

# Понятия и формули:

- извадка: наблюденията, които сме събрали, означени с  $X_1, X_2, \dots, X_n$
- извадъчно средно:  $\bar{X} = \sum_{i=1}^{i=n} \frac{X_i}{n}$
- извадъчна дисперсия:  $s^2 = \sum_{i=1}^{i=n} \frac{(X_i - \bar{X})^2}{n-1} = \frac{\sum_{i=1}^{i=n} [X_i^2] - n\bar{X}^2}{n-1}$
- извадъчно стандартно отклонение:  $s = \sqrt{s^2} = \sqrt{\sum_{i=1}^{i=n} \frac{(X_i - \bar{X})^2}{n-1}}$
- вариационен ред: стойностите от извадката, подредени по големина:  $X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}$
- размах (обхват): разликата между най-голямата и най-малката стойност в извадката:  $X_{(n)} - X_{(1)}$
- медиана: стойност, за която 50% от данните са по-малки от нея и 50% от данните са по-големи от нея
- $p$ -ти процентил ( $p \in (0, 100)$ ): стойност, за която  $p\%$  от данните са по-малки от нея и  $(100 - p)\%$  от данните са по-големи от нея
- долен и горен квартил ( $Q_L, Q_U$ ): стойност, за която 25% и 75% съответно от данните са по-малки от нея
- интерквартилен размах: разликата между горния и долния квартил:  $Q_U - Q_L$
- lower and upper hinge: използват се като приближение на долния и горния квартил
- силно отличаващо се наблюдение (outlier): стойност, която силно се отличава от извадката - необикновено голямо или необикновено малко наблюдение в сравнение с останалите
- кутия с мустаци: графичен начин за представяне на данните, използвайки медианата, lower and upper hinge, интерквартилния размах
- z-score на наблюдение  $X_i$ :  $\text{z-score}(X_i) = \frac{X_i - \bar{X}}{s}$

**Зад.1** Намерете извадъчното средно на следните данни: 2, 9, 11, 5, 6

Отговор: 6.6

**Зад.2** Намерете медианата на следните данни: 9, 2, 7, 11, 14

Отговор: 9

**Зад.3** Намерете медианата на следните данни: 9, 2, 7, 11, 14, 6

Отговор: 8

**Зад.4** Дадени са  $n = 5$  наблюдения: 0, 5, 1, 1, 3. Намерете:

а)  $\bar{x}$ ; б)  $m$ ; в)  $s$ .

Отговор: а) 2; б) 1; в) 2

**Зад.5** Дадени са  $n = 8$  наблюдения: 3, 1, 5, 6, 4, 4, 3, 5. Намерете:

а) обхвата; б)  $\bar{x}$ ; в)  $s^2$ .

Отговор: а) 5; б) 3.875; в) 2.4107

**Зад.6** Дадени са  $n = 10$  наблюдения: 3, 5, 4, 6, 10, 5, 6, 9, 2, 8. Намерете: а)  $\bar{x}$ ; б)  $m$ ; в)  $s$ .

Отговор: а) 5.8; б) 5.5; в) 2.573

**Зад.7** Разглеждаме извадка от  $n = 10$  наблюдения: 1, 1, 0, 15, 2, 3, 4, 0, 1, 3. Вижда се, че наблюдението  $x = 15$  е необичайно голямо. Пресметнете z-score за това наблюдение и направете заключение.

Отговор: 2.71

**Зад.8** Намерете медианата, долния и горния квартил на следните данни:

8, 7, 1, 4, 6, 6, 4, 5, 7, 6, 3, 0

а) пресметнете  $\bar{x}$  и стандартното отклонение  $s$

б) пресметнете z-score за най-малкото и най-голямото наблюдение. Можем ли да заключим, че са необичайно малко и необичайно голямо?

Отговор:  $m = 5.5$ ,  $Q_L = 3.25$ ,  $Q_U = 6.75$  а)  $\bar{x} = 4.75$ ,  $s = 2.454$ ; б)  $-1.94$ ,  $1.32$ , не

**Зад.9** Намерете медианата, долния и горния квартил на следните данни:

19, 12, 16, 0, 14, 9, 6, 1, 12, 13, 10, 19, 7, 5, 8

Отговор:  $m = 10$ ,  $Q_L = 6$ ,  $Q_U = 14$

**Зад.10** Начертайте кутия с мустаци за данните от **Зад.7**. Има ли outlier?

Отговор: да

**Зад.11** Начертайте кутия с мустаци за следните данни и определете има ли outlier:

25, 22, 26, 23, 27, 26, 28, 18, 25, 24, 12

Отговор: да