WAS VERSTEHEN WIR UNTER DEM ERWARTUNGSWERT?

von DAVID CASSEL, Hewett School, Norwich, England

Originaltitel in "Teaching Statistics" Vol. 11 (1989), Nr. 2:
What do we mean by the Mean?
Übersetzung: A. Müller, Coburg

Zusammenfassung: Der Autor untersucht, was Schüler unter dem Ausdruck "Erwartungswert" (Mittelwert) verstehen und beschreibt Möglichkeiten, wie eine Einführung im Unterricht erfolgen könnte.

ZDM-Klassifikation: K 40

"Zu oft spielen wir mit Zahlen, ohne ihre empirische Bedeutung genügend zu beachten" (D. HUFF, How to Lie with Statistics).

Die meisten Kinder haben bis zum Alter von 12 oder 13 Jahren drei Typen von "Mittelwerten" kennengelernt. Untersuchungen in den USA (N.C.T.M.) und Umfragen in England (A.P.U., 1980) zeigen, daß die meisten Kinder das arithmetische Mittel, den Median und den Modalwert ausrechnen, aber mit den Begriffen wenig anfangen können. Nur einige waren in der Lage, das jeweilige Vorgehen zu erläutern. Es schien ein allgemeiner Verständnismangel, was die Bedeutung der Begriffe anlangt, zu bestehen.

Als meine 11-jährige Tochter neulich von der Schule nach Hause kam und mir voller Stolz erzählte, daß sie heute das Berechnen von Mittelwerten gelernt habe, fragte ich sie, was dieser Begriff über eine Zahlenreihe aussage. Sie antwortete: "Das ist die Zahl, um die alle anderen herumliegen." Daraufhin bat ich sie, den Mittelwert der Zahlen 1, 2, 2, 3, 3 und 19 auszurechnen. Sie gab die richtige Antwort. Als ich sie fragte, ob sie allen Ernstes behaupten wolle, daß diese Zahlen um 5 herumlägen, sah sie schnell ein, wie schwach ihre Erklärung war. Dies scheint ein allgemeines Problem zu sein, das sich auch nicht mit zunehmendem Alter bessert, wie Untersuchungen zeigen. Andere Autoren haben in dieser Zeitschrift über ähnliche Schwierigkeiten beim Begriff der Streuung berichtet.

Wir müssen uns fragen, vorausgesetzt die Maße sind richtig berechnet und richtig verwendet, ob es überhaupt nötig ist, die Begriffe besser zu verstehen. EHRENBERG (1982) nennt zwei verschiedene Fälle, in denen Maße wie Mittelwert und Streuung benutzt werden müssen:

(1) Als absolut beschreibende Maße bei einer einzelnen Datenreihe.
(2) Als relative Maße zwischen vergleichbaren Datenreihen.

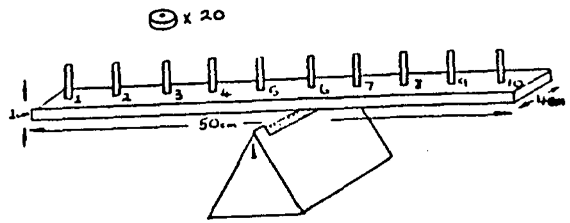
Man könnte folgern, daß es im zweiten Fall ausreicht zu wissen, wozu diese Maße verwendet werden und welche Fehler ihnen innewohnen können. Im ersten Fall dagegen braucht man ein tiefergehendes Verständnis für die Maße, um eine Aussage über die Eigenschaften der Daten machen zu können.

Ich muß zugeben, daß ich in meinem langjährigen Statistikunterricht auf allen Stufen bis kürzlich dem traditionellen Schulbuchvorgehen folgte, die einzelnen Maße anzugeben und dabei die jeweiligen Vorteile in den verschiedenen Situationen hervorzuheben. Wir fahren dann fort, den Erwartungswert und die Standardabweichung bei anderen Themen, wie das Testen von Hypothesen, als fast ausschließliche Maße für Zentralität und Streuung irgendeines Datensatzes zu benutzen. Dies wird gewöhnlich durch die Aussage gerechtfertigt, daß sie "mathematisch zuverlässiger" seien als die anderen Maße.

Um die wahre Aussagekraft des arithmetischen Mittels zu verstehen, muß seine Definition in bezug auf den erwarteten Wert betrachtet werden. WINTER (1983) regt in seinen Beiträgen über erste Vorstellungen eine direkte mathematische Analogie zum Schwerpunkt an. Besonders hilfreich ist sein Vorschlag, wie man den Schwerpunkt ausgeschnittener Schablonen von Verteilungen finden kann. In der Praxis hat dieses Vorgehen zwei Nachteile:

(1) Die Schwierigkeit, den Schwerpunkt auf einen speziellen numerischen Wert zu beziehen.
(2) Die Notwendigkeit, daß ein Schüler eine Vorstellung von einer stetigen Verteilung besitzen muß.

Um diesen Gedankengang für jüngere Schüler, die nur Erfahrungen mit diskreten Verteilungen haben, zu verdeutlichen, halte ich den Gebrauch einer einfachen Waage für sehr wertvoll.



