position measures

mode: The most common value in a sample

arithmetic mean: $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$

median: $x_{med} = x_{0.5}$ (or the central value of the ordered sample)

p-quantile: $x_p = \begin{cases} x_{(\lfloor np \rfloor + 1)} & \text{if } np \notin \mathbb{N} \\ \frac{1}{2}(x_{(np)} + x_{(np+1)}) & \text{if } np \in \mathbb{N} \end{cases}$

Comments:

$x_{(i)}$: the $i$te Element in the *ordered* sample (i.e. $x_{(1)} \leq x_{(2)} \leq \ldots \leq x_{(n)}$)

$\lfloor np \rfloor$ is the largest integer less than or equal to np

Measures of dispersion

span: $r = x_{(n)} - x_{(1)}$

quartile difference: $qd = x_{0.75} - x_{0.25}$

sample variance: $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$

sample standard deviation: $s = \sqrt{s^2}$

Graphic representation

Histogram: On the horizontal axis, the class boundaries are $c_0, \ldots, c_k$ worn. above each interval $(c_{j-1}, c_j]$ a rectangle is drawn with width $d_j = c_j - c_{j-1}$ (Class width) and height $g_j = \frac{f_j}{c_j - c_{j-1}}$. ($f_j$: relative frequency).

The **simple box plot** consists of a box with the horizontal boundary lines $x_{0.75}$ and $x_{0.25}$ and the connecting lines (Whiskern) from the quartiles to the corresponding extreme values. The median is marked by a horizontal line in the box.

In the case of a **refined box plot** the connecting lines are (Whisker) drawn only to the outermost value, which is not greater than $z_o = x_{0.75} + \frac{3}{2}(x_{0.75} - x_{0.25})$ or is not smaller than $z_u = x_{0.25} - \frac{3}{2}(x_{0.75} - x_{0.25})$. Each observation value $x_i \notin [z_u, z_o]$ (Outlier) will be marked with a „$*$“ or „$\circ$“.