

Univerzita Karlova v Praze
Matematicko-fyzikální fakulta

BAKALÁŘSKÁ PRÁCE



Jan Kubů

Testování hypotéz ve finančních časových řadách

Katedra pravděpodobnosti a matematické statistiky
Vedoucí bakalářské práce: RNDr. Jitka Zichová, Dr.

Studijní program: Matematika
Studijní obor: Finanční matematika
Praha 2012

Rád bych na tomto místě poděkoval vedoucí mé bakalářské práce, RNDr. Jitce Zichové, Dr., za cenné připomínky a podněty k této práci, poskytnutí potřebné literatury a její čas strávený na konzultacích.

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

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

V PRAZE DNE 13. 5. 2012

JAN KUBŮ

Název práce: Testování hypotéz ve finančních časových řadách

Autor: Jan Kubů

Katedra: Katedra pravděpodobnosti a matematické statistiky

Vedoucí bakalářské práce: RNDr. Jitka Zichová, Dr., Katedra pravděpodobnosti a matematické statistiky

e-mail vedoucího: zichova@karlin.mff.cuni.cz

Abstrakt: Finanční data mají často podobu časových řad, v takových případech se jejich analýza provádí za pomoci statistických metod pro časové řady. V práci jsou popsány vybrané parametrické a neparametrické testy hypotézy náhodné procházky. Testy jsou konstruovány proti obecným alternativám vzájemné korelovanosti hodnot, ale i proti alternativám trendu a alternativám cyklické struktury dat. Práce poskytuje teoretická východiska těchto testů a také jejich aplikace na reálná finanční data.

Klíčová slova: Časová řada, testování hypotéz, hypotéza náhodné procházky.

Title: Hypotheses Testing in Financial Time Series

Author: Jan Kubů

Department: Department of probability and mathematical statistics

Supervisor: RNDr. Jitka Zichová, Dr., Department of probability and mathematical statistics

Supervisor's e-mail address: zichova@karlin.mff.cuni.cz

Abstract: Financial data often take the form of time series. In such cases, their analysis is performed using statistical methods for time series. The thesis describes selected parametric and nonparametric tests of random walk hypothesis. Tests are designed against common mutual correlation alternatives but also against trend and cyclic data structure alternatives. The thesis provides the theoretical basis of these tests and their application to real financial data.

Keywords: Time series, hypothesis testing, random walk hypothesis.

Obsah

Úvod	8
1 Základní pojmy a metody	10
1.1 Časová řada	10
1.2 Korelační analýza	11
1.3 Testování hypotéz	12
2 Statistické testování	14
2.1 Hypotéza náhodné procházky	14
2.2 Odhad autokorelační funkce	15
2.3 Testování hypotézy náhodné procházky	17
2.3.1 Autokorelační testy	17
2.3.2 Spektrální testy	18
2.3.3 Neparаметrické testy	20
2.4 Trend v časové řadě	22
2.4.1 Spektrální hustota finanční časové řady s trendem	24
2.5 Testování hypotézy náhodné procházky proti alternativě trendu v časové řadě	25
2.5.1 Test „ T^* “ hypotézy náhodné procházky proti alternativě trendu .	25
2.5.2 Spektrální test „ f_0 “ hypotézy náhodné procházky proti alternativě trendu	27
3 Aplikace testů na reálná finanční data	28
3.1 Přepočet výnosů	28
3.2 Analyzovaná data	29
3.2.1 Index PX	29

3.2.2	Kurz CZK/USD	29
3.2.3	Simulace	29
3.3	Testy	30
3.3.1	Autokorelační testy	33
3.3.2	Spektrální analýza	33
3.3.3	Neparametrické testy	35
3.3.4	Test hypotézy náhodné procházky proti alternativě trendu v časové řadě	36
3.3.5	Shrnutí	36
	Literatura	38
	Seznam tabulek	39
	Seznam obrázků	40

Úvod

Finanční data mají často charakter časové řady. Jejich zkoumání v takových případech provádíme pomocí metod analýzy časových řad. Často nás zajímá existence nějaké zákonitosti, která řídí vývoj dat. Její nalezení může být užitečné například pro predikce budoucího vývoje, nebo při vysvětlování chování dat v minulosti. Stochastickým modelem může být rostoucí nebo klesající trend v časové řadě, cyklické opakování vývoje řady, případně jen vysoká míra korelace mezi hodnotami navzájem. Proti tomu stojí hypotéza náhodné procházky, která předpokládá, že se hodnoty řady chovají zcela náhodně a nijak na sobě vzájemně nezávisí. V práci se budeme zabývat právě testováním této hypotézy proti různým alternativám.

Testy jsou převzaty z knihy [9], kde je uveden jejich stručný přehled. V textu práce jsou provedena odvození vybraných vlastností testů, která v [9] nejsou zpracována. V případě složitějších důkazů odkazujeme na specializovanou časopiseckou literaturu.

V první kapitole uvedeme definice základních pojmů z oblasti časových řad, korelační analýzy a testování hypotéz, které pak budeme při testování používat. Ve druhé kapitole se zaměříme na teoretická východiska pro analýzu dat. Nejprve uvedeme definici hypotézy náhodné procházky v té podobě, kterou budeme testovat. Seznámíme se s bodovými odhady autokorelačních funkcí a jejich vlastnostmi, ze kterých pak vyjdeme při konstrukci autokorelačních, spektrálních a trendových testů. V sekci o autokorelačních testech odvodíme jak základní autokorelační test, zkoumající jen korelaci hodnot, které po sobě bezprostředně následují, tak i testy, které berou v úvahu více různých korelačních koeficientů. V sekci o spektrální analýze uvedeme definici spektrální hustoty a pomocí jejího odhadu sestavíme test hypotézy náhodné procházky proti hypotéze cyklů konkrétní periody, ale i komplexní test proti alternativě cyklů různých period.

Než se pustíme do testování alternativy trendu, uvedeme ještě jeden neparametrický test hypotézy náhodné procházky. V sekci o trendu v časové řadě uvedeme na jednom příkladě chování kovariance dvou hodnot časové řady s trendem. Poté v návaznosti na předchozí část o spektrální analýze odvodíme spektrální hustotu časové řady, která splňuje naši hypotézu trendu v časové řadě. Následně díky znalosti rozdělení logaritmického věrohodnostního poměru vektoru autokorelačních koeficientů za platnosti hypotézy trendu vůči hypotéze náhodné procházky zkonstruujeme test této hypotézy proti hypotéze trendu. Pak s využitím výsledků ze sekce o spektrální analýze odvodíme test proti hypotéze trendu v časové řadě o nulové frekvenci.

V poslední kapitole využijeme testy z druhé kapitoly k analýze dvou reálných finančních časových řad, která byla provedena s využitím softwaru Wolfram Mathematica 8.0. Jedna

časová řada je sestavena z hodnot kurzu české koruny k americkému dolaru za roky 2004-2011. Druhá časová řada je posloupnost hodnot burzovního indexu PX pražské burzy ve stejném časovém období. Pro srovnání jsme také simulovali řadu, která splňuje naši hypotézu náhodné procházky, a aplikovali naše testy také na ni. Výsledky všech testů jsou pak uvedeny ve shrnutí na konci práce.

Kapitola 1

Základní pojmy a metody

V této kapitole připomeneme pojmy a postupy související se statistickou analýzou časových řad, které budeme používat v dalším textu. Většina definic je uvedena podle knihy [2]. Symboly \mathbb{Z} , \mathbb{R} budeme značit množinu celých respektive reálných čísel.

1.1 Časová řada

Definice 1.1: *Stochastický proces*

Stochastický proces $\{Y_t, t \in T\}$ je množina náhodných veličin nebo vektorů na stejném pravděpodobnostním prostoru (Ω, \mathcal{A}, P) indexovaná pomocí hodnot t z množiny $T \subset \mathbb{R}$ interpretovaných jako čas.

Definice 1.2: *Časová řada*

Časová řada je posloupnost hodnot určité veličiny (nebo veličin v případě vícerozměrných časových řad) pozorovaná v určitém časovém intervalu s určitou frekvencí záznamu (každý obchodní den, v okamžicích transakcí, měsíčně apod.). Lze ji považovat za realizaci stochastického procesu s celočíselným časem.

Definice 1.3: *Stacionarita*

Stacionarita časové řady $\{X_t\}$, $t = 1, 2, \dots$ znamená, že chování této řady je v jistém smyslu stochasticky ustálené.

Striktní stacionarita znamená, že pravděpodobnostní chování příslušného stochastického procesu je invariantní vůči posunům v čase, tj. pravděpodobnostní rozdělení vektoru $(X_{t_1}, \dots, X_{t_k})$ je stejné jako rozdělení vektoru $(X_{t_1+h}, \dots, X_{t_k+h})$ pro libovolné reálné číslo h , přirozené číslo k a celá čísla t_1, \dots, t_k .

(Slabá) stacionarita je méně omezující než striktní stacionarita, neboť stačí, aby příslušný proces byl invariantní vůči posunům v čase pouze v rámci momentů do druhého řádu, tj. pro každé s a t

$$E(X_t) = \mu = konst; \quad (1.1a)$$

$$cov(X_s, X_t) = E(X_s - \mu)(X_t - \mu) = cov(X_{s+h}, X_{t+h}) \quad \forall h \in \mathbb{Z}, \quad (1.1b)$$

speciálně

$$var(X_t) = \sigma^2 = konst. \quad (1.1c)$$

Definice 1.4: *Bílý šum*

Bílý šum je posloupnost navzájem nekorelovaných náhodných veličin s nulovou střední hodnotou a konstantním kladným rozptylem.

Definice 1.5: *Striktní bílý šum*

Striktní bílý šum je posloupnost navzájem nezávislých náhodných veličin s nulovou střední hodnotou a konstantním kladným rozptylem.

