# Idősorelemzés gyakorlat

Biztosítási és pénzügyi matematika mesterszak
Alkalmazott matematikus mesterszak
2019/2020 tavaszi félév

## Játékszabályok

- A gyakorlatokról maximum 3-szor lehet hiányozni. Aki többször hiányzik, nem kap gyakjegyet.
- 100 + x pontot lehet szerezni a félév során:
  - · 32 pont: kis beadandó feladatok (4p · 8) – **B** jelű példák
  - · 28 pont: nagy beadandó feladat
  - · 40 pont: ZH: V.12., D 0-803 Szabó József előadó
  - · x pont: szorgalmi feladatok – **SZ** jelű példák
- A ZH-n minimálisan teljesíteni kell 30 %-ot. 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. GyakUV írása esetén legfeljebb 2-est lehet szerezni.
- A ZH-kon használható: **számológép** és egy legfeljebb A4-es méretű lapra KÉZZEL írott "puska".
- Kis beadandók: Mindegyik maximálisan 4 pontot ér, a legjobb 8-at veszem figyelembe. A beadandóknál több feladatot is kihirdetek, amik közül ízlés szerint válogathattok. A beadandók célja, hogy folyamatosan tanuljatok, gyakoroljatok, ezért fix határidőig lehet őket benyújtani.
- Nagy beadandó: fix határidővel hirdetem ki
- Szorgalmik: a pótZH időpontjáig bármikor, bármennyit be lehet adni.
- A beadandóknál nem tilos, sőt, bizonyos fokig még kívánatos is a közös munka/konzultáció hallgatótársaiddal – az viszont elvárás, hogy **a gondolataidat önállóan írd le**! Amennyiben nyilvánvaló másolás gyanúja merül fel, az érintettek 0 pontban részesülnek.
- Osztályozás:

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

## Infók a gyakvezetőről



| | |
|---|---|
| Név | Varga László, *óraadó* |
| Munkahely | Morgan Stanley, Model Risk Management |
| Tanszék | Valószínűségelméleti és Statisztika Tanszék (ELTE TTK) |
| E-mail | `vargal4@cs.elte.hu` |
| Honlap | `vargal4.elte.hu` |



## Ajánlott irodalom

- Shumway-Stoffer: Time series analysis and its applications, with R examples
  Egy lebutított verziója ingyenesen letölthető: `http://www.stat.pitt.edu/stoffer/tsa4/tsaEZ.pdf`
- Brockwell-Davis: Introduction to time series and forecasting

---

Szimulációkhoz használt szoftver/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

- Letöltési helye: `https://cran.r-project.org/`
- Szövegszerkesztésre ajánlott szoftver: RStudio; letöltési helye: `https://www.rstudio.com/products/rstudio/download3/`

---

*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 $\delta$ 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 $\varepsilon_t$ fehér zaj folyamat normális eloszlású. Határozd meg a $\delta$ paraméter és a fehér zaj $\sigma^2$ szórásnégyzet paraméterének ML-becslését az első $n$ megfigyelés, mint $n$ elemű minta alapján! Torzítatlan, illetve konzisztens a $\delta$-ra 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 $\varepsilon_t$ 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 $\alpha, \beta$ valós paraméterek és $\varepsilon_t$ 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 $\alpha$ valós paraméter, $U$ és $V$ egymástól független, 0 várható értékű és $\tau$ 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 $\varepsilon_t$ 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 $X_t = \sin(2\pi Ut)$ $(t = 1, 2, \ldots)$, 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 $P\left(X_2 < -\frac{\sqrt{3}}{2}, X_3 < -\frac{\sqrt{3}}{2}\right)$!

**B1.)** **[II.25.]** Tekintsük a következő $(X_t)_{t=1,2,\ldots}$ folyamatot: $X_t = ae^{X_{t-1}} + \varepsilon_t$, ahol $a$ valós paraméter, $\varepsilon_t$ normális eloszlású fehér zaj 0 várható értékkel és ismeretlen $\sigma$ szórással, $X_0 = x_0 \in \mathbb{R}$. Vezesd le az ismeretlen $a$ paraméter ML-becslését!

