MATEMATICKO-FYZIKÁLNÍ FAKULTA
Univerzita Karlova

BAKALÁŘSKÁ PRÁCE

Jakub Kárný

Autokorelace v časových řadách

Katedra pravděpodobnosti a matematické statistiky

Vedoucí bakalářské práce: RNDr. Jitka Zichová, Dr.
Studijní program: Matematika
Studijní obor: Finanční matematika

Praha 2021

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

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

V .... dne ....

Podpis autora

Zde bych chtěl velmi poděkovat vedoucí mé bakalářské práce RNDr. Jitce Zichové, Dr. za doporučené materiály ke studiu, rady a obohacující vysvětlení náročnějších pasáží. Rovněž bych jí rád poděkoval za její asistenci při shánění poznatků pro některá odvozování a doporučení pro sepsání práce.

Název práce: Autokorelace v časových řadách

Autor: Jakub Kárný

Katedra: Katedra pravděpodobnosti a matematické statistiky

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

Abstrakt: Práce se zabývá autokorelační strukturou časových řad, konkrétně i procesu AR(1). Je odvozen odhad rozptylu výběrových autokorelací a jsou dokázány jeho asymptotické vlastnosti. V simulační studii je generován normálně rozdělený bílý šum a AR(1). V těchto řadách zkoumáme rychlost konvergence odhadu rozptylu výběrových autokorelací. Dále vyšetřujeme empirickou hladinu a sílu některých testů nekorelovanosti časové řady.

Klíčová slova: Autokorelace, výběrová autokorelace, momenty, časová řada, autoregresní proces 1.řádu, statistický test, hypotéza náhodné procházky

Title: Serial correlations in time series

Author: Jakub Kárný

Department: Department of Probability and Mathematical Statistics

Supervisor: RNDr. Jitka Zichová, Dr., Department of Probability and Mathematical Statistics

Abstract: The subject of the thesis is the autocorrelation structure of time series. AR(1) process is studied as a special example. An estimator of the variance of the sample autocorrelation is derived and its asymptotic properties are proved. We investigate the convergence of the variance estimates of sample autocorrelations in some simulated series. Further, the empirical significance level and power of selected autocorrelation tests are calculated.

Keywords: Serial correlation, sample autocorrelation, moments, autoregressive process of the first order, statistical test, random walk hypothesis

Obsah

Úvod

Finanční data (např. výnosy a ceny obchodovatelných cenných papírů) jsou běžně v podobě časové řady, tedy pozorovaných hodnot obvykle v pravidelných časových intervalech. V této práci se budeme zabývat zkoumáním časových řad z hlediska jejich autokorelační struktury. Odvodíme odhad rozptylu výběrových autokorelací a jejich asymptotické vlastnosti a ukážeme vztah výběrových autokorelací a testových statistik testů o nekorelovanosti časové řady.

V první kapitole zadefinujeme základní pojmy a představíme autoregresní proces prvního řádu AR(1) jako příklad stacionárního korelovaného procesu. Vyjádříme AR(1) v kauzálním tvaru a odvodíme rozptyl, autokovariance a autokorelace veličin z AR(1) a jejich kvadrátů, které využijeme při dalších odvozováních.

V kapitole 2 za daných předpokladů sestrojíme nestranný odhad rozptylu výběrových autokorelací a dokážeme některé asymptotické vlastnosti tohoto odhadu. Na závěr kapitoly v numerické ilustraci ukážeme odlišnosti limitních hodnot pro rozptyl výběrových autokorelací v nekorelovaném procesu a v AR(1).

Třetí kapitola pojednává o hypotéze náhodné procházky a jejích statistických testech. Propojíme teoretické poznatky o výběrových autokorelacích jako náhodné veličině s testovými statistikami zmíněných testů. Zmíníme rovněž transformaci dat z časových řad, aby výběrové autokorelace splňovaly předpoklad pro spolehlivé statistické testy.

Na závěr v praktické části simulujeme vybrané modely časových řad bílého šumu a AR(1). Vyšetřujeme rychlost konvergence odhadu rozptylu výběrových autokorelací a budeme zkoumat empirickou hladinu a sílu vybraných testů hypotézy náhodné procházky.

Kapitola 1

Náhodný proces

V této kapitole představíme vybrané pojmy z teorie náhodných procesů a z analýzy časových řad. Budeme zkoumat stacionaritu náhodného procesu a autokorelace, které charakterizují vzájemnou závislost veličin náhodného procesu. Rovněž se zaměříme na odhady autokorelací. Dále rozebereme autoregresní proces prvního řádu, který je jednoduchým příkladem stacionárních procesů.

1.1 Základní pojmy

V prvních dvou definicích zavedeme pojmy náhodný proces a časová řada jako jeho pozorovaná realizace.

Definice 1 (Náhodný proces, Prášková, 2016, Definice 1.1, str. 9)**.** *Nechť* $(\Omega, \mathcal{A}, P)$ *je pravděpodobnostní prostor,* $T \subset \mathbb{R}$ *a* (S,E) *je obecný měřitelný prostor. Množina náhodných veličin* $\{X_t, t \in T\}$ *definovaných na pravděpodobnostním prostoru s hodnotami v* S *se nazývá náhodný proces.*

Poznámka. Pravděpodobnostní prostor $(\Omega, \mathcal{A}, P)$ je složen z trojice tvořené množinou elementárních jevů, σ-algebrou jejích podmnožin a nezápornou, měřitelnou funkcí $P : \mathcal{A} \to [0, 1]$ představující pravděpodobnost.

Množina T je indexová množina reprezentující čas, který může být spojitý nebo diskrétní. Dále budeme uvažovat procesy s diskrétním časem ($T = \mathbb{Z}$).

Obecný měřitelný prostor (stavový prostor) (S,E) je dvojice množiny hodnot náhodných veličin X_t a σ-algebry jejích podmnožin. Uvažujeme $(S, E) = (\mathbb{R}, B)$, což znamená prostor reálných čísel s borelovskou σ-algebrou.

Definice 2 (Časová řada, Prášková, 2016, Definice 1.1, str. 9)**.** *Nechť* $T = \{1, 2, \ldots, n\}$*, pak definujeme časovou řadu* $x_1, x_2, \ldots, x_n$ *jako realizaci náhodného procesu s diskrétním časem, kde* x_t *je hodnota zaznamenaná v čase* $t \in T$*.*

Významnou charakteristikou náhodných procesů je jejich stacionarita. Ta reprezentuje stabilitu procesu v čase a definujeme několik typů.

Definice 3 (Silná stacionarita, Prášková, 2016, Definice 1.7, str. 11)**.** *Silně (striktně) stacionární proces definujeme jako náhodný proces* $\{X_t, t \in T\}$*, pro který platí, že pro libovolné* $n \in \mathbb{N}$*, pro libovolná reálná* $x_1, \ldots, x_n$ *a pro libovolná* $t_1, \ldots, t_n$ *a* τ *taková, že* $t_k \in T, t_{k+\tau} \in T, 1 \leq k \leq n$*, platí*

$$F_{t_1,\ldots,t_n}(x_1, \ldots, x_n) = F_{t_1+\tau,\ldots,t_n+\tau}(x_1, \ldots, x_n).$$

Definice 3 říká, že sdružená distribuční funkce pro náhodné veličiny ze striktně stacionárního procesu nezávisí na časovém posunu. Z toho dále vyplývá, že charakteristiky náhodných veličin X_t (střední hodnota, rozptyl, kovariance, korelace) také nezávisí na časovém posunu (Taylor, 2007).

Definice 4 (Slabá stacionarita, Prášková, 2016, Definice 1.8, str. 11)**.** *Slabě stacionární náhodný proces definujeme jako náhodný proces* $\{X_t, t \in T\}$*, pro který platí, že má konstantní střední hodnotu* $\mu_t = \mu$ *(pro libovolné* $t \in T$*) a kovariance závisí pouze na časovém posunu* τ *mezi dvěma náhodnými veličinami* $X_t, X_{t+\tau}$ *z náhodného procesu,* $t \in T, \tau > 0$*.*

Z definicí striktní a slabé stacionarity je zřejmé, že striktní stacionarita implikuje slabou stacionaritu. Striktní stacionaritou je dáno stejné rozdělení veličin X_t, tedy i jejich stejné konečné střední hodnoty. Dále je dáno stejné sdružené rozdělení dvojic veličin $X_t, X_{t+\tau}$, takže pro libovolné τ mají stejné kovariance, které závisejí pouze na časovém posunu. Opačná implikace obecně neplatí, protože konečnou střední hodnotou a kovariancí není jednoznačně určeno pravděpodobnostní rozdělení.

Pouze slabě stacionární, normálně rozdělený proces je zároveň i striktně stacionární. Ze slabé stacionarity dostáváme pro náhodné vektory $(X_{t_1}, \ldots, X_{t_n})$ a $(X_{t_1+\tau}, \ldots, X_{t_n+\tau})$ stejnou konečnou střední hodnotu a varianční matici $\forall n \in \mathbb{N}$, libovolná $t_1, \ldots, t_n, \tau$. Těmito charakteristikami je pak normální rozdělení určeno jednoznačně (Prášková, 2016).

Důležitým příkladem stacionárního procesu je bílý šum.

Definice 5 (Bílý šum, Prášková, 2016, str. 51)**.** *Bílý šum definujeme jako posloupnost nekorelovaných náhodných veličin s nulovou střední hodnotou a konečným kladným rozptylem.*

Poznámka. V textu budeme častěji uvažovat pojem striktní bílý šum, kterým myslíme posloupnost nezávislých náhodných veličin s momenty jako v definici 5. Budeme předpokládát i jejich stejné rozdělení.

Závislostní strukturu veličin náhodných procesů vyjadřujeme pomocí autokorelací, tedy korelací dvojic $X_t, X_{t+\tau}$ $\forall t \in T, \tau > 0$.

Definice 6 (Autokorelace,Taylor, 2007, str. 17)**.** *Nechť* $\{X_t\}$ *je náhodný proces s konečnými druhými momenty náhodných veličin* X_t $\forall t \in T$*, pak autokorelaci řádu* τ *definujeme předpisem*

$$\rho(X_t, X_{t+\tau}) = \frac{\sigma(X_t, X_{t+\tau})}{\sqrt{varX_t}\sqrt{varX_{t+\tau}}}, \tau \geq 1, \tag{1.1}$$

kde $\sigma(X_t, X_{t+\tau})$ *je kovariance těchto dvou náhodných veličin.*

V silně i slabě stacionárním procesu značíme autokorelace ρ_τ, závisejí pouze na velikosti časového posunu τ.

