



# STATOKOS JEGYZET: LEÍRÓ STATISZTIKA

StatOkos – Statisztikai és Módszertani Adatbázis

2018

# LEÍRÓ STATISZTIKA

szerzők: Kazinczi Csaba, Holczer Adrienn, Alter Emese

## Tárgymutató

- Skálatípusok, mérési szintek
- Középérték mutatók
- Szóródási mutatók
- Eloszlás

## **Skálatípusok, mérési szintek**

### Nominális változók

- névleges, minőségi különbség az egyes kategóriák között
- a különbség az adatok között nem mérhető, a különböző kategóriák nem rendezhetők sorrendbe
- a két kategórás nominális változót (pl.: nem) dichotóm változóknak nevezzük
- ilyen változó például: nem (nő/férfi), egyetemi szak, stb.
- elemzéskor használható eljárások: leíró statisztikai eljárások, pl.: eloszlás vagy a leggyakoribb érték (módusz) vizsgálata

### Ordinális változók

- a különbség inkább minőségi, mint mennyiségi, nem mérhető pontosan
- az egyes kategóriák sorba rendezhetők
- ilyen változó például: jövedelemkategóriák, iskolai végzettség, versenyen elért helyezés
- használható középérték mutatók: medián, módusz

### Metrikus változók

- mennyiségi különbség az egyes értékek között
- az értékek közti távolság mérhető, nem változik
- matematikai műveletek elvégzésére is alkalmas
- 2 altípus: intervallum-és arányskála
- intervallum skála: nincs természetes 0 pont
- arányskála: rendelkezik természetes 0 ponttal
- középérték mutatók: átlag, medián és módusz egyaránt használható

## **Középérték mutatók**

### Átlag

- csak metrikus változók esetében alkalmazható
- kiugró értékek könnyen torzíthatják, ilyenkor érdemes mediánt használni
- képlet

### Medián

- középső érték
- ordinális és metrikus változók esetén alkalmazható

- ha az átlag az eloszlás miatt nem nyújtana megfelelő információt a minta tulajdonságáról, helyette használhatjuk a mediánt (egyres nemparaméteres próbák is a mediánnal számolnak átlag helyett)

### Módusz

- a leggyakrabban szereplő érték
- lehet 1 érték, ekkor unimodális módusznak nevezzük
- amikor több érték is ugyanannyi alkalommal szerepel, és ezek szerepelnek a legtöbbször, több módusunk lesz, ezt multimodális módusznak nevezzük
- nominális, ordinális és metrikus változók esetén is alkalmazható

### **Skálátípusok és középérték mutatók**

A skála típusa	Középérték mutatók (használható-e)		
	Átlag	Medián	Módusz
Nominális	NEM	NEM	<b>IGEN</b>
Ordinális	NEM	<b>IGEN</b>	<b>IGEN</b>
Metrikus	<b>IGEN</b>	<b>IGEN</b>	<b>IGEN</b>

### **Szóródási mutatók**

- a középérték körüli koncentrálódást, az ingadozás mértékét adják meg

### A terjedelem

- a legnagyobb és legkisebb érték különbsége
- metrikus skálák esetében használható

### A szórás

- a középértéktől való eltérés átlagát adja meg
- a középérték +- 1 szórás között található a legtöbb érték, az ezen kívül eső értékek kiugrónak minősülnek
- kiszámításának lépései: 1: átlag kiszámítása, 2: Egyes adatok kivonása az átlagból, 3: Kapott különbségek négyzetre emelése, 4: Négyzetre emelt értékek átlagának kiszámítása, 5: Gyökvonás a kiszámított átlagból

### **Eloszlás**

#### Normáleloszlás

- Gauss-görbe vagy haranggörbe néven is ismert
- természetes változók esetén szokott előfordulni
- átlag +-1 szórás távolságra esik a minta 68%-a, átlag +-2 szórásra pedig a minta 95%-a
- a paraméteres próbák előfeltétele a normáleloszláshoz való közelítés
- nem normális eloszlás: ferdeség és csúcsosság

#### Ferdeség

- az eloszlás x tengelyen való jobbra vagy balra tolódása
- balra tolódás esetén pozitív ferdeség
- jobbra tolódás esetén negatív ferdeség

- a módusz, medián és átlag torzításával jár

#### Csúcsosság

- a normáeloszlástól y tengelyen való eltérés
- a normáeloszlás csúcsosságának értéke nulla
- normáeloszlásnál nagyobb csúcsosság: leptokurtikus, normáeloszlásnál kisebb mértékű csúcsosság, laposság: platykurtikus csúcsosság
- normáeloszláskor tapasztalható csúcsosság: mezokurtikus