**B2.)** **[II.25.]** Tekintsük a következő $(X_t)_{t\in\mathbb{Z}}$ folyamatot: $X_t = \varepsilon_t\eta_t$, ahol $\varepsilon_t$ és $\eta_t$ egymástól és időben is független folyamatok, $\varepsilon_t$ fehér zaj 0 várható értékkel és $\sigma$ szórással, $P(\eta_t = 0) = P(\eta_t = 2) = \frac{1}{2}$ $\forall t$-re. Gyengén stacionárius az $X_t$ folyamat? Határozd meg az autokovariancia és autokorreláció függvényét!

**B3.)** **[III.3.]** Nézd végig a honlapomon található **R** bevezetőt és az alapján oldd meg **R** segítségével a következő feladatot! A kódot küldd el az E-mail címemre!
Egy magyarkártya-csomagból visszatevéssel húzunk 4 lapot. Szimulációval számold ki, hogy milyen eséllyel húzunk pontosan két zöld színű lapot! Legalább mennyi ismétlésszámot ajánlanál, hogy a valódi valószínűséget legalább 0,5%-os pontossággal közelítsük? Válaszodat a konvergencia sebességét bemutató alkalmas ábrával támaszd alá!

**B4.)** **[III.3.]** Nézd végig a honlapomon található **R** bevezetőt és az alapján oldd meg **R** segítségével a következő feladatot! A kódot küldd el az E-mail címemre!
Szimulálj 10000 darab 1000 elemű mintát exponenciális eloszlás szerint $\lambda = 2$ paraméterrel! Határozd meg az összes mintára az ML-becslést (jel. $\lambda_{ML}$)! Tanultuk bevezető statisztikából, hogy bizonyos regularitási feltételek esetén az ML-becslés aszimptotikusan normális. Ennek fényében tekints az ML-becslésekre mint mintára normális eloszlásból és becsüld meg a normális eloszlás ismeretlen paramétereit! Készíts hisztogramot 50 töréspontal a $\lambda_{ML}$ értékekből, majd ábrázold a hisztogrammal együtt a normális eloszlás sűrűségfüggvényét!

**SZ1.)** 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.)** 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 $\varepsilon_t$ 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 $\varepsilon_t \sim WN(0; 1)$. Becsüljük vissza az $\omega$ és $\alpha$ 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.)** Legyen $X_t = \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \varepsilon_t$, ahol $\varepsilon_t$ 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!

**9.)** 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?

**10.)** 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)!

**11.)** 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.

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

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!

**B5.)** **[III.10]** Tekintsük az $X_t = \alpha X_{t-1} + \alpha^2 X_{t-2} + \varepsilon_t + \beta\varepsilon_{t-1}$ folyamatot, ahol $\varepsilon_t \sim WN(0, 2^2)$. Az $\alpha \in (0;1)$ és $\beta \in (0;1)$ paraméterek mely értékeire lesz $X_t$ okozati ARMA(2;1), illetve AR(1) folyamat?

**B6.)** **[III.31]** Tekintsük a következő tapasztalati becsléseket egy 1000 elemű minta alapján:

| Lag | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|---|
| Tap.-i autokov. | 0.620 | -0.200 | 0.005 | -0.013 | 0.005 | -0.012 | -0.028 | 0.017 |
| Tap.-i parciális autokorr. | 1 | -0.323 | -0.107 | -0.060 | -0.021 | -0.028 | -0.070 | -0.017 |

Milyen modellt választanál az ARMA családból a minta modellezésére? Becsüld meg a paramétereit (az autokovariancia függvény elméleti értékei alapján)!

**B7.)** **[V.12.]** Legyen $X_t = 0,3X_{t-1} + 0,04X_{t-2} + \varepsilon_t$, ahol $\varepsilon_t$ fehér zaj 0 várható értékkel és $\sigma_\varepsilon$ véges szórással.

a.) Stacionárius/kauzális a folyamat?

b.) Invertálható a folyamat?

c.) Állítsd elő a folyamat $MA(\infty)$ reprezentációját! (1+1+2= 4p)

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

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

**SZ5.)** Legyen $X_1, \ldots, X_n$ 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.)** 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)

---

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

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

b.) Számold ki $X_t$ várható értékét!

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

d.) Határozzuk meg $X_t$ szórásnégyzetét és 4. momentumát! Milyen paraméterértékek esetén léteznek ezek a mennyiségek?

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

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

**14.)** Illesszünk megfelelő GARCH modellt a(z)

a.) NYSE;

b.) GBP-CHF;

idősorra! Vizsgáljuk meg, mennyire jó a modell!

