MATEMATICKO-FYZIKÁLNÍ FAKULTA
Univerzita Karlova

BAKALÁŘSKÁ PRÁCE

Jan Škurek

Expektily a jejich odhady

Katedra pravděpodobnosti a matematické statistiky

Vedoucí bakalářské práce: RNDr. Šárka Hudecová, Ph.D.
Studijní program: Finanční matematika

Praha 2023

Prohlašuji, že jsem tuto bakalářskou práci vypracoval(a) samostatně a výhradně s použitím citovaných pramenů, literatury a dalších odborných zdrojů. Tato práce nebyla využita k získání jiného nebo stejného titulu.

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 má právo na uzavření licenční smlouvy o užití této práce jako školního díla podle §60 odst. 1 autorského zákona.

V .... dne ....

Podpis autora

Na tomto místě bych chtěl poděkovat vedoucí mé bakalářské práce RNDr. Šárce Hudecové, Ph.D. za věnovaný čas, cenné rady a ochotu, během psaní této práce. Dále bych chtěl poděkovat své rodině za podporu během celého studia.

Název práce: Expektily a jejich odhady

Autor: Jan Škurek

Katedra: Katedra pravděpodobnosti a matematické statistiky

Vedoucí bakalářské práce: RNDr. Šárka Hudecová, Ph.D., Katedra pravděpodobnosti a matematické statistiky

Abstrakt: Tato bakalářská práce se zaměřuje na studium expektilů jako alternativního přístupu k tradičním kvantilům. Expektily se stávají stále populárnějšími jako míra rizika v různých oblastech, včetně finančního a pojišťovacího sektoru. V práci jsou uvedeny základní vlastnosti a jednotlivá odvození jsou podrobně popsána. V rámci odhadu expektilu je vedle definice výběrového expektilu představena parametrická metoda odhadu expektilu. Praktická část je věnována odhadům výběrového expektilu, parametrického odhadu momentovou metodou a porovnáním obou metod na nasimulovaném výběru z exponenciálního rozdělení.

Klíčová slova: kvantily, expektily, parametrické odhady

Title: Expectiles and their estimates

Author: Jan Škurek

Department: Department of Probability and Mathematical Statistics

Supervisor: RNDr. Šárka Hudecová, Ph.D., Department of Probability and Mathematical Statistics

Abstract: This bachelor's thesis focuses on studying expectiles as an alternative approach to traditional quantiles. Expectiles are becoming increasingly popular as risk measures in various fields, including finance and insurance sectors. The thesis presents basic properties and derivations of expectiles in detail. In addition to the definition of sample expectiles, a parametric estimation method for expectiles is introduced. The practical part is devoted to the estimation of sample expectiles, parametric estimation using the method of moments, and a comparison of both methods on a simulated sample from the exponential distribution.

Keywords: quantiles, expectiles, parametric estimates

Obsah

Úvod

V této bakalářské práci se budeme zabývat expektily, které jsou relativně novým pojmem v rámci matematické statistiky. Expektily byly poprvé navrženy v Newey a Powell (1987). Tato literatura je stále jednou z nejvíce citovaných v oblasti expektilů. Expektily jsou, podobně jako kvantily, charakteristikami rozdělení náhodné veličiny a v posledních letech se začínají využívat jako míry rizika.

V první kapitole se podíváme na kvantily a ukážeme si jejich základní vlastnosti. Pro pochopení expektilů a jejich vlastností je užitečné nejprve pochopit, co jsou kvantily a jak fungují.

Ve druhé kapitole si zadefinujeme expektil a ukážeme si jeho základní vlastnosti. Odvodíme si explicitní vyjádření pro rovnoměrné a exponenciální rozdělení. Tento tvar následně využijeme pro parametrický odhad expektilu pomocí momentové metody.

V praktické části budeme zkoumat, jak přesné odhady poskytuje definice pro výběrový expektil v porovnání s parametrickou metodou na simulovaných náhodných výběrech z exponenciálního rozdělení o různých rozsazích.

Vlastním přínosem práce je podrobné odvození vztahů pro expektily a detailní popsání jednotlivých kroků v důkazech. Dále je vlastním přínosem praktická část, zejména parametrický odhad expektilů.

1. Kvantily

V této kapitole se budeme zabývat jedním z klíčových konceptů statistiky a analýzy dat - kvantily. Kvantily jsou důležitým nástrojem pro porozumění rozdělení dat a popis jejich variabilit. Mají široké využití například pro tvorbu intervalů spolehlivosti a identifikaci extrémních hodnot. Kvantily poskytují informace o centrálních charakteristikách datového souboru, jako jsou například medián nebo kvartily.

1.1 Základní pojmy

Nejprve si zadefinujeme kvantilovou funkci pro náhodnou veličinu X s distribuční funkcí F_X.

Definice 1. *Nechť F_X je distribuční funkce reálné náhodné veličiny X a $\alpha \in (0{,}1)$. Funkce*

$$F_X^{-1}(\alpha) = \inf\{x : F_X(x) \geq \alpha\} \tag{1.1}$$

se nazývá kvantilová funkce *náhodné veličiny X.*

Mnohdy nás zajímá hodnota kvantilové funkce pro konkrétní hodnotu α. Tuto hodnotu budeme nazývat α-kvantilem.

Definice 2. *Nechť $\alpha \in (0{,}1)$.* α-kvantil $u_X(\alpha)$ *rozdělení F_X je kterékoli reálné číslo splňující*

$$\lim_{h \to 0}\Big(F_X(u_X(\alpha) - h)\Big) \leq \alpha \text{ a } F_X(u_X(\alpha)) \geq \alpha.$$

Poznámka. Definic kvantilu je více, tato jej neurčuje vždy jednoznačně. $F_X^{-1}(\alpha)$ je vždy jeden z α-kvantilů.

Hodnoty α-kvantilu nám udávají hodnotu, pod kterou se nachází α-podíl hodnot v daném rozdělení.

Některé hodnoty kvantilu mají speciální názvy, například u hodnoty $u_X(0.5)$ mluvíme o mediánu, pro $u_X(0.25)$ a $u_X(0.1)$ o prvním kvartilu a decilu.

Věta 1. *(Omelka) Nechť X má distribuční funkci F_X, pak*

$$F_X^{-1}(\alpha) = \arg\min_{\beta} \mathsf{E}\, r_\alpha(X - \beta), \tag{1.2}$$

kde α je pevné z intervalu $(0{,}1)$. r_α se nazývá ztrátová funkce a má následující tvar

$$r_\alpha(x) = (\alpha - \mathbb{I}_{(-\infty,0)}(x)) \cdot x. \tag{1.3}$$

Funkcí $\mathbb{I}_A$ je myšlena indikátorová funkce, která má pro množinu A tento tvar:

$$\mathbb{I}_A(x) = \begin{cases} 1, & \text{pro } x \in A, \\ 0, & \text{pro } x \notin A. \end{cases}$$

Důkaz. Inspiraci si bereme z předlohy a některé kroky více rozepíšeme. Nejprve si zavedeme derivaci naší ztrátové funkce. Dostáváme

$$\frac{\partial}{\partial x} r_\alpha(x) = \alpha - \mathbb{I}_{(-\infty,0)}(x) = \begin{cases} \alpha - 1, & \text{pro } x < 0, \\ \alpha, & \text{pro } x > 0, \end{cases}$$

pro $x \neq 0$. Dodefinujeme v bodě $x = 0$, $r'_\alpha(0) = 0$. Dále si označme $M(\beta) = \mathsf{E}\ r_\alpha(X - \beta)$ funkci, kterou budeme minimalizovat. Rovnou si funkci $M(\beta)$ zderivujeme podle β

$$M'(\beta) = \frac{\partial}{\partial \beta} \mathsf{E}\ r_\alpha(X - \beta).$$

Rozepíšeme si E jako integrál a zaměníme derivaci a integrál dle Leibnizova pravidla

$$M'(\beta) = \frac{\partial}{\partial \beta} \int_{-\infty}^{\infty} r_\alpha(x - \beta) dF_X(x) = \int_{-\infty}^{\infty} \frac{\partial}{\partial \beta} r_\alpha(x - \beta) dF_X(x).$$

Upravíme vnitřek integrálu

$$\begin{aligned} M'(\beta) &= \int_{-\infty}^{\infty} -(\alpha - 1)\mathbb{I}_{(-\infty,0)}(x - \beta) - \alpha \mathbb{I}_{(0,\infty)}(x - \beta) dF_X(x) \\ &= (1 - \alpha) \int_{-\infty}^{\beta} dF_X(x) - \alpha \int_{\beta}^{\infty} dF_X(x) \\ &= (1 - \alpha) F_X(\beta) - \alpha(1 - F_X(\beta)) \\ M'(\beta) &= F_X(\beta) - \alpha. \end{aligned}$$

Pro $\beta < F_X^{-1}(\alpha)$ máme

$$M'(\beta_-) = F_X(\beta_-) - \alpha \leq F_X(\beta) - \alpha < 0,$$

$$M'(\beta_+) = F_X(\beta_+) - \alpha = F_X(\beta) - \alpha < 0.$$

Využili jsme, že distribuční funkce je neklesající a zprava spojitá. Protože $\beta < F_X^{-1}(\alpha)$ tak víme, že $\beta \leq \sup\{x : F_X(x) < \alpha\}$, tedy $F_X(\beta) < \alpha$, pro $\beta < F_X^{-1}(\alpha)$.