Definice 1.6: *Lineární proces*

Lineární proces $\{X_t\}$, $t = 1, 2, \dots$ definujeme předpisem

$$X_t = \varepsilon_t + \psi_1 \varepsilon_{t-1} + \psi_2 \varepsilon_{t-2} + \dots = (1 + \psi_1 B + \psi_2 B^2 + \dots) \varepsilon_t = \psi(B) \varepsilon_t, \quad (1.2)$$

kde $\{\varepsilon_t\}$, $t = 1, 2, \dots$ je bílý šum, ψ_i , $i = 1, 2, \dots$ jsou parametry a B je operátor časového posunu. Lineární proces $\{X_t\}$, $t = 1, 2, \dots$ existuje právě tehdy, když řada na pravé straně vztahu (1.2) konverguje. Proto předpokládáme konvergenci mocninné řady $\psi(z) = \sum_{k=1}^{\infty} \psi_k z^k$ pro $|z| \leq 1$.

1.2 Korelační analýza

Definice 1.7: *Kovariance*

Kovariance náhodných veličin X a Y , pro které existují rozptyly, je

$$cov(X, Y) = E[(X - EX)(Y - EY)] = E(XY) - EXEY. \quad (1.3)$$

Definice 1.8: *Korelační koeficient*

Korelační koeficient náhodných veličin X a Y s konečnými kladnými rozptyly je

$$\rho_{X,Y} = \frac{cov(X, Y)}{\sqrt{var(X)}\sqrt{var(Y)}}. \quad (1.4)$$

Definice 1.9: *Autokovarianční funkce*

Autokovarianční funkce (autokovariance) časové řady $\{X_t\}$, $t = 1, 2, \dots$ pro zpoždění τ se definuje jako

$$\gamma_\tau = \text{cov}(X_t, X_{t+\tau}) = E(X_t - \mu)(X_{t+\tau} - \mu), \quad \tau = \dots, -1, 0, 1, \dots \quad (1.5)$$

Definice 1.10: *Autokorelační funkce*

Autokorelační funkce (autokorelace) časové řady $\{X_t\}$, $t = 1, 2, \dots$ pro zpoždění τ se definuje jako

$$\rho_\tau = \frac{\gamma_\tau}{\gamma_0} = \frac{\gamma_\tau}{\sigma^2} = \frac{\text{cov}(X_t, X_{t+\tau})}{\text{var}(X_t)}, \quad \tau = \dots, -1, 0, 1, \dots \quad (1.6)$$

1.3 Testování hypotéz

Definice 1.11: *Nulová hypotéza*

Nulová hypotéza označovaná jako H_0 je domněnka o určitém pravděpodobnostním modelu, která má být ověřena.

Definice 1.12: *Alternativní hypotéza*

Alternativní hypotéza označovaná jako H_1 zahrnuje zbývající možnosti v daném modelu.

Definice 1.13: *Testová statistika*

Testová statistika je vhodná funkce náhodných proměnných, jejíž pravděpodobnostní rozdělení za platnosti H_0 známe.

Definice 1.14: *Hladina významnosti testu*

Hladina významnosti testu je předem zvolená (malá) pravděpodobnost chyby prvního druhu. Je to pravděpodobnost, že test zamítne H_0 , přestože H_0 platí. Značíme ji α , volíme obvykle $\alpha = 0,05$.

Definice 1.15: *Kritický obor testu*

Kritický obor je množina hodnot testové statistiky, pro které zamítáme H_0 . Jeho konstrukce vychází z pravděpodobnostního rozdělení testové statistiky a ze zvolené hladiny významnosti testu.

Definice 1.16: *Statistický test*

Testová statistika spolu s příslušným kritickým oborem definují statistický test.

Definice 1.17: *Síla testu*

Doplňek pravděpodobnosti chyby druhého druhu, tedy pravděpodobnosti, že test nezamítne H_0 , když H_0 neplatí, do jedné, se nazývá síla testu. Jde tedy o pravděpodobnost, že test zamítne H_0 , pokud H_0 neplatí.

Kapitola 2

Statistické testování

V této kapitole se věnujeme hypotéze náhodné procházky a různým testům pro její ověření. Vycházíme z knihy [9].

2.1 Hypotéza náhodné procházky

Ve finanční analýze se zpracovávají například časové řady výnosů z cenných papírů, kurzů cenných papírů, měnových kurzů apod. Taková data nelze považovat za realizace nezávislých stejně rozdělených náhodných veličin. Testují se proto obecnější hypotézy o charakteru pravděpodobnostního modelu, z něhož data pocházejí.

Jedna z možností je nahradit stejné rozdělení konstantní střední hodnotou a nezávislost nekorelovaností. V tomto smyslu zavedeme definici hypotézy náhodné procházky.

Definice 2.1: *Hypotéza náhodné procházky pro časovou řadu X_t , $t = 1, 2, \dots$*

má tvar

$$H_0 : EX_t = EX_{t+\tau}, \quad cov(X_t, X_{t+\tau}) = 0$$

pro všechna t a $\tau > 0$.

Předpoklad nekorelovanosti zajišťuje, že hodnoty řady pozorované před časem t nikterak neovlivňují předpovědi do budoucna, konstruované v čase t . Hypotéza H_0 nepožaduje stacionaritu procesu X_t , rozptyl X_t se může v čase měnit. Uvedená definice hypotézy náhodné procházky byla použita např. v práci [3].

Pro testování hypotézy náhodné procházky byly navrženy různé přístupy související mimo jiné s různými typy alternativ. Alternativními hypotézami mohou být například hypotéza (cenového) trendu nebo hypotéza cykličnosti dat, které budou formulovány v dalším textu.

Máme-li k dispozici více testů hypotézy náhodné procházky, může se stát, že pro stejná data dojde u některých testů k zamítnutí nulové hypotézy a u některých nikoli. V principu se tento problém řeší tak, že ještě před testováním odhadneme nejpravděpodobnější

alternativní hypotézu a následně najdeme test, který má proti této alternativě velkou sílu. Výsledek tohoto testu pak akceptujeme jako správný.

2.2 Odhad autokorelační funkce

K testování hypotézy náhodné procházky potřebujeme znát odhad autokorelační funkce a jeho vlastnosti. Pro danou stacionární časovou řadu délky n se používají následující odhady.

Definice 2.2: *Odhad autokovarianční funkce*

$$C_\tau = \frac{1}{n} \sum_{t=1}^{n-\tau} (X_t - \bar{X})(X_{t+\tau} - \bar{X}) \quad \tau = 0, 1, \dots, n-1 \quad (2.1)$$

kde

$$\bar{X} = \frac{1}{n} \sum_{t=1}^n X_t \quad (2.2)$$

je odhad střední hodnoty.

Definice 2.3: *Odhad autokorelační funkce*

$$R_\tau = \sum_{t=1}^{n-\tau} (X_t - \bar{X})(X_{t+\tau} - \bar{X}) / \sum_{t=1}^n (X_t - \bar{X})^2 \quad \tau = 0, 1, \dots, n-1 \quad (2.3)$$

Nahrazením náhodné veličiny X_t pozorovanou hodnotou x_t získáme číselnou hodnotu odhadnuté autokorelace, kterou budeme značit r_τ .

Rozdělení odhadu autokorelační funkce

Při hledání rozdělení odhadu autokorelační funkce vyjdeme z asymptotických výsledků pro $n \rightarrow \infty$ v případě autokorelovaného lineárního procesu. Předpokládáme, že X_t má konečný rozptyl, ale ne nutně normální rozdělení. Finanční časová řada nemusí být realizací lineárně generovaného procesu. Výsledky pro lineární proces budou však užitečné při zpracování řad, které byly přepočteny tak, aby se proces stal přibližně lineárním.

Tvrzení 2.4:

Nechť X_t je lineární proces popsáný v definici 1.6 a ρ_τ korelační koeficient veličin $X_t, X_{t+\tau}$ $\tau > 0$. Pak má náhodná veličina $\sqrt{n}(R_\tau - \rho_\tau)$ asymptotické rozdělení $N(0, \omega_{\tau\tau})$, kde

$$\omega_{\tau\tau} = \sum_{v=-\infty}^{\infty} (\rho_v^2 + \rho_v \rho_{v+2\tau} + 2\rho_\tau^2 \rho_v^2 - 4\rho_\tau \rho_v \rho_{v+\tau}). \quad (2.4)$$

Pro $\tau \neq \xi$ je asymptotická kovariance veličin $\sqrt{n}R_\tau$ a $\sqrt{n}R_\xi$ dána předpisem

$$\omega_{\tau\xi} = \sum_{v=-\infty}^{\infty} (\rho_v \rho_{v+\xi-\tau} + \rho_v \rho_{v+\tau+\xi} + 2\rho_\tau \rho_\xi \rho_v^2 - 2\rho_\tau \rho_v \rho_{v+\xi} - 2\rho_\xi \rho_v \rho_{v+\tau}). \quad (2.5)$$

Náhodný vektor $\sqrt{n}(R_1 - \rho_1, \dots, R_k - \rho_k)$ $k > 1$ má asymptoticky mnohorozměrné normální rozdělení s výše uvedenými momenty.

Důkaz:

Lze najít ve článku [1]. □

Tvrzení 2.5:

Nechť X_t je posloupnost nezávislých stejně rozdělených náhodných veličin. Pak má náhodná veličina $\sqrt{n}R_\tau$, $\tau > 0$, asymptoticky rozdělení $N(0, 1)$ a pro $\tau \neq \xi$ jsou veličiny R_τ a R_ξ asymptoticky nezávislé.

Důkaz:

Podle tvrzení 2.4 má $\sqrt{n}(R_\tau - \rho_\tau)$ asymptotické rozdělení $N(0, \omega_{\tau\tau})$. Z nezávislosti veličin X_t plyne $\rho_\tau = 0$ pro $\tau > 0$ a tedy $\sqrt{n}R_\tau$ má asymptotické rozdělení $N(0, \omega_{\tau\tau})$. Dle vztahu (2.4) je pro $\tau > 0$ $\omega_{\tau\tau} = \rho_0^2 = 1$ a $\sqrt{n}R_\tau$ má asymptotické rozdělení $N(0, 1)$.

Pro $\tau \neq \xi$ mají podle tvrzení 2.4 veličiny $\sqrt{n}R_\tau$ a $\sqrt{n}R_\xi$ asymptoticky mnohorozměrné normální rozdělení $N_k(\mathbf{0}, \mathbf{\Omega}_k)$. $\mathbf{\Omega}_k$ je jednotkovou maticí, protože její prvky $\omega_{\tau\xi}$ jsou podle (2.5) a vzhledem k nezávislosti veličin X_t nulové až na diagonální prvky, které jsou rovny 1 podle předcházející části důkazu. Protože u mnohorozměrného normálního rozdělení implikuje nekorelovanost veličin jejich nezávislost, jsou R_τ a R_ξ asymptoticky nezávislé. □

Nyní se podívejme na situaci u časových řad konečné délky n . Odhady autokorelační funkce nejsou neustranné, mají vychýlení řádu $\frac{1}{n}$.

