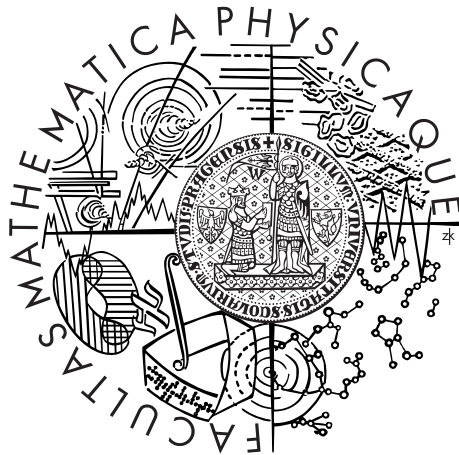


Univerzita Karlova v Praze
Matematicko-fyzikální fakulta

DIPLOMOVÁ PRÁCE



Bc. Eva Mináriková

Modelování operačního rizika

Katedra pravděpodobnosti a matematické statistiky

Vedoucí diplomové práce: RNDr. Lucie Mazurová, Ph.D.

Studijní program: Matematika

Studijní obor: Finanční a pojistná matematika

Praha 2012

V prvom rade by som sa chcela poďakovať vedúcej diplomovej práce RNDr. Lucii Mazurovej, Ph.D. za samotné vedenie tejto práce a poskytnutie cenných rád. Veľmi pekne ďakujem všetkým, ktorí pri mne stáli a podporovali ma počas celého môjho vysokoškolského štúdia.

Prehlasujem, že som túto diplomovú prácu vypracovala samostatne a výhradne s použitím citovaných prameňov, literatúry a ďalších odborných zdrojov.

Beriem na vedomie, že sa na moju prácu vzťahujú práva a povinnosti vyplývajúce zo zákona č. 121/2000 Sb., autorského zákona v platnom znení, predovšetkým skutočnosť, že Univerzita Karlova v Prahe má právo na uzavretie licenčnej zmluvy o použití tejto práce ako školského diela podľa §60 odst. 1 autorského zákona.

V Prahe dňa 26. novembra 2012

Bc. Eva Mináriková

Názov práce: Modelování operačního rizika

Autor: Bc. Eva Mináriková

Katedra: Katedra pravděpodobnosti a matematické statistiky

Vedúci diplomovej práce: RNDr. Lucie Mazurová, Ph.D., Katedra pravděpodobnosti a matematické statistiky

Abstrakt: V predloženej práci sa najprv zoznámime s pojmom operačné riziko, spôsobom, akým je definované direktívami *Basel II* a *Solventnosť II* a následne metódami stanovenými týmito direktívami na výpočet *kapitálovej požiadavky na straty plynúce z operačného rizika*. V druhej časti práce sa zameriame na spôsob, akým je možné straty z operačného rizika modelovať. Predstavíme si *teóriu extrémnych hodnôt*, ktorá popisuje možnosti, akými pristúpiť k modelovaniu rozdelenia dát, ktoré nastávajú zriedkavo, ale nadobúdajú vysoké hodnoty, čo je charakteristickou vlastnosťou dát o stratách z operačného rizika. Zameriame sa predovšetkým na prístup založený na prekročení medze, pri ktorom rozdelenie excesov modelujeme pomocou *zovšeobecneného Paretovho rozdelenia*. Poznatky z tejto teórie v práci uplatníme pri modelovaní rozdelenia nasimulovaných dát. Posledným krokom je testovanie úspešnosti modelovania rozdelenia dát o stratách pomocou predstavených metód.

Kľúčové slová: Operačné riziko, Teória extrémnych hodnôt, Zovšeobecnené rozdelenie extrémnych hodnôt, Prekročenie medze, Zovšeobecnené Pareto rozdelenie

Title: Operational risk modelling

Author: Bc. Eva Mináriková

Department: Department of Probability and Mathematical Statistics

Supervisor: RNDr. Lucie Mazurová, Ph.D., Department of Probability and Mathematical Statistics

Abstract: In the present thesis we will firstly familiarize ourselves with the term of operational risk, it's definition presented in the directives *Basel II* and *Solvency II*, and afterwards with the methods of calculation *Capital Requirements for Operational Risk*, set by these directives. In the second part of the thesis we will concentrate on the methods of modelling operational loss data. We will introduce the *Extreme Value Theory* which describes possible approaches to modelling data with significant values that occur infrequently; the typical characteristic of operational risk data. We will mainly focus on the model for threshold exceedances which utilizes *Generalized Pareto Distribution* to model the distribution of those excesses. The theoretical knowledge of this theory and the appropriate modelling will be applied on simulated loss data. Finally we will test the ability of presented methods to model loss data distributions.

Keywords: Operational Risk, Extreme Value Theory, Generalized Extreme Value Distribution, Threshold Exceedances, Generalized Pareto Distribution

Obsah

Úvod	1
1 Operačné riziko	2
1.1 Úvod	2
1.2 Operačné riziko v bankovníctve	3
1.2.1 Basel II	3
1.2.2 Metódy výpočtu podľa Basel II	4
1.3 Operačné riziko v poisťovníctve	8
1.3.1 Solventnosť II	8
1.3.2 Metódy výpočtu podľa Solventnosti II	9
2 Teória extrémnych hodnôt	13
2.1 Úvod	13
2.2 Prístup založený na blokových maximách	13
2.2.1 Zovšeobecnené rozdelenie extrémnych hodnôt	14
2.2.2 Metóda blokových maxim	16
2.2.3 Úroveň návratu a doba návratu	17
2.3 Prístup založený na prekročení medze	17
2.3.1 Zovšeobecnené Paretovo rozdelenie	17
2.3.2 Modelovanie strát prekračujúcich medzu	21
2.3.3 Testovanie úspešnosti modelovania excesov	22
2.3.4 Správanie sa rozdelenia excesov pri zvýšení medze	29
2.3.5 Hillova metóda	31
2.3.6 Modelovanie chvostu rozdelenia pôvodných dát	34
2.3.7 Odhadovanie vysokých kvantilov	38
3 Záver	41

Úvod

Každý z nás sa každý deň stretáva s rôznymi rizikami. Takmer každá činnosť človeka je rizikami zaťažená. Napriek tomu, že pojem riziko je často používaný, definovať ho vôbec nie je jednoduché. Čo si teda pod slovom riziko máme predstaviť? Podľa [7] sa slovo riziko vzťahuje k situáciám, v ktorých je možné, ale nie isté, že nastane nejaká neželaná udalosť. Vo všeobecnosti sa dá tiež povedať, že so slovom riziko sa spájajú viaceré nekonzistentné a nejednoznačné významy, teda jeho definícia sa líši podľa oblasti jeho používania.

Nielen ľudia, ale aj podniky, a teda aj finančné inštitúcie, sú spojené s rizikami. Ide prevažne o iné riziká, ale podstata zostáva rovnaká. V prípade rizík, ktorým sa dá predísť, sa toho snažia docieľiť a v prípade rizík, ktorým sa aj pri najlepšej snahe zabrániť nedá, sa snažia na ne čo najviac pripraviť a tým sa ochrániť pred následkami strát z rizikových javov plynúcich.

V poslednej dobe sa vo sfére finančných inštitúcií začal čoraz viac skloňovať najmä pojem operačné riziko. Práve preto sa mu v tejto práci budeme venovať. Tento záujem súvisí aj s jeho zaradením medzi základné druhy rizík v direktívach Basel II a Solventnosť II a so snahou o jeho systematické a regulované riadenie.

V prvej časti tejto práce sa budeme venovať práve operačnému riziku v zmysle, v akom ho chápú tieto direktívy, požiadavkám na dáta o stratách z operačného rizika, kvantitatívnym a kvalitatívnym požiadavkám na výpočet solventnosti a predstavíme si jednotlivé metódy výpočtu kapitálovej požiadavky na straty z neho plynúce.

Náplňou tejto diplomovej práce však nie je vytvorenie kompletného interného modelu na výpočet kapitálovej požiadavky na solventnosť. Naopak, práca si za cieľ dáva popísať jeden z možných prístupov k modelovaniu rozdelenia dát o stratách z operačného rizika, ktoré vstupuje do interného modelu. K tomu slúži teória extrémnych hodnôt popísaná v druhej časti tejto práce.

V rámci teórie extrémnych hodnôt si predstavíme v krátkosti prístup založený na blokových maximách, ktorý sa však v praxi na modelovanie dát z operačného rizika priveľmi nepoužíva. Slúži však ako teoretický základ pre prístup k modelovaniu rozdelenia dát založenom na prekročení medze, ktorému sa budeme venovať podrobnejšie.

V teórii extrémnych hodnôt sa na modelovanie rozdelenia dát používa zovšeobecnené Paretovo rozdelenie, ktoré si taktiež predstavíme. Popísané teoretické prístupy k modelovaniu sú všetky otestované pri praktickom použití na nasimulovaných dátach v závere tejto práce.

Kapitola 1

Operačné riziko

1.1 Úvod

Každý podnik sa stretáva s určitými druhmi operačného rizika nezávisle na poli jeho pôsobnosti. Takými rizikami sú napríklad narušenie činnosti z dôvodu podvodu, zlyhania výpočtovej techniky, povodne, zemetrasenia, či iných prírodných katastrof. Avšak existujú aj viac špecifické druhy operačného rizika, ktoré sa týkajú iba podnikov s určitým zameraním. Ďalej sa budeme venovať predovšetkým takým druhom operačného rizika, ktorým sú vystavené najmä inštitúcie pôsobiace vo finančnom sektore.

Inštitúcie pôsobiace vo finančnom sektore sa s operačným rizikom stretávajú od samotného vzniku tohto sektoru, avšak aktívne ho začali sledovať a merať iba pred niekoľkými rokmi. Cieľom týchto kontrol a meraní je predovšetkým zachovanie solventnosti a ratingu podniku, zabezpečenie jeho prežitia, ale aj ďalšie. Toto je dôvod prečo sa regulácia operačného rizika stala podstatnou súčasťou podnikovej politiky v inštitúciách poskytujúcich finančné služby, predovšetkým v bankách a poisťovniach.

Napriek tomu, že pojem operačné riziko je skloňovaný čoraz častejšie, jeho definícia stále nie je jednoznačná. Vo všeobecnosti je možné povedať, že ide o riziko spojené s operáciami v podniku. Avšak ani jeden z týchto pojmov nie je dostatočne jasný. V tomto kontexte sa dá riziko chápať ako možnosť, že niečo nedopadne podľa očakávaní a operáciu je možné chápať ako funkciu podniku spojenú s jeho činnosťou. Operačným rizikom môžeme taktiež myslieť riziko, ktoré nie je ani kreditné, ani trhové. Bazilejský výbor pre bankový dohľad (*Basel Committee on Banking Supervision*) chápe operačné riziko ako riziko spojené s chybami v procese vnútornej kontroly a s chybami v podnikovej politike. Podľa [5], str. 5, Asociácia britských bankérov (*The British Bankers Association*) definuje operačné riziko nasledovne:

Definícia 1.1.1. *Operačné riziko je riziko spojené s chybami ľudí, nečestnými kriminálnymi aktivitami, neadekvátnymi procedúrami a kontrolami. Ďalej riziko zapríčinené technickými chybami, zlyhaniami systému, riziko "nebankové", ale vyplývajúce z obchodných rozhodnutí, právne riziko, riziko vyplývajúce z obchodných vzťahov, z nenaplnenia regulačnej požiadavky alebo zo škodlivého vplyvu na reputáciu banky, ďalej ide o riziko spojené s prírodnými katastrofami a teroristickými útokmi.*

Najväčším problémom, ktorý súvisí s reguláciou operačného rizika, je nedostatok verejne prístupných dát. To je v kontraste s trhovým a kreditným rizikom, o ktorých máme k dispozícii dostatok dát a informácií. Situácia sa začala zlepšovať postupným zhromažďovaním a zdieľaním historických dát o stratách spojených s operačným rizikom. V súčasnosti mnohé organizácie pracujú na budovaní databáz obsahujúcich dáta o historických udalostiach a škodách súvisiacich s operačným rizikom a súčasne budujú systémy, ktoré zaznamenávajú a reportujú nové udalosti súvisiace s týmto rizikom. Podľa [5], jedným z hlavných zdrojov takýchto dát sú v súčasnosti databázy Kvantitatívnych dopadových štúdií Bazilejského výboru (*Quantitative Impact Studies of Basel Committee*) a databázy Federálnej rezervnej banky v Bostone (*Federal Reserve Bank of Boston*). Ďalším problémom je spôsob zjednotenia dostupných informácií z rôznych podnikov a odvetví pre zostavovanie modelu konkrétneho podniku. Tento problém je tiež známy pod pojmom "scaling" a je ďalej popísaný v [5].

1.2 Operačné riziko v bankovníctve

V histórii bankovníctva sa vyskytlo viacero prípadov, kedy riziko, dnes známe ako operačné, stálo za bankrotom, alebo v lepších prípadoch, za veľkými problémami finančných inštitúcií. Jedným z odstrašujúcich a zároveň jedným z najznámejších príkladov je prípad pádu banky Barings (*Barings Bank*) v roku 1995, za ktorým stál jeden z jej zamestnancov, *Nick Leeson*, ktorý prostredníctvom neautorizovaného špekulatívneho investovania s *future* kontraktmi stratil približne 960 miliónov Eur. Tento príklad, popísaný aj v [5], je jedným z mnohých dôvodov, prečo je potrebné operačné riziko monitorovať.

Iniciatívy v rámci vytvorenia medzinárodne platných požiadaviek na riadenie operačného rizika v bankách sa ujal Bazilejský výbor pre bankový dohľad, keď do svojho regulátórneho návrhu *Basel II: International Convergence of Capital Measurement and Capital Standards: A Revised Framework*, známeho pod skráteným názvom *Basel II*, pridal k preberaným rizikám, trhovému a kreditnému, riziko operačné.

Nasledujúci text tejto sekcie je založený na textoch [5] a [4].

1.2.1 Basel II

Basel II je trojpilierový koncept vyvinutý Bazilejským výborom pre bankový dohľad zaoberajúci sa otázkou určenia minimálnej kapitálovej požiadavky pre banky. Ide o druhý súbor odporúčení pre bankové právo a reguláciu. Cieľom Bazilejského výboru bolo vytvoriť medzinárodný štandard, ktorý by bol využívaný pri tvorbe samotných regulácií o minimálnej kapitálovej požiadavke. Dokument Basel II má za cieľ určiť metodiku výpočtu, ktorá by viedla k presnejšiemu a komplexnejšiemu zobrazeniu solventnosti bánk. Jedným z rozdielov oproti prvému vydaniu je pridanie novej kategórie rizík, a to rizika operačného. Hlavným rozdielom v porovnaní s kreditným a trhovým rizikom, ktoré sú merané a regulované už dlhší čas je, že operačné riziko neprináša pre banku žiadnu výhodu. Vzniká pri nesprávnom fungovaní každodennej činnosti banky. Niektoré druhy operačného rizika majú banky pod dohľadom už dlhšiu dobu, s inými sa zoznámili len nedávno. Je samozrejmosťou, že banky sa snažia eliminovať operačné riziko, ale aj napriek

ich najlepšej snahe sa operačnému riziku nevedia úplne vyhnúť. To bol jeden z dôvodov zahrnutia operačného rizika do direktívy. V rámci [4] na strane 149 je operačné riziko definované nasledovne:

Definícia 1.2.1. *Operačné riziko je riziko straty vyplývajúcej z nedostatočnosti, alebo zlyhania vnútorných procesov, osôb, alebo systémov a z vonkajších udalostí.*

Poznámka 1.2.1. *Podľa definície 1.2.1 operačné riziko zahŕňa právne riziko, ale nespadá pod neho riziko strategické a reputačné.*

Dokument Basel II sa skladá z troch pilierov. Prvý pilier sa zaoberá výpočtom minimálnych kapitálových požiadaviek pre kreditné, trhové a operačné riziko. Kapitálová požiadavka pre operačné riziko je explicitná a samostatná. Koncept nestanovuje iba jednu možnosť výpočtu, ale udáva hneď niekoľko spôsobov. Jedným z nich je budovanie vlastných vnútorných modelov podnikového profilu operačného rizika. Táto voľnosť pri výpočte minimálnej kapitálovej požiadavky stojí za potrebou rozsiahleho systému dohľadu. Práve ním sa zaoberajú ďalšie dva piliere konceptu.

Druhý pilier rieši otázku dohľadu nad kapitálovou primeranosťou a interným systémom riadenia rizík. Tento pilier bankám nariaďuje sledovanie všetkých udalostí týkajúcich sa operačného rizika a zapojenie vnútorných kontrolných procesov, ktoré budú transparentné pre bankový dohľad. Tieto nariadenia majú pomôcť bankám, aj dohľadu lepšie pochopiť minulé a prípadné budúce oblasti strát plynúce z udalostí súvisiacich s operačným rizikom, čo môže redukovať množstvo a závažnosť týchto udalostí.

Tretí pilier vyžaduje od inštitúcií, aby zverejňovali dôležité informácie ostatným účastníkom trhu. Tento pilier je často označovaný ako trhová disciplína.

Dá sa povedať, že cieľom Basel II je povzbudiť banky k tomu, aby venovali viac pozornosti samotnému riziku, jeho meraniu, monitorovaniu, zmierňovaniu a aby o ňom zverejňovali požadované informácie.

1.2.2 Metódy výpočtu podľa Basel II

Teraz si popíšeme základné metódy výpočtu kapitálovej požiadavky pre operačné riziko. V [4] sú predstavené tri základné metódy so stúpajúcou náročnosťou výpočtu a citlivosťou na riziko, a to:

1. Základný indikátor (*Basic Indicator Approach*),
2. Štandardizovaný prístup (*Standardised Approach*),
3. Pokročilé prístupy (*Advanced Measurement Approaches (AMA)*).

Bankám je odporučené, aby sa v spektre metód pohybovali v smere k náročnejším metódam na základe vyvinutia sofistikovanejších výpočetných systémov a praktík. Bazilejský výbor určuje pre prechod k štandardizovanému prístupu a pre používanie pokročilých prístupov kvalifikačné podmienky definované v [4]. Pri prechode na štandardizovaný prístup je od banky požadované, aby členovia jej správnej rady a senior manažéri boli aktívne zapojení do dohľadu nad riadením operačného rizika. Ďalej je požadovaná existencia zdravého, nezávislého systému riadenia operačného rizika a venovanie dostatočných finančných zdrojov

na použitie prístupu v majoritných líniach podnikania, v oblasti kontroly a auditu a ďalšie. V prípade, že banka používa k výpočtu kapitálovej požiadavky pre operačné riziko štandardizovaný prístup a chce prejsť na pokročilé prístupy, bude od nej vyžadované, aby výsledky merania operačného rizika boli zahrnuté do každodenného riadenia rizika, pravidelné hlásenie vystaveniu sa operačnému riziku a zároveň hlásenie strát z neho plynúcich správnej rade a senior manažérom, dobre zdokumentovaný systém riadenia operačného rizika, pravidelné preskúmanie riadiacich a meracích procesov interným a externým auditorom, početné požiadavky na históriu údajov o stratách a ďalšie.

