

VILNIAUS UNIVERSITETAS  
MATEMATIKOS IR INFORMATIKOS FAKULTETAS

Magistro baigiamasis darbas

Lietuvos ir pasaulio gyventojų skaičiaus  
kitimo prognozavimas

Forecasting population changes in Lithuania and the  
World

Rita Norbutaitė

VILNIUS 2020

**MATEMATIKOS IR INFORMATIKOS FAKULTETAS**  
**MATEMATIKOS INSTITUTAS**  
**MATEMATINĖS ANALIZĖS KATEDRA**

Darbo vadovas Asist. Dr. Aldona Skučaitė \_\_\_\_\_

Darbo recenzentas \_\_\_\_\_

Darbas apgintas \_\_\_\_\_

Darbas įvertintas \_\_\_\_\_

Registravimo Nr. \_\_\_\_\_

Atidavimo į katedrą data \_\_\_\_\_

# Lietuvos ir pasaulio gyventojų skaičiaus kitimo prognozavimas

## Santrauka

Darbe pristatomi eksponentinis ir logistinis populiacijos kitimo modeliai bei pritaikomi Lietuvos ir pasaulio populiacijoms. Naudojantis laiko eilučių teorija yra parenkama populiacijos kitimo norma ir apskaičiuojama jos prognozė. Remiantis Žemės resursų gausa ir žmonijos vartojimo įpročiais įvertinamas maksimalus galimas gyventojų skaičius. Pateikiama Lietuvos ir pasaulio populiacijų prognozė 80 metų į priekį.

**Raktiniai žodžiai:** populiacija, populiacijos modeliai, eksponentinis modelis, logistinis modelis, ARIMA, prognozė.

## Forecasting population changes in Lithuania and the World

### Abstract

In this work exponential and logistic population models are described and applied to populations of Lithuania and the World. Population growth rate is calculated and predicted using time series. Based on Earth's resources and human consumption, the maximum possible population is estimated. The projections of Lithuanian and World populations for 80 years are made.

**Key words:** population, population models, exponential model, logistic model, ARIMA, forecast.

# Turinys

<b>1</b>	<b>Įvadas</b>	<b>3</b>
<b>2</b>	<b>Populiacijų kitimo modeliai</b>	<b>4</b>
2.1	Pagrindiniai žymėjimai ir sąvokos . . . . .	4
2.2	EkspONENTINIS modelis . . . . .	4
2.3	Logistinis modelis . . . . .	6
<b>3</b>	<b>Populiacijos kitimo normos prognozavimas</b>	<b>9</b>
3.1	Grafikas . . . . .	9
3.2	Transformavimas . . . . .	9
3.3	Stacionarumas . . . . .	9
3.4	Autokoreliacinės ir dalinės autokoreliacinės funkcijos tyrimai . . . . .	11
3.5	Kriterijai modelio tinkamumui įvertinti . . . . .	12
3.6	Liekanų analizė . . . . .	12
3.7	Prognozė . . . . .	13
<b>4</b>	<b>Pasaulio populiacijos prognozė</b>	<b>14</b>
4.1	Bendra apžvalga . . . . .	14
4.2	Pasaulio metinės populiacijos kitimo normos prognozavimas . . . . .	16
4.3	Logistinio populiacijos kitimo modelio pritaikymas pasaulio populiacijai . . .	21
<b>5</b>	<b>Lietuvos populiacijos prognozė</b>	<b>22</b>
5.1	Bendra apžvalga . . . . .	22
5.2	Lietuvos metinės populiacijos kitimo normos prognozavimas . . . . .	23
5.3	Logistinio populiacijos kitimo modelio pritaikymas Lietuvos populiacijai . . .	27
<b>6</b>	<b>Rezultatai ir išvados</b>	<b>29</b>
<b>7</b>	<b>Literatūra</b>	<b>30</b>

# 1. Įvadas

2011 m. pasaulio populiacija peržengė 7 milijardų ribą ir šiuo metu pasaulyje gyvena daugiau negu 7,7 mlrd. žmonių. Per pastaruosius 50 metų žmonių skaičius padidėjo maždaug 2 kartus: nuo 3,626 mlrd. 1969 m. iki 7,713 mlrd. žmonių 2019 metais [16]. Pastaraisiais metais vis dar yra pastebimas pasaulio populiacijos didėjimas, nors nebe toks spartus [15]. Lietuvoje didžiausia populiacija užfiksuota 1992 m. - kiek daugiau negu 3,7 mln., tačiau pastaruosius 28 metus žmonių mažėja - 2019 m. pradžioje gyveno 2,8 mln. - ir vis dar išlieka mažėjimo tendencija [9].

Nors apskaičiuoti, kiek žmonių gyvena kiekvienu momentu, kai žmonės pasaulyje nuolat gimsta ir miršta, nėra paprastas uždavinys, tačiau daug įdomesnis ir sudėtingesnis klausimas yra kiek žmonių gyvens ateityje. Atsakymas į šią klausimą ne tik patenkina mūsų smalsumą, bet ir padeda spręsti socialines problemas, pvz., senatvės pensijų dydžio klausimą ar vaikų darželių kiekį. Žvelgiant giliau, populiacijos tyrimai padeda prognozuoti verslo rinkų dydį bei ateities pelnus. Šio darbo tikslas yra aprašyti pagrindinius populiacijos kitimo modelius ir palyginti eksponentinį ir logistinį populiacijos kitimo modelius, pateikti pasaulio ir Lietuvos gyventojų prognozes bei palyginti jas su esamomis Eurostat'o ir Jungtinių tautų organizacijos prognozėmis.

## 2. Populiacijų kitimo modeliai

Šiame skyriuje aprašysime pagrindinius modelius, jų privalumus ir trūkumus.

### 2.1. Pagrindiniai žymėjimai ir sąvokos

Populiacijos dydį laiko momentu  $t$  pažymėkime  $P_t$ ,  $t \geq 0$ . Dydis

$$\Delta P_{t,1} = P_{t+1} - P_t \quad (1)$$

yra vadinamas **absoliučiu** populiacijos pokyčiu per vieną laiko tarpą (dažniausiai metus) [7]. Santykis

$$r_{t,1} = \frac{\Delta P_{t,1}}{P_t}, \quad t \geq 0, P_t \neq 0 \quad \forall t,$$

yra vadinamas **santykiniu** populiacijos pokyčiu per vienerius metus arba **metine populiacijos kitimo norma** [7]. Populiacija, kurios metinė kitimo norma  $r_{t,1} = r$  yra konstanta, yra vadinama **stabilia**. Atskiru atveju, jei populiacijos dydis laikui bėgant nesikeičia, t.y.  $r = 0$ , populiacija vadinama **stacionaria** [14].

### 2.2. Eksponentinis modelis

Dar 1751 metais JAV politikas ir išradėjas Benjamin'as Franklin'as parašė publikaciją "Žmonių gausėjimo stebėjimai, šalių apgyvendinimai ir pan." (angl. *Observations concerning the Increase of Mankind, Peopling of Countries, etc.*), kurioje tyrinėjo tuometinės britų kolonijos Amerikoje populiacijos pokyčius. Pagrįsdamas tuo, kad jauna populiacija yra žymiai judresnė ir labiau kintanti negu senos pastovios šalys, nustatė, kad Amerikos gyventojų populiacija didėja eksponentiškai ir per 25 metus padvigubėja [4]. Praėjus daugiau negu 250 metų nuo B. Franklin'o publikacijos literatūroje eksponentinis populiacijos kitimo modelis vis dar yra aprašomas ir pristatomas kaip paprasčiausias ir nesunkiai interpretuojamas. Eksponentinis populiacijos kitimo modelis yra aprašomas diferencialine lygtimi

$$\frac{dP_t}{dt} = RP_t \quad (2)$$

su pradine sąlyga  $P_0$ , kai  $t = 0$  [7]. Eksponentinio populiacijos kitimo modelio parametras  $R$  yra vadinamas **momentine populiacijos kitimo norma** [7].

Išsprendę (2) lygtį gauname:

$$\begin{aligned} \frac{dP_t}{P_t} &= Rdt \\ \int \frac{dP_t}{P_t} &= \int Rdt \\ P_t &= Ce^{Rt}. \end{aligned}$$

Pasinaudoję pradine sąlyga gauname eksponentinio modelio lygties sprendinį:

$$P_t = P_0 e^{Rt}. \quad (3)$$

Metinė ir momentinė populiacijos kitimo normos yra susijusios lygybe  $R = \ln(1 + r)$ , kai  $r \neq -1$  [7]. Iš (1) lygybės turime, kad

$$P_{t+1} = P_t + \Delta P_{t,1} = P_t \left( 1 + \frac{\Delta P_{t,1}}{P_t} \right) = P_t (1 + r_{t,1}) = P_t (1 + r).$$

Tęsdami lygybę toliau, gauname, kad

$$P_{t+1} = P_{t-1} (1 + r)^2 = \dots = P_0 (1 + r)^{(t+1)} = P_0 e^{(t+1) \ln(1+r)}. \quad (4)$$

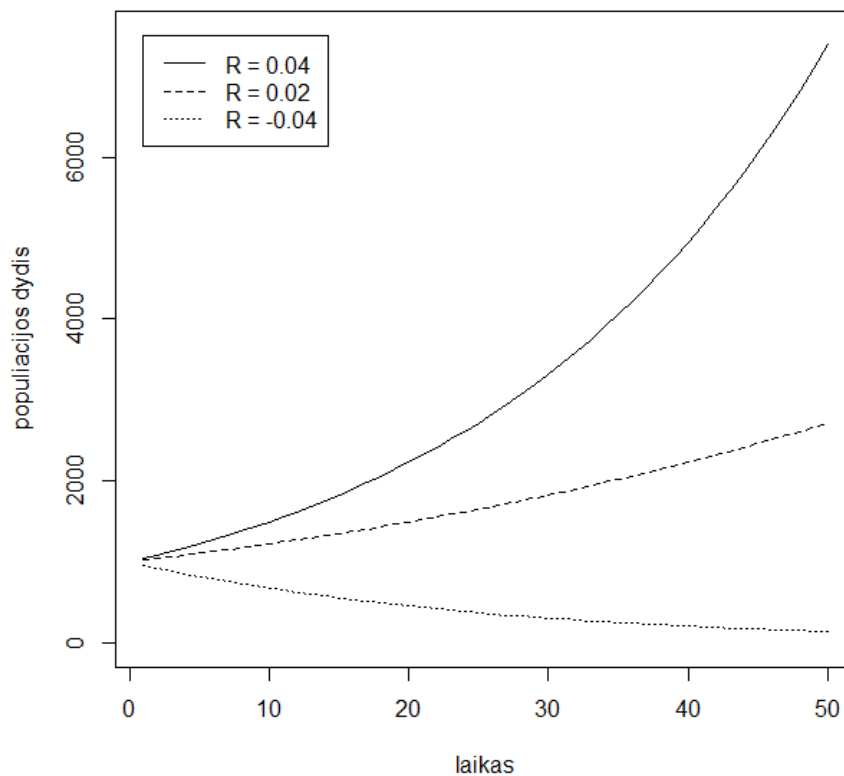
Kita vertus, iš (3) lygybės turime:

$$P_{t+1} = P_0 e^{R(t+1)}. \quad (5)$$

Kairės (4) ir (5) lygybių pusės sutampa, vadinasi, ir dešinės lygybių pusės yra lygios. Todėl

$$R = \ln(1 + r).$$

Praktikoje įprastai  $r \in (-0,04; 0,04)$  [7], todėl  $r \neq -1$  apribojimo galime nepaisyti. Jei  $r$  yra "pakankamai mažas", galime taikyti Teiloro skleidinį nulinio aplinkoje. Turime, kad  $R = \ln(1 + r) = r + o(r^2) \approx r$  [7].