$M(\beta)$ je spojitá a má zápornou derivaci, pak je klesající na $(-\infty, F_X^{-1}(\alpha))$. Podobně pro $\beta > F_X^{-1}(\alpha)$

$$M'(\beta_-) = F_X(\beta_-) - \alpha \geq F_X(F_X^{-1}(\alpha)) - \alpha \geq 0,$$

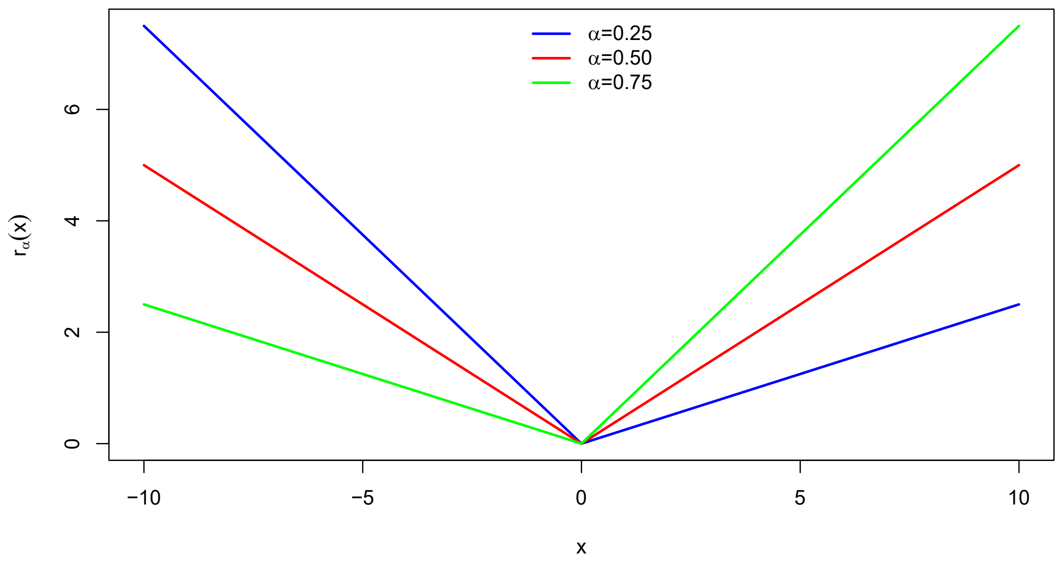
$$M'(\beta_+) = F_X(\beta_+) - \alpha \geq F_X(F_X^{-1}(\alpha)) - \alpha \geq 0.$$

Zde jsem využili, že $F_X(F_X^{-1}(\alpha)) = F_X(\inf\{x : F_X(x) \geq \alpha\}) \geq \alpha$. $M(\beta)$ má nezápornou derivaci a tedy je neklesající na intervalu $(F_X^{-1}(\alpha), \infty)$.

Celkem se globální minimum $M(\beta)$ nachází v bodě $F_X^{-1}(\alpha)$. Ukázali jsme, že rovnost (1.2) platí.

□

Na obrázku 1.1 si můžeme všimnout, jak vypadá ztrátová funkce r_α pro hodnoty α, které odpovídají hodnotám prvního až třetího kvartilu.



Obrázek 1.1: Ztrátová funkce r_α pro konkrétní hodnoty α.

Věta 2. *Nechť* $\boldsymbol{X}$ *je náhodná veličina s distribuční funkcí* $F_X(x)$*. Pokud je* F_X *spojitá a rostoucí, pak platí*

$$F_X(F_X^{-1}(\alpha)) = \alpha,$$

pro $\alpha \in (0,1)$*.*

Důkaz. Důkaz plyne z definice kvantilové funkce.

□

Tato věta nám říká, že pokud je distribuční funkce F_X rostoucí a spojitá, pak se kvantilová funkce rovná inverzní funkci distribuční funkce a zároveň α-kvantily jsou určeny jednoznačně a platí $F_X^{-1}(\alpha) = u_X(\alpha)$.

1.2 Odhady kvantilů

Nyní se podíváme na situaci, kdy máme náhodný výběr $X_1, ..., X_n$ z rozdělení F_X. Pro kvantily jsme si ukázali platnost rovnosti (1.2). Toto využijeme i pro definování výběrového α-kvantilu.

Definice 3. *Nechť* $X_1, ..., X_n$ *je náhodný výběr z rozdělení* F_X*,* $\alpha \in (0,1)$ *a* r_α *je ztrátová funkce definovaná v* (1.3)*, pak číslo splňující*

$$\widehat{u}_n(\alpha) = \arg\min_\beta \frac{1}{n} \sum_{i=1}^{n} r_\alpha(X_i - \beta), \tag{1.4}$$

nazveme výběrovým α-kvantilem.

Z definice plyne, že výběrový α-kvantil je na rozdíl od α-kvantilu určen jednoznačně, protože ztrátová funkce r_α je konvexní a zdola omezená, pak existuje právě jedno globální minimum.

Poznámka. Pro $\alpha = 0.5$ dostaneme výběrový medián.

Věta 3. *(Kulich, str. 61) Nechť $\alpha \in (0,1)$. Nechť $X_1, ..., X_n$ je náhodný výběr z rozdělení, které má distribuční funkci F_X spojitou a rostoucí na nějakém okolí bodu $u_X(\alpha)$.*

(i) Potom $\widehat{u}_n(\alpha) \xrightarrow[n \longrightarrow \infty]{P} u_X(\alpha)$.

(ii) Pokud navíc existuje hustota f_X, která je spojitá a nenulová v bodě $u_X(\alpha)$, pak

$$\sqrt{n}[\widehat{u}_n(\alpha) - u_X(\alpha)] \xrightarrow[n \longrightarrow \infty]{D} \mathcal{N}\left(0, \frac{\alpha(1-\alpha)}{f_X^2(u_X(\alpha))}\right).$$

Důkaz. Jednotlivé kroky důkazu jsou popsány v práci Kulich, str. 62.

□

Poznámka. Značením $\xrightarrow[n \longrightarrow \infty]{P}$ máme na mysli konvergenci v pravděpodobnosti, pro $n \to \infty$, a $\xrightarrow[n \longrightarrow \infty]{D}$ myslíme konvergenci v distribuci pro $n \to \infty$.

Odhad α-kvantilu pomocí výběrového α-kvantilu není jediná možnost. Další možností, jak můžeme odhadovat α-kvantily, je pomocí parametrického postupu, kdy dosazujeme odhadnuté parametry do kvantilové funkce daného rozdělení. Parametry můžeme odhadovat například metodou maximální věrohodnosti nebo momentovou metodou. Ukážeme si odhadování parametrů pomocí momentové metody. Zbytek kapitoly jsme převzali z Kulich, str. 45, a doplnili o příklady pro odhady α-kvantilu.

Mějme náhodný výběr $X_1, ..., X_n$ z rozdělení s hustotou $f(x; \boldsymbol{\theta_X})$, kde tvar funkce $f(;)$ je známý a $\boldsymbol{\theta_X}$ je neznámý parametr, který leží v parametrickém prostoru $\boldsymbol{\Theta} \subseteq \mathbb{R}^d, d \geq 1$. Budeme předpokládat, že $\mathsf{E}\,|X|^d < \infty$. Pro zjednodušení budeme uvažovat pouze ta rozdělení, jejichž $\mathsf{E}\ X_i$ je funkcí parametru θ_X, případně pro $d = 2$ parametry jsou $\mathsf{E}\ X_i$ a $Var X_i$ funkcemi parametru $\boldsymbol{\theta_X}$. Dále budeme předpokládat, že distribuční funkce je absolutně spojitá a ryze monotónní.

Máme tedy model

$$\mathcal{F} = \{\text{rozdělení s hustotou } f(x; \boldsymbol{\theta}), \boldsymbol{\theta} \in \boldsymbol{\Theta} \subseteq \mathbb{R}^d\}.$$

Nejprve se podíváme na případ, kdy máme pouze $d = 1$ parametr. Máme $\mathsf{E}\ X_i = g(\theta_X)$, pokud je funkce g ryze monotónní, pak ji můžeme zinvertovat a dostáváme $\theta_X = g^{-1}(\mathsf{E}\ X_i)$. Víme, že $\overline{X}_n$ je konzistentní odhad $\mathsf{E}\ X_i$, a tak se nabízí odhadnout parametr θ_X jako $\widehat{\theta}_n = g^{-1}(\overline{X}_n)$.

Ukažme si to na příkladu, kdy máme náhodný výběr $X_1, ..., X_n$ z exponenciálního rozdělení s parametrem $\lambda > 0$, $Exp(\lambda)$, s hustotou $f_X(x) = \lambda e^{-\lambda x}$, pro $x \geq 0$ Víme, že $\mathsf{E}\ X_i = \frac{1}{\lambda}$, pak momentovým odhadem parametru λ je $\widehat{\lambda} = \frac{1}{\overline{X}_n}$.