Definice 7 (Výběrová autokorelace, Taylor, 2007, str. 24)**.** *Nechť $x_1, \dots, x_n$ je časová řada, $\bar{x} = \frac{1}{n}\sum_{t=1}^{n} x_t$, pak výběrovou autokorelaci řádu τ definujeme jako*

$$r_{\tau,X} = \frac{\sum_{t=1}^{n-\tau}(x_t - \bar{x})(x_{t+\tau} - \bar{x})}{\sum_{t=1}^{n}(x_t - \bar{x})^2}, \tau \geq 1. \tag{1.2}$$

Poznámka. Výběrová autokorelace je realizací náhodné veličiny

$$R_{\tau,X} = \frac{\sum_{t=1}^{n-\tau}(X_t - \bar{X})(X_{t+\tau} - \bar{X})}{\sum_{t=1}^{n}(X_t - \bar{X})^2}, \tau \geq 1. \tag{1.3}$$

$\bar{X}$ označuje výběrový průměr náhodných veličin $X_1, \dots, X_n$.

1.2 Autoregresní proces

V této části zavedeme autoregresní proces prvního řádu, některé jeho vlastnosti a odvodíme rozptyl, autokovarianci a autokorelaci tohoto náhodného procesu.

Definice 8 (Autoregresní proces 1. řádu, Prášková, 2016, str. 58)**.** *Náhodnou posloupnost $\{X_t, t \in \mathbb{Z}\}$ definovanou vztahem*

$$X_t = \varphi X_{t-1} + \epsilon_t, \tag{1.4}$$

kde ϵ_t je striktní bílý šum, nazýváme autoregresní proces prvního řádu (označujeme AR(1)).

Dále budeme bez újmy na obecnosti předpokládat pro veličiny procesu AR(1)

$$\mathsf{E}\, X_t = 0 \quad \forall t \in \mathbb{Z}.$$

Ukážeme, že proces AR(1) lze zapsat ve formě lineárního procesu, to znamená jako nekonečnou lineární kombinaci veličin $\{\epsilon_t\}$. Využijeme operátor posunutí B definovaný v Taylor (2007, str. 20):

$$B^k X_t = X_{t-k}, k \geq 0. \tag{1.5}$$

Odtud dosadíme za X_{t-1} v (1.4). Dostáváme

$$X_t - \varphi B X_t = \epsilon_t,$$
$$X_t = \frac{1}{1 - \varphi B}\epsilon_t.$$

Pro $|\varphi| < 1$ lze koeficient u ϵ_t chápat jako součet nekonečné geometrické řady a vyjádřit tak X_t jako

$$X_t = \sum_{i=0}^{\infty} (\varphi B)^i \epsilon_t.$$

Z definice operátoru posunutí B máme $B^i \epsilon_t = \epsilon_{t-i}$ a dostáváme

$$X_t = \sum_{i=0}^{\infty} \varphi^i \epsilon_{t-i}. \tag{1.6}$$

Jak jsme ukázali výše, AR(1) lze zapsat jako lineární proces pouze pro $|\varphi| < 1$. Budeme tedy dále klást na parametr φ tento předpoklad. Podle Prášková (2016, věta 5.4, str. 56) odtud plyne slabá stacionarita autoregresního procesu 1. řádu.

Pro $|\varphi| > 1$ nemá nekonečná řada součet, nelze tudíž zapsat AR(1) jako lineární proces. Takový náhodný proces není stacionární.

Pro alespoň slabě stacinonární proces AR(1) odvodíme rozptyl, autokovarianci a autokorelaci. Rozptyl spočítáme ze vzorce (1.4). S využitím stacionarity AR(1) dostaneme

$$\operatorname{var} X_t = \varphi^2 \operatorname{var} X_t + \operatorname{var} \epsilon_t + 2\sigma(\varphi X_{t-1}, \epsilon_t).$$

Kovarianci díky nulovým středním hodnotám X_t, ϵ_t zapíšeme jako

$$\sigma(\varphi X_{t-1}, \epsilon_t) = \mathsf{E}\,(\varphi X_{t-1}\epsilon_t) = \varphi\,\mathsf{E}\,(X_{t-1}\epsilon_t).$$

X_{t-1} rozepíšeme jako lineární proces ze vzorce (1.6), máme tedy s využitím nezávislosti ϵ_t

$$\mathsf{E}\,(X_{t-1}\epsilon_t) = \mathsf{E} \sum_{i=0}^{\infty} \varphi^i \epsilon_{t-1-i}\epsilon_t = 0. \tag{1.7}$$

Po úpravě dostaneme rozptyl AR(1)

$$\operatorname{var} X_t = \frac{\operatorname{var} \epsilon_t}{1-\varphi^2}. \tag{1.8}$$

Vidíme, že rozptyl nezávisí na časovém kroku τ. Pro určení autokovariance $\sigma(X_t, X_{t+\tau})$ vynásobíme vzorec (1.4) pro čas $t+\tau$ veličinou X_t a aplikujeme operátor střední hodnoty

$$\mathsf{E}\,(X_t X_{t+\tau}) = \varphi\,\mathsf{E}\,(X_t X_{t+\tau-1}) + \mathsf{E}\,(X_t \epsilon_{t+\tau}).$$

S předpokladem $\mathsf{E}\, X_t = 0\ \forall t$ a analogickým postupem jako při odvození vzorce (1.7) můžeme psát

$$\sigma(X_t, X_{t+\tau}) = \varphi\sigma(X_t, X_{t+\tau-1}).$$

Po τ krocích pak dostaneme s využitím stacionarity AR(1)

$$\sigma(X_t, X_{t+\tau}) = \varphi^\tau \sigma(X_t, X_t) = \varphi^\tau \operatorname{var} X_t. \tag{1.9}$$

Po dosazení do vzorce pro autokorelaci (1.1) máme autokorelaci řádu τ pro AR(1)

$$\rho(X_t, X_{t+\tau}) = \frac{\varphi^\tau \operatorname{var} X_t}{\operatorname{var} X_t} = \varphi^\tau. \tag{1.10}$$

Dále předpokládejme, že striktní bílý šum je normálně rozdělený ($\epsilon_t \sim N(0,1)$). Pak je i proces $\{X_t\}$ normálně rozdělený. Odvodíme rozptyl, autokovarianci a autokorelaci pro kvadráty procesu AR(1). Máme

$$\operatorname{var} X_t^2 = \mathsf{E}\, X_t^4 - (\mathsf{E}\, X_t^2)^2.$$

Nejprve určíme čtvrtý moment X_t, a to s pomocí vyjádření AR(1) ve vzorci (1.4). Po roznásobení dostaneme

$$\mathsf{E}\, X_t^4 = \mathsf{E}\,(\varphi^4 X_{t-1}^4 + 4\varphi^3 X_{t-1}^3 \epsilon_t + 6\varphi^2 X_{t-1}^2 \epsilon_t^2 + 4\varphi X_{t-1}\epsilon_t^3 + \epsilon_t^4).$$

Vzhledem k možnosti zapsání $X_{t-\tau}$ jako lineární proces, jsou veličiny $X_{t-\tau}, \epsilon_t$ nezávislé $\forall \tau > 0$. Dle tvrzení (Kulich, 2018, tvrzení 3.6, str. 19) jsou mocniny nezávislých náhodných veličin $X_{t-\tau}, \epsilon_t$ nezávislé. S využitím tohoto poznatku a vlastností normované normální veličiny máme

$$\mathsf{E}\, X_t^4 = \varphi^4\, \mathsf{E}\, X_{t-1}^4 + 6\varphi^2\, \mathsf{E}\, X_{t-1}^2 + 3.$$

Za $\mathsf{E}\, X_{t-1}^2$ můžeme dosadit při $\mathsf{E}\, X_t = 0$ a $\operatorname{var} \epsilon_t = 1$ ze vzorce (1.8)

$$\operatorname{var} X_t = \mathsf{E}\, X_t^2 = \frac{1}{1-\varphi^2}. \tag{1.11}$$

Ze slabé stacionarity AR(1) máme díky předpokladu $\epsilon_t \sim N(0,1)$ striktní stacionaritu, tudíž stejné $\mathsf{E}\, X_t^4$ $\forall t$. Dostaneme pak

$$\mathsf{E}\, X_t^4 - \varphi^4\, \mathsf{E}\, X_t^4 = 6\varphi^2 \frac{1}{1-\varphi^2} + 3.$$

Po úpravách máme

$$\mathsf{E}\, X_t^4 = \frac{6\varphi^2 + 3(1-\varphi^2)}{1-\varphi^2} \frac{1}{1-\varphi^4} = \frac{3\varphi^2 + 3}{(1-\varphi^2)^2(1+\varphi^2)} = \frac{3}{(1-\varphi^2)^2}. \tag{1.12}$$

Nyní dosadíme (1.11) a (1.12) do $\operatorname{var} X_t^2$ a dostáváme

$$\operatorname{var} X_t^2 = \frac{3}{(1-\varphi^2)^2} - \frac{1}{(1-\varphi^2)^2} = \frac{2}{(1-\varphi^2)^2}. \tag{1.13}$$

Nyní odvodíme autokovarianci kvadrátů veličin AR(1), kterou lze díky stacionaritě procesu a $\mathsf{E}\, X_t = 0$ zapsat ve tvaru

$$\sigma(X_t^2, X_{t+\tau}^2) = \mathsf{E}\,(X_t^2 X_{t+\tau}^2) - (\operatorname{var} X_t)^2. \tag{1.14}$$

Veličinu $X_{t+\tau}^2$ budeme iterativně vyjadřovat pomocí (1.4) až do času t

$$X_{t+\tau}^2 = \varphi^2 X_{t+\tau-1}^2 + 2\varphi X_{t+\tau-1}\epsilon_{t+\tau} + \epsilon_{t+\tau}^2 = \cdots =$$
$$= \varphi^{2\tau} X_t^2 + 2\varphi \sum_{i=1}^{\tau} \varphi^{2i-2} X_{t+\tau-i}\epsilon_{t+\tau-(i-1)} + \sum_{i=0}^{\tau-1} \varphi^{2i}\epsilon_{t+\tau-i}^2.$$

Po vynásobení veličinou X_t^2 a aplikaci střední hodnoty dostaneme

$$\mathsf{E}\,(X_t^2 X_{t+\tau}^2) = \varphi^{2\tau}\,\mathsf{E}\; X_t^4 + 2\varphi \sum_{i=1}^{\tau} \varphi^{2i-2}\,\mathsf{E}\,(X_t^2 X_{t+\tau-i}\epsilon_{t+\tau-(i-1)}) +$$
$$+ \sum_{i=0}^{\tau-1} \varphi^{2i}\,\mathsf{E}\,(X_t^2 \epsilon_{t+\tau-i}^2).$$

Po dosazení (1.12) a využití tvrzení 3.6 z Kulich (2018, str. 19) a (1.6) máme

$$\mathsf{E}\,(X_t^2 X_{t+\tau}^2) = \varphi^{2\tau}\frac{3}{(1-\varphi^2)^2} + (\sum_{i=0}^{\tau-1}\varphi^{2i})\,\mathsf{E}\; X_t^2.$$

Po sečtení konečné geometrické posloupnosti dosadíme do vzorce (1.14). Střední hodnota kvadrátů veličin z AR(1) je rovna rozptylu ze vzorce (1.11). Tím dostaneme

$$\sigma(X_t^2, X_{t+\tau}^2) = \frac{3\varphi^{2\tau}}{(1-\varphi^2)^2} + \frac{1-\varphi^{2\tau}}{(1-\varphi^2)^2} - \frac{1}{(1-\varphi^2)^2} = \frac{2\varphi^{2\tau}}{(1-\varphi^2)^2}. \tag{1.15}$$

Autokorelaci kvadrátů veličin AR(1) získáme ihned dosazením (1.15) a (1.13) do (1.1)

$$\rho_{\tau,X^2} = \varphi^{2\tau}. \tag{1.16}$$

Kapitola 2

Výběrová autokorelace

Ve druhé kapitole se podrobněji podíváme na výběrové autokorelace $R_{\tau,X}$ z hlediska variability. Nejprve odvodíme odhad rozptylu výběrových autokorelací, poté se budeme zabývat asymptotickými vlastnostmi tohoto odhadu.

