Quadratisch? Praktisch? Gut? Pseudoerklärung und Bedeutungsarmut im Unterricht über Varianz und Standardabweichung

RAPHAEL DIEPGEN, BOCHUM

Zusammenfassung: *Der Autor kritisiert die in den Schulbüchern zumeist übliche Einführung von Varianz und Standardabweichung als unmodern, uninteressant, unehrlich, wenig beziehungshaltig und damit kaum bildungswirksam; er skizziert Alternativen dazu.*

1 Einleitung

Zu denen, die von den großen und seit PISA auch öffentlich etwas mehr wahrgenommenen Defiziten und Zumutungen des milliardenverschlingenden deutschen Mathematikunterrichtes profitieren, gehört das Nachhilfeschulgewerbe. Von einem führenden Unternehmen dieses Gewerbes erhielt ich unlängst den Auftrag, für die Nachhilfeschüler der Sekundarstufe I ein kleines Nachschlageheftchen zum Thema Stochastik zu konzipieren. Als Grundlage wurden mir die Stochastikkapitel aller nach Ansicht des Nachhilfeunternehmens hinreichend weit verbreiteten Mathematikschulbuchreihen für Realschule, Gymnasium und Gesamtschule zur Verfügung gestellt. Das Studium dieser Unterlagen hinterließ bei mir immer wieder das Gefühl, man sollte, bevor man auf die in diesen Schulbüchern vorgegebene Art Stochastik unterrichte, lieber ganz darauf verzichten; und schon gar nicht könne man von dieser Art Stochastikunterricht erhoffen, dass er dem Mathematikunterricht aus seiner Krise helfe. Betrachten wir als Beispiel die in diesen Schulbüchern – und damit wohl auch im normalen Stochastikunterricht – übliche Einführung von Varianz und Standardabweichung.

2 Varianz und Standardabweichung in der üblichen Darstellung

Nach Behandlung von arithmetischem Mittelwert (und ggf. Median) von statistisch erhobenen Werten eines Merkmals wird nach deren „Streuung“ gefragt. (Dabei wird diese Streuung implizit sofort als Streuung um den Mittelwert interpretiert, nicht etwa als Unterschiedlichkeit der Einzelwerte untereinander.) Beantwortet wird die Frage durch den Vorschlag, die Differenzen zwischen den Einzelwerten und dem Mittelwert zu betrachten. Da es hier aber offensichtlich nur auf die Größe der Differenzen, also auf das Ausmaß der Abweichungen ankomme, nicht aber auf deren Richtung, sei es doch sicherlich sinnvoll, die noch störenden Vorzeichen dieser Differenzen durch Quadrieren zum Verschwinden zu bringen und schließlich die mittlere quadratische Abweichung der Einzelwerte von ihrem Mittelwert als Maß für die Streuung zu verwenden. Das Ganze wird dann formal als Varianz definiert und an einigen mehr oder minder (un)interessanten Beispieldatensätzen durchgerechnet. Ohne weitere Begründung folgt dann zumeist der Hinweis, statt der Varianz benutze man als Streuungsmaß häufig auch deren Quadratwurzel, genannt Standardabweichung. (Gelegentlich wird dabei der Eindruck erweckt, das Radizieren bei der Standardabweichung kompensiere gleichsam das Quadrieren bei der Varianz.) Auch diese Standardabweichung wird dann an einigen Beispielen formal durchgerechnet. Danach wendet sich der Unterricht wieder anderen Themen zu.

Auch die typische Darstellung in den Schulbüchern für die Sekundarstufe II scheint diesem Muster zu folgen.

