

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

DIPLOMOVÁ PRÁCE

Modely volatility



Vedoucí diplomové práce:

Mgr. Ondřej Vencálek, Ph.D.

Rok odevzdání: 2012

Vypracovala:

Bc. Martina Vrzalová

AME, 2. ročník

Prohlášení

Prohlašuji, že jsem diplomovou práci zpracovala samostatně pod vedením pana Ondřeje Vencálka a výhradně s použitím uvedené literatury.

V Olomouci dne 14. 12. 2012

.....

Poděkování

Ráda bych poděkovala vedoucímu diplomové práce panu Ondřeji Vencálkovi za jeho trpělivost a čas, který mi věnoval v době konzultací. Poděkování si zaslouží i má rodina, která mě po celý čas mého studia podporovala.

Obsah

1. Úvod	5
2. Analýza časových řad	7
2.1. Časové řady	7
2.2. Dělení časových řad.....	7
2.3. Dekompozice	8
2.4. Box – Jenkinsonova metodologie.....	9
2.4.1. Autokorelační funkce	10
2.4.2. Parciální autokorelační funkce	10
2.4.3. Bílý šum	11
2.4.4. Podmíněné rozdělení	11
2.4.5. Autoregresní proces AR(p)	12
2.4.6. Proces klouzavých součtů řádu MA(q).....	14
2.4.7. Smíšený proces ARMA.....	14
2.4.8. ARIMA(p,d,q).....	15
2.4.9. SARIMA (p,d,q) x (P,D,Q)	16
3. Finanční časové řady	18
3.1. Výnosy z aktiv.....	18
3.2. Leptokurtické rozdělení.....	19
3.3. Podmíněná heteroskedasticita.....	20
3.4. Volatilita	21
3.5. Míry volatility	22
4. Modely volatility.....	24
5. Lineární modely volatility	28
5.1. ARCH (Autoregressive Conditional Heteroscedasticity).....	28
5.1.1. ARCH (1)	31
5.1.2. Simulace modelu ARCH (1).....	32
5.1.3. Vliv změny parametrů	33
5.2. GARCH(p,q) (Generalized ARCH).....	36
5.2.1. GARCH (1,1).....	38
5.2.2. Simulace modelu GARCH (1,1)	39
5.2.3. Vliv změny parametrů	40

5.3 Odhady parametrů v modelech volatility	41
5.3.1 Věrohodnostní funkce	41
6. Výstavba modelů volatility	44
6.1 Výstupy softwaru R	44
6.2 Zajištění stacionarity časové řady	45
6.2.1 Augmented Dickey-Fuller test (ADF)	45
6.2.2. Stacionarizace časové řady	46
6.3 Určení úrovněvého modelu	47
6.4 Ověření kvality úrovněvého modelu	48
6.4.1. Významnost parametrů	48
6.4.2 Testování autokorelace reziduí	50
6.4.3. Normalita reziduí	52
6.4.4 Testování přítomnosti heteroskedasticity v časové řadě	53
6.4.5 Určení modelu volatility a jeho verifikace	55
6.4.6 Předpovědi podmíněného rozptylu	58
7. Aplikace modelů volatility na data nefinančního charakteru	62
8. Závěr	68
Použité zdroje	69

1. Úvod

V poslední době roste význam modelování volatility ve finanční sféře. Bez uvažování volatility, vyjadřující nejistotu na trhu, je téměř nemožné kvalitně predikovat budoucí vývoj cen. Donedávna byla volatilita odhadována pomocí směrodatné odchylky či rozptylu, jejichž hodnoty neodpovídaly skutečnosti, nereagovaly na dynamické změny na finančním trhu. Cílem tedy bylo hledat způsob předpovědí, které budou brát v potaz více informací a věrněji odhadovat budoucí vývoj nejistoty na trhu.

Jedním z řešení jsou právě modely volatility. Počátky modelů volatility se datují do roku 1982, kdy Engle přišel s návrhem modelu ARCH. Následně byl model zobecněn modelem GARCH a dnes se již můžeme dočíst o desítkách a možná i více dalších modifikací modelů.

ARCH modely používají k modelování volatility podmíněného rozptylu. U mnoha časových řad, zejména s vysokou frekvencí měření se totiž často vyskytuje v čase proměnlivý rozptyl, který je v rozporu s předpoklady klasických modelů, kde byl předpokládán v čase konstantní rozptyl. Kombinací modelů podmíněné střední hodnoty a podmíněného rozptylu pak můžeme dosáhnout kvalitnějších předpovědí cen finančních aktiv.

Cílem mé práce je seznámit čtenáře se základními lineárními modely volatility typu ARCH a GARCH, jejich vlastnostmi a ukázat postup při výstavbě těchto modelů za pomoci softwaru R.

V druhé kapitole připomenu základní pojmy z oblasti časových řad a popíšu některé modely Box-Jenkinsonovy metodologie. U autoregresního modelu odvodím základní charakteristiky a dokážu tak předpoklad v čase neměnného podmíněného rozptylu.

Následující kapitola bude mít za úkol popsat charakteristické vlastnosti finančních časových řad, kterými jsou např. leptokurtické rozdělení, výskyt podmíněné heteroskedasticity atd. Budou také zmíněny jednodušší míry volatility a poukázáno na jejich nevhodnost.

Obsahem čtvrté kapitoly je popis principu fungování modelů volatility. Mimo jiné se budu snažit poukázat na nadřazenost podmíněné předpovědi nad nepodmíněnou.

Následovně budou detailně popsány modely ARCH a GARCH. Průběh procesů bude demonstrován na simulaci procesu. Zaměřím se také na zkoumání souvislosti mezi volbou parametrů a výskytem volatility.

Šestá kapitola tvoří souhrn kroků potřebných k výstavbě modelů volatility. Postup bude ukázán na příkladě časové řady směnných kurzů CZK/EUR.

Modely volatility jsou převážně spojeny s analýzou finančních či ekonomických časových řad. V této části se pokusím aplikovat teorii modelů volatility na vysokofrekvenční data z oblasti energetiky.

2. Analýza časových řad

V této kapitole se připomenou základní pojmy používané v oblasti analýzy časových řad. Uvedu princip Box–Jenkinsonovy metodologie a základní modely spolu s jejich vlastnostmi. Pro vypracování přehledu bylo použito zejména zdrojů [2] a [6].

2.1. Časové řady

Náhodným procesem nazveme systém náhodných veličin $\{y_t, t \in T\}$ definovaných na témže pravděpodobnostním prostoru (Ω, A, P) , kde

Ω je množina všech možných výsledků náhodného jevu,

A jevové pole, tzn. systém podmnožin Ω , splňující vlastnosti

$$\begin{aligned} B \in A &\Rightarrow B^c \in A \\ A_n \in A &\Rightarrow \cup A_n = \sum_n P(A_n) \end{aligned}$$

a T je množina časových okamžiků, ve kterých je prováděno měření. Výsledkem realizace náhodného procesu je časová řada.

Časovou řadu tvoří chronologicky uspořádaná data, která získáváme pozorováním určité veličiny v čase.

Časové řady jsou předmětem zkoumání nejrůznějších statistických analýz ekonomických jevů. Cílem analýzy časových řad je porozumět mechanismu, který generuje data a najít vhodný model, na základě něhož budeme moci co nejkvalitněji předpovídat budoucí směr vývoje pozorované veličiny.

2.2 Dělení časových řad

I. Dle frekvence pozorování

a) *Dlouhodobé časové řady* - pozorování je prováděno v ročních, někdy i delších intervalech.

b) *Krátkodobé časové řady* – měření v úsecích kratších než jeden rok, např. měsíc, týden.

c) *Vysokofrekvenční časové řady* – pozorování prováděno týdně, denně, každou hodinu apod. Vykazují specifické vlastnosti, které budou zmíněny v následující kapitole.

II. Dle typu ukazatele

a) Intervalové časové řady – zagregované údaje za jednotku času, např. produkce za měsíc.

b) Okamžikové časové řady – hodnoty této časové řady jsou naměřeny k určitému časovému okamžiku.

2.3. Dekompozice

Jednou z metod používaných k rozboru ekonomických časových řad je klasický dekompoziční přístup, kdy našim cílem je proniknout hlouběji do podstaty časové řady tak, že dokážeme eliminovat jednotlivé složky časové řady. Poté lze jednotlivé složky zkoumat a předpovídat jejich budoucí vývoj zvlášť. Dekompozice spočívá v modelování systematické, nenáhodné složky.

Existují dva způsoby dekompozice časových řad, a to *aditivní* $y = T + S + C + \varepsilon$ a *multiplikativní* $y = TSC\varepsilon$, kde

T je trend časové řady, znázorňuje tendenci dlouhodobého vývoje,

S značí sezónnost dat, tedy pravidelné odchylky od trendu s periodou kratší než jeden rok,

C je označením pro cyklus – dlouhodobé kolísání s proměnlivou periodou a amplitudou a

ε zahrnuje vlivy, které nelze modelem vysvětlit, tedy náhodnou složku.

Typ dekompozice volíme v závislosti na typu časové řady.

K modelování trendu používáme různých regresních funkcí (lineární, kvadratická, exponenciální atd.), u nichž poté odhadujeme parametry. Pokud je funkce lineární v parametrech, využijeme odhadovací metody nejmenších čtverců, která nám zaručí dobré vlastnosti odhadu. Vhodnost zvolené regresní funkce lze ověřit např. graficky, vykreslením skutečných a modelovaných hodnot nebo můžeme použít další míry vhodnosti, mezi které patří index determinace, reziduální součet čtverců, střední čtvercová chyba MSE a další.

2.4. Box – Jenkinsonova metodologie

Na rozdíl od předchozího dekompozičního přístupu, kde byl kladen důraz na zkoumání systematické nenáhodné složky za předpokladu, že pozorování jsou vzájemně nekorelovaná, Box–Jenkinsonův přístup tvoří modely na základě náhodné složky a připouští, že rezidua mohou být vzájemně korelovaná. Důležitým nástrojem tohoto přístupu je tedy korelační analýza spolu s grafickými nástroji, s jejichž pomocí jsme schopni vyšetřit závislost mezi pozorováními prostřednictvím chyb.

Metody Box – Jenkinsonovy metodologie flexibilně reagují na změny v charakteru modelované časové řady, avšak mají i několik nevýhod. Mezi tyto nevýhody patří fakt, že aplikace metod v praxi je často časově i finančně náročná, předpoklady modelů nebývají často splněny a ke konstrukci kvalitního modelu je zapotřebí alespoň 50 pozorování.

Jedním z předpokladů aplikace Box-Jenkinsonovy metodologie je, že pro výstavbu modelů je potřeba stacionarity časové řady. Tady již vzniká problém, jelikož mnohé časové řady dat reálného světa tento předpoklad nespĺňují. Avšak existují nástroje, s jejichž pomocí lze časovou řadu převést na stacionární. Abychom zajistili stacionaritu časové řady, můžeme užít transformace časové řady (např. logaritmická transformace pro kladná data) či převést úroveň časové řady pomocí diferencování.

Stacionaritou řady rozumíme ustálené pravděpodobnostní chování řady v čase. Pro účely modelů stačí uvažovat tzv. *slabou stacionaritu*, která klade požadavek na konstantnost střední hodnoty a rozptylu a invarianci kovarianční struktury vůči časovým posunům.

Stochastický proces $\{y_t, t \in T\}$, který splňuje následující vlastnosti

$$\mu_t = E y_t, \forall t \in T$$

$$\text{cov}(y_s, y_t) = \text{cov}(y_{s+h}, y_{t+h}) \forall t, s \in T$$

nazýváme slabě stacionární.

V případě normálního stochastického procesu implikuje slabá stacionarita striktní stacionaritu, což znamená, že pravděpodobnostní rozdělení náhodného vektoru $(y_{t_1}, \dots, y_{t_k})$ je stejné jako rozdělení náhodného vektoru $(y_{t_1+h}, \dots, y_{t_k+h})$.

Stochastický proces $\{y_t\}$ označujeme za *invertibilní*, jestliže jej můžeme přepsat do autoregresního tvaru, tedy můžeme-li jeho současnou hodnotu vyjádřit pomocí hodnot minulých a současně hodnoty bílého šumu.

K analýze a identifikaci modelů stochastického procesu využíváme autokorelační a parciální autokorelační funkce a jejich identifikačních bodů.

2.4.1. Autokorelační funkce

Pro stacionární časovou řadu definujeme autokovarianční *funkci* γ_k

$$\gamma_k = \text{cov}(y_t, y_{t+k})$$

a autokorelační funkci
$$\zeta_k = \frac{\gamma_k}{\gamma_0},$$

pro $k = \dots, -1, 0, 1, \dots$, kde γ_0 je rozptyl časové řady.

Funkční hodnoty autokorelační funkce lze vykreslit pomocí *korelogramu*, který nám může posloužit ke zjištění, zda časová řada má charakter bílého šumu viz. 2.4.3.

Autokovarianční i autokorelační funkce jsou sudé funkce, proto nám stačí se při jejich zkoumání omezit na body $k \geq 0$.

Identifikačním bodem nazýváme takový bod k_0 , od nějž dále je autokorelační či parciální autokorelační funkce nulová. Podle tvaru autokorelační funkce a jejího identifikačního bodu jsme schopni identifikovat vhodný typ modelu.

2.4.2. Parciální autokorelační funkce

Parciální autokorelační funkce ζ_{kk} je další důležitou charakteristikou časové řady. Představuje parciální korelační koeficient mezi y_t a y_{t+k} při pevných hodnotách $y_{t+1}, \dots, y_{t+k-1}$. Parciální korelační koeficient vyjadřuje tedy korelaci y_{t+k} a y_t očištěnou o vliv veličin mezi nimi. [8]

2.4.3. Bílý šum

Proces bílého šumu je nejjednodušší formou stochastického procesu $\{y_t\}, t \in T$, jehož pravděpodobnostní rozdělení má konstantní střední hodnotu a konstantní rozptyl. Autokorelační a parciální autokorelační funkce splňují tyto podmínky

$$\begin{aligned} \zeta_k &= 1 \text{ pro } k = 0 & \zeta_{kk} &= 1 \text{ pro } k = 0 \\ \zeta_k &= 0 \text{ pro } k \neq 0 & \zeta_{kk} &= 0 \text{ pro } k \neq 0. \end{aligned}$$

Znamená to tedy, že ACF a PACF bílého šumu jsou identicky nulové, jedná se o stacionární proces. Pro $k = 0$ se rovná $\zeta_k = 1$, jelikož $\gamma_k = \text{cov}(y_t, y_{t+k}) = \text{cov}(y_t, y_t) = 0$.

V případě normálního rozdělení, nazýváme proces bílého šumu *gaussovským procesem* bílého šumu.

2.4.4. Podmíněné rozdělení

Podmíněná střední hodnota je střední hodnota náhodné veličiny X_t za podmínky, že náhodná veličina v časech $t-1, t-2, \dots$ nabyla konkrétních hodnot. Narozdíl od nepodmíněné střední hodnoty, která byla vyjádřena číslem, se zde jedná o náhodnou veličinu závislou na volbě podmínky, je tedy funkcí podmínky. [6]

K modelování podmíněné střední hodnoty slouží modely podmíněné střední hodnoty, např. AR, ARMA, ARIMA.

Na základě modelů podmíněné střední hodnoty bývají často tvořeny předpovědi např. cen finančních aktiv. Není však zohledněna rizikovost, jelikož modely předpokládají v čase neměnný rozptyl (homoskedasticitu).

Podmíněný rozptyl je rozptyl náhodné veličiny X_t za podmínky, že náhodná veličina v časech $t-1, t-2, \dots$ nabyla konkrétních hodnot. Opět se jedná o náhodnou veličinu. Podmíněný rozptyl se nazývá funkcí skedastickou, a to *homoskedastickou*, je-li podmíněný rozptyl v čase neměnný a *heteroskedastickou*, je-li podmíněný rozptyl měnlivý

v čase. Na základě předpokladu heteroskedastického podmíněného rozptylu pracují modely volatility.

2.4.5. Autoregresní proces AR(p)

V této kapitole nejprve uvedu nejjednodušší AR(1) proces, na jehož základě budou demonstrovány vlastnosti procesu. Následně bude uvedena i obecná forma procesu AR(p).

Autoregresním procesem 1. řádu AR(1) nazveme proces $\{X_t\}$ takový, že

$$\forall t \in T \quad x_t = \phi_1 x_{t-1} + \varepsilon_t, \quad (1)$$

kde $\varepsilon_t \sim WN(0, \sigma^2)$.

Při splnění podmínky stacionarity $|\phi_1| < 1$ můžeme přepsat do tvaru stacionárního lineárního procesu

$$x_t = \varepsilon_t + \phi_1 \varepsilon_{t-1} + \phi_1^2 \varepsilon_{t-2} + \dots, \quad (2)$$

ke kterému dospějeme následujícím způsobem

$$x_t = \phi_1 x_{t-1} + \varepsilon_t$$

$$x_{t-1} = \phi_1 x_{t-2} + \varepsilon_{t-1}$$

...

Dosadíme-li do vztahu (1)

$$x_t = \varepsilon_t + \phi_1 \varepsilon_{t-1} + \phi_1^2 \varepsilon_{t-2} + \dots + \phi_1^n \varepsilon_{t-n} + \phi_1^{n+1} \varepsilon_{t-(n+1)}. \quad (3)$$

Abychom dostali nekonečnou řadu, potřebujeme vyrušit poslední člen vztahu (3). Aby se poslední člen pro $n \rightarrow \infty$ blížil k 0, musí být splněna podmínka stacionarity a pak dostáváme vztah (2).

Nyní se podívejme na základní charakteristiky procesu AR(1), které můžeme vypočítat ze vztahu (2) s využitím vlastností bílého šumu $\varepsilon_t \sim WN(0, \sigma^2)$

Nepodmíněná střední hodnota je nulová.

$$Ex_t = E(\varepsilon_t + \phi_1 \varepsilon_{t-1} + \phi_1^2 \varepsilon_{t-2} + \dots) = E\varepsilon_t + \phi_1 E\varepsilon_{t-1} + \phi_1^2 E\varepsilon_{t-2} + \dots = 0.$$

Podmíněná střední hodnota je v čase proměnlivá.

$$\begin{aligned} E(x_t | x_{t-1}) &= E(\phi_1 x_{t-1} + \varepsilon_t | x_{t-1}) = E(\phi_1 x_{t-1} | x_{t-1}) + E(\varepsilon_t | x_{t-1}) = \phi_1 E(x_{t-1} | x_{t-1}) + E(\varepsilon_t | x_{t-1}) = \\ &= \phi_1 x_{t-1} + E(\varepsilon_t) = \phi_1 x_{t-1}. \end{aligned}$$

Nepodmíněný rozptyl nezávisí na čase, je konstantní.

$$\begin{aligned} \text{var } x_t &= \text{var}(\varepsilon_t + \phi_1 \varepsilon_{t-1} + \phi_1^2 \varepsilon_{t-2} + \dots) \\ &= \text{var}(\varepsilon_t) + \phi_1^2 \text{var}(\varepsilon_{t-1}) + \dots \\ &= \sigma^2 (1 + \phi_1^2 + \phi_1^4 + \dots) = \sigma^2 \frac{1}{1 - \phi_1^2}. \end{aligned}$$

Podobně vypočteme i podmíněný rozptyl, který je u procesu AR v čase neměnný. Při výpočtu využijeme idey, že známe-li x_{t-1} , můžeme $\phi_1 x_{t-1}$ považovat za konstantu a počítáme podmíněný rozptyl

$$\text{var}(x_t | x_{t-1}) = \text{var}(\phi_1 x_{t-1} + \varepsilon_t | x_{t-1}) = \text{var}(\varepsilon_t | x_{t-1}) = \sigma^2.$$

Výpočtem charakteristik AR(1) procesu jsme poukázali na skutečnost, že proces je podmíněně homoskedastický. Podmíněný rozptyl je konstantní, což právě u finančních časových řad často neplatí.

