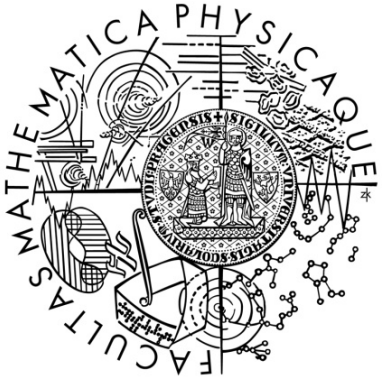
Univerzita Karlova v Praze
Matematicko-fyzikální fakulta

BAKALÁŘSKÁ PRÁCE



Anastasia Tyuleneva

Odhady metodou maximální věrohodnosti a jejich aproximace

Katedra pravděpodobnosti a matematické statistiky

Vedoucí bakalářské práce: Mgr. Vadym Omelchenko

Studijní program: Matematika

Studijní obor: FMMAT

Praha 2013

V této části své práce chtěla bych poděkovat Mgr. Vadymu Omelchenku, vedoucímu bakalářské práce, za jeho pomoc a za mnoho cenných rad získaných při psaní této práce a při vypracovávání praktické části.

Prohlašuji, že jsem tuto bakalářskou práci vypracovala samostatně a výhradně s použitím citovaných pramenů, literatury a dalších odborných zdrojů.

Beru na vědomí, že se na moji práci vztahují práva a povinnosti vyplývající ze zákona č. 121/2000 Sb., autorského zákona v platném znění, zejména skutečnost, že Univerzita Karlova v Praze má právo na uzavření licenční smlouvy o užití této práce jako školního díla podle § 60 odst. 1 autorského zákona.

V ….....dne..... podpis

Název práce: Odhady metodou maximální věrohodnosti a jejich aproximace

Autor: Anastasia Tyuleneva

Ústav: Katedra pravděpodobnosti a matematické statistiky

Vedoucí bakalářské práce: Mgr. Vadym Omelchenko

Abstrakt: Metoda maximální věrohodnosti je jedna z nejoptimálnějších a nejpřesnějších metod, kterých lze použít pro odhady rozdělení a parametru. V této práci se seznámíme s plusy a mínusy této metody a porovnáme ji s jinými odhadovými modely. V teoretické části uvedeme důležité pojmy a věty pro definování obecného postupu při odhadování parametru a pro práci s realnými daty. V praktické části aplikujeme MMV na vzorových rozděleních pro nalezení neznámých parametrů. Na závěr aplikujeme tuto metodu na reálných datech cen a výnosu EEX AG, Germani. A taktéž ji porovnáme s jinými modely pro odhadování rozdělení a parametru a vybereme nejlepší rozdělení z nabízených. Vsechny testy a odhady budou prováděny pomoci softwaru Mathematica.

Klíčová slova: odhady parametru, Metoda Maximální věrohodnosti, MMV, Stabilní rozdělení, Charakteristická funkce, Test dobry shody, Rao-Cramer.

Title: Maximum likelihood estimators and their approximations

Author: Anastasia Tyuleneva

Department: Department of Probability and Mathematical Statistics

Supervisor: Mgr. Vadym Omelchenko

Abstract: Maximum likelihood estimators method is one of the most effective and accurate methods that was used for estimation distributions and parameters. In this work we will find out the pros and cons of this method and will compare it with other estimation models. In the theoretical part we will review important theorems and definitions for creating common solution algorithms and for processing the real data. In the practical part we will use the MLE on the case study distributions for estimating the unknown parameters. In the final part we will apply this method on the real price data of EEX A. G, Germani. Also we will compare this method with other typical methods of estimation distributions and parameters and chose the best distribution. All tests and estimators will be provided by Mathematica software.

Keywords: parametr estimates, Maximum Likelihood estimators, MLE, Stable distribution, Characteristic function, Pearson's chi-squared test, Rao-Crámer.

.

Obsah

Úvod

V této práci se seznámíme s Metodou maximální věrohodnosti a její aproximací v praxi. Řekneme hlavní myšlenku této metody a její základní cíle. Budeme mluvit o plusech a mínusech, také budeme uvažovat o spojení s jinými metodami, jakými jak metoda nejmenších čtverců.

Zavedeme několik důležitých pojmů, definic a vět, který nám umožní pracovat s různými veličinami a dovolí aplikovat na nich naši metodu. Sestavíme obecný postup realizace odhadu parametru a použijeme ho na konkrétních příkladech pro konkrétní rozdělení.

Budeme mluvit o situacích, ve kterých není možné postupovat podle obecných vzorců, ale kde musíme pro dosazení rezultátu jít přes charakteristickou funkci a jí příslušné vlastnosti.

Odvodíme odhady parametru pro různé velečiny, které budou mít přesně zadanou pravděpodobnostní funkci a pro různý rozsah těchto velečin, s pomocí MMV(metoda maximální věrohodnosti). Také si ukážeme aplikaci metody na datech, ve kterých je pravděpodobnostní funkce těžko definovatelná a jak se tento problém dá obejít. Provedeme srovnávací analýzu dosazených výsledků s oběma problematikami.

Na závěr uvidíme, jak se metoda realizuje v praxi na reálných veličinách: v oceňovaní cen finančních dat.

1. Metoda Maximální věrohodnosti

Před tím než začneme mluvit o tom, co je metoda maximální věrohodnosti a jak ji můžeme aproximovat, měly bychom se vrátit až k samému začátku. První člověk, který začal brát v úvahu danou metodu, metodu maximální věrohodnosti, byl známý anglický statistik, evoluční biolog, eugenik a genetik R. A. Fisher [1] . Během 1912- 1922 let Ronald Fisher otevřel pro svět statistiky nový horizont možností. Tento postup už byl použít a zmíněn v dřívějších pracích různých vědců. Například stručný popis metody můžeme najít u Gaussu, Laplace, Lambert a tak dále. Vývoj MMV (metoda maximální věrohodnosti) můžeme sledovat a i v jiných dílech různých autorů a jak ukazuje praxe, tato témata nabírají větší a větší popularity mezi lidmi a každý den se objeví nová informace o maximální věrohodnosti odhadů a nový způsob použití této metody.

1.1. Idea Metody Maximální Věrohodnosti

Hlavní myšlenka metody maximální věrohodnosti vychází z odhadu neznámých parametrů, které maximalizují hodnoty námi použité pravděpodobnosti údaji podle statistického modelu. Výhoda MMV je v poskytovaní spolehlivého a více efektivního odhadu z dat. Metoda maximální věrohodnosti je natolik univerzální, že se hodí pro většinu statistických modelů a pro různé typy dat a udává výsledek co nejvíc odpovídající realitě. Bohužel ale i tato metoda není ideální. Jak víme, máme svobodu ve výběru dat a modelů pro práci, ale nemůžeme je využít, pokud nejsou splněny podmínky pro použití MMV.

Co to znamená? K využití této metody a výpočtu potřebujeme vědet přesnou formu rozdělení, resp. pravděpodobnostní funkce a to je ten hlavní problém. Protože neexistuje vždy konkrétní distribuce a co je nejhorší, někdy bude docela obtížné

definovat počáteční rozdělení a bez přesného popisu získané odhady nebudou odpovídat získaným datům.

Navíc, nevíme dopředu, jak moc jsou komplikované parametry a věrohodnosti funkce zvoleného modelu, čímž můžeme dojít k situaci, ve které nebude existovat analytické řešení a pro hledaní maxima bude nutné použít nějaké jiné metody.

Bez ohledu na mínusy, MMV zůstává docela univerzální pro práci s daty. Pro pevnou sadu dat a základní pravděpodobnostní model s pomocí maximální věrohodnosti dostaneme hodnoty parametrů modelu, které budou tvořit údaje "blíže" k realitě. Příklady použití metody maximální věrohodnosti ve statistikách můžeme najít v následujících modelech, jako jsou:

- lineární modely a zobecněné lineární modely
- faktorová analýza
- modelování strukturálních rovnic
- testování hypotéz
- diskrétní modely výběru.

Zajímavý je ten fakt, že odhady MV (maximální věrohodnosti) jsou specifické pro daný konkrétní typ rozdělení, ale značení metody může být o hodně širší. Myšlenka stojí na tom, že proces získánívání odhadů pro jednu distribuci lze rozšířit na "podobnou" distribuci. Tyto metody se nazývají kvazi-nebo pseudo-MMV [2].

1.2. Vztah metody maximální věrohodnosti a Metoda nejmenších čtverců

Metoda maximální věrohodnosti se používá k nalezení způsobů výpočtu a jenom pak ukazuje, jaké vlastnosti má tato metoda vzhledem k nějaké větší třídě distribucí. Například, MMV i regrese s normálními distribuovanými chybami nám dává metodu nejmenšího čtverce, která už má "Dobré" vlastnosti, ale i chyby, které nemají normální

rozdělení. Samozřejmě efektivita nebude tak perfektní jako samotné využívaní MMV, ale s jiné strany dostaneme výsledek, což nám mohlo by vyhovovat. Mezi oběma metodami existuje inverzní vztah. MNČ (Metodu nejmenších čtverců) můžeme použít jako výpočetní postup, který nám dovoluje najít odhady MV a pomáhá v sestavení testů. Tento technicky trik nazýváme pomocnou regresí.

1.3. Základní pojmy

Abychom lépe pochopily, jak funguje MMV, ukážeme její aproximaci na malém jednoduchém příkladu:

Příklad 1. Chceme odhadnout "binomický" parametr p, to znamená, že chceme vědět, jaká je pravděpodobnost v jednom pokusu. Provedeme 100 pokusů, z nich 80 bude úspěšných a ostatních 20 neúspěšný. Dostaneme pravděpodobnost pozorovaní k úspěchu v 100 pokusech popsané pravděpodobnostní funkce binomického rozdělení, pro které platí p je pevný parametr a k pochybuje od 0 do 100:

$$\binom{100}{k} p^k (1-p)^{100-k}$$

V našem případě víme jen parametr k, k = 80, potřebujeme odhadnout p. Dosadíme k do pravděpodobnostní funkce čímž dostaneme funkce proměnné p . Pak budeme psát

$$L(p) = \binom{100}{80} p^{80} (1-p)^{20}$$

danou funkci budeme nazývat *funkcí věrohodnosti* a hodnota $\hat{p}$, která maximalizuje tuto funkci, se nazývá *maximální věrohodný odhad*. Tímto způsobem metoda vybere vhodný odhad parametru p, pro které platí, že pravděpodobnost získaní dat je nejvyšší. Samozřejmě skutečná hodnota parametru p je pevná a funkce $L(p)$ je funkce

neznámého parametru [3]. Pro maximalizaci funkce využijeme logaritmus, s pomocí kterého je náš vzorec

$$\ln L(p) = \ln \begin{pmatrix} 100 \\ 80 \end{pmatrix} + 80 \ln p + 20 \ln(1 - p) .$$

Protože logaritmus je rostoucí funkce na celém oboru hodnot $L(p)$ hodnoty maxima $L(p)$ a $ln L(p)$ budou ve stejných bodech, to znamená že maximum bude v těch bodech, pro kterých platí, že derivace podle parametru p funkce $ln L(p)$ rovná se nule

$$\frac{\partial \ln L(p)}{\partial p} = \frac{80}{p} - \frac{20}{1 - p} = 0 .$$

Odtud $\hat{p} = 0.8$. Pro tento případ, nalezený odhad je maximálně věrohodným odhadem.

Popíšeme celou metodu ještě jeden krát, ale tentokrát obecně [3].

Defenice 1. *Nechť náhodný vektor* $\boldsymbol{X}=(X_1,\ldots,X_n)$ *má sdruženou hustotu* $p(\boldsymbol{x},\theta)$ *kde* $\theta \in \Omega$ *a* $X_1,\ldots,X_n$ *je náhodný výběr. Při pevné hodnotě* $\boldsymbol{x}$ *se funkce* $p(\boldsymbol{x},\theta)$ *jakožto funkce* θ *nazývá věrohodností funkce. Hodnota* $\hat{\theta}$ *parametru* θ, *která maximalizuje věrohodností funkci* $p(\boldsymbol{x},\theta)$ *pro dané* $\boldsymbol{X}=x$, *se nazývá maximálně věrohodný odhad parametru* θ.

Pro následující práci s MMV a odhadem parametru θ, potřebujeme několik základních vět pro danou metodu, které by nám dovolily dělat vážné předpoklady a potvrdily by správnost dané metody.

1.4. Konsistentní odhad

Jedna ze základních vlastností, která nám říká o tom, co je "dobry" odhad nebo ne- konsistence. Konsistence se udává s růstem *n*, sám odhad se bude blížit ke skutečné hodnotě odhadovaného parametru [2, 4].

Defenice 2. *Nechť θ je jednorozměrný parametr, $X_1,\ldots,X_n$ je výběr z nějakého rozdělení, které závisí na parametru θ a je definován odhad $T_n = g(\mathrm{X}_1,\ldots,\mathrm{X}_n)$ pro každé přirozené n. Pak odhad T_n je konsistentní, pokud $T_n \to \theta$ podle pravděpodobnosti pro $n \to \theta$.*

Poznámka 1. Jestliže $ET_n^2 < \infty$ pro každé n, pak platí vztahy $E(T_n) \to \theta$ a $\lim_{n\to\infty} var(T_n)=0$, které nám říká že T_n je *konsistentním odhadem parametru* θ.

Konsistentní odhad potřebujeme po případ, kdy bude velmi obtížné vypočítat maximálně věrohodný odhad. Ale s pomocí nějakého $\sqrt{n}-$ konsistentního odhadu $\tilde{\theta}_n$, se da aproximovat jako maximálně věrohodný odhad $\hat{\theta}_n$, kde $\tilde{\theta}_n$ je $\sqrt{n}$ *-konsistentním odhadem* parametru θ, je-li posloupnost $\sqrt{n}(\tilde{\theta}_n - \theta)$ omezena v pravděpodobnosti. Pak je to jasně vidět, že $\sqrt{n}-$ *konsistentní odhad* [4] je také *konsistentní*, to znamená, že platí $\tilde{\theta}_n \xrightarrow{P} \theta$. Dál s pomocí Newtonovy-Prahsonovy metody dostaneme rovnici tvaru

$$L'(\tilde{\theta}_n) + (\theta - \tilde{\theta}_n)L''(\tilde{\theta}_n) = 0.$$

Jejím řešením bude $\theta = \delta_n$, kde

$$\delta_n = \tilde{\theta}_n - \frac{L'(\tilde{\theta}_n)}{L''(\tilde{\theta}_n)}. \tag{1.1}$$

1.5. Raova- Cramérova věta

Nechť θ je jednorozměrný parametr a nechť náhodný vektor $\mathbf{X} = (X_1, \ldots, X_n)'$ má hustotu $p(\boldsymbol{x}, \theta)$ vzhledem k nějaké σ-konečně míře μ. Nechť $T = T(\mathbf{X})$ je nestranný odhad pro parametrickou funkci $g(\theta)$.

Defenice 3. *Nechť* $\{p(\boldsymbol{x}, \theta), \theta \in \Omega\}$ *je systém hustot a je regulární, a jsou splněny podmínky:*

- *Množina* Ω *je neprázdná a otevřená.*
- *Množina* $M = \{\boldsymbol{x}: p(\boldsymbol{x}, \theta) > 0\}$ *nezávisí na parametru* θ.
- *Pro skoro všechna* $x \in M$ *existuje konečná parciální derivace*

$$p'(\boldsymbol{x}, \theta) = \frac{\partial p(\boldsymbol{x}, \theta)}{\partial \theta}$$

- *Pro všechna* $\theta \in \Omega$ *platí* $\int_M p'(\boldsymbol{x}, \theta) d\mu(\boldsymbol{x}) = 0$

Integrál

$$J_n(\theta) = \int_M \left[\frac{p'(\boldsymbol{x}, \theta)}{p(\boldsymbol{x}, \theta)} \right]^2 p(\boldsymbol{x}, \theta) d\,\mu(\boldsymbol{x})$$

je konečný a kladný, je $J_n(\theta)$ *Fisherova míra informace* [4, 5].

Při prací s odhady parametru nás bude zajímat rozptyl, protože je jedním z nedůležitých ukazatelů, který nám říká o kvalitě odhadu. Pro každé rozdělení se kterým budeme pracovat, potřebujeme nějakou omezující podmínku, dolní mez, která nás omezí a nedovolí nám jít za ni.

Proto použijeme Rao-Cramerovu mez, která vytvoří takový odhad rozptylu, který bude nejmenší dosažitelný a je na dolní hranici.

Věta 1 *(Raova-Cramérova)* *Nechť T je nestranný odhad pro parametrickou funkci $g(\theta)$, a platí $\mathrm{E}T^2 < \infty$ pro každý $\theta \in \Omega$. Pokud jsou splněny předpoklady:*

- *$\{p(\boldsymbol{x}, \theta), \theta \in \Omega\}$ je systém hustot a je regulární.*
- *Derivace $g'(\theta)$ existuje v každém bodě $\theta \in \Omega$.*
- *Platí $\frac{d}{d\theta}\int_M T(\boldsymbol{x}) p(\boldsymbol{x}, \theta) d\mu(\boldsymbol{x}) = \int_M T(\boldsymbol{x}) p'(\boldsymbol{x}, \theta) d\mu(\boldsymbol{x})$.*

Pak pro každé $\theta \in \Omega$ platí

$$\mathrm{E}[T - g(\theta)]^2 \geq \frac{[g'(\theta)]^2}{J_n(\theta)} \tag{1.2}$$

1.6. Rao-Cramérova dolní mez

Pokud T je nestranným odhadem parametru θ a jsou splněny všechny předpoklady Rao-Cramérovy věty, pak se T nazývá nestranný *regulární odhad*. Jen z $g(\theta) = \theta$ a z vztahu (1.2) vyplývá, že

$$\mathrm{var}\, T \geq \frac{1}{J_n(\theta)} \tag{1.3}$$

Číslu $\frac{1}{J_n(\theta)}$ se říká *dolní Raova Cramérova* mez pro rozptyl regulárního nestranného odhadu [2, 4].

Eficienci e regulárního nestranného odhadu T definujeme jako

$$e = \frac{1}{J_n(\theta)\, \mathrm{var}\, T} \tag{1.4}$$

Jinými slovy je to podíl dolní Raovy-Cramérovy meze a skutečného rozptylu odhadu T. Proto platí

$$0 < e \leq 1 \tag{1.5}$$

V případě $e = 1$ se odhad T nazývá *eficientní*.

Pokud budeme mít několik nestranných odhadů parametru, eficience nám ukáže, který z nich má nejmenší rozptyl, tím se nejvíc přiblížíme k odhadovanému parametru a to je *nejlepším nestranným odhadem.*

Věta 2 *Nechť T je nestranný eficientní odhad a nechť U je regulární nestranný odhad parametru θ. Má-li U eficienci e, pak korelační koeficient mezi náhodnými veličinami T a U je roven $\sqrt{e}$.*

1.7. Fisherova míra informace

Funkce $J_n(\theta)$, jak bylo už řečeno v definice 3, se nazývá Fisherova míra informace. V práci s odhadem pro konkrétní rozdělení ji budeme docela často používat, proto má smysl zmínit pár hlavních pojmů týkajících této míry.

