Idősorelemzés gyakorlat

Biztosítási és pénzügyi matematika mesterszak
Matematikus / alkalmazott matematikus mesterszak
2022/2023 tavaszi félév

Játékszabályok

- Legfeljebb 4 gyakorlatról lehet hiányozni. Ez fölött nem kapsz gyakjegyet.
- 100 + x pontot lehet szerezni a félév során:
  - 50 pont: számos, programozást is igénylő beadandó feladat, amiket a félév során hirdetek ki a honlapomon a beadandók linken
  - 50 pont: ZH: V.22.
  - x pont: szorgalmi feladatok – **SZ** jelű példák, fix határidőig
- Tudnivalók a beadandókról
  - Összességében teljesíteni kell 40%-ot, azaz 20 pontot. Ha ezt nem éred el, nem kapsz aláírást.
  - A beadandóknál elvárás az önálló munkavégzés. Amennyiben nyilvánvaló másolás gyanúja merül fel, az érintettek 0 pontban részesülnek.
- Tudnivalók a ZH-ról
  - Minimálisan teljesíteni kell 30 %-ot, azaz a 15 pontot. Ha a ZH sikertelen, nem írod meg, vagy javítani szeretnél, akkor vizsgaidőszak első hetén lesz lehetőség a pótZH megírására vagy a javításra. A jobbik eredményt veszem figyelembe, azaz nem lehet rontani. Sikertelen pótZH esetén gyakUV-t kell írnod, a beadandókért kapott pontok megmaradnak.
  - A ZH-n használható: órai és egyéb jegyzetek, minden online elérhető anyag. Nem használható: hallgatótársaid, más emberek online segítsége.
- Osztályozás:

elégtelen (1)	0	–	34,99
elégséges (2)	35	–	49,99
közepes (3)	50	–	64,99
jó (4)	65	–	79,99
jeles (5)	80	–	∞^3

Infók a gyakvezetőről

Név	Varga László, *óraadó*
Munkahely	Citi, MQA
Tanszék	Valószínűségelméleti és Statisztika Tanszék (ELTE TTK)
E-mail	`vargala4@gmail.com`
Honlap	`vargal4.elte.hu`

Ajánlott irodalom

- Shumway-Stoffer: Time series analysis and its applications, with R examples
  Egy lebutított verziója ingyenesen letölthető innen
- Brockwell-Davis: Introduction to time series and forecasting

---

A félév során idősormodellezésre/szimulációkhoz használt programnyelv: **R**

- Statisztikai modellezésre, adatok elemzésére kiválóan alkalmas programnyelv
- Nyílt forráskódú, ma már alig van probléma, feladat, aminek a megoldására ne lenne valamilyen package – akár több is
- Jelenleg a legelterjedtebb matematikai célú programnyelv

- R nyelv/szoftver letöltési helye
- Szövegszerkesztésre ajánlott szoftver: RStudio; letöltési helye
- R MarkDown: egyszerű szövegszerkesztő és kódot is futtató package, rövid bevezető: `https://rmarkdown.rstudio.com/lesson-1.html`

Amennyiben az adott feladat nem jelzi külön, t mindig az egész számok halmazát futja be, azaz $t \in \mathbb{Z}$.

1.) Legyen X_t *véletlen bolyongás drift*-tel: $X_t = \delta + X_{t-1} + \varepsilon_t \quad (t = 1, 2, \ldots)$, ahol δ valós paraméter, $P(X_0 = 0) = 1$ és $\varepsilon_t \sim WN(0, \sigma^2)$.

a.) Fejezzük ki X_t-t a fehér zaj folyamat segítségével!

b.) Határozzuk meg X_t várható érték, autokovariancia és autokorreláció függvényét! Gyengén stacionárius a folyamat?

c.) Szimuláljunk egy ilyen folyamatot normális innovációk, $\delta = -1, 0, 1$, valamint $\sigma = 1, 5, 10$ esetén, majd ábrázoljuk X_t-t, a trendet és az autokorreláció függvényét!

d.) Van-e olyan $\varphi(X_t, X_{t-1}, \ldots)$ transzformáció, amivel származtatott folyamat már stacionárius?

e.) Tegyük fel, hogy a ε_t fehér zaj folyamat normális eloszlású. Határozd meg a δ paraméter és a fehér zaj σ^2 szórásnégyzet paraméterének ML-becslését az első n megfigyelés, mint n elemű minta alapján! Torzítatlan, illetve konzisztens a δ és σ paraméterekre adott becslés?

