Korrelációk

6 & 7. fejezet

Tartalom

- Időben nem változó, stohasztikus folyamatok
- Korrelációk
  - Definició(k), rövid- és hosszútávú korrelációk
- Időbeli korrelációk gazdasági adatokban

- Példák

Időben állandó folyamatok

- Mit jelent, hogy állandó? Nem változó?
  - Több definició
- Mit jelent a korreláltság
  - Több definició
- Eszköz: statisztikailag megfigyelhető mennyiségek, azaz kumulánsok
- Átlag v. várható érték:

$$E\{x(t)\} = \int_{-\infty}^{\infty} x \cdot PDF(x)dx$$

- Autokorrelációs függvény:

Joint PDF

$$E\{x(t_1)x(t_2)\} = \int_{-\infty}^{\infty} x_1 x_2 PDF(x_1, x_2; t_1, t_2)dx$$

Időben állandó stohasztikus folyamatok

- Általában $PDF(x_1, \ldots, x_n; t_1, \ldots, t_n)$ kell, hogy teljesen leirjunk egy folyamatot, de a gyakorlatban 2-pont függvényeket vizsgálunk
- Definició: $x(t)$ időben állandó stohasztikus ha a PDF-je invariáns az időbeli eltolásokra
- Példák
  - Széles értelemben vett IÁSF
    - $E\{x(t)\} = \mu$
    - $E\{x(t_1)x(t_2)\} = R(t_1, t_2) = R(t_2 - t_1) = R(\tau)$
    - $E\{x^2(t)\} = R(0) \quad \Rightarrow \quad \text{variance } R(0) - \mu^2 \text{ állandó}$
  - Aszimptotikusan IÁSF
    - $x(t_i + c)$ független c –től nagy c értékekre
  - N-ed rendű IÁSF
    - $PDF(x_1, \ldots, x_n; t_1, \ldots, t_n) = PDF(x_1, \ldots, x_n; t_1 + c, \ldots, t_n + c)$ csak $n \leq N$ esetén igaz
  - Egy intervallumban is értelmezhető ugyanez

Korrelációk

- $R(t_1, t_2)$ az átlagra érzékeny, ami levonható
- Autokovarianca: $C(t_1, t_2) = R(t_1, t_2) - \mu(t_1)\mu(t_2)$
- Konstans folyamatokban $C(\tau) = R(\tau) - \mu^2$
- Legyen most $\mu = 0, \sigma^2 = R(0) = 1 \;\Rightarrow\; R = C$
- Tipikus időskála – C or R memóriája

- Állandó folyamatokra $\int_0^\infty R(\tau) = \begin{cases} < \infty \\ \infty \\ \text{undet.} \end{cases}$
- Ha véges, akkor létezik τ_c korrelációs idő

Korrelációk – példák

- $R(\tau) = e^{\tau/\tau_0}$ memória $\tau_c = \tau_0$
- $R(\tau) = e^{\tau^\nu/\tau_c}$ memória $\tau_c = \frac{\tau_0^{\frac{1}{\nu}}}{\nu} \Gamma\left(\frac{1}{\nu}\right)$
- $R(\tau) \sim \tau^{\eta-1}$ ha $0 < \eta \leq 1$ τ_c divergens

0^{th} approximation: there is a correlation up to τ_c

- Formálisan, poziti korrelációt feltételezve

$$E\{S_n^2\} \approx n\left(E\{n_i^2\} + \underbrace{2\sum_{k=1}^{n} E\{x_i x_{i+k}\}}\right)$$

const. rövidtávú korr.

∞ , hosszútávú korr. , nincs tipikus időskála

Rövidtávú korrelációk

- Példa: Brown mozgás v(t) függvénye $\rightarrow$ exponenciális autokorr. fv.
- A frekvenciatérben (Fourier-térben) is lehet vizsgálódni

$$R(\tau) = \sigma^2 e^{-\frac{|\tau|}{\tau_c}} \Rightarrow \; S(f) = \mathcal{F}\big(R(\tau)\big) = \frac{2\sigma^2\tau_c}{1 + (2\pi f \tau_c)^2}$$

- Amikor $f \ll \frac{1}{2\pi\tau_c}$ a spektrum frekvenciafüggetlen
- Időablakra ami lényegesen nagyobb, mint a τ_c a folyamat fehérzaj
- Ennek integrálja a Wiener-folyamat aminek a spektruma

$$S(f) \sim \frac{1}{f^2}$$

Hosszútávú korrelációk

- Hatványfüggvényszerű autokorr. fv. → hosszútávú korrelációk

$$S(f) \sim \frac{const.}{|f|^{\eta}} \qquad 0 < \eta < 2$$

- Sok helyen megfigyelhető
  - $\eta = 0$: fehérzaj
  - $\eta = 2$: Wiener-folyamat
  - $\eta \approx 1$: $1/f$ zaj, amit megfigyeltek diódák és tranzisztorok áramingadozásában, közlekedési jelenségekig, de minden zene és beszéd is ilyen, ld. *1/f noise in music and speech*, October 1975 Nature 258(5533):317
  - A cikket itt le lehet tölteni

