

UNIVERZA NA PRIMORSKEM  
FAKULTETA ZA MATEMATIKO, NARAVOSLOVJE IN  
INFORMACIJSKE TEHNOLOGIJE

Zaključna naloga

**Primerjava modela z Brownovim gibanjem in ARIMA  
modelom na primeru napovedovanja cen delnic.**

(Forecasting stock price with Brownian motion and ARIMA model.)

Ime in priimek: Nataša Kunc

Študijski program: Matematika v ekonomiji in financah

Mentorica: doc. dr. Arjana Brezigar Masten

**Koper, september 2015**

## Ključna dokumentacijska informacija

Ime in PRIIMEK: Nataša KUNC

Naslov zaključne naloge: Primerjava modela z Brownovim gibanjem in ARIMA modelom na primeru napovedovanja cen delnic

Kraj: Koper

Leto: 2015

Število listov: 44

Število slik: 13

Število tabel: 2

Število referenc: 18

Mentorica: doc. dr. Arjana Brezigar Masten

Ključne besede: napovedovanje cen delnic, Brownovo gibanje, ARIMA model, ARIMA metodologija, simulacija, slučajni hod

Math. Subj. Class. (2010):

### Izvleček:

V zaključni nalogi se ukvarjam z napovedovanjem cen delnic. V uvodu predstavljam nekaj razlogov, zakaj se cene gibljejo naključno. V nadaljevanju obravnavam dva najbolj poznana modela napovedovanja cen delnic, to sta Brownovo gibanje in ARIMA model. Oba teoretično opisujem in izpeljem model oziroma predstavim metodologijo. V zadnjem delu zaključne naloge simuliram oba modela na primeru cen delnice The Coca-Cola Company. Napovedane cene z Brownovim gibanjem veliko odstopajo od dejanskih cen in tako neuspešno napovedujejo gibanje cene. Simulacijo ARIMA modela zaradi neuspešnega iskanja pravih ARIMA modela nisem mogla izvesti, saj so statistični testi pokazali, da imam opraveka s slučajnim hodom. Ugotavljam, da se cene delnic gibljejo naključno, to naključnost pa je zelo težko modelirati in prav tako napovedati cene.

## Key words documentation

Name and SURNAME: Natasa KUNC,

Title of final project paper: Forecasting stock price with Brownian motion and ARIMA model

Place: Koper

Year: 2015

Number of pages: 44

Number of figures: 13

Number of tables: 2

Number of references: 18

Mentor: Assist. Prof. Arjana Brezigar Masten, PhD

Keywords: forecasting stock price, Brownian motion, ARIMA model, ARIMA methodology, simulation, random walk

Math. Subj. Class. (2010):

### **Abstract:**

In my final project paper I deal with forecasting stock prices. In the introduction I introduce some of the reasons why prices move randomly. In the following stage of my final project paper I am discussing two of the most popular model, Brownian motion and the ARIMA model. Both models I describe theoretically and present the methodology. The last part of the final project I simulate the two models in the case of share price The Coca-Cola Company. Announced prices of Brownian motion much from the actual price and so unsuccessfully predict price movement. Simulating the ARIMA model, I could not be conducted due to the search proper ARIMA model, as the statistical tests showed that I have something to do with random walk. I note that the stock prices move randomly and this randomness is very difficult to model and predict future prices.

## Zahvala

Zahvaljujem se mentorici, dr. Arjani Brezigar Master, in mag. Radu Pezdirju za pomoč, čas, nasvete in usmerjanje pri izdelavi zaključne naloge.

Posebno bi se zahvalila svoji družini in prijateljem za vse spodbude, podporo in popestritve med študijem.

# Kazalo vsebine

<b>1</b>	<b>Uvod</b>	<b>1</b>
<b>2</b>	<b>Brownovo gibanje</b>	<b>3</b>
2.1	Matematični opis Brownovega gibanja . . . . .	4
2.1.1	Razlaga Brownove definicije na primeru delniškega trga . . . . .	5
2.2	Izpeljava modela . . . . .	6
2.3	Gibanje cen delnic z Brownovim gibanjem . . . . .	9
<b>3</b>	<b>ARIMA MODEL</b>	<b>11</b>
3.1	Stacionarnost . . . . .	12
3.1.1	Transformiranje podatkov . . . . .	16
3.2	Splošna definicija ARIMA modela . . . . .	17
3.2.1	Integriranost časovnih vrst . . . . .	17
3.2.2	Avtoregresijski model (AR) . . . . .	18
3.2.3	Drseča sredina ali premikajoče povprečje (MA) . . . . .	18
3.2.4	ARMA model . . . . .	18
3.3	ARIMA metodologija . . . . .	19
3.3.1	Identifikacija modela . . . . .	20
3.3.2	Ocenitev ARIMA modela . . . . .	20
3.3.3	Ovrednotev modela . . . . .	21
3.3.4	Napovedovanje . . . . .	21
<b>4</b>	<b>Simulacija</b>	<b>22</b>
4.1	Analiza časovne vrste . . . . .	23
4.2	Napovedovanje cen delnice z Brownovim gibanjem . . . . .	24
4.3	ARIMA model . . . . .	28
4.4	Rezultati . . . . .	34
<b>5</b>	<b>Zaključek</b>	<b>35</b>
<b>6</b>	<b>Literatura in viri</b>	<b>36</b>

## Kazalo tabel

1	Teoretični vzorci AC-grafa . . . . .	20
2	Osnovne statistične lastnosti časovne vrste cen delnice . . . . .	24

## Kazalo slik

1	Korelogram belega šuma . . . . .	13
2	Korelogram slučajnega sprehoda . . . . .	14
3	Graf gibanja cen delnice The Coca-Cola Company (KO) od 25. 11. 2014 do 31. 6. 2015. . . . .	23
4	Izračun donosa posameznih cen v programu Excel . . . . .	25
5	Izračun povprečnega donosa vseh donosov ter varianco v programu Excel	26
6	Izračun letnega donosa in variance v programu Excel . . . . .	27
7	Izračun $V_1$ , $V_2$ in napovedane cene ( $S$ ) v programu Excel . . . . .	28
8	Graf napovedanih cen delnice z Brownovim gibanjem in dejanske cene .	29
9	Avtokorelacijska in delno avtokorelacijska funkcija . . . . .	30
10	Korelogram logaritmiranih cen . . . . .	31
11	Graf prve diference logaritemske časovne vrste cen . . . . .	32
12	Korelogram logaritmiranih cen . . . . .	32
13	Izpis Dickey-Fullerjevega testa v STATA . . . . .	33

## Seznam kratic

*AR* Autoregressive model

*MA* Moving average

*AC* Autocorrelation

*PAC* Partial Autocorrelation

*LB* Ljung-Box test

*ADF* Autocorrelation Function

*OLS* ordinary least squares

*AIC* Akaike Information Criterion

*SC* Schwarz Criterion

*i.i.d.* independent and identical distribution



# 1 Uvod

V zadnjih desetletjih je modeliranje finančnih trgov vzbudilo veliko zanimanja med akademiki in v finančni industriji po vsem svetu. Finančni trg je zelo kompleksen sistem, ki vključuje veliko faktorjev iz psiholoških, družbenih in političnih vidikov splošne gospodarske uspešnosti. Gospodarsko stanje tako najboljše ponazarja delniški trg. Z drugimi besedami je borza ogledalo gospodarstva. Cene delnice so določene po osnovni gospodarski teoriji povpraševanja in ponudbe. [9] Na dolgi rok vplivajo na ceno delnice fundamentalni elementi podjetja, kot so uspešnost podjetja, dobiček, gospodarske novice ... Na kratki rok pa vpliva bolj sentiment vlagateljev. Sentiment vlagateljev, ki se nanaša na psihologijo udeležencev na trgu, neposredno vpliva na ravnovesje med ponudbo in povpraševanjem. Kot rezultat se cena delnice posamezne družbe konstantno spreminja, medtem ko se njeni fundamentalni elementi spreminjajo veliko počasneje. [9]

Predvidljivost cen delnic je tako eden najpomembnejših in privlačnih funkcij za raziskovalce in investitorje. Zaradi razlogov, ki so večina finančni, so cene delnic najbolj analizirani ekonomski podatki v zadnjih štiridesetih letih. [8]

Gibanje cene delnice kot funkcija časa predstavlja finančno časovno vrsto in vsebuje elemente negotovosti, ki zahteva uporabo statističnih metod za njeno analizo. Analiza časovnih vrst je pomemben del statistike, ki analizira podatke za preučevanje njenih karakteristik in pomaga pri napovedovanju prihodnjih vrednosti časovne vrste. [7]

Številni finančni modeli so sestavljeni iz sistematičnega sestavljenega dela, ki vključuje različne merljive dejavnike in naključno komponento, ki predstavlja netržne oziroma resnično naključne dejavnike:

$$Y = f(y_t) + E, \quad (1.1)$$

kjer funkcija  $f(y_t)$  predstavlja funkcijo, ki jo definira posamezni model in s pomočjo  $f(y_t)$  oceni koeficiente modela na podlagi podanih podatkov  $y_t$ .  $E$  predstavlja naključje in je običajno najpomembnejši izraz v modelu. Modele lahko razdelimo v dve vrsti, in sicer po indeksiranju časa  $t$ . Poznamo diskretne modele (čas je število) in zvezne modele (čas v intervalih). Med diskretne modele uvrščamo ARIMA model, med zvezne pa Brownovo gibanje. [7]

Modeli diskretne časovne vrste opisujejo obnašanje stohastičnega gibanja v obliki funkcije kot

$$X_t = f(X_{t-1}, \dots, X_{t-p}, \epsilon_t, \epsilon_{t-1}, \dots, \epsilon_{t-q}), \quad (1.2)$$

kjer je  $\epsilon_i$  slučajna spremenljivka. Ti modeli so predmet uveljavljenega področja analize časovnih vrst v časovni domeni. Modeli z zveznim časom se izražajo v diferenčnih enačbah, ki jih definiramo za posamezen model. [7]

Namen zaključne naloge je teoretično spoznati dva najbolj uporabljena modela za napovedovanje cen delnic, Brownovo gibanje in ARIMA model, ter na podlagi podane teorije na praktični primer aplicirati napovedovanje delnice podjetja The Coca-Cola Company.

V prvem poglavju zaključne naloge bom predstavila Brownovo gibanje. Opisala bom njegovo zgodovino, matematično definicijo in izpeljala model, ki ga bom uporabila v poglavju Simulacije. V drugem poglavju bom na podoben način kot Brownovo gibanje definirala ARIMA model. Tudi tega bom uporabila v poglavju Simulacije. V zadnjem poglavju Simulacije bom zgoraj navedena modela simulirala na časovni vrsti cen delnice The Coca-Cola Company. Simulacijo Brownovega gibanja bom izvedla v programu Excel, simulacijo ARIMA modela pa v programu STATA. Za konec bom dobljene cene iz obeh modelov primerjala z dejanskimi cenami in ocenila njihovo točnost.