2.1 Odhad rozptylu výběrových autokorelací

Odhad rozptylu výběrových autokorelací $R_{\tau,X}$ jako náhodných veličin provedeme za námi zadaných předpokladů dle Taylor (2007, str. 118)

P1 $\sigma(X_t, X_{t+\tau}) = 0 \quad \forall t \in T \quad \forall \tau > 0$,

P2 $\mathsf{E}\, X_t = 0 \quad \forall t \in T$,

P3 sdružená hustota f_X náhodných veličin $X_1, \ldots, X_n$ pro libovolné $n \in \mathbb{N}$ a libovolné hodnoty $x_1, \ldots, x_n \in \mathbb{R}$ je symetrická ve smyslu rovnosti sdružených hustot $f_X(x_1, \ldots, x_n) = f_X(|x_1|, \ldots, |x_n|)$.

Předpoklad **P3** určuje hypotézu H_S a předpoklad **P1** je ekvivaletní s hypotézou H_0, obě definované v Taylor (2007, str.118), a při označení $x = (x_1, \ldots, x_n)^\top$ platí

$$H_S : f(x) = f(|x|) \quad \forall x \in \mathbb{R}^n \implies H_0 : \sigma(X_t, X_{t+\tau}) = 0 \quad \forall t \in T \quad \forall \tau > 0. \tag{2.1}$$

Zformulujeme a poté dokážeme tvrzení, které ukáže, že předpoklad **P3** splňuje nekorelovaný náhodný proces z normálního rozdělení.

Tvrzení 1. *Nechť $\{X_t\}$ je nekorelovaný náhodný proces z normálního rozdělení ($X_t \sim N(0, \sigma_t^2)$, $\sigma_t^2 > 0 \quad \forall t \in T$). Označme $x = (x_1, \ldots, x_n)^\top$ a $|x| = (|x_1|, \ldots, |x_n|)^\top$, pak pro libovolné $n \in \mathbb{N}$ je sdružená hustota f_X symetrická ve smyslu rovnosti $f_X(x) = f_X(|x|)$ $\forall x \in \mathbb{R}^n$.*

Důkaz. Z předpokladu normálního rozdělení a nekorelovanosti veličin X_t vyplývá, že sdružené rozdělení náhodného vektoru $X = (X_1, \dots, X_n)^\top$ je n-rozměrné normální s nulovým vektorem středních hodnot a varianční maticí

$$\Sigma = \begin{pmatrix} \sigma_1^2 & 0 & 0 & \dots & 0 \\ 0 & \sigma_2^2 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \sigma_n^2 \end{pmatrix}.$$

Pro takto zadané n-rozměrné normální rozdělení máme hustotu

$$f_X(x) = \frac{1}{(2\pi)^{\frac{n}{2}} \sqrt{|\Sigma|}} \operatorname{exp}(-\frac{1}{2}\, x^\top \Sigma^{-1} x), x \in \mathbb{R}^n.$$

Spočteme determinant varianční matice, její inverzi a $x^\top \Sigma^{-1} x$:

$$\begin{aligned} \sqrt{|\Sigma|} &= \prod_{t=1}^{n} \sigma_t, \\ \Sigma^{-1} &= \operatorname{diag}(\frac{1}{\sigma_1^2}, \frac{1}{\sigma_2^2}, \dots, \frac{1}{\sigma_n^2}), \\ x^\top \Sigma^{-1} x &= \sum_{t=1}^{n} \frac{x_t^2}{\sigma_t^2}. \end{aligned}$$

Po dosazení dostaneme přehlednější vzorec pro hustotu n-rozměrného normálního rozdělení za daných předpokladů

$$f_X(x) = \frac{1}{(2\pi)^{\frac{n}{2}} \prod_{t=1}^{n} \sigma_t} \operatorname{exp}(-\frac{1}{2} \sum_{t=1}^{n} \frac{x_t^2}{\sigma_t^2}), x \in \mathbb{R}^n. \tag{2.2}$$

Předpokládaná nekorelovanost je ve sdruženém normálním rozdělení ekvivalentní nezávislosti. Sdruženou hustotu vektoru X jsme tedy dostali jako součin marginálních hustot složek X.

Zkoumáme symetrii ve smyslu rovnosti sdružených n-rozměrných hustot v bodech x a $|x|$ $\forall x \in \mathbb{R}^n$. Snadno v (2.2) nahlédneme, že rovnost $f_X(x) = f_X(|x|)$ $\forall x \in \mathbb{R}^n$ platí.

□

Vzhledem k nulovým středním hodnotám náhodných veličin X_t můžeme přepsat vzorec pro výběrovou autokorelaci v definici 7

$$r_{\tau,X} = \frac{\sum_{t=1}^{n-\tau} x_t\, x_{t+\tau}}{\sum_{t=1}^{n} x_t^2}, \tau \geq 1. \tag{2.3}$$

I zde je $r_{\tau,X}$ realizací nahodné veličiny $R_{\tau,X}$ tak jako ve vzorci (1.3). Dále označme

$$\alpha_\tau = \operatorname{var}(R_{\tau,X}). \tag{2.4}$$

Odhad rozptylu výběrových autokorelací bude založen na předpokladu symetrie mnohorozměrné hustoty.

Nechť $x_t^*, 1 \leq t \leq n$ je posloupnost, pro kterou platí, že každý prvek je buď x_t, nebo $-x_t$. Za předpokladu **P3** o symetrii mají takové posloupnosti stejnou sdruženou hustotu, jsou stejně pravděpodobné a způsobují různé realizace náhodné veličiny $R_{\tau,X}$ při pevném časovém kroku τ. Mějme podmíněné diskrétní rozdělení náhodné veličiny $R_{\tau,X}$ za podmínky $|X_t| = |x_t|, t = 1, \ldots, n$ a v následujícím tvrzení odvodíme rozptyl tohoto rozdělení.

Tvrzení 2 (Taylor, 2007, str. 119)**.** *Nechť platí předpoklady **P1**, **P2**, **P3**. Pak podmíněné diskrétní rozdělení náhodné veličiny $R_{\tau,X}$ za podmínky $|X_t| = |x_t|$, $t = 1, \ldots, n$ má rozptyl*

$$a_\tau = \frac{\sum_{t=1}^{n-\tau} x_t^2\, x_{t+\tau}^2}{(\sum_{t=1}^{n} x_t^2)^2}, \tag{2.5}$$

který je nestranným odhadem α_τ ze vzorce (2.4).

Důkaz. Označme $m_t = |x_t|, M_t = |X_t|, 1 \leq t \leq n$ a n-rozměrné vektory m, M

$$m = (m_1, \ldots, m_n)^\top = (|x_1|, \ldots, |x_n|)^\top,$$
$$M = (M_1, \ldots, M_n)^\top = (|X_1|, \ldots, |X_n|)^\top.$$

Budeme uvažovat X_t jako spojitou náhodnou veličinu. Definujme diskrétní náhodnou veličinu π_t předpisem

$$X_t = \pi_t\, M_t \tag{2.6}$$

s hodnotami -1, 1. Tímto dostáváme $X_t^2 = \pi_t^2\, M_t^2 = M_t^2$. Zavedeme n-rozměrný vektor

$$\pi = (\pi_1, \ldots, \pi_n)^\top.$$

Pravděpodobnost každé možné realizace vektoru π je určena integrálem z hustoty f_X náhodného vektoru X přes oblast $(x_j > 0\, \forall j \in J, x_k < 0\ \forall k \notin J)$. Množinou J myslíme libovolnou podmnožinu indexové množiny $\{1, 2, \ldots, n\}$ (včetně prázdné podmnožiny). Uvažujeme pouze nenulové hodnoty, nulu vylučujeme vzhledem k její nulové pravděpodobnosti u spojitého rozdělení. Z předpokladu **P3** o symetrii hustoty plyne, že každý takový integrál se rovná integrálu přes oblast $(x_j > 0, 1 \leq j \leq n)$.

Dostaneme 2^n možných realizací vektoru π, všechny se stejnou pravděpodobností $1/2^n$. Pak zřejmě musí libovolná složka $\pi_t, t = 1, \ldots, n$ nabývat obou hodnot se stejnou pravděpodobností $\mathrm{P}(\pi_t = 1) = \mathrm{P}(\pi_t = -1) = 1/2$. Náhodné veličiny $\pi_1, \ldots, \pi_n$ jsou tudíž nezávislé, stejně rozdělené.

Nyní můžeme podmíněný rozptyl zapsat jako $\mathrm{var}(R_{\tau,X}|M = m)$. Budeme ho počítat z definice podmíněného rozptylu:

$$\mathrm{var}(R_{\tau,X}|M = m) = \mathsf{E}\,(R_{\tau,X}^2|M = m) - (\mathsf{E}\,(R_{\tau,X}|M = m))^2.$$

Přepíšeme veličinu $R_{\tau,X}$ za pomocí (1.3) a (2.3):

$$R_{\tau,X} = \frac{\sum_{t=1}^{n-\tau} \pi_t\, \pi_{t+\tau}\, M_t\, M_{t+\tau}}{\sum_{t=1}^{n} M_t^2}.$$

Při $M = m$ můžeme psát

$$\mathsf{E}\,(R_{\tau,X}|M = m) = \frac{\sum_{t=1}^{n-\tau} \mathsf{E}\,(\pi_t\, \pi_{t+\tau})\, m_t\, m_{t+\tau}}{\sum_{t=1}^{n} m_t^2},$$

$$\mathsf{E}\,(\pi_t\, \pi_{t+\tau}) = \mathsf{E}\ \pi_t\ \mathsf{E}\ \pi_{t+\tau} = (\mathsf{E}\ \pi_t)^2 = (-1 \cdot \frac{1}{2} + 1 \cdot \frac{1}{2})^2 = 0. \tag{2.7}$$

Z toho plyne, že $\mathrm{var}(R_{\tau,X}|M = m) = \mathsf{E}\,(R_{\tau,X}^2|M = m)$. Toto spočítáme na následujících řádcích:

$$\mathsf{E}\,(R_{\tau,X}^2|M = m) = \frac{\mathsf{E}\,(\sum_{t=1}^{n-\tau} \pi_t\, \pi_{t+\tau}\, m_t\, m_{t+\tau})^2}{(\sum_{t=1}^{n} m_t^2)^2}.$$