3 Kritik

Macht ein solcher Unterricht Sinn? Meine Zweifel:

1. Ein solcher Unterricht erschöpft sich in definitorischen Formeln und deren rechenmäßiger Einübung. Er zeigt keine Zusammenhänge auf, er stiftet keine Bezüge. Er verankert die neuen Begriffe nicht in einem bedeutungshaltigen Netzwerk; sie dürften als isolierte Konzepte auch bald wieder vergessen sein. Insbesondere: Es gibt in diesem Unterricht überhaupt nichts zu entdecken. Er ist langweilig. Die bloße Kenntnis der unreflektierten Formeln für Varianz und Standardabweichung hat kaum Bildungswert, schon gar keinen mathematischen. Angesichts der mathematischen Trivialität dieser bloßen Formeln fragt sich überdies, warum sie eigentlich auf Vorrat im Mathematikunterricht eingeführt werden, und nicht bei Gelegenheit in den Anwendungsfächern, die dieser Formeln – angeblich – bedürfen.

2. Den Schülern werden „Begründungen“ für die Begriffsbildung suggeriert, die weder den Schüler überzeugen können, noch die tatsächlichen Gründe benennen: Dass man die Vorzeichen der Differenzen durch Quadrieren – warum denn dann nicht gleich auch durch Potenzieren mit beliebigen geraden Ex-

ponenten? – statt durch den einzig sich überzeugend anbietenden Übergang zum Absolutbetrag zum Verschwinden bringen soll, das kann doch im Ernst auch den Schüler der Sekundarstufe I nicht überzeugen; ihm bleibt nur ein resigniert hinnehmendes Achselzucken und wieder die Verwunderung über die Unverständlichkeit mathematischer Sprachspiele. Nach irgendeinem ernsthaften Begründungsversuch dafür, warum denn eigentlich eine Abweichung des Betrages 2 genauso viel zählen soll wie vier Abweichungen des Betrages 1, sucht der Schüler vergeblich. Gelegentlich findet man vor allem in den Büchern, die auch die mittlere absolute Abweichung vom Mittelwert behandeln, die ,,Begründung", man wolle – warum auch immer – durch das Quadrieren extreme Werte stärker gewichten – so als sei die Ausreißerempfindlichkeit eine erwünschte Eigenschaft. Da erscheint es dann fast noch besser, dass für den folgenden Übergang von der Varianz zur Standardabweichung ersichtlich überhaupt gar keine ,,Begründung" mehr angegeben wird.

3. Der Unterricht ist eigentlich eher rückwärts gewandt denn zukunftsorientiert. Die quadratische Begrifflichkeit der klassischen Kleinst-Quadrate-Statistik hatte vor allem historisch in Zeiten mangelnder Rechenkapazitäten große Bedeutung. Heute dagegen sind angesichts moderner Computer die rechentechnischen Vorteile der ,,quadratischen" L^2-Statistik gegenüber der ,,absoluten" L^1-Statistik weitgehend irrelevant.[1] Von daher erscheint es etwas anachronistisch, im Statistikunterricht die dem Schüler naheliegende und algorithmisch inzwischen beherrschbare L^1-Begrifflichkeit – hier etwa die mittlere oder mediale absolute Abweichung vom Median – zugunsten einer ,,komischen" L^2-Begrifflichkeit zurückzustellen, nur weil diese L^2-Begrifflichkeit früher angesichts der Rechenprobleme die einzig praktikable, damit aber nicht unbedingt die sachlich angemessenere war. Das Quadratische war nur früher das Praktische und ist damit heute nicht mehr unbedingt das Gute.

4 Vorschläge

Wenn sich denn der schulische Statistikunterricht zunächst weiterhin an der klassischen L^2-Statistik orientieren soll, obwohl möglicherweise eine Umorientierung hin auf die moderne computerbasierte L^1-Statistik die interessanteren Perspektiven eröffnet – nämlich vor allem auf numerische und algorithmische Fragestellungen –, so muss er dem Schüler auch schon auf der Sekundarstufe I für diese L^2-Begrifflichkeit sinnstiftende Argumente anbieten. Dies gilt insbesondere, wenn man den Mathematikunterricht angesichts des fraglichen Allgemeinbildungswertes vieler seiner Inhalte wenigstens als Schule vernünftigen Argumentierens versteht; hier verbietet es sich, mathematische Konzepte ohne Begründung oder gar mit irreführender Pseudobegründung einzuführen. Wie könnte eine halbwegs tragfähige Begründung für den Unterricht über Varianz und Standardabweichung aussehen?