Věta 3 *Nechť systém hustot $\{p(\boldsymbol{x}, \theta), \theta \in \Omega\}$ je regulární. Pokud pro skoro všechny $\boldsymbol{x} \in M$ vzhledem k μ existuje*

$$p''(\boldsymbol{x},\theta) = \frac{\partial^2 p(\boldsymbol{x},\theta)}{\partial \theta^2}$$

a jestliže pro všechna $\theta \in \Omega$ platí

$$\int_M p''(\boldsymbol{x},\theta)\, d\mu(\boldsymbol{x}) = \boldsymbol{0},$$

pak

$$J_n(\theta) = -\int_M \frac{\partial^2 p(\boldsymbol{x},\theta)}{\partial \theta^2}\, p(\boldsymbol{x},\theta)\, d\mu(\boldsymbol{x}).$$

Poznamka 1. Jsou-li splněny předpoklady *věty 3* bude platit

$$J_n(\theta) = -\mathrm{E}\frac{\partial^2 \ln p(\boldsymbol{X}, \theta)}{\partial \theta^2}.$$

Věta 4 *Nechť* $X_1,\dots,X_n$ *je náhodný výběr, který má hustotu* $g(x,\theta)$ *vzhledem k* $\sigma-$*konečné míře* μ. *Předpokládáme, že systém hustot* $\{g(x,\theta),\theta\in\Omega\}$ *je regulární a má Fisherovu míru informace* $J(\theta)$. *Pak náhodný vektor* $\mathbf{X}=\left(X_1,\dots,X_n\right)'$ *má hustotu*

$$p_n(x_1,\dots,x_n,\theta)=g(x_1,\theta)\dots g(x_n,\theta)$$

vzhledem k míře $\mu_n=\mu\times\cdots\times\mu$. *Systém hustot* $\{p_n(\boldsymbol{x},\theta),\theta\in\Omega\}$ *bude také regulární a pro jeho Fisherovu míru informace* $J_n(\theta)$ *platí*

$$J_n(\theta)=n\,J(\theta).$$

1.8. Princip invariance pro maximálně věrohodné odhady

Teď nás bude zajímat případ, když vektor $\boldsymbol{X}$ má hustotu $p(\boldsymbol{x},\theta)$ a $\theta\in\Omega\subset\mathbb{R}_m$. Nechť h je taková funkce, která zobrazuje a Ω na $\Omega^*\subset\mathbb{R}_m$. Pak ke každému $\theta\in\Omega$ bude přirazen $\theta^*\in\Omega^*$ předpisem $\theta^*=h(\theta)$.

Nechť $U(\theta^*)=\{\theta:\theta\in\Omega,h(\theta)=\theta^*\}$. Definujeme $M(\boldsymbol{x},\theta^*)=\sup_{\theta\in U(\theta^*)}p(\boldsymbol{x},\theta)$, kde M jako funkce θ^* je *věrohodností funkce indukovaná parametrickou funkcí h.* Hodnotu $\hat{\theta}^*$ maximalizující $M(\boldsymbol{X},\theta^*)$ nazýváme *maximální věrohodný odhad parametrické funkce h* [4].

Věta 5 *Nechť* $\hat{\theta}$ *je maximálně věrohodný odhad parametru* θ, *pak* $h(\hat{\theta})$ *je maximálně věrohodný odhad parametru funkce* $h(\theta)$.

Dal budeme považovat θ za jednorozměrný parametr.

Předpoklady

1. *Nechť* Ω *je parametrický prostor obsahující neprázdný otevřený interval* ω *tak, že hodnota parametru* θ_0 *patří* ω.

2. *Nechť* $\boldsymbol{X} = (X_1, \ldots, X_n)^T$, *kde* X_i *nezávislé stejně rozdělené náhodné veličiny s hustotou* $p(x, \theta)$ *vzhledem k nějaké* σ *– konečně míře* μ.

3. *Nechť* $\mathrm{M} = \{x: p(x, \theta) > 0\}$ nezávisí na θ.

4. *Nechť* $\theta_1, \theta_2 \in \Omega$. *Pak* $p(x, \theta_1) = p(x, \theta_2)$ *pro skoro všechny* μ *platí právě tehdy, když* $\theta_1 = \theta_2$.

Věta 6 *Pokud* $n \to \infty$, *pak bude platit pro každé pevné* $\theta \in \Omega$, *ze* $\theta \neq \theta_0$

$$P_{\theta_0}\{p(\mathbf{X}, \theta_0) = p(\mathbf{X}, \theta)\} \to 1. \tag{1.6}$$

1.9. Odhad jednorozměrného parametru

Nechť θ je jednorozměrný parametr, pak pro jednoduchost výpočtu označíme funkce $L(\boldsymbol{x}, \theta) = \ln p(\boldsymbol{x}, \theta)$ jako funkce proměnné θ, které při pevném $\boldsymbol{x}$ budeme se nazývat *logaritmickou věrohodností funkcí.*

Rovnice s *logaritmickou věrohodností funkcí* jsou obyčejné a lehčeji řešitelné než věrohodností funkce. Ale maximální značení těchto funkcí bude stejné pro stejné argumenty.

Věta 7 *Pokud jsou splněny předpoklady* 1-4. *Na intervalu* ω *existuje* $p'(x, \theta) = \frac{\partial p(x, \theta)}{\partial \theta}$ *pro skoro všechna x. Pak pro každé* $\varepsilon > 0$, $n \to \infty$ *platí, že s pravděpodobností, která konverguje k jedné, má věrohodností rovnice*

$$\frac{\partial L(\mathbf{X}, \theta)}{\partial \theta} = 0$$

takový kořen $\hat{\theta}_n = \hat{\theta}_n(\boldsymbol{X})$, *a platí* $\left|\hat{\theta}_n - \theta_0\right| < \varepsilon$.

Věta 8 *Nechť* $\{p_n(x,\theta), \theta \in \Omega\}$ *je regulární systém hustot s Fisherovou mírou* $J(\theta)$ *a nechť platí předpoklady 1-4 a pro* $\theta_0 \in \omega$ *platí, že je skutečná hodnota parametru a jsou splněny následující předpoklady:*

1. *Pro všechna* $\theta \in \omega$ *a skoro všechna* $x \in M$ *existuje derivace*

$$p'''(x,\theta) = \left| \frac{\partial^3 \ln p(x,\theta)}{\partial \theta^3} \right| \leq \mathrm{H}(x)$$

.

2. *Pro všechna* $\theta \in \omega$ *platí* $\int_M p''(x,\theta)\, d\mu(x) = 0.$

3. *Existuje nezáporná měřitelná funkce* $H(x)$, *že* $\mathrm{E}_{\theta_0} H(x) < \infty$ *a pro skoro všechna* $x \in M$ *a pro všechna* θ *taková, že* $|\theta - \theta_0| < \varepsilon$ *pro nějaké dostatečně malé* $\varepsilon > 0$ *platí*

$$\left| \frac{\partial^3 \ln p(x,\theta)}{\partial \theta^3} \right| \leq \mathrm{H}(x)$$

.

Pak platí následující tvrzení.

1. *Jestli* $n \to \infty$, *pak*

$$\frac{1}{\sqrt{n}} L'(\theta_0) \xrightarrow{d} N[0, J_n(\theta_0)]. \tag{1.7}$$

2. *Existuje-li pro každé dostatečně velké n a pro každou hodnotu* $\boldsymbol{X}$ *takový kořen* $\hat{\theta}_n$ *věrohodnosti rovnice, že* $\hat{\theta}_n$ *je konsistentním odhadem parametru* θ_0, *pak*

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{d} N\left[0, \frac{1}{J(\theta_0)}\right] \tag{1.8}$$

Věta 9 *Nechť jsou splněny předpoklady věty 8, to pokud* $\tilde{\theta}_n$ *je* $\sqrt{n}$ *-konsistentním odhadem, pak*

$$\sqrt{n}(\delta_n - \theta_0) \xrightarrow{d} N[0, \frac{1}{J(\theta_0)}] \tag{1.9}$$

Poznámka 2. Odhad δ_n je docela blízko k označení maximálně věrohodného odhadu $\hat{\theta}_n$, pokud z předpokladu *věty 9* bude platit

$$\sqrt{n}(\delta_n - \hat{\theta}_n) \xrightarrow{P} 0.$$

2. Teorie potřebná pro práci s odhadem.

2.1. Charakteristická funkce.

Pro práci s rozděleními, které nemají přesně definovanou hustotu, jako stabilní rozdělení, nemůžeme přimo aplikovat metodu maximální věrohodnosti v tom tvaru, ve kterém jsme ji definovaly. Proto budeme používat charakteristickou funkci [3, 6], a abychom mohly pracovat dál, následuje pár definic a vět, které budou pro nás užitečné.

Defenice 4. *Charakteristická funkce náhodné veličiny X je funkce* $\psi(t) = \mathrm{E}e^{itX}$.

Pro diskrétní náhodné veličiny

$$\psi(t) = \sum_j e^{itx_j} \cdot P[X = x_j],$$

pro spojité náhodné veličiny s hustotou $p(x)$

$$\psi(t) = E[e^{itx}] = \int_{-\infty}^{\infty} e^{itx} p(x)\,dx,$$

kde $t \in R, i = \sqrt{1}$ *– imaginární jednotka.*

Protože $\left|e^{itx}\right| = 1, \forall t \in R,$ charakteristická funkce bude existovat pro jakékoli reálné náhodné veličiny.

Defenice 5. *Charakteristickou funkci jednoznačně určuje distribuční funkce F náhodné veličiny X a obsahuje celou informaci o rozdělení.*

Věta 10 *(Vlastnosti charakteristické funkce)* *Nechť X je náhodná veličina a ψ její charakteristická funkce. Potom*

- $\psi(t)$ *existuje pro každé* $t \in R$;
- $\psi(0) = 1$;
- $|\psi(t)| \leq 1$.

Defenice 6. *Nechť* $\mathbf{X} = (x_1, \ldots, x_n)$ *je náhodný výběr, pak můžeme určit s pomocí výběrové statistiky charakteristickou funkci*

$$\hat{\psi}(t) = \frac{1}{N}\sum_{j=1}^{N} e^{itx_j} = \frac{1}{N}\sum_{j=1}^{N} (\cos(t x_j) + i \sin(t x_j)).$$

2.2. Test dobré shody

Velmi často potřebujeme z daných údajů zjistit typ rozdělení, závislost či nezávislost, zda se napozorované četnosti významně liší od četnosti, kterou jsme očekávali na základě položené hypotézy. Dále ukážeme, že jistý speciální poměr vytvořený v průběhu testu dobré shody [4] je asymptoticky roven rozdělení χ^2.

Defenice 7. *Nechť* $A_1, \cdots, A_k$ *jsou neslučitelné jevy, z nichž v průběhu náhodného pokusu musí nastat pravě jeden. Nechť* $P(A_i) = p_i > 0$, *pro* $i = 1, \ldots, k$, *kde*

$$0 < p_i < 1, \; p_1 + \cdots + p_k = 1.$$

Náhodný pokus je opakován n-krát, X_i *je počet výskytu jevu* A_i *v těchto n opakovaných náhodných pokusech. Sdružené rozdělení veličin* $X_1, \cdots, X_k$ *je multinomické. Budeme ho značit* $M(n; p_1, \cdots, p_k)$, *resp.* $M(n; \mathbf{p})$ *a daným vzorcem*

$$P(X_1 = x_1, \ldots, X_k = x_k) = \frac{n!}{x_1! \ldots x_k!} p_1^{x_1} \ldots p_k^{x_k}$$

pro

$$x_i = 0,1,\ldots,n \quad (i = 1,\ldots,k), \quad x_1 + \cdots + x_k = n.$$

Poznámka 3. Z předchozí definice vyplývá, že pokud máme jenom jeden náhodný jev A_i a jednu náhodnou veličinu X_i, tak dostaneme klasický případ binomického rozdělení $\mathrm{Bi}(n,p_i)$.

Pomocí této poznámky můžeme zjistit charakteristiky pro jednotlivé veličiny X_i. Z čehož vyplývá následující věta

Věta 11 *Pro multinomické rozdělení* $M(n,\boldsymbol{p})$ *platí*

$$\mathrm{E}\,X_i = np_i, \quad \mathrm{var}\,X_i = np_i(1-p_i), \quad 1 \le i \le k,$$

$$\mathrm{cov}(X_i, X_j) = -np_ip_j, \quad 1 \le i \ne j \le k.$$

Důkaz Víc ve svazku [4], str. 268.

Věta 12 *Nechť* $\boldsymbol{X} = (X_1,\cdots,X_k)' \sim M(n;\ p_1,\cdots,p_k)$, *pak Pearsonova statistika* [7]

$$\chi^2 = \sum_{i=1}^{k} \frac{(X_i - np_i)^2}{np_i} \qquad (2.1)$$

má při $n \to \infty$ *asymptotickém rozdělení* χ^2_{k-1}.

Poznámka 4. Vzorec (2.1) lze upravit na tvar

$$\chi^2 = \sum_{i=1}^{k} \frac{X_i^2}{np_i} - n. \qquad (2.2)$$

Hodnoty veličin $X_1,\cdots,X_k$ se nazývají *empirická četnost* a číslům $np_1,\cdots,np_k$ se říká *teoretická četnost.*

Tato věta 12 je velmi důležitá pro nás, protože nám dovoluje testovat hypotézy o shodě teoretických a empirických četností. Je jasné, ze čím více se budou empirické

výsledky lišit od předpokladu, tím větší bude hodnota testové statistiky. Ve chvíli, když dostane $\chi^2 \geq \chi^2_{k-1}(\alpha)$, zamítneme hypotézu H_0.

Test dobré shody je asymptotický a velmi užitečný pro práci s dostatečné velkým rozsahem *n*. Teoreticky musí platit, že $np_i \geq 5$ pro každé $1 \leq i \leq k$.

Testovaní dat v této práci bude provedeno pomocí příslušné funkce, která už je reprezentovaná v softwaru Matematica.

3. Aproximace

3.1. Rozdělení I typu. Diskrétní rozdělení

3.1.1. Binomické rozdělení

Nechť $\mathbf{X} \sim Bi(n, p)$, $0 < p < 1$. Pak [4] má sdruženou hustotu vzhledem k parametru p

$$f(\boldsymbol{x}, p) = \prod_{i=1}^{n} \left(\binom{n}{x_i} p^{x_i} (1-p)^{n-x_i} \right), \text{ pro } x \neq 0, x \neq n.$$

V případě $x = 0, x = n$, maximálně věrohodný odhad parametru p nebude existovat na intervalu (0,1).

Implementaci MMV začínáme tím, že převedeme na logaritmický tvar rovnice

$$\ln f(\boldsymbol{x}, p) = n^2 \ln(1-p) + \ln \prod_{i=1}^{n} \binom{n}{x_i} + \sum_{i=1}^{n} (x_i) \ln p - \sum_{i=1}^{n} (x_i) \ln(1-p),$$

po derivaci podle parametru p dostaneme

$$\frac{\partial \ln f(\boldsymbol{x}, p)}{\partial p} = \frac{-n^2}{1-p} + \frac{\sum_{i=1}^{n} (x_i)}{p} + \frac{\sum_{i=1}^{n} (x_i)}{1-p}$$

maximalizujeme rovnice tak, že

$$\frac{\partial \ln f(\boldsymbol{x}, p)}{\partial p} = 0$$

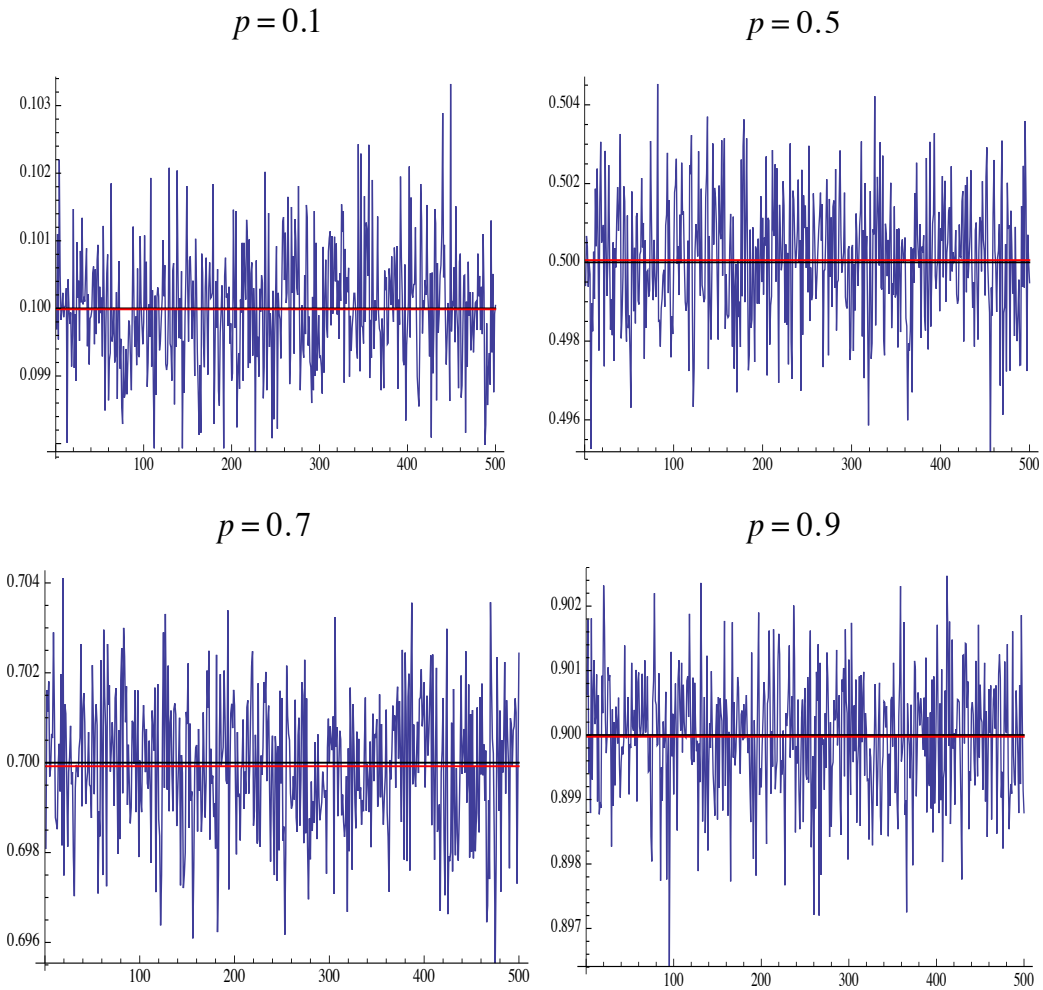
a dostaneme

$$\hat{p} = \frac{\sum_{i=1}^{n} (x_i)}{n^2} = \frac{\overline{x}}{n}$$

$\hat{p}$ je maximálně věrohodný odhad. Fišerova míra informací v případě Binomického rozdělení se bude rovnat

$$J_n(p) = \frac{n}{p(1-p)}.$$

Ověříme platnost metody, nasimulujeme data z Binomického rozdělení pro různé parametry $p = \{0.1, 0.5, 0.7, 0.9\}$ o rozsahu n=1000.



Obrázek 1. Odhady parametru MMV pro Binomické rozdělení $Bi(n, p)$. Černá přímka je parametr *p,* červená přímka je střední hodnota dosáhnutých odhadů.

Z grafů je možné vidět, že jednotlivé odhady parametrů velmi těsně kolísají kolem zadaných parametrů, černé přímky. Střední hodnota EX, červená přímka, soubor vygenerovaných dat a k nim připočtených příslušných odhadů je skoro stejná s hledaným parametrem *p*. S těchto pozorovaní můžeme vidět, že MMV funguje.