2.) Legyen $X_t = \frac{1}{4}\varepsilon_{t-1} + \frac{1}{2}\varepsilon_t + \frac{1}{4}\varepsilon_{t+1} \quad (t = 1, 2, \ldots)$, ahol ε_t fehér zaj.

a.) Határozzuk meg X_t várható érték, autokovariancia és autokorreláció függvényét! Gyengén stacionárius a folyamat?

b.) Szimuláljunk egy fehér zajt, készítsük el hozzá X_t-t, majd ábrázoljuk a két idősort közös diagramban! Értelmezzük a látottakat!

3.) Legyen $X_t = \alpha + \beta t + \varepsilon_t \quad (t = 1, 2, \ldots)$, ahol α, β valós paraméterek és ε_t fehér zaj.

a.) Határozzuk meg X_t várható érték, autokovariancia és autokorreláció függvényét! Gyengén stacionárius a folyamat?

b.) Van-e olyan $\varphi(X_t, X_{t-1}, \ldots)$ transzformáció, amivel származtatott folyamat már stacionárius?

4.) Legyen $(X_t)_{t \in \mathbb{R}} = U \sin(2\pi\alpha t) + V \cos(2\pi\alpha t)$, ahol α valós paraméter, U és V egymástól független, 0 várható értékű és τ szórású val. változók.

a.) Határozzuk meg X_t várható érték, autokovariancia és autokorreláció függvényét! Gyengén stacionárius a folyamat?

b.) Legyen $Y_t = X_t + \varepsilon_t$, ahol ε_t fehér zaj. Szimuláljunk egy fehér zajt $D\varepsilon_t = 1, 2, 4, 8$ esetén, készítsük el hozzá Y_t-t, ha $\tau = 2$ és $\alpha = 1$, majd ábrázoljuk X_t és Y_t idősorokat! Értelmezzük a látottakat!

5.) Legyen $X_t = \sin(2\pi U t)$ $(t = 1, 2, \ldots)$, ahol $U \sim E(0;1)$. Mutassuk meg, hogy X_t folyamat gyengén stacionárius, de erősen *nem* stacionárius!
Útmutatás az erős stacionaritás vizsgálatához: számoljuk ki a következő két valószínűséget: $P\left(X_1 < -\frac{\sqrt{3}}{2}, X_2 < -\frac{\sqrt{3}}{2}\right)$ és $P\left(X_2 < -\frac{\sqrt{3}}{2}, X_3 < -\frac{\sqrt{3}}{2}\right)$!

SZ1.) [III.20.] Legyen $(X_t)_{t\in\mathbb{R}}$ stacionárius Gauss-folyamat (azaz minden $k \in \mathbb{Z}_+$-ra és $t_1, \ldots, t_k \in \mathbb{R}$-re $(X_{t_1}, \ldots, X_{t_k})$ együttesen normális eloszlású) és $Y_t = e^{X_t}$.

a.) Határozd meg az X_t folyamat momentumgeneráló függvényét, azaz $M_{X_t}(s) = E(e^{sX_t})$-t!

b.) Mutasd meg, hogy az Y_t folyamat gyengén stacionárius! (1+1= 2p)

SZ2.) [III.20.] Legyen $X_t = at + W_t$ $(0 \le t \in \mathbb{R})$, ahol a ismeretlen valós paraméter, W_t Wiener-folyamat. Határozd meg az a paraméter ML-becslését egy azonos időközönként rendelkezésünkre álló n elemű minta alapján, azaz legyen $\delta > 0$ fix, $t_i - t_{i-1} = \delta$ $(i = 1, 2, \ldots, n)$, $t_0 = 0$; a tapasztalati mintánk pedig $x_{t_1}, x_{t_2}, \ldots, x_{t_n}$ – azonos időközönként "nézünk rá" a folyamatra. (2p)
Útmutatás: Könnyen látható, hogy $X_{t_j} = X_{t_{j-1}} + a\delta + W_{t_j} - W_{t_{j-1}}$ $(j = 1, 2, \ldots, n)$, ebből pedig már megállapítható az $X_{t_j} | X_{t_{j-1}}$ feltételes eloszlás.

