď môže byť dobrým modelom pre popis škôd do hodnoty $x_{0,875} = 10$. Teda škody vyššie ako 10 sú popísané problematicky.

V našom prípade teda môžeme tento pomer vyjadriť takto:

$$\lim_{x\to\infty}\frac{S_X(x)}{S_Y(x)} = \lim_{x\to\infty}\frac{\frac{\beta^\alpha}{(x+\beta)^\alpha}}{e^{-\delta x}} = \lim_{x\to\infty}\frac{\beta^\alpha e^{\delta x}}{(x+\beta)^\alpha} = \infty$$

Vo všeobecnosti platí, že ak pomer dvoch funkcií prežitia ide do nekonečna, znamená to, že rozdelenie v čitateli má ťažší chvost. Ak pomer ide do nekonečna, funkcia prežitia v čitateli konverguje k nule pomalšie vzhľadom na menovateľ.

Ďalšia funkcia miery rizika podľa (2.5) pre Paretovo rozdelenie je

$$h_X(x) = \frac{\alpha}{x+\beta}$$

z ktorej vyplýva, že pre ľubovoľné parametre Paretovho rozdelenia je klesajúcou funkciou, čo potvrdzuje fakt, že funkcia prežitia konverguje k nule pomalšie. Funkcia $h_Y(x)$ vyjadruje mieru rizika pre exponenciálne rozdelenie. Je to konštantná miera rizika, teda rozdelenie nesie vyššiu mieru rizika týkajúcu sa pravých koncov rozdelenia.

To, že platí vlastnosť definovaná vzťahom (2.6), dokážeme vyjadrením funkcie prežitia pomocou $h_X(x) = \frac{\alpha}{x+\beta}$. Platí

$$S_X(x) = e^{-\int_0^x \frac{\alpha}{t+\beta}\mathrm{dt}} = e^{[-\alpha\ln(t+\beta)]_0^x} = e^{-[\alpha\ln(x+\beta)-\alpha\ln\beta]} = e^{-\ln(x+\beta)^\alpha}e^{\ln\beta^\alpha}$$

$$S_X(x) = \frac{\beta^\alpha}{(x+\beta)^\alpha}$$

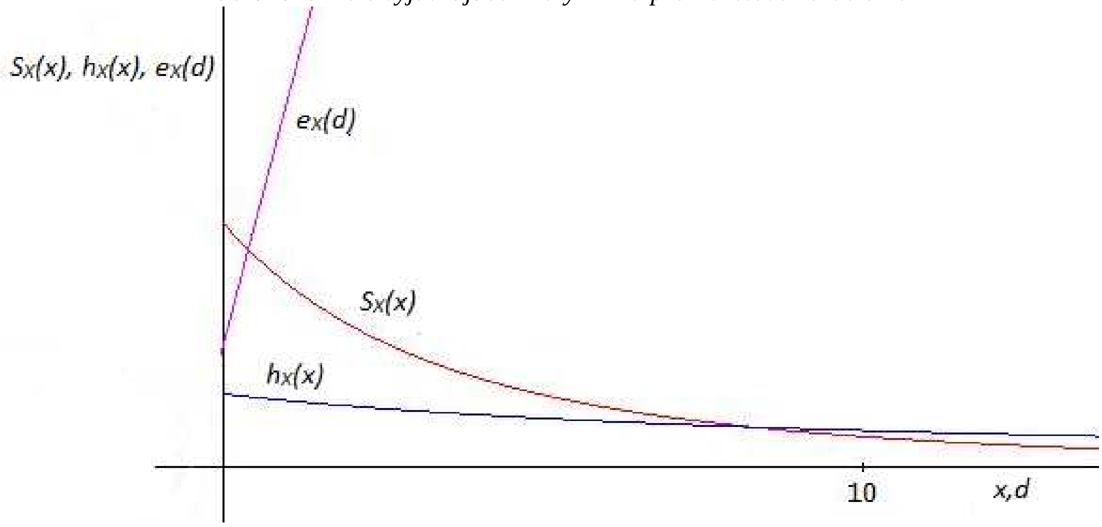
Funkcia strednej nadbytočnej straty je pre Paretovo rozdelenie podľa (2.7)

$$e_X(d) = E(X-d/X>d) = \frac{\int_d^\infty (x-d)f_X(x)\mathrm{dx}}{1-F_X(d)} = \frac{\int_d^\infty \frac{(x-d)\cdot 3000}{(x+10)^4}\mathrm{dx}}{\left(\frac{10}{10+d}\right)^3}$$

$$e_X(d) = 0{,}5d + 5$$

Nasledujúci obrázok (Obr. 1) znázorňuje priebeh funkcie prežitia, funkcie intenzity rizika a funkcie nadbytočnej straty pre Paretovo rozdelenie. Tieto funkcie sme odvodili v predošlej časti, kde sme zároveň dokázali, že Paretovo rozdelenie je rozdelenie s ťažkým pravým chvostom. Tento fakt je zrejmý aj z obrázku, v ktorom sú funkcia prežitia a funkcia intenzity rizika klesajúcimi funkciami a funkcia nadmernej straty naopak rastúcou funkciou.