Ve jmenovateli je konstanta, takže dále budeme upravovat pouze čitatele:

$$\begin{aligned}\mathsf{E}\,(\sum_{t=1}^{n-\tau} \pi_t\, \pi_{t+\tau}\, m_t\, m_{t+\tau})^2 &= \mathsf{E}\,(\sum_t \sum_s \pi_t\, \pi_s\, \pi_{t+\tau}\, \pi_{s+\tau}\, m_t\, m_s\, m_{t+\tau}\, m_{s+\tau}) = \\ &= \sum_t \mathsf{E}\,(\pi_t\, \pi_{t+\tau}\, m_t\, m_{t+\tau})^2 + \sum_t \sum_{s \neq t} m_t\, m_s\, m_{t+\tau}\, m_{s+\tau}\ \mathsf{E}\,(\pi_t\, \pi_s\, \pi_{t+\tau}\, \pi_{s+\tau}).\end{aligned}$$

Podrobněji se podíváme na střední hodnotu:

$$\begin{aligned}\mathsf{E}\,(\pi_t\, \pi_{t+\tau}\, m_t\, m_{t+\tau})^2 &= \mathsf{E}\,(\pi_t^2\, \pi_{t+\tau}^2\, m_t^2\, m_{t+\tau}^2) = m_t^2\, m_{t+\tau}^2, \\ \mathsf{E}\,(\pi_t\, \pi_s\, \pi_{t+\tau}\, \pi_{s+\tau}) &= \mathsf{E}\ \pi_t\ \mathsf{E}\ \pi_s\ \mathsf{E}\ \pi_{t+\tau}\ \mathsf{E}\ \pi_{s+\tau} = (\mathsf{E}\ \pi_t)^4 = 0.\end{aligned}$$

Konečně dostáváme

$$\mathrm{var}(R_{\tau,X}|M = m) = \mathsf{E}\,(R_{\tau,X}^2|M = m) = \frac{\sum_{t=1}^{n-\tau} m_t^2\, m_{t+\tau}^2}{(\sum_{t=1}^{n} m_t^2)^2}. \tag{2.8}$$

Vzhledem k tomu, že $x_t^2 = \pi_t^2\, m_t^2$, kde vždy platí $\pi_t^2 = 1$, dostáváme $x_t^2 = m_t^2$. Díky tomu máme v (2.8) a_τ ze vzorce (2.5). Musíme ještě dokázat nestrannost tohoto odhadu pro α_τ.

Vezměme a_τ jako pozorovanou hodnotu náhodné veličiny A_τ a dokážeme rovnost $\mathsf{E}\ A_\tau = \alpha_\tau$. Víme díky (2.6) a vztahům odvozeným výše v důkazu následující

$$\begin{aligned}A_\tau &= \frac{\sum_{t=1}^{n-\tau} M_t^2\, M_{t+\tau}^2}{(\sum_{t=1}^{n} M_t^2)^2}, \\ \mathsf{E}\,(A_\tau|M = m) &= \frac{\sum_{t=1}^{n-\tau} m_t^2\, m_{t+\tau}^2}{(\sum_{t=1}^{n} m_t^2)^2} = \mathsf{E}\,(R_{\tau,X}^2|M = m) \quad \forall m.\end{aligned}$$

Protože máme rovnost podmíněných středních hodnot při všech možných realizacích $M = m$, platí s pravděpodobností 1, že $\mathsf{E}\,(A_\tau|M) = \mathsf{E}\,(R^2_{\tau,X}|M)$, což díky vlastnostem podmíněné střední hodnoty znamená rovnost

$$\mathsf{E}\ A_\tau = \mathsf{E}\ R^2_{\tau,X}.$$

Víme, že $\mathrm{var}(R_{\tau,X}) = \mathsf{E}\ R^2_{\tau,X} - (\mathsf{E}\ R_{\tau,X})^2$, kde platí

$$\mathsf{E}\ R_{\tau,X} = \mathsf{E}\,(\mathsf{E}\ R_{\tau,X}|M),$$

což pro všechny realizace M je rovno nule, jak jsme dokázali v (2.7). Tedy máme vzhledem k (2.4)

$$\mathsf{E}\ A_\tau = \mathsf{E}\ R^2_{\tau,X} = \mathrm{var}(R_{\tau,X}) = \alpha_\tau.$$

□

2.2 Asymptotické vlastnosti odhadu rozptylu výběrových autokorelací

V této části se zaměříme na odvození asymptotických vlastností odhadu rozptylu výběrových autokorelací a_τ. Budeme uvažovat odhad založený na n pozorováních časové řady, který budeme dále značit $a_{\tau,n}$. Protože se budeme zabývat asymptotickými vlastnostmi, zavedeme tři typy konvergence, které využijeme v samotných vlastnostech či jejich odvozování. Definice jsou převzaty z Kulich (2018, str.31).

Definice 9 (konvergence v pravděpodobnosti)**.** *Říkáme, že posloupnost náhodných vektorů $\{X_n\}_{n=1}^\infty$ konverguje v pravděpodobnosti k náhodnému vektoru X pro n rostoucí nade všechny meze ($X_n \xrightarrow{p} X$) právě když*

$$\forall \epsilon > 0 : \lim_{n\to+\infty} P[||X_n - X|| > \epsilon] = 0.$$

Definice 10 (konvergence skoro jistě)**.** *Říkáme, že posloupnost náhodných vektorů $\{X_n\}_{n=1}^\infty$ konverguje skoro jistě k náhodnému vektoru X pro n rostoucí nade všechny meze ($X_n \xrightarrow{sj} X$) právě když*

$$P[\lim_{n\to+\infty} ||X_n - X|| = 0] = 1.$$

Definice 11 (konvergence v distribuci)**.** *Říkáme, že posloupnost náhodných vektorů $\{X_n\}_{n=1}^\infty$ konverguje v distribuci k náhodnému vektoru X pro n rostoucí nade všechny meze ($X_n \xrightarrow{D} X$) právě když*

$$\lim_{n\to+\infty} F_{X_n}(x) = F_X(x) \quad \forall x,$$

kde je distribuční funkce $F_X(x)$ spojitá.

Poznámka. Uvedeme vztahy mezi definovanými konvergencemi dle tvrzení 7.1 z Kulich (2018, str.31):

$$X_n \xrightarrow{sj} X \implies X_n \xrightarrow{p} X \implies X_n \xrightarrow{D} X. \tag{2.9}$$

Pro odvozování asymptotických vlastností budeme předpokládat, že časová řada je pozorovanou realizací striktně stacionárního a ergodického procesu dle definice 3 a definice 6.1 z Prášková (2016, str.77) a dále dle Taylor (2007, str.121) má X_t nulovou střední hodnotu a konečnou špičatost. Nepředpokládáme, že náhodný proces je nekorelovaný, obecně tedy pracujeme se vzájemně závislými veličinami. Označíme $b_{\tau,n} = na_{\tau,n}$, kde $a_{\tau,n}$ je definováno vzorcem (2.5), a $b_{\tau,n}$ tak zapíšeme ve tvaru

$$b_{\tau,n} = \frac{\frac{1}{n}\sum_{t=1}^{n-\tau} x_t^2\, x_{t+\tau}^2}{(\frac{1}{n}\sum_{t=1}^{n} x_t^2)^2}. \tag{2.10}$$

Uvažujme $b_{\tau,n}$ jako pozorovanou hodnotu náhodné veličiny

$$B_{\tau,n} = \frac{\frac{1}{n}\sum_{t=1}^{n-\tau} X_t^2\, X_{t+\tau}^2}{(\frac{1}{n}\sum_{t=1}^{n} X_t^2)^2}. \tag{2.11}$$

Nyní vyslovíme první asymptotickou vlastnost, kterou posléze dokážeme.

Tvrzení 3. *Nechť máme striktně stacionární a ergodický náhodný proces* $\{X_t\}$ *spojitě rozdělený s nulovou střední hodnotou, kladným rozptylem a konečnou špičatostí, pak pro* $n \to \infty$ *platí*

$$B_{\tau,n} \xrightarrow{sj} \beta_\tau = \frac{\mathsf{E}[X_t^2 X_{t+\tau}^2]}{(\mathsf{E}[X_t^2])^2}. \tag{2.12}$$

Důkaz. Za uvedených předpokladů platí podle Birkhoffovy ergodické věty v Seidler (2020, věta 3.1, str. 22) a následně podle poznámky 3.4 v Seidler (2020, str. 35) pro $n \to \infty$ konvergence

$$\frac{1}{n}\sum_{t=1}^{n-\tau} X_t^2\, X_{t+\tau}^2 \xrightarrow{sj} \mathsf{E}\,[X_t^2 X_{t+\tau}^2], \tag{2.13}$$

$$(\frac{1}{n}\sum_{t=1}^{n} X_t^2)^2 \xrightarrow{sj} (\mathsf{E}\,[X_t^2])^2. \tag{2.14}$$

Dále dokážeme, že limita podílu je v našem případě podíl limit. Zapíšeme čitatel a jmenovatel výrazu (2.11) do vektoru

$$\begin{pmatrix} \frac{1}{n}\sum_{t=1}^{n-\tau} X_t^2\, X_{t+\tau}^2 \\ \\ (\frac{1}{n}\sum_{t=1}^{n} X_t^2)^2 \end{pmatrix}.$$

Nyní dle tvrzení 7.4 v Kulich (2018, str.32) vidíme, že konverguje-li vektor skoro jistě po složkách, tak jako v (2.13) a (2.14), konverguje celý vektor skoro jistě k vektoru s limitami složek

$$\begin{pmatrix} \mathsf{E}\,[X_t^2 X_{t+\tau}^2] \\ \\ (\mathsf{E}\,[X_t^2])^2 \end{pmatrix}.$$

Následně ve smyslu tvrzení 7.3 z Kulich (2018, str.32) definujeme transformační funkci g:

$$g\begin{pmatrix} Y \\ Z \end{pmatrix} = \frac{Y}{Z},$$

která je spojitá až na $z = 0$. Vzhledem k tomu, že máme

$$Y = \frac{1}{n}\sum_{t=1}^{n-\tau} X_t^2\, X_{t+\tau}^2, \tag{2.15}$$

$$Z = (\frac{1}{n}\sum_{t=1}^{n} X_t^2)^2, \tag{2.16}$$

je transformace g spojitá skoro všude na nosiči rozdělení prvků časové řady. Spojitá tato transformace nebude kvůli kvadrátu v druhé složce vektoru pouze, když $x_t = 0\ \forall t$. Pak z tvrzení 7.3 v Kulich (2018, str.32) plyne, že konvergence vektoru implikuje konvergenci jeho transformace g. Tím je dokázáno, že limita podílu je podíl limit, a první asymptotická vlastnost je tím potvrzena.

