



**MATEMATICKO-FYZIKÁLNÍ
FAKULTA**
Univerzita Karlova

DIPLOMOVÁ PRÁCE

Karolína Böhmová

Poissonovská autoregrese

Katedra pravděpodobnosti a matematické statistiky

Vedoucí diplomové práce: RNDr. Šárka Hudecová, Ph.D.

Studijní program: Matematika

Studijní obor: Pravděpodobnost, matematická statistika
a ekonometrie

Praha 2019

Prohlašuji, že jsem tuto diplomovou práci vypracoval(a) 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 má právo na uzavření licenční smlouvy o užití této práce jako školního díla podle §60 odst. 1 autorského zákona.

V dne

Podpis autora

Tímto bych chtěla poděkovat RNDr. Šárce Hudecové, Ph.D. za její vstřícnost, ochotu a cenné rady při psaní této práce.

Název práce: Poissonovská autoregrese

Autor: Karolína Böhmová

Katedra: Katedra pravděpodobnosti a matematické statistiky

Vedoucí diplomové práce: RNDr. Šárka Hudecová, Ph.D., Katedra pravděpodobnosti a matematické statistiky

Abstrakt: Práce se zabývá modelováním časových řad počtů událostí pomocí INGARCH modelů. Hlavní důraz je kladen na lineární INARCH model. Jsou zde odvozené jeho vlastnosti a metody odhadu parametrů modelu (metoda maximální věrohodnosti, metoda nejmenších čtverců a její modifikace), které jsou později porovnané pomocí simulací. Uvedené jsou i vlastnosti a odhad metodou maximální věrohodnosti INGARCH(1,1) modelu. Krátce jsou diskutované lineární INGARCH modely vyšších řádů a nelineární INGARCH modely. Použití modelů je ilustrované na několika časových řadách počtu nehod.

Klíčová slova: INARCH model, INGARCH model, Poissonovská autoregrese, časové řady počtů

Title: Poisson autoregression

Author: Karolína Böhmová

Department: Department of Probability and Mathematical Statistics

Supervisor: RNDr. Šárka Hudecová, Ph.D., Department of Probability and Mathematical Statistics

Abstract: This thesis deals with INGARCH models for a count time series. Main emphasis is placed on a linear INARCH model. Its properties are derived. Several methods of estimation are introduced - maximum likelihood method, least squares method and its modifications - and later compared in a simulation study. Main properties and maximum likelihood estimation for INGARCH(1,1) model are stated. Higher order linear INGARCH models and nonlinear INGARCH models are discussed briefly. An application of the presented models on time series of car accidents is given.

Keywords: INARCH model, INGARCH model, Poisson autoregression, count time series

Obsah

Úvod	2
1 Základní pojmy	3
1.1 Poissonovo rozdělení	3
1.2 Základní vlastnosti procesů	3
1.3 Podmíněná střední hodnota	4
1.4 Markovův řetězec	5
1.5 Autoregresní proces	6
1.6 Martingalové diference	7
2 Lineární INARCH(p) proces	9
2.1 Definice modelu	9
2.2 Momenty procesu	11
2.3 INARCH jako AR proces	15
2.4 Stacionarita a ergodicita	16
2.5 Odhad parametrů	16
2.5.1 Volba řádu modelu	16
2.5.2 Metoda maximální věrohodnosti	16
2.5.3 Metoda nejmenších čtverců	22
2.5.4 Metoda momentů	26
2.5.5 Diagnostika	26
3 Možná zobecnění INARCH modelu	27
3.1 Lineární INGARCH(p,q) model	27
3.1.1 Proces INGARCH(1,1)	28
3.1.2 Odhad parametrů modelu INGARCH(1,1)	29
3.2 Nelineární modely	30
4 Praktické použití	32
4.1 Implementace v R	32
4.2 Simulace	32
4.3 Reálná data	34
Závěr	42
Seznam použité literatury	43

Úvod

V dnešní době se v praxi často setkáváme s potřebou modelovat diskrétní časové řady (myšleno časové řady diskrétních náhodných veličin). V mnoha případech k tomu můžeme použít teorii pro spojité řady. To ale neplatí například pro řady s malými hodnotami. Zároveň nemůžeme použít ani zobecněné lineární modely, protože tyto řady jsou často silně korelované. Začalo se proto přemýšlet, jak modifikovat známé modely, aby se daly použít i na diskrétní řady.

Jedním z typů diskrétních řad jsou časové řady počtu událostí. Je přirozené vytvořit pro ně model, který pracuje s Poissonovo rozdělením. V literatuře se objevuje několik takových modelů (McKenzie, 2003), my se zaměříme na INARCH model, který předpokládá podmíněné Poissonovo rozdělení s podmíněnou střední hodnotou závislou na minulých pozorováních.

V první kapitole si připomeneme některé základní pojmy. Druhou kapitolu věnujeme právě lineárním INARCH modelům. Rozebereme jejich vlastnosti a vztah ke standardním AR modelům. Představíme i několik různých metod odhadu parametrů modelu. Ve třetí kapitole se podíváme na možná zobecnění, mezi které patří například INGARCH model nebo nelineární INARCH model. V poslední kapitole pak budeme popsanou metodologii ilustrovat na reálných datech. Uvedeme též simulační studii pro porovnání jednotlivých metod odhadu.

1. Základní pojmy

V první kapitole se zaměříme na vysvětlení vybraných pojmů, které využijeme v dalších částech této práce.

1.1 Poissonovo rozdělení

Náhodná veličina Y má *Poissonovo rozdělení*, jestliže nabývá hodnoty k z množiny $\{0, 1, 2, \dots\}$ s pravděpodobností

$$P(Y = k) = \frac{\lambda^k}{k!} e^{-\lambda},$$

kde $\lambda > 0$. Značíme $Y \sim \text{Po}(\lambda)$.

Toto rozdělení patří mezi rozdělení exponenciálního typu, neboť pravděpodobnostní funkci můžeme přepsat do tvaru

$$P(Y = k) = \exp \left\{ \frac{k\zeta - b(\zeta)}{\varphi} + c(k, \varphi) \right\}.$$

V případě Poissonova rozdělení $\zeta = \log \lambda$, $\varphi = 1$, $b(\zeta) = e^\lambda$ a $c(k, \varphi) = -\log k!$.

Specifickou vlastností Poissonova rozdělení je rovnost střední hodnoty a rozptylu

$$EY = \text{var } Y = \lambda.$$

Uvedme i druhý moment Poissonova rozdělení, neboť ho využijeme v následujících kapitolách

$$EY^2 = \lambda^2 + \lambda.$$

Ještě můžeme zmínit, že pro vysoké hodnoty λ se dá Poissonovo rozdělení aproximovat normálním rozdělením $\mathcal{N}(\lambda, \lambda)$. Pro celočíselné λ toto tvrzení vychází z centrální limitní věty použité na nezávislé stejně rozdělené náhodné veličiny X_1, X_2, \dots s rozdělením $\text{Po}(1)$

$$\frac{Y - \lambda}{\sqrt{\lambda}} \stackrel{\mathcal{D}}{=} \frac{1}{\sqrt{\lambda}} \sum_{i=1}^{\lambda} (X_i - 1) \xrightarrow[\lambda \rightarrow \infty]{\mathcal{D}} \mathcal{N}(0, 1).$$

Využíváme tu vlastnost Poissonova rozdělení, podle které součet nezávislých náhodných veličin X_1, \dots, X_n , kde $X_i \sim \text{Po}(\lambda_i)$, má Poissonovo rozdělení s parametrem $\lambda = \lambda_1 + \dots + \lambda_n$.

1.2 Základní vlastnosti procesů

Abychom mohli zkoumat některé vlastnosti procesu, potřebujeme, aby bylo jeho chování v jistém smyslu ustálené. Definujeme proto pojem striktní a slabé stacionarity.

Řekneme, že náhodný proces $\{Y_t, t \in T\}$ je *striktně stacionární*, jestliže pro libovolné $n \in \mathbb{N}$, pro libovolná y_1, \dots, y_n a pro libovolná t_1, \dots, t_n a h taková, že $t_k \in T$, $t_k + h \in T$, $1 \leq k \leq n$, platí

$$P(Y_{t_1} = y_1, \dots, Y_{t_n} = y_n) = P(Y_{t_1+h} = y_1, \dots, Y_{t_n+h} = y_n).$$

Důsledkem této vlastnosti je, že všechna Y_t mají stejné rozdělení.

Náhodný proces $\{Y_t, t \in T\}$ s konečnými druhými momenty se nazývá *slabě stacionární*, jestliže má konstantní střední hodnotu $EY_t = \mu$ pro všechna $t \in T$ a autokovarianční funkce $R(s, t)$ je funkcí pouze $t - s$.

Striktně stacionární proces s konečnými druhými momenty je i slabě stacionární.

1.3 Podmíněná střední hodnota

V celé této práci budeme pracovat s podmíněným rozdělením, připomeňme si proto definici některých základních pojmů.

Nechť $X \in \mathcal{L}_1(\Omega, \mathcal{A}, \mathbb{P})$ a $\mathcal{F} \subset \mathcal{A}$ je σ -algebra. *Podmíněná střední hodnota X při \mathcal{F}* je náhodná veličina $E(X|\mathcal{F})$ splňující

1. $E(X|\mathcal{F}) \in \mathcal{L}_1(\Omega, \mathcal{A}, \mathbb{P}_{/\mathcal{F}})$,
2. $\forall B \in \mathcal{F} : \int_B X d\mathbb{P} = \int_B E(X|\mathcal{F}) d\mathbb{P}$.

Podmíněnou střední hodnotu X při Y pak definujeme jako

$$E(X|Y) = E(X|\sigma(Y)) \text{ s.j.,}$$

podmíněnou pravděpodobnost jevu $A \in \mathcal{A}$ za podmínky Y jako

$$\mathbb{P}(A|Y) = E(1_A|Y),$$

a *podmíněný rozptyl X při Y* pro $X \in \mathcal{L}_2(\Omega, \mathcal{A}, \mathbb{P})$ jako

$$\text{var}(X|Y) = E\left([X - E(X|Y)]^2 | Y\right).$$

Uvedme si i pár vlastností podmíněné střední hodnoty a podmíněného rozptylu, které později využijeme.

Nechť $X, X_1, X_2 \in \mathcal{L}_1(\Omega, \mathcal{A}, \mathbb{P})$, $a, b, c \in \mathbb{R}$. Pak pro podmíněnou střední hodnotu platí:

1. $E(E(X|Y)) = EX$,
2. $E(aX_1 + bX_2 + c | Y) = aE(X_1|Y) + bE(X_2|Y) + c$ s.j.,
3. $E(f(Y)X|Y) = f(Y)E(X|Y)$ s.j.,
4. $E(E(X|Y) | f(Y)) = E(X|f(Y))$ s.j.

Navíc pokud $X \in \mathcal{L}_2(\Omega, \mathcal{A}, \mathbb{P})$, platí

$$\text{var}(X) = \text{var}(E(X|Y)) + E(\text{var}(X|Y)).$$

Důkazy a další vlastnosti můžeme nalézt například v knize Lachout (2004) nebo Anděl (2011).

1.4 Markovův řetězec

Důležitou třídou procesů jsou Markovovy řetězce. Charakteristickou vlastností těchto procesů je, že pokud známe současnou hodnotu, budoucí hodnoty procesu závisí pouze na této současné hodnotě.

Základní vlastnosti Markovových řetězců uvádí například Prášková a Lachout (1998), pokročilejšími vlastnostmi se zabývají mimo jiné Meyn a Tweedie (1993).

Proces $\{Y_t, t \in \mathbb{N}_0\}$ se nazývá *Markovův řetězec s množinou diskrétních stavů* S , jestliže

$$P(Y_t = i_t | Y_{t-1} = i_{t-1}, \dots, Y_0 = i_0) = P(Y_t = i_t | Y_{t-1} = i_{t-1})$$

pro všechna $t \geq 1$ a všechna $i_t, i_{t-1}, \dots, i_0 \in S$ taková, že pravděpodobnosti $P(Y_{t-1} = i_{t-1}, \dots, Y_0 = i_0)$ jsou nenulové.

Podmíněné pravděpodobnosti

$$p_{ij}(t, t+n) = P(Y_{t+n} = j | Y_t = i)$$

se nazývají *pravděpodobnosti přechodu ze stavu i v čase t do stavu j v čase $t+n$* , neboli pravděpodobnosti přechodu n -tého řádu.

Jestliže tyto pravděpodobnosti nezávisí na čase t , říkáme, že je Markovův řetězec *homogenní*. Pravděpodobnosti přechodu pak budeme značit $p_{ij}^{(n)}$. Pokud zapíšeme pravděpodobnosti přechodu 1.řádu do matice, dostaneme *matici pravděpodobností přechodu*

$$\mathbf{P} = \{p_{ij}^{(1)}\}_{i,j \in S}.$$

Ještě zmiňme, že *počátečním rozdělením* \mathbf{p} Markovova řetězce se myslí rozdělení Y_0 , tedy $\mathbf{p} = \{p_j, j \in S\}$, kde $p_j = P(Y_0 = j)$.

Mějme homogenní Markovův řetězec $\{Y_t\}$ a pravděpodobnostní rozdělení $\boldsymbol{\pi} = \{\pi_j, j \in S\}$ na množině S . Potom $\boldsymbol{\pi}$ se nazývá *stacionární rozdělení* řetězce $\{Y_t\}$, jestliže platí

$$\boldsymbol{\pi}^\top = \boldsymbol{\pi}^\top \mathbf{P}.$$

Každý Markovův řetězec s konečnou množinou stavů má nějaké stacionární rozdělení, Markovův řetězec s nekonečnou množinou stavů ale stacionární rozdělení mít nemusí. Předpokládejme, že stacionární rozdělení existuje a počáteční rozdělení je rovno tomuto rozdělení. Markovův řetězec je pak striktně stacionární.

Dále řekneme, že homogenní Markovský proces $\{Y_t, t \in \mathbb{N}_0\}$ je *ergodický*, jestliže platí

$$\forall i, j \in S : |p_{ij}^{(n)} - \pi_j| \longrightarrow 0, \quad n \rightarrow \infty,$$

a *silně ergodický*, jestliže platí

$$\sup_{i \in S} \sum_{j \in S} |p_{ij}^{(n)} - \pi_j| \longrightarrow 0, \quad n \rightarrow \infty.$$

Pro stacionární ergodické procesy platí například silný zákon velkých čísel (Stout, 1974), tedy

$$\frac{1}{n} \sum_{t=0}^n f(Y_t) \xrightarrow[n \rightarrow \infty]{} \mathbf{E}f(Y_0)$$

pro libovolnou funkci f takovou, že $\mathbf{E}|f(Y_t)| < \infty$. Totéž platí i pro funkce více proměnných jako $\Psi(Y_t, \dots, Y_{t+p})$ pro nějaké $p \in \mathbb{N}$.

Markovův řetězec se dá ještě zobecnit na *Markovův řetězec řádu p* . Ten je definován jako proces splňující

$$P(Y_t = i_t | Y_{t-1} = i_{t-1}, \dots, Y_0 = i_0) = P(Y_t = i_t | Y_{t-1} = i_{t-1}, \dots, Y_{t-p} = i_{t-p})$$

pro všechna $t \geq p$ a všechna $i_t, i_{t-1}, \dots, i_0 \in S$ taková, že pravděpodobnosti $P(Y_{t-1} = i_{t-1}, \dots, Y_0 = i_0)$ jsou nenulové.

Můžeme si všimnout, že pokud transformujeme Markovův řetězec řádu p na

$$\Phi_t = (Y_t, \dots, Y_{t-p+1})^\top,$$

získáme standardní Markovův řetězec. Všechny vlastnosti uvedené v této kapitole proto platí i pro Markovovy řetězce vyšších řádů.

1.5 Autoregresní proces

Nejprve připomeňme definici bílého šumu. Proces $\{\epsilon_t, t \in \mathbb{Z}\}$ je *bílým šumem*, jestliže jde o posloupnost nekorelovaných náhodných veličin s nulovou střední hodnotou a konečným kladným rozptylem σ^2 . Tento proces značíme $WN(0, \sigma^2)$. Nyní můžeme definovat autoregresní proces.

Mějme proces $\{Y_t, t \in \mathbb{Z}\}$, pro který platí

$$Y_t = a + b_1 Y_{t-1} + \dots + b_p Y_{t-p} + \epsilon_t,$$

kde $\{\epsilon_t\}$ je proces bílého šumu $WN(0, \sigma^2)$. Proces $\{Y_t\}$ pak nazýváme *autoregresní proces řádu p* , značíme $AR(p)$.

Vezměme si polynom $b(z) = 1 - b_1 z - \dots - b_p z^p$. Pokud má tento polynom všechny kořeny vně jednotkového kruhu, proces $AR(p)$ je slabě stacionární. Tato podmínka je například splněná, pokud $b_i \geq 0$ pro každé i a $b_1 + \dots + b_p < 1$.

Nejjednodušším příkladem je proces $AR(1)$ ve tvaru

$$Y_t = a + b_1 Y_{t-1} + \epsilon_t.$$

Za podmínky $|b_1| < 1$ je proces slabě stacionární a střední hodnota procesu je pro každé t rovna

$$EY_t = \frac{a}{1 - b_1}.$$

Můžeme snadno vyjádřit i autokovarianční funkci

$$R(k) = b_1^k \frac{\sigma^2}{1 - b_1^2}, \quad k \geq 0$$

a autokorelační funkci

$$r(k) = b_1^k, \quad k \geq 0.$$

Pokud rozšíříme autoregresní proces o zpožděné hodnoty bílého šumu, získáme tzv. *ARMA model*.

Mějme proces $\{Y_t, t \in \mathbb{Z}\}$, pro který platí

$$Y_t = a + b_1 Y_{t-1} + \dots + b_p Y_{t-p} + \epsilon_t + c_1 \epsilon_{t-1} + \dots + c_q \epsilon_{t-q},$$

kde $\{\epsilon_t\}$ je proces bílého šumu $WN(0, \sigma^2)$. Proces $\{Y_t\}$ pak nazýváme *proces ARMA*, konkrétně $ARMA(p, q)$.

Tento proces je stacionární, jestliže kořeny polynomu $b(z) = 1 - b_1 z - \dots - b_p z^p$ jsou vně jednotkového kruhu a zároveň polynomy $b(z)$ a $c(z) = 1 + c_1 z + \dots + c_q z^q$ nemají společné kořeny.

