

**UNIVERZITA PARDUBICE  
FAKULTA EKONOMICKO-SPRÁVNÍ  
ÚSTAV SYSTÉMOVÉHO INŽENÝRSTVÍ  
A INFORMATIKY**

**Modelování finančních časových řad  
pomocí vybraného stochastického modelu**

**DIPLOMOVÁ PRÁCE**

**AUTOR PRÁCE: Bc. Eva Srkalová**

**VEDOUCÍ PRÁCE: Ing. Jan Panuš, Ph.D.**

**2012**

**UNIVERSITY OF PARDUBICE  
FACULTY OF ECONOMICS AND ADMINISTRATION  
INSTITUTE OF SYSTEM ENGINEERING  
AND INFORAMTICS**

**Modeling financial time series using the  
selected stochastic model**

**DIPLOMA THESIS**

**AUTHOR: Bc. Eva Srkalová**

**SUPERVISOR: Ing. Jan Panuš, Ph.D.**

**2012**

Univerzita Pardubice  
Fakulta ekonomicko-správní  
Akademický rok: 2010/2011

## ZADÁNÍ DIPLOMOVÉ PRÁCE

(PROJEKTU, UMĚLECKÉHO DÍLA, UMĚLECKÉHO VÝKONU)

Jméno a příjmení: **Bc. Eva SRKALOVÁ**  
Osobní číslo: **E08927**  
Studijní program: **N6209 Systémové inženýrství a informatika**  
Studijní obor: **Informatika ve veřejné správě**  
Název tématu: **Modelování finančních časových řad pomocí vybraného  
stochastického modelu**  
Zadávající katedra: **Ústav systémového inženýrství a informatiky**

### Zásady pro vypracování:

Charakteristika vybrané/vybraných časové/časových řady/řad  
Modelování vybrané/vybraných časové/časových řady/řad  
Závěry získané modelováním

Rozsah grafických prací:

Rozsah pracovní zprávy:

Forma zpracování diplomové práce: **tištěná/elektronická**

Seznam odborné literatury:

**ARLT, Josef, ARLTOVÁ, Markéta.** *Finanční časové řady : Vlastnosti, metody modelování, příklady a aplikace.* Grada Publishing, Praha, 2003. 220 s. ISBN 80-247-0330-0.

**ARLT, Josef, ARLTOVÁ, Markéta.** *Ekonomické časové řady : Vlastnosti, metody modelování, příklady a aplikace.* Grada Publishing, Praha, 2007. 288 s. ISBN 978-80-247-1319-9.

**ČAPEK, Jan.** *Modelování ekonomických a sociálních procesů : pro kombinovanou formu studia.* Univerzita Pardubice, Pardubice, 2006. 103 s. ISBN 80-7194-838-1.

Vedoucí diplomové práce:

  
**Ing. Jan Panuš, Ph.D.**

Ústav systémového inženýrství a informatiky

Datum zadání diplomové práce: **4. října 2010**

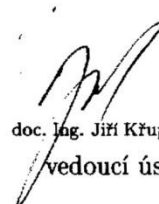
Termín odevzdání diplomové práce: **6. května 2011**



doc. Ing. Renáta Myšková, Ph.D.

děkanka

L.S.

  
doc. Ing. Jiří Křupka, Ph.D.

vedoucí ústavu

V Pardubicích dne 4. října 2010

Prohlašuji:

Tuto práci jsem vypracovala samostatně. Veškeré literární prameny a informace, které jsem v práci využila, jsou uvedeny v seznamu použité literatury.

Byla jsem seznámena s tím, že se na moji práci vztahují práva a povinnosti vyplývající ze zákona č. 121/2000 Sb., autorský zákon, zejména se skutečností, že Univerzita Pardubice má právo na uzavření licenční smlouvy o užití této práce jako školního díla podle § 60odst. 1 autorského zákona, a s tím, že pokud dojde k užití této práce mnou nebo bude poskytnuta licence o užití jinému subjektu, je Univerzita Pardubice oprávněna ode mne požadovat přiměřený příspěvek na úhradu nákladů, které na vytvoření díla vynaložila, a to podle okolností až do jejich skutečné výše.

Souhlasím s prezenčním zpřístupněním své práce v Univerzitní knihovně Univerzity Pardubice.

V Pardubicích dne 19. 3. 2012

Bc. Eva Srkalová

## **Poděkování**

Děkuji panu Ing. Janu Panušovi, Ph.D., vedoucímu práce, za odborné vedení, cenné připomínky a odborné rady, trpělivost a pomoc při zpracování diplomové práce.

## **Abstrakt**

Diplomová práce se zabývá modelováním vybraných finančních časových řad pomocí vybraného stochastického procesu a jejich předpovědí. Jsou zde přiblíženy termíny jako model a modelování. V další kapitole jsou popsány časové řady, finanční časové řady, jejich předpoklady a charakteristiky. Čtvrtá část popisuje Boxovu-Jenkinsovu metodologii, její konstrukci a modely stacionárních a nestacionárních časových řad. Závěr této práce je věnován experimentální části, která se zabývá modelováním indexu S&P 500 a kurzu GBP/USD pomocí procesů ARIMA. Cílem této diplomové práce je nalezení vhodného modelu a následná predikce daných finančních časových řad.

## **Summary**

Diploma thesis deals with the modeling of financial time series using the selected stochastic model and its predictions. Terms are approximated as a model and modeling. The next chapter describes the time series, financial time series, their assumptions and characteristics. The fourth section describes the methodology of Box-Jenkins, its structure and models of stationary and non-stationary time series. The conclusion of this work is devoted to the experimental section, which deals with modeling using the index S&P 500 and course of GBP/USD using ARIMA processes. The aim of this thesis is to find a suitable model and subsequent prediction of the financial time series.

## **Klíčová slova**

Modelování, finanční časové řady, Box-Jenkinsova metodologie, stochastický proces, AR, MA, ARMA, ARIMA, predikce

## **Keywords**

Modeling, financial time series, Box-Jenkins methodology, stochastic process, AR, MA, ARMA, ARIMA, prediction

# Obsah

<b>1</b>	<b>Úvod .....</b>	<b>9</b>
<b>2</b>	<b>Model, modelování.....</b>	<b>11</b>
2.1	Model.....	11
2.2	Klasifikace modelů .....	11
2.3	Modelování .....	12
<b>3</b>	<b>Časové řady .....</b>	<b>14</b>
3.1	Problémy analýzy časových řad .....	15
3.1.1	Problémy s volbou časových bodů pozorování.....	15
3.1.2	Problémy s kalendářem .....	16
3.1.3	Problémy s délkou časových řad.....	16
3.1.4	Problém s nesrovnatelností jednotlivých měření .....	16
3.2	Základní přístupy k analýze časových řad.....	16
3.2.1	Dekompozice časové řady.....	16
3.2.2	Boxova-Jenkinsova (B-J) metodologie .....	19
3.2.3	Lineární kauzální (faktorové) modely.....	19
3.2.4	Spektrální analýza časových řad .....	19
3.3	Předpovědi v časových řadách.....	20
3.4	Finanční časové řady, jejich předpoklady a vlastnosti .....	20
3.4.1	Hypotéza efektivního trhu.....	21
3.4.2	Jednoduché výnosy aktiva.....	21
3.4.3	Předpoklad normality .....	23
3.4.4	Předpoklad linearity .....	24
3.4.5	Pákový efekt.....	24
3.4.6	Předpoklad autokorelace .....	25
3.4.7	Předpoklad heteroskedasticity.....	25
3.4.8	Vliv mikrostruktury trhu na vlastnosti finančních časových řad ...	25
<b>4</b>	<b>Box-Jenkinsova metodologie.....</b>	<b>27</b>
4.1	Základní pojmy B-J metodologie .....	27
4.1.1	Stochastický proces .....	27
4.1.2	Stacionarita.....	28



4.1.3	Autokorelační funkce (ACF).....	28
4.1.4	Parciální autokorelační funkce (ACF) .....	29
4.1.5	Proces bílého šumu .....	30
4.1.6	Lineární proces.....	30
4.2	Lineární modely stacionárních časových řad.....	31
4.2.1	MA – proces klouzavých součtů .....	31
4.2.2	AR – autoregresní proces .....	35
4.2.3	ARMA – smíšený proces .....	37
4.3	Lineární modely nestacionárních časových řad.....	37
4.3.1	Proces náhodné procházky („RandomWalk Proces“). .....	38
4.3.2	Procesy ARIMA.....	38
4.4	Modely stacionárních a nestacionárních sezonních časových řad.....	40
4.5	Konstrukce modelu v Box-Jenkinsově metodologii.....	40
<b>5</b>	<b>Experimentální část .....</b>	<b>42</b>
5.1	Index S&P 500 (Standard & Poor’s 500) .....	43
5.1.1	Identifikace modelu.....	44
5.1.2	Odhad parametrů modelu a následná verifikace .....	50
5.1.3	Prezentace dosažených výsledků .....	55
5.1.4	Predikce finanční časové řady indexu S&P 500 pomocí modelu ..	56
5.2	Vývoj kurzu GBP/USD .....	57
5.2.1	Identifikace modelu.....	57
5.2.2	Odhad parametrů modelu a následná verifikace .....	62
5.2.3	Prezentace dosažených výsledků .....	67
5.2.4	Predikce finanční časové řady kurzu GBP/USD pomocí modelu..	67
<b>6</b>	<b>Závěr .....</b>	<b>69</b>
	<b>Použitá literatura .....</b>	<b>71</b>
	<b>Seznam grafů.....</b>	<b>73</b>
	<b>Seznam obrázků.....</b>	<b>74</b>
	<b>Seznam tabulek .....</b>	<b>74</b>
	<b>Seznam zkratk .....</b>	<b>75</b>

# 1 Úvod

Nalezení vhodného modelu je základním cílem modelování časových řad. Vybraný vhodný model nám následně umožní porozumět mechanismu, na kterém vznikly hodnoty dané časové řady a pochopit vazby a podmínky, které působily na vznik těchto hodnot. Na základě změn těchto podmínek nebo vazeb lze modelovat jejich vliv na změny ve vývoji časové řady. Jedním z dalších cílů modelování časových řad je využití získaných poznatků při předpovědích budoucího chování takovýchto řad.

Předmětem této diplomové práce je modelování dat z finančního trhu pomocí časových řad a testování principů, což nám umožní posoudit, zda experimentálně získaná data vyhovují předpokladu, který jsme před provedením testování učinili. Pro modelování byly vybrány dvě finanční časové řady, a to konkrétně denní vývoj indexu S&P 500 za rok 2011 a měnového kurzu GBP/USD od počátku roku 2003 do konce roku 2011.

Pojmy model a modelování jsou popsány ve druhé kapitole. Samotným časovým řadám je věnována kapitola 3, kde jsou vysvětleny jednotlivé přístupy, které se dnes využívají v oblasti analýz ekonomických a finančních řad. Jsou zde popsány problémy analýzy časových řad, základní přístupy k analýze časových řad, předpovědi v časových řadách a především finanční časové řady, jejich předpoklady a vlastnosti. V části zaměřující se na předpoklady finančních časových řad, jsou objasněny následující předpoklady – normalita, linearita, pákový efekt, autokorelace a heteroskedasticita. Nedílnou součástí je i vysvětlení vlivu mikrostruktury trhu na některé vlastnosti finančních časových řad.

Jelikož v současné době má v této oblasti analýzy ekonomických a finančních řad nejpropracovanější přístup k charakteristice, popisu a modelování těchto řad Box-Jenkinsova metodologie, je jí věnována celá čtvrtá kapitola. V první řadě je tato metodologie popsána, je vysvětlen její princip, výhody a nevýhody. Především jsou zde popsány základní pojmy této metodologie - stochastický proces, stacionarita, autokorelační funkce, parciální autokorelační funkce, proces bílého šumu a lineární proces. V neposlední řadě jsou zde vysvětleny a popsány lineární modely stacionárních časových řad v jednotlivých řádech, jako například AR, MA i smíšený model ARMA, dále lineární modely nestacionárních časových řad, což je proces náhodné procházky a proces ARIMA, a také je zde zmínka o modelech stacionárních a nestacionárních sezonních časových řad. Konečná část této kapitoly se věnuje konstrukci modelu dané metodologie, která je identifikace modelu, odhad parametrů modelu a jeho verifikace.

V našem případě je modelování provedeno na reálných datech námi vybraných nestacionárních finančních časových řad s konkrétními charakteristickými vlastnostmi, pro něž je vhodné použít smíšené procesy ARIMA. Na základě pochopení jednotlivých principů jejich vlastností a vazeb mezi nimi je na konkrétních příkladech prokázáno zvolení vhodného modelu a následně jeho ověření buď subjektivními metodami či statistickými testy. V konečné fázi jsou získané výsledky zapsány do formy rovnice vhodného modelu a ukázána jejich předpověď do budoucna. U burzovního indexu je predikce na 10 dní a u měnového kurzu na 1 rok. Tomuto se věnujeme v páté kapitole této práce s názvem experimentální část.

Veškeré námi provedené analýzy, testy a následné předpovědi jsou zpracovány v programu STATISTICA.

## 2 Model, modelování

### 2.1 Model

Z latinského „modulus“ (míra, takt) ve spojení s „modus“ (kopie, vzor) pochází slovo model. [8]

Pomocí některých citací z publikací ze současné doby vymezíme pojem model.

- „*Struktura symbolů a operačních pravidel, o kterých se předpokládá, že se shodují se sadou relevantních faktů (points) v existující struktuře nebo procesu*“. [11]
- „*Teoretická a zjednodušená reprezentace reálného světa. Je to izomorfní konstrukce reality nebo očekávané reality*“. [6]
- „*...Především ke stejné situaci můžeme najít velmi mnoho matematických modelů. Člověk si z nich musí vybrat ten nejvhodnější – nejpriléhavější k dané situaci. Model nesmí být příliš složitý, abychom s ním mohli matematicky pracovat. Nemusí se v každém směru podobat realitě, ..., stačí, když dá věrný opis reality v tom směru, který je důležitý pro danou úlohu. ...Hledání odpovídajícího matematického modelu přivede člověka k hlubšímu pochopení zkoumaných jevů tím, že ho nutí logicky domyslet všechny možnosti, jasně definovat použité pojmy, vzít v úvahu všechny činitele, tedy i ty rozhodující*“. [18]
- „*Model je tedy abstrakcí získaný obraz určitým vlastností objektu, který je co možná jednoduchý. „Kvalita“ modelu je přitom daná přesností vystižení objektu, jednoduchostí a stupněm vystižení zákonitostí, které platí pro objekt*“. [12]

### 2.2 Klasifikace modelů

#### Deterministické a stochastické modely

Matematické modely rozdělujeme na deterministické a stochastické (pravděpodobnostní) z hlediska typu informací, které má k dispozici rozhodovatel a poté i analytik. Ve stochastických modelech se některé procesy nebo hodnoty řídí zákony pravděpodobnostního charakteru, kdežto v deterministických modelech jsou všechny informace o zkoumaném systému s jistotou známy.

Příkladem deterministického procesu může být výroba na výrobní lince, stochastickým procesem např. přeprava nákladu po veřejné dopravní síti, která podléhá náhodným vlivům. [13]

## **Statistické a dynamické modely**

Dynamický model používáme v případě, že ve zkoumaném systému hraje významnou roli čas, resp. časová proměnná (teplota, populace, napětí, produkce, zaměstnanost, ceny atd.). V opačném případě v analýze použijeme statický model. [8], [13]

Dynamické modely jsou nejčastěji představovány diferenčními rovnicemi, také i obyčejnými diferenciálními rovnicemi, parciálními diferenciálními rovnicemi, evolučními rovnicemi a prostředky umělé výpočetní inteligence. [8]

Při analýze problému je nutné najít určitý kompromis, pokud jde o co nejjednodušší řešení a o co nejděrnější zachycení skutečností. Obecně je nejjednodušší analýza deterministických statistických modelů, opačně řešení úloh, ve kterých se vyskytují dynamické a stochastické procesy, je nejsložitější. [13]

## **Mikroekonomické a makroekonomické modely**

Dělení podle druhu či velikosti systému, který modely popisují. Mikromodely se týkají hlavně podniků, individuálních trhů výrobků a služeb, domácností, spotřebitelů apod. Makromodely používáme především k analýze vývoje celého národního hospodářství. [13]

## **Účel grafických modelů**

U deskriptivního modelu je snahou o co nejděrnější popis reality, např. mapa, ale také organizační uspořádání organizace.

Normativní model nám přibližuje představu o tom, jak by to mělo být, např. žádoucí stav inovačního systému, strategie rozvoje vzdělání.

Vysvětlující model je teorie, která má snahu vysvětlit určitý jev, např. zvýšení počtu rozvodů.

Posledním modelem, který se snaží simulovat různé možné stavy v budoucnosti, je prognostický model. [22]

## **2.3 Modelování**

Pojem modelování je odvozen od pojmu model a zahrnuje tedy lidské činnosti, ve kterých model vystupuje. Z dostatečně širokého, časového i věcného hlediska můžeme ve vymezeních najít společné rysy pro oba pojmy. Tyto společné rysy jsou shrnuty a citovány dle [8] viz níže.

- „Modelování se týká vztahu lidí k objektům jejich zájmu, v němž člověk má dominantní úlohu a chce realizovat svůj záměr... Modelování se proto týká soustavy člověk- objekt – model.“
- „S modelováním se lze setkat na různé úrovni a v různé interpretaci všude tam, kde se vyskytuje Homo sapiens – člověk myslící, a to v celé historii a v celém životě od nejtělejšího mládí a do nejvyššího stáří.“
- „Modelování vždy sleduje snahu usnadnit dosažení cíle, který si člověk vytýčil, a to nepřímou, oklikou... Je atributem člověka.“
- „V modelování se využívají nejrůznější lidské schopnosti k realizaci řady činností specifických i všeobecných, vědomých i podvědomých, rutinních i tvůrčích, fyzických i duševních, dovolených i zakázaných, k jedinému cíli, kterým je efektivní, nepřímé (prostřednictvím modelu) vyřešení úkolu.“
- „Modelování je lidskou schopností řešit úspěšně úkoly a problémy na objektech, o jejichž struktuře, vlastnostech a chování máme jen neúplné znalosti a k řešení máme jen omezené podmínky.“
- Modelování je založeno na vztahu dvou nezávislých objektů – primárního problémového a sekundárního modelového, který umožňuje z řešení problému nebo obecněji z realizace zájmu na jednom objektu myšlenkovými procesy odvozovat řešení na druhém objektu. Základem k tomu je vztah podobnosti objektů. Podstatnými jsou tedy vztahy nezávislosti, preference a podobnosti.“
- „Úloha vytvoření modelu je úloha mnohoznačná. Ke stejnému problému na objektu lze vytvořit libovolné množství modelů různé podstaty, složitosti, náročnosti. Existuje však dolní hranice, dané objektivně řešeným problémem a horní hranice, daná možnostmi lidí a prostředků, které lidé pro řešení mají.“

Z výše psaného výkladu plyne, že se modelování dotýká všech oborů lidských činností. To ale neznamená, že úroveň, význam a pojetí modelování jsou v různých oborech stejné. Spíše naopak se řada charakteristik modelování více či méně liší. Jsou to např. tradice uvědomělého modelování, míra uvědomění si modelování jak specificky lidské činnosti, úroveň vymezení a formalizace modelování, úroveň a rozsah týmové práce, apod. [8]

### 3 Časové řady

Časové řady jsou důležitými statistickými daty, díky kterým můžeme zkoumat dynamiku jevů v čase. Mají hlavní význam pro analýzu příčin, které ovlivňovaly jejich chování a na tyto jevy působily v minulosti, tak i pro předvídání jejich vývoje v budoucnu.

Příkladem řad jsou seizmický záznam v geofyzice, řada nejvyšších (nejnižších) denních teplot v meteorologii, změny počtu jedinců nějaké populace v demografii, vývoj koncentrace nečistot v ekologii, vývoj rozvodovosti v sociologii nebo vývoj cen v ekonomii. [15]

#### Definice

*„Časová řada jsou věcně a prostorově srovnatelná pozorování (dat), která jsou jednoznačně uspořádána z hlediska času ve směru minulost – přítomnost.“ [9]*

#### Statistické (stochastické) a deterministické řady

Statistické řady jsou zatíženy nejistotou. Jedná se o řadu typu (3.1):

$$y(t) = \beta_0 + \beta_1 t + \varepsilon(t); \quad (3.1)$$

kde  $\varepsilon(t)$  je náhodná veličina s nulovou střední hodnotou a konečným rozptylem a  $t$  označuje čas.

Deterministické řady jsou řady, u kterých lze jednoznačně popsat jejich chování nějakým matematickým vzorcem, lze zkonstruovat její přesnou předpověď. Např. řada typu (3.2):

$$y(t) = \sin(2\pi ft). \quad (3.2)$$

Cílem analýzy časové řady je určení modelu, podle kterého jsou generována sledovaná data. Znalostí modelu můžeme předpovídat budoucí vývoj systému a do určité míry i řídit a optimalizovat chování systému pomocí vhodné volby vstupních parametrů a počátečních podmínek.

V minulosti se jako první sledovaly řady astronomických a meteorologických pozorování, kdežto v současnosti se aplikace zaměřují především spíše do ekonomické oblasti.

K analýze časových řad převládá zpočátku deterministický přístup, který přetrvával ještě během první čtvrtiny dvacátého století, a to i přesto že byl často kritizován pro neschopnost vysvětlit nepravidelnosti v amplitudách i ve vzdálenostech mezi lokálními extrémy časových

řad. Obrovským pokrokem v rozvoji disciplíny se představoval nový stochastický přístup, díky němuž lze popsat většinu reálných časových řad z ekonomické praxe. [8], [15]

### **Typy časových řad**

- Časová řada intervalových ukazatelů – jejich velikost přímo úměrně závisí na zvolené délce intervalu. Data se musí převést na srovnatelné hodnoty.
- Časová řada okamžikových ukazatelů – u těchto řad se ukazatel vztahuje k přesně definovanému okamžiku. Z tohoto vyplývá, že hodnota ukazatele nezávisí na délce intervalu, za který je sledován. Práce s těmito časovými řadami je tedy složitější.
- Časová řada odvozených charakteristik – tyto řady jsou odvozovány na základě absolutních údajů okamžikových nebo intervalových. Např. časové řady součtové nebo časové řady poměrných čísel. [10]

## **3.1 Problémy analýzy časových řad**

V životě máme mnoho nepravidelných a nespolehlivých procesů. Struktura těchto procesů se nám mění přímo před očima, a proto při analýze časových řad bychom neměli některé problémy opomíjet.

### **3.1.1 Problémy s volbou časových bodů pozorování**

Časové řady jsou tvořené měřením (pozorováním), která jsou prováděna v diskrétních časových bodech (okamžicích). Určité z nich jsou už diskrétní svou povahou (např. řady úhrnné produkce nějaké zemědělské plodiny za jednotlivé roky), jiné je potřeba předem diskretizovat. Mohou vznikat několika způsoby:

- diskretizací spojité časové řady (např. cena určitého zboží na daném trhu),
- akumulací (agregací) hodnot sledované veličiny za dané časové období (např. roční výroba průmyslového podniku),
- průměrováním hodnot uvažované veličiny v daném časovém intervalu (např. řada průměrných denních teplot).

