

WIE KANN MAN MIT EINEM RECHENBLATT DAS VERARBEITEN  
BIVARIATER DATEN LEHREN?

nach JEREMY J. LAGEARD

Originaltitel in "Teaching Statistics" Vol. 10 (1988), Nr. 1:

Using a Spreadsheet to Teach Coding of Bivariate Data

Übersetzung: K. Röttel, Buxheim

Zusammenfassung: Bei bestimmten Transformationen der Zahlen zweier Datenreihen ändern sich die Werte der Korrelationskoeffizienten nicht. Schüler können sich von jener Gesetzmäßigkeit bequem überzeugen, wenn sie bei ihren Arbeiten zwei vom Verfasser entwickelte Formblätter benutzen. Die Schüler lernen zudem verstehen, daß gewisse Transformationen eine Vereinfachung der Rechnungen bewirken.

ZDM-Klassifikation: K80, K84

Diese Arbeit entstand beim Unterrichten der linearen Korrelation nach dem neuen (englischen) Oxford-A-Lehrplan (1).

Die Schüler, die Sätze bivariater Daten aus mehreren Projekten besaßen, erkannten schnell, daß Korrelationsrechnungen leicht von Hand vorgenommen werden konnten, wenn es höchstens zehn Datenpaare waren; für größere Datensätze wurden die Rechnungen aber bald undurchführbar.

Schon immer vereinfacht man längere Rechnungen, indem man die Variablen umrechnet, und einige Lehrbücher (z.B. (2)) zeigen, daß eine lineare Transformation einer oder beider Variablen den Wert des Korrelationskoeffizienten nicht verändert. Manche Schüler sehen in solchen Transformationen aber zusätzliche Erschwernisse und vermeiden sie, wo immer möglich.

Das in Abb. 1 (Transparent von Acornsoft) gezeigte Rechenblatt wurde zusammen mit den Aufgabenblatt (Abb. 2) und dem Ergebnisblatt (Abb. 3) verwendet, um den Schülern eine einfache Methode des Umrechnens von Daten vorzuführen. Die Schüler arbeiteten in Zweiergruppen und benutzten die Schulcomputer.

	.....A	.....B	.....C	.....D	.....E	.....F
1	X	Y	X <sup>2</sup>	Y <sup>2</sup>	X·Y	
2						
3	6	2	36	4	12	
	8	3	64	9	24	
5	10	4	100	16	40	
6	12	5	144	25	60	
7	14	12	196	144	168	
8	-----		-----		-----	
9	50	26	540	198	304	
10	=====	=====	=====	=====	=====	
11						
12	$\bar{X}$	10		Regression	Y bez. X	
13	$\bar{Y}$	5,2		Y =	0,5 X + 0,2	
14	Var(X)	8				
15	$\sigma(X)$	2,82843				
16	Var(Y)	12,56				
17	$\sigma(Y)$	3,51401				
18	Cov(X,Y)	8,8				
19	r(X,Y)	0,3779				

Abb. 1: Rechenblatt

Die Absicht bei diesem Zugang zu Korrelationsrechnungen war, die Schüler dank der Schnelligkeit und Genauigkeit des Computers in die Lage zu versetzen, die Wirkungen der Transformationen zu beobachten.

Korrelation anhand des Rechenblattes

Dieses Formblatt soll die gesamte Korrelationsrechnung für die zwei Variablen X und Y aufnehmen:

- (1) Vervollständigen Sie das Blatt, in dem Sie alle notwendigen Werte in die Plätze B12 bis B17 schreiben. Um die Standardabweichung  $\sigma(X)$  für B15 zu erhalten, verwende man  $\sqrt{\text{Var}(X)}$  von B14.
- (2) Alle weiteren Fragen betreffen zwei Variablen X und Y, deren Werte X 2 4 6 8 10 lauten.

Y 3 5 6 9 12

Auch nach einer Transformation der Werte kehre man für die nächste Frage wieder zu diesen Ausgangswerten zurück. Ermitteln Sie den Wert des Korrelationskoeffizienten (r) für folgende Fälle:

- (a) Die oben gegebenen X- und Y-Werte.
- (b) Die um 3 vermehrten X-Werte.
- (c) Die verdoppelten X-Werte.
- (d) Transformieren Sie die X-Werte nach der Formel  $2X + 3$ .
- (e) Transformieren Sie die X-Werte nach der Formel  $AX + B$  mit selbstgewählten Werten für A und B.
- (f) Erklären Sie die Wirkung einer solchen Transformation auf den Korrelationskoeffizienten r.
- (g) Transformieren Sie die Y-Werte linear, wobei die geg. X-Werte jedoch ungeändert bleiben sollen.
- (h) Nehmen Sie eine lineare Transformation sowohl der X- als auch der Y-Werte mit unterschiedlichen Formeln vor.
- (i) Fassen Sie Ihre bisherigen Erkenntnisse zusammen.
- (j) Führen Sie eine nichtlineare Transformation wie etwa  $X^2 + 3$  für die X-Werte durch.  
Was sagen Sie zum Ergebnis?
- (k) Wie könnten Transformationen verwendet werden, um die Rechnungen von r zu vereinfachen?

