

Statisztikai becslések Konfidencia intervallum



Szűcs Mónika, Griechisch Erika, Rárosi Ferenc
SZTE ÁOK-TTIK Orvosi Fizikai és Orvosi Informatikai Intézet

Utoljára frissítve: 2018. augusztus 31.



1. A leckében előforduló fogalmak, jelölések

Megfigyelési, mintavételi egység: A vizsgálat alanya vagy tárgya. A mintavételi egységeken végezzük el a vizsgálatokat, méréseket, megfigyeléseket. Orvosi gyakorlatban a leggyakoribb mintavételi egység egy ember, egy szerv, egy vérminta, más vizsgálatok esetén gyakori példa egy család, egy település, egy erdő, egy állat, egy faj, egy iskolai osztály.

Paraméter: Az az érték, amely leírja a populáció valamely tulajdonságát (nem ismerünk)

Populáció: Az összes lehetséges mintavételi egységet tartalmazó halmaz, amiről szeretnénk információt nyerni a vizsgálat, kísérlet során. A populáció nagyon ritkán egy jól meghatározott véges halmaz, mint például egy ritka trópusi betegségben szenvedő 200 beteg (rajtuk kívül senki más nem fertőzött meg ez a betegség, és már nem is fog). Gyakoribb azonban hogy a populáció nagysága végtelen, nem lehet az összes múlt, jelen és jövőbeli egyedet beazonosítani. *Biostatistikában, és ebben a leckében is populáció alatt nem magukat a mintavételi egységek sokaságát, hanem az azokon mért értékeket értjük.*

Minta: A populáció egy részhalmaza, azon mintavételi egységek (egyedek) melyeken a vizsgálat ténylegesen megtörtént. *Biostatistikában, és ebben a leckében is minta alatt nem magukat a megvizsgált mintavételi egységeket, hanem az azokon mért értékeket értjük.*

A populáció minden elemére történő vizsgálata leggyakrabban nem lehetséges, vagy nem gazdaságos, a mintában szereplő egyedek adataiból próbálunk információt nyerni a populációról. A mintaadatok alapján számolt értéket vagy intervallumot, amely közelíti a populáció megfelelő paraméterét nevezük **becslésnek**.

Jelölések:

- μ a vizsgált populáció ismeretlen átlaga
- σ a vizsgált populáció ismeretlen szórása
- n a minta elemszáma
- x_1, x_2, \dots, x_n a vizsgált minta

2. Pontbecslés

A **pontbecslés** egyetlen számérték, amelyet a populáció paraméterének közelítésére használunk

- A mintaátlag (\bar{x}) egy becslése a populációátlagnak (μ).
$$\mu \approx \bar{x} = \frac{\sum_{i=1}^n x_i}{n}$$

SZÉCHENYI 2020



MAGYARORSZÁG
KORMÁNYA

Európai Unió
Európai Szociális
Alap



BEFEKTETÉS A JÖVŐBE

- A minta szórása (sd) egy becslése a populáció szórásának (σ).
$$\sigma \approx \text{sd} = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}}$$

A becslés pontosságának jellemzésére használjuk a becslés szóródását, amit **standard hibának** hívunk. Leggyakrabban a mintaátlag standard hibáját (**standard error of mean, SEM**) szoktuk használni, melynek jelentése a mintaátlagok szóródása az ismeretlen populációátlag körül.
$$\text{SEM} = \frac{\text{sd}}{\sqrt{n}}$$

3. Intervallumbecslés, konfidencia intervallum

- **Intervallumbecslés:** két olyan szám, amely közé becsljük populáció ismeretlen paraméterét.
- **Konfidencia (megbízhatósági) intervallum:** olyan mintából számolt intervallum, amely $1 - \alpha$ valószínűséggel tartalmazza a populáció valódi (ismeretlen) paraméterét. Az α -t hibaszintnek, $1 - \alpha$ értéket megbízhatósági szintnek, vagy konfidencia szintnek nevezzük. Leggyakrabban 90, 95 vagy 99%-os konfidencia (szintű) intervallumot szoktunk számolni.

3.1. Konfidencia intervalluma populáció átlagára (μ)

Az $(1 - \alpha)$ megbízhatósági szintű konfidencia intervallum a populáció átlagára:

$$\left(\bar{x} - t_{\alpha} \cdot \frac{\text{sd}}{\sqrt{n}} ; \bar{x} + t_{\alpha} \cdot \frac{\text{sd}}{\sqrt{n}} \right)$$

ahol t_{α} a Student t eloszlás α -hoz és $n - 1$ szabadságfokhoz tartozó kritikus értéke.

A leggyakrabban ezt a konfidencia intervallumot használjuk. Ebben az esetben a populáció szórását (σ) nem ismerjük, a mintából számolt szórást (sd) és átlagot (\bar{x}) használjuk az intervallum meghatározásához.

Felhasznált irodalom

- Reiczigel Jenő, Harnos Andrea, Solymosi Norbert: Biostatistika nem statisztikusoknak, Pars Kft. (2014)
- Reiczigel Jenő: Válogatott fejezetek a biostatistikából, SZIE ÁOTK, (2005), <https://docplayer.hu/279567-Reiczigel-jeno-valogatott-fejezetek-a-biostatistikabol.html> (hozzáférés: 2018.08.31.)

Jelen tananyag a Szegedi Tudományegyetemen készült az Európai Unió támogatásával.
Projekt azonosító: EFOP-3.4.3-16-2016-00014