Pokud je možnost volby, pak se doporučuje volit kompromisní řešení. Velká hustota časových bodů měření umožňuje dobře vystihnout charakteristické rysy časové řady, ale je možné, že nastanou potíže při výpočtech. V každém případě se však snažíme volit ekvidistantní (tj. stejně časově vzdálené) intervaly mezi sousedními pozorováními. [8], [15]



### 3.1.2 Problémy s kalendářem

Při analýze ekonomických časových řad je možný nepříznivý projev nepravidelností v kalendáři (např. různá délka kalendářních měsíců, čtyři nebo pět víkendů v měsíci, různý počet pracovních dní v měsíci nebo pohyblivé svátky (např. velikonoce)). Tento problém je možné odstranit korekcí – zavedení standardního měsíce. Dále pak určité krátkodobé nepravidelnosti lze odstranit pomocí akumulace (např. použití kvartálních dat namísto měsíčních).

Dalším problémem souvisejícím s kalendářem mohou být i neočekávané události, např. výpadek energie, volby, výkyvy počasí, změny směrnic, havárie, změna zákonů atd. [8], [15]

### 3.1.3 Problémy s délkou časových řad

Délka řady znamená celkový počet pozorování (měření) v časové řadě, nikoliv jako časové rozpětí mezi prvním a konečným měřením. Počet měření výrazně ovlivňuje „množství informace“, která řada obsahuje. Ovšem zde neexistuje žádná přímá závislost. Např. zdvojnásobení počtu pozorování neznamená zdvojnásobení informací. B-J metoda vyžaduje délku alespoň 50 měření. Paradoxně u dlouhých časových řad je nebezpečí, že by se mohl změnit charakter řady v průběhu času. [8], [15]

### 3.1.4 Problém s nesrovnatelností jednotlivých měření

Pomocí příslušných cenových indexů je nutné upravit ekonomické údaje vzhledem ke změnám cen. Nesrovnatelnosti mohou nastat i například neúplností dat v různých přehledech (např. daná hodnota v jednom roce se týká 89 podniků a v druhém roce 92 podniků atd.) [8]

## 3.2 Základní přístupy k analýze časových řad

Výběr správné metody závisí na celé řadě faktorů: účelu analýzy, typu sledované řady, na zkušenostech, statistice, jakož i dostupnosti výpočetní techniky a statistického softwaru.

### Stručná charakteristika čtyř nejčastěji používaných přístupů k analýze časových řad

Budeme uvažovat časovou řadu  $y_1, y_2, \dots, y_n$ .

#### 3.2.1 Dekompozice časové řady

Dekompozice časových řad je proces kvalifikace složek časových řad. Je založena na práci se systematickými složkami časové řady. U tohoto přístupu předpokládáme, že časová řada obsahuje systematické složky. Jsou to:

- trend  $Tr(t)$ ;
- sezónní složka  $Sz(t)$ ;
- cyklická složka  $C(t)$ ;
- náhodná (reziduální, stochastická) složka  $\varepsilon(t)$ .

Časová řada nemusí obsahovat všechny výše uvedené složky najednou. Časovou řadu je možné chápat jako trend, který na sebe „nabaluje“ periodické složky (sezónní a cyklická) a náhodnou složku (bílý šum). Dekompozice (rozklad) časové řady provádíme s nadějí, že se v jednotlivých složkách rozkladu snáze podaří identifikovat pravidelné chování řady než v původní nerozložené řadě.

Dekompoziční metody pracují pouze se systematickými složkami (trendem, sezónní a cyklickou složkou). Využíváme tedy zpravidla metody regresní analýzy. [1], [8], [15]

### **Trend**

Trendová složka odráží dlouhodobé změny v chování časové řady, např. zachycuje dlouhodobý růst či pokles. Trend vznikl důsledkem působení sil, které systematicky působí stejným směrem. Například pokud sledujeme prodej v daném průmyslovém odvětví, bývají těmito silami např. změny ve výši příjmů obyvatelstva, technologické změny ve výrobě, změny v populaci, atd.

Trend je většinou možné popsat matematickou funkcí v celé délce časové řady. Při popisu trendu je důležité zachycení tendence pohybu časové řady, nejde tedy o to, zda časová řada krátkodobě klesá či roste. Trend má relativní charakter, např. změna klimatu, která je pro zemědělce dlouhodobá, je pro geologa krátkodobá epizoda apod. [1], [8]

### **Sezónní složka**

Sezónní složka popisuje periodické změny v časové řadě, které se odehrávají v rámci jednoho časového úseku (kalendářního roku) a každý úsek se opakuje. Příkladem sezónních změn je důsledek střídání ročních období. Nejčastěji vidíme sezónnost u čtvrtletních a měsíčních časových řad. Sezónní složka se nemůže vyskytovat u ročních časových řad, přestože se tato složka pravidelně v časové řadě opakuje, protože by mohla v průběhu let měnit svůj charakter. [1], [8]

## Cyklická složka

Tato složka popisuje dlouhodobé fluktuace kolem trendu. Charakterizuje tedy např. fázi růstu a pak poklesu, která je mnohem větší než jeden rok. U ekonomických řad je cyklická složka spojována se střídáním hospodářských cyklů. Tuto složku je velmi těžké popsat, protože působí dlouhodobě. Její perioda se může pohybovat v násobcích let, proto když je krátká časová řada, cyklickou složku nemusíme rozeznat a také se její charakter může v čase měnit. [1], [8]

## Reziduální (náhodná, stochastická) složka

Pokud z časové řady odstraníme předcházející tři složky, zůstane náhodná složka. Je představena náhodnými fluktuacemi (pohyby) v průběhu časové řady, které nemají systematický (rozpoznatelný) charakter jako předcházející složky. Zahrnuje např. chyby v měření údajů, chyby ve statickém zpracování dat (zaokrouhlování), atd. Předpokladem je, že reziduální složka má charakter bílého šumu, tj. že je tvořena hodnotami nezávislých náhodných veličin s nulovou střední hodnotou a nějakým konstantním rozptylem. Bílý šum pokrývá frekvenční pásmo od 0 do  $\infty$ . [8], [15]

Rozklad (dekompozice) časové řady může být dvojího druhu:

### a) Aditivní

Ve tvaru (3.3):

$$y(t) = Tr(t) + Sz(t) + C(t) + \varepsilon(t), \quad (3.3)$$

$t = 1, 2, \dots, n$ , kdy všechny složky měříme ve stejných jednotkách jako  $y(t)$ .

### b) Multiplikativní

Ve tvaru (3.4):

$$y(t) = Tr(t) * Sz(t) * C(t) * \varepsilon(t), \quad (3.4)$$

$t = 1, 2, \dots, n$ , kdy jenom trend měříme ve stejných jednotkách jako  $y(t)$  a zbylé složky jsou bezrozměrné veličiny.

Po logaritmické transformaci přechází multiplikativní rozklad na aditivní, aniž by se statistické vlastnosti bílého šumu změnily. Nemusí být obsaženy všechny složky u všech vyšetřovaných časových řad. [8], [15]

### 3.2.2 Boxova-Jenkinsova (B-J) metodologie

Předpokládejme časovou řadu slabě stacionární. Reziduální složka je základním prvkem při konstrukci modelu. Uvedeme typické příklady modelu:

- Model klouzavých součtů 1. řádu ve tvaru (3.5):

$$y(t) = \varepsilon(t) + k\varepsilon(t-1), \quad (3.5)$$

kde  $k$  je nějaká reálná konstanta a  $\varepsilon(t)$  reprezentuje tzv. bílý šum.

- Autoregresní model 1. řádu ve tvaru (3.6):

$$y(t) = \varepsilon(t) + ky(t-1). \quad (3.6)$$

B-J metodologie může modelovat i časové řady s výrazným trendovým, a nebo sezónním charakterem, kde trend i sezónní složka mohou být stochasticky modelovány na rozdíl od dekompozice.

Modely B-J jsou mnohem flexibilnější na rozdíl od modelů dekompozičních – lépe se adaptují na změny v průběhu časové řady.

V tomto případě se pro analýzu časové řady využívá metod korelační analýzy. Tyto metody umožňují zkoumat závislosti mezi jednotlivými měřeními dané časové řady. [15]

### 3.2.3 Lineární kauzální (faktorové) modely

Tyto modely tvoříme tak, že naměřené hodnoty sledované časové řady vysvětlíme pomocí tzv. faktorových časových řad. Jednoduchý ekonometrický model ve tvaru (3.7):

$$c(t) = \alpha + \beta c(t-1) + \gamma x(t) + \delta p(t) + \varepsilon(t), \quad (3.7)$$

kde  $c(t)$  jsou výdaje obyvatelstva na nákup spotřebního zboží v roce  $(t)$ . Tyto výdaje vysvětlujeme pomocí výdajů  $c(t-1)$  v roce bezprostředně předcházejícím a pomocí peněžních příjmů  $x(t)$  obyvatelstva a cenového indexu  $p(t)$  spotřebního zboží v roce  $(t)$ .  $\alpha$ ,  $\beta$ ,  $\gamma$  a  $\delta$  jsou parametry modelu a  $\varepsilon(t)$  je bílý šum. [7]

### 3.2.4 Spektrální analýza časových řad

V této analýze se považuje časová řada za „směs“ sinusovek a kosinusovek s různými amplitudami a frekvencemi. Například pomocí speciálního statistického nástroje periodogramu můžeme získat představu o intenzitě zastoupení jednotlivých frekvencí v časové řadě. V této analýze je možné nahrazovat časovou řadu pojmem signál.

V této analýze časových řad budeme využívat hlavně Fourierovu analýzu. [8], [15]

### 3.3 Předpovědi v časových řadách

Tyto předpovědi mohou být dvojího druhu:

- bodová předpověď – jedná se o hodnotu představující odhad budoucí hodnoty časové řady v jistém budoucím okamžiku;
- předpovědní interval – tento interval je analogií intervalu spolehlivosti.

Abychom mohli předpovědi vytvářet, je zapotřebí metod buď kvalitativních, nebo kvantitativních.

- Kvalitativní metody (např. metoda Delfi) jsou většinou založeny na názoru odborníků (expertů), a proto mají pouze subjektivní charakter.
- Kvantitativní metody vycházejí z objektivních matematických postupů. Předpokládáme, že charakter zkoumané časové řady se v budoucnu nemění.

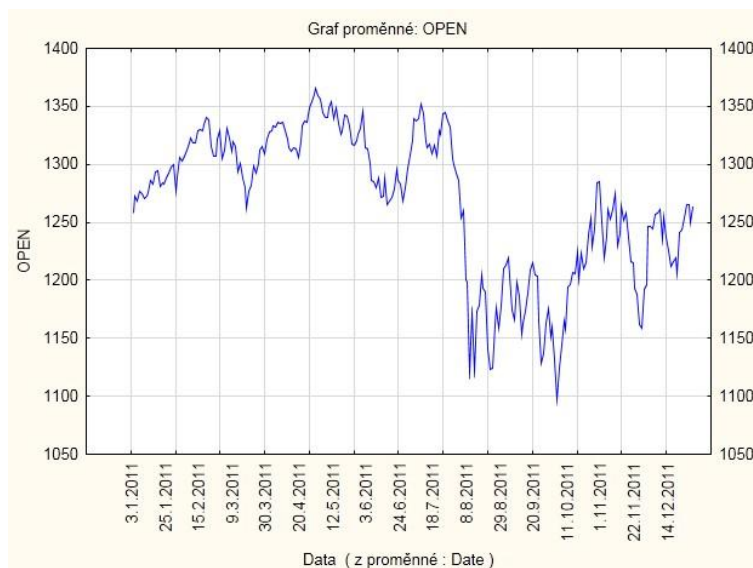
Správný výběr předpovědní techniky závisí na spoustě faktorů, například na požadované formě předpovědi, požadované přesnosti, charakteru vstupních dat a jejich dostupnosti a na časovém horizontu předpovědi.

Musíme vzít v úvahu všechny zkonstruované předpovědi pro správné posouzení kvality předpovědi. V praxi se nejčastěji používají míry kvality předpovědi (součet čtvercových chyb SSE, průměrná čtvercová chyba MSE, průměrná absolutní odchylka MAD). Míry SSE a MSE ve srovnání s MAD posuzují přísněji velké chyby předpovědi než malé. [8], [15]

### 3.4 Finanční časové řady, jejich předpoklady a vlastnosti

Součástí tržního systému je finanční trh, kde se setkává nabídka a poptávka peněz a kapitálu. Základní informací finančních trhů je cena, která se sleduje v určité časové frekvenci a tím tvoří časové řady. V porovnání s jinými ekonomickými časovými řadami mají finanční časové řady některé specifické vlastnosti a tvarové odlišnosti dané mikrostrukturou finančních trhů. Hlavním rysem finančních časových řad je vysoká časová frekvence jednotlivých hodnot, zaznamenávají se nejčastěji v denní frekvenci. Z toho vyplývá, vedle systematických faktorů mají na dynamiku časových řad celkem značný vliv i faktory nystematického charakteru, což se projevuje v jejich poměrně vysoké a proměnlivé variabilitě. U složek způsobených systematickými faktory se nejvíce projevuje trend a cyklická složka, kdežto sezónní složka není tak výrazná. [4]

Následující graf 1 ukazuje typický průběh finanční časové řady indexu S&P 500 v denní frekvenci za rok 2011. O indexu S&P 500 více v kapitole 5.1.



**Graf 1: Průběh finanční časové řady indexu S&P 500**

*Zdroj: vlastní zpracování pomocí sw. STATISTICA na základě dat z [19]*

### 3.4.1 Hypotéza efektivního trhu

Základním východiskem pro fungování finančního trhu je hypotéza efektivního trhu. Její formulace zní: *“Za předpokladu, že ceny plně zahrnují očekávání a informace všech účastníků trhu, jsou jejich změny nepredikovatelné.”* [4]

S touto hypotézou efektivního trhu je spojena idea modelu vývoje cen aktiv, který je znám pod pojmem martingál. Model lze popsat následujícím způsobem: *„Jestliže  $P_t$  představuje cenu aktiva v čase  $t$ , potom očekávaná cena v čase  $t + 1$  je cena v čase  $t$ , za podmínky znalosti všech cen aktiva v minulosti, tj. v čase  $t - 1$ ,  $t - 2$ , ...“.* [citováno z 4] Z hlediska tvorby predikcí to znamená, že nejlepší předpovědí zítřejší ceny je cena dnešní. Jestliže stochastický proces  $\{P_t\}$  je časová řada znázorňující cenu nějakého aktiva, potom formálně lze pojem martingál zapsat pomocí podmíněné střední hodnoty (3.8):

$$E[P_{t+1}|P_t, P_{t-1}, \dots] = P_t. \quad (3.8)$$

### 3.4.2 Jednoduché výnosy aktiva

V analýzách časových řad se setkáváme i s modelováním transformací těchto řad. U finančních časových řad se pracuje s tzv. výnosy aktiva. Jednoduché výnosy jsou relativní přírůstky cen, kde takto transformovaná časová řada má lepší vlastnosti, resp. stacionaritu.

Z publikace [4] víme, že cena aktiva nemůže být menší než nula. Minimální dosažitelný jednoduchý výnos aktiva je tedy (3.9):

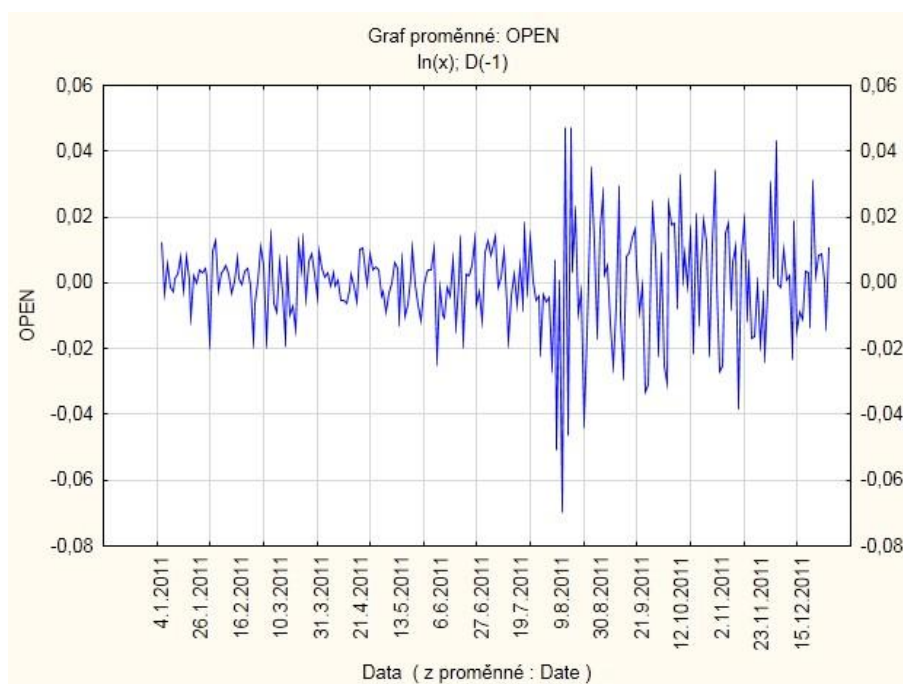
$$R_t = \frac{P_t - P_{t-1}}{P_{t-1}} = -1 \text{ a poté } R_t + 1 = \frac{P_t}{P_{t-1}}, \quad (3.9)$$

kde  $P_t$  je cena aktiva v čase  $t$ . Takto definované relativní přírůstky ceny, by měly mít rozdělení nezáporné náhodné veličiny. V tomto případě je nasnadě rozdělení logaritmicko-normální. Poté logaritmus náhodné veličiny s logaritmicko-normálním rozdělením má rozdělení normální, a tedy jednoduchý výnos  $R_t + 1$ , resp. jeho logaritmus značíme (3.10):

$$r_t = \ln(R_t + 1) = \ln P_t - \ln P_{t-1} = p_t - p_{t-1}. \quad (3.10)$$

Nejmenší jednoduchý výnos je -1, protože cena aktiva nemůže být záporná. Nevýhodou jednoduchých výnosů oproti logaritmickým je, že pokud bereme normální rozdělení, tak jsou právě pravděpodobné i hodnoty menší než -1. Proto se pracuje s výnosy logaritmickými, u kterých není toto omezení.

Graf 2 zobrazující vypočtené logaritmy koeficientů růstu původních hodnot indexu S&P 500 je následně znázorněn.



**Graf 2: Logaritmy výnosů řady indexu S&P 500**

*Zdroj: vlastní zpracování pomocí sw. STATISTICA na základě dat z [19]*

### 3.4.3 Předpoklad normality

Jedním ze základních předpokladů je, že logaritmy výnosů mají normální rozdělení s konstantní střední hodnotou a konstantním rozptylem. Charakteristikou tohoto rozdělení je, že je symetrické, takže jeho šikmost je rovna nule a jeho špičatost je rovna číslu tři.

V publikaci [4] je dokázáno, že u burzovních indexů je možné vidět odhad šikmosti jako záporné číslo. S rostoucí agregací dat má zešikmení tendenci zvětšovat se. Také odhad špičatosti je ve všech případech denních výnosů číslo větší než 3, z toho vyplývá, že výnosy mají špičatější rozdělení než normální.

U denních kurzů ve srovnání s burzovními indexy není však zešikmení výnosů tak jednostranné. I špičatost ani variabilita není tak vysoká jako u burzovních indexů. Dá se říci, že rozdělení výnosů mají „tlustší“ konce než rozdělení normální, tzn. četnost výskytu extrémně vysokých kladných či záporných výnosů je vyšší než za předpokladu normality.

Také je známo, že s rostoucí agregací se měnlivost skutečných hodnot přibližuje měnlivosti hodnot normálního rozdělení.

Tyto vlastnosti finančních časových řad jsou známy delší dobu, proto se začalo hledat pravděpodobnostní rozdělení, které by dané vlastnosti charakterizovalo lépe než normální rozdělení. Nejprve se navrhlo tzv. stabilní rozdělení, jedná se o rozdělení, které zachycuje zejména výraznou špičatost a „tlusté“ konce rozdělení četností výnosů. Má nekonečný rozptyl a momenty řádu vyššího než dva. Protože ale toto rozdělení má nekonečný rozptyl, hledalo se jiné rozdělení, které by rozptyl mělo konečný a přesto momenty vyššího řádu. Jedná se o Studentovo rozdělení  $t$ .

Populárním rozdělením v poslední době se stala tzv. směsice rozdělení, vyznačuje se tím, že logaritmus výnosů má podmíněné normální rozdělení s podmíněným rozptylem, který se mění v závislosti na čase. Výsledkem je nepodmíněné špičaté rozdělení s „tlustými konci“, které má konečný rozptyl a momenty vyššího řádu.

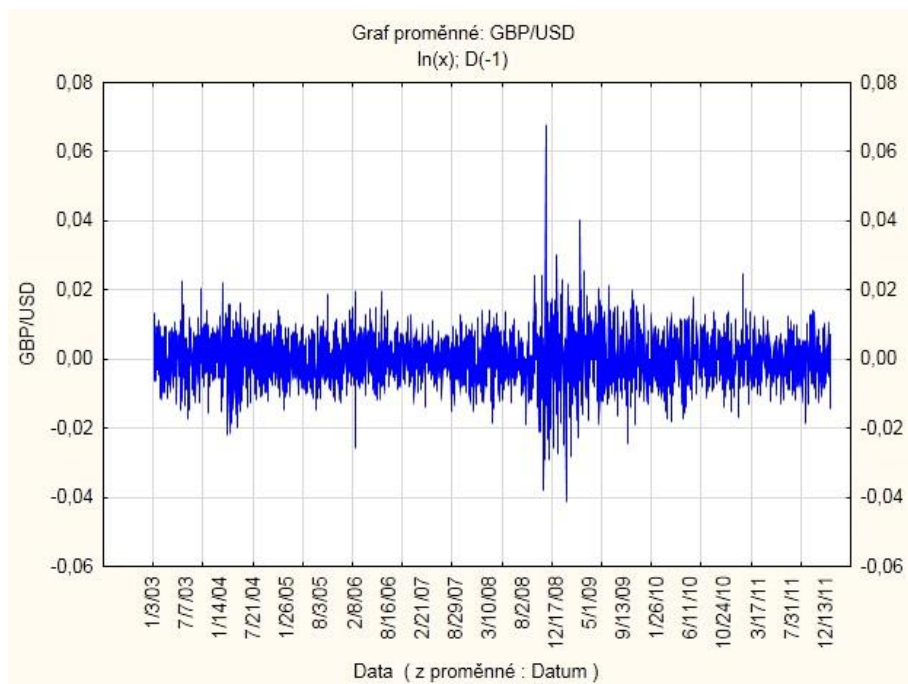
Z výše uvedeného vyplývá, že mezi základní charakteristické vlastnosti výnosů finančních časových řad patří leptokurtické pravděpodobnostní rozdělení, které je ve srovnání s normálním rozdělením špičatější s tlustšími konci. Toto rozdělení má také proměnlivou volatilitu. [4]



### 3.4.4 Předpoklad linearity

U klasických analýz finančních časových řad je předpoklad takový, že logaritmy výnosů jsou nekorelované nebo nezávislé stejně rozdělené náhodné veličiny s nulovou střední hodnotou a konstantním rozptylem. Pokud jde o nekorelovanost, jedná se o proces bílého šumu, u nezávislosti jde o proces striktního bílého šumu.

