## Tradicionális statisztikai módszerek

### Regresszió

#### Lineáris regresszió

Ha több folytonos változó lineáris kapcsolatban van egymással, akkor az egyik csoport segítségével (magyarázó változók) előre jelezhetjük a másik csoport értékét (eredmény változók). Szükségünk van a függő és független változó kiválasztására, de ez nem jelent oksági kapcsolatot! (1)

Az összefüggés segítheti a megértését a kapcsolatnak és legfőképp releváns előrejelzéseink lehetnek. Sajnos ez nem ilyen egyszerű.

Kétváltozós eset:

Y az eredmény változó

X a magyarázó változó

és a regressziós együtthatók, míg az a véletlen változó.

A lehető legkisebb hibájú becslés a cél. A hibáról feltételezzük, hogy független a magyarázó változótól és átlaga nulla. A becslés csak tökéletes kapcsolat esetén lenne hibamentes.

Becslési módszerek pl.:

* Legkisebb négyzetek módszere
* Momentumok módszere
* Legnagyobb valószínűségek módszere

##### Legkisebb négyzetek módszere

A **legkisebb négyzetek módszere** a mérések matematikai feldolgozásában használt eljárás. Nevét arról kapta, hogy az eltérések négyzetösszegét igyekszik minimalizálni.

A módszer érzékeny a nagyon kilógó adatokra. Egy kilógó adat az egész eljárás eredményét megváltoztathatja, hamis képet adva az adatsorról. Különböző statisztikai tesztekkel szűrik az adatsort, hogy ne maradjanak benne mérési hibák. A kilógó adatokat elhagyják, vagy a kívülállókra kevésbé érzékeny módszerekkel alternatív becsléseket végeznek. Ilyen például a súlyozott regresszió, amiben a kívülálló adatok súlyát, és ezzel befolyását is csökkentik.

##### Elaszticitás

A rugalmasság mérőszáma. Azt fejezi ki, hogy az x magyarázó változó 1% változása hány %-os változást okoz az eredmény változóban.

### Mozgó átlagolás

A mozgóátlagok számítása az idősorok hosszabb távú elemzésének legegyszerűbb módja. Csak annyit tűzünk ki célul, hogy átlagolással kiszűrjük a durva, egészen rövid távú ingadozásokat. Végrehajtása matematikailag igen egyszerű, jól alkalmazkodik az idősor jelleméhez. Hátránya hogy az idősor megrövidül, és jól kell, megválasztani az átlagolandó tagok számát különben torzít. Simítja az idősort de az extrém értékek erősen befolyásolják. Mindig páratlan számú elemet tudunk átlagolni mert az adott tagnak az átlagát úgy számítjuk hogy vesszük az adatott tagot és előtte egy utána ugyan annyi számú további tagot.

Lehet számtani vagy kronologikus átlagot számolni.

### ARIMA modell

AutoReggressieve Integrated Moving Average

Gyakran használják idősorok elemzéséhez és előrejelzéséhez

## Irodalom jegyzék:

<https://www.mateking.hu/statisztika-2/regresszioszamitas/>

<https://hu.wikipedia.org/wiki/Line%C3%A1ris_regresszi%C3%B3>

<http://psycho.unideb.hu/munkatarsak/balazs_katalin/stat1/stat1ora4.pdf>

<https://hu.wikipedia.org/wiki/Legkisebb_n%C3%A9gyzetek_m%C3%B3dszere>

<http://www.cs.bme.hu/nagyadat/bartok-ferenc.pdf>

<http://ilias.gdf.hu/data/ilias-ha/lm_data/lm_9370/index.html>

Wiki\_LkNegy, 1

1. **Wikipedia. [Online] https://hu.wikipedia.org/wiki/Legkisebb\_n%C3%A9gyzetek\_m%C3%B3dszere.**