Sie besteht aus einem 50 cm langen Holzlineal, in das Holzstäbchen in einem Abstand von 5 cm gesteckt sind und das auf einem verschiebbaren Drehpunkt aufliegt. Die von einer hohlen Stahlstange abgesägten schweren Scheiben passen auf die von 1 bis 10 numerierten Stäbchen. Um den Mittelwert von 2, 3, 5, 8, und 10 zu finden, wurden die Scheiben auf die entsprechenden Holzstäbchen gesteckt und der Drehpunkt so lange verschoben, bis der Schwerpunkt gefunden war. Der Zahlenwert wurde auf der Skala abgelesen. Arbeitsblätter, die diesen Aufbau begleiten, befähigen die Schüler, nicht nur Erwartungswerte zu berechnen, sondern auch die Auswirkungen von extremen Werten sowie die Ergebnisse bei Addition und Subtraktion von Werten zu erarbeiten. Die Vorstellung des "Gleichgewichtspunktes" einer Zahlenreihe, die die Schüler mitnehmen, ist nicht nur mathematisch korrekt, sondern bringt auch keine Verwirrung bei ungünstigen Meßwerten. EHRENBERG weist darauf hin, daß höckerförmige Verteilungen leichter zusammenzufassen sind als schiefe. Muß man die Verwendung des Erwartungswertes als Summenwert in einigen Fällen beanstanden, so ist dies bei der Interpretation als Gleichgewichtslage nicht der Fall. Die Nützlichkeit eines Maßes, daß "empfindlich" gegenüber extremen Werten ist, steht außer Frage. Ein weiterer kleiner Vorteil bei der Verwendung dieser Apparatur ist, daß jüngere Schüler bereitwilliger die Vorstellung annehmen, daß der Erwartungswert keiner der gegebenen Zahlenwerte zu sein braucht und auch keine ganze Zahl sein muß, selbst wenn alle Zahlenwerte ganz sind.

Angespornt durch den Erfolg beim Einsatz dieser Vorrichtung versuchte ich, eine Möglichkeit zu finden, diese Apparatur zur Veranschaulichung des Begriffes der Standardabweichung zu benutzen. HART (1985) weist darauf hin, daß die Hauptschwierigkeit bei der Standardabweichung die Erklärung ist, warum man Abweichungsquadrate benötigt. Nach meiner Erfahrung beachten die meisten Studenten das Vorzeichen der Abweichung einfach nicht und empfinden das Quadrieren als eine unnötige Erschwernis. Wiederum schien die Betrachtung der mathematischen Definition der Varianz der beste Weg zu sein, um eine physikalische Analogie zu finden. Ich hatte festgestellt, daß sich der Balken auf Grund seiner Flexibilität an den Enden nach unten neigte, wenn Gewichte auf die Vorrichtung gelegt wurden. Das brachte mich auf den Gedanken, daß die Standardabweichung vielleicht dadurch gemessen werden könnte, daß man diese Ablenkung am Ende der Waage auf einer Skala feststellt. Diskussionen mit Physikern verliefen in dieser Hinsicht enttäuschend. Der einzige Weg, das 2. Moment zu demonstrieren, schien der zu sein, die Waage um ihren Drehpunkt drehen zu lassen und die Drehgeschwindigkeit oder den Drehwiderstand zu messen. Es kommt hinzu, daß die Ablenkung umso größer wird, je mehr Gewichte aufgelegt werden - unabhängig von ihrer Position. Obwohl ich den Begriff der mittleren Ablenkung abgelehnt habe, habe ich die Ähnlichkeit damit dennoch als hilfreich empfunden, um zu zeigen, daß das Verhältnis der Ablenkungen "anders als linear" ist.

Wichtig erscheint hier die Feststellung, daß die Standardabweichung ein Maß der Streuung um den Erwartungswert darstellt und die höhere "Empfindlichkeit" von der Berücksichtigung der Quadrate der Abweichungen herrührt. EHRENBERG versteht den Erwartungswert als "Brennpunkt" eines Datensatzes, an den wir folgende Fragen stellen können:

(1) Wie sind die einzelnen Zahlen damit zu vergleichen?

(2) Sind die Zahlen gleichmäßig um ihn gestreut?