Obecný autoregresní proces AR(p) má tvar

$$x_t = \phi_1 x_{t-1} + \dots + \phi_p x_{t-p} + \varepsilon_t,$$

kde ε_t je bílý šum a ϕ_1, \dots, ϕ_p jsou parametry. Přepíšeme-li pomocí operátoru zpětného posunutí

$$\varepsilon_t = \phi(B)x_t = (1 - \phi_1 B - \dots - \phi_p B)x_t.$$

Proces AR(p) je stacionární, leží-li všechny kořeny $\phi(B)$ vně jednotkového kruhu a je vždy invertibilní, protože je zadán v požadovaném tvaru ke splnění podmínky invertibility.

Autokorelační funkce procesu má tvar kombinace klesajících geometrických posloupností a sinusoid s geometricky klesající amplitudou bez existence identifikačního bodu. U PACF naopak identifikační bod existuje a je roven řádu procesu.

2.4.6. Proces klouzavých součtů řádu MA(q)

MA(q) model zapisujeme ve tvaru

$$x_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}$$

nebo pomocí operátoru zpětného posunutí $x_t = \theta(B)\varepsilon_t$, kde $\theta(B) = 1 + \sum_{j=1}^q \theta_j B^j$.

MA(q) je vždy stacionární. Proces MA(q) nazveme invertibilním, leží-li kořeny $\theta(B)$ vně jednotkového kruhu.

ACF procesu má identifikační bod v bodě rovnému řádu procesu q. Pro PACF, která je ve tvaru omezené geometrické posloupnosti či sinusoidy s geometricky klesající amplitudou, identifikační bod neexistuje.

Proces MA(1) je zadán vztahem $x_t = \varepsilon_t + \theta_1 \varepsilon_{t-1}$.

Podmínkou invertibility, za jejíž platnosti lze proces přepsat do autoregresního tvaru je $|\theta_1| < 1$.

2.4.7. Smíšený proces ARMA(p,q)

Proces ARMA (p,q) je kombinací procesů AR(p) a MA(q). Uvádí se ve tvaru

$$x_t = \varphi_1 x_{t-1} + \dots + \varphi_p x_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}.$$

Proces ARMA(p,q) lze zapsat pomocí operátoru zpětného posunutí

$$\varphi(B)x_t = \theta(B)\varepsilon_t.$$

Podmínka stacionarity je shodná s podmínkou stacionarity u AR(p) a podmínka invertibility je stejná jako u procesu MA(q).

Jeli $q \geq p$, pak prvních $q - p$ členů ACF nezapadá do charakteristického průběhu ACF, který po prvních $q - p$ členech má podobu lineární kombinace exponenciálně klesajících pohybů a sinusoid s geometricky klesající amplitudou. Pro $p > q$ bude mít ACF pro $k > p - q$ tvar podobný tvaru ACF procesu MA(q). PACF procesu ARMA je po prvních $p - q$ hodnotách funkce omezená křivkou, která je lineární kombinací klesajících

geometrických posloupností a sinusoid s geometricky klesající amplitudou. Pro ACF ani PACF procesu ARMA neexistuje identifikační bod. [2]

2.4.8 ARIMA(p,d,q)

Proces ARIMA se řadí mezi integrované procesy, kterými lze modelovat nestacionární časové řady. Nestacionarita časové řady může být následkem v čase měnící se střední hodnoty nebo rozptylu. Jedná se v podstatě o proces ARMA(p,q), který je aplikován na řadu d-tých diferencí původní řady.

Proces ARIMA(p,d,q), zapíšeme-li jej pomocí operátoru zpětného posunutí je ve tvaru

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

kde $w_t = \Delta^d x_t$, $\Delta^d = (1-B)^d$ lze takto zapsat pomocí operátoru zpětného posunutí a označuje d-té diference, x_t původní časovou řadu.

Při výstavbě modelů ARIMA se nejprve převede pomocí diferencí nestacionární řada na stacionární, a teprve poté se identifikuje na časové řadě w_t proces ARMA(p,q).

Důležitou poznámkou je, že integrovaný smíšený proces x_t , při řadu diferencování $d > 0$ již nemusíme centrovat, jeho střední hodnota je nulová.

U ekonomických časových řad se doporučuje aplikovat na časovou řadu tzv. logaritmickou transformaci, která stabilizuje časovou řadu z hlediska rozptylu. S její pomocí dosáhneme rozdělení podobného normálnímu, které je robustní vzhledem k odlehlým bodům. [14] Docílíme toho, že časová řada bude generována šoky, které jsou způsobeny nesystematickými vlivy, je tedy charakteru bílého šumu. Logaritmická transformace je odůvodnělá zejména v případě ekonomických časových řad, kde pozorování bývají kladná čísla. Je-li logaritmicky transformovaná časová řada stacionární, jedná se o integrovaný proces I(0), není-li stacionární, je třeba provést první diference. Je-li z grafu prvních diferencí již vidět stacionarita, jedná se o integrovaný proces I(1).

Určení řadu diferencování probíhá na základě vykreslení časové řady a jejich diferencí, zhodnocení ACF, kdy ACF v případě nestacionarity klesá velmi pomalu. Využit lze také testu jednotkových kořenů, který bude přiblížen v kapitole 6. 2. 1. V praxi se užívá diferencí nižších řádů, prvního či druhého řadu.

2.4.9 SARIMA (p,d,q) x (P,D,Q)

SARIMA je modelem Box – Jenkinsonovy metodologie, který umožňuje modelovat sezónní časové řady.

Pro vysvětlení jeho složení uvažujme dva modely. [2]

- Model $\Phi(B^L)\Delta_L^D x_t = \Theta(B^L)\eta_t$, (4)

kde t odpovídá pozorováním v rámci periody L ,

$\Phi(B^L) = 1 - \Phi_1 B^L - \Phi_2 B^{2L} - \dots - \Phi_P B^{PL}$ je sezónní autoregresní operátor řádu P ,

$\Theta(B^L) = 1 + \Theta_1^L + \Theta_2^{2L} + \dots + \Theta_Q^{QL}$ sezónní operátor klouzavých součtů řádu Q ,

$\Delta_L = 1 - B^L$ sezónní diferenční operátor, kde L sezónní perioda, např. máme-li měsíční data, $L=12$.

Na model (4) můžeme pohlížet jako na model ARIMA v rámci jedné periody L , např. jednoho měsíce.

Jelikož náhodné složky, označené v tomto případě η_t , jsou v jednotlivých periodách korelovány, neboť předpokládáme vztah mezi sezónami, uvažujme druhý model pro η_t také ve formě ARIMA procesu

- $\varphi(B)\Delta^d \eta_t = \theta(B)\varepsilon_t$, (5)

kde ε_t je bílý šum, Δ^d operátor běžného diferencování.

Pak spojíme-li dohromady modely (4) a (5) dostaneme model SARIMA(p,d,q) x (P,D,Q)_L

$$\varphi(B)\Phi(B^L)\Delta^d \Delta_L^D x_t = \theta(B)\Theta(B^L)\varepsilon_t.$$

Při výstavbě SARIMA modelů se postupuje tak, že nejprve sestavíme časovou řadu $w_t = \Delta^d \Delta_L^D x_t$ a určíme řády běžného a sezónního diferencování d, D . Pokud z vykreslené ACF řady w_t vidíme, že v bodech autokorelační funkce rovných $L, 2L, 3L..$ nabývá autokorelační funkce významných hodnot a hodnota ACF v těchto bodech pomalu klesá, je nutné sezónně diferencovat, tedy $D=1$. Víme totiž, že autokorelační funkce v bodech $L, 2L,..$ odpovídá odhadnutým hodnotám ACF modelu (4). Stačí zkontrolovat prvních 4 L bodů. Pokud navíc ACF i mezi body $L, 2L, 3L,..$ klesá pomalu, téměř lineárním tempem, musíme použít i běžné difference, tedy $d=1$. Řády diferencování d, D mívají nejčastěji hodnotu 0 či 1.

Máme-li stanovené hodnoty d , D , snažíme se identifikovat model $\varphi(B)\Phi(B^L)w_t = \theta(B)\Theta(B^L)\varepsilon_t$ dle ACF a PACF.

Většinou se užívá homogenních modelů SARIMA(p,d,0)x(P,D,O) nebo SARIMA(0,d,q)x(0,D,Q). Identifikace modelů (p,d,q)x(P,D,Q) je již složitější záležitostí jak z hlediska odhadu velkého počtu parametrů, tak z hlediska identifikace, kdy určení řádů dosti závisí na zkušenostech statistika. [2]

3. Finanční časové řady

V dnešní době je značná část číselných údajů z finanční sféry udávána ve formě časových řad. Finanční časové řady mohou např. vyjadřovat růst kurzů akcií, kurzů měn a dalších. Většinou se jedná o časové řady, které jsou měřeny s vyšší frekvencí než ostatní časové řady, např. každou hodinu, každý den. Bylo ukázáno, že vysokofrekvenční časové řady vykazují zvláštní vlastnosti, proto je třeba brát je v potaz a rozvíjet nové, lepší způsoby jejich zpracování.

V této kapitole se seznámíme s charakteristickými znaky finančních, ale i obecně krátkodobých časových řad a způsobu jejich zpracování.

3.1. Výnosy z aktiv

Mnoho finančních studií pracuje s časovými řadami koeficientů výnosů aktiv, namísto časových řad cen aktiv. Jedním z důvodů je, že výnos z aktiva představuje pro investora kompletní přehled o investiční příležitosti, a tím druhým, že časová řada výnosů aktiv mívá lepší statistické vlastnosti.

Hrubý jednoduchý výnos nebo-li koeficient růstu je definován takto [7]

$$1 + R_t = \frac{P_t}{P_{t-1}},$$

kde R_t představuje relativní přírůstek ceny aktiva, $R_t = \frac{P_t - P_{t-1}}{P_{t-1}}$,

P_t cenu aktiva v čase t a P_{t-1} cenu aktiva v čase $t-1$. Výnos můžeme po vynásobení 100 interpretovat jako %-ní změnu hodnoty z času $t-1$ do času t . Výnos aktiva by měl být nezáporný, proto se uvažuje log-normální rozdělení. Má-li výnos log-normální rozdělení, pak logaritmus výnosu

$$r_t = \ln(1 + R_t) = \ln P_t - \ln P_{t-1}$$

má rozdělení normální. Můžeme si všimnout, že ve výpočtu r_t je zahrnuta logaritmická transformace a následně první diference, díky níž časová řada výnosů většinou bývá stacionární.

3.2 Leptokurtické rozdělení

Při práci s finančními řadami předpokládáme, že logaritmy výnosů jsou nekorelované, identicky rozdělené náhodné veličiny, s konstantní střední hodnotou a rozptylem, popř. s normálním rozdělením s nulovou střední hodnotou a konstantním rozptylem. Tento předpoklad zajišťuje jisté výhodné statistické vlastnosti, avšak ukazuje se, že často nebývá splněn. Často můžeme u časové řady logaritmů výnosů pozorovat, že střední hodnota není nulová.

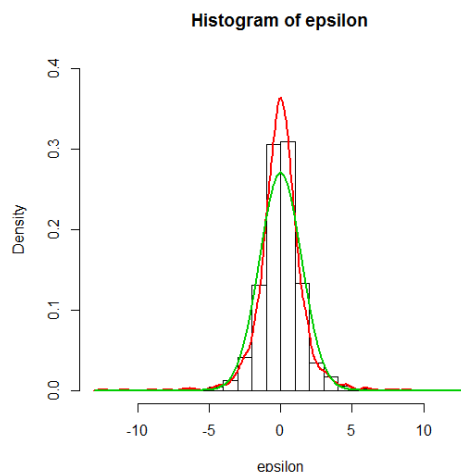
Navíc ani předpoklad nekorelovanosti, který v případě normálního rozdělení splývá s předpokladem nezávislosti, nebývá splněn. Logaritmy výnosů mohou být korelované a pokud se na první pohled zdá, že korelované nejsou, lze často pozorovat významnou korelaci čtverců hodnot, která může být odstraněna použitím vhodných modelů volatility

Na selhání předpokladu normality upozornil již 1963 Mandelbrot a poté řada dalších. Oblast neplatnosti základního předpokladu mnoha teorií je neustále zkoumána, autoři studií se snaží nalézt taková rozdělení, které by dokázalo lépe popsat chování finančních časových řad. Pro zajímavost uvádím některá z navrhovaných rozdělení chyb jako je Studentovo rozdělení, GED rozdělení (Generalized Error Distribution), Laplaceovo rozdělení, L-stabilní rozdělení apod. [15]

Rozdělení logaritmů výnosů je špičatější než pravděpodobnostní rozdělení normálního rozdělení, viz. obr. 1, tzn.

$$K(x) = E\left[\frac{(X - \mu_x)^4}{\sigma_x^4}\right] > 3,$$

kde $K(x)$ je charakteristika špičatosti rozdělení, μ_x střední hodnota rozdělení, σ_x směrodatná odchylka rozdělení. Toto rozdělení má tzv. těžké konce („heavy tails“), které by se daly vysvětlit pravděpodobnějším výskytem extrémně malých i velkých hodnot. Tato vlastnost je tím pozorovatelnější, čím je vyšší frekvence měření. [1] Příčinou tlustých konců rozdělení je v čase měnící se podmíněný rozptyl.



Obr.1 Jádrový odhad hustoty denní finanční řady kurzu CZK/SKK a hustota normálního rozdělení.

Poznámka

Jádrový odhad hustoty je neparametrickou metodou odhadu hustoty, kdy dopředu nepředpokládáme určitý tvar hustoty. Metoda pracuje na základě klouzavého váženého průměru.

V softwaru R se vytvoří jádrový odhad hustoty na základě příkazu `density()`. Při odhadu hustoty záleží také na typu zvoleného jádra a jeho šířce, jež lze nastavit v argumentech `kernel` a `bw`. Nezměníme-li v R typ jádra, R automaticky použije gaussovské jádro s `bw` odhadnutou dle Silverova pravidla na základě vztahu

$$bw = 0,9 \frac{\min(\sigma, h)}{\sqrt[5]{n}},$$

kde σ je odhadnutá směrodatná odchylka z dat,

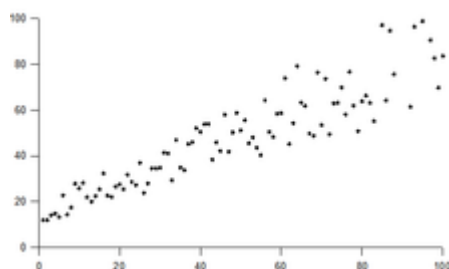
$$h = \frac{x_{0,75} - x_{0,25}}{1,34},$$

kde čítec představuje mezikvartilové rozpětí

a n je rozsah výběru.

3.3 Podmíněná heteroskedasticita

Heteroskedasticita je pojmem vyjadřující vlastnost, že chyby modelu jsou navzájem korelované s měnícím se rozptylem. Tato skutečnost může být způsobena např. závislostí rozptylu na některém z regresorů v modelu. Nejjednodušším způsobem, jak detekovat heteroskedasticitu, je vykreslit si rezidua modelu v závislosti na proměnné. [4] Růst rozptylu reziduí v závislosti na proměnné je ihned patrný, viz. obr. 2.



Obr.2 Heteroskedasticita, dle [16]

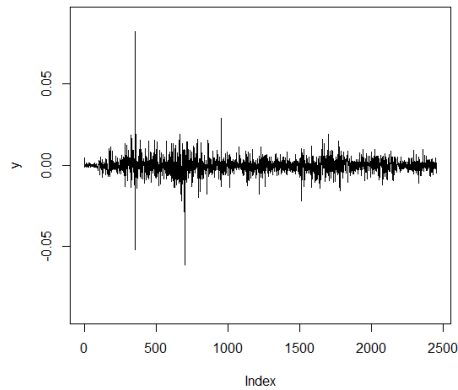
Podmíněná heteroskedasticita znamená proměnlivost podmíněného rozptylu v čase. Bývá často pozorovanou vlastností právě u vysokofrekvenčních časových řad a souvisí s měnící se nejistotou (volatilitou) na trhu v čase. Důsledkem podmíněné heteroskedasticity je nepodmíněné rozdělení, které je charakteristické větší špičatostí a tlustšími chvosty. Její prokázání pomocí testů heroskedasticity nás vede k oprávněné aplikaci modelů volatility nebo-li modelů podmíněné heteroskedasticity.

3.4 Volatilita

Volatilita je termínem pro náhodné autokorelované změny rozptylu v čase. Tato vlastnost bývá často pozorována zejména u vysokofrekvenčních časových řad. Může se stát, že např. u denní časové řady nebude volatilita pozorovatelná a projeví se až u dat s vyšší frekvencí, např. hodinovou. V oblasti finančních časových řad termín vypovídá o úrovni rizika výnosů z finančních aktiv. Období s vyšší volatilitou, spojená s nejistotou na trhu finančních aktiv, jsou střídána obdobími s nižší variabilitou. Volatilita v okamžiku t , závisí na volatilitě v okamžicích předchozích.

Typickou vlastností volatility je její výskyt ve shlucích, tento jev bývá nazýván shlukování volatility (*volatility clustering*). Pod pojmem shlukování volatility rozumíme střídání období s vysokou volatilitou a obdobími méně volatilními, viz obr. 3. Vysoká volatilita v čase t s velkou pravděpodobností způsobí vysokou volatilitu v čase $t+1$.

Skoky ve vývoji volatility mohou být následkem skutečnosti, že na finančních trzích se často neobchoduje každý den. To způsobuje akumulaci volatility, která se projeví skokem až v následujícím obchodním dnu.



Obr.3 Shluky volatility.

Další vlastností volatility je její návrat k dlouhodobému průměru, tzv. mean reversion. Volatilita nediverguje k nekonečnu, je často stacionární a osciluje kolem svého dlouhodobého průměru, tj. nepodmíněného rozptylu (long run variance). [10]

Při vyšetřování volatility se také může objevit tzv. *leverage effect* – pákový efekt. Jedná se o asymetrický efekt, který je následkem skutečnosti, že kladné a záporné šoky se promítají do vývoje volatility nesymetricky. K zachycení této vlastnosti je lepší použít nelineárních modelů volatility, lineární modely pracují totiž s druhými mocninami šoků, tudíž nejsou schopny asymetrický efekt zachytit.

Je známo, že investor by se měl při rozhodování o investicích rozhodovat nejen na základě výnosnosti, ale měl by zohlednit i další faktory jako je likvidita, a především rizikovitost. Proto se stává analýza volatility jednou z důležitých částí analýzy rizika, oceňování aktiv, optimalizace portfolia a řízení rizik.

Pokusy o vysvětlení příčin volatility nedávají prozatím uspokojující výsledky. Ví se však, že volatilita je odpovědí na nečekané nové události např. v ekonomickém vývoji, politickém vývoji a vývoji na trzích. Avšak vývoj těchto událostí je vzhledem k provázanosti díky globalizaci téměř nepředvídatelný.

