

UNIVERZITA PALACKÉHO V OLOMOUCI
PŘÍRODOVĚDECKÁ FAKULTA
KATEDRA MATEMATICKÉ ANALÝZY A APLIKACÍ MATEMATIKY

DIPLOMOVÁ PRÁCE

Nelineární metody analýzy EKG záznamů



Vedoucí diplomové práce: **RNDr. Tomáš Fürst, Ph.D.**

Vypracovala: **Bc. Markéta Mraznicová**

Studijní program: N1101 Matematika

Studijní obor Matematika a její aplikace

Forma studia: prezenční

Rok odevzdání: 2015

BIBLIOGRAFICKÁ IDENTIFIKACE

Autor: Bc. Markéta Mraznicová

Název práce: Nelineární metody analýzy EKG záznamů

Typ práce: Diplomová práce

Pracoviště:

Katedra matematické analýzy a aplikací matematiky

Vedoucí práce: RNDr. Tomáš Fürst, Ph.D.

Rok obhajoby práce: 2015

Abstrakt:

Práce se zabývá dvěma významnými nelineárními metodami, které se využívají k analýze variability srdeční frekvence. V první části byla podrobně popsána metoda víceškálové entropie (Multiscale Entropy Analysis), jejíž výsledkem je závislost hodnot MSE na škálovém faktoru. V dalších kapitolách byla popsána metoda DFA (Detrended Fluctuation Analysis), u které je předmětem zájmu tzv. škálový koeficient, podle kterého lze signály rozdělit do několika skupin. Obě metody byly následně aplikovány na dva datové soubory reálných EKG signálů (pacientů se synkopou a bez synkopy). Veškeré výsledky byly otestovány pomocí statistických nástrojů a porovnány s časovou a spektrální analýzou stejných EKG signálů.

Klíčová slova:

EKG, víceškálová entropie, MSE, DFA, synkopa.

Počet stran: 54

Počet příloh: 14

Jazyk: český

BIBLIOGRAPHICAL IDENTIFICATION

Author: Markéta Mraznicová

Title: Nonlinear methods of ECG analysis

Type of thesis: Master's

Department:

Department of Mathematical Analysis and Application of Mathematics

Supervisor: RNDr. Tomáš Fürst, Ph.D.

The year of presentation: 2015

Abstract:

The thesis deals with two significant nonlinear methods which are used for heart rate variability analysis. In the first part, the method of Multiscale Entropy (MSE) analysis is described, the result of which is the dependency of MSE values on the scale factor. In the next chapter the method of Detrended Fluctuation Analysis is introduced which gives a scale coefficient. According to this coefficient, signals can be divided into several groups. Both methods are applied to two data sets of real ECG signals (patients with syncope and without syncope). The results are tested by standard statistical tools and compared with time domain and frequency domain analysis of the same ECG signals.

Key words:

ECG, Multiscale Entropy Analysis, Detrended Fluctuation Analysis, syncope.

Number of pages: 54

Number of appendices: 14

Language: Czech

Prohlášení

Prohlašuji, že jsem diplomovou práci zpracovala samostatně pod vedením Tomáše Fürsta a všechny použité zdroje jsem uvedla v seznamu literatury na konci práce.

V Olomouci dne 8. dubna 2015.

Obsah

Úvod	7
1 Multiscale entropy analysis (MSE)	9
1.1 Entropie a její původ	9
1.2 Entropie u biologických signálů	10
1.3 Vzorková entropie	11
1.3.1 Výpočet přibližné entropie	13
1.3.2 Výpočet vzorkové entropie	15
1.3.3 Implementace funkce vzorkové entropie v programu Matlab	18
1.4 Víceškálová entropie – MSE	21
1.4.1 Výpočet víceškálové entropie	21
1.4.2 Implementace funkce víceškálové entropie v programu Matlab	22
1.5 Aplikace víceškálové entropie na analýzu EKG záznamů	26
1.5.1 Popis datových souborů	26
1.5.2 Výsledné hodnoty MSE reálných EKG signálů	27
1.5.3 Statistické testování hodnot MSE reálných EKG signálů .	38
2 Detrended fluctuation analysis (DFA)	44
2.1 Výpočet DFA	44
2.2 Implementace funkce DFA v programu Matlab	46
2.3 Výsledné hodnoty DFA reálných EKG záznamů	48
2.4 Statistické testování hodnot DFA reálných EKG signálů	53
3 Srovnání metod MSE a DFA s časovou a spektrální analýzou posloupnosti RR intervalů	55
3.1 Časová analýza	55
3.2 Spektrální analýza	57
Závěr	59
Přílohy	61
Literatura	75

Poděkování

Mé poděkování patří především vedoucímu mé diplomové práce Tomáši Fürstovi za vstřícný přístup, odbornou pomoc a faktické rady a připomínky, které mi velmi pomohly při zpracování celé práce.

Úvod

Lidské srdce je jedním z nejdůležitějších orgánů lidského těla, o jehož pochopení se pokoušejí lékaři a vědci po celém světě. Na povrchu tohoto životně důležitého svalu vzniká elektrický signál, který lze zachytit přístrojem zvaným elektrokardiograf a pořídit záznam zvaný elektrokardiogram (zkráceně EKG). Na EKG lze z pohledu teorie signálů nahlížet jako na časovou řadu, na kterou lze aplikovat množství matematických postupů.

Hlavním cílem této práce je aplikovat dvě matematické nelineární metody na reálná data pacientů s tzv. synkopou, které budu srovnávat s kontrolní skupinou zdravých lidí. Synkopa je náhlá krátkodobá ztráta vědomí způsobena nedostatkem okysličené krve v mozku. Může nastat zcela bez varování i u zdravých jedinců a nemusí mít v budoucím životě žádné následky. EKG záznam těchto pacientů je přirozeně předmětem zájmu lékařů, protože pravděpodobně obsahuje informaci o tom, jestli synkopa signalizuje počínající potíže se srdcem, které mohou mít v budoucnu vážné kardiologické následky, jako například infarkt myokardu. Nastává ovšem otázka, jakým způsobem tuto informaci z EKG záznamu nejlépe vyčíst.

V posledních dvaceti letech se experimentuje s mnoha lineárními i nelineárními charakteristikami variability srdeční frekvence a mnoho výzkumníků se domnívá, že vhodné matematické metody zpracování EKG signálů mohou přinést daleko více informací než prosté sledování středních hodnot či rozptylu délky po sobě jdoucích RR intervalů, případně vizuální inspekce EKG záznamu.

Proto jsem se v rámci bakalářské práce seznámila s touto problematikou a v rámci této práce se budu snažit již získané znalosti rozšířit a aplikovat na stejný vzorek dat, což je jeden z cílů a přínosů mé diplomové práce

Analýzu variability srdeční frekvence lze rozdělit na dvě velké skupiny – na lineární a nelineární metody. V předchozí práci jsem se zabývala první oblastí, tedy lineárními metodami – speciálně časovou analýzou řady RR intervalů a spektrální analýzou reprezentovanou tzv. Lombovým periodogramem. Tato metoda tvořila stěžejní část práce. V této práci se budu věnovat modernějšímu přístupu k analýze

dynamických systémů, který představují nelineární metody, speciálně víceškálovou entropií (Multiscale Entropy Analysis – MSE) a metodou DFA (Detrended Fluctuation Analysis).

Dalšími cíli práce je podrobně obě metody popsat z matematického hlediska a vytvořit algoritmy v programu Matlab. Na jednoduchých příkladech signálů ukáží použitelnost těchto funkcí s různými kombinacemi parametrů, abych se následně mohla věnovat nejdůležitější části této práce – aplikaci a srovnání metod MSE a DFA na reálných EKG signálech se synkopou a zdravých EKG signálech. Mým cílem bude zjistit, zda se skupina zdravých pacientů významně liší od skupiny pacientů, kteří prodělali synkopu. Případně najít, v čem se liší.

Dalším cílem bude zjistit, zda se dají obě výše uvedené metody (MSE a DFA) použít ke stratifikaci skupiny pacientů se synkopou, přičemž optimální by bylo, pokud bychom mohli obě metody použít k odlišení synkop s kardiologickou etiologií od synkop způsobených neurologicky či prostým snížením tlaku.

1. Multiscale entropy analysis (MSE)

V první části této kapitoly se zaměřím na pochopení pojmu informační entropie, který má motivaci ve fyzice a využívá se také v matematice. V dalších částech se budu věnovat entropii u biologických signálů. Vysvětlím matematické základy a definuji více druhů entropie, abych se v závěru mohla věnovat stěžejnímu cíli této kapitoly – víceškálové entropii (z anglického Multiscale Entropy, dále budeme používat zkratku MSE). Podrobně popíši její odvození a budu ji aplikovat na analýzu EKG záznamů.

1.1. Entropie a její původ

Entropie je velmi důležitý pojem vyskytující se v mnoha oblastech vědy, např. ve fyzice, teorii pravděpodobnosti, informatice a samozřejmě i v matematice, viz [11]. Původ slova „entropie“ pochází z řeckého termínu, který v překladu znamená „směrem k“ či „transformace“.

Poprvé byl tento termín použit v souvislosti s termodynamikou německým fyzikem a matematikem Rudolfem Clausiem roku 1865, viz [11]. Pomocí tohoto termínu je možné interpretovat druhý termodynamický zákon a vyjadřuje fakt, že některé procesy jsou nevratné. Princip růstu entropie pak postuluje, že celková entropie izolovaného systému nemůže nikdy klesnout. Pouze stoupá ke své maximální hodnotě. Jakmile dosáhne své maximální hodnoty, zůstává entropie konstantní, viz [10].

Tímto pojmem se později inspiroval rakouský fyzik a filozof Ludwig Boltzmann, který roku 1870 definoval entropii ze statistického pohledu jakožto výsledek zkoumání statistického chování mikroskopických částic systému. Boltzmann po mnoha letech ukázal, že termodynamická a statistická entropie (doposud zkoumané zcela odděleně) se od sebe liší pouze jistou konstantou, která je od té doby známa jako *Boltzmannova konstanta*, viz [10].

Ve své práci se budu zabývat tzv. *informační entropií*, jež byla formulována roku 1948 americkým matematikem C. E. Shannonem, a tedy se můžeme setkat také s názvem *Shannonova entropie*. Obecně bývá termín „entropie“ často chápán

jako míra neurčitosti (neusprádanosti) systému, viz [11]. Pokud je entropie vysoká, považujeme daný systém za neuspořádaný – lze říct i chaotický či náhodný, viz [17]. Na informační entropii můžeme ale také nahlížet jako na „míru informace (překvapení) systému“. Vezměme si jako takový systém například anglicky psaný text, který budeme zkoumat po jednotlivých písmenech. Pokud ve větě uvidíme například písmeno „t“, je malým překvapením, že by mohlo následovat písmeno „h“ (jedná se o velmi časté spojení, viz určitý člen „the“). Jednalo by se tedy o nízkou entropii. Na druhou stranu písmeno „x“ téměř nikdy po písmenu „t“ nenajdeme, tedy jde o systém s vysokou entropií, viz [4]. Čtení takové textu pak tedy představuje náhodný jev x , který nastává s pravděpodobností $p(x)$. Abychom mohli entropii textu vypočítat, je tedy nutné znát pravděpodobnost výskytu jednotlivých písmen. Pokud se jedná o písmeno, které není příliš očekávané, je míra informace velká. Pokud je pravděpodobnost výskytu nějakého znaku rovna 1 (na 100% očekáváme dané písmeno, tedy nenese žádnou informaci), je pak jeho entropie rovna 0. Tedy čím očekávanější je znak, tím je entropie nižší, viz [9].

1.2. Entropie u biologických signálů

Biologický signál je signál, který vzniká v živém organismu a dělí se do několika skupin – bioelektrický (např. elektrokardiologický signál – EKG, elektroencefalografický signál – EEG), biomagnetický (v praxi se téměř nevyužívá), bioimpedanční (např. fonokardiografický signál – srdeční ozvy), biomechanický (např. krevní tlak, respirační frekvence) a biochemický (např. Ph tkání), viz [7]. Ve své diplomové práci se budu zabývat první skupinou – bioelektrickými signály, do které patří elektrokardiografický záznam (zkráceně EKG). V této podkapitole budu využívat zejména poznatků ze dvou odborných článků [1] a [2].

Již několik let se přední vědci věnují tématu vyčíslení složitosti (neuspořádanosti) fyziologických časových řad, jako je například srdeční rytmus. Tradiční metody (např. přibližná entropie) určují míru pravidelnosti časové řady hodnocením vzhledu opakujících se menších úseků této řady. Na druhou stranu ale nelze

jednoznačně určit přímočarou souvislost mezi pravidelností a složitostí. Složitost lze intuitivně chápat jako „významnou strukturální bohatost“. Zde tedy nastává základní nesoulad, jelikož tyto tradiční algoritmy přiřazují vyšší složitost jistým patologickým projevům než zdravé dynamice systému, jež se projevuje dlouhými korelacemi. Zdravý organismus totiž nevykazuje zcela pravidelnou dynamiku, ale spíše „zdravou míru“ neuspořádanosti. Podle algoritmů zabývajících se složitostí (např. přibližná entropie) se entropie zvyšuje na základě množství přeházení (neuspořádání) a je maximální pro zcela náhodný systém. Na druhou stranu, podle nynějších studií, nemusí růst entropie znamenat zvýšení složitosti systému. Viz [1] a [2].

Pokud tedy budeme tyto postupy, jež byly používány doposud, aplikovat na fyziologickou časovou řadu, můžeme dostat zavádějící výsledky (popsané výše). Proto se zvýšená pozornost věnuje měření, které přiřadí minimální hodnoty entropie jak deterministickému (předvídatelnému), tak náhodnému (zcela nepředvídatelnému) signálu.

Motivaci tohoto přístupu poskytují tři základní hypotézy, viz [2]:

- složitost biologického signálu se odráží ve schopnosti neustále se přizpůsobovat a fungovat ve stále se měnícím prostředí,
- biologický systém musí být schopen pracovat na více prostorových a časových škálách, tedy jeho složitost je také víceúrovňová,
- složitost může být ovlivněna mnoha faktory, jako je například věk či genetická predispozice k nemoci, které mohou snížit schopnost přizpůsobení se daného jedince.

1.3. Vzorková entropie

Víceškálová entropie (MSE) je poměrně nová metoda měření složitosti konečně dlouhé časové řady. Tento postup může být aplikován jak na fyzikální, tak i na fyziologické soubory dat. Abychom mohli MSE definovat, musíme zavést některé

další základní druhy entropie, budeme čerpat především z odborných článků [1], [2] a [13].

Informační entropie $H(X)$ diskrétní náhodné veličiny X , tzv. *Shannonova entropie*, je dána vztahem

$$H(X) = - \sum_{x_i \in \Theta} p(x_i) \log_b p(x_i), \quad (1)$$

kde Θ je obor hodnot náhodné veličiny X a $p(x_i)$ představuje pravděpodobnost, že X nabyde hodnoty x_i . S použitím základem logaritmu b souvisí jednotka této entropie. Nejběžněji se využívá $b = 2$ a jednotkou je *bit*, $b = e$ s jednotkou *nat* a *dit* pro $b = 10$.

Pro časovou řadu, která představuje realizaci stochastického procesu, skládajícího se z indexované sekvence n náhodných veličin (X_1, \dots, X_n) je sdružená entropie definovaná vztahem

$$H_n = H(X_1, X_2, \dots, X_n) = - \sum_{x_1 \in \Theta_1} \dots \sum_{x_n \in \Theta_n} p(x_1, \dots, x_n) \log p(x_1, \dots, x_n), \quad (2)$$

kde $p(x_1, \dots, x_n)$ je sdružená pravděpodobnost, že $X_1 = x_1, \dots, X_n = x_n$.

Stav systému X_n v určitém okamžiku n je částečně určen jeho historií X_1, \dots, X_{n-1} . Nicméně, každý další nový stav s sebou přináší další nové informace. Průměrná hodnota tvorby těchto informací, nazývaná *Kolmogorova-Sinaiova (KS) entropie*, je používaný parametr při popisování dynamiky systému. Uvažujme D -dimenzionální dynamický systém (generovaný systémem X_n). Předpokládejme, že fázový prostor \mathbb{R}^D systému je rozdělen do hyperkrychli o obsahu ε^D a stav systému je měřen v časových intervalech δ . Nechť $p(k_1, k_2, \dots, k_n)$ označuje sdruženou pravděpodobnost, že se systém nachází v hyperkrychli k_1 v čase $t = \delta$, v hyperkrychli k_2 v čase $t = 2\delta$, atd. až v hyperkrychli k_n v čase $t = n\delta$. *KS-entropie* celého systému je potom definována jako

$$\begin{aligned} H_{KS} &= - \lim_{\delta \rightarrow 0} \lim_{\varepsilon \rightarrow 0} \lim_{n \rightarrow \infty} \frac{1}{n\delta} \sum_{k_1, \dots, k_n} p(k_1, \dots, k_n) \log p(k_1, \dots, k_n) = \\ &= \lim_{\delta \rightarrow 0} \lim_{\varepsilon \rightarrow 0} \lim_{n \rightarrow \infty} \frac{1}{n\delta} H_n. \end{aligned} \quad (3)$$

Tato rovnice má pouze omezené použití pro odhady entropie konečně dlouhé časové řady.

1.3.1. Výpočet přibližné entropie

Pro analýzy konečně dlouhých časových řad (původně od 100 do 5000 hodnot) byla S. M. Pincusem zavedena skupina parametrů, používaná velmi široce v oblasti fyziologie a medicíny, pojmenovaná *přibližná entropie* $ApEn$ (z angl. Approximate Entropy), která se vypočítá sledováním opakování stejně dlouhých úseků časové řady, jež jsou si podobné (viz [2]). Nižší hodnoty $ApEn$ jsou přiřazovány více pravidelným časovým řadám (menší složitost systému). Naopak vyšší hodnoty $ApEn$ jsou spojovány s nepravidelnými, méně předvídatelnými, časovými řadami (vyšší složitost systému).

K výpočtu přibližné entropie je nutné předem specifikovat tři parametry:

- n je délka časové řady,
- m je délka sekvence, jež je určena ke srovnání,
- r představuje toleranci přijetí shody.

Formální výpočet této metody lze rozdělit do několika kroků, viz [2] a [12]: Uvažujme časovou řadu délky n , $\{x_1, x_2, \dots, x_n\}$.

- Vytvoříme postupně m -tice $U_m(1), U_m(2), \dots, U_m(n - m + 1)$ následujícím předpisem:

$$U_m(i) = \{x_i, x_{i+1}, \dots, x_{i+m-1}\}, \quad \text{pro } 1 \leq i \leq n - m + 1. \quad (4)$$

Takto vzniklé m -tice představují m po sobě jdoucích hodnot x začínajících na i -té pozici.

- Určíme vzdálenost $d[U_m(i), U_m(j)]$ dvou sekvencí $U_m(i)$ a $U_m(j)$ jako maximum rozdílu jejich prvků (tzv. *Čebyševova vzdálenost*):

$$d[U_m(i), U_m(j)] = \max_{0 \leq k \leq m-1} (|x(i+k) - x(j+k)|). \quad (5)$$

- Vypočítáme N_i , což představuje počet indexů j , pro $j \in \{1, \dots, n - m\}$, $j \neq i$, pro které platí:

$$d[U_m(i), U_m(j)] \leq r. \quad (6)$$

$N_i^m(r)$ definujeme následovně:

$$N_i^m(r) = \frac{N_i}{n - m}. \quad (7)$$

Jedná se tedy o podíl m -tic v celé časové řadě, které jsou blíže dané m -tici $U_m(i)$ než je tolerance r .

- Položme $N^m(r)$ jako průměr $N_i^m(r)$:

$$N^m(r) = \frac{1}{n - m + 1} \sum_{i=1}^{n-m+1} N_i^m(r). \quad (8)$$

$N^m(r)$ představuje tedy podíl párů m -tic, které jsou si blíže než je tolerance r .

- Nyní definujeme funkci $\Phi^m(r)$ předpisem:

$$\Phi^m(r) = \frac{1}{n - m + 1} \sum_{i=1}^{n-m+1} \ln N_i^m(r). \quad (9)$$

- Nyní zvýšíme délku segmentu na $m + 1$ a stejným způsobem vypočítáme vypočítáme $\Phi^{m+1}(r)$. Hodnota $\Phi^m(r)$ představuje pravděpodobnost, že dvě m -tice jsou si blíže než je tolerance r a podobně $\Phi^{m+1}(r)$ vyjadřuje pravděpodobnost blízkosti pro $(m + 1)$ -tice.

