Az idősorelemzés alapjai

Ferenci Tamás tamas.ferenci@medstat.hu

Utoljára frissítve: 2023. május 12.

- Az idősor fogalma, jelentősége, idősorelemzés
 - Az idősor fogalma
 - Az idősorelemzés közgazdasági jelentősége, története
 - Idősorelemzési iskolák, a módszerek felosztása

Idősorok jellemzői a sokaságban

- Az idősor fogalma, jelentősége, idősorelemzés
 - Az idősor fogalma
 - Az idősorelemzés közgazdasági jelentősége, története
 - Idősorelemzési iskolák, a módszerek felosztása
- Idősorok jellemzői a sokaságban

- Az idősor fogalma, jelentősége, idősorelemzés
 - Az idősor fogalma
 - Az idősorelemzés közgazdasági jelentősége, története
 - Idősorelemzési iskolák, a módszerek felosztása
- Idősorok jellemzői a sokaságban

- Mi a sokaság és a minta? ismétlés stat 2-ből
- ("Sokaságban valszám kell, mintánál statisztika"
- Idősor minta értelemben: időben rendezett megfigyelésel

$$\{5492, 5640, \ldots, 7317\}$$
 ,

általános jelölésse

$$\{y_1, y_2, \ldots, y_T\}$$

- Mi a sokaság és a minta? ismétlés stat 2-ből
- ("Sokaságban valszám kell, mintánál statisztika")
- Idősor minta értelemben: időben rendezett megfigyelésel

$$\{5492, 5640, \ldots, 7317\}$$
 ,

általános jelölésse

$$\{y_1, y_2, \ldots, y_T\}$$

- Mi a sokaság és a minta? ismétlés stat 2-ből
- ("Sokaságban valszám kell, mintánál statisztika")
- Idősor minta értelemben: időben rendezett megfigyelések

$$\{5492, 5640, \ldots, 7317\}$$
,

általános jelöléssel

$$\{y_1, y_2, \ldots, y_T\}$$

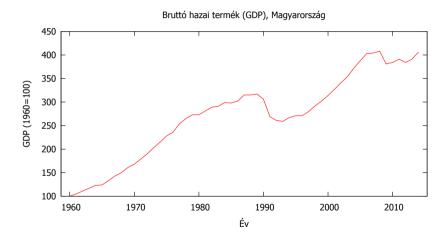
- Mi a sokaság és a minta? ismétlés stat 2-ből
- ("Sokaságban valszám kell, mintánál statisztika")
- Idősor minta értelemben: időben rendezett megfigyelések

$$\{5492, 5640, \ldots, 7317\}$$
,

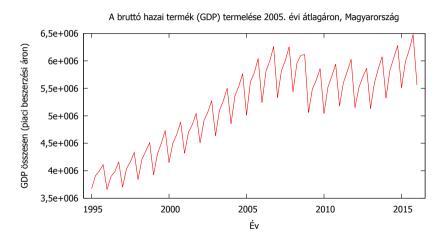
általános jelöléssel

$$\{y_1, y_2, \ldots, y_T\}$$

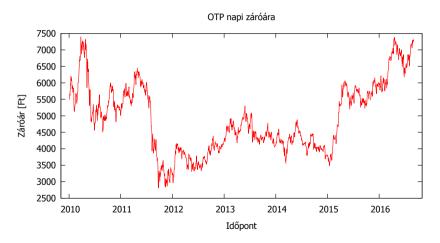
Példa: magyar GDP (hosszú éves)



Példa: magyar GDP (negyedéves)



Példa: OTP napi záróár



$$\{Y_t, t \in N\}$$

- Az idősor egy időpontban tehát egy valószínűségi változó
- Itt az N tehát egy rendezett halmaz
- Valszámos szó: sztochasztikus folyamat
- ullet Közgázban jellemzően N diszkrét, sőt véges: $N=\{1,2,\ldots,T\}$
- A minta tehát ennek egy realizációja (természetesen), itt a neve: trajektória

$$\{Y_t, t \in N\}$$

- Az idősor egy időpontban tehát egy valószínűségi változó
- Itt az N tehát egy rendezett halmaz
- Valszámos szó: sztochasztikus folyamat
- ullet Közgázban jellemzően N diszkrét, sőt véges: $N=\{1,2,\ldots,T\}$
- A minta tehát ennek egy realizációja (természetesen), itt a neve: trajektória