□

Vzorec (2.12) pro β_τ dokážeme díky vlastnostem kovariance a předpokládaným nulovým středním hodnotám upravit, jak je ukázáno v následujícím tvrzení.

Tvrzení 4. *Nechť máme striktně stacionární a ergodický náhodný proces $\{X_t\}$ s nulovou střední hodnotou, k_X označuje konečnou špičatost prvků X_t a ρ_{τ,X^2} označuje autokorelaci náhodného procesu $\{X_t^2\}$, pak lze β_τ zapsat jako*

$$\beta_\tau = 1 + (k_X - 1)\rho_{\tau,X^2}. \tag{2.17}$$

Důkaz. Využijeme vlastnosti kovariance pro veličiny náhodného procesu $\{X_t^2\}$. Použijeme tedy vztah

$$\sigma(X_t^2, X_{t+\tau}^2) = \mathsf{E}\,[X_t^2 X_{t+\tau}^2] - \mathsf{E}\,[X_t^2]\,\mathsf{E}\,[X_{t+\tau}^2].$$

Díky striktní stacionaritě náhodného procesu $\{X_t\}$ jsou druhé momenty stejné a dostaneme

$$\sigma(X_t^2, X_{t+\tau}^2) = \mathsf{E}\,[X_t^2 X_{t+\tau}^2] - (\mathsf{E}\,[X_t^2])^2.$$

Když dosadíme do vzorce (2.12) za $\mathsf{E}\,[X_t^2 X_{t+\tau}^2]$, dostaneme

$$\beta_\tau = \frac{\sigma(X_t^2, X_{t+\tau}^2) + (\mathsf{E}\,[X_t^2])^2}{(\mathsf{E}\,[X_t^2])^2} = \frac{\sigma(X_t^2, X_{t+\tau}^2)}{(\mathsf{E}\,[X_t^2])^2} + 1. \tag{2.18}$$

Protože se přičtená 1 vyskytuje ve vzorci (2.17) i (2.18), dále se budeme zabývat jen podílem kovariance a střední hodnoty kvadrátu v (2.18). Ve vztahu (2.17) rozepíšeme špičatost a autokorelaci

$$k_X - 1 = \frac{\mathsf{E}\,[X_t - \mathsf{E}\ X_t]^4}{(\mathrm{var}\, X_t)^2} - 1 = \frac{\mathsf{E}\,[X_t^2]^2}{(\mathsf{E}\,[X_t^2])^2} - 1 = \frac{\mathsf{E}\,[X_t^2]^2 - (\mathsf{E}\,[X_t^2])^2}{(\mathsf{E}\,[X_t^2])^2},$$

$$\rho_{\tau,X^2} = \frac{\sigma(X_t^2, X_{t+\tau}^2)}{\sqrt{\mathrm{var}\, X_t^2}\sqrt{\mathrm{var}\, X_{t+\tau}^2}} = \frac{\sigma(X_t^2, X_{t+\tau}^2)}{\mathrm{var}(X_t^2)} = \frac{\sigma(X_t^2, X_{t+\tau}^2)}{\mathsf{E}\,[X_t^2]^2 - (\mathsf{E}\,[X_t^2])^2}.$$

Druhou rovnost v upravování $k_X - 1$ můžeme napsat díky předpokládané nulové střední hodnotě. V úpravě ρ_{τ,X^2} dle definice 6 můžeme psát druhou rovnost díky shodným rozptylům.

Po vynásobení $k_X - 1$ a ρ_{τ,X^2} dostaneme hodnotu, kterou máme i v (2.18).

□

Druhou asymptotickou vlastnost můžeme odvodit díky tomu, že jsme již v tvrzení 2 dokázali, že $a_{\tau,n}$ je nestranným odhadem rozptylu výběrových autokorelací $\mathrm{var}(R_{\tau,X})$.

Tvrzení 5. *Nechť $a_{\tau,n}$ je nestranným odhadem rozptylu výběrových autokorelací, pak*

$$n\, var(R_{\tau,X}) \rightarrow \beta_\tau. \tag{2.19}$$

Důkaz. Ze vztahu (2.9) vidíme, že z konvergence skoro jistě plyne konvergence v distribuci. Z tvrzení 3 víme, že platí $B_{\tau,n} = nA_{\tau,n} \xrightarrow{sj} \beta_\tau$. Tedy platí i konvergence v distribuci $nA_{\tau,n} \xrightarrow{D} \beta_\tau$.

Dále v důsledku Helly – Brayovy věty v Kulich (2018, tvrzení 7.2.ii, str.32) víme, že konvergence v distribuci je ekvivaletní s konvergencí středních hodnot. Potom přímo dostáváme

$$\mathsf{E}\,[nA_{\tau,n}] = n\,\mathsf{E}\,[A_{\tau,n}] = n\,\mathrm{var}(R_{\tau,X}) \rightarrow \mathsf{E}\ \beta_\tau = \beta_\tau.$$

□

Dokázali jsme (2.19) za předpokladu nestranného odhadu $a_{\tau,n}$ rozptylu výběrových autokorelací podle tvrzení 2. Předpoklady tvrzení 2 splňuje například bílý šum z rozdělení $N(0,1)$. Pokud nejsou splněny předpoklady tvrzení 2 **P1**, **P2**, **P3**, odhady $a_{\tau,n}$ a $b_{\tau,n}$ mohou být vychýlené. To je například v situaci, když je proces korelovaný, tedy není splněno **P1**, což dle (2.1) implikuje nesplnění mnohorozměrné symetrie hustoty v **P3**.

Jako příklad náhodného procesu, který není nekorelovaný, ale je stacionární při $|\varphi| < 1$, jsme uvedli autoregresní proces prvního řádu v definici 8. Mějme AR(1) se striktním bílým šumem $\{\epsilon_t\}$ z normovaného normálního rozdělení. Spočteme pro něj hodnotu β_τ pomocí (2.17). Při $\mathsf{E}\, X_t = 0$ můžeme psát špičatost veličin X_t jako $k_X = \mathsf{E}\, X_t^4/(\operatorname{var} X_t)^2$. Dosadíme odvozené vzorce (1.12) a (1.11), za autokorelaci kvadrátů veličin AR(1) dosadíme ze vzorce (1.16). Pro AR(1) lze psát β_τ jako

$$\beta_\tau = 1 + 2\varphi^{2\tau}. \tag{2.20}$$

Dle Kendall a Stuart (1976, str.472) ale $n \operatorname{var}(R_{\tau,X})$ v AR(1) nekonverguje k výše spočtené hodnotě β_τ, nýbrž

$$n \operatorname{var}(R_{\tau,X}) \rightarrow \lambda_\tau = \frac{(1+\varphi^2)(1-\varphi^{2\tau})}{1-\varphi^2} - 2\tau\varphi^{2\tau} \neq \beta_\tau. \tag{2.21}$$

Následující tabulky ilustrují na různých hodnotách φ a τ odlišnost (2.20) a (2.21). Hodnoty v tabulce jsou zaokrouhleny na 5 desetinných míst.

	$\varphi = 0.1$	$\varphi = 0.5$	$\varphi = 0.9$
$\tau = 1$	0,99	0,75	0,19
$\tau = 2$	1,0197	1,3125	0,6517
$\tau = 5$	1,0202	1,65527	2,71791
$\tau = 10$	1,0202	1,66665	5,93661
$\tau = 20$	1,0202	1,66667	8,79457
$\tau = 50$	1,0202	1,66667	9,52341
$\tau \to \infty$	1,0202	1,66667	9,52632

Tabulka 2.1: Hodnoty výrazu $\lambda_\tau = \frac{(1+\varphi^2)(1-\varphi^{2\tau})}{1-\varphi^2} - 2\tau\varphi^{2\tau}$

	$\varphi = 0.1$	$\varphi = 0.5$	$\varphi = 0.9$
$\tau = 1$	1,02	1,5	2,62
$\tau = 2$	1,0002	1,125	2,3122
$\tau = 5$	1	1,00195	1,69736
$\tau = 10$	1	1	1,24315
$\tau = 20$	1	1	1,02956
$\tau = 50$	1	1	1,00005
$\tau \to \infty$	1	1	1

Tabulka 2.2: Hodnoty výrazu $\beta_\tau = 1 + 2\varphi^{2\tau}$

Kapitola 3

Statistické testy o autokorelacích

V této kapitole rozebereme testování hypotézy o nekorelovanosti časové řady a úpravu dat, která je nutná pro splnění předpokladu o rozptylu výběrových autokorelací ve zmíněných testech.

3.1 Testování nekorelovanosti v časové řadě

Základem pro statistické testy o autokorelacích, ve kterých se snažíme obvykle rozhodnout, zda platí hypotéza o nekorelovanosti časové řady, je výběrová autokorelace jako náhodná veličina ze vzorce (1.3). Testové statistiky jsou založeny na R_τ. Potřebujeme u nich vždy znát alespoň asymptotické pravděpodobnostní rozdělení, které souvisí s asymptotickým rozdělením R_τ.

Podrobněji jsme se zabývali autoregresním procesem prvního řádu při $|\varphi| < 1$, kdy je slabě stacionární a lze ho zapsat ve tvaru lineárního procesu. Ukážeme proto asymptotické pravděpodobnostní rozdělení R_τ, formulaci hypotézy H_0 a 3 různé testové statistiky pro test nulovosti autokorelací u lineárního procesu.

Uvedeme tvrzení z Taylor (2007, str.25) o asymptotickém mnohorozměrném normálním rozdělení pro vektor výběrových autokorelací $R = (R_1, \ldots, R_k)^\top$.

Tvrzení 6. *Nechť máme za předpokladu* $\mathsf{E}\, X_t = 0\ \forall t$ *lineární proces v obecném tvaru*

$$X_t = \sum_{j=0}^{\infty} b_j \epsilon_{t-j},$$

kde jsou ϵ_t *nezávislé, stejně rozdělené náhodné veličiny s konečným rozptylem a pro* b_j *platí konvergence řad* $\sum_{j\geq 0} |b_j|$ *a* $\sum_{j\geq 0} j b_j^2$*. Pak asymptoticky platí pro výběrové autokorelace*

$$\sqrt{n}(R_1 - \rho_1, \ldots, R_k - \rho_k)^\top \sim N_k(0_k, W_k),$$

kde 0_k je k-rozměrný nulový vektor středních hodnot a W_k je varianční matice, jejíž prvky jsou určeny celou posloupností autokorelací $\{\rho_\tau\}_{\tau>0}$ lineárního procesu.

