$S \neq \sigma$ – oder: Standardabweichung ist nicht gleich Standardabweichung

Dr. Guido Pinkernell

Einige Notizen zum Gebrauch der TI-83 und TI-89 bei der Bestimmung von Verteilungskennwerten.

Zwei Werte für die Standardabweichung

GTR sowie CAS erlauben, auf Knopfdruck die wesentlichen Kennwerte einer Verteilung festzustellen. Am Beispiel der folgenden Liste, die die Körpergrößen von zehn zufällig ausgewählten Sechsjährigen im cm darstellen soll

$$L_1 = \{\,123{,}8;\ 115{,}2;\ 120{,}9;\ 112{,}4;\ 115{,}8;\ 115{,}4;\ 118{,}9;\ 111{,}9;\ 108{,}5;\ 121{,}6\,\},$$

sieht das so aus:

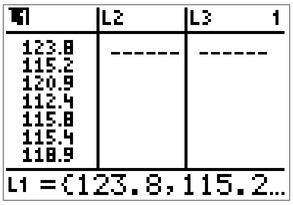

Abb. 1

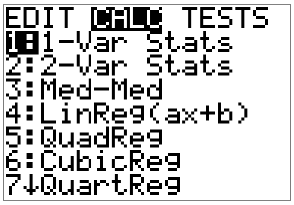

Abb. 2

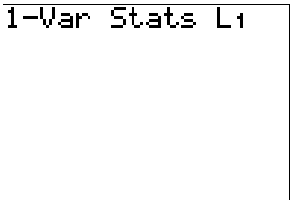

Abb. 3

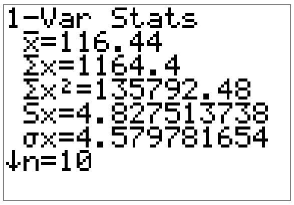

Abb. 4

Der TI-83 bietet mit $S \approx 4{,}828$ und $\sigma \approx 4{,}58$ zwei Standardabweichungswerte an, von denen der zweite Wert aufgrund seiner Bezeichnung und durch Nachrechnen als die geläufige Standardabweichung

$$\sigma = \sum_{i=1}^{n} \frac{(\mu - x_i)^2}{n} \tag{1}$$

identifiziert werden kann. Welche Bedeutung hat dagegen der Wert S? Er muss eine wichtigere Rolle als σ spielen, denn der TI-89 gibt im Gegensatz zum TI-83 unter den Verteilungskennwerten nur noch S aus. σ fehlt hier ganz:

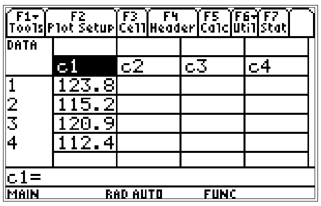

Abb. 5

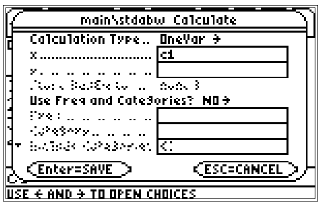

Abb. 6

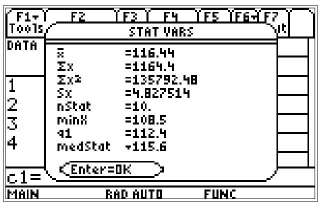

Abb. 7

Wozu zwei Standardabweichungen?

Der Schlüssel zur Lösung des Problems liegt darin, dass die beiden beschriebenen Prozeduren die Kennwerte von *Häufigkeits-* und nicht Wahrscheinlichkeitsverteilungen ermitteln. Genauer sind es Stichproben, deren Verteilungskennwerte berechnet werden. Stichproben werden erhoben, um Aussagen über die Grundgesamtheit machen zu können. Im Beispiel könnte die Stichprobe deshalb erhoben worden sein, um eine Verteilung der Körpergrößen bei sechsjährigen Jungen zu erstellen. Da man annehmen kann, dass die betrachtete Größe normalverteilt ist, reicht es, den Mittelwert und die Streuung zu beziffern. (Dabei ist der Stichprobenumfang von 10 Jungen ziemlich klein. Mit wachsendem Stichprobenumfang darf man Kennwerte erwarten, die die „wirkliche“ Grundgesamtheit immer besser beschreiben. Wir kommen darauf zurück.)