1 pav.: Eksponentinis populiacijos kitimo modelis,  $P_0 = 1000$ ,  $t = 50$ .

Eksponentinio populiacijos kitimo modelio pavyzdžiai yra pavaizduoti 1 paveiksle.

Didžiausias šio modelio trūkumas - jį taikant populiacijos dydis neapibrėžtai auga, o tai įmanoma tik tada, jei yra begaliniai maisto ir gyvenimo vietos išteklių. Kadangi realybėje tokia aplinka neegzistuoja, eksponentinis modelis yra taikomas tik prognozuojant baigtinį (ir dažniausiai trumpą) laiko tarpą [14].

### 2.3. Logistinis modelis

Susidūrus su neapibrėžto augimo problema eksponentiniame modelyje, buvo pradėta ieškoti tikslesnio populiacijos įverčio. Belgų matematikas Pierre François Verhulst 1838 metais leidinyje "Matematikos ir fizikos leidinys" (pranc. *Correspondance Mathématique et Physique*) pristatė patobulintą populiacijos augimo modelio lygtį. Terminas "logistinis" pirmą kartą pavartotas tik 1920 metais, kai po R. Pearl'o ir L. J. Reed'o, publikacijos P.F. Verhulst'o darbas buvo vėl prisimintas [1]. R. Pearl ir L. J. Reed nebuvo susipažinę su P.F. Verhulst'o lygtimi, kai leidinyje publikavo savo skaičiavimus, todėl kai kuriuose šaltiniuose logistinis modelis dar yra vadinamas Pearl-Verhulst lygtimi [1]. logistinis populiacijos kitimo modelis yra aprašomas diferencialine lygtimi

$$\frac{dP_t}{dt} = RP_t \left(1 - \frac{P_t}{K}\right), \quad (6)$$

čia  $K$  - prisisotinimo koeficientas, t.y. maksimalus populiacijos dydis, o visi kiti parametrai apibrėžti taip pat, kaip ir eksponentiniame modelyje [14].

Išspręsimė (6) lygtį:

$$\frac{dP_t}{dt} = RP_t \left(1 - \frac{P_t}{K}\right),$$

Iš

$$\frac{dP_t}{P_t \left(1 - \frac{P_t}{K}\right)} = Rdt$$

turime:

$$\int \frac{dP_t}{P_t \left(1 - \frac{P_t}{K}\right)} = \int Rdt.$$

Pritaikę neapibrėžtųjų koeficientų metodą, turime:

$$\int \frac{dP_t}{P_t} + \int \frac{dP_t}{K - P_t} = \int Rdt.$$

Galiausiai

$$\ln(P_t) - \ln(K - P_t) = Rt + \ln(C).$$

Tada

$$P_t = \frac{CK}{e^{-Rt} + C}. \quad (7)$$

Įstatę pradinę reikšmę randame konstantos  $C$  išraišką:

$$P_0 = \frac{CK}{1 + C} \Rightarrow C = \frac{P_0}{K - P_0}.$$

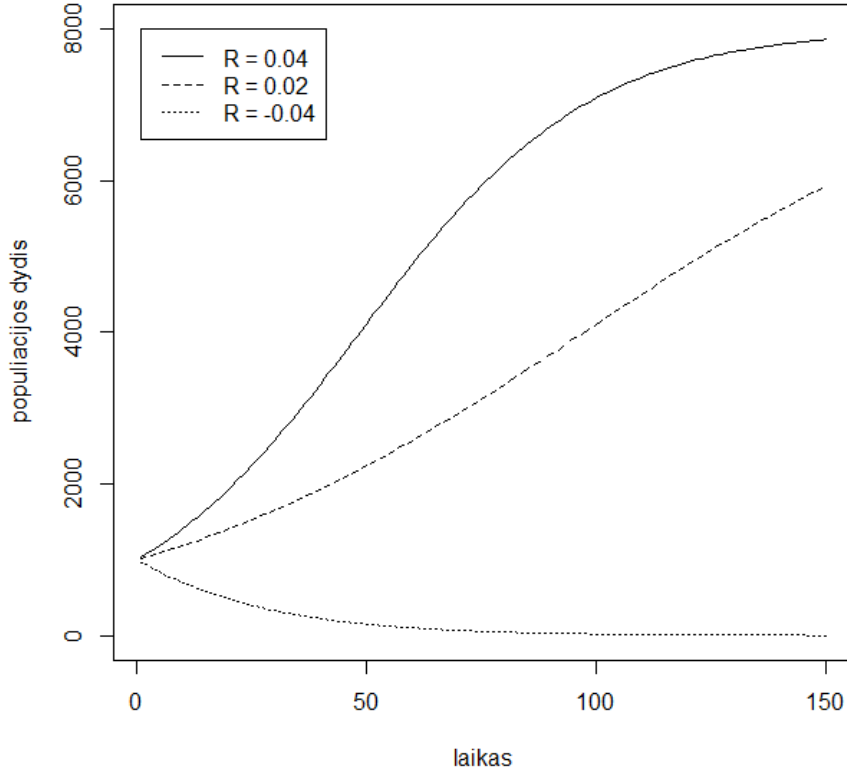
Įstatome  $C$  išraišką į (7) lygtį:

$$P_t = \frac{\frac{P_0}{K - P_0} K}{e^{-Rt} + \frac{P_0}{K - P_0}}.$$

Pertvarkius reiškinių gauname logistinio populiacijos kitimo modelio lygties sprendinį

$$P_t = \frac{P_0 K}{(K - P_0) e^{-Rt} + P_0}. \quad (8)$$





2 pav.: Logistinis populiacijos kitimo modelis,  $P_0 = 1000$ ,  $K = 8000$ ,  $t = 150$ .

2 paveiksle galime pamatyti kaip grafiškai atrodo logistinis modelis.

Kai  $R > 0$ , logistinis modelis įgauna S formos kreivę, dar vadinama sigmoidiniu (angl. *sigmoidal*) augimu [14]. Kai  $R < 0$ ,  $t \rightarrow \infty$ , kreivė artėja link nulio. Taip pat įsitikinsime, kad kai  $R < 0$ , didėjant  $K$ ,  $P_t$  mažėja, kai  $t$  yra fiksuotas. Pertvarkome (8) lygybę:

$$P_t = \frac{P_0}{\left(1 - \frac{P_0}{K}\right) e^{-Rt} + \frac{P_0}{K}} = \frac{P_0}{e^{-Rt} - \frac{P_0}{K} (e^{-Rt} - 1)}.$$

Kadangi  $R < 0$ , tai

$$e^{-Rt} - 1 > 0.$$

Kai  $K \rightarrow \infty$ ,

$$\frac{P_0}{K} \rightarrow 0.$$

Iš čia turime, kad kai  $K \rightarrow \infty$ ,

$$\frac{P_0}{K} (e^{-Rt} - 1) \downarrow 0,$$

todėl vardiklyje iš  $e^{-Rt}$  atimame vis mažiau ir mažiau, dėl to

$$e^{-Rt} - \frac{P_0}{K} (e^{-Rt} - 1)$$

didėja, kai  $K$  didėja. Iš čia turime, kad  $P_t$  didėja, kai  $K \rightarrow \infty$ ,  $R < 0$ ,  $t = const$ .

Logistinis modelis yra pranašesnis už eksponentinį tuo, kad turi populiacijos augimo ribą, kurios populiacija peržengti negali, taigi, yra išsprendžiama begalinio populiacijos augimo

problema. Tačiau logistinis modelis yra aprašomas naudojant tą patį  $R$  visiems laiko periodams  $t$ , t.y. yra daroma prielaida, kad populiacija stabili, kai realybėje stabilių populiacijų beveik nepasitaiko. 3 skyriuje apžvelgsime populiacijos kitimo normų  $r$  ir  $R$  prognozavimo metodus.

### 3. Populiacijos kitimo normos prognozavimas

Šiame skyriuje nagrinėsime, kaip keičiasi metinė populiacijos norma  $r$  ir pateiksime jos prognozę.

Populiacijos pokytis per vienerius metus priklauso nuo to, kiek žmonių per metus gimė, mirė ir migravo [15]. Kadangi žmonių gimstamumas ir mirtingumas yra atsitiktiniai dydžiai, metinės populiacijos kitimo normos  $r$  prognozavimui naudosime laiko eilučių (ang. *time series*) modelius. Modeliuodami remsimės hipoteze, kad populiacijos prieaugis yra pats nuo savęs priklausomas, t.y., jei, pvz., tam tikru laiko tarpu buvo mažas žmonių prieaugis, tai ši žmonių karta susilauks palyginus nedaug palikuonių, t.y. po kurio laiko populiacijos prieaugis taip pat bus mažas. Todėl naudosime ARIMA laiko eilučių modelį. ARIMA modelį naudoja Eurostat prognozuodamas populiacijos gimstamumo normą [3], šiame darbe jį taikysime prognozuodami populiacijos pokytį. Toliau pateikiami 7 modelio pritaikymo ir prognozavimo žingsniai, aprašyti šaltinyje [5].

#### 3.1. Grafikas

Kad galėtume susidaryti pirmą išpūdį apie tai, kokie yra duomenys, reikia įvertinti juos grafiškai. Šitame žingsnyje nubrėžiame grafiką, pastebime neįprastas ar išsišokančias reikšmes, ieškome to priežasčių (pvz., staigus ekonominės, politinės ar socialinės situacijos pokytis). Taip pat įvertiname, ar duomenys turi tendą bei sezoniškumą.

#### 3.2. Transformavimas

Jei yra reikalinga, pritaikome duomenų transformaciją. Transformacija yra reikalinga tuomet, jei tai parodo tikslesnį duomenų vaizdą, pvz., jei analizuosime, kiek vištos padeda kiaušinių per mėnesį, tai nevienodas dienų skaičius mėnesiuose gali iškreipti duomenis. Taip pat, jei pastebime besikeičiančią dispersiją laikui bėgant, pritaikome Box-Cox transformaciją. Box-Cox transformacija yra apibrėžiama taip:

$$w_t = \begin{cases} \log(y_t), & \text{jei } \lambda = 0, \\ \frac{(y_t^\lambda - 1)}{\lambda}, & \text{kitur,} \end{cases}$$

čia  $y_t$  yra duomenų eilutė,  $\lambda$  - transformacijos parametras [13].  $\lambda$  turėtų būti pasirinktas toks, kad nauja eilutė  $w_t$  turėtų tą pačią dispersiją visiems  $t$ . Jei duomenis pasirenkame transformuoti, modelį taikome ir prognozės ieškome transformuotai eilutei  $w_t$ , tačiau po to reikia nepamiršti eilutės "transformuoti atgal", t.y. pasinaudojant atvirkštine funkcija gauti prognozę pagal originalią eilutės  $y_t$  skalę [5].

