Begründungsprobleme im Statistikunterricht. Werkstattbericht eines Schulbuchautors

Raphael Diepgen, Bochum

Zusammenfassung: Es werden einige Begründungslücken diskutiert, die sich im neuen Unterricht über Beschreibende Statistik in Klasse 11 des Gymnasiums in Nordrhein-Westfalen seines nichtprobabilistischen Charakters wegen ergeben.

1. Neu in den Lehrplänen der gymnasialen Oberstufe: Beschreibende Statistik

Nordrhein-Westfalen renoviert die Mathematik-Lehrpläne für die gymnasiale Oberstufe. Insbesondere für die Jahrgangsstufe 11 zeichnen sich Veränderungen ab: Es sollen - ohne Differenzierung in Grund- und Leistungskurs - gleich alle drei Themenbereiche der weiteren Oberstufe repräsentiert werden: Lineare Algebra/Geometrie, Stochastik und Analysis. Und zwar durch die obligatorischen Themen Koordinatengeometrie, Beschreibende Statistik und Differenzialrechnung ganz-rationaler Funktionen. Es bedarf also neuer Schulbücher. Als "Stochastikautor" erhielt ich von meinem Verlag den Auftrag, im neuen 11-er Band (vgl. Bielig-Schulz u.a. 1999) einer etablierten Schulbuchreihe die Kapitel über die - für die gymnasiale Oberstufe in Nordrhein-Westfalen neue - Beschreibende Statistik zu verfassen.

Hat Beschreibende Statistik auch einen eher langweiligen Ruf - man assoziiert damit zunächst das wahrlich öde und anspruchslose Berechnen von Mittelwerten, Varianzen und Korrelationen, so schien dies dennoch ein reizvoller Auftrag: Wann erhält ein Schulbuchautor schon einmal die Gelegenheit, einen für die Schule im wesentlichen neuen Inhalt darzustellen, relativ frei von der Rücksicht auf jahrzehntelang verfestigte Unterrichtstraditionen? Auch die Vorgaben durch den neuen Lehrplanentwurf waren nur knapp - so knapp, dass sie hier wörtlich wiedergegeben seien (Stand 10.10.1997):

"*Beschreibende Statistik*

- *Erfassen, Darstellen und Aufarbeiten statistischer Daten*
- *Statistische Kenngrößen (Mittelwerte, Streuungsmaße)*
- *Interpretieren und Bewerten von Kenngrößen*
- *Ausgleichsgerade, Regression, Korrelation*

In vielen Publikationen der letzten Jahre wurde gefordert, Studienanfänger sollten über statistische Kenntnisse verfügen. Dies ist der Hauptgrund dafür, die Beschreibende Statistik in der Jahrgangsstufe 11 zu verankern. Die Schülerinnen und Schüler sollen lernen, mit großen und unübersichtlichen Datenmengen vernünftig umzugehen und sie durch Kenngrößen zu charakterisieren. Beim Umgang mit Daten, Diagrammen und Graphen soll der kritische Vernunftgebrauch bei ihnen geschärft werden.

Stochastik in der Schule 19 (1999), Nr. 3, 39-46.

Verknüpfungen mit den anderen Themen der Jahrgangsstufe sind bei der Behandlung von Ausgleichsgeraden möglich. Auch erleichtern Kenntnisse in der Beschreibenden Statistik den Zugang zu stochastischen Problemstellungen in den nachfolgenden Jahrgangsstufen.

Die Beschreibende Statistik eignet sich in besonderer Weise für die Durchführung kleinerer Projekte aus dem Umfeld der Schule. Im Experiment können Antworten auf wirklichkeitsnahe Fragen gegeben werden. Der Einsatz des Computers, etwa in Form von Tabellenkalkulation, erleichtert und beschleunigt die Verarbeitung von Daten und wird daher nachdrücklich empfohlen."