$$\{Y_t, t \in N\}$$

- Az idősor egy időpontban tehát egy valószínűségi változó
- Itt az N tehát egy rendezett halmaz
- Valszámos szó: sztochasztikus folyamat
- Közgázban jellemzően N diszkrét, sőt véges: $N = \{1, 2, \dots, T\}$
- A minta tehát ennek egy realizációja (természetesen), itt a neve: trajektória

$$\{Y_t, t \in N\}$$

- Az idősor egy időpontban tehát egy valószínűségi változó
- Itt az N tehát egy rendezett halmaz
- Valszámos szó: sztochasztikus folyamat
- Közgázban jellemzően N diszkrét, sőt véges: $N = \{1, 2, \dots, T\}$
- A minta tehát ennek egy realizációja (természetesen), itt a neve: trajektória

$$\{Y_t, t \in N\}$$

- Az idősor egy időpontban tehát egy valószínűségi változó
- Itt az N tehát egy rendezett halmaz
- Valszámos szó: sztochasztikus folyamat
- ullet Közgázban jellemzően N diszkrét, sőt véges: $N=\{1,2,\ldots,T\}$
- A minta tehát ennek egy realizációja (természetesen), itt a neve: trajektório

$$\{Y_t, t \in N\}$$

- Az idősor egy időpontban tehát egy valószínűségi változó
- Itt az N tehát egy rendezett halmaz
- Valszámos szó: sztochasztikus folyamat
- ullet Közgázban jellemzően N diszkrét, sőt véges: $N=\{1,2,\ldots,T\}$
- A minta tehát ennek egy realizációja (természetesen), itt a neve: trajektória

- Mi a különbség 1406 lakás ára és az OTP 1406 napi záróárfolyama között?
- Mi van emögött?
- A függetlenség feltevés (tarthatósága)!

- Mi a különbség 1406 lakás ára és az OTP 1406 napi záróárfolyama között?
- Mi van emögött?
- A függetlenség feltevés (tarthatósága)!

- Mi a különbség 1406 lakás ára és az OTP 1406 napi záróárfolyama között?
- Mi van emögött?
- A függetlenség feltevés (tarthatósága)!

- ullet Elég, ha ismerjük az *egyes időpontok* eloszlásait? (F_{Y_1} , F_{Y_2} stb.)
- Ami teljes mértékben leírja az idősort: az összes időpont együttes eloszlása

$$F_{Y_1,Y_2,...,Y_T}$$

 (Igazából keresztmetszetnél is ez volt, csak a függetlenségi feltevés miatt esett szét egyváltozósokra)

- ullet Elég, ha ismerjük az *egyes időpontok* eloszlásait? (F_{Y_1} , F_{Y_2} stb.)
- Ami teljes mértékben leírja az idősort: az összes időpont együttes eloszlása:

$$F_{Y_1,Y_2,\dots,Y_T}$$

 (Igazából keresztmetszetnél is ez volt, csak a függetlenségi feltevés miatt esett szét egyváltozósokra)

- ullet Elég, ha ismerjük az *egyes időpontok* eloszlásait? (F_{Y_1} , F_{Y_2} stb.)
- Ami teljes mértékben leírja az idősort: az összes időpont együttes eloszlása:

$$F_{Y_1,Y_2,...,Y_T}$$

 (Igazából keresztmetszetnél is ez volt, csak a függetlenségi feltevés miatt esett szét egyváltozósokra)

- Azaz: 1 dimenziós adatra vett 100 megfigyelés vs. 100 dimenziósra vett 1 megfigyelés!
- (Ezért nem segít az sem, ha hosszabbítjuk az idősorunkat!)
- Hogyan lehet egyáltalán így bármi becsülni…?
- ullet 1 megfigyelésből? Az érdekes lesz... ightarrow további feltevésekre lesz majd szükség $^{\circ}$

- Azaz: 1 dimenziós adatra vett 100 megfigyelés vs. 100 dimenziósra vett 1 megfigyelés!
- (Ezért nem segít az sem, ha hosszabbítjuk az idősorunkat!)
- Hogyan lehet egyáltalán így bármi becsülni…?
- ullet 1 megfigyelésből? Az érdekes lesz... o további feltevésekre lesz majd szükség!

- Azaz: 1 dimenziós adatra vett 100 megfigyelés vs. 100 dimenziósra vett 1 megfigyelés!
- (Ezért nem segít az sem, ha hosszabbítjuk az idősorunkat!)
- Hogyan lehet egyáltalán így bármi becsülni...?
- ullet 1 megfigyelésből? Az érdekes lesz... o további feltevésekre lesz majd szükség!