- V posledním kroku definujeme vzorkovou entropii:

$$ApEn(m, r, n) = \Phi^m(r) - \Phi^{m+1}(r). \quad (10)$$

Později došlo Richmanem a Moormanem k modifikaci tohoto algoritmu na *vzorkovou entropii SampleEntropy*. V následujícím odstavci popíšeme myšlenku a odvození výpočtu této metody.

1.3.2. Výpočet vzorkové entropie

V tomto odstavci čerpáme především z článku autorů tohoto algoritmu, [8]. Vzorková entropie má značné výhody oproti předešlému pojmu $ApEn$:

- vynechává se porovnání vzorového úseku se sebou samým,
- pravděpodobnost je vypočítána jako přirozený logaritmus podmíněné pravděpodobnosti místo použití vzorce při odhadu poměru podmíněné pravděpodobnosti.

Tím je algoritmus méně závislý na délce časové řady. Nicméně, tento postup přiřadí opět vyšší hodnoty entropie jistým patologickým časovým řadám vyjádřených menší složitostí dynamiky systému (např. bílý šum), než časovým řadám odvozených ze zdravé funkčnosti fyziologického systému.

Ve zkratce lze říci, že tato procedura vypočítá záporný logaritmus podmíněné pravděpodobnosti, že dvě sekvence si jsou podobné podobné v $m + 1$ bodech za podmínky, že jsou si podobné v m bodech.

Formální výpočet této metody lze, stejně jako v předchozím odstavci, rozdělit do několika kroků, viz [2] a [8]:

Uvažujme časovou řadu délky n , $\{x_1, x_2, \dots, x_n\}$.

- Vytvoříme postupně m -tice $U_m(1), U_m(2), \dots, U_m(n - m + 1)$ následujícím předpisem:

$$U_m(i) = \{x_i, x_{i+1}, \dots, x_{i+m-1}\}, \quad \text{pro } 1 \leq i \leq n - m + 1. \quad (11)$$

Takto vzniklé m -tice představují m po sobě jdoucích hodnot x začínající na i -té pozici.

- Určíme vzdálenost $d[U_m(i), U_m(j)]$ dvou sekvencí $U_m(i)$ a $U_m(j)$ jako maximum rozdílu jejich prvků (tzv. *Čebyševova vzdálenost*):

$$d[U_m(i), U_m(j)] = \max_{0 \leq k \leq m-1} (|x(i+k) - x(j+k)|). \quad (12)$$

- Vypočítáme N_i , což představuje počet indexů j , pro $j \in \{1, \dots, n - m\}$, $j \neq i$, pro které platí:

$$d[U_m(i), U_m(j)] \leq r. \quad (13)$$

$N_i^m(r)$ definujeme následovně:

$$N_i^m(r) = \frac{N_i}{n - m - 1}. \quad (14)$$

Jedná se tedy o podíl m -tic v celé časové řadě, které jsou blíže dané m -tici $U_m(i)$ než je tolerance r .

- Položme $N^m(r)$ jako průměr $N_i^m(r)$:

$$N^m(r) = \frac{1}{n - m} \sum_{i=1}^{n-m} N_i^m(r). \quad (15)$$

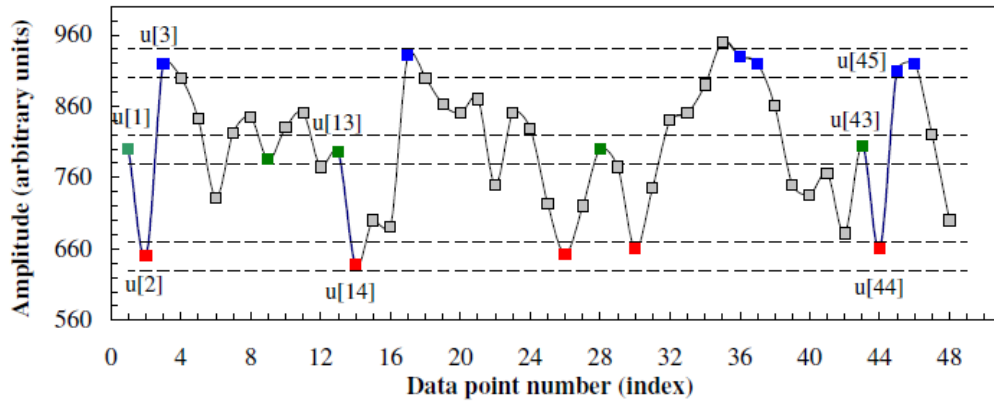
$N^m(r)$ představuje tedy podíl párů m -tic, které jsou si blíže než je tolerance r .

- Nyní zvýšíme délku segmentu na $m + 1$ a stejným způsobem vypočítáme vypočítáme $N_i^{m+1}(r)$ a $N^{m+1}(r)$. Hodnota $N^m(r)$ představuje pravděpodobnost, že dvě m -tice jsou si blíže než je tolerance r a podobně $N^{m+1}(r)$ vyjadřuje pravděpodobnost blízkosti pro $(m + 1)$ -tice.
- V posledním kroku definujeme vzorkovou entropii:

$$SampleEntropy(m, r, n) = -\ln \left(\frac{N^{m+1}(r)}{N^m(r)} \right). \quad (16)$$

Nulová hodnota *SampleEntropy* značí, že se jedná o časovou řadu, jež se po částech opakuje s danou tolerancí. Nízké hodnoty *SampleEntropy* představují časovou řadu s jistými podobnostmi, oproti tomu vysoké hodnoty ukazují na náhodnou, či chaotickou řadu.

Příklad použití vzorkové entropie pro pravidelný i zcelanáhodný signál bude uveden v dalších kapitolách po představení algoritmu této metody.



Obrázek 1: Znázornění použití algoritmu vzorkové entropie SampleEntropy.

Na obr. 1 (převzatý z [2]) můžeme vidět ilustrace vzorkové entropie. V tomto případě je délka sekvence, jež je určena ke srovnání, rovna $m = 2$ a tolerance $r = 20$. Čárkované vodorovné čáry značí rozsah tolerance u prvních tří hodnot, tedy $u[1] \pm r$, $u[2] \pm r$ a $u[3] \pm r$. Dále můžeme konstatovat, že dvě datové hodnoty si navzájem odpovídají, jestliže absolutní rozdíl mezi nimi nepřesáhne r . V grafu jsou tyto body znázorněny zelenou barvou pro hodnotu $u[1]$, červenou barvou pro hodnotu $u[2]$ a modrou barvou pro hodnotu $u[3]$. Lze vypočítat, že v celé časové řadě se vyskytnou celkem 2 shody po sobě jdoucího zeleného a červeného bodu - $(u[13], u[14])$ a $(u[43], u[44])$ a pouze jedna shoda po sobě jdoucího červeného, zeleného a modrého bodu - $(u[43], u[44], u[45])$. Proto počet odpovídajících si 2-komponentních sekvencí je 2 a počet 3-komponentních sekvencí je 1. Stejný výpočet provedeme také pro další vzorové sekvence o dvou a třech složkách - $(u[2], u[3])$ a $(u[2], u[3], u[4])$ a počty shodných sekvencí přičteme k těm předchozím. Postup opakujeme pro všechno možné vzorové sekvence $(u[3], u[4], u[5]), \dots, (u[n-2], u[n-1], u[n])$, abychom dostali poměr všech 2-komponentních a 3-komponentních shodných sekvencí. Přibližná entropie je pak přiřazený logaritmus tohoto poměru a vyjadřuje pravděpodobnost, že si sekvence, navzájem odpovídající si v prvních dvou datových bodech, budou odpovídat i v dalším bodě.

1.3.3. Implementace funkce vzorkové entropie v programu Matlab

V tomto odstavci popíši vlastní programový kód vytvořený v programu Matlab, který vypočítá hodnoty vzorkované entropie a který bude dále použit k vypočítání víceškálové entropie. Dále uvedu názornou ukázkou použití této funkce na dvou rozdílných signálech. Veškerý programový kód s komentáři je uveden na přiloženém CD.

Funkce `SampleEntropy.m` obsahuje tři vstupní parametry:

- *signal* - časová řada (zadaná jako sloupcový vektor),
- *m* - délka sekvence určena ke srovnání,
- *r* - tolerance přijetí shody (procentuální podíl směrodatné odchylky časové řady).

Výstupním parametrem je hodnota vzorkové entropie *SampleEntropy* vypočítána na základě postupu uvedeného v předchozím odstavci.

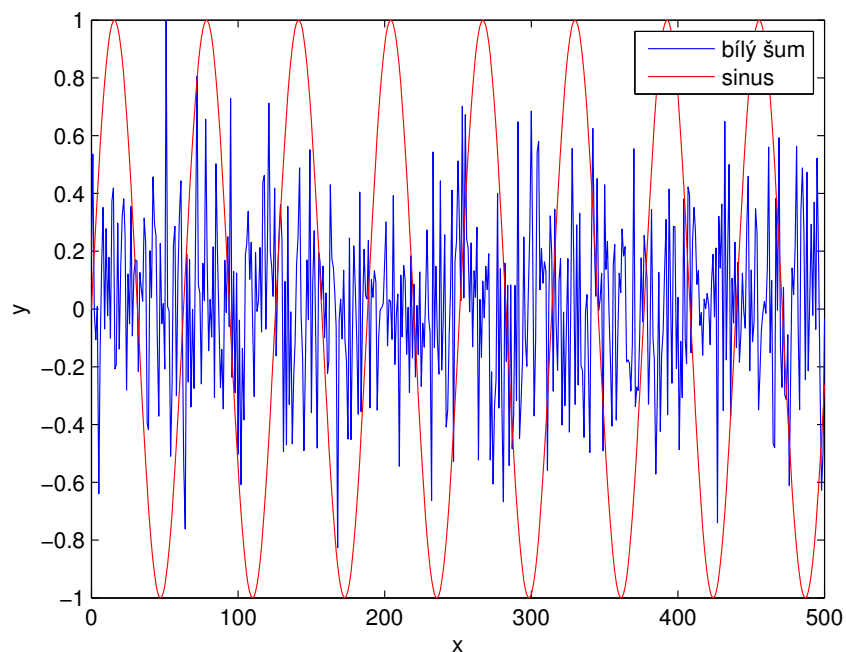
Pro názornou ukázkou použití této metody jsem zvolila dva různé signály. První, zcela náhodný signál představuje bílý šum, což je náhodný signál, jehož hustota spektrálního výkonu je konstantní. Hodnoty této časové řady byly vytvořeny pomocí funkce *randn* a následně normalizovány pomocí následujícího příkazu:

$$signal = signal ./ mean(signal).$$

Druhý signál tvoří hodnoty funkce $\sin(x)$. Oba signály mají stejnou délku $n = 1000$ hodnot. Zobrazení těchto signálů můžeme vidět na obr. 2. Z tohoto grafu je zcela patrné, že bílý šum představuje zcela nepravidelný, tedy nepředvídatelný signál, zatímco sinus je funkce periodická s periodou 2π , tedy předvídatelná na intervalech této délky.

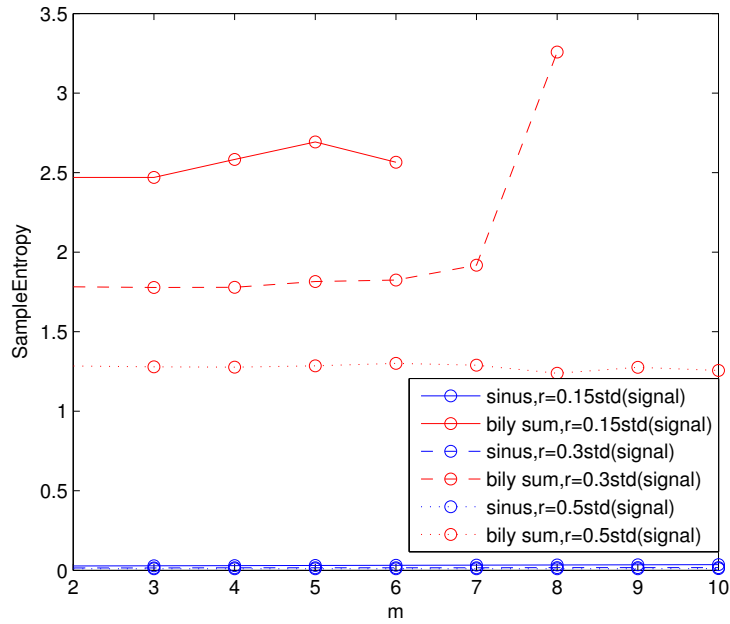
Na oba signály aplikujeme metodu `SampleEntropy` s následujícími vstupními parametry:

- $m = 2, \dots, 10$,
- $r = 0.15std(signal), r = 0.3std(signal), r = 0.5std(signal)$.

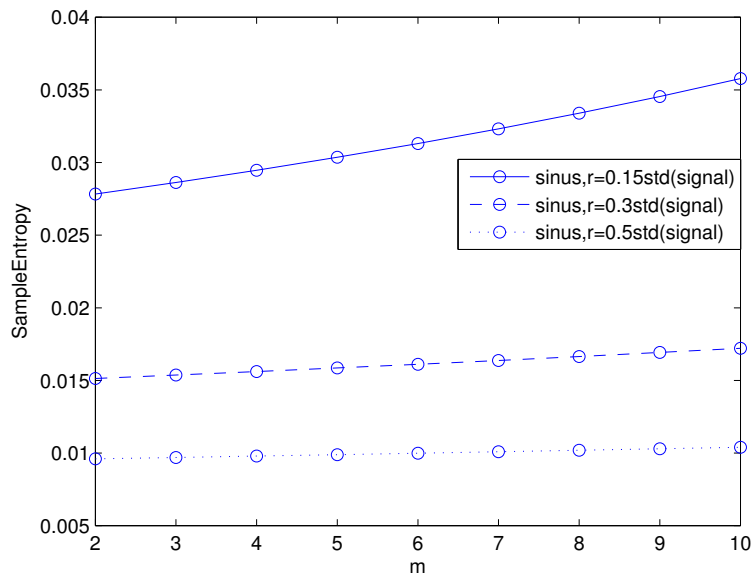


Obrázek 2: Znázornění signálů funkce sinus a bílého šumu.

Výstupním parametrem je hodnota vzorkové entropie `SampleEntropy`. Srovnání naměřených hodnot vzorkové entropie obou signálů pro různá m a r můžeme pozorovat na obr. 3. Je patrné, že pro všechny hodnoty m (porovnávají se tedy úseky délky 2 a více) je hodnota vzorkové entropie vyšší pro náhodný bílý šum, zatímco hodnoty entropie funkce sinus jsou blízké hodnotě 0. Hodnoty `SampleEntropy` klesají s rostoucí tolerancí r . U bílého šumu je pro $r = 0.15std(signal)$ hodnota entropie pro parametr $m = 7$ rovna nekonečnu. Tedy lze usoudit, že se jedná o zcela náhodný signál bez jakýchkoli pravidelností či podobností. Je logické, že pro zbylé hodnoty parametru m vyšla entropie `NaN(NotANumber)`. Stejná situace nastala i pro toleranci $r = 0.3std(signal)$ od hodnoty $m = 9$. Při bližším pozorování, na obr. 4, si lze všimnout, že hodnoty `SampleEntropy` funkce sinus jsou blízké hodnotě 0 pro všechny hodnoty m , tedy signál se při porovnávání stále většího vzorku dat, s rostoucí tolerancí r , jeví jako velmi pravidelný.



Obrázek 3: Srovnání hodnot vzorkové entropie signálů funkce sinus a bílého šumu.



Obrázek 4: Srovnání hodnot vzorkové entropie funkce sinus.

1.4. Víceškálová entropie – MSE

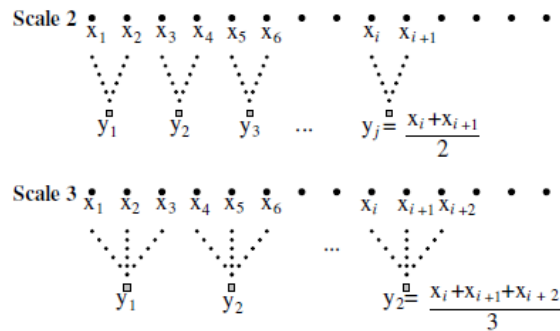
Fyziologické systémy jsou ovlivňovány vzájemně na sebe působícími mechanismy, které pracují na více časových a prostorových škálách. Výstupní proměnné takových systémů (např. EKG signál) často vykazují složité kolísání, jež může dát informaci o skrytých dynamikách systému.

Z tohoto důvodu představil Y.-C. Zhang obecný přístup, ve kterém bere v úvahu více časových škál v systému. Entropie tak může být počítána a zkoumána v různých úrovních signálu samostatně. Tento postup přináší vyšší hodnoty entropie pro korelované signály ve srovnání s nekorelovanými, viz [2].

1.4.1. Výpočet víceškálové entropie

V tuto chvíli mám dostatek podkladů k popisu a definování stěžejní části této kapitoly - *víceškálové entropie (MSE)*. Uvažujme jednodimenzionální diskrétní časovou řadu $\{x_1, \dots, x_i, \dots, x_n\}$, ze které zkonstruujeme novou zhlazenou časovou řadu $\{y^\tau\}$ odpovídající škálovému faktoru $\tau \in \mathbb{N}$. Nejprve rozdělíme původní časovou řadu na nepřekrývající se okna délky τ , následně vypočítáme průměrné hodnoty ve všech oknech, jak můžeme vidět na obr.5 (převzatý z [2]). Obecně lze každý prvek nové časové řady $\{y^\tau\}$ vypočítat pomocí vztahu

$$y_j^\tau = \frac{1}{\tau} \sum_{i=(j-1)\tau+1}^{j\tau} x_i, \quad 1 \leq j \leq \frac{n}{\tau}. \quad (17)$$



Obrázek 5: Schématické znázornění vytvoření časové řady $\{y^\tau\}$.

Pro škálu 1 (scale 1) odpovídá nová časová řada $\{y^{(1)}\}$ původní časové řadě. Délka každé zhlazené časové řady je rovna délce původní řady dělené škálovým faktorem τ .

V posledním kroku vypočítáme vzorkovou entropii (SampleEntropy) pro každou přeškálovanou řadu jako funkci parametru τ . Tento proces se nazývá *víceškálová entropie* a výsledkem je závislost vzorkované entropie SampleEntropy na škálovém faktoru τ .

Pro hodnocení se využívá zejména křivka hodnot MSE v závislosti na škálovém faktoru.

1.4.2. Implementace funkce víceškálové entropie v programu Matlab

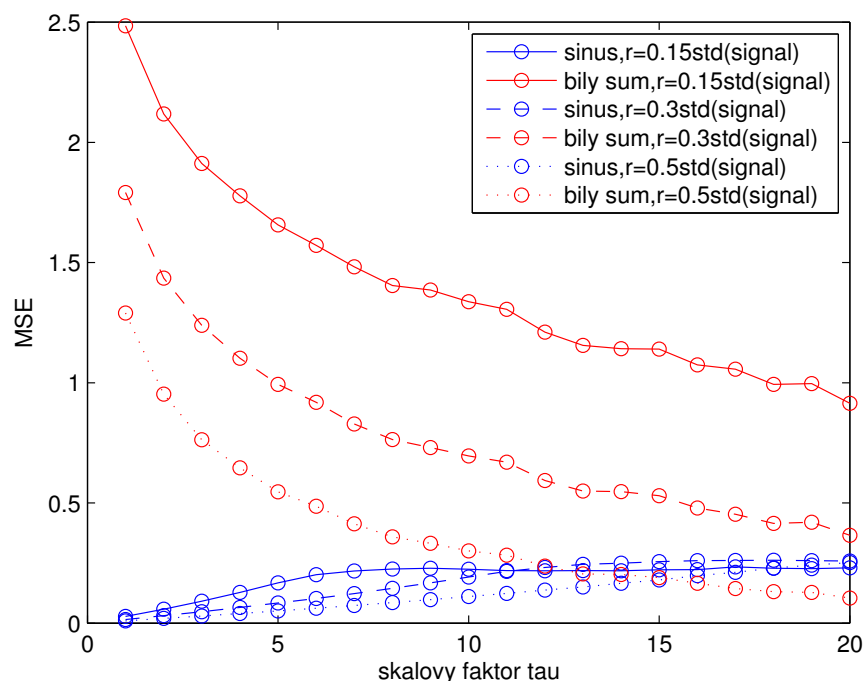
Nyní se zaměřím na postup implementace funkce víceškálové entropie v programu Matlab, jež využívá vnořenou funkci SampleEntropy (popsaná v odstavci 1.3.2.). Následně uvedu příklad srovnání víceškálové entropie pro různé druhy šumů. Veškerý programový kód s komentáři je uveden na přiloženém CD.

