

Idősor statisztikai elemzése S&P 500 index példáján

Szmolenicki Flórián, 2025. június

SZTE - Idősorok statisztikai elemzése

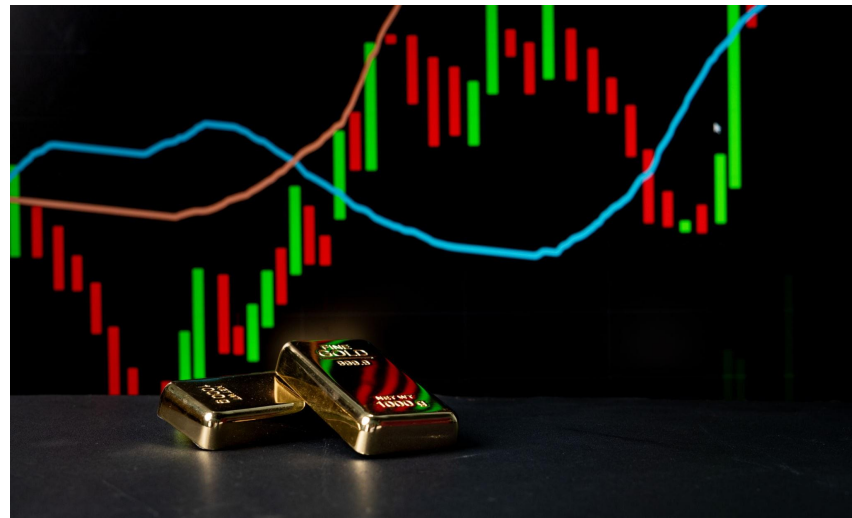
Mi az S&P 500?

- Az S&P 500 index az USA 500 legnagyobb, nyilvánosan jegyzett vállalatok teljesítményét követi.
- Piaci kapitalizációval súlyozott index.
- Szektorálisan diverzifikált: technológia, egészségügy, pénzügy, ipar stb.
- Pénzügyi indikátor: gyakran használják a gazdasági hangulat mérésére.



Miért vizsgáljuk az S&P 500-at?

- **Erőteljes trend és ciklikus mozgás** jellemzi.
- Jelentős **volatilitási klaszterek** figyelhetők meg (pl. válság, FED döntések idején).
- Alkalmas példája a **valós pénzügyi idősortani modellezésnek**.



Elemzés célja

Két fő kérdésre keressük a választ:

1. Hogyan modellezhető az árfolyam- és hozam adatok viselkedése? → ARIMA és GARCH modellekkel.
2. Hogyan hatnak makrogazdasági események, mint a FOMC vagy NFP jelentések? → hozam és volatilitás reakciók elemzése.

Használt módszerek

- ARIMA modellek (trend és autokorreláció vizsgálat)
- GARCH modellek (heteroszkedasztikus volatilitás)
- Eseményhatás-vizsgálat (t-próba)