Speciálním případem tohoto procesu je $ARMA(1,1)$ ve tvaru

$$Y_t = a + b_1 Y_{t-1} + \epsilon_t + c_1 \epsilon_{t-1}.$$

Střední hodnota procesu je

$$EY_t = \frac{a}{1 - b_1 - c_1}$$

a autokovarianční funkce

$$R(0) = \frac{1 + 2 b_1 c_1 + c_1^2}{1 - b_1^2} \sigma^2, \quad (1.1)$$

$$R(k) = b_1^{k-1} \frac{(c_1 + b_1)(1 + b_1 c_1)}{1 - b_1^2} \sigma^2, \quad k \geq 1. \quad (1.2)$$

Vlastnosti AR i ARMA modelů jsou dokázané a více rozepsané v knize Cipra (2008).

1.6 Martingalové diference

Martingalové diference jsou obdobou bílého šumu. Posloupnost $\{D_t, t \in \mathbb{N}_0\}$, kde $D_t \in \mathcal{L}_1(\Omega, \mathcal{A}, \mathbb{P})$ pro každé t , se nazývá *posloupnost martingalových diferencí*, jestliže

$$E(D_{t+1} | D_t, \dots, D_0) = 0 \text{ s.j.}$$

pro každé t .

Stejně jako v případě bílého šumu jde o posloupnost nekorelovaných náhodných veličin, tedy

$$ED_t D_s = 0, \quad t, s \in \mathbb{N}_0,$$

s nulovou střední hodnotou, nijak ale neomezujeme rozptyl D_t .

Pokud tato posloupnost splňuje silný zákon velkých čísel ve smyslu kapitoly 1.4, považujeme posloupnost za ergodickou.

Výhodou posloupnosti martingalových diferencí je, že oproti bílému šumu pro ni za určitých podmínek (Billingsley (1961)) platí centrální limitní věta.

Věta 1. *Necht $\{D_t, t \in \mathbb{N}_0\}$ je striktně stacionární a ergodická posloupnost martingalových diferencí taková, že $ED_0^2 < \infty$. Pak platí*

$$\frac{1}{\sqrt{n}} \sum_{t=1}^n D_t \xrightarrow[n \rightarrow \infty]{\mathcal{D}} \mathcal{N}(0, ED_0^2).$$

Tuto větu můžeme rozšířit i na k -rozměrné posloupnosti martingalových diferencí. Předpokládejme, že $\{D_t\}$ je striktně stacionární a ergodická posloupnost

vektorů s konečnými druhými momenty. Pro $c^\top D_t$, kde $c \in \mathbb{R}^k$ libovolné, jsou všechny předpoklady věty (1) splněné, tedy

$$\frac{1}{\sqrt{n}} \sum_{t=1}^n c^\top D_t \xrightarrow[n \rightarrow \infty]{\mathcal{D}} c^\top X,$$

kde $c^\top X \sim \mathcal{N}(0, c^\top \Sigma c)$ a Σ je kovarianční matice vektoru D_0 . Z Cramér-Woldovy věty je pak za použití vlastnosti mnohorozměrného normálního rozdělení vidět, že

$$\frac{1}{\sqrt{n}} \sum_{t=1}^n D_t \xrightarrow[n \rightarrow \infty]{\mathcal{D}} \mathcal{N}(0, \text{var} D_0).$$

2. Lineární INARCH(p) proces

Existují různé metody, jak modelovat časové řady počtů událostí. Řady s vysokými hodnotami můžeme aproximovat veličinami s normálním rozdělením, což nám umožňuje použít například standardní AR model. S řadami s nízkými počty je to ale složitější. V literatuře se objevují různé modely, kterými se dají tyto řady modelovat. Jmenujme například INAR, INARCH nebo DAR modely (McKenzie, 2003). My se v této práci budeme zabývat právě INARCH modely.

2.1 Definice modelu

Mějme proces $\{Y_t, t \in \mathbb{N}_0\}$ a předpokládejme, že podmíněně svou minulostí má Y_t Poissonovo rozdělení se střední hodnotou lineárně závislou na předchozích p hodnotách procesu, tedy Y_{t-1}, \dots, Y_{t-p} . Formálně můžeme tento model zapsat ve tvaru

$$Y_t | \mathcal{F}_{t-1} \sim \text{Po}(\lambda_t), \quad \lambda_t = a + b_1 Y_{t-1} + \dots + b_p Y_{t-p}, \quad t \geq p, \quad (2.1)$$

kde $a > 0$, $b_1, \dots, b_p \geq 0$ jsou parametry modelu a $\mathcal{F}_t = \sigma(Y_s, 0 \leq s \leq t)$ je σ -algebra generovaná náhodnými veličinami Y_0, \dots, Y_t . Značíme INARCH(p). Z vlastností Poissonova rozdělení víme, že

$$\mathbb{E}(Y_t | \mathcal{F}_{t-1}) = \text{var}(Y_t | \mathcal{F}_{t-1}) = \lambda_t.$$

Podmínkou Poissonova rozdělení je kladná střední hodnota. Aby bylo λ_t pro všechna t kladné, museli jsme omezit $a > 0$ a $b_1, \dots, b_p \geq 0$. Tyto podmínky spolu s faktem, že $Y_t \geq 0$ s.j., zaručí, že střední hodnota bude vždy kladná.

Marginální rozdělení procesu INARCH(p) se nedá explicitně vyjádřit. Někteří autoři se proto rozhodli marginální rozdělení aproximovat, například pomocí pravděpodobností přechodu procesu (Weiß, 2010).

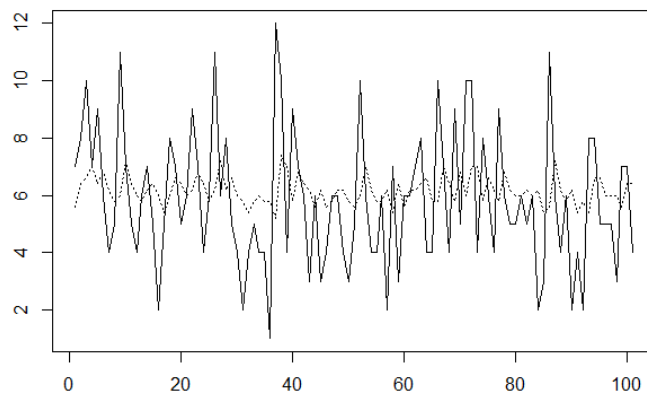
Můžeme si všimnout, že model (2.1) svou strukturou připomíná zobecněný lineární model s identickou linkovou funkcí a lineárním prediktorem. V našem případě jsou ale jako regresory zpožděné hodnoty procesu, nemůžeme tedy čekat nezávislost prvků vektoru odezvy.

Příklad. Nejjednodušší model, který spadá do třídy INARCH modelů, je model INARCH(1) ve tvaru

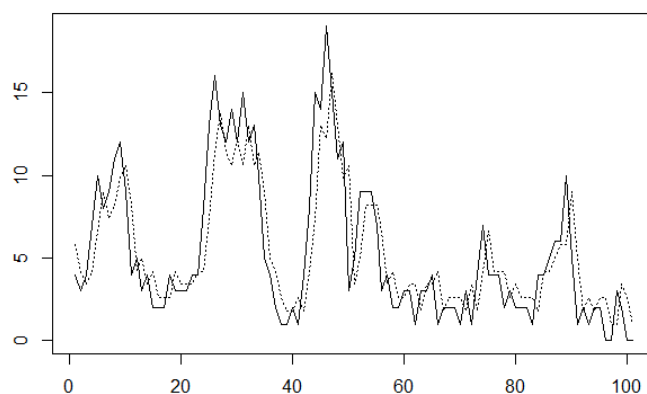
$$Y_t | \mathcal{F}_{t-1} \sim \text{Po}(\lambda_t), \quad \lambda_t = a + b_1 Y_{t-1}, \quad t \geq 1. \quad (2.2)$$

Hodnota parametru b_1 ovlivňuje určitou hladkost procesu. Čím je hodnota parametru nižší, tím se průběh procesu víc blíží bílému šumu (viz obrázky 2.1 a 2.2). Pokud by byl parametr b_1 dokonce nulový, znamenalo by to, že hodnoty procesu nejsou závislé na minulých hodnotách a $Y_t \sim \text{Po}(a)$ tvoří přímo proces bílého šumu.

Tomuto procesu se podrobně věnuje například Weiß (2010).



Obrázek 2.1: Realizace a podmíněná střední hodnota procesu INARCH(1) s parametry $a = 5$ a $b_1 = 0.2$.



Obrázek 2.2: Realizace a podmíněná střední hodnota procesu INARCH(1) s parametry $a = 1$ a $b_1 = 0.8$.

2.2 Momenty procesu

Předtím, než se pustíme do vyjadřování základních momentů, je dobré si uvědomit, že proces splňující model (2.1) je vlastně Markovův řetězec řádu p , neboť pro každé $t \geq p$ a $k_0, \dots, k_{t-1} \in \mathbb{N}_0$ platí

$$\begin{aligned} P(Y_t = k_t | Y_{t-1} = k_{t-1}, \dots, Y_0 = k_0) &= \\ &= \frac{(a + b_1 k_{t-1} + \dots + b_p k_{t-p})^{k_t}}{k_t!} e^{-(a + b_1 k_{t-1} + \dots + b_p k_{t-p})} \\ &= P(Y_t = k_t | Y_{t-1} = k_{t-1}, \dots, Y_{t-p} = k_{t-p}), \end{aligned}$$

takže markovská vlastnost je splněná. Navíc si můžeme všimnout, že tyto pravděpodobnosti nezávisí na čase t , jde tedy o homogenní Markovův řetězec.

Prozatím předpokládejme, že je proces $\{Y_t, t \in \mathbb{N}_0\}$ slabě stacionární. Později dokážeme, že za určitých podmínek je proces splňující model (2.1) dokonce striktně stacionární.

Věta 2. *Nechť $\{Y_t, t \in \mathbb{N}_0\}$ je slabě stacionární proces splňující model (2.1) a $b_1 + \dots + b_p < 1$. Pak pro střední hodnotu procesu $\{Y_t\}$ platí*

$$EY_t = \frac{a}{1 - b_1 - \dots - b_p}. \quad (2.3)$$

Důkaz. Střední hodnota procesu se dá rozepsat jako

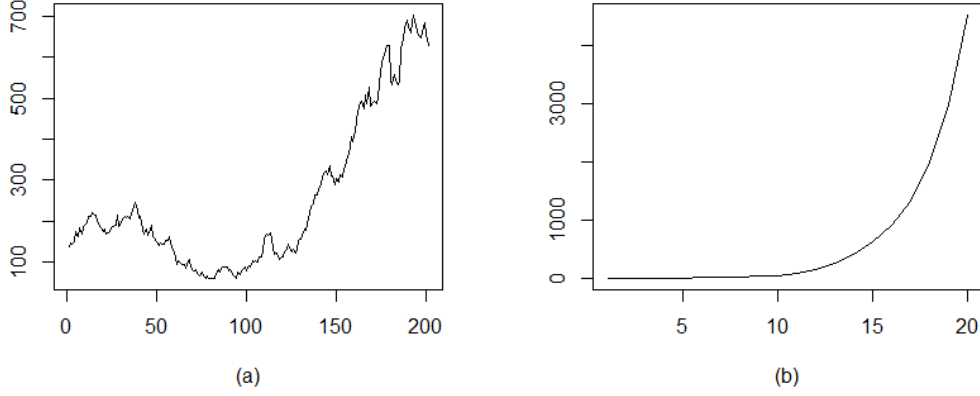
$$\begin{aligned} EY_t &= E(E(Y_t | Y_{t-1}, \dots, Y_{t-p})) = \\ &= E\lambda_t = \\ &= E(a + b_1 Y_{t-1} + \dots + b_p Y_{t-p}) = \\ &= a + b_1 EY_{t-1} + \dots + b_p EY_{t-p} = \\ &= a + (b_1 + \dots + b_p) EY_t, \end{aligned} \quad (2.4)$$

kde poslední rovnost vyplývá z předpokladu slabé stacionarity, a to z vlastnosti, že slabě stacionární proces má v čase konstantní střední hodnotu. Výslednou střední hodnotu získáme vyjádřením EY_t z rovnice (2.4). □

Z vyjádření střední hodnoty (2.3) nám vyplývá další podmínka pro parametry modelu. Protože $Y_t \geq 0$ s.j., tak i $EY_t \geq 0$. Aby střední hodnota tuto nerovnost splňovala, musíme omezit parametry $b_1 + \dots + b_p < 1$.

Věta 3. *Nechť $\{Y_t, t \in \mathbb{N}_0\}$ je slabě stacionární proces splňující model (2.1) a pro parametry modelu platí $b_1 + \dots + b_p < 1$. Pak hodnoty druhého momentu EY_t^2 a smíšených momentů $EY_t Y_{t-1}, \dots, EY_t Y_{t-p+1}$ řeší soustavu rovnic*

$$\begin{aligned} EY_t^2 &= \left(1 + a + a \sum_{k=1}^p b_k\right) EY_t + \sum_{k=1}^p b_k^2 EY_t^2 + 2 \sum_{k=1}^{p-1} b_k b_{k+1} EY_t Y_{t-1} + \\ &\quad + 2 \sum_{k=1}^{p-2} b_k b_{k+2} EY_t Y_{t-2} + \dots + 2 b_1 b_p EY_t Y_{t-p+1} \\ EY_t Y_{t-1} &= a EY_t + b_1 EY_t^2 + b_2 EY_t Y_{t-1} + \dots + b_p EY_t Y_{t-p+1} \\ &\quad \vdots \\ EY_t Y_{t-p+1} &= a EY_t + b_1 EY_t Y_{t-p+2} + \dots + b_{p-2} EY_t Y_{t-1} + b_{p-1} EY_t^2 + b_p EY_t Y_{t-1}. \end{aligned}$$



Obrázek 2.3: Realizace procesu INARCH(1) při porušení podmínky $b_1 < 1$
(a) $a = 1$, $b_1 = 1$; (b) $a = 1$, $b_1 = 1.5$.

Další smíšené momenty lze odvodit rekurzivně ze vztahu

$$\mathbf{E}Y_t Y_{t-k} = a\mathbf{E}Y_t + b_1\mathbf{E}Y_t Y_{t-k+1} + b_2\mathbf{E}Y_t Y_{t-k+2} + \dots + b_p\mathbf{E}Y_t Y_{t-k+p},$$

kde $k \geq p$.

Důkaz. Obecně můžeme pro $t \geq p$, $t \geq k > 0$ vyjádřit $\mathbf{E}Y_t Y_{t-k}$ jako

$$\begin{aligned} \mathbf{E}Y_t Y_{t-k} &= \mathbf{E}(\mathbf{E}(Y_t Y_{t-k} | Y_{t-1}, \dots, Y_0)) = \\ &= \mathbf{E}(Y_{t-k} \mathbf{E}(Y_t | Y_{t-1}, \dots, Y_0)) = \\ &= \mathbf{E}(Y_{t-k} (a + b_1 Y_{t-1} + \dots + b_p Y_{t-p})) = \\ &= a \mathbf{E}Y_{t-k} + b_1 \mathbf{E}Y_{t-k} Y_{t-1} + \dots + b_p \mathbf{E}Y_{t-k} Y_{t-p}. \end{aligned}$$

Jednotlivé rovnosti ze znění věty pak získáme nahrazením k konkrétní hodnotou a záměnou časů náhodných veličin u jednotlivých momentů, k čemuž využijeme vlastnosti slabé stacionarity.

Speciálním případem je moment $\mathbf{E}Y_t^2$, kde k odvození použijeme vyjádření druhého momentu Poissonova rozdělení

$$\begin{aligned} \mathbf{E}Y_t^2 &= \mathbf{E}(\mathbf{E}(Y_t^2 | Y_{t-1}, \dots, Y_{t-p})) = \\ &= \mathbf{E}(\lambda_t^2 + \lambda_t) = \\ &= \mathbf{E}((a + b_1 Y_{t-1} + \dots + b_p Y_{t-p})^2 + (a + b_1 Y_{t-1} + \dots + b_p Y_{t-p})) = \\ &= \psi + \sum_{k=1}^p b_k^2 \mathbf{E}Y_{t-k}^2 + 2 \sum_{k=1}^{p-1} b_k b_{k+1} \mathbf{E}Y_{t-k} Y_{t-k-1} + \\ &\quad + 2 \sum_{k=1}^{p-2} b_k b_{k+2} \mathbf{E}Y_{t-k} Y_{t-k-2} + \dots + 2 b_1 b_p \mathbf{E}Y_{t-1} Y_{t-p} = \\ &= \psi + \sum_{k=1}^p b_k^2 \mathbf{E}Y_t^2 + 2 \sum_{k=1}^{p-1} b_k b_{k+1} \mathbf{E}Y_t Y_{t-1} + 2 \sum_{k=1}^{p-2} b_k b_{k+2} \mathbf{E}Y_t Y_{t-2} + \\ &\quad + \dots + 2 b_1 b_p \mathbf{E}Y_t Y_{t-p+1}, \end{aligned}$$

kde

$$\begin{aligned}
\psi &= a^2 + 2a \sum_{k=1}^p b_k \mathbf{E}Y_{t-k} + a + \sum_{k=1}^p b_k \mathbf{E}Y_{t-k} = \\
&= a \left(1 - \sum_{k=1}^p b_k \right) \mathbf{E}Y_t + 2a \sum_{k=1}^p b_k \mathbf{E}Y_t + \left(1 - \sum_{k=1}^p b_k \right) \mathbf{E}Y_t + \sum_{k=1}^p b_k \mathbf{E}Y_t = \\
&= a \mathbf{E}Y_t + a \sum_{k=1}^p b_k \mathbf{E}Y_t + \mathbf{E}Y_t = \\
&= \left(1 + a + a \sum_{k=1}^p b_k \right) \mathbf{E}Y_t.
\end{aligned}$$

□

Jak je vidět z věty 3, není snadné explicitně vyjádřit druhý moment nebo smíšené momenty, natož rozptyl nebo autokovarianční funkci. Můžeme si ale všimnout, že rozptyl

$$\begin{aligned}
\text{var } Y_t &= \mathbf{E}(\text{var}(Y_t | Y_{t-1}, \dots, Y_{t-p})) + \text{var}(\mathbf{E}(Y_t | Y_{t-1}, \dots, Y_{t-p})) = \\
&= \mathbf{E}\lambda_t + \text{var } \lambda_t = \mathbf{E}Y_t + \text{var } \lambda_t
\end{aligned}$$

je větší nebo rovný střední hodnotě $\mathbf{E}Y_t$ pro každé t , protože rozptyl λ_t je vždy nezáporný. Dochází zde k tzv. overdisperzi. Kdybychom předpokládali, že má proces marginální Poissonovo rozdělení, museli bychom také předpokládat rovnost střední hodnoty a rozptylu, což v praxi moc často nenastává. Použití INARCH modelu s podmíněným Poissonovo rozdělením je proto v tomto směru výhodnější.