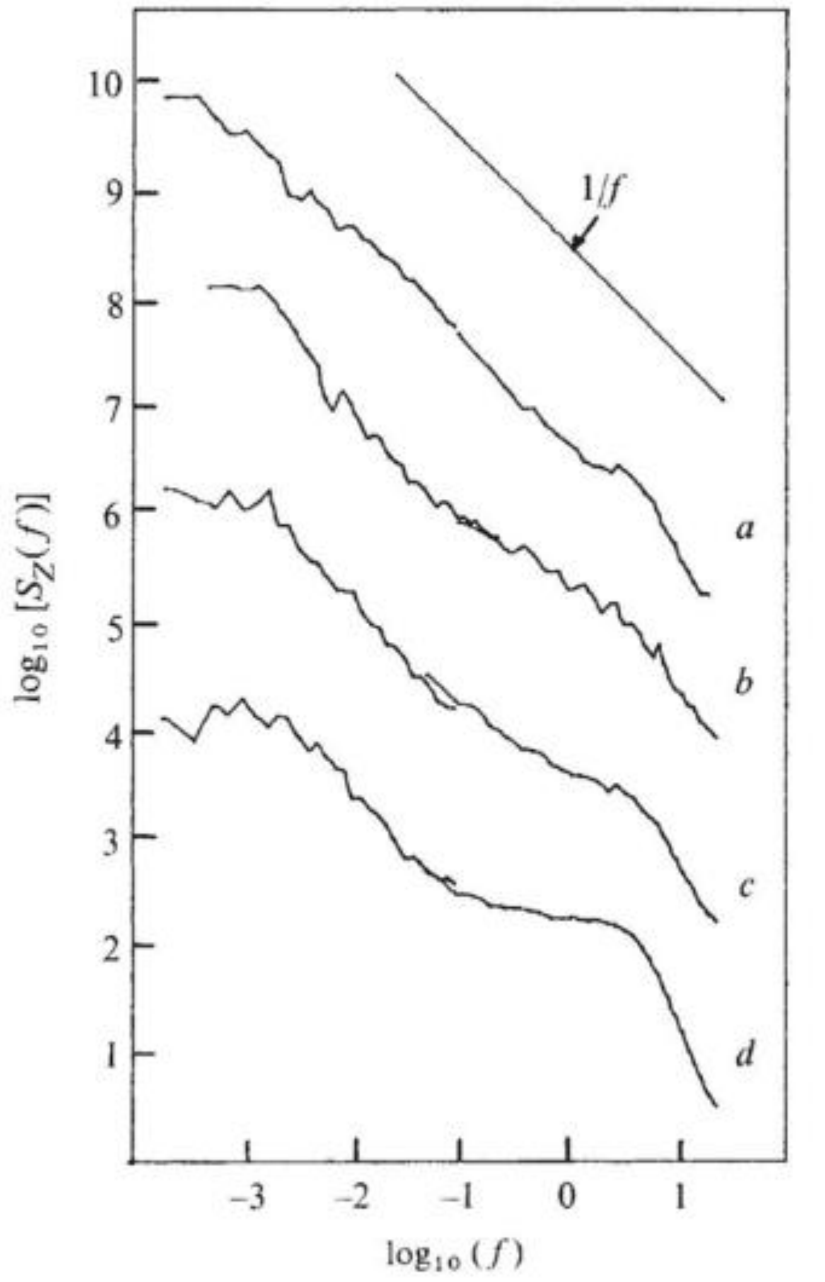


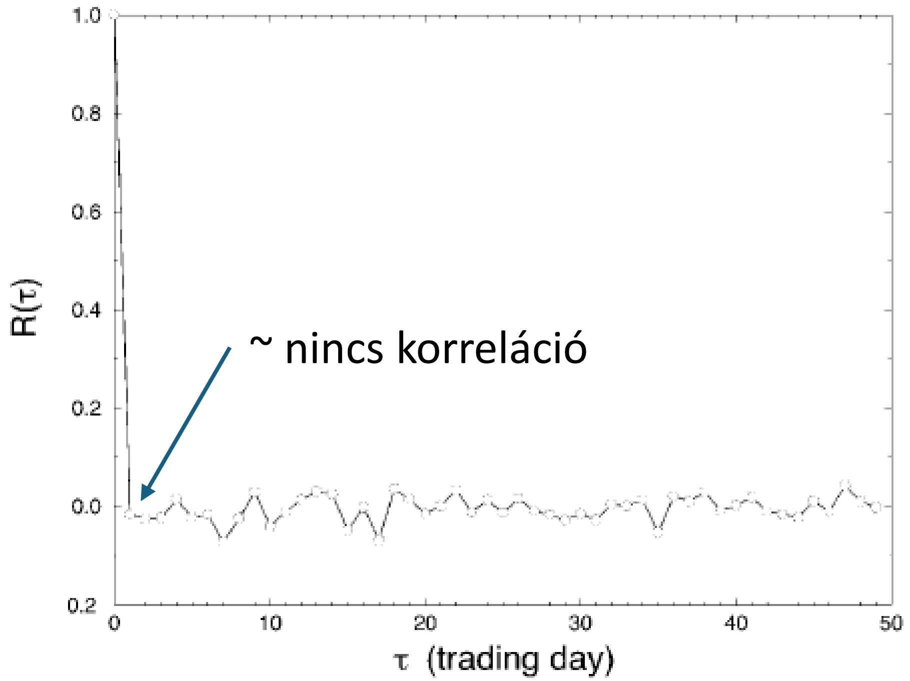
Fig. 3 Power spectra of pitch fluctuations, $S_Z(f)$ against f, for four radio stations: *a*, classical; *b*, jazz and blues; *c*, rock; *d*, news and talk.

