i	Klasse $(x_i^u; x_i^o)$	Häufigkeit h_i	Mitte $\frac{x_i^u + x_i^o}{2}$	Breite $x_i^o - x_i^u$	Dichte n_i^*	rel. Häufig. f_i $\frac{n_i}{\text{ges. Zahl}}$	kumuliert F_i $\sum_{n=1}^{i} f_i$	Werte kum. $q_{ges.}$ $\sum$ Mittelwert $\cdot$ Häufigk.	Anteil q_i $h_i \cdot \frac{(x_k^o + x_k^u)}{2} \cdot \frac{1}{q_{ges.}}$	kumuliert Q_i

p-Quantil (klassierte Daten): p-Quantil in Klasse $k \Rightarrow$

$$\bar{x}_p = x_k^u + (x_k^o - x_k^u)\frac{p - F(x_k^u)}{F(x_k^o) - F(x_k^u)}$$

Spannweite:

$$s_{max} = \max\{x_i\} - \min\{x_i\}$$

Empirische Stichprobenvarianz, Standardabweichung, Variationskoeffizient:

$$s^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2, \qquad s = \left|\sqrt{s^2}\right|, \qquad v_x = \frac{s}{\bar{x}}$$

Mittlere absolute Abweichung vom (beliebig wählbaren) Lageparameter $\bar{x}$:

$$d = \frac{1}{n}\sum_{i=1}^{n}|x_i - \bar{x}|$$

Gini-Koeffizient:

$$G = 1 - \sum_{i=1}^{n} f_i(Q_i + Q_{i-1})$$

Rang · Wert z.B. Einkommen

$$G = \frac{2 \cdot (\sum i \cdot x_{[i]}) - (n+1) \cdot \sum x_{[i]}}{n \cdot \sum x_{[i]}}$$

Summe Werte

normiert:

$$G^* = \frac{n}{n-1} \cdot G$$

Kovarianz:

$$s_{xy} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})$$

$$\Rightarrow s_{xx}(= s^2) = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})(x_i - \bar{x}) = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

Lineare Regression, Pearson-Korrelationskoeffizient: $y = a{\cdot}x{+}b$ (linearer Zusammenhang)

x	y	$x_i - \bar{x}$	$y_i - \bar{y}$	$(x_i - \bar{x})^2$	$(y_i - \bar{y})^2$	$(x_i - \bar{x})(y_i - \bar{y})$

$$a = \frac{s_{xy}}{s_{xx}}, \qquad b = \bar{y} - a\bar{x} \qquad r_{xy} = \frac{s_{xy}}{\sqrt{s_{xx}} \cdot \sqrt{s_{yy}}}$$

Bestimmtheitsmaß:

$$R^2 := \frac{\sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} = 1 - \frac{\sum_{i=1}^{n}(e_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}$$

mit der unerklärten Variation der Residuen e_i

$$\sum_{i=1}^{n}(e_i)^2$$

.