



Közgazdasági adatok elemzése

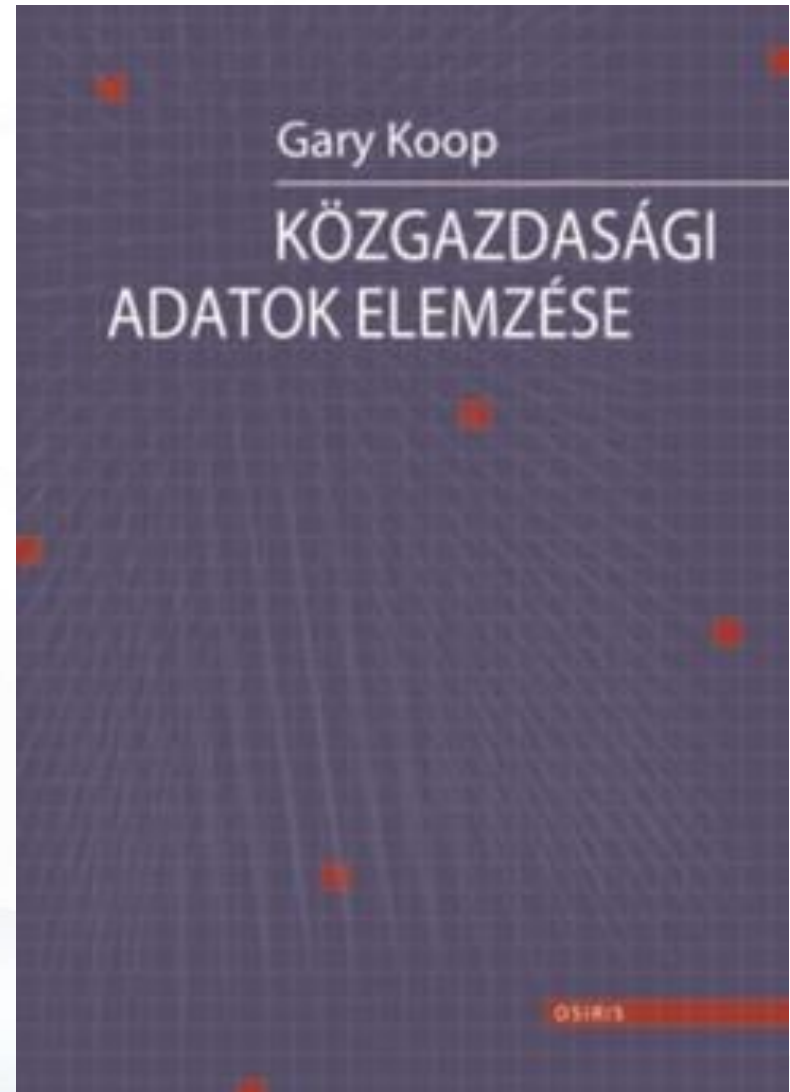
1. óra

Bemutatókozás, a kurzus menetének ismertetése,
alapfogalmak ismétlése, rövid elméleti bevezető

- Tóth Marcell László, PhD-hallgató
- Regionális- Tudományi és Közpolitikai Tanszék
- toth.marcell@sze.hu

- (Közgazdasági) adatok elemzése
- MS Excel program, és annak ismerete szükséges
- A tananyag: korreláció, regresszió, és ezek különböző változatai

- Gary Koop: Közgazdasági adatok elemzése (Osiris Kiadó, Budapest, 2008)



- Középérték
- (Számítási közép, mértani átlag, harmonikus közép, négyzetes közép)
- A vizsgált adatsort „leginkább leíró” egyetlen szám.
- =ÁTLAG

- Átlagtól való átlagos eltérés
- Az átlaggal együtt van értelme használni.
- A szórás négyzetét olyan gyakran használják a valószínűségszámításban és a matematikai statisztikában, hogy önálló fogalomként, mint szórásnégyzet vagy variancia is szoktak rá utalni.
- =SZÓRÁS

- **Módusz:** egy adatsor vagy sorozat leggyakrabban előforduló eleme
- =MÓDUSZ

- **Medián:** helyzeti középérték. Az adatok számértékű sorba-rendezése utáni középső elem.
- =MEDIÁN

- Két változó közötti összefüggések vizsgálatának egyik legfontosabb módszere.
- A korrelációt „ r ”-el jelöljük.
- Értéke -1 és 1 közé esik.
- =KORREL

- Kettőnél több változó közötti összefüggések vizsgálatának eszköze.
- Ha csak két változó van, akkor egyváltozós regresszióról beszélhetünk. Ha több, akkor többváltozós regresszió.
- Ok-okozati összefüggések feltárására is jó.

Köszönöm a figyelmet!



Közgazdasági adatok elemzése

2. óra

Egyváltozós regresszió, és annak típusai, regressziós „egyenes” és trendvonal, tömbképlet, korreláció és regresszió kapcsolata

- Átlag (=ÁTLAG)
- Szórás (= SZÓRÁS)
- Módusz (=MÓDUSZ)
- Medián (=MEDIÁN)
- Korreláció (=KORREL)

- Kettőnél több változó közötti összefüggések vizsgálatának eszköze.
- Ha csak két változó van, akkor egyváltozós regresszióról beszélhetünk. Ha több, akkor többváltozós regresszió.
- Függő és független (magyarázó) változó.
- Ok-okozati összefüggések feltárására is jó.