Poznámka. Varianční matice W_k má na diagonále asymptotické rozptyly jednotlivých prvků $\sqrt{n}R_\tau$

$$\omega_{\tau\tau} = \lambda_0 + \lambda_{2\tau} + 2(\lambda_0\rho_\tau^2 - 2\lambda_\tau\rho_\tau) \tag{3.1}$$

a mimo diagonálu asymptotické kovariance mezi veličinami $\sqrt{n}R_\tau$ a $\sqrt{n}R_\xi$

$$\omega_{\tau\xi} = \lambda_{\xi-\tau} + \lambda_{\xi+\tau} + 2(\lambda_0\rho_\tau\rho_\xi - \lambda_\tau\rho_\xi - \lambda_\xi\rho_\tau), \tag{3.2}$$

kde

$$\lambda_i = \sum_{j=-\infty}^{+\infty} \rho_j\rho_{i+j} \tag{3.3}$$

vyjadřuje závislost na posloupnosti všech autokorelací (Taylor, 2007, str. 136).

Důkaz. Viz Anderson a Walker (1964, str.1296-1303).

□

Hypotézu náhodné procházky formulujeme dle Taylor (2007, str.133) ve tvaru

$$H_0* : \mathsf{E}\ X_t = \mathsf{E}\ X_{t+\tau}, \tag{3.4}$$
$$\sigma(X_t, X_{t+\tau}) = 0 \quad \forall t \in T \quad \forall\tau > 0, \tag{3.5}$$

kde (3.5) je ekvivalentní s $\rho(X_t, X_{t+\tau}) = 0\ \forall t\ \forall\tau > 0$. Zde není předpokládaná stacionarita procesu $\{X_t\}$. V případě alespoň slabě stacionárního procesu stačí testovat hypotézu

$$H_0 : \rho_\tau = 0 \quad \forall\tau > 0. \tag{3.6}$$

Nadále budeme předpokládat alespoň slabě stacionární proces.

Za platnosti H_0 z předpisu (3.6) lze asymptotické mnohorozměrné normální rozdělení z tvrzení 6 díky vyjádření prvků matice W_k v (3.1), (3.2) a (3.3) psát jako

$$\sqrt{n}(R_1, \ldots, R_k)^\top \sim N_k(0_k, I_k), \tag{3.7}$$

kde I_k je jednotková matice řádu k. Ta vznikne z W_k vynulováním všech autokorelací ve vzorcích (3.2) a (3.1) po dosazení (3.3) až na vyjádření λ_0 ve vzorci (3.1). Při $\rho_0 = 1$ dle vzorce (3.3) zůstane každý diagonální prvek jednotkový.

Pro každý jednotlivý prvek $R_\tau, \tau = 1, \ldots, k$ máme z (3.7) po vydělení $\sqrt{n}$ normální rozdělení

$$R_\tau \sim N(0, \frac{1}{n}). \tag{3.8}$$

Dostali jsme asymptotické rozdělení výběrové autokorelace R_τ, díky kterému za platnosti H_0 a splnění předpokladů tvrzení 6 dokážeme určit asymptotické rozdělení testových statistik.

Nyní si uvedeme tři testové statistiky založené na (3.8) pro testy H_0, jak jsou popsány v Taylor (2007, str.137-138). První jednoduchý test je prováděn za pomoci testové statistiky $\sqrt{n}R_1$, která se skládá pouze z jedné výběrové autokorelace. Myšlenka tohoto testu je následující. Najdeme-li statisticky významnou r_τ pro některé $\tau > 0$, zamítneme H_0. Zpravidla nás zajímá korelovanost sousedních veličin v náhodném procesu, proto zkoumáme r_1.

Za platnosti H_0 má $\sqrt{n}R_1$ asymptotické normované normální rozdělení, proto H_0 zamítáme na hladině $\alpha \in (0{,}1)$ právě tehdy, když

$$\sqrt{n}|r_1| > u_{1-\frac{\alpha}{2}}, \tag{3.9}$$

kde r_1 je vypočítaná hodnota výběrové autokorelace R_1 a $u_{1-\frac{\alpha}{2}}$ je kvantil normovaného normálního rozdělení. Pro $\alpha = 0{,}05$ je $u_{1-\frac{\alpha}{2}} = 1{,}96$.

Pokud chceme pracovat s více výběrovými autokorelacemi, zavedeme testovou statistiku

$$\boldsymbol{\alpha}(k, N_0) = \sum_{i=N_0}^{k} \frac{k!}{i!(k-i)!}\alpha^i(1-\alpha)^{k-i}, \alpha \in (0,1). \tag{3.10}$$

Definujeme ji jako pravděpodobnost, že N_0 nebo více výběrových autokorelací z celkového počtu k je statisticky významných na základě testu typu (3.9) na hladině α za platnosti H_0. Testová statistika $\boldsymbol{\alpha}(k, N_0)$ je součtem pravděpodobností binomického rozdělení $Bi(k, \alpha)$. Hypotézu H_0 lze zamítnout, pokud $\boldsymbol{\alpha}(k, N_0) \leq \alpha$, α je zvolená hladina testu.

Potřebujeme mít pevně dané hodnoty k a N_0. Volíme $k = 28$ a $N_0 = 4$ podle Taylor (2007, str.137-138). Pak dostaneme pro $\alpha = 0{,}05$, že přibližná hodnota pravděpodobnosti $\boldsymbol{\alpha}(28, 4)$ je 0,049. Je-li tedy N_r počet statisticky významných výběrových autokorelací z celkových $r_1, \ldots, r_{28}$ na hladině $\alpha = 0{,}05$,
H_0 zamítáme, když $N_r \geq 4$.

Poslední testová statistika, kterou uvedeme, je založená na celém vektoru
$R = (R_1, \ldots, R_k)^\top$ a má tvar

$$Q_k = n \sum_{\tau=1}^{k} R_\tau^2 = nR^\top R. \tag{3.11}$$

Za platnosti H_0 máme pro $\sqrt{n}R$ rozdělení z (3.7). Ze zápisu statistiky Q_k ve vzorci (3.11) je díky (3.8) ihned patrné, že $Q_k \sim \chi^2_k$.

Hypotézu H_0 zamítáme na hladině α, když po dosazení vypočítaných hodnot $r_1, \ldots, r_k$ platí

$$Q_k > \chi^2_k(1-\alpha),$$

kde $\chi^2_k(1-\alpha)$ je příslušný kvantil rozdělení χ^2_k. Pro $k = 28$ a $\alpha = 0{,}05$ je

$$\chi^2_k(1 - \alpha) = 41{,}34.$$

Představili jsme testy a jejich statistiky pro testování hypotézy H_0 založené na tom, že R_τ má podle tvrzení 6 asymptoticky normální rozdělení s var $R_\tau = 1/n$.

Pro nelineární nekorelovaný proces neplatí tvrzení 6 a var R_τ může být výrazně větší než $1/n$ (Taylor, 2007, str. 25). Za pomoci tvrzení 2 a (2.10) odhadujeme $n \operatorname{var} R_\tau$ veličinou b_τ. Pokud je b_τ výrazně větší než 1, naznačuje to, že var R_τ může být větší než $1/n$. Na základě uvedených testů tedy nelze korektně zamítat hypotézu nekorelovanosti H_0. Můžeme pouze tvrdit, že veličiny X_t nejsou nezávislé, stejně rozdělené (Taylor, 2007, str. 25, 126).

3.2 Úprava dat z časové řady

V praxi pozorujeme časové řady s nenulovou střední hodnotou, předpokládejme proto $\mathsf{E}\, X_t = \mu$. Střední hodnotu odhadneme výběrovým průměrem $\overline{X}$ a pozorovanou řadu $x_1, \ldots, x_n$ centrujeme. Výběrové autokorelace počítáme podle (1.3), což změní i odhad $\operatorname{var}(R_{\tau,X})$ ze vzorce (2.5). Vztah (2.10) pro b_τ pak dle Taylor (2007, str.124) přepíšeme do tvaru

$$b_\tau = \frac{n \sum_{t=1}^{n-\tau} (x_t - \bar{x})^2 (x_{t+\tau} - \bar{x})^2}{(\sum_{t=1}^{n} (x_t - \bar{x})^2)^2}. \tag{3.12}$$

Při praktické analýze časových řad se ukázalo, že b_τ z (3.12) jsou výrazně větší než 1 (Taylor, 2007, str.125). Z toho plyne, že $R_{\tau,X}$ mají rozptyl větší než $1/n$. Použití testů nekorelovanosti není v takové situaci vhodné. Vysoký rozptyl veličin $R_{\tau,X}$ může být způsoben známou skutečností, že finanční časové řady nejsou stacionární v rozptylu a mají nekonstantní podmíněný rozptyl při známé minulosti. Potřebujeme tak transformovat časovou řadu, aby se homogenizoval její rozptyl.

Data dále potřebujeme upravit tak, abychom dostali

$$\operatorname{var}(R_\tau) \simeq 1/n \quad \forall \tau > 0.$$

Uvažujme podle Taylor (2007, str.68) faktorizaci procesu s nekonstantním podmíněným rozptylem ve tvaru $X_t = \mu + V_t U_t$, kde $\{U_t\}$ je striktní bílý šum s nulovou střední hodnotou a jednotkovým rozptylem a V_t je podmíněná směrodatná odchylka veličiny X_t při známé minulosti $\{X_{t+\tau}; \tau > 0\}$. Pak máme

$$U_t = \frac{X_t - \mu}{V_t}. \tag{3.13}$$

Realizace v_t náhodné veličiny V_t nejsou pozorovatelné, ale dle Taylor (2007, str.127) lze vytvořit aproximaci realizací u_t veličin U_t odhadem střední hodnoty μ výběrovým průměrem a vhodnou predikcí hodnoty v_t ze známé minulosti označenou jako $\hat{v}_t$. Pak dostaneme s využitím (3.13) transformované hodnoty y_t ve tvaru

$$y_t = \frac{x_t - \bar{x}}{\hat{v}_t} \simeq \frac{x_t - \mu}{\hat{v}_t} = u_t \frac{v_t}{\hat{v}_t}. \tag{3.14}$$

Zde vidíme, že hodnoty u_t jsou vhodně aproximovány pomocí hodnot y_t, pokud je $\hat{v}_t$ kvalitní predikce.

V Taylor (2007, str.128) je počáteční predikce z času $t = 21$ a dále je predikce definována rekurzí

$$\hat{v_{21}} = 1{,}253 \sum_{t=1}^{20} \frac{|x_t - \bar{x}|}{20},$$

$$\hat{v}_t = (1-\gamma)\hat{v_{t-1}} + 1{,}253\gamma|x_t - \bar{x}|,$$

kde γ je parametr pro danou časovou řadu, který závisí na její konkrétní povaze. V Taylor (2007, str. 128) je doporučeno pro řady výnosů z akcií $\gamma = 0{,}04$ a v ostatních případech $\gamma = 0{,}1$.

Pro $t = 21, \dots, n$ pak transformujeme pozorovanou časovou řadu podle (3.14) a výběrové autokorelace počítáme podle předpisu

$$r_{\tau,Y} = \frac{\sum_{t=21}^{n-\tau}(y_t - \bar{y})(y_{t+\tau} - \bar{y})}{\sum_{t=21}^{n}(y_t - \bar{y})^2}, \tag{3.15}$$

a při $n^* = n - 20$ upravená data y_t dávají požadované $\mathrm{var}(R_{\tau,Y}) \simeq 1/n^*$.