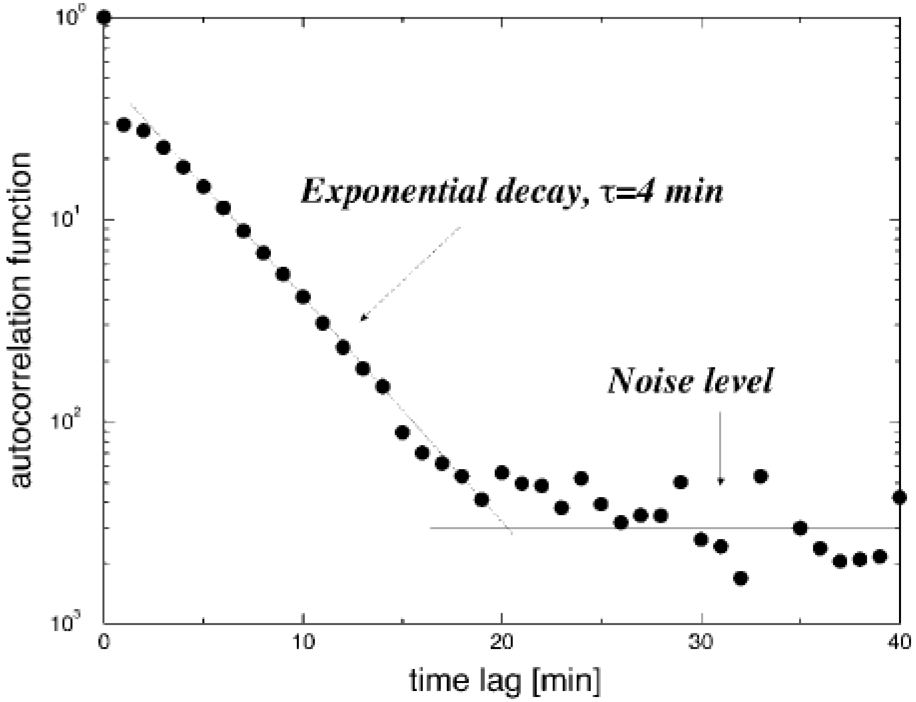
Összehasonlitás

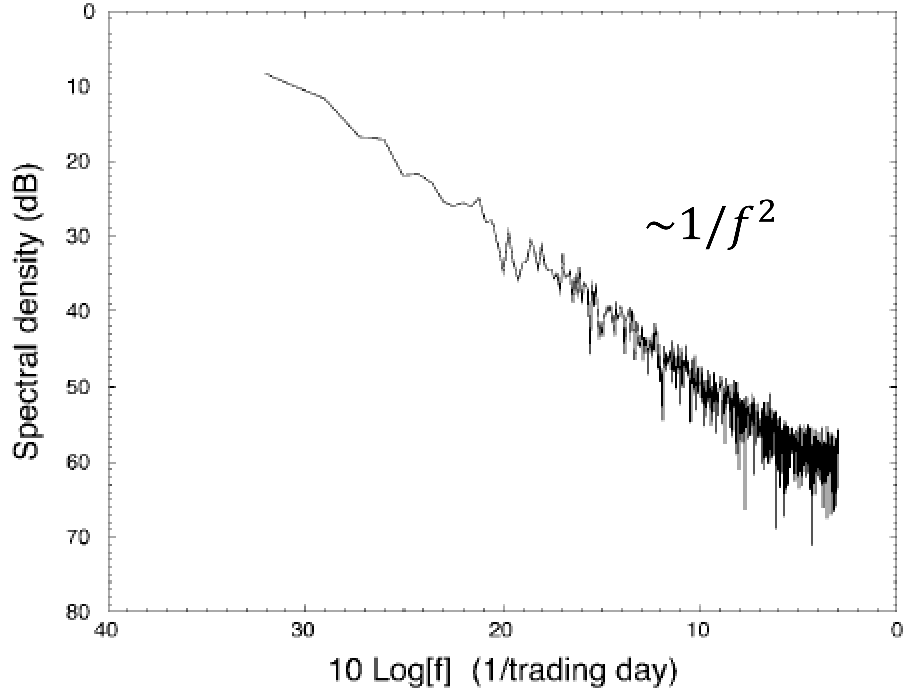
- Ha τ_c jellemzi a rendszert, akkor
$$f(x_1, \dots, x_{n-1}; t_1, \dots, t_{n-1}|x_n; t_n) = f(x_{n-1}; t_{n-1}|x_n; t_n)$$
- Markov folyamat. A legegyszerűbb példa
$$f(x_1, x_2, x_3; t_1, t_2, t_3) = f(x_1, t_1) f(x_1; t_1|x_2; t_2) f(x_2; t_2|x_3; t_3)$$
- Vagyis első- és másodrendű feltételes valószinűségek jellemzik a folyamatokat. Olyan folyamatok amelyekben nincs tipikus időskála, nem Markov folyamatok.
- A nem-markovi folyamatok esetén minden rendű valószinűség kell
- Fontos következmény, hogy különböző $1/f$ zajok nem tudhatók be ugyananannak a folyamatnak, hacsak nem tudunk többet.

Gazdasági adatok

- Autokorreláció és spektrum
  - Az ár logaritmusának időbeli mozgása
- Véletlen séta vagyis fehér zaj szerű viselkedés látszik
- A korrelációs hossz kevesebb, mint egy kereskedelmi nap
- Gyakori méréssel kereskedelmi perc szintű korreláció kimutatható







Egyéb eszközök

- Az autokorrelációs fv. és a spektrum nem túl érzékeny a hosszú távú korrelációkra
- A szórás időfejlődése többet árul el
- Általában $\sigma(t) \sim t^{\nu}$ ahol $\nu = 1/2$ ha a árak függetlenek
- „recent investigation” : 30 kereskedelmi perctől 100 kereskedelmi napig terjedő időablakban: $\nu \approx 0.5$
- Piacfüggő lehet. Más vizsgálatok szerint:
  - New York Composite Index: $\nu \approx 0.52$
  - DAX index: $\nu \approx 0.53$
  - MIB index: $\nu \approx 0.57$