Obr. 1:Funkcie vyjadrujúce miery rizika pre Paretovo rozdelenie



3.2 Funkcie určujúce mieru rizika pre niektoré ďalšie rozdelenia individuálnej škody

Tak ako sme odvodili funkcie popisujúce miery rizika pre Paretovo rozdelenie a dokázali, že dané rozdelenie je s ťažkým pravým chvostom, analogicky uvedený postup možno zopakovať pre ďalšie rozdelenie popisujúce výšku individuálnej škody. Rozdiel je iba v numerickej náročnosti. V nasledujúcej tabuľke uvádzame prehľad niektorých rozdelení aj s príslušnými funkciami.

Tabuľka 2: Funkcie popisujúce miery rizika pre niektoré rozdelenia

Rozdelenia s ľahkými chvostmi			
Rozdelenie	*Distribučná funkcia*	*Funkcia prežitia*	*Funkcia intenzity rizika*
Exponenciálne	$F_X(x) = 1 - e^{-\delta x}$	$S_X(x) = e^{-\delta x}$	$h_X(x) = \delta$
Gama	$F_X(x) = 1 - e^{-\beta x} \sum_{i=0}^{\alpha-1} \frac{(\beta x)^i}{i!}$	$S_X(x) = e^{-\beta x} \sum_{i=0}^{\alpha-1} \frac{(\beta x)^i}{i!}$	$h_X(x) = \frac{\beta^\alpha x^{\alpha-1}}{\Gamma(\alpha) \sum_{i=0}^{\alpha-1} \frac{(\beta x)^i}{i!}}$
Rozdelenia s ťažkými chvostmi			
Weibulovo	$F_X(x) = 1 - e^{-cx^\gamma}$	$S_X(x) = e^{-cx^\gamma}$	$h_X(x) = c\gamma x^{\gamma-1}$
Lognormálne	$F_X(x) \approx \Phi\left(\frac{\ln x - \mu}{\sigma}\right)$	$S_X(x) \approx 1 - \Phi\left(\frac{\ln x - \mu}{\sigma}\right)$	$h_X(x) = \frac{\frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\ln(x-\mu)^2}{\sigma}\right)}}{1 - \Phi\left(\frac{\ln x - \mu}{\sigma}\right)}$

Funkcia nadbytočnej straty pre dané rozdelenia nie je uvedená, pretože sa v niektorých prípadoch nedá všeobecne vyjadriť tak, ako napr. pri Paretovom rozdelení. V takýchto prípadoch sa používajú alternatívne postupy na jej vyjadrenie.

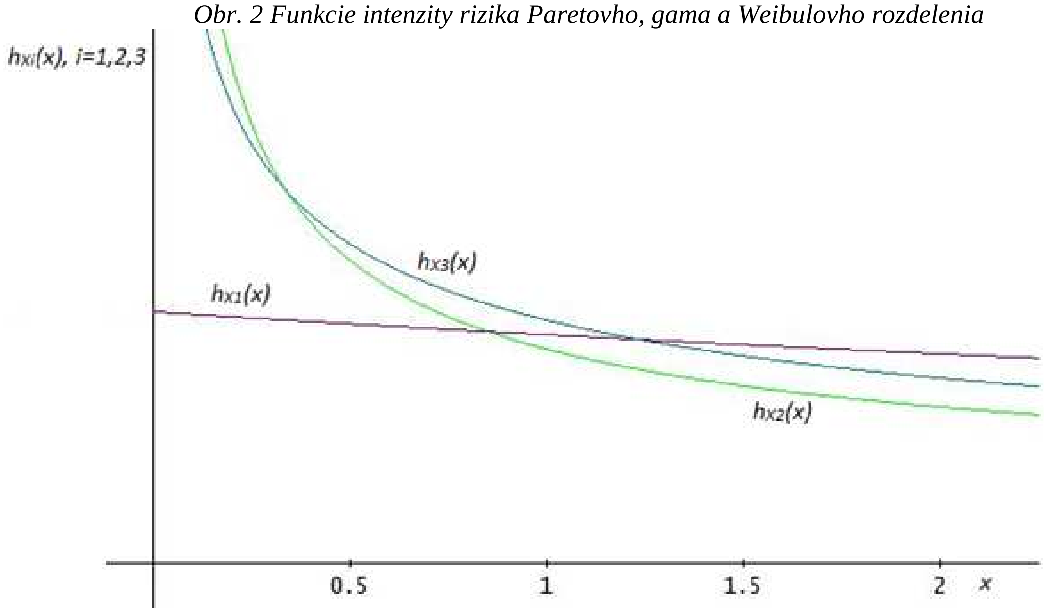
Na priblíženie mier rizika v tomto požijeme uvedené premenné s dvoma parametrami ako modely individuálnej výšky škody, pričom všetky premenné budú mať rovnakú strednú hodnotu a disperziu ako Paretovo rozdelenie uvedené v 3.1.