3.1.2. Poissonovo rozdělení

Nechť $\mathbf{X} \sim Po(\lambda)$, $\lambda > 0$. Pak má [3] sdruženou hustotu vzhledem k parametru λ

$$P(\boldsymbol{x}) = \prod_{i=1}^{n} \frac{e^{-\lambda}\,\lambda^{x_i}}{x_i!}$$

Pro poissonovo rozdělení funkce věrohodnosti má tvar

$$L(\lambda) = \prod_{i=1}^{n} \frac{e^{-\lambda}\,\lambda^{x_i}}{x_i!}$$

Pak implementaci MMV zahájíme tím, že převedeme na logaritmický tvar rovnice

$$\ln L(\lambda) = -n\lambda + \ln\lambda \sum_{i=1}^{n} x_i - \sum_{i=1}^{n} \ln x_i\,,$$

po derivaci podle parametru λ dostaneme

$$\frac{\partial \ln L(\lambda)}{\partial \lambda} = -n + \frac{\sum_{i=1}^{n}(x_i)}{\lambda}$$

maximalizujeme rovnici tak, že

$$\frac{\partial \ln L(\lambda)}{\partial \lambda} = 0$$

a dostaneme

$$\hat{\lambda} = \frac{\sum_{i=1}^{n}(x_i)}{n} = \overline{x}$$

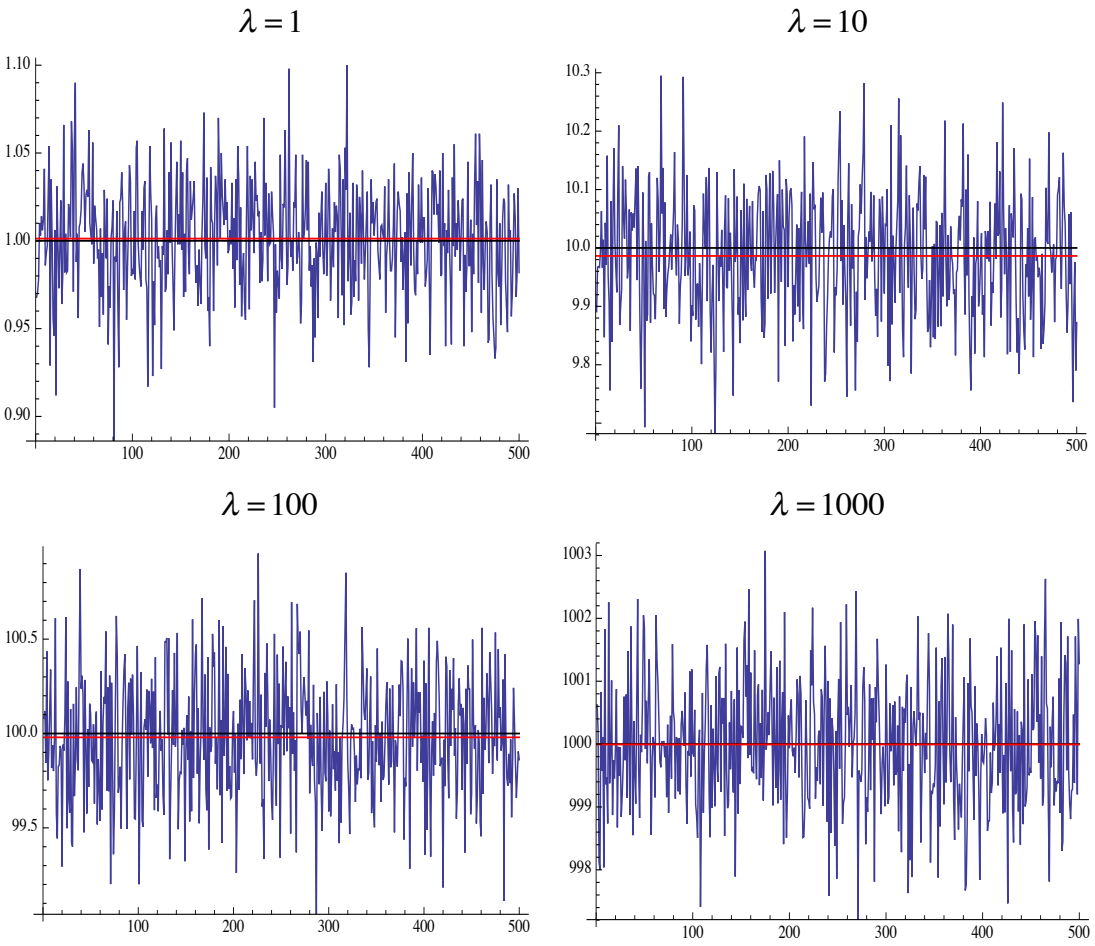
$\overline{x}$ je maximálně věrohodný odhad. Přičemž pro Poissonovo rozdělení platí, že

$$\mathrm{E}X = \mathrm{var}\, X = \lambda = \overline{x}\ .$$

Fišerova míra informací v případě Poissonova rozdělení se bude rovnat

$$J_n(p) = \frac{n}{\lambda}\ .$$

Ted‘ provedeme simulaci metody maximální věrohodnosti na příkladě Poissonova rozdělení pro různé parametry λ o rozsahu n=1000 .



Obrázek 2. Odhady parametru MMV pro Poissonovo rozdělení $Poi(\lambda)$. Černá přímka je parametr λ, červena přímka je střední hodnota dosáhnutých odhadů.

Z porovnaní grafů vidíme, že čím víc je parametr λ, tím je frekvence kolísaní odhadu kolem přímky $y = \lambda$ nebude tak hustý, jako při $\lambda = 1$, ale rozsáhlejší a velmi blízký k značení odhadnutého parametru. Jak ukazuje červená přímka, střední značení odhadu se docela málo liší od počátečních zadaných parametru.

3.2. Rozdělení II typu. Spojité rozdělení

3.2.1. Normální rozdělení se známým rozptylem

Nechť $f(\boldsymbol{x},\theta)$ je [3] sdružená hustota normálního rozdělení $N(\theta,1)$

$$f(\boldsymbol{x},\theta)=\prod_{i=1}^{n}\frac{1}{\sqrt{2\pi}}e^{\frac{(x_i-\theta)^2}{2}}=\left(\frac{1}{\sqrt{2\pi}}\right)^n\exp\left(-\frac{1}{2}\sum_{i=1}^{n}(x_i-\theta)^2\right),$$

a vzhledem k tomu, že

$$ln\,f(\boldsymbol{x},\theta)=-\frac{1}{2}ln\,2\pi-\frac{1}{2}\sum_{i=1}^{n}(x_i-\theta)^2,$$

dostáváme

$$\frac{\partial^3 \ln f(\boldsymbol{x},\theta)}{\partial^3\theta}=0.$$

Postupujeme podle metody a dostáváme do následujícího bodu

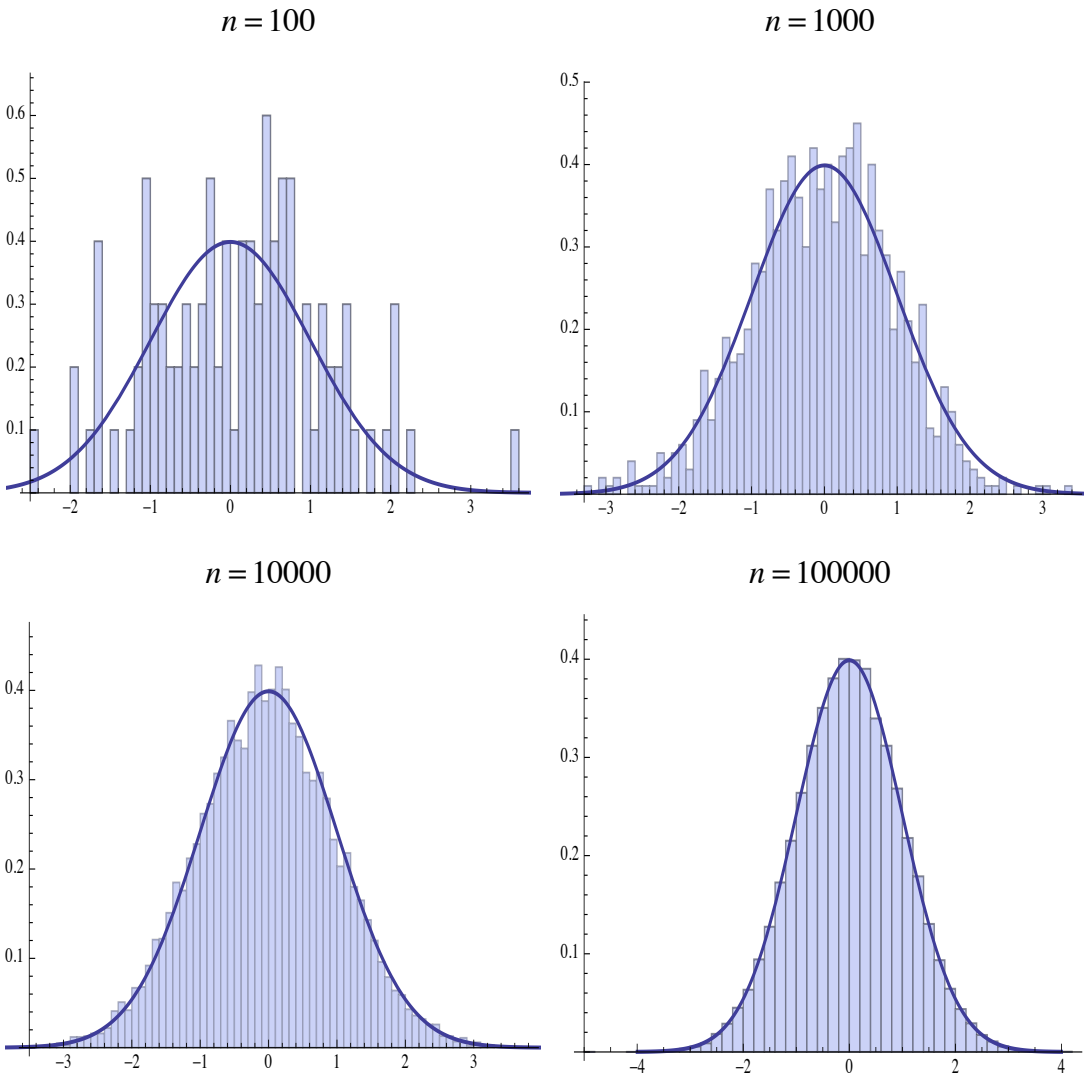
$$\frac{\partial ln\,f(\boldsymbol{x},\theta)}{\partial\theta}=\sum_{i=1}^{n}(x_i-\theta)=0,$$

a má kořen $\hat{\theta}_n=\overline{X}$. To je konsistentní odhad parametru θ.

Přitom $\overline{X}\sim\mathrm{N}\left(\theta,\frac{1}{n}\right)$ *a* $J(\theta)=1$ podle Rao-Cramerovy věty a eficientních odhadů bude platit konvergence podle distribuce $\sqrt{n}(\overline{X}-\theta)\xrightarrow{d}N(0,1)$ a pro každé *n* je to Normální rozdělení $N(0,1)$.

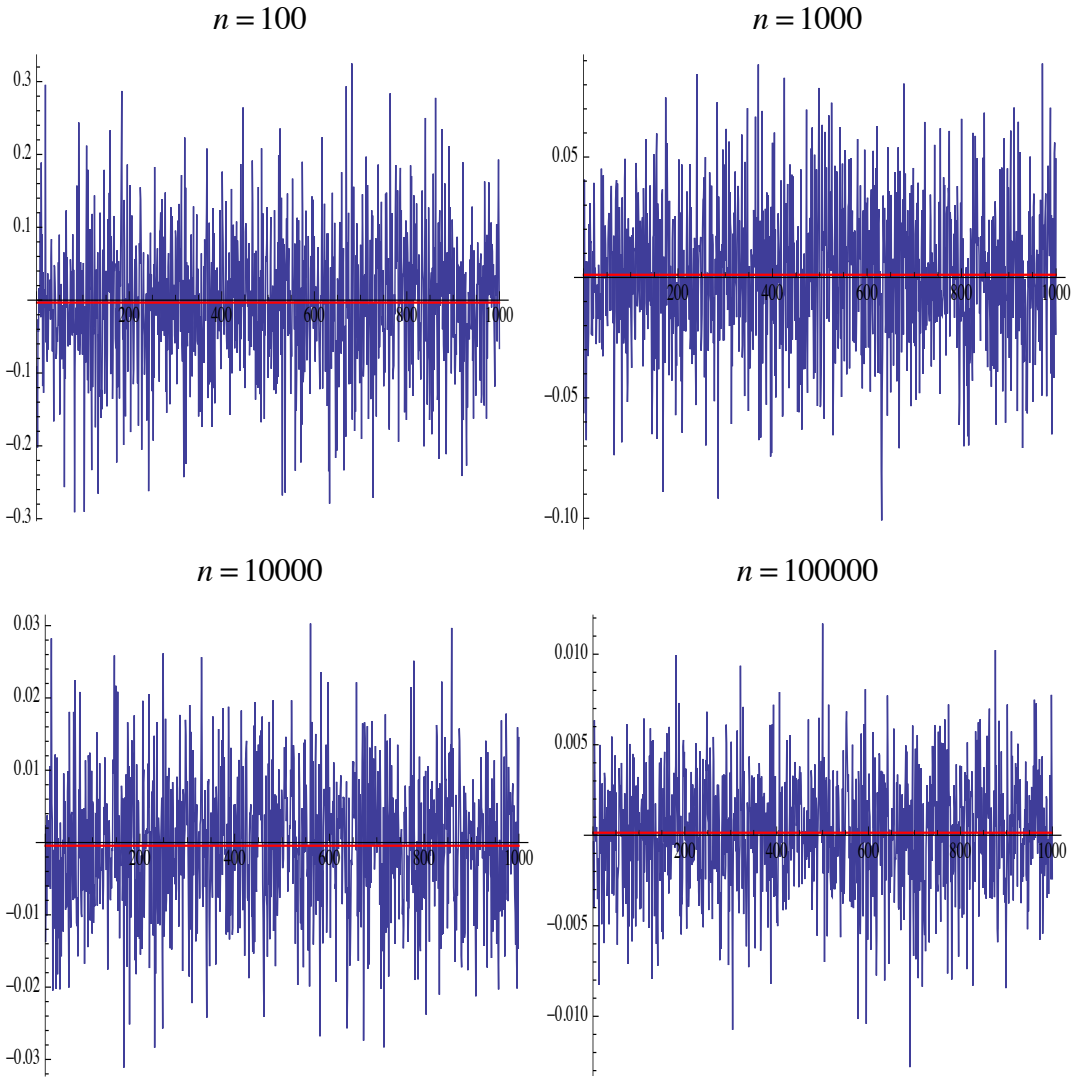
Odvodíme odhad pro data $\boldsymbol{X}$ s normálním rozdělením pro různé *n*. Porovnáme data, která jsme vygenerovaly s pomocí předem zadaného parametru $\theta=0$ a odhadu parametru θ s pomocí aplikace Metody maximální věrohodnosti.

Na obrázku můžeme vidět platnost předpokladu. S rostoucím *n* značení, které byla dosáhnuta při využití konsistentního odhadu, se velmi blíží k základním značením Normálního rozdělení $N(0,1)$.



Obrázek 3. Odhad parametru MMV pro Normální rozdělení $N(\theta,1)$. Modrá přímka je základní značení pro data s rozdělením $N(0,1)$, světle modrá značení jsou data s rozdělením $N(\theta,1)$. Parametr θ je předem odhadnutý parametr MMV.

Normální rozdělení je jedno z nejlepších ukázkových rozdělení pro platnost metody maximální věrohodnosti. Při generaci více odhadů parametru θ s rostoucím n dostáváme lepší představu a přesnější, že odhadnutý parametr θ je konsistentní odhad a střední značení $n\theta$ parametru konverguje k 0 , čímž se ukazuje platnost konvergence podle distribuce k základnímu rozdělení $N(0,1)$.



Obrázek 4. Odhady parametru MMV pro Normální rozdělení $N(\theta,1)$. Červená přímá představuje střední značení odhadnutých parametru θ pro různé n. Modré skoky jsou odhady parametru θ s rozdělením $N(\theta,1)$ s využitím metody maximální věrohodnosti.

Z porovnaní dat je jasně vidět, že s rostoucím *n* značení odhadu budou maximálně přiblížené k našemu předpokladu. Pro malé značení *n* a i z obrázku 3 a obrázku 4 ukazují rozsáhlé skoky mezi odhady, ale stále kolísají kolem odhadnutého parametru θ.

3.2.2. Exponenciální rozděleni.

Exponenciální rozdělení $\boldsymbol{Ex(\theta)}$ se sdruženou [4] hustotou

$$f(\boldsymbol{x},\theta)=\prod_{i=1}^{n}\frac{1}{\theta}e^{\frac{x_i}{\theta}}=\left(\frac{1}{\theta}\right)^n\exp\left[\frac{\sum_{i=1}^{n}x_i}{\theta}\right]\ \text{pro}\ x>0,\theta>0$$

které má střední hodnotu rovnou θ a rozptyl θ^2 s $J(\theta)=\frac{1}{\theta^2}$.

Aplikujeme MMV na exponenciální rozdělení, dostaneme

$$ln\,f(\boldsymbol{x},\theta)=n\ln\frac{1}{\theta}+\frac{\sum_{i=1}^{n}(x_i)}{\theta},$$

$$\frac{\partial ln\,f(\boldsymbol{x},\theta)}{\partial\theta}=\frac{n}{\theta}-\frac{\sum_{i=1}^{n}(x_i)}{\theta^2}=0,$$

má věrohodností rovnice

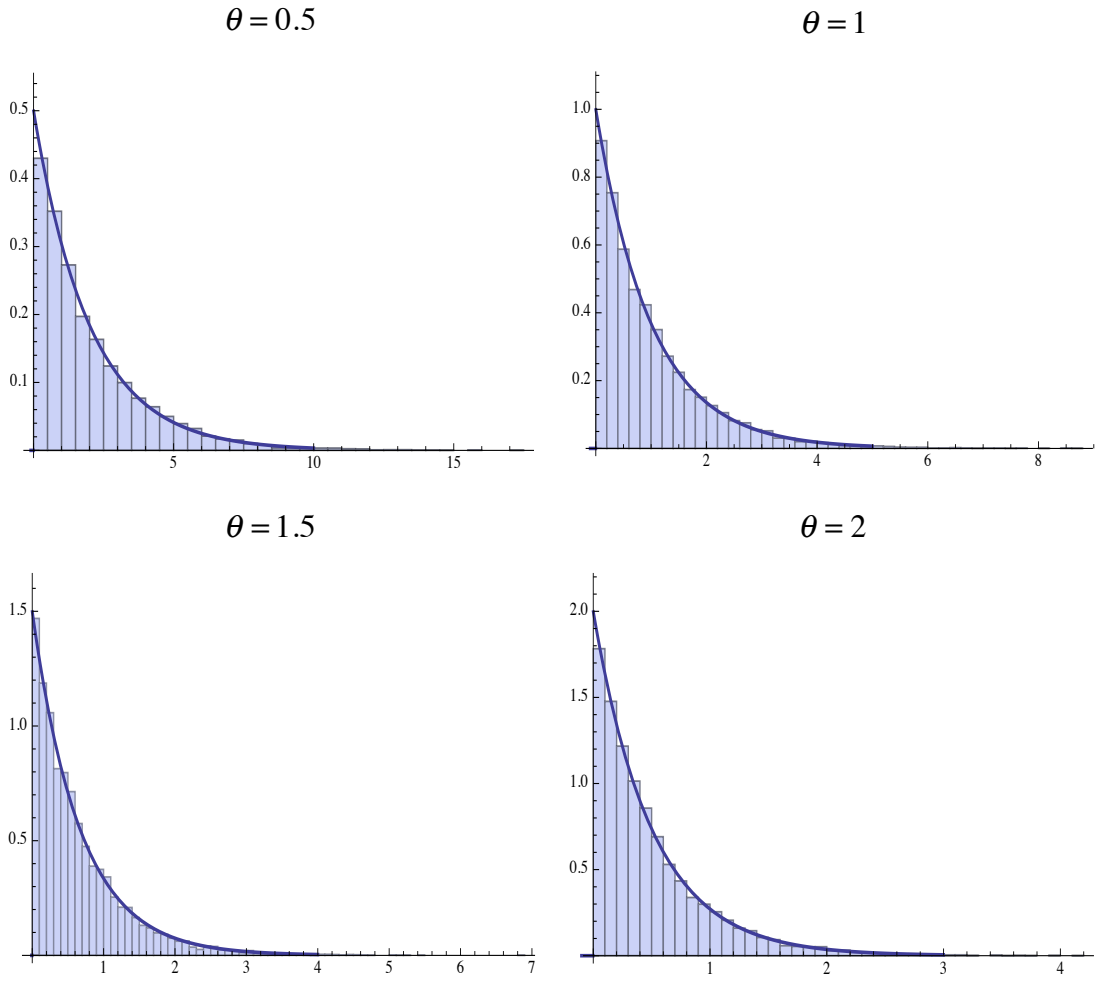
$$\theta-\bar{X}=0$$

a má kořen $\hat{\theta}_n=\bar{X}$. Pro $\bar{X}$ platí ze $\mathrm{E}X(\bar{X})=\theta$, $\mathrm{var}(\bar{X})=\frac{\theta^2}{n}$ proto $\hat{\theta}_n$ je konsistentním odhadem parametru θ a z věty 8 vyplývá $\sqrt{n}(\bar{X}-\theta)\sim N(0,\theta^2)$.

