

# A stacionaritás és az ergodicitás fogalma

Ferenci Tamás  
tamas.ferenci@medstat.hu

Utoljára frissítve: 2023. május 9.

# Tartalom

- 1 A stacionaritás fogalma, szükségessége
- 2 Idősor-jellemzők mintából becslése

# Az alapprobléma

- Egyetlen realizáció problémája
- Hogyan lehet egyetlen mintából bármit megbecsülni? Nyilván sehogy...
- És ezen ráadásul – ugyebár – a hosszabb megfigyelés sem segít
- A megoldás: valamilyen plusz-feltevés kell!

# Jön a jótündér

- Mondjuk egy jótündér megszűgja nekünk, hogy minden időpontban (ami ugye egy valószínűségi változó) ugyanaz a várható érték:

$$\mu_t \equiv \mu$$

- ( $A \equiv$  természetesen azt jelenti, hogy  $\forall t$ -re)
- Ez tehát egyfajta *eltolásinvariancia*; írhattuk volna, teljesen egyenértékűen azt is, hogy  $\mu_t = \mu_s$ , vagy azt is, hogy  $\mu_t = \mu_{t+h}$  (most elég nyakatekert, de később ez lesz a jobb)
- Ekkor már ezt az – immár létező – közös várható értéket, igaz *csak ezt*, de ezt meg tudjuk becsülni mintából!
- Ehhez – de csak ehhez! – „összeönthetők” a különböző időponthoz tartozó megfigyelések (igazából ehhez még valami további dolog is kell, de erről majd később, most fogadjuk el, hogy megvalósítható)
- Hiszen:

$$\hat{\mu} = \frac{1}{T} \sum_{t=1}^T y_t$$

- Menjünk tovább...

## További jótündérek

- Ha  $\sigma_t^2 \equiv \sigma^2$  akkor létezik közös szórásnégyzet, és becsülhető:  $\widehat{\sigma^2} = \frac{1}{T} \sum_{t=1}^T (y_t - \widehat{\mu})^2$  (torzított, de aszimptotikusan torzítatlan, továbbá konzisztens)
- Menjünk tovább...
- Ha létezik az  $F_{Y_i} \equiv F$  közös eloszlás, akkor  $\widehat{F}$  becsülhető (pl. hisztogrammal)
- Menjünk tovább...
- Két változónál is az eltolásinvarianciát akarjuk megtartani:  $\gamma_{t,s}$  legyen ugyanaz mint  $\gamma_{t+h,s+h}$
- De gondoljuk végig, ez magyarul azt jelenti, hogy  $\gamma_{t,s}$  csak a  $t - s$ -től függ!
- Ez esetben becsülhető:  $\widehat{\gamma}_k = \frac{1}{T-k} \sum_{t=1}^{T-k} (y_t - \widehat{\mu})(y_{t+k} - \widehat{\mu})$
- Menjünk tovább...
- $F_{Y_t, Y_s} = F_{Y_{t+h}, Y_{s+h}}$ , minden  $s, t, h$ -ra, akkor a kétváltozós eloszlás komplettül becsülhető

## És végül...

- Az utolsó, immár tényleg mindent lefedő szint:  $F_{Y_{t_1}, Y_{t_2}, \dots, Y_{t_k}} = F_{Y_{t_1+h}, Y_{t_2+h}, \dots, Y_{t_k+h}}$  minden értelmes  $k$ -ra,  $t_1, t_2, \dots, t_k$ -ra és  $h$ -ra
- Ennek a neve: **erős stacionaritás**
- Ez az a feltevés, ami a keresztmetszethez hasonló becslhetőséget tesz lehetővé annak ellenére is, hogy idősorban vagyunk
- A keresztmetszet azért volt egyszerűbb, mert a függetlenséget feltettük, de itt most erről szó nincs: pont az a lényeg, hogy úgy teremtettük meg a keresztmetszethez hasonló becslhetőséget, hogy semmilyen függetlenséget nem kellett feltételeznünk – szerencsére, mert annak ugye nem lenne értelme (persze a függetlenség implikálja az erős stacionaritást)
- Viszont: rengeteget követel, elméleti kezeléshez jó, de gyakorlatban nagyon nehéz ellenőrizni a teljesülését
- Éppen ezért gyengítsük kicsit...