Funkce mse.m pracuje na základě postupu popsáního v předchozím odstavci. Obsahuje čtyři vstupní parametry:

- *signal* – časová řada (zadaná jako sloupcový vektor),
- *m* – délka sekvence určena ke srovnání,
- *r* – tolerance přijetí shody (procentuální podíl směrodatné odchylky původní časové řady),
- *factor* – maximální hodnota škálového faktoru τ .

Funkce obsahuje dvě vnořené funkce, jejichž výstupy jsou závislé na škálovém faktoru τ . Výsledkem první funkce zhlazeni.m je přeškálovaná časová řada a následně druhá funkce SampleEntropy.m vypočítá hodnoty vzorkované entropie pro každou přeškálovanou řadu.

Stejně jako u vzorkové entropie, aplikuji metodu víceškálové entropie nejprve na srovnání funkce sinus a bílého šumu. Vypočítané hodnoty *MSE* jsou znázorněny na obr. 6. Výsledky nejsou nijak překvapivé, jelikož bílý šum, je náhodný

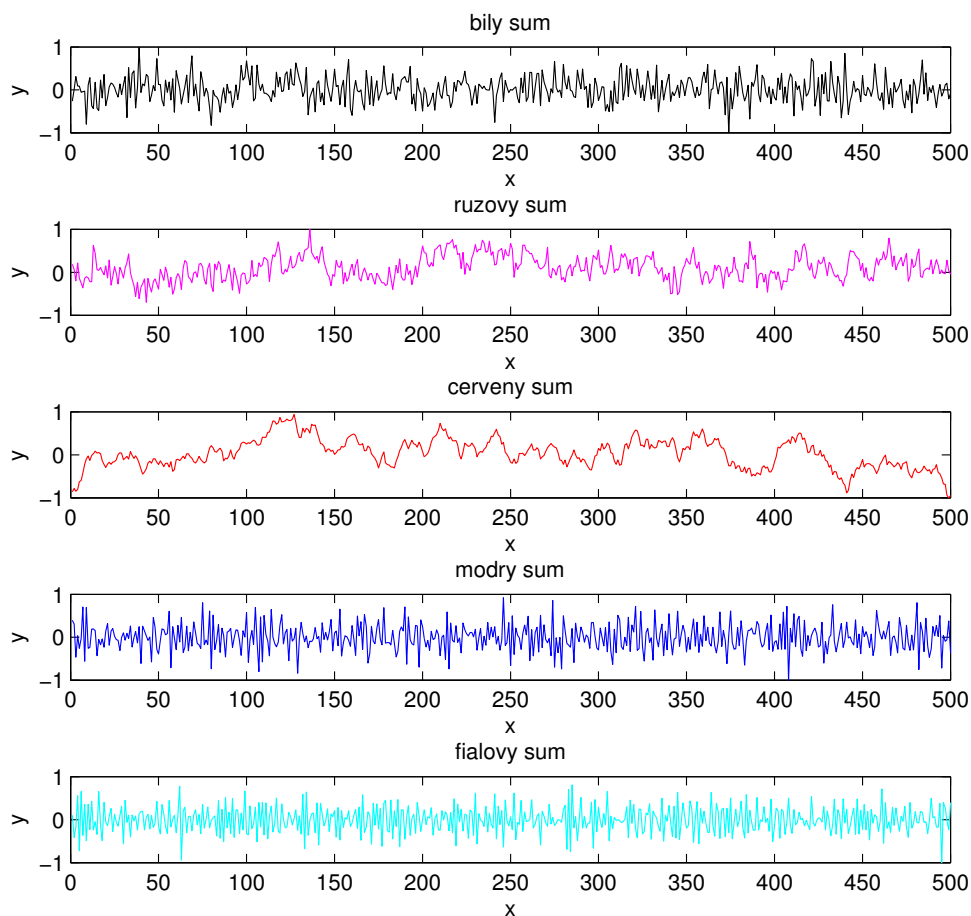


Obrázek 6: Srovnání hodnot víceškálové entropie funkce sinus a bílého šumu.

signál, jsou jeho hodnoty MSE vyšší než hodnoty MSE funkce sinus na všech časových škálách.

Další aplikaci MSE uvedu na srovnání několika základních signálů, viz [14]:

- **bílý šum** - náhodný signál, jehož hustota spektrálního výkonu je konstantní,
- **růžový šum** (neboli $1/f$ šum) - signál, jehož hustota spektrálního výkonu je nepřímo úměrná frekvenci signálu,
- **červený šum** (neboli $1/f^2$ šum) - signál, jehož hustota spektrálního výkonu je nepřímo úměrná druhé mocnině frekvence signálu,
- **modrý šum** - signál, jehož hustota spektrálního výkonu je přímo úměrná frekvenci signálu,
- **fialový šum** - signál, jehož hustota spektrálního výkonu je přímo úměrná druhé mocnině frekvence signálu,

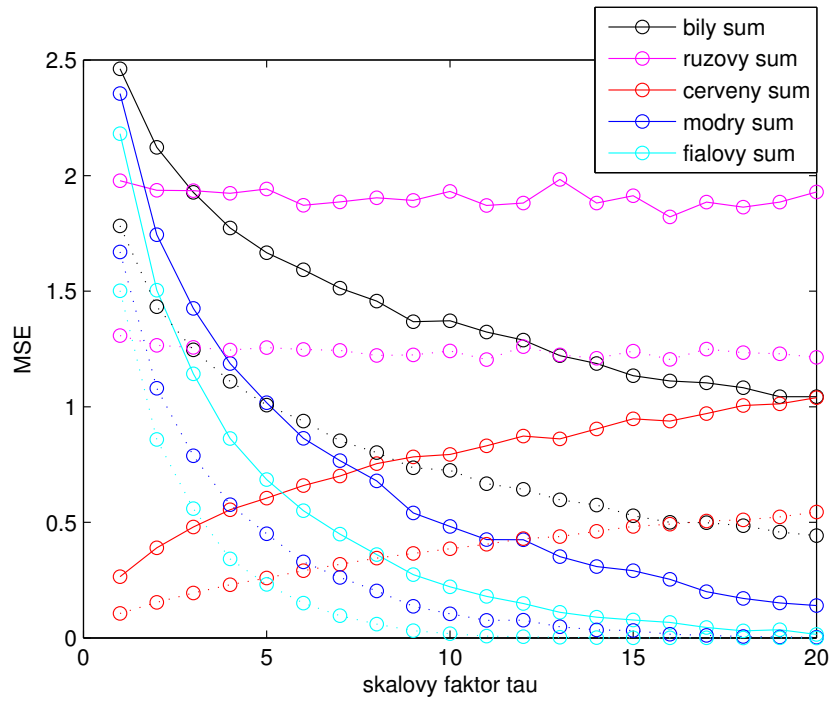


Obrázek 7: Základní rozdělení šumů.

Jednotlivé grafické znázornění těchto signálů pro 500 vzorků je uvedeno na obr. 7. Podrobný popis všech druhů šumů i s možností poslechu viz [14].

Na výše uvedené signály, o délce 10000 hodnot, jsem aplikovala funkci `mse.m` s následujícími parametry:

- $m = 2$,
- $r = 0.15std(signal)$,
- $factor = 20$.



Obrázek 8: Víceškálová entropie základních.

Vývoj entropie těchto signálů v závislosti na škálovém faktoru τ a různých hodnotách můžeme pozorovat na obr. 8.

1.5. Aplikace víceškálové entropie na analýzu EKG záznamů

V této kapitole uvedu praktické použití a vyhodnocení metody víceškálové entropie na reálných EKG signálech.

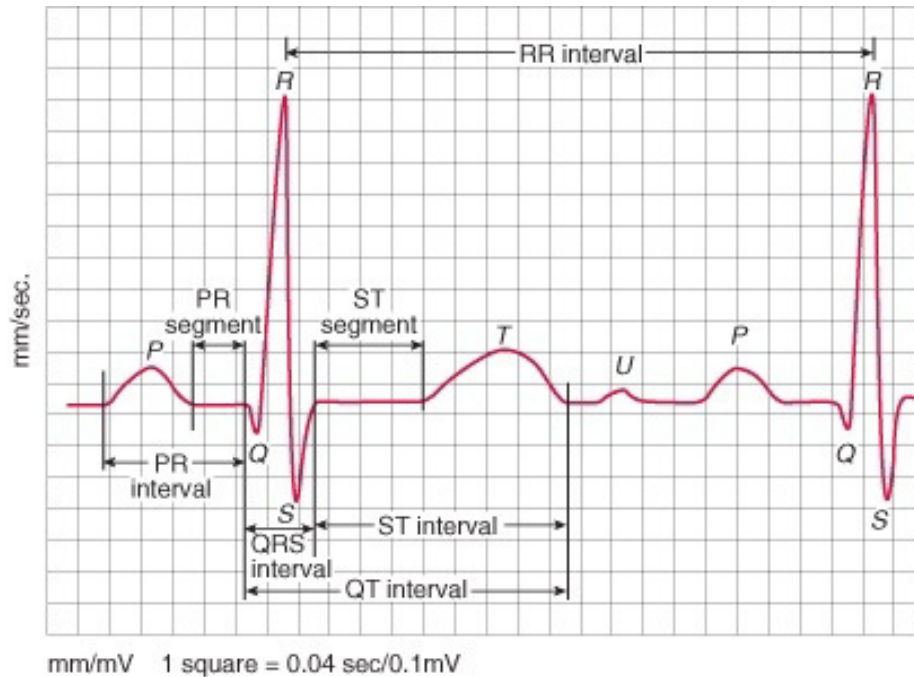
1.5.1. Popis datových souborů

Pro účely své diplomové práce mám k dispozici dva datové soubory. První soubor obsahuje 22 EKG signálů pacientů se synkopou. Synkopa je náhlá, krátkodobá ztráta vědomí, jež nastane následkem nedostatku okysličené krve v mozku - tzv. hypoxie. Synkopa může nastat zcela bez varování u kohokoli, například po dlouhém stání či silném citovém vypětí, viz [19]. Druhý datový soubor obsahuje EKG signály 22 zdravých pacientů. Tento soubor jsem převzala z databáze portálu [16]. Všechny EKG záznamy byly získány během 24 hodinového měření přístrojem zvaným *EKG holter*, který snímá srdeční aktivitu po celou tuto dobu a během všech aktivit a činností, které pacienti během dne provozují.

Jedním z cílů a také přínosů této diplomové práce je aplikování metody víceškálové entropie na srovnání těchto dvou datových souborů. Budu testovat hypotézu, že hodnoty funkce víceškálové entropie v závislosti na škálovém faktoru τ zdravých lidí jsou vyšší než u pacientů se synkopou. Tedy že EKG záznam zdravých pacientů je méně pravidelný (více složitý) než u nemocných.

Stejně jako ve své bakalářské práci, aplikuji tuto metodu na posloupnost tzv. *RR intervalů*, což jsou časové intervaly mezi následujícími stahy srdce. Na EKG záznamu se jedná o délku mezi po sobě jdoucími QRS komplexy, jak můžeme vidět na Obr. 9.

Získání RR intervalů z reálného EKG záznamů je popsáno v mé bakalářské práci, viz [5].



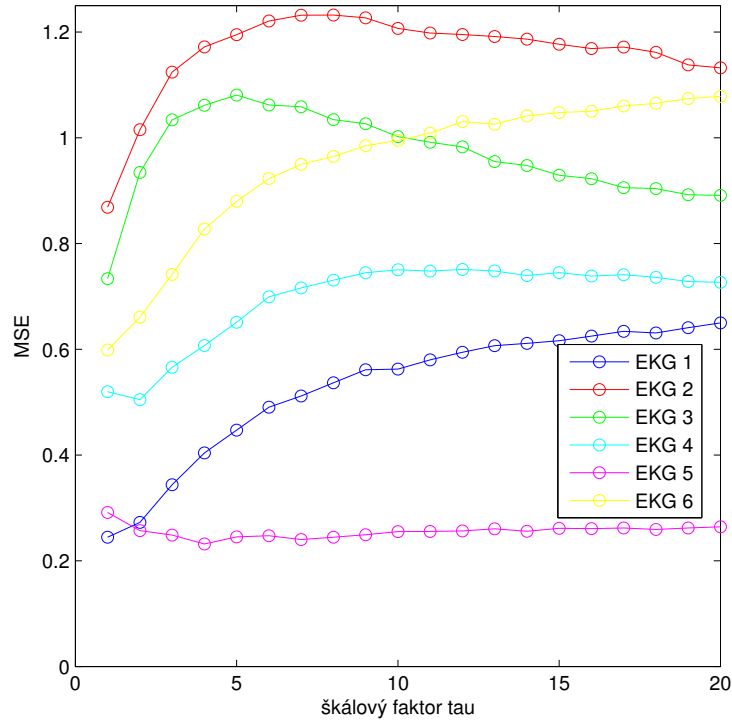
Obrázek 9: Znázornění RR intervalu v EKG záznamu.

1.5.2. Výsledné hodnoty MSE reálných EKG signálů

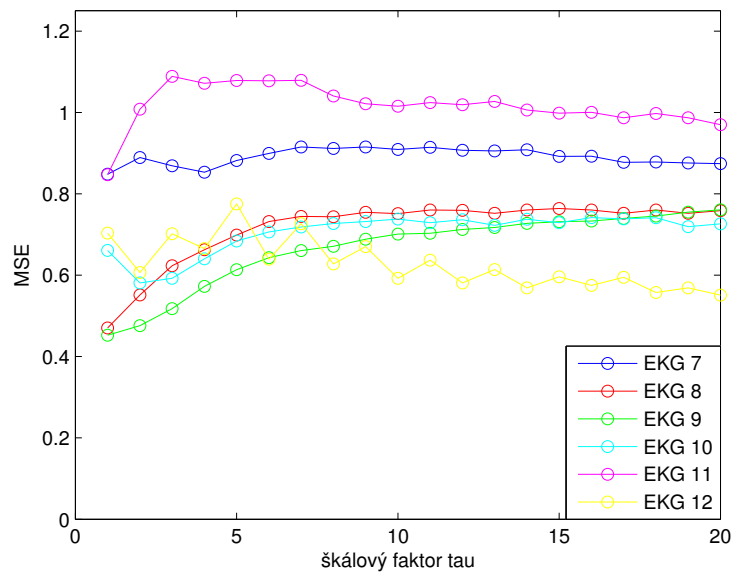
Pro biologické signály se ukázalo, že je rozumné používat metodu MSE s následující parametry, viz [2]:

- $m = 2$,
- $r = 0.15std(signal)$,
- $factor = 20$.

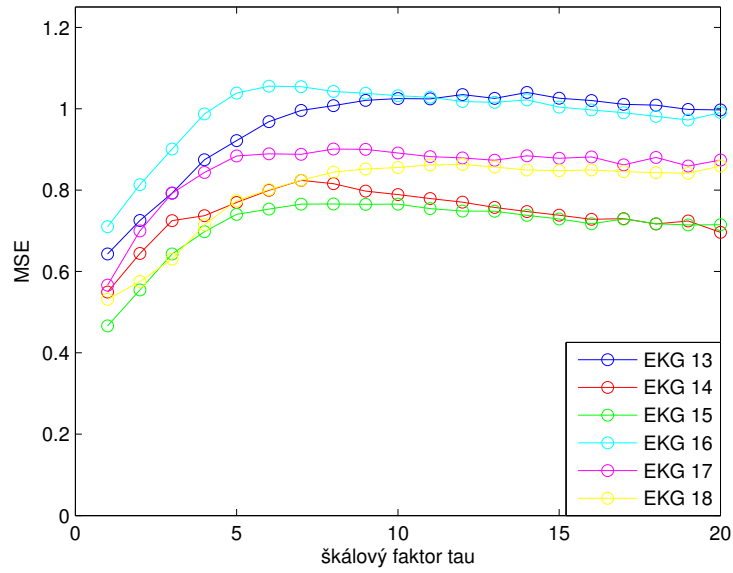
Na všechny EKG signály aplikujeme algoritmus popsany v předchozích odstavcích s těmito parametry. Vypočítané hodnoty souboru EKG se synkopou v závislosti na škálovém faktoru $\tau = 1, \dots, 20$ jsou uvedeny na grafech na obr. 10, 11, 12 a 13. Ze získaných hodnot je vypočítána průměrná hodnota a střední chyba průměru MSE. Tyto výsledky jsou zaznamenány v grafu na obr. 14. Hodnoty MSE podle škálového faktoru všech EKG signálů jsou uvedeny v kapitole Přílohy na konci diplomové práce v tabulkách 10, 11, 12 a 13.



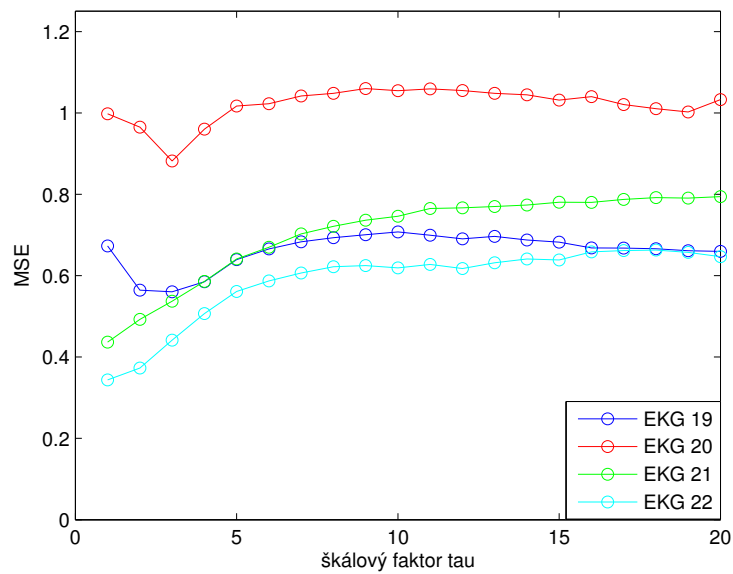
Obrázek 10: Znázornění hodnot MSE EKG signálů se synkopou.



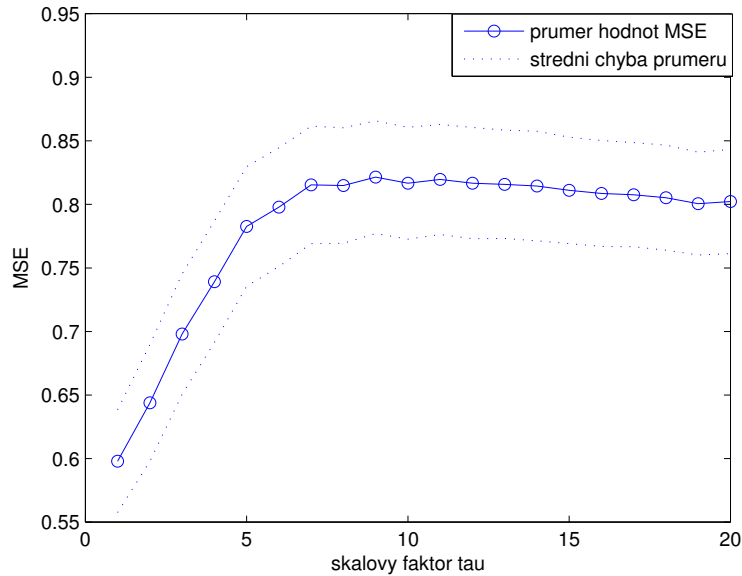
Obrázek 11: Znázornění hodnot MSE EKG signálů se synkopou.



Obrázek 12: Znázornění hodnot MSE EKG signálů se synkopou.



Obrázek 13: Znázornění hodnot MSE EKG signálů se synkopou.