2. Erste Reaktionen zum Lehrplan und dessen Begründung

Natürlich regten sich in mir spontane Einwände: Seit wann werden in der Schule ganze Themen auf bloßen Zuruf aus den Hochschulen obligatorisch - so als hätte man noch nie die leidvolle Erfahrung gemacht, dass solche mit den Moden oder Partikularinteressen wechselnden Zurufe aus der Professorenschaft einer kritischen Nachprüfung allzu oft nicht standhalten? Ist das "vernünftige" Umgehen mit großen Datenmengen nicht eher eine Frage inhaltlicher denn mathematischer Kompetenz - und daher kaum ein im Mathematikunterricht von Mathematiklehrern isoliert anzustrebendes Lernziel?

Wie soll die Verknüpfung mit den anderen Themen der Jahrgangsstufe hergestellt werden, wenn erstens - wie zu erwarten - die Beschreibende Statistik üblicherweise vor der Analysis behandelt wird und daher die statistikspezifische Minimierung von Fehlerquadratsummen eben nicht mit dem neuen Instrumentarium der Differenzialrechnung bewerkstelligt werden kann, sondern nur mit den aus der Mittelstufe gewohnten Werkzeugen der Scheitelpunktsuche bei Parabeln, und wenn zweitens in der Koordinatengeometrie gerade nicht jene Konzepte der vektoriellen analytischen Geometrie erarbeitet werden, mit deren Hilfe man Beschreibende Statistik beleuchten könnte, etwa das der Nullkorrelation entsprechende Konzept der Orthogonalität von Vektoren?

Wieso sollen deskriptivstatistische Konzepte wie Mittelwert, Varianz und Korrelation den Zugang zu den stochastischen Problemstellungen in Klasse 12/13 erleichtern, wenn doch die stochastische Modellierung dieser Größen viel zu komplex ist, um mit den bescheidenen Mitteln der Binomialverteilung in Klasse 12/13 bewerkstelligt werden zu können?

Und wieso empfiehlt sich das "Experiment" gerade im nichtprobabilistischen, also ohne Wahrscheinlichkeiten arbeitenden Unterricht über Beschreibende Statistik, obwohl erstens nach gängiger wissenschaftstheoretischer Auffassung gerade die Gleichwahrscheinlichkeiten herstellende Randomisierung wesentliches Merkmal von Experimenten ist und obwohl zweitens die Regressionsrechnung Domäne gerade jener Wissenschaften ist, die in diesem Sinne nicht experimentieren können und daher ersatzweise auf die bloße Beobachtung des "Korrelierens" von experimentell nicht manipulierbaren Größen angewiesen sind?

Und dass gerade im Experiment - und dies auch noch im Rahmen "kleiner Projekte im Umfeld der Schule" - Antworten auf "wirklichkeitsnahe" Fragen gefunden werden können, diese Vorstellung erscheint mir denn doch nach den jahrzehntelangen Debatten über die mangelnde Realitätsnähe experimenteller Humanwissenschaften wie etwa der Psychologie mehr als frommer Wunsch denn als erfahrungsgetränktes Wissen der Lehrplankommission.

Dessen eingedenk, dass Lehrpläne auch politische Funktion haben und daher nur bescheidenen Ansprüchen an logische Stringenz genügen können, ließ ich mich von diesen Fragen nicht abschrecken und ging an die Arbeit. Manche der Fragen holten mich während der Arbeit aber immer wieder ein

Der gute Mathematiklehrer verfolgt ebenso wie der bemühte Schulbuchautor das "schöne" Ziel, die gelehrten Dinge möglichst überzeugend zu begründen. Denn gerade der Mathematikunterricht gilt als Schule vernünftigen Argumentierens, und da sind alle Begründungslücken schmerzhaft - für den Lehrer vermutlich mehr als für die Schüler. Für den etablierten Analysis- und Algebraunterricht haben alle längst gelernt, mit den Frustrationen aus fehlenden Begründungen zu leben. Für den neuen Statistikunterricht steht es indessen noch aus, sich an solchen Enttäuschungen abzuarbeiten. Und die ersten, die diese Enttäuschungen erleben müssen, sind die Schulbuchautoren. Davon sei hier kurz berichtet in der Hoffnung, man könne die Enttäuschungen Fremder nutzen, eigene zu vermeiden - und sei es auch nur, indem man die eigenen Erwartungen herunterschraubt.

3. Kritische Fragen

Warum quadratisch?