V reálu se stává, že často není splněna podmínka nulové střední hodnoty, ale může se také stát, že nemusí být splněna podmínka nekorelovanosti logaritmů výnosů, z čeho vyplývá, že logaritmy výnosů mohou být lineárně závislé. K vyjádření závislosti se používají modely ARMA. Nelineární forma závislosti se projevuje shlukováním volatility, na což upozornil Mandelbrot [17]. Jedná se o tendenci volatility finančních trhů objevovat se ve shlucích nízkých a vysokých volatilit (malé cenové změny jsou následovány změnami malými a velké velkými). V následujícím grafu 3 vidíme shlukování volatility, kde úseky s nízkým rozptylem se střídají s úseky s rozptylem vysokým.



**Graf 3: Logaritmy výnosů časové řady kurzu GBP/USD**

*Zdroj: vlastní zpracování pomocí sw. STATISTICA na základě dat z [14]*

### 3.4.5 Pákový efekt

Pákový efekt („leverage effect“) znamená, že negativní šoky působí na volatilitu jinak, než je tomu tak u pozitivních šoků ve stejné absolutní výši. Akumulace informace se vytváří ve dnech, kdy se neobchoduje. Právě to se projevuje zvýšenou volatilitou v následujících

obchodních dnech. Další zvýšení volatility způsobuje také pravidelné zveřejňování důležitých informací. Volatilita se snižuje při dobrých zprávách, kdežto při špatných zprávách na trhu se zvětšuje. [4]

#### 3.4.6 Předpoklad autokorelace

Změna autokorelační struktury souvisí se změnou volatility v časové řadě, takže nelineární modely lze interpretovat také jako modely proměnlivé volatility. Předpoklad výskytu autokorelace je možné otestovat pomocí Durbin - Watsonova testu. [4]

#### 3.4.7 Předpoklad heteroskedasticity

Heteroskedasticita vymezuje vlastnost, že logaritmus výnosu má normální rozdělení s měnícím se rozptylem v závislosti na čase. V analýze finančních časových řad je heteroskedasticita velmi typická. V případě porušení homoskedasticity, mluvíme o heteroskedasticitě. Jedná se tedy o situaci, kdy je porušena podmínka konečného a konstantního rozptylu náhodných složek.

Následující model (3.11) znázorňuje podmíněnou heteroskedasticitu:

$$(\ln X_t - \ln X_{t-1})^2 = \alpha + \rho(\ln X_{t-1} - \ln X_{t-2})^2 + u_t, \quad (3.11)$$

kde  $X_t$ ,  $X_{t-1}$  vyjadřují hodnoty v časové řadě při změně času  $t$  o jednotku. Parametr  $\alpha$  se vypočítá pomocí metody nejmenších čtverců a  $u_t$  je náhodná složka. Pokud je parametr  $\rho$  roven nule, nemůžeme mluvit o heteroskedasticitě.

K zjištění heteroskedasticity se používá celá řada statistických testů. Jednoduchým a nejpoužívanějším testem je Spearmanův test korelace pořadí nebo také Whiteův test, či testování podmíněné heteroskedasticity založené na principu Lagrangeových multiplikátorů. [4]

#### 3.4.8 Vliv mikrostruktury trhu na vlastnosti finančních časových řad

Mikrostruktura trhu, respektive aspekty organizačního a technického charakteru, mohou vlastnosti finančních časových řad modifikovat. Většinou u dlouhých časových řad se vliv těchto aspektů ignoruje, jinak se s vlivem musí počítat.

*Nesynchronní obchodování* nejvíce ovlivňuje charakter finančních časových řad. Tento efekt vzniká, když se u hodnot časových řad, především u řad cen akcií, předpokládá, že jsou generovány v intervalech stejně dlouhých. Ve skutečnosti tomu tak není, protože existují

právě dny, ve kterých se neobchoduje. Pokud bychom toto neakceptovali, může být způsobena zdánlivá autokorelace v časové řadě logaritmů denních výnosů. (Autokorelace je způsobena právě zadržováním informace v neobchodních dnech.)

Hlavní a nejdůležitější vlastnost finančního trhu je jeho *likvidita*. O její zabezpečení se starají obchodníci, kteří na objednávku buď nakupují, nebo prodávají. Cena nákupu a prodeje se liší, a to právě také komplikuje situaci s finančními časovými řadami.

*Transakční cena* je jednou z dalších komplikací. Tato cena se nemusí ztotožňovat s cenami předešlými, nevíme tedy, z jaké ceny počítat výnos. Také náhodný pohyb ceny může způsobit zdánlivou variabilitu v časové řadě nebo autokorelaci. Opět v případě dlouhých časových řad vliv aspektu nebude tak výrazný jako u krátkých.

Také *cenové shlukování*, které je způsobeno cenami udávanými nespojitě, ačkoliv se jedná o spojitou veličinu, může modifikovat charakter finančních časových řad. [4]

V této diplomové práci se nebudeme uvedeným aspektům mikrostruktury trhu dále věnovat. Jedná se o speciální problematiku přesahující rámec této práce, avšak musíme mít na paměti, že vytvořené závěry především v experimentální části mohou být z výše uvedených důvodů relativizovány<sup>1</sup>.

---

<sup>1</sup> Relativizmus - názor o podmíněnosti (relativitě) poznání

## 4 Box-Jenkinsova metodologie

B-J přístup bere za základní prvek konstrukce modelu časové řady reziduální složku. Tato složka může být tvořena závislými (korelovanými) náhodnými veličinami. B-J metodologie tedy může zpracovávat časové řady s navzájem závislými pozorováními. Hlavně vyšetřuje tyto závislosti pomocí korelační analýzy. Je zavádějící, myslet si, že tato metodologie věnuje přehnanou pozornost pouze náhodné složce a nemůže modelovat řady se zjevným trendovým nebo sezónním charakterem. Tyto řady dokáže B-J metodologie analyzovat pomocí tzv. integrovaných modelů ARIMA a tzv. sezónních modelů, v nichž sezónní nebo trendová složka může být stochasticky modelována. [7]

### Výhody a nevýhody tohoto přístupu

Nejprve začneme s pozitivními rysy:

- Modely mají možnost rychlé adaptace na změny v průběhu časové řady, proto je B-J metodologie úspěšná i v těch případech, kde dekompozice selhává.
- B-J přístup je systematický a může být tedy plně automatizován
- Přístup také vykazuje nejlepší výsledky při analýze ekonomických časových řad

Hlavní nevýhody B-J metodologie:

- Tento model je vhodný pouze pro časové řady o délce nejméně 50 pozorování.
- Praktické aplikace B-J přístupu jsou mnohem náročnější než aplikace dekompozičních metod.
- Výsledné modely, zejména modely s větším počtem parametrů se obtížně interpretují.

### 4.1 Základní pojmy B-J metodologie

Cílem této části je objasnit některé důležité pojmy pro modelování různých typů ekonomických a finančních časových řad.

#### 4.1.1 Stochastický proces

Můžeme říci, že stochastický proces je v čase uspořádaná řada náhodných veličin  $\{X(s, t), s \in S, t \in T\}$ , kde  $S$  je výběrový prostor a  $T$  je indexní řada. Časovou řadu chápeme jako realizaci stochastického procesu. V dalších kapitolách budeme uvažovat, že indexní řada je řadou celých čísel a že předpokládané stochastické procesy, které budeme značit jako  $\{X_t\}$ , mají nespojitý charakter. Jejich realizace, tedy časové řady budeme značit jako  $X_t$ . [4]

### 4.1.2 Stacionarita

Stacionarita časové řady znamená, že chování této řady je stochasticky ustálené. Převážně se rozlišuje tzv. striktní stacionarita a slabá stacionarita.

*Striktní stacionarita* – pravděpodobnostní chování příslušného stochastického procesu je invariantní vůči posunům v čase.

*Slabá stacionarita* – příslušný stochastický proces má konstantní střední hodnotu, konstantní rozptyl a kovarianční strukturu druhého řádu invariantní vůči posunům v čase.

V této diplomové práci budeme pracovat pouze se slabou stacionaritou, budeme ji jednoduše nazývat stacionaritou. Stacionární proces je tedy rovnoměrně vyvážený (tj. s konstantním rozptylem) kolem konstantní úrovně (tj. má konstantní střední hodnotu). Závislost mezi oběma jeho pozorováními závisí pouze na jejich vzájemné časové vzdálenosti, nezávisí tedy na jejich skutečném časovém umístění v řadě.

Stochastický proces můžeme značit jako normální nebo gaussovský, pokud jeho sdružené pravděpodobnostní rozdělení je normální. Poté můžeme říci, že striktní a slabá stacionarita jsou ekvivalentními pojmy, protože normální rozdělení je jednoznačně charakterizováno prvními dvěma obecnými momenty. V našem případě budeme předpokládat právě gaussovské stochastické procesy.

Stacionární stochastický proces je tedy proces s náhodnými veličinami, které mají konstantní (nulovou) střední hodnotu a konečný rozptyl.

V B-J metodologii lze analyzovat pouze stacionární časové řady. Většinou si k převodu z nestacionární časové řady na stacionární pomáháme různými transformacemi (nejčastěji diferencováním). [4], [7], [8]

### 4.1.3 Autokorelační funkce (ACF)

V případě stacionárního stochastického procesu  $\{X_t\}$  lze vyjádřit autokovarianční funkci (4.1) mezi veličinami  $X_t$  a  $X_{t-1}$  jako:

$$\gamma_k = C(X_t, X_{t-k}) = E(X_t - \mu)(X_{t-k} - \mu). \quad (4.1)$$

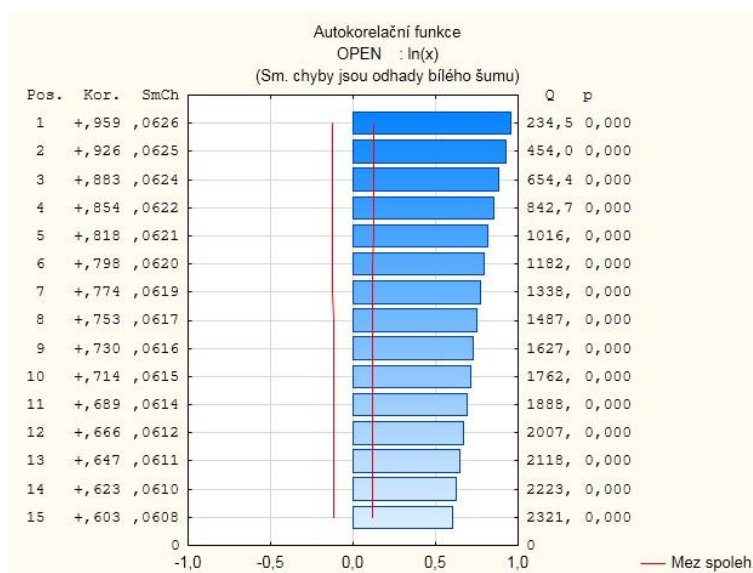
Hodnota  $\gamma_0$  je rozptylem hodnot časové řady  $\sigma^2$ . Po normování získáme autokorelační funkci (4.2), kterou je možné vyjádřit jako:

$$\rho_k = \frac{C(X_t, X_{t-k})}{\sqrt{D(X_t)}\sqrt{D(X_{t-k})}} = \frac{\gamma_k}{\gamma_0} = \frac{\gamma_k}{\sigma^2}. \quad (4.2)$$

Vlastnosti autokorelační funkce:

- $\rho_0 = 1$ , pro  $k > 0$
- $|\gamma_k| \leq \gamma_0; |\rho_k| \leq 1$ , pro  $k > 0$
- $\gamma_k = \gamma_{-k}$  a  $\rho_k = \rho_{-k}$  pro všechna  $k$ , funkce je tedy symetrická kolem  $k = 0$ . [4]

Graf autokorelační funkce se nazývá korelogram. Příklad korelogramu si ukážeme na následujícím grafu 4, který zobrazuje ACF logaritmované časové řady indexu S&P 500.



**Graf 4: ACF logaritmované časové řady**

*Zdroj: vlastní zpracování pomocí sw. STATISTICA na základě dat z [19]*

#### 4.1.4 Parciální autokorelační funkce (ACF)

Parciální autokorelace podávají informaci o korelaci veličin  $X_t$  a  $X_{t-k}$  očištěnou o vliv veličin, které leží mezi nimi.

Parciální regresní koeficient  $\phi_{kk}$  v autoregresi  $k$ -tého řádu vyjadřuje parciální autokorelaci se zpožděním  $k$  (4.3). Obecnou teorii parciálních korelací můžeme najít např. v učebnici<sup>2</sup>. V této práci uvedeme jen odpovídající vzorce podle monografie [7]. Platí:

<sup>2</sup> Anděl, J.: *Matematická statistika*. Praha, SNTL/Alfa 1978.

$$\phi_{kk} = \frac{|P_k^*|}{|P_k|}, \quad (4.3)$$

kde  $||$  označuje determinant matice,  $P_k$  je matice autokorelací tvaru (4.4):

$$P_k = \begin{bmatrix} 1 & \rho_1 \cdots & \rho_{k-1} \\ \rho_1 & 1 \cdots & \rho_{k-2} \\ \vdots & \ddots & \vdots \\ \rho_{k-1} & \rho_{k-2} \cdots & 1 \end{bmatrix} \quad (4.4)$$

a matice  $P_k^*$  (4.5), která vznikne předefinováním posledního sloupce předcházející matice, má tvar:

$$P_k^* = \begin{bmatrix} 1 & \rho_1 \cdots & \rho_1 \\ \rho_1 & 1 \cdots & \rho_2 \\ \vdots & \ddots & \vdots \\ \rho_{k-1} & \rho_{k-2} \cdots & \rho_k \end{bmatrix}. \quad (4.5)$$

#### 4.1.5 Proces bílého šumu

Definice procesu citována z [4]: „*Jestliže je stochastický proces  $\{a_t\}$  řadou nekorelovaných náhodných veličin jednoho pravděpodobnostního rozdělení s konstantní střední hodnotou  $E(a_t) = \mu_a$  (obvykle nulovou), konstantním rozptylem  $D(a_t) = \sigma_a^2$  a  $\gamma_k = C(a_t, a_{t-k}) = 0$ , pro všechna  $k \neq 0$ .*“ Z této definice plyne, že proces bílého šumu  $\{a_t\}$  je stacionární s autokorelační funkcí  $\rho_k$  a parciální autokorelační funkcí  $\phi_{kk}$ . Charakteristické pro tento proces je, že ACF a PACF jsou identicky nulové. Přesto, že se tento proces v podstatě nevyskytuje, je důležitým prvkem při výstavbě modelů časových řad.

Bílý šum je tedy takový náhodný proces, pro který platí, že procesy  $\varepsilon_t$  a  $\varepsilon_{t-1}$  jsou nezávislé a bílý šum pokrývá frekvenční pásmo od 0 do  $\infty$ . [8]

#### 4.1.6 Lineární proces

„*Každý stacionární proces, který neobsahuje deterministickou složku (složka, která je na základě minulosti perfektně predikovatelná – střední hodnota, resp. konstanta, nějaká periodická, polynomiální či exponenciální funkce časové proměnné  $t$ ), může být vyjádřen jako lineární kombinace řady nekorelovaných stejně rozdělených náhodných veličin (dokázal Wold H.)*“, definice citována dle publikace [4]. Lineární proces (Woldova reprezentace) se tedy označuje jako lineární kombinace popsaná v předešlé definici.

Lineární proces je definován jako nekonečná řada (4.6):

$$y_t = \varepsilon_t + \psi_1 \varepsilon_{t-1} + \psi_2 \varepsilon_{t-2} + \dots, \quad (4.6)$$

kde  $\varepsilon_t$  je bílý šum s rozptylem  $\sigma_\varepsilon^2$  a  $\psi_j$  jsou parametry. Při zápisu lineárního procesu se obvykle používá tzv. operátoru zpětného posunutí  $B$ , který je definován jako  $By_t = y_{t-1}$  a  $B^j y_t = y_{t-j}$  (což je několikanásobná aplikace). S použitím tohoto operátoru lze rovnici (4.6) zapsat ve tvaru (4.7):

$$y_t = \psi(B)\varepsilon_t, \quad (4.7)$$

kde

$$\psi(B) = 1 + \psi_1 B + \psi_2 B^2 + \dots = 1 + \sum_{j=1}^{\infty} \psi_j B^j. \quad (4.8)$$

Lineární proces existuje pouze, pokud nekonečná řada náhodných veličin na pravé straně rovnice (4.6) konverguje podle kvadratického středu. Tato podmínka současně zaručuje stacionaritu lineárního procesu a nulovou jeho střední hodnotu.

V B-J metodologii mají praktický význam speciální lineární modely s konečným počtem nenulových parametrů. Přitom parametry volíme tak, aby výsledné modely splňovaly podmínky pro stacionaritu a invertibilitu. Nesmíme opomenout, že existují i nestacionární modely ARIMA, které si přiblížíme v pozdější kapitole.

Speciální případy B-J metodologie:

- Proces klouzavých součtů MA(q)
- Autoregresní proces AR(p)
- Smíšený proces ARMA(p,q).

## 4.2 Lineární modely stacionárních časových řad

V této části si stručně shrneme vlastnosti procesů AR, MA a ARMA, jakožto základní stavební prvky B-J metodologie. Budeme se soustředit na modely nižších řádů, které se v praktických aplikacích používají nejčastěji. Čerpáno z publikací [4], [7], [8].

### 4.2.1 MA – proces klouzavých součtů

Nejprve si ukážeme model obecně a poté pro první i druhý řád.



#### 4.2.1.1 Proces klouzavých součtů $MA(q)$

Proces  $MA(q)$  má následující tvar (4.9):

$$y_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q}, \quad (4.9)$$

nebo s operátorem zpětného posunutí  $B$  jako (4.10):

$$y_t = \theta(B) \varepsilon_t, \quad (4.10)$$

kde

$$\theta(B) = 1 + \sum_{j=1}^q \theta_j B^j \quad (4.11)$$

je operátor klouzavých součtů.  $\varepsilon_t$  je bílý šum a  $\theta_j$  jsou parametry. Střední hodnota tohoto procesu je nulová, a tedy jeho rozptyl je roven (4.12):

$$\sigma_y^2 = \gamma_0 = (1 + \theta_1^2 + \dots + \theta_q^2) \sigma_\varepsilon^2 \quad (4.12)$$

a jeho autokorelační funkce má tvar (4.13):

$$\rho_k = \frac{\theta_k + \theta_1 \theta_{k+1} + \dots + \theta_{q-k} \theta_q}{1 + \theta_1^2 + \dots + \theta_q^2} \quad (4.13)$$

pro  $k = 1, \dots, q$  a  $\rho_k = 0$  pro  $k > q$ .

#### 4.2.1.2 Proces $MA(1)$

Tento proces  $MA(1)$  má tvar (4.14):

$$y_t = \varepsilon_t + \theta_1 \varepsilon_{t-1}, \quad (4.14)$$

s podmínkou invertibility  $|\theta| < 1$ .

Jeho autokorelační funkce má tvar (4.15):

$$\rho_1 = \frac{\theta_1}{1 + \theta_1^2} \quad (4.15)$$

$\rho_k = 0$  pro  $k > 1$ .

Kvadratická rovnice (4.16) má jen jeden kořen, který splňuje podmínku invertibility.

$$\rho_1 \theta_1^2 - \theta_1 + \rho_1 = 0 \quad (4.16)$$

Vztah (4.17) pro výpočet parametru  $\theta$  na základě znalosti autokorelační funkce  $\rho$  je:

$$\theta_1 = \frac{1 - \sqrt{1 - 4\rho_1^2}}{2\rho_1}. \quad (4.17)$$

#### 4.2.1.3 Proces MA(2)

Proces MA(2) má tvar (4.18):

$$y_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2}, \quad (4.18)$$

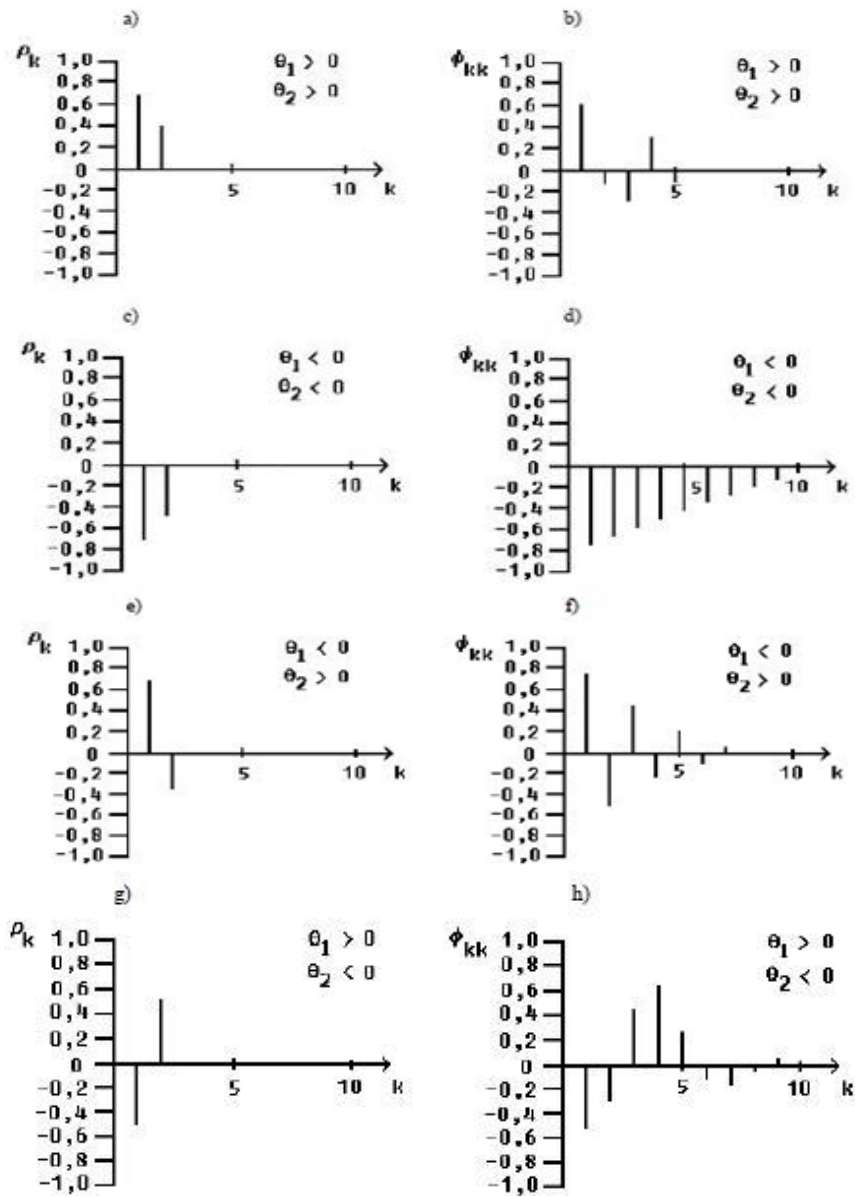
s podmínkou invertibility  $\theta_2 + \theta_1 > -1$ ,  $\theta_2 - \theta_1 > -1$ ,  $-1 < \theta_2 < 1$ .