## 2 Brownovo gibanje

Zgodovinsko je Brownovo gibanje povezano z analizo gibanja, ki se v času premika neenakomerno, zato je za kratko časovno obdobje težko določiti smer gibanja. Pomembno vlogo igra v teorijah slučajnih procesov, ker predstavlja manj teoretičnih in aplikativnih problemov. Brownovo gibanje omogoča enostavne omejitve modelov, na katere se lahko izvede veliko izračunov.

Brownovo gibanje je dobilo ime po škotskem botaniku Robertu Brown (1773–1858), ki je opisoval nepredvidljivo gibanje drobnih organskih delcev v tekočini. Leta 1900 je Louis Bachelier (1870–1946) predstavil Brownovo gibanje na dinamičnem modelu cen delnic, ampak je njegovo odkritje bilo pozabljeno vse do šestdesetih let prejšnjega stoletja.

Največji razvoj Brownovega gibanja je naredil Albert Einstein (1879–1955), ki je leta 1905 zgradil verjetnostni model za opisovanje položaja delcev. Ugotovil je, da položaj delca v času  $t$  glede na začetni položaj  $x_0$  določi gostoto, ki ustreza toplotni enačbi oziroma Gaussovi. Isto leto je poljski fizik Smoluchowski opisal diskretno različico Brownovega gibanja z uporabo slučajnega sprehoda.

Leta 1923 je N. Wiener (1894–1964) temeljito zgradil slučajno funkcijo Brownovega gibanja. Njegovi ideji je sledilo aktivno teoretično raziskovanje v matematiki. Paul Lévy (1886–1971) je z drugimi matematiki odkril veliko verjetij v Brownovem gibanju in predstavil prvo obliko stohastičnih diferencialnih enačb, študijo, ki jo je kasneje sistematiziral K. Itô (1915–2008). Njegovo zbrano delo, objavljeno leta 1948, je tako znano kot Itôvi stohastični izračuni. [1]

## 2.1 Matematični opis Brownovega gibanja

Slučajni proces Brownovega gibanja označimo z  $B(t)$  za čas  $t \geq 0$ . V grafu bo horizontalna os predstavljala čas in vertikalno vrednost  $B(t)$  v času  $t$ . Brownovo gibanje nekateri imenujejo tudi Wienerjev proces. Je eden najbolj uporabnih stohastičnih procesov in implicira verjetnostno teorijo. [2]

**Definicija 2.1.** Brownovo gibanje  $B(t), t \geq 0$  je stohastični proces z naslednjimi značilnostmi.

1.  $B(0) = 0$ .
2. Neodvisnost povečanj: za vsak časovni interval  $0 = t_1 < t_2 < \dots < t_n$  so slučajne vrednosti  $[B(t_0), B(t_2) - B(t_1), B(t_3) - B(t_2), \dots, B(t_n) - B(t_{n-1})]$  neodvisne in enako porazdeljene.
3. Normalna porazdelitev: za vsak  $0 < s < t$  je  $B(t) - B(s)$  normalno porazdeljen s pričakovano vrednostjo 0 in varianco  $t - s$

$$B(t) - B(s) \sim N(0, t - s).$$

4. Zveznost:  $B(t)$  za  $t > 0$  je zvezna funkcija po  $t$ .

Gostota Brownovega gibanja:

$$f_{B(t)}(x) = \frac{1}{\sqrt{2\pi t}} \exp -\frac{x^2}{2t}.$$

Navedimo še nekaj matematičnih lastnosti Brownovega gibanja, ki sledijo iz definicije:

1.  $B(t)$  je Gaussov proces, to pomeni, da ima za vse  $0 \leq t_1 \leq t_2 \leq \dots \leq t_k$  slučajni vektor  $\mathbb{Z} = (B(t_1), B(t_2), \dots, B(t_k)) \in \mathbb{R}$  normalno porazdelitev.
2.  $B_t$  ima stacionarni prirastek, to pomeni, da ima premik  $(B_{t+h} - B_t)$  za  $h \geq 0$  in vsak  $t$  enako porazdelitev z  $E(B_{t+h} - B_t) = 0$  in  $Var(B_{t+h} - B_t) = h$ .
3. Brownovo gibanje je martingal.
4.  $Cov(B_s, B_t) = \min(s, t)$ . [3]

### 2.1.1 Razlaga Brownove definicije na primeru delniškega trga

V tem poglavju bom predstavila Brownovo definicijo na podlagi delniškega trga in s tem obrazložila, zakaj se Brownovo gibanje lahko uporablja za cene delnic.

1.  $B(0) = 0$ , če ceno postavimo na vrednost 0. (Nima pomembne vloge.)
2. Neodvisnost povečanj se zdi dovolj razumljiva lastnost Brownovega gibanje vsaj na dolgi rok. Če bi gledali ceno delnice za vsako sekundo, bi lahko rekli, da ta lastnost ne velja, vendar če jo spremljamo enkrat dnevno, lahko vidimo, da se giblje neodvisno od prejšnjega dne. Ravno zaradi te neodvisnosti gibanja cen je na delniškem trgu priljubljeno Brownovo gibanje. Ekonomisti to lastnost imenujejo hipoteza učinkovitega trga (Efficient Market Hypothesis) ali hipoteza slučajnega sprehoda (Random Walk Hypothesis).
3. Lastnost normalne porazdelitve Brownovega gibanja lahko apliciramo na spremembe cen delnic. To lahko trdimo po centralnem limitnem izreku, ki skupaj z neodvisnostjo predstavlja normalno porazdelitev. [3]

Drugi razlog normalne porazdelitve je enostavna uporaba. Normalna porazdelitev uporablja enostavno funkcijo, s katero se nekateri izračuni poenostavijo. Dobra lastnost normalne porazdelitve je tudi, da je kumulativna vsota neodvisnih slučajnih spremenljivk normalno porazdeljena. Lastnost normalne porazdelitve ignorira majhno verjetnost negativne spremembe cen delnic kljub velikim negativnim spremembam. [3]

Lastnost normalne porazdelitve govori tudi o konstantni varianci na intervalih iste dolžine. V tem primeru glede na spremembo cen delnic predpostavka ni dobra, saj spremembe cen niso enako volativne na vseh intervalih iste dolžine.

4. Zadnja predpostavka kumulative je sicer matematična abstrakcija zbranih podatkov, ampak je smiselna. Cene delnic se spreminjajo vsak dan oziroma tudi iz sekundo v sekundo za majhne premike. Tako se cene spreminjajo zvezno. [3]

Zdaj poznamo razloge za uporabe Brownovega gibanja pri cenah delnice. V poglavju Izpeljava modela bom vse našete definicije in lastnost pretvorila v strukturirano enačbo, s katero bom kasneje ocenila napovedane cene izbrane delnice.

## 2.2 Izpeljava modela

Za izpeljavo Brownovega modela z vsemi lastnostmi Brownovega gibanja začnemo s standardnim Brownovim gibanjem. Za standardno Brownovo gibanje najprej vzamemo interval  $[0, T]$  in ga razdelimo na  $N$  mnogo podintervalov:  $0 = t_0 < t_1 < \dots < t_N = T$  in rekurzivno izračunamo: [5]

$$\Delta B(t_i) = B(t_i) - B(t_{i-1}) = \epsilon_i \sqrt{\Delta t_i} \quad (2.1)$$

z  $\epsilon \sim \phi(0, 1)$  in i.i.d. slučajnimi spremenljivkami ter  $\Delta t_i = t_i - t_{i-1}$ . [5]

Iz enačbe (2.1) sledi, da je tudi  $\Delta B(t_i)$  sama porazdeljena normalno z

$$E(\Delta B(t_i)) = 0,$$

$$sd(\Delta B(t_i)) = \sqrt{\Delta t_i},$$

$$var(\Delta B(t_i)) = \Delta t_i.$$

Da bi izračunali spremembo na celotnem intervalu  $B(T) - B(0)$ , lahko vsoto vseh  $N$  podintervalov z dolžino  $\Delta t$  zapišemo kot:

$$N = \frac{T}{\Delta t}.$$

Torej

$$B(T) - B(0) = \sum_{i=1}^N \epsilon_i \sqrt{\Delta t_i}, \quad (2.2)$$

kjer je za vsak  $\epsilon_i (i = 1, 2, \dots, N)$  porazdeljen  $\phi(0, 1)$  in so neodvisni med seboj. [4] Če sledimo iz enačbe (2.2), je  $B(T) - B(0)$  normalno porazdeljen z

$$E(B(T) - B(0)) = 0,$$

$$var(B(T) - B(0)) = N\Delta t = T,$$

$$sd(B(T) - B(0)) = \sqrt{T}.$$

Če spremembo  $\Delta t$  limitiramo k ničli ( $\Delta t \rightarrow 0$ ), dobimo  $\Delta x = \alpha \Delta t$ . V diferenčni enačbi lahko to zapišemo kot: [5] [4]

$$dX(t) = \alpha dt. \quad (2.3)$$

Naslednja stopnja razširitve Brownovega gibanja je aritmetično Brownovo gibanje ali Brownovo gibanje z zdrsom, za kar se uporablja tudi izraz generalizirani Wienerjev proces. [5] Splošnemu Brownovemu gibanju se dodata dve konstanti  $\alpha$  (stopnja zdrsa) in  $\sigma$  (stopnja variance). Tako dobimo naslednjo diferenčno enačbo:

$$dX(t) = \alpha dt + \sigma dB(t). \quad (2.4)$$

Desno stran enačbe (2.4) lahko razdelimo na dva dela. Prvi del  $\alpha dt$ , ki implicira, da ima  $X$  pričakovano stopnjo zrsa  $\alpha$  na enoto časa. Drugi del  $\sigma dB(t)$  dodaja vrednosti  $X$  šum oz. variabilnost. [4] Za majhne časovne intervale ( $\Delta t \rightarrow 0$ ) lahko s pomočjo enačb (2.1) in (2.4) zapišemo:

$$\Delta X = \alpha \Delta t + \sigma \epsilon \sqrt{\Delta t} \quad (2.5)$$

in je tako kot pri enačbi (2.1)  $\epsilon \sim \psi(0, 1)$ . Ravno tako je  $\Delta X$  porazdeljen normalno z lastnostmi [4]:

$$E(\Delta X) = \alpha \Delta t,$$

$$var(\Delta X) = \sigma^2 \Delta t,$$

$$sd(\Delta X) = \sigma \sqrt{\Delta t}.$$

Če želimo izračunati vrednost  $X$  v neki točki  $t$ , moramo enačbo 3 integrirati:

$$\int_0^t dX(s) = \alpha \int_0^t dt + \sigma \int_0^t dB(s). \quad (2.6)$$

Integrale izračunamo in dobimo:

$$X(t) = X(0) + \alpha t + \sigma B(t). \quad (2.7)$$

Glede na enačbo (2.7) lahko aritmetično Brownovo gibanje postavimo v diskreten čas na interval  $[0, T]$  z danimi parametri  $\alpha$  in  $\sigma$ . Interval  $[0, T]$  razdelimo na  $N$  delčkov dolžine  $T/N$ . Začnemo pri  $X(0) = 0$  in integriramo za  $i = 1, \dots, N$  korakov in dobimo rekurzivno enačbo: [5] [4]

$$X(i) = X(i-1) + \alpha \frac{T}{N} + \sigma \epsilon_i \sqrt{\frac{T}{N}}. \quad (2.8)$$

Diskretno časovno gibanje  $X(i)$ , definirano v enačbi, je torej slučajni proces s stopnjo zdrs  $\alpha$  in konstantno varianco  $\sigma^2$  na enoto časa ter normalno porazdelitvijo, za katero pa na sliki vidimo, da ni stacionarna, čeprav so premiki neodvisni in stacionarni. Te trditve povedo, da so pretekle informacije irelevantne za prihodnje vrednosti gibanja. [5]

Naslednja razširitev Brownovega gibanja se imenuje geometrijsko Brownovo gibanje, kjer uporabimo Itovo lemo.