Analog zum Schwerpunktzentrum ist der Erwartungswert der angemessenste Brennpunkt einer Datenreihe. Selbst wenn wir keine absolute Analogie für die Standardabweichung finden, aber zeigen können, daß sie unzertrennlich mit dem Erwartungswert verbunden ist, so haben wir wenigstens eine Rechtfertigung dafür, daß wir die Standardabweichung lieber benutzen als andere Maße. Hilfreich scheint mir WAINWRIGHTs (1984) Vorschlag zu sein, absolute und quadratische Abweichungen für geeignete Datensätze aufzulisten und zu zeigen, daß der Erwartungswert die quadratische Abweichung minimiert, der Median dagegen die absolute Abweichung. Obwohl dies der Standardabweichung keine physikalische Bedeutung gibt, zeigt es trotzdem, warum absolute Abweichungen weniger geeignet sind, wenn diese auf den Erwartungswert bezogen werden. LOOSEN und andere (1985) verwendeten in Experimenten Datensätze, die lose um einen Mittelwert und dicht um Ausreißer verteilt waren und vergleichen die Standardabweichungen. Der Waagebalken könnte benutzt werden. An geeigneten Datensätzen kann man nämlich sehen, daß dort, wo die Werte gehäuft um nicht zentrale Punkte liegen, die Standardabweichung unsere instinktive Vorstellung, daß die Daten nicht weit streuen, nicht bestätigt. Diese Experimente zeigen weiter die Wichtigkeit der Standardabweichung als ein Maß für die Abweichung vom Mittelwert und weniger für die Heterogenität.

Das Vorgehen, das ich als am erfolgreichsten erachte, um Mittelwert und Standardabweichung im Unterricht zu behandeln, ist folgendes:

(1) Die Einführung des Begriffes Erwartungswert als "Gleichgewichtspunkt" eines Datensatzes an der Balkenwaage und die Verdeutlichung des Begriffes 'Standardabweichung' als eine nichtlineare Kombination von Abweichungen. WINTERs Experimente zeigen, daß mit stetig verteilten Daten ähnlich verfahren werden kann.
(2) Die Nutzung von WAINWRIGHTs Untersuchungen über quadratische und absolute Abweichungen und seine Ableitungen zum Auffinden der Standardabweichung einer Stichprobe.

(3) Die Nutzung der Beispiele von LOOSEN u.a., um die Bedeutung der Standardabweichung als ein Maß der Abweichung vom Mittelwert hervorzuheben.

Schließlich sind die Waageexperimente gut für jüngere Schüler, und sie bilden einen geeigneten Ausgangspunkt für schwierigere Aufgaben. EHRENBERG meint, daß das Erlernen der Standardabweichung keine schwierige Formel erfordert. Er schlägt für die Varianz vor:

$\mathrm{Var}(x) = E(x^2) - (E(x))^2$, wenn mit $E(x)$ der Erwartungswert der Zufallsgröße x bezeichnet wird.

Verständlicher und leichter anwendbar ist die Darstellung

Var (x) = Erwartungswert [(Differenz jedes Einzelwertes vom Erwartungswert)²]

Mit den üblichen Bezeichnungen schreibt man dafür kurz:

$\mathrm{Var}(x) = E\,[(x - E(x))^2]$.

Dies zeigt die Möglichkeit, die Standardabweichung auf niedrigerem Niveau zu unterrichten, wünschenswert für 16-jährige, die einen der vielen nichtmathematischen Leistungskurse besuchen und die Standardabweichung verwenden müssen.

Literatur

(1) Verweisungen

A.P.U. (1980): Mathematical Development-Secondary Survey Report No. 1. Assessment of Performance Unit (D.E.S.).

EHRENBERG A.S.C. (1982): A Primer in Data Reduction. John Wiley.

HART, A. E. (1984): How should we teach the S.D.? Teaching Statistics, Vol. 6, No. 1, S. 24-25.

LOOSEN, F., LIOEN, M. and LACANTE, M. (1985): The S.D.: Some drawbacks of an intuitive approach. Teaching Statistics Vol. 7, No. 1, S. 2-5.

N.C.T.M. Results from the Second Mathematics Assessment of Educational Progress. N.C.T.M. (U.S.).

WAINWRIGHT, S.J. (1984): How should we teach the S.D. - revisited. Teaching Statistics Vol. 6, No. 3, S. 86-87.

WINTER, H. (1983): Why Teach Descriptive Techniques in Schools. Proceedings of the 2nd International Conference on Teaching Statistics. Vol. 1, S. 127-143.

(2) Sonstige Literatur

BEYTH-MAROM, R. (1983): Choosing an Appropriate Summary Statistic. Teaching Statistics Vol. 5, No. 3, S. 70-74.

GADGE, M. W. (1982): Balancing a skew distribution. Teaching Statistics Vol. 4, No. 2, S. 46-48.

GORDON, T. (1986): Is the S.D. tied to the mean? Teaching Statistics Vol. 8, No. 2, S. 40-42.

HART, A. E. (1983). The non-standard deviation. Teaching Statistics Vol. 5, No. 1, S. 16-20.

Mathematik lehren. Themenheft Mittelwerte. Heft 8, Februar 1985.

SYKES, A. W. (1981): An alternative approach to the mean. Teaching statistics Vol. 3, No. 3, S. 82-86.

WINTER, H. (1983): Zur beschreibenden Statistik in der Sekundarstufe I. In: Dörfler, Fischer, R. (Hrsg.): Stochastik im Schulunterricht. Wien: Hölder-Pichler-Tempsky, Stuttgart: Teubner (1983) S. 279-304.

WINTER, H. (1985): Mittelwerte - eine grundlegende mathematische Idee. In: Mathematik lehren, Heft 8, Februar 1985, S. 4-15.