Nyní si ukážeme, jak můžeme parametricky odhadnout α-kvantil. V parametrickém modelu předpokládáme, že $u_X(\alpha) = h(\alpha, \theta_X)$. Využijeme momentovou metodu pro odhad θ_X a dostáváme

$$\tilde{u}_n(\alpha) = h(\alpha, \widehat{\theta}_n).$$

Pojďme se podívat na názorný příklad odhadu kvantilu rozdělení $Exp(\lambda)$ pomocí momentové metody. Víme, že momentovým odhadem parametru λ je $\widehat{\lambda} = \frac{1}{\overline{X}_n}$. Kvantilovou funkci $Exp(\lambda)$ získáme jako inverzní funkci k F_X dle věty 2, protože F_X je spojitá a ryze monotónní. Distribuční funkce má tvar

$$F_X(x) = \begin{cases} 0, & \text{pro } x < 0, \\ 1 - e^{-\lambda x}, & \text{pro } x \geq 0. \end{cases}$$

Inverzní funkce k $F_X(x)$ má tvar

$$F_X^{-1}(\alpha) = -\frac{\ln(1-\alpha)}{\lambda},$$

pro $\alpha \in (0,1)$. F_X^{-1} je s korespondujícím značením výše naše funkce h. Pro odhad $\tilde{u}_n(\alpha)$ dostáváme

$$\tilde{u}_n(\alpha) = h(\alpha,\widehat{\lambda}) = -\frac{\ln(1-\alpha)}{\widehat{\lambda}} = -\overline{X}_n \ln(1-\alpha),$$

pro $\alpha \in (0,1)$.

Zavedli jsme značení $\tilde{u}_n(\alpha)$ pro parametrický odhad α-kvantilu a $\widehat{u}_n(\alpha)$ pro empirický odhad α-kvantilu.

Nyní si odhad momentovou metodou rozšíříme na $d = 2$ parametry. Víme, že $\overline{X}_n$ a S_n^2 jsou konzistentní odhady $\mathsf{E}\ X_i$ a $VarX_i$. Vyjádříme si podobně jako pro případ $d = 1$ parametr $g(\boldsymbol{\theta_X}) = (\mathsf{E}\ X_i, VarX_i)^T$. Máme soustavu 2 rovnic o 2 neznámých, pokud je g prostá, dostaneme $\boldsymbol{\theta_X} = g^{-1}(\mathsf{E}\ X_i, VarX_i)$. Pokud je g^{-1} spojitá, pak je $\widehat{\boldsymbol{\theta}}_n = g^{-1}(\overline{X}_n, S_n^2)$ konzistentní odhad.

Pro lepší pochopení si to ukážeme na příkladu, kdy odhadujeme parametry θ_1 a θ_2 rovnoměrného rozdělení $R(\theta_1,\theta_2)$, pro $-\theta_1 \neq \theta_2$. Máme soustavu rovnic

$$\begin{aligned} \mathsf{E}\ X_i &= \frac{\theta_1 + \theta_2}{2} \\ VarX_i &= \frac{(\theta_2 - \theta_1)^2}{12}. \end{aligned}$$

Z první rovnice si vyjádříme $\theta_1 = 2\,\mathsf{E}\ X_i - \theta_2$ a dosadíme do druhé rovnice

$$VarX_i = \frac{(\theta_2 - 2\,\mathsf{E}\ X_i + \theta_2)^2}{12} = \frac{(\theta_2 - \mathsf{E}\ X_i)^2}{3}.$$

Po umocnění na $\frac{1}{2}$ máme

$$\sqrt{VarX_i} = \frac{|\theta_2 - \mathsf{E}\ X_i|}{\sqrt{3}}.$$

Jelikož $\theta_2 > \mathsf{E}\ X_i$, tak můžeme odstranit absolutní hodnotu. Vyjádříme si θ_2 a dostáváme

$$\begin{aligned} \theta_2 &= \mathsf{E}\ X_i + \sqrt{3VarX_i} \\ \theta_1 &= \mathsf{E}\ X_i - \sqrt{3VarX_i}. \end{aligned}$$

Momentové odhady $\widehat{\theta}_1$ a $\widehat{\theta}_2$ parametru θ_1 a θ_2 jsou

$$\widehat{\theta}_1 = \overline{X}_n - \sqrt{3S_n^2}$$
$$\widehat{\theta}_2 = \overline{X}_n + \sqrt{3S_n^2}.$$

Pro nalezení parametrického odhadu $u_n(\alpha)$ budeme postupovat obdobně jako v předchozím příkladě. Víme, že distribuční funkce rozdělení $R(\theta_1,\theta_2)$ má tvar

$$F_X(x) = \begin{cases} 0, & \text{pro } x < \theta_1, \\ \frac{x-\theta_1}{\theta_2-\theta_1}, & \text{pro } \theta_1 \leq x \leq \theta_2, \\ 1, & \text{pro } x > \theta_2. \end{cases}$$

Vidíme, že je na intervalu (θ_1,θ_2) spojitá a rostoucí, tedy její kvantilová funkce bude rovna inverzní funkci distribuční funkce, která má tento tvar

$$u_X(\alpha) = \alpha\theta_2 + \theta_1(1-\alpha),$$

pro $\alpha \in (0,1)$. Parametrický odhad $\tilde{u}_n(\alpha)$ momentovou metodou získáme dosazením

$$\tilde{u}_n(\alpha) = \alpha\widehat{\theta}_2 + (1-\alpha)\widehat{\theta}_1 = \alpha(\overline{X}_n + \sqrt{3S_n^2}) + (1-\alpha)(\overline{X}_n - \sqrt{3S_n^2})$$
$$\tilde{u}_n(\alpha) = \overline{X}_n - (2\alpha+1)\sqrt{3S_n^2},$$

pro $\alpha \in (0,1)$.

Poznámka. V případě kdy $\mathsf{E}\ X_i$ není funkcí parametru θ_X, pak hledáme odhad parametru pomocí vyšších momentů. Pro více parametrů je postup analogický. Některé vlastnosti odhadu se tímto postupem mohou změnit.

2. Expektily

V této kapitole se budeme zabývat expektily, což je koncept v teorii pravděpodobnosti a matematické statistiky, který rozšiřuje tradiční pojetí kvantilů. Článek Newey a Powell (1987) je často citován a využíván jako primární zdroj k expektilům. V rámci této kapitoly budeme také brát Newey a Powell (1987) jako hlavní zdroj informací.

2.1 Motivace

Expektily mohou mít více interpretací, ale v naší práci se zaměříme pouze na jednu z nich, a to pomocí metody nejmenších čtverců. Mějme náhodnou veličinu X s konečnou střední hodnotou. Pro nalezení střední hodnoty budeme chtít najít takovou hodnotu β, která minimalizuje $\mathsf{E}\,(X-\beta)^2$. Když si definujeme ztrátovou funkci $L(x-\beta) = (x-\beta)^2$, pak střední hodnotu můžeme přepsat jako

$$\mathsf{E}\ X = \arg\min_{\beta} \mathsf{E}\ L(X-\beta).$$

Ztrátová funkce $L(x-\beta)$ má stejné váhy bez ohledu na hodnotu x, což znamená, že každá hodnota x přispívá stejnou měrou k celkové hodnotě ztrátové funkce.

Pokud bychom chtěli upravit váhy v závislosti na hodnotě x, mohli bychom použít upravenou ztrátovou funkci

$$\rho_\alpha(x-\beta) = (1-\alpha)(x-\beta)^2 \mathbb{I}_{(-\infty,0)}(x-\beta) + \alpha(x-\beta)^2 \mathbb{I}_{(0,\infty)}(x-\beta),$$

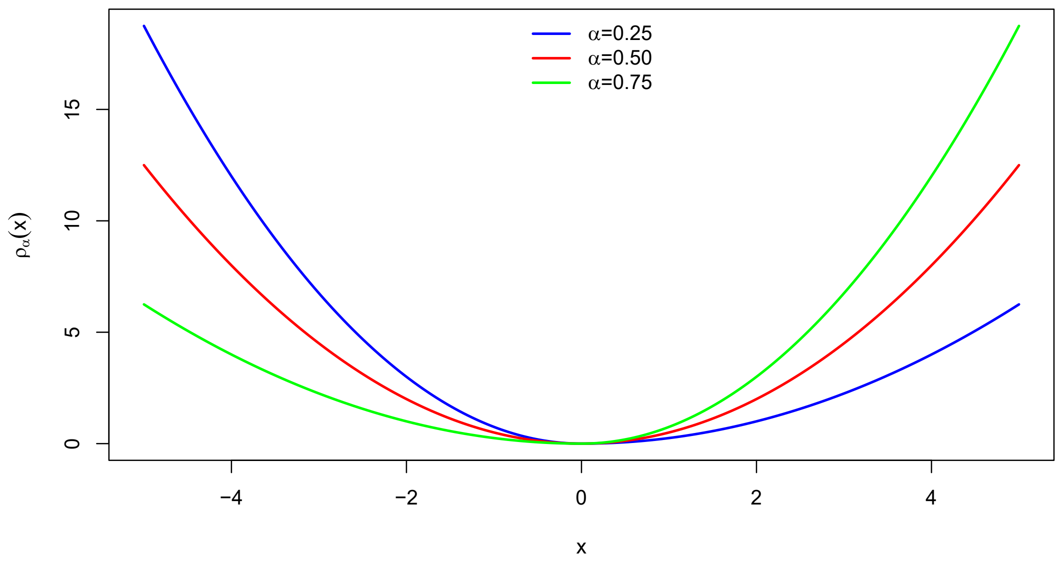
kde $\alpha \in (0,1)$ je parametr, který určuje poměr vah kladných a záporných odchylek od hodnoty β. Tato upravená ztrátová funkce poskytuje flexibilitu v nastavení vah v závislosti na hodnotách x. Pokud α přiblížíme hodnotě 1, kladné odchylky budou mít vyšší váhu než záporné odchylky, a naopak, pokud α přiblížíme hodnotě 0, záporné odchylky budou mít vyšší váhu než kladné odchylky. S pomocí této ztrátové funkce si definujeme expektily.