Exponenciální rozdělení je jiným dobrým příkladem spojitého rozdělení, na které lze lehce aplikovat Metodu maximální věrohodnosti. Proto využijeme už naznačený

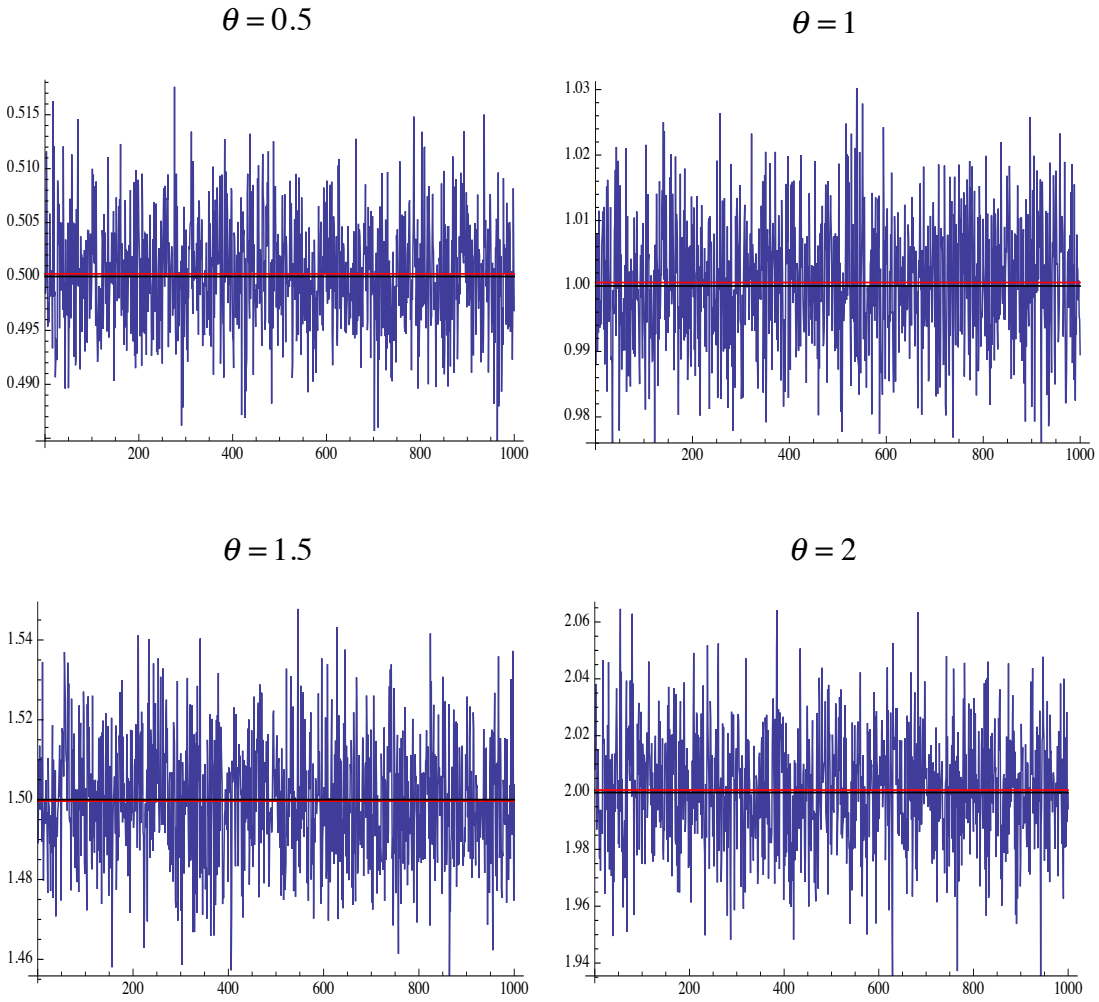
postup použití metody na data o rozsahu $n = 10000$, které budou mít exponenciální rozdělení a modelujeme odhady pro různé parametry $\theta = \{0.5, 1, 1.5, 2\}$.

Na obrázku 5 sledujeme data s už odhadnutým parametrem θ a v jakém vztahu se nachází data s rozdělením $\boldsymbol{Ex(\theta)}$, které má předpokládáný parametr, který jsme chtěly odhadnout. S rostoucím parametrem θ, se histogram začíná víc podobat přímce.



Obrázek 5. Odhad parametru MMV pro Exponenciální rozděleni $\boldsymbol{Ex(\theta)}$. Modrá přímka je základní značení dat s exponenciálním rozdělením, modrý histogram je značení dat s rozdělením $\boldsymbol{Ex(\theta)}$. Parametr θ je předem odhadnutý parametr pomocí MMV.

Následující obrázek 6 potvrzuje naše předpoklady o vlivu metody na odhady parametru. Roste parametr θ, skoky jednotlivých odhadů prestávaji být chaotickými a frekvence kolísaní kolem hledaného parametru začínají být napraveny k střední hodnotě, která se skoro rovná předpokládanému značení odhadu. Pro malé značení θ MMV taky ukazuje věrohodné výsledky a ještě jednou potvrzuje její správnost.



Obrázek 6. Odhady parametru MMV pro Exponenciální rozdělení $\boldsymbol{Ex(\theta)}$. Červená přímka představuje střední značení odhadnutých parametrů θ. Modré skoky jsou odhady parametru θ s rozdělením $\boldsymbol{Ex(\theta)}$ s využitím metody maximální věrohodnosti. Černá přímka je konstantní značení parametru, které chceme odhadnout.

Dosud jsme probíraly rozdělení se známou hustotou, na kterých se velmi snadno dá aplikovat metoda maximální věrohodnosti.

Ale co když se dostaneme do situace, kdy nebudeme vědět předem hustotu rozdělení nebo budeme muset pracovat s daty, která jsou definována rozdělením, co nemá přesný tvar pravděpodobnostní funkce. V takovém případě už nemůžeme využít přimo MMV, proto musíme najít jinou cestu.

Jiný způsob, jak bychom mohly vyřešit problém, je využití charakteristické funkce pro definovaní pravděpodobnostní funkce. Přiřadíme takové těžké případy do rozdělení III typu a ukážeme na nich možnost obcházení problému, přičemž dostaneme stejně dobré a věrohodné výsledky.

3.3. Rozdělení III typu.

3.3.1. Stabilní rozdělení

Stabilní rozdělení jsme vynechali jako samostatný typ, protože jedním z hlavních momentu je ten fakt, že některá rozdělení patřící k tomuto druhu mají nekoneční rozptyl, například Cauchy rozdělení, Pareto rozdělení pro parametr $0 < \alpha < 2$. Tento fakt nám dovoluje možnost použít odhadový algoritmus s chybami, které mají rozdělení s << těžkými chvosty >>.

Skupina stabilních rozdělení je docela široká, protože obsahuje hodně praktických realizovaných rozdělení. Přičemž neexistence pravděpodobnostní funkce u tohoto typu (výjimky Normální rozdělení, Caushy rozdělení, Levi rozdělení) nám dovoluje pracovat se všemi možnými podtypy tohoto druhu rozdělení bez ohledu na apriorní výběr nejlepší analytické představy pravděpodobnostní funkce.

Defenice 8. *Náhodná veličina X má stabilní rozdělení tehdy a jen tehdy, pokud pro všechny* $n > 1$, *existují konstanty* $c_n > 0$ *a* $d_n \in \mathbb{R}$ *takové, že*

$$X_1 + \cdots + X_n \overset{d}{=} c_n X + d_n \tag{3.1}$$

kde $X_1, \ldots, X_n$ *jsou nezávisle náhodné veličiny.*

Všechna stabilní rozdělení [6] mohou být definována pomocí charakteristické funkce, která jak víme z definice 5, jednoznačné určuje distribuční funkce a obsahuje plnou informaci o rozdělení, které jsme používali.

Defenice 9. *Stabilní rozdělení* $S(\alpha, \sigma, \beta, \mu)$, *resp.* $S_\alpha(\sigma, \beta, \mu)$ *jejíž distribuční funkce* $\Phi(x)$ *lze definovat pomocí charakteristické funkce*

$$\ln \varphi(t) = \begin{cases} i\mu t - \sigma^\alpha |t|^\alpha \left(1 + i\beta \, sign(t) tg \frac{\pi\alpha}{2}\right), \alpha \neq 1 \\ i\mu t - \sigma |t| \left(1 + i\beta \, sign(t) \frac{2}{\pi} \ln |t|\right), \quad \alpha = 1 \end{cases} \tag{3.2}$$

kde $0 < \alpha \leq 2, -1 \leq \beta \leq 1, \sigma > 0, -\infty < \mu < \infty$ *jsou neznáme parametry.*

Každé stabilní rozdělení [8] je jednoznačně reprezentováno pomocí jeho parametru:

1. Parametr α, pro něhož platí $0 < \alpha \leq 2$, budeme nazývat charakteristickým ukazatelem stabilního rozdělení $S_\alpha(\sigma, \beta, \mu)$.
2. Parametr β, platí $-1 \leq \beta \leq 1$ je šikmost nebo parametr asymetričnosti.
3. Parametr σ je parametrem variance, resp. rozptyl.
4. Parametr μ, definovaný na intervalu $-\infty < \mu < \infty$, představuje sebou střední hodnotu.

Parametr α odpovídá jaký bude pohyb chvostu rozdělení. Pokud $0 < \alpha < 2$, to

$$\lim_{x \to \infty} x^\alpha P(X > x) = C_\alpha \frac{1+\beta}{2} \sigma^\alpha \tag{3.3}$$

$$\lim_{x \to \infty} x^\alpha P(X < -x) = C_\alpha \frac{1-\beta}{2} \sigma^\alpha \tag{3.4}$$

kde

$$C_\alpha = \left(\int_0^\infty x^{-\alpha} \sin x \, dx \right)^{-1} = \begin{cases} \dfrac{1-\alpha}{\Gamma(2-\alpha)\cos\dfrac{\pi\alpha}{2}}, & \alpha \neq 1, \\ \dfrac{2}{\pi}, & \alpha = 1. \end{cases} \tag{3.5}$$

Pro $\alpha = 2$ z (3.2) dostaneme

$$\varphi(\alpha) = e^{i\mu\alpha - \frac{\alpha^2}{2}(2\sigma^2)}, \tag{3.6}$$

kde $\varphi(\alpha)$ charakteristická funkce Normálního rozděleni $N(\mu, 2\sigma^2)$.

Věta 13 *Pro libovolnou stabilní náhodnou veličenu X existuje takové číslo* $\alpha \in (0,2]$ *dvě nezávislé veličiny* X_1, X_2 *a konstanty A, B, D, C pro něž platí*

$$AX_1 + BX_2 = CX + D \tag{3.7}$$

kde $C^\alpha = A^\alpha + B^\alpha$.

Věta 14 *Pro libovolnou náhodnou stabilní veličinu X existuje takové číslo* $\alpha \in (0,2]$ *pro které platí, že* c_n *z definice 8 se rovná* $n^{1/\alpha}$ [8].

Podle parametru stability α lze posoudit jaké bude stabilní rozdělení a kde ho jde použít.

Příklad 2.

1) $\alpha < 1$, absolutní stabilní rozdělení, podle předchozí věty dostáváme

- $\alpha = 0.1$, $n = 3$

$$X_1 + X_2 + X_3 \overset{d}{=} 3^{1/0.1} X = 3^{10} X = 59049X$$

- $\alpha = 0.1$, $n = 10$

$$X_1 + X_2 + \ldots + X_{10} \overset{d}{=} 10^{1/0.1} X = 10^{10} X = 10000000000X$$

- $\alpha = 0.01$, $n = 3$

$$X_1 + X_2 + X_3 \overset{d}{=} 3^{1/0.01} X = 3^{100} X (= 5.14 * 10^{47} X)$$

Z posledního přikladu můžeme jasně vidět, že takovou závislost pro $\alpha \in (0,1)$ mezi veličinami můžeme aplikovat spíše ve fyzice, než ve financích.

2) $\alpha \in (1,2]$, absolutní stabilní rozdělení, podle předchozí věty dostáváme

- $\alpha = 1.5$, $n = 3$

$$X_1 + X_2 + X_3 \overset{d}{=} 3^{1/1.5} X = 2.0800838231X$$

- $\alpha = 2$, $n = 10$

$$X_1 + X_2 + \ldots + X_{10} \overset{d}{=} 10^{1/2} X = 3.1622776602 X$$

Čím blíž parametr α k značení 2, tím bude Stabilní rozdělení velmi podobné Normálnímu rozdělen, a toto značení je využitelné více ve financích.

Existují jenom tři dobře známá rozdělení [9] patřicí ke stabilním rozdělením:

Normální Rozdělení

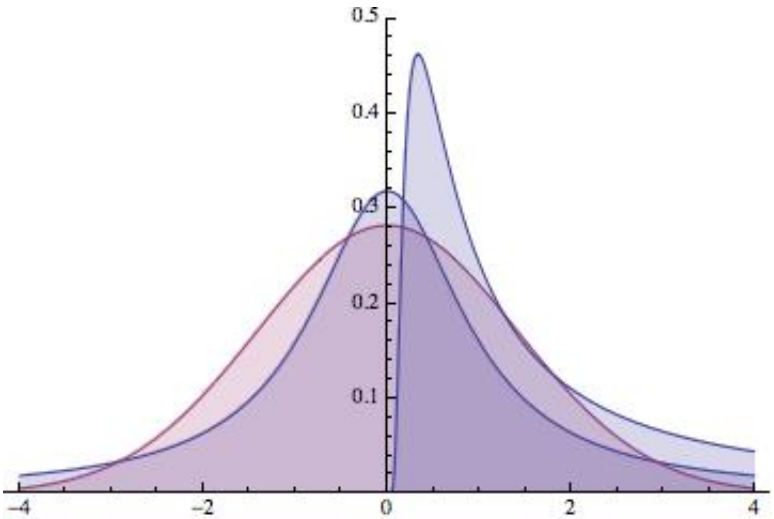
$\alpha = 2, \quad X \sim S_2(\sigma, 0, \mu)$

Cauchy Rozdělení

$\alpha = 1, \quad X \sim S_1(\sigma, 0, \mu)$

Levy Rozdělení

$\alpha = 0.5, \quad X \sim S_{0.5}(\sigma, 1, \mu)$



Obrázek 7. Příklad rozdělení $N(0,1)$, $Cauchy(1,0)$, $Levy(1,0)$ reprezentovaný přes Stabilní rozdělení pro parametry $\alpha = \{0.5, 1, 2\}$ a $\beta = \{0, 1\}$.

Jenom pro tato 3 rozdělení by mělo smysl aplikovat MMV přimo. Akorát obecný tvar hustoty stabilní náhodné veličiny s libovolnými parametry není známý a tím komplikuje proces provedení odhadu parametru.

Budeme počítat [10] , že všechny stabilní rozdělení jsou spojitá a mají derivace všech řádů. Protože nemůže použít přímo metodu maximální věrohodnosti, budeme aplikovat jinou variantu odhadu parametru.

Podíváme se na regresní rovnice [6] typu

$$y = Z\theta + \varepsilon \tag{3.8}$$

kde $Z = \begin{bmatrix} f_1(z_{11}) & \cdots & f_p(z_{1p}) \\ \vdots & \ddots & \vdots \\ f_1(z_{N1}) & \cdots & f_p(z_{Np}) \end{bmatrix}$ matice značení regresních funkci, $\theta = (\theta_1,\ldots,\theta_p)^T$ vektor neznámých parametru, které chceme odhadnout, p počet neznámých parametrů, N počet prováděných pokusů, $f_i(z)$ známe funkci a z_{ij} vstupní data. Vektor $y = (y_1,\ldots,y_N)^T$ udává výsledek a vektor $\varepsilon = (\varepsilon_1,\ldots,\varepsilon_N)^T$ je bílý šum. Přičemž budeme předpokládat, že bílý šum představuje navzájem nezávisle stejné rozdělení náhodných velečin s unimodální hustotou $\psi(x)$, patřící do typu stabilních rozdělení, pro která platí

$$\mathrm{E}(\varepsilon_i) = 0, \quad D(\varepsilon_i) = \sigma^2.$$

Hlavní myšlenka spočívá v tom, že s použitím vstupních dat budeme odhadovat vektor s neznámými parametry z regresní rovnice. Pro indikaci charakteristická funkce Stabilního rozdělení použijeme empiricko-charakteristickou funkci. Z realizace $x_1,\ldots,x_N$ a náhodných veličin lze najít empirické značení odhadů charakteristické funkce

$$\hat{\varphi}(t) = \frac{1}{N}\sum_{j=1}^{N} e^{itx_j} = \frac{1}{N}\sum_{j=1}^{N}(\cos(tx_j) + i\sin(tx_j)). \tag{3.9}$$

Podle zákona Velkých čísel empirické značení odhadů bude konsistentním. Empirické značení odhadů charakteristické funkce nám dovolují odříznout ty části dat, které zhoršují výsledek odhadu. Rozdíl charakteristické funkce Stabilního rozdělení a empirických značeni ukazují, jak silné jsou projevy odchylky v datech od normality a jaký vliv mají na vstupní data. Z obrázku 8 vidíme, že čím menší odhadový parametr, α tím je rozdíl vetší, protože chyby v datech mají silnější vliv na obraz grafu, na rychlost výpočtu a i na přesnost odhadu. Jenomže se zvyšováním předpokládaný pro parametr α se rozdíl skoro rovná nule a grafy představují obraz normálního rozdělení.