Navíc pokud je splněna hypotéza H_S (z předpokladu **P3**) pro náhodný proces $\{X_t\}$, tak je splněna i pro proces $\{Y_t\}$ při zanedbání rozdílu mezi průměrem a střední hodnotou. Veličinu $n^*(\mathrm{var}(R_{\tau,Y}))$ lze tedy odhadnout pomocí

$$b_\tau^* = (n-20)\frac{\sum_{t=21}^{n-\tau}(y_t - \bar{y})^2(y_{t+\tau} - \bar{y})^2}{(\sum_{t=21}^{n}(y_t - \bar{y})^2)^2}, \tag{3.16}$$

který, jak bylo prokázáno v praxi, nabývá hodnot blízkých 1 (Taylor, 2007, str.128).

V tabulkách z Taylor (2007, str. 125,130) porovnáme mediány a průměry odhadů $n\,\mathrm{var}(R_{\tau,X})$ z (3.12) a $n^*\,\mathrm{var}(R_{\tau,Y})$ z (3.16).

	Medián					Průměr				
	$\tau=1$	$\tau=2$	$\tau=3$	$\tau=4$	$\tau=5$	$\tau=1$	$\tau=2$	$\tau=3$	$\tau=4$	$\tau=5$
b_τ	1.81	1.43	1.43	1.47	1.36	1.79	1.46	1.47	1.43	1.44
b_τ^*	1.27	1.06	1.02	1.05	1.00	1.36	1.09	1.08	1.04	0.99

Tabulka 3.1: Mediány a průměry b_τ, b_τ^* přes 15 řad výnosů akcií amerických firem

Vidíme v tabulce 3.1 již na hodnotách mediánů a průměrů b_τ, b_τ^* spočtených přes relativně malý vzorek finančních časových řad, každá délky $n = 2750$, že se b_τ a b_τ^* liší řádově v desetinách. Rovněž vidíme, že průměrné hodnoty a mediány b_τ výrazně překračují 1, zatímco stejné hodnoty pro b_τ^* se pohybují až na $\tau = 1$ blízko 1. Hodnoty b_τ výrazně větší než 1 ukazují, že pro řady výnosů z akcií (v tomto případě 15 amerických firem) je $\mathrm{var}(R_\tau)$ větší než $1/n$, což odpovídá situaci popsané na konci kapitoly 3.1. Navíc dle Taylor (2007, str. 57) se řady výnosů z akcií řídí spíše nelineárními modely než lineárními.

Kapitola 4

Simulační studie

V simulační studii se zaměříme na generování časových řad a výpočet průměrů a výběrových směrodatných odchylek vypočítaných hodnot $b_{\tau,n}$ veličiny $B_{\tau,n}$ z (2.11), které odhadují $n\,\mathrm{var}(R_{\tau,X})$. Budeme vyšetřovat rychlost konvergence těchto odhadů. Dále určíme empirickou hladinu a sílu testů se statistikami (3.9), (3.10) a (3.11).

Budeme zkoumat model bílého šumu z $N(0,1)$ a proces AR(1) s totožným bílým šumem. Vždy budeme generovat 1000 realizací dané časové řady pro různé délky a různé volby parametru φ v modelu AR(1). Simulace a výpočty budou prováděny v softwaru Mathematica.

První model je bílý šum $\{\epsilon_t\}_{t=1}^n$ z $N(0,1)$. Délky časové řady n bereme 50, 100, 500, 1000 a časový posun τ volíme 1, 2, 5, 10, 20 podobně jako v tabulkách 2.1 a 2.2.

Jako druhý model máme proces AR(1) definovaný vztahem (1.4), kde $\{\epsilon_t\}$ je bílý šum z $N(0,1)$. Hodnoty n a τ volíme jako v předchozím modelu. Navíc bereme různé hodnoty autoregresního parametru φ, a to stejně jako v tabulkách 2.1 a 2.2, tedy $\varphi \in \{0{,}1; 0{,}5; 0{,}9\}$.

4.1 Přesnost odhadů

V první části simulace se zaměříme na rychlost konvergence odhadů $n\,\mathrm{var}(R_{\tau,X})$, tedy veličin $B_{\tau,n}$. Střední hodnotu veličiny $B_{\tau,n}$ odhadneme průměrem hodnot $b_{\tau,n}$ spočítaným přes 1000 realizací jednotlivých typů časových řad. Budeme zkoumat konvergenci $\mathsf{E}\, B_{\tau,n}$ k β_τ dle tvrzení 3 a Helly – Brayovy věty v Kulich (2018, tvrzení 7.2.ii, str.32) pomocí rychlosti přiblížení průměru odhadů $b_{\tau,n}$ k β_τ. Z (2.17) plyne, že $\beta_\tau = 1$ v případě bílého šumu z $N(0,1)$. Hodnoty β_τ v procesech AR(1) s uvedenými volbami parametru φ a časového posunu τ jsou zapsány v tabulce 2.2. Také bychom měli pozorovat klesající výběrovou směrodatnou odchylku odhadu $b_{\tau,n}$ s rostoucím počtem pozorovování (rostoucí délkou n časové řady).

	$\tau = 1$	$\tau = 2$	$\tau = 5$	$\tau = 10$	$\tau = 20$
$n = 50$	0.946	0.917	0.868	0.774	0.558
$n = 100$	0.974	0.962	0.930	0.888	0.774
$n = 500$	0.990	0.995	0.984	0.975	0.951
$n = 1000$	0.995	0.996	0.991	0.990	0.978

Tabulka 4.1: Průměry $b_{\tau,n}$ v bílém šumu z $N(0,1)$

	$\tau = 1$	$\tau = 2$	$\tau = 5$	$\tau = 10$	$\tau = 20$
$n = 50$	0.257	0.252	0.250	0.241	0.192
$n = 100$	0.179	0.183	0.189	0.187	0.175
$n = 500$	0.090	0.087	0.090	0.085	0.086
$n = 1000$	0.062	0.062	0.062	0.062	0.064

Tabulka 4.2: Směrodatné odchylky $b_{\tau,n}$ v bílém šumu z $N(0,1)$

Pro normálně rozdělený bílý šum platí ekvivalence se striktním bílým šumem. Díky nezávislosti veličin z náhodného procesu jsou splněny předpoklady **P1**, **P2**, **P3** a $B_{\tau,n}$ je v důsledku tvrzení 2 a (2.10) nestranný odhad $n\,\text{var}(R_{\tau,X})$. Podle 5 pak $n\,\text{var}(R_{\tau,X})$ konverguje k $\beta_\tau = 1$. Z tabulky 4.1 lze vidět, že se hodnoty průměrů se zvyšujícím se n pro všechny zvolené časové posuny τ blíží k 1. Nízké hodnoty průměrů v krátkých časových řadách při relativně velkém časovém posunu jsou způsobeny nedostatečným počtem pozorování. Také vidíme v tabulce 4.2 že při rostoucím n se výběrové směrodatné odchylky odhadů zmenšují.

		$\tau = 1$	$\tau = 2$	$\tau = 5$	$\tau = 10$	$\tau = 20$
$\varphi = 0.1$	$n = 50$	0.955	0.919	0.867	0.761	0.564
	$n = 100$	0.989	0.961	0.939	0.879	0.775
	$n = 500$	1.010	0.994	0.984	0.976	0.951
	$n = 1000$	1.014	0.995	0.991	0.991	0.977
$\varphi = 0.5$	$n = 50$	1.301	0.992	0.841	0.749	0.574
	$n = 100$	1.437	1.061	0.924	0.870	0.774
	$n = 500$	1.480	1.107	0.985	0.975	0.951
	$n = 1000$	1.494	1.116	0.991	0.987	0.976
$\varphi = 0.9$	$n = 50$	1.909	1.551	0.997	0.682	0.476
	$n = 100$	2.177	1.835	1.241	0.889	0.707
	$n = 500$	2.501	2.184	1.576	1.151	0.946
	$n = 1000$	2.556	2.245	1.628	1.182	0.982

Tabulka 4.3: Průměry $b_{\tau,n}$ v AR(1)

V modelu AR(1) při splnění předpokladů tvrzení 3 platí stejně jako u bílého šumu, že $B_{\tau,n}$ se s rostoucím n blíží k β_τ. To lze zjistit porovnáním tabulek 4.3 a 2.2 a v souhrnné tabulce 4.5. Nejsou ale splněny předpoklady **P1** a **P3** z tvrzení 2, a tedy již neplatí, že $B_{\tau,n}$ je nestranný odhad pro $n\,\text{var}(R_{\tau,X})$.

Naopak je vidět z porovnání tabulek 4.3 a 2.1 a ze souhrnné tabulky 4.6, že se průměry odhadů $b_{\tau,n}$ výrazně liší od hodnot limity λ_τ pro $n\,\text{var}(R_{\tau,X})$.

		$\tau = 1$	$\tau = 2$	$\tau = 5$	$\tau = 10$	$\tau = 20$
$\varphi = 0.1$	$n = 50$	0.248	0.251	0.248	0.233	0.200
	$n = 100$	0.180	0.186	0.187	0.187	0.184
	$n = 500$	0.090	0.086	0.090	0.085	0.085
	$n = 1000$	0.064	0.063	0.063	0.062	0.064
$\varphi = 0.5$	$n = 50$	0.374	0.269	0.241	0.236	0.200
	$n = 100$	0.322	0.229	0.191	0.192	0.182
	$n = 500$	0.166	0.122	0.093	0.094	0.092
	$n = 1000$	0.117	0.084	0.069	0.067	0.069
$\varphi = 0.9$	$n = 50$	0.487	0.465	0.365	0.262	0.240
	$n = 100$	0.478	0.429	0.304	0.247	0.214
	$n = 500$	0.400	0.385	0.326	0.227	0.169
	$n = 1000$	0.298	0.287	0.245	0.178	0.133

Tabulka 4.4: Směrodatné odchylky $b_{\tau,n}$ v AR(1)

		$\tau = 1$	$\tau = 2$	$\tau = 5$	$\tau = 10$	$\tau = 20$
$\varphi = 0.1$	$n = 50$	0.955	0.919	0.867	0.761	0.564
	$n = 100$	0.989	0.961	0.939	0.879	0.775
	$n = 500$	1.010	0.994	0.984	0.976	0.951
	$n = 1000$	1.014	0.995	0.991	0.991	0.977
	$\beta_\tau(n \to \infty)$	1.002	1	1	1	1
$\varphi = 0.5$	$n = 50$	1.301	0.992	0.841	0.749	0.574
	$n = 100$	1.437	1.061	0.924	0.870	0.774
	$n = 500$	1.480	1.107	0.985	0.975	0.951
	$n = 1000$	1.494	1.116	0.991	0.987	0.976
	$\beta_\tau(n \to \infty)$	1.500	1.125	1.002	1	1
$\varphi = 0.9$	$n = 50$	1.909	1.551	0.997	0.682	0.476
	$n = 100$	2.177	1.835	1.241	0.889	0.707
	$n = 500$	2.501	2.184	1.576	1.151	0.946
	$n = 1000$	2.556	2.245	1.628	1.182	0.982
	$\beta_\tau(n \to \infty)$	2.620	2.312	1.697	1.243	1.030