Autokorelační funkce tohoto procesu má tvar (4.19):

$$\rho_1 = \frac{\theta_1(1 + \theta_2)}{1 + \theta_1^2 + \theta_2^2} \quad \rho_2 = \frac{\theta_2}{1 + \theta_1^2 + \theta_2^2} \quad (4.19)$$

$\rho_k = 0$  pro  $k > 2$ .

Na následujícím obrázku 1 vidíme ACF a PACF tohoto procesu. PACF buď exponenciálně klesá, nebo má tvar exponenciálně klesající sinusoidy.



Obrázek 1: ACF a PACF procesu MA(2)

Zdroj: publikace [5]

## 4.2.2 AR – autoregresní proces

Ukázku procesu začneme nejprve modelem obecně, a poté pro první i druhý řád.

### 4.2.2.1 Proces AR(p)

Autoregresní proces AR(p) má následující tvar (4.20):

$$y_t = \varepsilon_t + \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \dots + \varphi_p y_{t-p}, \quad (4.20)$$

nebo s operátorem zpětného posunutí  $B$  jako (4.21):

$$\varphi(B)y_t = \varepsilon_t, \quad (4.21)$$

kde

$$\varphi(B) = 1 - \sum_{j=1}^p \varphi_j B^j \quad (4.22)$$

je autoregresní operátor. Střední hodnota tohoto procesu AR(p) je nulová a jeho autokorelační funkce vyhovuje následující soustavě diferenčních rovnic (4.23):

$$\rho_k = \varphi_1 \rho_{k-1} + \varphi_2 \rho_{k-2} + \dots + \varphi_p \rho_{p-k}, \quad k > 0. \quad (4.23)$$

Parametry se vypočítají pomocí Yueovy - Walkerovy soustavy rovnic:

$$\rho_1 = \varphi_1 + \dots + \varphi_p \rho_{p-1} \quad (4.24)$$

$$\rho_2 = \varphi_1 \rho_1 + \varphi_2 + \dots + \varphi_p \rho_{p-2}$$

.....

$$\rho_p = \varphi_1 \rho_{p-1} + \varphi_2 \rho_{p-2} + \dots + \varphi_p.$$

Díky této soustavě lineárních rovnic (4.24), resp. jejím řešením, určíme parametry  $\varphi$  vyjádřené pomocí  $\rho$  autokorelační funkce.

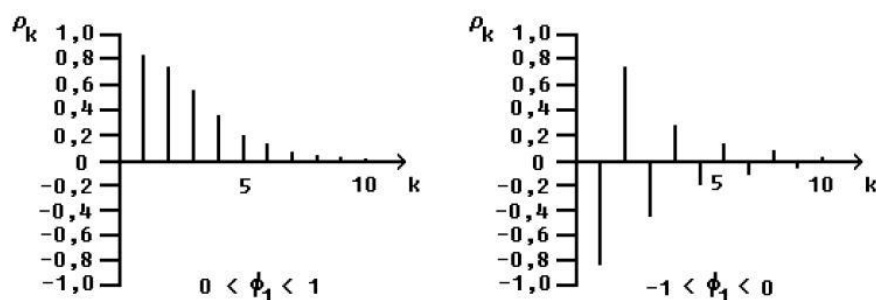
### 4.2.2.2 Proces AR(1)

Tento proces AR(1) má následující tvar (4.25):

$$y_t = \varepsilon_t + \varphi_1 y_{t-1}, \quad (4.25)$$

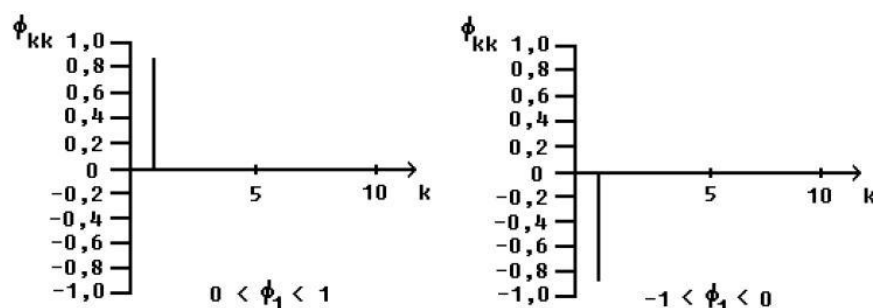
s podmínkou stacionarita  $|\varphi_1| < 1$ . Speciálně je autokorelační funkce  $\varphi_1 = \rho_1$ .

Na následujícím obrázku 2 vidíme tvary autokorelační funkce procesu AR(1), a další obrázek 3 znázorňuje charakteristické tvary parciální autokorelační funkce tohoto procesu.



Obrázek 2: ACF procesu AR(1)

Zdroj: publikace [5]



Obrázek 3: PACF procesu AR(1)

Zdroj: publikace [5]

### 4.2.2.3 Proces AR(2)

Tento autokorelační proces AR(2) má tvar (4.26):

$$y_t = \varepsilon_t + \varphi_1 y_{t-1} + \varphi_2 y_{t-2} \quad (4.26)$$

s podmínkou stacionarity  $\varphi_2 + \varphi_1 < 1$ ,  $\varphi_2 - \varphi_1 < 1$ ,  $-1 < \varphi_2 < 1$ .

Yuleova - Walkerova soustava rovnic pro proces AR(2) má tvar (4.27):

$$\rho_1 = \varphi_1 + \varphi_2 \rho_1 \quad (4.27)$$

$$\rho_2 = \varphi_1 \rho_1 + \varphi_2$$

s řešením parametrů (4.28):

$$\varphi_1 = \frac{\rho_1(1 - \rho_2)}{1 - \rho_1^2} \quad \varphi_2 = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2}. \quad (4.28)$$

### 4.2.3 ARMA – smíšený proces

Opět nejprve začneme ukázkou procesu obecně, a poté pro prvním řádem obou složek.

#### 4.2.3.1 Proces ARMA(p,q)

Smíšený proces řádu p a q, který značíme jako ARMA(p,q) zobrazuje následující rovnice (4.29):

$$y_t = \varphi_1 y_{t-1} + \dots + \varphi_p y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}, \quad (4.29)$$

nebo s operátorem zpětného posunutí B jako (4.30):

$$\varphi(B)y_t = \theta(B)\varepsilon_t. \quad (4.30)$$

Podmínka stacionarity smíšeného procesu je totožná s podmínkou stacionarity procesu AR(p) a stejně tak podmínka invertibility je stejná jako podmínka invertibility procesu MA(q). Střední hodnota smíšeného procesu je nulová a jeho autokorelační funkce  $\rho_k$  vyhovuje podobné soustavě diferenčních rovnic jako v případě autoregresního procesu. Tato funkce má tvar:

$$\rho_k = \varphi_1 \rho_{k-1} + \varphi_2 \rho_{k-2} + \dots + \varphi_p \rho_{p-k}, \quad k > q. \quad (4.31)$$

#### 4.2.3.2 Proces ARMA(1,1)

Tento proces má tvar (4.32):

$$y_t = \varphi_1 y_{t-1} + \varepsilon_t + \theta_1 \varepsilon_{t-1}, \quad (4.32)$$

s podmínkou stacionarity  $|\varphi_1| < 1$  a s podmínkou invertibility  $|\theta_1| < 1$ .

Soustava (4.31) bude mít pro tento proces tvar (4.33):

$$\rho_k = \varphi_1 \rho_{k-1}, \quad k > 1. \quad (4.33)$$

## 4.3 Lineární modely nestacionárních časových řad

V předchozí kapitole jsme se zabývali pouze stacionárními procesy. Především v ekonomické praxi se však velmi často setkáváme s časovými řadami tvořenými nestacionárními stochastickými procesy. Tato nestacionarita procesu může být způsobena v čase se měnící střední hodnotou procesu či v čase se měnícím rozptylem procesu.

### 4.3.1 Proces náhodné procházky („RandomWalk Proces“)

Proces náhodné procházky je zvláštní případ autokorelačního procesu AR(1), kde  $\phi_1 = 1$ .

Tento proces má tvar (4.34):

$$X_t = X_{t-1} + a_t. \quad (4.34)$$

Tento proces můžeme vyjádřit i vztahem pomocí operátoru zpětného posunutí jako (4.35):

$$(1 - B)X_t = a_t \quad (4.35)$$

a v případě jeho modifikace jako:

$$(1 - B)^d X_t = a_t, \quad d = 2, 3, \dots \quad (4.36)$$

Tento modifikovaný proces (4.36) se nazývá integrovaný proces řádu  $d$  a označuje se jako  $I(d)$ . Integrovaným se nazývá proto, že jeho první diference je proces bílého šumu. Proces náhodné procházky se tvoří kumulováním náhodných veličin tvořících proces bílého šumu.

Podmíněná střední hodnota procesu se v čase mění a podmíněný rozptyl je konstantní v čase. [2], [3]

### 4.3.2 Procesy ARIMA

Procesy ARIMA umožňují popis procesů, u nichž dochází ke změnám úrovně a tyto změny mají nesystematický náhodný charakter, jak je to běžné u většiny časových řad v praxi. Tyto modely modelují stochasticky vedle náhodných fluktuací i trend. Při konstrukci těchto modelů již nepožadujeme stacionaritu modelované časové řady, ale řada musí být převeditelná na stacionární přechodem k prvním nebo vyšším diferencím. Příkladem je řada, která vykazuje stacionaritu až na náhodné změny ve velikosti své úrovně, lze stacionarizovat přechodem k řadě prvních diferencí.

Má tvar (4.37):

$$\Delta y_t = y_t - y_{t-1}. \quad (4.37)$$

Časovou řadu, vykazující stacionaritu až na náhodné změny ve sklonu úrovně, můžeme stacionarizovat přechodem k řadě druhých diferencí.

Integrovaný smíšený model ARIMA(p,d,q) je možné definovat jako:

$$\varphi(B)w_t = \theta(B)\varepsilon_t, \quad (4.38)$$

kde

$$w_t = \Delta^d y_t \quad (4.39)$$

je  $d$ -tá diference modelovaného procesu  $y_t$  a (4.38) je stacionární model ARMA(p,q) pro proces  $w_t$ . Jde o to, že se nejprve provede stacionarizace pomocí vhodné diference modelované řady a vzniklá stacionární řada se modeluje pomocí modelu ARMA.

Model ARIMA (p,d,q) souhrnně zapíšeme jako (4.40):

$$\varphi(B)(1 - B)^d y_t = \theta(B)\varepsilon_t. \quad (4.40)$$

Operátor

$$v(B) = \varphi(B)(1 - B)^d \quad \rho_k = \varphi_1 \rho_{k-1} \quad k > 1 \quad (4.41)$$

nazýváme zobecněný autoregresní operátor. [7]

### **Transformace časové řady a určení řádu diferencování**

Cílem transformace časové řady je linearizovat řadu tak, abychom ji mohli analyzovat modelem ARIMA nebo jeho speciálními případy (AR, MA, ARMA). Touto linearizací dosáhneme především toho, že náhodné šoky generující řadu mají opravdu charakter bílého šumu s konstantním rozptylem a často navíc s normálním rozdělením či jemu podobným. Pro ekonomické resp. finanční časové řady je často vhodná transformace logaritmická.

#### Postup pro určení řádu diferencování $d$ :

- Použije se grafický záznam řady. Pokud existují pochybnosti o stacionaritě řady, zakreslíme dále řadu prvních či druhých diferencí a subjektivně posoudíme stacionaritu těchto diferencovaných časových řad.
- Pokud je zapotřebí rozhodnout na základě objektivnějších metod, použijeme odhad autokorelační funkce. Je důležité sestavit ACF dané řady a diferencí této řady. Pokud hodnoty ACF klesají přibližně lineárním tempem, pak je to příznakem toho, že alespoň jeden kořen autoregresního operátoru je velice blízký jedné a je tedy třeba provést další diferencování



- Dalším krokem je použít metodu, při níž se posuzuje velikost odhadnutého rozptylu dané časové řady a rozptylů jejich diferencí. Za  $d$  se zvolí ta hodnota, která dává nejmenší odhadnutý rozptyl. Platí totiž, že při postupném diferencování hodnoty odhadnutých rozptylů klesají, dokud není stacionarita dosažena a pak opět začnou růst.

Časovou řadu není vhodné diferencovat vícekrát, než je nezbytně nutné. Jinak by mohlo nastat přediferencování řady. Většinou používáme řád diferencování  $d$  menší než 2. [7]

#### 4.4 Modely stacionárních a nestacionárních sezonních časových řad

Důležitou vlastností mnoha krátkodobých ekonomických časových řad je sezonnost. Sezonní složkou časových řad rozumíme periodické kolísání mající systematický charakter. Toto kolísání se odehrává během jednoho kalendářního roku a každý rok se ve stejné nebo změněné podobě opakuje. U denních finančních časových řad se může také projevit pravidelné periodické kolísání (ve dnech během týdne). Tradičním předpokladem je, že sezonní složka časové řady má pravidelný deterministický charakter. Skutečnost bývá však komplikovanější, např. v případě ekonomických časových řad je rozumné předpokládat, že sezonní složka má stochastický charakter. Takový to typ systematickosti můžeme dále klasifikovat na stacionární a integrovaný. [2]

Druhy sezonních procesů:

##### Sezonní stacionární procesy

- Sezonní autoregresní proces SAR(P)
- Sezonní proces klouzavých průměrů SMA(Q)
- Smíšený proces SARMA(p,q)(P,Q)

##### Sezonní nestacionární procesy

- Integrované procesy
- Integrovaný proces SARIMA(p,d,q)(P,D,Q)

Podrobnější popisy a konstrukce modelů jsou uvedeny v publikacích [2], [3], [4], [7].

#### 4.5 Konstrukce modelu v Box-Jenkinsově metodologii

V této části si velmi zjednodušeně ukážeme, jak lze pro danou stacionární řadu systematicky ve třech fázích vybudovat příslušný model. Fázemi B-J metodologie jsou

identifikace modelu, odhad parametrů modelu a verifikace modelu. Složitější a podrobnější popisy fází konstrukce modelů v B-J metodologii zaznamenávají mnohé publikace [2], [3], [4], [7], proto je zbytečné se dopodrobna rozepisovat i v této práci. Pro čtenáře je především důležité nastínění této problematiky.

### **Identifikace modelu**

Základním úkolem je vybrat typ modelu (AR, MA či ARMA) a explicitně určit řád modelu. Velmi zjednodušeným postupem je, že si nejprve ověříme stacionaritu a případně data stacionarizujeme. Dále zkontrolujeme nulovost střední hodnoty, pokud by byla nenulová, je nutné řadu vycentrovat. Vlastní identifikace je založena na analýze odhadů autokorelační a parciální autokorelační funkce. Tyto odhady mohou být navzájem silně korelované, doporučuje se tedy netrvat na jednoznačném určení řádu modelu, ale vyzkoušet více modelů. Závěrem je spočtení počátečních odhadů parametrů. Tyto odhady bývají velmi hrubé, a proto se využívají jako výchozí hodnoty při odhadování parametrů modelu v následující etapě.

### **Odhad parametrů modelu**

Odhady parametrů již identifikovaného modelu postupně upřesňujeme pomocí iteračních postupů. Klasickou metodou je metoda nejmenších nelineárních čtverců. V praxi se uplatňuje podmíněná metoda nejmenších nelineárních čtverců nebo metoda nepodmíněná nejmenších nelineárních čtverců. Určení odhadů parametrů v nynější době provádíme pomocí softwareových produktů. Závěrečným krokem této části je určení přesnosti získaných odhadů parametrů.

### **Verifikace (ověřování) modelu**

Cílem verifikace modelu je potvrzení správnosti a pravdivosti navrženého modelu. Pro verifikaci se nejčastěji používají metody, jako metoda přeparametrizování modelu, metoda odhadnutých reziduí nebo metoda založená na Portmanteau testu. Například jsou-li zjištěny závažnější odchylky modelu od původní časové řady, opakujeme postup či zvýšíme model (na místo MA vybudujeme ARMA), případně zvýšíme řád modelu, a pak opět hledáme odchylky u původní řady. Tento uvedený příklad se týká metody přeparametrizování modelu.

## 5 Experimentální část

Analýzy zabývající se jakoukoli kvantitativní analýzou finančních dat, především statistickým zpracováním finančních dat, mezi které patří klasická popisná statistická analýza, statistická identifikace, odhad modelu a verifikace modelu, statistické testování různých finančních principů a modelování předpovědí v rámci vytvořeného modelu, nazýváme analýzy finančních časových řad, nebo také ekonometrie finančních trhů, finanční ekonometrie a kvantitativní analýza finančních dat. Zpracování těchto finančních dat se provádí v námi zvoleném softwaru s využitím všech známých informací. Následně probíhá vyhodnocení získaných výsledků z pohledu jejich platnosti teorie v praxi a dopad na reálná finanční data.

Cílem této části je pro vybraná finanční data sestrojít vhodný teoretický model a ten použít na reálná data. Sestrojení daného modelu umožňuje porozumět mechanismu časové řady a pochopit podmínky a vazby, které jsou mezi těmito hodnotami. Na základě těchto modelů, podmínek a předpokladů lze simulovat vliv působící změny ve vývoji finanční časové řady. My se zde budeme zabývat pouze základními B-J. modely, přestože se v nynější době pracuje s jejich upravenými modely, např. ARCH, GARCH, EGARCH...atd. Je nutné ověřit adekvátnost zvolených metod a postupů. V závěru bude popsána prognóza finančních časových řad.

Finanční časové řady se vyskytují běžně kolem nás, ale hlavním zdrojem pro tyto řady jsou statistické úřady, centrální banky, finanční instituce, finanční agentury, penzijní fondy, pojišťovny, banky, poradenské firmy, burzy a mnoho dalších.

Pro naši experimentální část byly vybrány dvě finanční časové řady. První řadou jsou denní hodnoty burzovního indexu S&P 500 za rok 2011. Další zkoumanou finanční časovou řadou je vývoj kurzu GBP/USD od počátku roku 2003 do konce roku 2011.

Veškeré výpočty a modelování jsou prováděny ve statistickém softwaru STATISTICA.

### Software STATISTICA

Společnost StatSoft byla založena v roce 1984 v USA. Později pak vznikaly pobočky firmy po celém světě. Stala se tak jedním z předních světových dodavatelů na poli softwaru pro analýzu dat.

Produkty řady STATISTICA našly uplatnění v celé řadě nejrůznějších výrobních podniků, v oblasti vzdělávání a finančních institucí. Tento produkt získal cenná uznání jak od běžných

uživatelů, tak od odborníků. Od vydání první verze tohoto systému v roce 1993 je hodnocena nejlépe ve své kategorii ve všech nezávislých hodnoceních. [21]

STATISTICA je systém, který slouží více jak 600 000 uživatelům po celém světě pro podporu strategického i výkonného řízení financí, vztahů se zákazníky, kvality, rizik, při výzkumu, ve výuce a vývoji. Využívá k tomu rozsáhlou nabídku analytických metod, grafických výstupů, nástrojů pro automatizaci, spolupráci, reportování, správu dat, atd. [20]

Existuje mnoho způsobů, jak je možné data analyzovat. Na svých oficiálních webových stránkách společnost StatSoft odkazuje např. na metodu Credit Scoring, Fraud Detection a další. Pro naše účely byla vybrána analýza časových řad a testování hypotéz. Studenti Univerzity Pardubice si po přihlášení do univerzitního intranetu mohou tento software bezplatně stáhnout. Licenci je nutné každoročně obnovovat. Díky tomu bylo možné provést modelování v tomto programu.

## **5.1 Index S&P 500 (Standard & Poor's 500)**

Tento index je jedním ze tří nejsledovanějších amerických indexů. Považuje se za standardní měřítko výkonnosti amerického akciového trhu. Obsahuje akcie 500 nyní již výhradně amerických společností, které reprezentující dané odvětví, mají velký podíl na trhu a jsou dostatečně stabilní. Převážná většina akcií v indexu S&P 500 zahrnutých je obchodována na NYSE, ostatní na Nasdaq a zlomek na AMEX.

S&P 500 vznikl v roce 1957, ale tvůrci propočítali jeho hodnoty zpětně do roku 1926. Změny cen akcií společností s vyšší tržní kapitalizací budou mít větší vliv na jeho hodnotu než změny cen akcií malých společností, protože je vážen podle tržní kapitalizace.

Hlavní výhodou indexu S&P 500 je, že představuje asi 70 % celkové tržní kapitalizace amerického akciového trhu a také to, že je vhodně diverzifikován<sup>3</sup>. Díky tomu je dostatečně americká ekonomika reprezentována. Ovšem jeho nevýhodou je, že nezahrnuje akcie zahraničních společností, které na americkém akciovém trhu tvoří nezanedbatelnou část.

Při zavedení indexu v roce 1957 byla jeho hodnota propočítána na 45 bodů, od roku 1998 do současnosti se pohybuje v rozmezí 1000 až 1500 bodů. Roční zhodnocení přesahuje 7 %. Dlouhodobého rekordu 1 527 bodů dosáhl index v březnu 2000, ten byl překonán opět v říjnu 2007 (1565 b).

