Warum die Standardabweichung im Allgemeinen größer ist als die mittlere absolute Abweichung

Sei x_i eine Merkmalsausprägung von $i = 1 \ldots n$ Merkmalsausprägungen und $\bar{x}$ die durchschnittliche Merkmalsausprägung, dann ist die mittlere absolute Abweichung

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

Unter der Voraussetzung, dass alle Merkmalsausprägungen der Grundgesamtheit erhoben werden, ist die Standardabweichung

$$\sqrt{\frac{1}{n} \cdot \sum_{i=1}^{n} \left(x_i - \bar{x} \right)^2}$$

Mit konkreten Daten wird man finden, dass die Standardabweichung größer ist als die mittlere absolute Abweichung, sie auch gleich groß sein kann, aber niemals kleiner. Es scheint folgende Ungleichung zu gelten:

(1) $$\frac{1}{n} \cdot \sum_{i=1}^{n} \left| x_i - \bar{x} \right| \le \sqrt{\frac{1}{n} \cdot \left(x_i - \bar{x} \right)^2}$$

Dass dies tatsächlich der Fall ist, sei im Folgenden bewiesen.

Zur Vereinfachung der Beweisführung wird die Abweichung $\left| x_i - \bar{x} \right|$ als a_i bezeichnet:

(2) $$\left| x_i - \bar{x} \right| = a_i$$

Die zu beweisende Ungleichung lautet dann:

(3) $$\frac{1}{n} \cdot \sum_{i=1}^{n} a_i \le \sqrt{\frac{1}{n} \cdot \sum_{i=1}^{n} a_i^2}$$

Die Ungleichung wird quadriert und vereinfacht:

$$\frac{1}{n^2} \cdot \left(\sum_{i=1}^{n} a_i \right)^2 \le \frac{1}{n} \cdot \sum_{i=1}^{n} a_i^2$$

$$\frac{1}{n} \cdot \left(\sum_{i=1}^{n} a_i \right)^2 \le \sum_{i=1}^{n} a_i^2$$

(4) $$\left(\sum_{i=1}^{n} a_i \right)^2 \le n \cdot \sum_{i=1}^{n} a_i^2$$

Anhand von Ungleichung (4) kann zunächst gezeigt werden, wann aus der Ungleichung eine Gleichung wird. Es liegt der Gedanke nahe, dass dies der Fall ist, wenn alle Abweichungen gleich groß sind. Wird diese Abweichung einfach mit a bezeichnet, dann bedeutet in Ungleichung (4) der Ausdruck $\sum_{i=1}^{n} a$, dass eine Summe aus n Summanden zu bilden ist, die aber mit a alle gleich groß sind. Die Summe ist also einfach das n-Fache des einzelnen Summanden:

$$\sum_{i=1}^{n} a = n \cdot a$$

Die gleiche Überlegung auf $\sum_{i=1}^{n} a^2$ angewendet:

Warum die Standardabweichung im Allgemeinen größer ist als die mittlere absolute Abweichung

$$\sum_{i=1}^{n} a^2 = n \cdot a^2$$

Setzt man diese beiden Ausdrücke in (4) ein, gilt in der Tat das Gleichheitszeichen:

$$(n \cdot a)^2 \le n \cdot n \cdot a^2$$

$$a^2 \cdot n^2 \le a^2 \cdot n^2$$

$$a^2 = a^2$$

Im Folgenden wird nun die Voraussetzung gleicher Abweichungen wieder aufgehoben und es wird mithilfe des Verfahrens der vollständigen Induktion bewiesen, dass Ungleichung (4) für alle Werte von n gilt, wobei n eine natürliche Zahl ist. Hierzu wird $n = 2$ in Ungleichung (4) eingesetzt. Es ergibt sich folgender Induktionsanfang:

$$\left(\sum_{i=1}^{2} a_i\right)^2 \le 2 \cdot \sum_{i=1}^{2} a_i^2$$

$$(a_1 + a_2)^2 \le 2 \cdot (a_1^2 + a_2^2)$$

$$a_1^2 + 2a_1a_2 + a_2^2 \le 2a_1^2 + 2a_2^2$$

$$0 \le a_1^2 - 2a_1a_2 + a_2^2$$

$$0 < (a_1 - a_2)^2$$

Für $a_1 \ne a_2$ kann die Differenz von a_1 und a_2 positiv oder negativ sein, das Quadrat der Differenz ist aber in jedem Fall positiv. Damit ist Ungleichung (4) für den Fall $n = 2$ und $a_1 \ne a_2$ bewiesen.