**Lema 2.2.** Če je  $X(t)$  zvezen stohastični proces, potem velja enačba:

$$dX(t) = a(X, t)dt + b(X, t)dB(t), \quad (2.9)$$

kjer je  $B(t)$  Brownovo gibanje in so  $a(X, t)$  in  $b(X, t)$  funkcije, odvisne od  $X$  in  $t$ .

Geometrijsko Brownovo gibanje je v splošnem Itov proces, kjer je  $a(X, t) = \alpha X(t)$  in  $b(X, t) = \sigma X(t)$ . Tako dobimo diferencialno enačbo za geometrijsko Brownovo gibanje [5]:

$$dX(t) = \alpha X(t)dt + \sigma X(t)dB(t). \quad (2.10)$$



## 2.3 Gibanje cen delnic z Brownovim gibanjem

V tem poglavju bom predstavila slučajni proces gibanja cen delnic. Že na začetku sem z definicijo Brownovega gibanja razložila, zakaj lahko Brownovo gibanje apliciramo na gibanje cen delnic. Obstaja namreč več razširitev Brownovega gibanja: standardno, aritmetično in geometrijsko Brownovo gibanje.

Najboljši model za gibanje cen delnic je geometrijsko Brownovo gibanje. Gibanje cen delnic ne more slediti standardnemu Brownovemu gibanju, saj cena delnice ne more biti nikoli negativna, kar pa standardno Brownovo gibanje dopušča. [3]

Cena delnice ravno tako ne more slediti aritmetičnemu Brownovemu gibanju zaradi konstantne stopnje zdrsa oziroma konstantne mere rasti in konstantne stopnje variance. Če ima delnica konstantno mero rasti, potem lahko investitor zahteva odstotni donos na delnico neodvisno od njene cene. Očitno je, da je za modeliranje cen delnic predpostavka o konstantni meri rasti in variance napačna. Tudi pri aritmetičnem Brownovem gibanju je verjetnost, da lahko cene delnice zavzamejo negativno vrednost, pozitivna. [5]

Ker je predpostavka o konstantni meri napačna, jo je treba nadomestiti s konstantnim pričakovanim donosom. Če  $S(t)$  označuje ceno delnice v času  $t$ , potem je pričakovana rast cene delnice za neki konstantni parameter  $\mu$  enaka  $\mu S$ . To pomeni, da je za kratek časovni interval  $\Delta t$  pričakovana rast cene  $S$  enaka  $\mu S \Delta t$ . Parameter  $\mu$  je torej pričakovani donos delnice v decimalnem zapisu. [4]

Če cena delnice ne bi bila volatilna, bi lahko model zapisali kot:

$$dS = \mu S dt. \quad (2.11)$$

Če to enačbo integriramo za časovni interval  $[0, T]$ , dobimo

$$S(T) = S(0)e^{\mu T}, \quad (2.12)$$

kjer sta  $S(T)$  in  $S(0)$  cene delnice v času 0 in  $T$ . Enačba (2.12) nam pove, da cena delnice stalno raste za stopnjo  $\mu$  na enoto časa [4]. V praksi so cene volatilne in zato moramo v model vključiti volatilitnost. Volatilitnost delnice lahko izračunamo kot standardni odklon spremembe v kratkem časovnem obdobju proporcionalno s ceno delnice. Tako dobimo model [4]:

$$dS = \mu S dt + \sigma S dB(t)$$

oz.

$$\frac{dS}{S} = \mu dt + \sigma dB(t). \quad (2.13)$$

Zgornja enačba je najbolj razširjena in uporabna za modeliranje gibanja cen delnic. Parameter  $\sigma$  predstavlja volatilitnost delnice,  $\mu$  pa pričakovani donos na delnico. [4]

Za kasnejše izračune moram enačbo (2.13) prestaviti v diskretni čas. Tako dobim enačbo

$$\frac{\Delta S}{S} = \mu \Delta t + \sigma \epsilon \sqrt{\Delta t} \quad (2.14)$$

oz.

$$\Delta S = \mu S \Delta t + \sigma S \sqrt{\Delta t} \epsilon.$$

Spremenljivka  $\Delta S$  predstavlja spremembo cene delnice  $S$  v časovnem obdobju  $\Delta$  in  $\epsilon$  ima standardno normalno porazdelitev  $\psi(0, 1)$ . Parameter  $\mu$  je pričakovani donos na enoto delnice in  $\sigma$  predstavlja volatilitnost delnice. Za enostavni izračun napovedovanja cen delnic bosta parametra  $\mu$  in  $\sigma$  konstanti. [4]

Leva stran enačbe (2.14) predstavlja relativno spremembo cene delnice v času  $\Delta t$ . Desno stran enačbe (2.14) pa lahko razdelimo na dva dela, in sicer na  $\mu \Delta t$  in  $\sigma \epsilon \sqrt{\Delta t}$ . Prvi del  $\mu \Delta t$  predstavlja pričakovani donos vrednosti delnice,  $\sigma \epsilon \sqrt{\Delta t}$  pa stohastično komponento donosa. Varianca stohastične komponente je enaka  $\sigma^2 \Delta t$ . [4]

Enačba (2.14) prikazuje normalno porazdelitev  $\frac{\Delta S}{S}$  s pričakovano vrednostjo  $\mu \Delta t$  in standardnim odklonom  $\sigma \sqrt{\Delta t}$ , torej

$$\frac{\Delta S}{S} \sim \phi(\mu \Delta t, \sigma^2 \Delta t). \quad (2.15)$$

Ker je cilj zaključne naloge napovedati prihodnjo ceno delnice, je treba relativno spremembo cene skozi čas spremeniti v rekurzivno obliko [4]

$$\frac{\Delta S}{S} = \frac{S(t-1) - S(t)}{S(t)}. \quad (2.16)$$

Tako dobimo enačbo

$$\frac{S(t-1) - S(t)}{S(t)} = \mu \Delta t + \sigma \epsilon \sqrt{\Delta t}. \quad (2.17)$$

Če je  $t = 0$ , dobimo

$$\frac{S(1) - S(0)}{S(0)} = \mu \Delta t + \sigma \epsilon \sqrt{\Delta t} \quad (2.18)$$

oziroma

$$S(1) = S(0) + \mu S(0) \Delta t + \sigma S(0) \epsilon \sqrt{\Delta t}. \quad (2.19)$$

Enačba (2.19) je naša končna enačba modela, ki jo bom v poglavju Simulacija uporabila za napovedovanje cen delnic z Brownovim gibanjem.

## 3 ARIMA MODEL

ARIMA model<sup>1</sup> je napredna ekonometrična modelirna tehnika. ARIMA gleda v zgodovino časovne vrste podatkov in uprizori prilegajočo optimizacijo ob upoštevanju zgodovinske avtokorelacije, popravi problem nestacionarnosti podatkov. Skozi celoten postopek sestave model popravlja napovedovalno napako.

ARIMA model je bil prvič predstavljen leta 1900. Poznan je postal šele po letu 1970, ko sta George Box in Gwilym Jen v publikaciji *Time Series Analysis: Forecasting and Control* razvila celoten pristop, ki združuje ustrezne informacije, potrebne za razumevanje in uporabo tega modela. [10] Nekateri ARIMA model tako predstavljajo tudi kot Box-Jenkins model. ARIMA model uporablja ARIMA metodologijo. Poudarek teh metod ni gradnja ene enačbe, ampak analiza verjetnosti ali stohastike, saj je filozofija ekonomskih časovnih vrst, da podatki govorijo sami zase. Za razliko od regresijskih modelov, v katerih se odvisna spremenljivka  $Y_t$  pojasnjuje z regresorji, ARIMA model časovno vrsto  $Y_t$  pojasnjuje s preteklostjo ali odlogi, kot so vrednosti  $Y_t$ , in s stohastičnimi napakami. Zaradi te lastnosti ARIMA model imenujejo tudi neteoretični model,<sup>2</sup> ker ne temelji na nobeni ekonomski teoriji, ampak na simultanih enačbah. [11]

ARIMA model je sestavljen iz treh parametrov  $p$ ,  $d$  in  $q$ , zato lahko model razdelimo na tri podkategorije:

1. Integriranost časovnih vrst  $d$  = stopnja integriranosti časovne vrste. Integriranost predstavlja stabilnost oziroma naredi podatke stacionarne in ergotske, zaradi česar je lažje napovedati prihodnje vrednosti.
2. Avtoregresivni model (AR):  $p$  = število lastnih odlogov. AR-komponenta povezuje sedanje vrednosti in predhodne vrednosti  $p$ . Vrednost  $p$  imenujemo odlog ali spomin procesa AR.
3. Drseča sredina (MA):  $q$  = število odlogov slučajnih napak. MA-proces povezuje sedanjo vrednost s predhodnimi napakami  $q$ . Vrednost  $q$  se imenuje red ali spomin procesa MA. [14]

Pred obravnavo celotnega ARIMA modela je treba obrazložiti nekaj ekonometričnih pogojev in predstaviti njegovo metodologijo.

---

<sup>1</sup>Ang. Autoregressive integrated moving average

<sup>2</sup>Ang. atheoretic models: model, ki ni povezan s teorijo.

