

Einfache Datensätze für unterschiedliche statistische Kenngrößen¹

LAWRENCE M. LESSER, USA

¹ Das Original erschien in Teaching Statistics (Volume 33, Number 1, Spring 2011; S. 9–11).
Originaltitel: Simple Data Sets for Distinct Basic Summary Statistics
Übersetzung, Bearbeitung und Kürzung: J. MEYER

Zusammenfassung: Der Artikel stellt einfache Datensätze bereit, für die die wesentlichen statistischen Kenngrößen verschieden sind.

1 Einleitung

Eine unglückliche Wahl von Eingangsparametern kann dazu führen, dass Lernende Fehlkonzepte unerkannt verstärken oder es jedenfalls für Lehrende schwierig machen zu erkennen, ob die Lernenden die richtige Antwort auch auf dem richtigen Weg bekommen haben. Wenn zum Beispiel eine Schülerin oder ein Schüler auf die Frage, was der Umfang eines Rechtecks mit den Seitenlängen 3 und 6 sei, die richtige Antwort „18“ gibt, kann es sein, dass einfach 3 mal 6 gerechnet wurde. Dies Beispiel ist nicht unrealistisch, da es eine Reihe von Lernenden gibt, die Flächeninhalte schlecht von Umfängen unterscheiden können.

Auch in der beschreibenden Statistik ist es sinnvoll, mit Datensätzen zu arbeiten, die für die unterschiedlichen Kenngrößen auch unterschiedliche numerische Werte liefern. Der Datensatz $\{3, 4, 5, 5, 8\}$ ist hier keine gute Wahl, da arithmetisches Mittel, Median, Modalwert, Spannweite und Anzahl der Daten alle den gleichen Wert „5“ liefern. Außerdem könnte die Schülerin oder der Schüler den Eindruck gewinnen, dass das arithmetische Mittel immer im Datensatz enthalten sein müsse.

2 Kriterien

Ein Datensatz soll einen eindeutigen Modalwert haben, und die Häufigkeit, mit der der Modalwert auftritt, soll vom Modalwert verschieden sein (das ist zum Beispiel bei $\{1, 2, 2, 3, 4, 5\}$ nicht der Fall).

Ferner sollen Minimum, Maximum, Spannweite, arithmetisches Mittel, Median, Modalwert und Anzahl der Daten alle verschieden sein.

Schließlich soll das arithmetische Mittel nicht im Datensatz enthalten sein.

Falls die Anzahl der Daten gerade ist, soll der Median das arithmetische Mittel der beiden „inneren Daten“ sein.

3 Datensätze

Ein Datensatz muss die Mindestlänge 6 haben (sonst können die im letzten Abschnitt benannten Kriterien nicht alle erfüllt sein).

Ein Beispiel eines 6-elementigen Datensatzes ist $\{1, 2, 3, 4, 4, 16\}$ mit den Kenngrößen

Anzahl der Daten	6
Minimum	1
Maximum	16
Spannweite	15
Median	3,5
arithmetisches Mittel	5
Modalwert	4
Häufigkeit des Modalwerts	2.

Ein Beispiel eines 7-elementigen Datensatzes ist $\{1, 2, 3, 4, 6, 6, 13\}$ mit den Kenngrößen

Anzahl der Daten	7
Minimum	1
Maximum	13
Spannweite	12
Median	4
arithmetisches Mittel	5
Modalwert	6
Häufigkeit des Modalwerts	2.

Andere Datensätze sind

$\{1, 3, 3, 4, 6, 13\}$, $\{1, 3, 3, 4, 7, 12\}$,
 $\{1, 3, 3, 4, 8, 11\}$, $\{2, 3, 3, 4, 7, 11\}$,
 $\{1, 3, 3, 4, 6, 8, 10\}$.

4 Erweiterung

Man kann das Verständnis der Lernenden auch dadurch überprüfen, dass man sie selber Datensätze finden lässt, die bestimmte vorgegebene Kennzahlen haben. Man kann auch Lernende erläutern lassen, weshalb es für bestimmte Kennzahlen keinen Datensatz geben kann (etwa Median 3, arithmetisches Mittel 4 und eindeutiger Modalwert 5).

Selbstverständlich lassen sich die vorzugebenden Kennzahlen erweitern durch Quartile oder die Standardabweichung.

Anschrift des Verfassers

Lawrence M. Lesser
The University of Texas at El Paso, USA
lesser@utep.edu