1. Der Bezug zur ,,quadratischen" Struktur der Normalverteilung scheidet hier wohl aus. Er könnte überdies auch in den rein deskriptivstatistischen Kontexten des Statistikunterrichtes auf der Sekundarstufe I kaum überzeugen, geht es dort doch zunächst darum, die Unterschiedlichkeit von Daten zu charakterisieren, nicht aber darum, Parameter einer ,,quadratischen" Normalverteilung zu schätzen. Die allermeisten Schüler der Sekundarstufe I werden überdies in ihrem ganzen Leben die (quadratische Struktur der) Normalverteilung ohnehin nie kennen lernen.

2. Die gelegentlich in Schulbüchern erwähnte physikalische Interpretation der Varianz als Trägheitsdrehmoment in Bezug auf den Massenschwerpunkt liefert selbstverständlich keine Begründung für diesen quadratischen statistischen Begriff. Die interessante Entdeckung dieser Beziehung könnte den Unterricht aber sicherlich bereichern – zumal dann, wenn man zuvor das arithmetische Mittel als Massenschwerpunkt interpretiert hat. Diese Entdeckung hätte vielleicht auch allgemeinbildenden Wert, insofern sie exemplarisch demonstriert, dass die Mathematisierung sehr verschiedener Gegenstandsbereiche – hie Physik, dort Statistik – zu identischen Strukturen führen kann, dass also umgekehrt identische mathematische Strukturen in sehr verschiedenen Wirklichkeitsbereichen ihre Anwendung finden können.

3. Die überzeugendste und natürlichste Begründung

[1] Diese Begrifflichkeit bezieht sich auf die Darstellung einer Zufallsvariablen X als Vektor $(x_1, x_2, ..., x_n)$ in einem n-dimensionalen Koordinatensystem, das durch die n Messobjekte aufgespannt wird. Bei L^2 ist die Länge eines solchen Vektors die übliche euklidische Länge und ergibt sich daher nach Pythagoras als Wurzel aus der Quadratsumme $x_1^2 + x_2^2 + ... + x_n^2$. Erfasst die Zufallsvariable Y die Abweichungen vom Mittelwert in X, so entspricht die euklidische Länge von Y gerade dem $\sqrt{n}$-fachen der Standardabweichung in X. Bei dieser L^2-Normierung gilt insbesondere eine Äquivalenz zwischen Nullkorrelation und Orthogonalität; dies ist die Basis der klassischen linearen Statistik. Demgegenüber ist die Länge eines Vektors in L^1 die Summe der Absolutbeträge $|x_1| + |x_2| + ... + |x_n|$; dieser Länge entspricht also bei einem auf den Median 0 zentrierten Vektor das n-fache der mittleren absoluten Abweichung vom Median.