## 3.1 Stacionarnost

Za uporabo modela časovnih vrst mora biti časovna vrsta stacionarna.

**Definicija 3.1.** Časovna vrsta  $Y_t$  je stacionarna, če sta prvi in drugi moment  $Y_t$  konstantni

1.  $E(y_t) = \mu_y$  za vsak  $t \in T$ ,
2.  $E((y_t - \mu_y)(y_{t-h} - \mu_y)) = \gamma_h$  za vsak  $t \in T$  in za vsa cela števila  $h$  tako da  $t - h \in T$ ,

v nasprotnem primeru je nestacionarna. [15]

Prvi pogoj govori o tem, da morajo vse vrednosti časovne vrste, če je časovna vrsta stacionarna, imeti konstantno pričakovano vrednost. Z drugimi besedami: vrednosti časovne vrste se nenehno gibajo okoli konstantnega povprečja, ki je neodvisen od časa in ne sledi nikakršnemu trendu. Drugi pogoj za stacionarnost pravi, da je tudi varianca neodvisna od časa. Če v drugi pogoj vstavimo  $h = 0$ , je varianca  $\sigma_y^2 = E((y_t - \mu_y)^2) = \gamma_0$  neodvisna od  $t$ . Še več, kovarianca  $E(y_t - \mu_y)(y_{t-h} - \mu_y) = \gamma_h$  je neodvisna od  $t$ , ampak samo od razdalje  $h$  med dvema vrednostma časovne vrste. [15]

Če časovna vrsta ni stacionarna, lahko preučujemo njeno obnašanje le za časovno obdobje, ki se obravnava. Za vsak niz podatkov časovne vrste je potrebna posamezna analiza in zato je ni mogoče posplošiti na druga časovna obdobja. Za namene napovedovanja taka nestacionarna časovna vrsta nima torej nobene praktične vrednosti. Med nestacionarne časovne vrste spada tudi časovna vrsta cen delnic. [11]

Stacionarnost časovne vrste lahko preverimo na več načinov:

1. Uporabljamo lahko grafično analizo. Če graf časovne vrste ne namiguje na neki trend, potem je časovna vrsta stacionarna, v nasprotnem primeru je nestacionarna. [11]
2. Avtokorelacija in avtokorelogram je najenostavnejši test stacionarnosti. Najprej moramo izračunati avtokorelacijsko funkcijo.

Avtokorelacijska funkcija (ACF)<sup>3</sup> časovne vrste  $Y_t$  odloga  $h$  je

$$\rho_h = \frac{\gamma_h}{\gamma_0} = \text{Corr}(Y_{t+h}, Y_t), \quad (3.1)$$

kjer je

$$\gamma_h = \text{Cov}(Y_{t+h}, Y_t), \quad (3.2)$$

$$\gamma_0 = \text{var}(Y_t). \quad (3.3)$$

---

<sup>3</sup>Ang. Autocorrelation Function

Ker imamo v praksi samo vzorec časovne vrste, moramo izračunati vzorčno avtokorelacijo. Izračunamo jo kot

$$\tilde{\rho}_h = \frac{\tilde{\gamma}_h}{\tilde{\gamma}_0}, \quad (3.4)$$

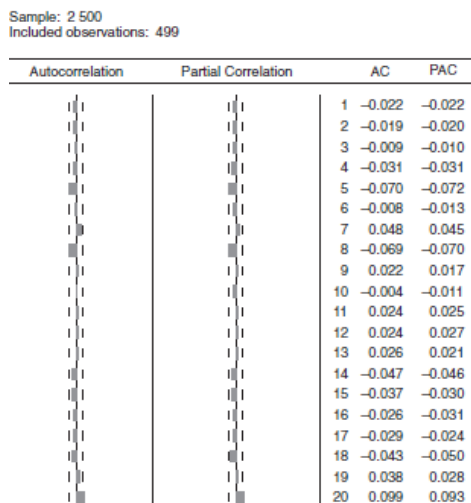
kjer

$$\tilde{\gamma}_h = \frac{1}{T} \sum_{t=h+1}^T (y_t - \bar{y})(y_{t-h} - \bar{y}), \quad (3.5)$$

$$\tilde{\gamma}_0 = \frac{1}{T} \sum_{t=1}^T (y_t - \bar{y})^2, \quad (3.6)$$

kjer je  $\bar{y} = \frac{1}{T} \sum_{t=1}^T y_t$  vzorčno povprečje. Če je časovna vrsta stacionarna, potem vzorčna avtokorelacija hitro pade z naraščajočim  $h$ . Graf avtokorelacije imenujemo korelogram.<sup>4</sup> [15] [11] Z drugimi besedami avtokorelacija preverja naključnost nabora podatkov. Če se avtokorelacija giblje okoli nič za vse odloge, potem je časovna vrsta naključna in s tem stacionarna. Časovna vrsta ni naključna in s tem ne stacionarna, če je ena ali več avtokorelacij precej različna od nič. Primer stacionarne časovne vrste je beli šum,<sup>5</sup> nestacionarna pa slučajni hod.<sup>6</sup>

Korelograma belega šuma in slučajnega hoda sta prikazana na slikah 1 in 2. [13]

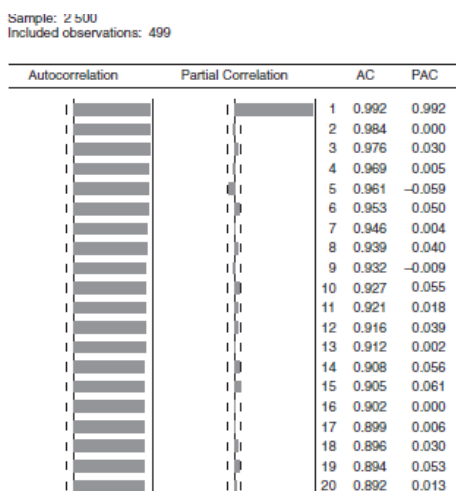


Slika 1: Korelogram belega šuma

<sup>4</sup>Ang. corelogram

<sup>5</sup>Ang. White noise

<sup>6</sup>Ang. random walk



Slika 2: Korelogram slučajnega sprehoda

Na slikah avtokorelacije sta poleg črte, ki predstavlja nič, tudi dve črtkani črti, ki predstavljata 95-odstotni interval zaupanja. Bartlett je pokazal, da če je časovna vrsta popolno naključna (kaže beli šum), je vzorec avtokovariacijskih koeficientov  $\rho_h$  porazdeljen približno normalno s pričakovano vrednostjo 0 in varianco ena deljeno z velikostjo vzorca.

$$\bar{\rho}_h \sim N(0, 1/n). \quad (3.7)$$

95-odstotni interval zaupanja je

$$\bar{\rho}_h \pm 1,96\sqrt{\frac{1}{n}} \quad (3.8)$$

oziroma

$$P(\bar{\rho}_h - 1,96\sqrt{\frac{1}{n}} \leq \rho_h \leq \bar{\rho}_h + 1,96\sqrt{\frac{1}{n}}) = 0,95. \quad (3.9)$$

Če 95-odstotni interval zaupanja vsebuje 0, sprejmemo ničelno hipotezo, da je pravi  $\rho_h$  enak 0 pri stopnji značilnosti 5 %. Če pa ne vsebuje 0, potem ničelno hipotezo, da je  $\rho_h$  enak 0, zavrnamo. [12] [11] Intervale zaupanja lahko preverimo za vse odloge posebej, ampak ker je to zamudno, obstajata tudi Box-Pierce Q statistika in Ljung-Box (LB) statistika. Obe testirata skupno hipotezo, da je avtokorelacijski koeficient enak 0. Q statistika je definirana kot

$$Q = n \sum_{h=1}^m \bar{\rho}_h^2, \quad (3.10)$$

kjer je  $n$  velikost vzorca in  $m$  število odlogov. Q-statistika se po navadi uporablja za testiranje belega šuma v časovni vrsti. Ljung-Box statistika se je izkazala za

boljšo, ker je statistično močnejša. LB-statistika je definirana kot

$$LB = n(n+2) \sum_{h=1}^m \frac{\bar{\rho}_h^2}{n-h}. \quad (3.11)$$

Obe statistiki, Q in LB, sta porazdeljeni s hi-kvadrat porazdelitvijo. Ničelno hipotezo, da so vsi avtokorelacijski koeficienti, zavrnamo, ko

$$Q > \chi_{1-\alpha, m}^2 \quad (3.12)$$

oziroma

$$LB > \chi_{1-\alpha, m}^2 \quad (3.13)$$

pri stopnji značilnosti  $\alpha$  in z  $m$  stopnjami prostosti. [16] [11]

3. Enotska rešitev: test enotske rešitve je v zadnjih letih postal eden bolj priljubljenih testov stacionarnosti. Najenostavnejši primer testa enotske rešitve je avtoregresijski model prvega reda in je oblike

$$Y_t = \alpha_1 Y_{t-1} + \epsilon_t, \quad (3.14)$$

kjer je  $\alpha_1$  neka konstanta in  $\epsilon \sim N(0, 1)$ . Enačbo preuredimo in vpeljemo operatorje odloga (L), za katerega velja

$$L^0 y_t = y_t,$$

$$L y_t = y_{t-1},$$

$$L^j y_t = y_{t-j},$$

$$L^{-j} y_t = y_{t+j}.$$

Če enačbo (3.14) pretvorimo v enačbo z operatorjem odloga, dobimo

$$y_t - \alpha_1 y_{t-1} = \epsilon_t, \quad (3.15)$$

$$(\alpha_0 L^0 - \alpha_1 L^1) y_t = \epsilon_t. \quad (3.16)$$

Enačbo odlogov lahko zapišemo kot funkcijo polinomov

$$\alpha(L) y_t = \epsilon_t. \quad (3.17)$$

Ideja enotske rešitve je preveriti, ali je časovna vrsta stacionarna ali ne. Vemo, da je beli šum stacionarna časovna vrsta. Zapišemo ga kot

$$y_t = y_{t-1} + \epsilon_t. \quad (3.18)$$

Polinom odlogov mora torej biti enak 0. Tako iščemo ničlo polinoma

$$\alpha(L) = 0. \quad (3.19)$$

Vemo, da če je  $\alpha_1 = 1$ , dobimo iz enačbe (3.14) slučajni hod, ki ni stacionaren. Rešitev funkcije odlogov mora torej biti

$$|\alpha_1| > 1. \quad (3.20)$$

Če želimo izvedeti, koliko je  $\alpha_1$ , jo moramo oceniti iz časovne vrste. Če je  $\bar{\alpha}_1 = 1$ , pomeni, da je časovna serija nestacionarna, in obratno.