Tvrzení 2.6:

Nechť X_t je posloupnost nezávislých stejně rozdělených náhodných veličin. Pak platí

$$ER_\tau = -\frac{n - \tau}{n(n - 1)}, \quad (2.6a)$$

$$\text{var}R_\tau < \frac{1}{n}, \quad 0 < \tau < n, \quad (2.6b)$$

$$\text{cov}(R_\tau, R_\xi) < 0, \quad 0 < 2\tau < 2\xi \leq n. \quad (2.6c)$$

Důkaz:

Je uveden ve článku [6]. □

Vychýlení odhadu R_τ v případě lineárního procesu nemá triviální vyjádření. Informace lze najít v článku [4].

2.3 Testování hypotézy náhodné procházky

2.3.1 Autokorelační testy

Autokorelační test „ r_1 “

Populární a jednoduchý test používá jen první autokorelační koeficient. H_0 je zamítnuta na hladině 5% pokud

$$\sqrt{n}|r_1| > u(0, 975) = 1,96, \quad (2.7)$$

kde dolní mez je 97,5%-ní kvantil rozdělení $N(0, 1)$. Test je založen na tvrzení 2.5. Lze jej použít, pokud jediná předpokládaná závislost je mezi hodnotami řady, které následují bezprostředně po sobě.

Autokorelační test „ N_r “

Pokud uvažujeme více koeficientů r_τ , $\tau = 1, 2, \dots$, některé mohou být statisticky významné i za platnosti H_0 . Mějme zjištěny autokorelace r_1, \dots, r_k a označme

$$N_r = \sum_{\tau=1}^k \mathbb{I}(\sqrt{n}|r_\tau| > 1,96), \quad (2.8)$$

kde $\mathbb{I}(\sqrt{n}|r_\tau| > 1,96)$ je 1, pokud je nerovnost splněna, a 0 jinak.

Za platnosti H_0 se počet statisticky významných výběrových autokorelací N_r řídí binomickým rozdělením s parametry k a $p = P(\sqrt{n}|r_1| > 1,96) = 0,05$. Najdeme-li k danému počtu k číslo N_0 takové, že

$$\alpha(k, N_0) = \sum_{i=N_0}^k \binom{k}{i} 0,05^i * 0,95^{k-i} \simeq 0,05, \quad (2.9)$$

je kritický obor testu H_0 určen nerovností $N_r \geq N_0$, tedy zamítáme H_0 na hladině přibližně 5%, pokud $N_r \geq N_0$. Uvedme některé možné volby k a N_0 :

k	17	28	67	81
N_0	3	4	7	8
$\alpha(k, N_0)$	0,0503	0,0491	0,0497	0,0495

Tabulka 2.1: Některé přípustné hodnoty parametrů N_0, k

Autokorelační test „ Q_k “

Přirozenější způsob jak použít k korelačních koeficientů v jedné testové statistice je následující.

Označme

$$Q_k = n \sum_{\tau=1}^k r_{\tau}^2 \quad (2.10)$$

Za platnosti H_0 nastává na základě tvrzení 2.5

$$n(R_1^2 + \dots + R_k^2) \sim \chi_k^2. \quad (2.11)$$

Na 5%-ní hladině zamítáme H_0 pro $Q_k \geq \chi_k^2(0,95)$, kde dolní mez je 95%-ní kvantil χ^2 rozdělení o k stupních volnosti.

2.3.2 Spektrální testy

Ani jeden z předcházejících testů nebyl konstruován proti specifické alternativní hypotéze. Spektrální analýza je dalším možným přístupem vedle studia autokorelací. Je vhodná pro testování hypotézy náhodné procházky proti alternativě cyklického chování časové řady.

Model lze psát ve tvaru

$$X_t = \sum_{j=1}^J \alpha_j \cos(\omega_j t - \beta_j) + \epsilon_t, \quad (2.12)$$

kde α_j, β_j a ω_j jsou konstanty a ϵ_t jsou nekorelované náhodné veličiny. Cyklus j má potom frekvenci ω_j a periodu $\frac{2\pi}{\omega_j}$ časových jednotek.

Definice 2.7: Spektrální hustota

je funkce

$$s(\omega) = \sigma^2/(2\pi)[1 + 2 \sum_{\tau=1}^{\infty} \rho_{\tau} \cos(\tau\omega)] \quad 0 \leq \omega \leq 2\pi \quad (2.13)$$

kde $\sigma^2 = \text{var}(X_t)$.

Platí

$$\int_0^{2\pi} s(\omega) d\omega = \sigma^2, \quad (2.14)$$

$$s(\omega) = s(2\pi - \omega). \quad (2.15)$$

Proto stačí uvažovat pouze frekvence od 0 do π . V cyklickém modelu bude mít graf funkce $s(\omega)$ vrcholy s frekvencí ω_j . Pokud však bude platit hypotéza náhodné procházky, $s(\omega)$ bude konstantní pro všechna ω_j . K testování H_0 potřebujeme odhad funkce $s(\omega)$.

Odhadem pro

$$f(\omega) = 2\pi s(\omega)/\sigma^2 \quad (2.16)$$

je

$$\hat{f}(\omega) = 1 + 2 \sum_{\tau=1}^{M-1} \lambda_{\tau} r_{\tau} \cos(\tau\omega) \quad (2.17)$$

s kladným a monotonně rostoucím λ_{τ} zajišťujícím konzistentní odhady. Tzv. Parzenovy váhy jsou pro fixní M definovány předpisem

$$\lambda_{\tau} = 1 - 6\tau^2(M - \tau)/M^3 \quad 0 < \tau \leq M/2 \quad (2.18a)$$

$$\lambda_{\tau} = 2(M - \tau)^3/M^3 \quad M/2 \leq \tau < M. \quad (2.18b)$$

Rovnice (2.17) ukazuje, že spektrální odhady jsou lineárními funkcemi prvních $M-1$ autokorelací. V článku [7] je ukázáno, že odhady $\hat{f}(\omega_1)$ a $\hat{f}(\omega_2)$ jsou korelované právě tehdy, když

$$|\omega_1 - \omega_2| \leq \frac{3\pi}{M}. \quad (2.19)$$

Uvažujme M , které je násobkem 4. Testy jsou založeny na odhadech (2.17) v bodech $\omega = 0, 4\pi/M, 8\pi/M, \dots, \pi = M\pi/M$. Je možné standardizovat tyto odhady za použití asymptotické teorie pro výběrové autokorelace. Získáme

$$f_j = [\hat{f}(4j\pi/M) - 1] / \sqrt{\{4 \sum_{\tau=1}^{M-1} [\lambda_{\tau} \cos(4j\tau\pi/M)]^2/n\}} \quad (2.20)$$

$j = 0, 1, \dots, M/4$.

Spektrální test „ f_w “ pro týdenní cykly

Za platnosti H_0 jsou f_j nezávislé s rozdělením $N(0, 1)$ pro velká n . Velmi častými jsou týdenní cykly, kdy je $\omega = 2\pi/5$ a klademe $j = M/10$. Testovou statistiku (2.20) pro $j = M/10$ označíme

$$f_w = [\hat{f}(\frac{2}{5}\pi) - 1] / \sqrt{\{4 \sum_{\tau=1}^{M-1} [\lambda_{\tau} \cos(\frac{2}{5}\pi\tau)]^2 / n\}} \quad (2.21)$$

V knize [9] je doporučen jednostranný test, kde na pětiprocentní hladině zamítáme H_0 , když $f_w > u(0, 95) = 1,65$, kde dolní mez je 95%-ní kvantil $N(0, 1)$.

Spektrální test „ N_s “

Někteří autoři, jako například v článku [7], sčítají statisticky významné koeficienty f_j . Obdobně jako u autokorelačního testu „ N_r “ označíme

$$N_s = \sum_{j=0}^{M/4} \mathbb{I}(|f_j| > 1,96). \quad (2.22)$$

Testová statistika N_s má za platnosti H_0 binomické rozdělení. V analogii se vztahem (2.9) najdeme k pevně zvolenému M takové číslo N_{s0} , že platí pro $k = M/4$

$$\sum_{i=N_0}^k \binom{k}{i} 0,05^i * 0,95^{k-i} \simeq 0,05,$$

H_0 zamítáme v případě, že je $N_s \geq N_{s0}$. Například pro $M = 100$ je nejlepší zvolit $N_{s0} = 4$. Pro $M = 200$ je nejlepší $N_{s0} = 6$.

2.3.3 Neparametrické testy

Protože finanční časové řady nemají normální a často ani stacionární rozdělení, je vhodné použití neparametrických testů. Jejich nejpoužívanějším zástupcem je test založený na iteracích.

Iterační test

Definice 2.8: Iterace

Kladná iterace

je sekvence po sobě jdoucích kladných hodnot.

Nulová iterace

je sekvence po sobě jdoucích nulových hodnot.

Záporná iterace

je sekvence po sobě jdoucích záporných hodnot.

Definice 2.9: *Znaménko* x_t^*

Znaménko hodnoty časové řady definujeme předpisem

$$\begin{array}{lll} x_t^* = 1 & \text{pro} & x_t > 0, \\ x_t^* = 0 & \text{pro} & x_t = 0, \\ x_t^* = -1 & \text{pro} & x_t < 0. \end{array}$$

Definice 2.10: *Změna znaménka* h_t

Změna znaménka v časové řadě je

$$\begin{array}{lll} h_t = 0 & \text{pro} & x_t^* = x_{t+1}^*, \\ h_t = 1 & & \text{jinak.} \end{array}$$

Je-li $h_t = 1$, značí to novou iteraci. Proto

$$H = 1 + \sum_{t=1}^{n-1} h_t \quad (2.23)$$

je počet všech iterací časové řady. Je to náhodná veličina, pro kterou platí následující tvrzení.