Příklad. Mějme proces $\{Y_t\}$ splňující model (2.2) a $a > 0$, $0 \leq b_1 < 1$. Střední hodnota tohoto procesu je

$$\mathbf{E}Y_t = \frac{a}{1 - b_1}.$$

Podle věty 3 můžeme vyjádřit druhý moment

$$\mathbf{E}Y_t^2 = \frac{1 + a + ab_1}{1 - b_1^2} \mathbf{E}Y_t$$

a smíšený moment

$$\begin{aligned}
\mathbf{E}Y_t Y_{t-1} &= a \mathbf{E}Y_t + b_1 \mathbf{E}Y_t^2 = a \mathbf{E}Y_t + b_1 \frac{1 + a + ab_1}{1 - b_1^2} \mathbf{E}Y_t = \\
&= (1 - b_1)(\mathbf{E}Y_t)^2 + \frac{b_1}{1 - b_1^2} \mathbf{E}Y_t + b_1 \frac{a}{1 - b_1} \mathbf{E}Y_t = \\
&= (\mathbf{E}Y_t)^2 + \frac{b_1}{1 - b_1^2} \mathbf{E}Y_t.
\end{aligned}$$

Podobně můžeme vyjádřit i ostatní smíšené momenty

$$\begin{aligned}
 \mathbf{E}Y_t Y_{t-k} &= a \mathbf{E}Y_t + b_1 \mathbf{E}Y_t Y_{t-k+1} = \\
 &= a \mathbf{E}Y_t + b_1 (a \mathbf{E}Y_t + b_1 \mathbf{E}Y_t Y_{t-k+2}) = \\
 &= a(1 + b_1) \mathbf{E}Y_t + b_1^2 \mathbf{E}Y_t Y_{t-k+2} = \\
 &= \dots = a(1 + b_1 + \dots + b_1^{k-1}) \mathbf{E}Y_t + b_1^k \mathbf{E}Y_t^2 = \\
 &= a \frac{1 - b_1^k}{1 - b_1} \mathbf{E}Y_t + b_1^k \mathbf{E}Y_t^2 = \\
 &= (1 - b_1^k) (\mathbf{E}Y_t)^2 + \frac{b_1^k}{1 - b_1^2} \mathbf{E}Y_t + b_1^k (\mathbf{E}Y_t)^2 = \\
 &= \frac{b_1^k}{1 - b_1^2} \mathbf{E}Y_t + (\mathbf{E}Y_t)^2.
 \end{aligned}$$

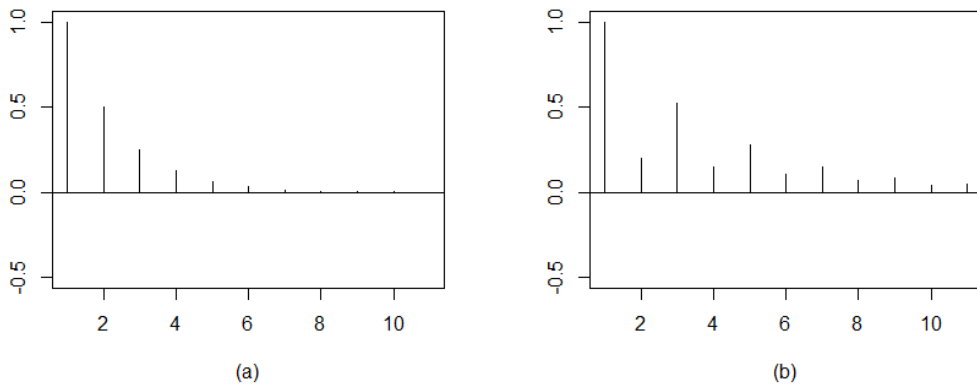
Teď když máme vyjádřené tyto momenty, můžeme vyjádřit i autokovarianční funkci

$$\begin{aligned}
 R(0) &= \text{var } Y_t = \mathbf{E}Y_t^2 - (\mathbf{E}Y_t)^2 = \frac{1}{1 - b_1^2} \mathbf{E}Y_t, \\
 R(k) &= \mathbf{E}Y_t Y_{t-k} - (\mathbf{E}Y_t)^2 = \frac{b_1^k}{1 - b_1^2} \mathbf{E}Y_t, \quad k > 0
 \end{aligned}$$

a autokorelační funkci

$$\begin{aligned}
 r(0) &= 1, \\
 r(k) &= \frac{R(k)}{R(0)} = b_1^k, \quad k > 0.
 \end{aligned}$$

Všimněme si, že korelace je pro libovolné k vždy kladná, protože $b_1 \geq 0$. Z toho vyplývá, že INARCH(1) modelem můžeme modelovat jen řady s kladnými korelacemi. Toto omezení se obecně týká všech lineárních INARCH modelů.



Obrázek 2.4: Autokorelační funkce (a) procesu INARCH(1) s $b_1 = 0.5$; (b) procesu INARCH(2) s $b_1 = 0.1$ a $b_2 = 0.5$.

2.3 INARCH jako AR proces

Proces splňující model (2.1) se dá také vyjádřit pomocí náhodné chyby ϵ_t

$$Y_t = a + b_1 Y_{t-1} + \dots + b_p Y_{t-p} + \epsilon_t, \quad (2.5)$$

kde $\epsilon_t = Y_t - \lambda_t$.

Věta 4. *Nechť $\{Y_t, t \in \mathbb{N}_0\}$ je slabě stacionární proces splňující model (2.1) a pro parametry modelu platí $b_1 + \dots + b_p < 1$. Pak proces náhodných chyb $\{\epsilon_t\}$ je bílý šum a proces $\{Y_t\}$ je AR proces řádu p .*

Důkaz. Abychom dokázali, že je proces $\{\epsilon_t\}$ bílý šum, musíme ukázat nulovou střední hodnotu, rozptyl konstantní v čase a nekorelovanost. Jako první ověříme nulovou střední hodnotu

$$\begin{aligned} \mathbf{E}\epsilon_t &= \mathbf{E}(Y_t - \lambda_t) = \mathbf{E}(\mathbf{E}(Y_t - \lambda_t | Y_{t-1}, \dots, Y_{t-p})) = \\ &= \mathbf{E}(\mathbf{E}(Y_t | Y_{t-1}, \dots, Y_{t-p}) - \lambda_t) = 0. \end{aligned}$$

Dále ověříme konstantní rozptyl

$$\begin{aligned} \text{var } \epsilon_t &= \text{var}(\mathbf{E}(\epsilon_t | Y_{t-1}, \dots, Y_{t-p})) + \mathbf{E}(\text{var}(\epsilon_t | Y_{t-1}, \dots, Y_{t-p})) = \\ &= 0 + \mathbf{E}(\text{var}(Y_t - \lambda_t | Y_{t-1}, \dots, Y_{t-p})) = \\ &= \mathbf{E}(\text{var}(Y_t | Y_{t-1}, \dots, Y_{t-p})) = \\ &= \mathbf{E}\lambda_t = \mathbf{E}(Y_t - \epsilon_t) = \mathbf{E}Y_t = \\ &= \frac{a}{1 - b_1 - \dots - b_p}. \end{aligned}$$

Nakonec ověříme nekorelovanost náhodných chyb

$$\begin{aligned} \text{cov}(\epsilon_t, \epsilon_{t+k}) &= \mathbf{E}\epsilon_t \epsilon_{t+k} = \mathbf{E}(\mathbf{E}(\epsilon_t \epsilon_{t+k} | \mathcal{F}_{t+k-1})) \\ &= \mathbf{E}(\epsilon_t \mathbf{E}(\epsilon_{t+k} | \mathcal{F}_{t+k-1})) = \mathbf{E}(0) = 0 \end{aligned}$$

pro $k \geq 0$.

Tímto jsme ověřili, že proces $\{\epsilon_t\}$ je bílý šum, tudíž proces $\{Y_t\}$ je AR proces řádu p . □

Je dobré si uvědomit, že veličina ϵ_t je sice nekorelovaná se všemi zpožděnými hodnotami Y_{t-1}, Y_{t-2}, \dots

$$\text{cov}(Y_{t-k}\epsilon_t) = \mathbf{E}Y_{t-k}\epsilon_t = \mathbf{E}(Y_{t-k} \mathbf{E}(\epsilon_t | Y_{t-1}, \dots, Y_0)) = 0, \quad t \geq k \geq 0,$$

není na nich ale nezávislá, protože součet λ_t , které nabývá reálných hodnot, a ϵ_t musí dát celé číslo.

Příklad. Mějme proce $\{Y_t\}$ splňující model (2.2). Jak bylo uvedeno výše

$$\begin{aligned} \mathbf{E}Y_t &= \frac{a}{1 - b_1}, \\ R(k) &= \frac{b_1^k}{1 - b_1^2} \mathbf{E}Y_t, \quad k \geq 0 \end{aligned}$$

pro každé t .

Můžeme si všimnout, že vzorce pro střední hodnotu i pro autokovarianční funkci odpovídají vzorcům střední hodnoty a autokovarianční funkce AR(1) procesu, viz kapitola 1, což souhlasí s větou 4.

2.4 Stacionarita a ergodicita

V kapitole 2.2 jsme předpokládali slabou stacionaritu procesu $\{Y_t\}$, bez které bychom nemohli odvodit základní charakteristiky procesu. Nyní si ukážeme, že pokud platí $b_1 + \dots + b_p < 1$ a počáteční rozdělení procesu je stacionárním rozdělením, je proces definovaný vztahem (2.1) dokonce striktně stacionární. Navíc si ukážeme, že je tento proces i silně ergodický.

Věta 5. *Nechť $\{Y_t, t \in \mathbb{N}_0\}$ je proces splňující model (2.1) a pro parametry modelu platí $b_1 + \dots + b_p < 1$. Pak je proces $\{Y_t\}$ silně ergodický a má právě jedno stacionární rozdělení. Navíc $EY_t^k < \infty$ pro $k > 0$.*

Důkaz. Viz Zhu a Wang (2011). □

Předpokládejme, že stacionární rozdělení z věty 5 je počátečním rozdělením procesu $\{Y_t\}$. Pak je proces $\{Y_t\}$ striktně stacionární. Pro obecnější model je důkaz striktní stacionarity uveden i ve Ferland a kol. (2006).

Po celý zbytek druhé kapitoly uvažujme striktně stacionární proces.

2.5 Odhad parametrů

Mějme proces $\{Y_t, t \in \mathbb{N}_0\}$ splňující pro známé p model

$$Y_t | \mathcal{F}_{t-1} \sim \text{Po}(\lambda_t), \quad \lambda_t = a + b_1 Y_{t-1} + \dots + b_p Y_{t-p}, \quad t \geq p, \quad (2.6)$$

kde $a > 0$, $b_1, \dots, b_p \geq 0$ a $b_1 + \dots + b_p < 1$.

Na základě pozorování Y_0, \dots, Y_n chceme odhadnout parametry a, b_1, \dots, b_p . Pro přehlednost označme vektor odhadovaných parametrů $\theta = (a, b_1, \dots, b_p)^\top$.

2.5.1 Volba řádu modelu

Ve všech metodách odhadu uvedených v této kapitole se předpokládá, že známe řád modelu. V praxi tomu tak ale není. K určení p se využívají metody obdobné jako pro klasický AR(p) proces, například analýza parciální autokorelační funkce. Jako p pak zvolíme nejvyšší zpoždění, pro které je ještě parciální autokorelační funkce nenulová, příp. statisticky významná. Vlastnosti parciální autokorelační funkce a výpočet výběrové parciální autokorelační funkce jsou uvedené v Cipra (2008).

2.5.2 Metoda maximální věrohodnosti

Nejprve využijeme naší znalost podmíněného rozdělení Y_t a k odhadu použijeme metodu podmíněné maximální věrohodnosti (dále jen metodu maximální věrohodnosti).

Vycházíme z pravděpodobnostní funkce

$$\begin{aligned}
& \mathbf{P}_\theta(Y_n = y_n, \dots, Y_p = y_p | Y_{p-1} = y_{p-1}, \dots, Y_0 = y_0) = \\
& = \mathbf{P}_\theta(Y_n = y_n | Y_{n-1} = y_{n-1}, \dots, Y_0 = y_0) \\
& \quad \cdot \dots \cdot \mathbf{P}_\theta(Y_p = y_p | Y_{p-1} = y_{p-1}, \dots, Y_0 = y_0) \\
& = \mathbf{P}_\theta(Y_n = y_n | Y_{n-1} = y_{n-1}, \dots, Y_{n-p} = y_{n-p}) \\
& \quad \cdot \dots \cdot \mathbf{P}_\theta(Y_p = y_p | Y_{p-1} = y_{p-1}, \dots, Y_0 = y_0) \\
& = \prod_{t=p}^n \frac{\lambda_t^{y_t}(\theta)}{y_t!} e^{-\lambda_t(\theta)}.
\end{aligned}$$

Odvozenou pravděpodobnostní funkci pak použijeme jako podmíněnou věrohodnostní funkci

$$L_n(\theta) = \prod_{t=p}^n \frac{\lambda_t^{Y_t}(\theta)}{Y_t!} e^{-\lambda_t(\theta)}$$

a po zlogaritmování získáme log-věrohodnostní funkci

$$\ell_n(\theta) = \sum_{t=p}^n [Y_t \log \lambda_t(\theta) - \lambda_t(\theta) - \log Y_t!].$$

Pro zjednodušení si ještě označme $l_t(\theta) = Y_t \log \lambda_t(\theta) - \lambda_t(\theta) - \log Y_t!$.

Hledáme odhad parametru θ , označme $\hat{\theta}_n$, který maximalizuje $L_n(\theta)$, respektive $\ell_n(\theta)$. Za tímto účelem odvodíme skórovou funkci

$$U(\theta; Y_t) = \frac{\partial l_t(\theta)}{\partial \theta} = Y_t \frac{1}{\lambda_t(\theta)} \frac{\partial \lambda_t(\theta)}{\partial \theta} - \frac{\partial \lambda_t(\theta)}{\partial \theta} = \left(\frac{Y_t}{\lambda_t(\theta)} - 1 \right) Z_t,$$

kde $Z_t = \frac{\partial \lambda_t(\theta)}{\partial \theta} = (1, Y_{t-1}, \dots, Y_{t-p})^\top$, a skórovou statistiku

$$U_n(\theta) = \sum_{t=p}^n U(\theta; Y_t).$$

Všimněme si, že Z_t nezávisí na parametru θ .

Odhad $\hat{\theta}_n$ pak najdeme vyřešením soustavy rovnic $U_n(\theta) = 0$, přesněji řešíme rovnice

$$\begin{aligned}
\sum_{t=p}^n \frac{Y_t}{\lambda_t(\theta)} &= n \\
\sum_{t=p}^n \frac{Y_t Y_{t-1}}{\lambda_t(\theta)} &= \sum_{t=p}^n Y_{t-1} \\
&\vdots \\
\sum_{t=p}^n \frac{Y_t Y_{t-p}}{\lambda_t(\theta)} &= \sum_{t=p}^n Y_{t-p}.
\end{aligned}$$

Tento odhad je určen jednoznačně, neboť pozorovaná Fisherova informační matice je pozitivně definitní. Tento fakt je vidět přímo z maticového zápisu pozorované informační matice (viz dále), neboť jde o součin regulární matice se sama

sebou. Ještě poznamenejme, že toto tvrzení platí, jen pokud jsme napozorovali alespoň p nenulových pozorování. Pro Fisherovu informační matici dostáváme

$$\begin{aligned}
I_n(\theta) &= -\frac{1}{n} \sum_{t=p}^n \frac{\partial^2 l_t(\theta)}{\partial \theta \partial \theta^\top} = \\
&= -\frac{1}{n} \sum_{t=p}^n \left(-\frac{Y_t}{\lambda_t^2(\theta)} \frac{\partial \lambda_t(\theta)}{\partial \theta} \frac{\partial \lambda_t(\theta)}{\partial \theta^\top} + \left(\frac{Y_t}{\lambda_t(\theta)} - 1 \right) \frac{\partial^2 \lambda_t(\theta)}{\partial \theta \partial \theta^\top} \right) = \\
&= \frac{1}{n} \sum_{t=p}^n \frac{Y_t}{\lambda_t^2(\theta)} Z_t Z_t^\top,
\end{aligned} \tag{2.7}$$

neboť $\frac{\partial^2 \lambda_t(\theta)}{\partial \theta \partial \theta^\top}$ je nulová matice. Fisherovu informační matici můžeme ještě dál přepsat do tvaru

$$I_n(\theta) = \frac{1}{n} \mathbb{Z}^\top \mathbb{Z},$$

kde

$$\mathbb{Z} = \begin{pmatrix} \sqrt{\frac{Y_p}{\lambda_p^2(\theta)}} Z_p^\top \\ \vdots \\ \sqrt{\frac{Y_n}{\lambda_n^2(\theta)}} Z_n^\top \end{pmatrix}.$$

Označme skutečnou hodnotu parametru θ jako $\theta_0 = (a^0, b_1^0, \dots, b_p^0)^\top$. Můžeme ověřit, že podmíněná střední hodnota skórové funkce je pro každé t nulová

$$\begin{aligned}
\mathbb{E}(U(\theta_0; Y_t) | Y_{t-1}, \dots, Y_{t-p}) &= \\
&= \mathbb{E} \left[\left(\frac{Y_t}{\lambda_t(\theta_0)} - 1 \right) Z_t \middle| Y_{t-1}, \dots, Y_{t-p} \right] = \\
&= \left(\frac{\mathbb{E}(Y_t | Y_{t-1}, \dots, Y_{t-p})}{\lambda_t(\theta_0)} - 1 \right) \frac{\partial \lambda_t(\theta_0)}{\partial \theta} = \\
&= 0.
\end{aligned} \tag{2.8}$$

Poslední rovnost vyplývá z toho, že $\mathbb{E}(Y_t | Y_{t-1}, \dots, Y_{t-p}) = \lambda_t(\theta_0)$. Stejně tak nepodmíněná střední hodnota je rovna nule

$$\mathbb{E}(U(\theta_0; Y_t)) = \mathbb{E}(\mathbb{E}[U(\theta_0; Y_t) | Y_{t-1}, \dots, Y_{t-p}]) = 0.$$