Postavimo hipotezo, da je  $\alpha(1) = 0$ . Za testiranje te hipoteze proti alternativni za stacionarnost je potrebna reparametralizacija modela. Dodati moramo  $y_{t-1}$  na obeh straneh in preurediti enačbo v regresijo

$$\Delta y_t = \phi y_{t-1} + \sum_{j=1}^{p-1} \alpha_j^* \Delta y_{t-j} + \epsilon_t, \quad (3.21)$$

kjer je  $\phi = -\alpha(1)$  in  $\alpha_j^* = -(\alpha_{j+1} + \dots + \alpha_p)$ . V tem modelu želimo torej testirati hipotezi

$$H_0 : \phi = 0,$$

$$H_1 : \phi < 0.$$

Testiranje te hipoteze imenujemo Dickey-Fuller (ADF) test, ki temelji na t-statistiki koeficienta  $\phi$ , ki je OLS cenilka enačbe (3.21). [15] [11]

### 3.1.1 Transformiranje podatkov

Veliko ekonomskih časovnih vrst se ne sklada s stacionarnostjo časovne vrste. Tako lahko časovno vrsto transformiramo tudi z namenom, da jo naredimo čim bolj stacionarno. Logaritemska transformacija nam lahko na primer pomaga stabilizirati varianco. Pri logaritemski transformaciji nam ostane še vedno trend, ki se ga lahko znebimo s prvo diferenco  $\Delta \log y_t = \log y_t - \log y_{t-1}$  in tako dobimo stopnjo spreminjanja časovne vrste. Če časovna vrsta sledi nekemu trendu in je njena varianca sorazmerna časovni vrsti, jo logaritemska transformacija in prva diferenca transformirata, da postane stacionarna, ki pa je pogoj za izpeljavo kakršnegakoli modela. [15]



## 3.2 Splošna definicija ARIMA modela

Splošna oblika ARIMA(p, d, q) modela je naslednja:

$$\alpha(L)\Delta^d y_t = m(L)\epsilon_t, \quad (3.22)$$

kjer je

$$\Delta^d y_t \quad (3.23)$$

komponenta integriranost časovne vrste odloga  $d$ ,

$$\alpha(L) = 1 - \alpha_1 L - \dots - \alpha_p L^p \quad (3.24)$$

komponenta avtoregresijskega modela odloga  $p$  in

$$m(L) = 1 + m_1 L + \dots + m_q L^q \quad (3.25)$$

komponenta premikajočega povprečja odloga  $q$ .

Vse navedene enačbe so podrobneje opredeljene v naslednjih podpoglavjih. [15]

### 3.2.1 Integriranost časovnih vrst

Večina časovnih vrst je v praksi nestacionarna, zato ARIMA model za odpravo nestacionarnosti uporablja integriranost časovnih vrst. Že prva diferenca ( $I(1)$ ) lahko nestacionarno časovno vrsto transformira v stacionarno. Če je časovna vrsta integrirana stopnje  $d$  ( $I(d)$ ) lahko torej pomeni, da smo uporabili prvo diferenco  $d$ -krat.

Diferenco prve stopnje lahko torej zapišemo kot

$$\Delta y_t = y_t - y_{t-1}, \quad (3.26)$$

diferenco stopnje  $d$  ( $I(d)$ ) pa

$$y_t \sim I(d). \quad (3.27)$$

Če je  $\Delta^d y_t$  stacionarna, potem je  $\Delta^{d-1} y_t$  nestacionarna. Če je  $I(d)$  časovna vrsta z  $d \geq 1$ , se imenuje enotska rešitev.<sup>7</sup> Za določitev stopnje  $d$  se v praksi uporabljajo enotsko rešitveni testi.<sup>8</sup> [15] [11]

---

<sup>7</sup>Ang. Unit root

<sup>8</sup>Ang. Unit root tests

### 3.2.2 Avtoregresijski model (AR)

Avtoregresijski model (AR)<sup>9</sup>  $y_t$  odloga  $p$  (AR(p)) lahko zapišemo kot

$$y_t = \alpha_1 y_{t-1} + \dots + \alpha_p y_{t-p} + \epsilon_t, \quad (3.28)$$

kjer je  $\epsilon_t \sim N(0, 1)$  in  $\alpha_t$  fiksni koeficienti. Z uporabo odlogov lahko zapišemo enačbo (3.28) kot

$$(1 - \alpha_1 L - \dots - \alpha_p L^p) y_t = \epsilon_t \quad (3.29)$$

ali

$$\alpha(L) y_t = \epsilon_t \quad (3.30)$$

z  $\alpha(L) = 1 - \alpha_1 L - \dots - \alpha_p L^p$ . [15] Avtoregresijski model je torej regresijski model, ki povezuje časovno vrsto z njegovo preteklostjo. [12] Lahko tudi rečemo, da podatki govorijo sami zase. [11]

### 3.2.3 Drseča sredina ali premikajoče povprečje (MA)

Premikajoče povprečje (MA)<sup>10</sup>  $y_t$  lahko zapišemo kot

$$y_t = \epsilon_t + m_1 \epsilon_{t-1} + \dots + m_q \epsilon_{t-q}, \quad (3.31)$$

kjer je  $\epsilon_t \sim N(0, 1)$  in  $m_t$  konstanta. V funkciji odlogov zapišemo enačbo (3.31) kot

$$y_t = (1 + m_1 L + \dots + m_q L^q) \epsilon_t \quad (3.32)$$

ali

$$y_t = m(L) \epsilon_t \quad (3.33)$$

z  $m(L) = 1 + m_1 L + \dots + m_q L^q$ . [15] Premikajoče povprečje je torej enostavna linearna kombinacija napak belega šuma.

### 3.2.4 ARMA model

ARMA model (ARMA(p, q)) je združenje modela AR z odlogom  $p$  in modela MA z odlogom  $q$ . ARMA model zapišemo kot

$$y_t = \alpha_1 y_{t-1} + \dots + \alpha_p y_{t-p} + \epsilon_t + m_1 \epsilon_{t-1} + \dots + m_q \epsilon_{t-q}, \quad (3.34)$$

kjer smo definirali postavke, definirane v predhodnih poglavjih. V funkciji odlogov lahko ARMA model zapišemo:

$$\alpha(L) y_t = m(L) \epsilon_t. \quad (3.35)$$

---

<sup>9</sup>Ang. Autoregressive proces

<sup>10</sup>Ang. Moving Average

### 3.3 ARIMA metodologija

Interpretacija ARIMA modela so specifikacije, ki so identične multivariantni regresijski analizi. Vendar obstaja več dodatnih specifikacij, značilnih samo za ARIMA model. Za identifikacijo ARIMA modela sta Akaike Information Criterion (AIC) in Schwarz Criterion (SC) parametra. Uporabljata se za določitev parametrov  $p$ ,  $d$  in  $q$ , ki so najbolj statistično značilni. [13]

SC nalaga večjo težo na ocenjene koeficiente kot AIC. V splošnem je model z najbližjo vrednosto SC in AIC tudi najprimernejši. Dodatni niz rezultatov, ki jih je treba omeniti pri ARIMA modelu, sta avtokorelacija in delna avtokorelacija. Če je avtokorelacija  $AC(1)$  večja od nič, pomeni, da je časovna vrsta prvega odloga zelo korelirana med seboj. Če  $AC$  približno geometrijsko upada, s povečanjem odklonov, to implicira, da časovna vrsta sledi nizkemu odlogu premikajoče sredine (MA). [13]

Če se vzorec  $AC$  ujema z avtoregresijo odloga manj kot  $k$ , potem je PAC na odlogu blizu 0.

Ljung-Box Q statistika testira vrsto ničelnih hipotez, da so vse avtokorelacije do vključno danega odloga enake nič. Tako je poleg Q-statistike izračunana tudi p-vrednost, za katere odloge zavrnilo ničelno hipotezo. Če so p-vrednosti zelo majhne oziroma manjše od 0,05, potem zavrnilo ničelno hipotezo in sklepamo, da obstaja avtokorelacija med spremenljivkami. [13]

Da bi našli pravi ARIMA model, je potrebno veliko izkušenj in prakse. V analizi bom za ocenitev ARIMA parametrov uporabila osnovna statistična orodja, in sicer AC, PAC, SC, Q-statistiko in AIC. [13]

Postopek iskanje pravega ARIMA modela se imenuje ARIMA metodologija oziroma Box-Jenkins metodologija. ARIMA metodologija vsebuje štiri korake. To so: [11]

1. Identifikacija modela: poiskati je treba primerne parametre  $p$ ,  $d$  in  $q$ . Pri tem nam pomagata korelogram in delni korelogram.
2. Ocena parametrov: to ima večina programov v programskih paketih.
3. Ovrednotev modela: ocenjene koeficiente ocenjenega ARIMA modela je potrebno preveriti na dejanskih podatkih, ali se ujemajo. Z danimi testi je treba preveriti, ali je dani model beli šum. Če je, potem je ocenjeni ARIMA model pravi, če ni, pa je treba začeti znova.
4. Napovedovanje.

### 3.3.1 Identifikacija modela

V identifikaciji bomo poiskali primerne vrednosti  $p$ ,  $d$  in  $q$ , kjer  $d$  predstavlja stopnjo difference,  $p$  število odlogov v AR-modelu in  $q$  število odlogov MA-modela. Te parametre bomo določili s pomočjo avtokorelacije (AC) in delne avtokorelacije (PAC).

Delni avtokorelacijski koeficient (PAC)<sup>11</sup> nam prav tako pove, kolikšna je avtokorelacija med členi v časovni vrsti glede na odlog  $h$ , le da tokrat iz vrednosti avtokorelacijskega koeficienta izločimo vpliv vmesnih členov. Izračun delnih avtokorelacijskih koeficientov je zapleten in ga ne bom obravnavala. V poglavju avtokorelacije sem definirala  $\bar{\sigma}_h$ , tako je treba definirati tudi delno avtokorelacijo  $z \bar{\sigma}_{hh}$ . [11]

Prvi korak identifikacije modela je preveriti stacionarnost, to storimo s testi stacionarnosti, opisanimi v poglavju Stacionarnost. Če časovna vrsta ni stacionarna, jo transformiramo. Pri transformaciji lahko uporabimo tudi diferenco časovne vrste. Po navadi je dovolj že prva diferenca. Če je časovna vrsta stacionarna po prvi diferenci, je parameter  $d$  ARIMA modela enak 1. Za določitev parametrov  $p$  in  $q$  uporabimo izris avtokorelacije (AC) in delne avtokorelacije (PAC). Opazimo, da imata AR( $p$ ) in MA( $q$ ) nasproten vzorec AC in PAC: v primeru AR( $p$ ) AC pada geometrijsko ali eksponentno in PAC odsekoma pade na določenem odlogu, medtem pa se z MA( $q$ ) dogaja ravno nasprotno. Teoretične vzorce AC in PAC lahko predstavimo v tabeli.