Gyenge hosszútávú korreláció
Nagyobb ν a kevésbé hatékony piacoknál

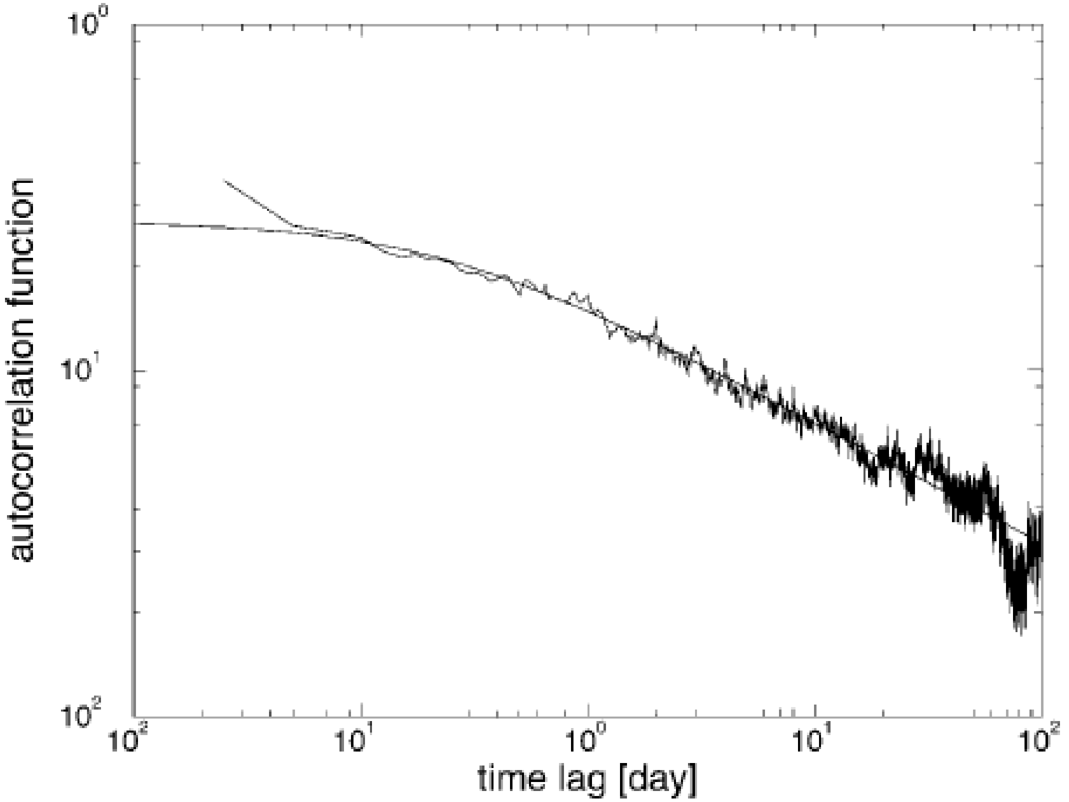
- HF S&P500 data $t < 30$ perc, superdiffuse $\nu \approx 0.8$
- Hosszabb távon $\nu = 0.53$

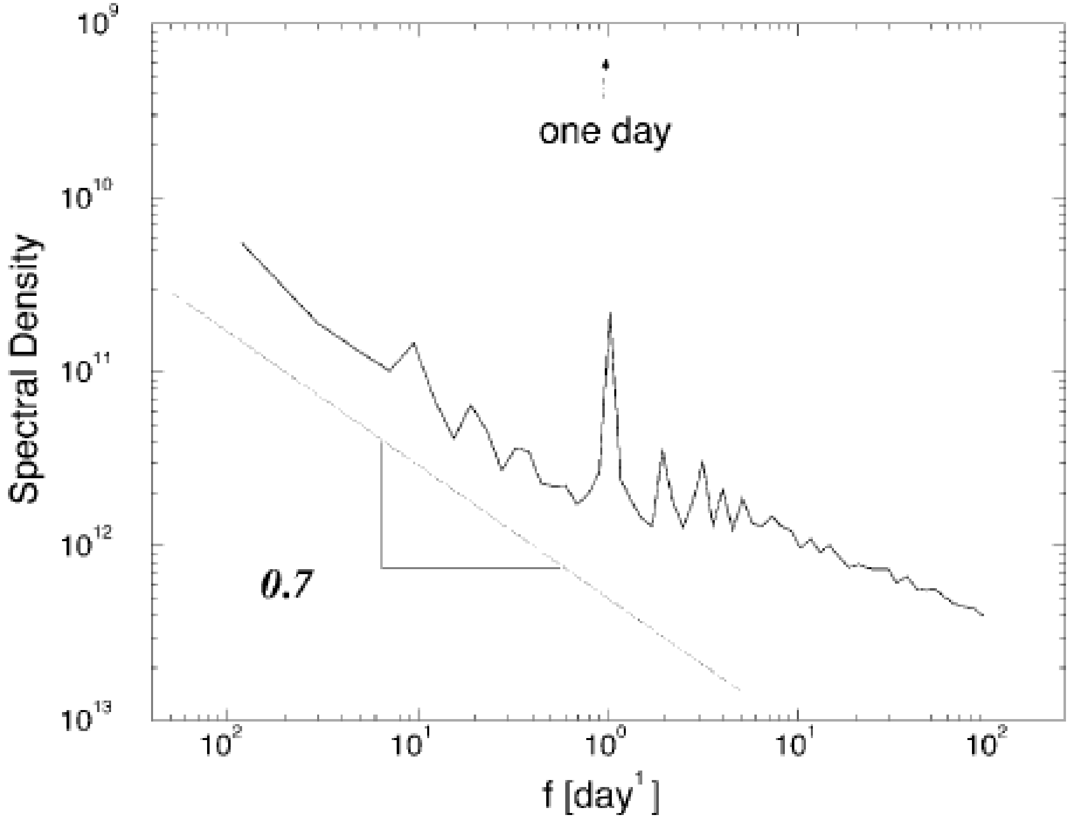
Magasabb rendű korrelációk – volatilitás

