

## Statisztikai alapfogalmak

Az adatsor leggyakoribb értéke a **módusz**. Hogyha például Bob matekjegyei ezek:

2, 3, 1, 4, 1, 2, 2, 3, 5, 2, 3, 2, 3, 2, 4, 3, 2, 4, 2, 4

Akkor egyszerűen meg kell számolni, hogy melyikből van a legtöbb, és az a matekjegy lesz a **módusz**. Most 2-esből van a legtöbb, így Bob matekjegyeinek a módusza 2. A **módusz** jele  $M_o$  és így most  $M_o=2$ .

Léteznek olyan eloszlások is, amelyeknek több módusza van. Hogyha például Bob jegyei:

1, 2, 2, 3, 5, 3, 3, 4, 2

Itt 2-esből és 3-asból ugyanannyi van, mindkettőből 3 darab. Ez egy kétmódusú **eloszlás**.

[Megnézem a kapcsolódó epizódot](#)

A **medián** a növekvő sorba rendezett adatsor középső értéke. Ha az adatsorban páros sok elem van, akkor nincs középső elem, ilyenkor a két középső elem átlagát vesszük.

Hogyha például Bob matekjegyei ezek:

2, 3, 1, 4, 1, 2, 2, 3, 5, 2, 3, 2, 3, 2, 4, 3, 2, 4, 2, 4

Akkor egyszerűen növekvő sorba kell rakni..

1, 1, 2, 2, 2, 2, 2, 2, 2, 2, 3, 3, 3, 3, 4, 4, 4, 4, 5

És aztán meg kell keresni melyik a középső. Most nincsen középső, mert páros sok elem van, így ilyenkor a két középen lévőt átlagoljuk:

1, 1, 2, 2, 2, 2, 2, 2, 2, **2, 3**, 3, 3, 3, 3, 4, 4, 4, 4, 5

Ezeknek az átlaga 2,5 vagyis a **medián** most 2,5. A **medián** jele  $M_e$ , így  $M_e=2,5$

[Megnézem a kapcsolódó epizódot](#)

Az átlagot úgy kapjuk meg, hogy az összes elemet összeadjuk, és aztán elosztjuk az elemek számával.

Jele:  $\bar{x}$

[Megnézem a kapcsolódó epizódot](#)

Az átlagtól való átlagos eltérés egyik legjobb mérőszáma a szórás. Hátránya, hogy egy kicsit ronda a szórás képlete. A szórást egy szigma nevű görög betűvel jelöljük.

$$\sigma = \sqrt{\frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2}{n}}$$

[Megnézem a kapcsolódó epizódot](#)

Az adatsor első felének a felezőpontja az alsó kvartilis.

Az alsó kvartilis jele:  $Q_1$

[Megnézem a kapcsolódó epizódot](#)

---

Az adatsor második felének a felezőpontja a felső kvartilis.

A felső kvartilis jele:  $Q_3$

[Megnézem a kapcsolódó epizódot](#)

---

A kvartilisek és a [medián](#) azt szemlélteti, hogyan oszlanak el az adatsorban szereplő adatok. Ezek segítségével készíthető el a doboz-ábra, vagy másnéven dobozdiagram. Szokás még sodrófa diagramnak is nevezni, és az angol elnevezést is gyakran használják, ami a box plot.

Egy sobarendezett adatsorban öt darab speciális negyedelőpontot fogunk használni. Az első az adatsor legkisebb értéke, ez a  $Q_0$ . Aztán a következő negyedelő az alsó kvartilis, ami  $Q_1$  utána jön a felezőpont vagyis a [medián](#), ezt  $Me$ -vel és  $Q_2$ -vel is jelöljük, végül a felső kvartilis, ami a  $Q_3$ . Az adatsor legnagyobb értéke pedig  $Q_4$ . A legnagyobb és a legkisebb érték különbsége a terjedelem, míg a két kvartilis különbségét félterjedelemnek vagy más néven interkvartilisnek hívjuk. Ezekből épül föl a doboz-ábra vagy másként dobozdiagram.

Előfordulhat, hogy az adatsorban kiugró értékek is szerepelnek. A kiugró érték az, ami az alsó kvartilisnél legalább a félterjedelem másfélszeresénél kisebb, vagy pedig a felső kvartilisnél legalább a félterjedelem másfélszeresénél nagyobb. Huh, ez elég bonyolult hangzik. De valójában nagyon egyszerű, csak nézd meg kapcsolódó epizódot és kiderül.

[Megnézem a kapcsolódó epizódot](#)

---

A relatív szórás azt mondja meg, hogy a szórás az átlagnak hány százaléka:

$$V = \frac{\sigma}{\bar{X}}$$

[Megnézem a kapcsolódó epizódot](#)

---