Tvrzení 2.11:

Nechť n_1 je počet kladných hodnot časové řady, n_2 počet nulových hodnot a n_3 počet záporných hodnot. Předpokládejme platnost hypotézy H_0^* : X_t^* je striktní bílý šum s realizacemi x_t^* .

Pak pro náhodnou veličinu H platí při pevných n_1, n_2, n_3

$$EH = n + 1 - \sum_{j=1}^3 n_j^2/n \quad (2.24)$$

a

$$\text{var} H = \left\{ \sum_{i=1}^3 n_i^2 \left(\sum_{j=1}^3 n_j^2 + n + n^2 \right) - 2n \sum_{j=1}^3 n_j^3 - n^3 \right\} / (n^3 - n). \quad (2.25)$$

Důkaz:

Viz práce [5].

□

Hypotéza H_0^* se testuje namísto původní hypotézy H_0 . Pro velká n je náhodná veličina H přibližně normálně rozdělená. Testová statistika tedy bude

$$K = (H - EH) / \sqrt{\text{var}H}. \quad (2.26)$$

H_0^* zamítáme na hladině 5%, pokud je $|K| > 1,96$. Záporné hodnoty K odpovídají malému počtu změn ve znaménku u řad s trendem. Kladná K budou v řadách s častým kolísáním hodnot. Pokud je alternativou pouze jedna ze zmíněných možností, provádíme jednostranný test. Nevýhodou testu je malá síla, způsobená ztrátou informace při přechodu od x_t k x_t^* .

Další neparametrické testy jsou popsány v knize [2].

2.4 Trend v časové řadě

Idea trendů v časové řadě je specifickou alternativou k hypotéze náhodné procházky. Trend je celkový pohyb pevným směrem nahoru nebo dolů. Existence trendů znamená, že hodnoty řady plně a okamžitě neodrážejí nové informace. Trendy způsobují pozitivní autokorelaci, která by ale s rostoucím rozestupem mezi hodnotami řady měla klesat. Tedy

$$\rho_\tau > 0, \rho_\tau > \rho_{\tau+1}, 1 \gg \rho_\tau, \forall \tau > 0. \quad (2.27)$$

Parametrické vyjádření korelací nás vede k definici hypotézy trendu.

Definice 2.12: *Hypotéza trendu v časové řadě*

$$H_1 : \rho_\tau = Ap^\tau, \quad A > 0, 0 < p < 1, \tau > 0. \quad (2.28)$$

Volby A a p doporučené v knize [9] jsou $A = 0,03$ a $p = 0,95$.

Příklad 2.13:

Uvažujme situaci, kdy je náhodná veličina X_t součtem autoregresní trendové komponenty μ_t a nepředpovědatelného rezidua e_t .

$$X_t = \mu_t + e_t, \quad (2.29a)$$

$$\mu_t - \mu = p(\mu_{t-1} - \mu) + \eta_t. \quad (2.29b)$$

Předpokládáme, že

$$E[e_t] = E[\eta_t] = E[e_t e_{t+\tau}] = E[\eta_t \eta_{t+\tau}] = E[e_t \eta_s] = 0, \quad s, t, \tau > 0. \quad (2.30)$$

Z (2.29b) plyne

$$\mu_t - \mu = \sum_{k=0}^{\infty} p^k \eta_{t-k} \quad (2.31a)$$

a

$$\mu_{t+\tau} = \mu + p^\tau (\mu_t - \mu) + \sum_{k=0}^{\tau-1} p^k \eta_{t+\tau-k} \quad (2.31b)$$

pro $\tau > 0$. Potom

$$\text{cov}(X_t, X_{t+\tau}) = E[X_t X_{t+\tau}] - E[X_t]E[X_{t+\tau}] =$$

dle (2.29a)

$$\begin{aligned} &= E[(\mu_t + e_t)(\mu_{t+\tau} + e_{t+\tau})] - E[\mu_t + e_t]E[\mu_{t+\tau} + e_{t+\tau}] = \\ &= E[\mu_t \mu_{t+\tau} + \mu_t e_{t+\tau} + e_t \mu_{t+\tau} + e_t e_{t+\tau}] - \\ &\quad - (E[\mu_t] + E[e_t])(E[\mu_{t+\tau}] + E[e_{t+\tau}]) = \\ &= E[\mu_t \mu_{t+\tau}] + E[\mu_t e_{t+\tau}] + E[e_t \mu_{t+\tau}] + E[e_t e_{t+\tau}] - \\ &\quad - E[\mu_t]E[\mu_{t+\tau}] - E[\mu_t]E[e_{t+\tau}] - E[e_t]E[\mu_{t+\tau}] - \\ &\quad - E[e_t]E[e_{t+\tau}] = \end{aligned}$$

dle (2.31a) a (2.30)

$$\begin{aligned} &= E[\mu_t \mu_{t+\tau}] - E[\mu_t]E[\mu_{t+\tau}] = \\ &= \text{cov}(\mu_t, \mu_{t+\tau}) = \end{aligned}$$

dle (2.31b)

$$\begin{aligned} &= \text{cov}(\mu_t, (1 - p^\tau)\mu + p^\tau \mu_t + \sum_{k=0}^{\tau-1} p^k \eta_{t+\tau-k}) = \\ &= E[(\mu_t(1 - p^\tau)\mu + E[p^\tau \mu_t^2] + E[\mu_t \sum_{k=0}^{\tau-1} p^k \eta_{t+\tau-k}]) - \\ &\quad - E[\mu_t]E[(1 - p^\tau)\mu + p^\tau \mu_t + \sum_{k=0}^{\tau-1} p^k \eta_{t+\tau-k}]] = \end{aligned}$$

dle (2.31a) a (2.30)

$$\begin{aligned} &= p^\tau [E[\mu_t^2] - (E[\mu_t])^2] = \\ &= p^\tau \text{var}(\mu_t) \quad \forall \tau > 0. \end{aligned} \quad (2.32)$$

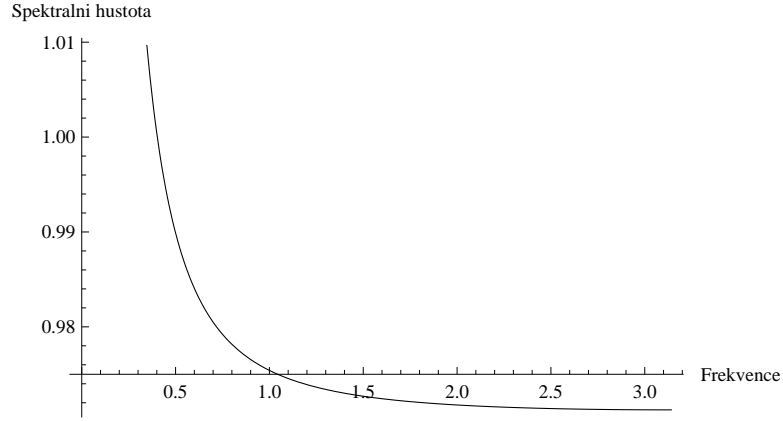
Tedy H_1 platí a vzhledem k (2.28), (1.6) a (2.32)

$$\begin{aligned} A &= \frac{\rho_\tau}{p^\tau} = \frac{\text{cov}(X_t, X_{t+\tau})}{\text{var}(X_t)} \frac{\text{var}(\mu_t)}{\text{cov}(X_t, X_{t+\tau})} = \\ &= \frac{\text{var}(\mu_t)}{\text{var}(X_t)}. \end{aligned} \quad (2.33)$$

2.4.1 Spektrální hustota finanční časové řady s trendem

Odvodíme ještě funkci spektrální hustoty pro hypotézu trendu v časové řadě při parametrickém vyjádření autokorelací (2.28). Značením $\Re[x]$ zde rozumíme reálnou část komplexního čísla x . Dle (2.16) a (2.13) lze psát

$$\begin{aligned} f(\omega) &= 1 + 2 \sum_{\tau=1}^{\infty} \rho_\tau \cos(\tau\omega) = \\ &= 1 + 2A \sum_{\tau=1}^{\infty} p^\tau \cos(\tau\omega) = \\ &= 1 + 2A \Re \left[\sum_{\tau=1}^{\infty} p^\tau (\cos(\tau\omega) + i \sin(\tau\omega)) \right] = \\ &= 1 + 2A \Re \left[\sum_{\tau=1}^{\infty} p^\tau e^{i\tau\omega} \right] = \\ &= 1 + 2A \Re \left[\sum_{\tau=1}^{\infty} (pe^{i\omega})^\tau \right] = \\ &= 1 + 2A \Re \left[\frac{pe^{i\omega}}{1 - pe^{i\omega}} \right] = \\ &= 1 + 2A \Re \left[\frac{p(\cos(\omega) + i \sin(\omega))}{1 - p(\cos(\omega) + i \sin(\omega))} \frac{1 - p(\cos(\omega) - i \sin(\omega))}{1 - p(\cos(\omega) - i \sin(\omega))} \right] = \\ &= 1 + 2A \frac{p \cos(\omega) - p^2 \cos^2(\omega) - p^2 \sin^2(\omega)}{1 - 2p \cos(\omega) + p^2 \cos^2(\omega) + p^2 \sin^2(\omega)} = \\ &= 1 + 2A \frac{p \cos(\omega) - p^2}{1 - 2p \cos(\omega) + p^2} = \\ &= 1 + A \frac{2(p \cos(\omega) - p^2) - (1 - 2p \cos(\omega) + p^2) + (1 - 2p \cos(\omega) + p^2)}{1 - 2p \cos(\omega) + p^2} = \\ &= 1 - A + \frac{A(1 - p^2)}{1 - 2p \cos(\omega) + p^2} \quad 0 \leq \omega \leq 2\pi. \end{aligned} \quad (2.34)$$



Obrázek 2.1: Spektrální hustota finanční časové řady s trendem pro $A = 0,03$ a $p = 0,95$

2.5 Testování hypotézy náhodné procházky proti alternativě trendu v časové řadě

Testy hypotézy náhodné procházky proti alternativě trendu jsou konstruovány s použitím teoretických rozdělení odhadu autokorelační funkce. Uvažujme k odhadů autokorelačních funkcí $\mathbf{R} = (R_1, R_2, \dots, R_k)$. Podle tvrzení 2.4 má náhodný vektor $\sqrt{n}(\mathbf{R} - \boldsymbol{\rho})$ asymptoticky k -rozměrné normální rozdělení $N_k(\mathbf{0}, \boldsymbol{\Omega}_k)$. Přitom za platnosti H_0 máme $\boldsymbol{\rho} = (0, 0, \dots, 0)$ a $\boldsymbol{\Omega}_k = \mathbf{I}_k$ je k -rozměrná jednotková matice. Naopak za platnosti H_1 je $\boldsymbol{\rho} = (Ap, Ap^2, \dots, Ap^k)$ a $\boldsymbol{\Omega}_k \simeq \mathbf{I}_k$, pokud je A malé. Podrobněji je problematika pojednána v článku [8].