---

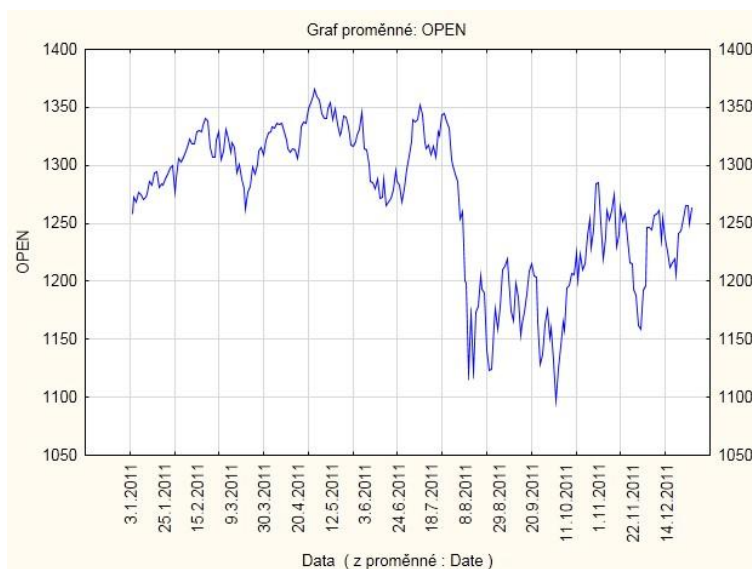
<sup>3</sup> Diverzifikace - rozlišování, rozčlenění, rozrůžňování; rozšíření počtu, členění

Vedle indexu S&P 500 existují ještě další indexy z produkce Standard & Poor's, například S&P 400 Mid Cap, který monitoruje vývoj akcií firem se středně velkou tržní kapitalizací, nebo S&P 600 Small Cap, který sleduje vývoj akcií firem s nízkou tržní kapitalizací. [19]

Jak bylo výše uvedeno, pro naše modelování byla vybrána finanční časová řada tvořena denními hodnotami indexu S&P 500 očištěnými o víkendy za rok 2011.

### 5.1.1 Identifikace modelu

Prvním krokem počátku analýzy bylo grafické zobrazení dané finanční časové řady indexu. Finanční časová řada byla znázorněna pomocí spojnicového grafu 5, který zobrazuje její charakteristické vlastnosti a rysy. Jednotlivé hodnoty této řady jsou vyznačeny osami grafu. Na ose x (horizontální osa) jsou vyznačeny denní hodnoty a na ose y (vertikální osa) hodnoty indexu. V grafu je možné vidět, že po počáteční stagnaci zhruba do poloviny roku došlo k prudkému poklesu a následnému mírnému vzestupu, který pokračoval až do konce roku.



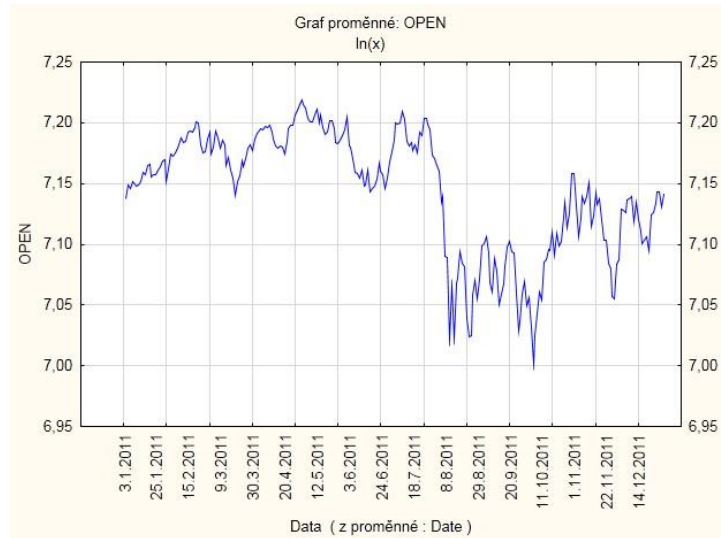
**Graf 5: Spojnicový graf znázorňující index S&P 500**

*Zdroj: vlastní zpracování pomocí sw. STATISTICA na základě dat z [19]*

Z grafu je také zřejmé, že časová řada je nestacionární a že neobsahuje sezónní složku. Řadu dále stabilizujeme logaritmickou transformací (přirozený logaritmus) danou vztahem (5.1):

$$x_t^* = \ln x_t, \quad (5.1)$$

kde  $x^*$  jsou hodnoty původní řady. Tuto transformovanou řadu zobrazuje následující graf 6.

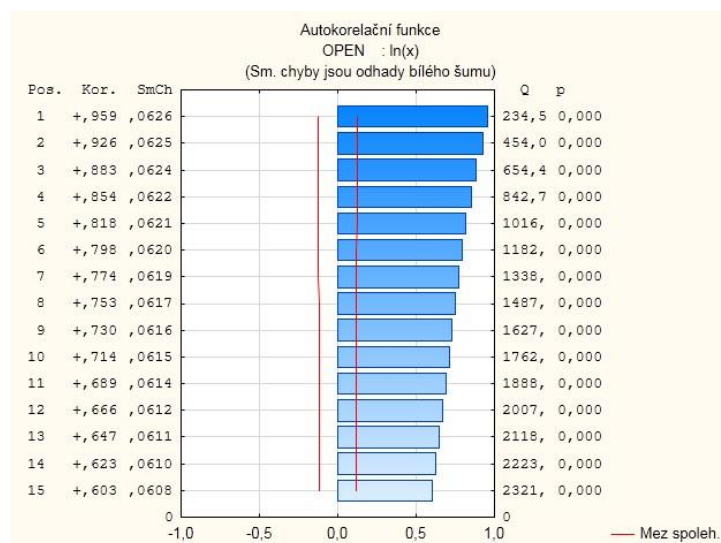


**Graf 6: Logaritmovaná časová řada indexu S&P 500**

*Zdroj: vlastní zpracování pomocí sw. STATISTICA na základě dat z [19]*

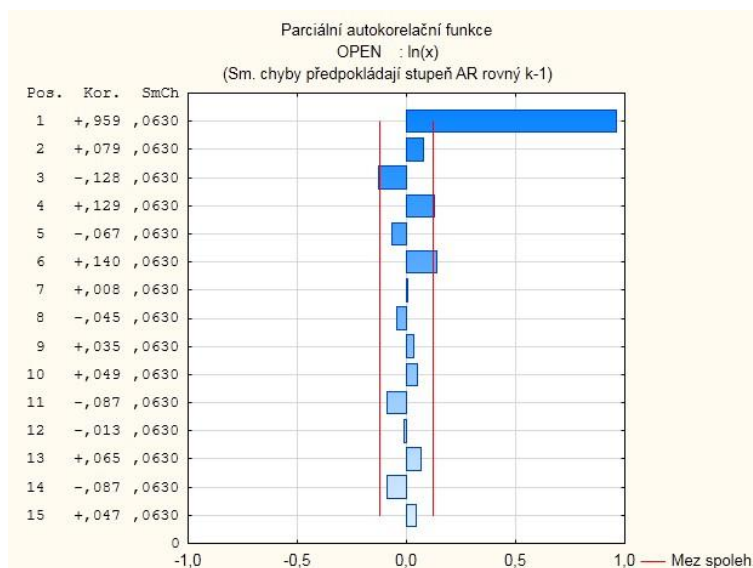
Logaritmovaná časová řada je zjevně nestacionární, což potvrzuje i tvar ACF a PACF [grafy 7, 8]. Hodnoty ACF klesají velmi pomalu, přibližně lineárním tempem, vykazují přítomnost deterministického a stochastického trendu. Stejně jako u PACF je první hodnota blízká jedné, předpokládáme tedy, že finanční časová řada je typu  $I(1)$ , a proto ji budeme stacionarizovat první diferencí. Tyto diference jsou dány vztahem (5.2), kde  $p$  označuje místo potřebné diference.

$$x_t^{**} = x_t^* - x_{t-p}^* \quad (5.2)$$



**Graf 7: ACF logaritmované časové řady indexu S&P 500**

*Zdroj: vlastní zpracování pomocí sw. STATISTICA na základě dat z [19]*

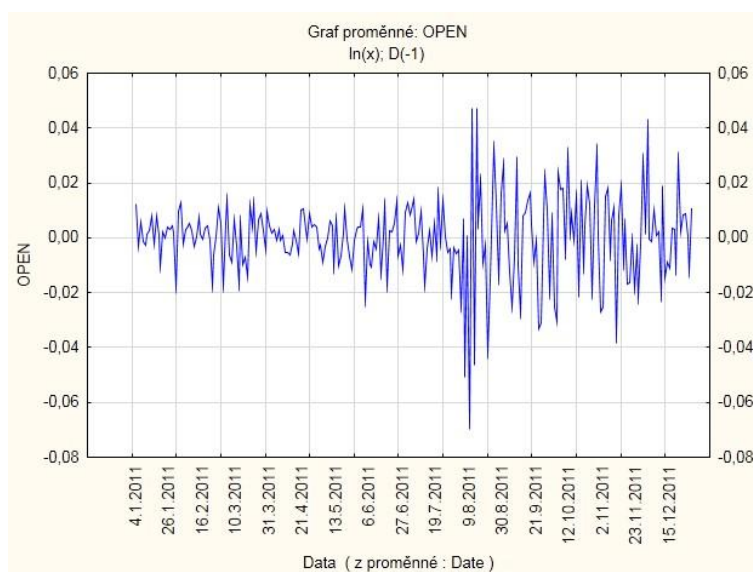


**Graf 8: PACF logaritmované časové řady indexu S&P 500**

*Zdroj: vlastní zpracování pomocí sw. STATISTICA na základě dat z [19]*

Po následné první diferenci je z grafu 9 patrné, že časová řada je stacionární. Proto není nutné řadu analyzovat druhou diferencí. Na tomto grafu je také vidět měnící se průběh volatility časové řady. Je možné říci, že čím je volatilita nižší, tím odhad modelu je přesnější. Také je zřejmé, že nesystematická složka vykazuje přítomnost podmíněné heteroskedasticity.

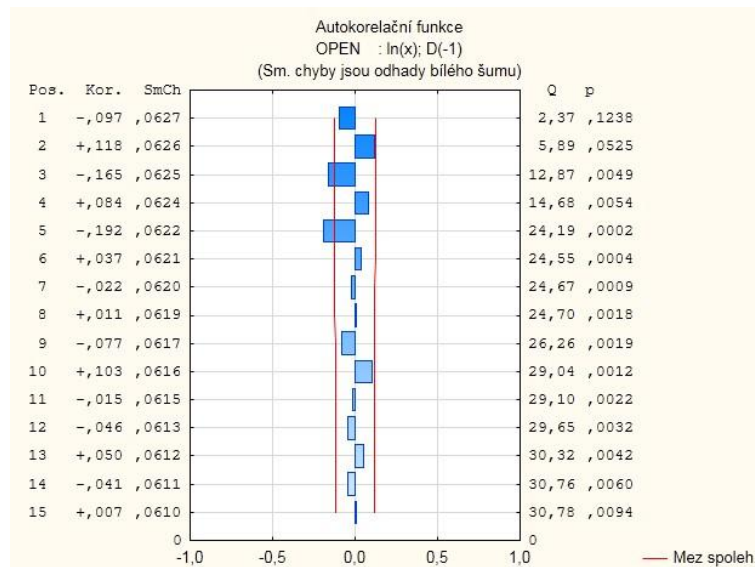
Pro kontrolu byla stacionarita potvrzena pomocí ACF a PACF [graf 10, 11]. Tyto funkce ukazují, že jsme se zbavili silných závislostí.



**Graf 9: Logaritmy výnosů řady indexu S&P 500**

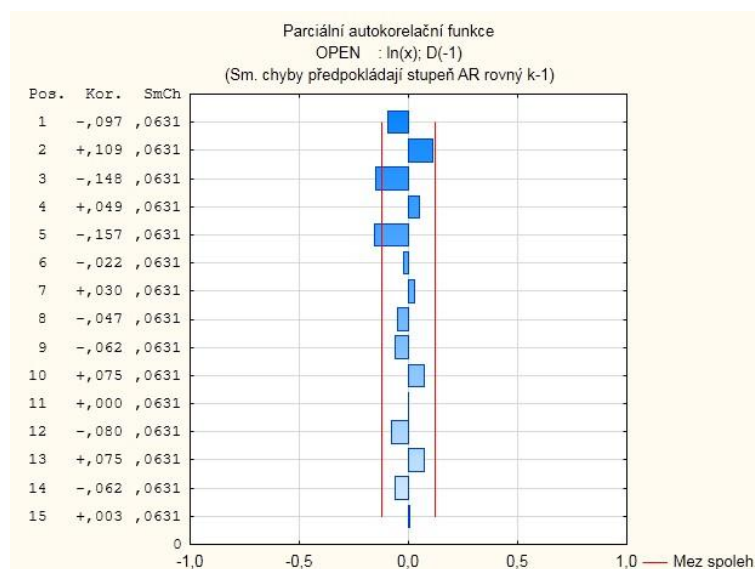
*Zdroj: vlastní zpracování pomocí sw. STATISTICA na základě dat z [19]*

V grafu 10 také vidíme, že autokorelační koeficienty překračují danou hladinu významnosti ( $\alpha = 0,05$ ), můžeme je tedy pokládat za nenulové, tzn. časová řada obsahuje systematickou složku. Přítomnost této složky vidíme ve sloupci pod písmenem  $Q$  (Box-Ljungova statistika) na pravé straně zobrazeného grafu ACF. Z tohoto grafu jsme si také ověřili předpoklad autokorelace pomocí autokorelační funkce. Pro toto ověření je také možnost použít Durbinův - Watsonův test.



**Graf 10: ACF první diference indexu S&P 500**

*Zdroj: vlastní zpracování pomocí sw. STATISTICA na základě dat z [19]*



**Graf 11: PACF první diference indexu S&P 500**

*Zdroj: vlastní zpracování pomocí sw. STATISTICA na základě dat z [19]*



Stacionaritu jsme určili pomocí grafů, tedy subjektivní metodou, která je velice jednoduchá a velmi efektivní. Určení a ověřování řádu diferencí je také možné ověřit testem poměrem rozptylů a testováním jednotkových kořenů (např. D-F testy, Phillipsovými - Perronovými testy, KPSS testy, atd.) [4], což v tomto případě není nutné.

Proměnná	Popisné statistiky								
	N platných	Průměr	Medián	Minimum	Maximum	Rozptyl	Sm.odch.	Šikmost	Špičatost
Logaritmy výnosů	251	0.000016	0.001279	-0.067520	0.044981	0.000208	0.014411	-0.536757	2.926853

**Tabulka 1: Popisné statistiky logaritmů výnosů indexu S&P 500**

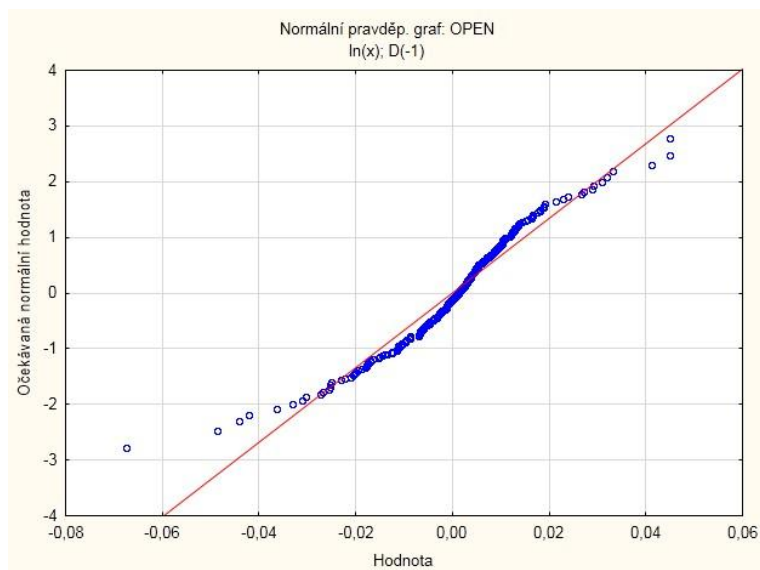
*Zdroj: vlastní zpracování pomocí sw. STATISTICA na základě dat z [19]*

V tabulce 1 můžeme vidět popisné statistiky logaritmů výnosů, tj. logaritmy koeficientů růstu indexu S&P 500. Z těchto popisných statistik můžeme vyčíst, že se jedná o 251 platných pozorování, která jsou rozložena v intervalu od -0,0675 do 0,04498. V tabulce také vidíme šikmost a špičatost. Z teorie víme, že symetrické rozdělení je charakteristické tím, že jeho šikmost je rovna nule a jeho špičatost je rovna číslu 3. O předpokladu normality u finančních časových řad je pojednáno v kapitole 3.4.

V našem případě je hodnota šikmosti  $SK_r = -0,536757$  záporné číslo. V literatuře [4] bylo dokázáno, že šikmost, vypočítaná z logaritmů výnosů finančních časových řad burzovních indexů, má s rostoucí agregací dat tendenci se zvětšovat, a toto budeme předpokládat i v našem případě. Vzhledem k tomu, že medián je malé číslo  $\tilde{x} = 0,001279$ , můžeme říci, že skutečné rozdělení je zešikmené tak, že záporné výnosy se objevují častěji než kladné výnosy.

Odhad špičatosti  $K_r = 2,926853$ , je velice blízký číslu 3, můžeme předpokládat normální rozdělení.

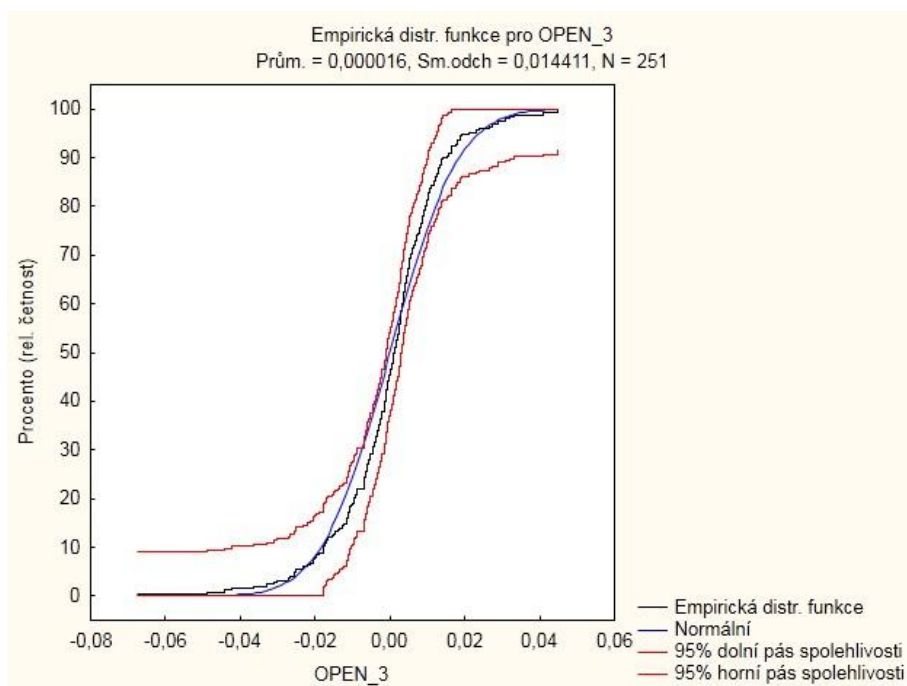
Pro ukázkou je znázorněn normální graf 12 pravděpodobnostního rozdělení logaritmů výnosů indexu. Na tomto grafu je zřetelné, že se nejedná zcela o normální rozdělení, což jsme si dokázali výše z hodnoty šikmosti. Toto rozdělení má „tlustší“ konce než normální rozdělení, a proto četnost výskytu extrémně vysokých kladných či záporných výnosů bude vyšší než za předpokladu normality. Můžeme proto říci, že výsledkem je nepodmíněné rozdělení s „tlustými“ konci.



**Graf 12: Normální rozdělení logaritmů výnosů indexu S&P 500**

*Zdroj: vlastní zpracování pomocí sw. STATISTICA na základě dat z [19]*

Na grafu 13 je možné vidět průběh empirické distribuční funkce logaritmů výnosů indexu.



**Graf 13: Empirická distribuční funkce logaritmů výnosů indexu S&P 500**

*Zdroj: vlastní zpracování pomocí sw. STATISTICA na základě dat z [19]*

### 5.1.2 Odhad parametrů modelu a následná verifikace

Dalším krokem B-J metodologie je odhad parametrů zkušného modelu a následná verifikace těchto parametrů i modelu. Nejprve odhadneme model ARIMA(1,1,0), tento model se skládá z první diference a modelu AR(1). Výsledky jsou zobrazeny v následující tabulce 2.

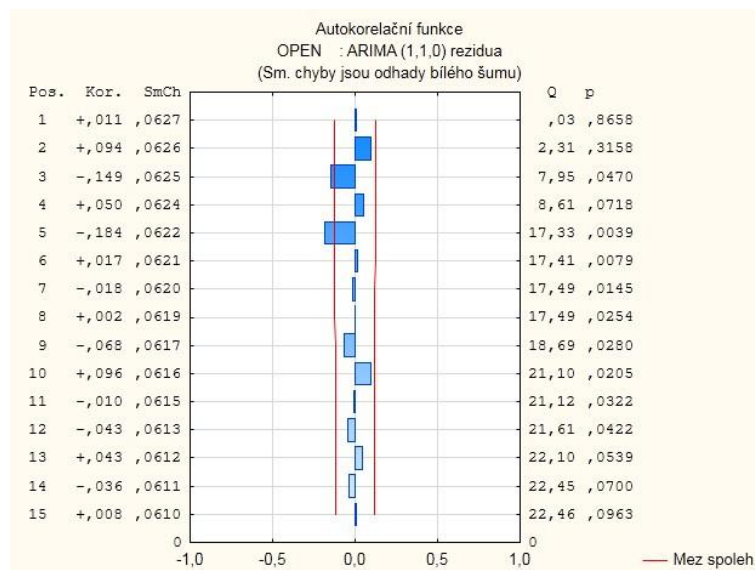
Vstup: OPEN (SP500-2011) Transformace: ln(x),D(1) Model:(1,1,0) PČ Rezid. = ,00021						
Paramet.	Param.	Asympt. SmCh	Asympt. t( 250)	p	Dolní 95% spol	Horní 95% spol
p(1)	-0,096781	0,063142	-1,53276	0,126601	-0,221138	0,027577

**Tabulka 2: Odhad parametru modelu ARIMA(1,1,0)**

*Zdroj: vlastní zpracování pomocí sw. STATISTICA na základě dat z [19]*

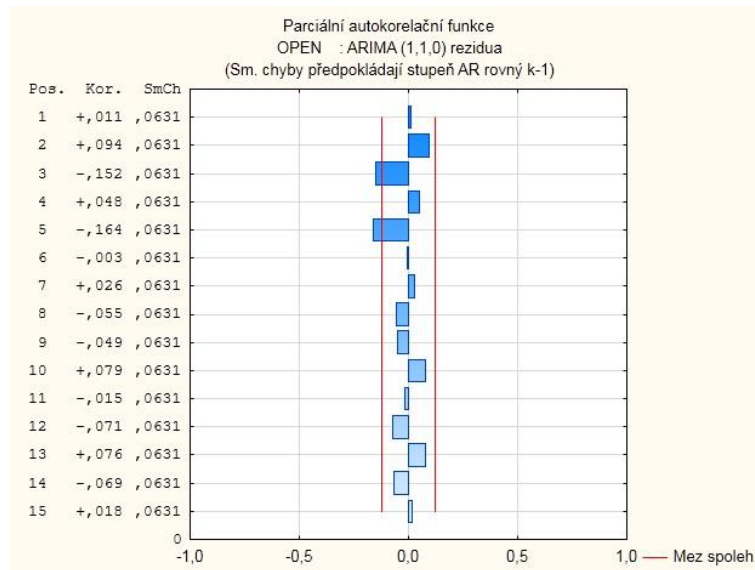
Významnost parametru otestujeme pomocí t-testu. Porovnáme-li hodnotu  $p = 0,126601$  s hladinou významnosti  $\alpha = 0,05$ , prokážeme, že střední hodnota a parametr nejsou různé od nuly. A také porovnáme-li hodnotu t-testu  $T_{\theta_1} = 1,53276$  s hodnotou  $t = 1,96$ , kde  $t$  je 95% kvantil rozdělení  $t$ , vidíme, že  $t$  je větší než testovací kritérium, a tedy parametr nemůžeme považovat za statisticky významný.