Dokážeme vyjádřit i podmíněný rozptyl skórové funkce pro každé t

$$\begin{aligned}
\text{var}(U(\theta_0; Y_t) | Y_{t-1}, \dots, Y_{t-p}) &= \\
&= \mathbb{E} \left(\frac{\partial l_t(\theta_0)}{\partial \theta} \frac{\partial l_t(\theta_0)}{\partial \theta^\top} \middle| Y_{t-1}, \dots, Y_{t-p} \right) = \\
&= \mathbb{E} \left(\left(\frac{Y_t}{\lambda_t(\theta_0)} - 1 \right)^2 Z_t Z_t^\top \middle| Y_{t-1}, \dots, Y_{t-p} \right) = \\
&= \mathbb{E} \left(\frac{Y_t^2}{\lambda_t^2(\theta_0)} - 2 \frac{Y_t}{\lambda_t(\theta_0)} + 1 \middle| Y_{t-1}, \dots, Y_{t-p} \right) Z_t Z_t^\top = \\
&= \left(\frac{\mathbb{E}(Y_t^2 | Y_{t-1}, \dots, Y_{t-p})}{\lambda_t^2(\theta_0)} - 2 \frac{\mathbb{E}(Y_t | Y_{t-1}, \dots, Y_{t-p})}{\lambda_t(\theta_0)} + 1 \right) Z_t Z_t^\top = \\
&= \left(\frac{\lambda_t^2(\theta_0) + \lambda_t(\theta_0)}{\lambda_t^2(\theta_0)} - 2 \frac{\lambda_t(\theta_0)}{\lambda_t(\theta_0)} + 1 \right) Z_t Z_t^\top = \\
&= \frac{1}{\lambda_t(\theta_0)} Z_t Z_t^\top.
\end{aligned}$$

K určení asymptotického rozdělení nám ještě zbývá vyjádřit očekávanou informační matici pro θ_0

$$\begin{aligned}
I(\theta_0) &= \text{var} U(\theta_0; Y_t) = \\
&= \mathbb{E}(\text{var}[U(\theta_0; Y_t) | Y_{t-1}, \dots, Y_{t-p}]) = \\
&= \mathbb{E} \left(\frac{1}{\lambda_t(\theta_0)} Z_t Z_t^\top \right).
\end{aligned}$$

Navíc můžeme ukázat, že

$$I(\theta_0) = \mathbb{E} \left(\frac{\partial l_t(\theta_0)}{\partial \theta} \frac{\partial l_t(\theta_0)}{\partial \theta^\top} \right) = \mathbb{E} \left(- \frac{\partial^2 l_t(\theta_0)}{\partial \theta \partial \theta^\top} \right) = \mathbb{E} \left(\frac{1}{\lambda_t(\theta_0)} Z_t Z_t^\top \right),$$

neboť

$$\begin{aligned}
\mathbb{E} \left(- \frac{\partial^2 l_t(\theta_0)}{\partial \theta \partial \theta^\top} \right) &= \mathbb{E} \left(\frac{Y_t}{\lambda_t^2(\theta_0)} Z_t Z_t^\top \right) \\
&= \mathbb{E} \left(\frac{\mathbb{E}(Y_t | Y_{t-1}, \dots, Y_{t-p})}{\lambda_t^2(\theta_0)} Z_t Z_t^\top \right) \\
&= \mathbb{E} \left(\frac{1}{\lambda_t(\theta_0)} Z_t Z_t^\top \right).
\end{aligned}$$

Předpokládejme, že parametry modelu jsou omezené nezápornými konstantami, přesněji

$$\delta \leq a \leq M_1, \quad 0 \leq b_1, \dots, b_p \leq M_2 < 1 \quad (2.9)$$

pro nějaká $\delta, M_1, M_2 > 0$. Parametrický prostor je pak kompaktní množina.

Věta 6. *Nechť proces $\{Y_t, t \geq 0\}$ splňuje model (2.6), parametry modelu splňují podmínky (2.9) a $\theta_n^* \xrightarrow[n \rightarrow \infty]{P} \theta_0$. Potom $I_n(\theta_n^*)$, definovaná v (2.7), i*

$$I_n^*(\theta_n^*) = \frac{1}{n} \sum_{t=p}^n \frac{1}{\lambda_t(\theta_n^*)} Z_t Z_t^\top \quad (2.10)$$

jsou konzistentními odhady $I(\theta_0)$.

Důkaz. Ukážeme zde jen konzistenci matice $I_n(\theta_n^*)$, neboť konzistence $I_n^*(\theta_n^*)$ se dokazuje obdobně.

Označme $i_{j,k}(\theta)$ prvek matice $I(\theta)$ a $i'_{j,k}(\theta)$ prvek matice $I_n(\theta)$. Naším cílem je ukázat, že pro každou dvojici indexů jde prvek matice $I_n(\theta)$ v pravděpodobnosti k odpovídajícímu prvku matice $I(\theta)$. Platí

$$\left| i'_{j,k}(\theta_n^*) - i_{j,k}(\theta_0) \right| \leq \left| i'_{j,k}(\theta_n^*) - i'_{j,k}(\theta_0) \right| + \left| i'_{j,k}(\theta_0) - i_{j,k}(\theta_0) \right|. \quad (2.11)$$

První sčítanec můžeme rozvést pomocí věty o střední hodnotě

$$\begin{aligned} \left| i'_{j,k}(\theta_n^*) - i'_{j,k}(\theta_0) \right| &= \left| \frac{1}{n} \sum_{t=p}^n \frac{\partial^2 l(\theta_n^*)}{\partial \theta_j \partial \theta_k} - \frac{1}{n} \sum_{t=p}^n \frac{\partial^2 l(\theta_0)}{\partial \theta_j \partial \theta_k} \right| \\ &\leq \frac{1}{n} \sum_{t=p}^n \sum_{l=1}^p \left| \frac{\partial^3 l(\tilde{\theta}_n)}{\partial \theta_j \partial \theta_k \partial \theta_l} \right| \left| \theta_{n,l}^* - \theta_{0,l} \right|, \end{aligned}$$

kde třetí derivace jsou ve tvaru

$$\left| \frac{\partial^3 l_t(\tilde{\theta}_n)}{\partial \theta_j \partial \theta_k \partial \theta_l} \right| = \frac{2Y_t}{\lambda_t^3(\tilde{\theta}_n)} Z_{t,j} Z_{t,k} Z_{t,l},$$

$Z_{t,j}$ je j -tý prvek vektoru $Z_t = (1, Y_{t-1}, \dots, Y_{t-p})^\top$ a $\tilde{\theta}_n$ je bod mezi θ_n^* a θ_0 , tedy $\tilde{\theta}_n$ konverguje v pravděpodobnosti k θ_0 .

Díky podmínkám (2.9) můžeme omezit $\lambda_t(\theta)$ pro každé θ

$$\lambda_t(\theta) \geq a \geq \delta,$$

protože b_1, \dots, b_p i Y_{t-1}, \dots, Y_{t-p} jsou nezáporné. Tím pádem můžeme omezit i třetí derivace

$$\left| \frac{\partial^3 l_t(\tilde{\theta}_n)}{\partial \theta_j \partial \theta_k \partial \theta_l} \right| \leq \frac{2Y_t}{\delta^3} Z_{t,j} Z_{t,k} Z_{t,l}.$$

Všimněme si, že omezení už nezávisí na hodnotě parametru.

Dostáváme tedy

$$\begin{aligned} \left| i'_{j,k}(\theta_n^*) - i'_{j,k}(\theta_0) \right| &\leq \frac{1}{n} \sum_{t=p}^n \sum_{l=1}^p \frac{2Y_t}{\delta^3} Z_{t,j} Z_{t,k} Z_{t,l} \left| \theta_{n,l}^* - \theta_{0,l} \right| \\ &= \sum_{l=1}^p \left| \theta_{n,l}^* - \theta_{0,l} \right| \frac{1}{n} \sum_{t=p}^n \frac{2Y_t}{\delta^3} Z_{t,j} Z_{t,k} Z_{t,l}. \end{aligned} \quad (2.12)$$

Proces $\{Y_t\}$ je stacionární a ergodický (věta 5), všechny momenty procesu $\{Y_t\}$ jsou konečné a Y_t je pro každé t skoro jistě nezáporné. Tím jsou splněny všechny předpoklady silného zákona velkých čísel pro ergodické procesy a

$$\frac{1}{n} \sum_{t=p}^n \frac{2Y_t}{\delta^3} Z_{t,j} Z_{t,k} Z_{t,l} \xrightarrow[n \rightarrow \infty]{s.j.} \mathbb{E} \left(\frac{2Y_p}{\delta^3} Z_{p,j} Z_{p,k} Z_{p,l} \right).$$

Zároveň θ_n^* konverguje v pravděpodobnosti k θ_0 . Dostáváme tedy, že pravá strana nerovnosti (2.12) konverguje v pravděpodobnosti k nule, a proto i levá strana nerovnosti konverguje v pravděpodobnosti k nule.

Druhý sčítanec pravé strany nerovnosti (2.11)

$$\left| i'_{j,k}(\theta_0) - i_{j,k}(\theta_0) \right| = \left| \frac{1}{n} \sum_{t=p}^n \frac{\partial^2 l(\theta_0)}{\partial \theta_j \partial \theta_k} - \mathbb{E} \left(\frac{\partial^2 l(\theta_0)}{\partial \theta_j \partial \theta_k} \right) \right|$$

závisí jen na pevném θ_0 , proto můžeme rovnou využít silný zákon velkých čísel pro ergodické procesy, čímž dokážeme konvergenci druhého sčítance pravé strany nerovnosti (2.11) v pravděpodobnosti k nule.

Pravá strana nerovnosti (2.11) konverguje v pravděpodobnosti k nule, proto i levá strana konverguje v pravděpodobnosti k nule, čímž jsme dokázali konzistenci odhadu $I_n(\theta_n^*)$. □

Následující věta popisuje asymptotické rozdělení odhadu $\hat{\theta}_n$.

Věta 7. *Nechť proces $\{Y_t, t \geq 0\}$ splňuje model (2.6) a parametry modelu splňují podmínky (2.9). Pak pro maximálně věrohodný odhad $\hat{\theta}_n$ platí*

$$(i) \quad \hat{\theta}_n \xrightarrow[n \rightarrow \infty]{a.s.} \theta_0,$$

$$(ii) \quad \sqrt{n} (\hat{\theta}_n - \theta_0) \xrightarrow[n \rightarrow \infty]{\mathcal{D}} \mathcal{N}(0, I^{-1}(\theta_0)).$$

Důkaz. Důkaz tvrzení (i) je uveden v Zhu a Wang (2011). My zde uvedeme jen hlavní část důkazu tvrzení (ii).

Z Lagrangeovy věty o střední hodnotě víme, že existuje nějaké θ_n^* mezi $\hat{\theta}_n$ a θ_0 takové, že

$$U_n(\hat{\theta}_n) = U_n(\theta_0) - nI_n(\theta_n^*) (\hat{\theta}_n - \theta_0).$$

Zároveň z konstrukce odhadu $\hat{\theta}_n$ vyplývá, že $U_n(\hat{\theta}_n) = 0$. Po úpravě pak dostaneme rovnost

$$\sqrt{n} (\hat{\theta}_n - \theta_0) = \frac{1}{\sqrt{n}} I_n^{-1}(\theta_n^*) \sum_{t=p}^n U(\theta_0; Y_t).$$

Matice $I_n(\theta_n^*)$ je konzistentním odhadem matice $I(\theta_0)$ (věta 6), která je pozitivně definitní (Zhu a Wang, 2011), takže existuje inverzní matice $I^{-1}(\theta_0)$ a s pravděpodobností jdoucí k jedné i $I_n^{-1}(\theta_n^*)$.

Můžeme si všimnout, že $\{U(\theta_0; Y_t), t \geq p\}$ je posloupnost martingalových diferencí, což plyne z rovnosti (2.8)

$$\begin{aligned} \mathbb{E}(U(\theta_0; Y_t) | U(\theta_0; Y_{t-1}), \dots, U(\theta_0; Y_p)) &= \\ &= \mathbb{E}(\mathbb{E}(U(\theta_0; Y_t) | Y_{t-1}, \dots, Y_0) | U(\theta_0; Y_{t-1}), \dots, U(\theta_0; Y_p)) = \\ &= 0. \end{aligned}$$

Tato posloupnost je také stacionární, ergodická a $\mathbb{E}(U^2(\theta_0; Y_t)) < \infty$. Tím jsme ověřili všechny podmínky věty 1, tedy platí

$$\frac{1}{\sqrt{n}} \sum_{t=p}^n U(\theta_0; Y_t) \xrightarrow[n \rightarrow \infty]{\mathcal{D}} \mathcal{N}(0, I(\theta_0)).$$

Tvrzení (ii) pak získáme použitím Cramér-Sluckého věty. □

Abychom mohli určit intervaly spolehlivosti nebo testovat hypotézy, musíme matici $I(\theta_0)$ konzistentně odhadnout. Vzhledem k větě 6 se nabízí buď

$$I_n(\hat{\theta}_n) = \frac{1}{n} \sum_{t=p}^n \frac{Y_t}{\lambda_t^2(\hat{\theta}_n)} Z_t Z_t^\top \quad (2.13)$$

nebo

$$I_n^*(\hat{\theta}_n) = \frac{1}{n} \sum_{t=p}^n \frac{1}{\lambda_t(\hat{\theta}_n)} Z_t Z_t^\top. \quad (2.14)$$

Příklad. Vezměme si INARCH(1) proces a odhadněme parametr $\theta = (a, b_1)^\top$. Podle odvození výše získáme odhad $\hat{\theta}_n$ vyřešením rovnic

$$\begin{aligned} \sum_{t=1}^n \frac{Y_t}{a + b_1 Y_{t-1}} &= n, \\ \sum_{t=1}^n \frac{Y_t Y_{t-1}}{a + b_1 Y_{t-1}} &= \sum_{t=1}^n Y_{t-1}. \end{aligned}$$

Přenásobením druhé rovnice parametrem b_1 a přičtením a -násobku první rovnice můžeme tuto soustavu ještě zjednodušit na

$$\begin{aligned} \sum_{t=1}^n \frac{Y_t}{a + b_1 Y_{t-1}} &= n, \\ \sum_{t=1}^n Y_t - b_1 \sum_{t=1}^n Y_{t-1} &= na. \end{aligned}$$

Ačkoli jsme rovnice znatelně zjednodušili, stále nedokážeme vyjádřit odhad explicitně a soustavu musíme řešit numericky.

Pozorovaná Fisherova informační matice má tvar

$$I_n(\theta) = \frac{1}{n} \sum_{t=1}^n \frac{Y_t}{\lambda_t^2(\theta)} \begin{pmatrix} 1 & Y_{t-1} \\ Y_{t-1} & Y_{t-1}^2 \end{pmatrix}$$

a očekávanou informační matici můžeme vyjádřit jako

$$I(\theta_0) = \mathbb{E} \left(\frac{1}{a^0 + b_1^0 Y_1} \begin{pmatrix} 1 & Y_1 \\ Y_1 & Y_1^2 \end{pmatrix} \right),$$

kde $\theta_0 = (a^0, b_1^0)^\top$ je skutečná hodnota parametru $\theta = (a, b_1)^\top$.

Z věty 7 pak získáme asymptotické rozdělení

$$\sqrt{n} (\hat{\theta}_n - \theta_0) \xrightarrow[n \rightarrow \infty]{\mathcal{D}} \mathcal{N}(0, I^{-1}(\theta_0)).$$

2.5.3 Metoda nejmenších čtverců

Další metodou, kterou se dá odhadnout parametr θ , je podmíněná metoda nejmenších čtverců (dále jen metoda nejmenších čtverců). Výhodou této metody

je, že Y_t nemusí mít podmíněně Poissonovo rozdělení. Stačí, když budeme znát správný tvar podmíněné střední hodnoty. Hledáme $\tilde{\theta}_n$, které minimalizuje funkci

$$S_n(\theta) = \sum_{t=p}^n (Y_t - \lambda_t(\theta))^2 = \sum_{t=p}^n (Y_t - a + b_1 Y_{t-1} + \dots + b_p Y_{t-p})^2.$$

Abychom našli odhad $\tilde{\theta}_n$, řešíme rovnici

$$\frac{\partial S(\theta)}{\partial \theta} = 2 \sum_{t=p}^n (Y_t - \lambda_t(\theta)) \frac{\partial \lambda_t(\theta)}{\partial \theta} = 0, \quad (2.15)$$

která se dá upravit do tvaru

$$\sum_{t=p}^n Y_t Z_t = \sum_{t=p}^n Z_t Z_t^\top,$$

kde $Z_t = \frac{\partial \lambda_t(\theta)}{\partial \theta} = (1, Y_{t-1}, \dots, Y_{t-p})^\top$. Odhad $\tilde{\theta}_n$ se pak dá vyjádřit jako

$$\tilde{\theta}_n = \left(\sum_{t=p}^n Z_t Z_t^\top \right)^{-1} \left(\sum_{t=p}^n Y_t Z_t \right). \quad (2.16)$$

V následující větě vyjádříme asymptotické rozdělení odhadu $\tilde{\theta}_n$. Stejně jako u metody maximální věrohodnosti budeme předpokládat kompaktní parametrický prostor.

Věta 8. *Nechť proces $\{Y_t, t \geq 0\}$ splňuje model (2.6) a parametry modelu splňují (2.9). Pak pro odhad metodou nejmenších čtverců $\tilde{\theta}_n$ platí*

$$(i) \quad \tilde{\theta}_n \xrightarrow[n \rightarrow \infty]{a.s.} \theta,$$

$$(ii) \quad \sqrt{n} (\tilde{\theta}_n - \theta_0) \xrightarrow[n \rightarrow \infty]{\mathcal{D}} \mathcal{N} \left(0, U_1^{-1} V_1 U_1^{-1} \right),$$

kde $U_1 = E(Z_t Z_t^\top)$ a $V_1 = E(\lambda_t(\theta_0) Z_t Z_t^\top)$.