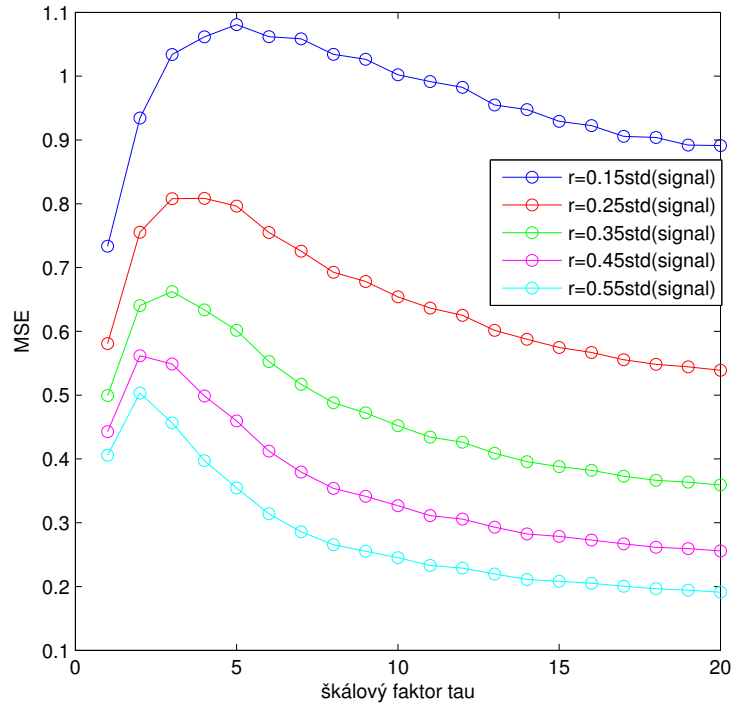
Obrázek 14: Znázornění průměrné hodnoty a střední chyby průměru souboru EKG se synkopou.

U všech EKG signálů byly použity stejné parametry pro výpočet víceškálové entropie. Pokud bychom zvyšovali toleranci r , budou se hodnoty MSE snižovat, ale tvar závislosti zůstane zachován. Pro názornost jsem na dva EKG signály (č. 3 a č. 4) aplikovala algoritmus s hodnotami tolerance

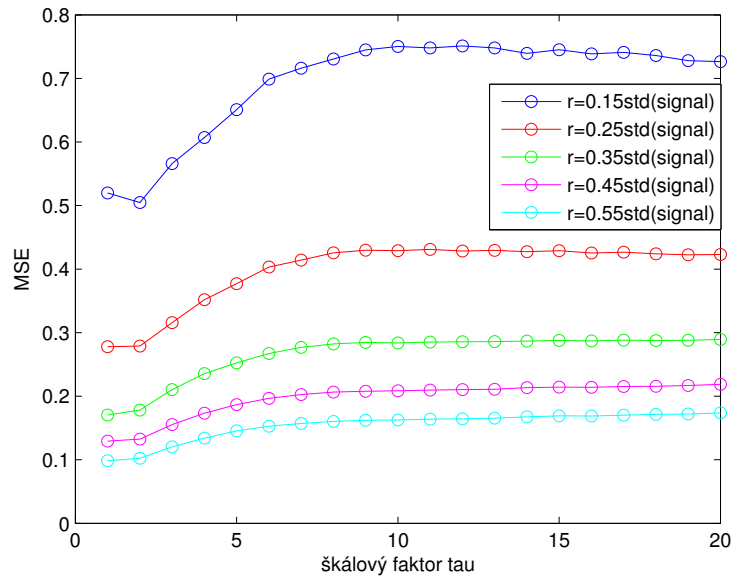
$$r = 0.15, 0.25, 0.35, 0.45, 0.55 \text{std}(\text{signal}). \quad (18)$$

Průběh hodnot MSE můžeme pozorovat na obr. 15 a 16.

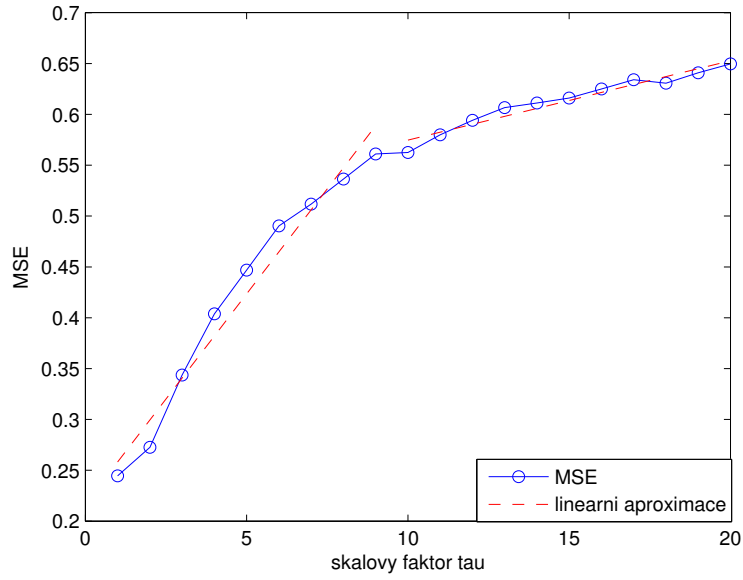
Pro lepší a názornější vyhodnocení jsem každou křivku MSE aproximovala dvěma lineárními polynomy, k jejichž výpočtu jsem využila metodu nejmenších čtverců. V programu matlab k tomu slouží dvě funkce (*polyval* a *polyfit*). První přímka aproximuje počáteční vývoj hodnot MSE v závislosti na škálovém faktoru a druhá přímka aproximuje koncový vývoj hodnot MSE, jak můžeme vidět na obr. 17 a 18. Hodnoty směrnic těchto přímek jsou uvedeny v tabulce 1.



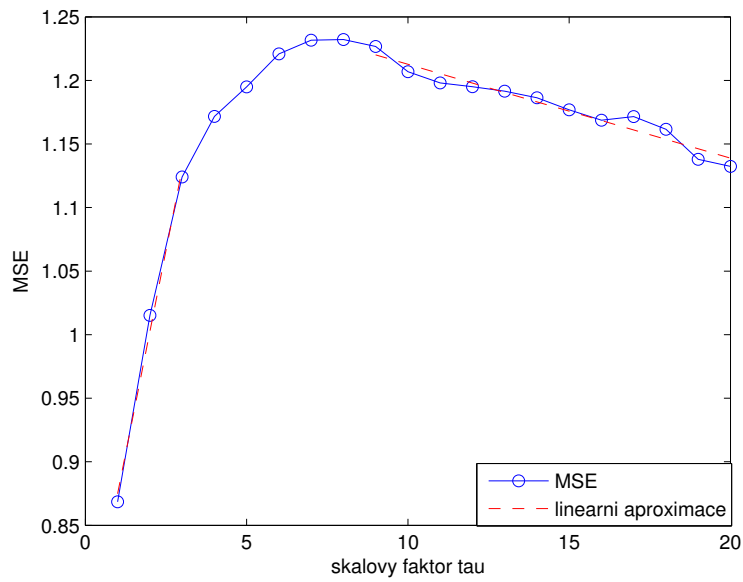
Obrázek 15: Hodnoty MSE EKG č. 3 pro vyšší hodnoty tolerance r .



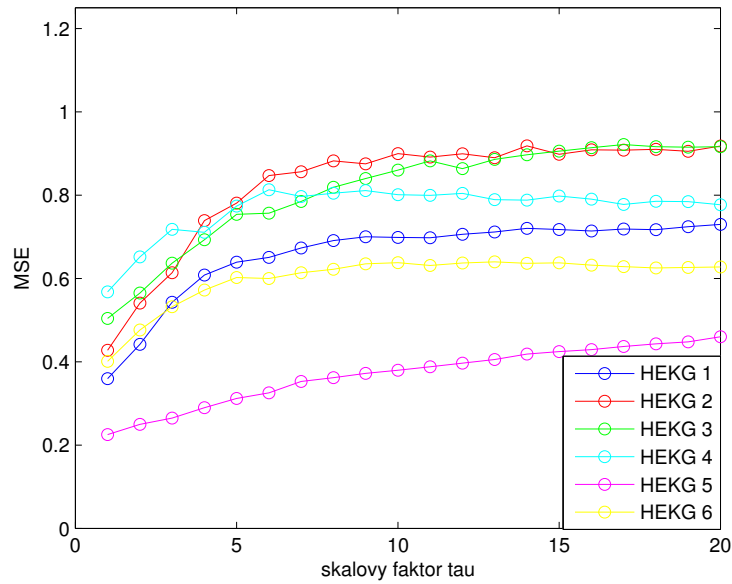
Obrázek 16: Hodnoty MSE EKG č. 4 pro vyšší hodnoty tolerance r .



Obrázek 17: Hodnoty MSE EKG č. 1 a lineární aproximace počátečních a koncových hodnot.

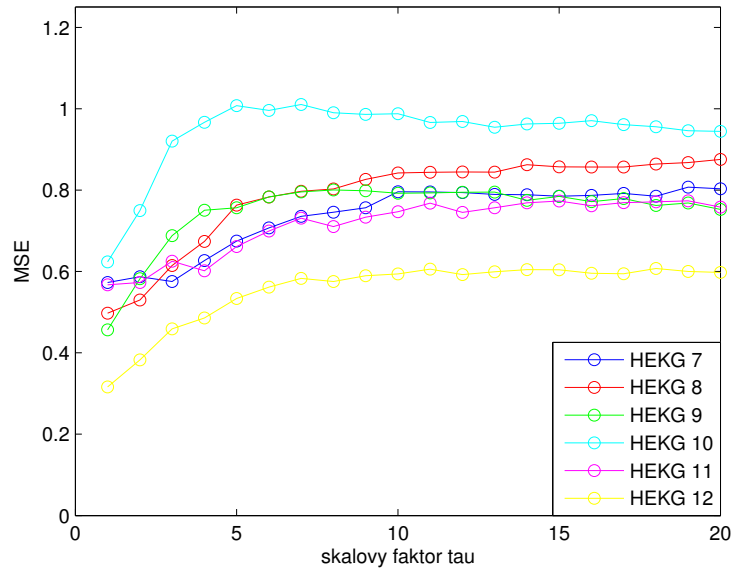


Obrázek 18: Hodnoty MSE EKG č. 2 a lineární aproximace počátečních a koncových hodnot.

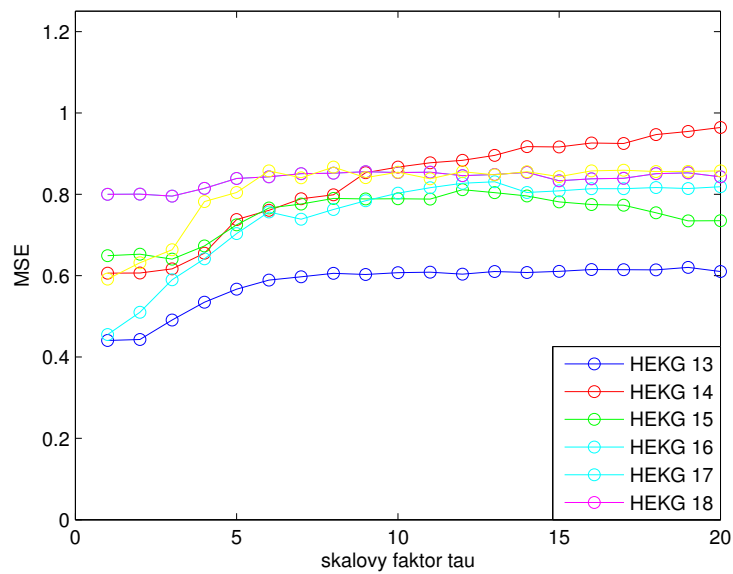


Obrázek 19: Znázornění hodnot MSE zdravých EKG signálů.

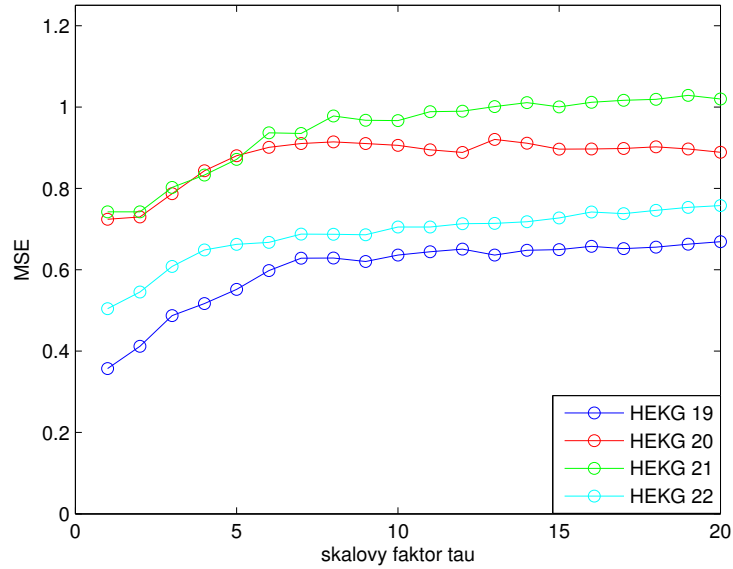
Stejnou aplikaci metody MSE s výše uvedenými parametry jsem provedla se souborem EKG signálů bez synkopy - zdravé EKG. Výsledné hodnoty v závislosti na škálovém faktoru jsou znázorněny na grafech na obr. 19, 20, 21 a 22. Ze všech hodnot víceškálové entropie jsem vypočítala průměrné hodnoty a střední chyby průměru pro každý škálový faktor τ . Průběh průměrných hodnot MSE je znázorněn v grafu na obr. 23. Dále jsem vypočítala lineární aproximace pro počáteční a koncové úseky hodnot MSE (stejně jako u souboru EKG se synkopou). Směrnice těchto přímk jsou uvedeny také v tabulce 1.



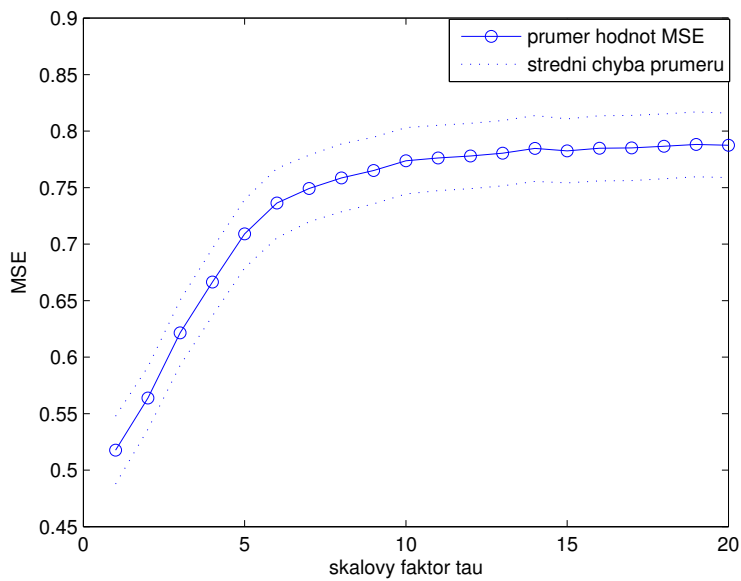
Obrázek 20: Znázornění hodnot MSE zdravých EKG signálů.



Obrázek 21: Znázornění hodnot MSE zdravých EKG signálů.



Obrázek 22: Znázornění hodnot MSE zdravých EKG signálů.



Obrázek 23: Znázornění průměrné hodnoty a střední chyby průměru souboru EKG bez synkopy.

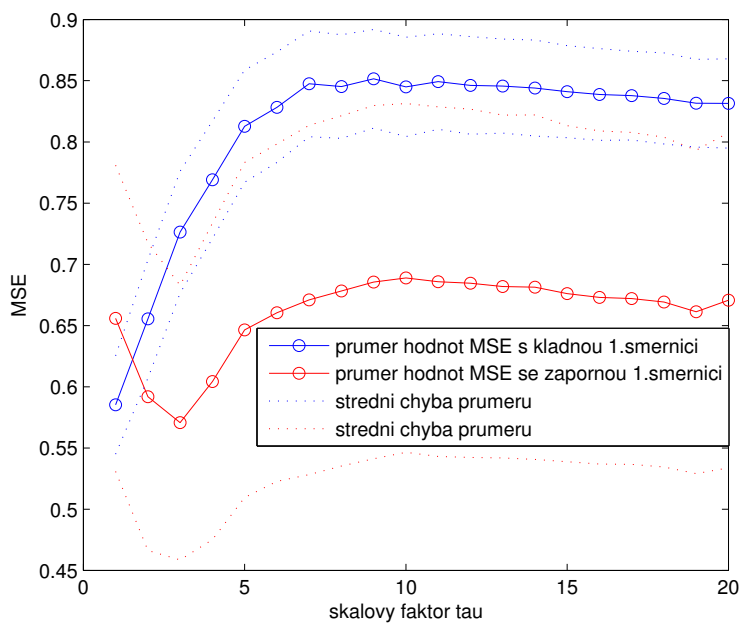
EKG	EKG se synkopou		EKG bez synkopy	
	1. směrnice	2. směrnice	1. směrnice	2. směrnice
1	0.0413	0.0078	0.0848	0.0022
2	0.1277	-0.0074	0.1007	0.0019
3	0.1502	-0.0135	0.0639	0.0059
4	0.0393	-0.0024	0.0453	-0.0025
5	-0.0187	0.0008	0.021	0.0074
6	0.0675	0.0066	0.057	-0.0018
7	0.0084	-0.0049	0.0151	0.0015
8	0.0512	0.0002	0.0614	0.0033
9	0.0405	0.0059	0.0988	-0.0047
10	-0.0344	-0.0005	0.12	-0.0023
11	0.1208	-0.0058	0.0137	0.0017
12	0.0202	-0.0084	0.0585	0.0004
13	0.0598	-0.0052	0.033	0.0012
14	0.0878	-0.0092	0.0157	0.0094
15	0.0691	-0.0051	0.0059	-0.0101
16	0.0831	-0.0057	0.0639	-0.0009
17	0.0778	-0.0025	0.0039	0
18	0.0624	-0.0013	0.0605	0.0008
19	-0.0565	-0.0046	0.0556	0.0029
20	-0.0582	-0.0046	0.0414	-0.0014
21	0.0449	0.0035	0.0332	0.0039
22	0.0568	0.0043	0.0496	0.0061

Tabulka 1: Hodnoty směrnic přímek lineárních aproximací počátečních a koncových hodnot MSE.

Z tabulky 1 je patrné, že 1. směrnice u souboru EKG bez synkopy je vždy kladná. Proto jsem rozdělila datový soubor se synkopou na dvě kategorie podle znaménka 1. směrnice na:

- **Kladná 1. směrnice:** EKG č. 1, 2, 3, 4, 6, 7, 8, 9, 11, 12, 13, 14, 15, 16, 17, 18, 21, 22,
- **Záporná 1. směrnice:** EKG č. 5, 10, 19, 20.

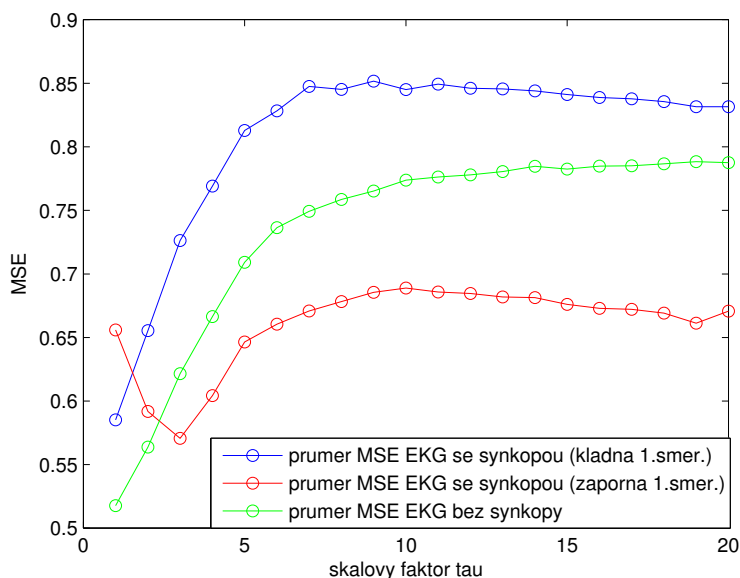
V těchto kategoriích jsem spočítala průměrné hodnoty a střední chyby průměru MSE pro jednotlivé škálové faktory τ . Tyto hodnoty jsou znázorněny na grafu na obr. 24.



Obrázek 24: Znázornění průměrných hodnot a směrodatných odchylek souboru EKG se synkopou rozděleného do dvou kategorií.