- A páronkénti árkülönbség rövid távú memóriával rendelkezik
- Ugyanakkor nem-lineáris függvények szerinti kombinációk már nem
- Más alapvető stoch. foly. lehet : volatilitás → szórás egy időablakban
- Más módszerekkle is lehet becsülni
- Miért érdekes:
  - Összeköthető a piacba aott időben érkező információ mennyiségével
    - Ha jön infó, megváltozik a kereskedés → nagy volatilitás
  - Később tárgyalandó modellekben lehet használni
  - A gyakorlatban egy fontos paraméter a befektetések kockázatának meghatározásában

Magasabb rendű korrelációk – volatilitás

- A volatilitás autokorrelációs függvénye jól leirható hatványfüggvénnyel
- A spektrum is jól jellemezheti a hosszútávú korrelációkat
- Nincs ellentmondás: ezek magasabb rendű valószinűségek





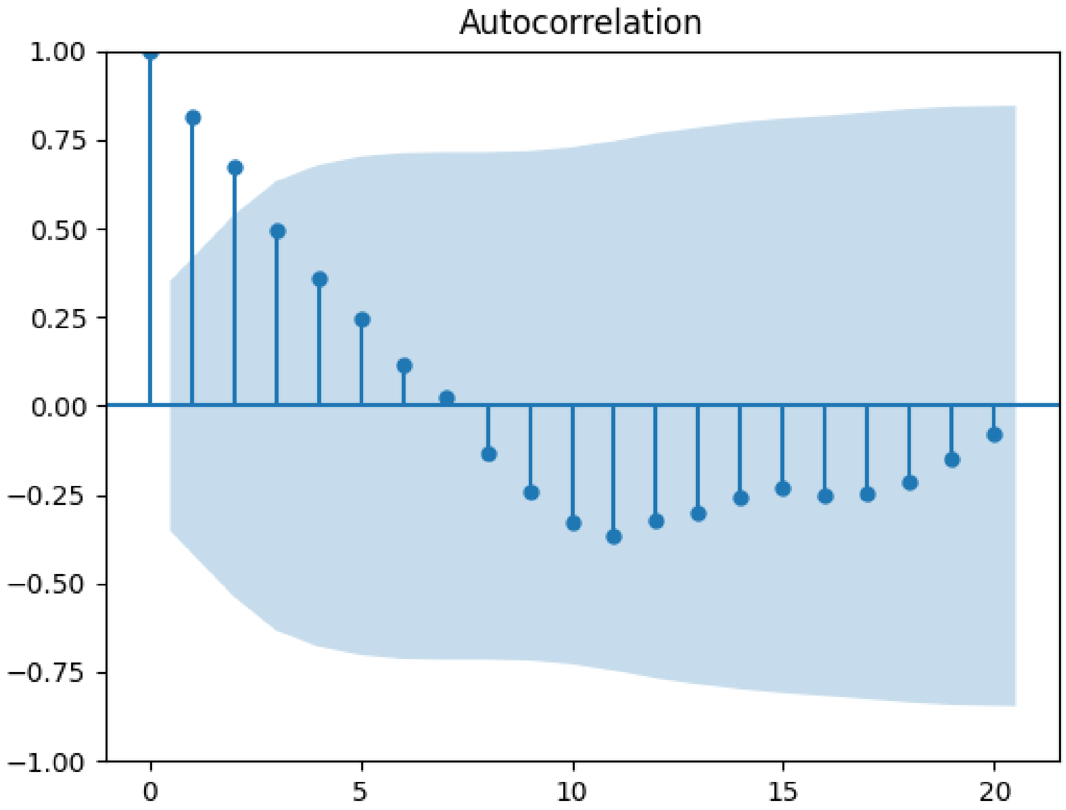
Összefoglalás

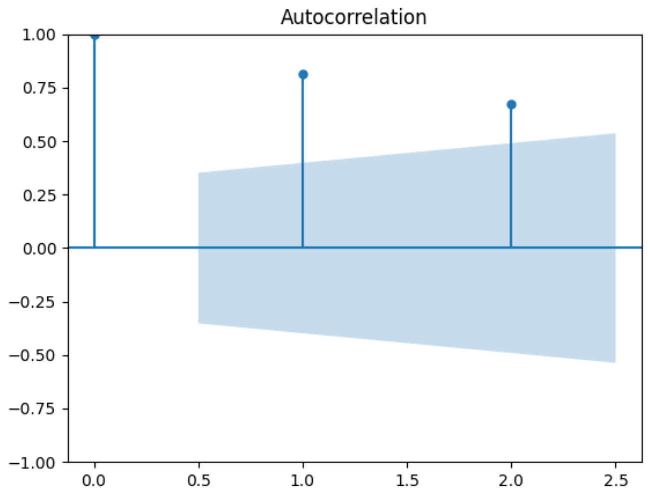
- Az árak stochasztikus dinamikája: véletlen bolyongás, rövidtávú korr.
- Leirhatóak az árváltozások időben állandó folyamatként?
- Empiria: nem, mert a volatilitás időfüggő
- Aszimptotikusan, nagy idősorokat vizsgálva, időben állandó