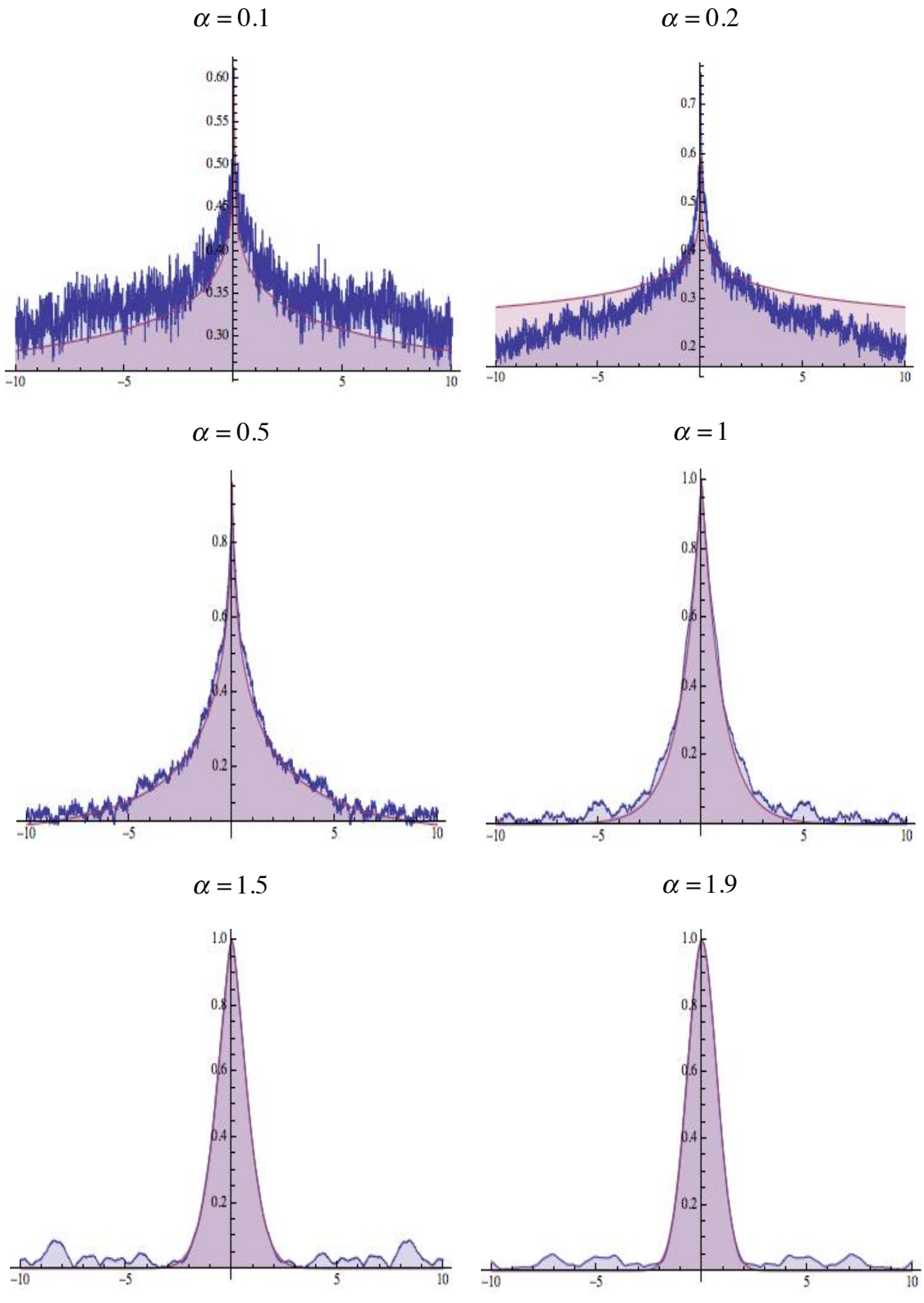
Při odhadu parametru α bude platit

$$\min_i \sum_{i=1}^{n} (|\hat{\varphi} - \varphi|)^2 \tag{3.10}$$

kde $\varphi = e^{-|t|^\alpha}$ a $\hat{\varphi} = \frac{1}{N}\sum_{j=1}^{N} \cos(tx_j)$. Výsledek rovnice (3.10) i bude hledaným odhadem parametru $\hat{\alpha}$. Přičemž proces minimalizace rozdílu empirické funkce i charakteristické funkce provádíme jako v přímé MMV, derivace podle neznámého parametru, což se bude rovnat nule. Tím samým najdeme extrémy funkce.

Po provedení několika experimentů, porovnaní efektivity nalezení neznámých parametrů přes charakteristickou funkci, víc tabulka 1. Došli jsme k výsledku a s ohledem na časovou náročnost a přesnost pro $\alpha \in (0,1)$ bylo by lepší používat tento postup, tedy nepřímou aplikaci Metody maximální věrohodnosti. Jakmile platí $\alpha \in (1,2)$ přijdeme na přímou aplikace MMV s využitím různých softwarů pro výpočet, pro nás to je Mathematica.

.



Obrázek 8. Porovnání značení charakteristické funkce pro stabilní rozdělení s empirickým značením pro různé parametry $\alpha = \{0.1, 0.2, 0.5, 1, 1.5, 1.9\}$

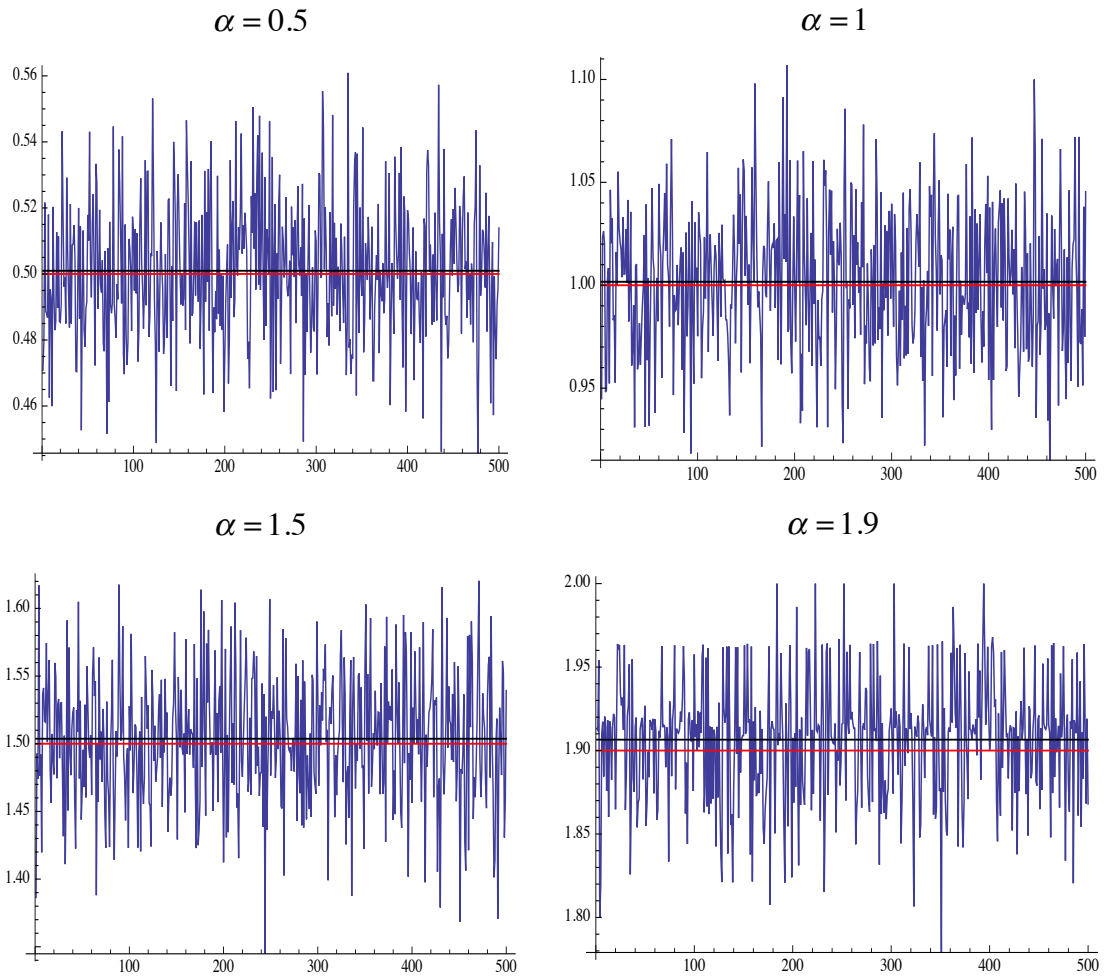
.

S pomocí programu Mathematica nasimulujeme soubor dat $\boldsymbol{X}$ o rozsahu n, které má Stabilní rozděleni pro určité parametry $\alpha, \beta, \mu, \sigma$, z nich se pak pokusíme odhadnout α pomocí MMV tam, kde to bude možné. Značení, kterých jsme dosáhli budou skoro stejná, víc Tabulka 1., Kdyby jsme nepočítali přes funkci, která už excituje v tomto softwaru, ale přesně přes ručně sestavovaný postup s využitím charakteristické funkce a empirických značení, která jsme popsali výše.

Z obrázku 9 jde jasně vidět velký rozdíl mezi odhady pro $\alpha = 0.5$ a pro $\alpha = 1.9$. Pro vetší α jsou odhadnuta značení, která už nemají rozsáhlé skoky mezi předchozím a následujícím nalezeným odhadem, blíží se k tvaru Normálního rozdělení a odchylky jsou minimální. Jenomže odhady s malým značením α nejsou pro velké *n* a už ted' bylo docela těžké jich dosáhnout, proto pro parametr $\alpha = 0.5$ jsme museli použít ne nabízenou funkce, ale postup s využitím charakteristické funkce. Očividné, [8] že pro $\alpha \in (0,1)$ z (3.3) a (3.4) chvosty těžší než pro $\alpha = 2$ ve smyslu rychlosti pochybu.

Pro $\alpha \in (0,1)$ časová náročnost spravování dat byla o dost vetší než pro $\alpha \in [1,2]$, bez ohledu na to, že pro interval $(0,1)$ jsme museli přejít na charakteristickou funkci, čím jsme se pokusili zlepšit čas a pravdivost odhadu. A to je ten problém. Proto když pracujeme se stabilními rozděleními, musíme někdy vybrat jaký postup použijeme pro zpracovaní dat, protože někdy v praxi nám bude záležet na času a i na těch malých rozdílech výsledku, které nám budou udávat ta nebo jiná zvolená metoda.

Pro výchozí asymptotické [8] výsledky z (3.3) a (3.4) má smysl připomenout rozdělení Pareto. Porovnaní těchto dvou rozdělení podle (3.3) a (3.4) udává, že na nekonečném intervalu Stabilní rozdělení se podobají rozdělení Pareto. V tomto smyslu «chvosty» kusů Stabilních rozdělení patří k paretovskému typu.



Obrázek 9. Odhady parametru MMV pro Stabilní rozděleni $S_\alpha(\sigma,\beta,\mu)$. Černá přímka představuje střední značení odhadnutých parametru α. Modré skoky jsou odhady parametru α s rozdělením $S_\alpha(\sigma,\beta,\mu)$ s využitím metody maximální věrohodnosti. Červená přímka je předpokládané značení parametru α.

Při generaci těchto dat jsme využili konkrétní zadané parametry $\alpha, \beta, \mu, \sigma$ a odhadovali jenom jeden parametr, pro nás nejvážnější parametr α. Ale když máme úplně náhodná data o hodně většího rozsahu, o kterých můžeme jenom předpokládat, že mají Stabilní rozděleni? Odpovědět na tuto otázku se pokusíme ve 4. části této práce.

α	$\hat{\alpha}_{Ch.f}$	$\hat{\alpha}_{MMV}$
0.1	0.102879	0.101529
0.2	0.197704	0.21303
0.5	0.502631	0.491368
1	0.984024	1.03301
1.5	1.4741	1.5007
1.9	1.83174	1.91795

Tabulka 1. Dosáhnuto značení odhadů pro různé parametry α s pomocí MMV $\hat{\alpha}_{MMV}$ a charakteristickou funkcí $\hat{\alpha}_{Ch.f}$

3.3.2. Paretovo rozdeleni

Paretovo rozdělení [8] je rozdělení s pravděpodobnostní funkcí a distribuční funkcí

$$f(x)=\frac{\alpha b^{\alpha}}{x^{\alpha+1}},\quad F(x)=1-\left(\frac{b}{x}\right)^{\alpha} \tag{3.11}$$

které je definováno pro všechny $x \geq b$ a patří do typu stabilních rozdělení. Pro parametr α platí, ze stačí pokud $\alpha > 0$. Jenom pokud $\alpha < 1$ Paretovo rozdělení má nekonečnou střední hodnotu a rozptyl. Pokud $1 < \alpha < 2$ pro toto rozdělení bude existovat střední hodnota $\frac{\alpha b}{\alpha - 1}$, ale rozptyl se stále bude rovnat nekonečnu. Jak už bylo řečeno výše, Stabilní rozdělení a rozdělení Pareto mají podobné «chvosty». Pokud $\alpha > 2$ pro hodnoty s Paretovským rozdělením bude existovat druhý moment

$$\mu_2' = \frac{\alpha b^2}{\alpha - 2}.$$

Při odhadech parametru Paretova rozdělení nemůžeme použít charakteristickou funkci, protože nemáme explicitní formu této charakteristické funkce. Ale to ne znamená, že nemůžeme použít postup odhadu parametru přes Metodu maximální věrohodnosti.

Rozděleni má sdruženou hustotu

$$f(x,\alpha,b)=\prod_{i=1}^{n}\frac{\alpha b^{\alpha}}{x_i^{\alpha+1}}=\alpha^n b^{\alpha n}\prod_{i=1}^{n}x_i^{-(\alpha+1)}.$$

Aplikujeme MMV na Paretovo rozdělení, dostaneme

$$\ln f(x,\alpha,b)=n\ln\alpha+\alpha n\ln b-(\alpha+1)\sum_{i=1}^{n}\ln x_i$$

a pak následující krok derivace podle parametru α a sestava věrohodností rovnice

$$\frac{\partial ln\ f(x,\alpha,b)}{\partial \alpha} = \frac{n}{\alpha} + n\ln b - \sum_{i=1}^{n} \ln x_i = 0$$

$$\frac{\partial ln\ f(x,\alpha,b)}{\partial b} = \frac{\alpha n}{b} = 0$$

Kořenem a zároveň odhadem věrohodností funkce je

$$\hat{\alpha} = \frac{n}{\sum_{i=1}^{n} \left(\ln x_i - \ln \hat{b} \right)}$$

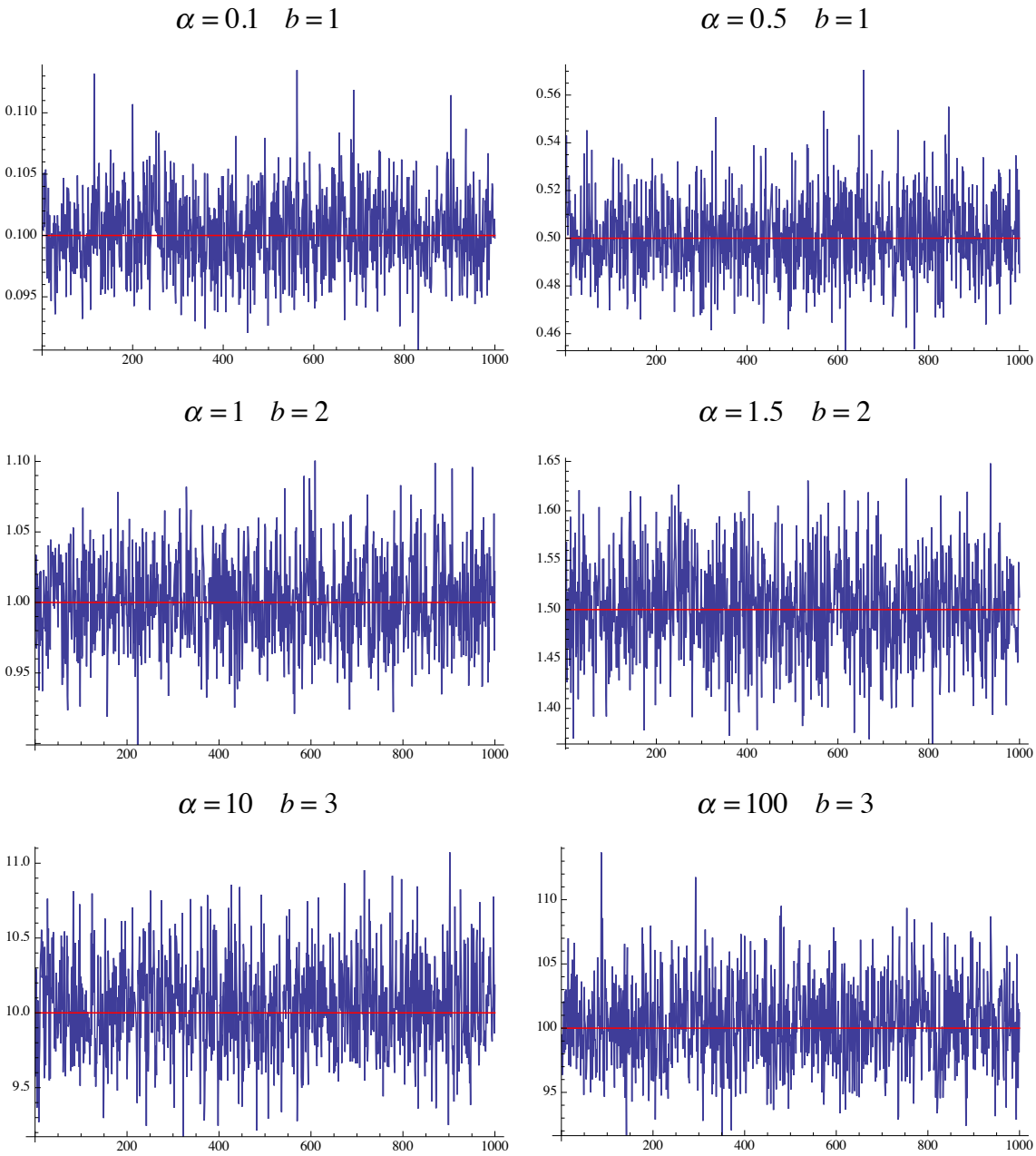
Jak vidíme, funkce $\ln f(x,\alpha,b)$ je monotónní rostoucí pro parametr b. Největší značení b je největším značením věrohodností funkce. Z tohoto důvodu pro $x \geq b$ bude platit

$$\hat{b} = \min_i x_i$$

Fišerova míra informací v případě Paretova rozdělení se bude rovnat

$$J_n(p) = \begin{pmatrix} \frac{\alpha}{b^2} & -\frac{1}{b} \\ -\frac{1}{b} & \frac{1}{\alpha^2} \end{pmatrix}$$

Po provedení několika odhadů pro hodnoty, které jsme vygenerovaly pro konkrétní parametry o rozsahu *n,* můžeme říct, že proces odhadu Metodou Maximální věrohodnosti byl docela snadný a odhady, které jsme dostali odpovídají našim předpokladům. Ale když se podíváme na střední hodnoty a rozptyl dat z Paretova rozdělení pro určité intervaly, výsledky budou docela nepříjemné a budou se skoro blížit k nekonečným značením, čímž se potvrzuje ten fakt, že Paretovo rozdělení je správně přeřazeno k typu Stabilních rozdělení s nekonečnými momenty.



Obrázek 10. Odhady parametru MMV pro Paretovo rozděleni. Černa přímka představuje střední značeni odhadnutých parametru α pro konkrétní parametr *b*. Modré skoky jsou *n* odhadu parametru α a *b* s pomoci metody maximální věrohodnosti. Červena přímka předpokládané značeni parametru α.

4. Aplikace Metody Maximální věrohodnosti v praxi

V předchozí části naší práce jsme hodně mluvili o aplikaci Metody maximální věrohodnosti na různých rozděleních Ted' se pokusíme to, co jsme zatím měli za předpoklad, převést do praxe a aproximovat na reálných datech.

Když dostaneme skutečná data, komodity pro konkrétní produkt, budeme chtít vědět základní informaci o těchto datech, například závisí-li cena v čase *t* na ceně v čase *t-1* nebo ne, jaké rozdělení mají a jaké jsou parametry. Bude-li existovat druhý, třetí a čtvrtý moment, jsou-li data normální či ne a tak dále.

Při odhadech a testovaní dat se budeme se držet následujících kroků

1. Použijeme model AR(q) s vhodným parametrem q z dané časové řady $\{X_t\}$ pro nalezení parametru z regresní rovnice.
2. S použím těchto parametrů najdeme rezidua.
3. S pomoci vhodný metody odhadneme rozděleni rezidua a příslušný danému rozděleni parametry.
4. Porovnání empiricky distribuční funkce a distribuční funkce rozdělení, vybereme nejlepší rozdělení pro naše data.
5. Ověří se vhodnost zvoleného modelu různými testy.

Popíšeme následující kroky podrobněji.

Pracujeme s časovou řadou komodit a cen na elektřinu v Německu, European Energy Exchange AG. Jsou to různá pozorování v různých časových intervalech. Pro modelování časových řad je nejvhodnější použití modelu AR(q).

Krok 1. AR(q) proces.