Důkaz. Stačí nám ověřit podmínky vět 3.1 a 3.2 článku Klimko a Nelson (1978) (neuvádíme podmínky, které jsou triviálně splněné díky nulovým druhým derivacím $\lambda_t(\theta)$):

- (a) první derivace $\lambda_t(\theta)$ existují a jsou spojité v θ ,
- (b) pro každé $i, j \leq p$: $E \left| (Y_p - \lambda_p(\theta_0)) \frac{\partial \lambda_t(\theta_0)}{\partial \theta_i} \right| < \infty$, $E \left| \frac{\partial \lambda_t(\theta_0)}{\partial \theta_i} \frac{\partial \lambda_t(\theta_0)}{\partial \theta_j} \right| < \infty$,
- (c) existují funkce $H^{(0)}(Y_{p-1}, \dots, Y_0)$ a $H_i^{(1)}(Y_{p-1}, \dots, Y_0)$ pro každé $i \leq p$ takové, že

$$|\lambda_p(\theta)| \leq H^{(0)}(Y_{p-1}, \dots, Y_0), \quad \left| \frac{\partial \lambda_p(\theta)}{\partial \theta} \right| \leq H_i^{(1)}(Y_{p-1}, \dots, Y_0)$$

pro všechny θ ,

- (d) $E(Y_t | Y_{t-1}, \dots, Y_0) = E(Y_t | Y_{t-1}, \dots, Y_{t-p})$ s.j. pro $t \geq p$,

$$(e) \ E\left((Y_p - \lambda_p)^2 \left| \frac{\partial \lambda_p(\theta_0)}{\partial \theta_i} \frac{\partial \lambda_p(\theta_0)}{\partial \theta_j} \right| \right) < \infty, \ i, j \leq p.$$

Podmínka (a) je splněná, neboť derivace $\lambda_t(\theta)$ nezávisí na θ , a (d) jsme ověřili už na začátku kapitoly 2.2.

Zvolme

$$\begin{aligned} H^{(0)}(Y_{p-1}, \dots, Y_0) &= M_1 + Y_{p-1} + \dots + Y_0, \\ H_i^{(1)}(Y_{p-1}, \dots, Y_0) &= Z_{p,i}, \quad i \leq p, \end{aligned}$$

kde $Z_{p,i}$ je i -tý prvek vektoru Z_p . Pak je splněná i podmínka (c).

Dále

$$\begin{aligned} E|(Y_p - \lambda_p(\theta_0)) Z_{p,i}| &= E(|Y_p - \lambda_p(\theta_0)| Z_{p,i}) \\ &\leq E((Y_p + \lambda_p(\theta_0)) Z_{p,i}) = \\ &= E(E(Y_p + \lambda_p | Y_{p-1}, \dots, Y_0) Z_{p,i}) = \\ &= E(2\lambda_p Z_{p,i}) = \\ &= 2E((a + b_1 Y_{p-1} + \dots + b_p Y_0) Z_{p,i}) < \infty, \end{aligned}$$

což plyne z faktu, že Y_p , $\lambda_p(\theta_0)$ i $Z_{p,i}$ jsou skoro jistě nezáporné, a z věty 3. Tím jsme dokázali první část (b). I druhá část (b) vyplývá z věty 3 a nezápornosti $Z_{p,i}$ pro každé i .

Zbývá ukázat (e). Pomocí vlastností Poissonova rozdělení si můžeme výraz přepsat na

$$\begin{aligned} E\left((Y_p - \lambda_p(\theta_0))^2 | Z_{p,i} Z_{p,j} \right) &= E\left(E\left((Y_p - \lambda_p(\theta_0))^2 | Y_{p-1}, \dots, Y_0\right) Z_{p,i} Z_{p,j}\right) = \\ &= E(\text{var}(Y_p | Y_{p-1}, \dots, Y_0) Z_{p,i} Z_{p,j}) = \\ &= E(\lambda_p(\theta_0) Z_{p,i} Z_{p,j}), \end{aligned}$$

což je podle Zhu a Wang (2011) také konečné pro všechna i, j .

Tímto jsme ověřili všechny podmínky vět 3.1 a 3.2 z Klimko a Nelson (1978), ze kterých přímo plyne dokazovaná věta. □

Tento odhad není tak eficientní jako odhad metodou maximální věrohodnosti, můžeme ho ale využít pro výpočet odhadu vážených nejmenších čtverců s neznámými vahami

$$\bar{\theta}_n = \left(\sum_{t=p}^n \frac{Z_t Z_t^\top}{\lambda_t(\bar{\theta}_n)} \right)^{-1} \left(\sum_{t=p}^n \frac{Y_t Z_t}{\lambda_t(\bar{\theta}_n)} \right).$$

I tento odhad má asymptoticky normální rozdělení, jak je uvedeno v následující větě.

Věta 9. *Nechť proces $\{Y_t, t \geq 0\}$ splňuje model (2.6). Pak pro modifikovaný odhad metodou nejmenších čtverců $\bar{\theta}_n$ platí $\sqrt{n}(\bar{\theta}_n - \theta_0) \xrightarrow[n \rightarrow \infty]{\mathcal{D}} \mathcal{N}(0, I^{-1}(\theta_0))$, kde $I(\theta)$.*

Důkaz. Viz Zhu a Wang (2011). □

Další možností, jak odhadnout parametr θ , je využít metodu vážených nejmenších čtverců. Za váhy w_t zvolme hodnoty

$$w_t = \frac{1}{1 + Z_t^\top Z_t}.$$

Odhad θ_n^* pak minimalizuje funkci

$$S^*(\theta) = \sum_{t=p}^n w_t (Y_t - \lambda_t(\theta))^2.$$

Stejným postupem jako při odvozování odhadu $\tilde{\theta}_n$ získáme i odhad

$$\theta_n^* = \left(\sum_{t=p}^n w_t Z_t Z_t^\top \right)^{-1} \left(\sum_{t=p}^n w_t Y_t Z_t \right).$$

I tento odhad má asymptoticky normální rozdělení.

Věta 10. *Nechť proces $\{Y_t, t \geq 0\}$ splňuje model (2.6). Pak pro odhad metodou vážených nejmenších čtverců θ_n^* platí*

$$(i) \theta_n^* \xrightarrow[n \rightarrow \infty]{a.s.} \theta,$$

$$(ii) \sqrt{n} (\theta_n^* - \theta_0) \xrightarrow[n \rightarrow \infty]{\mathcal{D}} \mathcal{N} \left(0, U_2^{-1} V_2 U_2^{-1} \right),$$

kde $U_2 = E(w_t Z_t Z_t^\top)$ a $V_2 = E(w_t \lambda_t(\theta_0) Z_t Z_t^\top)$.

Důkaz. Viz Zhu a Wang (2011). □

Příklad. Vraťme se zpět k INARCH(1) modelu. Tentokrát odhadneme parametr $\theta = (a, b_1)^\top$ metodou nejmenších čtverců.

Minimalizací

$$S_n(\theta) = \sum_{t=1}^n (Y_t - a + b_1 Y_{t-1})^2 \tag{2.17}$$

získáme odhad

$$\tilde{\theta}_n = \left(\sum_{t=1}^n Z_t Z_t^\top \right)^{-1} \left(\sum_{t=1}^n Y_t Z_t \right), \tag{2.18}$$

kde $Z_t = (1, Y_{t-1})^\top$. Konkrétně

$$a = \frac{\sum_{t=1}^n Y_{t-1}^2 \sum_{t=1}^n Y_t - \sum_{t=1}^n Y_{t-1} \sum_{t=1}^n Y_t Y_{t-1}}{n \sum_{t=1}^n Y_{t-1}^2 - (\sum_{t=1}^n Y_{t-1})^2}, \tag{2.19}$$

$$b_1 = \frac{n \sum_{t=1}^n Y_t Y_{t-1} - \sum_{t=1}^n Y_t \sum_{t=1}^n Y_{t-1}}{n \sum_{t=1}^n Y_{t-1}^2 - (\sum_{t=1}^n Y_{t-1})^2}. \tag{2.20}$$

Na rozdíl od metody maximální věrohodnosti můžeme odvodit i přesný tvar matic U a V a tím pádem i asymptotický rozptyl

$$U = E \begin{pmatrix} 1 & Y_{t-1} \\ Y_{t-1} & Y_{t-1}^2 \end{pmatrix} = \begin{pmatrix} 1 & \frac{a^0}{1-b_1^0} \\ \frac{a^0}{1-b_1^0} & \frac{1+a^0+a^0 b_1^0}{1-(b_1^0)^2} \frac{a^0}{1-b_1^0} \end{pmatrix},$$

$$\begin{aligned}
V &= \mathbb{E} \left(\lambda_t(\theta_0) \begin{pmatrix} 1 & Y_{t-1} \\ Y_{t-1} & Y_{t-1}^2 \end{pmatrix} \right) = \\
&= \frac{a^0}{1-b_1^0} \begin{pmatrix} 1 & \frac{a^0 + a^0 b_1^0 + b_1^0}{1-(b_1^0)^2} \\ \frac{a^0 + a^0 b_1^0 + b_1^0}{1-(b_1^0)^2} & \frac{a^0(1+a^0) + a^0 b_1^0(2a^0(1+b_1^0) + 3b_1^0 + 1) + (b_1^0)^3(2+2a^0 + (a^0)^2)}{(1-(b_1^0)^2)(1-(b_1^0)^3)} \end{pmatrix}, \\
U^{-1}VU^{-1} &= \begin{pmatrix} \frac{a^0}{1-b_1^0} \left(a^0(1+b_1^0) + \frac{1+2(b_1^0)^4}{1+b_1^0+(b_1^0)^2} \right) & -a^0(1+b_1^0) - \frac{(1+2b_1^0)(b_1^0)^3}{1+b_1^0+(b_1^0)^2} \\ -a^0(1+b_1^0) - \frac{(1+2b_1^0)(b_1^0)^3}{1+b_1^0+(b_1^0)^2} & (1-(b_1^0)^2) \left(1 + \frac{b_1^0(1+2(b_1^0)^2)}{a^0(1+b_1^0+(b_1^0)^2)} \right) \end{pmatrix}.
\end{aligned} \tag{2.21}$$

2.5.4 Metoda momentů

Stejně jako u klasického AR modelu můžeme i u modelu (2.6) odhadnout parametry modelu metodou momentů. K tomu využijeme vyjádření střední hodnoty (2.3) a rovnice pro smíšené momenty $\mathbb{E}Y_t Y_{t-1}, \dots, \mathbb{E}Y_t Y_{t-p}$ z věty 3. Pokud všechny momenty nahradíme patřičnými výběrovými momenty, tj. $\mathbb{E}Y_t$ nahradíme výběrovým průměrem $\frac{1}{n} \sum_{t=p}^n Y_t$ a $\mathbb{E}Y_t Y_{t-k}$ nahradíme výběrovým momentem $\frac{1}{n} \sum_{t=p}^n Y_t Y_{t-k}$ pro $k = 1, \dots, p$, získáme soustavu rovnic

$$\begin{aligned}
\frac{1}{n} \sum_{t=p}^n Y_t &= a + (b_1 + \dots + b_p) \frac{1}{n} \sum_{t=p}^n Y_t \\
\frac{1}{n} \sum_{t=p}^n Y_t Y_{t-1} &= a \frac{1}{n} \sum_{t=p}^n Y_t + b_1 \frac{1}{n} \sum_{t=p}^n Y_t^2 + \dots + b_p \frac{1}{n} \sum_{t=p}^n Y_t Y_{t-p+1} \\
&\vdots \\
\frac{1}{n} \sum_{t=p}^n Y_t Y_{t-p} &= a \frac{1}{n} \sum_{t=p}^n Y_t + b_1 \frac{1}{n} \sum_{t=p}^n Y_t Y_{t-p+1} + \dots + b_p \frac{1}{n} \sum_{t=p}^n Y_t^2.
\end{aligned}$$

Vyřešením této soustavy pak získáme odhad metodou momentů.

Můžeme si všimnout, že tato soustava rovnic je až na indexy některých sčítanců shodná se soustavou rovnic pro odhad metodou nejmenších čtverců (2.15). Vzhledem k tomu, že za podmínky $b_1 < 1$ je proces striktně stacionární i ergodický, má odhad získaný metodou momentů stejné asymptotické vlastnosti jako odhad metodou nejmenších čtverců (2.16).

2.5.5 Diagnostika

Po odhadnutí modelu je vždy vhodné zjistit, jak dobře sedí model na naše data. K tomu se nejčastěji využívají Pearsonova rezidua

$$r_t^P = \frac{Y_t - \hat{\lambda}_t}{\sqrt{\hat{\lambda}_t}},$$

kde $\hat{\lambda}_t = \lambda_t(\hat{\theta}_n)$ pro nějaký odhad $\hat{\theta}_n$.

Autokorelační funkce těchto reziduí nám může napovědět, že v modelu nejsou zohledněné všechny závislosti, a graf reziduí může odhalit trend, sezónnost nebo změnu v chování řady (Liboschik a kol., 2017).

Další možností jak určit, jestli je model pro naše data vhodný, jsou testy dobré shody. V tomto případě testujeme vztah podmíněné střední hodnoty a zpožděných hodnot procesu. Testům dobré shody se věnují například Fokianos a Neumann (2013).

3. Možná zobecnění INARCH modelu

Ne vždy je lineární INARCH(p) vhodný k modelování vybraných dat, proto existují různá zobecnění. Ve své práci se jimi zabývá například Fokianos (2012). Tato zobecnění jsou ale často náročnější na výpočet, proto se jejich použití nemusí v praxi vyplatit.

3.1 Lineární INGARCH(p, q) model

Do teď jsme modelovali podmíněnou střední hodnotu procesu jen pomocí zpožděných hodnot procesu. Zkusme předpokládat, že podmíněná střední hodnota lineárně závisí i na zpožděných podmíněných středních hodnotách.

Mějme proces $\{Y_t, t \in \mathbb{N}_0\}$, pro který platí

$$Y_t | \mathcal{F}_{t-1}^\lambda \sim \text{Po}(\lambda_t), \quad \lambda_t = a + \sum_{i=1}^q b_i Y_{t-i} + \sum_{j=1}^p c_j \lambda_{t-j}, \quad t \geq \max(p, q), \quad (3.1)$$

kde $a > 0$, $b_i \geq 0, i = 1, \dots, q$, $c_j \geq 0, j = 1 \dots p$ a

$$\mathcal{F}_t^\lambda = \sigma(Y_s, \lambda_u, 0 \leq s \leq t, 0 \leq u \leq \max(p, q) - 1)$$

je σ -algebra generovaná náhodnými veličinami Y_0, \dots, Y_t a $\lambda_0, \dots, \lambda_{\max(p, q) - 1}$. Tento model značíme INGARCH(p, q).

Stejně jako u INARCH modelu i zde platí

$$\mathbb{E}(Y_t | \mathcal{F}_{t-1}^\lambda) = \text{var}(Y_t | \mathcal{F}_{t-1}^\lambda) = \lambda_t. \quad (3.2)$$

Pokud parametry ještě víc omezíme

$$0 \leq \sum_{i=1}^q b_i + \sum_{j=1}^p c_j < 1, \quad (3.3)$$

je proces $\{Y_t, t \in \mathbb{N}_0\}$ striktně stacionární s konečnými druhými momenty (Ferland a kol. (2006)).

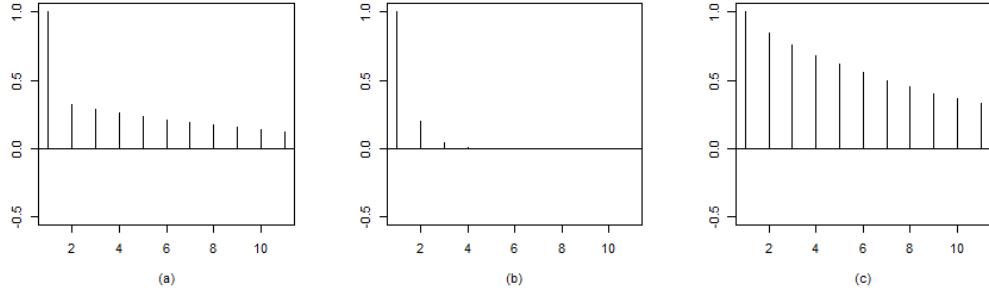
Model můžeme formulovat i pomocí náhodných chyb $\epsilon_t = Y_t - \lambda_t$

$$\begin{aligned} Y_t &= \lambda_t + \epsilon_t = \\ &= a + \sum_{i=1}^q b_i Y_{t-i} + \sum_{j=1}^p c_j (Y_{t-j} - \epsilon_{t-j}) + \epsilon_t = \\ &= a + \sum_{i=1}^{\max(p, q)} (b_i + c_i) Y_{t-i} + \epsilon_t - \sum_{j=1}^p c_j \epsilon_{t-j}, \quad t \geq \max(p, q), \end{aligned} \quad (3.4)$$

kde je třeba dodefinovat

$$\begin{aligned} b_i &= 0, \quad i = q+1, \dots, \max(p, q), \\ c_j &= 0, \quad j = p+1, \dots, \max(p, q), \end{aligned}$$

a stejně jako v důkazu věty 4 se dá ukázat, že model INGARCH(p, q) je zároveň model ARMA($\max(p, q), p$).



Obrázek 3.1: Autokorelační funkce procesu INGARCH(1,1) (a) s $b_1 = 0.2$ a $c_1 = 0.7$; (b) s $b_1 = 0.2$ a $c_1 = 0$; (c) s $b_1 = 0.7$ a $c_1 = 0.2$.

3.1.1 Proces INGARCH(1,1)

Nejpoužívanějším a zároveň nejjednodušším případem INGARCH modelu je model INGARCH(1,1). Tento model má tvar

$$Y_t | \mathcal{F}_{t-1}^\lambda \sim \text{Po}(\lambda_t), \quad \lambda_t = a + b_1 Y_{t-1} + c_1 \lambda_{t-1}, \quad t \geq 1, \quad (3.5)$$

kde $a > 0$, $b_1 \geq 0$, $c_1 \geq 0$ a $\mathcal{F}_t^\lambda = \sigma(Y_s, \lambda_0, 0 \leq s \leq t)$.

Za podmínky $0 < b_1 + c_1 < 1$ je proces striktně stacionární s konečnými momenty (Ferland a kol. (2006)), proto po zbytek této kapitoly předpokládáme, že je tato podmínka splněná.