- Azaz: 1 dimenziós adatra vett 100 megfigyelés vs. 100 dimenziósra vett 1 megfigyelés!
- (Ezért nem segít az sem, ha hosszabbítjuk az idősorunkat!)
- Hogyan lehet egyáltalán így bármi becsülni...?
- ullet 1 megfigyelésből? Az érdekes lesz... o további feltevésekre lesz majd szükség!

- 1 Az idősor fogalma, jelentősége, idősorelemzés
 - Az idősor fogalma
 - Az idősorelemzés közgazdasági jelentősége, története
 - Idősorelemzési iskolák, a módszerek felosztása

Idősorok jellemzői a sokaságban

Az idősorelemzés közgazdasági jelentősége

- A legtöbb közgazdasági adat igazából idősorként érhető el!
- Számos feladatnál időbeli a fókusz, gondoljunk a szó szoros értelmében vett előrejelzési kérdésekre

Az idősorelemzés közgazdasági jelentősége

- A legtöbb közgazdasági adat igazából idősorként érhető el!
- Számos feladatnál időbeli a fókusz, gondoljunk a szó szoros értelmében vett előrejelzési kérdésekre

- Eleinte: egyszerű determinisztikus módszerek (pl. dekompozíciós modellek már a XIX. században)
- Később regresszió is, de még tekintet nélkül az idősoros jellegr
- Cochrane és Orcutt mutatott rá először 1949-ben, hogy ez nem jó ötle
- Megindult a kutatás ennek figyelembevételére
- Box és Jenkins könyve 1970-ben fordulópont: sztochasztikus módszerek
- Korszerű eljárások és aktuális kérdések (nemstacionaritás, nemlinearitás, többváltozós módszerek, ARCH, ...)

- Eleinte: egyszerű determinisztikus módszerek (pl. dekompozíciós modellek már a XIX. században)
- Később regresszió is, de még tekintet nélkül az idősoros jellegre
- Cochrane és Orcutt mutatott rá először 1949-ben, hogy ez nem jó ötlet
- Megindult a kutatás ennek figyelembevételére
- Box és Jenkins könyve 1970-ben fordulópont: sztochasztikus módszerekk
- Korszerű eljárások és aktuális kérdések (nemstacionaritás, nemlinearitás, többváltozós módszerek, ARCH, ...)

- Eleinte: egyszerű determinisztikus módszerek (pl. dekompozíciós modellek már a XIX. században)
- Később regresszió is, de még tekintet nélkül az idősoros jellegre
- Cochrane és Orcutt mutatott rá először 1949-ben, hogy ez nem jó ötlet
- Megindult a kutatás ennek figyelembevételére
- Box és Jenkins könyve 1970-ben fordulópont: sztochasztikus módszerek
- Korszerű eljárások és aktuális kérdések (nemstacionaritás, nemlinearitás, többváltozós módszerek, ARCH, ...)

- Eleinte: egyszerű determinisztikus módszerek (pl. dekompozíciós modellek már a XIX. században)
- Később regresszió is, de még tekintet nélkül az idősoros jellegre
- Cochrane és Orcutt mutatott rá először 1949-ben, hogy ez nem jó ötlet
- Megindult a kutatás ennek figyelembevételére
- Box és Jenkins könyve 1970-ben fordulópont: sztochasztikus módszerek
- Korszerű eljárások és aktuális kérdések (nemstacionaritás, nemlinearitás, többváltozós módszerek, ARCH, ...)

- Eleinte: egyszerű determinisztikus módszerek (pl. dekompozíciós modellek már a XIX. században)
- Később regresszió is, de még tekintet nélkül az idősoros jellegre
- Cochrane és Orcutt mutatott rá először 1949-ben, hogy ez nem jó ötlet
- Megindult a kutatás ennek figyelembevételére
- Box és Jenkins könyve 1970-ben fordulópont: sztochasztikus módszerek
- Korszerű eljárások és aktuális kérdések (nemstacionaritás, nemlinearitás, többváltozós módszerek, ARCH, ...)

- Eleinte: egyszerű determinisztikus módszerek (pl. dekompozíciós modellek már a XIX. században)
- Később regresszió is, de még tekintet nélkül az idősoros jellegre
- Cochrane és Orcutt mutatott rá először 1949-ben, hogy ez nem jó ötlet
- Megindult a kutatás ennek figyelembevételére
- Box és Jenkins könyve 1970-ben fordulópont: sztochasztikus módszerek
- Korszerű eljárások és aktuális kérdések (nemstacionaritás, nemlinearitás, többváltozós módszerek, ARCH, ...)