Definujeme obecný stochastický proces [11], který na vstupu má bílý šum a jinak řečeno vystupuje jako lineární proces

$$\tilde{z}_t = a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + ... = a_t + \sum_{i=1}^{\infty} \psi_i a_{t-i}, \tag{4.1}$$

kde $\tilde{z}_t = z_t - \mu$ je odchylka procesu od počátečního stavu nebo, pokud je stacionární, od svého středního značení. Obecný lineární proces (4.1) nám dovoluje považovat $\tilde{z}_t$ za vážený součet přítomnosti a budoucnosti a_t. Bílý šum je posloupnost impulsu, který působí na celý systém. Tvoří ho posloupnosti nekorelovaných náhodných velečin s nulovým středním značením a stejným rozptylem

$$E[a_t] = 0, \quad \text{var}[a_t] = \sigma_a^2.$$

Protože náhodné vyléčení nekorelované jejich autokovarianční funkce musí mít tvar

$$\gamma_k = E[a_t a_{t+k}] = \begin{cases} \sigma_a^2, & k = 0, \\ 0, & k \neq 0. \end{cases} \tag{4.2}$$

Autokorelační funkce bílého šumu bude mít velmi jednoduchou formu

$$\rho_k = \begin{cases} 1, & k = 0, \\ 0, & k \neq 0. \end{cases} \tag{4.3}$$

Model (4.1) můžeme zapsat jinak, přesně řečeno, jako vážený součet minulých značení $\tilde{z}_t$ plus dodatečný impuls a_t

$$\tilde{z}_t = \pi_1 \tilde{z}_{t-1} + \pi_2 \tilde{z}_{t-2} + ... + a_t = \sum_{i=1}^{\infty} \pi_i \tilde{z}_{t-i} + a_t \tag{4.4}$$

Podíváme se na specifický příklad (4.4), který tvoří autoregresní proces, kde jenom první q je nenulový.

Tento model můžeme psát jako

$$\tilde{z}_t = \phi_1 \tilde{z}_{t-1} + \phi_2 \tilde{z}_{t-2} + ... + \phi_q \tilde{z}_{t-q} + a_t \tag{4.5}$$

kde $\phi_1, \phi_2, \ldots, \phi_q$ je konečný počet vážených parametrů. Proces (4.5) nazýváme procesem autoregresivního q řady nebo *AR(q)*.

Pro nás velké značeni budou hrát procesy *AR(1)* a *AR(2),* autoregresní $q = 1$ řadu a $q = 2$ řadu

$$\tilde{z}_t = \phi_1 \tilde{z}_{t-1} + a_t,$$

$$\tilde{z}_t = \phi_1 \tilde{z}_{t-1} + \phi_2 \tilde{z}_{t-2} + a_t.$$

Ted' rovnice (4.5) bude mít ekvivalentní zápis

$$\left(1 - \phi_1 B - \phi_2 B^2 - \ldots - \phi_q B^q\right) \tilde{z}_t = a_t$$

nebo

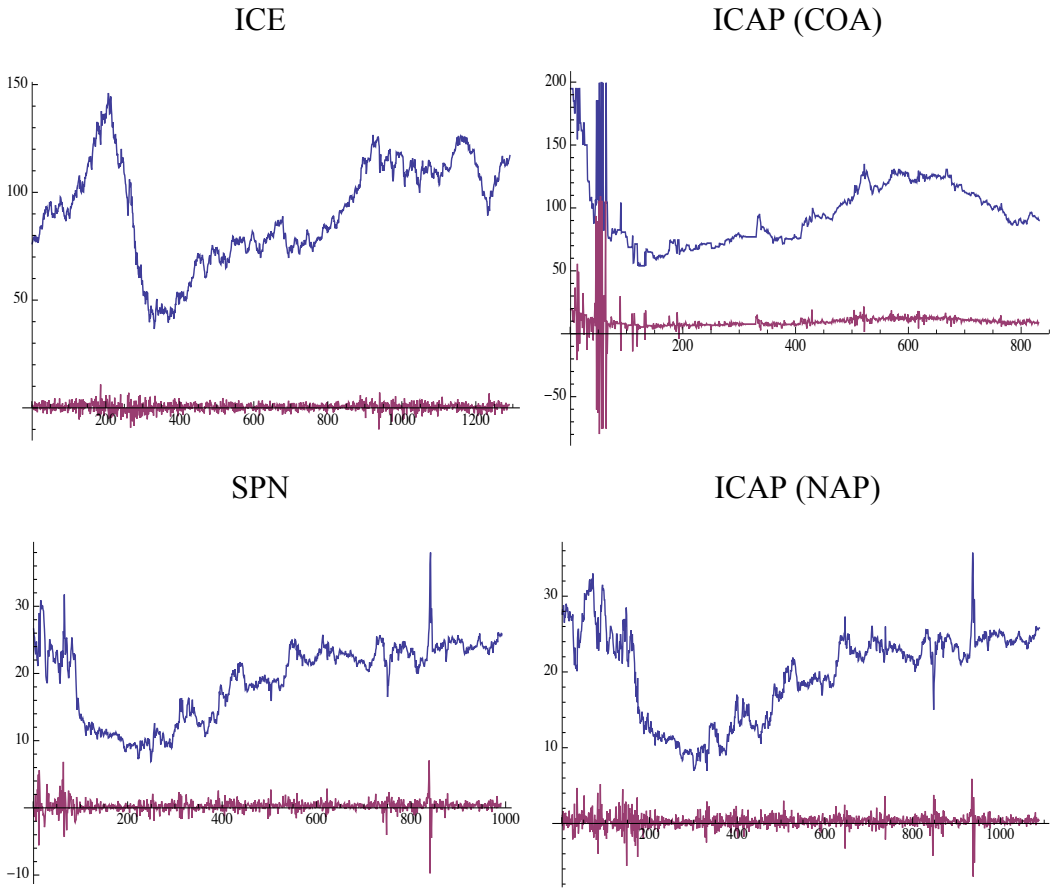
$$\phi(B) \tilde{z}_t = a_t \tag{4.6}$$

Víc tento proces a jeho vlastnosti jsou popsané v knize [11].

Při aplikace tohoto modelu na vstupní data jsme provedli několik experimentů pro různé parametry $q = \{1,2\}$. Z výsledku jsme nechali pro pokračovaní v následujících krocích parametry q, jako nejvhodnější parametr pro každou jednotlivou časovou řadu.

Krok 2. Rezidua.

Po použití procesu *AR(q)* jsme dosáhli residuálních veličin. Pokud se už teď podíváme na grafy počátečních velečin a residua, uvidíme, že mají podobny tvar a jejich histogramy jsou vzdálené, ale schází s normálním rozdělením, víc příloha 1. Na obrázku 11 jsme vykreslilí výsledky pro různé časové řady, ceny konkrétních podniků. Pohyby cen elektřiny, na kterých jsme prováděli experimenty najdete v příloze 1.



Obrázek 11. Ceny komodit a jejich residua ICE, ICAP(COA), SPN, ICAP(NAP).

Krok 3. Odhad rozdělení

V následujícím kroku aplikujeme MMV pomocí funkce `FindDistributionParametrs[]` pro odhad rozdělení a příslušných parametru. Z nabízených rozdělení vybereme nejlepší výsledek, ten, který by nejvíce odpovídal realitě a vyhovoval by nám. S pomocí zvoleného rozdělení provedeme ještě pár testů. Pravé teď nás bude zajímat ne jaké dostaneme rozdělení, ale přesnost odhadů parametrů a sám test který byl použít. V tomto bodě budeme porovnávat několik metod, které nám nabízí Mathematica :

- "*MaximumLikelihood*" metoda maximální věrohodnosti.
- "*MethodOfMoments*" momentová metoda.
- "*MethodOfCentralMoments"* metoda centrálních momentů .

Provádíme tento test s využitím Stabilního rozdělení. Protože přesně toto rozdělení ukáže i na malý rozdíl v odhadech parametru.

Parametr	MMV	MoM	MCM
α	1.53514	2	2
β	0.16194	0	0
μ	5.46532	5.36409	5.36409
σ	2.87073	5.10311	5.10311

Tabulka 1. Odhady parametru pomocí MMV, MoM, MCM pro ceny na elektřiny ze Stabilního rozdělení a s použitím AR(2) procesu.

V porovnaní s ostatními metodami maximální věrohodnosti odhad byl časově náročnější, ale přesnější než ostatní. Z toho důvodu že nás zajímá přesnost samozřejmě použijeme výsledek MMV.

Krok 4. Wassersteinova metrika.

V tomto kroku jsme chtěli porovnat empiricky distribuční funkce a distribuční funkce rozdělení, jinak řečeno chceme najít Wassersteinovu metriku [12].

Defenice 10. Wassersteinova metrika pro jedno dimensionální rozdělení je vzdálenost mezi distribučními funkcemi na daném metrickém prostoru M a rovná se

$$W(F,G) = \int_{-\infty}^{\infty} |F(z) - G(z)| dz. \tag{4.7}$$

Kvůli tomu, že Empirická distribuční funkce konverguje skoro jistě k distribuční funkci F, pro každou realizaci ξ, $F_n(z)$ je distribuční funkcí.

Tento rozdíl nám ukáže, nakolik se liší odhadnuté rozdělení s odhadnutým parametrem od empirické distribuční funkce a jak moc se liší navzájem. Čím menší je ta vzdálenost, tím přesnější odhad jsme dostali. Z těchto výsledků už bychom měli definitivně vybrat vyhovující vstupní data pro naše rezidua rozdělení, které bychom mohli použít i dál pro jiné věci, které už se netýkají této práce. V nejlepším případě by bylo, pokud by ten rozdíl byl skoro nulový.

q	Stabilní rozdělení	Normální rozdělení	Rovnoměrné rozdělení	Exponenciální rozdělení
2	0.726854	2.07042	201.001	1.53405
2	0.140869	0.197219	3.78089	0.810974
1	1.62036	6.75487	43.3773	5.22519
2	0.117627	0.264285	3.78947	0.268843
1	0.0997523	0.218277	2.66813	0.297494

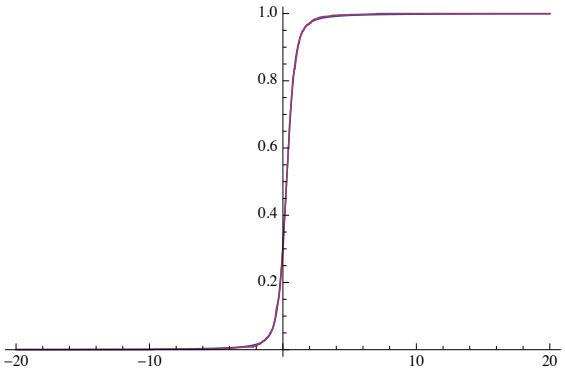
Tabulka 2. Wassersteinova metrika mezi Empirickou distribuční funkcí a distribučními funkcemi jiných rozdělení pro konkrétní *AR(q)* proces.

Z pozorovaných velečin vychází, že mají skoro Stabilní rozdělení a některé se chovají dokonce jako normální rozdělení, které je stejně ověřené odhadnutým stabilním rozdělením, které má příslušné parametry odpovídající Normálnímu rozdělení.

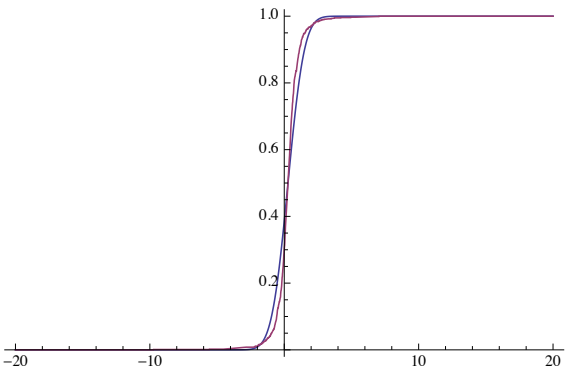
V této situaci je pro data bude lepší vybrat Stabilní rozdělení, protože je citlivěji než ostatní rozděleni. Navíc parametr α pro Stabilní rozdělení leží v intervalu $1 < \alpha \leq 2$, čímž ani nevyžaduje použití nepřímé aplikace MMV.

SPN

Stabilní rozdělení

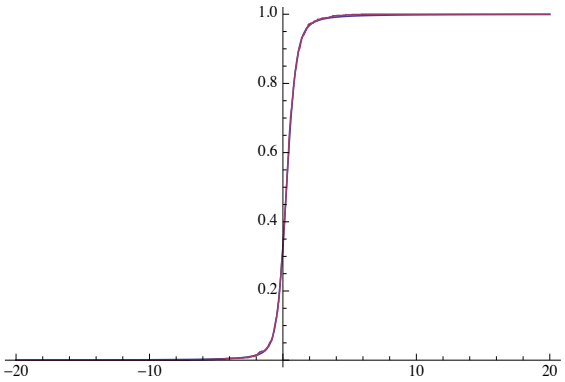


Normální rozdělení

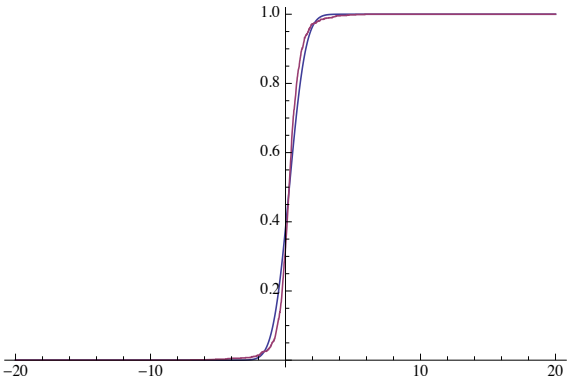


ICAP (NAT)

Stabilní rozdělení



Normální rozdělení



Obrázek 12. Wassersteinova metrika mezi Empirickou distribuční funkcí a distribučními funkcemi Stabilního a Normálního rozdělení pro SPN a ICAP (NAT).

Krok 5. Testování.

Protože samotná vizualizace nám nestačí, dostáváme se proto k poslednímu kroku, ověření vhodnosti zvoleného modelu různými testy. Pro tento krok použijeme testy dobré shody (`PearsonChiSquareTest[]`), který byl už popsán výše, test Kolmogorov-Smirnova (`KolmogorovSmirnovTest[]`) vic příloha 2 a (`DistributionFitTest[]`), který nám nabízí software Mathematica.

Vybereme 2 kandidáty na nejlepší odhad rozdělení. Pro všechna data jedním a zatím nejlepším je Stabilní rozděleni. Z provedených pokusů hodně testů bylo zamítnuto a i stabilní rozděleni i jiné druhé nejlepší rozdělení, ačkoli několik testů naopak ukázalo, že přesně Stabilní rozdělení je ten odhad, který potřebujeme.

Výsledek provedené práce je takový, že nejlepší metoda pro hledaní příslušného rozdělení s parametry pro vstupní data je Metoda Maximální věrohodnosti. Přestože je časová náročnost větší než u ostatních metod, výsledek a přesnost je o hodně lepší. Tato metoda je nejlépe aplikovatelná v praxi na finančních datech pro stabilní rozdělení. Protože přesně toto rozděleni spolu s MMV reaguje citlivěji na vstupní data nebo časovou řadu. A jak bylo řečeno, ze Stabilního rozdělení se lze dostat k jiným rozdělením a hlavně k Normálnímu rozdělení. Navíc pro všechny testy platilo, že stabilní rozdělení mělo odhadnutý parametr v intervalu $1 < \alpha \leq 2$ a jak už bylo řečeno pro finanční data parametr α automaticky vyloučil problémy, které by se mohli týkat aplikování MMV na Stabilní rozdělení.

test	Cena za elektřinu	ICE	ICAP (COA)	SPN	ICAP (NAT)
q	2	2	1	2	1
1	0	1	0	1	1
2	0	1	0	1	1
3	0	1	1	1	1
4	0	0	0	0	0
5	0	0	0	0	0
6	0	0	0	0	0

Tabulka 3. 1-3 test: I. Kandidát, ověřování Stabilního rozdělení pomocí PearsonChiSquareTest[], KolmogorovSmirnovTest[], DistributionFitTest[]. 4-6 test: II. Kandidát, ověřování druhého lepšího rozdělení dosud dosáhnutého pomocí PearsonChiSquareTest[], KolmogorovSmirnovTest[], DistributionFitTest[].

5. Závěr.

Cílem této bakalářské práce bylo posoudit možnosti odhadů parametrů pro různá rozdělení s pomocí konkrétní metody a to je metoda Maximální Věrohodnosti. Uvedli jsme plusy a mínusy tohoto modelu a naznačili oblasti, kde ho můžeme použít a to s ohledem na konkrétní podmínky. Poznali jsme, že MMV je velmi užitečná jak ve finanční oblasti, tak že se dá aplikovat v jiných modelech jako nepřímý náznak pro odhad maximální věrohodnosti.

V teoretické části byly nejprve uvedeny věty a definice, které byly spojeny s popisem odhadu parametru, jak ten odhad probíhá a jaký může být. Pak jsme popsali důležité pojmy, který se vyskytovaly v praktické i teoretické části a díky kterým jsme mohli definovat konkrétní postup pro aplikování MMV bud‘ přímo, nebo s využitím charakteristické funkce. Definované pojmy a vlastnosti Maximální Věrohodnosti nám dovolili pozorovat výsledky, které byly nalezeny v následující praktické části. Navíc test dobré shody, který byl stejně popsán v teoretické části, nám ještě pomohl posoudit výsledky experimentu.

V praktické i teoretické části byly předvedeny příklady diskrétních, spojitých a stabilních rozdělení a aproximací na nich Metoda Maximální věrohodnosti. Výsledky aproximace byly uvedeny v grafech i tabulkách ke každému příslušnému rozdělení. Nejvýznamnější aplikace MMV v těchto rozděleních je ten fakt, že existují rozdělení, na kterých nejde aplikovat tato přímo a proto jsme pro ně museli hledat jinou cestu, která nás přivedla k částečné myšlence MMV.

V praktické části byla ukázaná realizace MMV na reálných datech. Odhadovali jsme rozdělení a příslušné parametry vstupních dat. Po provedení experimentu jsme dosáhli takového výsledku, že velečiny měli stabilní rozdělení a někdy se přibližovali Normálnímu rozděleni, čímž jsme ověřili blízkost těchto dvou rozdělení a ověřili

definovaný fakt toho, že Normální rozdělení je speciální případ Stabilního rozděleni. Pak jsme ukázali,že použití MMV bylo nejvhodnější a díky ní jsme dosáhli lepších odhadů, než kdybychom používali jinou metodu odhadu parametru.

Seznam použité literatury

[1] Прохоров Ю., *Математический энциклопедический словарь*, Советская энциклопедия, Москва, 1988.

[2] Волков И. К., Зуев С. М.,.Цветкова Г. М. *Случайные процессы*, МГТУ им. Н. Э. Баумана, Москва, 1999.

[3] Rogalewicz V. *Pravděpodobnost a statistika pro inženýry*, CVUT, Praha, 1998.

[4] Andel J. *Zaklady matematicke statistiky*, MATFYZPRESS, Praha, 2011.

[5] Houda M. Using Metrics in Stability of Stochastic Programming Problems, casopis Acta Oeconomica Pragensia, VSE, Praha, Vol.13, No.1, 2005.

[6] Денисов В. И., Тимофеев В. С. *Устойчивые распределения и оценивание парамтров регрессионных зависимостей,* журнал Известия Томского политехнического университета , Vol 318 No. 2, 10-15, Томск, 2011.

[7] RNDr. Václav Kohout, vyukovy materiál dostupní na http://www.kmt.zcu.cz/person/Kohout/info_soubory/letnisem/SS/stat19.pdf.

[8] Shiryaev A. N. *Essentials of stochastic finance*. Facts, Models, Theory: World Scientific, FAZIS, Moscow, 1999.

