

In R i boxplot vengono costruiti nel modo seguente:

- Si calcolano il 25-esimo ed il 75-esimo percentile ( $Q_1$  e  $Q_3$ )
- Si costruisce la distanza interquartile:  $IQR = Q_3 - Q_1$  questa definisce la lunghezza del lato della scatola
- Gli estremi della scatola vengono disegnati in corrispondenza dei valori osservati più vicini a  $Q_1, Q_3$ .
- Tutti i valori che superano di  $1.5(Q_3 - Q_1)$  uno dei due estremi della scatola vengono considerati anomali, quindi i “baffi” del boxplot vengono tracciati a partire dagli estremi della scatola fino al valore osservato più piccolo minore di  $Q_1 - 1.5(Q_3 - Q_1)$ , per l'estremo inferiore, e fino al valore osservato più grande minore di  $Q_3 + 1.5(Q_3 - Q_1)$ .

## Iris setosa

Valori	Frequenze
1	1
1.1	1
1.2	2
1.3	7
1.4	13
1.5	13
1.6	7
1.7	4
1.9	2
Totale	50

$$Q_1 = 1.4 \quad Q_3 = 1.575 \approx 1.6 \quad \text{mediana} = 1.5, \quad \text{IQR} = 1.6 - 1.4 = 0.2$$

# Iris setosa

iris data: Petal Length setosa