3.5 Míry volatility

Jednou z nejvíce používaných měr volatility je *standardní směrodatná odchylka*. Bývá označována pojmem *historická volatilita*. Přestože není úplně správné odhadovat

budoucí riziko investice extrapolací historické volatility, zůstává tento jednoduchý ukazatel základním indikátorem rizikovosti. Někteří manažeři fondů jej i dnes nejčastěji zmiňují ve zprávách o svých fondech. [15]

Další možnou mírou volatility je *implikovaná volatilita*. Teorie výpočtu implikované volatility je založena na oceňování aktiv na základě známého ekonometrického modelu, Black-Scholesova vzorce, kde ceny kopírují geometrický Brownův pohyb. Tento přístup bývá kritizován, jelikož využívá specifického modelu a předpokladů, které nemusí ve skutečnosti platit. Modelovaná volatilita se tedy může značně lišit od skutečné. Zkušenosti ukázaly, že implikovaná volatilita výnosů aktiv bývá větší než volatilita získaná aplikací modelů volatility. [1]

Cílem modelů volatility je poskytnout lepší míry volatility, na jejichž základě bude moci být vytvořena i kvalitnější předpověď.

4. Modely volatility

Třetí kapitola se zabývá modely volatility. Jejich základ položil roku 1982 americký ekonom F. Robert Engle. Engle se snažil zjistit, jak mohou být snížena rizika pro investory při nepravidelných fluktuacích na finančních trzích.

Modely volatility se zabývají, narozdíl od ostatních modelů časových řad, modelováním náhodné složky na základě podmíněného rozptylu, tedy druhého momentu. Zkoumáme neautokorelovaná rezidua ε_t , jež u klasických modelů byla považována za efekty nevysvětlitelné modelem a dále se neanalyzovala.

U finančních časových řad obvykle pracujeme s časovou řadou logaritmů výnosů r_t . Po určení úrovněvého modelu podmíněné střední hodnoty označme zápisem ε_t rezidua úrovněvého modelu a považujme je za podmíněně heteroskedastický proces, tedy proces s měnícím se podmíněným rozptylem v čase, ve tvaru

$$\varepsilon_t = v_t \sqrt{h_t},$$

kde h_t je podmíněný rozptyl, $v_t \sim N_{t-1}(0,1)$ gaussovský bílý šum.

Je-li rozdělení v_t za podmínky informace dostupné v čase $t - 1$ normované normální, pak ε_t má podmíněné rozdělení $N_{t-1}(0, h_t)$, tedy normální rozdělení za podmínky dostupné v čase $t - 1$, ale s podmíněným rozptylem h_t měnícím se v čase. [6]

Proměnlivý podmíněný rozptyl nebyl v žádných klasických modelech ani modelech Box-Jenkinsonovy metodologie uvažován, proto se tyto modely nejeví jako vhodné při modelování volatility a je třeba využít modelů, kde není kladen předpoklad na podmíněnou homoskedasticitu.

Existují dvě třídy modelů volatility. I když se jedná v podstatě o nelineární modely, jelikož zachycují nelineární závislosti mezi veličinami stochastického procesu, z hlediska konkrétní funkční formy se rozdělují na modely lineární a nelineární. [6] Do své práce jsem kvůli obsáhlosti zahrnula pouze základní lineární modely volatility.

Nyní si na jednoduchém názorném příkladu ukážeme, proč je při modelování volatility důležité pracovat s podmíněným rozptylem.

Uvažujme, že známe průměrnou denní volatilitu (směrodatnou odchylku) indexu Standard & Poor's 500, s hodnotou 1%. Dále víme, že včerejší předpověď na základě předchozí informace byla 2% a volatilita dnešního dne činí 3%. Očividně se jedná o volatilní období, hodnota volatility se liší den ode dne a je vyšší než dlouhodobá nepodmíněná předpověď (1%). Dlouhodobá nepodmíněná předpověď směřuje k tomu, že obchodníci s akciemi budou překvapeni mírou nejistoty, kolísavostí cen akcií. Proto je důležité pracovat s podmíněnými hodnotami, na jejichž základě bude do předpovědi zahrnuto více informací z předchozích dnů, a to s jistými váhami příslušícím předchozím dnům. [13]

Pozn. Index Standard & Poors 500 je hodnotově vážený index vybraných 500 akcií na americkém trhu. Jeho základem je vážený průměr kurzů daných akcií, kde váha akcie je úměrná své tržní hodnotě. [17]

Jinými slovy, u některých ekonomických časových řad, můžeme pozorovat periody s vyšší a nižší volatilitou, které nemusí ovlivnit nepodmíněný rozptyl v dlouhodobém výhledu, který je jakousi střední hodnotou časové řady volatility. Tento rozptyl může být konstantní. Podmíněný rozptyl kolísá kolem nepodmíněného a může být podstatně vyšší. Proto předpovědi na základě podmíněného rozptylu mohou mít pro finanční analytiky či investory zásadní význam. Podmíněná předpověď bere známou současnost a minulé realizace časové řady s jistými váhami, proto je tato předpověď kvalitnější.

Uvažujme např. ARMA proces daný takto $x_t = \alpha_0 + \alpha_1 x_{t-1} + \varepsilon_t$, kde $\varepsilon_t \sim N(0, \sigma^2)$ a $\alpha_1 < 1$ kvůli zajištění stacionarity procesu. Nyní si ukážeme, že podmíněná předpověď se ukazuje být lepší než nepodmíněná. Budeme přitom sledovat rozptyl předpovědi, který je roven střední čtvercové chybě.

Podmíněná předpověď modelu v čase t+1 je daná střední hodnotou za podmínky znalosti informace v čase t

$$E(x_{t+1} | x_t) = \alpha_0 + \alpha_1 x_t.$$

Chyba předpovědi je ve tvaru $\varepsilon_{t+1} = x_{t+1} - \alpha_0 - \alpha_1 x_t$ a rozptyl podmíněné předpovědi vypadá takto

$$\text{var}(x_{t+1} | x_t) = \text{var}(\varepsilon_{t+1}) = \text{var}(x_{t+1} - \alpha_0 - \alpha_1 x_t) = \sigma^2, \quad (6)$$

Naopak nepodmíněná předpověď je ve tvaru střední hodnoty časové řady x_t . Protože

$$x_{t+1} = x_t \text{ dostáváme po dosazení do modelu } x_t = \alpha_0 + \alpha_1 x_{t-1} + \varepsilon_t$$

$$x_{t+1} = \alpha_0 + \alpha_1 x_t + \varepsilon_{t+1}$$

$$Ex_{t+1} = \alpha_0 + \alpha_1 x_t, \text{ protože } x_{t+1} = x_t, \text{ pak } E(x_t) = E(x_{t+1}) = x_t$$

$$x_t - \alpha_1 x_t = \alpha_0$$

$$x_t(1 - \alpha_1) = \alpha_0$$

$$Ex_{t+1} = \frac{\alpha_0}{1 - \alpha_1},$$

s odpovídajícím rozptylem chyby předpovědi

$$E\left[\left(x_{t+1} - \frac{\alpha_0}{1 - \alpha_1}\right)^2\right] = \frac{\sigma^2}{1 - \alpha_1^2}. \quad (7)$$

Nyní se podívejme na to, jak dospějeme k rovnosti (7).

Chyba předpovědi o jeden krok dopředu

$$f_t(1) = x_{t+1} - Ex_{t+1} = x_{t+1} - x_t = \varepsilon_{t+1},$$

chyba předpovědi o 2 kroky dopředu

$$f_t(2) = x_{t+2} - Ex_{t+2} = \varepsilon_{t+2} + \alpha_1 \varepsilon_{t+1},$$

kde

$$x_{t+1} = \alpha_0 + \alpha_1 x_t + \varepsilon_{t+1}$$

$$x_{t+2} = \alpha_0 + \alpha_1 x_{t+1} + \varepsilon_{t+2} = \alpha_0 + \alpha_1(\alpha_0 + \alpha_1 x_t + \varepsilon_{t+1}) + \varepsilon_{t+2} = \alpha_0 + \alpha_0 \alpha_1 + \alpha_1^2 x_t + \alpha_1 \varepsilon_{t+1} + \varepsilon_{t+2}$$

$$x_{t+2} = Ex_{t+2} + f_t(2)$$

Chyba předpovědi o j-kroků dopředu

$$f_t(j) = \varepsilon_{t+j} + \alpha_1 \varepsilon_{t+j-1} + \alpha_1^2 \varepsilon_{t+j-2} + \dots + \alpha_1^{j-1} \varepsilon_{t+1}.$$

$$\text{var } f_t(j) = \text{var}(\varepsilon_{t+j} + \alpha_1 \varepsilon_{t+j-1} + \alpha_1^2 \varepsilon_{t+j-2} + \dots + \alpha_1^{j-1} \varepsilon_{t+1}) = \sigma^2 + \alpha_1^2 \sigma^2 + \alpha_1^4 \sigma^2 + \dots + \alpha_1^{2(j-1)} \sigma^2$$

$$= \sigma^2 (1 + \alpha_1^2 + \dots + \alpha_1^{2(j-1)})$$

$$\lim_{j \rightarrow \infty} \text{var } f_t(j) = \frac{\sigma^2}{1 - \alpha_1^2}$$

a můžeme vidět, že za platnosti podmínky stacionarity $\alpha_1 < 1$, rozptyl podmíněné předpovědi (6) je menší jako rozptyl nepodmíněné předpovědi (7). [4]

$$\sigma^2 < \frac{\sigma^2}{1 - \alpha_1^2}.$$

5. Lineární modely volatility

Lineární modely volatility jsou charakteristické tím, že podmíněný rozptyl h_t je modelován na základě lineární funkce veličin $\varepsilon_{t-1}^2, \varepsilon_{t-2}^2, \dots, \varepsilon_{t-q}^2$.

Typy modelů se liší formulací podmíněného rozptylu. Mezi základní lineární modely patří ARCH a GARCH, které ve v následující kapitole proberu podrobněji. Dalšími modely s podmíněným rozptylem ve tvaru lineární funkce jsou např. IGARCH, FIGARCH, GARCH-M, viz. např. [4], [6], [15].

5.1 ARCH (Autoregressive Conditional Heteroscedasticity)

Uvažujme podmíněně heteroskedastický proces $\varepsilon_t = v_t \sqrt{h_t}$, $v_t \sim N_{t-1}(0,1)$ je bílý šum s podmíněným normálním rozdělením.

ARCH model je lineárním modelem, kde podmíněný rozptyl h_t je modelován autoregresním procesem řádu q , tedy AR(q).

$$h_t = \omega + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \varepsilon_{t-2}^2 + \dots + \alpha_q \varepsilon_{t-q}^2 = \omega + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2$$

kde $\omega, \alpha_i, i = 1, \dots, q$ jsou odhadovanými parametry v modelu a α_i představují jakési váhy šoků vstupujících do procesu.

Kladnost podmíněného rozptylu je zaručena podmínkami $\omega > 0, \alpha_i \geq 0$ pro $i = 1, 2, \dots, q$.

Modelujeme tedy rezidua úrovnového modelu pomocí podmíněně heteroskedastického procesu ε_t ,

$$\varepsilon_t = v_t \sqrt{\omega + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2},$$

kde v_t je bílý šum.

Model ARCH(q) lze vyjádřit v autoregresním tvaru procesu $\{\varepsilon_t^2\}$ tímto způsobem

$$\varepsilon_t^2 = \omega + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \varepsilon_{t-2}^2 + \dots + \alpha_q \varepsilon_{t-q}^2 + u_t, \quad (8)$$

kde $u_t = \varepsilon_t^2 - h_t = v_t^2 h_t - h_t = h_t(v_t^2 - 1)$.

ARCH(q) je ve tvaru AR(q) modelu čtverců reziduí procesu. Této skutečnosti se využívá při určování řádu q modelu ARCH(q). Identifikace modelu probíhá na základě PACF procesu $\{\varepsilon_t^2\}$, která se chová stejně jako PACF AR(q) procesu.

Model můžeme přepsat pomocí operátoru zpoždění B

$$\begin{aligned}\varepsilon_t^2 - \alpha_1 \varepsilon_{t-1}^2 - \dots - \alpha_q \varepsilon_{t-q}^2 &= \omega + u_t \\ \varepsilon_t^2 (1 - \alpha_1 B - \alpha_2 B^2 - \dots - \alpha_q B^q) &= \omega + u_t.\end{aligned}$$

Proces ARCH (q) je stacionární v kovariancích, jestliže

kořeny $(1 - \alpha_1 B - \alpha_2 B^2 - \dots - \alpha_q B^q) = 0$ leží vně jednotkového kruhu. [6]

Nyní si vypočteme základní podmíněné a nepodmíněné charakteristiky posloupnosti ε_t .

Střední hodnota procesu ε_t se vypočte

$$E\varepsilon_t = E\left[v_t \sqrt{\omega + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2}\right] = Ev_t E\left[\sqrt{\omega + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2}\right] = 0,$$

kde střední hodnotu součinu jsme zapsali jako součin středních hodnot, jelikož víme, že v_t je nezávislé na ε_t . Dále víme, že $v_t \sim WN(0,1)$ a navíc pokud $Ev_t v_{t-i} = 0$ pak také $E\varepsilon_t \varepsilon_{t-i} = 0, i \neq j$. [4]

Nepodmíněný rozptyl je konstantní v čase a vypočteme ho následovně

$$\begin{aligned}\text{var } \varepsilon_t &= E(\varepsilon_t^2) - (E(\varepsilon_t))^2 = E(\varepsilon_t^2) = E\left[v_t^2 \left(\sqrt{\omega + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2}\right)^2\right] = Ev_t^2 E\left(\omega + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2\right) \\ &= E\left(\omega + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2\right) = \omega + E\left(\sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2\right),\end{aligned}$$

protože jsme už vypočetli $E\varepsilon_t = 0$. Víme $Ev_t^2 = 1$, což plyne z rovnosti pro výpočet rozptylu $\text{var } v_t = Ev_t^2 + (Ev_t)^2$, kde $Ev_t = 0$. Tedy $\text{var } v_t = Ev_t^2 = 1$, z vlastností bílého šumu.

Dále označme $\text{var}(\varepsilon_t) = E(\varepsilon_t^2)$

$$E\varepsilon_t^2 - E\left(\sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2\right) = \omega$$

$$E\varepsilon_t^2 - \sum_{i=1}^q \alpha_i E(\varepsilon_{t-i}^2) = \omega.$$

Pokud je $\{\varepsilon_t\}$ stacionární, $\text{var } \varepsilon_t = \text{var } \varepsilon_{t-i} = E(\varepsilon_{t-i}^2)$,
proto můžeme vytknout

$$E(\varepsilon_t^2)\left(1 - \sum_{i=1}^q \alpha_i\right) = \omega$$

a po vydělení $\left[1 - (\alpha_1 + \dots + \alpha_q)\right]$ dostáváme

$$E(\varepsilon_t^2) = \frac{\omega}{1 - \alpha_1 - \alpha_2 - \dots - \alpha_q} = \text{var } \varepsilon_t.$$

Nyní se podívejme na tvar podmíněných charakteristik.

Podmíněná střední hodnota

$$\begin{aligned} E(\varepsilon_t | \varepsilon_{t-1}, \varepsilon_{t-2}, \dots) &= E\left[v_t \sqrt{\omega + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2} \mid \varepsilon_{t-1}, \varepsilon_{t-2}, \dots\right] = \\ &= Ev_t E\left[\sqrt{\omega + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2} \mid \varepsilon_{t-1}, \varepsilon_{t-2}, \dots\right] = 0, \end{aligned}$$

jelikož $Ev_t = 0$.

Dopad ARCH(q) modelu aplikovaného na proces ε_t můžeme vidět až při výpočtu podmíněného rozptylu.

$$\text{var}(\varepsilon_t | \varepsilon_{t-1}, \varepsilon_{t-2}, \dots) = E(\varepsilon_t^2 | \varepsilon_{t-1}, \varepsilon_{t-2}, \dots) - (E(\varepsilon_t | \varepsilon_{t-1}, \varepsilon_{t-2}, \dots))^2 = E(\varepsilon_t^2 | \varepsilon_{t-1}, \varepsilon_{t-2}, \dots),$$

protože jsme již ukázali, že podmíněná střední hodnota $E(\varepsilon_t | \varepsilon_{t-1}, \varepsilon_{t-2}, \dots) = 0$.

Dále

$$\begin{aligned} E(\varepsilon_t^2 | \varepsilon_{t-1}, \varepsilon_{t-2}, \dots) &= E\left[v_t^2 \left(\sqrt{\omega + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2}\right)^2\right] = Ev_t^2 E\left(\omega + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 \mid \varepsilon_{t-1}, \varepsilon_{t-2}, \dots\right) \\ &= E\left(\omega + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 \mid \varepsilon_{t-1}, \varepsilon_{t-2}, \dots\right) = \omega + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 = h_t, \end{aligned}$$

důsledkem toho, že pokud $E v_t v_{t-i} = 0, i = 1, \dots, q$ pak také $E \varepsilon_t \varepsilon_{t-i} = 0$. Podmíněný rozptyl ARCH(q) procesu závisí na všech q minulých hodnotách procesu ε_t , je tedy měnlivý v čase.

Model ARCH umožňuje zachytit vyskytující se shluky volatilit v časové řadě, ale také typický znak finančních časových řad, a to špičatější rozdělení ε_t s tlustými konci ve srovnání s normálním rozdělením.

Je-li potřeba volit ARCH model s vysokým počtem parametrů, přikláníme se raději k použití zobecněného GARCH modelu.

5.1.1. ARCH (1)

Nejjednodušší model ARCH (1) je typický tvarem podmíněného rozptylu

$$h_t = \omega + \alpha_1 \varepsilon_{t-1}^2. \quad (9)$$

Kvůli kladnosti podmíněného rozptylu musí splněny podmínky $\omega > 0$ a $\alpha_1 \geq 0$. Pro případ $\alpha_1 = 0$ by se jednalo o proces podmíněně homoskedastický, tzn. podmíněný rozptyl v čase by byl konstantní.

Model lze přepsat do autoregresního tvaru a to tak, že upravíme (9) přičtením ε_t^2 k oběma stranám a následným odečtením h_t .

$$\begin{aligned} \varepsilon_t^2 + h_t &= \omega + \alpha_1 \varepsilon_{t-1}^2 + \varepsilon_t^2 \\ \varepsilon_t^2 &= \omega + \alpha_1 \varepsilon_{t-1}^2 + \varepsilon_t^2 - h_t \end{aligned} \quad (10)$$

Dosadíme do (10) za $\varepsilon_t^2 = v_t^2 h_t$

$$\varepsilon_t^2 = \omega + \alpha_1 \varepsilon_{t-1}^2 + v_t^2 h_t - h_t = \omega + \alpha_1 \varepsilon_{t-1}^2 + h_t (v_t^2 - 1)$$

a dostaneme proces v autoregresním tvaru

$$\varepsilon_t^2 = \omega + \alpha_1 \varepsilon_{t-1}^2 + u_t, \text{ kde } u_t = \varepsilon_t^2 - h_t.$$

Víme, že $\varepsilon_t = v_t \sqrt{h_t}$, pak $E(\varepsilon_t^2) = E(h_t)$, protože

$$\varepsilon_t^2 = v_t^2 h_t$$

$$E(\varepsilon_t^2) = E(v_t^2) E(h_t) \text{ a } E(v_t^2) = \text{var}(v_t) = 1.$$

Podmíněná i nepodmíněná hodnota procesu $\{u_t\}$ jsou rovny nule. [6]

$$E(u_t) = E(\varepsilon_t^2 - h_t) = E(\varepsilon_t^2) - E(h_t) = E\varepsilon_t^2 - E\varepsilon_t^2 = 0.$$

$$E(u_t | u_{t-1}, u_{t-2}, \dots) = E(\varepsilon_t^2 | \varepsilon_{t-1}, \dots) - E(\varepsilon_t^2 | \varepsilon_{t-1}, \dots) = 0$$

Model nazveme stacionárním v kovariancích, jestliže pro $\alpha_1 < 1$ je nepodmíněná střední hodnota procesu $\{\varepsilon_t^2\}$, tj. nepodmíněný rozptyl procesu $\{\varepsilon_t\}$ ve tvaru $\text{var}(\varepsilon_t) = \frac{\omega}{1 - \alpha_1}$.