Tartalom

- Az idősor fogalma, jelentősége, idősorelemzés
 - Az idősor fogalma
 - Az idősorelemzés közgazdasági jelentősége, története
 - Idősorelemzési iskolák, a módszerek felosztása
- Idősorok jellemzői a sokaságban

Az idősorelemzési módszerek csoportosítása

- Időtartomány vs. frekvenciatartomány (csak az előbbivel fogunk most foglalkozni)
- Determinisztikus vs. sztochasztikus (definíciós kavarodások, most: a véletlennek van-e folyamatépítő szerepe; ld. később részletesen)
- Egyváltozós vs. többváltozós (+panel)

Az idősorelemzési módszerek csoportosítása

- Időtartomány vs. frekvenciatartomány (csak az előbbivel fogunk most foglalkozni)
- Determinisztikus vs. sztochasztikus (definíciós kavarodások, most: a véletlennek van-e folyamatépítő szerepe; ld. később részletesen)
- Egyváltozós vs. többváltozós (+panel)

Az idősorelemzési módszerek csoportosítása

- Időtartomány vs. frekvenciatartomány (csak az előbbivel fogunk most foglalkozni)
- Determinisztikus vs. sztochasztikus (definíciós kavarodások, most: a véletlennek van-e folyamatépítő szerepe; ld. később részletesen)
- Egyváltozós vs. többváltozós (+panel)

Tartalom

- Az idősor fogalma, jelentősége, idősorelemzés
 - Az idősor fogalma
 - Az idősorelemzés közgazdasági jelentősége, története
 - Idősorelemzési iskolák, a módszerek felosztása
- Idősorok jellemzői a sokaságban

Várhatóérték- és szórásnégyzet-függvény

Emlékezzünk rá, hogy egy adott időpontban az idősor egyszerűen egy valószínűségi változó, így definiálható várható értéke, szórásnégyzete, két ilyennek a kovarianciája; ez alapján:

• Várhatóérték-függvény ($\mu: N \to \mathbb{R}$):

$$\mu_t := \mathbb{E} Y_t$$

ullet Szórásnégyzet-függvény $(\sigma^2: \mathcal{N}
ightarrow \mathbb{R}_+)$:

$$\sigma_t^2 := \mathbb{D}^2 Y_t = \mathbb{E}(Y_t - \mathbb{E}Y_t)^2 = \mathbb{E}(Y_t - \mu_t)^2 = \mathbb{E}Y_t^2 - \mu_t^2$$

Várhatóérték- és szórásnégyzet-függvény

Emlékezzünk rá, hogy egy adott időpontban az idősor egyszerűen egy valószínűségi változó, így definiálható várható értéke, szórásnégyzete, két ilyennek a kovarianciája; ez alapján:

• Várhatóérték-függvény ($\mu: N \to \mathbb{R}$):

$$\mu_t := \mathbb{E} Y_t$$

• Szórásnégyzet-függvény ($\sigma^2: N \to \mathbb{R}_+$):

$$\sigma_t^2 := \mathbb{D}^2 Y_t = \mathbb{E}(Y_t - \mathbb{E}Y_t)^2 = \mathbb{E}(Y_t - \mu_t)^2 = \mathbb{E}Y_t^2 - \mu_t^2$$

• Autokovariancia-függvény (ACVF, $\gamma: N \times N \to \mathbb{R}$):

$$\gamma_{t,s} := \operatorname{cov}(Y_t, Y_s) = \mathbb{E}\left[\left(Y_t - \mathbb{E}Y_t\right)(Y_s - \mathbb{E}Y_s)\right] = \\ \mathbb{E}\left[\left(Y_t - \mu_t\right)(Y_s - \mu_s)\right] = \mathbb{E}\left(Y_t Y_s\right) - \mu_t \mu_s$$

- Nyilván $\gamma_{t,t} = \sigma_t^2$
- ullet Autokorrelációs függvény (ACF, $ho: \mathcal{N} imes \mathcal{N}
 ightarrow [-1,1])$

$$\rho_{t,s} := \operatorname{corr}(Y_t, Y_s) = \frac{\gamma_{t,s}}{\sigma_t \sigma_s}$$

Ne feledjük: mindezekben semmi sztochasztikus nincs, teljesen közönséges – determinisztikus – függvénvek!