Podobně jako v případě INARCH(1) modelu dokážeme vyjádřit hlavní momenty i INARCH(1,1) modelu. Střední hodnotu

$$\mathbf{E}Y_t = \frac{a}{1 - b_1 - c_1}, \quad (3.6)$$

můžeme vyjádřit stejným způsobem jako v důkazu věty (4), přičemž využijeme striktní stacionaritu procesu a rovnost

$$\mathbf{E}Y_{t-1} = \mathbf{E}\left(\mathbf{E}\left(Y_{t-1} | \mathcal{F}_{t-2}^\lambda\right)\right) = \mathbf{E}\lambda_{t-1}. \quad (3.7)$$

Autokovarianční funkci

$$R(0) = \frac{1 - 2b_1c_1 - c_1^2}{1 - (b_1 + c_1)^2} \mathbf{E}Y_t, \quad (3.8)$$

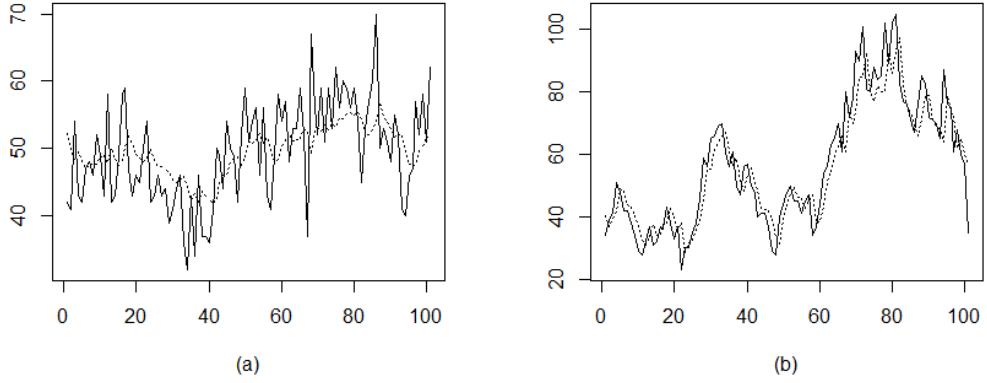
$$R(k) = (b_1 + c_1)^{k-1} \frac{b_1(1 - c_1(b_1 + c_1))}{1 - (b_1 + c_1)^2} \mathbf{E}Y_t \quad (3.9)$$

$$= (b_1 + c_1)^{k-1} R(1), \quad k \geq 1. \quad (3.10)$$

zase získáme použitím vlastností ARMA modelu na vyjádření (3.4) pro model INGARCH(1,1). Je-li c_1 nulové, dostaneme přirozeně momenty INARCH(1) modelu. Na obrázku 3.1 můžeme porovnat autokorelační funkce pro tři kombinace parametrů. Je vidět, že s vyšším součtem parametrů b_1 a c_1 autokorelační funkce klesá pomaleji, přičemž v případě nízké hodnoty b_1 je vidět skok mezi prvními dvěma hodnotami.

Můžeme si všimnout, že i v případě INGARCH(1,1) nastává overdisperze, tj. $\text{var } Y_t \geq \mathbf{E}Y_t$, neboť

$$\frac{1 - 2b_1c_1 - c_1^2}{1 - (b_1 + c_1)^2} \geq 1 \quad \Leftrightarrow \quad |b_1 + c_1| < 1. \quad (3.11)$$



Obrázek 3.2: Realizace a funkce střední hodnoty procesu INGARCH(1,1) s parametry (a) $a = 5$, $b_1 = 0.2$ a $c_1 = 0.7$; (b) $a = 5$, $b_1 = 0.7$ a $c_1 = 0.2$.

Na obrázku 3.2 jsou zobrazené časové řady pro dvě kombinace parametrů. Pro vysoké hodnoty parametru b_1 a nízké c_1 podmíněná střední hodnota procesu kopíruje časovou řadu. Naproti tomu časová řada s vyšším parametrem c_1 se spíše pohybuje kolem podmíněné střední hodnoty.

3.1.2 Odhad parametrů modelu INGARCH(1,1)

V této kapitole uvedeme jen odhad metodou maximální věrohodnosti. Předpokládejme $b_1 + c_1 < 1$. Chceme odhadnout parametr $\theta = (a, b_1, c_1)^\top$. Stejně jako v kapitole 2.5.2 si nejprve vyjádříme věrohodnostní a log-věrohodnostní funkci

$$L_n(\theta) = \prod_{t=1}^n \frac{\lambda_t^{Y_t}(\theta)}{Y_t!} e^{-\lambda_t(\theta)}, \quad (3.12)$$

$$\ell_n(\theta) = \sum_{t=1}^n l_t(\theta), \quad (3.13)$$

kde $l_t(\theta) = Y_t \log \lambda_t(\theta) - \lambda_t(\theta) - \log Y_t!$.

Derivací dostaneme skórovou statistiku

$$U_n(\theta) = \sum_{t=1}^n \left(\frac{Y_t}{\lambda_t(\theta)} - 1 \right) \frac{\partial \lambda_t(\theta)}{\partial \theta}, \quad (3.14)$$

kde $\frac{\partial \lambda_t(\theta)}{\partial \theta} = Z_t + c_1 \frac{\partial \lambda_{t-1}}{\partial \theta}$ a Z_t je definováno jako $Z_t = (1, Y_{t-1}, \lambda_{t-1})$.

Odhad $\hat{\theta}_n$ hledáme jako řešení soustavy rovnic $U_n(\theta) = 0$. Oproti INARCH modelu derivace $\frac{\partial \lambda_t(\theta)}{\partial \theta}$ závisí na parametru θ , což ztěžuje další odvozování a v praxi výpočet odhadů. Stále ale platí

$$\mathbb{E} \left(\frac{\partial \lambda_t(\theta)}{\partial \theta} \middle| \mathcal{F}_{t-1}^\lambda \right) = \mathbb{E} \left(Z_t + c_1 \frac{\partial \lambda_{t-1}}{\partial \theta} \middle| Y_{t-1}, \dots, Y_0, \lambda_0 \right) = \frac{\partial \lambda_t(\theta)}{\partial \theta}.$$

Pozorovaná Fisherova informační matice má tvar

$$I_n(\theta) = \frac{1}{n} \sum_{t=1}^n \frac{Y_t}{\lambda_t^2(\theta)} \frac{\partial \lambda_t(\theta)}{\partial \theta} \frac{\partial \lambda_t(\theta)}{\partial \theta^\top} + \frac{1}{n} \sum_{t=1}^n \left(\frac{Y_t}{\lambda_t(\theta)} - 1 \right) \frac{\partial^2 \lambda_t(\theta)}{\partial \theta \partial \theta^\top} \quad (3.15)$$

a stejně jako v kapitole 2.5.2 je očekávaná Fisherova matice pro $\theta_0 = (a^0, b_1^0, c_1^0)^\top$, skutečnou hodnotu parametru θ , rovna

$$I(\theta_0) = \mathbb{E} \left(\frac{\partial l_t(\theta_0)}{\partial \theta} \frac{\partial l_t(\theta_0)}{\partial \theta^\top} \right) = \mathbb{E} \left(- \frac{\partial^2 l_t(\theta_0)}{\partial \theta \partial \theta^\top} \right) = \mathbb{E} \left(\frac{1}{\lambda_t(\theta_0)} \frac{\partial \lambda_t(\theta_0)}{\partial \theta} \frac{\partial \lambda_t(\theta_0)}{\partial \theta^\top} \right).$$

Ještě si uvedme podmíněnou informační matici

$$I_n^*(\theta) = \frac{1}{n} \sum_{t=1}^n \frac{1}{\lambda_t(\hat{\theta}_n)} \frac{\partial \lambda_t(\theta)}{\partial \theta} \frac{\partial \lambda_t(\theta)}{\partial \theta}^\top. \quad (3.16)$$

Fokianos a kol. (2009) ukázali, že odhad $\hat{\theta}_n$ má podobné asymptotické vlastnosti jako maximálně věrohodný odhad INARCH modelu.

Věta 11. *Nechť proces $\{Y_t, t \geq 0\}$ splňuje model (3.5) a $0 < b_1^0 + c_1^0 < 1$. Pak pro maximálně věrohodný odhad $\hat{\theta}_n$ platí*

- (i) $\hat{\theta}_n \xrightarrow[n \rightarrow \infty]{a.s.} \theta$,
- (ii) $I_n^*(\hat{\theta}_n) \xrightarrow[n \rightarrow \infty]{a.s.} I(\theta)$,
- (iii) $\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow[n \rightarrow \infty]{\mathcal{D}} \mathcal{N}(0, I^{-1}(\theta_0))$.

Důkaz. Viz Fokianos a kol. (2009). □

Ferland a kol. (2006) uvádí i další konzistentní odhady $I(\theta_0)$.

3.2 Nelineární modely

Dalším možných zobecněním je povolení jiných než lineární funkcí. V této práci si uvedeme jen zobecnění INGARCH(1,1) modelu, ale v literatuře můžeme najít i nelineární INGARCH vyšších řádů.

Mějme proces $\{Y_t, t \in \mathbb{N}_0\}$, který splňuje model

$$Y_t | \mathcal{F}_{t-1}^\lambda \sim \text{Po}(\lambda_t), \quad \lambda_t = f(Y_{t-1}, \lambda_{t-1}), \quad t \geq 1, \quad (3.17)$$

kde $\mathcal{F}_t^\lambda = \sigma(Y_s, \lambda_0, 0 \leq s \leq t)$ je σ -algebra generovaná veličinami $\lambda_0, Y_0, \dots, Y_t$ a $f: \mathbb{N}_0 \times (0, \infty) \rightarrow (0, \infty)$ je až na vektor parametrů známá funkce, která splňuje

$$|f(y, \lambda) - f(y', \lambda')| \leq \alpha |\lambda - \lambda'| + \gamma |y - y'|, \quad \forall \lambda, \lambda' > 0 \forall y, y' \in \mathbb{N}_0 \quad (3.18)$$

pro nějaké $\alpha, \gamma \geq 0$, kde $\alpha + \gamma < 1$. Stacionaritou a ergodicitou tohoto procesu se zabývá Neumann (2011).

Fokianos a Tjøstheim (2012) se zabývali modely typu

$$\lambda_t = f(\lambda_{t-1}) + b(Y_{t-1}), \quad (3.19)$$

kde $f, b: \mathbb{R} \rightarrow (0, \infty)$ jsou až na vektor parametrů známé funkce, a dokazují, že za určitých podmínek je tento proces stacionární a odhad metodou maximální věrohodnosti je konzistentní a asymptoticky normální.

Mezi modely typu (3.19) patří i např. model

$$\lambda_t = d + \left(a + c \exp \left(-\gamma \lambda_{t-1}^2 \right) \right) \lambda_{t-1} + b Y_{t-1} \quad (3.20)$$

nebo i model INGARCH(1,1) z kapitoly 3.1.1. Fokianos (2015) uvádí i další typy nelineárních modelů.

Speciálním typem nelineárního modelu je log-lineární model

$$Y_t | \mathcal{F}_{t-1}^\lambda \sim \text{Po}(\lambda_t), \quad \log \lambda_t = a + b_1 \log(Y_{t-1} + 1) + c_1 \log \lambda_{t-1}, \quad t \geq 1. \quad (3.21)$$

V případě Poissonova rozdělení je volba logaritmicke funkce přirozená vzhledem k tomu, že jde o kanonickou linkovou funkci zobecněného lineárního modelu pro Poissonovo rozdělení.

Podmíněná střední hodnota a podmíněný rozptyl jsou pak ve tvaru

$$\lambda_t = e^a (Y_{t-1} + 1)^{b_1} \lambda_{t-1}^{c_1}. \quad (3.22)$$

Oproti INGARCH(1,1) modelu můžou mít parametry modelu kladnou i zápornou hodnotu. Navíc můžeme modelovat kladnou i zápornou korelaci procesu.

Navíc pokud $|b_1 + c_1| < 1$ pro b_1, c_1 se stejným znaménkem a $b_1^2 + c_1^2 < 1$ pro b_1, c_1 s různými znaménky, proces je striktně stacionární a odhad metodou maximální věrohodnosti je konzistentní a asymptoticky normální (Fokianos a Tjøstheim (2011)).

Další zobecnění pak nabízí modely, kde podmíněné rozdělení není Poissonovo, ale jiné. V praxi nejpoužívanější je negativně binomické (Zhu, 2010). Toto rozdělení umožňuje ještě výraznější overdispersi než modelovalo podmíněné Poissonovo rozdělení.

4. Praktické použití

V této kapitole vyzkoušíme metody uvedené v předchozích kapitolách na datech. Nejprve budeme simulovat několik INARCH modelů a porovnáme přesnost jednotlivých metod odhadu a poté se pokusíme modelovat tři reálné časové řady.

4.1 Implementace v R

Dříve než se pustíme do odhadování, uvedme si, jaké nástroje nám R nabízí pro práci s diskrétními časovými řadami.

Liboschik a kol. (2017) představili balíček `tscount`, pomocí kterého se mimo jiné dají odhadnout parametry modelu metodou maximální věrohodnosti. Tento balíček pracuje s lineárními a log-lineárními INGARCH modely za předpokladu Poissonova nebo negativně binomického podmíněného rozdělení.

Hlavní funkcí je v tomto balíčku funkce `tsglm`. Základními vstupy jsou modelovaná řada (`ts`), model (`model`), což je seznam indexů minulých hodnot procesu (`past_obs`) a podmíněných středních hodnot (`past_mean`), předpokládané podmíněné rozdělení procesu (`distr`) s možnostmi Poissonovo nebo negativně binomického rozdělení a linková funkce (`link`) s hodnotami identita nebo logaritmus. Výstupem jsou mimo jiné odhadnuté parametry modelu (`coefficients`), informační matice (`info.matrix`), odhadnuté podmíněné střední hodnoty λ_t (`fitted.values`) a rezidua (`residuals`).

Pro odhad parametrů metodami nejmenších čtverců můžeme využít funkci `lm`, kam můžeme zadat i případné váhy. Tato metoda je ale určena pro nezávislá stejně rozdělená data, proto odhady kovariance nejsou pro naše využití správné. Je tedy nutné vytvořit si pro odhad kovariance vlastní funkci.

4.2 Simulace

Budeme simulovat šest INARCH(1) modelů s parametry

- (a) $a = 1$ a $b_1 = 0.2$,
- (b) $a = 1$ a $b_1 = 0.5$,
- (c) $a = 1$ a $b_1 = 0.8$,
- (d) $a = 2.4$ a $b_1 = 0.2$,
- (e) $a = 1.5$ a $b_1 = 0.5$,
- (f) $a = 0.6$ a $b_1 = 0.8$.

Parametry modelů (d) - (f) jsou volené tak, aby byla střední hodnota procesu rovna 3.

Pro každý model si vytvoříme 1000 časových řad o 500 pozorování. Parametry modelu pak postupně odhadneme metodou maximální věrohodnosti (značme MLE), metodou nejmenších čtverců (CLS), metodou vážených nejmenších čtverců s neznámými vahami (CLSu) a metodou vážených nejmenších čtverců se známými

vahami (CLSw). Abychom porovnali přesnost odhadů pro různé počty pozorování, odhadneme parametry pomocí prvních 50, 100, 200 a všech 500 pozorování.

Odhady metodou maximální věrohodnosti jsou vždy uvnitř parametrického prostoru, neboť funkce `tsglm`, kterou k odhadu používáme, parametrický prostor zohledňuje. U metod nejmenších čtverců žádné omezení není, proto můžeme dostat i odhad se zápornými hodnotami. V tabulkách 4.6 a 4.7 uvádíme ve sloupci *out* počet odhadů mimo parametrický prostor. Je vidět, že s rostoucím počtem pozorování se odhady zpřesňují, a proto nedochází tak často k nepřipustným odhadům.

Abychom mohli metody lépe porovnat, položili jsme všechny záporné odhady parametru a rovny 10^{-6} , všechny záporné odhady parametru b_1 rovny nule a všechny odhady b_1 vyšší než 1 rovny $1 - 10^{-6}$.

Ze získaných odhadů můžeme pro jednotlivé parametry napočítat charakteristiky jako odhad vychýlení

$$bias = \frac{1}{K} \sum_{k=1}^K \hat{\xi}_{n,k} - \xi, \quad (4.1)$$

odhad směrodatné odchylky

$$sd = \sqrt{\frac{1}{K} \sum_{k=1}^K (\hat{\xi}_{n,k} - \bar{\xi}_K)^2} \quad (4.2)$$

nebo střední čtvercová chyba

$$MSE = \frac{1}{K} \sum_{k=1}^K (\hat{\xi}_{n,k} - \xi)^2, \quad (4.3)$$

kde $\hat{\xi}_{n,k}$ je odhad parametru k -té řady, $\bar{\xi}_K$ je výběrový průměr odhadů $\hat{\xi}_{n,k}$, $k = 1, \dots, K$, K je počet vygenerovaných řad pro daný model (v našem případě je K rovno 1000) a ξ je neznámý parametr (v našem případě a nebo b_1). Na základě těchto charakteristik porovnáme jednotlivé metody. Pro každou metodu jsou výsledné charakteristiky uvedené v tabulkách 4.6 a 4.7.

Nejdříve porovnejme metody MLE a CLS. Pro nízkou hodnotu b_1 (v našem případě $b_1 = 0.2$) a vyšší počet pozorování (200 nebo 500) jsou *bias*, *sd* i *MSE* odhadů těmito metodami skoro stejné. V některých případech jsou hodnoty pro metodu CLS dokonce nižší než pro MLE. S rostoucím b_1 se ale výsledky pro CLS výrazně zhoršují.

Metody MLE a CLSu jsou si, co se týče hodnot charakteristik, velmi podobné, hlavně pro vyšší počet pozorování. Obecně má MLE lepší hodnoty *bias* a u odhadů parametru a i *MSE*. Pro vyšší hodnoty parametru b_1 má CLSu zase nižší hodnoty *sd*. Hodnoty *MSE* jsou u odhadů parametru b_1 téměř shodné.

Ačkoli pro nízké hodnoty b_1 dává metoda CLS lepší výsledky než CLSu, u vyšších hodnot parametru b_1 je vidět výrazné zlepšení metody CLSu v porovnání s metodou CLS.

Metoda CLSw se vyznačuje odhady s velmi dobrým *bias*, zato hodnoty *sd* a *MSE* jsou v porovnání s ostatními metodami velmi vysoké. S tím souvisí i vysoký počet záporných odhadů parametru b_1 modelu (a) a (d). Výjimkou jsou hodnoty u modelu (f), kde tato metoda dává dobré hodnoty i *sd* a *MSE*. Je možné, že je to způsobené nízkou hodnotou parametru a .

	výběrový průměr	výběrový rozptyl
<i>PlzenRychlost2018</i>	2.488	5.547
<i>PlzenAlkohol2018</i>	0.975	1.167
<i>PrahaRychlost2018</i>	2.299	8.298

Tabulka 4.1: Výběrový průměr a rozptyl pozorovaných řad.

Ještě bychom měli zmínit, že odhadování metodou maximální věrohodnosti je pomalé. Odhadnout parametry 1000 řad o 500 pozorování trvá několik minut, kdežto odhadnout parametry jakoukoli z metod nejmenších čtverců trvalo nanejvýš pár sekund.