Špičatost rozdělení procesu $\{\varepsilon_t\}$ je u ARCH (1) modelu ve tvaru

$$KU_\varepsilon = \frac{E(\varepsilon_t^4)}{E(\varepsilon_t^2)^2} = \frac{3(1 - \alpha_1^2)}{1 - 3\alpha_1^2}. \quad (11)$$

Pokud by rezidua ε_t byla normálně rozdělená, výraz pro špičatost by nabýval hodnoty 3. Ukazuje se však, že špičatost bývá větší a platí, že pro $3\alpha_1^2 < 1$ je ve tvaru (11).

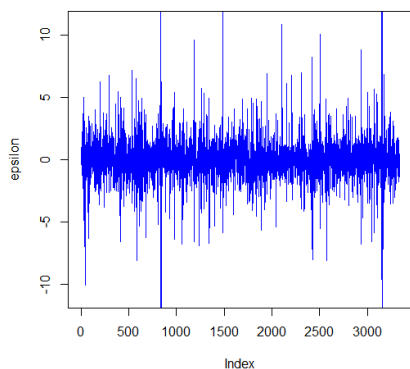
Pokud $\alpha_1 \notin \left(-\sqrt{\frac{1}{3}}, \sqrt{\frac{1}{3}}\right)$, blíží se k nekonečnu. [6]

5.1.2. Simulace modelu ARCH (1)

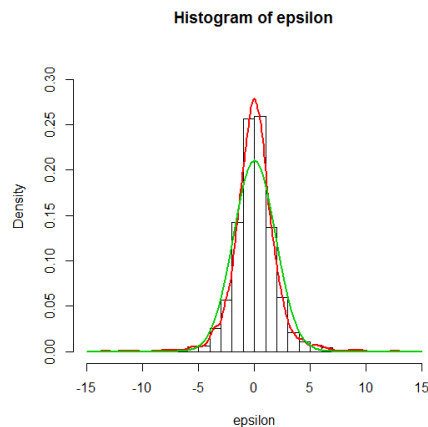
Nyní nasimulujeme ARCH proces $\varepsilon_t = v_t \sqrt{1 + 0,8\varepsilon_{t-1}}$ v programu R pomocí vygenerování 5000 hodnot bílého šumu z normovaného normálního rozdělení $v_t \sim N(0,1)$. Soubor s R-kódem k simulaci modelu ARCH(1) je uložen na přiloženém CD pod názvem simulace_ARCH(1).R.

Simulaci lze v také provést automaticky, pomocí funkce garchSim(), z balíčku fGarch.

Po vygenerování hodnot bílého šumu byl nadefinován tvar podmíněného rozptylu a hodnota podmíněného rozptylu v čase t=1 byla položena rovna jedné. Kvůli této volbě byla pak ze simulovaného procesu odstraněna první polovina hodnot, aby hodnoty procesu nebyly ovlivněny počáteční podmínkou.



Obr. 4 Proces ARCH(1) s parametrem $\alpha_1=0,8$ a absolutním členem $\omega=1$.

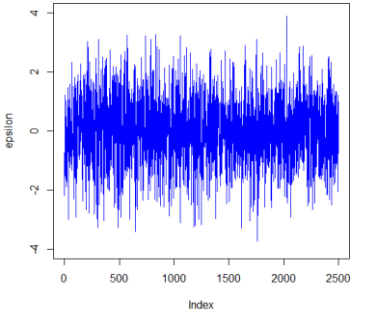
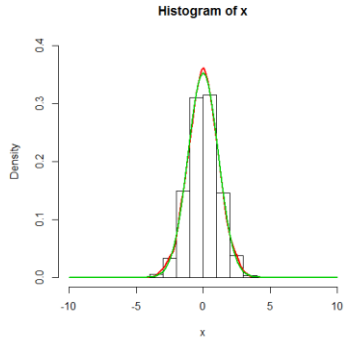
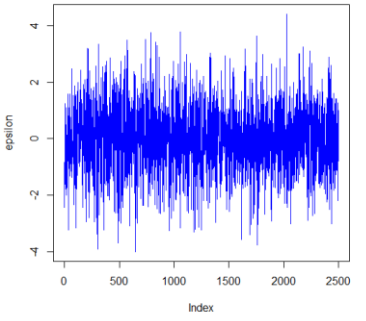
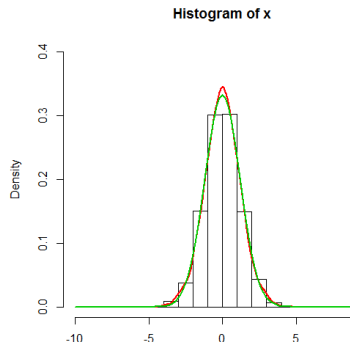
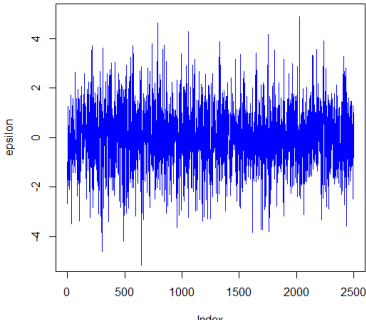
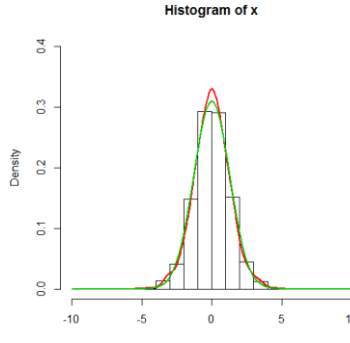
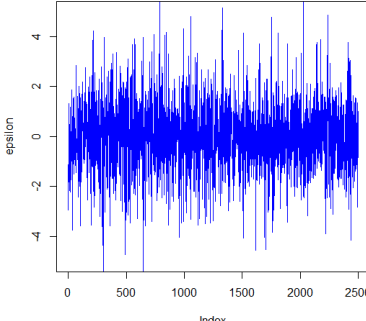
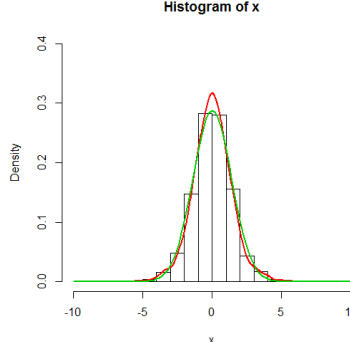


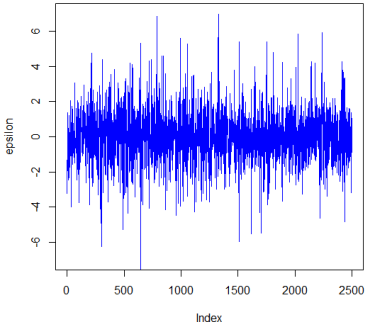
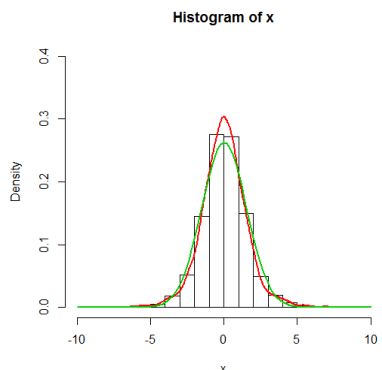
Obr. 5 Histogram ε_t , zeleně hustota normálního rozdělení, červeně jádrový odhad hustoty.

Na obr. 4 vidíte simulovaný proces ARCH(1). Na obr. 5 je vykreslen histogram ε_t a z něj odvozený jádrový odhad hustoty, který je vyznačen červeně. Zeleně je vykreslena hustota normálního rozdělení. Odhad špičatosti rozdělení ε_t je 8,2, což je hodnota vyšší než 3, jak tomu bývá u rozdělení normálního.

5.1.3. Vliv změny parametrů

Nyní se podíváme na to, jak proces ARCH(1) závisí na zvoleném parametru α_1 . Opět budu simulovat proces ARCH(1), stejně jako v podkapitole 5.1.2., akorát při různých hodnotách parametru. Budeme přitom sledovat, jak se vyvíjí odhad špičatosti hustoty.

α_1	Proces $\varepsilon_t = v_t \sqrt{h_t}, h_t = 1 + \alpha_1 \varepsilon_{t-1}^2$	Histogram s hustotou ε_t	KU_ε
$\alpha_1 = 0,2$			3,01
$\alpha_1 = 0,3$			3,19
$\alpha_1 = 0,4$			3,47
0,5			3,93

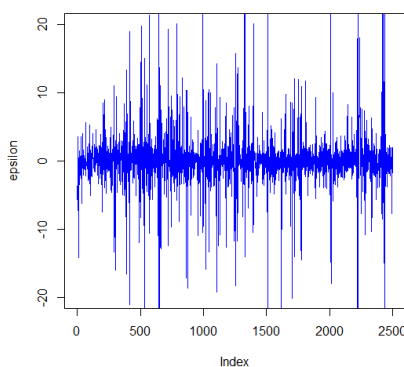
$\alpha_1 = 0,6$			4,68
------------------	---	--	------

Tabulka č. 1 Vliv parametrů na průběh procesu ARCH(1)

V tabulce č. 1. je shrnut průběh simulace ARCH(1) procesu při změně parametrů. Volené parametry α_1 najdete v prvním sloupci tabulky. Ve druhém sloupci jsou umístěny obrázky průběhu nasimulovaného procesu za použití příslušného parametru. Histogram procesu ε_t se zeleně vykreslenou hustotou normálního rozdělení a červeně vyznačeným jádrovým odhadem hustoty vidíte ve třetím sloupci. V posledním sloupci v tabulce naleznete odhadnutou hodnotu charakteristiky špičatosti.

Z tabulky č. 1 zřetelně vyplývá, že se zvyšujícím se parametrem α_1 vzrůstá volatilita procesu ε_t . Čím větší bude hodnota parametru α_1 , tím delší bude doba trvání šoku, který do procesu vstupuje skrz v_t . Se zvyšující se hodnotou parametru, také roste hodnota charakteristiky špičatosti.

Na obr. 6. uvádím pro zajímavost simulovaný proces ARCH(1) s podmíněným rozptylem ve tvaru $h_t = 1 + 1,4\varepsilon_{t-1}^2$, kde parametr $\alpha_1 > 1$, tedy není splněna podmínka stacionarity. Proces ARCH(1) vykazuje viditelné známky nestacionárního chování.



Obr. 6 Model ARCH(1), $\omega=1$, $\alpha_1=1,4$.

5.2. GARCH(p,q) (Generalized ARCH)

GARCH navrhl roku 1986 Bollerssev. Jedná se o zobecnění ARCH modelu, který je rozšířený o zpožděný podmíněný rozptyl. Zobecněný ARCH model najde své uplatnění, je-li třeba použít ARCH modely vysokého řádu q . Je-li q , vysoké, vzniká totiž problém odhadu velkého počtu parametrů. GARCH model se jeví jako model vhodnější, snadnější k indentifikaci i odhadu parametrů. Speciálně GARCH (0,q) je ekvivalentní modelu ARCH (q).

Narozdíl od ARCH modelu je zde podmíněný rozptyl h_t modelován pomocí smíšeného ARMA procesu. Proces reziduí $\{\varepsilon_t\}$ zde specifikujeme takto

$$\varepsilon_t = v_t \sqrt{h_t}, \text{ kde } v_t \text{ je bílý šum WN } (0,1),$$

$$h_t = \omega + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^p \beta_i h_{t-i}.$$

Můžeme také zapsat pomocí operátoru zpoždění

$$h_t = \omega + \alpha(B)\varepsilon_t^2 + \beta(B)h_t, \quad (12)$$

$$\text{kde } \alpha(B) = \alpha_1 B + \dots + \alpha_q B^q \text{ a } \beta(B) = \beta_1 B + \dots + \beta_p B^p.$$

Kladnost podmíněného rozptylu je zaručena při splnění podmínek

$$\omega > 0, \alpha_i > 0, \beta_i \geq 0, i = 1, \dots, p, i = 1, \dots, q.$$

Lze ukázat, že podmíněný rozptyl modelu GARCH(p,q) lze vyjádřit ve formě podmíněného procesu ARCH(∞), z čehož vyplývá skutečnost, že modely ARCH vysokého řádu lze nahradit modely GARCH. Od obou stran (12) odečteme $\beta(B)h_t$ a následně vyjádříme podmíněný rozptyl .

$$h_t - \beta(B)h_t = \omega + \alpha(B)\varepsilon_t^2$$

$$h_t(1 - \beta(B)) = \omega + \alpha(B)\varepsilon_t^2 = \frac{\omega}{1 - \beta(B)} + \frac{\alpha(B)}{1 - \beta(B)}.$$

Model GARCH (p,q) lze přepsat do tvaru procesu ARMA(m,p), kde $m = \max\{p, q\}$,
 $u_t = \varepsilon_t^2 - h_t$.

$$\varepsilon_t^2 = \omega + \sum_{i=1}^m (\alpha_i + \beta_i) \varepsilon_{t-i}^2 - \sum_{i=1}^p \beta_i v_{t-i} + u_t.$$

Specifickou vlastností ACF a PACF procesu $\{\varepsilon_t^2\}$ je, že odpovídá těmto funkcím v modelu ARMA (p,q). Řády procesu GARCH (p,q) se dají určit pomocí ACF a PACF procesu $\{\varepsilon_t^2\}$, podobným způsobem jako v Box – Jenkinsonově metodologii. [8]

Model GARCH (p,q) je stacionární v kovariancích, jsou-li kořeny $1 - \alpha(B) - \beta(B) = 0$ vně jednotkového kruhu.

Z vlastností bílého šumu $v_t \sim WN(0,1)$ a nezávislosti v_t plyne, že nepodmíněná střední hodnota

$$E(\varepsilon_t) = E(v_t \sqrt{\omega + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^p h_{t-i}}) = E v_t E \sqrt{\omega + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^p h_{t-i}} = 0.$$

Nepodmíněný rozptyl má tvar

$$\text{var}(\varepsilon_t) = E(\varepsilon_t^2) - (E(\varepsilon_t))^2 = E(\varepsilon_t^2) = E(v_t^2) E(\omega + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^p \beta_i h_{t-i})$$

$$E(\varepsilon_t^2) = \omega + \sum_{i=1}^q \alpha_i E \varepsilon_{t-i}^2 + \sum_{i=1}^p \beta_i E(h_{t-i})$$

Jelikož $\varepsilon_t = v_t \sqrt{h_t}$, $E(h_t) = E(\varepsilon_t^2)$, pak můžeme psát

$$E(\varepsilon_t^2) = \omega + \sum_{i=1}^q \alpha_i E \varepsilon_{t-i}^2 + \sum_{i=1}^p \beta_i E(\varepsilon_{t-i}^2)$$

$$E(\varepsilon_t^2) (1 - \sum_{i=1}^q \alpha_i - \sum_{i=1}^p \beta_i) = \omega$$

$$E(\varepsilon_t^2) = \frac{\omega}{1 - \alpha_1 - \dots - \alpha_q - \beta_1 - \dots - \beta_p} = \text{var}(\varepsilon_t).$$

Ve výpočtu nepodmíněného rozptylu jsme využili vlastnosti stacionarity procesu, kdy

$$E(\varepsilon_t^2) = E(\varepsilon_{t-i}^2).$$

Podmíněná střední hodnota je taktéž nulová

$$E(\varepsilon_t | \varepsilon_{t-1}, \varepsilon_{t-2}, \dots) = E(v_t) E \left(\sqrt{\omega + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^p \beta_i h_{t-i}} | \varepsilon_{t-1}, \varepsilon_{t-2}, \dots \right) = 0,$$

protože $E v_t = 0$.

Podmíněný rozptyl procesu $\{\varepsilon_t\}$ vypočteme

$$\text{var}(\varepsilon_t | \varepsilon_{t-1}, \varepsilon_{t-2}, \dots) = E(\varepsilon_t^2 | \varepsilon_{t-1}, \varepsilon_{t-2}, \dots) = h_t = \omega + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^p \beta_i h_{t-i},$$

jelikož pokud $E v_t v_{t-i} = 0$ pak $E \varepsilon_t \varepsilon_{t-i} = 0$.

5.2.1 GARCH (1,1)

Jedná ze o zobecnění modelu ARCH(1), kdy dojde k rozšíření modelu o zpožděný podmíněný rozptyl, v tomto případě $q=1$, o první zpoždění podmíněného rozptylu.

Podmíněný rozptyl má tedy tvar

$$h_t = \omega + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1}. \quad (13)$$

Opět uvažujeme podmínky, za jejichž splnění je podmíněný rozptyl kladný, a to $\omega > 0, \alpha_1 > 0, \beta_1 \geq 0$.

Proces GARCH(1,1) můžeme přepsat do tvaru smíšeného procesu ARMA(1,1)

$$\varepsilon_t^2 = \omega + (\alpha_1 + \beta_1) \varepsilon_{t-1}^2 + v_t - \beta_1 v_{t-1}.$$

Model GARCH(1,1) je stacionární v kovariancích, pokud $\alpha_1 + \beta_1 < 1$.

Nepodmíněný rozptyl procesu $\{\varepsilon_t\}$ je ve tvaru

$$\text{var}(\varepsilon_t) = \frac{\omega}{1 - \alpha_1 - \beta_1}.$$

Protože za stacionarity $\text{var}(\varepsilon_t^2) = E(\varepsilon_t^2) = E(h_t)$, dosadíme-li do vztahu (13) za $\omega = E(h_t)(1 - \alpha_1 - \beta_1)$, kde $E(h_t) = \frac{\omega}{1 - \alpha_1 - \beta_1}$, dostáváme vyjádření podmíněného rozptylu ve formě

$$h_t = (1 - \alpha_1 - \beta_1)Eh_t + \alpha_1\varepsilon_{t-1}^2 + \beta_1h_{t-1},$$

$\alpha_1, \beta_1, (1 - \alpha_1 - \beta_1)$ představují váhy, v součtu dávají 1. Podmíněný rozptyl v čase t je daný váženou kombinací nepodmíněného rozptylu ε_t a předchozích hodnot ε_{t-1}^2 a h_{t-1} . [18]

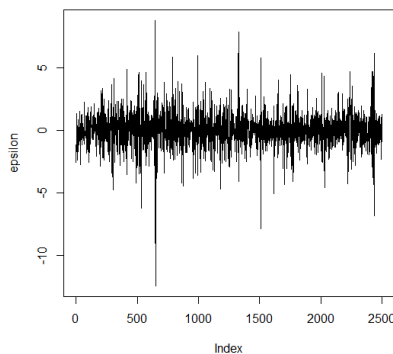
Špičatost pravděpodobnostního rozdělení procesu $\{\varepsilon_t\}$ je vždy větší než špičatost normálního rozdělení. U modelu GARCH (1,1) lze vyjádřit ve formě

$$KU_\varepsilon = \frac{3(1 - (\alpha_1 + \beta_1)^2)}{1 - (\alpha_1 + \beta_1)^2 - 2\alpha_1^2} \text{ pro } 3\alpha_1^2 + 2\alpha_1\beta_1 + \beta_1^2 < 1. \quad [6]$$

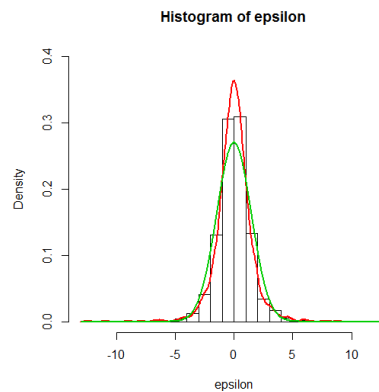
5.2.2 Simulace modelu GARCH (1,1)

Ukažme si simulaci modelu GARCH (1,1) v podobě $\varepsilon_t = v_t\sqrt{h_t}$, kde $v_t \sim WN(0,1)$ a $h_t = 0,3 + 0,6\varepsilon_{t-1}^2 + 0,3h_{t-1}$. R-kód simulace modelu GARCH(1,1) je uložen na cd, pod názvem simulace_GARCH(1,1).R.