Adatforrás

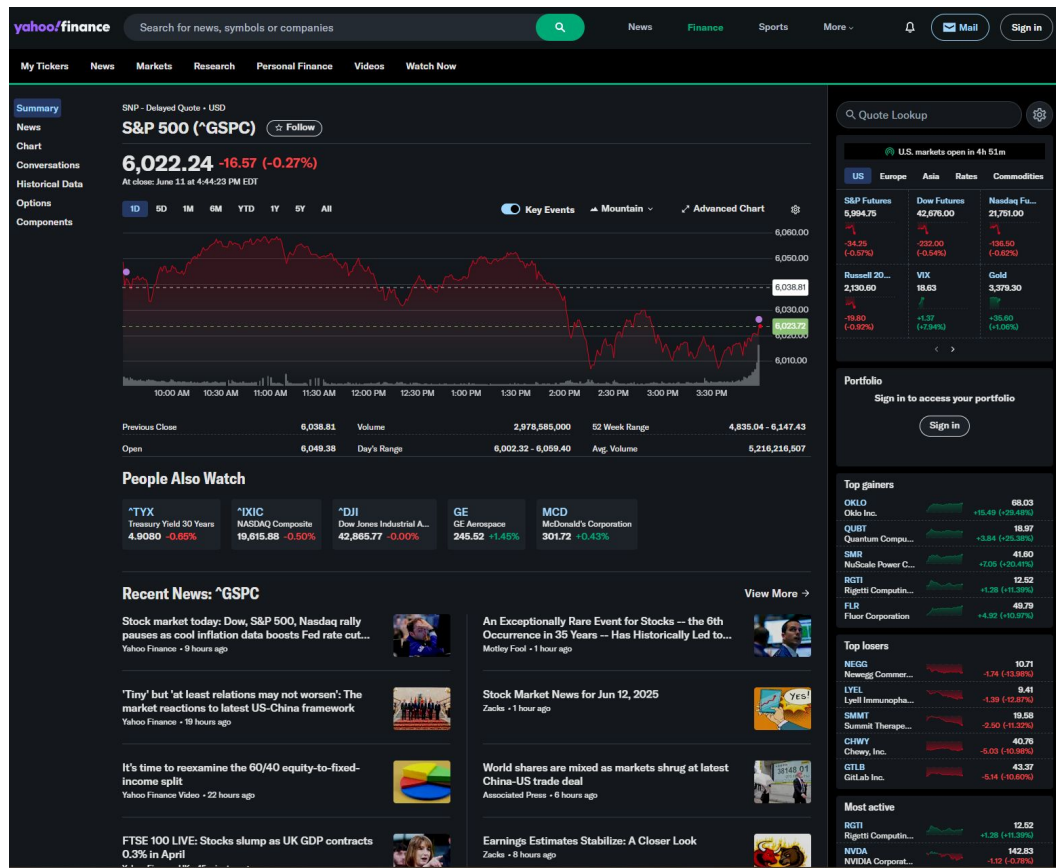
S&P 500 index (ticker: ^GSPC)

- Forrás: Yahoo Finance API

- Időtáv: 2019.01.01 – 2025.01.01

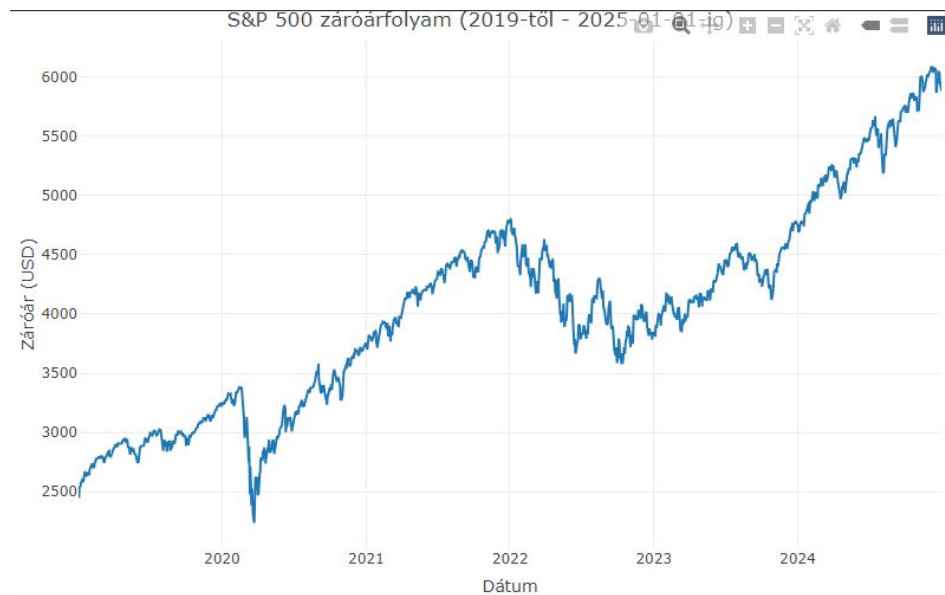
- Periodicitás: napi (daily)

- Töltés: `quantmod::getSymbols()`



A nyers adatok szerkezete

- GSPC.Open – Nyitóár
- GSPC.High – Napi maximum
- GSPC.Low – Napi minimum
- GSPC.Close – **Záróár (ezt használjuk)**
- GSPC.Volume – Forgalom
- GSPC.Adjusted – Osztalékkal korrigált záróár



Az elemzéshez a **záróárat (Close)** és az abból számolt **loghozamot** használjuk.