Jedním z hlavních cílů a přínosů mé diplomové práce je srovnání hodnot MSE souboru zdravých EKG signálů a souboru EKG se synkopou. Z tohoto důvodu jsem vytvořila graf na obr. 25. Zde jsou vykresleny průměrné hodnoty víceškálové entropie zdravých EKG signálů a průměrné hodnoty EKG signálů se synkopou rozdělené do dvou kategorií podle znaménka 1. směrnice.

Na obr. 25 můžeme pozorovat jistou podobnost dvou skupin signálů - EKG signály bez synkopy (znázorněny zelenou barvou) a EKG signály se synkopou s kladnou 1. směrnici – rostoucí počáteční vývoj hodnot MSE (znázorněné modrou barvou). Tuto podobnost dále otestuji pomocí statistického testování hypotéz s využitím Kruskal-Wallisova testu. Pokud se podaří prokázat podobnost modré a zelené skupiny EKG signálů, můžeme EKG signály z modré skupiny (s kladnou 1. směrnici) považovat za nerozlišitelné od zdravých. Hodnoty víceškálové entropie těchto signálů vzrůstají pro malé hodnoty časových škál ($\tau = 1, \dots, 7$) a následně se stabilizují na téměř konstantní hodnotu. Oproti tomu druhá skupina EKG signálů se synkopou (znázorněna červenou barvou) vykazuje do hodnoty $\tau = 3$



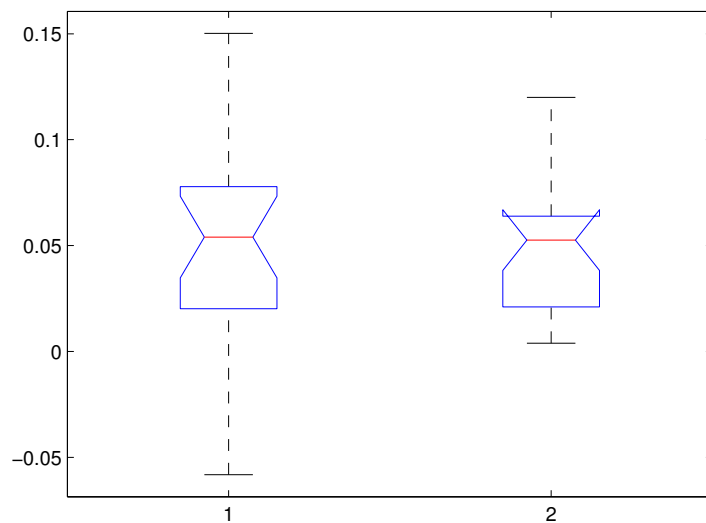
Obrázek 25: Srovnání průměrných hodnot souboru EKG se synkopou a bez synkopy.

klesající charakter. Dále lze pozorovat, že pro hodnoty škálového faktoru $\tau = 2, \dots, 20$ se hodnoty MSE EKG signálů s kladnou 1.směrnicí se synkopou (modré) pohybují nad hodnotami MSE EKG signálů se synkopou se zápornou 1. směrnicí (červené). Budu tedy testovat hypotézu, že hodnoty víceškálové entropie zdravých EKG signálů jsou vyšší ve srovnání s patologickými EKG signály (pacienti se synkopou z červené skupiny).

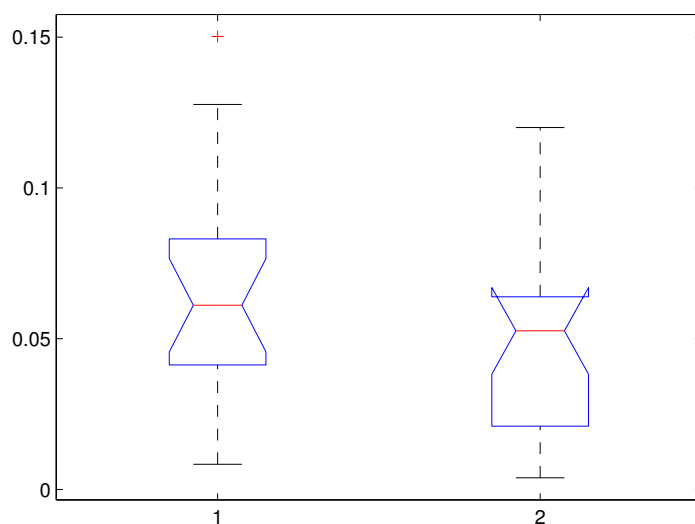
1.5.3. Statistické testování hodnot MSE reálných EKG signálů

V závěru této kapitoly uvedu statistické testování hypotéz popsanych v předchozím odstavci. Veškeré testy jsou provedeny na 5%-ní hladině významnosti. Budu využívat neparametrické testy (rozložení hodnot nevykazuje normální rozložení pravděpodobnosti) k porovnávání mediánů zkoumaných souborů (Kruskal-Wallisův test) a k porovnání rozptylů daných souborů (Levenův test). K testování jsem využila program Matlab.

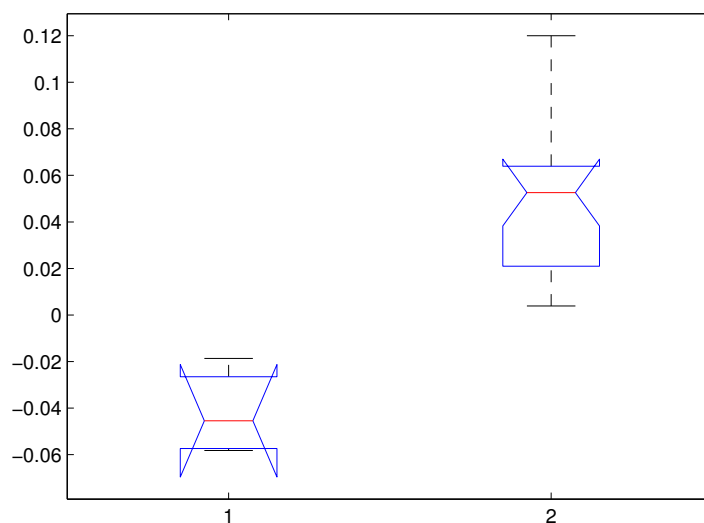
Abychom měli lepší představu o rozložení hodnot, jež budu testovat, vytvořila jsem krabicové diagramy hodnot 1. směrnice křivky MSE (viz tabulka 1).



Obrázek 26: Krabicový diagram hodnot 1.směrnice EKG signálů se synkopou (1) a bez synkopy (2).



Obrázek 27: Krabicový diagram hodnot 1.směrnice EKG signálů se synkopou s kladnou 1.směrnicí (1) a bez synkopy (2).



Obrázek 28: Krabicový diagram hodnot 1.směrnice EKG signálů se synkopou se zápornou 1.směrnicí (1) a bez synkopy (2).

Krabicový diagram skupiny EKG signálu se synkopou (č. 1) a souboru EKG bez synkopy (č. 2) je znázorněn na obr. 26, kde můžeme vidět velký rozptyl hodnot pro EKG signály se synkopou oproti Zdravým EKG signálům. Krabicový diagram skupiny EKG signálu se synkopou s kladnou 1. směrnicí(č. 1) a souboru EKG bez synkopy (č. 2) je znázorněn na obr. 27, kde je patrný vyrovnanější rozptyl hodnot v obou skupinách. Krabicový diagram skupiny EKG signálu se synkopou se zápornou 1. směrnicí (č. 1) a souboru EKG bez synkopy (č. 2) je znázorněn na obr. 28.

Aby bylo možné použít Kruskal-Wallisův test pro otestování rovnosti mediánů, je nutné otestovat rovnost rozptylů souborů pomocí Levenova testu. Pro Levenův test jsem použila funkci *Levenestest* (převzatá z oficiálních stránek programu Matlab, viz [15]). Pro Kruskal-Wallisův test byla použita funkce programu Matlab – *kruskalwallis*. Výsledky Levenova a Kruskal-Wallisova testu jsou p-hodnoty, které porovnám s hladinou významnosti. Pokud je p-hodnota menší než hladina významnosti, zamítáme nulovou hypotézu na hladině významnosti.

Nejprve se budu věnovat oběma celým zkoumaným souborům – EKG signálům

se synkopou a EKG signálům bez synkopy. Testuji nulovou hypotézu, že rozptyl hodnoty 1.směrnice přímký lineární aproximace prvních hodnot MSE souboru EKG se synkopou se rovná rozptylu těchto hodnot souboru EKG bez synkopy. Výsledná p-hodnota Levenova testu je

$$p = 0.0755 > 0.05, \quad (19)$$

tedy nulovou hypotézu o rovnosti rozptylů nelze zamítnout a lze použít Kruskal-Wallisův test k otestování shody mediánů zkoumaných souborů. Nyní tedy testujeme nulovou hypotézu, že medián hodnoty 1.směrnice přímký lineární aproximace prvních hodnot MSE souboru EKG se synkopou se rovná mediánu těchto hodnot souboru EKG bez synkopy. Výsledná p-hodnota Kruskal-Wallisova testu je

$$p = 0.851 > 0.05, \quad (20)$$

tedy nulovou hypotézu o rovnosti mediánů souboru EKG se synkopou a bez synkopy nelze zamítnout.

Nyní se zaměřím na testování souboru EKG se synkopou s kladnou 1.směrnicí lineární aproximace prvních hodnot MSE a souboru EKG signálů bez synkopy. Nejprve opět otestuji rovnost rozptylů. Výsledná p-hodnota Levenova testu je

$$p = 0.7074 > 0.05, \quad (21)$$

tedy nulovou hypotézu o rovnosti rozptylů nelze zamítnout. Dále jsem stejně otestovala rovnost mediánů zkoumaných souborů. Výsledná p-hodnota Kruskal-Wallisova testu je

$$p = 0.1574 > 0.05, \quad (22)$$

tedy nulovou hypotézu o rovnosti mediánů souboru EKG se synkopou s kladnou 1.směrnicí a bez synkopy nelze zamítnout.

Poslední testování jsem provedla na souboru EKG se synkopou se zápornou 1.směrnicí lineární aproximace prvních hodnot MSE a souboru EKG signálů bez synkopy. P-hodnota Levenova testu o rovnosti rozptylů je

$$p = 0.3383 > 0.05, \quad (23)$$

tedy nulovou hypotézu o rovnosti rozptylů nelze zamítnout. A konečně p-hodnota Kruskal-Wallisova testu o rovnosti mediánů vyšla

$$p = 0.0018 < 0.05, \quad (24)$$

z čehož plyne, že nulovou hypotézu o rovnosti mediánů souboru EKG se synkopou se zápornou 1. směrnici a souboru EKG bez synkopy zamítám na hladině významnosti 0.05.

MSE	Levenův test		Kruskal-Wall. test	
	p-hod.	H0	p-hod.	H0
se synkopou X bez synkopy	0.0755	nezam.	0.851	nezam.
kladná 1.směr. X bez synkopy	0.7074	nezam.	0.1574	nezam.
záporná 1.směr. X bez synkopy	0.3383	nezam.	0.0018	zam.

Tabulka 2: Přehled statistického testování pro hodnoty 1. směrnice lineární aproximace prvních hodnot MSE.

Aby bylo testování kompletní, otestovala jsem statistickou podobnost hodnot 2.směrnice – směrnice lineární aproximace posledních hodnot MSE (viz tabulka 1). Krabicový diagram těchto hodnot je uveden na obr. 43 v kapitole Přílohy. Výsledné p-hodnoty testů jsou uvedeny v tabulce 3 a potvrdily statisticky významný rozdíl 2. směrnice souboru EKG signálů se synkopou a bez synkopy.

MSE	Levenův test		Kruskal-Wall. test	
	p-hod.	H0	p-hod.	H0
se synkopou X bez synkopy	0.1437	nezam.	0.0183	zam.

Tabulka 3: Přehled statistického testování pro hodnoty 2. směrnice lineární aproximace posledních hodnot MSE.

Závěrem tohoto testování je, že se neprokázala statisticky významná rozdílnost EKG signálů se synkopou s kladnou 1. směrnici lineární aproximace prvních hodnot MSE a EKG signálů bez synkopy. Pacienty z této skupiny EKG signálů se synkopou lze považovat za nerozlišitelné od zdravých. Tímto jsem také dokázala hypotézu stanovenou v úvodu celé kapitoly, že hodnoty MSE jsou vyšší pro

zdravé EKG signály než pro jisté patologické signály. Tento výsledek je velice zajímavý a mohl by v budoucnu pomoci s využitím metody víceškálové entropie v praktické medicíně.

2. Detrended fluctuation analysis (DFA)

V této kapitole provedu odvození a aplikaci další nelineární metody, jež se využívá k hodnocení EKG signálů - *Detrended fluctuation analysis* (zkráceně DFA). V první části kapitoly popíši matematický princip a odvození této metody, pomocí kterého jsem implementovala funkci v programu Matlab. A v závěru kapitoly budu celý postup aplikovat na srovnání dvou skupin EKG signálů.

Současné studie, zabývající se analýzou EKG signálů, odhalují, že pod vlivem normálních podmínek kolísání srdečního rytmu vykazuje určitý druh dlouhých korelací. Oproti tomu, EKG signály pacientů s jistými kardiologickými potížemi (např. těžké městnavé srdeční selhání) ukazují selhání těchto dlouhých korelací. DFA rozlišuje, které kolísání signálu jsou způsobeny vlivem vnějších podnětů a které naopak odrážejí vnitřní dynamiku systému.

DFA je nová metoda určená k vyčíslení vlastnosti korelace v nestacionárních fyziologických časových řadách. DFA zkoumá časové řady, z nichž jsou lokálně odstraněny trendy pomocí metody klouzavých průměrů.

Metoda využívá tzv. power law metody (mocninné zákony), pomocí které se posuzuje charakter časové řady opakující se na různých časových škálách – tzv. soběpodobnost či fraktální struktura. Výsledkem celé analýzy je škálovací koeficient, který ukazuje na přítomnost sebepodobnosti. DFA je široce používána k rozlišování zdravých a patologických datových souborů, viz [3] a [6].

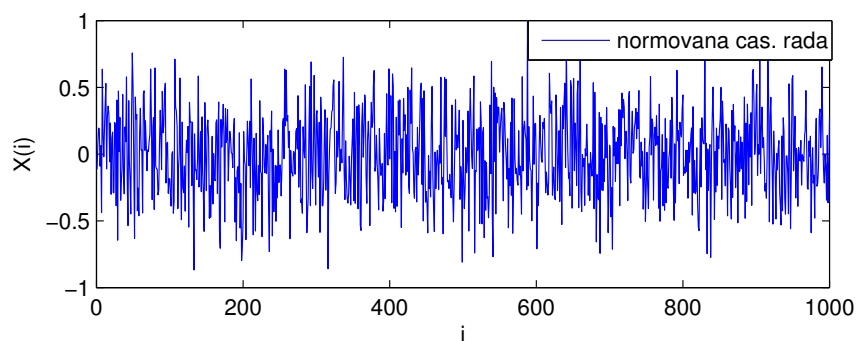
2.1. Výpočet DFA

V tomto odstavci popíši matematické odvození metody DFA, k čemuž využívám zejména poznatky od autora této metody Chung-Kang Penga, z odborného článku [6].

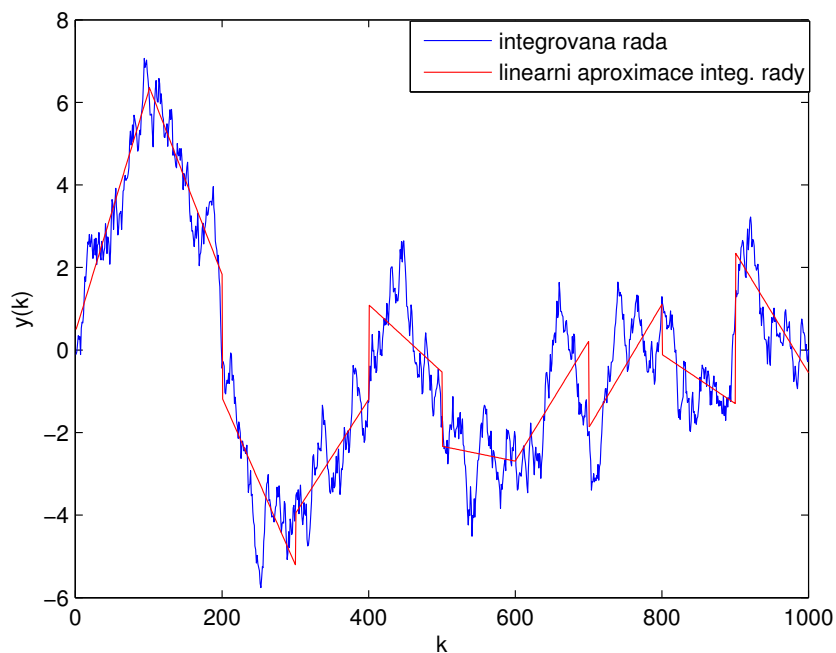
Uvažujme časovou řadu $X(i), i = 1, \dots, n$ délky n , viz obr. 29. Vypočítáme integrovanou řadu

$$y(k) = \sum_{i=1}^k (X(i) - \bar{X}), \quad (25)$$

kde \bar{X} je střední hodnota signálu. Dále rozdělíme integrovanou časovou $\{y(k)\}$ řadu na nepřekrývající se okna délky $\tau \in \mathbb{N}$ (stejně jako u metody MSE v podkapitole 1.4.1). V každém okně najdeme lineární aproximaci integrované řady $y_\tau(k)$ pomocí metody nejmenších čtverců (ta reprezentuje trend v každém okně), jak můžeme vidět na obr. 30.



Obrázek 29: Normovaná časová řada $X(i)$.



Obrázek 30: Znárodnění integrované řady a její lineární aproximace v jednotlivých oknech.

Nyní odstraníme lineární trend (tzv. detrendování) z integrované řady $\{y(k)\}$

odečtením lokálního trendu $y_\tau(k)$ v každém okně. Průměrná fluktuace $F(\tau)$ signálu okolo trendu je poté dána vztahem

$$F(\tau) = \sqrt{\frac{1}{n} \sum_{k=1}^n (y(k) - y_\tau(k))^2}. \quad (26)$$

Tento výpočet opakujeme pro všechny požadované časové škály τ . DFA vyjadřuje závislost $F(\tau)$ (průměrné kolísání signálu) na hodnotě τ . Obecně lze říci, že $F(\tau)$ se zvyšuje s rostoucím τ .

Pokud vyneseme tuto závislost s logaritmickým měřítkem na obou osách, dostaneme lineární závislost, kterou můžeme charakterizovat tzv. *škálovým faktorem* α (sklon lineární závislosti). Hodnoty parametru α a jeho význam se rozděluje následovně:

- $\alpha < 0.5$ – anti-korelovaný signál,
- $\alpha = 0.5$ – nekorelovaný signál (bílý šum),
- $\alpha > 0.5$ – pozitivní autokorelace v signálu,
- $\alpha = 1$ – růžový šum,
- $\alpha = 1.5$ – hnědý šum (náhodná procházka)

2.2. Implementace funkce DFA v programu Matlab

V tomto odstavci popíši postup implementace programového kódu funkce DFA v programu Matlab, který jsem vytvořila na základě matematického postupu popsaného v předchozí podkapitole. Algoritmus budu aplikovat na základní druhy šumů, abych otestovala, zda hodnoty koeficientu α odpovídají hodnotám z literatury. Programový kód s podrobnými komentáři je uveden na příloženém CD.

Funkce *DFA.m* pracuje pouze s jedním vstupním parametrem, kterým je normovaný signál (zadaný jako sloupcový vektor) a jako výstup uvádí dva parametry:

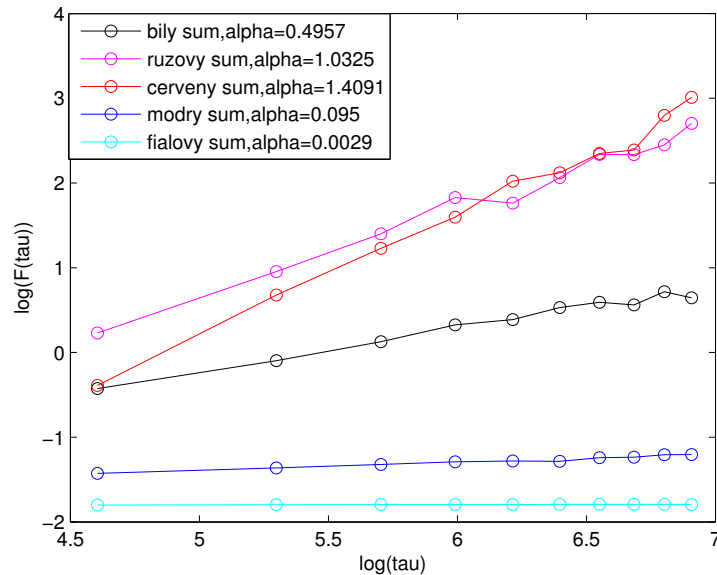
- α – škálový faktor uvádějící sklon lineární závislosti,
- $F(\tau)$ – průměrná fluktuace signálu.