Tabela 1: Teoretični vzorci AC-grafa

Oblika AC	Navedeni model
Eksponentno pada proti nič.	AR( $p$ ); kjer PAC pade na odlogu $p$ .
Izmenično je negativen in pozitiven ter pada proti nič.	AR( $p$ ); kjer PAC pade na odlogu $p$ .
Ena ali več črt, ostale so približno nič.	MA( $q$ ); kjer je $q$ odlog kjer AC postane nič.
Pade po nekaj začetnih odlogov.	ARMA ( $p$ , $q$ ).
Vse vrednosti so enake ali približno nič.	Vrednosti so naključne.
Ne pada proti nič.	Časovna vrsta ni stacionarna.

S pomočjo te tabele ter odlogov Akaike Information Criterion (AIC) in Schwarz Criterion (SC) ocenimo parametra  $p$  in  $q$ . [11] [16]

### 3.3.2 Ocenitev ARIMA modela

V prejšnjem koraku smo določili parametre  $d$ ,  $p$  in  $q$ . Naslednji korak je oceniti koeficiente modela na dejanskih podatkih. Ocenjevanje koeficientov ARIMA modela je precej kompliciran nelinearen ocenjevalni proces. Zato po navadi to ocenjevanje prepustimo raznim programom, kot je STATA, da nam izračuna ocenjeni model. Programi običajno uporabljajo oceno največjega verjetja.<sup>12</sup> [11] [16]

<sup>11</sup>Ang. Partial autocorrelation function

<sup>12</sup>Ang. Maximum likelihood estimation

### 3.3.3 Ovrednotev modela

Ko imamo ocenjen celoten ARIMA model, se moramo prepričati, ali se ocenjeni model prilagaja dejanskim podatkom. Preprosto orodje preverjanja prilagojenosti je, da preverimo, ali ostanki sledijo belemu šumu. To preverimo tako, da izrišemo AC in PAC ostankov. Če se avtokorelacijski koeficienti gibljejo v 95-odstotni intervalu zaupanja oziroma so približno nič, lahko rečemo, da ostanki sledijo belemu šumu. V nasprotnem primeru pomeni, da ocenjeni ARIMA model ni pravi in ga moramo ponovno identificirati. [11] [16]

### 3.3.4 Napovedovanje

V ocenjeni model vstavimo dejanske vrednosti in izračunamo prihodnje. To za nas opravijo različni programi. [11]

## 4 Simulacija

V poglavju Simulacija bom analizirala časovno vrsto cen delnice The Coca-Cola Company (KO), uporabila časovno vrsto za ocenitev Brownovega in ARIMA modela, napovedala prihodnje cene ter med seboj primerjala modela in dejanske cene.

To delnico sem izbrala zaradi likvidnosti. Likvidnost oziroma likvidno poslovanje pomeni sposobnost podjetja poravnati vse obveznosti ob zapadlosti. Če je podjetje likvidno, potem so cene delnice podjetja stabilne in ni nepričakovanih padcev ali vzponov cene delnic. Likvidnost podjetja smo preverili s P/E kazalnikom, ki pove, koliko evrov so vlagatelji pripravljeni plačati za evro dobička družbe oz. za kolikokrat tržna cena delnice presega zadnji letni dobiček na delnico družbe. Povprečni P/E kazalnik se giblje med 20–25. Izbrano podjetje (The Coca-Cola Company /KO/) ima vrednost P/E kazalnika 22,24, kar je ravno povprečje celotnega trga. Naslednji kazalnik izbire delnice je beta. Beta je mera volatilnosti oziroma sistematičnega tveganja posamezne naložbe, ki pove, koliko je gibanje donosnosti posamezne naložbe (ali portfelja) povezano s gibanjem trga. Beta vrednost 1, pomeni, da se tečaj delnice giblje skladno s gibanjem trga. Če je beta manjša od 1 pomeni, da je delnica manj volatilna od trga. Beta, večja od 1, pa nakazuje, da je delnica bolj volatilna od trga. Beta vrednost naše delnice je 1,03, kar pomeni, da je teoretično za 3 % bolj volatilna od trga, kar pa ni veliko. [17] [18]

Za izbrano podjetje The Coca-Cola Company (KO) sem pretekle cene delnice zbrala s spletne strani Yahoo Finance (2015). Opazovano obdobje odsega 149 delovnih dni, od 25. 11. 2014 do 31. 7. 2015. Za ceno sem izbrala prilagojeni zaključni tečaj.<sup>1</sup> Prilagojeni zaključni tečaj uporablja zaključni tečaj<sup>2</sup> kot izkodišče, ki ga popravi z upoštevanjem dejavnikov, kot so dividende, zaloge<sup>3</sup> in vmesne nove ponudbe delnic. Prilagojena cena zapiranja predstavlja natančnejši odsev vrednosti cene in je primernejša za nadaljnjo obravnavo. [18] [17]

V naslednjem poglavju bom cene delnice statistično analizirala in grafično prikazala.

---

<sup>1</sup>Ang. Adjusted Close Price

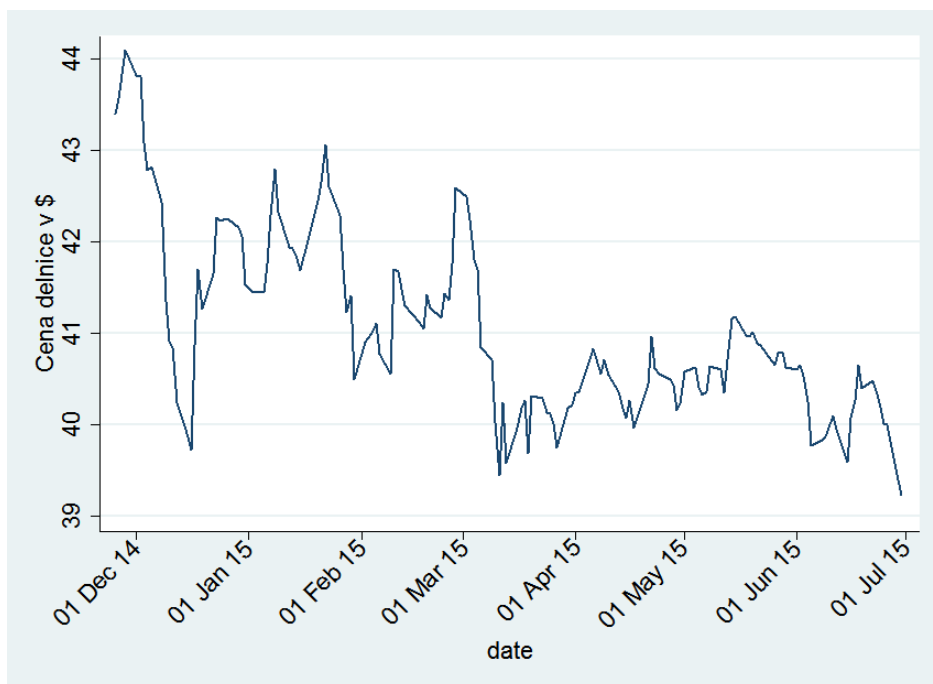
<sup>2</sup>Ang. Closing price

<sup>3</sup>Ang. Stock split

## 4.1 Analiza časovne vrste

Opisna statistika je skupina statističnih metod, ki se ukvarja s povzemanjem pridobljenih podatkov. Te metode iščejo opisne podatke o populaciji iz njenih sestavin, da bi ustvarile pregleden opis. V ta namen opisna statistika uporablja grafikone, tabele in statistične povzetke.

Grafični prikaz gibanja cene delnice The Coca-Cola Company (KO):



Slika 3: Graf gibanja cen delnice The Coca-Cola Company (KO) od 25. 11. 2014 do 31. 6. 2015.

Tabela 2: Osnovne statistične lastnosti časovne vrste cen delnice

Statistično orodje	Vrednost cene
Število observacij	149
Povprečje	41,01
Maksimalna vrednost	44,1
Minimalna vrednost	39,23
Standardni odklon	0,995
Varianca	0,990
Mediana	40,77

Iz tabele (4.1) lahko vidimo, da cene delnice niso volatilne ter se povprečje in mediana približno ujemata. Med maksimalno in minimalno ceno je razlike 4,87 \$.

Maksimalno vrednost je delnica dosegla 28. 11. 2014, minimalno pa 30. 7. 2015. Vse naštetost se sklada s prej predstavljenim grafom.

V naslednjih dveh poglavjih bom časovno vrsto cen delnic podjetja The Coca-Cola Company uporabila za napovedovanje prihodnjih vrednosti cen delnic z Brownovim in ARIMA modelom.



## 4.2 Napovedovanje cen delnice z Brownovim gibanjem

V tem poglavju bom izračunala napovedane cene izbrane delnice z Brownovim gibanjem. Za izračun cen bom uporabila program Excel. V poglavju 2.19 sem definirala model

$$S(1) = S(0) + \mu S(0)\Delta t + \sigma S(0)\epsilon\sqrt{\Delta t}. \quad (4.1)$$

Za lažji izračun napovedanih cen v Excelu bom enačbo (4.1) razdelila na več delov. Naj bo

$$V_1 = \epsilon_t, \quad (4.2)$$

kjer je  $\epsilon_t \sim N(0, 1)$ . In naj bo

$$V_2 = \frac{\Delta S}{S} \quad (4.3)$$

oziroma

$$V_2 = \mu\Delta t + \sigma\sqrt{\Delta t}V_1, \quad (4.4)$$

kjer je  $\mu$  donos in  $\sigma$  varianca.  $\Delta t$  izračunamo kot  $\Delta t = 1/n$ , kjer je  $n$  število observacij. Če v enačbo (4.1) vstavimo novi spremenljivki  $V_1$  in  $V_2$ , dobimo

$$S(1) = S(0)(1 + V_2). \quad (4.5)$$

Zdaj vse zgoraj navedene enačbe definirajmo v programu Excel.

Postopek izračunavanja prihodnjih vrednosti v programu Excel:

1. korak. V program Excel vstavimo tabelo vseh cen delnic z datumi in izračunamo donos delnic za vsak dan posebej po formuli, ki je prikazana na sliki.

	A	B	C	D	E	F	G
1	Date	P	Donos				
2	25.11.2014	43,4038					
3	26.11.2014	43,5661	0,003739		=(B3-B2)/B2		
4	28.11.2014	44,09727	0,012192				
5	1.12.2014	43,82185	-0,00625				
6	2.12.2014	43,81201	-0,00022				

Slika 4: Izračun donosa posameznih cen v programu Excel

Ko izračunamo vse donose, izračunamo povprečni donos na delnico in varianco.