6.) Tegyük fel, hogy ε_t független értékű zaj 0 várható értékkel és $\sigma_\varepsilon = 1$ szórással. Vizsgáljuk meg, hogy az alábbi folyamatok okozati folyamatok-e (azaz van stacionárius $MA(\infty)$ reprezentációjuk), továbbá, hogy invertálhatók-e (azaz létezik $AR(\infty)$ reprezentációjuk)!

a.) $X_t = 2 + 0,9X_{t-1} + \varepsilon_t$
b.) $X_t = 0,2X_{t-1} + 0,35X_{t-2} + \varepsilon_t$
c.) $X_t = 0,5X_{t-1} - 0,25X_{t-2} + \varepsilon_t$
d.) $X_t = X_{t-1} - 0,25X_{t-2} + \varepsilon_t$
e.) $X_t = \varepsilon_t + 0,5\varepsilon_{t-1}$
f.) $X_t = 0,5X_{t-1} + \varepsilon_t - 0,5\varepsilon_{t-1}$.
g.) $X_t = 0,9X_{t-1} + \varepsilon_t + 0,5\varepsilon_{t-1}$.

Határozzuk meg a következőket:

- a stacionárius megoldás várható értéke, szórása, autokorreláció-függvénye;
- a folyamat "inverze" ($MA(\infty)$ reprezentációja);
- a folyamat spektrális sűrűségfüggvénye.

Szimuláljunk ilyen folyamatokat **R**-rel! Határozzuk meg **R** függvények segítségével is az autokorrelációkat, parciális autokorrelációkat és ábrázoljuk a spektrális sűrűségfüggvényt!

7.) Generáljunk egy 1000 elemű mintát az $X_t = \omega + \alpha \cdot X_{t-1} + \varepsilon_t$ folyamatból, ahol $\omega = 2$, $\alpha = 0,6$ és $\varepsilon_t \sim IVN(0;1)$. Becsüljük vissza az ω és α paramétereket

- az `arima` függvénnyel különböző 'method'-ok esetén (ML-becslés, legkisebb négyzetes becslés);
- az `lm` függvénnyel, azaz lineáris regresszióként tekintve a feladatra.

Hasonlítsuk össze a becsléseket és értelmezzük a kapottakat! Próbáljunk magyarázatot találni a kapott furcsaságra!

8.) Tekintsük a 6. feladat b.) részében definiált idősort. Generáljunk $n = 100, 500, 1000, 5000, 10000$ elemű mintákat, majd becsüljük vissza az ismeretlen paramétereket

a.) regressziós (lineáris) modell illesztésével;
b.) momentum módszerrel (Yule-Walker egyenletek segítségével);
c.) maximum likelihood módszerrel – tegyük fel, hogy a hibák normális eloszlást követnek.

Értelmezzük a látottakat és vessük össze az **R** beépített paraméterbecslésével!

9.) Legyen $X_t = \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \varepsilon_t$, ahol ε_t véges szórású i.i.d. zaj. Bizonyítsuk be, hogy a folyamatnak akkor és csak akkor van stacionárius megoldása, ha $\alpha_1 + \alpha_2 < 1$, $\alpha_2 - \alpha_1 < 1$ és $|\alpha_2| < 1$ teljesül!

10.) Generáljunk egy zérus várható értékű, $\alpha_1 = 0,5$ és $\alpha_2 = 0,3$ paraméterű AR(2) folyamatból $n = 200$ elemű mintákat.

a.) Becsüljük vissza a minták paramétereit, majd végezzünk modelldiagnosztikát – vizsgáljuk meg, mennyire jó a becslés, illetve mennyire jó a modell!

b.) Játsszunk el azzal, hogy más modelleket illesztünk (például AR(1), AR(3), ARMA(1,1)), honnan vesszük észre, hogy ezek nem lesznek jók – ha egyáltalán észrevesszük.

c.) Vizsgáljuk meg szimulációval, hogy miként teljesít az AIC és a BIC információs kritérium a helyes rend kiválasztásában! Összehasonlításként nézzük meg az AIC/BIC értékeket, ha AR(1), AR(3), AR(4), ARMA(1,1) modelleket illesztünk a fenti AR(2)-ből generált mintákra! Mi történik, ha növeljük a mintaelemszámot?