Es wird nun angenommen, Ungleichung (4) gelte für alle Wert von n [Induktionsannahme]. Das ist dann der Fall, wenn die Ungleichung auch für $n + 1$ gilt [Induktionsbehauptung].

Induktionsschluss: Wenn die Anzahl der Abweichungen a_i von n auf $n + 1$ erhöht wird, kommt zu den bisherigen Abweichungen die Größe a_{n+1} hinzu. Nach dem Bildungsgesetz der linken Seite von Ungleichung (4) muss diese Abweichung quadriert werden, sodass die linke Seite von (4) nunmehr lautet:

$$\left(\sum_{i=1}^{n} a_i + a_{n+1}\right)^2$$

Auf der rechten Seite von Ungleichung (4) sind in der Klammer die Quadrate der Abweichungen zu addieren, es kommt also a_{n+1}^2 dazu. Die Summe der quadrierten Abweichungen ist mit der Anzahl der Merkmalsausprägungen zu multiplizieren, die jetzt $n + 1$ beträgt, sodass

$$(n+1) \cdot \left(\sum_{i=1}^{n} a_i^2 + a_{n+1}^2\right)$$

Für $n + 1$ ergibt sich also aus (4):

(5) $$\left(\sum_{i=1}^{n} a_i + a_{n+1}\right)^2 \le (n+1) \cdot \left(\sum_{i=1}^{n} a_i^2 + a_{n+1}^2\right)$$

In der Klammer auf der linken Seite kommt zur Summe aller a_i von a_1 bis a_n noch der nächste Summand a_{n+1} hinzu. Dies lässt sich in Summationsvorschrift aufnehmen, indem einfach bis $n + 1$ summiert wird. Es gilt:

$$\sum_{i=1}^{n} a_i + a_{n+1} = \sum_{i=1}^{n+1} a_i$$

Auf der rechten Seite von Ungleichung (5) lässt sich ebenfalls die Summationsvorschrift auf den neu hinzukommenden Summanden anwenden:

$$\sum_{i=1}^{n} a_i^2 + a_{n+1}^2 = \sum_{i=1}^{n+1} a_i^2$$

Beide Ausdrücke in Ungleichung (5) eingesetzt:

(6) $$\left(\sum_{i=1}^{n+1} a_i\right)^2 \leq (n+1) \cdot \sum_{i=1}^{n+1} a_i^2$$

Das aber ist Ungleichung (4), angewendet auf $n+1$. Ungleichung (4) gilt also auch für $n+1$ und damit für jeden Wert von n, da bei der Herleitung von (6) die Voraussetzung, n sei gleich 2, nicht verwendet wurde, sondern nur die Voraussetzung, $n+1$ sei genau um 1 größer als n, was evident ist.

Ungleichung (4) gilt somit für alle Werte von $n \geq 2$. Das Gleichheitszeichen gilt für $n = 1$, natürlich für den Trivialfall $n = 0$, und, wie oben gezeigt, wenn alle Werte von a gleich sind. Für $n \geq 2$ und $a_i \neq a$ ist in (4) nur das Ungleichheitszeichen gültig.

Ungleichung (4) ist in der Mathematik als „Ungleichung für das arithmetische und das quadratische Mittel" bekannt.

Um den statistischen Zusammenhang zwischen der Standardabweichung und der mittleren absoluten Abweichung wieder herzustellen, wird in die nunmehr bewiesene Ungleichung (4) Gleichung (2) eingesetzt, sodass:

$$\left(\sum_{i=1}^{n} \left|x_i - \bar{x}\right|\right)^2 \leq n \cdot \sum_{i=1}^{n} \left(x_i - \bar{x}\right)^2$$

Hieraus folgt:

$$\frac{1}{n^2} \cdot \left(\sum_{i=1}^{n} \left|x_i - \bar{x}\right|\right)^2 \leq \frac{1}{n} \cdot \sum_{i=1}^{n} \left(x_i - \bar{x}\right)^2$$

(7) $$\frac{1}{n} \cdot \sum_{i=1}^{n} \left|x_i - \bar{x}\right| \leq \sqrt{\frac{1}{n} \cdot \left(x_i - \bar{x}\right)^2}$$

Das aber ist Ungleichung (1), womit diese bewiesen ist.