Na obrázku č. 7 můžete vidět již nasimulovaný proces GARCH(1,1) a na obr. 8 histogram. Červená křivka vyznačuje jádrový odhad hustoty, který se vyznačuje vyšší špičatostí než hustota normálního rozdělení viz. zelená křivka v obr. 8. Vypočtená hodnota charakteristiky špičatosti je 9,6.



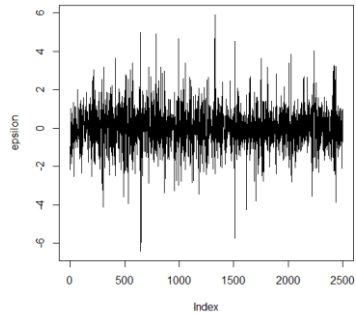
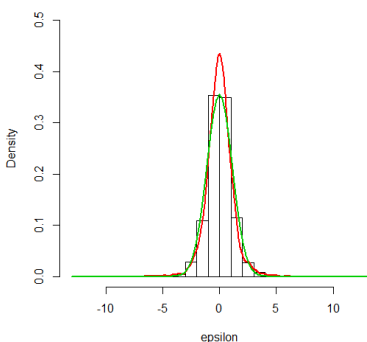
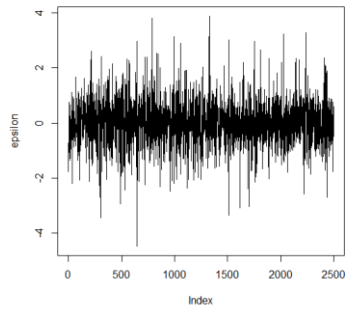
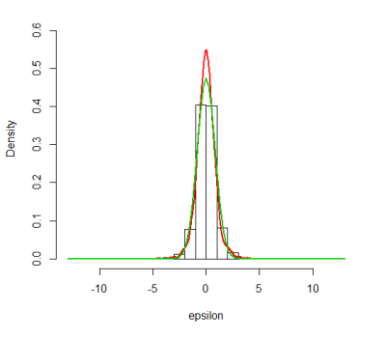
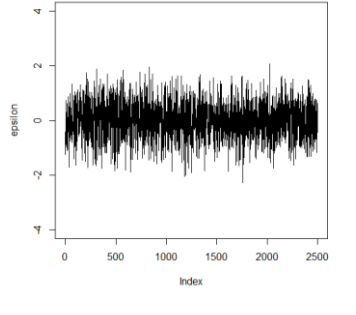
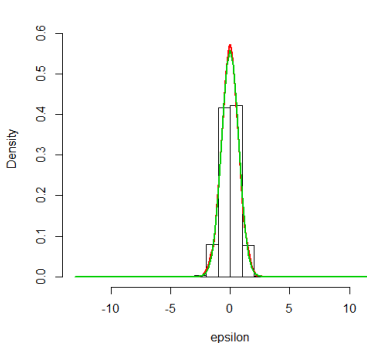
Obr. 7 Proces GARCH(1,1) s parametry $\alpha_1=0,6$, $\beta_1=0,3$ a absolutním členem $\omega=0,3$.



Obr.8 Histogram ε_t , zeleně hustota normálního rozdělení, červeně jádrový odhad hustoty.

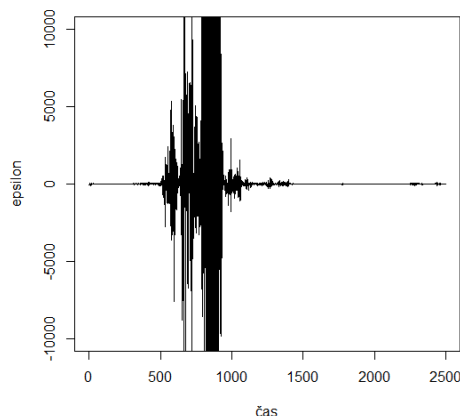
5.2.3. Vliv změny parametrů

Proces GARCH(1,1) je relativně citlivý na změnu hodnot parametrů α_1, β_1 , což můžete vidět v následující tabulce. S růstem obou parametrů, se proces GARCH stává nestabilním, tzn. volatilita ε_t v procesu roste s oběma parametry. Parametry tedy představují opět váhy dopadů a trvání výkyvů, které vstoupí do procesu.

$\omega = 0.3$ $\alpha_1 = 0.6$ $\beta_1 = 0.2$		<p>Histogram of epsilon</p> 	5,81
$\omega = 0.3$ $\alpha_1 = 0.6$ $\beta_1 = 0.01$		<p>Histogram of epsilon</p> 	4,70
$\omega = 0.3$ $\alpha_1 = 0.2$ $\beta_1 = 0.2$		<p>Histogram of epsilon</p> 	3,02

Tabulka č. 2 vliv změny parametrů u modelu GARCH(1,1)

Na obr. 9 je vykreslen nestacionární průběh procesu GARCH(1,1) s podmíněným rozptylem ve tvaru $h_t = 0,3 + 0,5\varepsilon_{t-1}^2 + 0,6h_{t-1}$, kdy není splněna podmínka stacionarity $\alpha_1 + \beta_1 > 1$.



Obr.9 Simulovaný proces GARCH(1,1) s $\omega = 0,3$ a parametry $\alpha_1 = 0,5$, $\beta_1 = 0,6$

5.3 Odhady parametrů v modelech volatility

K odhadování parametrů v modelech podmíněné heteroskedasticity se používá metoda maximální věrohodnosti. V této podkapitole jen stručně shrnu princip fungování odhadu pomocí věrohodnostní funkce.

Odhad parametrů probíhá metodou maximální věrohodnosti. Nejprve je třeba určit věrohodnostní funkci a dále princip odhadu spočívá v hledání takové hodnoty vektoru parametrů $\hat{\theta}$, která maximalizuje věrohodnostní funkci.

5.3.1 Věrohodnostní funkce

Nechť $\mathbf{X} = (X_1, \dots, X_n)$ je náhodný výběr s hustotou

a) $p(\mathbf{x}, \theta) = \prod_{i=1}^n p(x_i, \theta)$, pochází-li náhodný výběr z rozdělení diskrétního typu

b) $f(\mathbf{x}, \theta) = \prod_{i=1}^n f(x_i, \theta)$, je-li hustota náhodného výběrů spojitá,

kde θ představuje vektorový parametr.

Pak věrohodnostní funkce je funkce parametru θ ve tvaru

$$\text{ad a) } L(\theta) = \prod_{i=1}^n p(x_i, \theta),$$

$$\text{ad b) } L(\theta) = \prod_{i=1}^n f(x_i, \theta).$$

Zpravidla je výhodné maximalizovat logaritmickou věrohodnostní funkci, tedy logaritmus věrohodnostní funkce. [19]

Nyní uvažujme model finanční časové řady, který se skládá ze dvou částí, a to z úrovnového modelu a modelu volatility. [6]

$$\mathbf{X}_t = G(\mathbf{X}_t, \boldsymbol{\mu}) + \varepsilon_t, \quad (14)$$

kde $\mathbf{X}_t = (1, X_{t-1}, \dots, X_{t-p})'$, $t=1, \dots, T$

$G(\mathbf{X}_t, \boldsymbol{\mu})$ je skeleton autoregresního modelu s parametry $\boldsymbol{\mu}$, kde G je alespoň dvakrát spojitě derivovatelnou funkcí dle parametrů $\boldsymbol{\mu}$ a

$\{\varepsilon_t\} \sim N(0, h_t)$, kde $\varepsilon_t = v_t \sqrt{h_t}$, $v_t \sim N(0, 1)$ je bílý šum a h_t je podmíněný rozptyl modelu typu GARCH s vektorem parametrů $\boldsymbol{\varphi}$.

Z modelu (14) bude tedy potřeba odhadnout vektor parametrů $\boldsymbol{\theta} = (\boldsymbol{\mu}', \boldsymbol{\varphi}')$.

Nyní sestavme věrohodnostní funkci. Víme, že jestliže $v_t \sim N(0, 1)$, pak $\{\varepsilon_t\} \sim N(0, h_t)$. Logaritmická věrohodnostní funkce bude ve tvaru $L(\boldsymbol{\theta}) = \sum_{t=1}^T l_t(\boldsymbol{\theta})$, kde

$l_t(\boldsymbol{\theta})$ dostaneme následovně.

Hustota normálního rozdělení je ve tvaru

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}. \quad (15)$$

Nyní vypočteme její logaritmus

$$\ln f(x) = -\frac{1}{2} \ln(2\pi) - \frac{1}{2} \ln(\sigma^2) - \frac{(x-\mu)^2}{2\sigma^2} \quad (16)$$

a dosadíme do (16) za $x = \varepsilon_t$, $\mu = 0$, $\sigma^2 = h_t$.

Dostáváme

$$l_t(\boldsymbol{\theta}) = -\frac{1}{2} \ln(2\pi) - \frac{1}{2} \ln(h_t) - \frac{(\varepsilon_t - 0)^2}{2h_t} = -\frac{1}{2} \ln(2\pi) - \frac{1}{2} \ln(h_t) - \frac{\varepsilon_t^2}{h_t}.$$

Logaritmus věrohodnostní funkce považujeme za funkci neznámého parametru θ a hledáme takový odhad parametrů $\hat{\theta}$, který tuto funkci maximalizuje, tzn. řeší se

softwarově soustava rovnic.
$$\sum_{i=1}^T \frac{\partial l_i(\boldsymbol{\theta})}{\partial(\boldsymbol{\theta})} = 0 \quad (17)$$

Jelikož je podmínka (17) nelineární v parametrech, maximálně věrohodného odhadu se dosahuje pomocí zvolené iterativní optimalizační.

6. Výstavba modelů volatility

V následující kapitole bude shrnut postup při tvorbě modelů volatility.

K ilustraci jednotlivých kroků za pomoci softwaru R využijí časovou řadu směnných kurzů CZK/EUR za období 13.10.2010 - 12.10.2012. Data jsou volně stažitelná na webových stránkách České spořitelny www.csas.cz. Kód k příkladu, který se prolíná celou šestou kapitolou je uložen na cd pod názvem `czkeur.R` spolu s daty `czkeur2.csv`.

6.1 Výstupy softwaru R

K aplikaci ARCH či GARCH modelů budu používat funkci `garchFit` z balíčku `fGarch`, proto teď uvedu stručný popis výstupů R v rámci použití této funkce.

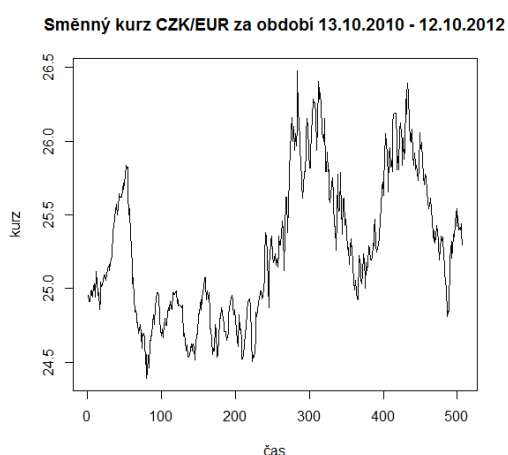
Označíme-li si objekt třídy `fGarch` např. `z=garchFit()`, můžeme později zobrazovat hodnoty komponent objektu, příp. vykreslit je do grafu

- `z@data` vrátí hodnoty dat (reziduí), na které jsme model aplikovali.
- `z@fit` zobrazí výsledky odhadu parametrů a vstupy a výsledky optimalizačního algoritmu. Mezi zobrazenými hodnotami najdeme řadu hodnot `$x` s hodnotami reziduí vstupujících do modelu a `$h` s hodnotami odhadnutého podmíněného rozptylu.
- `z@residuals` vypíše hodnotu reziduí, hodnoty jsou totožné s `$x` v rámci `z@fit`
- `z@fitted` znamená je označením pro vektor modelovaných středních hodnot. Pokud v argumentu funkce `garchFit()` zvolíme `include.mean=FALSE`, bude `z@fitted` řada nul.
- pomocí `z@h.t` můžeme vypsát vektor odhadnutých podmíněných rozptylů z modelu
- `z@sigma.t` zobrazí hodnoty odhadnuté volatility, směrodatné odchylky, jedná se o řadu odmocnin `z@h.t`

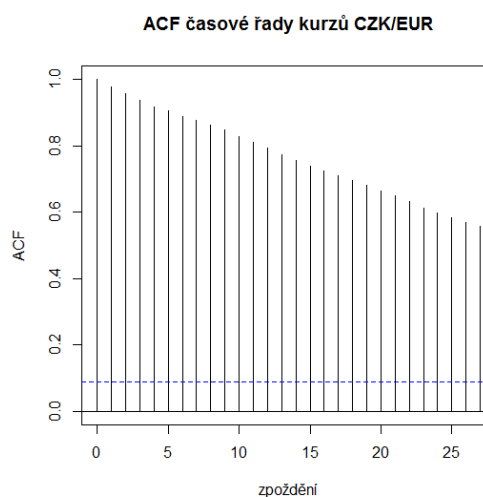
Pomocí funkce `residuals` (název objektu `fGarch`, `standartize = TRUE`), lze dostat standardizovaná rezidua modelu, tj. rezidua podělená odhadem směrodatné odchylky. Funkce `coef` (název objektu `fGarch`) slouží k vypsání odhadnutých hodnot parametrů.

6.2 Zajištění stacionarity časové řady

Stacionarita je cennou vlastností časové řady, které chceme před další analýzou časové řady dosáhnout. Můžeme ji ověřit zhodnocením grafu časové řady nebo analýzou vykreslené ACF, která v případě časové řady s jednotkovými kořeny klesá pomalu, téměř lineárně. Stacionaritu můžeme také otestovat, a to pomocí testu jednotkových kořenů.



Obr. 10 Časová řada kurzu CZK/EUR.



Obr. 11 ACF časové řady kurzů

Na obr. 10 vidíme časovou řadu kurzů CZK/EUR, která vykazuje nestacionární chování, což vypovídá i autokorelační funkce na obr. 11. ACF klesá pomalu, téměř lineárně a v prvním zpoždění je blízka jedné.

ACF vykreslíme v R pomocí příkazu `acf()`. ACF. Počet zpoždění, pro která chceme autokorelační funkci spočítat, nastavíme pomocí argumentu `lag.max`. V případě, že nenastavíme tuto hodnotu, ACF bude automaticky spočtena pro $\text{lag.max} = 10 \log_{10}(N/m)$ zpoždění, kde N je počet pozorování a m počet časových řad ($m=1$).

6.2.1 Augmented Dickey-Fuller test (ADF)

Dickey-Fullerův test patří mezi tzv. testy jednotkových kořenů (unit root test) a slouží k ověření hypotézy H_0 , že časová řada má jednotkový kořen, tedy není stacionární.

Již kapitole 2.4.5. bylo řečeno, že proces $AR(p)$ ve tvaru

$$X_t = \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + \varepsilon_t$$

je stacionární, jsou-li kořeny operátoru zpětného posunutí $\phi(B)$ vně jednotkového kruhu. Pomocí ADF můžeme tuto skutečnost otestovat. Kromě ADF testu je také možno použít testů KPSS (Kwiatkowski–Phillips–Schmidt–Shin test) či PP (Phillips-Perron test).

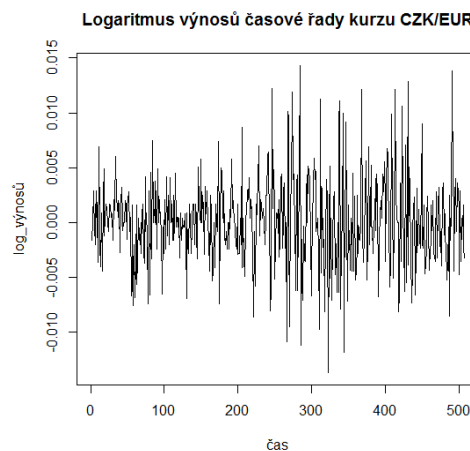
V softwaru R provedeme ADF test pomocí příkazu `adf.test()`, z balíčku `tseries`. Volba správného počtu zpoždění zahrnutých do testu (argument `k`) je otázkou. Při zvolení příliš malého `k` může zbývající autokorelace způsobit chybnost testu. Naopak zvolíme-li příliš velké `k` sníží se síla testu. [9]

Nyní aplikujeme ADF test na časovou řadu kurzů. V R. Neuvědeme-li jinak, je `k` nastaveno jako zaokrouhlená hodnota $((\text{length}(x)-1)^{(1/3)})$, kde $\text{length}(x)$ je délka vyšetřované časové řady. P-hodnota testu vyšla 0.47, tzn., že nulovou hypotézu o přítomnosti jednotkového kořenu nemůžeme zamítnout, bude tedy potřeba převést časovou řadu na stacionární.

6.2.2. Stacionarizace časové řady

Stacionarizace časové řady se provádí diferencováním. Používá se především diferencování nižších řádů, aby byl zachován ráz časové řady.

V našem případě je vhodné transformovat si časovou řadu kurzů CZK/EUR logaritmickou transformací a provést první diference. Dostaneme tak časovou řadu logaritmů výnosů.

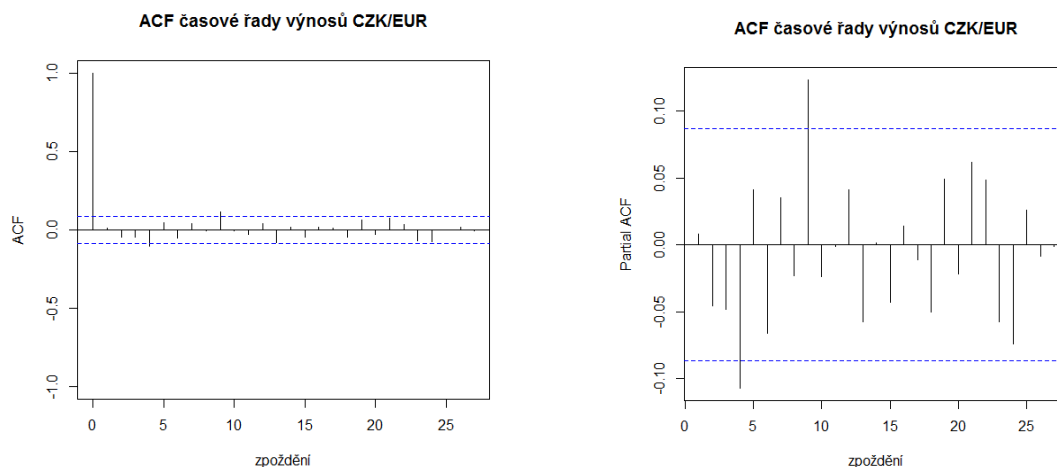


Obr. 12 Logaritmy výnosů.