Definice 4. *Nechť X je náhodná veličina s konečnou střední hodnotou, $\alpha \in (0{,}1)$. Potom*

$$\mu_X(\alpha) = \arg\min_{\beta} \mathsf{E}\ \rho_\alpha(X-\beta) \tag{2.1}$$

nazveme α-tým expektilem *náhodné veličiny X.*

Obrázek 2.1 nám ukazuje, jak vypadá ztrátová funkce ρ_α pro expektily pro konkrétní hodnoty α.



Obrázek 2.1: Ztrátová funkce ρ_α pro konkrétní hodnoty α.

Z definice expektilu (2.1) vyplývá, že hledáme hodnotu β, která minimalizuje očekávanou hodnotu funkce $\rho_\alpha(X-\beta)$. Pokusíme se vyjádřit $\mu_X(\alpha)$. Budeme postupovat podobně jako v důkazu (1).

Nejprve si označíme funkci $M(\beta) = \mathsf{E}\ \rho_\alpha(X-\beta)$. Poté zderivujeme podle β a E si napíšeme v integračním tvaru

$$M'(\beta) = \frac{\partial}{\partial\beta}\mathsf{E}\ \rho_\alpha(X-\beta) = \frac{\partial}{\partial\beta}\int_{-\infty}^{\infty}\rho_\alpha(x-\beta)dF_X(x).$$

Následně použijeme Leibnizovo integrační pravidlo a dopočítáme

$$\begin{aligned}
M'(\beta) &= \frac{\partial}{\partial\beta}\int_{-\infty}^{\infty}\rho_\alpha(x-\beta)dF_X(x) = \int_{-\infty}^{\infty}\frac{\partial}{\partial\beta}\rho_\alpha(x-\beta)dF_X(x) = \\
&= \int_{-\infty}^{\infty} 2(1-\alpha)(\beta-x)\mathbb{I}_{(-\infty,0)}(x-\beta) + 2\alpha(\beta-x)\mathbb{I}_{(0,\infty)}(x-\beta)dF_X(x) = \\
&= 2(1-\alpha)\int_{-\infty}^{\beta}\beta - xdF_X(x) + 2\alpha\int_{\beta}^{\infty}\beta - xdF_X(x) = \\
&= 2(\alpha-1)\int_{-\infty}^{\beta}x-\beta dF_X(x) - 2\alpha\int_{\beta}^{\infty}x-\beta dF_X(x).
\end{aligned}$$

Položíme $M'(\beta) = 0$ a dostáváme

$$(\alpha-1)\int_{-\infty}^{\beta}x-\beta dF_X(x) - \alpha\int_{\beta}^{\infty}x-\beta dF_X(x) = 0. \tag{2.2}$$

Vyjádřeme si $\mu_X(\alpha)$ z rovnosti výše. Nejprve rozepíšeme jednotlivé sčítance integrálů

$$\alpha\int_{\mu_X(\alpha)}^{\infty}x-\mu_X(\alpha)dF_X(x) = \alpha\int_{\mu_X(\alpha)}^{\infty}xdF_X(x) - \alpha\int_{\mu_X(\alpha)}^{\infty}\mu_X(\alpha)dF_X(x).$$

Stejným způsobem upravíme i na druhé straně rovnosti. Dále si převedeme na jednu stranu integrály s integrandem x a na stranu druhou s integrandem $\mu_X(\alpha)$

$$\alpha \int_{\mu_X(\alpha)}^{\infty} \mu_X(\alpha) dF_X(x) + (1-\alpha) \int_{-\infty}^{\mu_X(\alpha)} \mu_X(\alpha) dF_X(x) = \\ = \alpha \int_{\mu_X(\alpha)}^{\infty} x dF_X(x) + (1-\alpha) \int_{-\infty}^{\mu_X(\alpha)} x dF_X(x).$$

Na levé straně rovnosti vytkneme $\mu_X(\alpha)$

$$\mu_X(\alpha)\Big(\alpha \int_{\mu_X(\alpha)}^{\infty} dF_X(x) + (1-\alpha) \int_{-\infty}^{\mu_X(\alpha)} x dF_X(x)\Big) = \\ = \alpha \int_{\mu_X(\alpha)}^{\infty} x dF_X(x) + (1-\alpha) \int_{-\infty}^{\mu_X(\alpha)} dF_X(x).$$

Celkem tedy dostáváme

$$\mu_X(\alpha) = \frac{\alpha \int_{\mu_X(\alpha)}^{\infty} x dF_X(x) + (1-\alpha) \int_{-\infty}^{\mu_X(\alpha)} x dF_X(x)}{\alpha \int_{\mu_X(\alpha)}^{\infty} dF_X(x) + (1-\alpha) \int_{-\infty}^{\mu_X(\alpha)} dF_X(x)}. \tag{2.3}$$

Věta 4 (Newey a Powell, 1987, str. 823)**.** *Nechť X je náhodná veličina s konečnou střední hodnotou a distribuční funkcí $F_X(x)$, $\alpha \in (0,1)$. Pak je splněna následující rovnost*

$$\mu_X(\alpha) - \mathsf{E}\, X = \frac{2\alpha - 1}{1-\alpha} \int_{\mu_X(\alpha)}^{\infty} x - \mu_X(\alpha)\, dF_X(x), \tag{2.4}$$

kde $\mu_X(\alpha)$ je α-tý expektil.

Důkaz. Začneme úpravou levé strany

$$\mu_X(\alpha) - \mathsf{E}\, X = \mu_X(\alpha) - \int_{-\infty}^{\mu_X(\alpha)} x dF_X(x) - \int_{\mu_X(\alpha)}^{\infty} x dF_X(x).$$

Nyní upravme i pravou stranu

$$\frac{2\alpha - 1}{1-\alpha} \int_{\mu_X(\alpha)}^{\infty} x - \mu_X(\alpha) dF_X(x) = \frac{2\alpha - 1}{1-\alpha} \Big(\int_{\mu_X(\alpha)}^{\infty} x dF_X(x) - \int_{\mu_X(\alpha)}^{\infty} \mu_X(\alpha) dF_X(x) \Big).$$

Dále si uděláme rozdíl pravé a levé strany

$$\mu_X(\alpha) - \int_{-\infty}^{\mu_X(\alpha)} x dF_X(x) - \int_{\mu_X(\alpha)}^{\infty} x dF_X(x) - \\ - \frac{2\alpha - 1}{1-\alpha} \Big(\int_{\mu_X(\alpha)}^{\infty} x dF_X(x) - \int_{\mu_X(\alpha)}^{\infty} \mu_X(\alpha) dF_X(x) \Big).$$

Výrazy si rozšíříme a sečteme

$$\frac{\mu_X(\alpha) - \alpha\mu_X(\alpha) - \int_{-\infty}^{\mu_X(\alpha)} x dF_X(x) + \alpha \int_{-\infty}^{\mu_X(\alpha)} x dF_X(x)}{1-\alpha} - \\ - \frac{\alpha \int_{\mu_X(\alpha)}^{\infty} x dF_X(x) - 2\alpha \int_{\mu_X(\alpha)}^{\infty} \mu_X(\alpha) dF_X(x) + \int_{\mu_X(\alpha)}^{\infty} \mu_X(\alpha) dF_X(x)}{1-\alpha}.$$

Vyjádříme si $\alpha \int_{\mu_X(\alpha)}^{\infty} x dF_X(x)$ z (2.2) jako

$$\alpha \int_{\mu_X(\alpha)}^{\infty} x dF_X(x) = \alpha \int_{\mu_X(\alpha)}^{\infty} \mu_X(\alpha) dF_X(x) - (1-\alpha) \int_{-\infty}^{\mu_X(\alpha)} x - \mu_X(\alpha) dF_X(x).$$

Následně odsadíme a sečteme

$$\frac{\mu_X(\alpha) - \alpha\mu_X(\alpha) + \alpha \int_{\mu_X(\alpha)}^{\infty} \mu_X(\alpha) dF_X(x) + \alpha \int_{-\infty}^{\mu_X(\alpha)} \mu_X(\alpha) dF_X(x)}{1-\alpha} - \\ - \frac{\int_{-\infty}^{\mu_X(\alpha)} \mu_X(\alpha) dF_X(x) + \int_{\mu_X(\alpha)}^{\infty} \mu_X(\alpha) dF_X(x)}{1-\alpha}.$$

Spočítáme integrály

$$\frac{\mu_X(\alpha) - \alpha\mu_X(\alpha) + \alpha\mu_X(\alpha) - \alpha\mu_X(\alpha) F_X(\mu_X(\alpha))}{1-\alpha} - \\ - \frac{-\alpha\mu_X(\alpha) F_X(\mu_X(\alpha)) + \mu_X(\alpha) F_X(\mu_X(\alpha)) - \mu_X(\alpha) F_X(\mu_X(\alpha)) + \mu_X(\alpha)}{1-\alpha}.$$

Dostáváme

$$\frac{0}{1-\alpha}.$$

tedy levá a prává strana se rovnají.

□

Z předchozí věty si můžeme všimnout užitečných vlastností expektilů. Pokud do (2.4) dosadíme $\alpha = 0.5$, pak dostáváme

$$\mu_X(0.5) - \mathsf{E}\, X = 0.$$

Tedy 0.5-tý expektil je roven střední hodnotě náhodné veličiny X.