Augenscheinlich zielt der Lehrplanentwurf - insoweit ganz Abbilddidaktik - auf die klassische und weitverbreitete Kleinst-Quadrate-Statistik, also jene "parametrische" Statistik, die zu Recht oder zu Unrecht einen Großteil der Anwendungen dominiert. Die "wahren" Gründe aber, die Beschreibende Statistik seltsamerweise auf quadratischen und nicht auf absoluten Abweichungen aufzubauen, lassen sich in Klasse 11 schlechterdings nicht überzeugend darstellen. Schauen wir uns diese Gründe an:

1. Dass die Minimierung der Summe der quadratischen Abweichungen algorithmisch leichter fällt als die Minimierung der Summe der absoluten Abweichungen, dies kann im heutigen Computerzeitalter kaum noch überzeugen und wäre überdies wohl auch von der Sache her keine ausreichende Rechtfertigung, ohne weitere inhaltliche Begründung den folgenschweren Schritt von den absoluten zu den quadratischen Abweichungen zu tun. Gerade wenn man beim Vergleich von Mittelwert und Median den Schüler eindringlich darauf hingewiesen hat, welche schwerwiegenden Folgen die Wahl der einen oder der anderen Statistik haben kann - und das heißt: die Wahl der quadratischen oder der absoluten Abweichungen -, dann kann man ehrlicherweise in der Regressionsrechnung nicht mehr nur aus Gründen der rechnerischen Bequemlichkeit oder mathematischen Eleganz der einen Alternative den Vorzug geben.

2. Dass sich die mittlere quadratische Abweichung anders als die mittlere absolute Abwei-

chung in der analytischen Geometrie als euklidische Distanz oder Länge eines Vektors interpretieren lässt, mag des Mathematikers Bedürfnis befriedigen, Strukturgleichheit zu entdecken. Es ist aber kein inhaltlich überzeugendes Argument und könnte den Schüler der Klasse 11 wohl auch dann kaum überzeugen, könnte man es ihm überhaupt verständlich machen.

3. Gleiches gilt für die damit zusammenhängende Additivität des Varianzbegriffs, derer die mittlere absolute Abweichung entbehrt. Auch dieses rein innermathematische Argument kann inhaltlich nicht überzeugen, so beliebt bei vielen Anwendern auch das "Zerlegen" von Varianzen in durch dies und das "erklärte" und "unerklärte" Anteile sein mag.

4. Der quadratische Abweichungsbegriff legitimiert sich insbesondere mit Bezug zur "quadratischen" Struktur der Normalverteilung. Diese ist und bleibt dem Schüler der Klasse 11 aber unbekannt, ebenso wie der Zentrale Grenzwertsatz, der dieser Verteilung und damit dem "quadratischen" Ansatz überhaupt erst besondere Bedeutung verleiht.

5. Dass die Minimaleigenschaft der mittleren quadratischen - nicht aber absoluten - Abweichung äquivalent zu der Ausgleichseigenschaft ist, also dazu, dass die mittlere Abweichung null ist, dies lässt sich zwar dem mehr oder minder verblüfften Schüler der Klasse 11 beim Mittelwert auf diese und jene Weise demonstrieren. Diese Vorführung bleibt aber für den Schüler ziemlich irrelevant, da für ihn wohl einzig die Ausgleichseigenschaft des "ausgleichenden oder umverteilenden" Mittelwerts interessant und praktisch bedeutsam ist - also nur die eine Seite dieser Äquivalenz.

Für die zweidimensionale Regressionsrechnung ließe sich diese den quadratischen Ansatz legitimierende Äquivalenz aber nur nutzen, wenn man von der Regressionsgeraden Ausgleichseigenschaft forderte. Anders als beim Mittelwert ist diese Forderung hier aber kaum zu begründen: Beim Mittelwert folgt die Ausgleichseigenschaft aus der grundlegenden "praktischen" Idee, eine vorhandene Summe ungleicher Werte so umzuverteilen, dass die Summanden gleich werden. Aber warum um Gottes willen sollen bei einer Regressionsgeraden zur optimalen Prognose eines Kriteriums durch einen Prädiktor positive und negative Prognosefehler ausgeglichen sein? Die einzig sinnvolle Forderung ist hier, dass man mit der Regressionsgeraden im Schnitt möglichst gut trifft - und das heißt doch wohl zunächst, dass der mittlere absolute Vorhersagefehler minimal ist, ganz gleich, ob dann der mittlere Vorhersagefehler null ist oder nicht. (Dass man wie etwa gelegentlich in der Physik die Vorhersagefehler als "Messfehler" interpretieren und daher voraussetzen kann, sie seien als Summen sehr vieler voneinander unabhängiger Einflussgrößen auf lange Sicht symmetrisch um null verteilt, ist die Ausnahme.)