Abb. 2: Aufgabenblatt

	X					Y					r
	2	4	6	8	10	3	5	6	9	12	
(a)											
(b)	X + 3										
(c)	2 X										
(d)	2 X + 3										
(e)	<input type="checkbox"/> X + <input type="checkbox"/>										
(g)						<input type="checkbox"/> Y + <input type="checkbox"/>					
(h)	<input type="checkbox"/> X + <input type="checkbox"/>					<input type="checkbox"/> Y + <input type="checkbox"/>					
(j)	X <sup>2</sup> + 3										

Abb. 3: Blatt der Korrelationsergebnisse

Die zeitraubendste Aufgabe des Arbeitsblattes war das Ermitteln der Werte in (a), das die Vertrautheit mit der Software und natürlich die Kenntnis der statistischen Formeln verlangt. Da die Schüler das Rechenblatt schon bei ähnlichen Arbeiten zum Mittelwert und zur Varianz verwendet hatten, konnten sie entsprechende Vorkenntnisse verwerten.

### Ergebnisse aus dem Aufgabenblatt

Die nach 30 Minuten Arbeit bei den Schülern eingeholten Meinungen hören sich etwa so an:

- Lineare Transformationen haben keinen Einfluß auf den Wert der ursprünglichen Korrelation.
- Die Korrelation ändert sich nur, wenn man die Variable X quadriert.
- Wenn X oder Y quadriert wird, ändert sich der Korrelationskoeffizient, dagegen bleiben die Werte der Korrelationskoeffizienten unverändert, wenn X oder Y z.B. linear transformiert wird.

So konnten die Schüler genau und schnell mit dem Rechenblatt arbeiten; sie gewannen Zeit, um aus den Rechenergebnissen Schlüsse zu ziehen.

Eine Fehlerquelle trat jedoch auf: Die Schüler müssen die lineare Transformation der Originaldaten vornehmen. Fehler an dieser Stelle führen natürlich zu falschen Ergebnissen. Glücklicherweise waren die Ausgangsdaten und Transformationen (3 addieren, verdoppeln usw.) für die Schüler so einfach, daß Fehler gleich ins Auge fielen. Schnell sahen die Schüler, daß die ersten drei Werte von r identisch waren; so glaubten sie, daß angenommen werden muß, die Werte bleiben unverändert. Dieser Umrechnungsfehler führte einen Schüler zu falschen Werten für Frage (h), die nach beiden Variablen fragte, die nach Belieben des Anwenders unterschiedlich transformiert wurden. Dies verleitete den Schüler zur falschen Schlußfolgerung, daß bei Änderung einer Koordinate die Ergebnisse sich nicht ändern, doch bei Umrechnung beider Koordinaten oder beim Quadrieren einer Koordinate sich das Ergebnis ändert.

Andere Schüler mit arithmetischen Fehlern bei Umrechnungen waren so klug, die Werte noch einmal einzugeben, wenn sich eine Änderung von r ergab, weil sie sicher waren, daß es dieselben Werte sein mußten.

Nachdem so das Umrechnen von Daten als zulässig erkannt war, erfaßten die Schüler leicht den Grundsatz, daß solche Transformationen zur Datenvereinfachung durchgeführt werden und nicht, wie einige vermutet hatten, um alles zu verkomplizieren. Das elektronische Rechenblatt scheint allgemein ein brauchbares Mittel zu sein, einen großen Bereich statistischer Rechnungen zu veranschaulichen.

### Literatur

- 1 University of Oxford Delegacy of Local Examinations. Syllabuses for 1988.
- 2 BRYARS, D.A.: Advanced Level Statistics, p. 254, 1.ed. University Tutorial Press, 1983.

### Anmerkung des Übersetzers:

Der Autor spricht die Verwendung von Computern für die Berechnung von r an, führt aber diese Rechnung im "Rechenblatt" (er nennt es "spreadsheet") vor. Sollte die r-Berechnung tatsächlich im Oxford-A-Lehrplan enthalten sein, so ist selbstverständlich die Handrechnung durchzuführen. Zu diesem Zweck seien nachstehend die notwendigen Formeln genannt:

$$\bar{x} = \frac{1}{n} \cdot \sum x, \quad \text{Var}(x) = \overline{x^2} - (\bar{x})^2, \quad \sigma(x) = \sqrt{\text{Var}(x)}$$

$$\text{Cov}(X, Y) = \overline{X \cdot Y} - \bar{X} \cdot \bar{Y}$$

$$r(X, Y) = \frac{\text{Cov}(X, Y)}{\sigma(X) \cdot \sigma(Y)}$$