Der Mittelwert der Stichprobe wird genau so berechnet wie der Mittelwert der Grundgesamtheit, wenn man letzteren angesichts der Millionen Sechsjährigen tatsächlich vollständig erfassen könnte. Nämlich als

$$\mu = \sum_{i=1}^{n} \frac{x_i}{n} \tag{2}$$

Was die Standardabweichung betrifft, so bietet der TI-83 wie gesagt zwei Werte an, während der TI-89 sich gar nur noch auf den S-Wert beschränkt, was die Vermutung nahe legt, dass σ als Maß der Streuung einer Stichprobe womöglich ungeeignet ist. Und tatsächlich ergibt ein Blick in ein Statistiklehrbuch (Lienert 1994), dass der S-Wert die erste Wahl dafür ist, auf Grundlage von Stichproben die Standardabweichung der Grundgesamtheit zu bestimmen. Er heißt „empirische Standardabweichung“ und wird berechnet wird als

$$S = \sum_{i=1}^{n} \frac{(\mu - x_i)^2}{n-1} \tag{3}$$

Die empirische Standardabweichung S ist also wegen des Teilers $n-1$ grundsätzlich etwas größer als σ. Während mit der Formel (3) das Streuungsmaß der Stichprobe berechnet wird, liegt der Streuung der Grundgesamtheit die Formel (1) zugrunde. Dass man nun bei Stichproben die empirische Standardabweichung der σ -Standardabweichung vorzieht liegt daran, dass hier in der Regel die „wirkliche“ Streuung unterschätzt wird. Im Beispiel der Körpergrößen kann man also sagen, dass der der tatsächliche Mittelwert mit 116,44 wohl ganz gut beschrieben wird, während die tatsächliche Streuung besser mit $S \approx 4{,}828$ anzugeben ist als mit $\sigma \approx 4{,}58$. Man ist also meistens näher an der „wirklichen“ Standardabweichung, wenn man in der Stichprobe die Summe der Abweichungsquadrate durch $n-1$ teilt. Bei umfangreicheren Stichproben, also bei wachsendem n, nähern sich die empirische Standardabweichung und die σ - Standardabweichung einander an. Und das macht Sinn, denn je umfangreicher die Stichprobe ist, desto ähnlicher wird diese der Grundgesamtheit.

Woher kommt der Teiler n-1?

Warum S besser geeignet ist als σ wird in der Fachliteratur mithilfe des Begriffs „Erwartungstreue" erklärt und kann mit schätztheoretischen Mitteln bewiesen werden (Büchter und Henn 2005, S. 315). Ein anderer Erklärungsansatz nimmt Bezug auf die Freiheitsgrade einer Gleichung (Lienert 1994, S. 42). Soll nämlich eine Stichprobe die „wirklichen" Kennwerte der Grundgesamtheit liefern – was man als Idealfall ja von ihr erwartet – dann ist in der Gleichung des Mittelwertes (2) der Parameter μ als Mittelwert der Grundgesamtheit schon festgelegt. Die Stichprobenwerte x_i können bis auf den letzten „zufällig" gezogen werden. Der letzte Wert x_n dagegen muss einen bestimmten Wert annehmen, damit die Gleichung bei der idealen Stichprobe auch erfüllt ist. Es sind also n-1 Stichprobenwerte, die frei gewählt werden können. n-1 heißt demnach auch die Anzahl der Freiheitsgrade dieser Stichprobe. Da nun nur n-1 der Stichprobenwerte in der idealen Stichprobe wirklich frei sind, tut man bei der Berechnung der Standardabweichung so, als wenn der Stichprobenumfang nur n-1 beträgt. Deshalb wird die Summe der Abweichungsquadrate durch n-1 geteilt. Der n-te Stichprobenwert fällt dabei nicht unter Tisch, sondern wird in der Formel (3) weiter berücksichtigt. Ihre Abweichung vom Mittelwert wird, so kann man das sich erklären, auf die übrigen n-1 Abweichungen „verteilt", da die übrigen n-1 Stichprobenwerte diese letzte Abweichung ja „verursacht" haben.

Die Erklärung ist in dieser Kürze zugegebenermaßen unbefriedigend. Der Autor hat sich stattdessen einmal den Spaß gemacht, mittels des Rechners das Messen der Körpergröße von zehn zufällig ausgewählten Sechsjährigen zu simulieren. Zu jeder Stichprobe werden Mittelwert und σ sowie S berechnet und mit den entsprechenden Kennwerten der normalverteilten Grundgesamtheit gegenüber gestellt. Diese sind bekannt und lauten für sechsjährige Jungen μ=116,5 (cm) und σ =5,3 (Elemente d. Mathematik: LK Stochastik 2003 S. 229).