6. Es lassen sich ohne Verrenkungen kaum inhaltlich überzeugende Beispiele mit hinreichendem Allgemeinheitsgrad finden - zumindest ist es mir nicht gelungen -, bei denen für Schüler der Klasse 11 plausiblerweise die statistische Kleinst-Quadrate-Methode aus einer quadratischen Verlustfunktion folgen würde. Natürlich lassen sich entsprechende Wett- oder Spielsituationen konstruieren: Dies sind aber allemal wenig überzeugende Beispiele, weil sie ex post

auf eine bestimmte Mathematik hin künstlich konstruiert sind.

Warum Mittelwert statt Median?

Selbstverständlich lässt man die Schüler beim Vergleich von Mittelwert und Median entdecken, dass beide Kennwerte bei symmetrisch verteilten Daten übereinstimmen. Und selbstverständlich berichtet man von der insbesondere in der Physik beliebten Idee, bei wiederholter nichtperfekter Messung ein- und derselben Größe den Mittelwert der verschiedenen Messwerte als Annäherung an den einen "wahren" Wert zu verwenden, nämlich unter der mehr oder weniger plausiblen Annahme, die Messwerte seien auf lange Sicht symmetrisch um null verteilt.

Und da fragt der mitdenkende Schüler sofort: Warum soll man hier den Mittelwert verwenden und nicht den Median, wo doch beide angesichts der unterstellten Symmetrie auf lange Sicht übereinstimmen? Überzeugend beantworten lässt sich diese naheliegende Schülerfrage nicht: Denn in Klasse 11 ist auch andeutungsweise kaum vermittelbar, dass bei normalverteilten Größen der Mittelwert der effizientere Schätzwert ist. Auch hier mündet der Statistikunterricht in eine Schülerfrage, die er nicht beantworten kann.

Warum die Wurzel aus der Varianz?

Haben die Schüler den Begriff der mittleren quadratischen Abweichung vom Mittelwert, sprich Varianz, als zumindest nicht völlig abwegiges Streuungskonzept akzeptiert - etwa nachdem sie gesehen haben, dass dies gerade der Hälfte der mittleren quadratischen Abweichung der Messwerte voneinander entspricht -, wird von ihnen seltsamerweise sogleich auch noch verlangt, die Quadratwurzel daraus, also die Standardabweichung, als sinnvolles Streuungsmaß hinzunehmen. Für diesen aus Schülersicht höchst befremdlichen Übergang fehlt es an überzeugender Begründung im Unterricht: Der beliebte Hinweis, mit dieser Standardabweichung messe man die Streuung in denselben "Einheiten" wie den Mittelwert, ist rein formal und kann allenfalls in physikorientierten Kontexten erfreuen, in denen das "Rechnen mit Einheiten" eine wichtige Rolle spielt. Nein, der Übergang von dem "natürlichen" Begriff der Varianz s^2 zu dem "seltsamen" Begriff der Standardabweichung s erhält seinen Sinn erst im Hinblick auf probabilistische Konzepte wie die Normalverteilung und die Ungleichung von Tschebyschew. Gerade dieser sinnstiftende Bezug bleibt aber im nichtprobabilistischen Statistikunterricht verborgen. Kurzum: Der Standardunterricht über Beschreibende Statistik kann kaum zeigen, in welchem Sinne die Standardabweichung eine "Standardabweichung" ist. Entsprechendes gilt selbstverständlich für den Korrelationskoeffizienten als Wurzel aus dem Bestimmtheitsmaß.

Warum n-1 statt n?