#### 3.3. Stacionarumas

Šiame žingsnyje tiriamo **stacionarumą**. Literatūroje yra apibrėžiamas stacionarumas stipriąja (siaurąja) ir silpnąja (plačiąja prasme). Atsitiktinių dydžių seka  $\{X_t, t \in T\}$  vadi-

nama stacionaria **siauraja (stipriaja)** prasme, jei su visais  $k \in \mathbb{N}$ ,  $t_1, t_2, \dots, t_k$  ir  $h \in \mathbb{Z}$  vektorių  $(X_{t_1}, \dots, X_{t_k})$  ir  $(X_{t_1+h}, \dots, X_{t_k+h})$  pasiskirstymai sutampa. Seka  $\{X_t, t \in T\}$  vadinama stacionaria **plačiaja (silpnaja)** prasme, jeigu:

1.  $\mathbb{E}X_t^2 < \infty, \quad \forall t \in \mathbb{Z}$ ,
2.  $\mathbb{E}X_t = \mathbb{E}X_0, \quad \forall t \in \mathbb{Z}$ ,
3.  $r(s, t) = r(s + h, t + h), \quad \forall s, t, h \in \mathbb{Z}$ .

Čia  $r(s, t)$  yra kovariacinė funkcija, apibrėžiama lygybe

$$r(s, t) = Cov(X_s, X_t) = \mathbb{E}(X_s - \mathbb{E}X_s)(X_t - \mathbb{E}X_t), \quad s, t \in T.$$

Jei eilutė yra nestacionari, naudosime skirtuminių operatorių (angl. *differencing*) duomenų stabilizavimui. Skirtuminis operatorius yra apibrėžiamas lygybe  $\Delta X_t = X_t - X_{t-1} = (1 - \mathbb{B})X_t$ , čia  $\mathbb{B}$  - postūmio operatorius (angl. *backshift operator*),  $\mathbb{B}X_t = X_{t-1}$ . Operatorių kėlimo laipsniu procedūra apibrėžiama:  $\mathbb{B}^j X_t = X_{t-j}$ ,  $\Delta^j X_t = \Delta(\Delta^{j-1} X_t)$ ,  $\Delta^0 X_t = X_t$  ir  $\Delta^j X_t = (1 - \mathbb{B})^j X_t$  [8].

Ne visada stacionarumas yra akivaizdžiai matomas iš grafiko. Stacionarumui patikrinti naudosime KPSS testą. Testas remiasi prielaida, kad nagrinėjamos eilutės forma yra:

$$Y_t = \xi_t + e_t,$$

čia  $e_t$  yra stacionari eilutė, o  $\xi_t$  yra atsitiktinio klaidžiojimo (angl. *random walk*) seka, t.y.

$$\xi_t = \xi_{t-1} + v_t,$$

kur  $v_t$  yra nepriklausomi vienodai pasiskirstę atsitiktiniai dydžiai,  $\mathbb{E}v_t = 0$ ,  $\mathbb{D}v_t = \sigma_v^2$ . Jeigu  $\sigma_v = 0$ , tai  $\xi_t = \xi_0$  visiems  $t$  ir nagrinėjama eilutė  $Y_t$  yra stacionari. Iškeliame hipotezę:

$$\begin{cases} H_0 : \sigma_v^2 = 0; \\ H_A : \sigma_v^2 > 0. \end{cases}$$

Naudojantis regresiniu modeliu  $Y_t = \hat{\mu} + \hat{e}_t$  testo statistika apskaičiuojama pagal formulę:

$$KPSS = \frac{\sum_{t=1}^T S_t^2}{T^2 \hat{\sigma}_\infty^2},$$

čia  $T$  yra eilutės ilgis,  $S_t = \sum_{s=1}^t \hat{e}_s$  dalinė suma.  $\hat{\sigma}_\infty^2$  yra  $\hat{e}_t$  statistinis įvertinimas (angl. *estimator*), [6] šaltinis siūlo naudoti

$$\hat{\sigma}_\infty^2 = \frac{1}{T} \sum_{t=1}^T e_t^2 + \frac{2}{T^2} \sum_{s=1}^l \left(1 - \frac{s}{l+1}\right) \sum_{t=s+1}^T e_t e_{t-s},$$

kur  $l$  yra skirtumas tarp eilutės narių. Kritinės sritys yra pavaizduotos 1 lentelėje [6].

Reikšmingumo lygmuo	0,1	0,05	0,025	0,01
Kritinė reikšmė	0,347	0,463	0,574	0,739

1 lentelė: KPSS testo kritinės reikšmės. Šaltinis: [6]

### 3.4. Autokoreliacinės ir dalinės autokoreliacinės funkcijos tyrimai

Toliau įvertinsime autokoreliacinę (angl. *autocorrelation function (ACF)*) ir dalinę autokoreliacinę (angl. *partial autocorrelation function (PACF)*) funkcijas. Silpnai stacionarios sekos  $\{X_t, t \in T\}$  **autokoreliacinė** funkcija yra vadinama funkcija

$$\rho(k) = \frac{\sum_{t=k+1}^n (X_t - \bar{X}_t) (X_{t-k} - \bar{X}_t)}{\sum_{t=1}^T (X_t - \bar{X}_t)^2}.$$

Autokoreliacinė funkcija parodo koreliaciją tarp tos pačios eilutės narių, kurie skiriasi per atstumą  $k$ . Sekos  $\{X_t, t \in T\}$  **dalinė autokoreliacinė** funkcija yra vadinama

$$\alpha(k) = \text{Corr} \left( X_{t+k} - \hat{X}_{t+k}, X_t - \hat{X}_t \right),$$

čia  $\hat{X}_t$  žymi nario  $X_t$  prognozę [13]. Dalinė autokoreliacinė funkcija parodo koreliaciją tarp modelio liekanų (angl. *residuals*). Iš autokoreliacinės ir dalinės autokoreliacinės funkcijų galime daryti prielaidas apie duomenų modelį [5]. Stacionari seka  $\{X_t, t \in T\}$  yra vadinama **baltu triukšmu** (angl. *white noise*), jei  $\mathbb{E}X_t = 0$  ir

$$r(h) = \begin{cases} \sigma^2, & h = 0, \\ 0, & h \neq 0 [8]. \end{cases}$$

Jei sekos autokoreliacinės funkcijos 95% reikšmių yra intervale  $(\frac{-2}{\sqrt{T}}; \frac{2}{\sqrt{T}})$ , čia  $T$  yra laiko eilutės ilgis, sakome, kad seka yra baltas triukšmas [5]. Jeigu daugiau negu 5% reikšmių "išsoka" iš šio intervalo, galime daryti prielaidą, kad duomenys yra autokoreliuoti ir galime taikyti ARIMA modelį. Seka  $\{X_t, t \in \mathbb{Z}\}$  yra vadinama autoregresijos - slenkamojo vidurkio procesu, jeigu  $X_t$  yra stacionarus ir su visais  $t$  galioja

$$X_t - \phi_1 X_{t-1} - \dots - \phi_p X_{t-p} = Z_t + \theta_1 Z_{t-1} + \dots + \theta_q Z_{t-q},$$

čia  $Z_t$  yra baltas triukšmas, raidėmis  $\phi$  yra žymimi koeficientai prie autoregresinės dalies narių, o raidėmis  $\theta$  - prie slenkamojo vidurkio dalies narių. Tokia seka yra žymima  $ARMA(p, q)$  [8]. Seka  $\{X_t, t \in \mathbb{Z}\}$  yra vadinama autoregresijos - slenkamojo vidurkio integruotu procesu, jeigu

$$\phi(\mathbb{B})(1 - \mathbb{B})^d X_t = \theta(\mathbb{B})Z_t,$$

čia  $Z_t$  yra baltas triukšmas. Tokia seka žymima  $ARIMA(p, d, q)$ , kur  $p$  ir  $q$  yra  $ARMA(p, q)$  proceso parametrai. Kitaip tariant, jei seka  $(1 - \mathbb{B})^d X_t$  tenkina  $ARMA(p, q)$  lygtį, tai sakome, kad seka  $X_t$  yra  $ARIMA(p, d, q)$  [8]. Atsižvelgdami į autokoreliacinės ir dalinės autokoreliacinės funkcijos elgesį galime spręsti apie mūsų duomenų modelį.

1. Jei ACF reikšmės stipriai sumažėja po  $n$ -tojo nario, o PACF reikšmės yra palyginus mažos, tai galime daryti prielaidą apie  $ARIMA(0, d, n)$  modelį.
2. Jei ACF reikšmės yra palyginus mažos, o PACF reikšmės pradeda mažėti po  $n$ -tojo nario, tai galime daryti prielaidą apie  $ARIMA(n, d, 0)$  modelį.
3. Jei ir ACF, ir PACF nerodo palyginus didelių reikšmių su jokių parametru, reikia gilesnės analizės [11].

Šiame punkte pagal ACF ir PACF padarome pirmines prielaidas apie modelį.

### 3.5. Kriterijai modelio tinkamumui įvertinti

Toliau patikriname modelių tinkamumą duomenims. Naudodami Maksimalaus tikėtinumo metodą (angl. *maximum likelihood estimation, MLE*) galime rasti tokius modelio parametrus, kurie maksimizuoja tikimybę, kad modelis atitiks realybę. ARIMA modelyje MLE yra apibrėžiamas minimizuojant liekanų kvadratų sumą. Modelį vertinant programinės įrangos aplinka statistiniam skaičiavimui R, R išvestyje (angl. *output*) yra pateikiama logaritminė tikėtinumo reikšmė (angl. *log - likelihood*). Kuo didesnis logaritminis tikėtinumas, tuo labiau tikėtina, kad mūsų nagrinėjamas modelis atitinka tikrovę. Akaike informacijos kriterijus (angl. *Akaike's Information Criterion, AIC*) yra aprašomas pagal formulę

$$AIC = -2 \log(L) + 2(p + q + k + 1),$$

kur  $L$  yra modelio tikėtinumas, o  $k = 1$ , jei ARIMA modelis su poslinkiu (angl. *drift*), t.y. jei ARIMA modelio lygtyje yra konstanta, ir  $k = 0$ , kitu atveju. Jei duomenų eilutė yra gana trumpa, tuomet tikslesnis yra koreguotas Akaike informacijos kriterijus (angl. *corrected Akaike's Information Criterion, AICc*), apibrėžiamas lygybe

$$AICc = AIC + \frac{2(p + q + k + 1)(p + q + k + 2)}{T - p - q - k - 2},$$

ir Bajeso informacijos kriterijus (angl. *Bayesian Information Criterion, BIC*) apibrėžiamas kaip

$$BIC = AIC + (\log(T) - 2)(p + q + k + 1).$$

Kuo mažesni AIC, AICc ir BIC, tuo modelis tikslesnis. [5] šaltinis rekomenduoja prioritetizuoti AICc kriterijų.