Jak můžeme vidět na obr. 12, difference logaritmičsky transformované a diferencované řady již vypadají rozumněji, ve smyslu stacionarity. Naši domněnku ověříme opět pomocí Dickey-Fullerova testu. ADF s p-hodnotou 0,01 ukázal, že H_0 o přítomnosti jednotkových kořenů zamítáme a časová řada je tedy stacionární. V případě, že by tomu tak nebylo, mohli bychom přistoupit k druhým diferencím.

6.3 Určení úrovněvého modelu

V dalším kroku je potřeba vyšetřit průběh autokorelační a parciální autokorelační funkce. Cílem tohoto postupu je určit vhodný model Box-Jenkinsonovy metodologie tak, abychom získali časovou řadu reziduí, která nejsou korelovaná. Jelikož pracujeme s prvními diferencemi, je časová řada často stacionární a můžeme identifikovat model podmíněné střední hodnoty.

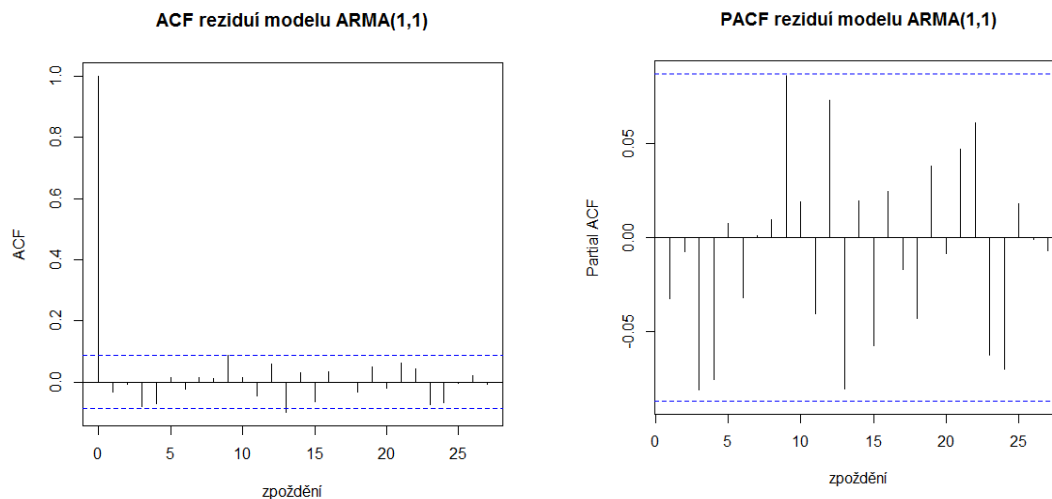


Obr. 13 ACF a PACF logaritmů výnosů.

Dle obrázku ACF by se dalo říct, že časová řada se chová jako proces bílého šumu, kde ACF je téměř nulová. Podle PACF to však říci nemůžeme, vidíme zde významnější parciální autokorelace. Zkusíme proto postupně, od jednoduššího ke složitějšímu aplikovat úrovněvý model. Preferujeme co nejmenší počet parametrů. Postupně byly vyzkoušeny

modely AR(1),MA(1) s i bez absolutního členu, ale parciální autokorelace se podařilo odstranit až u modelu ARMA(1,1) bez absolutního členu, ve tvaru

$$x_t = -0,92477x_{t-1} + 0,98115\varepsilon_{t-1} + \varepsilon_t.$$



Obr. 14 ACF a PACF reziduí z aplikovaného modelu ARMA(1,1) s parametry ar= -0,92477 a ma= 0,98115.

6.4 Ověření kvality úrovnového modelu

V dalším kroku je potřeba ověřit kvalitu zvoleného úrovnového modelu, z hlediska jak z hlediska významnosti parametrů, tak i autokorelace a normality reziduí.

6.4.1. Významnost parametrů

K ověření významnosti parametrů se využívá t-test, s jehož pomocí můžeme testovat hypotézu nulovosti jednotlivých parametrů $H_0 : b_j = 0$, kde nulová hypotéza říká, že testovaný parametr je nevýznamný.

Testovací statistika je ve tvaru

$$t = \frac{b_j - \beta_j^{H_0}}{\sigma_j} \sim t_{T-1},$$

kde b_j je odhad hodnoty parametru, β_j hypotetická hodnota je rovna nule, σ_j směrodatná odchylka odhadu, T počet pozorování.

V R-ku lze získat p-hodnotu testu významnosti parametru zadáním příkazu `summary()` po aplikaci konkrétního modelu a odhadu jeho parametrů. Funkce vrátí hodnoty odhadnutých parametrů, jejich směrodatnou odchylku, hodnotu testové statistiky, p-hodnotu testu (p-value), ale také hodnotu Akaikeho informačního kritéria pro daný model (AIC).

V případě časové řady logaritmu výnosů CZK/EUR, dostaneme tyto hodnoty

Coefficient(s):

```
Estimate Std. Error t value Pr(>|t|)
ar1 -0.92477  0.01855  -49.86 <2e-16 ***
ma1  0.98115  0.01103   88.95 <2e-16 ***,
```

kde můžeme vidět, že p-hodnota t-testu u obou parametrů je menší jak hladina významnosti 0,05, tudíž zamítáme hypotézu o nevýznamnosti parametrů.

Jelikož jsou hodnoty AR parametru i MA parametru významně nenulové, mohli bychom zkusit přidat další parametry. V tabulce č. 3 můžete vidět výstup funkce `summary()` po aplikaci modelů ARMA(2,1) a ARMA(1,2). Ukazuje se, že přidané parametry jsou nevýznamné.

ARMA(2,1)					ARMA(1,2)				
	Estimate	Std. Error	t value	Pr(> t)		Estimate	Std. Error	t value	Pr(> t)
ar1	0.66831	0.19387	3.447	0.000566	ar1	0.59726	0.22459	2.659	0.00783
ar2	-0.07134	0.04621	-1.544	0.122638	ma1	-0.59673	0.22549	-2.646	0.00814
ma1	-0.66429	0.19012	-3.494	0.000476	ma2	-0.06729	0.04630	-1.453	0.14609

Tab. 3 Výsledky aplikace modelů ARMA(2,1), ARMA(1,2) z hlediska významnosti parametrů.

K závěru, že ARMA(1,1) je nejvhodnějším modelem pro data, lze využít srovnání modelů pomocí Akaikeho informačního kritéria (AIC), které má podobu

$$AIC(M) = \ln \hat{\sigma}_\varepsilon^2 + 2 \frac{M}{T},$$

kde M je počet parametrů v modelu, tedy p+q,

T...počet reziduí získaných z modelu

$\hat{\sigma}_\varepsilon^2$...odhad reziduálního rozptylu modelu. [6]

Při srovnání modelů dle AIC preferujeme model s nejnižší hodnotou.

V příkladu směnného kurzů nejmenší hodnota vychází pro model ARMA(1,1), viz tab. 4.

Akaikeho informační kritérium (AIC)	
ARMA(1,1)	-4154.16
ARMA(2,1)	-4147.45
ARMA(1,2)	-4147.21

Tab.4 Hodnota AIC pro vybrané modely.

6.4.2 Testování autokorelace reziduí

Autokorelace vyjadřuje závislost různých pozorování skrze chyby. U analýzy ekonomických časových řad bývá tento problém častý. Autokorelaci můžeme odhalit již na základě grafu reziduí, kde kladná rezidua jsou následována kladnými a naopak. Vykreslíme-li si ACF a PACF reziduí, je autokorelace ještě zřetelnější.

Zhodnocení autokorelace reziduí můžeme ještě doplnit formálním testem. K otestování významnosti autokorelací slouží tzv. Portmanteau test. Zjistíme-li přítomnost významných autokorelací reziduí, je třeba přehodnotit volbu úrovnového modelu.

a) Box-Pierce test

Testová statistika Box-Pierce testu vypadá takto

$$Q_{BP} = T \sum_{k=1}^K \hat{r}_k^2,$$

kde \hat{r}_k je výběrová autokorelační funkce $\hat{r}_k = \frac{\sum_{t=k+1}^T \hat{\varepsilon}_t \hat{\varepsilon}_{t-k}}{\sum_{t=1}^T \hat{\varepsilon}_t^2}$,

$k = 1, \dots, K$ vyjadřuje zpoždění,

T ... počet pozorování,

$\hat{\varepsilon}_t$ odhadnutá rezidua, $t=1, \dots, T$

a za platnosti nulové hypotézy H_0 má pro velké T a K přibližně χ^2 rozdělení s $(K-p-q)$ stupni volnosti.

Testujeme hypotézu $H_0 : \zeta_1 = \zeta_2 = \dots = \zeta_K = 0$, která říká, že autokorelace pro zpoždění $k = 1, \dots, K$ jsou nevýznamné. [6]

V R se Box Pierce test provede příkazem `Box.test` z balíčku `stats`, kde v argumentech zvolíme za `type = („Box-Pierce“)`. Pokud se jedná o rezidua nějakého modelu, je potřeba zvolit argument `fitdf = p+q`, kde p je počet parametrů u autoregresních členů, a q počet parametrů u členů procesu klouzavých součtů. Zadáním argumentu `lag` definujeme zpoždění K , na jehož základě bude sestavena testová statistika. K můžeme volit jako \sqrt{T} , kde T délka analyzované řady [2] a mělo by platit, že `lag > fitdf = p+q`.

V ilustrativním příkladu jsme na základě analýzy ACF a PACF reziduí modelu, viz. obr. 14 neshledali významné autokorelace. Box-Pierce test tuto skutečnost potvrdil. Jelikož vyšetřujeme rezidua modelu ARMA(1,1), zadáme do argumentu `fitdf = 1+1=2` a zvolíme `lag = 22`, což je přibližně odmocnina z počtu reziduí \sqrt{T} . P-hodnota vypočtená v R pro `lag = 22` vychází 0.15, tudíž nezamítáme hypotézu, že autokorelace jsou nevýznamné.

b) Box-Ljung test

Box – Ljung test by měl být jakousi „vylepšenou“ modifikací Box-Pierce testu v případě, že chceme testovat hypotézu o nezáznamných autokorelacích ve zpoždění $k = 1, \dots, K$ pro malé rozsahy výběrů. [6] Testová statistika je ve tvaru

$$Q = T(T+2) \sum_{k=1}^K \frac{\hat{r}_k^2}{(T-k)} \stackrel{H_0}{\sim} \chi^2_{(K-p-q)},$$

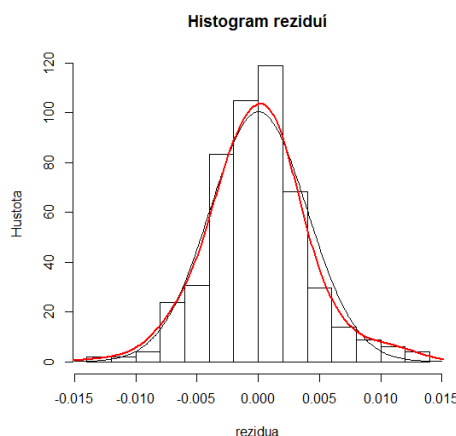
kde \hat{r}_k je výběrová autokorelační funkce, T .. počet pozorování a $k = 1, \dots, K$..zpoždění.

Počítáme-li Box-Ljung test v R, použijeme opět příkaz `Box.test`, ale za typ zvolíme `type=(„Ljung-Box“)`. V našem příkladu vyjde pro `lag=22` p-hodnota Ljung-Box testu 0.13, což opět znamená, že hypotézu nezamítáme.

6.4.3. Normalita reziduí

Normalitu nesystematických vlivů testujeme hlavně proto, že je důležitým předpokladem při mnoha statistických testech. K tomu, abychom zjistili, zda chyby splňují předpoklad normality, můžeme využít vykreslení histogramu reziduí.

Na obr. 13 můžete vidět histogram reziduí, červená křivka znázorňuje jádrový odhad hustoty a černá křivka hustotu normálního rozdělení se střední hodnotou a rozptylem stejnými jako u rozdělení reziduí. Již víme, že u finančních časových řad nebývá často předpoklad normality chyb splněn. Rozdělení reziduí bývá špičatější s těžkými konci, viz obr. 15. Vypočtený odhad špičatosti má hodnotu přibližně 4,0.



Obr. 15 Histogram reziduí modelu ARMA(1,1), červeně jádrový odhad hustoty, černě hustota normálního rozdělení.

Kromě vizuálního zhodnocení je však vhodné normalitu reziduí otestovat pomocí Jarque-Berova testu normality. Jarque-Berův test normality je založen na testování charakteristik šikmosti a špičatosti. Porovnává tyto charakteristiky reziduí s charakteristikami normálního rozdělení. Nulová hypotéza říká, že data pocházejí z normálního rozdělení. Testová statistika je sestavena ve formě

$$JB = \frac{T}{6} \left(SK^2 + \frac{(KU - 3)^2}{4} \right)^{H_0} \sim \chi^2_2,$$

kde $SK = \frac{\hat{\eta}_3}{\hat{\sigma}^3} = \frac{\frac{1}{T} \sum_{t=1}^T (x_t - \bar{x})^3}{\left(\frac{1}{T} \sum_{t=1}^T (x_t - \bar{x})^2 \right)^{\frac{3}{2}}}$ je výběrová šikmost,

$$KU = \frac{\hat{\eta}_4}{\hat{\sigma}^4} = \frac{\frac{1}{T} \sum_{t=1}^T (x_t - \bar{x})^4}{\left(\frac{1}{T} \sum_{t=1}^T (x_t - \bar{x})^2 \right)^2}$$
 je výběrová špičatost

a $\bar{x} = \frac{\sum_{t=1}^T x_t}{T}$ výběrový průměr.

K zamítnutí testované hypotézy může přispět nejen nenormalita, ale také skutečnost, že nesystematické složky nemají v čase konstantní rozptyl. [6]

V softwaru R se provádí pomocí příkazu `jarque.bera.test ()` z balíčku `tseries`. V příkladu směnných kurzů CZK/EUR byla prokázána nenormalita, p-hodnota testu $8,004 \cdot 10^{-7} < 0.05$, tedy zamítáme, že rezidua pochází z normálního rozdělení.

6.4.4 Testování přítomnosti heteroskedasticity v časové řadě

Dalším krokem výstavby modelu volatility je testování ARCH efektu řádu q v časové řadě. Potvrdí-li se heteroskedasticita, můžeme na základě tohoto zjištění aplikovat model volatility. Protože se ve své práci zabývám výhradně lineárními modely volatility, budeme k testování podmíněné heteroskedasticity využívat tzv. ARCH – LM test, který funguje na principu testování autokorelace čtverců reziduí se zpožděnými hodnotami čtverců reziduí.

Abychom mohli aplikovat model typu ARCH/GARCH, měla by se u testování reziduí úrovnového modelu prokázat podmíněná heteroskedasticita. Naopak po aplikaci modelu volatility, u testu čtverců standardizovaných reziduí by již heteroskedasticita prokázána být neměla.

Vezmeme-li si podmíněný rozptyl např. modelu ARCH(q) ve tvaru,

$$h_t = \omega + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \varepsilon_{t-2}^2 + \dots + \alpha_q \varepsilon_{t-q}^2.$$

nulová hypotéza testu předpokládá homoskedasticitu, která nastane v případě, budou-li

všechny parametry α_i nulové, tedy $\alpha_1 = \alpha_2 = \dots = \alpha_q = 0$.

Alternativní hypotéza předpokládá nenulovost alespoň jednoho z koeficientů $\alpha_1, \dots, \alpha_q$, tedy nekonstantní podmíněný rozptyl - výskyt heteroskedasticity. Testová statistika testu autokorelací reziduí má tvar

$$LM = TR^2 \sim \chi_q^2, \quad (18)$$

kde R^2 značí index determinace, T je počet pozorování.

Postup při testování podmíněné heteroskedasticity :

a) Nejprve určíme úrovnový model, ze kterého dostaneme rezidua $\hat{\varepsilon}_t$.

b) Sestavíme regresní model $\hat{\varepsilon}_t^2 = \hat{\omega} + \alpha_1 \hat{\varepsilon}_{t-1}^2 + \alpha_2 \hat{\varepsilon}_{t-2}^2 + \dots + \alpha_q \hat{\varepsilon}_{t-q}^2 + u_t$,

kde $u_t = \varepsilon_t^2 - h_t$ $t = q+1, \dots, T$.

Z uvedeného modelu získáme rezidua $\hat{u}_t = \hat{\varepsilon}_t^2 - \hat{\omega} - \hat{\alpha}_1 \hat{\varepsilon}_{t-1}^2 - \hat{\alpha}_2 \hat{\varepsilon}_{t-2}^2 - \dots - \hat{\alpha}_q \hat{\varepsilon}_{t-q}^2$

a spočteme

reziduální součet čtverců $RSC = \sum_{t=1}^T \hat{u}_t^2$,

celkový součet čtverců $S_t = \sum_{t=1}^T (\hat{u}_t - \bar{u})^2$,

Následně můžeme spočítat index determinace $R^2 = 1 - \frac{RSC}{S_t}$ a dosadit do testové statistiky (18).

V R lze ARCH test spočítat pomocí výše popsaného postupu nebo automaticky pomocí příkazu ArchTest() z balíčku FinTs.

U časové řady kurzů CZK/EUR byla u reziduí z modelu ARMA(1,1) prokázána podmíněná heteroskedasticita, p-hodnota ARCH testu pro lag=1 vyšla 0,02. Pro lag=5 a

pro lag=10 vyšly také hodnoty menší jak 0,05. Ve všech případech zamítáme hypotézu o homoskedasticitě a můžeme na rezidua aplikovat model volatility.

6.4.5 Určení modelu volatility a jeho verifikace

Při určování modelu volatility postupujeme od jednodušších, s méně parametry, ke složitějším. Můžeme využít graf ACF a PACF řady ε_t^2 a identifikovat model podobně jako u modelů Box-Jenkinsonovy metodologie.

Vhodnost modelu ověřujeme stejnými způsoby jako u úrovnových modelů s tím rozdílem, že se zaměřujeme na vyšetřování standardizovaných reziduí. Standardizovaná rezidua stanovíme vydělením reziduí odhadnutou směrodatnou odchylkou

$$\varepsilon_t = v_t \sqrt{h_t} \rightarrow \hat{v}_t = \frac{\hat{\varepsilon}_t}{\sqrt{\hat{h}_t}}$$

Zkoumáme tedy vhodnost modelu z hlediska významnosti parametrů, normality, autokorelace a v neposlední řadě provedeme také ARCH test, abychom zjistili, zda zvolený model zachytil a eliminoval podmíněnou heteroskedasticitu.

V případě více akceptovatelných modelů, můžeme v případě stejného počtu parametrů porovnat modely na základě hodnoty věrohodnostní funkce, kdy dáváme přednost modelu s maximální hodnotou. Je-li počet parametrů různý, můžeme přistoupit ke zhodnocení podle Akaikeho informačního kritéria. Alternativně lze zvolit Schwartzovo kritérium SC, Hannanovo-Quinnovo kritérium HQ.

V tab. č. (5) můžete vidět shrnutí modelů, které jsem aplikovala na řadu reziduí v příkladu.