11.) Elemezzük és illesszünk alkalmas ARMA modellt a 'gnp96.txt' fájlban található idősorra, ami az USA negyedéves, szezonális hatásoktól megszabadított reál (1996-os árakon számított) GNP adatait tartalmazza 1947 és 2002 között.

a.) Ábrázoljuk az idősort, és nézzük meg az ACF és PACF függvényét! Stacionáriusnak tűnik az idősor?

b.) Alkalmas transzformációval (logdifferenciák képzése) és/vagy regressziós modell illesztésével próbáljuk meg stacionáriussá tenni az idősort!

c.) A transzformált idősor ACF és PACF ábráinak szemrevételezése után próbáljunk meg alkalmas ARMA modellt illeszteni! AIC/BIC kritériumok alapján válasszuk ki a legjobbnak tűnő ARMA modellt, majd végezzünk modelldiagnosztikát (szignifikánsak-e a becsült paraméterek, a reziduálisok származhatnak-e fehér zaj folyamatból)!

12.) Az előző példához hasonlóan elemezzük és illesszünk alkalmas ARMA modellt az 'astsa' package-ben lévő 'gtemp' idősorra, ami 1880 és 2009 között a Föld átlaghőmérsékletét tartalmazza.

13.) Vizsgáljuk meg a különbség-előjel próbát! Legyen $X_1, \dots, X_n$ minta egy stacionárius folyamatból, és legyen Z_n annak a száma, ahányszor a folyamat értéke növekszik, azaz $Z_n = \sum_{i=2}^{n} Y_i$, ahol $Y_i = I(X_i > X_{i-1})$, $i = 2, 3, \dots, n$.

a.) Az összefüggő valószínűségi változó sorozatokra vonatkozó centrális határeloszlás-tétel segítségével mutassuk meg, hogy $\frac{Z_n - \frac{n-1}{2}}{\sqrt{\frac{n-1}{12}}} \xrightarrow[n\to\infty]{d} N(0;1)$!

b.) Generálj i.i.d. sorozatot $n = 10, 50, 100, 200, 500, 1000, 2000$ mintanagyságok esetén, majd szimulációk alapján becsüld meg az elsőfajú hiba valószínűségét!

c.) Generálj
- AR(1) folyamatból;
- driftes véletlen bolyongásból

mintákat $n = 10, 50, 100, 200, 500, 1000, 2000$ mintanagyságok esetén, majd szimulációk alapján becsüld meg a különbség-előjel próba erejét!

14.) Legyen X_t BL(1,0,1,1) folyamat, azaz $X_t = aX_{t-1} + \varepsilon_t + cX_{t-1}\varepsilon_{t-1}$, ahol $a, c \in \mathbb{R}$, $\varepsilon_t \sim i.i.d.\ N(0, \sigma^2)$.

a.) Mutasd meg, hogy amennyiben stacionárius a folyamat, akkor
- $M_1 = EX_t = \frac{c\sigma^2}{1-a}$
- $M_2 = EX_t^2 = \frac{2c^2\sigma^4 + \sigma^2 + 4ac\sigma^4 \frac{c}{1-a}}{1-a^2-c^2\sigma^2}$
- $R(1) = cov(X_t, X_{t-1}) = a \cdot \left[\frac{2c^2\sigma^4 + \sigma^2 + 4ac\sigma^4 \frac{c}{1-a}}{1-a^2-c^2\sigma^2} - \frac{c^2\sigma^4}{(1-a)^2}\right] + \frac{c^2\sigma^4}{1-a}$
- $R(k) = cov(X_t, X_{t-k}) = a^{k-1} \cdot R(1) \quad k = 2, 3, \dots$

b.) Szimulálj a folyamatból a paraméterek különböző értékei esetén! Vizsgáld meg a tapasztalati korreláció és autokorreláció függvényeket! Becsüld vissza az a és c együtthatókat momentum és ML-módszerrel!