### 3.6. Liekanų analizė

Atliekame liekanų analizę. Jei modelis yra parinktas teisingai, tuomet jo liekanos turi būti baltas triukšmas. Pirminį patikrinimą atliekame vertindami autokoreliacinės funkcijos reikšmes pagal 3.4 punkte aprašytą principą: jei sekos autokoreliacinės funkcijos 95% reikšmių yra intervale  $(\frac{-2}{\sqrt{T}}; \frac{2}{\sqrt{T}})$ , čia  $T$  yra laiko eilutės ilgis, sakome, kad seka yra baltas triukšmas. Tačiau vertinant kiekvieną autokoreliacinės funkcijos reikšmę atskirai išlieka 1

tipo klaidos tikimybė kiekvienai funkcijos reikšmei, todėl ilgesnėms eilutėms tikimybė, kad teigsime, jog liekanos yra baltas triukšmas, kai iš tikrųjų taip nėra. Bendrai visos eilutės liekanų analizei yra naudojami **portmanteau** testai, pavadinti pagal prancūzišką žodį, reiškiantį "lagaminas". Vienas iš testų yra **Box-Pierce** testas, aprašomas formule:

$$Q = T \sum_{k=1}^h r_k^2,$$

čia  $h$  yra maksimalus naudojamas skirtumas tarp eilutės narių, o  $T$  yra eilutės ilgis. Yra rekomenduojama naudoti  $h = 10$  nesezoniškoms eilutėms ir  $h = 2m$  sezoniškoms eilutėms, kur  $m$  - sezoniškumo periodas, tačiau testas neveikia korektiškai, kai  $h$  yra didelis, todėl jei eilutė yra pakankamai trumpa, t.y.  $\frac{T}{5} < 10$ , patartina naudoti  $h = \frac{T}{5}$ . Didelė  $Q$  reikšmė reiškia, kad autokoreliacinės funkcijos reikšmės yra pakankamai didelės, kad liekanų seka nebūtų baltas triukšmas. Panašus ir tikslesnis yra **Ljung-Box** testas, aprašomas formule

$$Q^* = T(T+2) \sum_{k=1}^h (T-k)^{-1} r_k^2.$$

Kaip ir  $Q$ , taip ir  $Q^*$  didelės reikšmės reiškia, kad modelio liekanos nėra baltas triukšmas. Kokio dydžio reikšmė yra "didelė"? Jei autokoreliacinė funkcija yra baltas triukšmas, tai  $Q$  ir  $Q^*$  yra pasiskirstę pagal  $\chi^2$  skirstinį su  $(h-K)$  laisvės laipsnių (angl. *degrees of freedom*), kur  $K$  yra parametrų skaičius naudotas modelyje [5].

### 3.7. Prognozė

Modelio prognozė yra apskaičiuojama pertvarkant ARIMA(p, d, q) narius. Tarkime, turime ARIMA(p, d, q) modelį, kuris užrašomas taip:

$$(1 - \mathbb{B})^d (X_t - \phi_1 X_{t-1} - \dots - \phi_p X_{t-p}) = Z_t + \theta_1 Z_{t-1} + \dots + \theta_q Z_{t-q},$$

tikslas yra pertvarkyti reiškinių taip, kad  $X_t$  būtų vienoje lygybės pusėje, o visi kiti nariai kitoje. Pvz., jei  $d = 1$ ,  $p = 2$ ,  $q = 3$ , tai lygtis atrodo taip:

$$(1 - \mathbb{B})(X_t - \phi_1 X_{t-1} - \phi_2 X_{t-2}) = Z_t + \theta_1 Z_{t-1} + \theta_2 Z_{t-2} + \theta_3 Z_{t-3},$$

atskliaudus ir sutraukus panašius narius turime:

$$X_t = (\phi_1 - 1)X_{t-1} + (\phi_2 - \phi_1)X_{t-2} - \phi_3 X_{t-3} + Z_t + \theta_1 Z_{t-1} + \theta_2 Z_{t-2} + \theta_3 Z_{t-3}.$$

Dabar  $t$  pakeičiame į  $T+1$  ir darydami prielaidą, kad turime  $T$  stebėjimų, galime prognozuoti  $T+1$  stebėjimą:

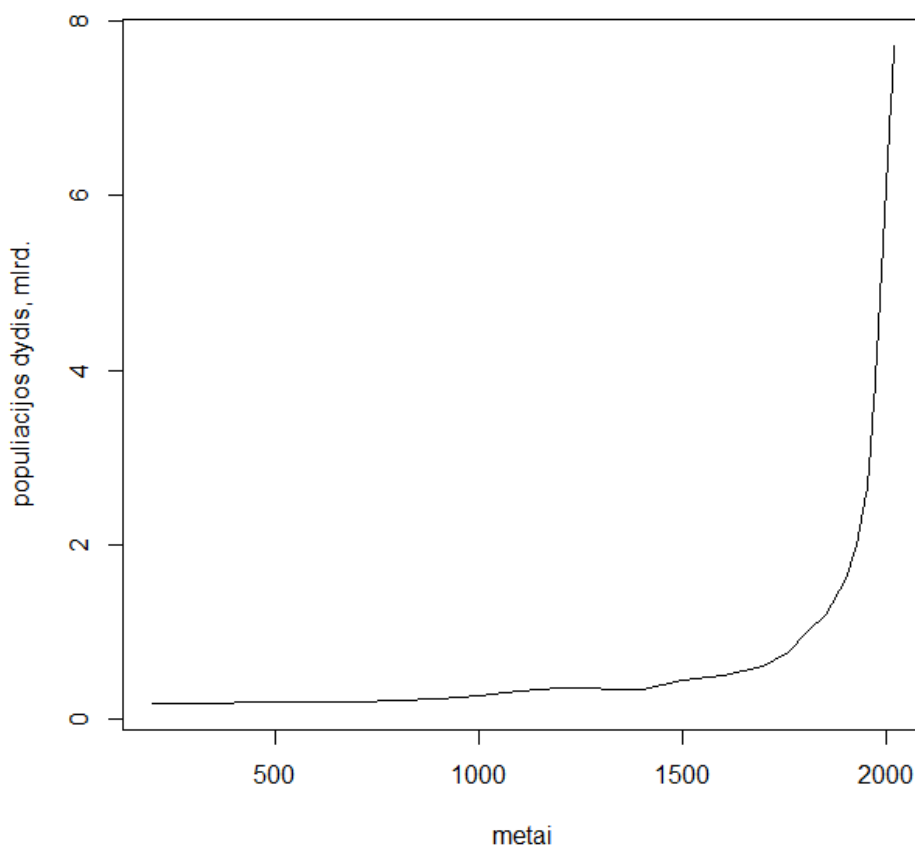
$$\hat{X}_{T+1|T} = (\phi_1 - 1)X_T + (\phi_2 - \phi_1)X_{T-1} - \phi_3 X_{T-2} + \theta_1 Z_T + \theta_2 Z_{T-1} + \theta_3 Z_{T-2}.$$

Kadangi  $Z_t$  yra baltas triukšmas, tai  $\hat{Z}_{T+1} = 0$ . Analogiškai galime prognozuoti kiek norime daug stebėjimų [5].

## 4. Pasaulio populiacijos prognozė

### 4.1. Bendra apžvalga

Populiacijos retai kada būna stabilios. Vienos populiacijos turi tendenciją mažėti, kitos - didėti. Šiuo metu pasaulio populiacija yra pati didžiausia kokia tik yra buvusi prieš tai ir vis didėja. 3 paveiksle pavaizduota pasaulio populiacijos dydis 0 - 2019 metais<sup>1</sup>. Populiacijos dydis iki 1950 metų yra tiesiškai ekstrapoliuotas atgal remiantis įvairiais rastaistoriniais šaltiniais apie žmonių populiacijos dydį įvairiais žmonijos laikotarpiais [16].



3 pav.: Pasaulio populiacijos dydis 0-2019 metais.

Pasiekti pirmąjį populiacijos milijardą žmonėms reikėjo daugiau negu 100 tūkst. metų [10], o per pastaruosius kiek daugiau negu 200 metų populiacija pasiekė ir likusius 6 milijardus. 2 lentelėje matome, kiek metų reikėjo kiekvienam Žemės milijardui. Nepaisant to, kad pasaulio istorijoje užfiksuotas ne viena pandemine liga ir ne vienas karas, nusinešę daugelio žmonių gyvybes, pasaulio populiacija laikui bėgant nemažėja. Akivaizdu, kad Žemė negali sutalpinti begalinio skaičiaus gyventojų, nes visiems pritrūks gyvenamo ploto, maisto, geriamo vandens, todėl kyla klausimas, kiek žmonių gali tilpti pasaulyje.

Žemės paviršiaus plotas yra 510 mln. km<sup>2</sup>, iš kurių sausuma sudaro 149 mln. km<sup>2</sup>

<sup>1</sup> duomenų šaltinis <https://www.worldometers.info/world-population/world-population-by-year/>



gyventojų skaičius	metai	kiek užtruko metų
1 mlrd.	1804 m.	-
2 mlrd.	1927 m.	123 m.
3 mlrd.	1960 m.	33 m.
4 mlrd.	1974 m.	14 m.
5 mlrd.	1987 m.	13 m.
6 mlrd.	1999 m.	12 m.
7 mlrd.	2011 m.	12 m.

2 lentelė: Kiek metų reikėjo kiekvienam Žemės milijardui.

(29,2%), o vanduo - 361 mln. km<sup>2</sup> (70,8%) [12]. Šiuo metu žmonių tankumas, t.y. vidutinis gyventojų skaičius kvadratiniam kilometre, yra 52 žm./km<sup>2</sup>.

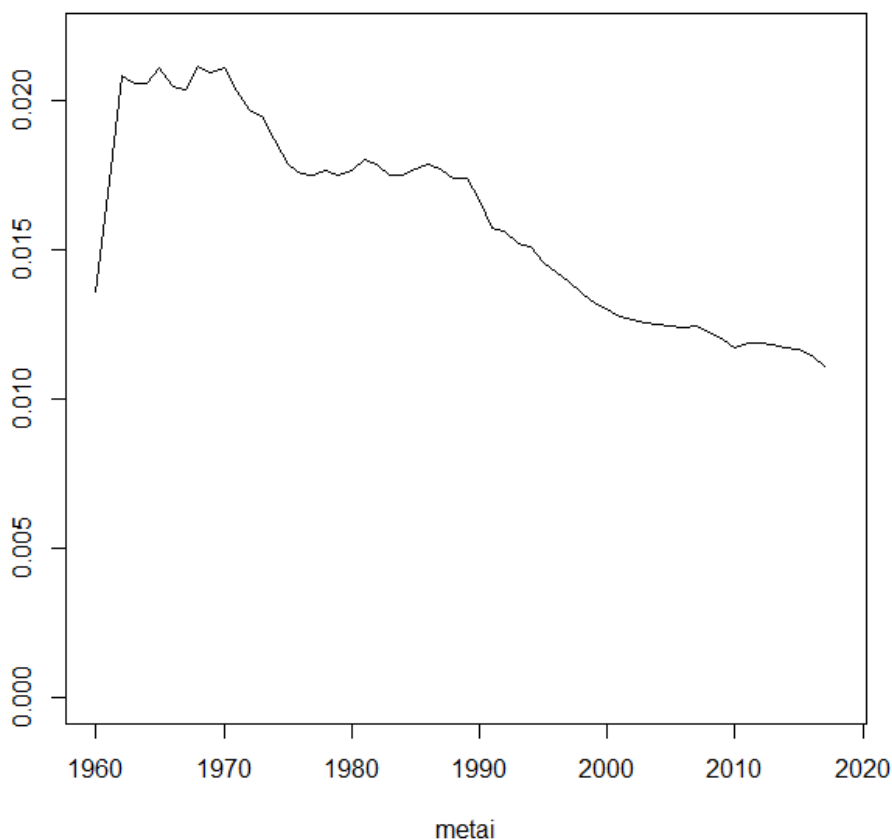
Kiek žmonių biologiškai gali gyventi pasaulyje buvo pradėta svarstyti dar XVII amžiuje. Olandų mokslininkas Antoni van Leeuwenhoek'as įvertino, kad jei tuometinės Olandijos gyventojų skaičių (1 mln.) ekstrapoliuosime visose tuo metu žinomose pasaulio teritorijose, pasaulio žmonių skaičius bus lygus 13 mlrd. žmonių. Nuo to laiko buvo daug maksimalaus žmonių skaičiaus pasaulyje įverčių, tačiau vieno bendro sutarimo iki šiol nėra rasta [10].