Obecně se dá říct, že metoda MLE je celkem spolehlivá a vhodná pro všechny hodnoty parametrů. Metodu CLS bychom doporučili, pokud očekáváme nízkou hodnotu parametru b_1 a máme dostatek pozorování. Pokud chceme použít metodu rychlejší než MLE, ale s podobnými vlastnostmi, můžeme použít metodu CLSu. Tato metoda se hodí hlavně na modely s vyšší hodnotou parametru b_1 . Pokud chceme odhad s nejnižším vychýlením zvolíme metodu CLSw.

Odhad metodou CLS má jako jediný explicitně vyjádřený asymptotický rozptyl. Skutečné hodnoty směrodatných odchylek pro jednotlivé počty pozorování a hodnoty parametrů uvádíme v tabulkách 4.6 a 4.7 ve sloupci sd^* . Obecně se dá říct, že odhady směrodatné odchylky celkem odpovídají skutečným směrodatným odchylkám, ačkoli konvergují poměrně pomalu.

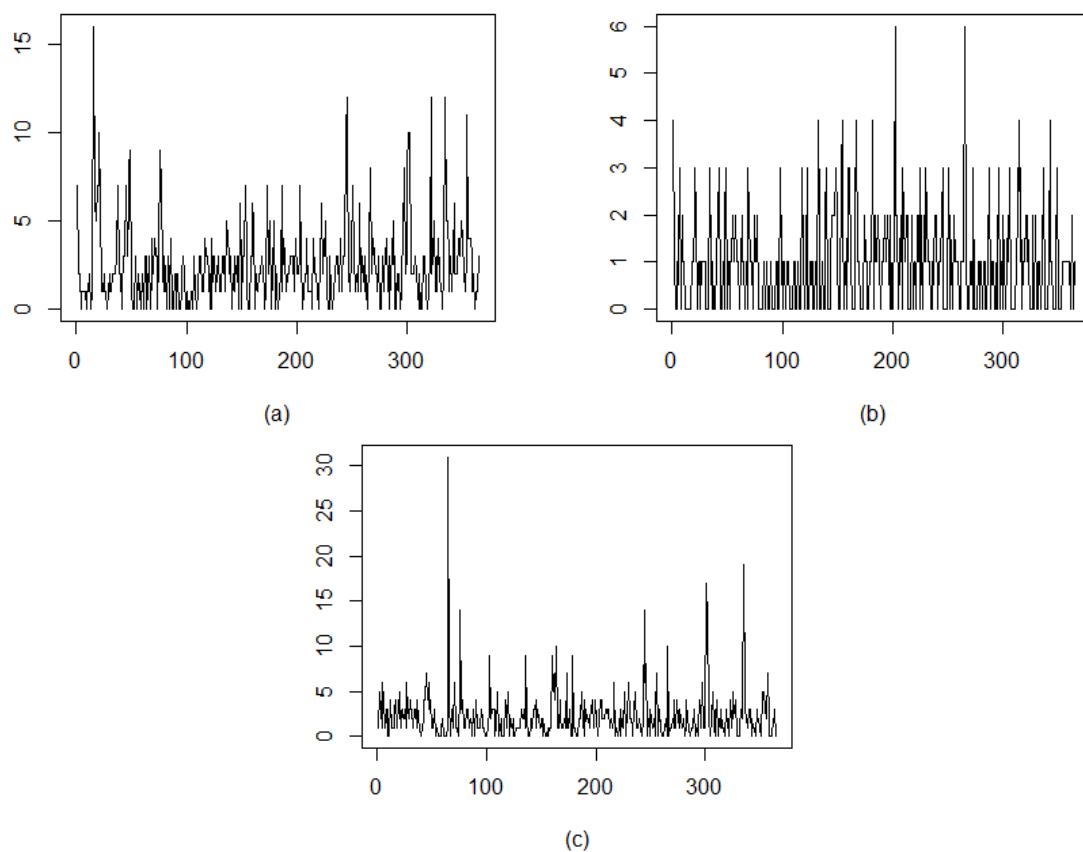
4.3 Reálná data

Použití modelů budeme ilustrovat na třech časových řadách popisujících denní počet nehod za rok 2018¹ (obr. 4.1). První časová řada popisuje nehody v Plzeňském kraji kvůli nepřiměřené rychlosti (*PlzenRychlost2018*), druhá řada popisuje nehody v Plzeňském kraji kvůli jízdě pod vlivem alkoholu (*PlzenAlkohol2018*) a třetí řada popisuje nehody v Praze kvůli nepřiměřené rychlosti (*PrahaRychlost2018*). V tabulce 4.1 uvádíme základní statistiky, na obrázcích 4.2 zase zobrazujeme autokorelační a parciální autokorelační funkci řad. Naším cílem je modelovat tyto řady pomocí INARCH, případně INGARCH modelu.

Začneme řadou *PlzenRychlost2018*. Na základě tvaru autokorelační a parciálně autokorelační funkce (obr. 4.2a) jsme se rozhodli tuto řadu modelovat pomocí INARCH(1) modelu, přičemž očekáváme, že parametr b_1 se bude pohybovat okolo hodnoty 0.3.

K odhadu parametrů využijeme metodu maximální věrohodnosti, jak bylo doporučeno v předchozí kapitole. Jak jsme předpokládali, odhadu b_1 je roven 0.3279 a téměř odpovídá výběrové korelaci pozorování v čase t s pozorováním v čase $t-1$, která je rovna 0.3287. Odhady spolu se směrodatnými chybami a intervaly spolehlivosti jsou uvedené v tabulce 4.2. Směrodatné chyby spočítané z odhadu očekávané informační matice (2.14) jsou na tři desetinná místa shodné se směrodatnými chybami spočítanými z pozorované informační matice (2.13), proto je jedno, jaký odhad informační matice zvolíme. Pro porovnání uvádíme i odhady metodami založenými na metodě nejmenších čtverců (tab. 4.3).

¹Zdroj: <https://www.irozhlas.cz/nehody>



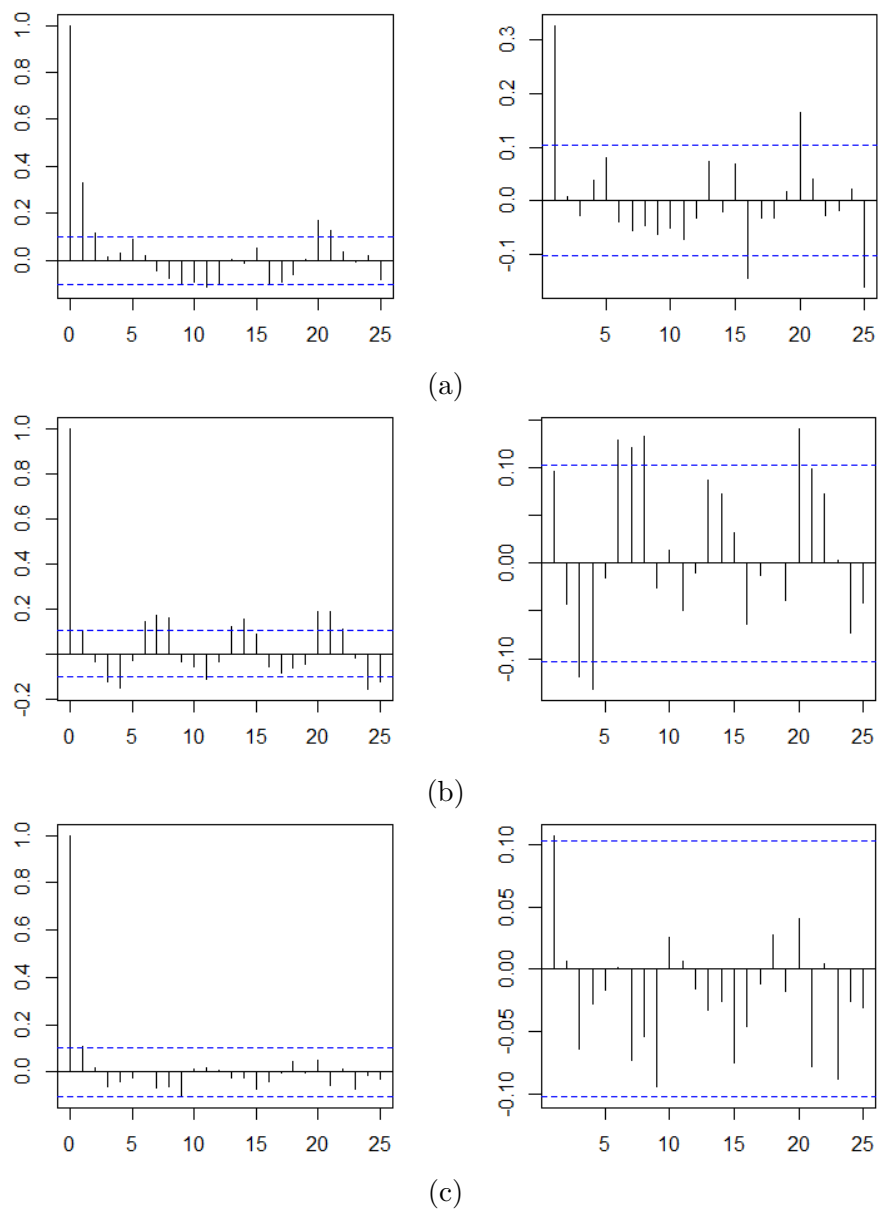
Obrázek 4.1: Počty nehod za rok 2018: (a) v Plzeňském kraji kvůli nepřiměřené rychlosti; (b) v Plzeňském kraji kvůli jízdě pod vlivem alkoholu; (c) v Praze kvůli nepřiměřené rychlosti.

	odhad	směrodatná chyba	interval spolehlivosti
a	1.674	0.113	(1.452;1.897)
b_1	0.328	0.041	(0.247; 0.409)

Tabulka 4.2: Odhad, směrodatná chyba spočítaná z pozorované informační matice (2.13) a interval spolehlivosti INARCH(1) modelu řady *PrahaRychlost2018*.

	MLE	CLS	CLS _u	CLS _w
a	1.6742	1.6602	1.6655	1.6640
b_1	0.3279	0.3278	0.3257	0.3240

Tabulka 4.3: Odhady metodou maximální věrohodnosti (MLE), nejmenších čtverců (CLS), nejmenších čtverců s neznámými vahami (CLS_u) a nejmenších čtverců se známými vahami (CLS_w) pro INARCH(1) model řady *PrahaRychlost2018*.



Obrázek 4.2: Autokorelační funkce a parciální autokorelační funkce řad (a) *PlzenRychlost2018*, (b) *PlzenAlkohol2018* a (c) *PrahaRychlost2018*.

	odhad	SE	int. spol.	p-hodnota
a	0.001	0.005	(-0.010; 0.011)	0.885
b_1	0.067	0.029	(0.011; 0.123)	0.019
b_2	-0.078	0.031	(-0.138; -0.017)	0.012
c_7	0.958	0.029	(0.901; 1.016)	0.000

Tabulka 4.4: Odhad, směrodatná chyba, interval spolehlivosti a p-hodnoty hypotéz, zda je parametr rovný nule, INGARCH modelu řady *PlzenAlkohol2018*.

Pokračujeme řadou *PlzenAlkohol2018*. Na průběhu autokorelační a parciální autokorelační funkce si můžeme všimnout dvou věcí. Zaprvé je u nich znát sedmidenní cyklus a za druhé čtvrtá i pátá hodnota (u parciální autokorelační funkce třetí a čtvrtá) jsou znatelně záporné. Po vzoru Liboschik a kol. (2017) zvolíme INGARCH model a zahrneme do modelu i o 7 dní zpožděné podmíněné střední hodnoty řady. Navíc kvůli záporným hodnotám zvolíme loglineární model se čtyřmi zpožděnými pozorováními. Předpokládáme tedy podmíněnou střední hodnotu řady, pro kterou platí

$$\log \lambda_t = a + b_1 \log (Y_{t-1} + 1) + b_2 \log (Y_{t-2} + 1) + b_3 \log (Y_{t-3} + 1) + b_4 \log (Y_{t-4} + 1) + c_7 \log \lambda_{t-7}, \quad (4.4)$$

kde Y_t je hodnota řady *PlzenAlkohol2018* v čase t a λ_t je podmíněná střední hodnota *PlzenAlkohol2018* v čase t . K odhadu použijeme znovu metodu maximální věrohodnosti.

Parametry b_1 , b_3 i b_4 jsou nesignifikantní, navíc p-hodnota Waldova testu, zda jsou parametry b_3 a b_4 rovny nule, vyšla 0.072, mohli bychom je proto vyřadit z modelu. Výsledné odhady spolu se směrodatnými chybami, intervaly spolehlivosti a p-hodnotami jsou v tabulce 4.4.

Způsob modelování sezónnosti použitý v tomto případě není jediný, který bychom mohli použít. Další možností je místo zpožděné podmíněné střední hodnoty přidat do modelu identifikátor dne v týdnu. Funkce `tsglm` toto umožňuje skrz vstupní parametr `xreg`.

Nakonec se podíváme na řadu *PrahaRychlost2018*. Podle autokorelační a parciální autokorelační funkce to vypadá, že pozorování v čase t nezávisí na minulých hodnotách řady. Zkusíme tedy řadu modelovat INARCH(1) modelem. Protože předpokládáme nízkou hodnotu parametru b_1 , odhadneme parametry metodou nejmenších čtverců. Jak bylo řečeno v předchozí kapitole, pro nízké b_1 a vysoký počet pozorování je tato metoda skoro stejně dobrá jako metoda maximální věrohodnosti. Odhady metodou CLS jsou uvedené v tabulce 4.5. P-hodnota u parametru b_1 je nižší než 0.5, na hladině 5 % proto zamítáme hypotézu, že je parametr b_1 roven 0. Pozorování v čase t tedy závisí na předchozí hodnotě, ačkoli ne tolik jako třeba u *PlzenRychlost2018*.

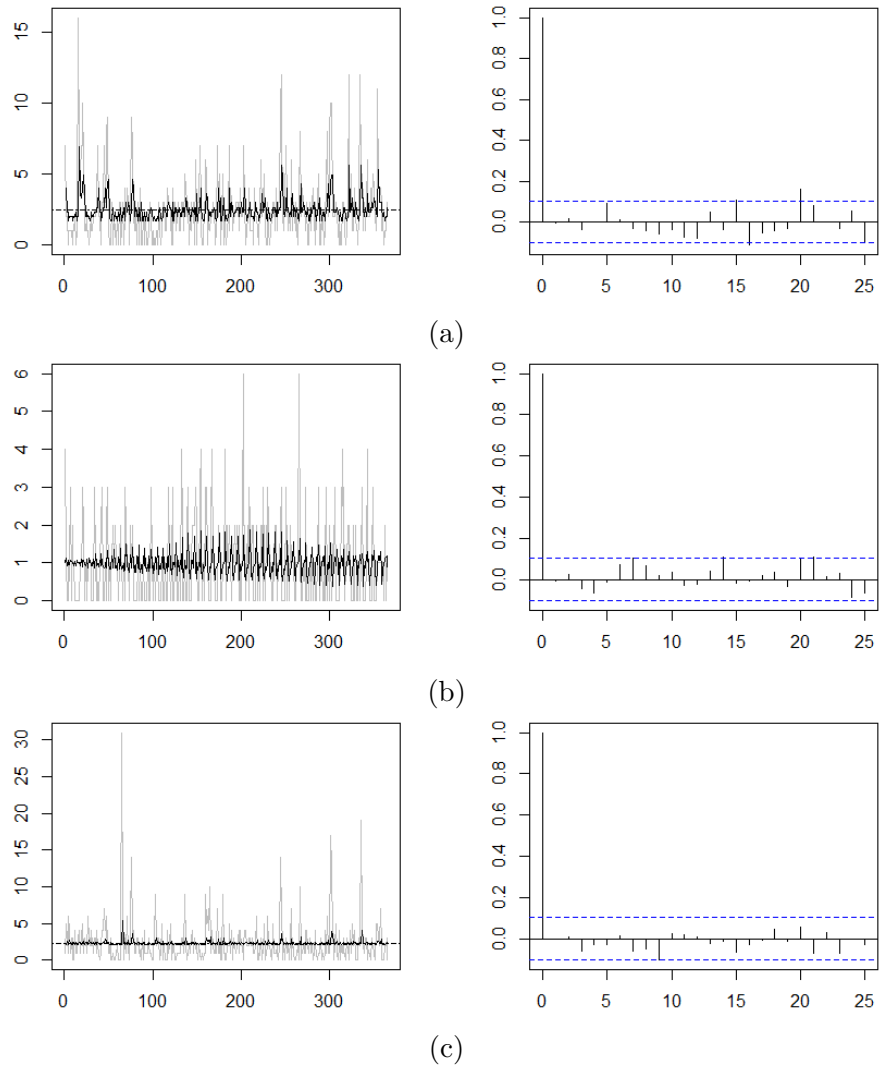
Výsledky modelů můžeme porovnat v grafech 4.3. V grafu řady *PlzenAlkohol2018* (graf 4.3b) si můžeme všimnout, že hodnoty podmíněné střední hodnoty jsou ze začátku velmi blízko a postupně se rozšiřují. To je pravděpodobně způsobené volbou počátečních hodnot podmíněné střední hodnoty a jejich derivací. Na podmíněných středních hodnotách je také zřejmý vliv zpožděné podmíněné střední hodnoty zahrnuté do modelu, která vytváří jistou periodicitu. V grafu

	odhad	SE (1)	SE (2)	int. spol. (2)	p-hodnota (2)
a	2.054	0.107	0.141	(1.777; 2.331)	0.000
b_1	0.108	0.035	0.053	(0.003; 0.212)	0.0434

Tabulka 4.5: Odhad, směrodatná chyba spočítaná (1) ze součinu tří matic z věty 8, přičemž každá z matic je odhadnuta jako průměr příslušných výrazů, (2) z explicitně vyjádřeného rozptylu odhadů metodou CLS (2.21), dále interval spolehlivosti a p-hodnoty hypotéz, zda je parametr roven nule, INARCH(1) modelu řady *PrahaRychlost2018*.

řady *PrahaRychlost2018* (graf 4.3c) je patrné, že parametr b_1 je nižší než v případě řady *PlzenRychlost2018* (graf 4.3a), protože podmíněná střední hodnota tolik nereaguje na vysoké výkyvy v řadě.

Po porovnání autokovarianční funkce reziduí s autokovariančními funkcemi samotných řad je vidět, že autokorelace reziduí je zřetelně nižší než autokorelace původní řady. Nejvýraznější rozdíl je u autokorelace nižších řádů. U reziduí modelu řady *PlzenAlkohol2018* je stále vidět mírná sezónnost, ale už není tak zřetelná jako u samotné řady.