Dále si ukážeme případ, kdy má náhodná veličina X střední hodnotu 0. Analogicky jako v předchozím případě pouze dosadíme $\mathsf{E}\, X = 0$ do (2.4) a máme

$$\mu_X(\alpha) = \frac{2\alpha - 1}{1-\alpha} \int_{\mu_X(\alpha)}^{\infty} x - \mu_X(\alpha) dF_X(x).$$

Věta 5 (Newey a Powell, 1987, str. 823)**.** *Předpokládejme, že $\mathsf{E}\, X$ je konečná, $s > 0$ a $t \in \mathbb{R}$. Pak pro každé $\alpha \in (0,1)$ má $\mu_X(\alpha)$ následující vlastnosti:*

1. *funkce $\mu(\alpha) : (0,1) \to \mathbb{R}$ je ryze rostoucí,*

2. *pro náhodnou veličinu $Y = sX + t$, pak pro α-tý expektil $\mu_Y(\alpha)$ náhodné veličiny Y platí $\mu_Y(\alpha) = s\mu_X(\alpha) + t$.*

Důkaz. Jednotlivé kroky důkazu jsou uvedeny v Newey a Powell (1987), str. 842.

□

Tato věta nám popisuje základní vlastnosti chování α-tého expektilu a jeho chování při lineární transformaci náhodné veličiny.

2.2 Příklady

V této části si ukážeme, jak se hledají expektily konkrétních rozdělení a porovnáme je s kvantily.

Nejprve se podíváme na rovnoměrné rozdělení na intervalu (θ_1,θ_2), tedy mějmě náhodnou veličinu $X \sim R(\theta_1,\theta_2)$ s hustotou $f_X(x) = \frac{1}{b-a}\mathbb{I}_{(\theta_1,\theta_2)}(x)$ a s následující distribuční funkcí

$$F_X(x) = \begin{cases} 0, & \text{pro } x < \theta_1, \\ \frac{x-\theta_1}{\theta_2-\theta_1}, & \text{pro } \theta_1 \leq x \leq \theta_2, \\ 1, & \text{pro } x > \theta_2. \end{cases}$$

Využijeme (2.2) a dosadíme distribuční funkci, respektive hustotu rovnoměrného rozdělení, pro nalezení expektilu

$$\alpha \int_{\mu_X(\alpha)}^{\theta_2} \frac{x - \mu_X(\alpha)}{\theta_2 - \theta_1} dx = -(1-\alpha) \int_{\theta_1}^{\mu_X(\alpha)} \frac{x - \mu_X(\alpha)}{\theta_2 - \theta_1} dx.$$

Vynásobíme obě strany $(\theta_2 - \theta_1)$

$$\alpha \int_{\mu_X(\alpha)}^{\theta_2} x - \mu_X(\alpha) dx = -(1-\alpha) \int_{\theta_1}^{\mu_X(\alpha)} x - \mu_X(\alpha) dx.$$

Vypočítáme integrály a zjednodušíme

$$\frac{\alpha}{2}(\theta_2^2 - 2\theta_2\mu_X(\alpha) + \mu_X(\alpha)^2) = \frac{\alpha-1}{2}(-\theta_1^2 + 2\theta_1\mu_X(\alpha) - \mu_X^2(\alpha)),$$

$$\alpha(\theta_2 - \mu_X(\alpha))^2 = (1-\alpha)(\mu_X(\alpha) - \theta_1)^2.$$

Obě strany rovnosti jsou nezáporné, můžeme je tedy umocnit na $\frac{1}{2}$

$$\sqrt{\alpha}(\theta_2 - \mu_X(\alpha)) = \sqrt{1-\alpha}(\mu_X(\alpha) - \theta_1).$$

Vyjádříme si $\mu_X(\alpha)$ jako

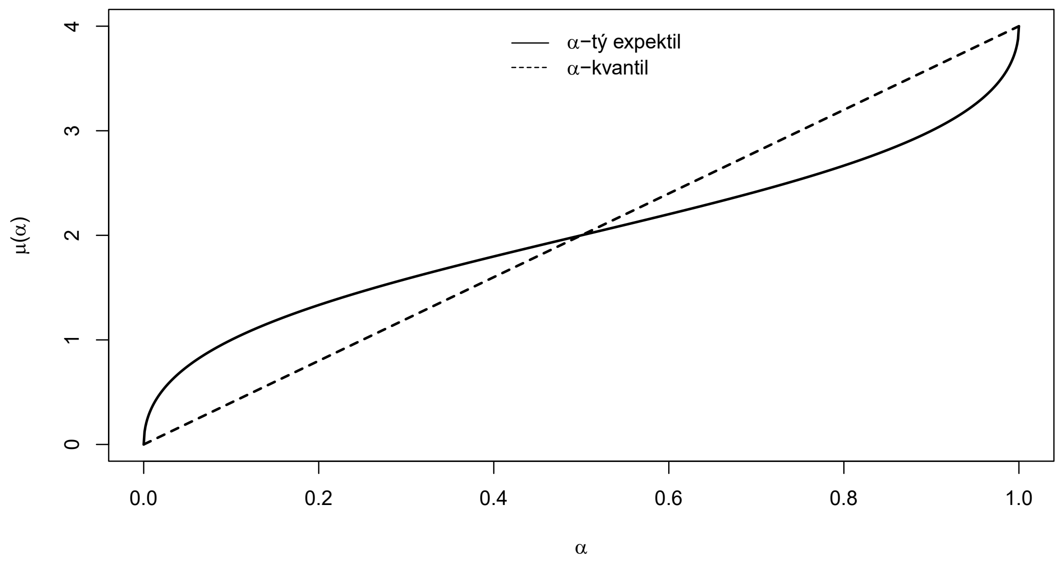
$$\mu_X(\alpha) = \frac{\sqrt{\alpha}\theta_2 + \sqrt{1-\alpha}\theta_1}{\sqrt{\alpha} + \sqrt{1-\alpha}}.$$

Můžeme si porovnat expektil s kvantilovou funkcí. Víme, že distribuční funkce je na intervalu (a,b) spojitá a rostoucí, tedy její kvantilová funkce bude rovna inverzní funkci distribuční funkce, která má tento tvar

$$u_X(\alpha) = \theta_2\alpha + \theta_1(1-\alpha),$$

pro $\alpha \in (0,1)$.

Můžeme si všimnout, že α-tý- expektil a α-kvantil se rovnají pro hodnotu $\alpha = \frac{1}{2}$ (viz. Obrázek 2.2). Toto platí nejen pro rovnoměrné rozdělení, ale i pro všechna symetrická rozdělení, kdy se medián rovná střední hodnotě.



Obrázek 2.2: Porovnání α-tého expektilu a α-kvantilu pro $R(0{,}4)$.

Dalším příkladem, který si rozebereme bude exponenciální rozdělení s parametrem $\lambda > 0$, $Exp(\lambda)$. Příklad je převzat z Bellini a Di Bernardino (2017) s tím, že si některé kroky detailněji popíšeme.

Víme, že pokud X je z rozdělení $Exp(\lambda)$, pak hustota tohoto rozdělení je $f_X(x) = \lambda e^{-\lambda x}, x \geq 0$, se střední hodnotou $\mathsf{E}\ X = \frac{1}{\lambda}$ a distribuční funkcí $F_X(x) = 1 - e^{-\lambda x}$. Pro výpočet α-tého expektilu využijeme rovnost (2.4)

$$\mu_X(\alpha) - \mathsf{E}\ X = \frac{2\alpha - 1}{1 - \alpha} \int_{\mu_X(\alpha)}^{\infty} x - \mu_X(\alpha) dF_X(x).$$

Spočítáme integrál

$$\int_{\mu_X(\alpha)}^{\infty} x - \mu_X(\alpha) dF_X(x) = \int_{\mu_X(\alpha)}^{\infty} x dF_X(x) - \mu_X(\alpha) \int_{\mu_X(\alpha)}^{\infty} dF_X(x).$$

Použijeme metodu per partes na první integrál

$$\begin{aligned}\int_{\mu_X(\alpha)}^{\infty} x dF_X(x) &= \int_{\mu_X(\alpha)}^{\infty} x f_X(x) dx = \int_{\mu_X(\alpha)}^{\infty} x(\lambda e^{-\lambda x}) dx = \\ &= \Big[-x e^{-\lambda x}\Big]_{\mu_X(\alpha)}^{\infty} + \int_{\mu_X(\alpha)}^{\infty} e^{-\lambda x} dx = \mu_X(\alpha) e^{-\lambda \mu_X(\alpha)} - \frac{e^{-\lambda \mu_X(\alpha)}}{\lambda} = \\ &= \frac{\lambda \mu_X(\alpha) + 1}{\lambda} e^{-\lambda \mu_X(\alpha)}.\end{aligned}$$

Snadno dopočítáme i druhý integrál

$$\int_{\mu_X(\alpha)}^{\infty} dF_X(x) = \Big[F_X(x)\Big]_{\mu_X(\alpha)}^{\infty} = 1 - F_X(\mu_X(\alpha)) = -e^{-\lambda \mu_X(\alpha)}.$$

Celkem po dosazení dostáváme

$$\mu_X(\alpha) - \frac{1}{\lambda} = \frac{2\alpha - 1}{1 - \alpha} \cdot \frac{\Big(\lambda \mu_X(\alpha) + 1 - \lambda \mu_X(\alpha)\Big) e^{-\lambda \mu_X(\alpha)}}{\lambda}.$$

Obě strany rovnice vynásobíme $\lambda e^{\lambda \mu_X(\alpha)-1}$

$$(\lambda \mu_X(\alpha) - 1)e^{\lambda \mu_X(\alpha)-1} = \frac{2\alpha - 1}{1-\alpha} \cdot e^{-1}.$$