Najzákladnejšou metódou výpočtu je metóda používajúca základný indikátor. Podľa [4] sa kapitálová požiadavka C_{BIA}^t pre rok t na základe tejto metódy vypočíta použitím vzťahu uvedeného aj v [2] na strane 465:

$$C_{BIA}^t = \frac{1}{Z_t} \sum_{i=1}^3 \alpha \max(G^{t-i}, 0),$$

kde

$$Z_t = \sum_{i=1}^3 \mathbb{I}[G^{t-i} > 0] \quad (1.1)$$

a G^{t-i} je hrubý ročný výnos v roku $t - i$ a $\alpha = 15\%^1$.

Poznámka 1.2.2. Indikátor $\mathbb{I}[y > x]$ pre $x, y \in \mathbb{R}$ z (1.1) definujeme nasledujúcim spôsobom:

$$\mathbb{I}[y > x] = \begin{cases} 1 & \text{pre } y > x, \\ 0 & \text{pre } y \leq x. \end{cases}$$

Túto definíciu budeme používať v celom texte tejto diplomovej práce.

To znamená, že kapitálová požiadavka C_{BIA}^t sa rovná fixným pätnástim percentám z priemerného hrubého ročného výnosu za posledné tri roky. Priemer sa vypočíta iba z jeho kladných hodnôt.

O niečo náročnejšou metódou je štandardizovaný prístup. Aby tento prístup banka mohla pri výpočte kapitálovej požiadavky pre operačné riziko používať musí splniť požiadavky uvedené v predchádzajúcom texte. Pri tomto prístupe k výpočtu kapitálovej požiadavky je činnosť banky rozdelená do ôsmich línií podnikania určených direktívou Basel II (viď tabuľka 1.1). Pre každú z týchto línií sa vypočíta vlastná čiastočná kapitálová požiadavka C_j^t pre rok t ako

$$C_j^t = \beta_j G_j^t,$$

kde $j = 1, \dots, 8$, G_j^t je prislúchajúci hrubý výnos v roku t k línií podnikania j a β_j je príslušný faktor², ktorého hodnoty pre jednotlivé línie podnikania sú zaznamenané v tabuľke 1.1. Celková kapitálová požiadavka C_{SA}^t pre rok t sa podľa [4] vypočíta ako trojročný priemer zo súčtu čiastočných kapitálových požiadaviek

$$C_{SA}^t = \frac{1}{3} \sum_{i=1}^3 \max \left[\sum_{j=1}^8 C_j^{t-i}, 0 \right], \quad (1.2)$$

¹Hodnota parametra α bola určená na základe dopadových štúdií Bazilejského výboru pre bankový dohľad.

²Hodnoty β_j sú určené rovnako ako hodnoty parametra α , na základe dopadových štúdií Bazilejského výboru pre bankový dohľad.

kde C_j^{t-i} je čiastočná kapitálová požiadavka v roku $t - i$ za líniu podnikania j . Vzťah (1.2) nájdeme v [2] na strane 465.

<i>Línia podnikania</i>	<i>Hodnota β_j</i>
1. Podnikové financovanie	$\beta_1 = 18\%$
2. Obchodovanie na finančných trhoch	$\beta_2 = 18\%$
3. Retailové bankovníctvo	$\beta_3 = 12\%$
4. Komerčné bankovníctvo	$\beta_4 = 15\%$
5. Platby a zúčtovanie	$\beta_5 = 18\%$
6. Služby z poverenia	$\beta_6 = 15\%$
7. Riadenie aktív	$\beta_7 = 12\%$
8. Retailové maklérsstvo	$\beta_8 = 12\%$

Tabuľka 1.1: Línie podnikania a príslušné hodnoty beta faktoru podľa [4].

Najnáročnejšou metódou výpočtu kapitálovej požiadavky pre operačné riziko je pokročilý prístup. Aj v tomto prípade musí banka splniť podmienky stanovené v direktíve Basel II, tiež uvedené v predchádzajúcom texte. Hlavnou myšlienkou pokročilého prístupu určovania minimálnej kapitálovej požiadavky z operačného rizika je, aby banky mali kapitálové rezervy, ktoré korešpondujú s ich individuálnym rizikovým profilom, teda banky majú držať o to viac vlastných finančných prostriedkov, čím väčšiemu riziku sú vystavené. Na rozdiel od prvých dvoch prístupov, ktoré sú dané explicitným vzorcom, pokročilé prístupy sú dané iba všeobecnými pokynmi. V rámci tohto prístupu je bankám dovolené vytvoriť si vlastné sofistikované vnútorné modely skutočných rizík zohľadňujúce akékoľvek stratégie banky, ktoré zmierňujú dopady týchto rizík. V tomto prípade je od bánk vyžadované vynaloženie značných finančných zdrojov do systému riadenia operačného rizika.

Bazilejský výbor pre bankový dohľad nešpecifikuje žiadne konkrétne modely, ani predpoklady pre pokročilý prístup, ale určuje pomerne veľké množstvo kvalitatívnych požiadaviek, ktoré musia byť splnené. Kvôli voľnosti, ktorá je v rámci tohto prístupu poskytnutá, banky musia pred zavedením svojho interného modelu preukázať, že je dostatočne spoľahlivý, vierohodný, detailný a vhodný na odhad neočakovaných strát. Ďalej musia ukázať, že je založený na kombinácii vnútorných a vonkajších dát o stratách a testovaní scenára. Bankou navrhnutá pokročilá metóda merania operačného rizika musí byť pred jej používaním schválená kontrolným orgánom a počas celej doby používania metódy musí podliehať pravidelným kontrolám kvality príslušnými orgánmi dohľadu. Ďalej je požadované pravidelné hlásenie vystavenia sa operačnému riziku a strát s ním spojených. Systém riadenia operačného rizika musí byť dobre zdokumentovaný a pravidelne posudzovaný ako vlastnými, tak aj externými auditori.

Ďalšími požiadavkami, ktoré musia byť splnené, sú požiadavky kvantitatívne. Výbor nešpecifikuje prístup ani pravdepodobnostné rozdelenie, ktoré má byť použité, ale banka musí byť schopná preukázať, že ňou zvolený prístup postihuje možnosť výskytu extrémne veľkých škôd, teda, že ide o rozdelenie s dostatočne ťažkým chvostom. Kapitál vypočítaný týmto prístupom má pokrývať riziko straty v priebehu jedného ročného obdobia pri hladine spoľahlivosti 99,9%.

Pri používaní pokročilých prístupov dáta o škodách z operačného rizika majú byť triedené podľa línií podnikania vymenovaných v tabuľke 1.1 a podľa siedmich

základných typov škodových udalostí, ktoré Bazilejský výbor v [4] vymedzuje nasledovne:

- interný podvod (neautorizované aktivity, krádeže a podvody, daňové úniky),
- externý podvod (krádeže informácií, podvody a zabezpečenie systému),
- zamestnávanie a bezpečnosť na pracovisku (vzťahy k zamestnancom, bezpečnosť prostredia, diskriminácia),
- klienti, produkty a obchodné praktiky (manipulácia na trhu, produktové chyby, nesprávny obchod),
- poškodzovanie hmotného majetku (prírodné katastrofy, terorizmus a vandalizmus),
- prerušenie činnosti a zlyhanie systému (zlyhanie softvéru a hardvéru),
- transakcie, obchodní partneri a predajcovia (transakcie, monitorovanie a reporting, získavanie zákazníkov, dokumentácia, správa klientských účtov, obchodní partneri, predajcovia a dodávatelia).

Teraz si ukážeme jeden z možných spôsobov výpočtu kapitálovej požiadavky pre operačné riziko podľa [2], ktorý je typický pre pokročilé prístupy. Majme dáta o škodách z operačného rizika za posledných $T \geq 5$ rokov

$$\{X_k^{t-i,b,l}\}_{i=1,\dots,T; b=1,\dots,8; l=1,\dots,7; k=1,\dots,N^{t-i,b,l}},$$

kde $X_k^{t-i,b,l}$ je k -ta škoda typu l v línii podnikania b v roku $t-i$ a $N^{t-i,b,l}$ označuje počet škôd typu l v línii b v roku $t-i$. Výška škôd $L^{t-i,b}$ v línii podnikania b v roku $t-i$ je súčtom všetkých škôd v línii b cez jednotlivé škodové typy

$$L^{t-i,b} = \sum_{l=1}^7 \sum_{k=1}^{N^{t-i,b,l}} X_k^{t-i,b,l}.$$

Potom celková výška škody za rok $t-i$ je daná ako súčet celkovej výšky škôd v jednotlivých líniiach podnikania cez všetky línie

$$L^{t-i} = \sum_{b=1}^8 L^{t-i,b}.$$

V rámci pokročilého prístupu je často problematické odhadnúť rozdelenie škôd L^t a spočítať vhodnú mieru rizika pre odhadnuté rozdelenie. Mieru rizika na hladine spoľahlivosti α , kde $\alpha \in [0, 99; 0, 999]$, označíme ρ_α . Kapitálová požiadavka C_{AMA}^t pre rok t sa vypočíta pokročilým prístupom ako

$$C_{AMA}^t = \rho_\alpha (L^t).$$

Keďže štruktúru združeného rozdelenia škôd spravidla nepoznáme, stanovujeme celkovú kapitálovú požiadavku C_{AMA}^t súčtom cez všetky línie podnikania

$$C_{AMA}^t = \sum_{b=1}^8 \rho_\alpha (L^{t,b}).$$

Za mieru rizika ρ_α sa často volí hodnota v riziku (*Value at Risk*) definovaná podľa [2], str. 38 v definícii 1.2.2.

Definícia 1.2.2. *Bud' $\alpha \in (0, 1)$, potom hodnotu*

$$\text{VaR}_\alpha = \inf\{l \in \mathbb{R} : P(L > l) \leq 1 - \alpha\} = \inf\{l \in \mathbb{R} : F_L(l) \geq \alpha\}$$

značíme ako hodnotu v riziku.

Poznámka 1.2.3. *Vidíme, že VaR_α je α -kvantil uvažovaného rozdelenia.*

Bazilejský výbor sa v direktíve Basel II nezaobrá iba samotným výpočtom kapitálovej požiadavky pre operačné riziko, ale aj požiadavkami o zbere a zázname dát o operačnom riziku. Výbor hovorí, že ide o základnú prerekvizitu vývoja funkčného a spoľahlivého systému na meranie tohto rizika. Pri stanovení kapitálovej požiadavky pokročilým prístupom majú byť použité interné dáta za minimálne päť rokov. Vnútorne dáta majú byť súhrnné a banka musí byť schopná ospravedlniť absenciu akejkoľvek materiálnej aktivity a vystavenie sa riziku tým, že by to nemalo dopad na celkový odhad rizika. Každá banka by si mala stanoviť dolnú medzu pre výšku škôd, typicky okolo 10000 Eur. Koncept dolnej medze je dôležitý v štatistickej analýze strát z operačného rizika, ktorou sa budeme zaoberať v ďalšom texte. Banky používajúce pokročilý prístup k určovaniu kapitálovej požiadavky sú povinné používať relevantné externé dáta, predovšetkým v prípade, že existuje predpoklad, že banka je vystavená zriedkavým, ale potencionálne vážnym rizikám. Relevantné externé dáta majú obsahovať dáta o skutočných veľkostiach strát, informácie o príčine a okolnostiach vzniku stratovej udalosti a ďalšie informácie, ktoré by mohli ostatným bankám pomôcť stanoviť relevantnosť daných dát.

1.3 Operačné riziko v poisťovníctve

Poistný sektor v oblasti vývoja kapitálových požiadaviek a v oblasti regulácie operačného rizika zaostáva za bankovým sektorom. Táto situácia sa však začala meniť, keď Medzinárodná asociácia orgánov dohľadu nad poisťovníctvom (*International Association of Insurance Supervisors*) iniciovala vytvorenie systému regulácie poisťovníctva na medzinárodnej úrovni. Na území Európskej únie Európska komisia (*European Commission*) prišla s projektom Solventnosť II (*Solvency II*), ktorý je reformou solventnostných štandardov pre poisťovníctvo aplikovateľných v rámci všetkých európskych poisťovní. Nasledujúce sekcie týkajúce sa direktívy Solventnosť II (*v plnom znení Directive 2009/138/EC of the European Parliament and of the Council of 25 November 2009 on the taking-up and pursuit of the business of Insurance and Reinsurance*) sú založené na texte [6].

1.3.1 Solventnosť II

Direktíva Solventnosť II stanovuje a harmonizuje reguláciu európskeho poisťovníctva. Zavádza systém dohľadu, ktorý zohľadňuje rizikový profil poisťovne a zjednocuje postupy s direktívou Basel II. Projekt má zohľadniť zmenu rizikových podmienok v oblasti poisťovníctva, ako sú napríklad nárast počtu katastrof, nárast volatility kapitálových trhov a rastúci význam operačného rizika. Hlavným cieľom je zvýšenie ochrany poistených a príjemcov poistného plnenia, zvýšenie stability celého odvetvia, a to vyššou kvalitou ohodnocovania rizík a efektívnou

alokáciou kapitálu. Aplikácia konceptu Solventnosť II by mala viesť k vytvoreniu jednotného trhu poisťných služieb v rámci Európskej únie.

V [6] je použitá na definovanie operačného rizika rovnaká definícia akú používa direktíva Basel II, teda definícia 1.2.1. Aj v tomto prípade operačné riziko zahŕňa právne riziko, ale nezahŕňa riziko strategické a reputačné.

Podobne ako Basel II, aj Solventnosť II je založená na troch pilieroch, z ktorých každý má samostatnú funkciu. Prvý pilier obsahuje kvantitatívne požiadavky na poisťovňu, pričom základom je adekvátne zobrazenie expozície poisťnému, kreditnému, trhovému a operačnému riziku. V tomto pilieri sú predstavené dva typy kapitálových požiadaviek, a to:

1. Kapitálová požiadavka na solventnosť (*Solvency Capital Requirement*),
2. Minimálna kapitálová požiadavka (*Minimum Capital Requirement*).

Kapitálová požiadavka na solventnosť je stanovená na základe štandardného, alebo interného modelu. Jej nesplnenie má za následok opatrenie dohľadného orgánu, ktoré je závislé na závažnosti nedostatku. Pokiaľ vlastné prostriedky poisťovne nedosahujú ani minimálnu kapitálovú požiadavku, je ohrozené jej ďalšie fungovanie a dochádza k významnému zvýšeniu rizika voči poisteným. V takomto prípade dochádza k okamžitému zakročeniu orgánu dohľadu, a to k odobratiu povolenia k poisťovacej činnosti, alebo nasadeniu ozdravného plánu. Minimálna kapitálová požiadavka sa počíta štvrťročne použitím jednoduchého a auditovateľného spôsobu, alebo použitím absolútneho minima, ktoré je v Solventnosti II stanovené zvlášť pre odvetvie životného a neživotného poistenia.

Druhý pilier má za cieľ stanoviť kvalitatívne požiadavky na vlastný systém riadenia rizík poisťovní a pravidlá pre efektívnu činnosť dohľadu pri kontrole splnenia požadovaných kvantitatívnych a kvalitatívnych požiadaviek poisťovňou.

Tretí pilier sa zaoberá trhovou disciplínou. Zameriava sa na požiadavky zverejňovania informácií pre verejnosť, orgány dohľadu a ostatných účastníkov trhu.

1.3.2 Metódy výpočtu podľa Solventnosti II

Od poisťovní je vyžadované držanie vlastného kapitálu vo výške, ktorá presahuje kapitálovú požiadavku na solventnosť. Direktíva Solventnosť II odporúča, aby poisťovne počítali výšku kapitálovej požiadavky na solventnosť minimálne jedenkrát ročne a aby s rovnakou frekvenciou o výsledku výpočtu informovali príslušné orgány dohľadu.

V nasledujúcom texte si predstavíme metódy výpočtu kapitálovej požiadavky na solventnosť. V direktíve [6] sú popísané dva prípustné spôsoby výpočtu tejto požiadavky, a to:

1. Štandardná formula (*Standard Formula*),
2. Interný model (*Internal Model*).

Najprv si popíšeme štandardnú formulu pre výpočet kapitálovej požiadavky na solventnosť. Podľa tohto modelu je výpočet kapitálovej požiadavky na solventnosť SCR založený na princípe "bottom-up", teda celková kapitálová požiadavka

vznikne nasčítaním viacerých čiastočných kapitálových požiadaviek. Štandardná formula na výpočet SCR má podľa [6] tvar

$$\text{SCR} = \text{SCR}_{\text{Basic}} + \text{SCR}_{\text{OperationalRisk}} - \text{Adj}, \quad (1.3)$$

kde $\text{SCR}_{\text{Basic}}$ je základná kapitálová požiadavka na solventnosť (*Basic Solvency Capital Requirement*), $\text{SCR}_{\text{OperationalRisk}}$ (*Solvency Capital Requirement for Operational Risk*) je čiastočná kapitálová požiadavka z operačného rizika a Adj je úprava zohľadňujúca kapacitu technických rezerv a odložených daní.