ARIMA modellezés

Jelölés: ARIMA(p, d, q)

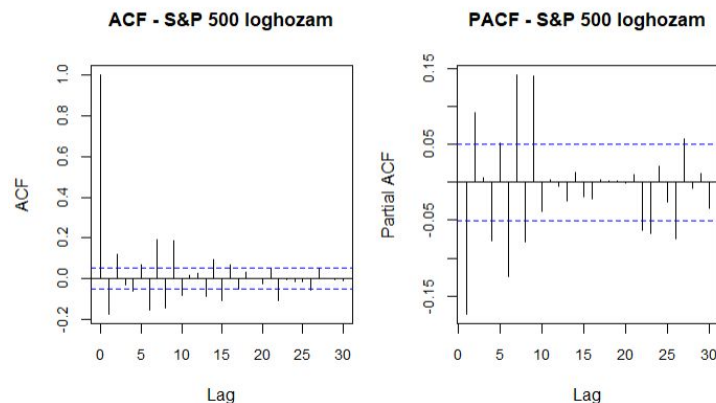
- **AR (p)**: autoregresszív tag – múltbeli értékeken alapul
- **I (d)**: integrált tag – hányadik differenciálás után lesz stacionárius
- **MA (q)**: mozgóátlag tag – múltbeli hibák hatása

Modell kiválasztás lépései

1. **Stacionaritás vizsgálata** → ha nem stacionárius: differenciálás ($d > 0$)
2. **ACF / PACF vizsgálat** → kezdeti modell struktúra meghatározása
3. **Modellek illesztése több konfigurációval**
4. **Diagnosztika és kiválasztás AIC / BIC / reziduum alapján**

S&P 500 idősor vizsgálata

1. Stacionaritás vizsgálata: az idősor nem stacionárius ezért differenciálás szükséges
2. ACF és PACF kezdeti vizsgálata:
 - a. ACF: gyors levágása MA(1) utal
 - b. PACF: lassan csökken és vannak kiugró értékek
3. Modell illesztése
 - a. ARIMA(2, 0, 1)
 - b. ARIMA(3, 0, 1)
4. Diagnosztika és kiválasztás AIC / BIC / reziduum alapján
 - a. AIC alapján az (2, 0, 1) model jobban illeszkedik
 - b. BIC alapján szintén a (2, 0, 1) -es model
 - c. (2, 0, 1) reziduum alapján volatilitás mutatkozik GARCH modell használatos



Modell finomhangolása

Az ACF a reziduumokon több lag szignifikáns, az ARIMA modell nem eliminálta teljesen az autokorrelációt.

Ljung–Box teszt A p-érték nagyon kicsi.

További modellek:

- (2, 0, 2)
- (3, 0, 2)
- (4, 0, 2)

Modell	AIC	BIC	σ^2	LLF	RMSE	MAE	ACF1
ARIMA(2,0,2)	-9553.86 ✓	-9521.55 ✓	0.0001555 ✓	4782.93	0.01245 ✓	0.00843 ✓	0.01819
ARIMA(3,0,2)	-9482.92	-9445.22	0.0001624	4748.46	0.01272	0.00846	0.00203
ARIMA(4,0,2)	-9551.06	-9507.98	0.0001555	4783.53 ✓	0.01244	0.00843	0.000046 ✓

GARCH

A pénzügyi idősorok (pl. részvényhozamok) volatilitása időben nem állandó. Ennek érzékeltetésére 21 napos gördülő ablakban számítjuk ki a loghozamok szórását, amely egy empirikus becslése a napi volatilitásának. Az így kapott sorozat jól szemlélteti a piaci turbulencia időbeli változását.

Ez alapján megalapozott a GARCH-modellek használatát, amelyek a volatilitás időbeli változását formalizáljuk.



GARCH modellezés

Miért nem elég az ARIMA?

- Az ARIMA modellek csak a **hozamok várható értékének** időbeli változását modellezik. De nem számolnak a **volatilitás klasztereződésével** (időben változó szórás)!
- Reziduumból láttuk, hogy:
 - Nem fehér zaj
 - Változó szórás

👉 Szükség van **GARCH (Generalized ARCH)** típusú modellre

Mi a GARCH modell?