- Miért kell több változó?
- Bonyolultabb eseménysor
- Makrogazdasági összefüggések
- Ha egy tényező több dologtól is függ

- Induljunk ki egy pontdiagramból.
- Két pont esetén: tökéletes az illeszkedés
- Három vagy több pont esetén: egyre nagyobb a valószínűsége, hogy az egyenes nem érint minden pontot
- Regressziós „egyenes” – lineáris, exponenciális, logaritmikus, polinomiális, hatványos, mozgóátlagos

- Az illeszkedés mérőszáma
- Szoros kapcsolatban van az X és Y közötti korrelációval (egyváltozós regresszió esetén R^2 megegyezik az X és Y közötti korreláció négyzetével)
- =RNÉGYZET
- Trendvonalnál is számítja az Excel

- =LIN.ILL
- Tömbképlet!
- =TREND
- Pontdiagram használata
- Regressziós egyenes felállítása, és az egyenlet megjelenítése

Köszönöm a figyelmet!



Közgazdasági adatok elemzése

3. óra

Többváltozós regresszió, és annak típusai, értelmezései,
tömbképlet, korrelációs mátrix

- Egyváltozós regresszió
- =LIN.ILL
- Pontdiagram
- Trendvonal
- Korreláció (=KORREL)

- Kettőnél több változó közötti összefüggések vizsgálatának eszköze.
- Ha csak két változó van, akkor egyváltozós regresszióról beszélhetünk. Ha több, akkor többváltozós regresszió.
- Függő és független (magyarázó) változó.
- Ok-okozati összefüggések feltárására is jó.

- Miért kell több változó?
- Bonyolultabb eseménysor
- Makrogazdasági összefüggések
- Ha egy tényező több dologtól is függ

- Egy függő változó
- Egynél több független/magyarázó változó
- Regressziós egyenes / trendvonal felírása akadályokba ütközik.

- Az illeszkedés mérőszáma
- Szoros kapcsolatban van az X és Y közötti korrelációval (egyváltozós regresszió esetén R^2 megegyezik az X és Y közötti korreláció négyzetével)
- $=R^2$
- 0 és 1 közötti érték

- =LIN.ILL
- Tömbképlet!
- =TREND
- Pontdiagram használata
- Regressziós egyenes felállítása, és az egyenlet megjelenítése

- Adatok / Elemzés / Adatelemzés

Köszönöm a figyelmet!



Közgazdasági adatok elemzése

4. óra

Regresszió kétértékű változókkal, dummy változó,
egyváltozós regresszió, többváltozós regresszió

- Egyváltozós regresszió
- =LIN.ILL

- Többváltozós regresszió
- Adatok / Elemzés / Adatelemzés

- Korreláció (=KORREL)

- Kettőnél több változó közötti összefüggések vizsgálatának eszköze.
- Ha csak két változó van, akkor egyváltozós regresszióról beszélhetünk. Ha több, akkor többváltozós regresszió.
- Függő és független (magyarázó) változó.
- Ok-okozati összefüggések feltárására is jó.

- Kértétékű változó
- A változó két értéket vehet fel
- Például:
 - Igen / Nem
 - Férfi / Nő
 - Valami van /nincs
 - 0 / 1

- Egyváltozós vagy többváltozós regressziós modell, ahol az egyik (vagy több) magyarázó változó dummy értékeket vesz fel
- Excelben ezeket a változókat az esetek döntő többségében 0 és 1 értékekkel látjuk el

- Előfordulhat olyan eset, hogy a függő változón (tehát az Y) a dummy.
- Ezeket a modelleket ez Excel nem tudja megfelelően kezelni, így ilyen számításokat nem végzünk jelen kurzus keretein belül.

Köszönöm a figyelmet!



Közgazdasági adatok elemzése

5. óra

Regresszió időbeli késleltetéssel, keresztmetszeti adatok,
idősorok, osztott késleltetésű modell

- Egyváltozós regresszió
- =LIN.ILL

- Többváltozós regresszió
- Adatok / Elemzés / Adatelemzés

- Korreláció (=KORREL)

- Kettőnél több változó közötti összefüggések vizsgálatának eszköze.
- Függő és független (magyarázó) változó.
- Ok-okozati összefüggések feltárására is jó.
- A magyarázó változó lehet keresztmetszeti adat, dummy változó és (bizonyos fajta) idősor is.

- Olyan statisztikai megfigyelések, amelynek elemeit egymást követő időpontokban (időszakokban) regisztrálták, és ez az időbeliség az adatok fontos tulajdonsága.
- Pl.: áruforgalom, szálláshelyek kihasználtsága, egyetemi kocsmák napi bevétele, GDP, stb.

- Eddig főleg keresztmetszeti adatokkal dolgoztunk.
- Idősor esetén lehetséges, hogy bizonyos késleltetéssel érezteti a hatását.
- Ha nem megfelelő idősort használunk, akkor a hamis regresszió esete állhat fent.

- Eddig főleg keresztmetszeti adatokkal dolgoztunk.
- Idősor esetén lehetséges, hogy bizonyos késleltetéssel érezteti a hatását.
- Ha nem megfelelő idősort használunk, akkor a hamis regresszió esete állhat fent.

- Fontos a késleltetés jó kiválasztása.
- Az idősoros adatok hatása lehet, hogy később jelentkezik, DE:
 - Ésszerű keretek között érdemes megválasztani, túl sok késleltetés nem feltétlenül jó
 - Figyelembe kell venni az adatsorunk hosszát, mert a késleltetés egyúttal az elemszám csökkentése is.

- $$Y_t = a + \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \dots + \beta_n X_{n-t} + e_t$$

Köszönöm a figyelmet!