Použijeme substituci $z = \lambda \mu_X(\alpha) - 1$ a dostáváme rovnost

$$ze^z = \frac{2\alpha - 1}{(1-\alpha)e}.$$

Jsme v situaci, kdy máme $ze^z = y$. Řešením je $z = W(y)$, kde W je Lambertova funkce. V našem případě $\frac{2\alpha-1}{(1-\alpha)e} = -\frac{2}{e} + \frac{1}{(1-\alpha)e} > -\frac{1}{e}$, pro $\alpha \in (0,1)$. Můžeme tedy použít Lambertovu W funkci pro nalezení řešení

$$z = W\left(\frac{2\alpha - 1}{(1-\alpha)e}\right).$$

Provedeme zpětnou substituci a dostáváme

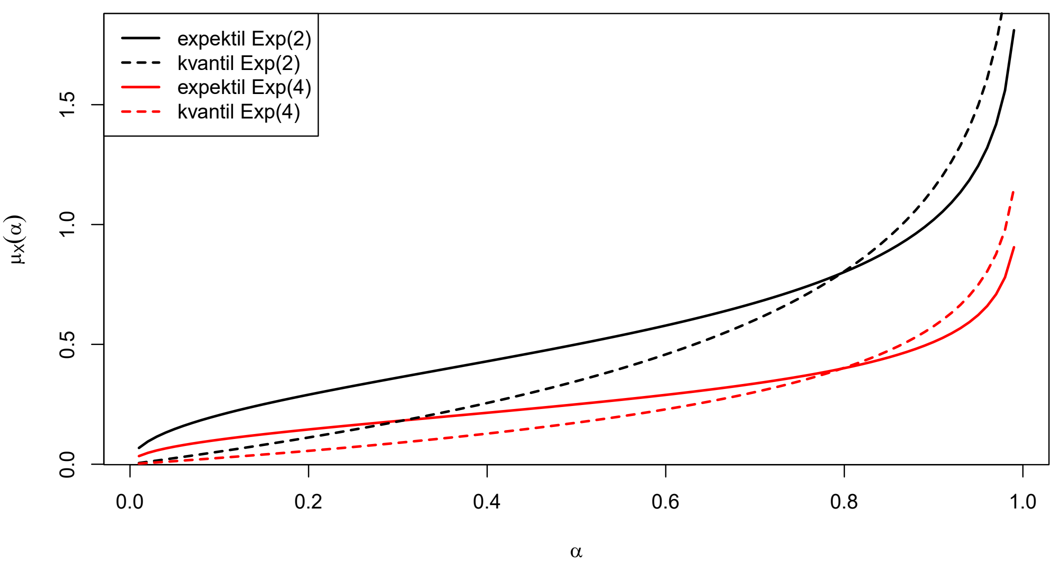
$$\mu_X(\alpha) = \frac{1}{\lambda}\left(W\left(\frac{2\alpha - 1}{(1-\alpha)e}\right) + 1\right),$$

pro $\alpha \in (0,1)$. Z přechozí kapitoly víme, že kvantilová funkce má následující tvar

$$F_X^{-1}(\alpha) = \frac{-\log(1-\alpha)}{\lambda},$$

pro $\alpha \in (0,1)$. Grafické porovnání expektilu a kvantilu pro exponenciální rozdělení s parametry $\lambda = 2$ a $\lambda = 4$ můžeme vidět na obrázku 2.3.

Poznámka. Podrobnosti Lambertovy W funkce jsou k nahlédnutí 3.3 v appendixu.



Obrázek 2.3: Porovnání expektilu a kvantilu pro exponenciální rozdělení s parametry 2 a 4.

2.3 Odhady expektilu

Pro definování výběrového expektilu opět nahradíme ztrátovou funkci v rovnosti (1.4) funkcí

$$\rho_\alpha(x) \equiv |\alpha - \mathbb{I}_{x<0}| \cdot x^2,$$

podobně, jako když jsme si definovali expektily.

Definice 5. *Nechť* $X_1,...,X_n$ *je náhodný výběr z rozdělení* F_X*,* $\alpha \in (0,1)$*, pak*

$$\widehat{\mu}_n(\alpha) = \arg\min_\beta \frac{1}{n} \sum_{i=1}^n \rho_\alpha(X_i - \beta)$$

nazveme α-tým výběrovým expektilem*.*

Podobně jako u výběrového α-kvantilu je i výběrový α-expektil jednoznačně určen. Ukážeme si větu, která nám vyjadřuje konzistenci odhadu výběrového expektilu.

Věta 6. *(Holzmann a Klar, 2016, str.4) Nechť* $X_1, X_2, ..., X_n$ *je náhodný výběr z rozdělení* $\boldsymbol{X}$ *s konečnou střední hodnotou a spojitou rostoucí distribuční funkcí. Potom pro* $\alpha_1, \alpha_2, \alpha \in (0,1)$*,* $\alpha_1 < \alpha_2$ *platí*

$$\sup_{\alpha_1 \le \alpha \le \alpha_2} |\widehat{\mu}_n(\alpha) - \mu_X(\alpha)| \xrightarrow[n\longrightarrow\infty]{P} 0.$$

Důkaz. Podrobné kroky důkazu jsou uvedeny v Holzmann a Klar (2016), str.12. □

Expektily můžeme podobně jako kvantily odhadovat parametricky. Postup je velice podobný právě parametrickému odhadu kvantilu. V parametrickém modelu předpokládáme, že $\mu_X(\alpha) = h(\alpha,\theta_X)$. Pro odhad parametru θ_X a dostáváme

$$\tilde{\mu}_n(\alpha) = h(\alpha,\widehat{\theta}_n),$$

kde $\widehat{\theta}_n$ je konzistentní odhad pomocí momentové metody. Pro $Exp(\lambda)$ jsme v si odvodili, že

$$\mu_X(\alpha) = \frac{1}{\lambda}\left(W\left(\frac{2\alpha - 1}{(1-\alpha)e}\right)+1\right),$$

pro $\alpha \in (0,1)$. Víme, že $\overline{X}_n$ je konzistentním odhadem $\mathsf{E}\ X_i = \frac{1}{\lambda}$. Celkem tedy dostáváme

$$\tilde{\mu}_n(\alpha) = \overline{X}_n\left(W\left(\frac{2\alpha - 1}{(1-\alpha)e}\right)+1\right),$$

pro $\alpha \in (0,1)$.

3. Praktická část

Nejprve se podrobně podíváme na odhad výběrového expektilu a poté se podíváme na parametrický odhad expektilu. Vše bude ilustrováno na nasimulovaných realizacích z exponenciálního rozdělení a doplněno grafickým srovnáním pro různé rozsahy realizace. Nakonec obě metody porovnáme. Veškeré simulace a výpočty byly prováděny pomocí statistického softwaru `R` (R Core Team, 2023).

3.1 Neparametrický odhad expektilu

Ukážeme si to na příkladu s exponenciálním rozdělením. Na základě nasimulovaných dat budeme zkoumat, jak se mění odhady střední kvadratické chyby, $\widehat{MSE}$, vychýlení, $\widehat{bias}$, a standardní odchylky, $\widehat{sd}$, pro různé rozsahy náhodného výběru.

Pro vyhodnocení a porovnání odhadů pro různé rozsahy budeme muset zvolit hodnoty α, ve kterých se bude porovnávat rozdíl mezi teoretickou hodnotou expektilu, $\mu_X(\alpha)$ a odhadu expektilu, $\widehat{\mu}_n(\alpha)$. Zvolme množinu bodů $\alpha_i = 0{,}05i$, pro $i \in \{1,...,19\}$, kde $\alpha_{19} = 0{,}95$ Máme celkem $m = 19$ dělících bodů.

Víme, že pro odhad $\widehat{\theta}_n$ parametru θ_X platí

$$bias(\theta) = \mathsf{E}\,(\widehat{\theta}_n - \theta_X).$$

Pro empirický odhad vychýlení v našem případě dostáváme

$$\widehat{bias}(\theta) = \frac{1}{m}\sum_{i=1}^{m}(\widehat{\mu}_n(\alpha_i) - \mu_X(\alpha_i)).$$

Pro jednodušší zápis si označíme $e_i = \widehat{\mu}_n(\alpha_i) - \mu_X(\alpha_i)$ a dosadíme pro odhad vychýlení

$$\widehat{bias}(e) = \frac{1}{m}\sum_{i=1}^{m} e_i = \frac{1}{19}\sum_{i=1}^{19} e_i = \overline{e_m}.$$

Analogickým postupem jako pro odhad $bias$ odhadneme sd

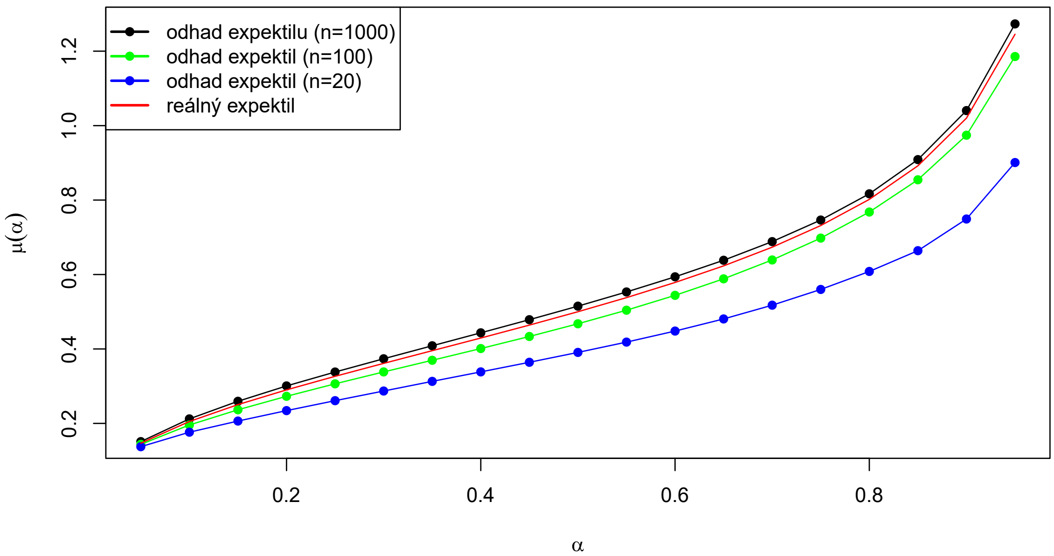
$$sd = \sqrt{\frac{1}{m-1}\sum_{i=1}^{m}(e_i - \overline{e_m})^2} = \sqrt{\frac{1}{m-1}\sum_{i=1}^{m}(e_i - \widehat{bias})^2}.$$