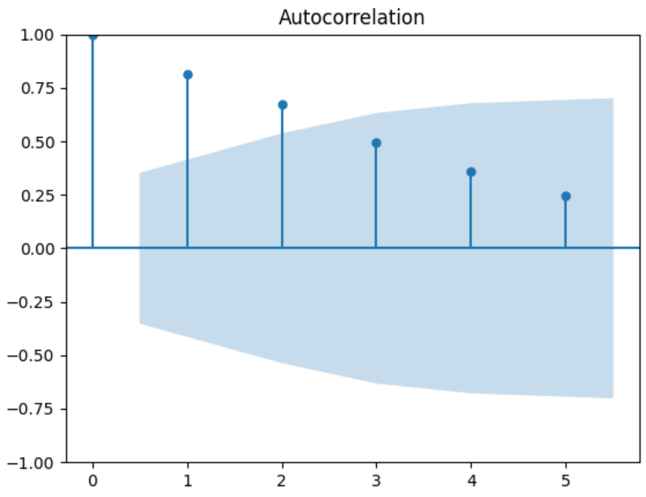
- Láttuk tehát:
  - Az árváltozások páronként korrelálatlanok
  - A rövidtávú memória néhány perc hosszú
  - Gyenge hosszútávú korreláció megfigyelhető a szórás időfejlődésében
  - A volatilitás hosszútávon, $1/f$ szerű spektrummaé korrelált