**SZ7.)** Stacionárius ARCH(1) folyamat esetén számold ki a folyamat lapultságát, azaz a kurtosis=kurtosis$_t$ = $\frac{E(X_t - EX_t)^4}{D^4 X_t} - 3$ mennyiséget! A paraméterek mely értéke esetén lesz a folyamat csúcsosabb a normális eloszlás csúcsosságához képest? (1p)

---

**15.)** Legyen $(X_n)_{n\in\mathbb{Z}}$ stacionárius folyamat $\mu$ 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!

**16.)** Legyen $X_t$ stacionárius AR(1) folyamat 0 várható értékkel és $\sigma$ 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 15.) 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!

**17.)** Legyen $X_t$ stacionárius MA(1) folyamat 0 várható értékkel és $\sigma$ 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!

**18.)** Legyen $X_t$ stacionárius AR($p$) folyamat 0 várható értékkel és $\sigma$ szórású innovációkkal. Határozzuk meg a folyamat egylépéses, legkisebb négyzetes előrejelzését!

**19.)** Legyen $X_t$ stacionárius AR(1) folyamat $\mu$ várható értékkel és $\sigma$ szórású innovációkkal. Az előrejelzési operátorok tulajdonságait felhasználva, határozzuk meg a folyamat $h$ lépéses, legkisebb négyzetes előrejelzését!

**20.)** 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!

**B8.)** **[IV.21.]** Legyen $X_1, X_2, \ldots, X_n$ minta az $X_t = -\pi + \varepsilon_t + 2\varepsilon_{t-1} - \varepsilon_{t-3}$ folyamatból, ahol $\varepsilon_t$ fehér zaj folyamat $\sigma$ szórással.

a.) Adj a minta első két eleme, $X_1$ és $X_2$ segítségével $h$ lépéses legkisebb négyzetes előrejelzést $h = 1, 2, \ldots$ esetén! Számold ki az előrejelzés hibáját!

b.) Tegyük fel, hogy $n = 1000$. Adj a minta első két eleme és utolsó két eleme segítségével becslést $X_{100}$-ra! (2,5+1,5= 4p)

**SZ8.)** Legyen $X_t$ AR(1) folyamat $\alpha$ 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)

**SZ9.)** Legyen $X_t$ stacionárius MA(1) folyamat 0 várható értékkel és $\sigma$ 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)

---

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

**22.)** 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?

**23.)** 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!

**24.)** 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

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

**25.)** 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)!

**26.)** 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!

**27.)** 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!

**28.)** Tekintsük azt a lineáris szűrőt, melynek együtthatói $a_i = \frac{1}{1+2q}$, ha $-q \leq i \leq q$ és $a_i = 0$, ha $|i| > q$.

a.) Mutassuk meg, hogy ez a szűrő torzítás nélkül átengedi az $m_t = c_0 + c_1 t$ lineáris trendet.

b.) Amennyiben $Z_t$, $t \in \mathbb{Z}$ független, 0 várható értékű és $\sigma$ szórású valószínűségi változók, akkor számoljuk ki az $U_t = \frac{1}{2q+1}\sum_{i=-q}^{q} Z_{t-i}$ lineárisan megszűrt folyamat várható értékét és szórását! Mire következtethetünk ebből nagy $q$ esetén?

**29.)** Mutassuk meg, hogy az $(a_{-2}, a_{-1}, a_0, a_1, a_2)^T = \frac{1}{9}(-1, 4, 3, 4, -1)^T$ együtthatókkal számolt lineáris szűrő torzítás nélkül átenged minden harmadfokú polinomot, továbbá eliminálja a 3 periódusú szezonális komponenseket!

**B9.)** **[IV.28.]** A Siófoki Aranyparton egy lángosárus forgalmának alakulása (bevétel, ezer Ft) havonta:

| Év | május | június | július | augusztus | szeptember |
|---|---|---|---|---|---|
| 2015 | 513 | 1052 | 2555 | 2118 | 634 |
| 2016 | 631 | 1246 | 2848 | 2249 | 831 |
| 2017 | 459 | 968 | 2648 | 2224 | 348 |
| 2018 | 642 | 1157 | 2245 | 1822 | 730 |

Jelölje az idősor elemeit $(x_t)_{t=1,2,\ldots,20}$. Néhány számítási eredmény:

$$\sum_{t=1}^{20} x_t = 27920 \qquad \sum_{t=1}^{20} t\cdot x_t = 292781 \qquad \sum_{t=1}^{20} t = 210 \qquad \sum_{t=1}^{20} t^2 = 2870$$

a.) Határozd meg a lineáris trend egyenletét!
b.) Határozd meg a szezonális eltéréseket!
c.) Adj pontbecslést a 2019. évi eladási mennyiségekre (mind az 5 hónapra)!