Obrázek 4.3: Vlevo pozorovaná řada (šedě) s odhadnutými podmíněnými středními hodnotami (plná čára) a nepodmíněnou střední hodnotou (čárkovaná), vpravo autokovarianční funkce reziduí modelu řady (a) *PlzenRychlost2018*, (b) *PlzenAlkohol2018* a (c) *PrahaRychlost2018*.

	# poz.	MLE						CLS						CLS _u						CLS _{sw}					
		bias	sd	MSE	out	bias	sd	sd*	MSE	out	bias	sd	MSE	out	bias	sd	MSE	out	bias	sd	MSE	out			
(a) $a = 1$	50	0.0230	0.2107	0.0449	0	0.0348	0.4752	0.2223	0.0506	0	0.0364	0.2212	0.0502	0	0.0332	0.2311	0.0545	0	0.0332	0.2311	0.0545	0			
	100	0.0080	0.1546	0.0239	0	0.0119	0.3343	0.1585	0.0252	0	0.0122	0.1576	0.0250	0	0.0089	0.1639	0.0269	0	0.0089	0.1639	0.0269	0			
	200	0.0089	0.1090	0.0120	0	0.0091	0.2358	0.1094	0.0120	0	0.0100	0.1091	0.0120	0	0.0090	0.1147	0.0132	0	0.0090	0.1147	0.0132	0			
(a) $b_1 = 0.2$	500	0.0017	0.0670	0.0045	0	0.0020	0.1489	0.0674	0.0045	0	0.0023	0.0670	0.0045	0	0.0015	0.0709	0.0050	0	0.0015	0.0709	0.0050	0			
	50	-0.0229	0.1278	0.0169	0	-0.0018	0.1450	0.1170	0.0137	115	-0.0036	0.1166	0.0136	114	0.0130	0.1329	0.0178	154	0.0130	0.1329	0.0178	154			
	100	-0.0120	0.1004	0.0102	0	-0.0063	0.1020	0.0955	0.0092	37	-0.0066	0.0947	0.0090	37	0.0030	0.1058	0.0112	64	0.0030	0.1058	0.0112	64			
(b) $a = 1$	200	-0.0091	0.0725	0.0053	0	-0.0086	0.0719	0.0724	0.0053	4	-0.0093	0.0714	0.0052	4	-0.0067	0.0809	0.0066	12	-0.0067	0.0809	0.0066	12			
	500	-0.0030	0.0461	0.0021	0	-0.0032	0.0454	0.0470	0.0022	0	-0.0034	0.0460	0.0021	0	-0.0025	0.0530	0.0028	0	-0.0025	0.0530	0.0028	0			
	50	0.0834	0.2948	0.0938	0	0.1047	0.4208	0.3044	0.1036	0	0.0919	0.2961	0.0960	0	0.0690	0.2984	0.0937	0	0.0690	0.2984	0.0937	0			
(b) $b_1 = 0.5$	100	0.0369	0.2041	0.0430	0	0.0487	0.2961	0.2172	0.0495	0	0.0417	0.2044	0.0435	0	0.0299	0.2076	0.0439	0	0.0299	0.2076	0.0439	0			
	200	0.0168	0.1361	0.0188	0	0.0265	0.2088	0.1464	0.0221	0	0.0208	0.1362	0.0190	0	0.0143	0.1399	0.0198	0	0.0143	0.1399	0.0198	0			
	500	0.0060	0.0835	0.0070	0	0.0096	0.1319	0.0908	0.0083	0	0.0069	0.0833	0.0070	0	0.0041	0.0861	0.0074	0	0.0041	0.0861	0.0074	0			
(c) $a = 1$	50	-0.0504	0.1447	0.0235	0	-0.0605	0.1403	0.1435	0.0242	3	-0.0542	0.1409	0.0228	3	-0.0402	0.1520	0.0247	4	-0.0402	0.1520	0.0247	4			
	100	-0.0209	0.0957	0.0096	0	-0.0271	0.0987	0.1013	0.0110	0	-0.0237	0.0955	0.0097	0	-0.0174	0.1059	0.0115	0	-0.0174	0.1059	0.0115	0			
	200	-0.0081	0.0661	0.0044	0	-0.0128	0.0696	0.0699	0.0050	0	-0.0102	0.0658	0.0044	0	-0.0066	0.0743	0.0056	0	-0.0066	0.0743	0.0056	0			
(c) $b_1 = 0.8$	500	-0.0029	0.0424	0.0018	0	-0.0047	0.0440	0.0459	0.0021	0	-0.0034	0.0424	0.0018	0	-0.0016	0.0473	0.0022	0	-0.0016	0.0473	0.0022	0			
	50	0.2980	0.5641	0.4066	0	0.3851	0.3343	0.5622	0.4641	1	0.3062	0.5303	0.3747	0	0.2702	0.5706	0.3983	0	0.2702	0.5706	0.3983	0			
	100	0.1312	0.3358	0.1299	0	0.2075	0.2352	0.3643	0.1756	0	0.1461	0.3218	0.1248	0	0.1172	0.3471	0.1341	0	0.1172	0.3471	0.1341	0			
(c) $b_1 = 0.8$	200	0.0557	0.2090	0.0467	0	0.1002	0.1659	0.2405	0.0678	0	0.0621	0.1990	0.0434	0	0.0475	0.2229	0.0519	0	0.0475	0.2229	0.0519	0			
	500	0.0161	0.1176	0.0141	0	0.0395	0.1048	0.1537	0.0252	0	0.0241	0.1177	0.0144	0	0.0180	0.1286	0.0168	0	0.0180	0.1286	0.0168	0			
	50	-0.0656	0.1073	0.0158	0	-0.0828	0.1285	0.1099	0.0189	2	-0.0674	0.1062	0.0158	0	-0.0600	0.1227	0.0186	1	-0.0600	0.1227	0.0186	1			
(c) $b_1 = 0.8$	100	-0.0300	0.0716	0.0060	0	-0.0461	0.0904	0.0767	0.0080	0	-0.0343	0.0702	0.0061	0	-0.0283	0.0826	0.0076	0	-0.0283	0.0826	0.0076	0			
	200	-0.0143	0.0485	0.0026	0	-0.0233	0.0637	0.0527	0.0033	0	-0.0159	0.0465	0.0024	0	-0.0125	0.0573	0.0034	0	-0.0125	0.0573	0.0034	0			
	500	-0.0043	0.0277	0.0008	0	-0.0085	0.0403	0.0345	0.0013	0	-0.0054	0.0279	0.0008	0	-0.0041	0.0338	0.0012	0	-0.0041	0.0338	0.0012	0			

Tabulka 4.6: Odhad vychýlení (*bias*), směrodatné odchytky (*sd*) a střední čtvercová chyba (*MSE*) odhadu jednotlivých parametrů INARCH(1) modelů s parametry (a) $a = 1, b_1 = 0.2$, (b) $a = 1, b_1 = 0.5$, (c) $a = 1, b_1 = 0.8$ metodami maximální věrohodnosti (MLE), nejmenších čtverců (CLS), nejmenších čtverců s neznámými vahami (CLS_u) a nejmenších čtverců se známými vahami (CLS_{sw}) pomocí 50, 100, 200 a 500 pozorování. Navíc počet odhadů mimo parametrický prostor (*out*) pro jednotlivé metody, parametry a počty pozorování a analytická směrodatná odchytka odhadu metodou CLS (*sd**).

	# poz.	MLE				CLS				CLS _u				CLS _w				
		<i>bias</i>	<i>sd</i>	<i>MSE</i>	<i>out</i>	<i>bias</i>	<i>sd</i>	<i>sd</i> *	<i>MSE</i>	<i>out</i>	<i>bias</i>	<i>sd</i>	<i>MSE</i>	<i>out</i>	<i>bias</i>	<i>sd</i>	<i>MSE</i>	<i>out</i>
$a = 2.4$	50	0.0666	0.4249	0.1848	0	0.1000	0.2957	0.4624	0.2236	0	0.0974	0.4626	0.2232	0	0.0845	0.5644	0.3254	0
	100	0.0331	0.3225	0.1050	0	0.0406	0.2081	0.3307	0.1109	0	0.0394	0.3298	0.1102	0	0.0242	0.3952	0.1566	0
	200	0.0210	0.2363	0.0562	0	0.0209	0.1468	0.2375	0.0568	0	0.0222	0.2366	0.0564	0	0.0239	0.2826	0.0803	0
	500	0.0099	0.1426	0.0204	0	0.0094	0.0927	0.1443	0.0209	0	0.0104	0.1425	0.0204	0	0.0119	0.1685	0.0285	0
$b_1 = 0.2$	50	-0.0268	0.1242	0.0161	0	-0.0044	0.1479	0.1121	0.0126	125	-0.0036	0.1125	0.0126	125	0.0391	0.1380	0.0206	197
	100	-0.0143	0.0983	0.0099	0	-0.0082	0.1040	0.0924	0.0086	39	-0.0078	0.0921	0.0085	39	0.0123	0.1174	0.0139	79
	200	-0.0092	0.0731	0.0054	0	-0.0086	0.0734	0.0728	0.0054	3	-0.0090	0.0723	0.0053	3	-0.0037	0.0910	0.0083	28
	500	-0.0033	0.0449	0.0020	0	-0.0032	0.0463	0.0455	0.0021	0	-0.0035	0.0448	0.0020	0	-0.0041	0.0594	0.0035	0
$a = 1.5$	50	0.1176	0.4242	0.1936	0	0.1414	0.2957	0.4351	0.2091	0	0.1292	0.4225	0.1950	0	0.0960	0.4562	0.2172	0
	100	0.0712	0.2767	0.0816	0	0.0846	0.2081	0.2903	0.0913	0	0.0738	0.2783	0.0828	0	0.0524	0.3038	0.0950	0
	200	0.0406	0.1927	0.0388	0	0.0486	0.1468	0.2029	0.0435	0	0.0434	0.1924	0.0389	0	0.0338	0.2138	0.0468	0
	500	0.0135	0.1204	0.0147	0	0.0184	0.0927	0.1299	0.0172	0	0.0149	0.1204	0.0147	0	0.0092	0.1316	0.0174	0
$b_1 = 0.5$	50	-0.0462	0.1360	0.0206	0	-0.0540	0.1479	0.1363	0.0215	2	-0.0501	0.1328	0.0201	2	-0.0369	0.1564	0.0258	4
	100	-0.0270	0.0908	0.0090	0	-0.0326	0.1040	0.0949	0.0101	0	-0.0290	0.0909	0.0091	0	-0.0198	0.1099	0.0125	0
	200	-0.0152	0.0617	0.0040	0	-0.0180	0.0734	0.0656	0.0046	0	-0.0163	0.0617	0.0041	0	-0.0123	0.0765	0.0060	0
	500	-0.0054	0.0385	0.0015	0	-0.0072	0.0463	0.0418	0.0018	0	-0.0061	0.0384	0.0015	0	-0.0035	0.0473	0.0022	0
$a = 0.6$	50	0.1114	0.3034	0.1044	0	0.2200	0.2957	0.3260	0.1546	1	0.1413	0.2965	0.1078	0	0.1057	0.2950	0.0981	0
	100	0.0497	0.1896	0.0384	0	0.1105	0.2081	0.2205	0.0608	0	0.0611	0.1824	0.0370	0	0.0413	0.1817	0.0347	0
	200	0.0214	0.1141	0.0135	0	0.0589	0.1468	0.1467	0.0250	0	0.0269	0.1116	0.0132	0	0.0167	0.1135	0.0131	0
	500	0.0057	0.0673	0.0046	0	0.0223	0.0927	0.0927	0.0091	0	0.0079	0.0670	0.0046	0	0.0042	0.0693	0.0048	0
$b_1 = 0.8$	50	-0.0596	0.1228	0.0186	0	-0.1020	0.1479	0.1260	0.0263	2	-0.0744	0.1191	0.0197	0	-0.0539	0.1233	0.0181	1
	100	-0.0276	0.0765	0.0066	0	-0.0505	0.1040	0.0862	0.0100	1	-0.0333	0.0751	0.0067	0	-0.0231	0.0789	0.0068	0
	200	-0.0117	0.0492	0.0026	0	-0.0254	0.0734	0.0582	0.0040	0	-0.0148	0.0488	0.0026	0	-0.0091	0.0526	0.0028	0
	500	-0.0033	0.0309	0.0010	0	-0.0086	0.0463	0.0370	0.0014	0	-0.0038	0.0307	0.0010	0	-0.0018	0.0340	0.0012	0

Tabulka 4.7: Odhad vychýlení (*bias*), směrodatné odchyly (*sd*) a střední čtvercová chyba (*MSE*) odhadu jednotlivých parametrů INARCH(1) modelů s parametry (d) $a = 2.4, b_1 = 0.2$, (e) $a = 1.5, b_1 = 0.5$, (f) $a = 0.6, b_1 = 0.8$ metodami maximální věrohodnosti (MLE), nejmenších čtverců (CLS), nejmenších čtverců s neznámými vahami (CLS_u) a nejmenších čtverců se známými vahami (CLS_w) pomocí 50, 100, 200 a 500 pozorování. Navíc počet odhadů mimo parametrický prostor (*out*) pro jednotlivé metody, parametry a počty pozorování a analytická směrodatná odchylnka odhadu metodou CLS (*sd**).

Závěr

V této práci jsme představili lineární INARCH model, který je určen pro modelování časových řad počtu událostí. Ukázalo se, že výhodou tohoto modelu je, že se nemusíme omezovat jen na řady, kde je střední hodnota rovna rozptylu, jak je tomu u Poissonova rozdělení. Obecně je tento model vhodný pro stacionární řady s vyšším rozptylem než střední hodnotou a kladnou autokorelační funkcí.

Odvodili jsme základní momenty a vlastnosti procesu splňujícího INARCH model a ukázali jsme, že INARCH model spadá do AR modelů. Představili jsme i několik metod odhadu parametrů modelu - metodu maximální věrohodnosti, metodu nejmenších čtverců, metodu vážených nejmenších čtverců s neznámými vahami, metodu vážených nejmenších čtverců se známými vahami a metodu momentů. Všechny metody odhadu jsme pak pomocí simulací porovnali. Obecně nejlepší se prokázala metoda maximální věrohodnosti. Dobré výsledky dávala i metoda nejmenších čtverců s neznámými vahami.

Dále jsme uvedli různá zobecnění lineárního INARCH modelu. Protože tento model předpokládá jen závislost podmíněné střední hodnoty na minulých pozorování, přidali jsme po vzoru ARMA modelu i závislost na podmíněné střední hodnotě, čímž jsme dostali INARCH model. Také jsme uvedli nelineární modely, speciálně log-lineární model, které nás neomezují jen na řady s kladnou autokorelační funkcí.

Nakonec jsme představené modely využili k modelování počtu nehod v Praze a v Plzeňském kraji, přičemž jsme využili jak lineární, tak nelineární modely.

Další možná rozšíření této práce by mohla zahrnovat například modely s podmíněným negativně binomickým rozdělením nebo modely s vysvětlujícími proměnnými. Také je tu prostor k rozšíření diagnostických metod o testy dobré shody.

Seznam použité literatury

- ANDĚL, J. (2011). *Základy matematické statistiky*. Matfyzpress, Praha.
- BILLINGSLEY, P. (1961). The Lindeberg-Lévy Theorem for Martingales. *Proceedings of the American Mathematical Society*, **12**(5), 788–792.
- CIPRA, T. (2008). *Finanční ekonometrie*. Ekopress, Praha.
- FERLAND, R., LATOUR, A. a ORAICHI, D. (2006). Integer-valued GARCH processes. *Journal of Time Series Analysis*, **27**, 923–942.
- FOKIANOS, K. (2012). Count time series models. In RAO, T. S., RAO, S. S. a RAO, C. R., editors, *Handbook of Statistics. Time Series Analysis: Methods and Applications*, volume 30, pages 315–347. Elsevier.
- FOKIANOS, K. (2015). Statistical analysis of count time series models: A GLM perspective. In DAVIS, R. A., HOLAN, S. H., LUND, R. a RAVISHANKER, N., editors, *Handbook of Discrete-Valued Time Series*, pages 3–28. Chapman & Hall.
- FOKIANOS, K. a NEUMANN, M. H. (2013). A goodness-of-fit test for Poisson count processes. *Electronic Journal of Statistics*, **7**(0), 793–819.
- FOKIANOS, K. a TJØSTHEIM, D. (2011). Log-linear poisson autoregression. *Journal of Multivariate Analysis*, **102**(3), 563–578.
- FOKIANOS, K. a TJØSTHEIM, D. (2012). Nonlinear poisson autoregression. *Annals of the Institute of Statistical Mathematics*, **64**, 1205–1225.
- FOKIANOS, K., RAHBEK, A. a TJØSTHEIM, D. (2009). Poisson autoregression. *Journal of the American Statistical Association*, **104**(488), 1430–1439.
- KLIMKO, L. A. a NELSON, P. I. (1978). On conditional least squares estimation for stochastic processes. *The Annals of Statistics*, **6**(3), 629–642.
- LACHOUT, P. (2004). *Teorie pravděpodobnosti*. Nakladatelství Karolinum, Praha.
- LIBOSCHIK, T., FOKIANOS, K. a FRIED, R. (2017). tscount: An R Package for Analysis of Count Time Series Following Generalized Linear Models. *Journal of Statistical Software*, **82**(5).
- MCKENZIE, E. (2003). Discrete variate time series. In *Stochastic Processes: Modelling and Simulation*, volume 21 of *Handbook of Statistics*, pages 573 – 606. Elsevier.
- MEYN, S. P. a TWEEDIE, R. L. (1993). *Markov chains and stochastic stability*. Springer-Verlag, London.
- NEUMANN, M. H. (2011). Absolute regularity and ergodicity of Poisson count processes. *Bernoulli*, **17**(4), 1268–1284.

- PRÁŠKOVÁ, Z. a LACHOUT, P. (1998). *Základy náhodných procesů I*. Karolinum, Praha.
- STOUT, W. F. (1974). *Almost sure convergence*. Probability and mathematical statistics. Academic Press, New York.
- WEISS, C. H. (2010). The INARCH(1) model for overdispersed time series of counts. *Communications in Statistics - Simulation and Computation*, **39**(6), 1269–1291.
- ZHU, F. a WANG, D. (2011). Estimation and testing for a Poisson autoregressive model. *Metrika*, **73**, 211–230.
- ZHU, F. (2010). A negative binomial integer-valued GARCH model. *Journal of Time Series Analysis*, **32**(1), 54–67.