• Autokovariancia-függvény (ACVF, $\gamma: N \times N \to \mathbb{R}$):

$$\gamma_{t,s} := \operatorname{cov}(Y_t, Y_s) = \mathbb{E}\left[\left(Y_t - \mathbb{E}Y_t\right)(Y_s - \mathbb{E}Y_s)\right] = \\ \mathbb{E}\left[\left(Y_t - \mu_t\right)(Y_s - \mu_s)\right] = \mathbb{E}\left(Y_t Y_s\right) - \mu_t \mu_s$$

- Nyilván $\gamma_{t,t} = \sigma_t^2$
- Autokorrelációs függvény (ACF, $\rho: N \times N \rightarrow [-1, 1]$)

$$\rho_{t,s} := \operatorname{corr}(Y_t, Y_s) = \frac{\gamma_{t,s}}{\sigma_t \sigma_s}$$

Ne feledjük: mindezekben semmi sztochasztikus nincs, teljesen közönséges – determinisztikus – függvények!



• Autokovariancia-függvény (ACVF, $\gamma: N \times N \to \mathbb{R}$):

$$\gamma_{t,s} := \operatorname{cov}(Y_t, Y_s) = \mathbb{E}\left[\left(Y_t - \mathbb{E}Y_t\right)\left(Y_s - \mathbb{E}Y_s\right)\right] = \\ \mathbb{E}\left[\left(Y_t - \mu_t\right)\left(Y_s - \mu_s\right)\right] = \mathbb{E}\left(Y_t Y_s\right) - \mu_t \mu_s$$

- Nyilván $\gamma_{t,t} = \sigma_t^2$
- Autokorrelációs függvény (ACF, $\rho: N \times N \rightarrow [-1,1]$):

$$\rho_{t,s} := \operatorname{corr}(Y_t, Y_s) = \frac{\gamma_{t,s}}{\sigma_t \sigma_s}$$

Ne feledjük: mindezekben semmi sztochasztikus nincs, teljesen közönséges – determinisztikus – függvények!



• Autokovariancia-függvény (ACVF, $\gamma: N \times N \to \mathbb{R}$):

$$\gamma_{t,s} := \operatorname{cov}(Y_t, Y_s) = \mathbb{E}\left[\left(Y_t - \mathbb{E}Y_t\right)\left(Y_s - \mathbb{E}Y_s\right)\right] = \\ \mathbb{E}\left[\left(Y_t - \mu_t\right)\left(Y_s - \mu_s\right)\right] = \mathbb{E}\left(Y_t Y_s\right) - \mu_t \mu_s$$

- Nyilván $\gamma_{t,t} = \sigma_t^2$
- Autokorrelációs függvény (ACF, $\rho: N \times N \rightarrow [-1, 1]$):

$$\rho_{t,s} := \operatorname{corr}(Y_t, Y_s) = \frac{\gamma_{t,s}}{\sigma_t \sigma_s}$$

Ne feledjük: mindezekben semmi sztochasztikus nincs, teljesen közönséges – determinisztikus – függvények!



- Úgy viszonyul az ACF-hez mint a sima (keresztmetszeti) parciális korreláció a korrelációhoz: bizonyos változókon keresztül terjedő hatásokat szűrjük (lineárisan)
- De melyikeket
- Ami közbeesik: t és s közti korreláció (s > t), szűrve a $t+1, t+2, \ldots, s-2, s-1$ időpontokon keresztül terjedő hatásokat
- Kiszámítható az ACF-ek ismeretében egyszerű mátrixműveletekkel
- (Avagy: a korreláció kijön egy olyan regresszióból, aminek egyetlen magyarázó változója van, a parciális korrelációhoz pedig hozzá kell adni a közbenső időpontokat is)

- Úgy viszonyul az ACF-hez mint a sima (keresztmetszeti) parciális korreláció a korrelációhoz: bizonyos változókon keresztül terjedő hatásokat szűrjük (lineárisan)
- De melyikeket?
- Ami közbeesik: t és s közti korreláció (s > t), szűrve a $t+1, t+2, \ldots, s-2, s-1$ időpontokon keresztül terjedő hatásokat
- Kiszámítható az ACF-ek ismeretében egyszerű mátrixműveletekke
- (Avagy: a korreláció kijön egy olyan regresszióból, aminek egyetlen magyarázó változója van, a parciális korrelációhoz pedig hozzá kell adni a közbenső időpontokat is)

