



**STATOKOS JEGYZET: TÖBBVÁLTOZÓS STATISZTIKA:
DISZKRIMINANCIA ANALÍZIS ÉS LOGISZTIKUS REGRESSZIÓ**

StatOkos – Statisztikai és Módszertani Adatbázis

• 2018

TÖBBVÁLTOZÓS STATISZTIKA: DISZKRIMINANCIA ANALÍZIS ÉS LOGISZTIKUS REGRESSZIÓ

szerzők: Kazinczi Csaba, Alter Emese

A diszkriminancia analízis

- A diszkriminancia analízis a prediktív jellegű eljárások közé sorolandó.
- Fontos kiindulópontja, hogy a populációnknak vannak bizonyos állandó tulajdonságai, melyek akkor is megmaradnak, ha a tagok között bizonyos mértékű fluktuáció van.
- A diszkriminancia analízis esetében a függő változónk egy csoportosításra alkalmas változó, a független változónk pedig olyan paraméteres változók, melyek valamilyen módon jellemzik a minta tagjait, így az ezekből kinyerhető információk alapján nagy valószínűséggel helyesen be tudjuk sorolni őket a függő változó által megadott kategóriákba.
- Lehet két/többváltozós: Kétváltozós esetén a függő változó dichotóm, vagyis 2 csoportra osztja a mintát, több kategória esetén többváltozós diszkriminancia analízisről beszélhetünk.

A logisztikus regresszió

- Célja a diszkriminancia analízishez hasonlóan a minta tagjainak a kategorikus függő változó által meghatározott csoportjaiba való helyes besorolás megadott független változók alapján, ebben az esetben azonban nem csak paraméteres változók lehetnek független változók.
- 2 fő típusa: Bináris logisztikus regresszió: Ekkor a függő változó dichotóm, vagyis 2 csoportba sorolhatjuk a személyeket a független változók alapján. Többváltozós logisztikus regresszió: A függő változó 2-nél több kategóriával rendelkezik, melyekbe besorolhatjuk a személyeket a független változók alapján.