

Adatfeldolgozási alapismeretek

1. rész

Bevezetés, elemi
formalizmus

Jelölések, konvenciók

- „Izé” = statisztika
- Fél-formálisan:
az adatok alapján számolt bármely
függvényt *statisztikának* szokás nevezni

Jelölések, formalizmus

Adatokat sorszámukkal azonosítjuk:

$$x_1, x_2, \dots$$

Azaz az *első* adatot x_1 , a *második* adatot x_2 jelöli, és így tovább. Az adatok számát n -el jelöljük.

Jelölések: példa

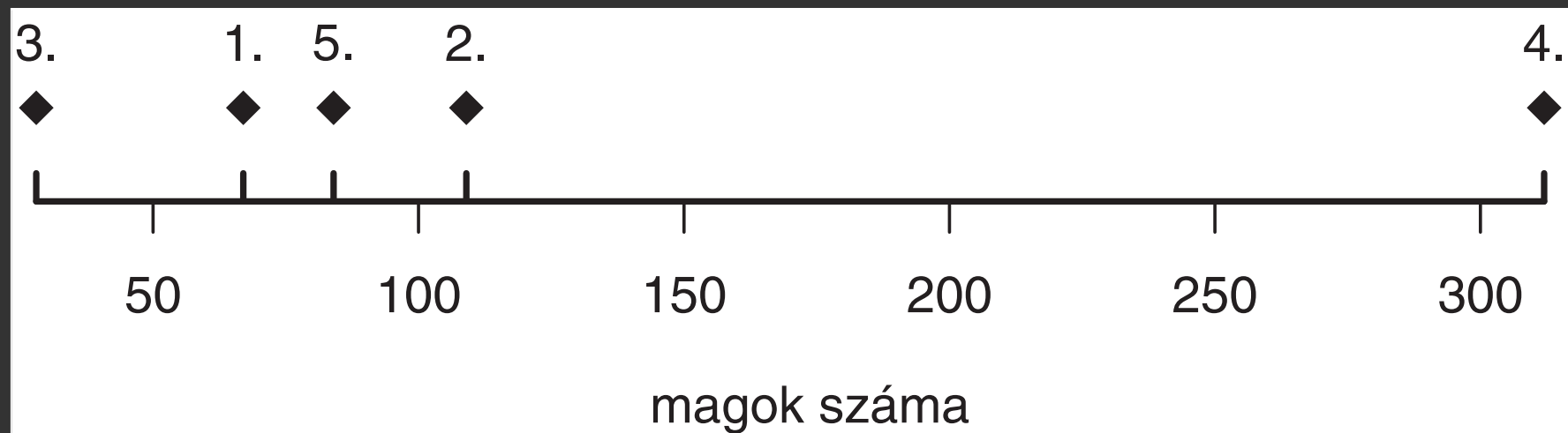
A vizsgálat során 5 mérést végzünk:

$$x_1 = 67, \quad x_2 = 109, \quad x_3 = 28,$$

$$x_4 = 312, \quad x_5 = 84$$

Példa

Cirsium arvense magok száma a fészek-
virágzatban:



Jelölések, formalizmus

Gyakran van szükség arra, hogy az adatokat *sorba rendezzük* nagyság szerint. Az ilyen sorba rendezett adatokból számolt statisztikákat ***rendstatisztikának*** nevezzük.

A sorba rendezett adatok jelölése többféleképpen is történhet.

Jelölések: adatvektor

Egy kísérletből származó adatokat gyakran adatvektor formájában állítjuk össze:

$$\mathbf{X} = (x_1, x_2, \dots, x_n)$$

Sorba rendezett adatok

- Csökkenő:

$$\mathbf{x}^{\downarrow} = (x_{[1]}, x_{[2]}, \dots, x_{[n]})$$

Sorba rendezett adatok

- Növekvő:

$$\mathbf{x}^{\uparrow} = (x_{(1)}, x_{(2)}, \dots, x_{(n)})$$

Jelölések: példa

$$x_{(1)} = 28, \quad x_{(2)} = 67, \quad x_{(3)} = 84,$$

$$x_{(4)} = 109, \quad x_{(5)} = 312$$

Sorba rendezett adatok

(Tipográfiaiailag) bonyolultabb jelölés:

$$x_{1:n}, x_{2:n}, \dots, x_{(n-1):n}, x_{n:n}$$