15.) Legyen X_t ARCH(1) folyamat, azaz $X_t = \sigma_t\varepsilon_t$, ahol $\sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2$ és ε_t gauss-i fehér zaj.

a.) Határozd meg az $X_t|\mathcal{F}_{t-1}$ feltételes eloszlást, ahol $\mathcal{F}_t = \sigma(X_s : s \le t)$! Ez alapján határozd meg az EX_t értéket!

b.) Határozd meg a $\text{cov}(X_{t+h}, X_t)$ mennyiséget, ha $h > 0$!

c.) Mutasd meg, hogy X_t^2 AR(1) folyamat!

d.) Mutasd meg, hogy $D^2X_t = \frac{\alpha_0}{1-\alpha 1}$ és $EX_t^4 = \frac{3\alpha_0^2(1-\alpha_1^2)}{(1-\alpha_1)^2(1-3\alpha_1^2)}$! Milyen paraméterértékek esetén léteznek ezek a mennyiségek?

e.) Határozzuk meg a paraméterek momentum becslését!

f.) Határozzuk meg a paraméterek ML-becslését!

Szimuláljunk ARCH(1) folyamatokat **R**-rel és az egyes feladatrészek eredményeit is vizsgáljuk meg! Becsüljük vissza a paramétereket!

16.) Szimuláljunk 1000 elemű mintát GARCH(1,1) folyamatból, ha a paraméterek $\alpha_0 = 0,01$, $\alpha_1 = 0,2$, $\beta = 0,75$, az innovációk pedig standard normális eloszlásúak. Illesszünk ARMA(1,1), ARCH(1) és GARCH(1,1) modelleket a mintára, majd értékeljük a kapott eredményeket!

17.) Alkalmas transzformációkat követően illesszünk megfelelő GARCH modellt a(z)

a.) NYSE Composite index (záróárfolyamok);

b.) EUR-HUF devizaárfolyam (közép)

idősorra! Vizsgáljuk meg, mennyire jó a modell!

18.) A Ljapunov-exponens segítségével állapítsd meg, hogy az alábbi folyamatok mely paraméterértékek esetén erősen stacionáriusak!

a.) $X_t = aX_{t-1} + \varepsilon_t$, ahol $a \in \mathbb{R}$, $\varepsilon_t \sim WN(0, \sigma^2)$

b.) $X_t = \sigma_t\varepsilon_t$, ahol $\sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2$, $\alpha_0, \in \mathbb{R}_+$, $\alpha_0, \in \mathbb{R}_+$, ε_t független értékű $N(0,1)$ eloszlású zaj

c.) $X_t = \sigma_t\varepsilon_t$, ahol $\sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2$, $\alpha_0 \in \mathbb{R}_+$, $\alpha_1 \in \mathbb{R}_+$, ε_t független értékű $E(0,1)$ eloszlású zaj

SZ3.) **[IV.3.]** Legyen $X_t = 0,4X_{t-1} + 0,45X_{t-2} + \varepsilon_t + \varepsilon_{t-1} + 0,25\varepsilon_{t-2}$, ahol ε_t fehér zaj. Van stacionárius "megoldása" (MA reprezentációja) a folyamatnak? Ha van, akkor állítsd elő! (2p)

SZ4.) **[IV.3.]** Mutasd meg, hogy az $X_t = \varepsilon_t + \beta\varepsilon_{t-1}$ MA(1) folyamat parciális autokorreláció függvénye $\rho_k = \frac{-(-\beta)^k}{1+\beta^2+\dots+\beta^{2k}}$. (2p)

SZ5.) **[IV.17.]** Legyen $X_1, \dots, X_n$ fehér zajból származó minta. Mutasd meg, hogy

a.) a Bartlett-tételben (lásd elméleti összefoglaló) szereplő $\mathbf{W}$ mátrix egységmátrix!

b.) Készíts konfidenciaintervallumot a $\hat{r}(h)$, $h \in \mathbb{Z}$ empirikus autokorrelációra! (1+1= 2p)