Základnú kapitálovú požiadavku na solventnosť $\text{SCR}_{\text{Basic}}$ podľa [6] vypočítame ako

$$\text{SCR}_{\text{Basic}} = \sqrt{\sum_{i,j=1}^5 c_{i,j} \cdot \text{SCR}_i \cdot \text{SCR}_j}, \quad (1.4)$$

kde $c_{i,j}$ je pre $\forall i, j \in \{1, \dots, 5\}$ prvok korelačnej matice³

$$\{C\}_{i,j} = \begin{pmatrix} 1 & 0,25 & 0,25 & 0,25 & 0,25 \\ 0,25 & 1 & 0,25 & 0,25 & 0,5 \\ 0,25 & 0,25 & 1 & 0,25 & 0 \\ 0,25 & 0,25 & 0,25 & 1 & 0 \\ 0,25 & 0,5 & 0 & 0 & 1 \end{pmatrix} \quad (1.5)$$

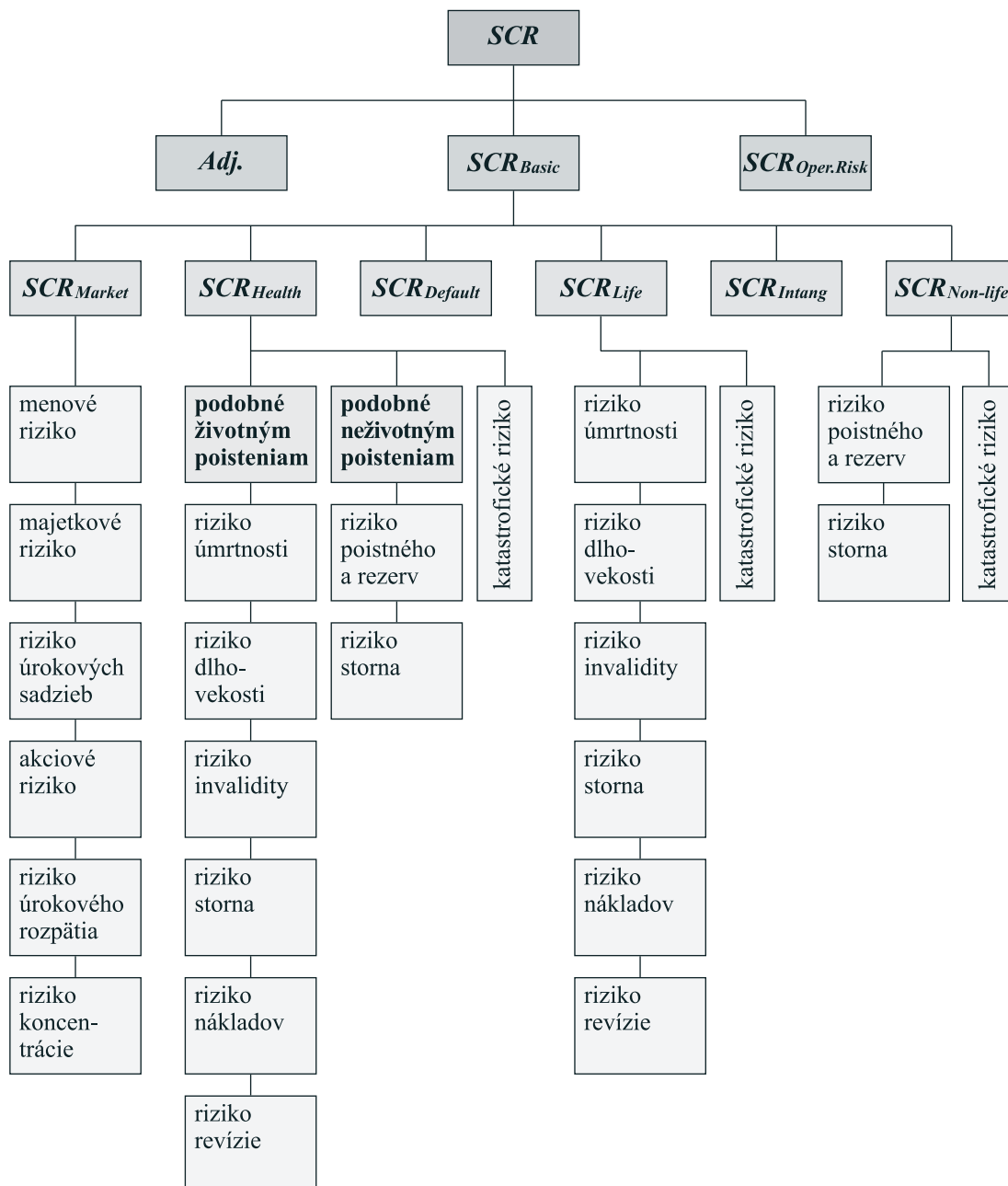
a SCR_i , resp. SCR_j , je kapitálová požiadavka vypočítaná pre modul

- trhového rizika (*Market Risk*) pre $i = 1$,
- rizika zo zlyhania protistrany (*Default Risk*) pre $i = 2$,
- krytia životných upisovacích rizík (*Life Risk*) pre $i = 3$,
- krytia zdravotných upisovacích rizík (*Health Risk*) pre $i = 4$,
- krytia neživotných upisovacích rizík (*Non-life Risk*) pre $i = 5$.

Čiastočné kapitálové požiadavky na solventnosť SCR_i sa vypočítajú analogicky ako základná kapitálová požiadavka na solventnosť $\text{SCR}_{\text{Basic}}$. Každý z použitých rizikových modulov sa ďalej delí do submodulov, ktoré využijeme pri tomto výpočte. Celková štruktúra rizikových modulov a submodulov kapitálovej požiadavky na solventnosť je zobrazená na obrázku 1.1.

Čiastočná kapitálová požiadavka z operačného rizika $\text{SCR}_{\text{OperationalRisk}}$ podľa direktívy Solventnosť II zahŕňa operačné riziká, ktoré neboli pokryté v základnej kapitálovej požiadavke na solventnosť $\text{SCR}_{\text{Basic}}$. Kapitálová požiadavka má byť kalibrovaná tak, aby bolo zaistené, že všetky kvantifikovateľné riziká, ktorým je poisťovňa vystavená, sú brané do úvahy. Má pokrývať ako súčasný, tak aj celý nový obchod poisťovne. V súvislosti s novým obchodom by mali byť pokryté iba neočakávané straty. Čiastočná kapitálová požiadavka z operačného rizika $\text{SCR}_{\text{OperationalRisk}}$ má korešpondovať s hodnotou v riziku základných vlastných fondov poisťovne na hladine spoľahlivosti 99,5% v priebehu ročného obdobia.

³Hodnoty korelačných koeficientov sú predpísané direktívou Solventnosť II. Uvedené sú v [6] na strane 124.



Obr. 1.1: Štruktúra kapitálovej požiadavky na solventnosť.

Čiastočná kapitálová požiadavka z operačného rizika $SCR_{OperationalRisk}$ by nemala prekračovať 30% zo základnej kapitálovej požiadavky na solventnosť SCR_{Basic} .

Teoreticky by štandardná formula mala odrážať rizikový profil väčšiny poisťovní. V niektorých špecifických prípadoch však tento prístup neodráža adekvátne niektoré rizikové profily. Aj z dôvodu výskytu takýchto prípadov direktíva Solventnosť II ponúka možnosť výpočtu kapitálovej požiadavky na solventnosť SCR použitím interných modelov. V prípade, že sa poisťovňa rozhodne pre takýto spôsob výpočtu, má podľa [6] na výber, či interné modely použije pri výpočte celej kapitálovej požiadavky SCR, alebo ich bude používať iba na výpočet kapitálovej požiadavky pre niektoré rizikové moduly, či submoduly používané pri výpočte SCR štandardnou formulou, výpočet kapitálovej požiadavky z operačného rizika $SCR_{OperationalRisk}$, alebo úpravy zohľadňujúcej kapacitu technických rezerv a odložených daní Adj. Pri využití parciálnych interných modelov sa zohľadňujú špecializácie poisťovne a výnimočné prípady. Cieľom použitia interných modelov je zlepšenie *risk managementu* poisťovne, zlepšenie odrazu rizikového profilu poisťovne v kapitálovej požiadavke na solventnosť a zníženie nákladov na kapitál. Interné modely, podobne ako pokročilé prístupy výpočtu kapitálových požiadaviek v dokumente Basel II, podliehajú schváleniu príslušným dozorným orgánom. Navrhované modely sú orgánom dohľadu schválené iba v prípade, že systémy poisťovne na identifikáciu, meranie, monitorovanie, riadenie a informovanie o rizikách sú adekvátne a navyše navrhované interné modely spĺňajú všetky požiadavky špecifikované direktívou Solventnosť II. Pri interných modeloch sa uplatňuje princíp "*top-down*". Ide o princíp, podľa ktorého sa najprv stanoví celková kapitálová požiadavka na solventnosť SCR a následne na to sa alokuje kapitál na čiastkové segmenty. Celková kapitálová požiadavka sa stanoví pomocou stochastického modelovania, simulácie vhodnej ekonomickej cieľovej funkcie. Najčastejšie sa za túto funkciu berie výkaz ziskov a strát.

Kapitola 2

Teória extrémnych hodnôt

2.1 Úvod

Straty z operačného rizika sa vyskytujú v rôznych frekvenciách a s rôznou významnosťou. V nasledujúcej časti sa budeme zaoberať rizikami, ku ktorým dochádza s malou frekvenciou, ale dopad strát s nimi spojených je významný. Jediný výskyt takejto straty môže spôsobiť inštitúcii veľké problémy, ktoré by mohli viesť až k jej zániku. Najväčším problémom pri riadení takýchto extrémnych rizík je, že množstvo dát je značne obmedzené vzhľadom k ich zriedkavému výskytu. Ďalším problémom je, že v histórii ešte neprišlo ku škodám z niektorých potencionálne veľmi významných rizík.

Modelovaním a analýzou javov, ktoré nastávajú zriedkavo a nadobúdajú vysoké hodnoty, sa zaoberá teória extrémnych hodnôt. Tá sa v posledných rokoch stala silným nástrojom pre odhad rizika. V rámci riadenia a modelovania operačných rizík sú pozorovanými javmi finančné straty vyplývajúce z jednotlivých rizík. Text týkajúci sa teórie extrémnych hodnôt sa opiera o výklad tejto teórie v [2].

V rámci teórie extrémnych hodnôt rozoznávame dva základné prístupy odlišujúce sa základným pohľadom na dáta. Tradičnejší prístup *blokových maxim*, popísaný v sekcii 2.2, spracováva dáta po blokoch a uvažuje iba ich maximum, čím dochádza k výraznému zníženiu počtu efektívnych pozorovaní. Naproti tomu, druhý prístup *prekročenia medze*, popísaný v sekcii 2.3, využíva všetky dáta, ktoré prekračujú stanovenú medzu. Tento prístup je uprednostňovanejší pre efektívnejšie využitie dostupných pozorovaní.

2.2 Prístup založený na blokových maximách

Pre zjednodušenie najprv zavedieme nasledujúce značenie. Ak máme reálnu náhodnú veličinu X a distribučnú funkciu $F : \mathbb{R} \rightarrow [0, 1]$, potom pod označením $X \sim F$ budeme rozumieť, že

$$\forall x \in \mathbb{R} : P(X \leq x) = F(x).$$

Ak budeme ďalej hovoriť o postupnosti $\{X_i^F\}_{i \in \mathbb{N}}$, budeme mať na mysli postupnosť navzájom nezávislých, rovnako rozdelených náhodných veličín takých, že

$$\forall i \in \mathbb{N} : X_i^F \sim F.$$

Hodnoty danej postupnosti reprezentujú finančné straty, v našom prípade, straty z operačného rizika. Ďalej označme blokové maximum

$$M_{k,n} = \max(X_k^F, X_{k+1}^F, \dots, X_{k+n-1}^F) \text{ pre } k, n \in \mathbb{N}.$$

Blokové maximá budeme používať ako reprezentantov extrémnych strát, ktoré nás zaujímajú.

2.2.1 Zovšeobecnené rozdelenie extrémnych hodnôt

Klasická teória extrémnych hodnôt sa zaoberá limitnými rozdeleniami vhodne normovaných blokových maxim $M_{1,n}$, pri $n \rightarrow \infty$. Pokiaľ sú tieto rozdelenia nedegenerované, tak sú zo skupiny zovšeobecných rozdelení extrémnych hodnôt (*Generalized Extreme Value Distribution*).

Definícia 2.2.1. *Distribučná funkcia zovšeobecného rozdelenia extrémnych hodnôt je podľa [2], str. 265 daná nasledovne:*

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

kde $1 + \xi x > 0$.

Poznámka 2.2.1. *Premennú ξ v definícii 2.2.1 nazývame parametrom tvaru (Shape Parameter).*

Poznámka 2.2.2. *Všimnime si, že pre zafixované x platí*

$$\lim_{\xi \rightarrow 0} H_\xi(x) = H_0(x).$$

Parametrizácia v 2.2.1 je teda spojitá v ξ .

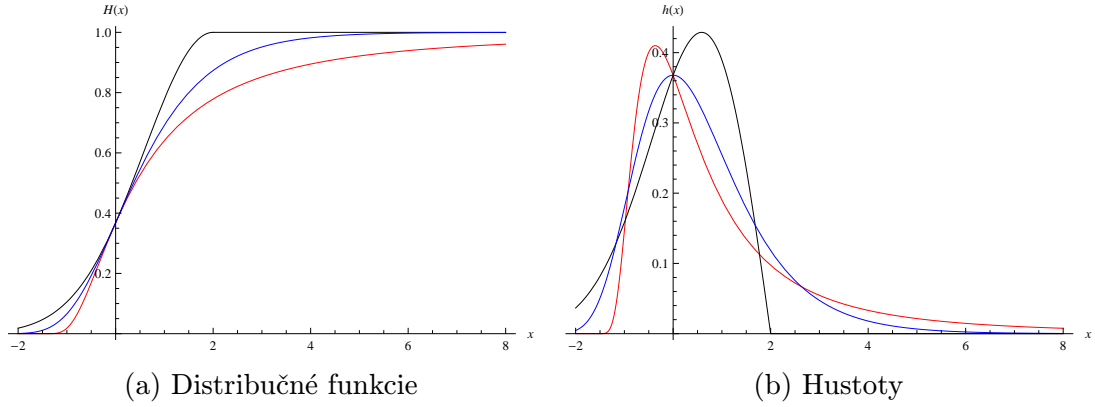
Poznámka 2.2.3. *Rozdelenie definované v 2.2.1 sa dá zovšeobecniť zavedením parametru polohy (Location Parameter) $\mu \in \mathbb{R}$ a parametru merítka (Scale Parameter) $\sigma \in \mathbb{R}^+$ nasledovne:*

$$H_{\xi,\mu,\sigma}(x) := H_\xi\left(\frac{x - \mu}{\sigma}\right). \quad (2.1)$$

Špeciálne prípady zovšeobecného rozdelenia extrémnych hodnôt majú rôzne pomenovania:

- Fréchetovo rozdelenie pre $\xi > 0$,
- Gumbelovo rozdelenie pre $\xi = 0$,
- Weibullovo rozdelenie pre $\xi < 0$.

Na obrázku 2.1a sú zobrazené grafy distribučných funkcií a na obrázku 2.1b grafy hustôt zovšeobecného rozdelenia extrémnych hodnôt pre rôzne hodnoty parametra ξ . Všimnime si, že hustota Weibullovoho rozdelenia má konečný pravý koncový bod.



Obr. 2.1: Zovšeobecnené rozdelenie extrémnych hodnôt. Červenou Fréchetovo rozdelenie ($\xi = 0, 5$), modrou Gumbelovo rozdelenie ($\xi = 0$) a čiernou Weibullovo rozdelenie ($\xi = -0, 5$).

Lemma 2.2.1. *Majme distribučnú funkciu $F : \mathbb{R} \rightarrow [0, 1]$ a postupnosť $\{X_i^F\}_{i \in \mathbb{N}}$. Potom pre blokové maximum $M_{k,n}$ tejto postupnosti platí, že $M_{k,n} \sim F^n$, kde $k, n \in \mathbb{N}$.*

Dôkaz.

$$\begin{aligned} \forall x \in \mathbb{R} : \mathbb{P}(M_{k,n} \leq x) &= \mathbb{P}(X_k^F \leq x, \dots, X_{k+n-1}^F \leq x) = \\ &= \prod_{i=k}^{k+n-1} \mathbb{P}(X_i^F \leq x) = F^n(x). \end{aligned}$$

□

Limitné správanie normovaných maxím charakterizuje veta 2.2.1 z [1], str. 121.

Veta 2.2.1. (Fisher-Tippet). *Uvažujme blokové maximum $M_{1,n}$ postupnosti $\{X_i^F\}_{i \in \mathbb{N}}$, kde F je distribučná funkcia. Pokiaľ existujú postupnosti konštánt $c_n \in \mathbb{R}^+$ a $d_n \in \mathbb{R}$ pre $\forall n \in \mathbb{N}$ také, že*

$$\lim_{n \rightarrow \infty} \mathbb{P}\left(\frac{M_{1,n} - d_n}{c_n} \leq x\right) = \lim_{n \rightarrow \infty} F^n(c_n x + d_n) = H(x),$$

pre nejakú nedegenerovanú distribučnú funkciu H , potom H je distribučná funkcia zovšeobecneného rozdelenia extrémnych hodnôt.

Náznak dôkazu vety 2.2.1 sa nachádza v [1] na strane 122.

Definícia 2.2.2. *Ak existuje nedegenerovaná distribučná funkcia H podľa vety 2.2.1, hovoríme, že distribučná funkcia F patrí do sféry príťažlivosti (Maximum Domain of Attraction) rozdelenia s distribučnou funkciou H . Značíme $F \in \text{MDA}(H)$.*

Poznámka 2.2.4. *Pri splnení podmienok vety 2.2.1 je parameter ξ distribučnej funkcie H určený jednoznačne. Parameter polohy μ a parameter merítka σ závisia na zvolených postupnostiach normujúcich konštánt c_n a d_n . Postupnosti c_n , d_n je možné zvoliť tak, že limitou je štandardná distribučná funkcia H_ξ , kde $\mu = 0$ a $\sigma = 1$.*

2.2.2 Metóda blokových maxím

Majme postupnosť pozorovaní $\{X_i^F\}_{i \in \{1, \dots, N\}}$ pre $N \in \mathbb{N}$ takých, že

$$F \in \text{MDA}(H_{\xi, \mu, \sigma}). \quad (2.2)$$

Dáta rozdelíme do $m = N/n$ blokov s veľkosťou $n \in \mathbb{N}$ a budeme sledovať maximá týchto blokov¹. Získame teda novú postupnosť $\{M_{i \cdot n + 1, n}\}_{i \in \{0, \dots, m-1\}}$. Zo vzťahu 2.2 a z vety 2.2.1 plynie, že

$$\forall i \in \{0, \dots, m-1\} : M_{i \cdot n + 1, n} \stackrel{n, N \rightarrow \infty}{\sim} H_{\xi, \mu, \sigma},$$

takže pravdepodobnostné rozdelenie náhodných veličín $M_{i \cdot n + 1, n}$ môžeme pre dostatočnú veľkosť blokov aproximovať rozdelením s distribučnou funkciou $H_{\xi, \mu, \sigma}$.

K odhadu parametrov ξ , μ a σ je možné využiť napríklad metódu maximálnej vierohodnosti. Najprv odvodíme hustotu $h_{\xi, \mu, \sigma}$ zovšeobecneného rozdelenia extrémnych hodnôt s distribučnou funkciou $H_{\xi, \mu, \sigma}$.

$$\begin{aligned} h_{\xi, \mu, \sigma}(x) &= \frac{dH_{\xi, \mu, \sigma}(x)}{dx} = \frac{d}{dx} H_{\xi} \left(\frac{x - \mu}{\sigma} \right) = \\ &= \frac{d}{dx} \left(\exp \left[- \left(1 + \frac{\xi(x - \mu)}{\sigma} \right)^{-\frac{1}{\xi}} \right] \right) = \\ &= \frac{1}{\sigma} \left(\exp \left[- \left(1 + \frac{\xi(x - \mu)}{\sigma} \right)^{-\frac{1}{\xi}} \right] \right) \left(1 + \frac{\xi(x - \mu)}{\sigma} \right)^{-\frac{1}{\xi} - 1}. \end{aligned}$$

Logaritmickej vierohodnostnej funkcia má teda tvar:

$$\begin{aligned} l(\xi, \mu, \sigma; M_{1, n}, M_{n+1, n}, \dots, M_{N-n+1, n}) &= \\ &= \sum_{i=0}^{m-1} \ln h_{\xi, \mu, \sigma}(M_{i \cdot n + 1, n}) = \\ &= -m \ln \sigma - \left(1 + \frac{1}{\xi} \right) \sum_{i=0}^{m-1} \ln \left(1 + \frac{\xi(M_{i \cdot n + 1, n} - \mu)}{\sigma} \right) \\ &\quad - \sum_{i=0}^{m-1} \left(1 + \frac{\xi(M_{i \cdot n + 1, n} - \mu)}{\sigma} \right)^{-\frac{1}{\xi}}. \end{aligned}$$

Logaritmickej vierohodnostnej funkciu maximalizujeme za podmienok

$$\sigma > 0 \text{ a } \forall i \in \{0, \dots, m-1\} : \sigma > \xi(\mu - M_{i \cdot n + 1, n}).$$

Poznámka 2.2.5. *Vzhľadom k obmedzujúcim podmienkam pre maximalizáciu funkcie l ide o iregulárny optimalizačný problém kvôli tomu, že tieto podmienky závisia na vstupných dátach. Podľa [2] sa dá ukázať, že pre $\xi > -1/2$ sú odhady získané maximalizáciou logaritmickej vierohodnostnej funkcie konzistentné.*

Poznámka 2.2.6. *Pre dostatočne veľké n sa dá podľa [2] predpokladať, že hodnoty blokových maxím sú nezávislé bez ohľadu na závislosť či nezávislosť pôvodných pozorovaní.*

¹Predpokladajme, že N je deliteľné n .