Maksimalus skaičius žmonių pasaulyje yra vertinamas lyginant Žemės biologinę talpą (angl. *biocapacity*) - biologiškai produktyvų žemės ir vandens plotą, atsirandantį iš miškų, ganyklų, pasėlių ir tvenkinių, reikalingų gaminti visą biomasę, kurią naudoja žmonės, ir absorbuoti visas žmonių pagamintas atliekas, įskaitant išmetamą CO<sub>2</sub> - ir žmogaus ekologinį pėdsaką (angl. *ecological footprint*) - matą, kuris nurodo, kiek Žemės generuojamų resursų žmogus sunaudoja. Jei Žemės biologinė talpa yra didesnė arba lygi visų gyventojų ekologiniam pėdsakui, tai žmogaus resursų naudojimas yra apibrėžiamas kaip tvarus (angl. *sustainable*) ir ilgalaikis tokio žmonių kiekio gyvenimas yra patogus. Jei Žemės biologinė talpa yra mažesnė negu žmogaus ekologinis pėdsakas, tai toks žmonijos vystymasis yra nedarnus ir gali lemti ekologinę katastrofą. Šiuo metu vidutinė Žemės biologinė talpa vienam gyventojui yra 1,78 gha/žm. (globalių hektarų vienam žmogui), o vidutinis žmogaus ekologinis pėdsakas yra 2,23 gha/žm. [2]. Šiuo metu žmonijos vartojimas nėra tvarus.

Darant skirtingas prielaidas apie dabarties ir ateities žmonių gyvenimo būdą ir Žemės resursų naudojimą [10] straipsnyje yra išskiriami 94 skirtingi įverčiai, varijuojantys nuo 2 iki 1024 mlrd. gyventojų. Dažniausias įvertis yra 8 mlrd. gyventojų, t.y. vos 300 mln. daugiau negu kad Žemėje gyvena dabar. Kitas gana dažnas įvertis yra 16 mlrd. gyventojų. Skaičiavimams naudosime  $K = 10$  mlrd. žmonių, kaip "Pesimistinį" žmonijos scenarijų, atsižvelgiant į dabartinio vartojimo mastus ir tai, kad dabartinis žmonijos gyvenimas nėra tvarus, ir  $K = 16$  mlrd. žmonių, kaip "Optimistinį" scenarijų, su lūkesčiu, kad žmogus gali gyventi ir palikdamas mažesnę ekologinį pėdsaką negu dabar.

## 4.2. Pasaulio metinės populiacijos kitimo normos prognozavimas

Pasaulio metinė populiacijos kitimo norma pavaizduota 4 paveiksle<sup>2</sup>.



4 pav.: Metinė pasaulio populiacijos kitimo norma 1960-2017 metais.

Pastebime, kad didžiausias pasaulio populiacijos augimo normos šuolis buvo 1960-1962 metais, kai 1962 metais norma peržengė 2% ribą ir išsilaikė iki 1972 metų. Nors pasaulio populiacijos kitimo norma laikui bėgant mažėja, tačiau nėra mažesnė už 0, tai žmonių pasaulyje vis dar daugėja, nors nebe taip sparčiai.

Sezoniškumo ir kintančios dispersijos duomenyse nepastebime, todėl transformacija yra nebūtina.

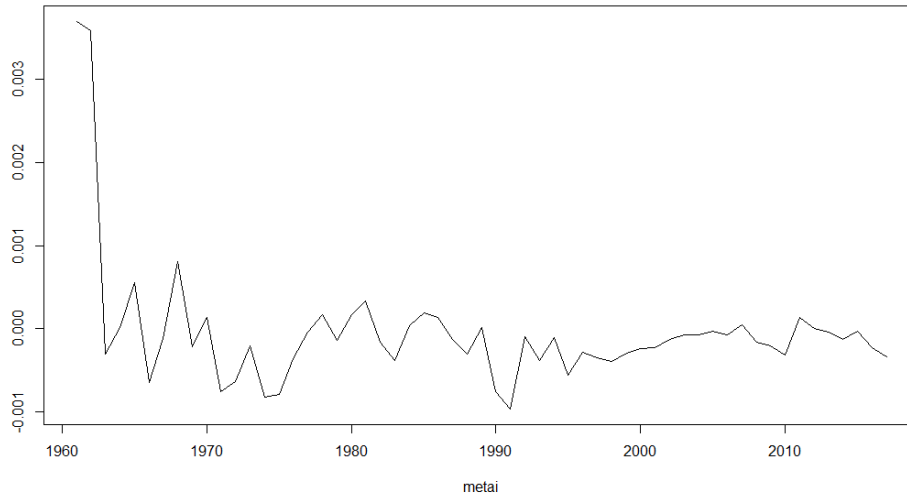
Kadangi yra pastebimas trendas, akivaizdu, kad duomenys nėra stacionarūs. Mūsų spėjimą patvirtina ir atliktas KPSS testas, detali informacija pateikta 3 lentelėje.

Skirtuminio operatoriaus eilė	KPSS testo reikšmė	p-reikšmė
0	1,3777	0,01
1	0,3468	0,1

3 lentelė: Pasaulio populiacijos kitimo normos KPSS testo rezultatai.

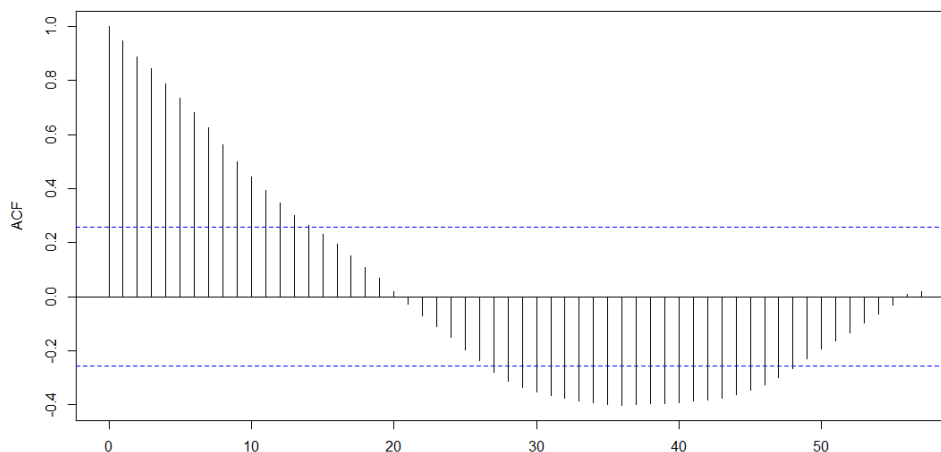
<sup>2</sup>skaičiavimams naudojame tikslus duomenis pasiekiamus adresu <https://data.worldbank.org/indicator/SP.POP.TOTL>

3 lentelėje matome, kad eilutės  $p$ -reikšmė (angl.  $p$ -value) yra mažesnė už 0,01, o tai yra mažiau už standartinę 0,05, todėl atmetame testo nulinę hipotezę ir teigiame, kad eilutė nėra stacionari. Pritaikius skirtuminių operatorių vieną kartą (žr. 5 paveikslą),  $p$ -reikšmė lygi 0,1, taigi, nulinės hipotezės neatmetame, eilutė (pritaikius skirtuminių operatorių vieną kartą) stacionari.

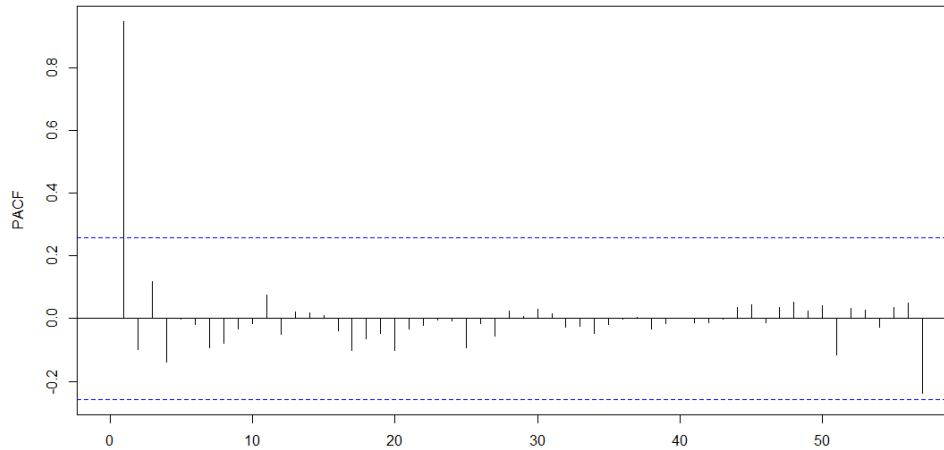


5 pav.: Metinė pasaulio populiacijos kitimo norma 1960-2017 metais pritaikius skirtuminių operatorių vieną kartą.

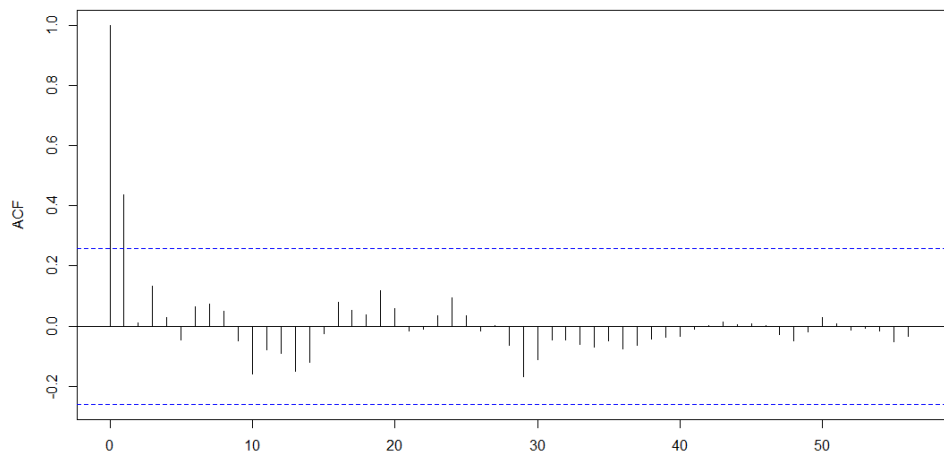
6 ir 7 paveiksluose yra pavaizduotos autokoreliacinė ir dalinė autokoreliacinė funkcijos. Kadangi mūsų nagrinėjama laiko eilutė pritaikius skirtuminių operatorių vieną kartą yra stacionari, tai modeliui nustatyti pasinaudosime 8 grafiku.



6 pav.: Pasaulio populiacijos kitimo normos autokoreliacinė funkcija.



7 pav.: Pasaulio populiacijos kitimo normos dalinė autokoreliacinė funkcija.



8 pav.: Pasaulio populiacijos kitimo normos pritaikius skirtuminį operatorių autokoreliacinė funkcija.