dürfte sich aus der Kleinst-Quadrate-Eigenschaft des arithmetischen Mittels ergeben, die es zuvor dann freilich herauszuarbeiten gilt, was seltsamerweise im Standardstatistikunterricht nicht geschieht, obwohl es auch auf der Sekundarstufe I ohne Weiteres möglich wäre: Der Mittelwert erscheint dann beispielsweise als optimale, d.h. verlustminimierende Wette bei Wetten mit quadratischer Verlustfunktion, und die Varianz bemisst dabei den – minimierten – durchschnittlichen Verlust ,,auf lange Sicht". Hier lässt sich die Frage nach der Varianz im Unterricht in paradigmatischen Spielsituationen motivieren: Statistisch erhoben sind beispielsweise von den Schülern *n* Merkmalswerte – etwa die Körpergrößen der Schüler selbst –, die jeweils auf einem Zettel notiert wurden; aus dem Gefäß mit diesen Zetteln wird ,,blind" gezogen; gewettet wird jeweils zuvor auf den zu ziehenden Merkmalswert, und man gewinnt jeweils einen Geldbetrag von x Cent minus der quadratischen Abweichung des getippten von dem gezogenen Merkmalswert in Cent (und nicht in Quadratcent – um die Pseudorechtfertigung des Übergangs von der Varianz zur Standardabweichung in manchen Schulbüchern zu karikieren). Was ist in diesem Spiel auf lange Sicht die beste Wette (Antwort: das arithmetische Mittel), und wie groß müsste x mindestens sein, damit sich das Spiel auf lange Sicht lohnt (Antwort: mindestens so groß wie die Varianz)? In einem solchen Unterricht gibt es etwas zu entdecken – das arithmetische Mittel minimiert quadratische Verluste (und, das sollten die Schüler im Unterricht dann natürlich auch gleich entdecken, der Median ,,absolute" oder lineare Verluste), wobei die Varianz den durchschnittlichen Verlust auf lange Sicht beziffert –, und ein solcher Unterricht klärt auf und fördert Kritikfähigkeit: Die Angemessenheit von Mittelwert und Varianz bemisst sich für den Schüler nach diesem Unterricht daran, ob die jeweilige statistische Situation durch eine quadratische Verlustfunktion angemessen modelliert erscheint – oder eben nicht. Fazit: Es erscheint mir kaum sinnvoll, im Unterricht die Varianz einzuführen, wenn man nicht die grundlegende Kleinst-Quadrate-Eigenschaft des Mittelwerts thematisiert. Die Minimalitätseigenschaften von Mittelwert und Median lassen sich mit elementaren Mitteln auch schon auf der Sekundarstufe I herausarbeiten.

4. Wählt man diesen spieltheoretischen Rahmen, können die Schüler auch sogleich darin die für die klassische varianzanalytische Statistik maßgebliche Zerlegbarkeit einer Gesamtvarianz in additive Varianzanteile entdecken – sicherlich eine wesentliche Ursache für die Beliebtheit des Varianzbegriffs in den empirischen Humanwissenschaften. Im Beispiel: Zu den oben auf den Zetteln notierten Körpergrößen wird auch das jeweilige Geschlecht notiert, und das Spiel wird dadurch modifiziert, dass man nach Ziehung des Zettels vor Abgabe seines Tipps gegen Zahlung eines Betrages von y Cent nach dem auf dem Zettel notierten Geschlecht fragen kann, auf dass man dann nach Erhalt der Antwort natürlich den geschlechtsspezifischen Mittelwert zur Wette nutzt. Wie groß darf y sein, damit sich das Nachfragen auf lange Sicht noch lohnt? Hier werden dann die Schüler entdecken können, dass die Gesamtvarianz (also der durchschnittliche Verlust in dem Wettspiel ,,Wette auf die individuelle Körpergröße des Probanden ohne Kenntnis des Geschlechts") gleich der Summe aus der durch das Geschlecht ,,erklärten" Varianz (also dem durchschnittlichen Verlust in dem Wettspiel ,,Wette auf die Durchschnittskörpergröße der Geschlechtsgruppe des Probanden ohne Kenntnis des Geschlechts") und der ,,unerklärten" Varianz (also dem durchschnittlichen Verlust in dem Wettspiel ,,Wette auf die individuelle Körpergröße in Kenntnis des Geschlechts des Probanden") ist. Und dasselbe bei ,,absoluter" oder linearer Verlustfunktion ausprobierend können die Schüler entdecken, dass diese Additivität der durchschnittlichen Verluste der Teilspiele zum durchschnittlichen Verlust des Gesamtspiels nur gilt, solange es sich um ,,quadratische" Verluste handelt.

