**BSc Matematika Alapszak, 2017.**

*Matematikai Intézet,*
*Természettudományi Kar,*
*Eötvös Loránd Tudományegyetem.*

# Idősorok és többdimenziós statisztika

- ***Óraszám (ea+gy):*** 2 + 2
- ***Specializáció:*** elemző
- ***Kredit (ea+gy):*** 3 + 2
- ***Számonkérés:*** kollokvium + gyak. jegy
- ***Tárgykód (ea, gy):*** idosor1e0_m17ea, idosor1e0_m17ga
- ***Ajánlott félév:*** 5
- ***Státusz:*** kötelező

## Tantárgyfelelős

- Márkus László, Valószínűségelméleti és Statisztika Tanszék, Matematikai Intézet.

## Előfeltételek

### *A gyakorlat előfeltételei:*

- *Erős:* Leíró és matematikai statisztikaE-e (leiros1e0_m17ea)
- *Erős:* Analízis3G-m (analiz3m0_m17ga) *vagy*
  Analízis3G-a (analiz3a0_m17ga) *vagy*
  Kalkulus3E-e (kalkul3e0_m17ea)

### *Az előadás előfeltételei:*

- *Gyenge:* a gyakorlat

## Megjegyzések

- ***Pótlási lehetőség:*** A félév végén, indokolt esetben, a gyakorlatvezető döntése alapján egy javító zárthelyi dolgozat írására van lehetőség.

### A tematikát kidolgozta:

- Márkus László, Valószínűségelméleti és Statisztika Tanszék, Matematikai Intézet.

## Szükséges előismeretek

- **Valószínűségszámításból:** Várható érték, momentumok, szórásnégyzet, kovariancia és korreláció. Nevezetes eloszlások. Konvergenciatípusok. Nagy számok törvényei. Centrális határeloszlás tétel. Feltételes valószínűség és feltételes várható érték.
- **Leíró és matematikai statisztikából:** Statisztika alapfogalmai: középértékek, kvantilisek, szóródási mérőszámok. Idősorok alapfogalmai. korrelációszámítás. Mintavétel alapfogalmai. Statisztikai becslések, konfidenciaintervallumok. Hipotézisvizsgálat alapfogalmai, normális eloszlásra vonatkozó próbák, chinégyzet próbák. Lineáris regresszió.
- **Kalkulus3-ból:** Fourier sorok. A Fourier sorok alkalmazásai.

## A tantárgy célkitűzése

A tárgy célja az idősorok alapvető fogalmainak, elemzésük legegyszerűbb módszereinek bemutatása, valamint a többdimenziós statisztika legelemibb alapismereteinek elsajátíttatása a hallgatókkal. Az akkreditált tematikában szereplő tananyag keretein belül elsősorban az alapvető fogalmakat, tételeket, módszereket tárgyaljuk igen részletesen. Ezért a tárgy elsősorban az elemző specializáció igényeit tartja szem előtt.

# Irodalom

***Ajánlott:***

- **Feller W.:** *Bevezetés a valószínűségszámításba és alkalmazásaiba.* Műszaki Kiadó, 1978.
- **S. Karlin-H. Taylor:** *Sztochasztikus folyamatok.* Gondolat Kiadó, 1985.
- **Michelberger-Szeidl-Várlaki:** *Alkalmazott folyamatstatisztika és idősor analízis.* Typotex, 2001.
- **Brockwell, Davis:** *Introduction to time series and forecasting.* Springer. 1996.
- **Rozanov, Yu. A.:** *Probability Theory, Stochastic Processes and Mathematical Statistics.* 1997
- **C. R. Rao:** *linear statistical inference and its applications.* Wiley, 1965.