**B10.)** **[IV.28]** Legyen $Y_t = -6 + \varepsilon_t - 6\varepsilon_{t-1}$, ahol $\varepsilon_t$ fehér zaj 0 várható értékkel és 2 szórással, valamint legyenek $a$ és $b$ valós konstansok. Ha $X_t = (a+bt)s_t + Y_t$, ahol $s_t$ szezonális komponens 12 periódussal, akkor mutasd meg, hogy $\nabla_{12}^2 X_t$ folyamat stacionárius, és határozd meg az autokovariancia függvényét!

**B11.)** **[V.5]** Határozd meg azt az $1 + \alpha B + \beta B^2 + \gamma B^3$ alakú lineáris szűrőt – azaz keresd meg az alkalmas $\alpha, \beta, \gamma$ valós együtthatókat –, amely torzítás nélkül átenged bármely lineáris trendet és minden 2 periódusú szezonális komponenst eliminál!

**SZ10.)** Tekintsük a 23. feladatban szereplő ARIMA modellt. Számold ki az autokorreláció függvényt! (2p)

**SZ11.)** Legyen $m_t = \sum_{i=0}^{p} c_i t^i$ polinomiális trend, ahol $0 \neq c_i \in \mathbb{R}$. Számítsd ki a $\nabla^p m_t$ és $\nabla^{p+1} m_t$ értékeket! Mire lehet ebből következtetni? (1p)

**SZ12.)** Bizonyítsd be, hogy az $\{a_i\}_{i\in\mathbb{Z}}$ együtthatókkal számolt lineáris szűrő akkor és csak akkor enged át torzítás nélkül egy tetszőleges $k$ fokú polinomot, ha $\sum_{i=-\infty}^{\infty} a_i = 1$ és minden $r = 1, 2, \ldots, k$ esetén $\sum_{i=-\infty}^{\infty} i^r a_i = 0$! (2p)

**SZ13.)** Adj meg (készíts) egy olyan lineáris szűrőt, amely torzítás nélkül átenged bármely másodfokú polinomiális trendet és minden 4 periódusú szezonális komponenst eliminál! (2p)

---

**30.)** 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.

**31.)** Határozzuk meg a következő folyamatok spektrális sűrűségfüggvényét, ha pedig nincsen, akkor a spektrális eloszlásfüggvényét:
a.) fehér zaj folyamat $\sigma$ 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 $\alpha$ valós paraméter, $U$ és $V$ egymástól független, 0 várható értékű és $\tau$ 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; $\alpha$ ismeretlen valós paraméter.

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

**32.)** 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 $\tau$ szórású valószínűségi változók; $Y_t = \varepsilon_t + 2,5\varepsilon_{t-1}$, ahol $\varepsilon_t$ fehér zaj $\sigma$ szórással; továbbá legyen $Z_t = X_t + Y_t$. Határozd meg a $Z_t$ folyamat autokovariancia függvényét és a spektrális eloszlásfüggvényét!

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

**34.)** 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!

**B12.)** **[V.12.]** Legyen $X_t = \sum_{i=1}^{N} a_i\cos(\omega_i t + U_i)$, ahol minden $i = 1, 2, \ldots, N$ esetén $\omega_i \neq 0$ és $a_i$ valós konstansok, $U_i \sim E(0; 2\pi)$ függetlenek. Határozd meg az $X_t$ folyamat spektrális eloszlásfüggvényét/sűrűségfüggvényét!

**SZ14.)** 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)

**SZ15.)** 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: $f_Z(x) = \int_{-\pi}^{\pi} f_X(x-y)f_Y(y)\,dy$. (2p)

**SZ16.)** 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 $r(1) = \operatorname{cor}(X_t, X_{t+1})$ mennyiséget! (2p)

**SZ17.)** 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! (1p)

---