Saját példa

- Erzsébet tér CO mérés
- Napi mintavétel
- Időtolás (lag) 2, 5, 20 nap

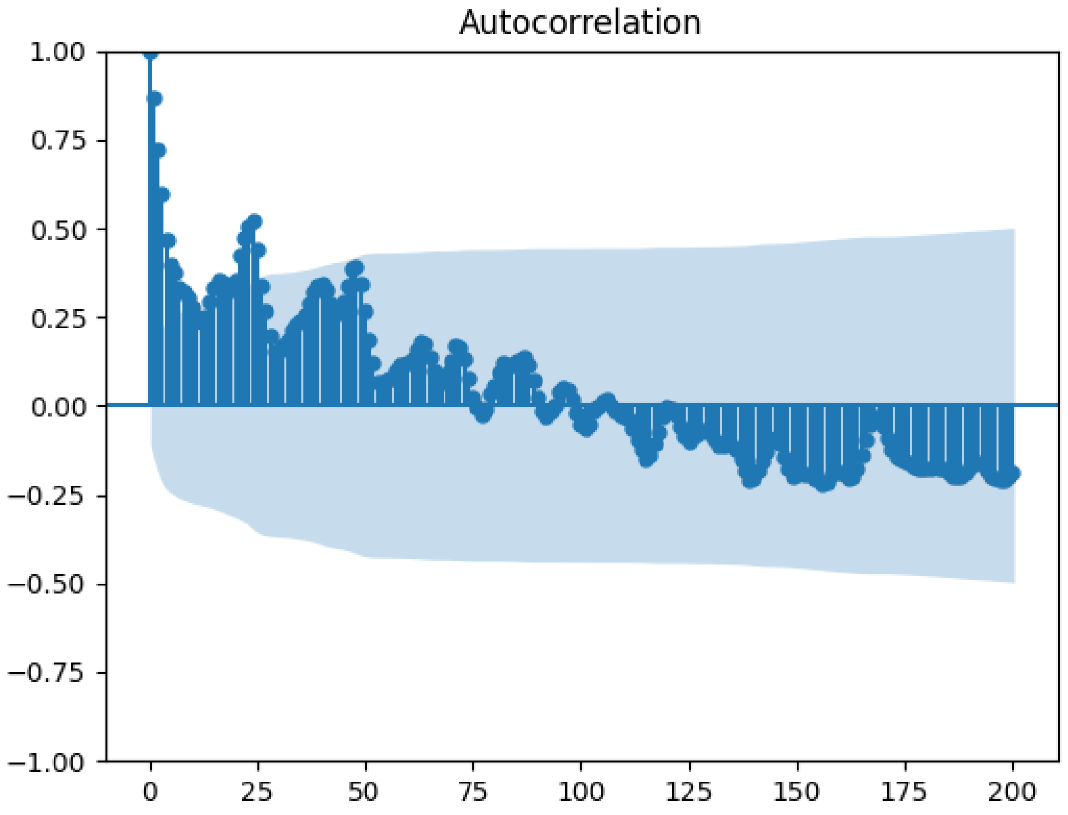


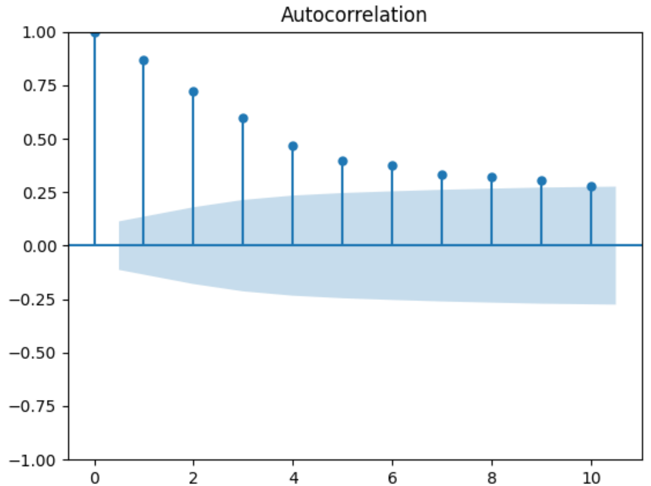


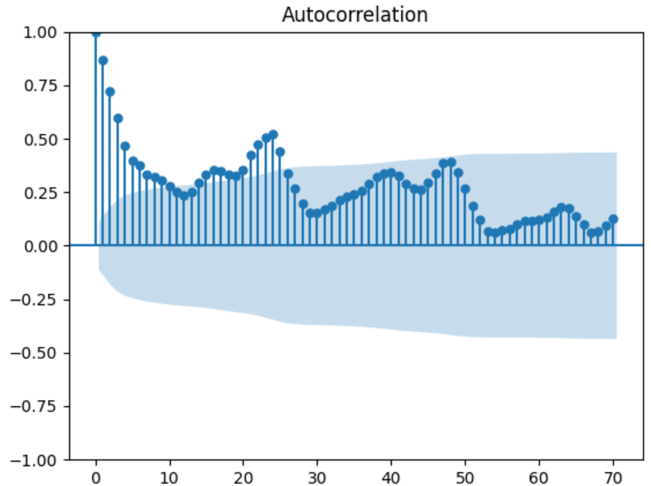


Saját példa

- Erzsébet tér CO mérés
- Órás mintavétel
- Időtolás (lag) 10, 70, 200 óra







```
import statsmodels.api as sm
from statsmodels.graphics import tsaplots
import matplotlib.pyplot as plt


#define data
x =
[278,277,251,262,268,297,284,289,305,306,304,304,265,257,270,266,308,285,302,277,300,347,370,337,375,469,345,297,285,303,301,303,304,306,264,2
55,234,223,218,234,248,218,232,221,237,310,345,285,298,312,296,286,262,281,329,364,483,361,325,306,287,300,306,255,291,304,298,216,242,271,23
5,223,252,227,229,222,230,241,231,294,338,284,285,320,267,244,223,213,212,212,218,220,217,253,248,254,280,234,226,237,240,246,264,322,394,337,
357,311,244,232,211,205,205,229,203,215,226,252,260,258,284,291,290,284,260,270,308,454,488,380,312,278,238,242,191,197,202,235,216,245,253,2
77,406,391,381,439,294,259,248,264,227,404,282,297,282,234,197,199,165,205,125,117,142,180,183,183,227,242,206,282,271,162,216,222,210,338,29
9,186,171,172,146,115,116,134,166,144,146,121,154,202,324,497,546,384,311,163,146,165,131,142,140,129,104,81,61,76,105,95,117,119,114,136,152,
179,192,215,192,151,116,102,104,122,125,146,206,223,224,252,194,246,240,187,192,268,205,217,193,142,159,138,125,118,122,115,120,118,126,144,1
93,188,196,175,148,128,124,103,103,103,98,92,103,126,132,142,141,138,117,120,117,130,140,151,203,269,230,221,216,175,154,181,167,165,165,169,1
66,172,182,188,177,180,173,176,169,171,173,254,265,304,269]
print(len(x))
y = [296,281,295,248,258,295,233,222,152,178,132,171,252,291,323,267,313,351,377,414,371,355,341,325,316,342,317,336,342,337,237]
#calculate autocorrelations
sm.tsa.acf(y)


#plot autocorrelation function
fig = tsaplots.plot_acf(y, lags=20)
plt.show()
```