Ve funkci *DFA.m* nejprve definujeme vektor hodnot $\tau = 100 : 100 : 1000$. Následně je volána vnořená funkce *dfa_vypocet.m* se vstupními parametry:

- *signal* - normovaný signál (zadaný při volání funkce *DFA.m*),
- τ - jednotlivé hodnoty vektoru τ .

Tato funkce vypočítá integrovanou řadu $y(k)$ a pomocí implementovaných funkcí *polyfit* a *polyval* vypočítá lineární aproximaci integrované řady. Konečným krokem je výpočet průměrné fluktuace signálu $F(\tau)$ podle vztahu (26). Výstupem této podfunkce jsou právě hodnoty $F(\tau)$.

Nejprve aplikuji tuto metodu na základní druhy šumů, jež jsem uvedla v kapitole 1.4.2. Výslednou logaritmickou závislost $F(\tau)$ na τ i se škálovými koeficienty α můžeme pozorovat na obr. 31. Hodnoty škálového koeficientu α korelují s hodnotami uvedenými v předchozí podkapitole.



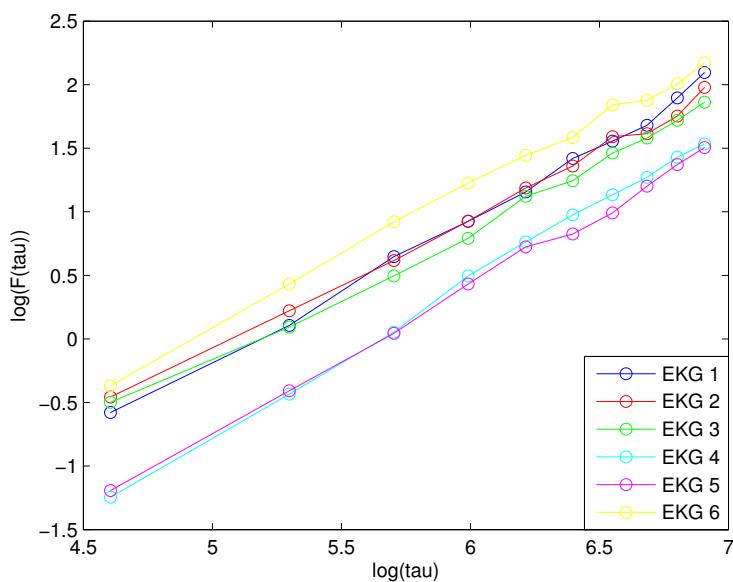
Obrázek 31: Logaritmická závislost metody DFA pro různé druhy šumů.

2.3. Výsledné hodnoty DFA reálných EKG záznamů

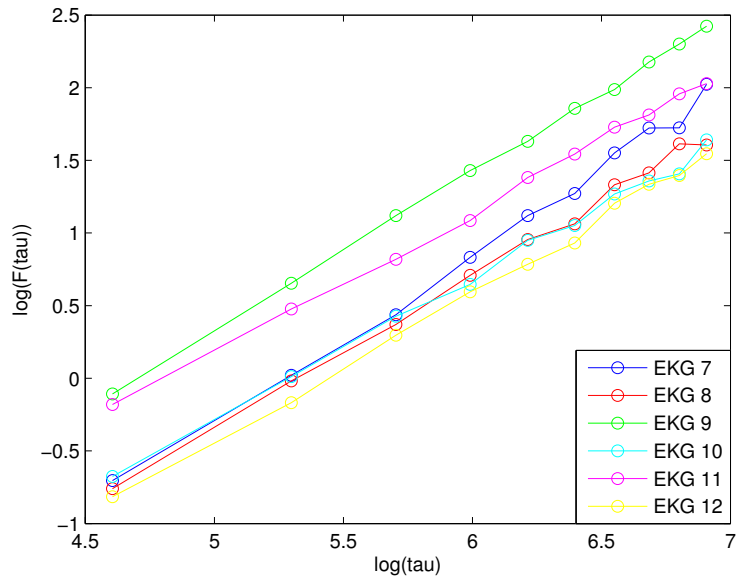
V této kapitole uvedu postup aplikace metody DFA a výsledné hodnoty obou databázových souborů EKG záznamů, tedy 22 pacientů se synkrou a 22 pacientů bez synkopy (popsané v podkapitole 1.5.1). V závěru odstavce otestuji podobnost souborů podle hodnot škálového koeficientu α .

V rámci této metody nemusíme zadávat žádné specifické parametry, vše je totiž předdefinované uvnitř implementované funkce, jak bylo popsáno v předchozím odstavci.

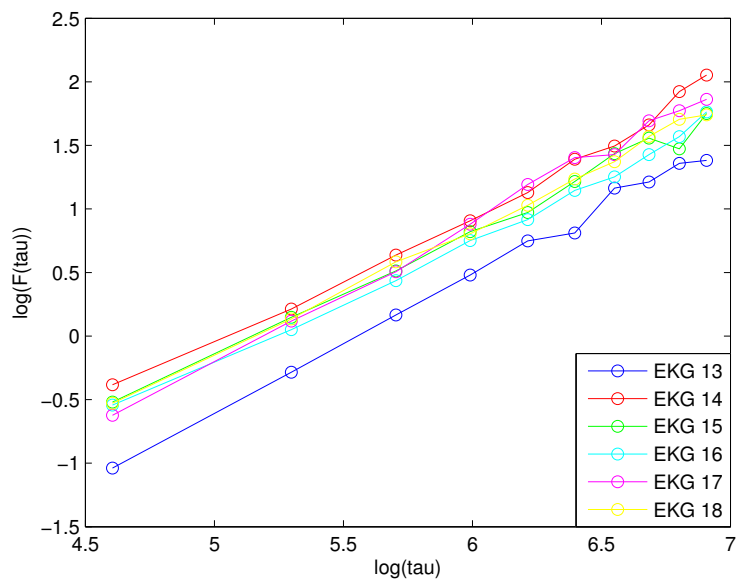
Výstupem této metody je lineární závislost logaritmované hodnoty průměrné fluktuace signálu $F(\tau)$ na logaritmu škálového faktoru τ a především směrnice této lineární závislosti – koeficient α . Tyto výsledné lineární závislosti pro oba soubory pacientů jsou znázorněny na obr. 32 – 39 a hodnoty škálového koeficientu α jsou uvedeny v tabulce 4.



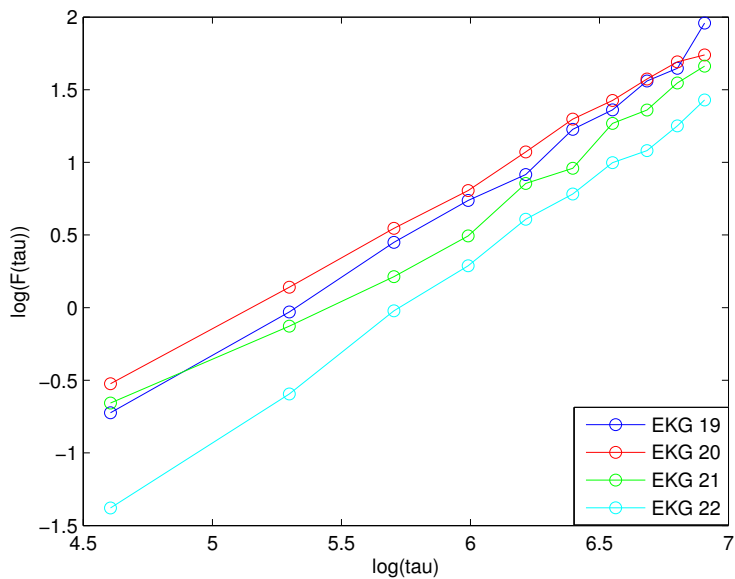
Obrázek 32: DFA EKG signálů se synkrou.



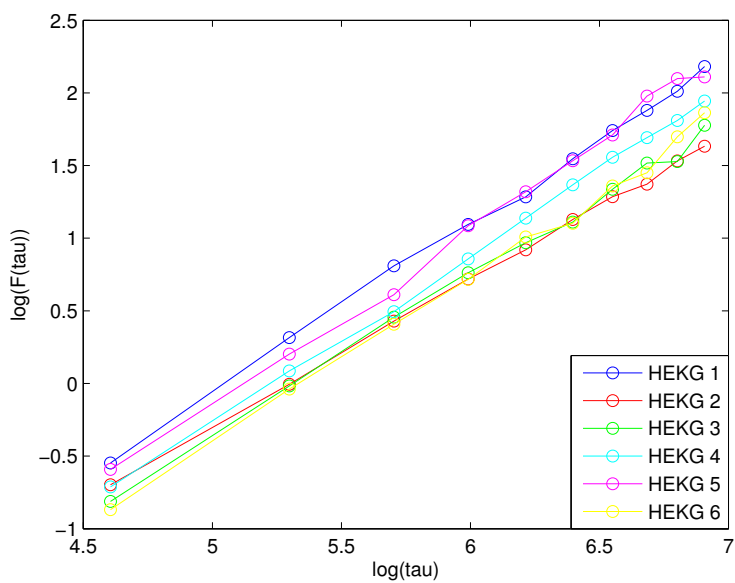
Obrázek 33: DFA EKG signálů se synkopou.



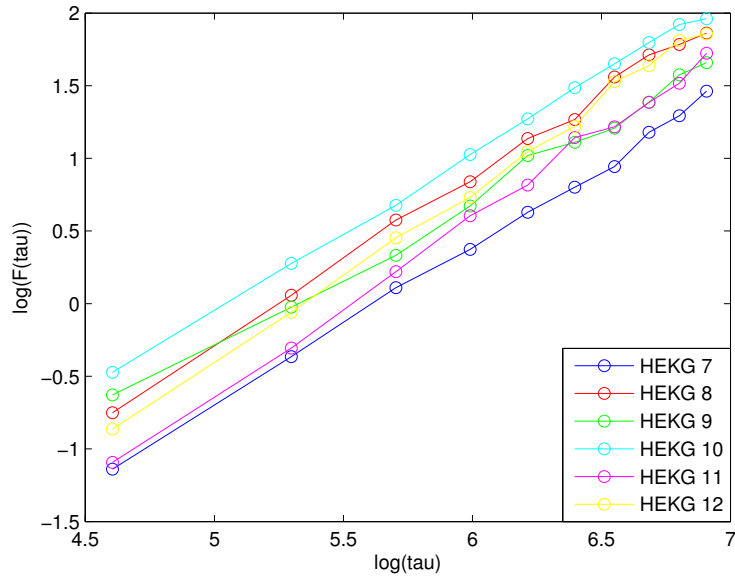
Obrázek 34: DFA EKG signálů se synkopou.



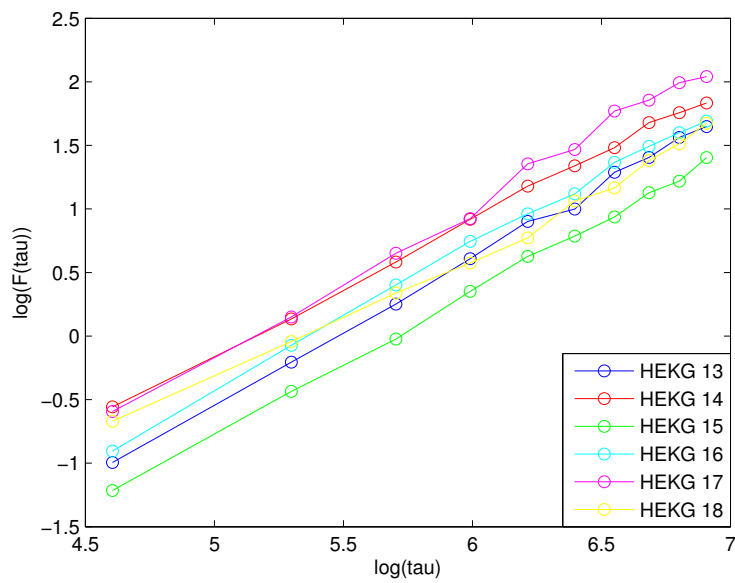
Obrázek 35: DFA EKG signálů se synkopou.



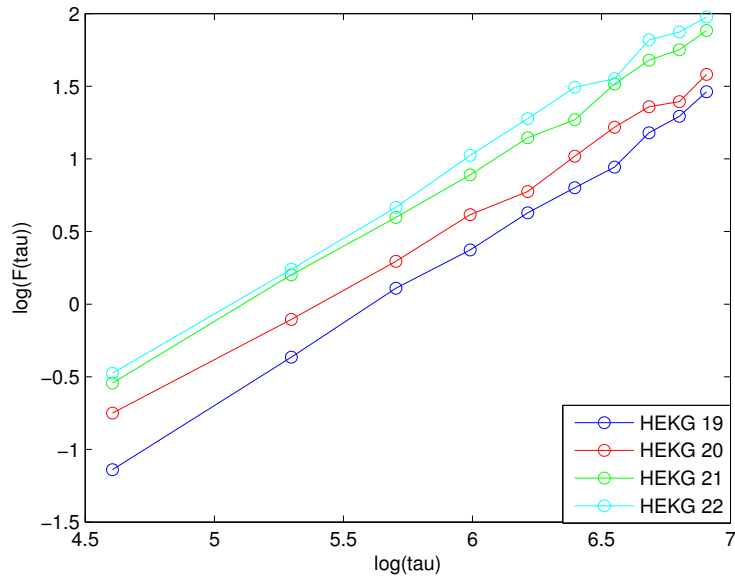
Obrázek 36: DFA EKG signálů bez synkopy.



Obrázek 37: DFA EKG signálů bez synkopy.



Obrázek 38: DFA EKG signálů bez synkopy.



Obrázek 39: DFA EKG signálů bez synkopy.

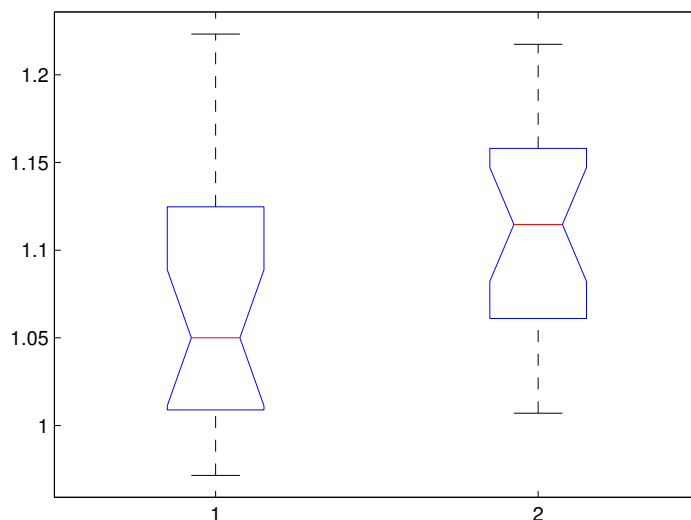
EKG se synkopou				EKG bez synkopy			
EKG	alpha	EKG	alpha	EKG	alpha	EKG	alpha
1	1.0147	12	1.0302	1	1.1579	12	1.67
2	1.0387	13	1.0749	2	1.0104	13	1.158
3	1.0397	14	1.0497	3	1.0895	14	1.061
4	1.2232	15	0.9716	4	1.1615	15	1.1244
5	1.1614	16	0.9825	5	1.74	16	1.1293
6	1.0816	17	1.0997	6	1.1489	17	1.1891
7	1.1737	18	1.0001	7	1.1049	18	1.007
8	1.0504	19	1.1247	8	1.1514	19	1.1049
9	1.0913	20	1.0089	9	1.0077	20	1.0129
10	0.9795	21	1.0353	10	1.0836	21	1.0505
11	0.9755	22	1.212	11	1.22	22	1.0812
průměr		1.0702		průměr		1.1123	
SE		0.0158		SE		0.0141	

Tabulka 4: Hodnoty koeficientu α všech EKG signálů.

2.4. Statistické testování hodnot DFA reálných EKG signálů

V této podkapitole se budu věnovat statistickému zhodnocení výsledků metody DFA zkoumaných souborů. Stejně jako v podkapitole 1.5.3 jsem k testování využila matematický program Matlab a v něm implementované funkce pro Levenův a Kruskal-Wallisův test.

Z tabulky 4 jsem nevyzozorovala žádnou podobnost hodnot mezi koeficienty α metody DFA souboru EKG se synkopou a bez synkopy. Z tohoto důvodu provedu statistické testování pouze pro celé zkoumané soubory.



Obrázek 40: Krabicový diagram hodnot koeficientu α EKG signálů se synkopou (1) a bez synkopy (2).

Na obr. 40 je znázorněný krabicový diagram hodnot koeficientu α pro soubor EKG se synkopou (č. 1) a soubor EKG bez synkopy (č. 2).

Nejprve jsem testovala hypotézu o rovnosti rozptylů daných souborů pomocí Levenova testu. Nulovou hypotézou je tvrzení, že rozptyl hodnot koeficientu α metody DFA souboru EKG se synkopou je roven rozptylu koeficientu α souboru EKG bez synkopy. Výsledná p - hodnota testu je

$$p = 0.6477 > 0.05, \quad (27)$$

tedy nulovou hypotézu o rovnosti rozptylů nelze zamítnout na 5%-ní hladině významnosti. Dále jsem otestovala rovnost mediánů zkoumaných souborů s využitím Kruskal-Wallisova testu. P-hodnota tohoto testu je rovna

$$p = 0.0514 > 0.05, \quad (28)$$

tedy nulovou hypotézu o rovnosti mediánů koeficientu α souboru EKG se synkopou a bez synkopy nelze zamítnout na 5%-ní hladině významnosti. Ovšem tato p-hodnota je hraniční, neboť není mnohonásobně větší než hladina významnosti (líší se pouze o 0.0014).

DFA	Levenův test		Kruskal-Wall. test	
	p-hod.	H0	p-hod.	H0
se synkopou X bez synkopy	0.6477	nezam.	0.0514	nezam.
kladná 1.směr. X bez synkopy	0.8374	nezam.	0.057	nezam.
záporná 1.směr. X bez synkopy	0.342	nezam.	0.7788	nezam.

Tabulka 5: Přehled statistického testování hodnot koeficientu α metody DFA.

Sice se mi nepodařila prokázat přímá souvislost použitelnosti metody DFA k odlišení zdravých pacientů a pacientů se synkopou s možnými budoucími kardiologickými potížemi, ale p-hodnoty prvních dvou testovaných skupin jsou hraniční, což může být způsobeno malým množstvím dat.

3. Srovnání metod MSE a DFA s časovou a spektrální analýzou posloupnosti RR intervalů

V této kapitole stručně popíši základní myšlenky časové a spektrální analýzy RR intervalů EKG signálů, čímž jsem se podrobně zabývala ve své bakalářské práci, viz [5]. Provedu statistické testování na 5%-ní hladině významnosti. Cílem této kapitoly je zjištění faktu, jestli je možné skupinu EKG se synkopou nějakým způsobem rozdělit na dvě kategorie, z nichž jedna bude statisticky významně odlišná od EKG signálů zdravých lidí.

3.1. Časová analýza

Základní myšlenkou této metody je identifikace časové polohy QRS komplexů, z nichž se následně vypočítají RR intervaly. Časovou analýzu můžeme rozdělit na dvě podkategorie – *statistickou a geometrickou*. *Geometrická metoda* využívá grafické znázornění RR intervalů, oproti tomu *statistická metoda* užívá frekvenční charakteristiky založené na statistických aparátech. Mezi základní charakteristiky časové analýzy patří:

- AVNN (ms): průměr všech RR intervalů,
- SDNN (ms): směrodatná odchylka všech RR intervalů,
- SDANN (ms): směrodatná odchylka průměru RR intervalů ve všech 5-minutových segmentech 24-hodinového záznamu,
- RMSSD (ms): odmocnina průměru druhých mocnin rozdílů mezi po sobě jdoucími RR intervaly,
- PNN50 (%): procentuální zastoupení rozdílů mezi po sobě jdoucími NN intervaly, které se liší o více než 50 ms.