2.2.3 Úroveň návratu a doba návratu

Rozdelenie blokových maxim aproximované distribučnou funkciou zovšeobecného rozdelenia extrémnych hodnôt môžeme využiť napríklad pri určovaní úrovne návratu (*Return Level*) a doby návratu (*Return Period*). Pri určovaní úrovne návratu stanovujeme medzu, ktorá bude v priemere prekročená so zvolenou frekvenciou. Pri určovaní doby návratu, naopak určíme medzu a zisťujeme priemernú frekvenciu javu presahujúceho danú medzu.

Definícia 2.2.3. *Nech H označuje skutočnú distribučnú funkciu blokových maxim $\{M_{i \cdot n+1, n}\}_{i \in \{0, \dots, m-1\}}$. Potom úroveň návratu $r_{n, k}$ pre $k \geq 1$ v sekvencii blokov veľkosti n je*

$$r_{n, k} = q_{1-\frac{1}{k}}(H),$$

kde $q_{1-\frac{1}{k}}(H)$ značí $(1 - \frac{1}{k})$ -kvantil rozdelenia s distribučnou funkciou H .

Úroveň návratu $r_{n, k}$ môžeme chápať, ako úroveň, ktorá je prekročená v priemere jedenkrát za k blokov s rozsahom n . V prípade, že parametre distribučnej funkcie $H_{\xi, \mu, \sigma}$ zavedené v (2.1) odhadneme parametrami $\hat{\xi}, \hat{\mu}, \hat{\sigma}$, odhad úrovne návratu $\hat{r}_{n, k}$ bude mať tvar:

$$\hat{r}_{n, k} = q_{1-\frac{1}{k}}(H_{\hat{\xi}, \hat{\mu}, \hat{\sigma}}) = H_{\hat{\xi}, \hat{\mu}, \hat{\sigma}}^{-1} \left(1 - \frac{1}{k} \right) = \hat{\mu} + \frac{\hat{\sigma}}{\hat{\xi}} \left(\left(-\ln \left(1 - \frac{1}{k} \right) \right)^{-\hat{\xi}} - 1 \right).$$

Definícia 2.2.4. *Nech H označuje skutočnú distribučnú funkciu blokových maxim $\{M_{i \cdot n+1, n}\}_{i \in \{0, \dots, m-1\}}$. Potom doba návratu indexu postupnosti i takého, že $M_{i \cdot n+1, n} > u$ pre $u \in \mathbb{R}$ v sekvencii blokov veľkosti n je*

$$k_{n, u} = \frac{1}{1 - H(u)}.$$

Doba návratu $k_{n, u}$ vyjadruje, že v $k_{n, u}$ blokoch s rozsahom n očakávame výskyt jedného bloku, v ktorom dôjde k prekročeniu úrovne u . Dobu návratu môžeme odhadnúť ako

$$\hat{k}_{n, u} = \frac{1}{1 - H_{\hat{\xi}, \hat{\mu}, \hat{\sigma}}(u)}.$$

2.3 Prístup založený na prekročení medze

Ako bolo spomenuté, prístup založený na blokových maximách má nevýhodu, pre ktorú strácame veľké množstvo pozorovaní, ktoré máme k dispozícii. Počet pozorovaní je v praxi často obmedzený a z tohto dôvodu sa častejšie používa prístup založený na prekročení medze, ktorý používa všetky dáta, ktoré sú významné v tom zmysle, že prekračujú stanovenú medzu.

2.3.1 Zovšeobecnené Paretovo rozdelenie

Významnú rolu pri analýze dát, ktoré prekračujú stanovenú medzu, má zovšeobecnené Paretovo rozdelenie, ktoré sa v teórii extrémnych hodnôt používa na modelovanie ich rozdelenia.

Definícia 2.3.1. Podľa [2], str. 275, definujeme zovšeobecnené Paretovo rozdelenie (Generalized Pareto Distribution) pomocou jeho distribučnej funkcie $G_{\xi,\beta}$ s predpisom:

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

kde $\beta > 0$ a

$$x \geq 0 \text{ pre } \xi \geq 0$$

a

$$0 \leq x \leq -\frac{\beta}{\xi} \text{ pre } \xi < 0.$$

Poznámka 2.3.1. V definícii 2.3.1 parameter ξ nazývame parametrom tvaru (Shape Parameter) a β je parameter merítka (Scale Parameter).

Zovšeobecnené Paretovo rozdelenie zahŕňa niekoľko špeciálnych typov rozdelení:

- exponenciálne rozdelenie pre $\xi = 0$,
- Paretovo rozdelenie typu II pre $\xi < 0$,
- Paretovo rozdelenie (parametre α , κ viď v poznámke 2.3.2) pre $\xi > 0$, kde $\alpha = \frac{1}{\xi}$ a $\kappa = \frac{\beta}{\xi}$.

Poznámka 2.3.2. Paretovo rozdelenie má distribučnú funkciu

$$F(x) = 1 - \left(\frac{\kappa}{\kappa + x}\right)^\alpha \text{ pre } \alpha > 0, \kappa > 0 \text{ a } x \geq 0.$$

Podobne ako pri zovšeobecnenom rozdelení extrémnych hodnôt, pre zafixované x je zovšeobecnené Paretovo rozdelenie spojité v ξ , teda platí, že:

$$\lim_{\xi \rightarrow 0} G_{\xi,\beta}(x) = G_{0,\beta}(x).$$

Veta 2.3.1. Buď X reálna náhodná veličina so zovšeobecneným Paretoovým rozdelením s distribučnou funkciou $G_{\xi,\beta}$. Potom

$$\mathbb{E}(X) = \begin{cases} \frac{\beta}{1-\xi} & \text{pre } \xi < 1, \\ \infty & \text{pre } \xi \geq 1. \end{cases}$$

Dôkaz. Pre $\xi \neq 0$ a $\xi \neq 1$:

$$\begin{aligned}
\mathbb{E}(X) &= \int_0^\infty \left(y \frac{dG_{\xi,\beta}(y)}{dy} \right) dy \stackrel{\text{per partes}}{=} [yG_{\xi,\beta}(y)]_0^\infty - \int_0^\infty G_{\xi,\beta}(y) dy = \\
&= \left[y \left(1 - \left(1 + \frac{\xi y}{\beta} \right)^{-\frac{1}{\xi}} \right) \right]_0^\infty - \left[y - \frac{\beta + \xi y}{\xi - 1} \left(1 + \frac{\xi y}{\beta} \right)^{-\frac{1}{\xi}} \right]_0^\infty = \\
&= \left[\frac{(\beta + y) \left(1 + \frac{\xi y}{\beta} \right)^{-\frac{1}{\xi}}}{\xi - 1} \right]_0^\infty = \\
&= \lim_{y \rightarrow \infty} \frac{(\beta + y) \left(1 + \frac{\xi y}{\beta} \right)^{-\frac{1}{\xi}}}{\xi - 1} - \frac{\beta}{\xi - 1} = \\
&= \frac{\beta^{\frac{1}{\xi}}}{\xi - 1} \lim_{y \rightarrow \infty} \left(\frac{\beta + y}{(\beta + \xi y)^{\frac{1}{\xi}}} \right) - \frac{\beta}{\xi - 1} \stackrel{L'Hospital}{=} \\
&= \frac{\beta^{\frac{1}{\xi}}}{\xi - 1} \lim_{y \rightarrow \infty} \left(\frac{1}{(\beta + \xi y)^{\frac{1}{\xi} - 1}} \right) - \frac{\beta}{\xi - 1} = \\
&= \begin{cases} \frac{\beta}{1 - \xi} & \text{pre } \xi < 1, \\ \infty & \text{pre } \xi > 1. \end{cases}
\end{aligned}$$

Pre $\xi = 1$ dostaneme analogickým spôsobom výsledok $\mathbb{E}(X) = \infty$. Pre $\xi = 0$ ide o exponenciálne rozdelenie s parametrom $\lambda = \frac{1}{\beta}$ a vieme, že exponenciálne rozdelenie má strednú hodnotu $\frac{1}{\lambda}$, teda $\mathbb{E}(X) = \beta$. □

Poznámka 2.3.3. Buď X reálna náhodná veličina so zovšeobecneným Paretoovým rozdelením s distribučnou funkciou $G_{\xi,\beta}$. Potom

$$\mathbb{E}(X^2) = \begin{cases} \frac{\beta^2}{(1-\xi)^2} \left(\frac{2-2\xi}{1-2\xi} \right) & \text{pre } \xi < \frac{1}{2}, \\ \infty & \text{pre } \xi \geq \frac{1}{2}. \end{cases}$$

Pre $\xi = 0$ ide o exponenciálne rozdelenie s parametrom $\lambda = \frac{1}{\beta}$ a vieme, že exponenciálne rozdelenie má druhý moment $\frac{2}{\lambda^2}$, teda $\mathbb{E}(X^2) = 2\beta^2$.

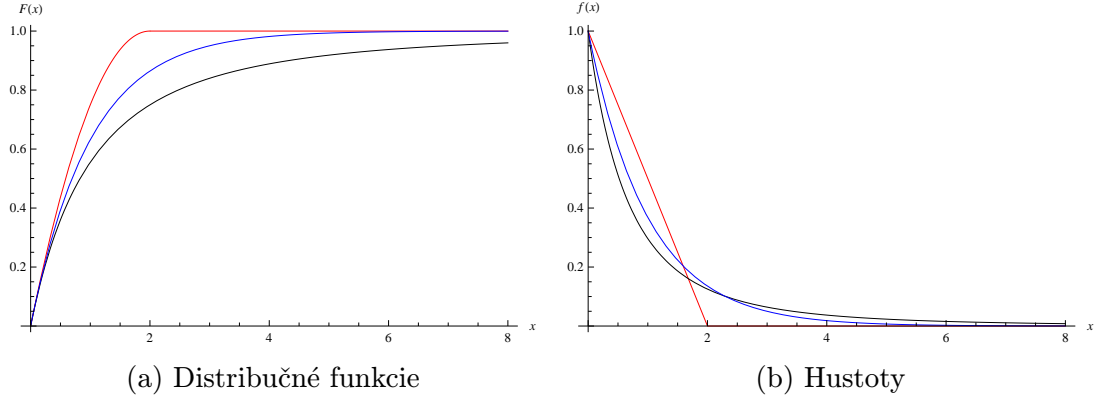
Na obrázku 2.2a sú zobrazené grafy distribučných funkcií a na 2.2b grafy hustôt zovšeobecneného Paretoého rozdelenia pre rôzne hodnoty parametru ξ . Vo všetkých prípadoch je hodnota parametra $\beta = 1$.

Teraz zavedieme distribučnú funkciu F_u opisujúcu rozdelenie strát za podmienky, že dané straty prekročili medzu u .

Definícia 2.3.2. Nech F je distribučná funkcia a X je reálna náhodná veličina, pre ktorú platí $X \sim F$. Rozdelenie excesov presahujúcich medzu $u \in \mathbb{R}$ má distribučnú funkciu

$$\begin{aligned}
F_u(x) &= P(X - u \leq x \mid X > u) = \frac{P(X - u \leq x \wedge X > u)}{P(X > u)} = \\
&= \frac{P(u < X \leq x + u)}{P(X > u)} = \frac{F(x + u) - F(u)}{1 - F(u)},
\end{aligned}$$

pre $0 \leq x < x_F - u$, kde x_F je pravý koncový bod distribučnej funkcie F .



Obr. 2.2: Zovšeobecnené Paretovo rozdelenie. Červenou Paretovo rozdelenie typu II ($\xi = -0,5$), čiernou Paretovo rozdelenie ($\xi = 0,5$) a modrou exponenciálne rozdelenie ($\xi = 0$).

Lemma 2.3.1. *Pre zovšeobecnené Paretovo rozdelenie s distribučnou funkciou $G_{\xi,\beta}$ má rozdelenie jeho excesov nad medzou u distribučnú funkciu*

$$F_u(x) = G_{\xi,\beta(u)}(x),$$

kde $\beta(u) = \beta + \xi u$ a pre x platia rovnaké podmienky ako v definícii 2.3.1.

Dôkaz.

$$\begin{aligned} F_u(x) &= \frac{G_{\xi,\beta}(x+u) - G_{\xi,\beta}(u)}{1 - G_{\xi,\beta}(u)} = \frac{\left(1 + \frac{\xi u}{\beta}\right)^{-\frac{1}{\xi}} - \left(1 + \frac{\xi(x+u)}{\beta}\right)^{-\frac{1}{\xi}}}{\left(1 + \frac{\xi u}{\beta}\right)^{-\frac{1}{\xi}}} = \\ &= 1 - \left(\frac{\beta + \xi(x+u)}{\beta + \xi u}\right)^{-\frac{1}{\xi}} = 1 - \left(1 + \frac{\xi x}{\beta + \xi u}\right)^{-\frac{1}{\xi}} = \\ &= G_{\xi,\beta+\xi u}(x) = G_{\xi,\beta(u)}(x). \end{aligned}$$

□

Z lemmatu 2.3.1 je zrejmé, že rozdelenie excesov zovšeobecného Paretoho rozdelenia zostáva zovšeobecným Paretoovým rozdelením s rovnakým parametrom tvaru ξ , ale parameter merítka β rastie lineárne s medzou u .

Definícia 2.3.3. *Nech X je náhodná veličina s konečnou strednou hodnotou. Funkcia strednej výšky excesov nad medzou u je definovaná ako*

$$e(X, u) = \mathbb{E}(X - u \mid X > u).$$

Funkcia strednej výšky excesov $e(X, u)$ vyjadruje strednú hodnotu rozdelenia s distribučnou funkciou F_u ako funkciu medze u .

Lemma 2.3.2. *Majme reálnu náhodnú veličinu X , pre ktorú platí $X \sim G_{\xi,\beta}$, kde $\xi < 1$. Potom platí, že*

$$e(X, u) = \frac{\beta(u)}{1 - \xi} = \frac{\beta + \xi u}{1 - \xi},$$

kde

$$u \geq 0 \text{ pre } 0 \leq \xi < 1$$

a

$$0 \leq u \leq -\frac{\beta}{\xi} \text{ pre } \xi < 0.$$

Dôkaz. Dôkaz je zrejmý podľa vety 2.3.1 a lemmatu 2.3.1. Obmedzujúce podmienky pre premennú u vyplývajú z definície 2.3.1. □

Charakteristickou vlastnosťou zovšeobecneného Paretoho rozdelenia teda je, že stredná výška jeho excesov je lineárnou funkciou medze u .

Veta 2.3.2 z [2], str. 277, nám ukazuje silnú vlastnosť zovšeobecneného Paretoho rozdelenia a to, že je limitným rozdelením pre všetky rozdelenia excesov s distribučnou funkciou F_u odvodených od distribučných funkcií F ležiacich v sfére príťažlivosti $\text{MDA}(H_\xi)$, kde H_ξ je zovšeobecnené rozdelenie extrémnych hodnôt, pre limitný rast medze u .

Veta 2.3.2. (Pickands-Balkema-de Hann). *Nech F je distribučná funkcia. Potom existuje kladná funkcia $\beta(u)$ taká, že platí*

$$\lim_{u \rightarrow x_F} \sup_{0 \leq x < x_F - u} |F_u(x) - G_{\xi, \beta(u)}(x)| = 0,$$

práve vtedy, keď

$$\exists \xi \in \mathbb{R} : F \in \text{MDA}(H_\xi),$$

kde x_F je pravý koncový bod distribučnej funkcie F .

Náznač dôkazu vety 2.3.2 je možné nájsť v [1] na strane 159. Významnosť vety 2.3.2 je vidieť predovšetkým v spojení s poznámkou 2.3.4.

Poznámka 2.3.4. *Podľa [2] väčšina bežných spojitých pravdepodobnostných rozdelení leží v sfére príťažlivosti H_ξ .*

2.3.2 Modelovanie strát prekračujúcich medzu

V nasledujúcej časti budeme využívať predpoklad 2.3.1, ktorý je idealizáciou, ale vzhľadom k vete 2.3.2 a poznámke 2.3.4 ide o aproximáciu, ktorej chyba ide limitne k nule.

Predpoklad 2.3.1. *Nech F je distribučná funkcia rozdelenia s pravým koncovým bodom x_F , potom budeme predpokladať, že pre dostatočne veľkú medzu u platí, že*

$$F_u(x) = G_{\xi, \beta}(x)$$

pre $0 \leq x \leq x_F - u$, $\beta > 0$ a $\xi \in \mathbb{R}$.

