

Centrális Határeloszlás-tétel (CHT)

Matematikai Statisztika

Czibik Lóránt Patrik

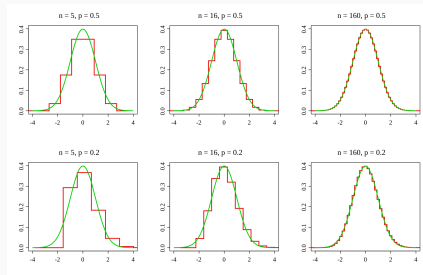
2025. február 24.

Miskolci Egyetem

Bevezetés

A Centrális Határeloszlás-tétel lényege:

- Nagyméretű minták átlaga közelít a normális eloszláshoz.
- Fontos a statisztikában és a valószínűségszámításban.
- Megmagyarázza, hogy miért találkozunk olyan gyakran normális eloszlású adatokkal a valós életben.



ábra 1: A normális eloszlás közelítése binomiális eloszlásokkal

A tétel definíciója

A Centrális Határeloszlás-tétel definíciója

Definíció: Legyenek $\xi_1, \xi_2, \dots, \xi_n$ független, azonos eloszlású valószínűségi változók, melyek várható értéke μ és szórása σ . Ekkor a megfelelően normált összeg:

$$Z_n = \frac{\sum_{i=1}^n \xi_i - n\mu}{\sigma\sqrt{n}} \quad (1)$$

$n \rightarrow \infty$ esetén standard normális eloszlást követ: $Z_n \sim N(0, 1)$.

Alkalmazások és példák

- **Statisztikai elemzés:** Segít megérteni, hogyan viselkednek az átlagok nagyobb minták esetén, például közvélemény-kutatásokban.
- **Minőségellenőrzés:** Gyárakban segít ellenőrizni, hogy a termékek mérete és minősége egyenletes maradjon.
- **Pénzügy:** Használják annak elemzésére, hogyan változnak az árfolyamok és egyéb pénzügyi mutatók az idő múlásával.

Magyarázatot ad arra, hogy sok természetes és mesterséges jelenség **miért** követ **normális eloszlást**:

- Az emberek magassága
- Az emberek napi kalóriabevitele
- Egy gyártósoron készült alkatrészek mérete (pl. csavarok hossza)
- A repülőgépek érkezési késései

R nyelv és használt csomagok

Az R egy nyílt forráskódú programozási nyelv statisztikai számításokhoz és adatelemzéshez.

Miért használtam

- Komplex műveletek egyszerűen megvalósíthatóak
- Könnyen bővíthető csomagokkal
- Korábbi tapasztalat

A ggplot2 csomag a grafikonok készítésére szolgál, ebben az esetben például hisztogramokhoz és sűrűségfüggvényekhez.

Miért használtam

- A mintaátlagok eloszlásának szemléltetésére hisztogramokkal és sűrűségfüggvényekkel.
- A normális eloszlás és az empirikus adatok összehasonlítására.
- Lehetővé teszi az ábrák testreszabását (pl. színek, jelmagyarázat).

A dplyr csomag megkönnyíti az adatok átalakítását és feldolgozását, különösen nagyobb adatállományok esetén.

Miért használtam

- Segít az adatok gyors és egyszerű átalakításában.
- A mintaátlagok kiszámításánál és az adatok előkészítésénél használtam.

Szimuláció R-ben

A szimulációk felépítése

- Véletlen számok generálása különböző eloszlásokból.
- Mintavételezés és átlagok kiszámítása.
- Mintaátlagok ábrázolása hisztogram formájában.
- Az adott eloszlás és a normális eloszlás sűrűségfüggvényeinek ábrázolása összehasonlítóképpen.

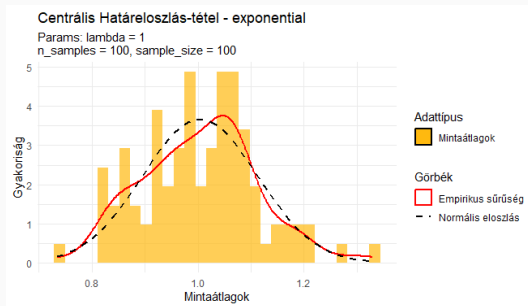
Megjegyzés

A felhasználó konzolos bemenetről, tetszőleges paraméterekkel is indíthat szimulációt az (1) opcióval, a (2) opcióval az exponenciális demo, a (3) opcióval a binomiális demo és a (4) opcióval az egyenletes demo futtatható. A jobb oldalon látható képeket a program generálta.

Szimulációs példa 1 | Exponenciális eloszlás

Példa: Szimuláljuk a CHT-t exponenciális eloszlású adatokkal.

- Generáljunk **100** mintát exponenciális eloszlásból ($\lambda = 1$).
- Minden minta **100** elemű.
- Számítsuk ki minden minta átlagát, eloszlásait ábrázoljuk hisztogramon.
- Ábrázoljuk a sűrűségfüggvényt és a normális eloszlást.

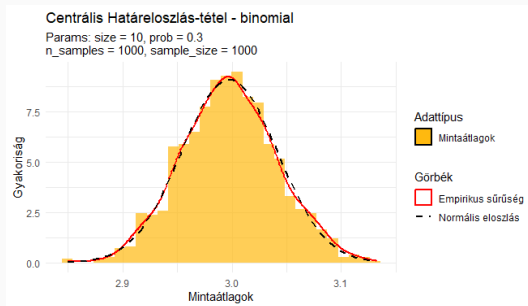


ábra 2: CHT exponenciális eloszlás példa

Szimulációs példa 2 | Binomiális eloszlás

Példa: Szimuláljuk a CHT-t binomiális eloszlású adatokkal.

- Generáljunk **1000** mintát binomiális eloszlásból ($n = 10, p = 0.3$).
- Minden minta **1000** elemű.
- Számítsuk ki minden minta átlagát, eloszlásait ábrázoljuk hisztogramon.
- Ábrázoljuk a sűrűségfüggvényt és a normális eloszlást.

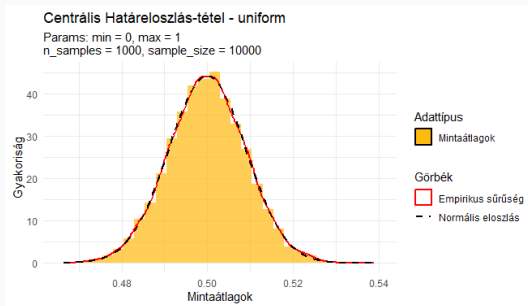


ábra 3: CHT binomiális eloszlás példa

Szimulációs példa 3 | Egyenletes eloszlás

Példa: Szimuláljuk a CHT-t egyenletes eloszlású adatokkal.

- Generáljunk **1000** mintát egyenletes eloszlásból ($[0, 1]$).
- Minden minta **10000** elemű.
- Számítsuk ki minden minta átlagát, eloszlásait ábrázoljuk hisztogramon.
- Ábrázoljuk a sűrűségfüggvényt és a normális eloszlást.



ábra 4: CHT egyenletes eloszlás példa

Szimulációs példa 4 | Interaktív szimuláció

Adjátok meg a paramétereket, és futtassunk azokkal szimulációt!

Megadandó értékek:

- Eloszlás típusa (pl. poisson, khí-négyzet, gamma)
- Eloszlás paraméterei ($param \geq 0$)
- Minták száma ($n_samples \geq 100$)
- Mintánkénti elemszám ($sample_size \geq 100$)

Forráskód

A forráskód elérhető a <https://github.com/CLorant/mat-stat> linken.

Köszönöm a figyelmet!