[9] Nolan J. P. *Stable Distributions . Models for Heavy Tailed Data,* Math/Stt Department American University, Washington, DC, 2009.

[10] Gnedenko B. V. *Lokální limitní věta pro součet nezávislých, stejně rozdělených nahodných veličin,* Časopis Pokroky matematiky,fyziky a astronomie, Vol.1 No. 1, 3-13, 1956. http://dml.cz/dmlcz/137253

[11] Бокс Дж., Дженкинс Г. *Анализ временных рядов, прогноз и управление: Пер. с англ.* // Под ред. В.Ф. Писаренко. Мир, Москва, 1974.

[12] Chouda M. *Using Metrics in stability of stochastic Programming Problems*, Acta Oeconomica Pragensia, Vol. 13, No.1, 2005.

Seznam tabulek

Seznam obrázků

Seznam použitých zkratek

MMV metoda maximální věrohodnosti.

MLE maximum Likelihood estimators.

MV maximální věrohodnost.

MNČ metoda nejmenších čtverců.

EX střední hodnota.

AR(q) autoregresní model s parametrem q.

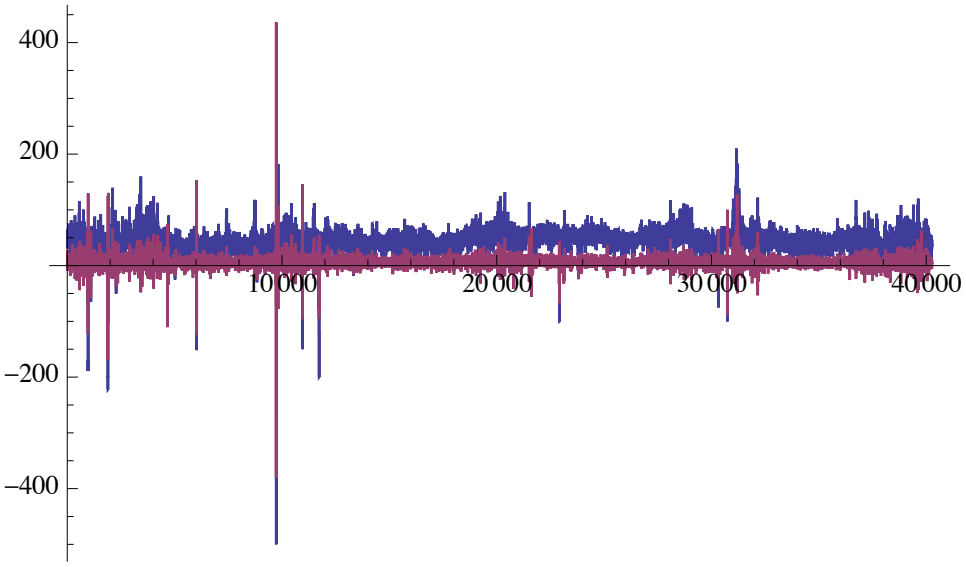
MoM "MethodOfMoments" momentová metoda.

MCM "MethodOfCentralMoments" metoda centrálních momentů.

Příloha.

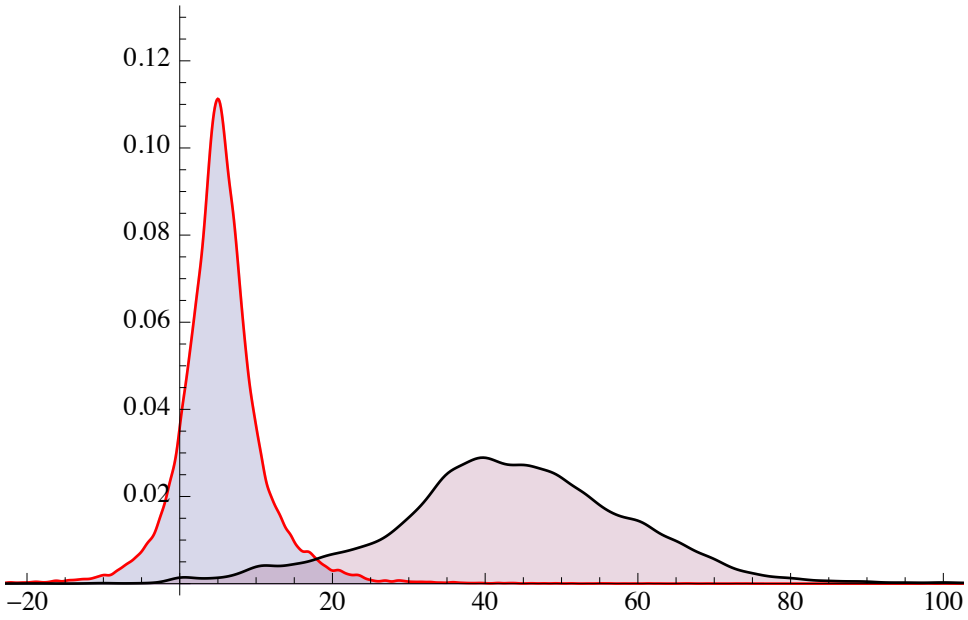
Přiloha 1. Grafy.

1.1. Ceny na elektřinu.

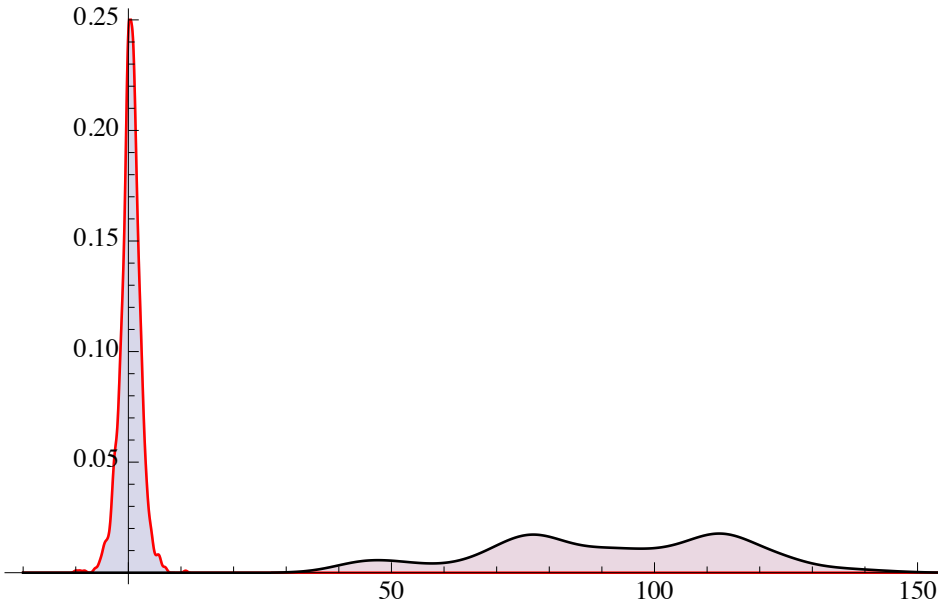


Obrázek 1. Ceny elektřiny a residua k těmto cenám vypočteny pomocí *AR(2)*.

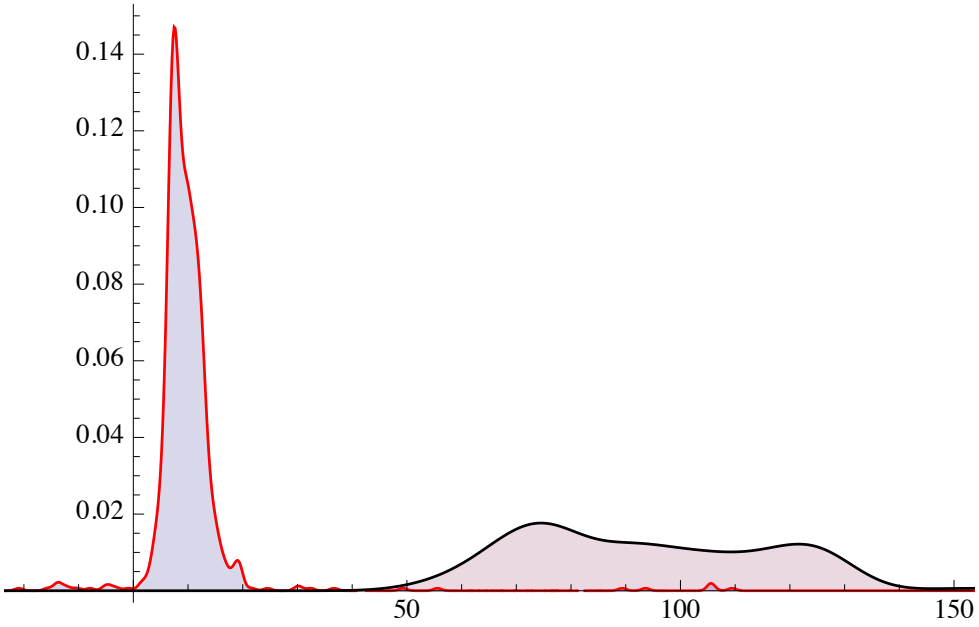
1.2. Histogramy cen.



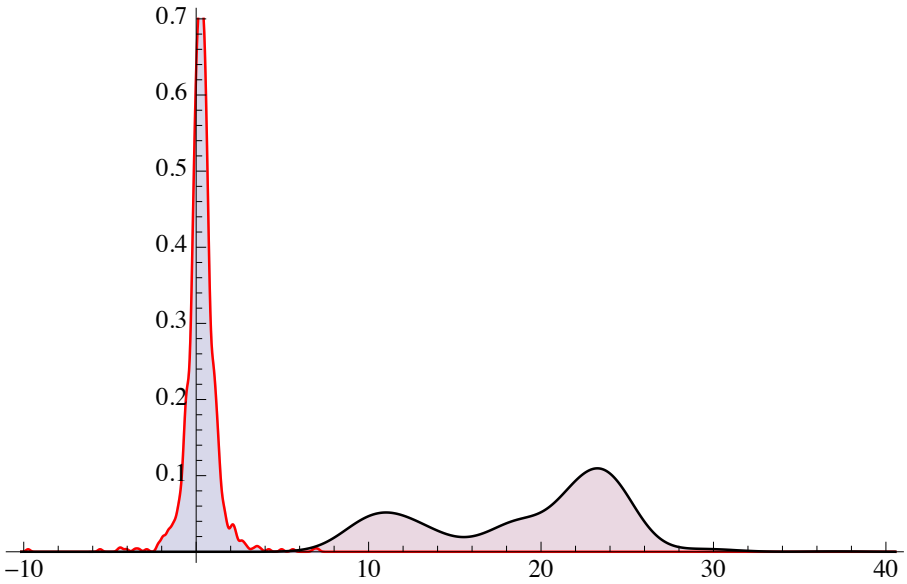
Obrázek 2. Histogramy cen na elektřinu a residua vypočteny pomocí *AR(2)*.



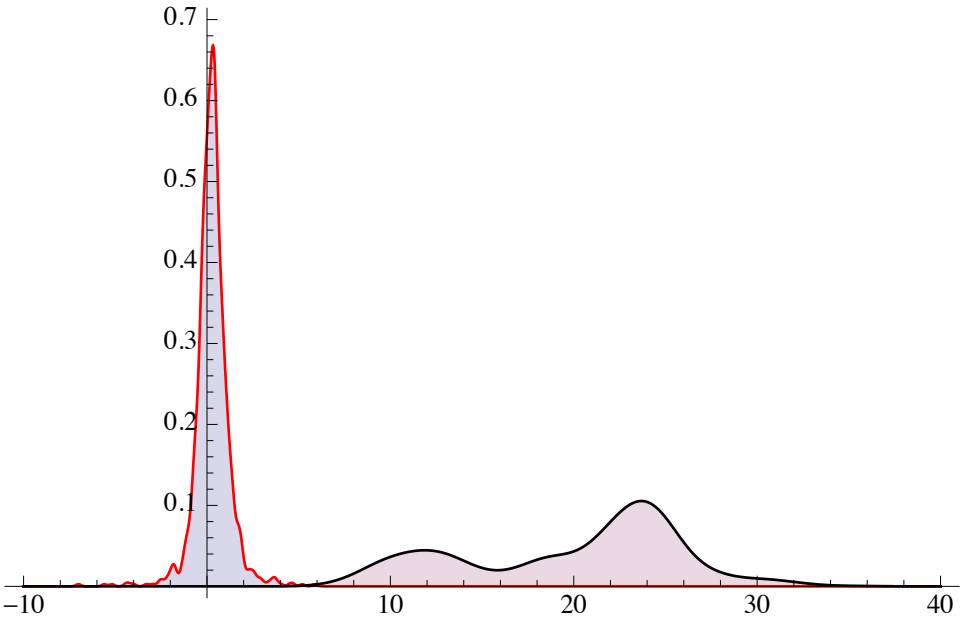
Obrázek 3. Histogramy cen komodit a residua pro ICE vypočteny pomocí *AR(2)*.



Obrázek 4. Histogramy cen komodit a residua pro ICAP(COA) vypočteny pomocí *AR(1)*.

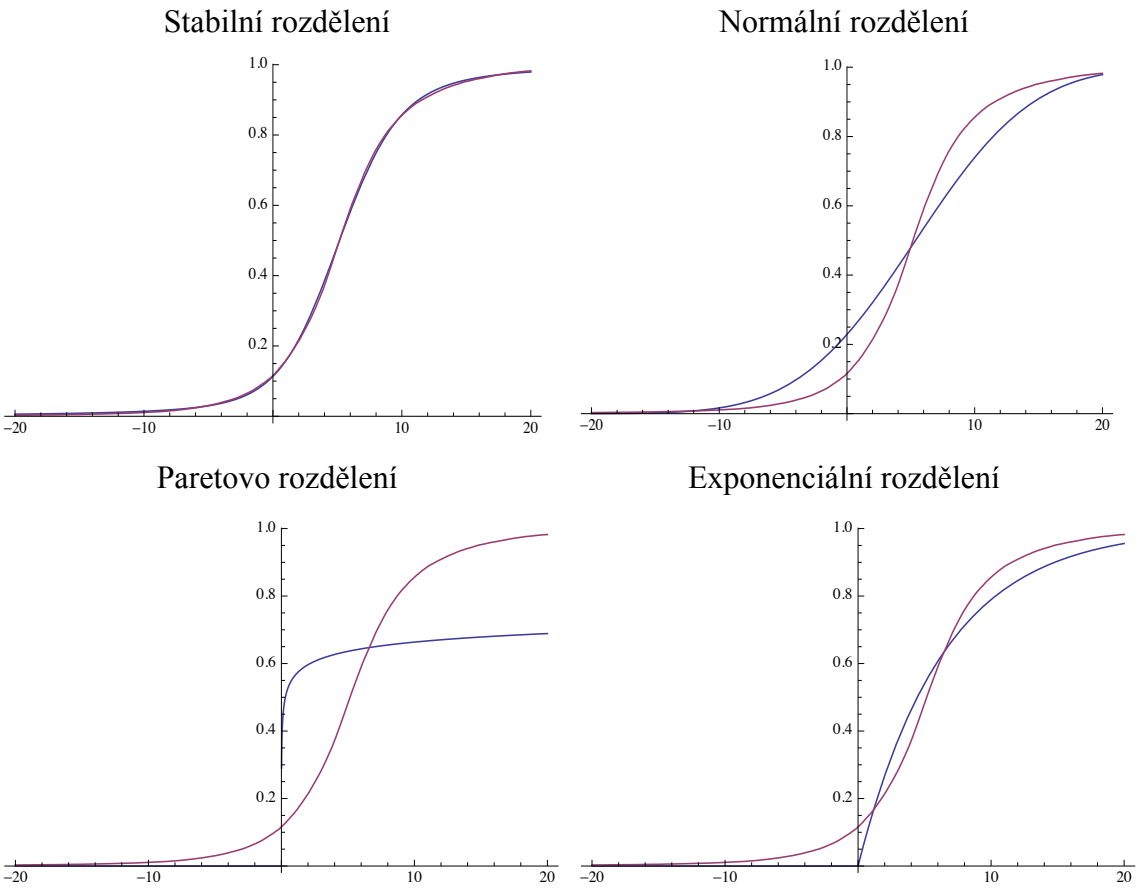


Obrázek 5. Histogramy cen komodit a residua pro SPN vypočteny pomocí *AR(2)*.



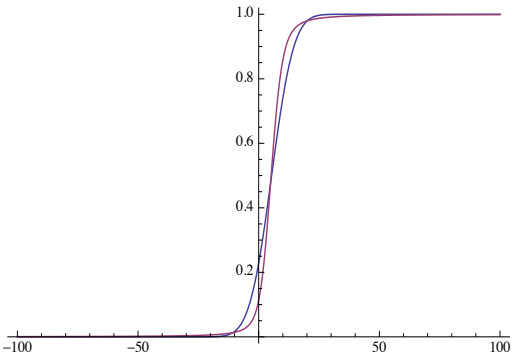
Obrázek 6. Histogramy cen komodit a residua pro ICAP(NAT) vypočteny pomocí *AR(1)*.

1.3. Porovnaní distribučních funkci.

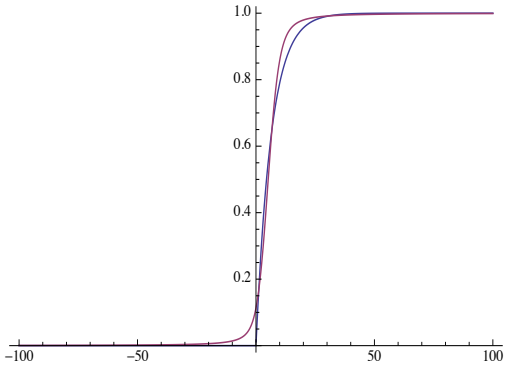


Obrázek 7. Porovnaní Empirické distribuční funkce a distribuční funkce F_n příslušného rozdělení pro časovou řadu cen za elektřinu .

Normální rozdělení

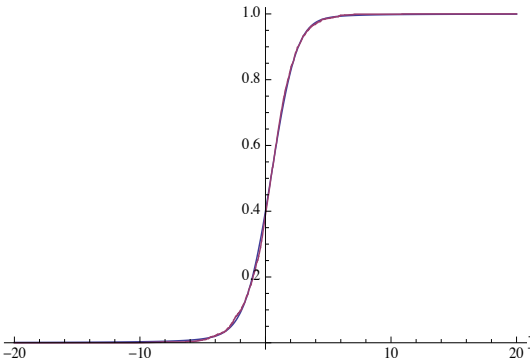


Exponenciální rozdělení

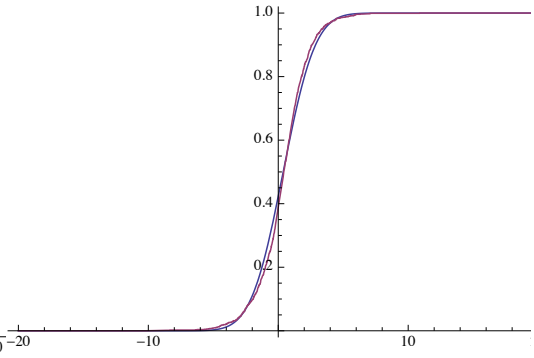


Obrázek 8. Porovnání distribuční funkce Stabilního rozdělení s distribučními funkcemi Normálního a Exponenciálního rozdělení pro časovou řadu cen za elektřinu.