Majme pozorovania o stratách x_1, \dots, x_N s rozdelením daným distribučnou funkciou F , pre ktorú platí predpoklad 2.3.1 a $N \in \mathbb{N}$. Náhodná veličina

$$N_u = \sum_{i=1}^N \mathbf{I}[x_i > u]$$

určuje počet hodnôt prekračujúcich stanovenú medzu $u \geq 0$. Označme $\tilde{x}_1, \dots, \tilde{x}_{N_u}$ pozorovania presahujúce medzu u . Nakoniec majme novú postupnosť y_1, \dots, y_{N_u} takú, že

$$\forall i \in \{1, \dots, N_u\} : y_i = \tilde{x}_i - u.$$

Teraz môžeme odhadnúť parametre distribučnej funkcie $G_{\xi, \beta}$ metódou maximálnej vierohodnosti z pozorovaní y_1, \dots, y_{N_u} . Logaritmickej vierohodnostná funkcia má tvar

$$\begin{aligned} l(\xi, \beta; y_1, \dots, y_{N_u}) &= \sum_{i=1}^{N_u} \ln \frac{dG_{\xi, \beta}}{dx}(y_i) = \\ &= \sum_{i=1}^{N_u} \ln \left(\frac{1}{\beta} \left(1 + \xi \frac{y_i}{\beta} \right)^{-\frac{1+\xi}{\xi}} \right) = \\ &= -N_u \ln \beta - \left(1 + \frac{1}{\xi} \right) \sum_{i=1}^{N_u} \ln \left(1 + \xi \frac{y_i}{\beta} \right). \end{aligned}$$

Parametre ξ a β získame maximalizáciou tejto funkcie s obmedzujúcimi podmienkami

$$\beta > 0 \text{ a } \forall i \in \{1, \dots, N_u\} : 1 + \xi \frac{y_i}{\beta} > 0.$$

Ďalšou možnou metódou pre odhad parametrov distribučnej funkcie $G_{\xi, \beta}$ je momentová metóda. Z vety 2.3.1 a poznámky 2.3.3 poznáme prvý a druhý teoretický moment m_1 a m_2 zovšeobecneného Paretoho rozdelenia pre $\beta > 0$ a pre $\xi < \frac{1}{2}$

$$\begin{aligned} m_1 &= \frac{\beta}{1 - \xi}, \\ m_2 &= \frac{\beta^2}{(1 - \xi)^2} \left(\frac{2 - 2\xi}{1 - 2\xi} \right). \end{aligned} \tag{2.3}$$

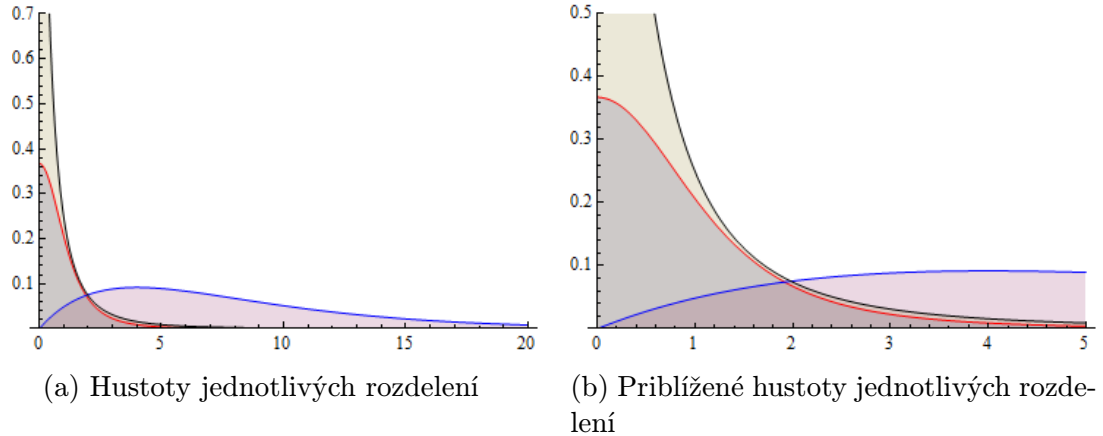
Na základe pozorovaní y_1, \dots, y_{N_u} odhadneme momenty \hat{m}_1 a \hat{m}_2 . Po porovnaní s teoretickými momentmi môžeme z (2.3) vyjadriť odhady parametrov $\hat{\xi}$ a $\hat{\beta}$:

$$\begin{aligned} \hat{\beta} &= \hat{m}_1 (1 - \hat{\xi}) = \frac{\hat{m}_1 \hat{m}_2}{2(\hat{m}_2 - \hat{m}_1^2)}, \\ \hat{\xi} &= \frac{\hat{m}_2 - 2\hat{m}_1^2}{2(\hat{m}_2 - \hat{m}_1^2)}. \end{aligned} \tag{2.4}$$

2.3.3 Testovanie úspešnosti modelovania excesov

V tejto časti ukážeme, ako dobre sa dajú modelovať excesy nad danou hranicou postupom popísaným v predchádzajúcej sekcii 2.3.2 pre rôzne rozdelenia s distribučnou funkciou $F \in \text{MDA}(H_\xi)$, kde H_ξ je distribučná funkcia zovšeobecneného rozdelenia extrémnych hodnôt a $\xi \in \mathbb{R}$. Tieto rozdelenia budú:

- Studentovo t-rozdelenie s 3 stupňami voľnosti,
- Gamma rozdelenie s parametrami $\alpha = 2$ a $\beta = 4$,
- Paretovo rozdelenie s parametrami $\alpha = 1$ a $\kappa = 2$.



Obr. 2.3: Zovšeobecnené Paretovo rozdelenie. Červenou Studentovo t-rozdelenie s 3 stupňami voľnosti, modrou Gamma rozdelenie s parametrami $\alpha = 2$ a $\beta = 4$ a čiernou Paretovo rozdelenie s parametrami $\alpha = 1$ a $\kappa = 2$.

Poznámka 2.3.5. Na obrázku 2.3 sú zobrazené grafy hustôt podkladových rozdelení, ktoré budeme používať na testovanie úspešnosti modelovania.

Počas testovania budeme využívať predpoklad 2.3.1. Pre objektívne porovnanie kvality modelu sa vyhneme subjektívnemu určovaniu medze u , keďže voľba predovšetkým silno ovplyvňuje množstvo pozorovaní, z ktorých odhadujeme parametre nášho modelu. Preto namiesto generovania dopredu stanoveného počtu dát, stanovenia odhadu vhodnej medze u a následnou selekciou tých dát, ktoré medzu u prekračujú, zafixujeme počet presahov zvolenej medze u a nagerujeme práve N_u presahov z podkladového rozdelenia. Takýto prístup k modelovaniu využili aj autori v článku [3].

Náš model sme testovali na počtoch excesov z intervalu [15, 621], pričom pre nižšie hodnoty sme volili viac vzoriek. Presahovú medzu volíme pre každé rozdelenie ako 90% kvantil (predpokladáme, že táto medza je dostatečne vysoká pre platnosť predpokladu 2.3.1).

Na základe takto nagerovaných dát môžeme začať s odhadom distribučnej funkcie excesov $G_{\xi,\beta}$. Parametre ξ a β odhadujeme pomocou metódy maximálnej vierohodnosti a pomocou momentovej metódy popísanej v sekcii 2.3.2.

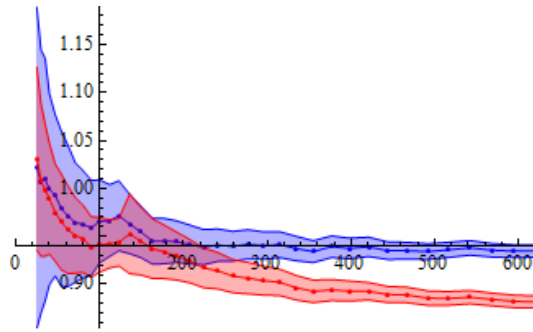
Vzhľadom k náhodnosti vstupov merania budeme modelovanie opakovať stokrát pre všetky kombinácie podkladového rozdelenia F , počtu dát prekračujúcich danú medzu u a metódy odhadu parametrov ξ a β .

Na obrázku 2.4 sú grafy, ktoré zobrazujú priemerné hodnoty nami odhadnutých parametrov $\hat{\xi}$ a $\hat{\beta}$ distribučnej funkcie excesov $G_{\xi,\beta}$ zo 100 simulácií pre rôzne počty nagerovaných excesov z jednotlivých podkladových rozdelení a pre obe používané metódy odhadu parametrov. Odhady hodnôt $\hat{\xi}$ a $\hat{\beta}$ sú zobrazené spolu s konfidenčným intervalom, ktorý je zostrojený na základe centrálnej limitnej vety a z vlastností normálneho rozdelenia vyplýva, že skutočná hodnota parametru bude s pravdepodobnosťou 99,7% ležať v tomto intervale.

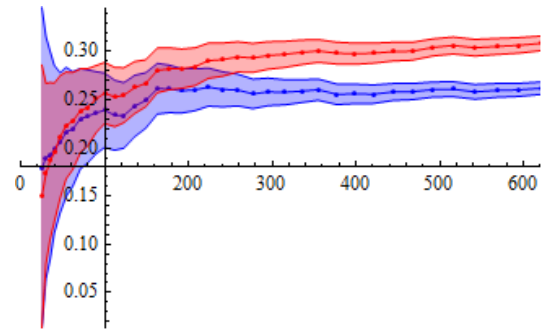
Z grafov na obrázku 2.4 je vidieť, že rozptyl hodnôt parametrov rozdelenia excesov sa s rastúcim počtom excesov znižuje. Ďalej je z nášho testovania modelovania excesov zreteľné, že s narastajúcim počtom excesov sa priemerná hodnota odhadov parametrov ustáli. Je vidieť, že hodnoty parametrov sa stabilizujú pre

zhruba 200 presahov nad medzou u .

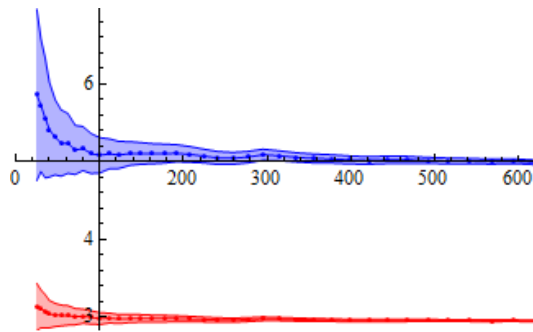
Na základe [2] vieme, že Studentovo t-rozdelenie s $\nu \geq 1$ stupňami voľnosti a distribučnou funkciou F patrí do sféry príťažlivosti Fréchetovho rozdelenia. Pre rozdelenia zo sféry príťažlivosti Fréchetovho rozdelenia vieme, že normalizované maximá pozorovaní konvergujú k Fréchetovmu rozdeleniu s parametrom tvaru ξ . Excesy takéhoto rozdelenia nad dostatočne vysokou medzou u konvergujú k zovšeobecnenému Paretovmu rozdeleniu s rovnakým parametrom ξ . V prípade Studentovho t-rozdelenia s $\nu \geq 1$ stupňami voľnosti je ξ podľa [2] rovné $1/\nu$.



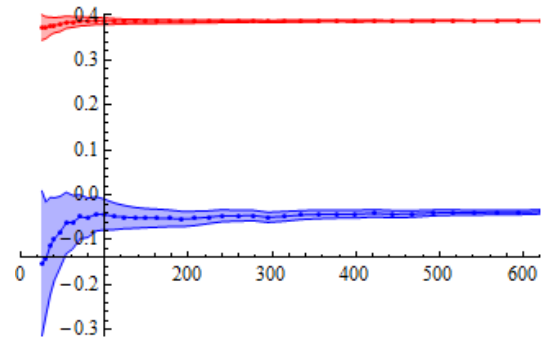
(a) Studentovo t-rozdelenie, parameter β



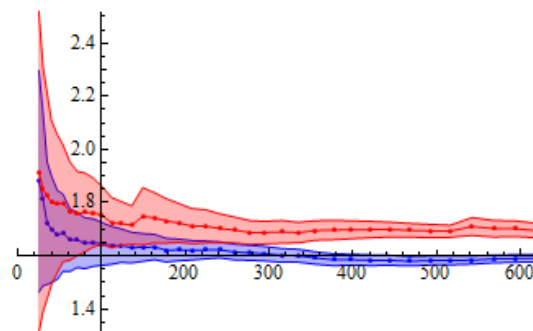
(b) Studentovo t-rozdelenie, parameter ξ



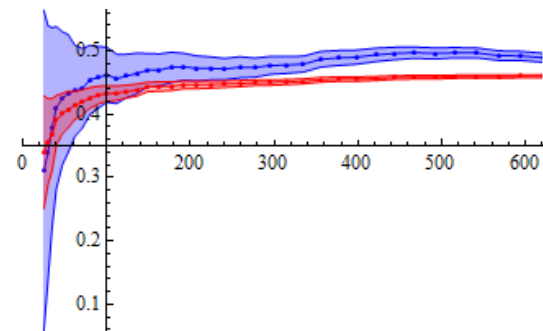
(c) Gamma rozdelenie, parameter β



(d) Gamma rozdelenie, parameter ξ



(e) Paretovo rozdelenie, parameter β



(f) Paretovo rozdelenie, parameter ξ

Obr. 2.4: Odhady parametrov ξ a β zovšeobecneného Paretovmu rozdelenia na základe dát zo Studentovho t-rozdelenia s 3 stupňami voľnosti, Gamma rozdelenia s parametrami $\alpha = 2$ a $\beta = 4$ a z Paretovmu rozdelenia s parametrami $\alpha = 1$ a $\kappa = 2$ metódou maximálnej vierohodnosti (modrou) a momentovou metódou (červenou). Na osi x je počet dát presahujúcich medzu u a na osi y je priemerná hodnota príslušného parametra.

Teda $F \in \text{MDA}(H_{1/\nu})$.

V našom prípade sme za podkladové rozdelenie brali Studentovo t-rozdelenie s 3 stupňami voľnosti, teda parameter tvaru zovšeobecneného rozdelenia ξ by sa mal rovnať $1/3$, čo je porovnateľné s naším odhadom tohto parametra ako momentovou metódou, tak aj metódou maximálnej vierohodnosti.

Na rozdiel od článku [3], kde bola úspešnosť modelu kvantifikovaná pomocou porovnávania skutočných a odhadnutých vysokých kvantilov rozdelenia, my budeme úspešnosť modelu kvantifikovať odchýlkou distribučnej funkcie odhadnutého zovšeobecneného Paretoho rozdelenia od distribučnej funkcie skutočného rozdelenia nagenovaných excesov a odchýlkou hustoty pravdepodobnosti zovšeobecneného Paretoho rozdelenia od hustoty skutočného rozdelenia nagenovaných excesov. Ako metriku rozdielnosti týchto funkcií sme zvolili supremálnu odchýlku definovanú v 2.3.4.

Definícia 2.3.4. *Majme dve spojité funkcie $f : \mathbb{R}_0^+ \rightarrow [0, 1]$ a $g : \mathbb{R}_0^+ \rightarrow [0, 1]$. Potom ich supremálnu odchýlku označíme ako*

$$E_{sup}(f, g) = \sup_{x \in \mathbb{R}_0^+} |f(x) - g(x)|.$$

Poznámka 2.3.6. *V nasledujúcich príkladoch budeme z dôvodu zložitosti výpočtu, odchýlku definovanú v 2.3.4, iba aproximovať. Určme veľkosť aproximačného kroku $s \in \mathbb{R}^+$ (v našom prípade používame $s = 0,01$) a majme množinu deliacich bodov*

$$D_s^k = \{i \cdot s; i \in \mathbb{N}_0 \wedge i \cdot s \leq k\},$$

kde $k \in \mathbb{R}_0^+$ je pravým koncovým bodom, do ktorého až odhad počítame. Maximálnu odchýlku budeme počítat ako

$$E_{max}(f, g, k, s) = \max_{x \in D_s^k} |f(x) - g(x)|.$$

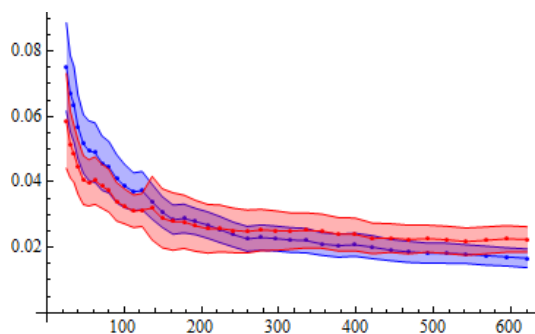
Ako pravý krajný bod k použijeme dostatočne vysoký kvantil rozdelenia excesov presahujúcich medzu u s distribučnou funkciou F_u vypočítanú na základe definície 2.3.2 so znalosťou podkladového rozdelenia F . V našom prípade sme zvolili $k = q_{0,99}(F_u)$.

V rámci nášho testovania budeme skúmať strednú hodnotu a rozptyl supremálnej odchýlky naprieč sto nezávislými modelovaniami. Zameriame sa jednak na odchýlku distribučnej funkcie rozdelenia excesov F_u a distribučnej funkcie nášho modelu, teda zovšeobecneného Paretoho rozdelenia $G_{\hat{\xi}, \hat{\beta}}$ a tiež na odchýlku hustoty pravdepodobnosti rozdelenia excesov s distribučnou funkciou F_u a hustoty zovšeobecneného Paretoho rozdelenia s distribučnou funkciou $G_{\hat{\xi}, \hat{\beta}}$.

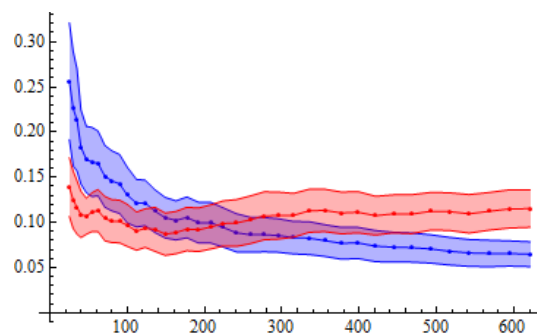
Stredné hodnoty supremálnych odchýlok spolu s konfidenčným intervalom, zostrojeným rovnakým spôsobom ako v prípade grafov na obrázku 2.4, sú pre rôzne hodnoty počtu excesov, metódy odhadu parametrov ξ a β a pre jednotlivé podkladové rozdelenia zobrazené na obrázku 2.5.

Na základe nášho modelovania nie je možné povedať, ktorá z metód odhadu parametrov ξ a β je lepšia pre všetky počty excesov a podkladové rozdelenia. Môžeme však povedať, že pre počet excesov $N_u \geq 250$ sa javí metóda maximálnej vierohodnosti vzhľadom k $E_{max}(f, g, k, s)$ pre všetky používané podkladové

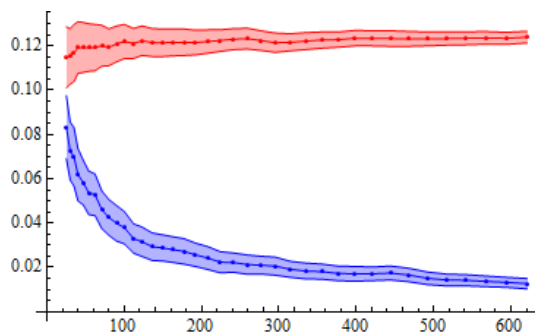
rozdelenia lepšia. V prípade Gamma rozdelenia je tento rozdiel zrejmy aj pre malý počet excesov. Ďalej môžeme povedať, že rozptyl $E_{max}(f, g, k, s)$ sa opäť so zväčšujúcim sa počtom excesov znižuje. Zúžovanie konfidenčného intervalu je zreteľnejšie v prípade odhadovania parametrov pomocou metódy maximálnej vierohodnosti.