2.5.1 Test „ T^* “ hypotézy náhodné procházky proti alternativě trendu

Alternativa H_1 má dva nespecifikované parametry A a p . Předpokládejme, že parametry jsou známé, označme je A^* a p^* . Uvažujme A^* malé a $\boldsymbol{\Omega}_k \simeq \mathbf{I}_k$. Alternativa tedy je $H_1^* : \rho_\tau = A^*p^{*\tau}$. Optimální test H_0 proti H_1^* využívá logaritmický věrohodnostní poměr

$$I = \log\{L(\mathbf{R}|H_1^*)/L(\mathbf{R}|H_0)\}, \quad (2.35)$$

kde $L(\mathbf{R}|\cdot)$ značí věrohodnostní funkci vektoru \mathbf{R} . S využitím předpokladu asymptotické normality získáváme

$$\begin{aligned} I &= \log L(\mathbf{R}|H_1^*) - \log L(\mathbf{R}|H_0) \simeq \\ &\simeq \log \prod_{\tau=1}^k \frac{\sqrt{n}}{\sqrt{2\pi}} \exp\left[-\frac{n}{2}(R_\tau - \rho_\tau)^2\right] - \log \prod_{\tau=1}^k \frac{\sqrt{n}}{\sqrt{2\pi}} \exp\left[-\frac{n}{2}R_\tau^2\right] = \\ &= \sum_{\tau=1}^k \left[\log \frac{\sqrt{n}}{\sqrt{2\pi}} - \frac{n}{2}(R_\tau - \rho_\tau)^2\right] - \sum_{\tau=1}^k \left[\log \frac{\sqrt{n}}{\sqrt{2\pi}} - \frac{n}{2}R_\tau^2\right] = \\ &= \frac{n}{2} \sum_{\tau=1}^k (2R_\tau\rho_\tau - \rho_\tau^2) = n \sum_{\tau=1}^k A^*p^{*\tau} R_\tau - \frac{n}{2} \sum_{\tau=1}^k A^{*2}p^{*2\tau} \Rightarrow \end{aligned}$$

$$\Rightarrow I \simeq nA^* \sum_{\tau=1}^k p^{*\tau} R_{\tau} + I_0, \quad (2.36)$$

kde I_0 je konstanta, nezávislá na \mathbf{R} . Při konkrétní hodnotě $\mathbf{R} = \mathbf{r}$ je tedy H_0 zamítána, pokud

$$T_{k,p^*} = \sum_{\tau=1}^k p^{*\tau} r_{\tau} \quad (2.37)$$

překročí hodnotu závislou na zvolené hladině testu a délce řady. Testový parametr p^* volíme subjektivně podle očekávané hodnoty p . Parametry k a p^* volíme ještě před výpočtem r_{τ} . Jedna z možností jak volit k a p^* je hledání hodnot vykazujících velkou sílu testu v případech, ve kterých A , p a n patří do množiny S pravděpodobných hodnot trendových parametrů a délek řady. Pro $0,01 \leq A \leq 0,04$; $0,8 \leq p \leq 0,975$ a $250 \leq n \leq 2000$ je v článku [8] ukázáno, že nejlepší volbou je $k = 30$ a $p^* = 0,92$ a $A^* = 0,03$.

Tvrzení 2.14:

Uvažujme náhodnou veličinu

$$T_{k,p^*} = \sum_{\tau=1}^k p^{*\tau} R_{\tau}. \quad (2.38)$$

Náhodná veličina

$$T^* = \frac{T_{k,p^*}}{\sqrt{\frac{1}{n} \sum_{\tau=1}^k p^{*2\tau}}} \quad (2.39)$$

má za platnosti H_0 asymptoticky rozdělení $N(0, 1)$.

Důkaz:

Dle tvrzení 2.4 má $\sqrt{n}(\mathbf{R} - \boldsymbol{\rho})$ asymptoticky rozdělení $N_k(\mathbf{0}, \boldsymbol{\Omega}_k)$. $\boldsymbol{\Omega}_k$ má prvky $\omega_{\tau\xi}$, $\tau, \xi = 1, 2, \dots, k$. Za platnosti H_0 máme $\boldsymbol{\rho} = (0, 0, \dots, 0)$ a $\boldsymbol{\Omega}_k = \mathbf{I}_k$. Z toho vyplývá, že $\sqrt{n}R_{\tau}$ má asymptoticky rozdělení $N(0, 1)$. Zřejmě je pro $n \rightarrow \infty$ $\text{var}(p^{*\tau} \sqrt{n}R_{\tau}) = p^{*2\tau} \text{var}(\sqrt{n}R_{\tau}) = p^{*2\tau}$. Náhodná veličina $\sqrt{n}T_{k,p^*}$ má proto asymptoticky rozdělení $N(0, \sum_{\tau=1}^k p^{*2\tau})$. Z toho plyne, že

$$T^* = \frac{T_{k,p^*}}{\sqrt{\frac{1}{n} \sum_{\tau=1}^k p^{*2\tau}}}$$

má asymptotické rozdělení $N(0, 1)$. □

V knize [9] je doporučeno provádět jednostranný test. H_0 tedy zamítáme na hladině 5%, když $T^* > u(0, 95) = 1,65$.

2.5.2 Spektrální test „ f_0 “ hypotézy náhodné procházky proti alternativě trendu

Testová statistika T^* je založena na váženém součtu odhadnutých autokorelací pro monotónně klesající váhy. Nahradíme-li váhy $p^{*\tau}$ Parzenovými vahami (2.18) zmíněnými v sekci 2.3.2, získáme statistiku o nulové frekvenci

$$f_0 = \sum_{\tau=1}^{M-1} \lambda_{\tau} r_{\tau} / \left(\sum_{\tau=1}^{M-1} \lambda_{\tau}^2 / n \right)^{1/2}. \quad (2.40)$$

Stejně jako v sekci 2.3.2 zamítáme H_0 na pětiprocentní hladině, pokud $f_0 \geq u(0, 95) = 1,65$.

Kapitola 3

Aplikace testů na reálná finanční data

3.1 Přepočet výnosů

Aby bylo možné aplikovat testy založené na asymptotických předpokladech o hodnotách časové řady, které byly uvedeny v kapitole 2, je třeba upravit data, aby je bylo možno považovat za řadu generovanou lineárním procesem. V knize [9] je doporučeno data centrovat a normovat pomocí následující transformace

$$y_t = \frac{x_t - \bar{x}}{\hat{v}_t}, \quad \bar{x} = \frac{1}{n} \sum_{t=1}^n x_t, \quad (3.1a)$$

$$\hat{v}_t = (1 - \gamma)\hat{v}_{t-1} + \frac{\gamma|x_{t-1} - \bar{x}|}{\delta}, \quad (3.1b)$$

$$\hat{v}_{21} = \frac{1}{20\delta} \sum_{t=1}^{20} |x_t - \bar{x}|. \quad (3.1c)$$

V [9] se dále doporučuje volba konstant $\delta \simeq 0,798$, $\gamma = 0,04$ pro akcie a $\gamma = 0,1$ jinak. Protože přepočet dat není pro prvních 20 hodnot definován, budeme počítat odhad korelačního koeficientu $r_{\tau,y}$ pro takto transformovaná data jen z hodnot y_t , $t > 20$.

$$r_{\tau,y} = \frac{\sum_{t=21}^{n-\tau} (y_t - \bar{y})(y_{t+\tau} - \bar{y})}{\sum_{t=21}^n (y_t - \bar{y})^2} \quad (3.2)$$

Tuto modifikaci vzorce r_{τ} použijeme ve všech testech, které budeme používat, aby byly testy počítány ze stejných časových řad a jejich výsledky byly lépe porovnatelné. Označme $n^* = n - 20$. Ve všech statistických testech aplikovaných na data pak nahrazujeme počet hodnot n počtem přepočtených hodnot n^* .

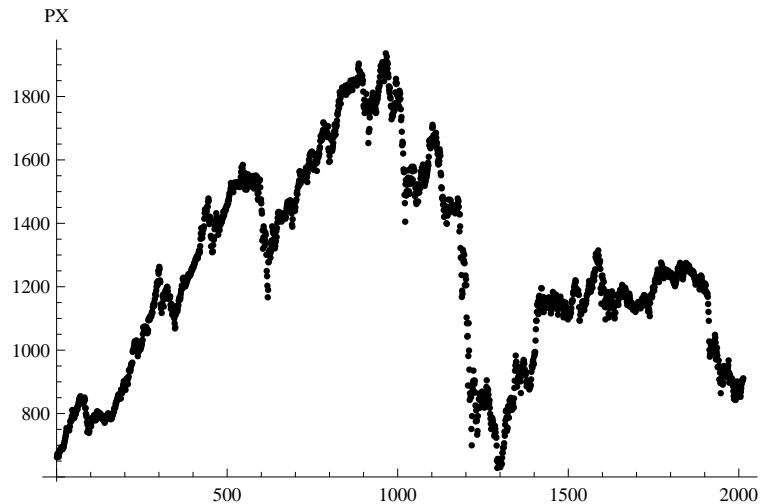
Jinou možností jak transformovat data x_t , $t = 1, 2, \dots$ je přepočítání na logaritmické výnosy

$$z_t = \log \frac{x_t}{x_{t-1}}, \quad t = 2, 3, \dots, n. \quad (3.3)$$

3.2 Analyzovaná data

3.2.1 Index PX

Burza cenných papírů Praha zveřejňuje od 20. 3. 2006 index blue chip emisí PX. Jedná se o cenový index nejspolehlivějších a nejvýnosnějších emisí obchodovaných na pražské burze. Před 20. 3. 2006 se nazýval tento index PX 50 a jeho historii index PX převzal. Na webové stránce [11] je zveřejněna časová řada hodnot indexu PX v čase uzavření každého obchodního dne od 7. 9. 1993. Pro analýzu vybereme časovou řadu hodnot indexu za roky 2004 - 2011 (obr. 3.1). Délka řady je 2013 hodnot. Více podrobností o indexu je možné najít na webových stránkách [11].