Podrobné zpracování metody a vypočítané hodnoty jsou uvedeny v mé bakalářské práci, viz [5]. V tabulce 6 jsou uvedeny p-hodnoty Levenova (pro testování rovnosti rozptylů) a Kruskal-Wallisova testu (pro testování rovnosti mediánů) pro

časovou analýzu řady RR intervalů souboru se synkopou a bez synkopy. Krabicové diagramy hodnot pro jednotlivé charakteristiky jsou uvedeny na obr. 44 – 48 v kapitole Přílohy.

Časová analýza	Levenův test		Kruskal-Wallisův test	
	p-hodnota	H0	p-hodnota	H0
AVNN	0.0304	zamítáme	0.5468	nezamítáme
SDNN	0.383	nezamítáme	0.0471	zamítáme
SDANN	0.5153	nezamítáme	0.5648	nezamítáme
RMSSD	0.0068	zamítáme	0.00005	zamítáme
PNN50	0.0006	zamítáme	0.0002	zamítáme

Tabulka 6: Přehled statistického testování charakteristik časové analýzy.

Pouze u jedné charakteristiky nelze zamítnout hypotézu o rovnosti mediánů – u SDANN. Jelikož u ostatních charakteristik je rozdíl obou skupin EKG signálů statisticky významný, není metoda časové analýzy vhodná pro rozlišení EKG signálů se synkopou na dvě kategorie, z nichž jedna by byla statisticky významně rozdílná zdravé skupině EKG signálů.

Časová analýza	Levenův test		Kruskal-Wallisův test	
	p-hodnota	H0	p-hodnota	H0
AVNN	0.7694	nezamítáme	0.1732	nezamítáme
SDNN	0.3415	nezamítáme	0.4959	nezamítáme
SDANN	0.5892	nezamítáme	0.9987	nezamítáme
RMSSD	0.1793	nezamítáme	0.0887	nezamítáme
PNN50	0.9201	nezamítáme	0.2016	nezamítáme

Tabulka 7: Přehled statistického testování charakteristik časové analýzy souboru EKG se synkopou rozděleného na EKG se zápornou a kladnou 1. směrnicí.

V tabulce 7 jsou dále uvedeny p-hodnoty testů pouze pro soubor EKG se synkopou, který jsem rozdělila na dvě podkategorie podle výsledků zpracovaných v kapitole Víceškalové entropie – na EKG s kladnou a zápornou 1. směrnicí lineární aproximace počátečních hodnot MSE. Časová analýza ukázala, že tyto skupiny EKG signálů jsou si statisticky podobné. Tedy pomocí této metody nedosáhneme

stejného rozdělení EKG se synkopou na dvě podkategorie, jak tomu je u metody MSE.

3.2. Spektrální analýza

Spektrální analýza převede RR intervaly do frekvenční oblasti, přičemž frekvenční spektrum srdečního rytmu lze zaznamenat v několika oblastech. Pro nás jsou nejdůležitější oblast vysoké (HF) a nízké frekvence (LF). Pomocí metody Lombova periodogramu jsem vypočítala spektrální charakteristiky, z nichž pro nás byla nejvýznamnější hodnota poměru LF/HF. Na základě této hodnoty jsem usoudila, že pacienti s EKG signály č. 2, 9, 16, 17 a 18 by neměli mít kardiologické potíže v budoucím životě. Celý tento postup je podrobně popsán v mé bakalářské práci, viz [5].

Pozoruhodným zjištěním a také přínosem této práce je fakt, že všechny EKG signály, jež jsem v rámci bakalářské práce zkonstatovala za zdravé, patří do skupiny EKG se synkopou s kladnou 1.směrnici – s rostoucím počátečním vývojem hodnot MSE. Tedy patří do skupiny, u které jsem usoudila, že má průběh MSE jako zdravé EKG signály.

Podrobné zpracování metody a vypočítané hodnoty jsou uvedeny v mé bakalářské práci, viz [5]. V tabulce 8 jsou uvedeny p-hodnoty Levenova (pro testování rovnosti rozptylů) a Kruskal-Wallisova testu (pro testování rovnosti mediánů) pro spektrální analýzu řady RR intervalů souboru se synkopou a bez synkopy. Krabicové diagramy hodnot pro jednotlivé charakteristiky jsou uvedeny na obr. 49 – 51 v kapitole Přílohy.

Spektrální analýza	Levenův test		Kruskal-Wallisův test	
	p-hodnota	H0	p-hodnota	H0
LF	0.0003	zamítáme	0.00001	zamítáme
HF	0	zamítáme	0	zamítáme
LF/HF	0.1115	nezamítáme	0.0002	zamítáme

Tabulka 8: Přehled statistického testování charakteristik spektrální analýzy.

V tabulce 9 jsou dále uvedeny p-hodnoty testů pouze pro soubor EKG se synkopou, který jsem rozdělila na dvě podkategorie podle výsledků zpracovaných v kapitole Víceškálová entropie – na EKG s kladnou a zápornou 1. směrníci lineární aproximace počáteční závislosti MSE. Spektrální analýza neukázala, že tyto skupiny EKG signálů jsou významně statisticky rozdílné. Tedy pomocí této metody nedosáhneme stejného rozdělení EKG se synkopou na dvě podkategorie, jak tomu je u metody MSE.

Spektrální analýza	Levenův test		Kruskal-Wallisův test	
	p-hodnota	H0	p-hodnota	H0
LF	0.1353	nezamítáme	0.2334	nezamítáme
HF	0.0599	nezamítáme	0.1058	nezamítáme
LF/HF	0.1115	nezamítáme	0.0738	nezamítáme

Tabulka 9: Přehled statistického testování charakteristik spektrální analýzy souboru EKG se synkopou rozděleného na EKG se zápornou a kladnou 1. směrníci.

Závěr

Cílem mé práce bylo nastudování a podrobný popis dvou nelineárních metod, víceškálové entropie a DFA, jenž se využívají mimo jiné při analýze variability srdeční frekvence. Obě metody jsem implementovala v programu Matlab podle matematického odvození. Použitelnost algoritmů obou metod jsem ilustrovala na jednoduchých umělých signálech s různou kombinací parametrů, protože je důležité rozumět tomu, jaký vliv mají jednotlivé parametry na výsledek. Tento vliv je dobré pochopit na umělých signálech, které máme zcela pod kontrolou.

Stěžejním přínosem mé diplomové práce je aplikování a srovnání metod MSE a DFA na reálných datech pacientů se synkopou a kontrolní skupině zdravých pacientů.

U metody víceškálové entropie se mi podařilo rozdělit soubor EKG pacientů se synkopou na dvě podskupiny podle počátečního průběhu hodnot MSE v závislosti na škálovém faktoru. U první skupiny (obsahující 18 pacientů), vykazující rostoucí počáteční vývoj hodnot MSE, se neprokázala rozdílnost s kontrolní skupinou zdravých pacientů, což bylo dokázáno pomocí statistického testování hypotéz na 5%-ní hladině významnosti. U druhé skupiny pacientů se synkopou s klesajícím počátečním vývojem hodnot MSE (obsahující 4 pacienty) se rozdílnost s kontrolní skupinou potvrdila.

Dále se mi podařilo prokázat, že hodnoty MSE zdravých pacientů jsou vyšší než hodnoty MSE patologických EKG signálů pro hodnoty škálového faktoru $\tau = 2, \dots, 20$, jak lze vidět na grafu na obr. 24.

Pozoruhodným zjištěním je, že mezi 18 EKG signály se synkopou, u kterých nebyla prokázána odlišnost od zdravých EKG signálů, se vyskytuje i pět EKG záznamů, které jsem v rámci své bakalářské práce označila za kardiologicky „zdravé“, tedy bez budoucích kardiologických potíží, na základě spektrální analýzy, respektive podle hodnoty poměru LF/HF.

Metoda víceškálové entropie se tedy ukázala jako velmi přínosná pro stanovení rozdílnosti normálních a „jiných“ EKG signálů v souboru pacientů po prodělání synkopy. Konstatování, že u této skupiny „jiných“ EKG signálů by v budoucím

životě mohly nastat kardiologické problémy je mimo rámec této diplomové práce a vyžadovalo by spolupráci s kardiology.

Tyto výsledky jsem se pokusila porovnat s etiologií synkop, ale bohužel datová sada neobsahuje jednoznačnou informaci o tom, zda-li synkopa byla kardiologické etiologie, či nikoliv. Bylo možné jen prozkoumat, zda-li stratifikace pacientů podle MSE nekoreluje s běžnými populačními charakteristikami (věk, pohlaví, ...). Zde žádná závislost nalezena nebyla a zůstává nám tedy otevřená hypotéza, že MSE stratifikuje skupinu pacientů se synkopou podle etiologie dané synkopy. To je hypotéza, jejíž případné potvrzení by mělo přímé klinické důsledky, neboť synkopy, které souvisejí se srdcem, vyžadují lékařskou pozornost a péči.

U metody DFA nelze jednoznačně stanovit použitelnost pro rozdělení EKG signálů se synkopou na dvě skupiny, z nichž jedna by byla statistivky významně odlišná od zdravých pacientů. Lze ovšem pozorovat hraniční p-hodnoty Kruskal-Wallisova testu, které by se mohli lépe projevit u většího datového souboru pacientů se synkopou.

Závěrem lze konstatovat, že tato práce by vyžadovala další otestování obou metod na větším souboru pacientů se synkopou, čímž by se prokázala například použitelnost metody víceškálové entropie k rozlišení EKG signálu se synkopou, a také by bylo možné lépe zhodnotit použitelnost metody DFA. Důležitým faktorem pro zhodnocení je také znalost etiologie synkop. Tato práce by mohla posloužit jako přínosný materiál pro budoucí experimenty a možná bude některá z metod v budoucnu implementována např. do EKG holterů, kde by z naměřeného EKG záznamu pacienta po synkopě mohla stanovit prediktor možných kardiologických problémů.

Přílohy

V této kapitole jsou uvedeny veškeré tabulky a obrázky, na které se odkazují v práci.

tau	EKG 1	EKG 2	EKG 3	EKG 4	EKG 5	EKG 6
1	0.2446	0.8685	0.7335	0.5197	0.2913	0.5986
2	0.2726	1.0152	0.9345	0.5047	0.257	0.6608
3	0.3437	1.124	1.034	0.566	0.2486	0.7415
4	0.4038	1.1717	1.0615	0.6071	0.2316	0.8273
5	0.4469	1.1949	1.0808	0.6509	0.2452	0.8798
6	0.4903	1.2208	1.0619	0.699	0.2474	0.9227
7	0.5117	1.2317	1.0585	0.716	0.2401	0.9496
8	0.5364	1.2321	1.0343	0.7305	0.2445	0.9646
9	0.5611	1.2267	1.0265	0.7448	0.2493	0.9849
10	0.5625	1.2069	1.0021	0.7504	0.255	0.9949
11	0.5798	1.198	0.9916	0.7479	0.2555	1.0087
12	0.5942	1.1951	0.9826	0.751	0.2562	1.0305
13	0.6066	1.1915	0.9549	0.748	0.2604	1.0256
14	0.6111	1.1864	0.9475	0.7393	0.2556	1.0413
15	0.6161	1.1769	0.9292	0.745	0.2612	1.0475
16	0.6249	1.1687	0.9225	0.7385	0.2607	1.0502
17	0.634	1.1716	0.9056	0.7409	0.2619	1.06
18	0.6307	1.1615	0.9038	0.736	0.259	1.0651
19	0.6409	1.1379	0.8921	0.728	0.2618	1.0741
20	0.6497	1.1323	0.8911	0.7265	0.264	1.0781

Tabulka 10: Hodnoty víceškálové entropie EKG signálů se synkopou.

tau	EKG 7	EKG 8	EKG 9	EKG 10	EKG 11	EKG 12
1	0.8481	0.4701	0.4524	0.6609	0.847	0.7036
2	0.8894	0.5513	0.4761	0.581	1.0079	0.6067
3	0.869	0.6231	0.5179	0.5922	1.0887	0.7019
4	0.8531	0.6625	0.5724	0.6404	1.0717	0.6652
5	0.882	0.6986	0.6133	0.6839	1.0786	0.7754
6	0.8992	0.732	0.6429	0.7065	1.0778	0.6393
7	0.9152	0.7444	0.6604	0.7185	1.079	0.7268
8	0.9115	0.7438	0.6711	0.7269	1.0404	0.6276
9	0.9151	0.7546	0.6885	0.7319	1.0213	0.6704
10	0.909	0.7512	0.7012	0.7381	1.0154	0.5923
11	0.9142	0.7604	0.7032	0.7291	1.0244	0.6368
12	0.9069	0.7593	0.7123	0.7364	1.0192	0.5811
13	0.9054	0.7521	0.7172	0.7226	1.0271	0.6137
14	0.9081	0.7602	0.7276	0.7377	1.0062	0.5688
15	0.8921	0.764	0.7315	0.729	0.9982	0.5962
16	0.8924	0.7601	0.7331	0.7415	1.0003	0.575
17	0.8774	0.7523	0.7397	0.7383	0.9869	0.5948
18	0.8783	0.7603	0.7453	0.7411	0.9976	0.5573
19	0.876	0.7522	0.755	0.7195	0.9868	0.5688
20	0.8743	0.7588	0.7606	0.726	0.97	0.5511

Tabulka 11: Hodnoty víceškálové entropie EKG signálů se synkopou.

tau	EKG 13	EKG 14	EKG 15	EKG 16	EKG 17	EKG 12
1	0.643	0.5492	0.4663	0.7099	0.5665	0.5314
2	0.7252	0.6443	0.5547	0.8134	0.7	0.5755
3	0.7929	0.7249	0.6432	0.9009	0.7921	0.6304
4	0.8746	0.7372	0.6979	0.9874	0.8433	0.7147
5	0.9217	0.7695	0.7402	1.0384	0.8839	0.7739
6	0.9682	0.7995	0.7533	1.0554	0.8894	0.8011
7	0.9961	0.8238	0.7655	1.0543	0.8879	0.8242
8	1.0076	0.8157	0.766	1.0427	0.9012	0.8443
9	1.0205	0.7977	0.765	1.038	0.9002	0.8521
10	1.025	0.7892	0.7655	1.0319	0.8911	0.8557
11	1.0242	0.7793	0.7544	1.0283	0.8823	0.862
12	1.0348	0.7703	0.7483	1.0177	0.8789	0.8634
13	1.0255	0.7576	0.7479	1.0158	0.8735	0.8569
14	1.0402	0.7472	0.738	1.0217	0.8843	0.8493
15	1.0258	0.7385	0.7289	1.0039	0.8784	0.8479
16	1.0201	0.7281	0.7174	0.997	0.8816	0.8493
17	1.0108	0.7298	0.7289	0.9903	0.8622	0.8453
18	1.009	0.7169	0.7173	0.9811	0.8805	0.8428
19	0.9983	0.7242	0.7145	0.9724	0.8592	0.8413
20	0.9972	0.6962	0.7148	0.9908	0.874	0.8592

Tabulka 12: Hodnoty víceškálové entropie EKG signálů se synkopou.

tau	EKG 19	EKG 20	EKG 21	EKG 22
1	0.6731	0.9981	0.4365	0.3439
2	0.5642	0.965	0.4924	0.3731
3	0.5602	0.8818	0.5372	0.4414
4	0.585	0.9603	0.5858	0.5068
5	0.6396	1.0168	0.6412	0.5612
6	0.6656	1.0226	0.6697	0.587
7	0.6832	1.0419	0.7025	0.6067
8	0.6935	1.0484	0.7212	0.6219
9	0.7008	1.0599	0.7364	0.6249
10	0.7077	1.0547	0.7459	0.6193
11	0.6996	1.0591	0.765	0.6277
12	0.6903	1.0553	0.7668	0.6176
13	0.6966	1.0481	0.7699	0.6317
14	0.6875	1.0446	0.7737	0.6409
15	0.6824	1.0316	0.7807	0.6386
16	0.6682	1.0215	0.78	0.6586
17	0.6678	1.0205	0.7875	0.6616
18	0.666	1.0105	0.792	0.6631
19	0.6611	1.0022	0.7905	0.6574
20	0.6598	1.033	0.7945	0.6469

Tabulka 13: Hodnoty víceškálové entropie EKG signálů se synkopou.

tau	HEKG 1	HEKG 2	HEKG 3	HEKG 4	HEKG 5	HEKG 6
1	0.3593	0.4276	0.5039	0.5678	0.225	0.4009
2	0.442	0.5409	0.5651	0.6516	0.2497	0.4762
3	0.543	0.6138	0.6368	0.7178	0.2648	0.532
4	0.6081	0.739	0.6931	0.7109	0.2899	0.5722
5	0.639	0.7809	0.7539	0.7737	0.3118	0.6022
6	0.6505	0.8468	0.7568	0.8128	0.3255	0.5999
7	0.6732	0.8561	0.7847	0.7966	0.3526	0.6138
8	0.691	0.8823	0.8189	0.8046	0.3619	0.622
9	0.7001	0.8752	0.84	0.8109	0.372	0.6352
10	0.6984	0.8996	0.8599	0.8011	0.3796	0.6377
11	0.6975	0.8914	0.8821	0.7993	0.3881	0.6313
12	0.7058	0.8992	0.8637	0.8043	0.3965	0.6371
13	0.7115	0.8899	0.886	0.7891	0.4054	0.6397
14	0.7203	0.9181	0.897	0.7881	0.4186	0.6361
15	0.7174	0.898	0.9057	0.7979	0.4242	0.6375
16	0.7141	0.9086	0.9138	0.7904	0.4292	0.632
17	0.7187	0.9078	0.9213	0.7776	0.4366	0.6283
18	0.717	0.9098	0.9163	0.7851	0.4429	0.6253
19	0.7243	0.9052	0.9152	0.7846	0.4476	0.6265
20	0.7298	0.918	0.9163	0.7768	0.46	0.6275

Tabulka 14: Hodnoty víceškálové entropie EKG signálů bez synkopy.

tau	HEKG 7	HEKG 8	HEKG 9	HEKG 10	HEKG 11	HEKG 12
1	0.5727	0.4974	0.4564	0.6235	0.5672	0.3161
2	0.5871	0.5298	0.5822	0.7496	0.5729	0.3826
3	0.5753	0.6145	0.6882	0.9204	0.6252	0.4591
4	0.6269	0.6738	0.7505	0.9665	0.6018	0.4855
5	0.6748	0.7632	0.7566	1.0078	0.6612	0.5334
6	0.7073	0.7829	0.7833	0.9961	0.6996	0.5616
7	0.7356	0.7972	0.7956	1.0106	0.7309	0.5832
8	0.7455	0.803	0.8006	0.9902	0.7104	0.5753
9	0.7566	0.8263	0.7983	0.9863	0.7334	0.5896
10	0.7958	0.8423	0.7924	0.9876	0.747	0.594
11	0.7956	0.8439	0.7931	0.9662	0.7678	0.6059
12	0.7939	0.8446	0.7948	0.9687	0.7451	0.5924
13	0.7895	0.8443	0.7952	0.9545	0.7567	0.5992
14	0.7886	0.8626	0.7747	0.9628	0.7688	0.6044
15	0.7848	0.857	0.7844	0.9642	0.7733	0.604
16	0.787	0.8567	0.7715	0.9706	0.7615	0.5957
17	0.7919	0.8566	0.7795	0.9611	0.7691	0.5944
18	0.7849	0.8641	0.7623	0.9559	0.7713	0.6074
19	0.8074	0.8679	0.7682	0.946	0.7737	0.6002
20	0.8032	0.8756	0.7525	0.9445	0.7581	0.5978

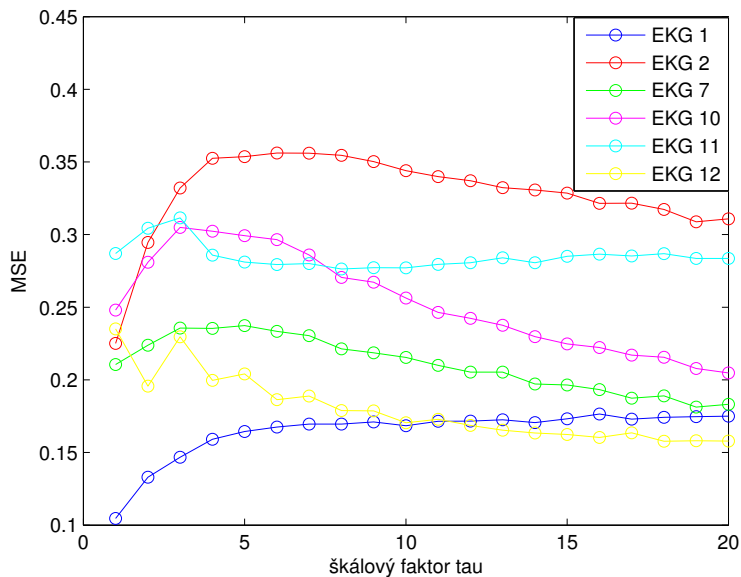
Tabulka 15: Hodnoty víceškálové entropie EKG signálů bez synkopy.