Viso išrinkome 7 galimus ARIMA modelius tolesniam nagrinėjimui. Modelių parinkimui, analizei ir vertinimui naudojome programinės įrangos aplinką statistiniam skaičiavimui ir grafiniam vaizdavimui R. Modelis fit1 yra gautas automatiškai R generuojant parametrus iš visų galimų, fit2 - analogiškai, tik nurodant įvestyje (angl. *input*) rinktis paprastesnius modelius, perrenkant mažiau pasirinkimų. Patikrinus fit1 modelio autoregresinės dalies parametras, buvo nustatyta, kad jis yra modeliui nereikšmingas, todėl buvo pasirinkti modelio fit3 parametrai. Naudojantis 3.4 skyriaus nurodymais, pastebime, kad 8 grafike 4 skirtumas dar yra gana ryškus, todėl parenkame fit4 modelio parametrus testavimui. Modeliai fit5, fit6 ir fit7 įtraukti į analizę viliantis, kad autoregresinės dalies parametrai gali būti reikšmingi modeliui, tačiau juos patikrinus, pastebime, kad taip nėra. Mažiausią *AICc* reikšmę turi modelis fit4, todėl jį pasirenkame kaip tinkamiausią. Taigi, metinės pasaulio populiacijos

kitimo normos prognozavimui naudosime

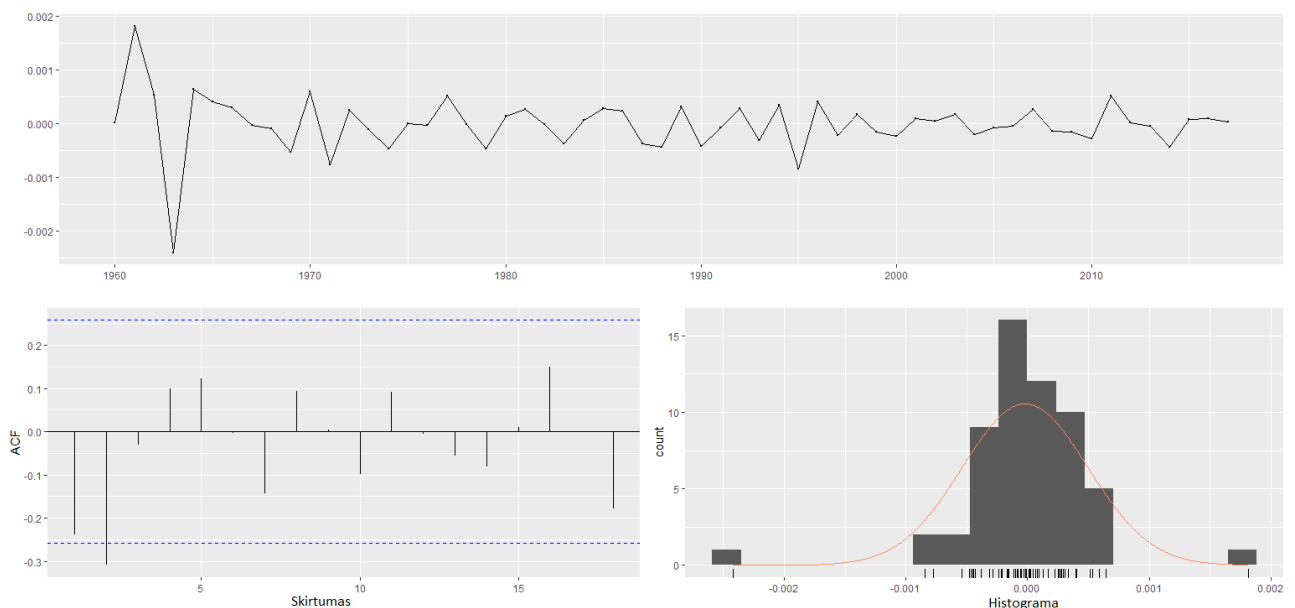
$$X_t = X_{t-1} + Z_t + 0,9106Z_{t-1} + 0,8003Z_{t-2} + 1,1045Z_{t-3} + 0,6645Z_{t-4}$$

Detali informacija pateikta 4 lentelėje.

Nr.	Modelis	Maks. tikėtinumai	$\sigma^2$	AIC	AICc	BIC
fit1	ARIMA(1, 1, 3)	344,51	$3,218 \times 10^{-7}$	-679,02	-677,84	-668,80
fit2	ARIMA(0, 1, 1)	340,38	$3,825 \times 10^{-7}$	-676,77	-676,55	-672,68
fit3	ARIMA(0, 1, 3)	340,52	$3,948 \times 10^{-7}$	-673,05	-672,28	-664,87
fit4	ARIMA(0, 1, 4)	347,06	$2,870 \times 10^{-7}$	-684,12	-682,94	-673,90
fit5	ARIMA(1, 1, 4)	348,04	$2,812 \times 10^{-7}$	-684,08	-682,40	-671,82
fit6	ARIMA(2, 1, 4)	348,09	$2,874 \times 10^{-7}$	-682,18	-679,89	-667,87
fit7	ARIMA(3, 1, 4)	348,28	$2,889 \times 10^{-7}$	-680,56	-677,56	-664,21

4 lentelė: ARIMA modelių su skirtingais parametrais palyginimas.

Atliekame liekanų analizę. 9 paveiksle pavaizduotos modelio liekanos - viso 3 grafikai. Viršutiniame pavaizduotos pačios liekanos, pastebime, kad jos šokinėja nedaug, vadinasi, galime spręsti, kad liekanos yra stacionarios. Apačioje kairėje esančiame grafike pavaizduota ACF, kadangi tik viena reikšmė išsišoka, vadinasi, galime teigti, kad liekanos yra nekoreliuotos. Apačioje dešinėje yra pateikta liekanų histograma ir Normaliojo skirstinio kreivė. Pastebime, kad liekanų pasiskirstymas panašus į Normalųjų dėsnį, tai tik patvirtina mūsų prielaidą, kad modelio liekanos yra baltos triukšmas. Norėdami įsitikinti savo įžvalgų teisingumą, atliekame Portmanteau testus, rezultatai pavaizduoti 5 lentelėje.



9 pav.: Liekanų analizė.

5 lentelėje pateiktų testų rezultatai sako, kad modelio fit4 liekanos nėra koreliuotos (p-reikšmė  $> 0,05$ ), todėl fit4 modelis yra tinkamas ir galime skaičiuoti prognozę.

Testas	Reikšmė	p-reikšmė
Box - Pierce	3,318	0,06852
Ljung - Box	21,087	0,1752

5 lentelė: Modelio fit4 Portmanteau testai.

Pagal modelį randame prognozę:

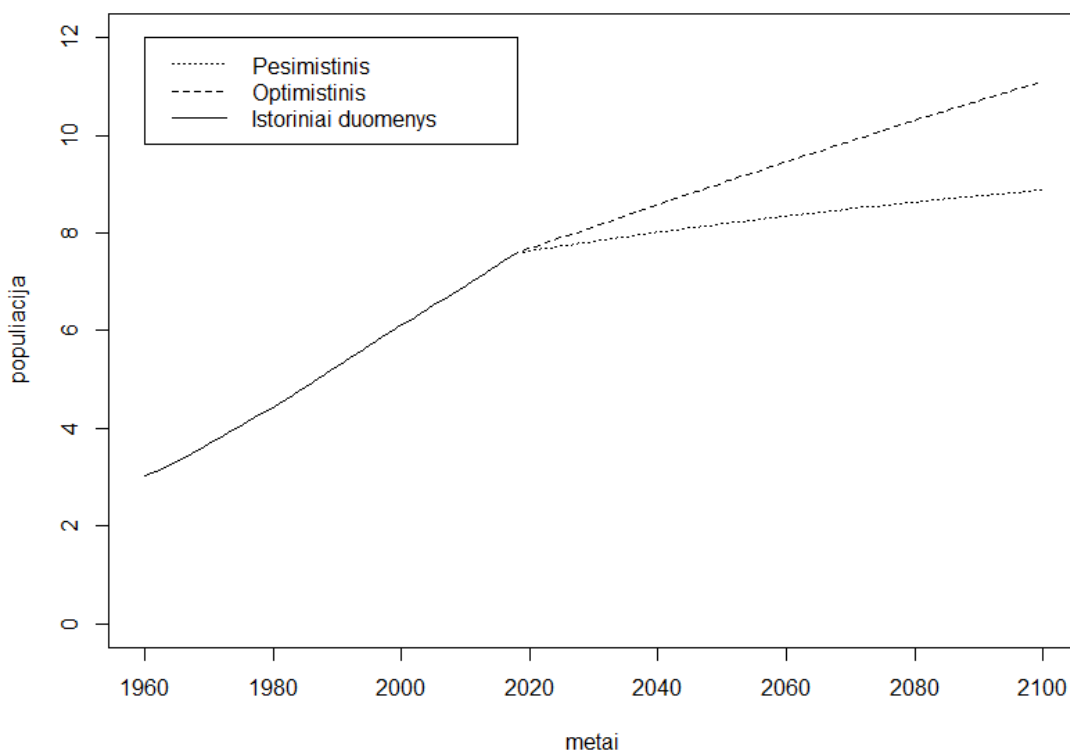
$$r_t = \begin{cases} 0,01098, & t = 2018, \\ 0,01116, & t = 2019, \\ 0,01124, & t = 2020, \\ 0,01126, & t \in [2021, 2099]. \end{cases}$$

Pastebime, kad prognozuojama kitimo norma teigiama, vadinasi, pasaulio populiacija nemažės.

### 4.3. Logistinio populiacijos kitimo modelio pritaikymas pasaulio populiacijai

Šiame skyriuje pagal 2.3, 4.1 ir 4.2 skyrių medžiagą rasime pasaulio populiacijos prognozę.

"Pesimistiniu" atveju pasaulyje 2030 metais gyvens 7,831 mlrd. žmonių, 2050 metais - 8,187 mlrd., o 2100 metais - 8,877 mlrd. žmonių. "Optimistiniu" atveju - 2030 metais gyvens 8,131 mlrd., 2050 m. - 9,021 mlrd., o 2100 m. - 11,096 mlrd. gyventojų. Prognozę galite pamatyti 10 grafike.



10 pav.: Pasaulio populiacijos prognozė 2019 - 2100 m.

Pasaulio populiacijos prognozių palyginimas pavaizduotas 6 lentelėje.