SZ6.) **[IV.17.]** Bizonyítsd be, hogy a fordulópont próba (turning point test) próbastatisztikája eloszlásban a standard normális eloszlásoz tart, amennyiben a minta i.i.d. fehér zajból származik, azaz teljesül a nullhipotézis! (2p)

19.) Legyen $(X_n)_{n\in\mathbb{Z}}$ stacionárius folyamat μ várható értékkel és $R(h)$ autokovariancia függvénnyel. Adjunk $0 < h \in \mathbb{Z}$ lépéses legkisebb négyzetes lineáris előrejelzést X_n segítségével, azaz határozzuk meg azon a és b értékeket, amikre az $MSE(a,b) := E([X_{n+h} - (a + bX_n)]^2)$ minimális! Határozzuk meg az így adódó $\hat{X}_{n+h} = \hat{a} + \hat{b}X_n$ becslés szórását!

20.) Legyen X_t stacionárius AR(1) folyamat 0 várható értékkel és σ szórású innovációkkal.

a.) Határozzuk meg a folyamat egylépéses, majd h lépéses legkisebb négyzetes előrejelzését és adjuk meg az előrejelzés hibáját! Vessük össze a kapottakat a 19.) feladat eredményével!

b.) Tegyük fel, hogy ismerjük a folyamat 1. és 3. megfigyelését, de a 2. megfigyelés hiányzik. Adjunk legkisebb négyzetes lineáris becslést X_2-re X_1 és X_3 segítségével, és határozzuk meg a becslés hibáját!

21.) Legyen X_t stacionárius MA(1) folyamat 0 várható értékkel és σ szórású innovációkkal. Tegyük fel, hogy ismerjük a folyamat 1. és 2. megfigyelését, de az 3. megfigyelés hiányzik. Adjunk legkisebb négyzetes lineáris becslést X_3-ra X_1 és X_2

segítségével, és határozzuk meg a becslés hibáját!

22.) Legyen X_t stacionárius AR(p) folyamat 0 várható értékkel és σ szórású innovációkkal. Határozzuk meg a folyamat egylépéses, legkisebb négyzetes előrejelzését!

23.) Legyen X_t stacionárius AR(1) folyamat μ várható értékkel és σ szórású innovációkkal. Az előrejelzési operátorok tulajdonságait felhasználva, határozzuk meg a folyamat h lépéses, legkisebb négyzetes előrejelzését!

24.) Határozzuk meg a legkisebb négyzetes legjobb előrejelzést a 'gnp96.txt' fájlban található idősor következő 10 értékére a legjobbnak bizonyult ARMA modell alapján a beépített `predict` függvény segítségével! Adjunk intervallumbecslést is az előrejelzésre! Ábrázoljuk az idősort az előrejelzésekkel együtt!

SZ7.) **[V.8.]** Legyen X_t AR(1) folyamat α paraméterrel és 1 szórású innovációval. Mutasd meg, hogy ekkor $D^2(\overline{X}_n) = \frac{\sigma^2}{n^2(1-\alpha^2)} \frac{n-n\alpha^2-2\alpha+2\alpha^{n+1}}{(1-\alpha)^2}$. (1 pont)

SZ8.) **[V.8.]** Legyen X_t stacionárius MA(1) folyamat 0 várható értékkel és σ szórású innovációkkal. Tegyük fel, hogy ismerjük a folyamat 1., 2., 4. és 5. megfigyelését, de a 3. megfigyelés hiányzik. Adjunk legkisebb négyzetes lineáris becslést X_3-ra X_1, X_2, X_4 és X_5 segítségével, és határozzuk meg a becslés hibáját! (1 pont)

25.) Tekintsük az $X_t = \alpha + \beta t + \varepsilon_t \quad (t = 1, 2, \ldots)$ folyamatot, ahol α, β valós paraméterek és ε_t fehér zaj. Milyen típusú ARIMA folyamat ez?

26.) Tekintsük az $X_t = \omega + \alpha X_{t-1} + \varepsilon_t$ AR(1) folyamatot, ahol $\varepsilon_t \sim WN(0, \sigma^2)$. Milyen folyamatot kapunk, ha 1-szer, 2-szer, stb. differenciáljuk?