Tabulka 4.5: Porovnání průměrů $b_{\tau,n}$ v AR(1) a β_τ

		$\tau = 1$	$\tau = 2$	$\tau = 5$	$\tau = 10$	$\tau = 20$
	$n = 50$	0.955	0.919	0.867	0.761	0.564
	$n = 100$	0.989	0.961	0.939	0.879	0.775
$\varphi = 0.1$	$n = 500$	1.010	0.994	0.984	0.976	0.951
	$n = 1000$	1.014	0.995	0.991	0.991	0.977
	$\lambda_\tau(n\to\infty)$	0.990	1.020	1.020	1.020	1.020
	$n = 50$	1.301	0.992	0.841	0.749	0.574
	$n = 100$	1.437	1.061	0.924	0.870	0.774
$\varphi = 0.5$	$n = 500$	1.480	1.107	0.985	0.975	0.951
	$n = 1000$	1.494	1.116	0.991	0.987	0.976
	$\lambda_\tau(n\to\infty)$	0.750	1.313	1.655	1.667	1.667
	$n = 50$	1.909	1.551	0.997	0.682	0.476
	$n = 100$	2.177	1.835	1.241	0.889	0.707
$\varphi = 0.9$	$n = 500$	2.501	2.184	1.576	1.151	0.946
	$n = 1000$	2.556	2.245	1.628	1.182	0.982
	$\lambda_\tau(n\to\infty)$	0.190	0.651	2.718	5.937	8.795

Tabulka 4.6: Porovnání průměrů $b_{\tau,n}$ v AR(1) a λ_τ

Tato odlišnost se výrazně projeví při větších hodnotách autoregresního parametru φ, to znamená při více korelovaných veličinách v procesu AR(1). Stejně jako u bílého šumu se výběrové směrodatné odchylky odhadů $b_{\tau,n}$ v AR(1) změnšují při zvětšujícím se n, viz tabulka 4.4.

4.2 Hladina a síla testů

V druhé části simulační studie budeme zkoumat empirickou hladinu a sílu testů z kapitoly 3.1 v generovaných časových řadách z první části simulace (modely bílého šumu a AR(1)). Hladina testu je pravděpodobnost, se kterou test zamítne platnou hypotézu. V teoretických popisech testů na nekorelovanost časové řady předpokládáme hladinu testu $\alpha = 0{,}05$. Naopak síla testu proti dané alternativě vyjadřuje pravděpodobnost zamítnutí neplatné hypotézy.

Empirické odhady hladiny a síly testů určíme jako relativní četnost řad, ve kterých byla hypotéza zamítnuta. Empirickou hladinu testu takto zjistíme v modelu splňujícím hypotézu o nekorelovanosti (například normálně rozdělený bílý šum, pro který je nekorelovanost ekvivaletní s nezávislostí veličin ϵ_t). Empirickou sílu testu zjišťujeme v modelu nesplňujícím hypotézu o nekorelovanosti (například korelovaný proces AR(1)).

$n = 50$	$n = 100$	$n = 500$	$n = 1000$
0.041	0.044	0.057	0.047

Tabulka 4.7: Empirická hladina testu se statistikou $\sqrt{n}|r_1|$

		Empirická síla testu
$\varphi = 0.1$	$n = 50$	0.052
	$n = 100$	0.131
	$n = 500$	0.563
	$n = 1000$	0.874
$\varphi = 0.5$	$n = 50$	0.890
	$n = 100$	0.998
	$n = 500$	1
	$n = 1000$	1
$\varphi = 0.9$	$n = 50$	1
	$n = 100$	1
	$n = 500$	1
	$n = 1000$	1

Tabulka 4.8: Empirická síla testu se statistikou $\sqrt{n}|r_1|$

Vidíme v tabulce 4.7, že jsme pro test (3.9) s rostoucí délkou n jednoho tisíc generovaných časových řad dostali stabilní empirickou hladinu testu blízkou předepsaným 5 %. Empirická síla testu v tabulce 4.8 pro časové řady výrazněji odlišné od bílého šumu je již pro krátkou délku řady blízká 1 nebo přímo 1. Stoprocetní síla testu zde znamená, že v simulovaných AR(1) byla ve všech 1000 realizacích zamítnuta nulová hypotéza H_0. V případě AR(1) s parametrem $\varphi = 0{,}1$, který není příliš vzdálený bílému šumu, je takový test přijatelně silný až při větším množství pozorovaných dat (délka $n = 1000$).

$n = 50$	$n = 100$	$n = 500$	$n = 1000$
0.012	0.022	0.034	0.044

Tabulka 4.9: Empirická hladina testu se statistikou $\boldsymbol{\alpha}(k, N_0)$

Test (3.10) založený na testové statistice $\boldsymbol{\alpha}(k, N_0)$ provádíme pro volby $k = 28$ a $N_0 = 4$. S rostoucí délkou časové řady sice empirická hladina testu roste dle tabulky 4.9, ale je stále nižší než předepsaná hladina 0,05. Dále lze v tabulce 4.10 vidět, že pro korelované procesy reprezentované AR(1) s parametry $\varphi = 0{,}1$ a $\varphi = 0{,}5$ není test dostatečně silný. Až s rostoucím n pro AR(1) s $\varphi = 0{,}5$ získává test na síle. Pro proces s nejvýraznější korelační strukturou je test uspokojivě silný.

Z tabulky 4.11 lze pozorovat, že test (3.11) pro $k = 28$ má podobné hodnoty empirické hladiny testu jako test (3.10), ale podle tabulky 4.12 je silnější. Ani to ale nezajišťuje dostatečnou sílu proti alternativě AR(1) s parametrem $\varphi = 0{,}1$. Obecně pro oba testy platí, že nejsou konstruovány jako silné testy proti konkrétní alternativě (Taylor, 2007, str. 138), což lze v tabulkách 4.10 a 4.12 snadno pozorovat. V případě alternativy AR(1) je silnější test (3.9).

		Empirická síla testu
$\varphi = 0.1$	$n = 50$	0.006
	$n = 100$	0.033
	$n = 500$	0.104
	$n = 1000$	0.131
$\varphi = 0.5$	$n = 50$	0.195
	$n = 100$	0.445
	$n = 500$	0.802
	$n = 1000$	0.888
$\varphi = 0.9$	$n = 50$	0.888
	$n = 100$	0.992
	$n = 500$	1
	$n = 1000$	1

Tabulka 4.10: Empirická síla testu se statistikou $\boldsymbol{\alpha}(k, N_0)$

$n = 50$	$n = 100$	$n = 500$	$n = 1000$
0.009	0.019	0.045	0.049

Tabulka 4.11: Empirická hladina testu se statistikou Q_k

		Empirická síla testu
$\varphi = 0.1$	$n = 50$	0.011
	$n = 100$	0.032
	$n = 500$	0.140
	$n = 1000$	0.314
$\varphi = 0.5$	$n = 50$	0.357
	$n = 100$	0.839
	$n = 500$	1
	$n = 1000$	1
$\varphi = 0.9$	$n = 50$	0.974
	$n = 100$	1
	$n = 500$	1
	$n = 1000$	1

Tabulka 4.12: Empirická síla testu se statistikou Q_k

Závěr

Cílem práce bylo odvodit odhad rozptylu výběrových autokorelací $\mathrm{var}(R_{\tau,X})$ a jeho asymptotické vlastnosti. Dále studovat asymptotické rozdělení výběrových autokorelací pro účely konstrukce testových statistik pro test hypotézy náhodné procházky a popsat tyto testové statistiky. Konečně v praktické části ověřit teoretické poznatky prostřednictvím simulační studie.

Nejprve jsme zavedli AR(1) jako konkrétní příklad stacionárního korelovaného procesu a odvodili potřebné momenty pro určení autokovariancí v AR(1) a kvadrátu tohoto procesu, což bylo důležité pro další odvozování. Potom jsme za daných předpokladů **P1**, **P2** a **P3** odvodili nestranný odhad A_τ pro $\mathrm{var}(R_{\tau,X})$. Pro jeho n násobek $B_{\tau,n}$ jsme našli limitu β_τ ve dvou ekvivaletních vyjádřeních v (2.12) a (2.17). Posloupnost $n\,\mathrm{var}(R_{\tau,X})$ má v nekorelovaném procesu limitu β_τ, v AR(1), nesplňujícím předpoklady **P1** a **P3**, má odlišnou limitu. To je ilustrováno v numerické studii.

Dále jsme popsali hypotézu náhodné procházky a její vybrané testy založené na autokorelacích. Nakonec jsme ukázali možný přístup k transformaci dat z časové řady, abychom dosáhli požadovaného rozptylu výběrových autokorelací pro spolehlivé testy náhodné procházky.

V simulační části jsme potvrdili, že v normálně rozděleném bílém šumu $b_{\tau,n}$ pro rostoucí délku řad konverguje k $\beta_\tau = 1$. Uspokojivé přiblížení nastává zejména u nižších hodnot časového posunu τ pro řady délky stovek pozorování. V AR(1) odhady $b_{\tau,n}$ konvergují k teoretické hodnotě β_τ, ale jsou vychýlené od skutečné limity pro $n\,\mathrm{var}(R_{\tau,X})$. Rovněž jsme v těchto modelech ukázali, že test nekorelovanosti (3.9) je velmi silný proti konkrétní alternativě AR(1) při předepsané hladině testu 0,05. Naopak testy (3.10) a (3.11) neprokázaly takovou sílu proti zmíněné alternativě. Při výrazné odlišnosti dat od nulové hypotézy a dostatečném množství pozorování jsou však uspokojivě silné. Empirická hladina všech 3 zkoumaných testů je nižší než teoretických 5 %.

Literatura

Anderson, T. W. a Walker, A. M. (1964). On the Asymptotic Distribution of the Autocorrelations of a Sample from a Linear Stochastic Process. *The Annals of Mathematical Statistics*, **35**(3).

Kendall, M. a Stuart, A. (1976). *The Advanced Theory of Statistics.* Třetí vydání. Griffin & Co., London.

Kulich, M. (2018). Základy teorie pravděpodobnosti. `https://www2.karlin.mff.cuni.cz/~zichova/FinMat/PravdepodobnostKulich.pdf`. Souhrn pro předmět Matematická statistika 1.

Prášková, Z. (2016). *Základy náhodných procesů II.* Druhé upravené vydání. Matfyzpress, Praha. ISBN 978-80-246-3529-3.

Seidler, J. (2020). *Přednášky o ergodické teorii.* Matfyzpress, Praha.

Taylor, S. (2007). *Modelling Financial Time Series.* Druhé vydání. World Scientific, Singapur. ISBN 978-981-277-084-4.

Seznam tabulek