	A	B	C	D	E	F	G
151							
152							
153							
154	DONOS	-0,00064			=AVERAGE(C3:C151)		
155	VOLATIVNOST	0,008933			=STDEV(C3:C151)		
156	N	149			=COUNT(B2:B151)		
157							

Slika 5: Izračun povprečnega donosa vseh donosov ter varianco v programu Excel

2. Povprečni donos je letni donos in ga je treba pretvoriti v dnevne. To storimo tako, da ga pomnožimo z  $\Delta t$ , ki smo ga v programu Excel označili s T. T izračunamo kot  $\frac{1}{(\text{število observacij})}$ . Volatilnost na letni ravni izračunamo kot  $(\text{letni donos})/\sqrt{T}$ .

	A	B	C	D	E	F	G
1		Letna osnova	Dnevna osnova				
2	Donos	-0,0006	0,0000		=B2*B5		
3	Volatilnost	0,0089	0,0007		=B3*SQRT(B5)		
4	N	149					
5	T	0,0067			=1/B4		
6	S(0)	39,23			zadnja cena delnice		
7							

Slika 6: Izračun letnega donosa in variance v programu Excel

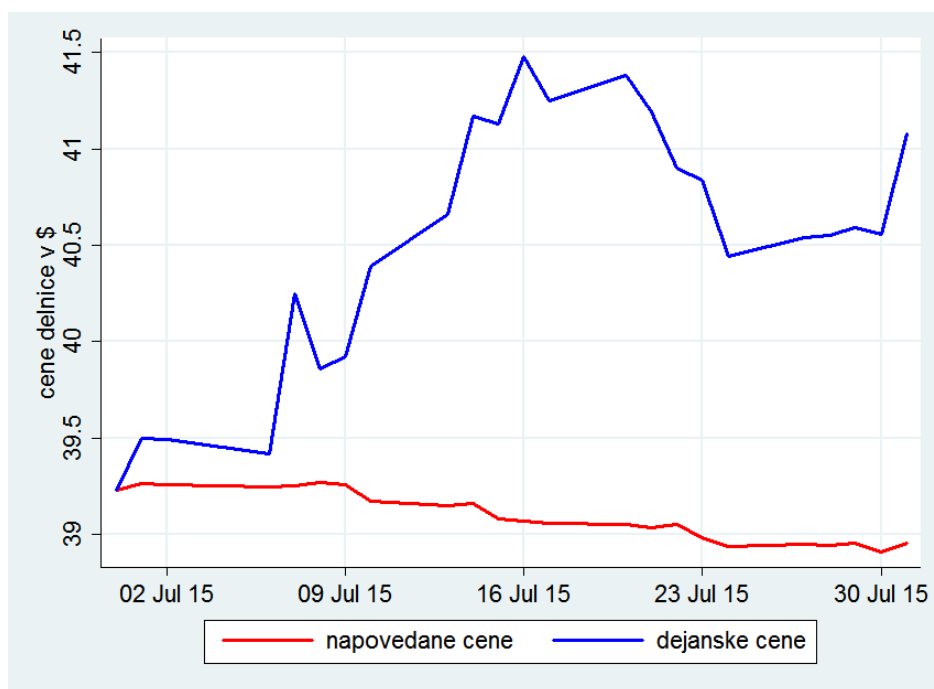
3. Še zadnja stvar, preden dobimo napovedano ceno delnice, je definiranje  $V_1$ ,  $V_2$  in izračun.

	A	B	C	D	E	F
1		Letna osnova	Dnevna osnova			
2	Donos	-0,0006	0			
3	Volatilnost	0,0089	0,0007			
4	N	149				
5	T	0,0067				
6	S(0)	39,23				
7						
8	Napoved					
9	t	1.8.2015				
10	v1	1,0835		=NORMSINV(RAND())		
11	v2	0,0008		=\$C\$2+\$C\$3*B10		
12	S	39,2609		=B6+B11*B6		

Slika 7: Izračun  $V_1$ ,  $V_2$  in napovedane cene ( $S$ ) v programu Excel

Tako naredimo za vse ostale napovedane cene.

Ko smo izračunali napovedane cene, jih lahko prikažemo grafično.



Slika 8: Graf napovedanih cen delnice z Brownovim gibanjem in dejanske cene

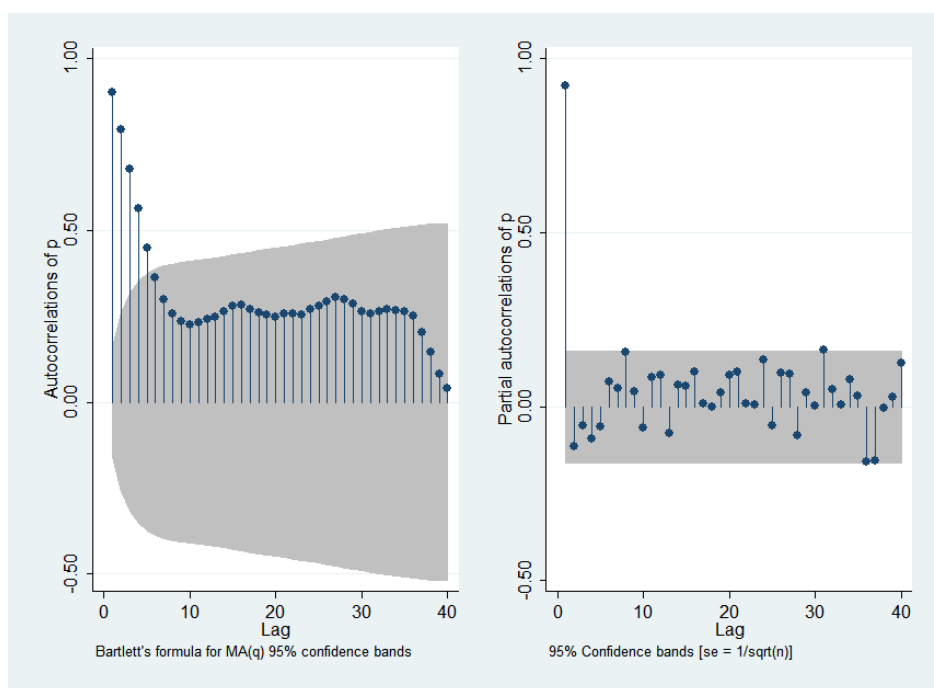
Iz grafa je razvidno, da so kar velika odstopanja med napovedanimi cenami in dejanskimi. Razlika med njima se vedno bolj veča. Brownovo gibanje je napovedalo negativni trend, ker cene padajo, a dejanske so se povzpele.

## 4.3 ARIMA model

ARIMA model bom simulirala s pomočjo programa STATA, pri čemer bom sledila ARIMA metodologiji.

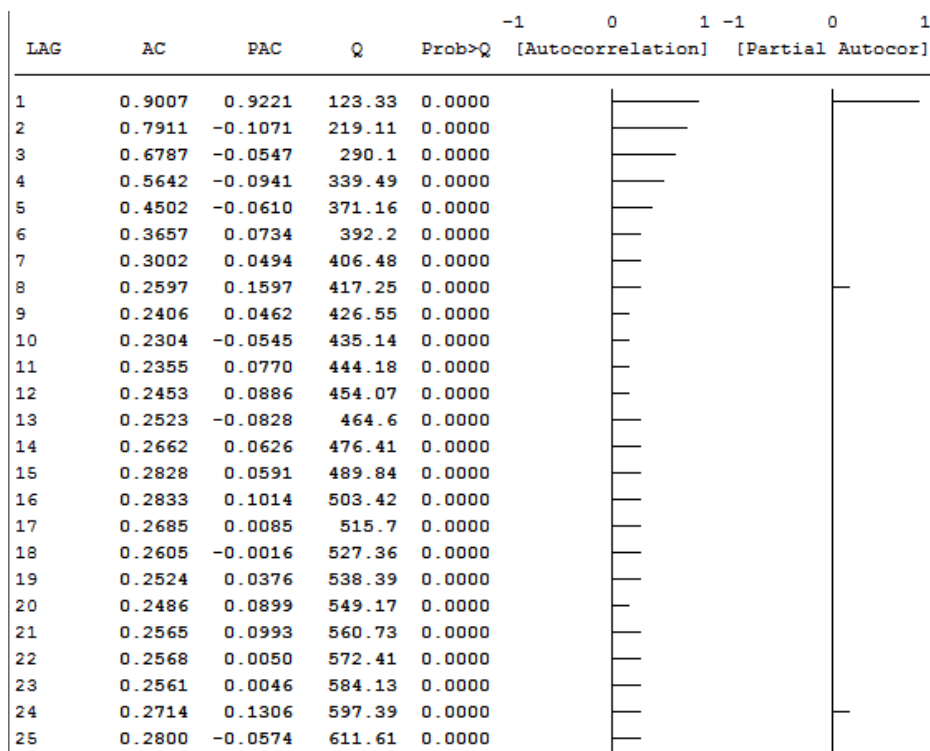
### 1. Identifikacija modela:

Predno začnemo identificirati parametre, je treba preveriti stacionarnost podane časovne vrste. Kot vidimo že s slike v poglavju Analiza časovne vrste, časovna vrsta cene delnice ni stacionarna in nakazuje na negativni trend. To potrjujeta tudi avtokorelacijska in delna avtokorelacijska funkcija.



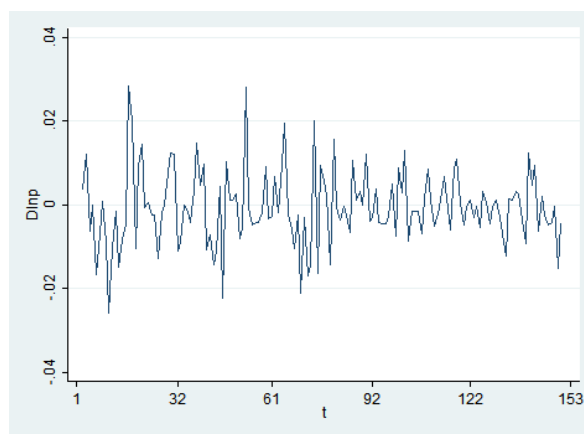
Slika 9: Avtokorelacijska in delno avtokorelacijska funkcija

Prvi pristop k odpravi nestacionarnosti je transformiranje podatkov. Najprej bom podatke logaritmirala, nato preverila AC in PAC.



Slika 10: Korelogram logaritmiranih cen

Iz korelograma vidimo, da logaritemska časovna vrsta cen ni stacionarna. Naslednji korak do stacionarnosti je prva diferenca.



Slika 11: Graf prve difference logaritemske časovne vrste cen

Z zgornjega grafa ne moremo videti, ali je časovna vrsta stacionarna, zato izrišemo korelogram.