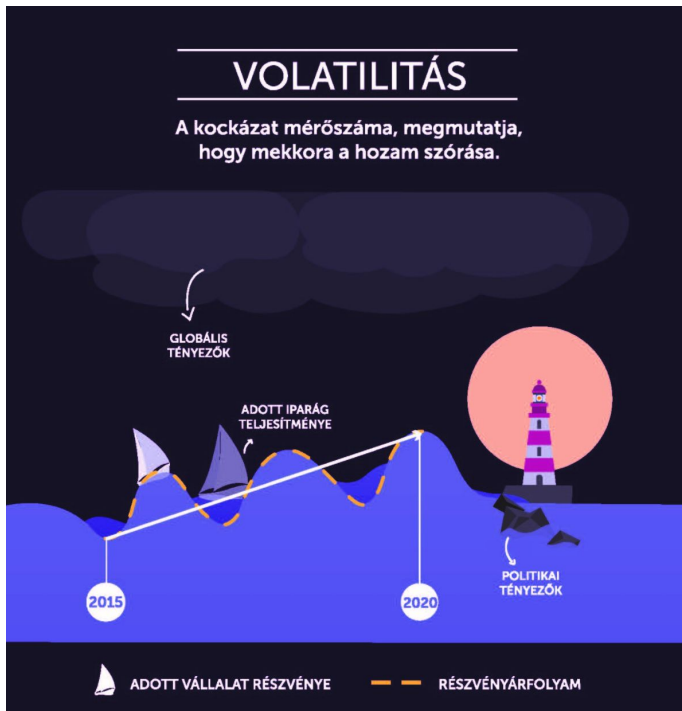
GARCH(p,q): a **reziduum szórásának időbeli dinamikáját** modellezi

- q = ARCH komponens \rightarrow múltbeli négyzetes hibák
- p = GARCH komponens \rightarrow múltbeli varianciák

Adatok előkészítése: diff log hozam 2 dolgozunk!

Alapmodell: **GARCH(1,1)**

Leggyakrabban használt: elegáns,
stabil és jól interpretálható

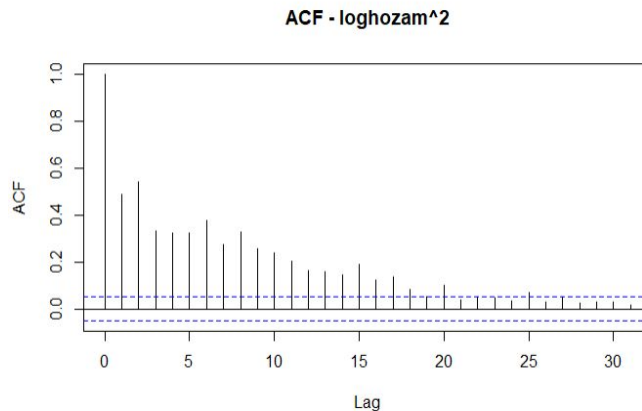


Modell kiválasztás lépései

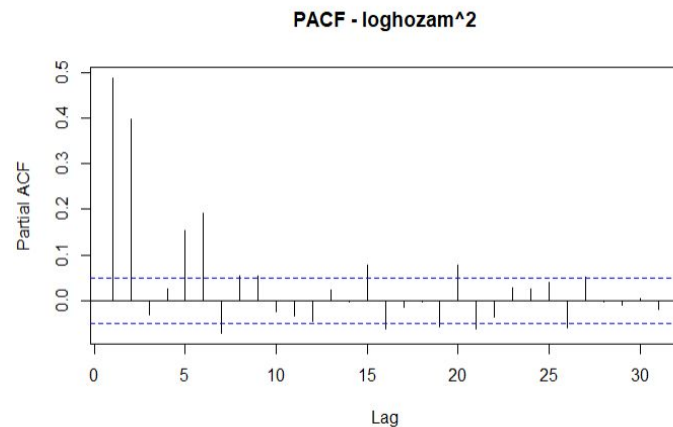
1. Adatok előkészítése ^2
2. ACF és PACF vizsgálata
3. Model illesztés
4. Modellek összehasonlítása

ACF és PACF

- **ACF:** Lassú lecsengés, fokozatos csökkenés. GARCH struktúra (GARCH komponens, azaz $p \geq 1$).



- **PACF:**
GARCH(1,1) Alapmodell, ha gyorsan csökken az ACF
GARCH(1,2) Ha PACF-ben több lag is szignifikáns
GARCH(2,1) Ha ACF-ben 2–3 lag is jelentős marad
GARCH(2,2) Ha mindkét sorozatban több szignifikáns érték látható.



Model illesztés és összehasonlítása

- GARCH(1,1)
- GARCH(1,2)
- GARCH(2,1)
- GARCH(2,2)

Értelmezése

- Akaike (AIC): → A modell illeszkedése + büntetés a paraméterszáma
- Bayes (BIC): → Szigorúbb büntetést ad a komplexitásra
- Shibata (SIC): → AIC továbbfejlesztett, kis mintákra optimalizált változata
- Hannan–Quinn (HQIC): → Átmenet az AIC és BIC között

Konklúzió: az értékek alapján a GARCH(1,1) tűnt a legjobbnak.