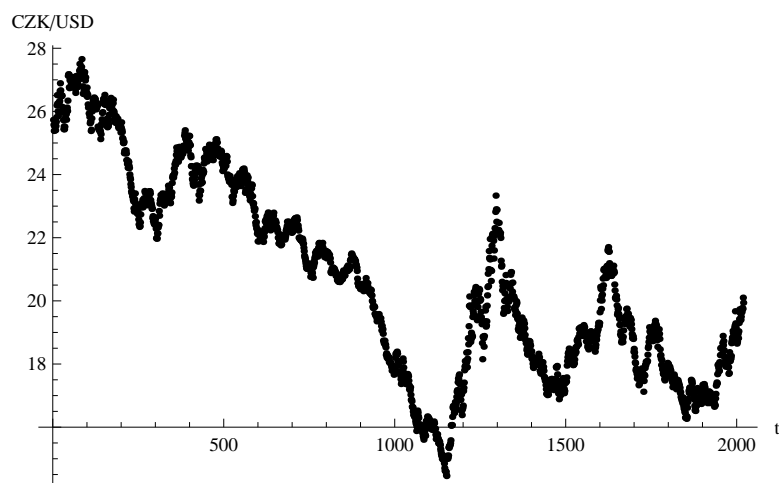
Obrázek 3.1: Časová řada indexu PX za roky 2004 - 2011

3.2.2 Kurz CZK/USD

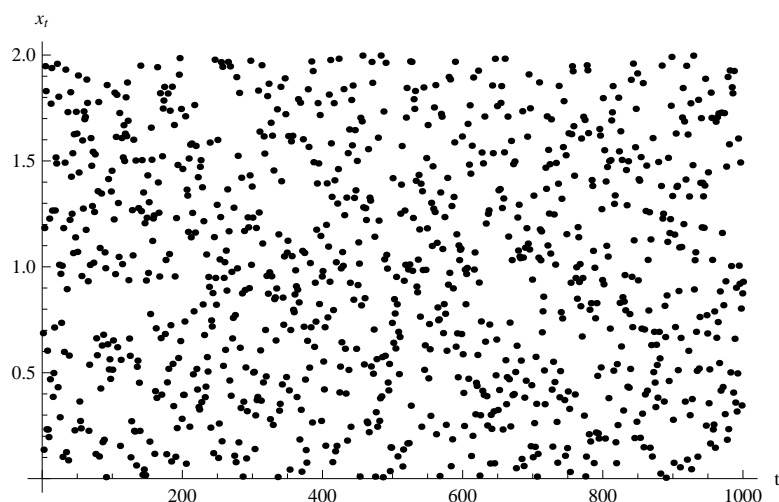
Česká národní banka zveřejňuje na webové stránce [10] každý den devizové kurzy české koruny. Časová řada hodnot kurzu české koruny k americkému dolaru ve 14:30 každého dne za období let 2004 - 2011 (obr. 3.2) je druhou řadou, kterou budeme analyzovat. Délka řady je 2019 hodnot.

3.2.3 Simulace

Pro srovnání použijeme také uměle generovanou řadu, která splňuje hypotézu náhodné procházky. Je to řada nezávislých veličin z rovnoměrného rozdělení na intervalu $(0, 2)$.



Obrázek 3.2: Časová řada cen USD v CZK za roky 2004 - 2011

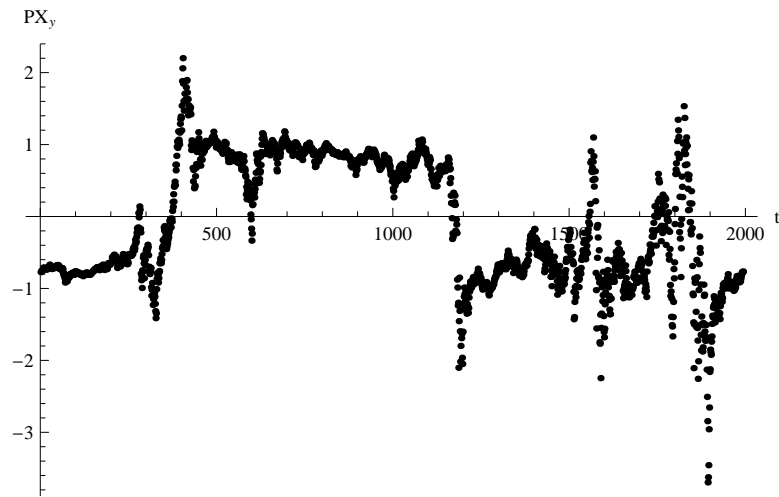


Obrázek 3.3: Řada náhodně generovaných hodnot délky 1000

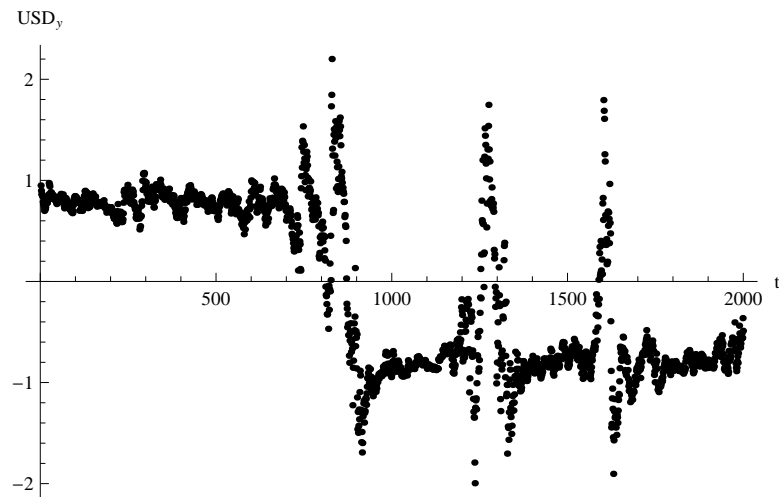
Byla generována v programu Wolfram Mathematica 8.0 pomocí funkce `RandomReal[]`. Před simulací byl generátor pseudonáhodných čísel nastaven na hodnotu 2012. Řadu označíme R a její délka je 1000 hodnot. Její graf vidíme na obrázku 3.3

3.3 Testy

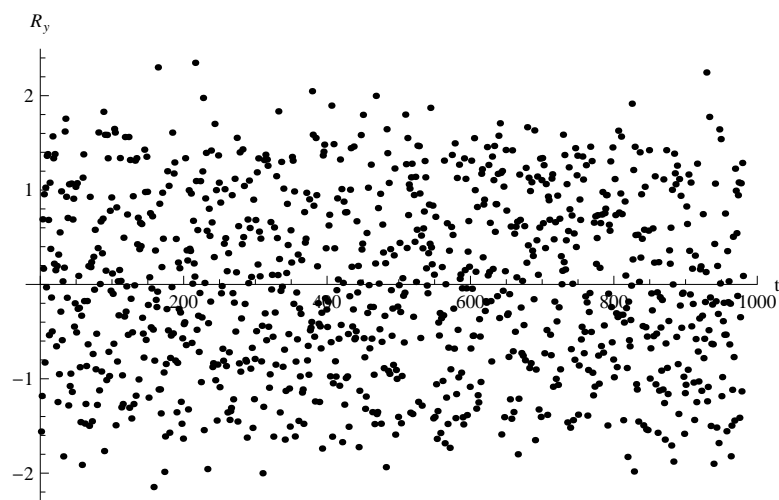
Na všechny časové řady aplikujeme testy, ve kterých je výpočet odhadu autokorelační funkce (2.3) nahrazen odhadem (3.2) z důvodů uvedených v sekci 3.1.1. Každou časovou řadu testujeme nejprve netransformovanou a poté s aplikací přepočtů (3.1) a (3.3). Průběhy řad transformovaných podle (3.1) vidíme na obrázcích 3.4 - 3.6. Průběhy řad transformovaných podle (3.3) jsou na obrázcích 3.7 - 3.9. Data, na která nebyl aplikován žádný přepočet, značíme $Data_x$. Data transformovaná dle (3.1) značíme $Data_y$ a logaritmické výnosy vypočítané z dat podle (3.3) značíme $Data_z$.



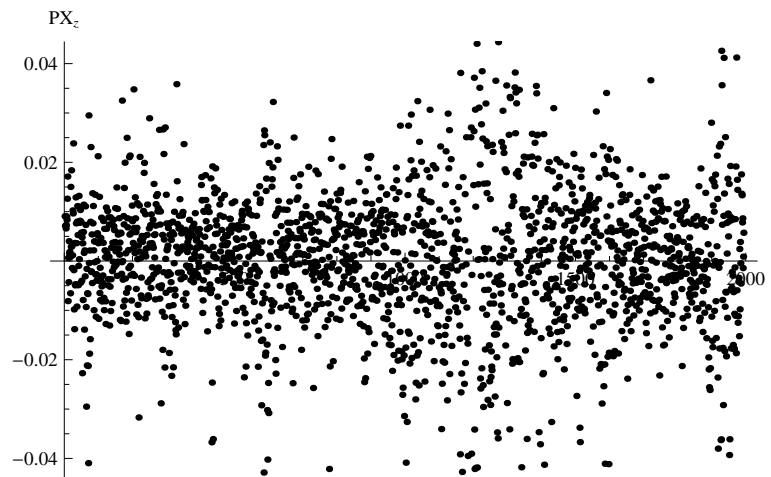
Obrázek 3.4: Časová řada indexu PX transformovaná dle (3.1)