5. Als – sicherlich weniger günstige – Alternative zum Weg über die Kleinst-Quadrate-Eigenschaft des Mittelwerts bleibt zur Sinnstiftung quadratischer Streuungsbegriffe die Ungleichung von Tschebyschow, die als Ungleichung über relative Häufigkeiten durchaus auch schon auf der Sekundarstufe I formuliert und begründet werden kann: Die relative Häufigkeit von Messwerten, die betragsmäßig von ihrem Mittelwert um mindestens t Standardabweichungen abweichen, beträgt höchstens $1/t^2$ – in dieser Formulierung ergibt sich dann auch sofort ein Sinn des Übergangs von der Varianz zur Standardabweichung. Dieser Übergang – und dabei dann auch der Name – lässt sich außerdem motivieren durch die grundlegende Idee, die Messung verschiedener Merkmale mit ,,willkürlich" definierten Einheiten vergleichbar zu machen, also durch die Idee der Standardisierung. Auch für die Schüler der Sekundarstufe I dürfte dies präzisierbar sein dadurch, dass man die ursprünglichen Messwerte in Standardwerte überführt, die jeweils den Mittelwert 0 und die Varianz 1 haben sollen. Dass dafür dann die Divi-

sion der Abweichungswerte durch die Wurzel aus der Varianz, genannt Standardabweichung, notwendig ist, lässt sich sofort einsehen. (Damit auch hier nicht der Eindruck entsteht, Standardisierung sei notwendig an einen „quadratischen“ Streuungsbegriff gebunden, könnte der Unterricht ergänzend die ebenso einfache Standardisierung aufgrund eines absoluten Streuungsbegriffes aufzeigen, nämlich die Division der absoluten Abweichungen vom Median durch die mediale absolute Abweichung vom Median.)

6. Auch den Schülern der Sekundarstufe I dürfte nachvollziehbar sein, dass man mittels einer Stichprobe des Umfangs n erhobene statistische Daten in einem $n-$dimensionalen Koordinatensystem darstellen kann, das durch die n Untersuchungseinheiten aufgespannt wird und in dem die untersuchten Merkmale oder Variablen durch Punkte bzw. Vektoren („Pfeile“) repräsentiert werden – jedenfalls für $n = 2$ und $n = 3$. In dieser – für die klassische lineare Statistik grundlegenden – Darstellungsweise können die Schüler dann entdecken, dass der Zähler der Standardabweichung gerade gleich der (euklidischen) Länge des entsprechenden (zentrierten) Vektors ist, dass also der Zähler der Varianz dem Satz des Pythagoras entsprechend das Quadrat einer Hypothenusenlänge als Summe quadrierter Kathetenlängen darstellt. Mit dieser Entdeckung bekommen die Schüler zumindest eine Ahnung davon, dass man Fragestellungen der „quadratischen“ Statistik in „geometrische“ überführen und dann als solche bearbeiten kann.