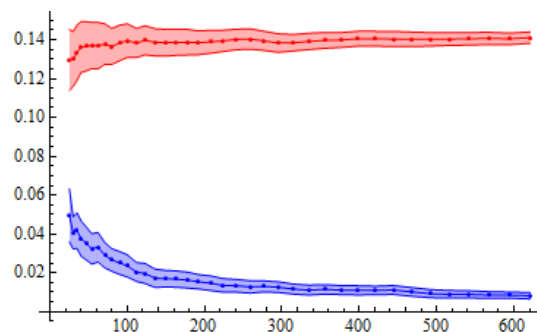
(a) Studentovo t-rozdelenie, odchýlka distribučných funkcií



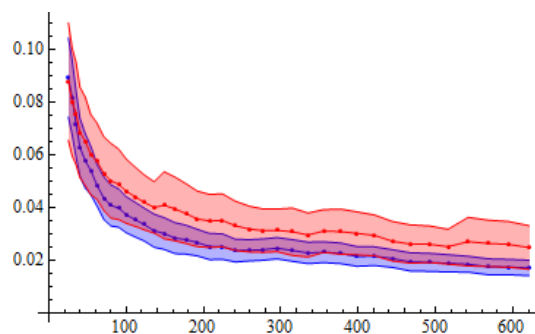
(b) Studentovo t-rozdelenie, odchýlka hustôt



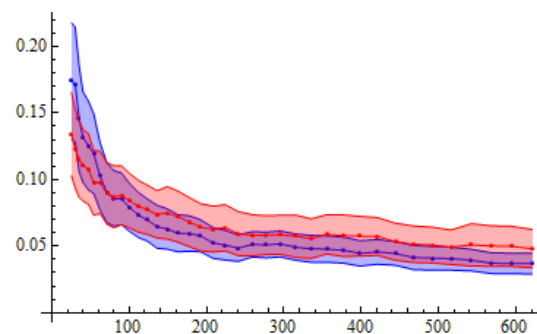
(c) Gamma rozdelenie, odchýlka distribučných funkcií



(d) Gamma rozdelenie, odchýlka hustôt



(e) Pareto rozdelenie, odchýlka distribučných funkcií



(f) Pareto rozdelenie, odchýlka hustôt

Obr. 2.5: Hodnoty stredných hodnôt maximálnych odchýlok $E_{max}(f, g, k, s)$ pre metódu maximálnej vierohodnosti (na grafoch modrou) a momentovú metódu (červenou) odhadu parametrov ξ a β zovšeobecneného Paretovho rozdelenia. Podkladovým rozdelením je Studentovo t-rozdelenie s 3 stupňami voľnosti, Gamma rozdelenie s parametrami $\alpha = 2$ a $\beta = 4$ a Pareto rozdelenie s parametrami $\alpha = 1$ a $\kappa = 2$. Na osi x je počet presahov nad medzou u a na osi y priemerná hodnota maximálne odchýlky.

Teraz sa budeme zaoberať tým, koľko excesov potrebujeme na to, aby zovšeobecnené Paretovo rozdelenie dostatočne dobre modelovalo rozdelenie excesov nad medzou u . K tomuto budeme používať postup odhadu modelu popísaný v sekcii 2.3.2 a Kolmogorovu-Smirnovu vetu 2.3.3 z [8], str. 87, ktorá popisuje rýchlosť konvergencie empirickej distribučnej funkcie \hat{F}_N ku skutočnej distribučnej funkcii F .

Veta 2.3.3. (Kolmogorov-Smirnov). *Nech F je spojitá distribučná funkcia, $X \sim F$ je náhodná veličina a \hat{F}_N je empirická distribučná funkcia*

$$\hat{F}_N(t) = \frac{1}{N} \sum_{i=1}^N \mathbf{I}[x_i \leq t],$$

kde x_1, \dots, x_N sú realizácie náhodnej veličiny X . Potom pre $\epsilon > 0$ a dostatočne veľké N platí, že

$$\mathbb{P}\{\sup_t |F(t) - \hat{F}_N(t)| > \epsilon\} = 2 \sum_{k=1}^{\infty} (-1)^{k-1} e^{-2\epsilon^2 k^2 N}.$$

Lemma 2.3.3. *Nech F je spojitá distribučná funkcia, $X \sim F$ je náhodná veličina a \hat{F}_N je jej empirická distribučná funkcia. Potom pre $\epsilon > 0$ a dostatočne veľké N platí, že*

$$\epsilon < \sqrt{\frac{\ln 2 - \ln \left[\mathbb{P}\{\sup_t |F(t) - \hat{F}_N(t)| > \epsilon\} \right]}{2N}}.$$

Dôkaz. Z vety 2.3.3 plynie, že

$$\mathbb{P}\{\sup_t |F(t) - \hat{F}_N(t)| > \epsilon\} < 2e^{-2\epsilon^2 N}. \quad (2.5)$$

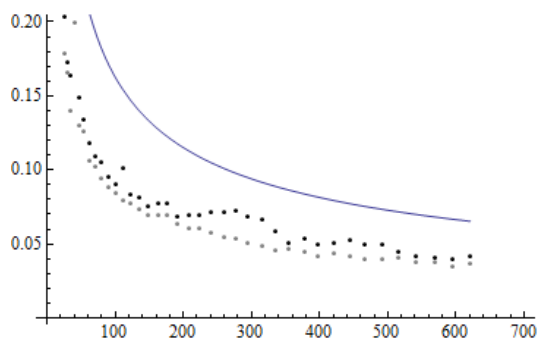
Odtiaľ dokazovaný vzťah plynie triviálne. □

Kolmogorovu-Smirnovu vetu využijeme tak, že budeme predpokladať, že rozdelenie excesov má práve zovšeobecnené Paretovo rozdelenie a budeme porovnávať distribučnú funkciu tohto rozdelenia (parametre odhadneme z excesov) a empirickú distribučnú funkciu danú priamo excesmi. Tým ukážeme, ako dobre zovšeobecnené Paretovo rozdelenie dokáže modelovať excesy našich dát, poprípade, aké množstvo excesov je potrebné pre dobrý odhad rozdelenia excesov.

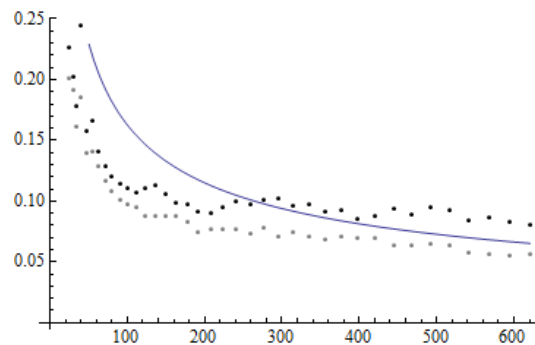
Zvolíme si zafixovanú pravdepodobnosť

$$\mathbb{P}\{\sup_t |F(t) - \hat{F}_{N_u}(t)| > \epsilon\} = 0,01,$$

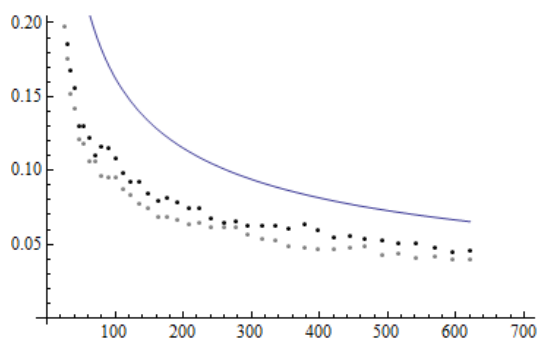
s ktorou odhad supremálnej odchýlky budeme počítať. Následne, na základe lemy 2.3.3 pre rôzne počty excesov N_u , vypočítame a vykreslíme grafy horných teoretických hraníc týchto odchýlok ϵ odhadnutej distribučnej funkcie rozdelenia $G_{\hat{\xi}, \hat{\beta}}$ a empirickej distribučnej funkcie odhadnutej priamo z hodnôt excesov. Takto skonštruovaný odhad by mal byť relatívne konzervatívny a to z toho dôvodu, že je nezávislý na pravdepodobnostnom rozdelení.



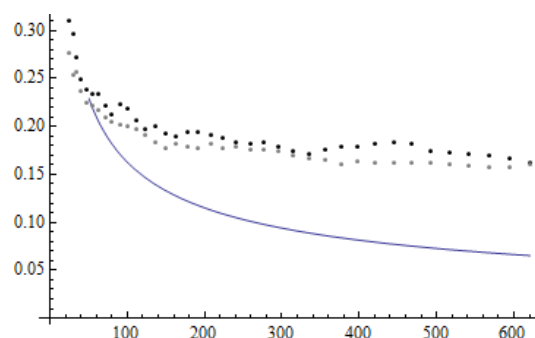
(a) Studentovo t-rozdelenie, metóda maximálnej vierohodnosti



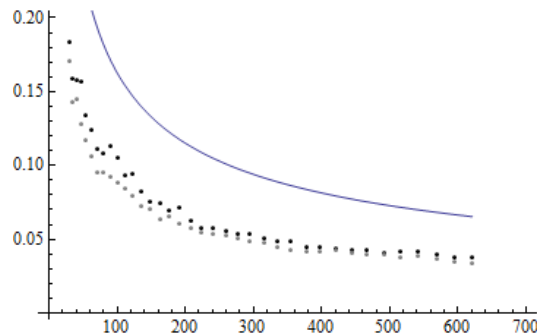
(b) Studentovo t-rozdelenie, momentová metóda



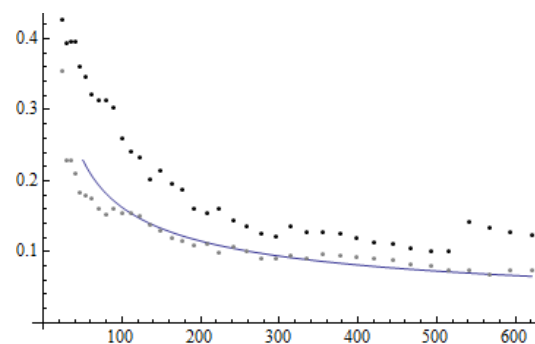
(c) Gamma rozdelenie, metóda maximálnej vierohodnosti



(d) Gamma rozdelenie, momentová metóda



(e) Paretovo rozdelenie, metóda maximálnej vierohodnosti



(f) Paretovo rozdelenie, momentová metóda

Obr. 2.6: Horná teoretická hranica supremálnej odchýlky ϵ a odhad skutočnej hranice ϵ pre konkrétne podkladové rozdelenie, tj. pre Studentovo t-rozdelenie s 3 stupňami voľnosti, Gamma rozdelenie s parametrami $\alpha = 2$ a $\beta = 4$ a pre Paretovo rozdelenie s parametrami $\alpha = 1$ a $\kappa = 2$. Pre metódu maximálnej vierohodnosti a momentovú metódu odhadu parametrov nášho modelu. Na osi x je počet dát prekračujúcich medzu u a na osi y je príslušná hodnota hornej teoretickej hranice supremálnej odchýlky a odhad skutočnej hranice supremálnej odchýlky.

Tento teoretický odhad hranice supremálnej odchýlky budeme porovnávať s odhadmi supremálnej odchýlky nameranými priamo z nasimulovaných dát. Tieto odhady teda budú 99% kvantilom postupnosti supremálnych odchýlok name-

raných v priebehu 100 nezávislých simulácií. To plyne z toho, že teoretický odhad hodnoty ϵ sme spočítali s pravdepodobnosťou

$$P\{\sup_t |F(t) - \hat{F}_{N_u}(t)| > \epsilon\} = 0,01.$$

Opäť budeme používať dáta zo Studentovho t-rozdelenia, Gamma rozdelenia a z Paretoovho rozdelenia.

Napočítané hodnoty teoretického odhadu hranice supremálnej odchýlky (na grafe plnou čiarou) spolu s príslušnými nameranými odhadmi (čierne bodky) sú zobrazené pre jednotlivé podkladové rozdelenia a metódy odhadu parametrov zovšeobecneného Paretoovho rozdelenia na obrázku 2.6. Na grafoch sú pre porovnanie zobrazené namerané odhady (sivé bodky) aj pre pravdepodobnosť

$$P\{\sup_t |F(t) - \hat{F}_{N_u}(t)| > \epsilon\} = 0,05.$$

Vidíme, že odhad hranice z nasimulovaných dát je v prípade momentovej metódy odhadu systematicky nad teoretickou hodnotou hranice supremálnej odchýlky na rozdiel od metódy maximálnej vierohodnosti. Preto by sa dalo povedať, že v našom prípade je odhad pomocou metódy maximálnej vierohodnosti vhodnejším. Z obrázku 2.6 sa ďalej dajú vyčítať hodnoty počtu presahov, ktoré potrebujeme na to, aby sme mohli povedať, že s pravdepodobnosťou 99% je supremálna odchýlka nášho modelu a skutočnej empirickej funkcie pre určité podkladové rozdelenie nanajvýš 0, 1. Napríklad pre Studentovo t-rozdelenie s 3 stupňami voľnosti potrebujeme v tomto prípade približne 100 excesov, ak sme pre odhad parametrov nášho modelu použili metódu maximálnej vierohodnosti. V prípade momentovej metódy by sme potrebovali takmer 200 presahov nad medzu u .

2.3.4 Správanie sa rozdelenia excesov pri zvýšení medze

Teraz sa pozrieme na správanie sa rozdelenia excesov, pokiaľ už poznáme rozdelenie prekročenia pre medzu u a zaujíma nás, ako vyzerá rozdelenie pre medzu v , kde $v \geq u$.

Veta 2.3.4. *Za predpokladu 2.3.1 pre všetky $v \geq u$, kde u je dostatočne veľké, platí, že*

$$F_v(x) = G_{\xi, \beta + \xi(v-u)}(x).$$

Dôkaz. *Z lemy 2.3.1 plynie, že*

$$F_v(x) = G_{\xi, \beta(v-u)}(x) = G_{\xi, \beta + \xi(v-u)}(x).$$

□

Vidíme, že rozdelenie excesov nad vyššou medzou má opäť zovšeobecnené Paretoovo rozdelenie s rovnakým parametrom ξ ako pri nižšej medze a parameter merítka vzrástol lineárne v závislosti na rozdielu $v - u$.

Lemma 2.3.4. *Majme za predpokladu 2.3.1 reálnu náhodnú veličinu X , pre ktorú platí, že $X \sim F$. Navyac parameter $\xi < 1$. Potom pre všetky $v \geq u$, kde u je dostatočne veľké, platí, že*

$$e(X, v) = \frac{\beta + \xi(v - u)}{1 - \xi} = \frac{\xi v}{1 - \xi} + \frac{\beta - \xi u}{1 - \xi},$$

kde

$$u \leq v < \infty \text{ pre } 0 \leq \xi < 1$$

a

$$u \leq v \leq u - \frac{\beta}{\xi} \text{ pre } \xi < 0.$$

Dôkaz. Priamo plynie z vety 2.3.4. Podmienky plynú z definície 2.3.1 a vety 2.3.1. \square

Práve linearita strednej výšky excesov sa používa pri určovaní toho, či je model zovšeobecneného Paretoovho rozdelenia pre rozdelenie excesov vhodný. Majme náhodnú veličinu X a jej realizácie, pozitívne hodnoty strát, x_1, \dots, x_n . Zostrojíme empirický odhad funkcie strednej výšky excesov $e(X, v)$ pre výšku medze v ako

$$\hat{e}_n(X, v) = \frac{\sum_{i=1}^n (x_i - v) \mathbb{I}[x_i > v]}{\sum_{i=1}^n \mathbb{I}[x_i > v]}.$$

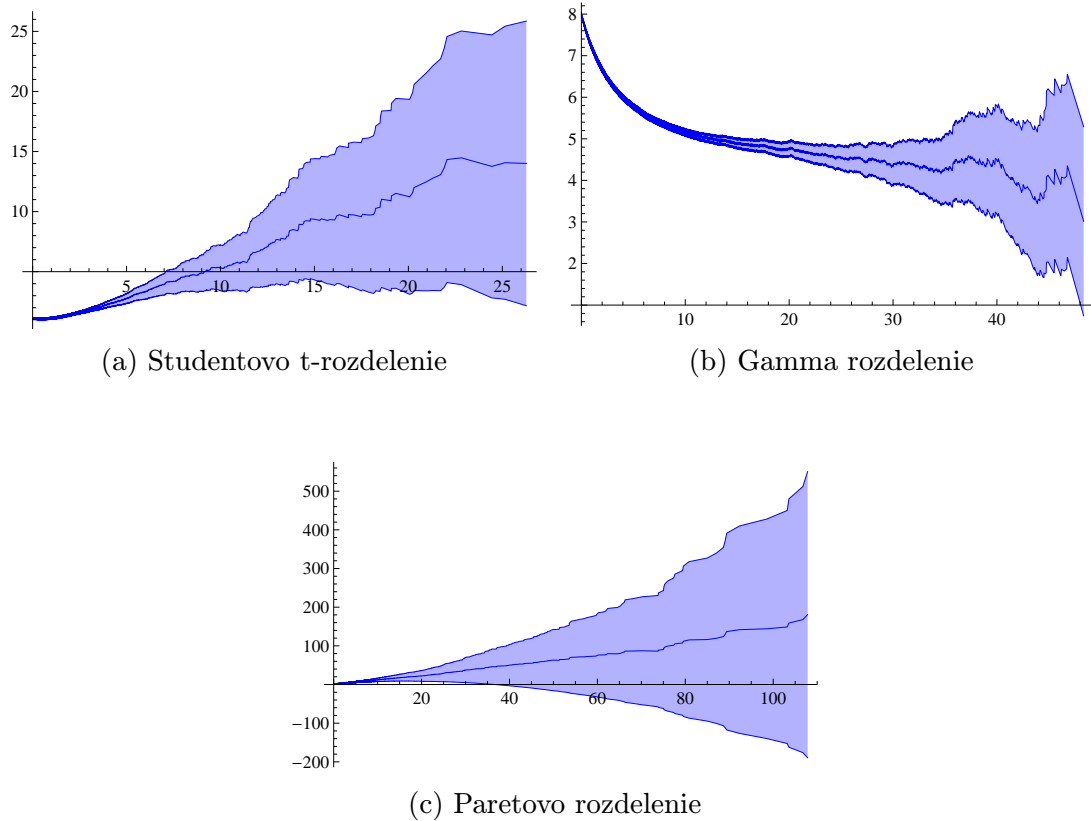
Empirický odhad graficky znázorníme vynesением bodu

$$(x_i, \hat{e}_n(X, x_i)), \text{ kde } i \in \{1, \dots, n\},$$

do grafu známeho pod anglickým názvom *Mean Excess Plot*. V prípade, že model zovšeobecneného Paretoovho rozdelenia je pre dáta vhodný, bude graf vykazovať lineárny trend pre vyššie hodnoty medze v . Práve miesto, kde začína byť trend grafu lineárny, môže slúžiť k stanoveniu vhodnej výšky medze. Začiatok lineárneho trendu volíme z toho dôvodu, že čím vyššiu medzu zvolíme pre odhad parametra modelu, tým menej pozorovaní nám pre tento odhad zostane.

V nasledujúcom príklade ukážeme využitie vyššie popísanej metódy odhadu medze u .

Príklad. Predpokladajme, že máme vygenerovaných $N = 150000$ pozorovaní x_1, \dots, x_N zo Studentovho t-rozdelenia s 3 stupňami voľnosti, Gamma rozdelenia s parametrami $\alpha = 2$ a $\beta = 4$ a Paretoovho rozdelenia s parametrami $\alpha = 1$ a $\kappa = 2$. Pre každé z týchto rozdelení postupne zostrojíme graf empirického odhadu stredných výšok excesov postupom, ktorý je popísaným v tejto sekcii. Takto zostrojené grafy spolu s konfidenčným intervalom sú zobrazené na obrázku 2.7. Z týchto grafov môžeme vyčítať vhodnú výšku medze u . Ako pomôcku pri hľadaní tejto vhodnej výšky môžeme použiť spomínaný konfidenčný interval zostrojený použitím centrálnej limitnej vety, ako tomu bolo v sekcii 2.3.3. Tentokrát však nepočítame stredné hodnoty a rozptyl z viacerých modelovaní, ale vychádzame z výpočtu strednej výšky excesov. My sme ako vhodnú výšku medze u pre Studentovo t-rozdelenie zvolili hodnotu 15, keďže vývoj grafu od tejto hodnoty sa už dá považovať za lineárny. V prípade Gamma rozdelenia je vhodné za medzu u na základe našej simulácie zobrať hodnotu 40, pri ktorej vidíme začiatok lineárneho trendu. Pre Pareto rozdelenie, na základe grafu na obrázku 2.7, by sme za medzu u určili hodnotu 80. Pri voľbe medze u je možné koncovú časť grafu zanedbať, keďže v tejto časti sú už hodnoty napočítavané z veľmi malého počtu dát.