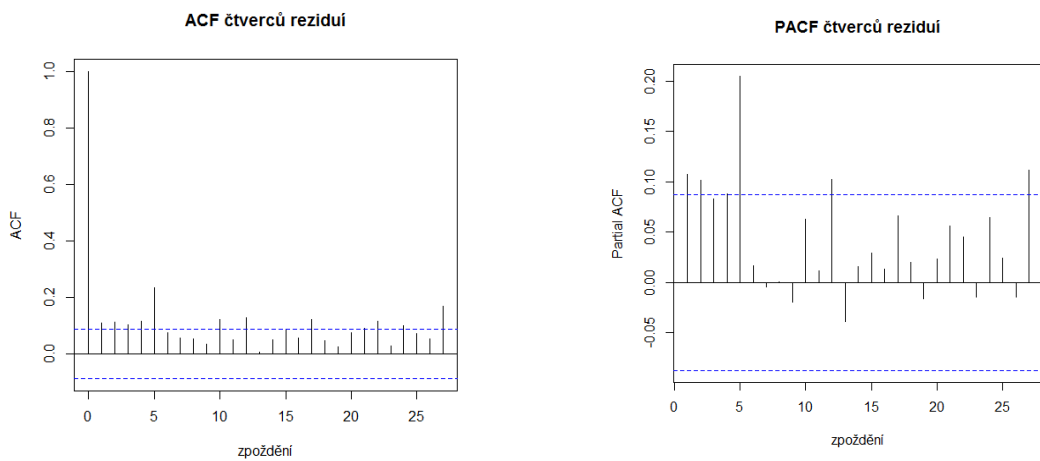
Typ modelu	Odhad parametrů	Významnost parametrů	Jarque-Bera	Ljung-Box lag=15	ARCH	AIC
ARCH(1)	omega 1.390 10 ⁻⁵ alpha1 0,1910	Parametry významné	3,05 · 10 ⁻⁶	0.09	7.7015 · 10 ⁻⁵	-8.2270
GARCH(1,1)	omega 4,35 · 10 ⁻⁷ alpha1 0,0762 beta1 0,8972	Parametry významné	0.0005	0.25	0.5474	-8.3034

Tab. 5 Určení modelu volatility.

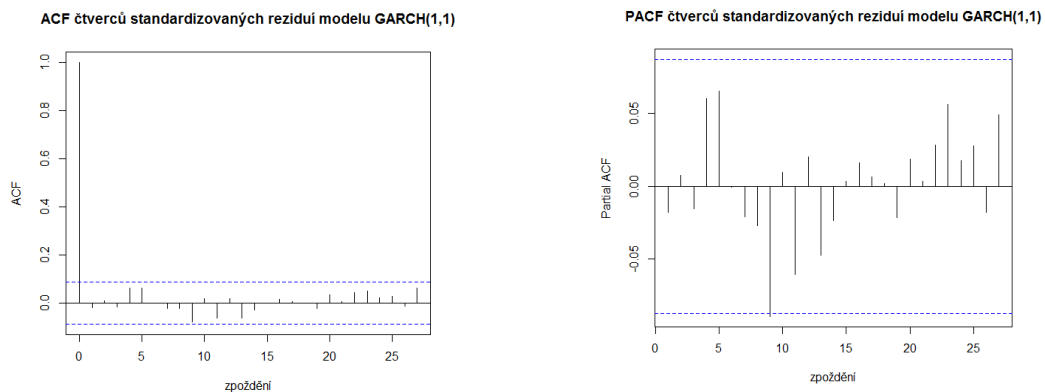
Jako první byl pomocí příkazu `garchfit` z balíčku `fGarch` aplikován model ARCH(1), kde parametry byly významné. Hypotézu o nevýznamných autokorelacích, jsme dle Ljung-Box testu nezamítli, ale při testu standardizovaných reziduí se ukázalo, že zamítáme hypotézu o homoskedasticitě. Proto zkusíme aplikovat zobecněný model GARCH (1,1), čemuž napovídá i tvar ACF a PACF ε_t^2 , viz. obr. 16.

Zde již můžeme dle ARCH-testu i ACF a PACF druhých mocnin standardizovaných reziduí (obr. 17) vidět, že heteroskedasticita byla s pomocí modelu eliminována. Aplikovali jsme tedy GARCH(1,1) model ve tvaru

$$\varepsilon_t = v_t \sqrt{0,0000004352 + 0,0762\varepsilon_{t-1} + 0,8972h_{t-1}}.$$



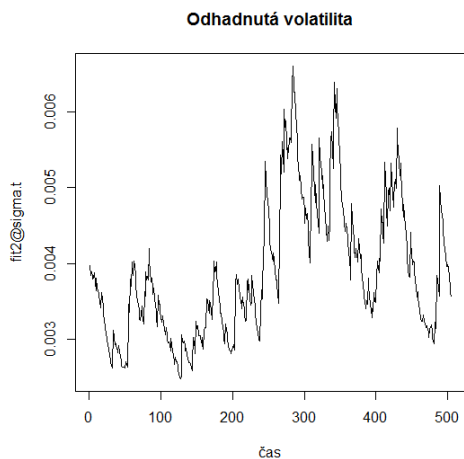
Obr. 16 ACF a PACF druhých mocnin reziduí.



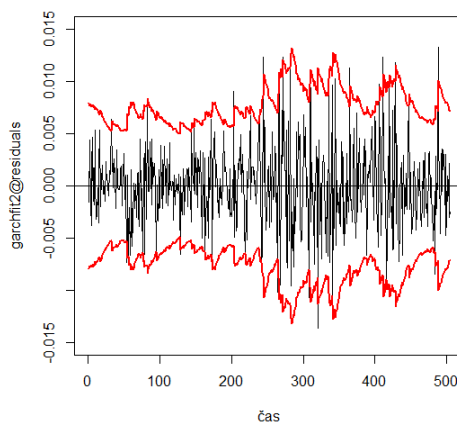
Obr. 17 ACF a PACF druhých mocnin standardizovaných reziduí modelu GARCH(1,1)

Při aplikaci modelů GARCH(2,1) a GARCH(1,2) byly přidány parametry nevýznamné.

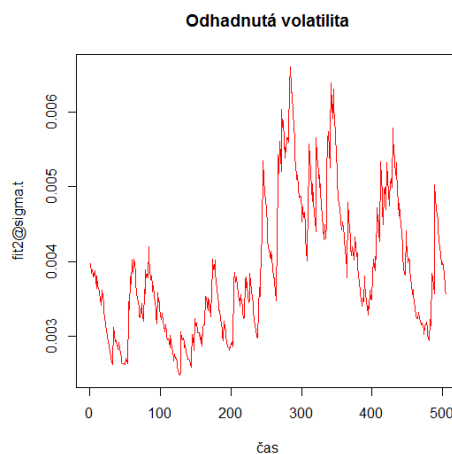
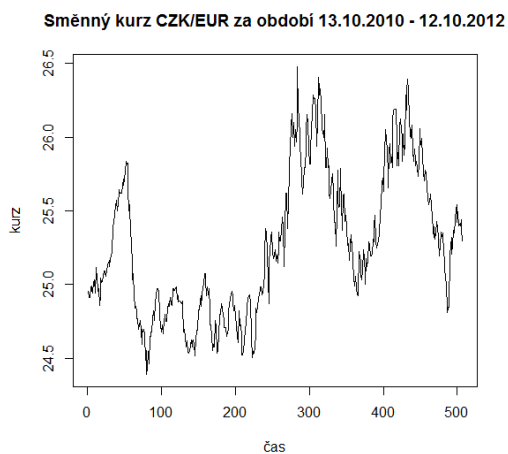
Odhady parametrů spolu s testy jejich významnosti, testy standardizovaných reziduí a kritérii pro volbu vhodného získáme v programu R pomocí příkazu `summary()`, kde do argumentu vložíme název objektu, jak jsme si jej pojmenovali při aplikaci příkazu `garchfit`, např. `fit=garchfit()`, pak `summary(fit)`.



Obr. 18 Odhadnutá volatilita (podmíněná směrodatná odchylka).



Obr.19 Interval spolehlivosti pro rezidua



Obr. 20 Časová řada kurzu CZK/EUR a odhadnutá volatilita z modelu GARCH(1,1).

Na obr. 20 si všimněme, že průběh z modelu odhadnuté podmíněné směrodatné odchylky v čase, vyjadřující volatilitu časové řady směnného kurzu CZK/EUR, téměř kopíruje tvar průběhu vývoje kurzu CZK/EUR.

Pomocí příkazu `garchFit` je možné zdlouhavý proces aplikování modelu ARMA(1,1) a poté GARCH(1,1) na rezidua modelu provést najednou. Použijeme k tomu příkaz

`garchFit (~arma(1,1)+garch(1,1),data,include.mean=FALSE,include.intercept=FALSE)`.

Odhady parametrů nevyjdou zcela stejně, ale velmi podobně viz tab. 6. Tento postup je nejen kratší, ale na jeho základě můžeme poté tvořit předpovědi nejen volatility, ale modelu jako celku.

~ARMA(1,1)	ar1 -0.9248 ma1 0.9812
~GARCH(1,1)	ω $4,352 \cdot 10^{-7}$ α_1 0.0762 β_1 0.8972
~ARMA(1,1)+GARCH(1,1)	ar1 -0.9032 ma1 0.9731 ω $4,547 \cdot 10^{-7}$ α_1 0.0815 β_1 8.909e-01

Tab. 6 Srovnání odhadů v R modelů zvlášť a jako celku .

Při aplikaci modelu GARCH na finanční časové řady se může stát, že hodnota parametrů v součtu je blízká jedné. Můžeme si všimnout, že tak je tomu i v našem případě, kde $\alpha_1 + \alpha_2 = 0,97$. V tomto případě by bylo vhodnější použít modelu lineárního modelu IGARCH, který bývá označován také jako proces integrovaný v rozptylu a můžeme ho definovat ve tvaru procesu ARIMA (0,1,1) takto

$$\varepsilon_t^2 = \omega + \varepsilon_{t-1}^2 + u_t - \beta_1 u_{t-1}.$$

6.4.6 Předpovědi podmíněného rozptylu

Jedním z důvodů, proč vlastně konstruuujeme modely volatility je možnost předpovídat její budoucí vývoj. Znalost předpovědi volatility je totiž zásadní pro

rozhodnutí v rámci investic, optimalizace portfolia či risk managementu. Na základě modelů ARCH, GARCH můžeme předpovídat budoucí hodnoty podmíněného rozptylu h_t . S čím větším předpovědním horizontem h chceme tvořit předpovědi, tím kvalitnější musí náš model být. Avšak je důležité poznamenat, že předpovědi s příliš velkým h nemají smysl, konvergují totiž k nepodmíněnému rozptylu procesu.

Předpověď podmíněného rozptylu o h kroků dopředu se vypočte jako podmíněná střední hodnota h_{T+h} za informace dostupné do času T (označme Ω_T), kdy je předpověď konstruována.

Tvar podmíněného rozptylu modelu GARCH(p,q) je

$$h_t = \omega + \alpha_1 \varepsilon_{t-1}^2 + \dots + \alpha_q \varepsilon_{t-q}^2 + \beta_1 h_{t-1} + \dots + \beta_p h_{t-p}.$$

Pro $t = T+h$ můžeme přepsat

$$h_{T+h} = \omega + \alpha_1 \varepsilon_{T+h-1}^2 + \dots + \alpha_q \varepsilon_{T+h-q}^2 + \beta_1 h_{T+h-1} + \dots + \beta_p h_{T+h-p}.$$

Předpověď podmíněného rozptylu s minimální MSE zapíšeme jako

$$\begin{aligned} \hat{h}_{T+h} &= E(h_{T+h} | \Omega_T) \\ &= \omega + \alpha_1 E(\varepsilon_{T+j}^2 | \Omega_T) + \dots + \alpha_q E(\varepsilon_{T+j}^2 | \Omega_T) + \beta_1 E(h_{T+j} | \Omega_T) + \dots + \beta_p E(h_{T+j} | \Omega_T) \end{aligned}$$

kde jsme provedli změnu v indexech, tzn. $j=h, h-1, \dots$

Při tvorbě předpovědí vycházíme ze vztahů

$$\bullet E(\varepsilon_{T+j}^2 | \Omega_T) = \hat{\varepsilon}_{T+j}^2 = E(h_{T+j} | \Omega_T) = \hat{h}_{T+j} \quad \text{pro } j \geq 1.$$

Tento výraz platí, jelikož

$$\text{var}(\varepsilon_{T+j} | \Omega_T) = E(\varepsilon_{T+j}^2 | \Omega_T) - (E(\varepsilon_{T+j} | \Omega_T))^2 = E(\varepsilon_{T+j}^2 | \Omega_T) = E(h_{T+j} | \Omega_T).$$

$$\bullet E(h_{T+j} | \Omega_T) = h_{T+j} \quad \text{a} \quad E(\varepsilon_{T+j}^2 | \Omega_T) = \varepsilon_{T+j}^2 \quad \text{pro } j \leq 0. \quad [6]$$

Uvažujme model GARCH (1,1) s tvarem podmíněného rozptylu

$$h_t = \omega + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1}$$

a předpokládejme, že je proces stacionární v kovariancích, tzn. $\alpha_1 + \beta_1 < 1$

$$\text{s nepodmíněným rozptylem } \sigma^2 = \frac{\omega}{1 - \alpha_1 - \beta_1}. \quad (20)$$

Předpověď v čase T o krok dopředu, h=1

$$\begin{aligned}\hat{h}_{T+1} &= \omega + \alpha_1 E(\varepsilon_T^2 | \Omega_T) + \beta_1 E(h_T | \Omega_T) \\ &= \omega + \alpha_1 h_T + \beta_1 h_T \\ &= \omega + (\alpha_1 + \beta_1) h_T\end{aligned}$$

a dosadíme z (20) za $\omega = \sigma^2(1 - (\alpha_1 + \beta_1))$

$$\begin{aligned}&= \sigma^2 - \sigma^2(\alpha_1 + \beta_1) + (\alpha_1 + \beta_1)h_T \\ &= \sigma^2 + (\alpha_1 + \beta_1)(h_T - \sigma^2).\end{aligned}$$

K předpovědi o dva kroky dopředu využijeme \hat{h}_{T+1}

$$\begin{aligned}\hat{h}_{T+2} &= \omega + \alpha_1 E(\varepsilon_{T+1}^2 | \Omega_T) + \beta_1 E(h_{T+1} | \Omega_T) \\ &= \omega + (\alpha_1 + \beta_1) \hat{h}_{T+1} \\ &= \sigma^2 + (\alpha_1 + \beta_1) (\hat{h}_{T+1} - \sigma^2)\end{aligned}$$

a dosadíme-li za \hat{h}_{T+1}

$$\begin{aligned}&= \sigma^2 + (\alpha_1 + \beta_1)(\sigma^2 + (\alpha_1 + \beta_1)(h_T - \sigma^2) - \sigma^2) \\ &= \sigma^2 + (\alpha_1 + \beta_1)^2 (h_T - \sigma^2)\end{aligned}$$

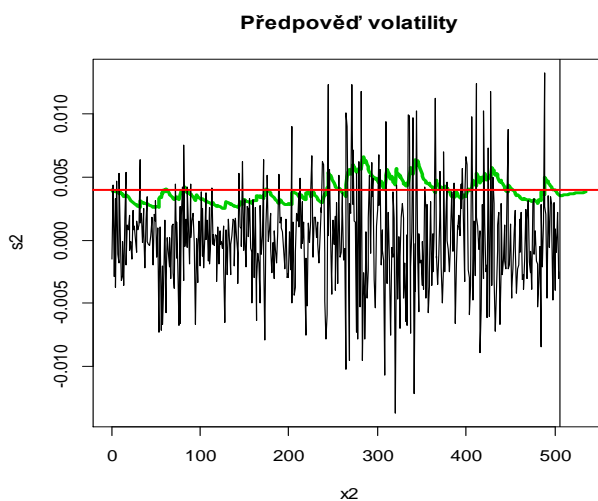
$$\begin{aligned}\hat{h}_{T+h} &= \omega + (\alpha_1 + \beta_1) \hat{h}_{T+h-1} \\ &= \sigma^2 + (\alpha_1 + \beta_1) (\hat{h}_{T+h-1} - \sigma^2) \\ &= \sigma^2 + (\alpha_1 + \beta_1)^h (h_T - \sigma^2)\end{aligned}$$

Ve vzorci pro předpověď určuje výraz $(\alpha_1 + \beta_1)$ rychlost konvergence předpovědi k nepodmíněnému rozptylu. [18]

V R můžeme tvořit předpovědi GARCH modelu pomocí příkazu predict(), kdy do argumentu zadáme název objektu typu fGarch a zadáním argumentu n.ahead určíme počet předpovědí.

Pro řadu směnných kurzů CZK/EUR jsem vytvořila předpověď podmíněné směrodatné odchylky na 30 dní dopředu. Z obrázku č. 21 můžete vidět, že předpověď na základě GARCH(1,1) příliš nepokračuje ve vývoji odhadnuté směrodatné odchylky, což je

zřejmě způsobeno hodnotou součtu parametrů blízkou jedné. Předpověď se rychle přibližuje k nepodmíněné směrodatné odchylce, která je vyznačena červenou linií.



Obr. 21 Předpověď volatility. Rezidua, zeleně odhadnutá volatilita a od okamžiku vyznačeného vertikální přímkou předpověď, červeně nepodmíněná směrodatná odchylka

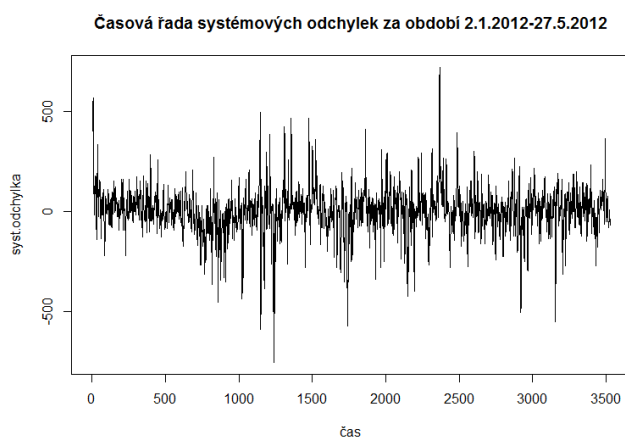
7. Aplikace modelů volatility na data nefinančního charakteru

V této části práce se budu snažit aplikovat teorii modelů volatility na časovou řadu nefinančního charakteru a zhodnotit, zda-li je použití modelu na místě či nikoliv. Budu pracovat s daty volně dostupnými na stránkách <http://www.ote-cr.cz/>. Použitá data i R-kód k této části práce jsou uloženy na cd pod názvy so.csv a OTE.R.

Jedná se o data zveřejněná akciovou společností OTE,a.s., která funguje za účelem organizování krátkodobého trhu s plynem a krátkodobého trhu s elektřinou a ve spolupráci s provozovatelem přenosové soustavy. Jedním z úkolů této společnosti, je vyhodnocovat odchylky odběru elektřiny a plynu na základě vyhodnocení zajistit zúčtování a vypořádání odchylek subjektů. Subjekty mají totiž smluvený odběr určitých MWh a je-li odběr nižší než smluvený, tedy odchylka záporná, znamená to pro distributory výdaje v podobě penále.