tau	HEKG 13	HEKG 14	HEKG 15	HEKG 16	HEKG 17	HEKG 18
1	0.4408	0.6062	0.6493	0.4551	0.8001	0.5914
2	0.443	0.6067	0.6526	0.5101	0.8004	0.6317
3	0.4911	0.6166	0.6405	0.5899	0.7957	0.6638
4	0.5349	0.6551	0.6729	0.6414	0.8145	0.7822
5	0.5666	0.738	0.7249	0.7034	0.8391	0.8044
6	0.5892	0.7606	0.7663	0.7567	0.8431	0.8577
7	0.5974	0.7889	0.7759	0.7386	0.8509	0.8405
8	0.6056	0.7987	0.79	0.7628	0.8518	0.8669
9	0.6029	0.8531	0.7888	0.784	0.8554	0.841
10	0.6073	0.8671	0.7891	0.8029	0.8534	0.8545
11	0.6087	0.8775	0.7884	0.8162	0.8542	0.8382
12	0.6037	0.8836	0.8113	0.8273	0.846	0.8557
13	0.6102	0.8959	0.804	0.8309	0.8487	0.8479
14	0.6077	0.9174	0.796	0.8043	0.854	0.8557
15	0.6104	0.9166	0.7815	0.8086	0.8331	0.8434
16	0.6149	0.9261	0.775	0.8139	0.8383	0.8575
17	0.6147	0.9251	0.7732	0.8136	0.8393	0.8591
18	0.6142	0.9469	0.7547	0.8165	0.8508	0.856
19	0.6205	0.9545	0.7349	0.8141	0.8529	0.8563
20	0.6101	0.9642	0.7351	0.8188	0.8432	0.8576

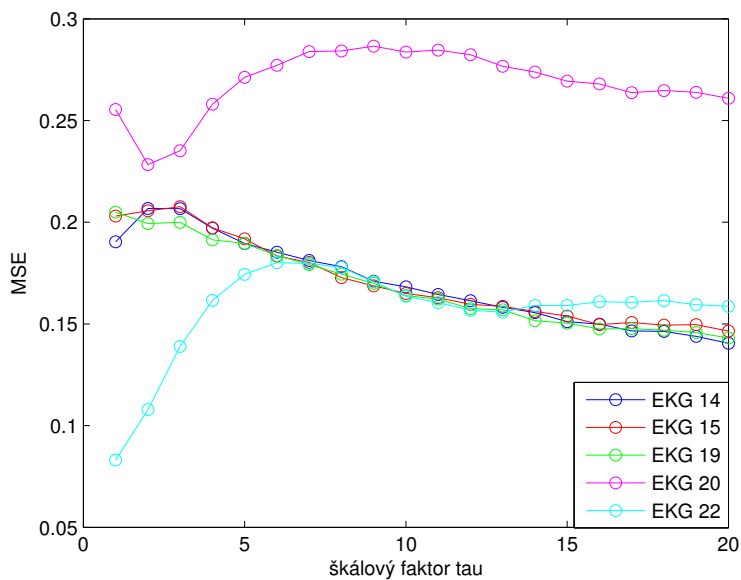
Tabulka 16: Hodnoty víceškálové entropie EKG signálů bez synkopy.

tau	HEKG 19	HEKG 20	HEKG 21	HEKG 22
1	0.357	0.7242	0.7423	0.5045
2	0.412	0.7299	0.7425	0.5452
3	0.4875	0.787	0.8026	0.608
4	0.5171	0.8433	0.833	0.6489
5	0.5516	0.8802	0.8715	0.6626
6	0.5981	0.901	0.9366	0.6673
7	0.6284	0.9104	0.9349	0.6875
8	0.6289	0.9143	0.978	0.6871
9	0.6204	0.9105	0.9674	0.6861
10	0.6361	0.9057	0.9669	0.7049
11	0.6443	0.8951	0.9886	0.7051
12	0.6506	0.8886	0.9897	0.7133
13	0.6362	0.9205	1.001	0.7141
14	0.6478	0.911	1.0111	0.7181
15	0.6495	0.8965	1.0003	0.7274
16	0.6577	0.8971	1.0116	0.7421
17	0.6519	0.8981	1.0166	0.738
18	0.6557	0.9021	1.019	0.7459
19	0.6628	0.8969	1.0287	0.7535
20	0.6689	0.8888	1.0198	0.758

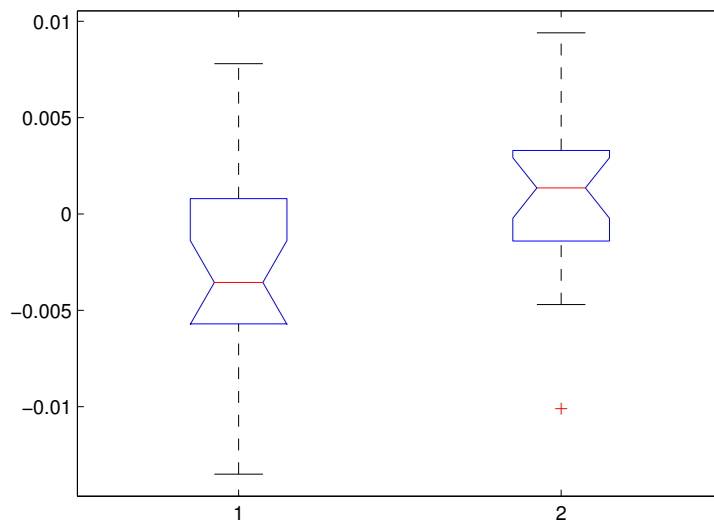
Tabulka 17: Hodnoty víceškalové entropie EKG signálů bez synkopy.



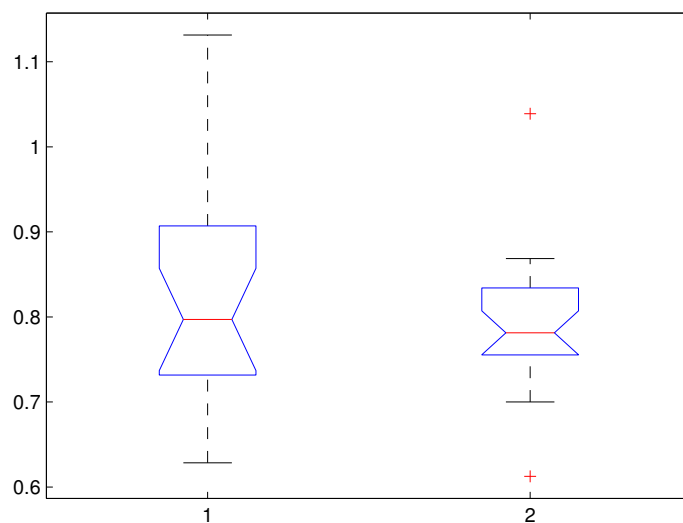
Obrázek 41: Hodnoty MSE vybraných EKG signálů se synkopou pro vyšší toleranci r .



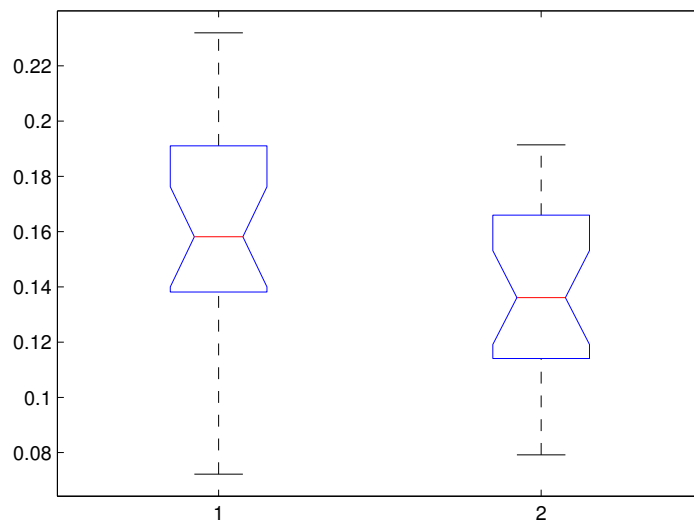
Obrázek 42: Hodnoty MSE vybraných EKG signálů se synkopou pro vyšší toleranci r .



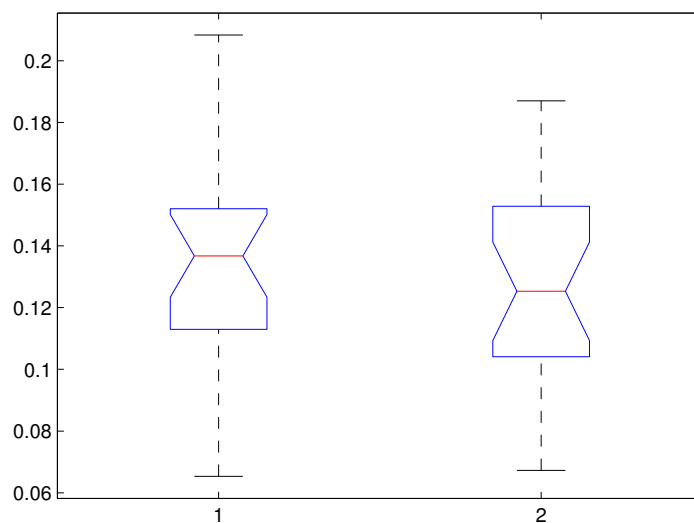
Obrázek 43: Krabicový diagram hodnot 2.směrnice MSE EKG signálů se synkopou (1) a bez synkopy (2).



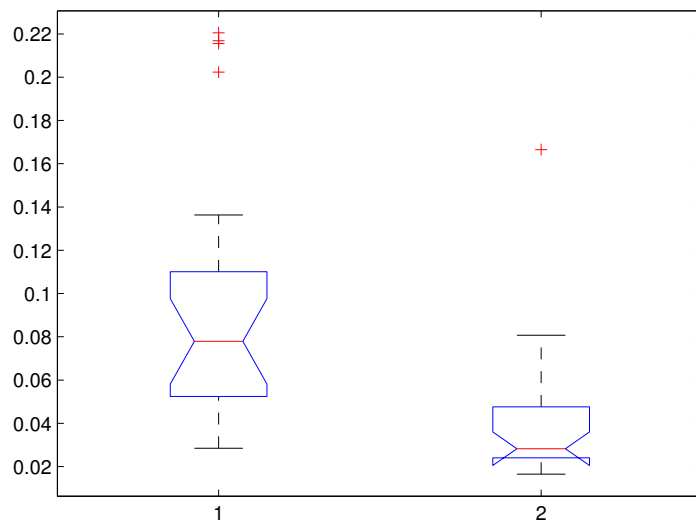
Obrázek 44: Krabicový diagram hodnot časové analýzy – charakteristika AVNN signálů se synkopou (1) a bez synkopy (2).



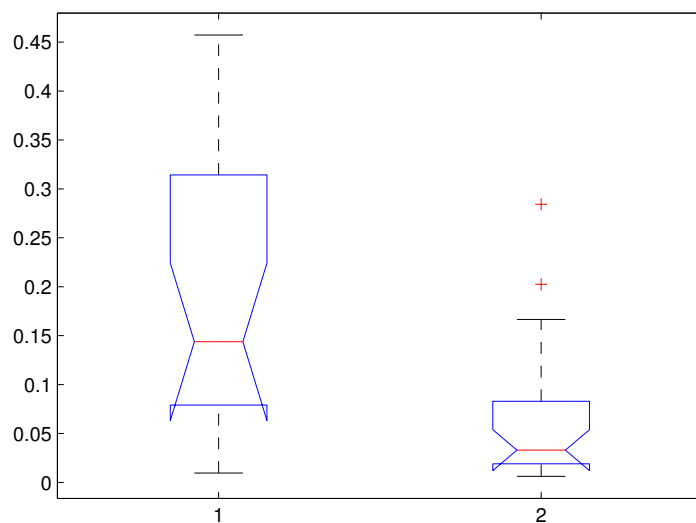
Obrázek 45: Krabicový diagram hodnot časové analýzy – charakteristika SDNN signálů se synkopou (1) a bez synkopy (2).



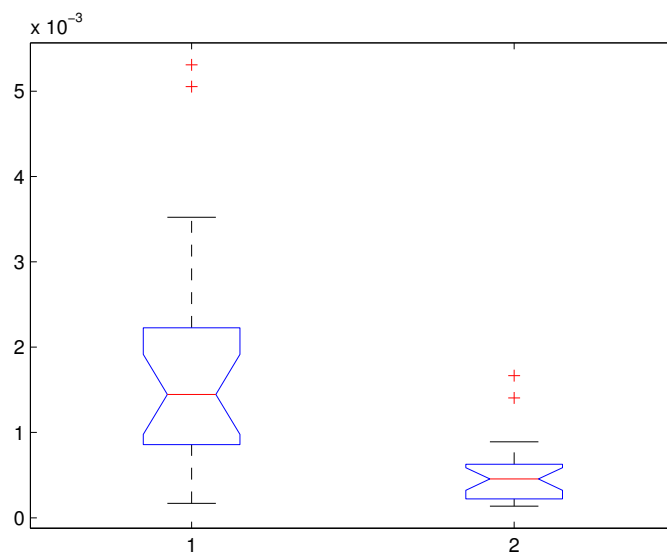
Obrázek 46: Krabicový diagram hodnot časové analýzy – charakteristika SDANN signálů se synkopou (1) a bez synkopy (2).



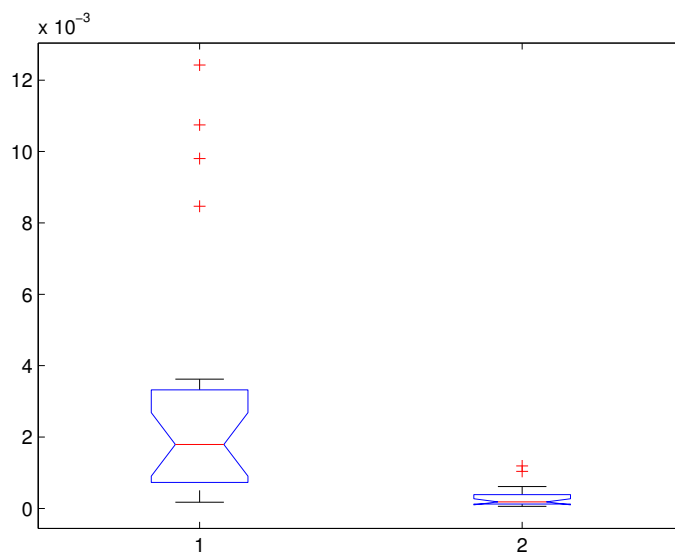
Obrázek 47: Krabicový diagram hodnot časové analýzy – charakteristika RMSSD signálů se synkopou (1) a bez synkopy (2).



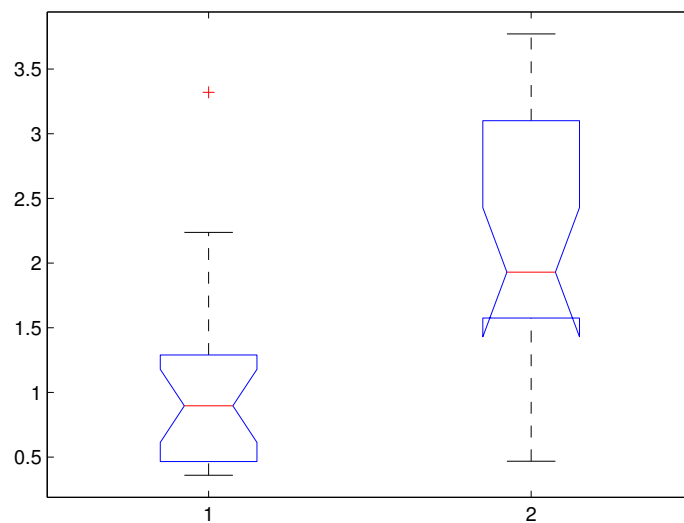
Obrázek 48: Krabicový diagram hodnot časové analýzy – charakteristika PNN50 signálů se synkopou (1) a bez synkopy (2).



Obrázek 49: Krabicový diagram hodnot spektrální analýzy – charakteristika LF signálů se synkopou (1) a bez synkopy (2).



Obrázek 50: Krabicový diagram hodnot spektrální analýzy – charakteristika HF signálů se synkopou (1) a bez synkopy (2).



Obrázek 51: Krabicový diagram hodnot spektrální analýzy – charakteristika LF/HF signálů se synkopou (1) a bez synkopy (2).

Literatura

- [1] Costa M., Goldberger A.L., Peng C.-K. Multiscale entropy analysis of physiologic time series. *Phys Rev Lett* 2002; 89:062102.
- [2] Costa M., Goldberger A.L., Peng C.-K. Multiscale entropy analysis of biological signals. *Phys Rev E* 2005; 71:021906.
- [3] Golinská A.K. Detrended Fluctuation Analysis (DFA) in biomedical signal processing: selected examples. *Studies in Logic, Grammar and Rhetoric*, 2012, 29, pp. 107-115.
- [4] Mackay D. J. C. *Information Theory, Inference, and Learning Algorithms*. Cambridge University Press, 2003.
- [5] Mraznicová, M. Lombův periodogram a jeho použití na frekvenční analýzu EKG signálů. Olomouc, 2013. Bakalářská práce (Bc.). UNIVERZITA PALACKÉHO V OLOMOUCI. Přírodovědecká fakulta.
- [6] Peng C.-K., Havlin S., Golberger A.L. Quantification of scaling exponents and crossover phenomena in nonstationary hearth time series. *Chaos*; 5(1), pp. 82-87,1995.
- [7] Penhaker, M. a kol.: *Lékařské diagnostické přístroje – učební texty*. VŠB TU Ostrava, Ostrava, 2004.
- [8] Richman J, Moorman R. Physiological time-series analysis using approximate entropy and sample entropy. *Am J Physiol Heart Circ Physiol*. 2000; 278, 2039-2049.
- [9] Urbaniková, N. „Entropie.“ *Encyklopedie lingvistiky*, ed. Kateřina Prokopová. Olomouc, Univerzita Palackého v Olomouci, 2013. [online 9. 2. 2015], dostupné z <http://oltk.upol.cz/encyklopedie/index.php5/Entropie>
- [10] En.wikipedia.org: Entropy [online 9. 8. 2014], dostupné z <http://en.wikipedia.org/wiki/Entropy>
- [11] Wikipedia.org: Entropie [online 9. 8. 2014], dostupné z <http://cs.wikipedia.org/wiki/Entropie>
- [12] En.wikipedia.org: Approximate entropy [online 14. 2. 2015], dostupné z http://en.wikipedia.org/wiki/Approximate_entropy
- [13] En.wikipedia.org: Information entropy [online 17. 10. 2014], dostupné z [http://en.wikipedia.org/wiki/Entropy_\(information_theory\)](http://en.wikipedia.org/wiki/Entropy_(information_theory))
- [14] En.wikipedia.org: colors of noise [online 2. 12. 2014], dostupné z http://en.wikipedia.org/wiki/Colors_of_noise

- [15] Mathworks.com: levene test [online 14. 3. 2015], dostupné z <http://www.mathworks.com/matlabcentral/fileexchange/3375-levenetest>
- [16] Physionet.org [online 5. 1. 2015], dostupné z <http://www.physionet.org/physiobank/database/nsr2db/>
- [17] Wikiskripta.eu: Entropie [online 9. 8. 2014], dostupné z <http://wikiskripta.eu/index.php/Entropie>
- [18] tzb-info.cz: Co je entropie [online 2. 11. 2014], dostupné z <http://www.tzb-info.cz/925-co-to-je-entropie>
- [19] Wikipedia.org: Synkopa [online 3. 3. 2012], dostupné z [http://cs.wikipedia.org/wiki/Synkopa_\(nemoc\)](http://cs.wikipedia.org/wiki/Synkopa_(nemoc))