Interquartile Range (IQR)

By
Dr Ravi Prakash Verma
Professor
Department of CSAI
ABESIT

Interquartile Range (IQR)

- The **Interquartile Range (IQR)** is a measure of statistical dispersion that describes the range within which the middle 50% of the data lies.
- It is the difference between the **third quartile (Q3)** and the **first quartile (Q1)**: IQR=Q3−Q1

Steps to Calculate the IQR:

1. **Order the data**: Arrange the dataset in ascending order.
2. **Find Q1 (First Quartile)**: The median of the lower half of the data (excluding the overall median if the number of values is odd).
3. **Find Q3 (Third Quartile)**: The median of the upper half of the data.
4. **Compute IQR**: Subtract **Q1** from **Q3**.

- **Example Calculation:**

Interquartile Range (IQR)

- **Example Calculation:**

Consider the dataset:
5, 7, 8, 12, 13, 15, 18, 21, 22, 25

1. **Q1** (First Quartile) → Median of **5, 7, 8, 12, 13** = **8**
2. **Q3** (Third Quartile) → Median of **15, 18, 21, 22, 25** = **21**
3. **IQR** = **Q3 - Q1** = **21 - 8** = **13**

Use of IQR:

- Used to detect **outliers**: Data points below **Q1 - 1.5 × IQR** or above **Q3 + 1.5 × IQR** are considered outliers.
- Helps understand the **spread** of the middle 50% of data, reducing the effect of extreme values.

Interquartile Range (IQR)

- **Example Calculation:**
- In the calculation of the **First Quartile (Q1)** and **Third Quartile (Q3)**, the dataset is divided into two halves **after finding the median**.
- The reason why only **5 elements** are chosen for Q1 in your example is because the total dataset consists of **10 values**, and when split into two equal halves, each half contains **5 elements**.

Interquartile Range (IQR)

- **Example Calculation:**

Understanding the Division of Data:

1. Arrange the dataset in ascending order:

$$5, 7, 8, 12, 13, 15, 18, 21, 22, 25$$

2. Find the **Median (Q2)**:
   - Since there are **10 values (even count)**, the median is the **average of the 5th and 6th values**:

$$Q2 = \frac{13 + 15}{2} = 14$$

3. **Split the dataset into two halves:**
   - **Lower half (left of median): 5, 7, 8, 12, 13 → Used to calculate Q1**
   - **Upper half (right of median): 15, 18, 21, 22, 25 → Used to calculate Q3**
4. **Find Q1** (Median of lower half):
   - The median of **5, 7, 8, 12, 13** is **8**
5. **Find Q3** (Median of upper half):
   - The median of **15, 18, 21, 22, 25** is **21**

Thus, **Q1 = 8** and **Q3 = 21**, and we compute **IQR = Q3 - Q1 = 21 - 8 = 13**.

Interquartile Range (IQR)

- **Example Calculation:**

Why Are Only 5 Elements Used in Q1?

- When the total **number of data points is even**, we **split the data equally** into two halves.
- If the total **number of data points is odd**, we **exclude the median** when splitting the dataset.

Interquartile Range (IQR)

- **Example Calculation:**

Let's consider an **odd-numbered dataset** and compute the **Interquartile Range (IQR)** step by step.

Example Dataset (Odd Count):

- **5, 7, 8, 12, 13, 15, 18, 21, 22** (**9 elements**)

Step 1: Find the Median (Q2)

- The **median** is the middle value when data is arranged in ascending order.
- Since there are **9 elements**, the **5th value** is the median: Q2=13Q2 = 13Q2=13

Interquartile Range (IQR)

- **Example Calculation:**

Step 2: Split the Data into Two Halves

- **Lower half (before the median): 5, 7, 8, 12**
- **Upper half (after the median): 15, 18, 21, 22**

Step 3: Find Q1 and Q3

- Q1 (First Quartile) → Median of the lower half:
  - Median of 5, 7, 8, 12 →

$$Q1 = \frac{7+8}{2} = 7.5$$

- Q3 (Third Quartile) → Median of the upper half:
  - Median of 15, 18, 21, 22 →

$$Q3 = \frac{18+21}{2} = 19.5$$

Interquartile Range (IQR)

- **Example Calculation:**

Step 4: Compute IQR

$$IQR = Q3 - Q1 = 19.5 - 7.5 = 12$$

Interquartile Range (IQR)

- **Example Calculation:**
- **Key Difference from an Even Dataset:**
  - When the total number of elements is even, we split the dataset into equal halves after computing the median.
  - When the total number of elements is odd, the median itself is excluded from both halves before computing Q1 and Q3.