Obr. 2.7: Graf empirického odhadu strednej výšky excesov pre rôzne hodnoty medze u spolu s konfidenčným intervalom pre Studentovo t-rozdelenie s 3 stupňami voľnosti, Gamma rozdelenie s parametrami $\alpha = 2$ a $\beta = 4$ a pre Paretovo rozdelenie s parametrami $\alpha = 1$ a $\kappa = 2$. Na osi x je hodnota medze u a na osi y je hodnota empirického odhadu strednej výšky excesov.

2.3.5 Hillova metóda

V závere predchádzajúcej sekcie bol popísaný jeden z možných spôsobov určenia vhodnej medze u použitej na modelovanie excesov, ktoré túto medzu prekračujú. Teraz popíšeme ďalší z možných spôsobov prístupu k určeniu výšky medze u . Tento postup je možné použiť iba v prípade, keď pre distribučnú funkciu F podkladových dát platí

$$F \in \text{MDA}(H_\xi) \text{ pre } \xi > 0,$$

kde H_ξ je distribučná funkcia zovšeobecneného rozdelenia extrémnych hodnôt z definície 2.2.1.

Definícia 2.3.5. *Nech L je kladná lebesgueovskú merateľná funkcia definovaná na intervale $(0, \infty)$. Potom L nazývame pomaly sa meniacou funkciou v nekonečne, ak platí, že*

$$\lim_{x \rightarrow \infty} \frac{L(tx)}{L(x)} = 1$$

pre všetky $t > 0$.

Veta 2.3.5. *Majme distribučnú funkciu F a pomaly sa meniacu funkciu v nekonečne L a $\xi > 0$. Potom*

$$F \in \text{MDA}(H_\xi) \Leftrightarrow 1 - F(x) = x^{-\frac{1}{\xi}} L(x).$$

Poznámka 2.3.7. Parameter $\frac{1}{\xi}$ vo vete 2.3.5 je tiež nazývaný index chvosta rozdelenia (Tail Index), často sa označuje ako α . Určuje stupeň klesania chvosta rozdelenia.

Veta 2.3.5 sa dá použiť pri dokazovaní, či distribučná funkcia daného rozdelenia patrí, alebo nepatrí do sféry príťažlivosti Fréchetovho rozdelenia.

Lemma 2.3.5. Nech F je distribučná funkcia Paretoho rozdelenia, ktoré je definované v definícii 2.3.2. Potom $F \in \text{MDA}(H_\xi)$, kde $\xi > 0$.

Dôkaz. Podľa vety 2.3.5 distribučná funkcia má mať nasledujúci tvar

$$F(x) = 1 - x^{-\alpha}L(x),$$

kde $L(x)$ je pomaly sa meniacou funkciou v nekonečne. Podľa definície 2.3.2 má distribučná funkcia Paretoho rozdelenia tvar

$$F(x) = 1 - \left(\frac{\kappa}{\kappa + x} \right)^\alpha.$$

Po jeho úprave dostávame požadovaný tvar distribučnej funkcie

$$F(x) = 1 - x^{-\alpha}L(x),$$

kde

$$L(x) = (\kappa^{-1} + x^{-1})^{-\alpha}.$$

Teraz stačí overiť, že $L(x)$ je pomaly sa meniacou funkciou v nekonečne. Čo overíme podľa definície 2.3.5.

$$\lim_{x \rightarrow \infty} \frac{L(tx)}{L(x)} = \lim_{x \rightarrow \infty} \frac{(\kappa^{-1} + (tx)^{-1})^{-\alpha}}{(\kappa^{-1} + x^{-1})^{-\alpha}} = \frac{\lim_{x \rightarrow \infty} \left(\frac{1}{\kappa} + \frac{1}{tx} \right)^{-\alpha}}{\lim_{x \rightarrow \infty} \left(\frac{1}{\kappa} + \frac{1}{x} \right)^{-\alpha}} = 1.$$

□

Teraz si ukážeme jeden z možných spôsobov odvodenia Hillovho odhadu vhodnej výšky medze u , ktorý vychádza z výpočtu podmienenej strednej hodnoty výšky logaritmu excesov nad logaritmom medze u .

Veta 2.3.6. Nech F je distribučná funkcia, pre ktorú platí

$$F \in \text{MDA}(H_\xi) \text{ pre } \xi > 0$$

a majme náhodnú veličinu $X \sim F$. Označme

$$e^*(\ln u) = \mathbb{E}(\ln X - \ln u | \ln X > \ln u).$$

Potom platí, že

$$e^*(\ln u) = \frac{1}{1 - F(u)} \int_u^\infty L(x)x^{-(\alpha+1)}dx.$$

Dôkaz. Nech $v < \infty$, potom

$$\begin{aligned}
\int_u^v (\ln x - \ln u) dF(x) &= [F(x)(\ln x - \ln u)]_u^v - \int_u^v \frac{1}{x} F(x) dx = \\
&= F(v)(\ln v - \ln u) - \int_u^v \frac{1}{x} F(x) dx = \\
&= - (1 - F(v))(\ln v - \ln u) - \int_u^v \frac{1}{x} F(x) dx + \int_u^v \frac{1}{x} dx = \\
&= - (1 - F(v))(\ln v - \ln u) + \int_u^v \frac{1}{x} (1 - F(x)) dx.
\end{aligned}$$

Pre $v \rightarrow \infty$ platí, že

$$\begin{aligned}
(1 - F(v))(\ln v - \ln u) &= \int_v^\infty dF(x) \cdot (\ln v - \ln u) \leq \\
&\leq \int_v^\infty (\ln x - \ln u) dF(x) = 0,
\end{aligned}$$

pretože predpokladáme, že

$$\int_0^\infty (\ln x - \ln u) dF(x) < \infty.$$

Pre $v \rightarrow \infty$ teda platí, že

$$\begin{aligned}
e^*(\ln u) &= \int_u^v (\ln x - \ln u) d\frac{F(x)}{1 - F(u)} = \\
&= \frac{1}{1 - F(u)} \int_u^v (\ln x - \ln u) dF(x) = \\
&= \frac{1}{1 - F(u)} \int_u^\infty \frac{1}{x} (1 - F(x)) dx = \\
&= \frac{1}{1 - F(u)} \int_u^\infty \frac{L(x)x^{-\alpha}}{x} dx = \frac{1}{1 - F(u)} \int_u^\infty L(x)x^{-(\alpha+1)} dx.
\end{aligned}$$

□

V ďalšom texte budeme využívať *Karamatovu vetu* z [2], str. 495.

Veta 2.3.7. (*Karamata*). Nech L je pomaly sa meniacia funkcia lokálne obmedzená na $[x_0, \infty)$ pre nejaké $x_0 \geq 0$. Potom pre $\kappa < -1$ platí

$$\int_x^\infty t^\kappa L(t) dt \sim -\frac{1}{\kappa + 1} x^{\kappa+1} L(x)$$

pre $x \rightarrow \infty$.

Pre dostatočne veľkú medzu u dostávame z 2.3.7 a viet 2.3.5 a 2.3.6 vzťah

$$e^*(\ln u) \sim \frac{L(u)u^{-\alpha}\alpha^{-1}}{1 - F(u)} = \alpha^{-1}.$$

Vidíme, že hodnota $e^*(\ln u)$ bude pre dostatečne veľkú medzu u konštantná. Skonstruujeme empirické odhady tejto hodnoty pre hodnoty medze $u = x_{1,n}, \dots, x_{n,n}$, kde $x_{k,n}$ je k -ta poradová štatistika hodnôt $\{x_1, \dots, x_n\}$. Empirické odhady hodnoty $e^*(\ln u)$ spočítame ako

$$\hat{e}^*(\ln u) = \frac{\sum_{i=1}^n (\ln x_i - \ln u) I[x_i > u]}{\sum_{i=1}^n I[x_i > u]}.$$

Budeme hľadať hodnoty $k_0 < k_1$, kde $k_0, k_1 \in \{1, \dots, n\}$ také, že empirické odhady $e^*(\ln u)$ pre $u \in \{x_{k_0,n}, \dots, x_{k_1,n}\}$ budú približne konštantné a k_0 bude dostatečne veľké (predpoklad veľkosti u) a k_1 bude dostatečne malé (pre empirický odhad potrebujeme dostatočný počet presahov).

V nasledujúcom príklade si ukážeme použitie vyššie popísanej metódy odhadu vhodnej presahovej medze u .

Príklad. Používame rovnaké rozdelenia a dáta ako v príklade na konci sekcie 2.3.4. V prípade Gamma rozdelenia nie je splnená počiatočná podmienka pre použitie Hillovej metódy a to, že distribučná funkcia rozdelenia musí patriť do sféry príťažlivosti Fréchetovho rozdelenia. Napriek tomu však toto rozdelenie v rámci príkladu používame a to, ako ilustráciu správania sa tejto metódy v prípade dát, ktoré spomínanú podmienku nespĺňajú. Výsledky príkladu sú zobrazené na obrázku 2.8 spolu s konfidenčným intervalom, zostrojeným rovnakým spôsobom, ako tomu bolo v prípade obrázku 2.7. Na základe týchto výsledkov by sme za vhodnú výšku medze u v prípade Studentovho t-rozdelenia zvolili hodnotu 15. Výška medze u pre Paretovo rozdelenie by bola určená ako 60. V prípade Gamma rozdelenia vidíme, že určovanie hodnoty medze u by bolo komplikovanejšie. Pravdepodobne by za vhodnú výšku medze bola zvolená hodnota okolo 45.

Keďže v tomto príklade sme používali rovnaké dáta ako v príklade na konci sekcie 2.3.4, tak aj výsledky z oboch použitých metód by mali byť porovnateľné. V našom prípade sme v oboch prípadoch za vhodnú medzu u pre Studentovo t-rozdelenie zvolili hodnotu 15. Pri Paretovom a Gamma rozdelení sa zvolené hodnoty medze u líšili. Aj na základe nášho príkladu je možné vidieť, že voľba medze u je problematickou a veľmi subjektívnou časťou v postupe modelovania rozdelenia excesov.

2.3.6 Modelovanie chvostu rozdelenia pôvodných dát

Pri modelovaní rozdelenia dát je často výsledkom model vhodne popisujúci dáta v oblastiach ich veľkého výskytu, ale tam, kde je pozorovaných dát málo, na chvoste rozdelenia, model väčšinou nepopisuje dáta dostatočne spoľahlivo. Jedným z dôvodov je práve počet pozorovaní, na ktorých základe odhadujeme parametre modelu.

Teraz si predstavíme alternatívny spôsob, ako k modelovaniu dát o škodách z operačného rizika pristúpiť, ktorý vhodne popisuje práve oblasť chvostu rozdelenia. Pri modelovaní škôd z operačného rizika nás predovšetkým zaujíma práve táto oblasť. V tomto prípade nebudeme modelovať excesy nad dostatočne vysokou medzou u , ako tomu bolo v postupe opísanom v sekcii 2.3.2, ale budeme modelovať rozdelenie pôvodných dát o stratách, ktoré túto medzu prekračujú.

Veta 2.3.8. *Majme distribučnú funkciu F . Za predpokladu 2.3.1 pre $x \geq u$ platí, že*

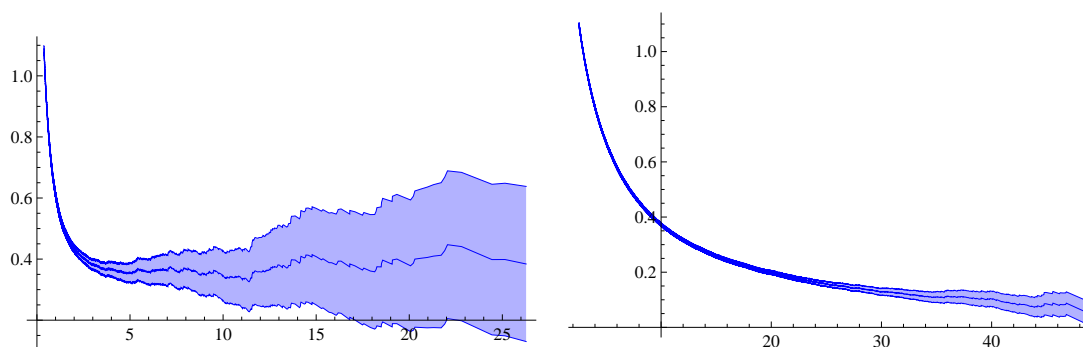
$$F(x) = 1 - (1 - F(u)) \left(1 + \frac{\xi(x - u)}{\beta}\right)^{-\frac{1}{\xi}}.$$

Dôkaz.

$$\begin{aligned} 1 - F(x) &= (1 - F(u)) \frac{1 - F(x)}{1 - F(u)} = (1 - F(u))(1 - F_u(x - u)) = \\ &= (1 - F(u))(1 - G_{\xi, \beta}(x - u)) = \\ &= (1 - F(u)) \left(1 + \frac{\xi(x - u)}{\beta}\right)^{-\frac{1}{\xi}}. \end{aligned}$$

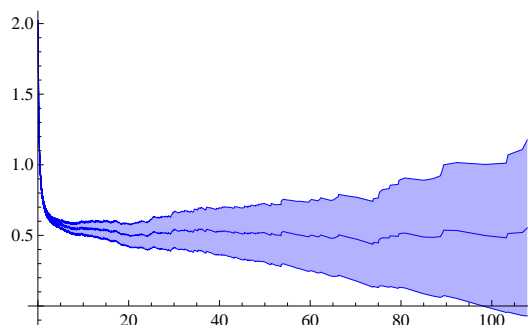
□

Z postupu odvodenia v dôkaze vety 2.3.8 je vidieť, že tento spôsob neprináša žiadne zlepšenie oproti modelu popísaného v sekcii 2.3.2. Dáva nám však návod na to, ako modelovať nielen excesy nad medzou u , ale aj celé škody, ktoré sú dostatočne vysoké, tj. prekračujú medzu u .



(a) Studentovo t-rozdelenie

(b) Gamma rozdelenie



(c) Paretovo rozdelenie

Obr. 2.8: Graf empirického odhadu podmienenej strednej výšky logaritmu excesov pre rôzne hodnoty logaritmu medze u spolu s konfidenčným intervalom pre Studentovo t-rozdelenie s 3 stupňami voľnosti, Gamma rozdelenie s parametrami $\alpha = 2$ a $\beta = 4$ a pre Paretovo rozdelenie s parametrami $\alpha = 1$ a $\kappa = 2$. Na osi x je hodnota medze u a na osi y je odhad podmienenej strednej výšky logaritmu excesov.

Veta 2.3.8 nám umožňuje vyjadriť hodnoty distribučnej funkcie daného rozdelenia v bodoch $x \geq u$ iba so znalosťou jej hodnoty v bode u a znalosťou parametrov rozdelenia $G_{\xi, \beta}$. Pre odhad distribučnej funkcie F v bodoch $x \geq u$ môžeme použiť odhad

$$\hat{F}(x) = 1 - \frac{N_u}{n} \left(1 + \frac{\hat{\xi}(x-u)}{\hat{\beta}} \right)^{-\frac{1}{\xi}},$$

kde parametre ξ a β môžeme odhadnúť metódou maximálnej vierohodnosti, prípadne momentovej metódy, ako bolo naznačené v sekcii 2.3.2 a hodnotu $F(u)$ môžeme odhadnúť empiricky ako $1 - \frac{N_u}{n}$.

Ďalším spôsobom, ako modelovať rozdelenie pôvodných dát o stratách z operačného rizika je použitím Hillovej metódy, ktorá bola opísaná v sekcii 2.3.5. Nasledujúcim textom budeme na túto kapitolu nadväzovať. Aj teraz budeme predpokladať, že podkladovým rozdelením dát, ktoré chceme modelovať, je rozdelenie s distribučnou funkciou F zo sféry príťažlivosti Fréchetovho rozdelenia.

Veta 2.3.9. *Nech F je taká distribučná funkcia, že $F \in \text{MDA}(H_\xi)$, kde $\xi > 0$ a H_ξ je distribučná funkcia zovšeobecneného rozdelenia extrémnych hodnôt. Nech x_1, \dots, x_n sú nezávislé rovnako rozdelené dáta z rozdelenia s distribučnou funkciou F . Potom odhad \hat{F} distribučnej funkcie F má pre $x \geq u$, kde u je dostatočne veľké, použitím Hillovej metódy tvar*

$$\hat{F}(x) = 1 - \frac{n-k}{n} \left(\frac{x}{x_{k,n}} \right)^{-\hat{\alpha}_{k,n}^{(H)}},$$

kde $\hat{\alpha}_{k,n}^{(H)} = \hat{e}^*(\ln x_{k,n})^{-1}$ a $\hat{e}^*(x)$ je definované v sekcii 2.3.5.

Dôkaz. Z vety 2.3.5 platí

$$1 - F(x) = L(x)x^{-\alpha} \text{ pre } x \geq u,$$

kde $\alpha > 0$ je index chvostu rozdelenia a $L(x)$ je pomaly sa meniaci funkcia v nekonečne.

Pre dostatočne veľké x nahradíme funkciu $L(x)$ konštantou C , teda

$$1 - F(x) = Cx^{-\alpha}.$$

V sekcii 2.3.5 sme ukázali, ako nájsť vhodnú hodnotu $k \in \{1, \dots, n\}$ pre odhad prevrátenej hodnoty indexu chvostu rozdelenia s distribučnou funkciou F

$$\hat{\alpha}^{-1} = \frac{1}{\hat{\alpha}_{k,n}^{(H)}} = \hat{e}^*(\ln x_{k,n}).$$

Konštantu C môžeme teraz vyjadriť ako

$$C = u^\alpha(1 - F(u)).$$

Teraz už zostáva iba odhadnúť výraz $1 - F(u)$. Tento výraz odhadneme empiricky ako

$$\frac{n-k}{n}.$$

Zložením týchto odhadov dostávame prezentovaný výsledok. □

Vzťah pre odhad chvostu plynúci z vety 2.3.9 sa dá upraviť do nasledujúceho tvaru

$$\hat{F}(x) = 1 - \frac{N_u}{n} \left(1 + \hat{\xi}^{(H)} \frac{x - u}{\hat{\xi}^{(H)} u} \right)^{-\frac{1}{\hat{\xi}^{(H)}}},$$

kde

$$\hat{\xi}^{(H)} = \frac{1}{\hat{\alpha}^{(H)}}.$$

Môžeme si všimnúť podobnosti so vzťahom pre odhad chvostu rozdelenia odvodeného z vlastností zovšeobecného Paretoho rozdelenia na začiatku tejto sekcie. Vidíme, že v tomto prípade chýba vo vzťahu parameter merítka zovšeobecného Paretoho rozdelenia β .