# Tematika

- Néhány jól ismert speciális folyamatot (Poisson, Yule, Galton-Watson folyamatok) ismertetünk elsőként. A modellleírások mellett véletlen időpontokban bekövetkező események segítségével is származtatjuk ezeket, és meghatározzuk az ehhez tartozó eloszlásokat. A Galton-Watson folyamatok esetében tárgyaljuk a kihalási valószínűség meghatározásának módját. E folyamatok bemutatásán keresztül érintjük a Markov folyamatok elméletének alapfogalmait is.
- Stacionárius folyamatok, alapfogalmainak bevezetésével folytatjuk a tárgyat. A gyenge, erős valamint a k-adredű stacionaritás és az ergodicitás fogalmát és kapcsolataikat tisztázzuk. Az összefüggési struktúra leírására bevezetjük az autokovariancia, autokorreláció és parciális autokorreláció függvényeket ismertetjük tulajdonságaikat, és megemlítjük a dinamikus kopulákat, mint egy újabban előtérbe került lehetőséget. Áttérve a frekvenciatartományra heurisztikus értelmezését adjuk a stacionárius idősor Fourier előállításának. Megadjuk a spektrálelőállítást egzakt formában is, kimondjuk a Herglotz tételt, de nem bizonyítjuk.
- Ezek után definiáljuk az autoregressziós (AR(p)), a mozgóátlag (MA(q)), valamint általánosan az integrált autoregressziós mozgóátlag ARIMA(p,d,q) folyamatot. Feltételt adunk a stacionárius megoldás létezésére, kiszámítjuk fontos speciális esetekben a szórást az autokovariancia függvényt, valamint a spektrálsűrűség függvényt. Rámutatunk, hogy az invertálható ARIMA folyamatok lineáris folyamatok. Bevezetjük az ARCH folyamatokat mint az egyik legegyeszerűbb nemlineáris folyamatot. Megvizsgáljuk az ARCH(1) folyamat stacionárius megoldása létezésének feltételét és a megoldás konstrukcióját, ha a megoldás szórása véges. Továbblépve, általánosan, sztochasztikus rekurziós egyenletek esetén Ljapunov-exponenssel adjuk a stacionárius megoldás létezésének feltételét és kimondjuk a Kesten-Vervaat-Goldie tételt reguláris változású eloszlással bíró stacionárius eloszlás létezéséről. Ezek után térünk vissza az ARCH(1) folyamathoz és áltlánosan második momentum végességének feltételezése nélkül is megadjuk a stacionárius megoldás létezésének feltételét. Ismertetjük a GARCH folyamatokra vonatkozó eredményeket is. Ugyancsak tárgyaljuk ezt a témakört a bilineáris folyamatokra is. A további nemlineáris modellek közül megemlítjük a véletlen együtthatós AR, és a SETAR modelleket.
- Idősorok becsléselméletének részeként először a várható érték becslésével foglalkozunk. A legjobb lineáris becslés kiszámítása után elemezzük az átlag viselkedését a spektrálmérték tulajdonságai függvényében. Megmutatjuk, hogy ha a $\{0\}$-ra koncentrált spektrálmérték pozitív, akkor az átlag még csak nem is konzisztens. Az autokorreláció függvény többféle becslését vizsgáljuk, és rámutatunk, hogy a torzítatlanság helyett fontosabb, hogy az elméletileg helyes pozítív definit tulajdonsággal rendelkezzen a becslés valamint, hogy a becslés szórása ne növekedjék a „lag", a késleltetés növekedtével. Számítjuk a becslés torzítását, szórását, és határeloszlását. Bemutatjuk a periodogrammot, mint a diszkrét spektrum becslését. A spektrálsűrűségfüggvény becslésére az ablakolás eljárását ismertetjük röviden.
- A többdimenziós normális eloszlást definiáljuk, megadjuk sűrűségfüggvényét majd rátérünk a várható érték és a szórásmátrix becsléseinek tulajdonságaira. A regresszió feladatával folytatjuk, megadjuk a legkisebb négyzetes becslést, kimondjuk a Gauss Markov tételt. A regresszió „jóságát" jellemző R2 érték és F-próba tárgyalása után rátérünk a magyarázó változók beválasztásának, vagy redukálhatóságának kérdéskörére, definiáljuk a toleranciát és alkalmazzuk a parciális korrelációt. A regresszió bizonytalanságát mérjük a Mahalanobis távolsággal. Outlierek detektálására bevezetjük a Cook távolság fogalmát. A továbbiakban még röviden megemlítjük az egyszempontú szórásanalízis feladatát, valamint kontingenciatáblák elemzését.