Eine statistische Überprüfung

Der Befehl `randNorm(116.5,5.3)` erzeugt eine normalverteilte Zufallszahl mit den genannten Kennwerten. Der Befehl `round(randNorm(116.5,5.3),1)` rundet diese Zahl auf eine Dezimalstelle.
Und `seq(round(randNorm(116.5,5.3),1),i,1,1Ø)` erzeugt eine Liste von zehn solchen Zufallszahlen. Das sind also die Körpergrößen der zehn aus der normalverteilten Grundgesamtheit zufällig ausgewählten Jungen. Diese Liste wird zur Bestimmung der Kennwerte mittels STO als „liste" abgespeichert. Die Befehle `mean(liste)` und `stdDev(liste)` geben den Mittelwert und die empirische Standardabweichung (also S) aus. Auf dem TI-89 muss der σ -Wert nachprogrammiert werden, und zwar wie folgt: `sqrt(sum((mean(t)-t)^2,1,dim(t))/dim(t))→σ(t)`.

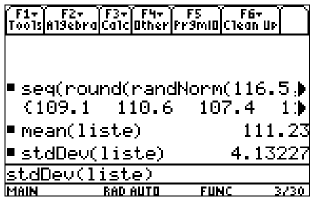

Abb. 8

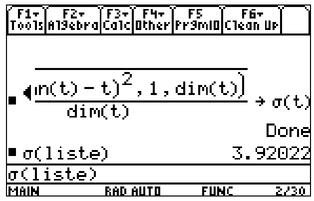

Abb. 9

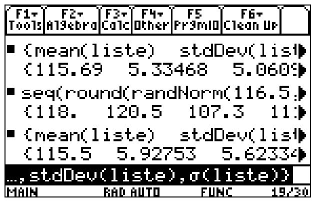

Abb. 10

Die Screenshots zeigen für die erste Stichprobe einen Mittelwert von 111,23, was dem tatsächlichen Mittelwert von 116,5 relative nahe kommt. Die empirische Standardabweichung liegt mit etwa 4,132 der tatsächlichen Standardabweichung von 5,3 näher als σ =3,92.

Eine Übersicht über die Kennwerte weiterer simulierter zufälliger Stichproben zeigt keine Präferenz für die empirische Standardabweichung:

Mittelwert	Empirische Standard-abweichung S	Standard-abweichung σ	S besser als σ
116,21	5,381	5,105	X
114,83	6,018	5,709	
117,69	7,314	6,938	
117,14	3,635	3,449	X
120,03	5,473	5,192	
115,69	5,335	5,061	X

Bei insgesamt 50 Stichproben sind es sogar nur 22 Fälle, in denen die empirische Standardabweichung als Näherungswert der tatsächlichen Standardabweichung besser geeignet war als der σ -Wert. Das ist weniger als die Hälfte. Bestätigt hat sich damit die Präferenz für den S-Wert bei Stichprobenerhebungen nicht. Würde dieses Thema im Unterricht besprochen werden, dann könnte man eine größere Anzahl an Simulationen in der Lerngruppe durchführen und zum Gegenstand eines Hypothesentests machen.

Konsequenzen für den Mathematikunterricht

Kaum ein Schulbuch, das ich überprüfen konnte, weist darauf hin, dass bei der Beschreibung von Grundgesamtheiten durch Stichproben die empirische Standardabweichung zu verwenden ist. In einem Buch ist die empirische Standardabweichung sogar sinngemäß als „σ -Standardabweichung für Stichproben" falsch definiert. Was ist also im Unterricht zu tun, wenn einerseits die empirische Standardabweichung nicht vorkommt, andererseits Rechner wie der TI-89 ein anderes Streuungsmaß gar nicht erst anbietet? Meine Vorschläge:

- Bei Problemstellungen, in denen mittels der Kennwerte von Stichproben auf die Grundgesamtheit geschlossen werden soll, sind ggf. beide Standardabweichungen zulässig. Lienert (1994, S. 42) schreibt, dass beide Werte in der wissenschaftlichen Literatur diskutiert werden.
- Bei der Bestimmung der Kennwerte von diskreten Wahrscheinlichkeitsverteilungen dagegen ist die Verwendung der empirischen Standardabweichung S unzulässig. Die Formel (3) würde im Vergleich zur korrekten Formel (1) zu niedrig sein. Wo er wie beim TI-89 nicht angeboten wird, ist es u. U. sinnvoll, σ wie oben für den TI-89 beschrieben nachzuprogrammieren.

Literatur

Andreas Büchter und Hans-Wolfgang Henn (2005): *Elementare Stochastik*. Berlin, Heidelberg, New York: Springer

Beat Eicke (2003): *Statistik*. Glarus: Pythagoras Lehrmittel

H. Griesel, H. Postel, F. Suhr (Hrsg.)(2003): *Elemente der Mathematik. Leistungskurs Stochastik*. Hannover: Schroedel

Gustav Lienert, Alexander von Eye (1994): *Erziehungswissenschaftliche Statistik*. Weinheim und Basel: Beltz

Autor:

Dr. Guido Pinkernell
Gymnasium Johanneum Lingen
http://gnetz.johanneum-lingen.de
E-Mail: guido.pinkernell@gmx.de