Odhad střední kvadratické odchylky můžeme spočítat takto

$$\widehat{MSE} = \widehat{sd}^2 + \widehat{bias}^2.$$

Nyní vše aplikujeme na následující situaci. Nechť máme 3 náhodné výběry $\boldsymbol{X} = (X_1, ..., X_n)$ z rozdělení $Exp(2)$ s různými rozsahy. První výběr má rozsah $n = 20$, druhý výběr má rozsah $n = 100$ a třetí výběr má rozsah $n = 500$. Tyto 3 výběry si nasimulujeme ve statistickém programu R. Hodnoty expektilu pro $\alpha \in \{0.05, 0.1, ..., 0.95\}$ jsou graficky zobrazeny na obrázku 3.1, kde je můžeme graficky porovnat s hodnotami skutečného expektilu.

Na základě odvození výše si spočítáme pomocí nasimulovaných dat $\widehat{bias}$, $\widehat{sd}$ a $\widehat{MSE}$. Jednotlivé výpočty jsme provedli pomocí statistického softwaru `R` a výsledky jsou uvedeny v tabulce 3.1.



Obrázek 3.1: Porovnání skutečného expektilu s nasimulovaným výběrovým expektilem pro konkrétní rozsahy n.

$\widehat{\mu}_n(\alpha)$	$\widehat{sd}$	$\widehat{bias}$	$\widehat{MSE}$
$n = 20$	0,0857	-0,1272	0,0235
$n = 100$	0,0130	-0,0291	0,0010
$n = 1000$	0,0049	0,0140	0,0002

Tabulka 3.1: Odhady $\widehat{sd}$, $\widehat{bias}$ a $\widehat{MSE}$ výběrového expektilu pro konkrétní rozsahy výběru.

3.2 Parametrický odhad expektilu

V této části budeme zkoumat přesnost parametrického odhadu expektilu na příkladě exponenciálního rozdělení $Exp(\lambda)$. Využijeme odhad parametrů pomocí momentové metody, kterou jsme si předvedli v předchozí kapitole. Máme

$$\mu_X(\alpha) = \frac{1}{\lambda}\left(W\left(\frac{2\alpha-1}{(1-\alpha)e}\right)+1\right),$$

pro $\alpha \in (0,1)$. Pro parametrický odhad momentovou metodou získáváme

$$\tilde{\mu}_n(\alpha) = \overline{X}_n\left(W\left(\frac{2\alpha-1}{(1-\alpha)e}\right)+1\right),$$

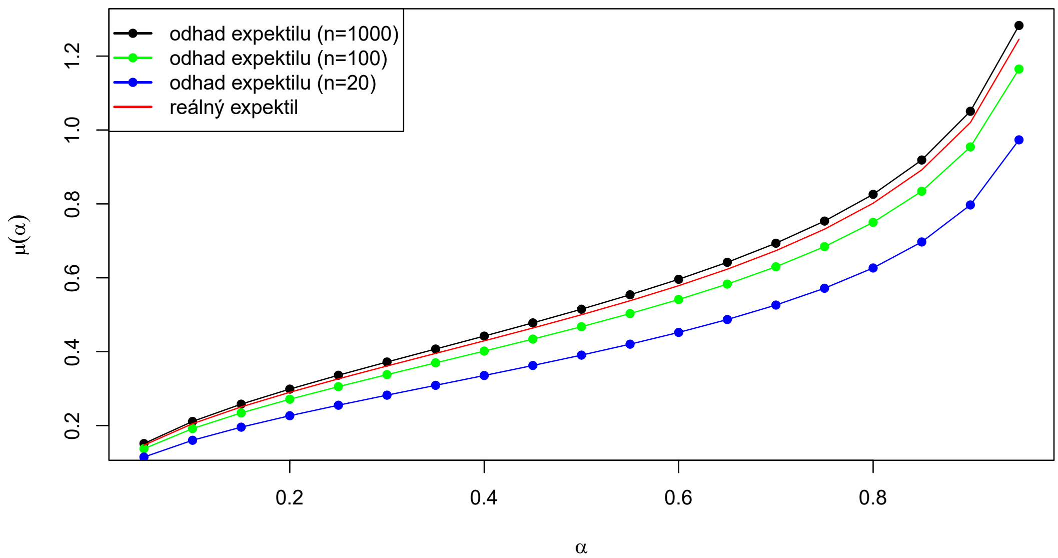
pro $\alpha \in (0,1)$

Expektily budeme počítat podobně jako v předchozí části. Vezmeme si stejnou množinu bodů $\alpha_i = 0.05i$, pro $i \in \{1, 2, ..., 19\}$, ve kterých budeme hodnoty odhadnutých expektilů porovnávat se skutečnými hodnotami expektilů.

Nechť máme náhodný $\boldsymbol{X} = (X_1, ..., X_n)$ s rozsahem $n \in \{20,100,1000\}$ z rozdělení $Exp(2)$. Využijeme odvození výše, do kterého budeme dosazovat vypočítané $\overline{X}_n$.

Pro získání odhadů $\widehat{bias}$, $\widehat{sd}$ a $\widehat{MSE}$ budeme postupovat stejně jako v předchozím příkladě, protože máme stejnou množinu bodů α_i, pro které nás budou zajímat hodnoty expektilu.

Grafické porovnání odhadů pro konkrétní rozsahy jsou k nahlednutí v obrázku 3.2. Výsledky zaokrouhlené na 4 desetinná místa můžeme vidět v tabulce 3.2.



Obrázek 3.2: Porovnání teoretického expektilu s nasimulovaným parametrickým odhadem expektilu pro konkrétní rozsahy n.

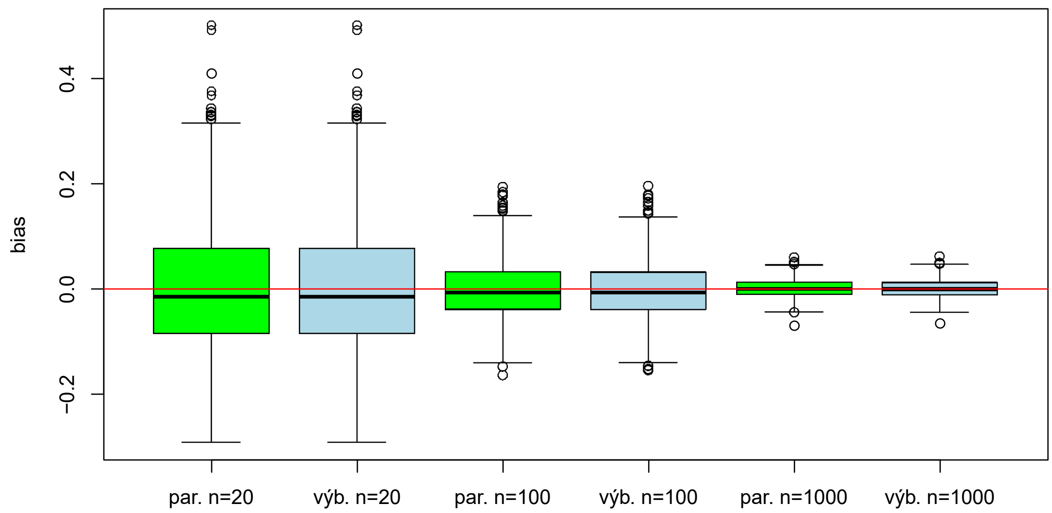
$\tilde{\mu}_n(\alpha)$	$\widehat{sd}$	$\widehat{bias}$	$\widehat{MSE}$
$n = 20$	0,0633	-0,1205	0,0185
$n = 100$	0,0187	-0,0356	0,0016
$n = 1000$	0,0087	0,0165	0,0003

Tabulka 3.2: Odhady $\widehat{sd}$, $\widehat{bias}$ a $\widehat{MSE}$ parametrického odhadu expektilu pro konkrétní rozsahy výběru.

3.3 Porovnání obou metod

Abychom mohli lépe porovnat obě metody odhadování expektilů, tak nám nebude stačit pouze jedna simulace, ale nasimulujeme si 1000 náhodných výběrů s rozsahy $n = 20$, $n = 100$ a $n = 1000$. Budeme porovnávat odhady $bias$ a MSE obou metod.

Nejprve se podíváme na odhad $bias$. Na obrázku 3.3 si můžeme všimnout, že se odhad vychýlení chová velmi podobně jak pro parametrickou metodu, tak pro výběrový odhad expektilu na námi nasimulovaných náhodných výběrech. Také můžeme vidět, jak se vychýlení snižuje s přibývajícím rozsahem náhodného výběru.