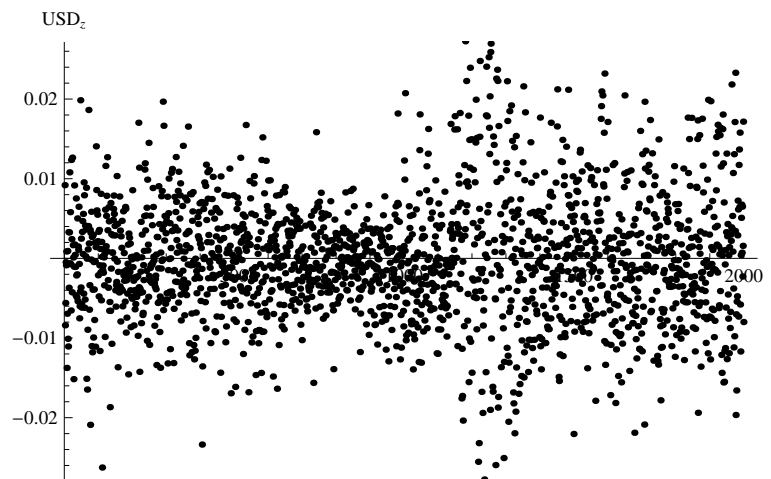
Obrázek 3.5: Časová řada cen USD transformovaná dle (3.1)



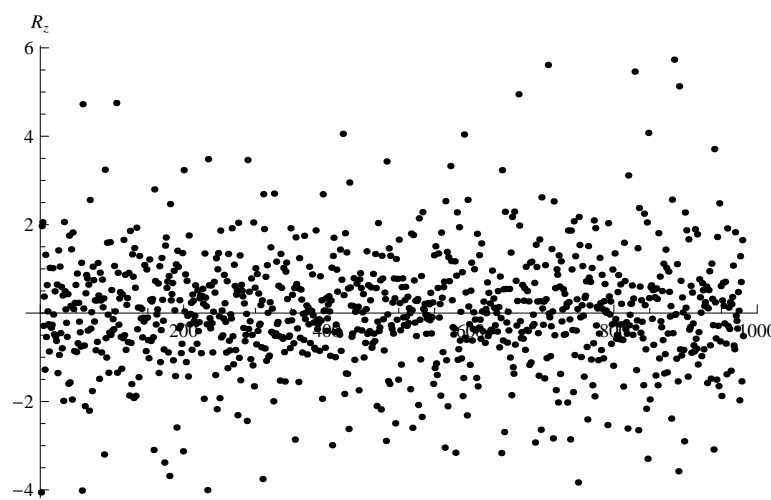
Obrázek 3.6: Časová řada náhodných hodnot R transformovaná dle (3.1)



Obrázek 3.7: Časová řada indexu PX transformovaná dle (3.3)



Obrázek 3.8: Časová řada cen USD transformovaná dle (3.3)



Obrázek 3.9: Časová řada náhodných hodnot R transformovaná dle (3.3)

3.3.1 Autokorelační testy

První test „ r_1 “ má za testovou statistiku samotný první autokorelační koeficient r_1 . Podle kritéria (2.7) zamítneme H_0 na hladině 5% pro $|r_1| > \frac{1,96}{\sqrt{n}}$.

Druhý autokorelační test „ N_r “, který počítá významné autokorelační koeficienty, zamítá H_0 na hladině 5% při dosažení počtu významných $N_r \geq N_0$ z prvních k autokorelací. N_r počítáme podle (2.8) a N_0 a k volíme podle (2.9), a tabulky 2.1. Pro test jsme zvolili $N_0 = 7$ a $k = 67$. Pak $\alpha(67, 7) = 0,0497$.

Další způsob jak rozhodnout o hypotéze náhodné procházky na základě hodnoty prvních k autokorelačních koeficientů je třetí autokorelační test „ Q_k “. Stejně jako v minulém testu volíme $k = 67$. Testovou statistiku Q_k počítáme podle (2.10). Pro $k = 67$ zamítáme H_0 na hladině 5% při dosažení $Q_{67} \geq \chi_{67}^2(0,95) = 87,11$. Shrnutí výsledků je uvedeno v tabulce 3.1.

	$\sqrt{n} r_1 $	H_0	N_r	H_0	Q_k	H_0
kritický obor	$> 1,96$		≥ 7		$\geq 87,11$	
R_x	0,321	nezamítáme	3	nezamítáme	68,93	nezamítáme
R_y	0,344	nezamítáme	3	nezamítáme	69,97	nezamítáme
R_z	15,62	zamítáme	4	nezamítáme	313,98	zamítáme
USD_x	44,6	zamítáme	67	zamítáme	111 051	zamítáme
USD_y	44,1	zamítáme	67	zamítáme	82 462,8	zamítáme
USD_z	2,2	zamítáme	9	zamítáme	106,859	zamítáme
PX_x	44,5	zamítáme	67	zamítáme	107 575	zamítáme
PX_y	43,98	zamítáme	67	zamítáme	76 228,9	zamítáme
PX_z	3,2	zamítáme	16	zamítáme	158,781	zamítáme

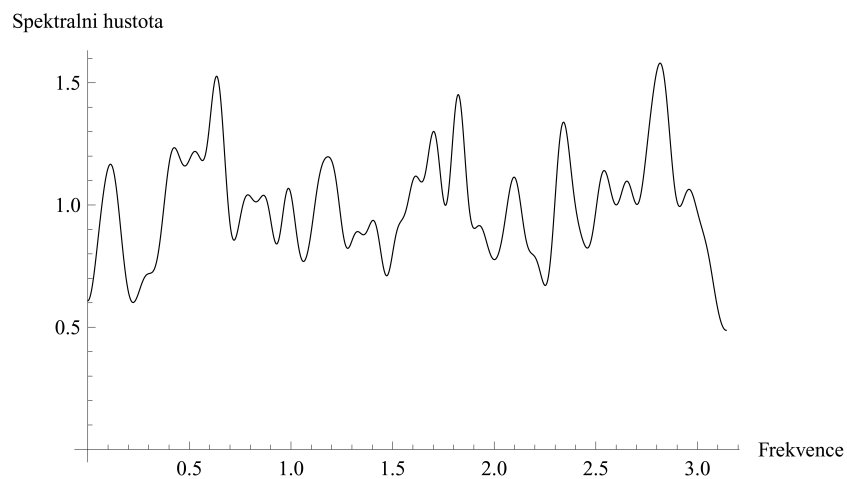
Tabulka 3.1: Výsledky autokorelačních testů

Pro řadu kurzů USD i indexů PX všechny testy zamítly hypotézu náhodné procházky a to i pro obě transformované verze. Naopak pro řadu R hypotéza zamítnuta nebyla až na testy „ r_1 “ a „ Q_k “ pro logaritmické výnosy. To ukazuje, že přepočtení na logaritmické výnosy není pro tuto řadu vhodný. O řadě totiž víme, že hypotézu H_0 splňuje, a hodnoty autokorelačních statistik to potvrzují.

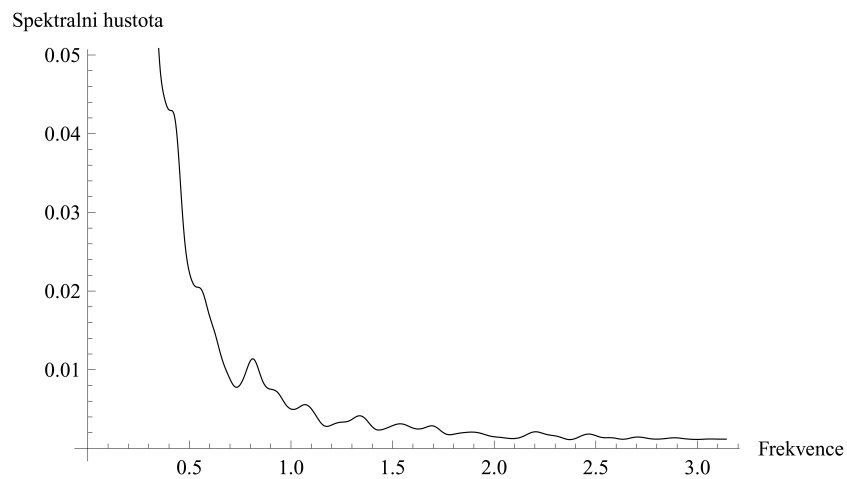
3.3.2 Spektrální analýza

Spektrální analýza vychází z předpokladu, že data mají tvar (2.12). Potom má spektrální hustota (2.13) lokální maxima v bodech, které odpovídají frekvenci cyklů v datech. Pomocí odhadu (2.17) její transformace (2.16) jsme získali grafy odhadnuté spektrální hustoty v obrázcích 3.10 - 3.12.

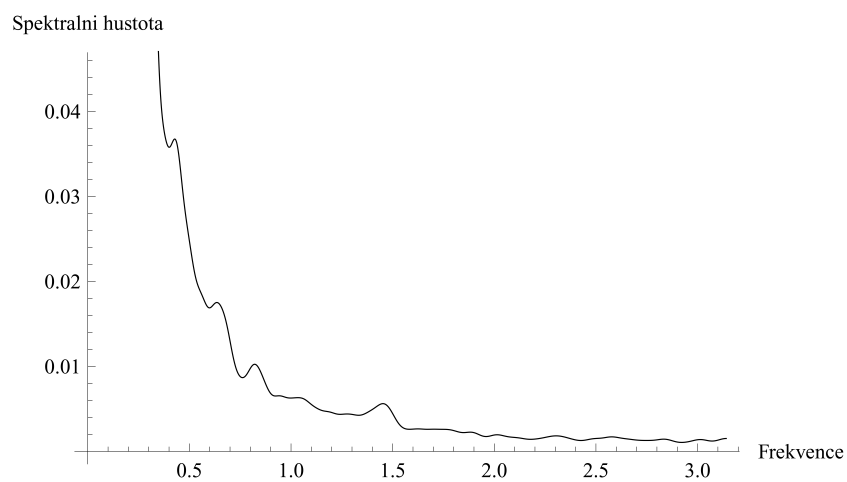
Jednou možností, jak testovat H_0 proti alternativě cyklického chování časové řady, je zkusit odhadnout frekvenci cyklů alternativní hypotézy. Spektrální test f_w uvažuje jako pravděpodobnou periodu cyklů v datech jeden týden, tedy 5 hodnot. H_0 pak zamítneme na hladině 5%, pokud $f_w > 1,65$. Hodnotu f_w počítáme podle (2.21).