- Úgy viszonyul az ACF-hez mint a sima (keresztmetszeti) parciális korreláció a korrelációhoz: bizonyos változókon keresztül terjedő hatásokat szűrjük (lineárisan)
- De melyikeket?
- Ami közbeesik: t és s közti korreláció (s>t), szűrve a $t+1, t+2, \ldots, s-2, s-1$ időpontokon keresztül terjedő hatásokat
- Kiszámítható az ACF-ek ismeretében egyszerű mátrixműveletekke
- (Avagy: a korreláció kijön egy olyan regresszióból, aminek egyetlen magyarázó változója van, a parciális korrelációhoz pedig hozzá kell adni a közbenső időpontokat is)

- Úgy viszonyul az ACF-hez mint a sima (keresztmetszeti) parciális korreláció a korrelációhoz: bizonyos változókon keresztül terjedő hatásokat szűrjük (lineárisan)
- De melyikeket?
- Ami közbeesik: t és s közti korreláció (s>t), szűrve a $t+1, t+2, \ldots, s-2, s-1$ időpontokon keresztül terjedő hatásokat
- Kiszámítható az ACF-ek ismeretében egyszerű mátrixműveletekkel
- (Avagy: a korreláció kijön egy olyan regresszióból, aminek egyetlen magyarázó változója van, a parciális korrelációhoz pedig hozzá kell adni a közbenső időpontokat is)

- Úgy viszonyul az ACF-hez mint a sima (keresztmetszeti) parciális korreláció a korrelációhoz: bizonyos változókon keresztül terjedő hatásokat szűrjük (lineárisan)
- De melyikeket?
- Ami közbeesik: t és s közti korreláció (s>t), szűrve a $t+1, t+2, \ldots, s-2, s-1$ időpontokon keresztül terjedő hatásokat
- Kiszámítható az ACF-ek ismeretében egyszerű mátrixműveletekkel
- (Avagy: a korreláció kijön egy olyan regresszióból, aminek egyetlen magyarázó változója van, a parciális korrelációhoz pedig hozzá kell adni a közbenső időpontokat is)

- Korrelogram: ACF és PACF együtt ábrázolva
- (Egyelőre úgy tűnik, hogy ez egy kétdimenziós függvény, ezt majd később árnyalni fogjuk és ezért nem is ábrázoljuk most még ténylegesen)
- Jelentősége: ha ismerjük nevezetes folyamatok elméleti korrelogramját, akkor egy minta mögötti, azt adó folyamatra következtethetünk az alapján, hogy a minta – empirikus – korrelogramia hogyan néz ki
- (Sajnos a gyakorlatban sokszor csak hozzávetőleges lehet)

- Korrelogram: ACF és PACF együtt ábrázolva
- (Egyelőre úgy tűnik, hogy ez egy kétdimenziós függvény, ezt majd később árnyalni fogjuk –
 és ezért nem is ábrázoljuk most még ténylegesen)
- Jelentősége: ha ismerjük nevezetes folyamatok elméleti korrelogramját, akkor egy minta mögötti, azt adó folyamatra következtethetünk az alapján, hogy a minta – empirikus – korrelogramia hogyan néz ki
- (Sajnos a gyakorlatban sokszor csak hozzávetőleges lehet)

- Korrelogram: ACF és PACF együtt ábrázolva
- (Egyelőre úgy tűnik, hogy ez egy kétdimenziós függvény, ezt majd később árnyalni fogjuk –
 és ezért nem is ábrázoljuk most még ténylegesen)
- Jelentősége: ha ismerjük nevezetes folyamatok elméleti korrelogramját, akkor egy minta mögötti, azt adó folyamatra következtethetünk az alapján, hogy a minta – empirikus – korrelogramja hogyan néz ki
- (Sajnos a gyakorlatban sokszor csak hozzávetőleges lehet)

- Korrelogram: ACF és PACF együtt ábrázolva
- (Egyelőre úgy tűnik, hogy ez egy kétdimenziós függvény, ezt majd később árnyalni fogjuk –
 és ezért nem is ábrázoljuk most még ténylegesen)
- Jelentősége: ha ismerjük nevezetes folyamatok elméleti korrelogramját, akkor egy minta mögötti, azt adó folyamatra következtethetünk az alapján, hogy a minta – empirikus – korrelogramja hogyan néz ki
- (Sajnos a gyakorlatban sokszor csak hozzávetőleges lehet)