Metai	JTO prognozė	Optimistinis scenarijus	Pesimistinis scenarijus
2050	9,735	9,021	8,187
2100	10,875	11,096	8,877

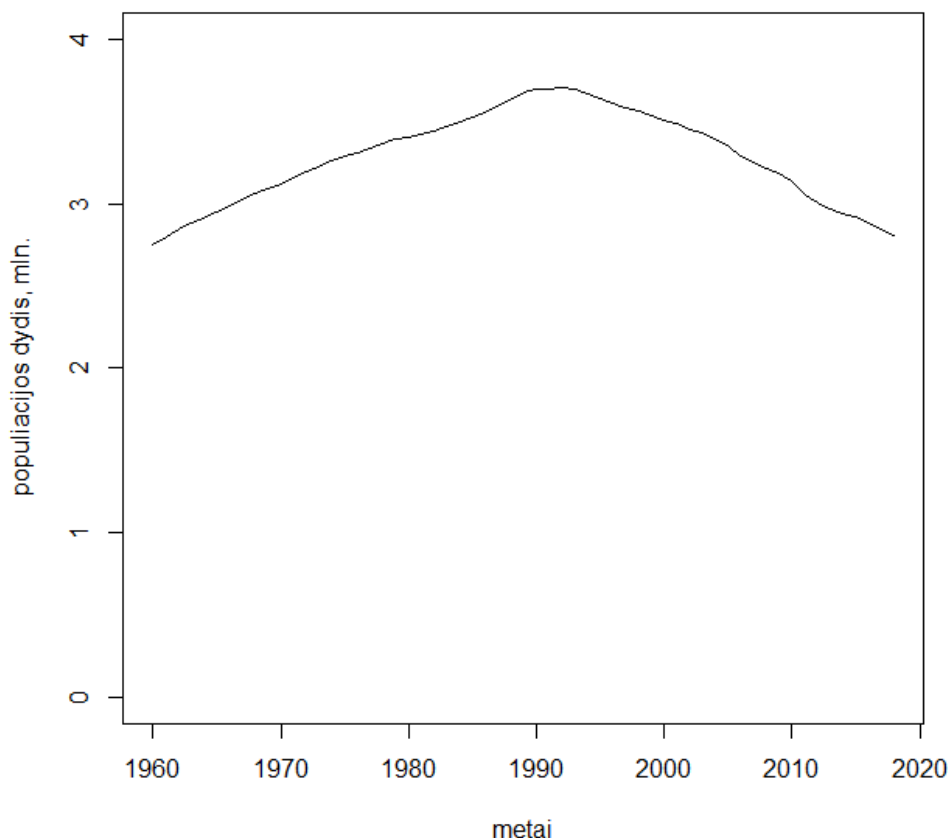
6 lentelė: Pasaulio populiacijos prognozės, mlrd. gyventojų.

Pastebime, kad Optimistinis scenarijus nedaug skiriasi nuo JTO skelbiamų prognozių.

## 5. Lietuvos populiacijos prognozė

### 5.1. Bendra apžvalga

Lietuvos populiacijos dydis yra pavaizduotas 11 paveiksle<sup>3</sup>.



11 pav.: Lietuvos populiacijos dydis 1960-2018 metais.

Daugiausia žmonių Lietuvoje gyveno 1992 metais - 3,7 mln., nuo to laiko žmonių skaičius pamažu mažėja. To priežastys gali būti mažėjantis gimstamumas, vėlyvos santuokos, taip pat emigracija.

Lietuvos populiacija nuo pasaulio skiriasi pagrindiniu principu - yra pasitraukimo galimybė nemirštant - emigracija. Taip pat Lietuvos gyventojų ekologinis pėdsakas palyginus nėra toks reikšmingas pasaulio kontekste, todėl net jei jis būtų ypatingai didelis, jį būtų galima kompensuoti kitų valstybių žmonių ekologija, dėl to maksimalaus kiekio prognozavimas matuojant Žemės resursų gausą ir vartojimo spartą nėra toks tikslus. Nepaisant to, Lietuva savo žmonių tankiu yra panaši į pasaulio populiaciją (Lietuvos respublikos plotas šiuo metu yra 65300 km<sup>2</sup>, o gyventojų tankis - 43 žm./km<sup>2</sup>. Lietuvos didžiausias kada nors buvęs gyventojų tankis yra 57 žm./km<sup>2</sup>), todėl Lietuvos maksimalų gyventojų skaičių nustatysime proporcingai pasaulio populiacijai. "Pesimistiniu" atveju naudosisime  $K = 4$  mln.

<sup>3</sup>tikslūs duomenys iš <https://ec.europa.eu/eurostat/data/database>



ir "Optimistiniu" -  $K = 7$  mln. žmonių. Tokiu atveju dėl logistinio modelio savybės "Pesimistiniu" atveju Lietuvoje gyvens daugiau žmonių negu "Optimistiniu", tačiau norėdami išlaikyti bendrumą pavadinimuose ir sąryšį su tuo, kad "Pesimistinis" byloja apie *mažesnę*  $K$ , pavadinimų nekeisime.

## 5.2. Lietuvos metinės populiacijos kitimo normos prognozavimas

Lietuvos metinės populiacijos kitimo norma yra pavaizduota 12 paveiksle.



12 pav.: Metinė Lietuvos populiacijos kitimo norma 1960-2017 metais.

Iš kreivės matome, kad populiacijos prieaugis Lietuvoje turi tendenciją mažėti, didžiausias prieaugis buvo užfiksuotas 1960 metais, o mažiausias - 2010 metais. Taip pat pastebime, kad prieaugis buvo teigiamas iki 1991 metų, nuo 1992 metų prieaugis yra neigiamas. 1988-1994 metų populiacijos prieaugio staigiam kritimui įtakos galėjo turėti nestabili politinė ir ekonominė Lietuvos situacija: pasitraukimas iš Sovietų sąjungos, ekonominė blokada, prieinamesnė emigracija; 2008-2010 metų populiacijos augimo kritimą galimai sąlygojo prasidėjusi krizė, o staigų augimą 2010-2012 metais galėjo lemti pagerėjusios motinystės atostogų sąlygos.

Sezoniškumas ar kintanti dispersija nėra pastebimi, todėl modeliavimui duomenų transformacija nėra reikalinga.

Atliekame KPSS stacionarumo testą, detali informacija pateikta 7 lentelėje.

Skirtuminio operatoriaus eilė	KPSS testo reikšmė	p-reikšmė
0	1,4248	0,01
1	0,048	0,1

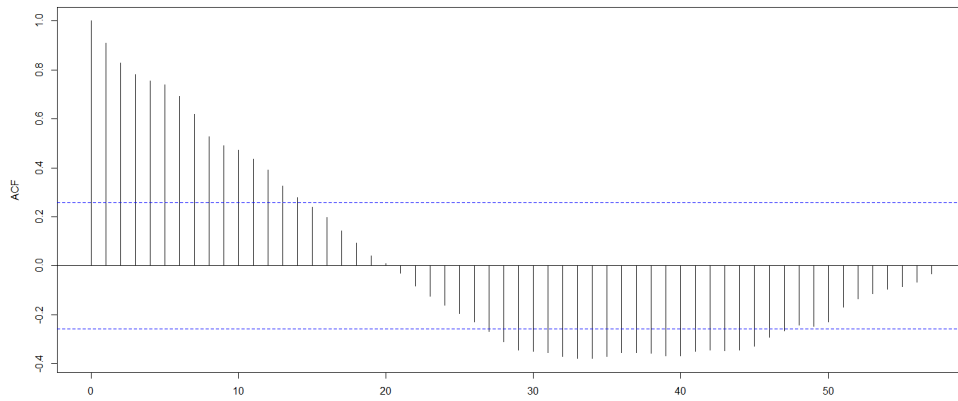
7 lentelė: Lietuvos populiacijos kitimo normos KPSS testo rezultatai.

7 lentelėje matome, kad eilutės p-reikšmė yra lygi 0,01, nulinę hipotezę atmetame - eilutė nėra stacionari, o panaudojus skirtuminį operatorių p-reikšmė yra 0,1, nulinės hipotezės atmesti nėra pagrindo, todėl teigiame, kad panaudojus skirtuminį operatorių vieną kartą (žr. 13 paveikslą) eilutė yra stacionari.

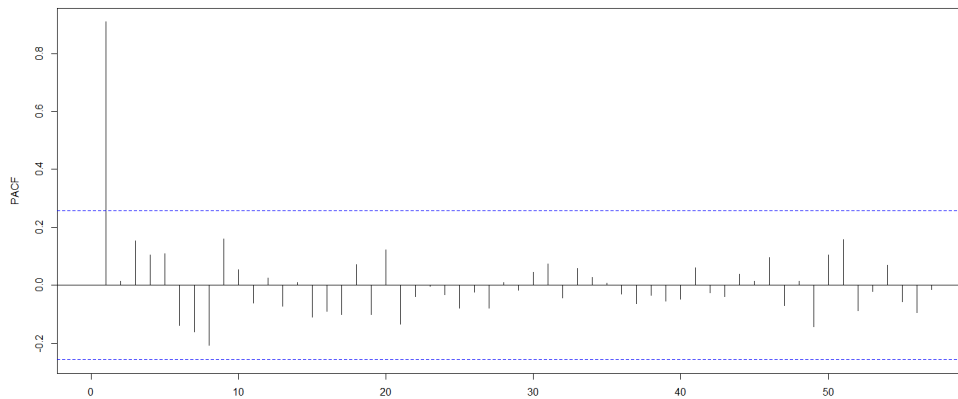


13 pav.: Metinė Lietuvos populiacijos kitimo norma panaudojus skirtuminį operatorių vieną kartą.

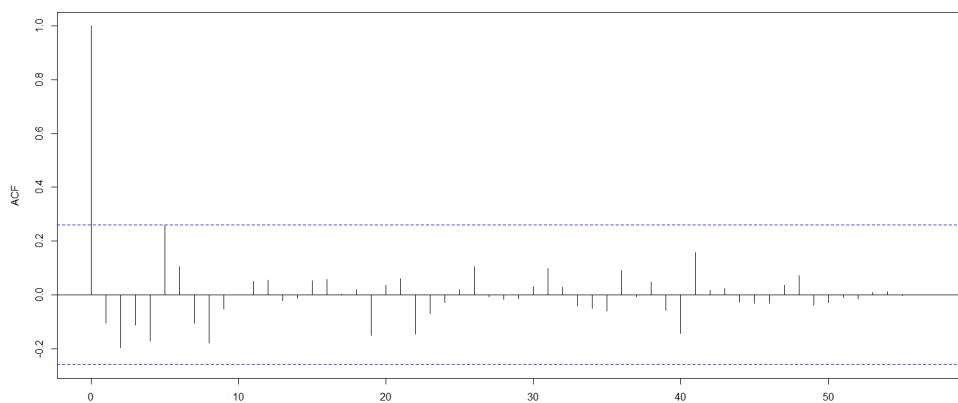
14 ir 15 pavaizduotos Lietuvos populiacijos metinės kitimo normos autokoreliacinė ir dalinė autokoreliacinė funkcijos. Analogiškai pasaulio populiacijos atvejui modeliui sudaryti pasinaudosime ir 16 grafiku.



14 pav.: Autokoreliacinės funkcijos grafikas.



15 pav.: Dalinės autokoreliacinės funkcijos grafikas.



16 pav.: Autokoreliacinės funkcijos grafikas panaudojus skirtuminį operatorių vieną kartą.