27.) Tekintsük az alábbi ARIMA(1,1,0) modellt: $(1-\phi B)(1-B)X_t = \varepsilon_t$, ahol $\varepsilon_t \sim WN(0, \sigma^2)$, $|\phi| < 1$, $DX_0 < \infty$, $t = 1, 2, \ldots$

a.) Oldjuk meg az ARIMA-egyenletet, azaz fejezzük ki az X_t-t a zaj segítségével!

b.) Szimuláljunk ilyen folyamatból és nézzük meg az autokorreláció függvényt! Becsüljük vissza a paramétereket! Teszteljük a modellt egységgyök teszttel is!

28.) Az általános lineáris modell paraméterének legkisebb négyzetes becslését felhasználva mutasd meg, hogy az $y_i = ax_i + b, \quad i = 1, 2, \ldots, n$ egyváltozós lineáris regressziós modell esetén az együtthatók legkisebb négyzetes becslése

$$\hat{a} = \frac{\sum_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y})}{\sum_{i=1}^{n}(x_i-\overline{x})^2} = \frac{\left(\sum_{i=1}^{n} x_i y_i\right) - n\cdot\overline{x}\cdot\overline{y}}{\left(\sum_{i=1}^{n} x_i^2\right) - n\cdot\overline{x}^2} \quad \text{és} \quad \hat{b} = \overline{y} - \hat{a}\overline{x}.$$

29.) Egy fagylaltárus negyedéves forgalmának alakulása (ezer gombóc):

Év	I. negyedév	II. negyedév	III. negyedév	IV. negyedév
2017	95	152	255	118
2018	102	146	248	124
2019	97	156	245	122

Jelölje az idősor elemeit $(x_t)_{t=1,2,\ldots,12}$. Néhány számítási eredmény:

$$\sum_{t=1}^{12} x_t = 1860 \qquad \sum_{t=1}^{12} t \cdot x_t = 12342 \qquad \sum_{t=1}^{12} t = 78 \qquad \sum_{t=1}^{12} t^2 = 650$$

a.) Határozd meg a lineáris trend egyenletét és értelmezd a paramétereket!

b.) Határozd meg a szezonális eltéréseket és értelmezd őket!

c.) Adj pont- és intervallumbecslést a 2020. évi eladási mennyiségekre (mind a 4 negyedévre)!

30.) Szabadítsuk meg az *astsa* package-ben lévő *jj* (Johnson&Jonhson) 1960 és 1980 közötti negyedéves adatsort a trendtől és a szezonális hatásoktól

a.) szezonindexek számolásával;

b.) az **R** nyelv *stl* parancsa (seasonal decomposition of time series by LOESS regression) segítségével;

c.) alkalmas differenciák képzésével;

d.) SARIMA modell illesztésével.

Illesszünk alkalmas ARMA modellt a reziduálisokra, majd adjunk előrejelzést az 1981-es év mind a 4 negyedévére!

31.) Legyen Y_t stacionárius folyamat 0 várható értékkel és legyenek a és b valós konstansok. Ha $X_t = a + bt + s_t + Y_t$, ahol s_t szezonális komponens 4 periódussal, akkor mutassuk meg, hogy $\nabla\nabla_4 X_t = (1-B)(1-B^4)X_t$ folyamat stacionárius, és fejezzük ki az autokovariancia függvényét Y_t folyamat $R_Y(h)$ autokovariancia függvényével!

SZ9.) **[V.22.]** Tekintsük a 27. feladatban szereplő ARIMA modellt. Számold ki az autokorreláció függvényt! (2p)

32.) Legyenek X_t és Y_t egymással korrelálatlan stacionárius folyamatok R_X és R_Y autokovariancia függvénnyel, valamint F_X és F_Y spektrális eloszlásfüggvénnyel. Mutassuk meg, hogy ekkor a $Z_t = X_t + Y_t$ folyamat autokovariancia függvényére $R_Z(h) = R_X(h) + R_Y(h)$ és spektrális eloszlásfüggvényére $F_Z(x) = F_X(x) + F_Y(x)$ teljesül minden $h \in \mathbb{Z}$ és $x \in \mathbb{R}$ esetén.