Dále to, že tento model nebude vhodný, nám indikují i grafy 14 a 15, zobrazující reziduální ACF a PACF indexu S&P 500 modelu ARIMA(1,1,0). Tyto grafy neukazují nesystematičnost (některé hodnoty leží vně tolerančních mezí).



**Graf 14: ACF reziduí indexu S&P 500 modelu ARIMA(1,1,0)**

*Zdroj: vlastní zpracování pomocí sw. STATISTICA na základě dat z [19]*



**Graf 15: PACF reziduí indexu S&P 500 modelu ARIMA(1,1,0)**

*Zdroj: vlastní zpracování pomocí sw. STATISTICA na základě dat z [19]*

Protože předchozí model nebyl vhodný, pokusíme se data namodelovat modelem ARIMA(0,1,1) – první diferenci jsme doplnili modelem o část MA(1). K této analýze použijeme obdobný postup odhadů parametru modelu a verifikaci jako u předešlého modelu.

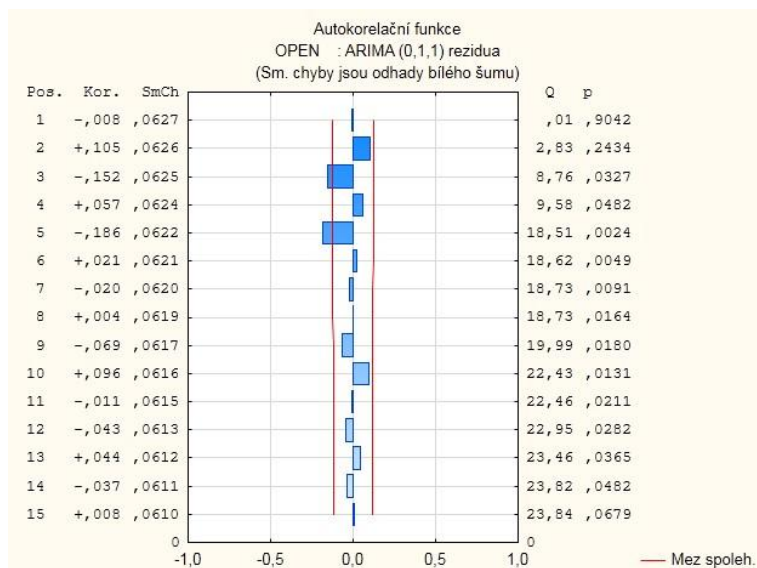
Vstup: OPEN (SP500-2011)						
Transformace: ln(x),D(1)						
Model:(0,1,1) PČ Rezid. = ,00021						
Paramet.	Param.	Asympt. SmCh	Asympt. t( 250)	p	Dolní 95% spol	Horní 95% spol
q(1)	0,081344	0,058659	1,386735	0,166758	-0,034184	0,196872

**Tabulka 3: Odhad parametru modelu ARIMA(0,1,1)**

*Zdroj: vlastní zpracování pomocí sw. STATISTICA na základě dat z [19]*

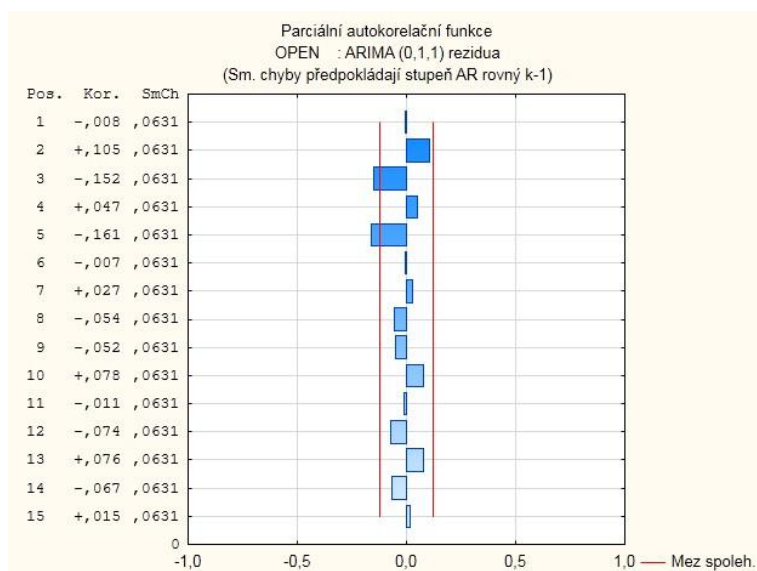
Tabulka 3 zobrazuje hodnoty parametru q(1). Významnost tohoto parametru otestujeme pomocí t-testu jako u předešlého modelu. Opět jsme došli k závěru, že testovaný parametr není statisticky významný. Hodnota t-testu (1,386735) je menší než testovací kritérium a i hodnota  $p = 0,166758$  je větší než hladina významnosti  $\alpha = 0,05$ .

Nevhodnost modelu nám indikují i následující grafy 16 a 17 zobrazující reziduální ACF a PACF indexu S&P 500 modelu ARIMA(0,1,1). Tyto reziduální grafy stále potvrzují existenci systematické složky (některé hodnoty leží vně tolerančních mezí).



**Graf 16: ACF reziduí indexu S&P 500 modelu ARIMA(0,1,1)**

*Zdroj: vlastní zpracování pomocí sw. STATISTICA na základě dat z [19]*



**Graf 17: PACF reziduí indexu S&P 500 modelu ARIMA(0,1,1)**

*Zdroj: vlastní zpracování pomocí sw. STATISTICA na základě dat z [19]*

Reziduální ACF i PACF mají některé hodnoty statisticky významně odlišné od nuly, a tedy daný model opět není vhodný, pokusíme se řadu prvních diferencí analyzovat sloučením obou modelů AR(1) a MA(1) a vytvořit tak smíšený model ARIMA(1,1,1). Pro modelování opět použijeme stejný postup odhadu parametrů a jejich kontroly.

Odhadnuté parametry zobrazuje následující tabulka 4.

Vstup: OPEN (SP500-2011)						
Transformace: ln(x),D(1)						
Model:(1,1,1) PČ Rezid. = ,00020						
Paramet.	Param.	Asympt. SmCh	Asympt. t( 249)	p	Dolní 95% spol	Horní 95% spol
p(1)	-0,871408	0,074319	-11,7252	0,000000	-1,01778	-0,725033
q(1)	-0,765596	0,092998	-8,2324	0,000000	-0,94876	-0,582434

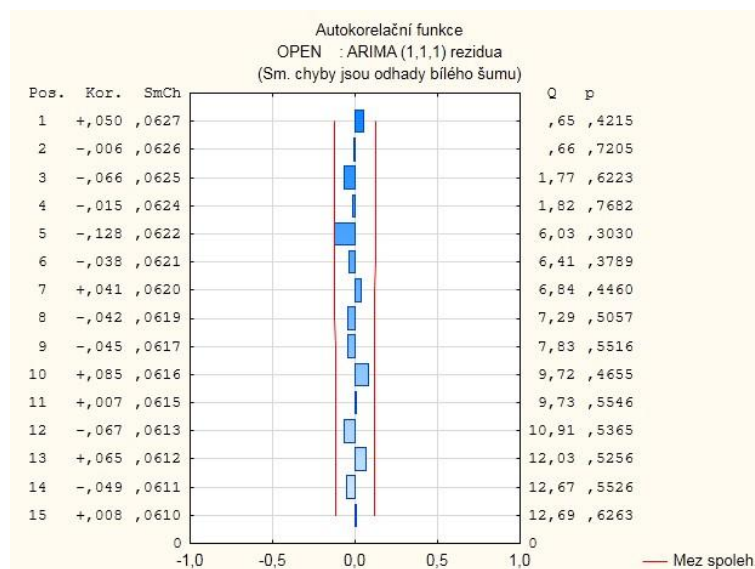
**Tabulka 4: Odhad parametrů modelu ARIMA(1,1,1)**

*Zdroj: vlastní zpracování pomocí sw. STATISTICA na základě dat z [19]*

Opět provedeme otestování parametrů  $p(1)$  a  $q(1)$ . U obou parametrů je hodnota  $p$  rovna nule. Když tuto hodnotu  $p$  porovnáme s hladinou významnosti  $\alpha = 0,05$ , prokážeme, že střední hodnota a parametr jsou různé od nuly, protože hodnota  $p$  je nižší než hladina významnosti.

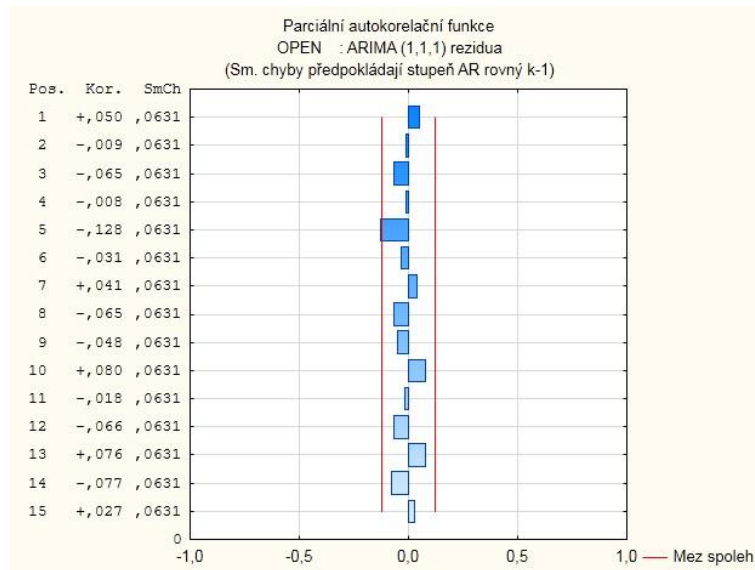
Dále porovnáme hodnoty t-testu každého parametru s hodnotou  $t = 1,96$ , kde  $t$  je 95% kvantil rozdělení  $t$ . Pro parametr  $p(1)$  máme hodnotu t-testu  $T_{01} = 11,7252$  a pro parametr  $q(1)$  je tato hodnota rovna 8,2324. Obě tyto hodnoty t-testu jsou vyšší než  $t = 1,96$ . [16] Výsledkem verifikace je tedy zjištění, že oba parametry mají statistickou významnost. Statisticky významné parametry software STATISTICA uvádí červenou barvou, což je vidět z tabulky 4 a i to nám dokazuje správnost provedených testů.

Vhodnost modelu nám ukazují i reziduální grafy ACF a PACF indexu S&P 500 [grafy 18, 19]. Hodnoty těchto reziduálních funkcí se snížily tak, že všechny leží uvnitř nebo na hranici intervalu spolehlivosti.



**Graf 18: ACF reziduí indexu S&P 500 modelu ARIMA(1,1,1)**

*Zdroj: vlastní zpracování pomocí sw. STATISTICA na základě dat z [19]*



**Graf 19: PACF reziduí indexu S&P 500 modelu ARIMA(1,1,1)**

*Zdroj: vlastní zpracování pomocí sw. STATISTICA na základě dat z [19]*

I tyto funkce, resp. statistickou významnost je možné verifikovat např. pomocí Portmanteaůva testu, který ověřuje nekorelovanost reziduí.

Na předešlém grafu 18 si opět všimněme výsledných hodnot ve sloupci pod písmenem  $Q$ , tyto hodnoty následně porovnáme s kritickou hodnotu  $\chi^2$ . Pro náš model má tedy  $Q$  statistika přibližně  $\chi^2$  rozdělení s  $K - p - q$  stupni volnosti. Nejprve si stanovíme  $K$ , což je vhodně zvolené číslo (doporučuje se blízké  $\sqrt{n}$ , kde  $n$  je počet pozorování). V našem případě  $n = 252$ , a tedy hodnota  $K$  bude rovna číslu 16. Následně vypočítáme stupně volnosti (5.3):

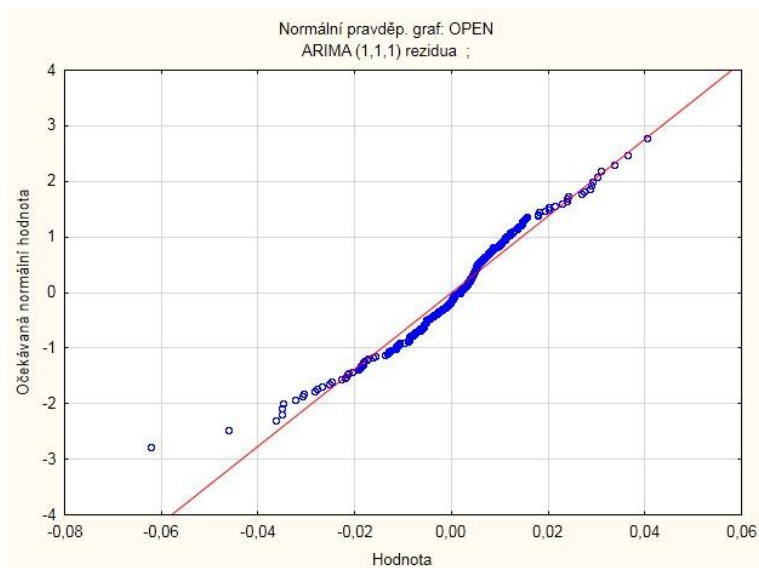
$$K - p - q = 16 - 1 - 1 = 14. \tag{5.3}$$

Kritická hodnota je

$$\chi^2_{14}(0,05) = 23,6848. \tag{5.4}$$

Vidíme, že hodnoty  $Q$  nepřesahují kritickou hodnotu (5.4), tento test opět potvrzuje správnost vybraného modelu.

Pokud porovnáme normální rozdělení reziduí indexu S&P 500 modelu ARIMA(1,1,1) [graf 20] a normální rozdělení logaritmu výnosů indexu [graf 12] vidíme, že se po modelování jednotlivé hodnoty více zarovnal k hladině významnosti.



**Graf 20: Normální rozdělení reziduí indexu S&P 500 modelu ARIMA(1,1,1)**

*Zdroj: vlastní zpracování pomocí sw. STATISTICA na základě dat z [19]*

### 5.1.3 Prezentace dosažených výsledků

V tomto bodě dosadíme odhadnuté parametry do rovnice vhodného modelu. Z předchozí teorie víme, že integrovaný model ARIMA(p,d,q) můžeme zapsat jako (5.5):

$$\varphi(B)(1 - B)^d y_t = \theta(B)\varepsilon_t \quad (5.5)$$

nebo jej lze vyjádřit jako (5.6):

$$y_t = \varphi_1 y_{t-1} + \dots + \varphi_p y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}. \quad (5.6)$$

Z kapitoly uvedené výše jsme zjistili, že vhodným modelem pro aplikaci na index S&P 500 se stal proces ARIMA(1,1,1). Jedná se o integrovaný model řádu I(1). Finanční časovou řadu indexu generovanou tímto procesem můžeme nazývat jako řadu s jedním jednotkovým kořenem.

Rovnice modelu ARIMA(1,1,1) má následující tvar (5.7):

$$y_t = \varphi_1 y_{t-1} + \varepsilon_t + \theta_1 \varepsilon_{t-1}. \quad (5.7)$$

Po dosazení našich výsledků má konkrétní model tvar (5.8):

$$y_t = -0,871408 y_{t-1} + \varepsilon_t - 0,765596 \varepsilon_{t-1}. \quad (5.8)$$



### 5.1.4 Predikce finanční časové řady indexu S&P 500 pomocí modelu

Následně je možné provést předpověď finanční časové řady pomocí modelu ARIMA(1,1,1). Z grafu 21 je patrná predikce budoucích 10 hodnot. Tedy předpověď na 10 dní dopředu. Tato předpověď je zobrazena červeně, zelenou barvou znázorňujeme 90% interval, ve kterém by měly nabývat hodnoty predikce. Původní řadu vykresluje graf modře.



**Graf 21: Predikce finanční časové řady indexu S&P 500**

*Zdroj: vlastní zpracování pomocí sw. STATISTICA na základě dat z [19]*

V následující tabulce 5 vidíme hodnoty predikce. První sloupec ukazuje data předpovědí dle daného modelu ARIMA(1,1,1), ve druhém a ve třetím sloupci jsou intervaly, které určují, v jakém rozmezí se může daný index pohybovat. Výpočet byl opět proveden pomocí statistického softwaru.

ČísloPřip.	Předpovědi; Model:(1,1,1) Sezónní posun:12 (SP500-2011) Vstup: OPEN Začátek původ. : 1 Konec původ. : 252		
	Předpově	Dolní 90,0000%	Horní 90,0000%
253	1259,574	1230,553	1289,279
254	1262,402	1223,538	1302,501
255	1259,937	1211,972	1309,801
256	1262,085	1207,531	1319,104
257	1260,213	1199,110	1324,429
258	1261,844	1195,410	1331,969
259	1260,423	1188,679	1336,497
260	1261,661	1185,315	1342,924
261	1260,582	1179,683	1347,029
262	1261,522	1176,518	1352,668

**Tabulka 5: Predikce finanční časové řady indexu S&P 500**

*Zdroj: vlastní zpracování pomocí sw. STATISTICA na základě dat z [19]*

## 5.2 Vývoj kurzu GBP/USD

Pro naše druhé modelování finanční časové řady byla jako nejvhodnější vybrána data čerpaná z webové stránky Citifinu, a to konkrétně kurz GBP/USD (britská libra vs. americký dolar). Data jsou sledována v denním intervalu a očištěna o víkendy od počátku roku 2003 do konce roku 2011.

Stejně kroky modelování provedené na předchozí finanční časové řadě indexu aplikujeme i na časovou řadu kurzu.

### 5.2.1 Identifikace modelu

Finanční časovou řadu kurzu GBP/USD znázorňuje spojnicový graf 22. Na ose x jsou vyznačeny denní hodnoty a na ose y hodnoty kurzu.

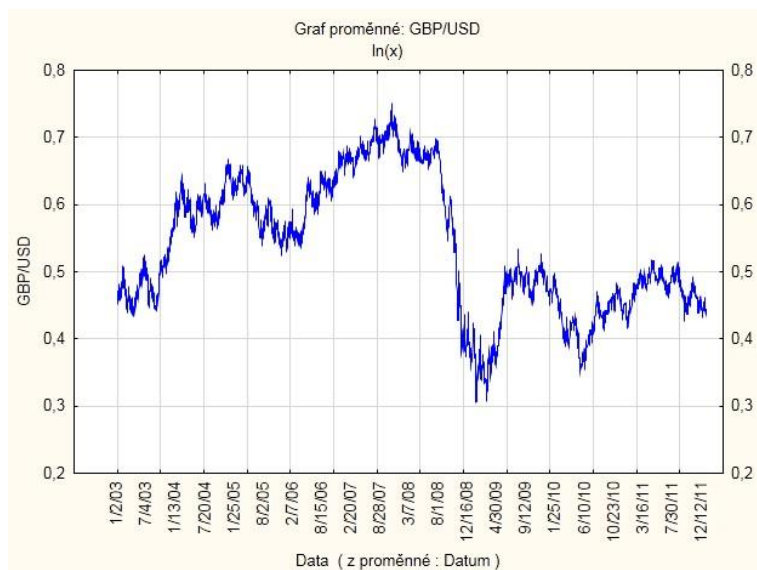


**Graf 22: Spojnicový graf znázorňující kurz GBP/USD**

*Zdroj: vlastní zpracování pomocí sw. STATISTICA na základě dat z [14]*

Z grafu 22 je zřejmé, že i tato finanční časová řada je nestacionární a že neobsahuje sezónní složku. Řadu dále stabilizujeme logaritmickou transformací.

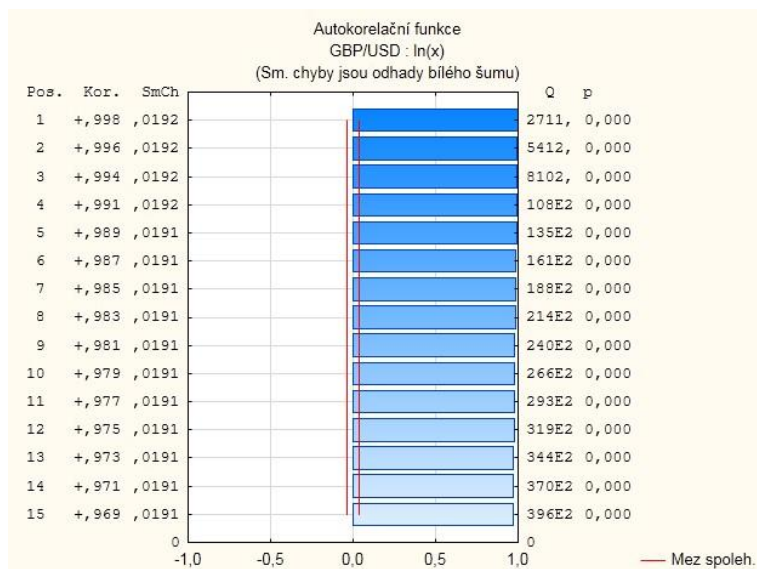
Tuto transformovanou řadu zobrazuje následující graf 23.



**Graf 23: Logaritmovaná časová řada kurzu GBP/USD**

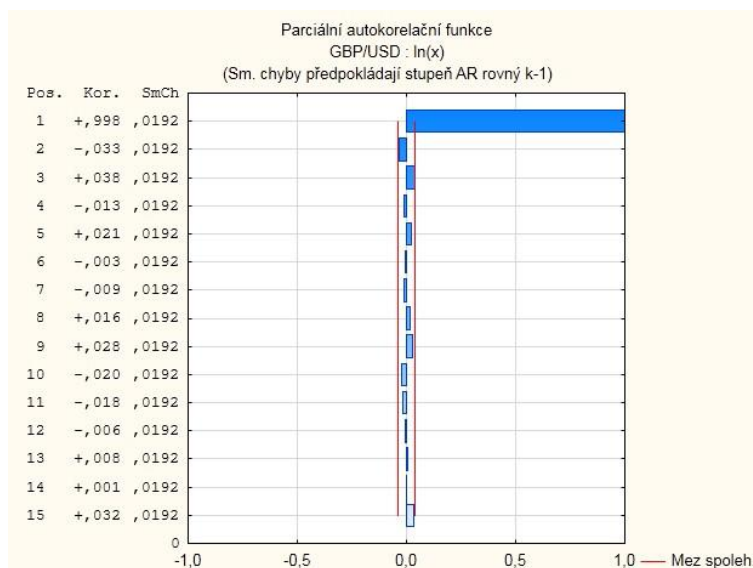
*Zdroj: vlastní zpracování pomocí sw. STATISTICA na základě dat z [14]*

Graf 24 funkce ACF a graf 25 funkce PACF také potvrzují nestacionaritu. Hodnoty ACF klesají velice pomalu a stejně jako u PACF je první hodnota blízká jedné, předpokládáme tedy, že finanční časová řada je typu  $I(1)$ , a proto ji budeme stacionarizovat první diferencí.