LAG	AC	PAC	Q	Prob>Q	-1	0	1	-1	0	1
					[Autocorrelation]	[Partial Autocor]				
1	0.0678	0.0679	.6951	0.4044						
2	0.0117	0.0081	.716	0.6991						
3	0.0435	0.0431	1.0051	0.8000						
4	0.0073	0.0018	1.0133	0.9078						
5	-0.1195	-0.1241	3.2299	0.6646						
6	-0.0996	-0.0908	4.7797	0.5724						
7	-0.1970	-0.1953	10.893	0.1434						
8	-0.0913	-0.0690	12.215	0.1419						
9	0.0226	0.0366	12.297	0.1971						
10	-0.0801	-0.0934	13.327	0.2059						
11	-0.0746	-0.1037	14.229	0.2206						
12	0.1280	0.0729	16.903	0.1533						
13	0.0030	-0.0704	16.905	0.2037						
14	-0.0070	-0.0646	16.913	0.2609						
15	-0.0352	-0.1082	17.12	0.3118						
16	0.0077	-0.0166	17.13	0.3773						
17	-0.0082	-0.0047	17.141	0.4449						
18	-0.0017	-0.0432	17.141	0.5134						
19	-0.0793	-0.0952	18.222	0.5076						
20	-0.0740	-0.1033	19.172	0.5107						
21	0.0447	-0.0090	19.521	0.5518						
22	0.0184	-0.0089	19.58	0.6093						
23	-0.0924	-0.1339	21.096	0.5752						
24	0.1076	0.0561	23.17	0.5098						
25	-0.0189	-0.1006	23.234	0.5639						

Slika 12: Korelogram logaritmiranih cen

Iz korelograma ne moremo razbrati, ali je časovna vrsta prve diference stacionarna, zato uporabimo Dickey-Fuller test stacionarnosti, ki testira ničelno hipotezo o stacionarnosti. Rezultati Dickey-Fullerjevega testa v STATA so prikazani na spodnji sliki.

```
. dfgls Dlnp

DF-GLS for Dlnp                                Number of obs = 134
Maxlag = 13 chosen by Schwert criterion

      DF-GLS tau      1% Critical      5% Critical      10% Critical
      [lags] Test Statistic Value Value Value
-----+-----+-----+-----+-----
      13      -4.895      -3.522      -2.783      -2.508
      12      -5.018      -3.522      -2.803      -2.526
      11      -5.116      -3.522      -2.821      -2.544
      10      -5.683      -3.522      -2.840      -2.561
      9       -5.737      -3.522      -2.858      -2.578
      8       -5.815      -3.522      -2.875      -2.594
      7       -6.532      -3.522      -2.891      -2.609
      6       -6.927      -3.522      -2.907      -2.623
      5       -6.463      -3.522      -2.922      -2.637
      4       -6.531      -3.522      -2.936      -2.650
      3       -6.297      -3.522      -2.949      -2.662
      2       -6.892      -3.522      -2.962      -2.673
      1       -8.380      -3.522      -2.973      -2.683

Opt Lag (Ng-Perron seq t) = 6 with RMSE .0082889
Min SC = -9.43157 at lag 1 with RMSE .0086315
Min MAIC = -7.347063 at lag 1 with RMSE .0086315
```

Slika 13: Izpis Dickey-Fullerjevega testa v STATA

STATA nam poleg Dickey-Fullerjevega testa predlaga tudi odloge Ng–Perron, Schwarz information criterion (SIC) in Akaike information criterion (MAIC). Hipotezo, da časovna vrsta ni stacionarna, zavrnilo, ko je DF-GLS test manjši od podanih kritičnih vrednosti. Iz prejšnje tabele lahko zavrnilo vse ničelne domneve, da časovna serija ni stacionarna. Test torej kaže, da je stacionarna pri 1 % kritične vrednosti.

Predlagane odloge, ki nam podajo Ng–Perron, SIC in MAIC (6, 1, 1), bomo uporabili za ocenitev ARIMA modela. Naš ARIMA model je lahko oblike ARIMA (1, 1, 0), ARIMA (6, 1, 0), ARIMA (0, 1, 1), ARIMA (0, 1, 6), ARIMA (1, 1, 1), ARIMA (1, 1, 6), ARIMA (6, 1, 1), ARIMA (6, 1, 6). Koeficienti ARIMA modelov ARIMA (1, 1, 1), ARIMA (0, 1, 6), ARIMA (0, 1, 1), ARIMA (1, 1, 0) ARIMA (6, 1, 0) so statistično neznačilni in se vsi gibljejo okoli nič. Koeficienti ARIMA modelov ARIMA (6,1,1), ARIMA (1, 1, 6) in ARIMA (6, 1, 6) pa imajo statistično značilne prve odloge in so približno ena. Vsi naši rezultati vseh ARIMA modelov namigujejo na slučajni hod.

Poglejmo ARIMA (1, 1, 0), ki ga lahko zapišemo kot

$$\Delta \log S_t = \beta_1 \Delta \log S_{t-1} + \epsilon_t, \quad (4.6)$$

Ko je  $\beta = 0$ , dobimo enačbo slučajnega hoda:

$$\log S_t = \log S_{t-1}. \quad (4.7)$$

Slučajni hod je nestacionarna časovna vrsta. Naša časovna vrsta cen delnice je torej slučajni hod, zaradi česar ne moremo napovedati cen delnic. Slučajni hod govori o tem, da je današnja cena delnice enaka jutrišnji plus neko naključno gibanje. Če je v času  $t = 0$  vrednost delnice  $y_0$ , potem je cena  $y_1$  enaka

$$y_1 = y_0 + \epsilon_1, \quad (4.8)$$

$$y_2 = y_0 + \epsilon_1 + \epsilon_2. \quad (4.9)$$

V splošnem lahko zapišemo, da je cena delnice v času  $t$  enaka

$$y_t = y_0 + \sum \epsilon_t, \quad (4.10)$$

torej je pričakovana vrednost cene delnice

$$E(y_t) = y_0, \quad (4.11)$$

kar pa v praksi ne drži. Zaradi te značilnosti lahko rečemo, da ima slučajni hod neskončen spomin. [11]



## 4.4 Rezultati

Naša simulacija napovedovanja cen delnice The Coca-Cola Company je bila neuspešna. Oba modela sta podala nenatančne in nezaželjene rezultate. Napovedane cene z Brownovim gibanjem so kazale velika odstopanja od dejanskih cen. Zadnja napovedana cena z Brownovim gibanjem je 38,95 \$, dejanska cena pa je 41,08 \$, kar pomeni, da je bila napoved napačna za 5,2 % dejanske cene. Na delniškem trgu pomeni napaka 5 % velike izgube. Zaradi 5,2-odstotnega odstopanja cene delnice ne moremo trditi, da je Brownov model uspešno napovedal gibanje cen. Napovedovanje cen pri ARIMA modelu je bilo še bolj neuspešno, saj smo se srečali z naključnim hodom, ki je nestacionaren in s katerim ni mogoče napovedati cen.

## 5 Zaključek

V zaključni nalogi sem obravnavala dva modela za napovedovanje cen delnic. Predstavila sem Brownovo gibanje in ARIMA model. Gibanje cen je zelo nepredvidljivo, to sem dokazala tudi s simulacijo obeh modelov na primeru cen delnice The Coca-Cola Company. Sama simulacija ni bila uspešna, saj je pri Brownovem gibanju bila razlika med napovedanimi in dejanskimi cenami 5,2-odstotna, kar lahko pri trgovanju z delnicami pomeni velike izgube.

Pri ARIMA modelu pa se je zataknilo že pri identifikaciji samega modela. Pri ARIMA modelu sem prišla do slučajnega hoda, kar pa pomeni, da se cene delnice gibljejo popolnoma naključno in jih ne moremo napovedati. Da bi rešili problem, ki je nastal s slučajnim hodom, bi morali poiskati še kakšne druge transformacije podatkov, katere napoved bi izračunali veliko lažje.

Namen zaključne naloge je bil teoretično spoznati modela in ju na podlagi simulacije primerjati in oceniti na uspešnost napovedovanja. Prvi namen smo popolnoma izpolnili, drugega pa ne, saj nam ni uspelo izračunati napovedanih cen za oba modela. S simulacijo napovedovanja smo spoznali, da so cene delnic zelo nepredvidljive in jih je zelo težko napovedati.

## 6 Literatura in viri

- [1] C. PARÉS in C. VÁZQUEZ, *Advances in Numerical Simulation in Physics and Engineering*. Prentice-Hall, 1981.
- [2] FIMA C. KLEBANER, *Introduction to Stochastic Calculus*, Imperial College Press, 2005.
- [3] STEVEN R. DUNBAR, *Stochastic Processes and Advanced Mathematical Finance*, University of Nebraska-Lincoln, 2014.
- [4] JOHN C. HULL, *Options, Futures and other Derivatives*, Upper Saddle River, New Jersey, 2009.
- [5] ARGIMIRO ARRATIA, *Computational Finance: An Introductory Course with R*, Atlantis Press, Spain, 2014.
- [6] UBBO F WIERSEMA, *Brownian Motion Calculus*, John Wiley & Sons, 2008.
- [7] JAMES E. GENTLE, *Wolfgang Karl Härdle*, Modeling Asset Prices. Humboldt-Universität zu Berlin, 2010.
- [8] JAMES E. GENTLE, *Wolfgang Karl Härdle*, Study of effectiveness of Time Series Modeling (ARIMA) in Forecasting Stock Prices. International Journal of Computer Science, Engineering and Applications (IJCSEA) Vol. 4, No. 2, April 2014.
- [9] XIAOLIAN ZHENG, *Ben M. Chen*, Stock Market Modeling and Forecasting. Springer-Verlag, London, 2013.
- [10] JOHNATHAN MUN, *Advanced Analytical Models*, Wiley, 2008.
- [11] D. GUJARATI, *Basic Econometrics*, McGraw Hill, Boston, 2003.
- [12] J. H. STOCK & M. W. WATSON, *Introduction to Econometrics*, Addison Wesley, Boston, 2003.
- [13] CHARLES W. CHASE, JR, *Demand-Driven Forecasting: A structural Approach to Forecasting*, John Wiley & Sons, Canada, 2009.
- [14] J. MUN, *Advanced Analytical Models*, Wiley, 2008..

- [15] HANLMUT LÜTKERPOHL, *Markus Krätzig*, Applied Time Series Econometrics. Cambridge University Press, 2004.
- [16] Engineering Statistics Handbook NIST SEMATECH, 2003.
- [17] *The Coca-Cola Company*, <http://finance.yahoo.com/>. (Datum ogleda: 25. 8. 2015.)
- [18] *Investopedija*, <http://www.investopedia.com/>. (Datum ogleda: 25. 8. 2015.)