Obrázek 3.3: Porovnání vychýlení parametrického odhadu expektilu a výběrového expektilu pro konkrétní rozsahy n.

Na rozdíl od $\widehat{bias}$ si $\widehat{MSE}$ porovnáme pomocí průměru. Z nasimulovaných výběrů nám vyšlo, že pro všechny rozsahy byl průměrný odhad MSE nižší u parametrického odhadu expektilu momentovou metodou než u výběrového expektilu, viz tabulka 3.3. Nižší hodnoty $\widehat{MSE}$ indikují přesnější odhady. Můžeme tedy říci, že parametrická metoda přesněji odhaduje expektil než výběrový expektil pro všechny rozsahy na námi nasimulovaných výběrech. Můžeme si z výsledků také všimnout, že se zvyšujícím n se snižuje $\widehat{MSE}$ pro obě metody.

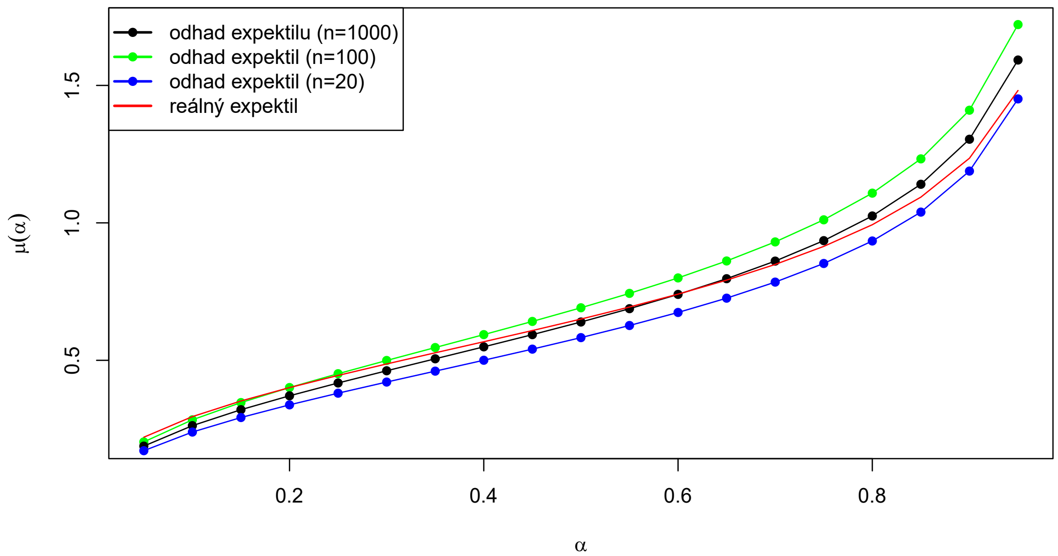
	$\widehat{\mu}_X(\alpha)$	$\tilde{\mu}_X(\alpha)$
$n = 20$	0,0538	0,0500
$n = 100$	0,0113	0,0102
$n = 1000$	0,0011	0,0010

Tabulka 3.3: Porovnání průměrných $\widehat{MSE}$ pro výběrový a parametrický odhad expektilu.

Musíme zmínit, že i když parametrická metoda vyšla lépe na nasimulovaných výběrech, tak je náchylná na typ rozdělení. Když bychom mylně předpokládali, že náhodný výběr je z rozdělení $Exp(\lambda)$, ale ve skutečnosti by náhodný výběr pocházel z rozdělení gamma s parametry $\Gamma(\frac{13}{10},\lambda)$, tak bychom velmi pravděpodobně dostali nepřesný odhad. V tabulce 3.4 můžeme vidět, jak nám dopadl odhad na nasimulovaném rozdělení a na obrázku 3.4 se můžeme podívat, jak se liší odhad expektilu od teoretického expektilu pro rozdělení $\Gamma(\frac{13}{10},2)$.

$\tilde{\mu}_n(\alpha)$	$\widehat{sd}$	$\widehat{bias}$	$\widehat{MSE}$
$n = 20$	0,0095	0,0023	0,0037
$n = 100$	0,0687	0,0594	0,0082
$n = 1000$	0,0399	-0,0602	0,0015

Tabulka 3.4: $\widehat{sd}$, $\widehat{bias}$ a $\widehat{MSE}$ pro konkrétní rozsahy výběru.



Obrázek 3.4: Porovnání skutečného expektilu s nasimulovaným výběrovým expektilem pro konkrétní rozsahy n.

Závěr

Bakalářská práce se věnovala základním vlastnostem expektilů.

Nejprve byly shrnuty základní vlastnosti kvantilů. Dále jsme si zadefinovali a shrnuli základní vlastnosti výběrového α-kvantilu. Ukázali jsme si parametrický odhad kvantilů pomocí momentové metody. Metoda byla ilustrovaná na příkladě rovnoměrného a exponenciálního rozdělení.

V další části jsme se podrobně zabývali expektily a jejich definicí. Představili jsme jednotlivá odvození a prohlédli si ilustrativní příklady expektilů na známých rozděleních. Dále jsme se zaměřili na výběrový expektil a popsali parametrický odhad expektilu.

Dalším přínosem je praktická část, kde jsme odhadovali vychýlení, standardní odchylku a také střední čtvercovou chybu výběrového α-tého expektilu pro náhodné výběry o různých rozsazích. Dále jsme provedli odhady pro parametrický odhad pomocí momentové metody a následně tyto dvě metody porovnali.

Seznam použité literatury

Bellini, F. a Di Bernardino, E. (2017). Risk management with expectiles. *The European Journal of Finance*, **23**(6), 487–506.

Holzmann, H. a Klar, B. (2016). Expectile asymptotics. *Electronic Journal of Statistics*.

Kulich, M. (2022). NMSA331 Matematická statistika 1. `https://www.karlin.mff.cuni.cz/~komarek/vyuka/2022_23/nmsa331/ms1.pdf`.

Newey, W. K. a Powell, J. L. (1987). Asymmetric least squares estimation and testing. *Econometrica: Journal of the Econometric Society*, pages 819–847.

Omelka, M. (2023). Course notes for NMST424. Dostupné z: `http://www.karlin.mff.cuni.cz/~omelka/Soubory/nmst424/nmst424_course-notes.pdf`.

R Core Team (2023). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. URL `https://www.r-project.org/`.

Veberič, D. (2010). Having fun with Lambert W (x) function. `https://arxiv.org/pdf/1003.1628.pdf`.

Seznam obrázků

Seznam tabulek

Appendix

Definice a popis vlastností Lambertovy funkce je převzata z Veberič (2010).

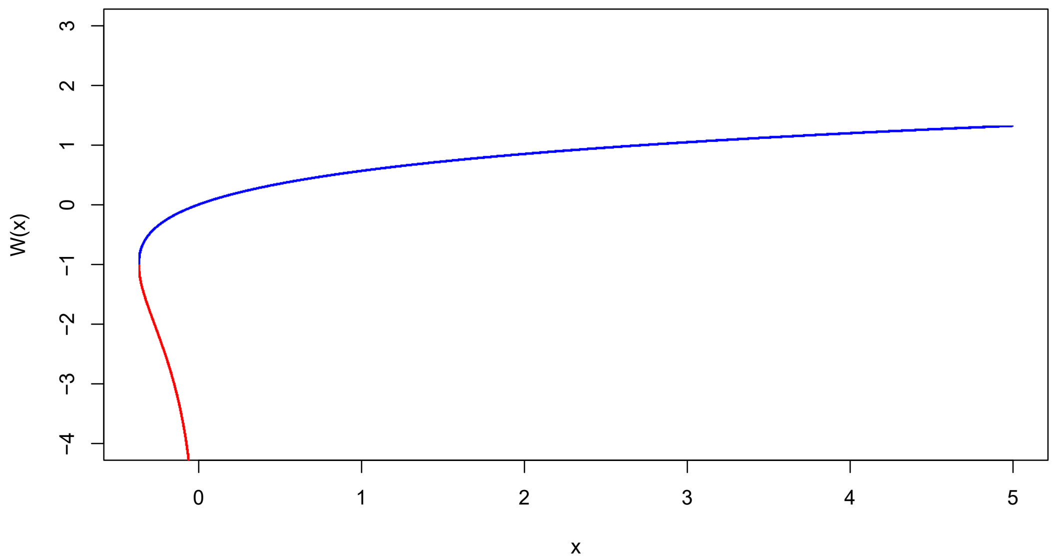
Definice 6. Lambertova $W(x)$ funkce *je definována jako inverzní funkce rovnice*

$$ye^y = x,$$

a její řešení je dáno vztahem

$$y = W(x).$$

Můžeme si všimnout, že zobrazení xe^x není prosté. Pokud ale rozdělíme xe^x v bodě $x = -\frac{1}{e}$ na 2 intervaly, pak na těchto 2 intervalech prosté je. Lambertova W funkce má tedy 2 větve, první se označuje $W_{-1}(x)$ a je definována, pro $x \in [-e^{-1},0]$. Druhá větev se označuje $W_0(x)$ a je definována pro $x \in [-e^{-1},\infty]$. Chování funkce $W(x)$ můžeme vidět na obrázku 3.5.



Obrázek 3.5: Červeně je zobrazena větev $W_{-1}(x)$ a modrou barvou větev $W_0(x)$.

A. Přílohy

Elektronická příloha obsahuje simulační studii použitou v bakalářské práci. Byla vytvořená ve statistickém softwaru `R`.