

Streuemaße - Grundwissen

Gegeben sei eine univariate statistische Erhebung mit

- einer Grundgesamtheit mit dem Erhebungsumfang n ,
- einem quantitativen Merkmal X mit m verschiedenen Merkmalsausprägungen $a_1; \dots; a_m$,
- der durch die Erhebung gewonnenen Urliste mit den Messwerten x_1, \dots, x_n , die sich in einer eindeutigen Reihenfolge, z.B. $x_1 \leq \dots \leq x_n$ nach steigender Größe anordnen lassen
- den absoluten Häufigkeiten $H(a_1), \dots, H(a_m)$ sowie
- den relativen Häufigkeiten $h(a_1), \dots, h(a_m)$ der einzelnen Merkmalsausprägungen
- den bekannten Mittelwerten x_{mod} , \tilde{x} und \bar{x} .



Als **Spannweite r** (der Urliste) bezeichnet man die Differenz von größtem (x_n , s.o.) und kleinstem (x_1 , s.o.) auftretenden Messwert in der Urliste, d.h.

$$r = x_n - x_1$$



Als **Mittlere Absolute Abweichung A_{abs}** oder \bar{d} (der Messwerte vom Mittelwert, meist dem Median \tilde{x} oder dem Arithmetischen Mittel \bar{x}) bezeichnet man die Zahl

$$A_{\text{abs}} = \frac{|x_1 - \tilde{x}| + \dots + |x_n - \tilde{x}|}{n} = \frac{|a_1 - \tilde{x}| \cdot H(a_1) + \dots + |a_m - \tilde{x}| \cdot H(a_m)}{n}$$

Berechnung mit absoluten Häufigkeiten

$$= \frac{|a_1 - \tilde{x}| \cdot h(a_1) + \dots + |a_m - \tilde{x}| \cdot h(a_m)}{1}$$

Berechnung mit relativen Häufigkeiten



Als **Mittlere Quadratische Abweichung s^2** oder **Varianz V** oder V_X (der Messwerte vom Mittelwert, meist dem Arithmetischen Mittel \bar{x}) bezeichnet man die Zahl

$$V = \frac{(x_1 - \bar{x})^2 + \dots + (x_n - \bar{x})^2}{n} = \frac{(a_1 - \bar{x})^2 H(a_1) + \dots + (a_m - \bar{x})^2 H(a_m)}{n}$$

Berechnung mit absoluten Häufigkeiten

$$= \frac{(a_1 - \bar{x})^2 h(a_1) + \dots + (a_m - \bar{x})^2 h(a_m)}{1}$$

Berechnung mit relativen Häufigkeiten

Als **Standardabweichung s** oder s_X (der Messwerte vom Mittelwert, meist dem Arithmetischen Mittel \bar{x}) bezeichnet man die Wurzel aus der mittleren quadratischen Abweichung, d.h.

$$s = \sqrt{V} \quad \text{oder} \quad s_X = \sqrt{V_X}$$

Weitere Erklärungen und Beispiele zu diesen Begriffen finden sich im Selbstlernprogramm ‚Beschreibende Statistik und explorative Datenanalyse‘ in Abschnitt 3.4.

Einen Vergleich der verschiedenen Kenngrößen und Übungsaufgaben finden sich im Selbstlernprogramm ‚Beschreibende Statistik und explorative Datenanalyse‘ in den Abschnitten 3.5 und 3.7.