Obrázek 3.10: Odhad spektrální hustoty řady R , splňující H_0



Obrázek 3.11: Odhad spektrální hustoty řady kurzů USD/CZK z let 2004 - 2011



Obrázek 3.12: Odhad spektrální hustoty řady hodnot indexu PX z let 2004 - 2011

Další způsob testování H_0 proti alternativě cyklického chování časové řady je počítání významných koeficientů f_j , podobně jako u testu „ N_r “. Před testováním je třeba zvolit, s jakou jemností bude test cykly v datech vyhledávat. Tu volíme pro všechny naše spektrální testy $M = 100$. H_0 na hladině 5% zamítneme, pokud testová statistika $N_s \geq 4$. Hodnotu N_s spočítáme podle vzorce (2.22). Výsledky jsou zapsány v tabulce 3.2.

	f_w	H_0	N_s	H_0
kritický obor	$> 1,65$		≥ 4	
R_x	-0,52	nezamítáme	1	nezamítáme
R_y	-0,5	nezamítáme	1	nezamítáme
R_z	-1,486	nezamítáme	15	zamítáme
USD_x	-6,184	nezamítáme	26	zamítáme
USD_y	-6,061	nezamítáme	26	zamítáme
USD_z	0,572	nezamítáme	2	nezamítáme
PX_x	-6,168	nezamítáme	26	zamítáme
PX_y	-6,07	nezamítáme	26	zamítáme
PX_z	0,969	nezamítáme	7	zamítáme

Tabulka 3.2: Výsledky spektrálních testů

U řad USD a PX a jejich transformací vidíme, že test, který bral v úvahu pouze týdenní frekvenci dat, nezamítl H_0 ani v jednom případě. Stalo se tak proto, že test má sice velkou sílu vůči alternativě týdenních cyklů, ale vůči ostatním alternativám je naopak slabý. Proto nedošlo k zamítnutí H_0 , ačkoli podle ostatních testů soudíme, že H_0 neplatí. Podle obrázků 3.11 a 3.12 naopak vidíme velkou podobnost spektrální hustoty finančních časových řad USD a PX se spektrální hustotou časové řady s trendem, která je na obrázku 2.1. Dále v textu uvidíme, že hypotéza cenového trendu je pro tyto řady skutečně lepší alternativou. Test „ N_s “ hypotézu H_0 pro řady USD a PX zamítl (až na logaritmické výnosy kurzu amerického dolaru), protože má větší sílu, než test „ f_w “. Logaritmické výnosy kurzu dolaru nebyly zamítnuty, protože tato transformace příliš znormovala hodnoty řady, takže ani test „ N_s “ nebyl dost silný, aby H_0 zamítl. Řada R a její přepočty nebyly opět podle očekávání zamítnuty ani jedním z testů kromě testu „ N_s “, který zamítl logaritmické výnosy náhodné řady R . Opět je to patrně způsobeno nevhodností logaritmického přepočtu.

3.3.3 Neparametrické testy

Jediným neparametrickým testem, který jsme zmínili, je iterační test „ K “. Ten předpokládá nulovou střední hodnotu veličin časové řady. Proto jej můžeme aplikovat pouze na centrovaná data. Hypotézu H_0 zamítáme, pokud $|K| > 1,96$. Testovou statistiku K počítáme podle vzorce (2.26).

Pro finanční řadu kurzů USD i pro finanční řadu indexů PX iterační test „ K “ H_0 na hladině 5% zamítá. Pro řadu R naopak hypotéza náhodné procházky zamítnuta nebyla.

	$ K $	H_0
kritický obor	$> 1,96$	
R_y	0,32	nezamítáme
USD_y	43,79	zamítáme
PX_y	42,85	zamítáme

Tabulka 3.3: Výsledky neparametrického testu

3.3.4 Test hypotézy náhodné procházky proti alternativě trendu v časové řadě

Testování H_0 proti alternativě trendu v časové řadě vychází z logaritmického věrohodnostního poměru (2.35). Volíme-li podle sekce 2.5.1 parametry $k = 30$ a $p^* = 0,92$, zamítáme H_0 na hladině 5%, pokud $T^* > 1,65$. Statistiku T^* počítáme podle vzorce (2.39) se zvolenými parametry.

Další možnost testování H_0 proti hypotéze trendu v datech vychází ze spektrální analýzy. Pokud do vzorce (2.20) dosadíme jako předpokládanou frekvenci cyklů v datech 0, získáme předpis (2.40) pro výpočet testové statistiky „ f_0 “. Na hladině 5% zamítneme H_0 pokud $f_0 \geq 1,65$. Výsledky vidíme v tabulce 3.4.

	T^*	H_0	f_0	H_0
kritický obor	$> 1,65$		$\geq 1,65$	
R_x	-0,864	nezamítáme	-1,192	nezamítáme
R_y	-0,963	nezamítáme	-1,246	nezamítáme
R_z	-6,159	nezamítáme	-3,04	nezamítáme
USD_x	196,73	zamítáme	301,5	zamítáme
USD_y	183,49	zamítáme	265,7	zamítáme
USD_z	1,699	zamítáme	1,119	nezamítáme
PX_x	195,7	zamítáme	298	zamítáme
PX_y	180,9	zamítáme	258,6	zamítáme
PX_z	1,112	nezamítáme	1,511	nezamítáme

Tabulka 3.4: Výsledky trendových testů

Oba testy zamítly H_0 pro obě finanční řady bez transformace i po centrování a normování. Kromě řady USD_z při testování testem T^* však nebyla hypotéza zamítnuta pro logaritmické výnosy. Přechod k logaritmickým výnosům odstranil příliš mnoho variability. Na obrázcích 3.7 a 3.8 je vidět, že trend se z řad zcela vytratil. Pro R i pro oba její přepočty nebyla H_0 podle očekávání zamítnuta.

3.3.5 Shrnutí

Zde uvádíme v tabulce hodnoty všech aplikovaných testových statistik. Tučně jsou hodnoty, které spadají do příslušného kritického oboru a H_0 proto zamítají.

	$\sqrt{n} r_1 $	N_r	Q_k	f_w	N_s	$ K $	T^*	f_0
k. o.	$> 1,96$	≥ 7	≥ 87.11	$> 1,65$	≥ 4	$> 1,96$	$> 1,65$	$\geq 1,65$
R_x	0,321	3	68,93	-0,52	1	N/A	-0,864	-1,192
R_y	0,344	3	69,97	-0,5	1	0,32	-0,963	-1,246
R_z	15,62	4	313,98	-1,486	15	N/A	-6,159	-3,04
USD_x	44,6	67	111 051	-6,184	26	N/A	196,73	301,5
USD_y	44,1	67	82 462,8	-6,061	26	43,79	183,49	265,7
USD_z	2,2	9	106,859	0,572	2	N/A	1,699	1,119
PX_x	44,5	67	107 575	-6,168	26	N/A	195,7	298
PX_y	43,98	67	76 228,9	-6,07	26	42,85	180,9	258,6
PX_z	3,2	16	158,781	0,969	7	N/A	1,112	1,511

Tabulka 3.5: Shrnutí výsledků všech aplikovaných testů

Vidíme, že se potvrdil předpoklad, že řada R splňuje hypotézu náhodné procházky H_0 . Finanční řada hodnot kurzu amerického dolaru k české koruně ani řada hodnot indexu PX pražské burzy naopak hypotézu H_0 nespĺňují.

Literatura

- [1] Anderson, T.W.; Walker, A.M.: On the Asymptotic Distribution of the Autocorrelations of a Sample from a Linear Stochastic Process. *Annals of Mathematical Statistics* 35(1964), 1296-1303.
- [2] Cipra, T.: *Finanční ekonometrie*. Ekopress, Praha, 2008.
- [3] Granger, C.W.J. and O. Morgenstern: *Predictability of Stock Market Prices*. Heath, Lexington, Massachusetts, 1970.
- [4] Lomnicki, Z.A.; Zaremba, S.K.: On the Estimation of Autocorrelation in Time Series. *Annals of Mathematical Statistics* 28(1957), 140-158.
- [5] Mood, A.M.: The Distribution Theory of Runs. *Annals of Mathematical Statistics* 11(1940), 367-392.
- [6] Moran, P.A.P.: Testing for Serial Correlation with Exponentially Distributed Variates. *Biometrika* 54(1967), 395-401.
- [7] Praetz, P.D.: Testing for a Flat Spectrum on Efficient Market Price Data. *Journal of Finance* 34(1979), 645-658.
- [8] Taylor, J.S.: Tests of Random Walk Hypothesis Against a Price-Trend Hypothesis. *Journal of Financial and Quantitative Analysis* 17(1982), 37-61.
- [9] Taylor, S.J.: *Modelling Financial Time Series*. Wiley, New York, 1995.
- [10] <http://www.cnb.cz>
- [11] <http://www.bcphp.cz/dokument.aspx?k=Burzovni-Indexy>.

Seznam tabulek

2.1	Některé přípustné hodnoty parametrů N_0, k	18
3.1	Výsledky autokorelačních testů	33
3.2	Výsledky spektrálních testů	35
3.3	Výsledky neparametrického testu	36
3.4	Výsledky trendových testů	36
3.5	Shrnutí výsledků všech aplikovaných testů	37

Seznam obrázků

2.1	Spektrální hustota finanční časové řady s trendem pro $A = 0,03$ a $p = 0,95$	25
3.1	Časová řada indexu PX za roky 2004 - 2011	29
3.2	Časová řada cen USD v CZK za roky 2004 - 2011	30
3.3	Řada náhodně generovaných hodnot délky 1000	30
3.4	Časová řada indexu PX transformovaná dle (3.1)	31
3.5	Časová řada cen USD transformovaná dle (3.1)	31
3.6	Časová řada náhodných hodnot R transformovaná dle (3.1)	31
3.7	Časová řada indexu PX transformovaná dle (3.3)	32
3.8	Časová řada cen USD transformovaná dle (3.3)	32
3.9	Časová řada náhodných hodnot R transformovaná dle (3.3)	32
3.10	Odhad spektrální hustoty řady R , splňující H_0	34
3.11	Odhad spektrální hustoty řady kurzů USD/CZK z let 2004 - 2011	34
3.12	Odhad spektrální hustoty řady hodnot indexu PX z let 2004 - 2011	34