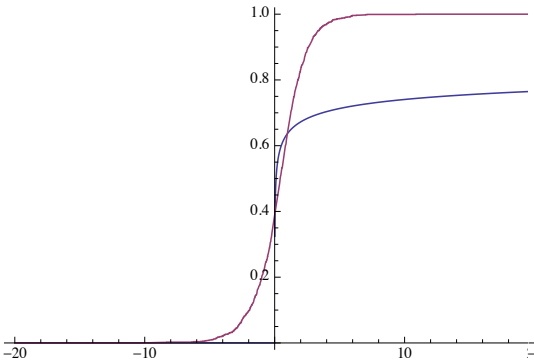
Stabilní rozdělení



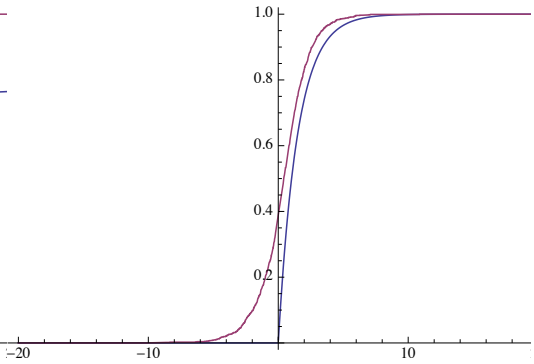
Normální rozdělení



Paretovo rozdělení

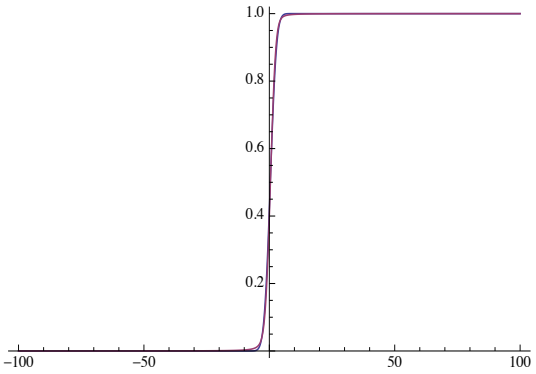


Exponenciální rozdělení

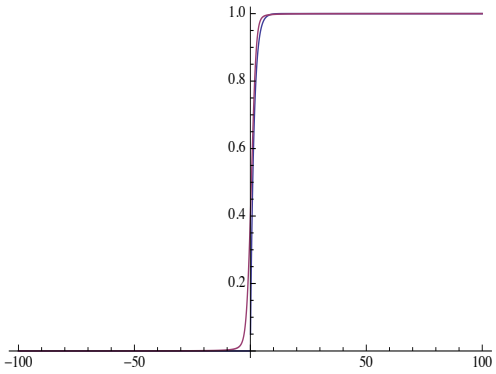


Obrázek 9. Porovnání Empirické distribuční funkce a distribuční funkce F_n příslušného rozdělení pro časovou řadu cen z komodit ICE.

Normální rozdělení

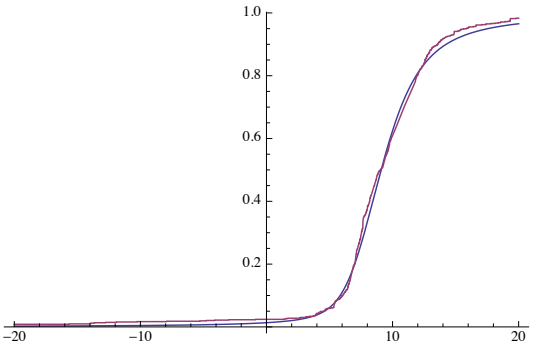


Exponenciální rozdělení

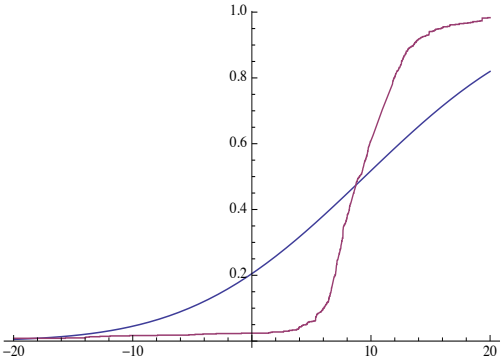


Obrázek 10. Porovnání distribuční funkce Stabilního rozdělení s distribučními funkcemi Normálního a Exponenciálního rozdělení pro časovou řadu cen z komodit ICE.

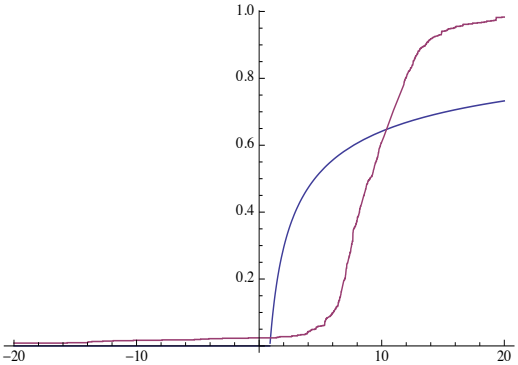
Stabilní rozdělení



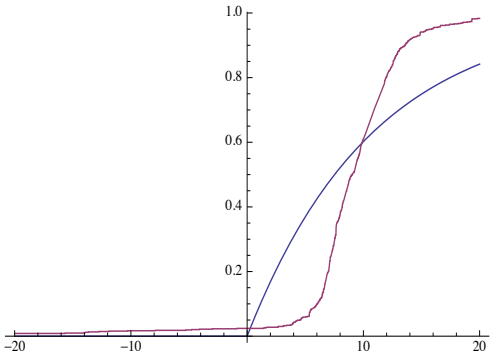
Normální rozdělení



Paretovo rozdělení

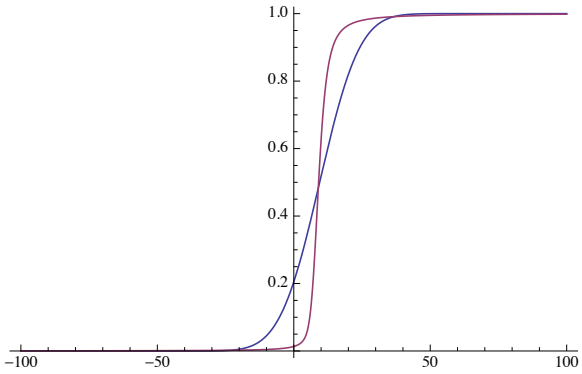


Exponenciální rozdělení

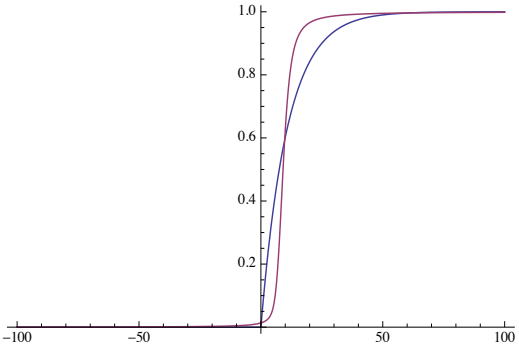


Obrázek 11. Porovnání Empirické distribuční funkce a distribuční funkce F_n příslušného rozdělení pro časovou řadu výnosu z komodit ICAP(COA).

Normální rozdělení

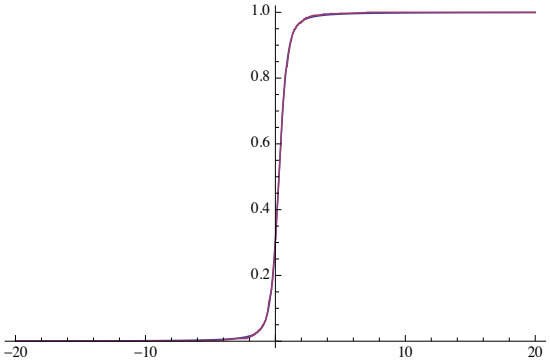


Exponenciální rozdělení

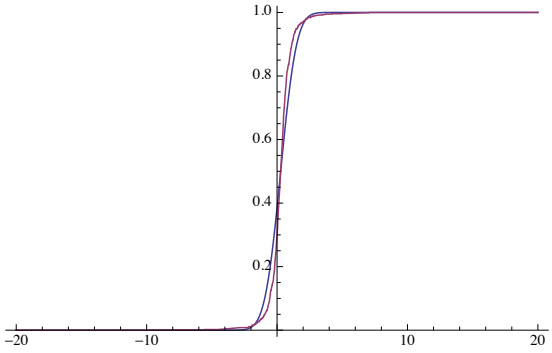


Obrázek 12. Porovnání distribuční funkce Stabilního rozdělení s distribučními funkcemi Normálního a Exponenciálního rozdělení pro časovou řadu výnosu z komodit ICAP(COA).

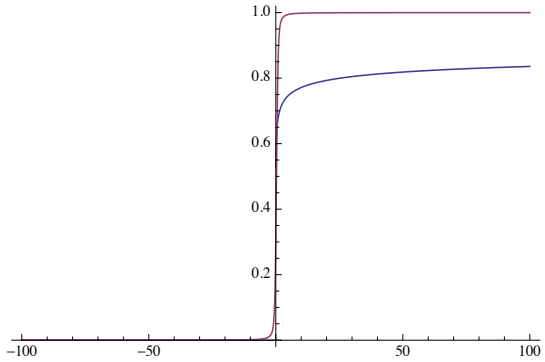
Stabilní rozdělení



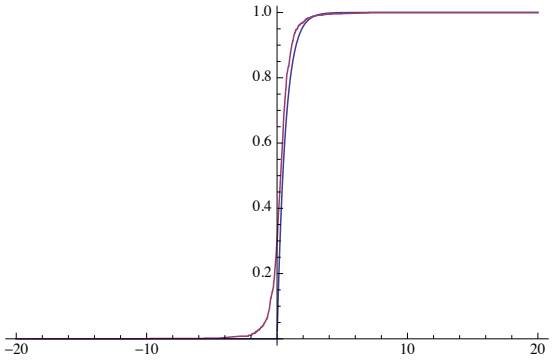
Normální rozdělení



Paretovo rozdělení

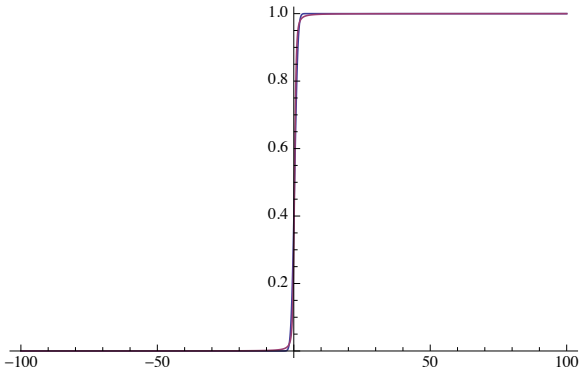


Exponenciální rozdělení

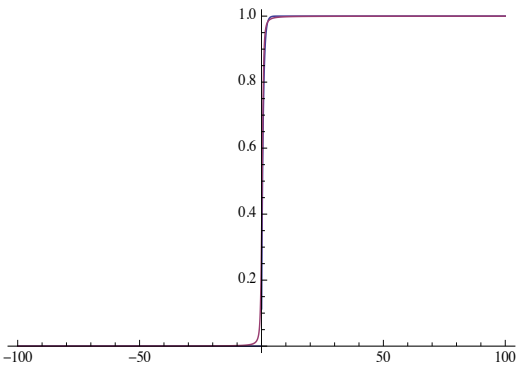


Obrázek 13. Porovnání Empirické distribuční funkce a distribuční funkce F_n příslušného rozdělení pro časovou řadu výnosu z komodit SPN.

Normální rozdělení

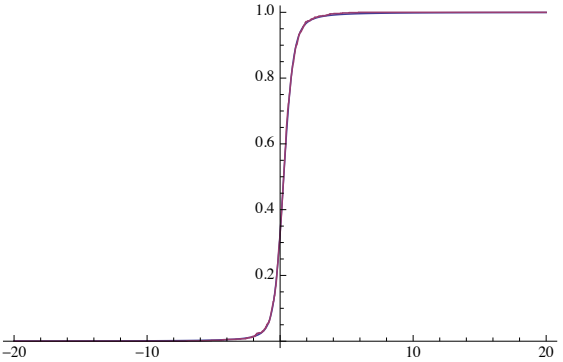


Exponenciální rozdělení

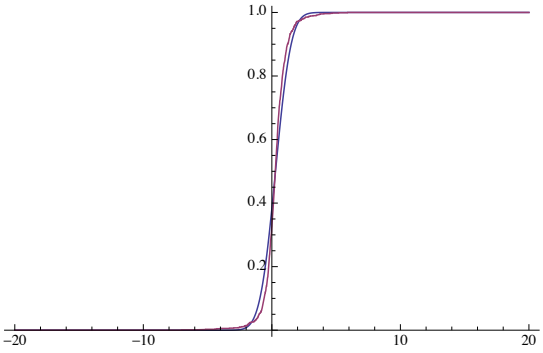


Obrázek 14. Porovnání distribuční funkce Stabilního rozdělení s distribučními funkcemi Normálního a Exponenciálního rozdělení pro časovou řadu výnosu z komodit SPN.

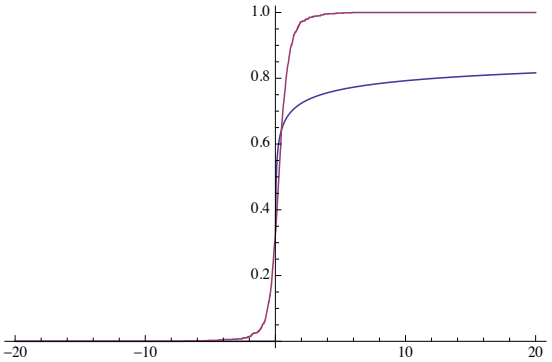
Stabilní rozdělení



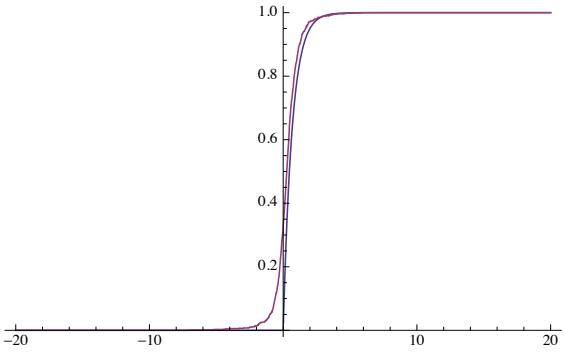
Normální rozdělení



Paretovo rozdělení

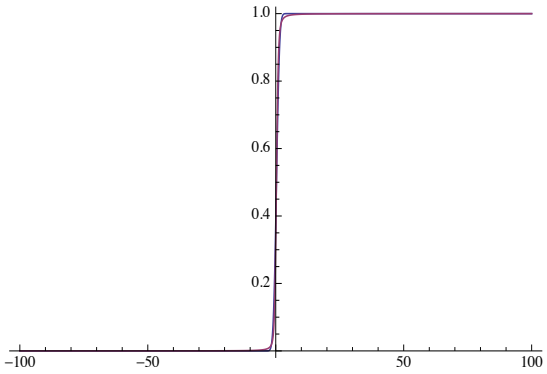


Exponenciální rozdělení

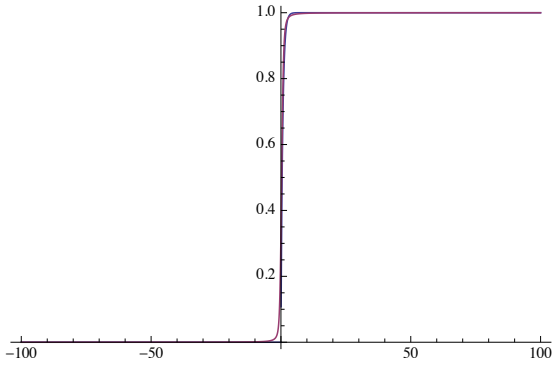


Obrázek 15. Porovnání Empirické distribuční funkce a distribuční funkce F_n příslušného rozdělení pro časovou řadu cen z komodit ICAP(NAT).

Normální rozdělení



Exponenciální rozdělení



Obrázek 16. Porovnání distribuční funkce Stabilního rozdělení s distribučními funkcemi Normálního a Exponenciálního rozdělení pro časovou řadu cen z komodit ICAP(NAT).

Přiloha 2. Kolmogorov-Smirnov test.

Předtím než přestoupíme k formulaci tohoto testu, uvedeme některá pomocná tvrzení [4].

Defenice 1. *Nechť $X_1,\ldots,X_m$ je náhodný výběr rozdělení s distribuční funkcí F, kde x je dané reálné číslo. Označíme náhodné veličiny*

$$\xi_i(x)=\begin{cases} 1, & X_i < x, \\ 0, & X_i \geq x \end{cases}$$

pro $i=1,\ldots,m$. Položíme

$$F_m(x)=\frac{1}{m}\sum_{i=1}^{m}\xi_i(x). \qquad (1.1)$$

Funkce $F_m(x)$ je empirická distribuční funkce.

Věta 1 *Pro každé x platí*

$$F_m(x) \to F(x) \text{ skoro jisté pro } m \to \infty.$$

To znamená, že s rostoucím *m* funkce $F_m(x)$ se bude blížit ke skutečně distribuční funkci $F(x)$.

Věta 2 *(Glivenkova) Označme* $D_m = \sup_x \left| F_m(x) - F(x) \right|$. *Pak platí*

$$P\left(\lim_{m\to\infty} D_m = 0\right) = 1.$$

Důkaz Víc ve svazku [4], str. 241.

Nechť $X_1,\ldots,X_m$ je náhodný výběr z rozdělení se spojitou distribuční funkci *F* a nechť $Y_1,\ldots,Y_m$ je na něm nezávislý náhodný výběr z rozdělení se spojitou distribuční funkcí *G*. Definujeme hypotézy, které pak budeme testovat

$$H_0 : F = G$$
$$H_1 : F \neq G.$$

F_m je empirickou distribuční funkcí prvního výběru a zároveň' druhého výběru. Podle vět 1 a 2 platí, ze funkce F_m a G_n pro rostoucí *m* a *n* blíží distribučním funkcím *F* a *G*. Označme

$$D_{m,n} = \sup_x \left| F_m(x) - G_n(x) \right|.$$

Pokud platí H_0, pak podle Glivenkovy věty skoro jisté při $m \to \infty, n \to \infty$.

Výsledek na kterém se pak dá založit test popisuje následující věta.

Věta 3 *(Smirnovova) Označme* $M = \frac{mn}{(m+n)}$. *Nechť*

$$K(\lambda) = 1 - 2\sum_{k=1}^{\infty} (-1)^{k+1} \exp\{-2k^2\lambda^2\}. \tag{1.2}$$

Pak pro každé $\lambda > 0$ *platí*

$$\lim_{m,n\to\infty} P(\sqrt{M}\, D_{m,n} < \lambda) = K(\lambda).$$

Praktické provedení Kolmogorovova-Smirnovava testu spočívá v tom, že se ve výběru $X_1,\ldots,X_m$ a $Y_1,\ldots,Y_m$ vypočtou empirické distribuční funkce F_m a G_n a veličina $D_{m,n}$. Jsou-li čísla *m* a *n* malá, porovná se $D_{m,n}$ s přesnými kritickými hodnotami $D_{m,n}(\alpha)$.

Pokud *m* a *n* jsou větší hodnoty použijeme větu 3. Přičemž položí se

$$\lambda_0 = \sqrt{M}\, D_{m,n}$$

a vypočítáme hodnotu $K(\lambda_0)$. Pokud dostaneme výsledek ze $K(\lambda_0) \geq 1-\alpha$, zamítneme H_0 na hladině, která se s rostoucími rozsahy výběru blíží číslu α. Při větších hodnotách *m* a *n* kritická hodnota pro veličinu $D_{m,n}$ budeme aproximovat $D^*_{m,n}(\alpha)$. Proto hypotézu H_0 zamítneme v případě, když $D_{m,n} \geq D^*_{m,n}(\alpha)$.

Příloha 3. Obsah přiloženého CD.

3.1. Bakalářská práce PDF.

3.2. Skripty.

Obsahuje 6 souboru postupně pro jednotlivý rozděleni, které byly popsaný v 3 časti práce.

3.3. Testy.

Obsahuje 5 soubory s jednotlivými testy pro příslušný data.

3.4. Data.

Data , které byly použity při testovaní.