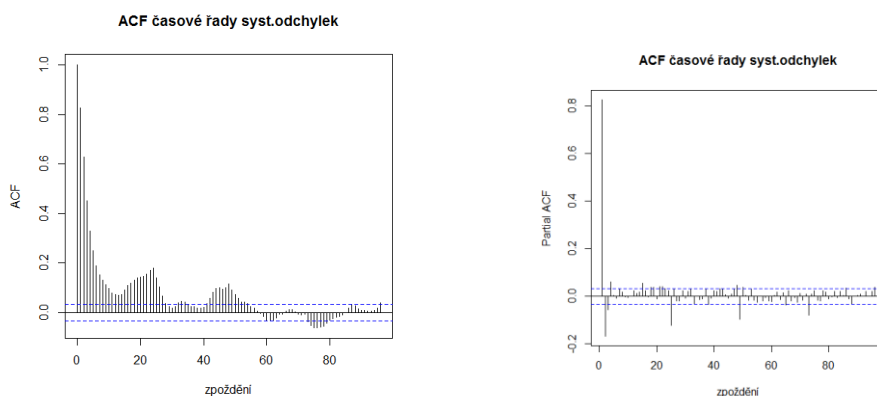
Konkrétně budu analyzovat časovou řadu systémových odchylek elektřiny v MWh za období 2.1.2012 – 27.5.2012. Časová řada je s hodinovou frekvencí pozorování o rozsahu 3528 údajů ze 147 dnů.

Průběh hodinové časové řady systematických odchylek můžete vidět na obr. 21.



Obr. 21 Časová řada systémových odchylek

Časová řada je dle Dickey-Fullerova testu (ADF) pro lag=5,10,15 stacionární.



Obr. 22 ACF,PACF časové řady systémových odchylek

Jak lze vidět z obr. 21, tak i ACF a PACF (viz. obr. 22) na první pohled vypovídají o sezónním charakteru časové řady. V bodech násobků 24 tyto funkce nabývají významnějších hodnot. Proto by se dalo uvažovat o stanovení úrovnového modelu pro sezónní časové řady SARIMA(p,d,q) x (P,D,Q)_L s periodou L=24. Parametry (p,d,q) odpovídají parametrům ARIMA modelu a slouží k modelování nesezónní složky a parametry (P,D,Q) odpovídají sezónnosti.

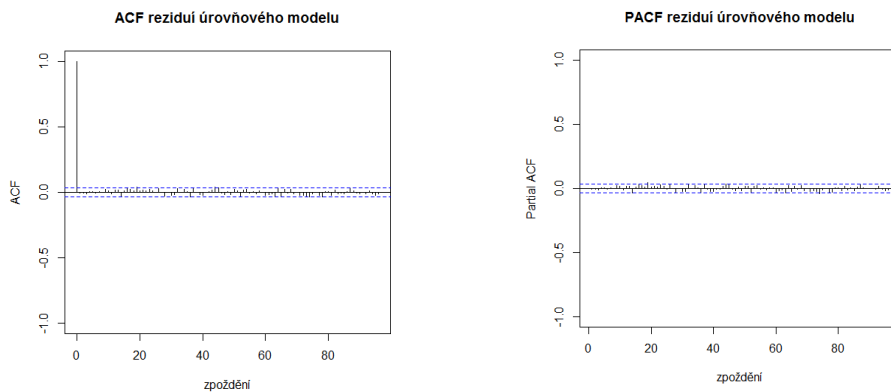
Modelovat budeme původní časovou řadu, tzn. d=0, jelikož je stacionární. Protože hodnoty funkcí ACF a PACF ve zpožděních 24,48,... nenabývají hodnoty blízké 1, stanovíme i D=0. Postupným přidáváním parametrů P, Q odstraníme závislost v sezónních zpožděních a následně volíme p, q. Díváme se přitom na významnost parametrů, počet parametrů, AIC.

Po mnoha pokusech byl zvolen model SARMA(1,0,2) x (1,0,1). Odhad modelu je uveden v tabulce č. 7. Ve třetím sloupci je vidět, že při přidání parametrů byly tyto parametry nevýznamné.

Model SARMA	Odhady parametrů (odhad směrodatné odchylky)		
(1,0,2)x(1,0,1)	ar1	0,71 (0,02)	(2,0,2)x(1,0,1) – ar2 nevýznamný
	ma1	0,28 (0,02)	(1,0,2)x(2,0,1) – sar2 nevýznamný
	ma2	0,15 (0,02)	(1,0,2)x(1,0,2) – sma2 nevýznamný
	sar1	0,90 (0,02)	
	sma1	0,80 (0,03)	

Tab. č. 7 Odhad modelu SARMA (1,0,2)x(1,0,1)

Na obr. 23 můžete vidět rezidua modelu a jejich ACF a PACF, kde lze vidět, že významné korelace byly eliminovány.



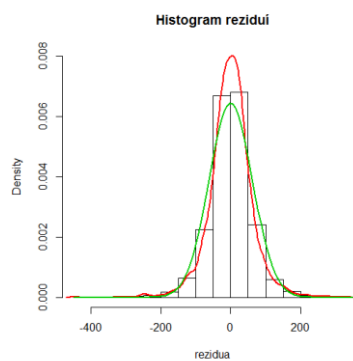
Obr. 23 ACF a PACF reziduí úrovněvého modelu.

Na rezidua úrovněvého modelu byly aplikovány Portmanteau test, Jarque-Berův test normality a ARCH test na přítomnost podmíněné heteroskedasticity. Výsledné p-hodnoty testů můžete vidět v tabulce č. 8.

Test	P-hodnota
Box-Pierce test	0.06231
Jarque-Bera test	$2,2 \cdot 10^{-16}$
ARCH-test	$2,2 \cdot 10^{-16}$

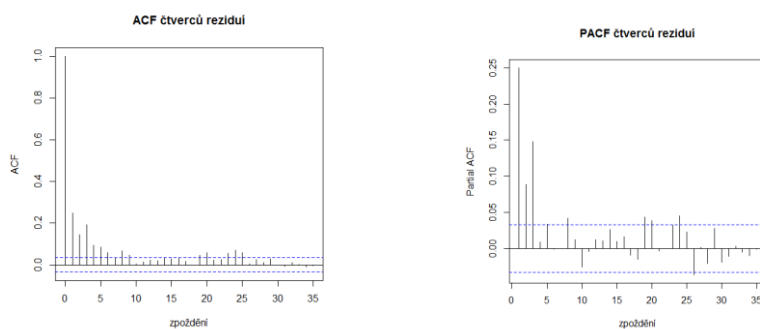
Tab. 8 P-hodnoty testů reziduí úrovněvého modelu.

Test hypotézy, že autokorelace jsou nevýznamné, ukazuje pro počet zpoždění zahrnutých do statistiky rovný odmocnině z délky řady reziduí (tj. lag=59), že nezamítáme hypotézu. P-hodnota testu normality svědčí pro nenormalitu reziduí, opět je tedy poukázáno na to, že rezidua úrovněvého modelu nepochází z normálního rozdělení viz. obr. 24, i když se jedná o vysokofrekvenční časovou řadu nefinančního charakteru. Rozdělení reziduí je špičatější, odhad špičatosti má hodnotu 6,53. Test na přítomnost podmíněné heteroskedasticity prokazuje přítomnost ARCH efektu pro počet zpoždění lag=1, viz tab. 8, ale i pro lag=5, 10, 20, 25.



Obr. 24 Histogram, zeleně hustota normálního rozdělení, červeně jádrový odhad hustoty.

Po ověření kvality úrovněového modelu přejdeme k identifikaci GARCH modelu dle ACF a PACF druhých mocnin reziduí.



Obr. 25 ACF, PACF druhých mocnin reziduí úrovněového modelu.

Při volbě modelu ARCH(1), ARCH(2) byly sice odhadnuté parametry významné, ale ARCH test standardizovaných reziduí prokázal, že podmíněná heteroskedasticita nebyla eliminována. Výsledky odhadu dalších modelů můžete vidět v tabulce č. 9.

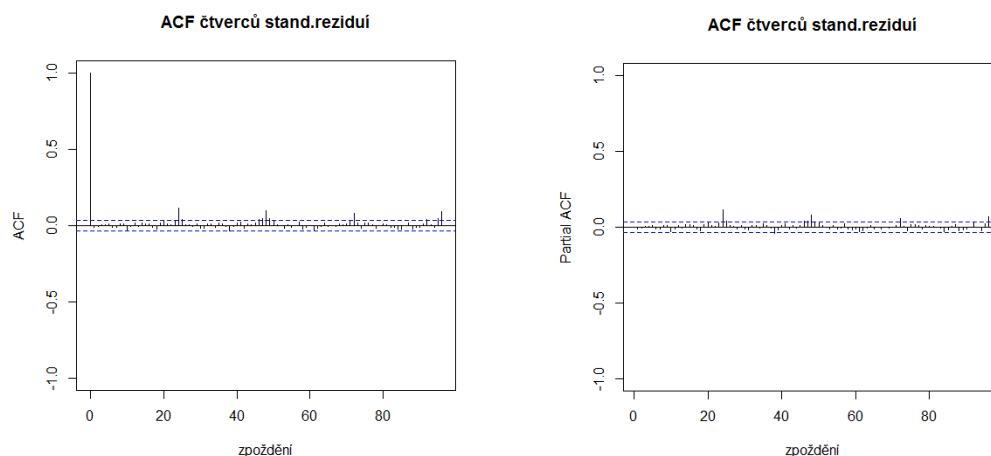
ARCH(3,0)	omega 1857,28 alfa 1 0,26 alfa 2 0,15 alfa 3 0,12	parametry významné AIC=10,95
GARCH(1,1)	omega 798,12 alfa 1 0,27 beta 1 0,53251	parametry významné AIC=10,94
GARCH(2,1)	omega 964,86 alfa 1 0,25 alfa 2 0,07 beta 1 0,44	parametry významné AIC=10,94

GARCH(1,2)	omega	795,7	Parametr beta 1 nevýznamný
	alfa 1	0,27	
	beta 1	0,53	
	beta 2	$1 \cdot 10^{-8}$	

Tab. 9 Odhady parametrů GARCH modelů.

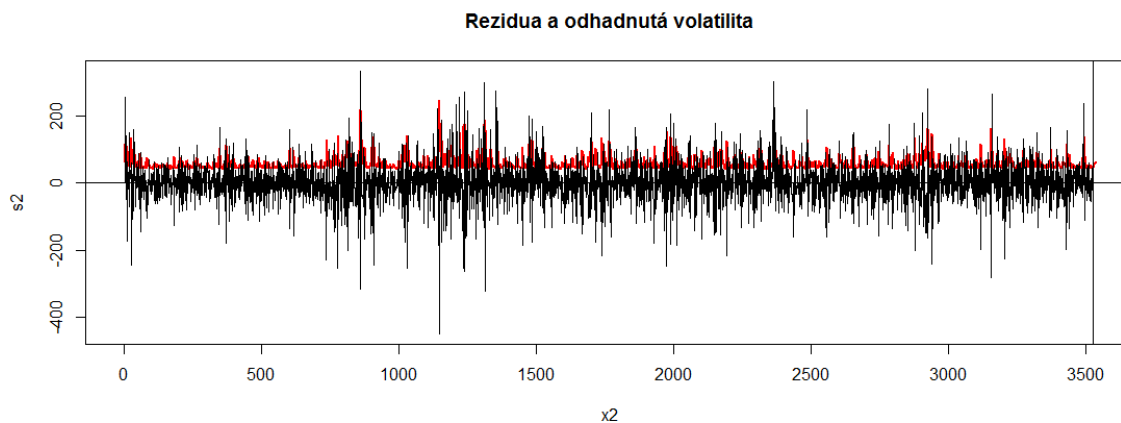
Na základě výsledků v tab. č. 9 jsem zvolila model GARCH(1,1). Oproti modelu ARCH(3) má menší počet parametrů a nižší hodnotu Akaikeho informačního kritéria. Model GARCH(2,1) nebyl zvolen i přestože parametry jsou významné, jelikož významnost parametru α_2 je příliš malá a zlepšení modelu není dle AIC.

Při verifikaci modelu svědčil Boxův test ve prospěch nevýznamnosti autokorelací standardizovaných reziduí. Dle testu podmíněné heteroskedasticity a ACF a PACF čtverců standardizovaných reziduí viz obr. 26 se však ukázalo, že podmíněnou heteroskedasticitu se nepodařilo odstranit úplně. Významnější hodnoty zůstávají v bodech násobků 24. Tato skutečnost by snad mohla být odstraněna pomocí modifikace modelu P-GARCH pro modelování sezónnosti ve volatilitě. O tomto modelu pojednává např. [20].



Obr. 26 ACF, PACF čtverců standardizovaných reziduí.

Pomocí aplikace sezónního modelu a modelu GARCH(1,1) se podařilo odhadnout volatilitu z dat viz. obr. 27, která celkem věrně kopíruje modelovaná rezidua s výjimkou odlehlých hodnot.



Obr. 27 Odhadnutá volatilita

Závěrem bych chtěla říct, že k vylepšení odhadu volatility by mohlo přispět nejen vyzkoušení aplikace modelu P-GARCH, přínosné by mohlo také být vyzkoušení některého z nelineárních modelů volatility. Analyzovaná data totiž nabývají kladných i záporných hodnot a je možné, že kladné a záporné šoky se promítají do podmíněného rozptylu různě, tzv. leverage efekt. Tato skutečnost nemůže být zachycena pomocí lineárních modelů, jelikož pracují s druhými mocninami šoků.

8. Závěr

Ve své práci jsem se snažila čtenáře seznámit s pojmem volatilita a způsoby modelování volatility. Byly představeny dva základní typy lineárních modelů ARCH a GARCH, u nichž byly ukázány charakteristické vlastnosti a odvozeny základní charakteristiky. Součástí teoretické části bylo také stručné popsání principu odhadu parametrů v těchto modelech a konstrukce předpovědí.

Cílem šesté kapitoly bylo názorně provést čtenáře postupem při výstavbě modelů volatility. K tomu bylo využito konkrétního příkladu zpracovaného v softwaru R, který se prolíná celou kapitolou.

Úkolem praktické části bylo pokusit se aplikovat GARCH model na časovou řadu z oblasti energetiky. Při určování úrovněvého modelu jsem se setkala s modely SARIMA. Identifikace vhodného SARIMA modelu, který by odstranil významné korelace bylo nesnadné, zejména kvůli mým nezkušenostem s modely tohoto typu, ale také časovou náročností odhadu parametrů v R. I když se podařilo získat řadu nekorelovaných reziduí potřebných k aplikaci GARCH modelu, nezdařilo se použitím tohoto modelu odstranit podmíněnou heteroskedasticitu úplně.

Psaní této práce bylo pro mne velice zajímavé. Způsoby modelování v čase proměnlivého rozptylu mě velice zaujaly. Proto je škoda, že vzhledem k rozsáhlosti teorie modelů volatility, se podařilo zahrnout do práce jen její malý zlomek. Zajímavou zkušeností pro mne byla také práce se sezónními modely SARIMA, se kterými jsem dříve nesetkala. Oceňuji také nové zkušenosti práce v softwaru R.

Použité zdroje

- [1] RUEY S. Tsay. *Analysis of financial time series* [online]. University of Chicago. Second edition. John Wiley & Sons, Inc.,2005. Dostupné online na http://cfa.goldenglobal.org.cn/uploadfile/append_file/%E8%B5%84%E6%96%99%E4%B8%8B%E8%BD%BD/CFA%E5%AD%A6%E4%B9%A0%E8%B5%84%E6%96%99/Analysis%20of%20Financial%20Time%20Series%202nd%20Edition.pdf, [citováno dne 9.12.2012].
- [2] CIPRA Tomáš. *Analýza časových řad s aplikacemi v ekonomii*. První vydání. Praha: SNTL – nakladatelství technické literatury,1986.
- [3] PAKOSTA Tomáš. *Analýza finančních časových řad typu ARCH*. Plzeň 2010. Bakalářská práce na Fakultě aplikovaných věd Západočeské univerzity na katedře Aplikovaných věd a informatiky. Vedoucí bakalářské práce RNDr. Blanka Šedivá Ph.D.
- [4] FEDDERKE Johannes. *Applied Time Series Analysis*. South Africa: University of Cape Town, 2003.
- [5] FARNSWORTH Grant V. *Econometrics in R* [online], dostupné na <http://cran.r-project.org/doc/contrib/Farnsworth-EconometricsInR.pdf>. [Citováno dne 4.9.2012].
- [6] ARLT Jiří, ARLTOVÁ Markéta. *Ekonomické časové řady*. 1.vyd. Příbram: Professional Publishing, 2009. ISBN 978-80-869446-85-6.
- [7] CIPRA Tomáš. *Finanční a pojistné vzorce*. První vydání. Praha : GRADA Publishing, a.s., 2006. ISBN 80-247-1633-X.
- [8] ARLT Josef, ARLTOVÁ Markéta. *Finanční časové řady*. První vydání. Praha: Grada Publishing a.s., 2003. ISBN 80-247-0330-0.
- [9] University of Washington. *Choosing the lag length for ADF Test* [online], dostupné na <http://faculty.washington.edu/ezivot/econ584/notes/unitrootLecture2.pdf>, [citováno dne 4.11.2012].

- [10] ŠTĚRBA Filip. *Modely typu ARCH a jejich využití k modelování volatility měnových kurzů*. [online] Česká bankovní asociace, Praha. Dostupné na <http://panda.hyperlink.cz/cestapdf/pdf07c3/sterba.pdf> .[Citováno dne 12.10.2012]
- [11] ARLT Josef. *Moderní metody modelování ekonomických časových řad*. 1.vyd. Praha: Grada Publishing, 1999. ISBN 80-7169-539-4.
- [12] Diethelm Wuertz et.al. *The fGarch Package*. [online], dostupné na <http://cran.r-project.org/web/packages/fGarch/fGarch.pdf>. [Citováno dne 8.11.2012]
- [13] ENGLE Robert, *The Use of ARCH/GARCH Models in Applied Econometrics*. [online]. Journal of Economic Perspectives, number 4, 2001. Dostupné na [http://www.stat.wharton.upenn.edu/~steele/Courses/434/434Context/GARCH/garch101\(ENGLE\).pdf](http://www.stat.wharton.upenn.edu/~steele/Courses/434/434Context/GARCH/garch101(ENGLE).pdf), [citováno dne 22.10.2012].
- [14] MILDE David. *Transformace dat*. [online] Dostupné na <http://ach.upol.cz/user-files/intranet/06-transformacedat-2011-1320420382.pdf>. [Citováno dne 4.12.2012]
- [15] POPELKA Jan. *Využití lineárních a nelineárních modelů volatility při analýze českých podílových fondů a akcií*. Praha 2007. Disertační práce na Fakultě informatiky a statistiky Vysoké školy ekonomické . Školitel : doc. Ing. Jiří Trešl, Csc.
- [16] WIKIPEDIA, The free Encyclopedia. 6 December 2012. [online]. Heteroscedasticity. Dostupné na <http://en.wikipedia.org/wiki/Heteroscedasticity>. [Citováno dne 28.9.2012].
- [17] WIKIPEDIA, Otevřená encyklopedie. Standard & Poor's. [online], 10. 10. 2012. Dostupné na http://cs.wikipedia.org/wiki/Standard_%26_Poor's, [citováno dne 4.9.2012].
- [18] REIDER ROB. *Volatility forecasting I: GARCH Models* [online].October,2009. Dostupné na http://cims.nyu.edu/~almgren/timeseries/Vol_Forecast1.pdf ,[citováno dne 24.11.2012].
- [19] KUNDEROVÁ Pavla, *Základy pravděpodobnosti a matematické statistiky*, 1. vydání, Univerzita Palackého v Olomouci, 2004

- [20] BESTER, C. Alan. *Seasonal Patterns in Futures Market Volatility A P-GARCH Approach*. [online] Duke University Durham, NC 1999. Dostupné na <http://public.econ.duke.edu/~get/browse/courses/200E/fall00/PROJECTS/previous/bester.pdf>, [citováno dne 13.12.2012].