Akaike	-6.455461
Bayes	-6.423736
Shibata	-6.455531
Hannan-Quinn	-6.443645

Akaike	-6.454135
Bayes	-6.418885
Shibata	-6.454222
Hannan-Quinn	-6.441007

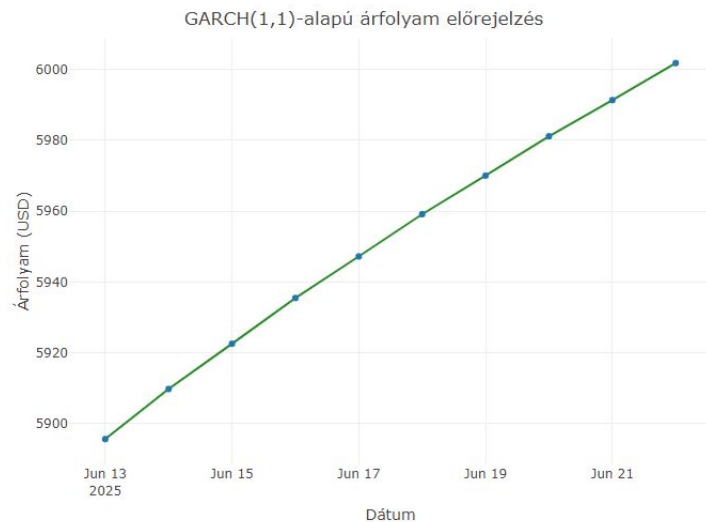
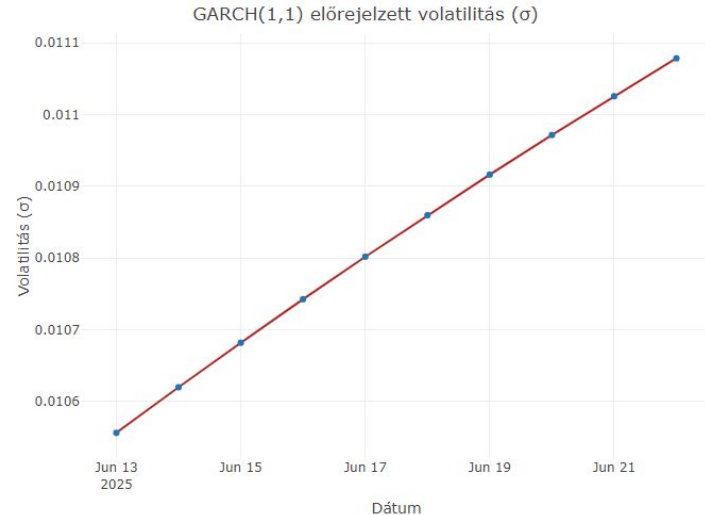
Akaike	-6.454841
Bayes	-6.419591
Shibata	-6.454928
Hannan-Quinn	-6.441713

Akaike	-6.454540
Bayes	-6.415765
Shibata	-6.454645
Hannan-Quinn	-6.440099

Előrejelzés

A piaci szereplők kockázatérzékelése az utóbbi időszakban növekedett, amely – ha tartósan fennáll – hatással lehet az opciók prémiumára és a pozícióméretezésre is.

A várakozások szerint az árfolyam mérsékelt ütemben emelkedik tovább, de a növekvő volatilitás figyelmeztetés lehet a piaci szereplők számára – ez megnövekedett bizonytalanságot vagy potenciális árfolyam-korrekciót jelezhet.



Különböző események hatása az idősorra

A FOMC kamatdöntések napjai körüli hatások

- a log hozam nem tér el szignifikánsan a megszokott értéktől.
- eseményeket megelőző és követő időszakok volatilitása statisztikailag nem különbözik egymástól.

A Non-Farm Payrolls (NFP)

- loghozamok átlaga nem tért el szignifikánsan a piaci referenciahozamtól.
- volatilitás-elemzés alapján nincs statisztikailag szignifikáns bizonyíték arra, hogy a volatilitás megváltozik ezeken a napokon. Bár az átlagos utólagos volatilitás enyhén alacsonyabb, ez a különbség nem elég nagy vagy konzisztens ahhoz, hogy érdemi következtetést vonjunk le.