V nasledujúcom príklade si ukážeme použitie oboch metód odhadu chvostu pôvodných dát predstavených v tejto sekcii a porovnáme nimi získaný odhad s pôvodným rozdelením dát, ktoré je pre nás známe.

Príklad. Chvost rozdelenia pôvodných dát odhadujeme na základe vzťahov odvodených v tejto sekcii. Odhad robíme z rôzneho počtu excesov nad medzou u . Pre každý z týchto počtov excesov simulovanie dát a modelovanie chvostu opakujeme stokrát. Podobne ako v sekcii 2.3.2, aj v tomto príklade úspešnosť modelovania budeme vyjadrovať pomocou hodnoty supremálnej odchýlky² definovanej v definícii 2.3.4. V tomto prípade využívame odchýlku distribučnej funkcie pôvodného podkladového rozdelenia a odhadu distribučnej funkcie chvostu rozdelenia pôvodných dát.

Za podkladové rozdelenia aj v tomto príklade použijeme Studentovo t-rozdelenie s 3 stupňami voľnosti, Pareto rozdelenie s parametrami $\alpha = 1$ a $\kappa = 2$ a Gamma rozdelenie s parametrami $\alpha = 2$ a $\beta = 4$. Napriek tomu, že posledné spomenuté rozdelenie nespĺňa podmienku, že ide o rozdelenie s distribučnou funkciou F zo sféry príťažlivosti Fréchetovho rozdelenia, ktorá bola požadovaná v Hillovej metóde odhadu chvostu rozdelenia pôvodných dát³. Gamma rozdelenie v tomto prípade používame preto, aby sme ukázali, ako Hillova metóda funguje v prípade dát, ktorých distribučná funkcia túto podmienku nespĺňa.

Na obrázku 2.9 sú zobrazené stredné hodnoty supremálnych odchýlok distribučnej funkcie pôvodného podkladového rozdelenia a odhadu distribučnej funkcie chvostu rozdelenia pôvodných dát spolu s konfidenčným intervalom, zostrojeným rovnakým spôsobom ako v prípade grafov na obrázku 2.4. Chvosty pôvodných rozdelení sú odhadnuté obidvomi metódami a pre rôzne počty excesov nad medzou u . Parametre zovšeobecného Paretoho rozdelenia sú odhadnuté pomocou metódy maximálnej vierohodnosti a tiež momentovej metódy.

Z obrázku 2.9 vidíme, že pokiaľ chceme porovnať predstavené metódy na odhad chvostov rozdelení pôvodných dát, tak metóda odhadu chvostu založená na vete 2.3.8 vychádza vo všetkých prípadoch lepšie.

V prípade Studentovho t-rozdelenia a Paretoho rozdelenia vidíme, že konfidenčný interval sa zužuje s rastúcim počtom excesov, na rozdiel od konfidenčného intervalu pri Gamma rozdelení. Na základe obrázku 2.9, a teda na základe našich

²V tomto príklade nebudeme používať priamo supremálnu odchýlku, ale iba jej zjednodušenie, odchýlku maximálnu z poznámky 2.3.6. Navyac množina deliacich bodov D_k^s neobsahuje body od 0, ale až od u . Hodnoty k a s sú rovnaké ako v poznámke 2.3.6.

³Z [2] vieme, že Gamma rozdelenie patrí do sféry príťažlivosti Gumbelovho rozdelenia.

simulácií môžeme povedať, že výsledky tejto metódy pri Studentovom t-rozdelení a Paretovom rozdelení sú dobré. V prípade Gamma rozdelenia, nespĺňajúcom podmienku použitia Hillovej metódy, výsledky nie sú také dobré.

Na základe obrázku 2.9 môžeme taktiež porovnať metódu maximálnej vierohodnosti a momentovú metódu pri použití Hillového odhadu chvostu rozdelenia. Tu sa javí byť lepšou metódou pre odhad parametrov metóda momentová.

2.3.7 Odhadovanie vysokých kvantilov

Lemma 2.3.6. *Majme distribučnú funkciu F . Potom za predpokladu 2.3.1 a pre $\alpha \geq F(u)$ platí, že*

$$q_\alpha(F) = \text{VaR}_\alpha = u + \frac{\beta}{\xi} \left(\left(\frac{1 - \alpha}{1 - F(u)} \right)^{-\frac{1}{\xi}} - 1 \right),$$

kde VaR_α je hodnota v riziku definovaná v definícii 1.2.2.

Dôkaz. Z vety 2.3.8 dostávame vzťah

$$\alpha = 1 - (1 - F(u)) \left(1 + \frac{\xi(q_\alpha(F) - u)}{\beta} \right)^{-\frac{1}{\xi}}.$$

Odtiaľ dostávame dokazovaný vzťah pre $q_\alpha(F)$. □

Podľa lemmatu 2.3.6 môžeme kvantil $q_\alpha(F)$ odhadnúť ako

$$\hat{q}_\alpha(F) = u + \frac{\hat{\beta}}{\hat{\xi}} \left(\left(\frac{1 - \alpha}{\frac{N_u}{n}} \right)^{-\frac{1}{\hat{\xi}}} - 1 \right).$$

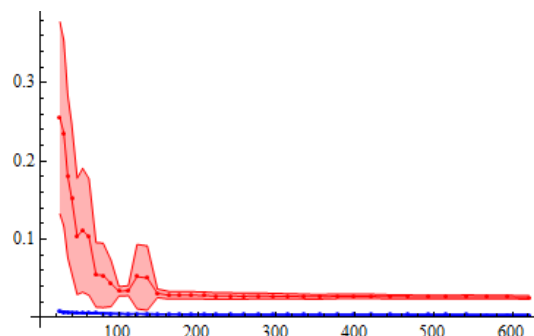
V nasledujúcom príklade sa ukáže úspešnosť odhadovania vysokých kvantilov pre Studentovo t-rozdelenie s 3 stupňami voľnosti, Gamma rozdelenie s parametrami $\alpha = 2$ a $\beta = 4$ a Paretovo rozdelenie s parametrami $\alpha = 1$ a $\kappa = 2$. Odhadnuté kvantily budeme porovnávať so skutočnou hodnotou kvantilov vypočítanou zo znalosti pôvodného rozdelenia dát.

Príklad. Odhady kvantilov $\hat{q}_\alpha(F)$ budeme robiť na základe vzťahu odvodeného v tejto sekcii, pričom používame odhad distribučnej funkcie z príkladu na konci sekcie 2.3.6 metódou založenou na vete 2.3.8. Parametre používaného zovšeobecného Paretovhov rozdelenia ξ a β odhadneme pomocou metódy maximálnej vierohodnosti aj pomocou momentovej metódy. Opäť budeme počítať hodnotu kvantilov pre rôzne počty excesov nad medzou u a pre každý z týchto počtov budeme robiť 100 testov, na základe ktorých vypočítame strednú hodnotu a konfidenčný interval.

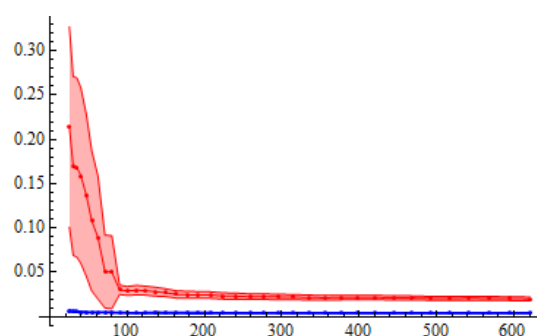
Stredné hodnoty 95% a 99% kvantilov vybraných podkladových rozdelení, Studentovo t-rozdelenie s 3 stupňami voľnosti, Paretovo rozdelenie s parametrami $\alpha = 1$ a $\kappa = 2$ a Gamma rozdelenie s parametrami $\alpha = 2$ a $\beta = 4$, pre metódu maximálnej vierohodnosti a momentovú metódu odhadu parametrov spolu s konfidenčným intervalom, zostrojeným rovnakým spôsobom ako tomu bolo v prípade

grafov na obrázku 2.4, spolu so skutočnou hodnotou kvantilov sú zobrazené na obrázku 2.10.

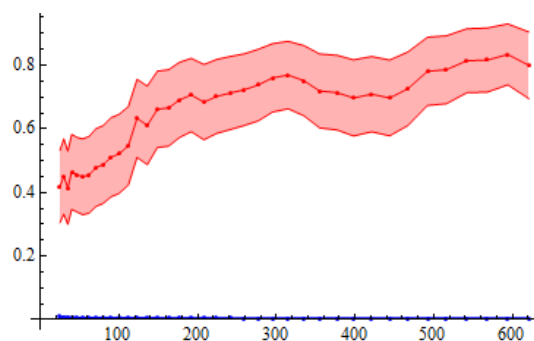
Na základe obrázku 2.10 môžeme povedať, že vo všetkých prípadoch sa metóda maximálnej vierohodnosti javí byť lepšou pre odhad kvantilu. Taktiež vidíme, že stredná hodnota kvantilu sa ustáľuje pri počte excesov okolo 300, samozrejme,



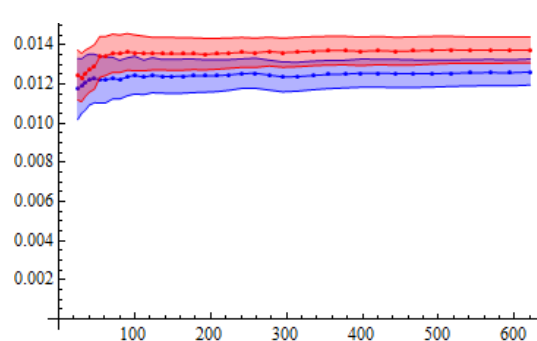
(a) Studentovo t-rozdelenie, metóda maximálnej vierohodnosti



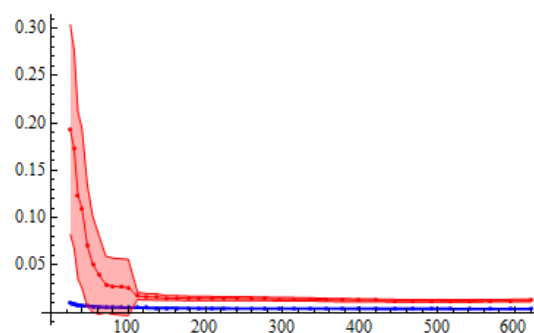
(b) Studentovo t-rozdelenie, momentová metóda



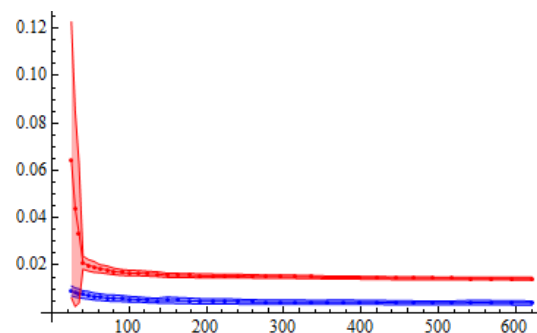
(c) Gamma rozdelenie, metóda maximálnej vierohodnosti



(d) Gamma rozdelenie, momentová metóda



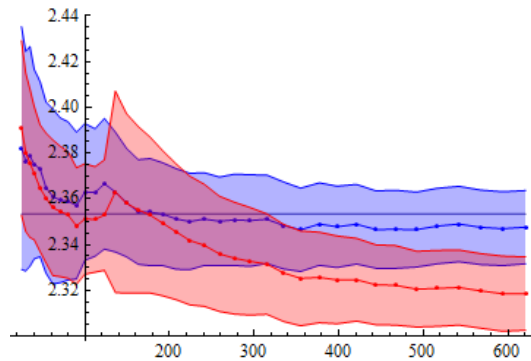
(e) Paretovo rozdelenie, metóda maximálnej vierohodnosti



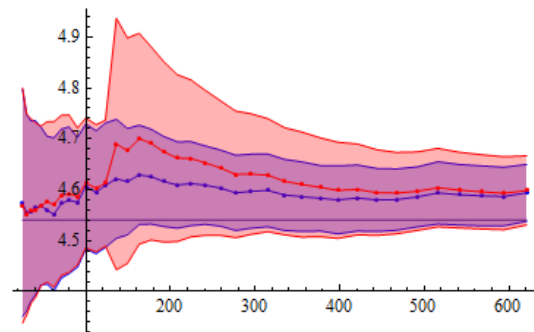
(f) Paretovo rozdelenie, momentová metóda

Obr. 2.9: Hodnoty stredných hodnôt maximálnych odchýlok $E_{max}(\hat{F}, F, k, s)$ pre obe predstavené metódy odhadu chvostov pôvodných dát (Hillova metóda na grafoch červenou) pre metódu maximálnej vierohodnosti a momentovú metódu odhadu parametrov. Na osi x je počet dát prekračujúcich medzu u a na osi y je priemerná hodnota maximálnej odchýlky.

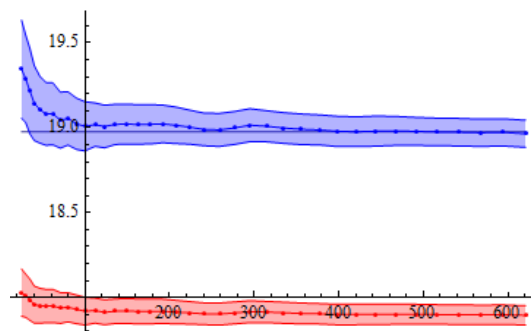
v závislosti na rozdelení a výške kvantilu. Z obrázku 2.10 je tiež vidieť, že pre 99% kvantil je pre ustálenie strednej hodnoty kvantilu potrebný vyšší počet excesov ako pri 95% kvantile. Ďalej môžeme povedať, že odhad kvantilov je úspešnejší pre nižšiu hodnotu kvantilu, tj. pre 95% kvantil.



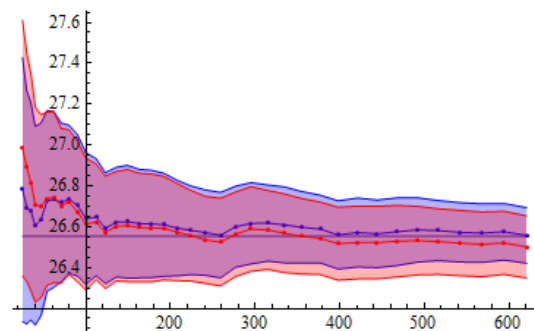
(a) Studentovo t-rozdelenie, 95% kvantil



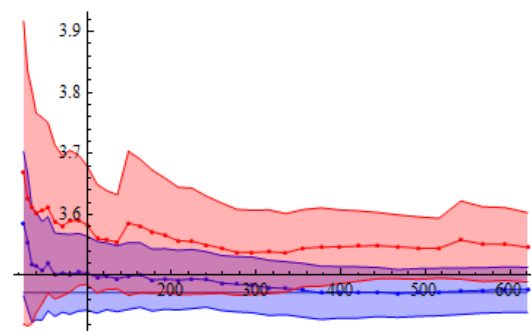
(b) Studentovo t-rozdelenie, 99% kvantil



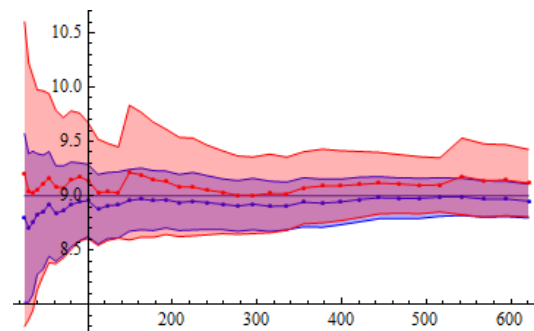
(c) Gamma rozdelenie, 95% kvantil



(d) Gamma rozdelenie, 99% kvantil



(e) Paretovo rozdelenie, 95% kvantil



(f) Paretovo rozdelenie, 99% kvantil

Obr. 2.10: Teoretická a odhadnutá hodnota 95% a 99% kvantilu pre Studentovo t-rozdelenie s 3 stupňami voľnosti, Gamma rozdelenie s parametrami $\alpha = 2$ a $\beta = 4$ a pre Paretovo rozdelenie s parametrami $\alpha = 1$ a $\kappa = 2$. Odhadnutá hodnota vypočítaná použitím momentovej metódy (na grafoch červenou) a metódy maximálnej vierohodnosti (na grafoch modrou) pre odhad parametrov ξ a β nášho modelu. Modrá vodorovná čiara v grafoch označuje skutočný kvantil jednotlivých podkladových rozdelení. Na osi x je počet dát prekračujúcich medzu u a na osi y sú hodnoty kvantilov.

Kapitola 3

Záver

V práci sme predstavili nielen pojem operačné riziko, direktívy, ktoré sú výsledkom iniciatívy vytvorenia medzinárodne platných požiadaviek na riadenie operačného rizika v bankách a poisťovniach, jednotlivé metódy výpočtu kapitálovej požiadavky na straty z operačného rizika, ale predovšetkým sme sa zamerali na spôsoby modelovania rozdelenia dát o stratách plynúcich z tohto rizika.

Popísali sme nielen ich teoretický základ, ale sme ich aj úspešne predviedli v rámci testovania úspešnosti týchto modelov na nasimulovaných dátach zo širšej škály podkladových rozdelení. Tieto sme vybrali tak, aby sme spomínané testovanie predviedli na rozdeleniach zo sfér príťažlivosti rôznych druhov rozdelení. Na základe prevedených simulácií prezentovaných v tejto práci sme prišli k záveru, že skúmané metódy sú v praxi dobre využiteľné.

Literatúra

- [1] Paul Embrechts, Claudia Klüppelberg, and Thomas Mikosch. *Modelling Extremal Events for Insurance and Finance*. Springer Verlag Berlin - Heidelberg, 1997.
- [2] Alexander J. McNeil, Rüdiger Frey, and Paul Embrechts. *Quantitative Risk Management: Concepts, Techniques, and Tools*. Princeton University Press, 2005.
- [3] Alexander J. McNeil and Thomas Saladin. The Peaks over Threshold Method for Estimating High Quantiles of Loss Distributions. *Proceedings of XXVIIth International Astin Colloquium*, 1997. <ftp://hubble2-1.math.ethz.ch/users/mcneil/cairns.pdf>.
- [4] Basel Committee on Banking Supervision. Basel II: International Convergence of Capital Measurement and Capital Standards: A Revised Framework. *Bank for International Settlements*, 2004. <http://www.bis.org/publ/bcbs107.pdf>.
- [5] Harry H. Panjer. *Operational Risk: Modeling Analytics*. John Wiley and Sons, Inc, 2006.
- [6] The European Parliament and the Council of the European Union. Directive 2009/138/EC of the European Parliament and of the Council of 25 November 2009 on the taking-up and pursuit of the business of Insurance and Re-insurance (Solvency II). *Bank for International Settlements*, 2009. <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2009:335:0001:0155:EN:PDF>.
- [7] Hansson O. Sven. Risk. *The Stanford Encyclopedia of Philosophy*, 2011. <http://plato.stanford.edu/archives/fall2011/entries/risk/>.
- [8] Vladimir N. Vapnik. *The Nature of Statistical Learning Theory*. Springer-Verlag New York, Inc., 1995.