7. Ein letzter Hinweis: Es ist zwar historisch verständlich, dass „Streuung“ immer sofort als Streuung um einen Wert aufgefasst wird – ging es doch bei Gauß, dem Vater dieser Begrifflichkeit, um die Fehlerstreuung von physikalischen Messwerten um einen „wahren“ Wert. Aber in den ganz anderen Kontexten statistischer Erhebungen, an denen sich der schulische Statistikunterricht typischerweise orientiert, ist dies mangels eines „wahren“ Wertes überhaupt nicht selbstverständlich; dies gilt insbesondere für den üblichen Statistikunterricht, solange man dort den Mittelwert bzw. Median nicht als beste Wette oder beste Vorhersage konzipiert und so nach den Vorhersagefehlern oder Residuen fragt. Hier geht es vielmehr zunächst um die Frage nach der Unterschiedlichkeit von statistisch erhobenen Daten. Die auch aus unverdorbener Schülersicht naheliegende Begriffsbildung müsste sich daher eigentlich zunächst auf die paarweisen Differenzen zwischen den individuellen Werten beziehen, nicht auf die Differenzen der Einzelwerte von einem statistisch aggregierten Mittelwert oder Median. Als sinnvolles Streuungsmaß ergäbe sich dann etwa der Mittelwert oder Median der n^2 quadratischen bzw. absoluten Differenzen der n individuellen Werte untereinander – schon bei mäßig großen Stichproben angesichts der mit dem Quadrat des Stichprobenumfangs wachsenden Zahl von zu verrechnenden Differenzen nur noch mühsam zu kalkulieren. Die Suche nach numerischer Vereinfachung lässt dann die Schüler für den quadratischen Streuungsbegriff – aber eben auch nur für diesen – entdecken, dass hier die mittlere (quadratische) Abweichung der individuellen Werte untereinander gleich der (doppelten) mittleren (quadratischen) Abweichung der Einzelwerte vom Mittelwert, genannt Varianz, ist. Beginnt man also den Unterricht – von der physikorientierten statistischen Tradition unbeeinflusst – mit der naheliegenden Betrachtung der Abweichungen der Einzelwerte voneinander, erlebt der Schüler, dass sich die quadratische Begrifflichkeit nicht unbedingt als sachlich angemessener, aber zumindest als numerisch besser bewältigbar erweist. (Dass man bei diesem Unterrichtskonzept auch noch den Nenner $n-1$ des erwartungstreuen Varianzschätzers, der den Schülern heute leider[2] schon auf ihrem Taschenrechner begegnet, aufgrund einfacher Repräsentativitätsüberlegungen verständlich machen kann – vgl. Diepgen (1999) –, ist ein netter Nebeneffekt.)

[2]Leider aus zweierlei Gründen. Erstens ist das Thema Erwartungstreue für einen anwendungsbezogenen Statistikunterricht ziemlich irrelevant; ihre eigentliche Bedeutung hat Erwartungstreue nicht im Praktischen, sondern im Theoretischen, insofern sich bestimmte mathematische Eigenschaften für die Klasse aller erwartungstreuen Schätzer beweisen lassen. Zweitens begegnet dem Schüler auf dem Taschenrechner, wie seltsamerweise auch dem Lehrer gelegentlich in der didaktischen Literatur, der Nenner $n-1$ typischerweise im Zusammenhang nicht mit der Varianz, sondern mit der „empirisch“ genannten Standardabweichung

$$s = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}}.$$

Hier macht dieser Nenner aber überhaupt keinen Sinn, weil er hier – anders als bei der Varianz – gar nicht zur Erwartungstreue führt.

P.S.

Es gelang mir nicht – wen wundert's –, das Nachhilfeunternehmen davon zu überzeugen, in dem Nachschlageheftchen für die Nachhilfeschüler die skizzierten – und manch' andere – Sinn- und Begründungslücken des üblichen Statistikunterrichts zu schließen. Durch Marktzwänge zum unternehmerischen Realismus gezwungen weiß man dort vermutlich nur zu gut, dass auch der Stochastikunterricht in der Praxis all zu oft nicht wirklich auf tieferes Sinnverständnis abzielt – schon gar nicht bei den schwächeren Schülern –, sondern eher auf die entfremdete Unterwerfung unter sinn-lose Regeln. Denn was sonst ist das eintrainierte Berechnen von Varianzen und Standardabweichungen ohne jedes tiefere Verständnis dafür, warum man die Abweichungen quadriert, warum man die gewichtete Abweichungsquadratsumme radiziert, und warum ggf. den Nenner $n-1$ benutzt – oder eben auch nicht?

Literatur

Diepgen, R. (1999): Warum $n-1$ und nicht n? Erwartungstreue – leicht gemacht. Stochastik in der Schule 19 (1), 10-13.

Anschrift des Verfassers
Dr. Raphael Diepgen
Fakultät für Psychologie
Ruhr-Universität Bochum
D-44780 Bochum
`raphael.diepgen@ruhr-uni-bochum.de`