# A gyenge stacionaritás

- Mint az erős stacionaritás, de
  - csak  $k = 1, 2$ -re
  - teljes eloszlás-egyezőség helyett csak első két momentumban egyezőséget követelünk meg
- Mi adódik ebből?
  - 1  $\mu_t \equiv \mu$
  - 2  $\sigma_t^2 \equiv \sigma^2$
  - 3  $\gamma_{t,s} \equiv \gamma_{t-s}$
- (Valójában a 2. feltétel redundáns, hiszen  $\gamma_{t,t} = \sigma_t^2$ )
- A továbbiakban, ha mást nem mondunk, stacionaritás alatt mindig ezt a **gyenge (vagy kovariancia-) stacionaritást** értjük

# A stacionaritás tesztelése

- A fentiek már adnak egy – teljesen szubjektív – módszert a stacionaritás tesztelésére: nézzünk rá az idősorra, az 1. és 2. feltétel megítélhető
- Ez a „grafikus tesztelés” persze abszolút szubjektív
- Később látni fogunk objektív módszert (statisztikai próbát) is



# Egy gondolat a mintából történő becslésekről

- A már látott mintából történő becsléseknél ( $\widehat{\mu}$ ,  $\widehat{\sigma}^2$ ,  $\widehat{\gamma}_k$  stb.) ne feledjük el, hogy mindezeket mind terheli a mintavételi ingadozás, az abból fakadó mintavételi hiba
  - ...konfidenciaintervallum szerkeszthető a valódi értékre
  - ...tesztelhető nevezetes értékre vonatkozó hipotézis
- Ez utóbbi tipikus példája az autokorrelálatlanság ( $H_0 : \rho_k = 0$  vs.  $H_1 : \rho_k \neq 0$ )
- Nem túl kis mintaméretnél már jó a normális approximáció:  $\widehat{\rho}_k \sim \mathcal{N}(\rho_k, 1/T)$ , így

$$\frac{\widehat{\rho}_k}{1/\sqrt{T}} \stackrel{H_0}{\sim} \mathcal{N}(0, 1)$$

# Autokorrelálatlanság tesztelése: Ljung–Box-teszt

- Nagyon sok esetben grafikusán is feltüntetik a korrelogramon az autokorrelációra vonatkozó  $\pm z_{1-\alpha/2} \frac{1}{\sqrt{T}}$  kritikus értékeket
- De vigyázat: ez a *teljes* autokorrelálatlanság ( $H_0 : \rho_1 = \rho_2 = \dots = \rho_M = 0$ ) tesztelésére nem alkalmas!
- (Többszörös összehasonlítások helyzete,  $\alpha$ -infláció)
- Legnépszerűbb teszt erre: Ljung–Box-teszt:

$$Q = T(T+2) \sum_{k=1}^M \frac{\hat{\rho}_k^2}{T-k} \stackrel{H_0}{\sim} \chi_M^2$$

- $M$  megválasztása kérdés (ha túl kicsi, elnézhetünk egy magasabbrendű autokorrelációt, ha túl nagy, eltérhetünk a  $\chi^2$  eloszlástól)!
- Tipikus alkalmazás majd: modelldiagnosztikában

# Autokorrelálatlanság tesztelése: LM-tesztek

- A Ljung–Box-teszttel komoly elméleti agályok vannak (ld. Maddala, 13.5 vagy Hayashi 2.10)
- Ennek ellenére teljesen általánosan használják...
- Lehetséges alternatíva: LM-elvű tesztek, például a Breusch–Godfrey-teszt a már említett modelldiagnosztikai helyzetben