Nagrinesime 4 modelius. Modelis fit1 sudarytas automatiškai nagrinėjant visus galimus variantus, fit2 - automatiškai, tačiau neperrenkant visų galimų variantų, o atsirenkant labiau tikėtinus. Modelis fit3 sudarytas remiantis ACF ir PACF diagramomis. Pastebime, kad

slenkančio vidurkio ir autoregresinės dalių parametrai modeliuose fit1 ir fit3 yra nereikšmingi, todėl parenkame modelio fit4 parametrus, tačiau jie taip pat nėra reikšmingi. Išvada: tinkamiausias modelis yra fit2, t.y. ARIMA(0,1,0), arba kitaip - atsitiktinis klaidžiojimas. Modelis yra užrašomas taip:

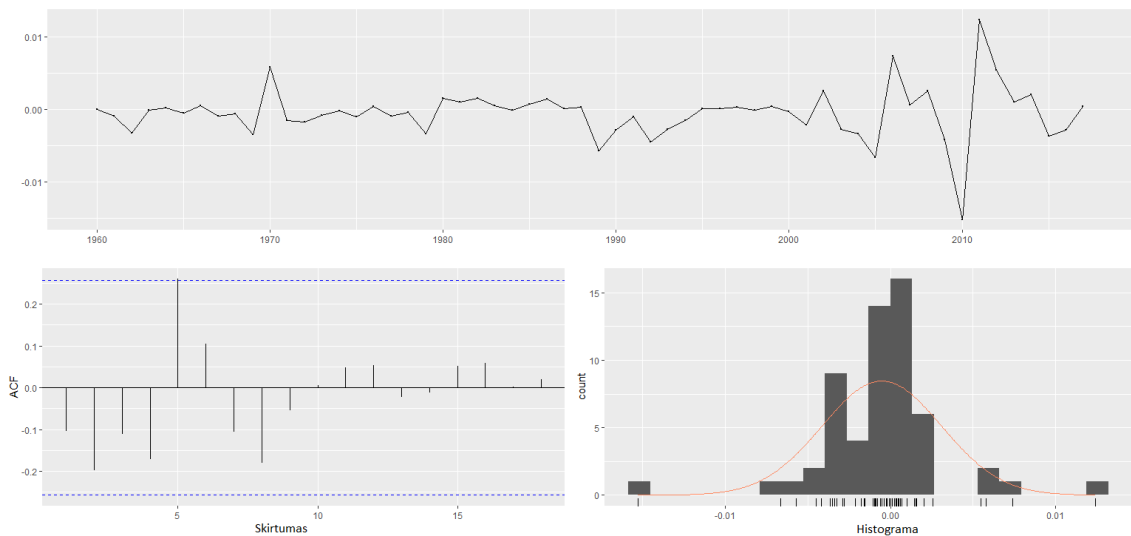
$$(1 - B)X_t = Z_t \Rightarrow X_t = X_{t-1} + Z_t.$$

Detali informacija pavaizduota 8 lentelėje.

Nr.	Modelis	Maks. tikėtinumai	$\sigma^2$	AIC	AICc	BIC
fit1	ARIMA(0, 1, 2)	242,68	$1,232 \times 10^{-5}$	-477,36	-476,59	-469,18
fit2	ARIMA(0, 1, 0)	239,27	$1,346 \times 10^{-5}$	-474,54	-474,32	-470,45
fit3	ARIMA(1, 1, 1)	240,30	$1,320 \times 10^{-5}$	-474,60	-474,15	-468,47
fit4	ARIMA(0, 1, 1)	239,39	$1,340 \times 10^{-5}$	-474,77	-474,55	-470,68

8 lentelė: ARIMA modelių su skirtingais parametrais palyginimas.

17 paveiksle pavaizduota liekanų analizė. Iš grafikų liekanos atrodo kaip baltas triukšmas, atliekame Portmanteau testus, rezultatai pavaizduoti 9 lentelėje. Matome, kad abiejų testų p-reikšmė yra daugiau už 0,05, vadinasi, modelio liekanos tikrai yra baltas triukšmas ir modelis yra tinkamas.



17 pav.: Liekanų analizė.

Testas	Reikšmė	p-reikšmė
Box - Pierce	12,634	0,2447
Ljung - Box	14,236	0,1625

9 lentelė: Portmanteau testai.

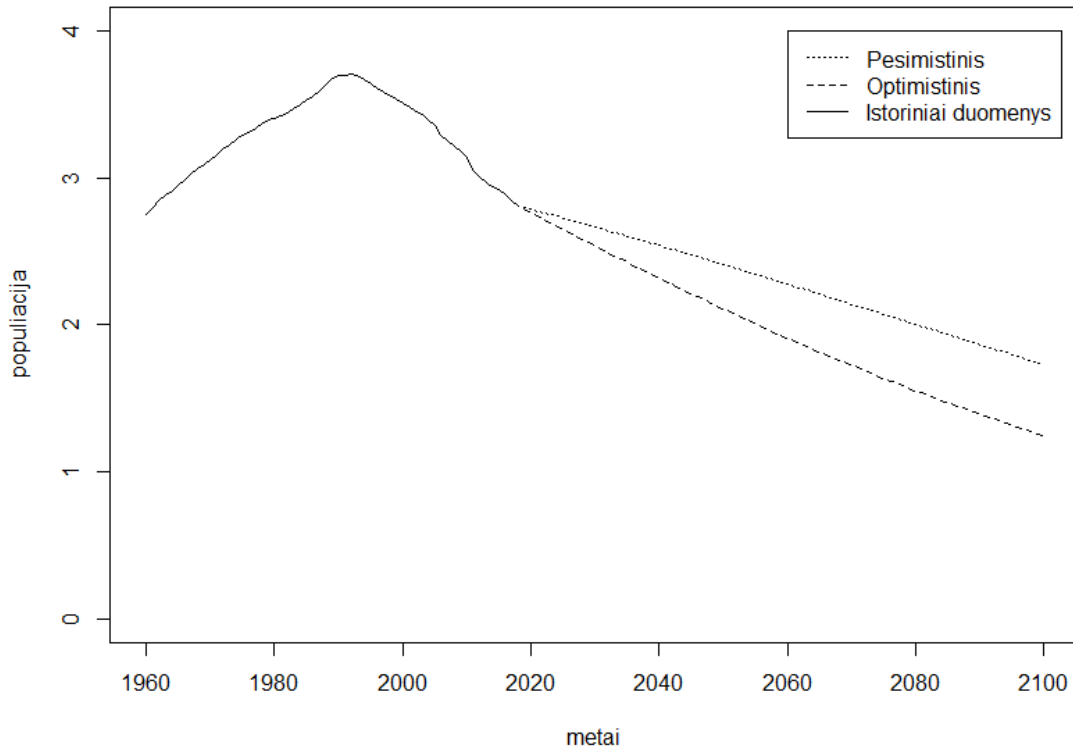
Kadangi taikome atsitiktinio klaidžiojimo modelį, pagal 3.7 skyriuje aprašytą metodiką turime, kad visi ateities spėjimai bus lygūs paskutiniam turimam stebėjimui. Vadinasi,

$$r_t = -0.0137, \quad t \in [2018, 2099].$$

### 5.3. Logistinio populiacijos kitimo modelio pritaikymas Lietuvos populiacijai

Šiame skyriuje pritaikysime 2.3 skyriuje aprašytą logistinį populiacijos kitimo modelį ir 5.1 ir 5.2 skyriuose minėtus populiacijos kitimo normą ir maksimalų populiacijos dydį.

"Pesimistiniu" atveju, Lietuvos populiacijos dydis 2030 metais bus 2,666 mln. gyventojų, 2050 metais - 2,410 mln. gyventojų, o 2100 metais - 1,729 mln. gyventojų. "Optimistiniu" atveju, 2030 metais Lietuvoje gyvens 2,536 mln., 2050 metais - 2,109 mln, o 2100 metais - 1,245 mln. gyventojų. Prognozę galite pamatyti 18 grafike.



18 pav.: Lietuvos populiacijos prognozė 2019 - 2100 m.

Lietuvos prognozių palyginimas yra pavaizduotas 10 lentelėje.

Metai	Eurostat prognozė	Optimistinis scenarijus	Pesimistinis scenarijus
2050	2,175	2,109	2,410
2100	1,625	1,245	1,729

10 lentelė: Lietuvos populiacijos prognozės, mln. gyventojų.

Nepaisant to, kad Lietuvos populiacijos kitimo normai prognozuoti buvo pasirinktas palyginus paprastas atsitiktinio klaidžiojimo modelis, darbe pateiktos prognozės nedaug skiriasi nuo Eurostat'o skelbiamų prognozių.

## 6. Rezultatai ir išvados

Darbe aprašyti eksponentinis ir logistinis populiacijos kitimo modeliai, populiacijos kitimo normos prognozavimo metodika naudojant ARIMA modelį. Logistiniame modelyje aprašytas maksimalus populiacijos dydis įvertintas lyginant Žemės biologinę talpą ir žmogaus ekologinį pėdsaką. Šiuo metu žmonijos vartojimas nėra tvarus ir ilgu laikotarpiu tai gali lemti ekologinę katastrofą, todėl prielaidos apie maksimalų žmonių skaičių pasaulyje priimtos su sąlyga, kad žmonija apribos savo vartojimo mastus. Atsižvelgiant į tai pasirinkti 2 galimi scenarijai: Optimistinis ir Pesimistinis. Lietuvos maksimalus žmonių skaičius parinktas sulyginant pasaulio ir Lietuvos gyventojų tankį.

Gautos populiacijos prognozės palygintos su JTO ir Eurostat'o skelbiamomis prognozėmis. Darbe pateikta pasaulio populiacijos Optimistinio scenarijaus prognozė yra palyginus artima JTO prognozei, manoma, kad 2100 metais pasaulyje gyvens netgi 11 mlrd. gyventojų. Lietuvos populiacijos atveju Pesimistinio scenarijaus prognozė yra artimesnė lyginamajai Eurostat'o prognozei, 2100 metais prognozuojama 1,7 mln. gyventojų.

## 7. Literatūra

- [1] J. S. Cramer. The origins of logistic regression. *Tinbergen Institute Working Paper*, 119(4):891–921, 2002.
- [2] M. Desvaux. The sustainability of human populations: How many people can live on earth? *Significance*, 4(3):102–107, 2007.
- [3] European Commission Eurostat, Directorate F: Social statistics, Unit F-2: Population and migration. Summary methodology of the 2018-based population projections, technical note, 2019.
- [4] B. Franklin. *Observations concerning the Increase of Mankind, Peopling of Countries, etc.* reprinted William Abbatt, Tarrytown, New York, 1918.
- [5] R. J. Hyndman and G. Athanasopoulos. *Forecasting: principles and practice, 2nd edition*. OTexts, Melbourne, Australia, 2018.
- [6] D. Kwiatkowski, P. C. B. Phillips, P. Schmidt, and Y. Shin. Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *Journal of Econometrics*, 54(1-3):159–178, 1992.
- [7] R. Lapinskas. *Trumpas matematinės demografijos kursas*. Vilniaus universiteto leidykla, Vilnius, Lietuva, 1998.
- [8] R. Leipus. Laiko eilutės. <http://klevas.mif.vu.lt/~remis/>, 2019. Žiūrėta: 2019-11-25.
- [9] Oficialiosios statistikos portalas. <http://osp.stat.gov.lt/>. Žiūrėta: 2019-11-05.
- [10] B. Pengra. One planet, How Many People? A Review of Earth’s Carrying Capacity. [https://na.unep.net/geas/archive/pdfs/geas\\_jun\\_12\\_carrying\\_capacity.pdf](https://na.unep.net/geas/archive/pdfs/geas_jun_12_carrying_capacity.pdf), 2012.
- [11] H. T. Pham and B.-S. Yang. Estimation and forecasting of machine health condition using arma/garch model. *Mechanical Systems and Signal Processing*, 24(2):546–558, 2010.
- [12] M. Pidwirny. Introduction to the oceans. *Fundamentals of Physical Geography, 2nd Edition*, 2006. <http://www.physicalgeography.net/fundamentals/8o.html>. Žiūrėta: 2019-12-11.
- [13] R. H. Shumway and D. S. Stoffer. *Time series analysis and its applications: with R examples*. Springer, New York, USA, 2017.
- [14] A. Tsoularis and J. Wallace. Analysis of logistic growth models. *Mathematical biosciences*, 179(1):21–55, 2002.



- [15] United Nations, Department of Economic and Social Affairs, Population Division. *World Population Prospects 2019: Highlights*. United Nations, New York, USA, 2019.
- [16] World population by year. <https://www.worldometers.info/faq/>. Žiūrēta: 2019-12-10.