**Graf 24: ACF logaritmované časové řady kurzu GBP/USD**

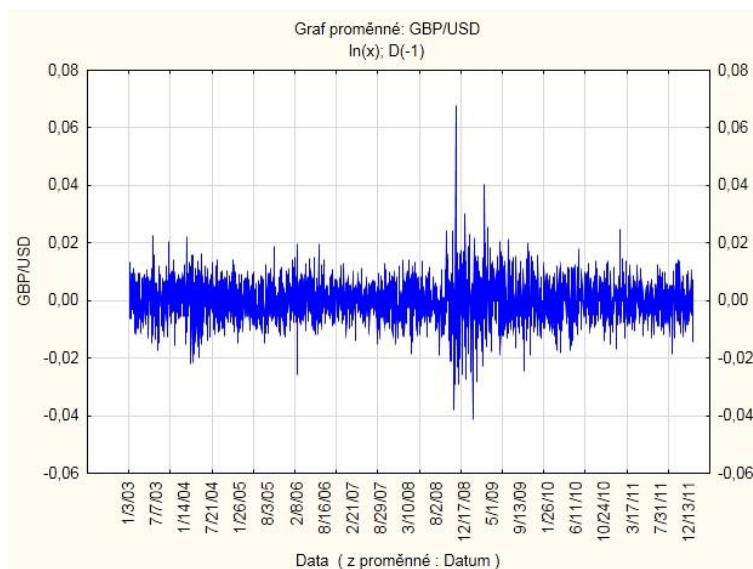
*Zdroj: vlastní zpracování pomocí sw. STATISTICA na základě dat z [14]*



**Graf 25: PACF logaritmované časové řady kurzu GBP/USD**

*Zdroj: vlastní zpracování pomocí sw. STATISTICA na základě dat z [14]*

Po následné první diferenci je z grafu 26 patrné, že časová řada je stacionární. Dále je na tomto grafu vidět měnící se průběh volatility časové řady. Také je zřejmé, že nesystematická složka modelu vykazuje přítomnost podmíněné heteroskedasticity.



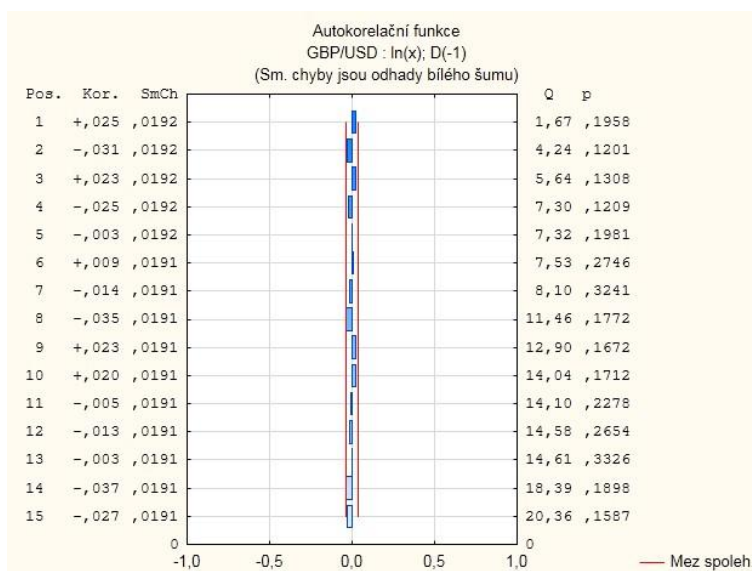
**Graf 26: Logaritmy výnosů řady kurzu GBP/USD**

*Zdroj: vlastní zpracování pomocí sw. STATISTICA na základě dat z [14]*

Kontrolu stacionarity potvrdíme pomocí ACF a PACF. Z grafu 27 funkce ACF je zřejmé, že autokorelační koeficienty leží uvnitř nebo na hranici tolerančních mezí. Podobné je tomu i u grafu 28 funkce PACF, nabízí se tedy možnost závěru, že by se mohlo jednat o model tvaru

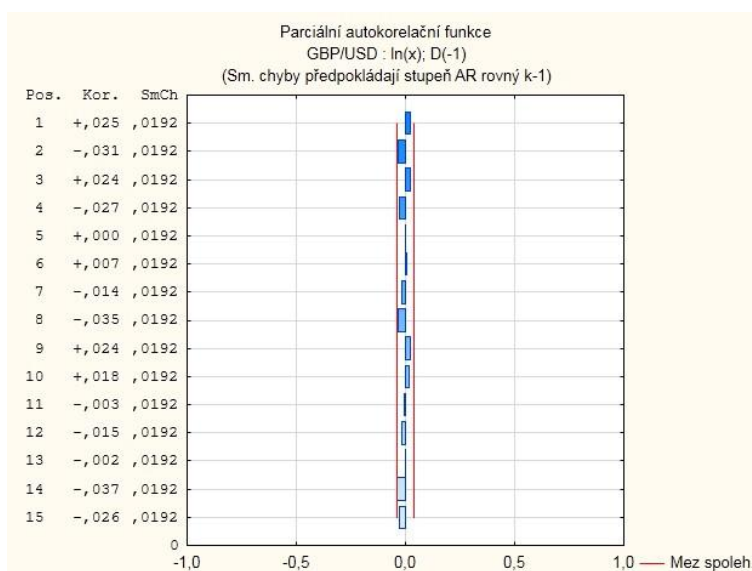
náhodné procházky. Tento proces se také nazývá integrovaný proces řádu jedna  $I(1)$ , protože jeho první diference je proces bílého šumu.

V našem případě přesto budeme časovou řadu analyzovat např. modely  $ARIMA(1,1,0)$ ,  $(0,1,1)$ ,  $(1,1,1)$ , atd. a pokusíme se zjistit, zda nenajdeme vhodnější model, kde by hodnoty ACF či PACF nebyly tak silně položeny na hranici tolerančních mezí.



**Graf 27: ACF první diference kurzu GBP/USD**

*Zdroj: vlastní zpracování pomocí sw. STATISTICA na základě dat z [14]*



**Graf 28: PACF první diference kurzu GBP/USD**

*Zdroj: vlastní zpracování pomocí sw. STATISTICA na základě dat z [14]*

Také i u kurzu, resp. u jeho logaritmů výnosů jsme si zjistili základní popisné statistiky, které jsou zobrazeny v tabulce 6.

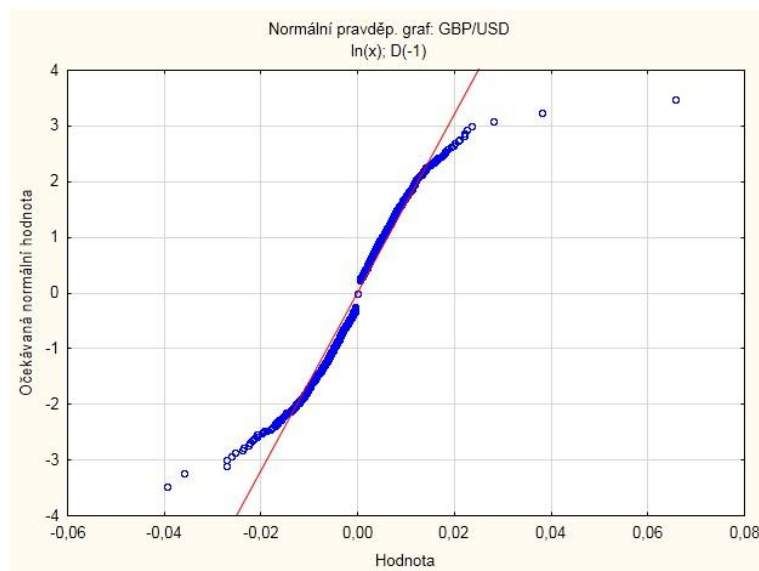
Proměnná	Popisné statistiky								
	N platných	Průměr	Medián	Minimum	Maximum	Rozptyl	Sm.odch.	Šikmost	Špičatost
Logaritmy výnosů	2719	-0,000009	0,00	-0,039525	0,065597	0,000036	0,006036	0,224627	8,193078

**Tabulka 6: Popisné statistiky logaritmů výnosů kurzu GBP/USD**

*Zdroj: vlastní zpracování pomocí sw. STATISTICA na základě dat z [14]*

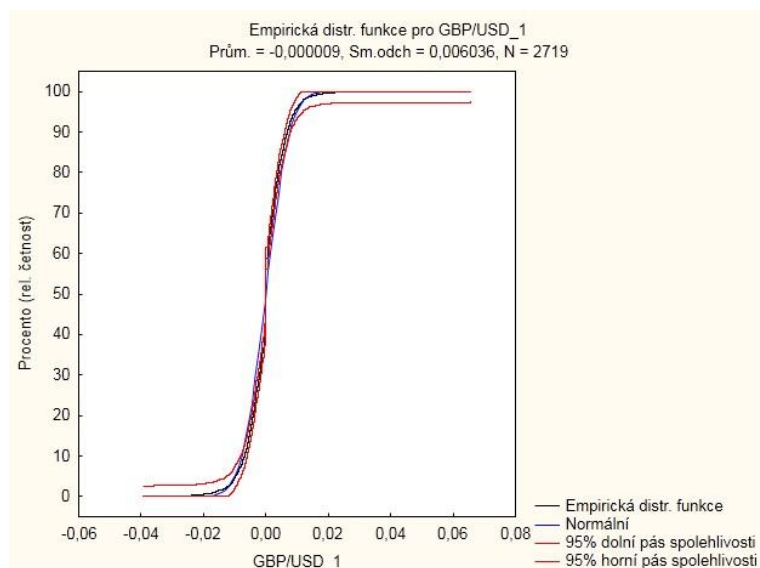
Z těchto popisných statistik vyčteme, že se jedná o 2719 platných pozorování, která jsou rozložena v intervalu od -0,039525 do 0,065597. Opět nás zajímají převážně hodnoty šikmosti a špičatosti, podle kterých rozhodujeme o normalitě. V našem případě je hodnota šikmosti  $SK_r = 0,224627$  a medián  $\tilde{x} = 0,00$ . Z hodnoty šikmosti vidíme, že zešikmení bude malé, protože hodnota se blíží k 0. Odhad špičatosti  $K_r = 8,193078$ , je poměrně vysoký. Můžeme proto říci, že výsledkem je nepodmíněné rozdělení, které je v porovnání s normálním rozdělením špičatější s „tlustšími“ konci.

Normální graf pravděpodobnostního rozdělení logaritmů výnosů [graf 29] a empirickou distribuční funkci [graf 30] kurzu zobrazují následující grafy.



**Graf 29: Normální rozdělení logaritmů výnosů kurzu GBP/USD**

*Zdroj: vlastní zpracování pomocí sw. STATISTICA na základě dat z [14]*



**Graf 30: Empirická distribuční funkce logaritmů výnosů kurzu GBP/USD**

*Zdroj: vlastní zpracování pomocí sw. STATISTICA na základě dat z [14]*

### 5.2.2 Odhad parametrů modelu a následná verifikace

Prvním krokem byl odhad procesu ARIMA(1,1,0), tento model se skládá z první diference a modelu AR(1). Výsledky jsou zobrazeny v tabulce 7.

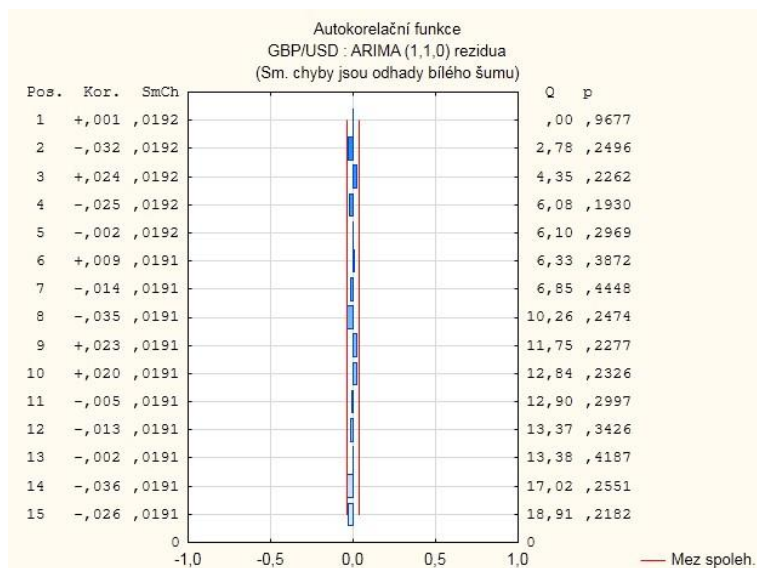
Vstup: GBP/USD (GBP-USD v Vývoj kurzů 03-11 (2))						
Transformace: ln(x),D(1)						
Model:(1,1,0) PČ Rezid. = ,00004						
Paramet.	Param.	Asympt. SmCh	Asympt. t( 2718)	p	Dolní 95% spol	Horní 95% spol
p(1)	0,024811	0,019186	1,293151	0,196069	-0,012810	0,062432

**Tabulka 7: Odhad parametru modelu ARIMA(1,1,0)**

*Zdroj: vlastní zpracování pomocí sw. STATISTICA na základě dat z [14]*

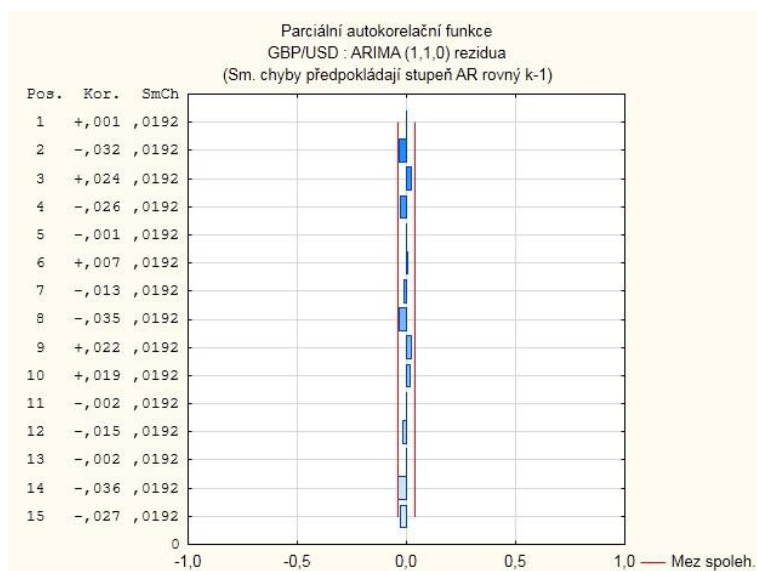
Významnost parametru jsme otestovali t-testem, protože hodnota  $p$  byla větší než hladina významnosti a naopak hodnota  $T_{01}$  menší než  $t$ , nemohl být tento parametr uznán za vhodný.

Grafy zobrazující reziduální ACF a PACF kurzu GBP/USD modelu ARIMA(1,1,0) [grafy 31, 32] sice indikují nesystematičnost, ale kvůli nevhodnosti parametru budeme dále modelovat kurz jiným procesem.



**Graf 31: ACF reziduí kurzu GBP/USD modelu ARIMA(1,1,0)**

*Zdroj: vlastní zpracování pomocí sw. STATISTICA na základě dat z [14]*



**Graf 32: PACF reziduí kurzu GBP/USD modelu ARIMA(1,1,0)**

*Zdroj: vlastní zpracování pomocí sw. STATISTICA na základě dat z [14]*

Novým modelem k analýze bude proces ARIMA(0,1,1), kde první diferenci doplníme o část MA(1). K této analýze opět použijeme obdobný postup odhadů parametru modelu a verifikaci jako u předešlých.

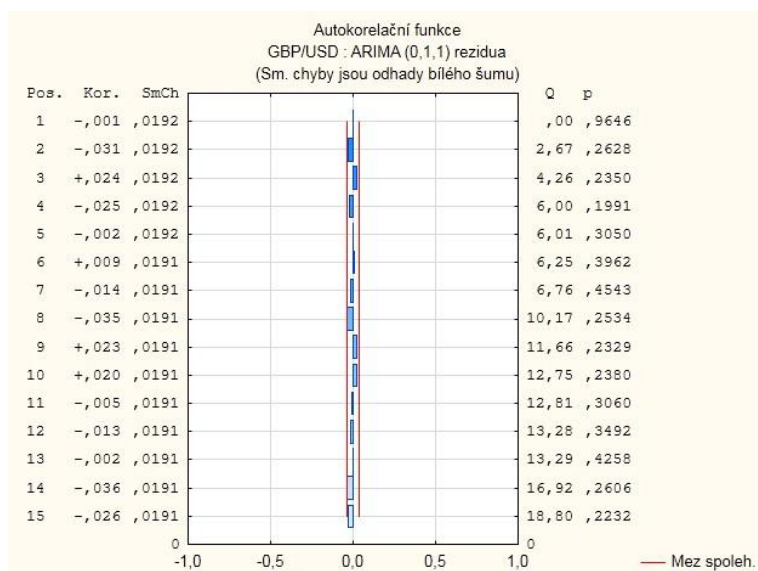
Tabulka 8 zobrazuje hodnoty parametru  $q(1)$ . Nevhodnost tohoto parametru, resp. modelu byla opět prokázána pomocí t-testu. ACF a PACF [grafy 33 a 34] vykazují podobné hodnoty jako u předchozího modelu.



Vstup: GBP/USD (GBP-USD v Vývoj kurzů 03-11 (2))						
Transformace: ln(x),D(1)						
Model:(0,1,1) PČ Rezid. = ,00004						
Paramet.	Param.	Asympt. SmCh	Asympt. t( 2718)	p	Dolní 95% spol	Horní 95% spol
q(1)	-0,026506	0,019840	-1,33598	0,181669	-0,065410	0,012398

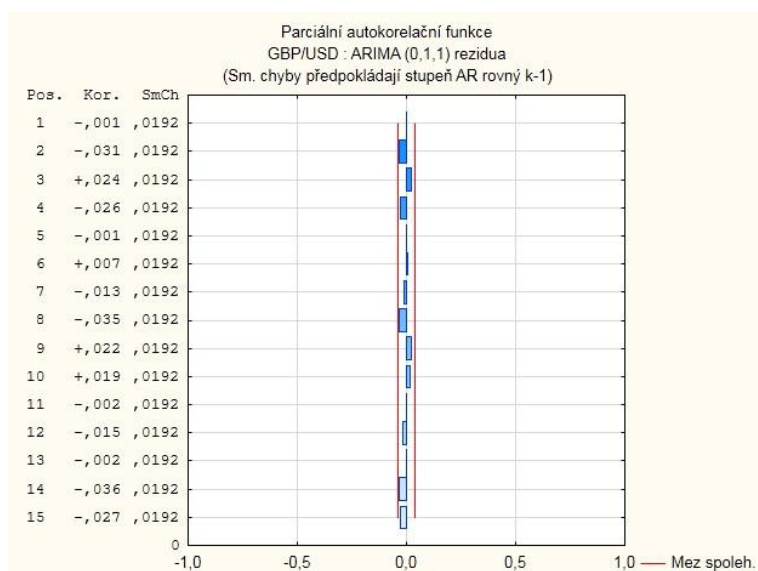
**Tabulka 8: Odhad parametru modelu ARIMA(0,1,1)**

*Zdroj: vlastní zpracování pomocí sw. STATISTICA na základě dat z [14]*



**Graf 33: ACF reziduí kurzu GBP/USD modelu ARIMA(0,1,1)**

*Zdroj: vlastní zpracování pomocí sw. STATISTICA na základě dat z [14]*



**Graf 34: PACF reziduí kurzu GBP/USD modelu ARIMA(0,1,1)**

*Zdroj: vlastní zpracování pomocí sw. STATISTICA na základě dat z [14]*

Nyní se pokusíme řadu prvních diferencí modelovat sloučením obou procesů AR(1) a MA(1), a vytvoříme tak smíšený model ARIMA(1,1,1). Pro analýzu použijeme stejných postupů a verifikací.

Odhadnuté parametry zobrazuje následující tabulka 9.

Vstup: GBP/USD (GBP-USD v Vývoj kurzů 03-11 (2))						
Transformace: ln(x),D(1)						
Model:(1,1,1) PČ Rezid. = ,00004						
Paramet.	Param.	Asympt. SmCh	Asympt. t( 2717)	p	Dolní 95% spol	Horní 95% spol
p(1)	-0,734493	0,140130	-5,24150	0,000000	-1,00927	-0,459720
q(1)	-0,768386	0,131959	-5,82292	0,000000	-1,02714	-0,509636

**Tabulka 9: Odhad parametrů modelu ARIMA(1,1,1)**

*Zdroj: vlastní zpracování pomocí sw. STATISTICA na základě dat z [14]*

Opět provedeme otestování parametrů  $p(1)$  a  $q(1)$ . U obou parametrů je hodnota  $p$  rovna nule. Po porovnání, hodnot  $p$  a  $\alpha = 0,05$  u obou parametrů, jsme dospěli k závěru, že střední hodnota a parametry jsou různé od nuly.

Dále porovnáme hodnoty t-testu každého parametru s hodnotou  $t = 1,96$ . [16] Obě tyto hodnoty t-testu jsou vyšší než daná hodnota  $t$ , proto výsledkem verifikace je zjištění, že oba parametry mají statistickou významnost.

Vhodnost modelu vykazují i reziduální graf 35 funkce ACF a graf 36 funkce PACF kurzu GBP/USD. Hodnoty těchto reziduálních funkcí leží uvnitř nebo na hranici intervalu spolehlivosti.