Um so verblüffender ist es dann, dass sich an einer üblicherweise als "viel zu schwierig" geltenden Stelle im Statistikunterricht eine ganz elementare Begründung geben lässt - nämlich dafür, dass man für die Schätzung einer Populationsvarianz tunlichst die Stichprobenvarianz nicht mit dem "natürlichen" Nenner n, sondern mit dem "unnatürlichen" Nenner n-1 verwenden sollte: Dazu sollte man allerdings den Streuungsbegriff nicht bei den quadratischen Abweichungen der n Messwerte vom Mittelwert ansetzen, sondern bei den n^2 quadratischen Abweichungen der n Messwerte voneinander. Es lässt sich dann mit elementaren Mitteln zeigen, dass die übliche Varianz, also die mittlere quadratische Abweichung der Messwerte vom Mittelwert, gerade die Hälfte der mittleren quadratischen Abweichung der Messwerte voneinander ist.

Nun sind in einer Stichprobe vom Umfang n gerade n Abweichungen der Messwerte voneinander a priori null, nämlich jeweils die Abweichungen eines Messwertes von sich selbst. Der relative Anteil dieser n A-priori-Nulldifferenzen unter den n^2 Abweichungen der Messwerte voneinander ist $n / n^2 = 1/n$. In einer Stichprobe mit kleinem n ist dieser relative Anteil $1/n$ vergleichsweise groß, in einer Population mit "unendlich" großem n ist dieser relative Anteil $1/n$ "unendlich klein", praktisch also null. Gegenüber der Population sind also in der Stichprobe die A-priori-Nulldifferenzen überrepräsentiert; es resultiert daraus eine systematische Unterschätzung der Populationsvarianz (also der Hälfte der mittleren quadratischen Abweichung der Messwerte voneinander) durch die Stichprobenvarianz mit Nenner n. Diese lässt sich verhindern, indem man auch in der Stichprobe den relativen Anteil der A-priori-Nulldifferenzen auf null bringt, diese also von Vornherein nicht berücksichtigt und nur noch mit den $n^2 - n = n(n-1)$ "nichttrivialen" Abweichungen der Messwerte voneinander arbeitet. Deren Mittelwert ist aber - wie wieder leicht zu zeigen - das Zweifache der Stichprobenvarianz mit Nenner n-1. Diese elementare und gänzlich nichtprobabilistische Begründung für den "erwartungstreuen" Varianzschätzer rekurriert lediglich auf ein basales Konzept von "Repräsentativität". Sie lässt sich im weiteren Unterrichtsverlauf bei der Kovarianz ein zweites Mal nutzen.

So schön diese überraschend elementare Begründung im Falle der erwartungstreuen Varianz- und Kovarianzschätzung auch sein mag: auch sie bleibt letztlich unvollständig. Denn für die statistische Praxis ist Erwartungstreue - anders als viele von Statistikprogrammen verführte Statistikanwender denken - eine eher unwichtige Kategorie; viel wichtiger als der Erwartungswert eines Schätzers, also seine Präzision auf lange Sicht, ist hier nämlich seine Streuung, also seine Stabilität und Verlässlichkeit. Die eigentliche Bedeutung der Erwartungstreue bleibt innermathematisch, nämlich als Fundament bestimmter theoretischer Generalisierungen über die Klasse aller erwartungstreuen Schätzer.

4. Fazit

Es bleiben die oben erwähnten Begründungslücken. Wo man aber dem Schüler gegenüber letztlich nicht überzeugend begründen kann, sollte man auch nicht so tun als ob. Soll Unterricht Argumentationsfähigkeit schulen, verträgt er keine Kaschierungen. Als Autor bleibt einem also nur, auch im Schulbuch hin und wieder die bestehenden Begründungsdefizite einzugestehen, wenigstens in Fußnoten, und vielleicht in der einen oder anderen Aufgaben anzudeuten, in welcher Richtung die fehlenden Begründungen zu suchen sind.

Literatur

Bielig-Schulz, G.; Diepgen, R.; Jahnke, T.; Lapport, G.; Wuttke, H. (1999): *Mathematik 11. Schuljahr.* Berlin: Cornelsen.

Dr. Raphael Diepgen
Ruhr-Universität Bochum Fakultät für Psychologie
D-44780 Bochum — e-mail: raphael.diepgen@ruhr-uni-bochum.de