33.) Határozzuk meg a következő folyamatok spektrális sűrűségfüggvényét, ha pedig nincsen, akkor a spektrális eloszlásfüggvényét:

a.) fehér zaj folyamat σ szórással;

b.) AR(1) folyamat: $X_t = \alpha X_{t-1} + \varepsilon_t$, ahol $1 > |\alpha| \in \mathbb{R}$, $\varepsilon_t \sim WN(0, \sigma^2)$;

c.) MA(1) folyamat: $X_t = \varepsilon_t + \alpha\varepsilon_{t-1}$, ahol $\alpha \in \mathbb{R}$, $\varepsilon_t \sim WN(0, \sigma^2)$;

d.) $X_t = U\sin(\alpha t) + V\cos(\alpha t)$, ahol α valós paraméter, U és V egymástól független, 0 várható értékű és τ szórású valószínűségi változók;

e.) $X_t = S_t + \alpha S_{t-D} + N_t$, ahol S_t és N_t stacionárius és egymástól független folyamatok 0 várható értékkel és f_S, valamint f_N spektrális sűrűségfüggvényekkel; D ismert pozitív egész szám; α ismeretlen valós paraméter.

Szimuláljunk az egyes folyamatokból különböző paraméterértékek esetén, ábrázoljuk a mintát, az autokorreláció függvényt, valamint a spektrális sűrűségfüggvényt (eloszlásfüggvényt)!

34.) Legyen $X_t = U\sin\left(\frac{\pi}{3}t\right) + V\cos\left(\frac{\pi}{3}t\right)$, ahol U és V egymástól független, 0 várható értékű és τ szórású valószínűségi változók; $Y_t = \varepsilon_t + 2{,}5\varepsilon_{t-1}$, ahol ε_t

fehér zaj σ szórással; továbbá legyen $Z_t = X_t + Y_t$. Határozd meg a Z_t folyamat autokovariancia függvényét és a spektrális eloszlásfüggvényét!

35.) Döntsük el, hogy az alábbi egész számokon értelmezett függvények lehetnek-e egy stacionárius folyamat autokovariancia függvényei:
a.) $R(h) = I(h = 0)$;
b.) $R(h) = I(h = 0) - 0,5 \cdot I(|h| = 2) - 0,25 \cdot I(|h| = 3)$;

36.) Legyen egy stacionárius folyamat spektrális sűrűségfüggvénye $f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$, $x \in [-\pi; \pi]$. Határozd meg a folyamat autokovariancia és autokorreláció függvényét!
Útmutatás: integráljuk parciálisan az autokovariancia függvény deriváltját, majd oldjuk meg az így adódó $R'(h) = -hR(h)$ differenciálegyenletet!

37.) Egy ARMA folyamat spektrális sűrűségfüggvénye a következő: $f(x) = \frac{2+2\cos x}{\pi(1,5625-2,5\cos x+\cos^2 x)}$. Határozd meg a folyamat rendjét és paramétereit!

SZ10.) **[V.22.]** Az egész számokon értelmezett $R(h) = I(h = 0) + cI(|h| = 1)$ függvény a c valós paraméter mely értékei esetén lehet egy stacionárius folyamat autokovariancia függvénye? (1p)

SZ11.) **[V.22.]** Legyenek X_t és Y_t stacionárius, egymástól független, 0 várható értékű folyamatok f_X és f_Y spektrális sűrűségfüggvényekkel, továbbá legyen $Z_t = X_t \cdot Y_t$. Mutasd meg, hogy a Z_t folyamat spektrális sűrűségfüggvénye a következőképp számolható: $f_Z(x) = \int\limits_{-\pi}^{\pi} f_X(x-y) f_Y(y)\, dy$. (2p)

SZ12.) **[V.22.]** Egy stacionárius X_t folyamat spektrális sűrűségfüggvénye a következő: $f(x) = 100 \cdot I\left(-\frac{\pi}{6} - 0,01 < x < -\frac{\pi}{6} + 0,01\right) + 100 \cdot I\left(\frac{\pi}{6} - 0,01 < x < \frac{\pi}{6} + 0,01\right)$. Határozd meg a folyamat lag 1 (egylépéses késleltetésű) autokorrelációját, azaz az $r(1) = \mathrm{cor}(X_t, X_{t+1})$ mennyiséget! (2p)