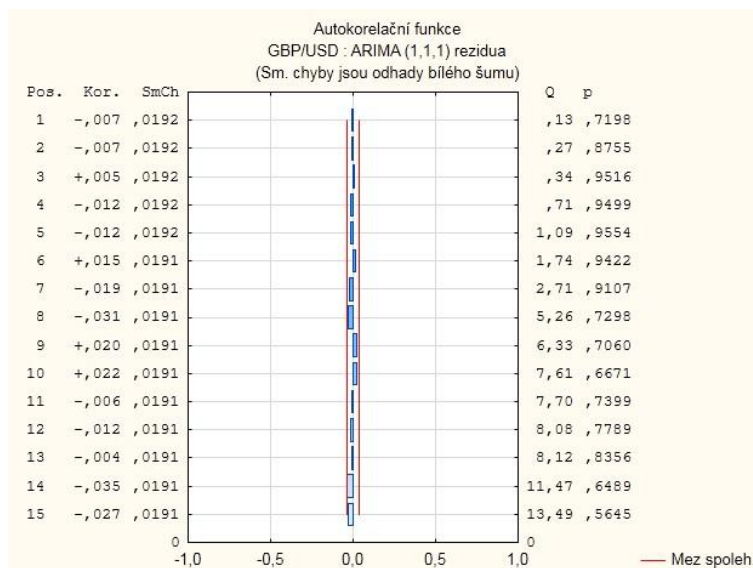
Významnost těchto funkcí zkontrolujeme pomocí Portmanteaůva testu. Nejprve stanovíme stupně volnosti (5.9), známe hodnotu  $n = 2720$  a určíme si  $K$  ( $K = 52$ ), potom tedy:

$$K-p-q= 52-1-1 = 50. \tag{5.9}$$

Kritická hodnota je

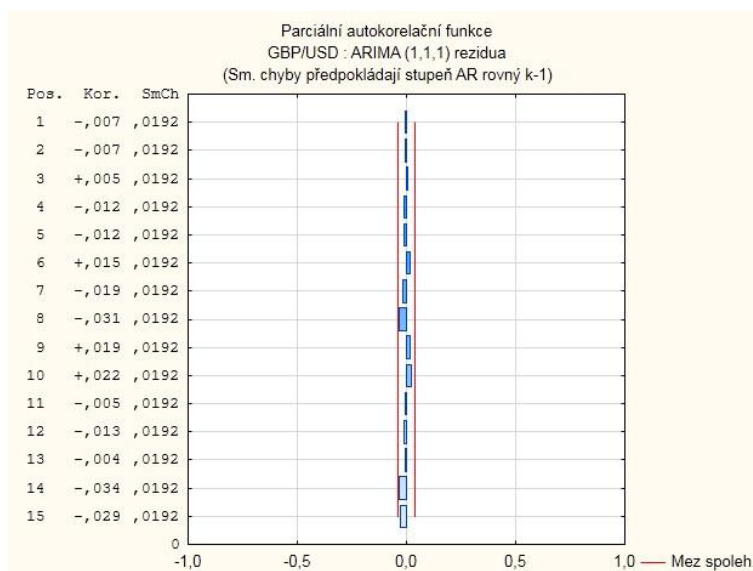
$$\chi_{50}^2(0,05) = 67,5084. \tag{5.10}$$

Z výpočtu vyplývá, že hodnoty  $Q$  nepřesahují kritickou hodnotu (5.10), tento test opět potvrzuje správnost vybraného modelu.



**Graf 35: ACF reziduí kurzu GBP/USD modelu ARIMA(1,1,1)**

*Zdroj: vlastní zpracování pomocí sw. STATISTICA na základě dat z [14]*



**Graf 36: PACF reziduí kurzu GBP/USD modelu ARIMA(1,1,1)**

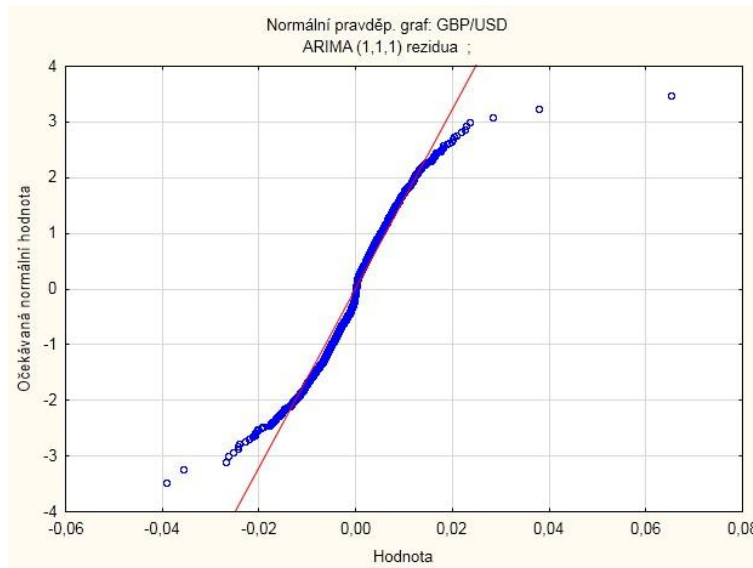
*Zdroj: vlastní zpracování pomocí sw. STATISTICA na základě dat z [14]*

I v tomto případě porovnáním grafu 37 normálního rozdělení reziduí kurzu modelu ARIMA(1,1,1) a grafu 29 normálního rozdělení logaritmu výnosů kurzu vidíme, že se po modelování prvních diferencí časové řady kurzu procesem ARIMA(1,1,1) jednotlivé hodnoty více zarovnaly k hladině významnosti.

Srovnáme-li ACF a PACF reziduí kurzu modelu ARIMA(0,1,0), tedy modelu náhodné procházky [grafy 27 a 28] s funkcemi ACF a PACF reziduí kurzu modelu ARIMA(1,1,1)

[grafy 35 a 36] dojdeme k závěru, že hodnoty funkcí modelu ARIMA(1,1,1) se více oddálily od hranice spolehlivosti směrem dovnitř, a tudíž daleko méně hodnot se této hranice dotýká.

Tyto poznatky vedou k ujištění, že model ARIMA(1,1,1) je vhodnější pro modelování finanční časové řady kurzu GBP/USD než proces náhodné procházky.



**Graf 37: Normální rozdělení kurzu GBP/USD modelu ARIMA(1,1,1)**

*Zdroj: vlastní zpracování pomocí sw. STATISTICA na základě dat z [14]*

### 5.2.3 Prezentace dosažených výsledků

Po předcházejícím zjištění získáme následující rovnici (5.11) vhodného modelu ARIMA(1,1,1). Rovnice modelu ARIMA(1,1,1) po do sazení potřebných parametrů má tvar

$$y_t = -0,734493y_{t-1} + \varepsilon_t - 0,768386\varepsilon_{t-1}. \quad (5.11)$$

### 5.2.4 Predikce finanční časové řady kurzu GBP/USD pomocí modelu

Pro tuto řadu si také provedeme predikci do budoucnosti na jeden rok. Tato předpověď je znázorněna v grafu 38 předpovědí a malý příklad hodnot jednotlivých predikcí po dnech jsou zapsány v tabulce 10.



**Graf 38: Predikce finanční časové řady kurzu GBP/USD**

*Zdroj: vlastní zpracování pomocí sw. STATISTICA na základě dat z [14]*

ČísloPříp.	Předpovědi; Model:(1,1,1) Sezónní posun:12 (GBP-USD v Vývoj kurzů 03-11 (2)) Vstup: GBP/USD Začátek původ. : 1 Konec původ. : 2720		
	Předpově	Dolní 90,0000%	Horní 90,0000%
2721	1,558208	1,542825	1,573746
2722	1,558055	1,535978	1,580450
2723	1,558168	1,531240	1,585569
2724	1,558085	1,526937	1,589868
2725	1,558146	1,523387	1,593698
2726	1,558101	1,520025	1,597132
2727	1,558134	1,517055	1,600325
2728	1,558110	1,514214	1,603278
2729	1,558128	1,511611	1,606076
2730	1,558115	1,509110	1,608710
2731	1,558124	1,506767	1,611232
2732	1,558117	1,504509	1,613635
2733	1,558122	1,502363	1,615951
2734	1,558119	1,500289	1,618177
2735	1,558121	1,498300	1,620331
2736	1,558119	1,496372	1,622415
2737	1,558121	1,494510	1,624439
2738	1,558120	1,492702	1,626404

**Tabulka 10: Predikce finanční časové řady kurzu GBP/USD**

*Zdroj: vlastní zpracování pomocí sw. STATISTICA na základě dat z [14]*

## 6 Závěr

Cílem této diplomové práce bylo modelování vybraných finančních časových řad, konkrétně burzovního indexu S&P 500 a měnového kurzu GBP/USD. Pomocí získaného vhodného modelu bylo možné provést i jejich predikci. Analýza byla provedena pomocí zpracované teorie, podle které jsme aplikovali hlavní postupy a metody na námi vybrané finanční časové řady. Díky těmto postupům a metodám byla ověřena a otestována data a následně byla určena predikce těchto dvou finančních časových řad.

V první řadě jsme vysvětlili základní pojmy model a modelování. Posléze byly definovány teorie řešící finanční časové řady. V kapitole časových řad bylo uvedeno rozdělení těchto časových řad, byly stručně popsány problémy analýzy časových řad, jejich základní přístupy a také předpovědi v časových řadách. Důležitou částí této kapitoly je podrobné definování finančních časových řad a jejich vlastností. Také zde byly objasněny předpoklady finančních časových řad, tato část je zaměřena na normalitu, linearitu, pákový efekt, autokorelaci a heteroskedasticitu. Nedílnou součástí bylo i vysvětlení vlivu mikrostruktury trhu na některé vlastnosti finančních časových řad.

Poslední teoretická část byla věnována Box-Jenkinsově metodologii, jejímu principu, výhodám i nevýhodám. Především zde byly popsány základní pojmy této metodologie, jako jsou stochastický proces, stacionarita, autokorelační funkce, parciální autokorelační funkce, proces bílého šumu a lineární proces. V neposlední řadě jsou zde vysvětleny a popsány lineární modely stacionárních časových řad v jednotlivých řádech, jako například AR, MA i smíšený model ARMA, dále lineární modely nestacionárních časových řad, což je proces náhodné procházky a proces ARIMA a také jsme se zde zmínily o modelech stacionárních a nestacionárních sezonních časových řad. Konečná část této kapitoly se věnuje konstrukci modelu dané metodologie, která je identifikace modelu, odhad parametrů modelu a jeho verifikace.

Základním cílem experimentální části bylo provedení analýzy burzovního indexu S&P 500 a měnového kurzu GBP/USD a jejich predikci. Veškeré výpočty, důležité pro tuto analýzu, byly zpracovány pomocí softwaru STATISTICA. Na základě analýz těchto finančních časových řad došlo k následujícím poznatkům, které jsou uvedeny níže. Postup může být rozdělen do několika na sebe navzájem navazujících kroků.

Nejprve byla provedena první analýza pro index S&P 500. Základním krokem byla identifikace modelu, kde byl zobrazen denní spojnicový graf tohoto indexu, ze kterého byl

patrný průběh indexu za rok 2011, dále byla provedena logaritmická transformace a zobrazeny její ACF a PACF, tento krok nás indikoval ke zjištění, že se nejedná o stacionární řadu. Bylo tedy nutné přistoupit k jejímu diferencování první diferencí, tento krok nám zajistil stacionaritu řady. Pro kontrolu stacionarity jsme si určili ACF a PACF logaritmů výnosů, tyto grafy také ukázaly přítomnost systematické složky. Byly získány i popisné statistiky, které nám učili například šikmost a špičatost, podle které se určuje normální rozdělení. Bylo zjištěno, že logaritmy výnosů této finanční řady pocházejí z rozdělení, které má „tlustší“ konce než normální rozdělení. Toto nám také dokázal normální pravděpodobnostní graf.

Dalším krokem bylo provedení odhadu parametrů modelu s následnou verifikací. Po provedení několika analýz bylo zjištěno, že proces ARIMA(1,1,1) tuto řadu modeluje nejlépe, což bylo dokázáno pomocí mnoha kontrol, například pomocí t-testů pro zjištění správnosti parametrů nebo pomocí Portmanteauva testu, který ověřil nekorelovanost reziduí ACF. Pomocí získaného vhodného modelu byla sestavena rovnice modelu, kde byly dosazeny odhadnuté parametry. Pro ukázkou byla znázorněna i predikce 10 budoucích hodnot a jejich graf.

Naprosto stejně bylo postupováno i u modelování finanční časové řady kurzu GBP/USD. Opět se jednalo o řadu nestacionární, kde bylo nutné ji stacionarizovat první diferencí, což bylo dokázáno pomocí ACF a PACF logaritmů výnosů této řady. Pomocí popisných statistik bylo zjištěno, že výsledkem je nepodmíněné rozdělení, které je v porovnání s normálním rozdělením špičatější s „tlustšími“ konci. ACF a PACF logaritmů výnosů vykazovaly, že autokorelační koeficienty leží uvnitř nebo na hranici tolerančních mezí. Nabízela se tedy možnost závěru, že by se mohlo jednat o model tvaru náhodné procházky. V našem případě přesto byla analyzována časová řada i modely ARIMA, kde by nebyly ACF či PACF, resp. autokorelační koeficienty tak silně položeny na hranici tolerančních mezí. Prozkoumáním různých procesů, byl nalezen vhodný model ARIMA(1,1,1), který opravdu ukazoval, že jeho autokorelační koeficienty ACF a PACF reziduí se více oddálily směrem dovnitř od hranic tolerančních mezí než tomu tak bylo u modelu náhodné procházky. Zjištěné skutečnosti byly opět verifikovány jako v předešlé analýze indexu, a tedy byla dokázána vhodnost tohoto modelu. Na závěr byla sestavena rovnice modelu a ukázána předpověď na rok dopředu.

V této práci tedy bylo ukázáno modelování jak burzovních indexů, tak i měnových kurzů, a jak finanční časové řady o malém počtu denních pozorování (1 rok) u indexu S&P 500, tak i časové řady o vysokém počtu denních pozorování (9 let) u kurzu GBP/USD.

## Použitá literatura

- [1] Analýza časových řad. [online]. [cit. 2012-02-29]. Dostupné z: <http://iastat.vse.cz/casovky/casovky2.htm>.
- [2] ARLT, Josef. Moderní metody modelování ekonomických časových řad. 1.vyd. Praha: Grada Publishing, 1999, 307 s. ISBN 80-716-9539-4.
- [3] ARLT, Josef, ARLTOVÁ, Markéta. Ekonomické časové řady: Vlastnosti, metody modelování, příklady a aplikace. 1. vyd. Praha : [s.n.], 2007. 288 s. ISBN 978-80-247-1319-9.
- [4] ARLT, Josef, ARLTOVÁ, Markéta. Finanční časové řady: [vlastnosti, metody modelování, příklady a aplikace]. 1. vyd. Praha: Grada, 2003, 220 s. ISBN 80-247-0330-0.
- [5] ARTL, Josef; ARTLOVÁ, Markéta; RUBLÍKOVÁ, Eva. Analýza ekonomických časových řad s příklady. 2. vyd. Skripta VŠE Praha, 148 str., 2004. ISBN 80-245-0777-3.
- [6] BILL, J. A., & HARDGRAVE, R. L. J. *Comparative politics: The quest for theory*. Columbus, OH: Charles E. Merrill, 1973.
- [7] CIPRA, T. *Analýza časových řad s aplikacemi v ekonomii*. Praha: SNTL/ALFA, 1986.
- [8] ČAPEK, Jan. *Modelování ekonomických a sociálních procesů: pro kombinovanou formu studia*. Vyd. 1. Pardubice: Univerzita Pardubice, 2006, 103 s. ISBN 80-719-4838-1.
- [9] Časová řada - Wikipedie. [online]. [cit. 2012-02-28]. Dostupné z: [http://cs.wikipedia.org/wiki/%C4%8Casov%C3%A1\\_%C5%99ada](http://cs.wikipedia.org/wiki/%C4%8Casov%C3%A1_%C5%99ada).
- [10] Časové řady. [online]. [cit. 2012-02-28]. Dostupné z: <http://homen.vsb.cz/~oti73/cdpast1/KAP10/KAP10.HTM>.
- [11] DEUTSCH, K. W. On communication models in the social sciences. *Public Opinion Quarterly*. 1952, 16, 356-380.
- [12] ERFURHT, H., JUST G. *Modelování a optimalizace chemických procesů*. Praha SNTL, 1979.
- [13] FÁBRY, Jan. *Matematické modelování*. Vyd. 1. Praha: Oeconomica, 2009, 146 s. ISBN 978-80-245-1266-2.



- [14] Historie vývoje kurzů | Kurzovní lístek měn - euro, koruna, libra, dolar | Citfin - Finanční trhy a.s. [online]. [cit. 2012-04-13]. Dostupné z: <http://www.citfin.cz/cz/kurzovni-zpravodajstvi/historie-vyvoje-kurzu.html>.
- [15] KŘIVÝ, Ivan. *Analýza časových řad*. Ostrava, 2006. Dostupné z: [http://informatika-osu.czechian.net/files/is/ancas/ANCAS\\_DiV.pdf](http://informatika-osu.czechian.net/files/is/ancas/ANCAS_DiV.pdf).
- [16] LINDA, Bohdan, KUBANOVÁ, Jana. *Statistické tabulky a vzorce*. Vyd. 3., dopl. Pardubice: Univerzita Pardubice, 2004, 44 s. ISBN 80-719-4657-5.
- [17] MANDELBROT, B. *The Variation of Certain Speculative Prices, Journal of Business* 36, 1963, 394-419.
- [18] RÉNYI, A. *Teorie pravděpodobnosti*. Academia, 1972.
- [19] Standard & Poors 500 - Burzy a indexy - Akcie.miras.cz. [online]. 2000-2012 [cit. 2012-04-11]. Dostupné z: <http://www.miras.cz/akcie/indexy-standard-poors.php>.
- [20] StatSoft: Produkty [online]. [cit. 2012-03-11]. Dostupné z: <http://www.statsoft.cz/produkty/>.
- [21] StatSoft: Profil společnosti [online]. [cit. 2012-03-11]. Dostupné z: <http://www.statsoft.cz/o-firme/profil-spolecnosti/>.
- [22] VESELÝ, Arnošt. *Metody a metodologie vymezení problému: (Strukturace, definice, modelování a formulace problému v policy analysis)*. 1. vyd. Praha, 2005. ISSN 1801-1640.

## Seznam grafů

Graf 1: Průběh finanční časové řady indexu S&P 500 .....	21
Graf 2: Logaritmy výnosů řady indexu S&P 500 .....	22
Graf 3: Logaritmy výnosů časové řady kurzu GBP/USD.....	24
Graf 4: ACF logaritmované časové řady .....	29
Graf 5: Spojnicový graf znázorňující index S&P 500 .....	44
Graf 6: Logaritmovaná časová řada indexu S&P 500 .....	45
Graf 7: ACF logaritmované časové řady indexu S&P 500.....	45
Graf 8:PACF logaritmované časové řady indexu S&P 500 .....	46
Graf 9: Logaritmy výnosů řady indexu S&P 500 .....	46
Graf 10: ACF první diference indexu S&P 500 .....	47
Graf 11:PACF první diference indexu S&P 500 .....	47
Graf 12: Normální rozdělení logaritmů výnosů indexu S&P 500 .....	49
Graf 13: Empirická distribuční funkce logaritmů výnosů indexu S&P 500.....	49
Graf 14: ACF reziduí indexu S&P 500 modelu ARIMA(1,1,0).....	50
Graf 15:PACF reziduí indexu S&P 500 modelu ARIMA(1,1,0) .....	51
Graf 16: ACF reziduí indexu S&P 500 modelu ARIMA(0,1,1).....	52
Graf 17: PACF reziduí indexu S&P 500 modelu ARIMA(0,1,1) .....	52
Graf 18: ACF reziduí indexu S&P 500 modelu ARIMA(1,1,1).....	53
Graf 19: PACF reziduí indexu S&P 500 modelu ARIMA(1,1,1) .....	54
Graf 20: Normální rozdělení reziduí indexu S&P 500 modelu ARIMA(1,1,1) .....	55
Graf 21: Predikce finanční časové řady indexu S&P 500 .....	56
Graf 22: Spojnicový graf znázorňující kurz GBP/USD.....	57
Graf 23: Logaritmovaná časová řada kurzu GBP/USD.....	58
Graf 24: ACF logaritmované časové řady kurzu GBP/USD .....	58
Graf 25:PACF logaritmované časové řady kurzu GBP/USD.....	59
Graf 26: Logaritmy výnosů řady kurzu GBP/USD .....	59
Graf 27: ACF první diference kurzu GBP/USD .....	60
Graf 28:PACF první diference kurzu GBP/USD.....	60
Graf 29: Normální rozdělení logaritmů výnosů kurzu GBP/USD.....	61
Graf 30: Empirická distribuční funkce logaritmů výnosů kurzu GBP/USD .....	62
Graf 31: ACF reziduí kurzu GBP/USD modelu ARIMA(1,1,0) .....	63

Graf 32: PACF reziduí kurzu GBP/USD modelu ARIMA(1,1,0).....	63
Graf 33: ACF reziduí kurzu GBP/USD modelu ARIMA(0,1,1).....	64
Graf 34: PACF reziduí kurzu GBP/USD modelu ARIMA(0,1,1).....	64
Graf 35: ACF reziduí kurzu GBP/USD modelu ARIMA(1,1,1).....	66
Graf 36: PACF reziduí kurzu GBP/USD modelu ARIMA(1,1,1).....	66
Graf 37: Normální rozdělení kurzu GBP/USD modelu ARIMA(1,1,1).....	67
Graf 38: Predikce finanční časové řady kurzu GBP/USD.....	68

## Seznam obrázků

Obrázek 1: ACF a PACF procesu MA(2).....	34
Obrázek 2: ACF procesu AR(1).....	36
Obrázek 3: PACF procesu AR(1).....	36

## Seznam tabulek

Tabulka 1: Popisné statistiky logaritmů výnosů indexu S&P 500.....	48
Tabulka 2: Odhad parametru modelu ARIMA(1,1,0).....	50
Tabulka 3: Odhad parametru modelu ARIMA(0,1,1).....	51
Tabulka 4: Odhad parametrů modelu ARIMA(1,1,1).....	53
Tabulka 5: Predikce finanční časové řady indexu S&P 500.....	56
Tabulka 6: Popisné statistiky logaritmů výnosů kurzu GBP/USD.....	61
Tabulka 7: Odhad parametru modelu ARIMA(1,1,0).....	62
Tabulka 8: Odhad parametru modelu ARIMA(0,1,1).....	64
Tabulka 9: Odhad parametrů modelu ARIMA(1,1,1).....	65
Tabulka 10: Predikce finanční časové řady kurzu GBP/USD.....	68

## Seznam zkratek

ACF	Autokorelační funkce
AMEX	American Stock Exchange
AR	Autoregresní modely
ARCH	Autoregresní podmíněná heteroskedasticita
ARIMA	Autoregresní integrovaný proces klouzavý součtů
ARMA	Autoregresní proces klouzavých součtů
B-J	Boxova-Jenkinsova
D-F	Dickeyov - Fullero
EGARCH	Exponenciální GARCH
GARCH	Zobecněný ARCH
GBP	Britská libra
KPSS	Kwiatkowski, Phillips, Schmidt a Shin
MA	Proces klouzavých součtů
MAD	Mean Absolute Deviation
MSE	Mean Squared Error
NYSE	New York Stock Exchange
PACF	Parciální autokorelační funkce
S&P	Standard & Poor's
SSE	Sum of Squared Errors
USA	Spojené státy americké
USD	Americký dolar