Tabuľka 3: Vyjadrenie mier rizika pre Paretovo rozdelenie (X_1), gama rozdelenie (X_2) a Weibulovo rozdelenie (X_3)

Rozdelenie	*Distribučná funkcia*	*Funkcia intenzity rizika*
$X_1 \sim Pa(3;10)$	$F_{X_1}(x) = 1 - \left(\frac{\beta}{\beta + x}\right)^{\alpha}$	$h_{X_1}(x) = \frac{\alpha}{x+\beta}$
$X_2 \sim \Gamma\left(\frac{1}{3};\frac{1}{15}\right)$	$F_{X_2}(x) = \int_0^t \frac{\left(\frac{1}{15}\right)^{\frac{1}{3}} x^{-\frac{2}{3}} e^{-\frac{1}{15}x}}{\Gamma\left(\frac{1}{3}\right)} \mathrm{dx}$	$h_{X_2}(x) = \frac{\frac{\left(\frac{1}{15}\right)^{\frac{1}{3}} x^{-\frac{2}{3}} e^{-\frac{1}{15}x}}{\Gamma\left(\frac{1}{3}\right)}}{1 - \int_0^t \frac{\left(\frac{1}{15}\right)^{\frac{1}{3}} x^{-\frac{2}{3}} e^{-\frac{1}{15}x}}{\Gamma\left(\frac{1}{3}\right)} \mathrm{dx}}$
$X_3 \sim W(0{,}478; 0{,}607)$	$F_{X_3}(x) = 1 - e^{-0{,}478x^{0{,}607}}$	$h_{X_3}(x) = \frac{0{,}478 \cdot 0{,}607 \cdot x^{0{,}607-1} \cdot e^{-0{,}478x^{0{,}607}}}{e^{-0{,}478x^{0{,}607}}}$

V tabuľke 3, v prípade Weibulovho rozdelenia sú hodnoty parametrov zaokrúhlené na tri desatinné miesta. Ich skutočná hodnota je $c = 0{,}4777582541$ a $\gamma = 0{,}6072483195$.

Obr. 2 Funkcie intenzity rizika Paretovho, gama a Weibulovho rozdelenia



Na obrázku 2 sú znázornené funkcie intenzity rizika pre rozdelenia uvedené v tabuľke 3. Obrázok je grafickým dôkazom, že Paretovo rozdelenie má najťažší koniec z porovnávaných rozdelení, nasleduje Weibulovo rozdelnie a gama rozdelenia klesá k nule najrýchlejšie. Dôkaz možno urobiť aj výpočtom limity podielu dvoch príslušných funkcií prežitia. Tento fakt korešponduje s hodnotami príslušných kvantilov. Napr. $VaR_{0,9}(X_1) = 11{,}5443468$, $VaR_{0,9}(X_2) = 14{,}7519996$ a $VaR_{0,9}(X_3) = 13{,}32779049$.

Teda možno konštatovať, že prezentované miery rizika, ktoré možno určiť v prípade znalosti distribučnej funkcie majú význam pri posudzovaní a riadení poistného rizika.

References

[1] Horáková, G., Mucha, V., 2008. *Teória rizika v poistení* I, II. Bratislava: Vydavateľstvo Ekonóm.

[2] Horáková, G., 2008. *Analýza rizikovosti portfólia poistných zmlúv neživotného poistenia*, 5. mezinárodní konference Rízení a modelovaní finančních rizik, VŠB-TU Ostrava, Ekonomická fakulta, Katedra financí.

[3] Klugman, S. A., Panjer, H. H., Willmot G. E., 2008: *Loss Models* (From Data to Decisions). New York: John Wiley & Sons.

[4] Panjer, H. H., 2006. *Operational Risk*. A John Wilye & Sons, Interscience Publication.

[5] Simanová, B., 2012. *Distribučná funkcia ako nástroj riadenia rizika*, 2. vedecký seminár doktorandov vo vednom odbore Kvantitatívne metódy v ekonómii, EU Bratislava, Fakulta hospodárskej informatiky.