Bringt der Storch die Babys? Korrelation und Kausalität im Unterricht

RAPHAEL DIEPGEN, BOCHUM

Zusammenfassung: *Anknüpfend an einen Beitrag in SiS 21 (2001) 2 diskutiert der Autor einige Aspekte des Problemfeldes Korrelation und Kausalität, mündend in einer Skizze eines Unterrichts über lineare Modelle zur kausalen Interpretation korrelativer Zusammenhänge.*

1 Einleitung

Warnungen vor der allzu voreiligen kausalen Interpretation von Korrelationen sollten im Statistikunterricht nicht fehlen – in der Schule ebenso wenig wie in der Hochschule. Freilich dürften den „unverdorbenen“ Schülern nach einer halbwegs vernünftigen Einführung des Korrelationsbegriffs kausale Überinterpretationen von Korrelationen zunächst eher fremd sein – wie auch dann die Warnungen davor, vergleichbar etwa den früher so beliebten Warnungen vor sexuellen Gefährdungen, die überhaupt (noch) nicht in der Lebenswelt von Kindern drohen. Sinnvoller und intellektuell reizvoller als solche Warnungen erscheint es allemal, die Konzepte Korrelation und Kausalität im Unterricht begrifflich ein wenig zu vertiefen.

2 Der Vorschlag von Matthews

Im letzten Heft von Stochastik in der Schule findet sich ein – des typischen Mangels deutscher Beiträge wegen aus Teaching Statistics übernommener – Beitrag von Robert Matthews (2001) mit dem symptomatischen Titel „Der Storch bringt die Babys zur Welt (p=0.008)“: Matthews präsentiert für 17 europäische Länder jeweils Daten für die Variablen „Fläche (km^2)“, „Störche (Paare)“, „Menschen (10^6)“ und „Geburtenrate (10^3/Jahr)“ – was die Störche angeht übrigens aus etwas dunkler Quelle, nämlich „persönlicher Mitteilung“ aus Kreisen der Königlichen Gesellschaft für den Schutz von Vögeln. Die aufgrund dieser Daten berechnete Korrelation von r=0.62 zwischen den Variablen „Geburtenrate“ und „Störche“ erweist sich – so Matthews – im t-Test mit einer Überschreitungswahrscheinlichkeit von p=0.008 als „hoch signifikant“ und liefert so für den Unterricht „ein nicht-triviales Beispiel einer Korrelation, die statistisch hoch signifikant ist, sich nicht unmittelbar erklären lässt und kausal doch unsinnig ist.

Tatsächlich“, so Matthews, „hat die klare Absurdität didaktischen Wert über den Korrelation/Verursachungs-Irrtum hinaus, da sie zu vermehrter Aufmerksamkeit gegenüber dem p-Wert zwingt und größere Anerkennung der Tatsache verlangt, dass die Ablehnung der Nullhypothese nicht die Richtigkeit der Ausgangshypothese impliziert.“

3 Kritik (destruktiv)

Wirklich? Wohl kaum! Die Frage der kausalen Interpretierbarkeit eines empirischen Korrelationskoeffizienten hat im Kern überhaupt nichts mit der inferenzstatistischen Problematik zu tun, ob dieser Korrelationskoeffizient „signifikant“ ist, welcher p-Wert mit ihm verbunden ist, mit welchen Annahmen über die entsprechende Korrelation in der Population er vereinbar ist. Es geht bei dieser Frage einzig darum, ob und inwieweit die durch eine Produkt-Moment-Korrelation – ganz gleich ob in einer Stichprobe oder ob in einer Population – charakterisierte Beziehung zwischen Variablen eine bestimmte kausale Interpretation zulässt.

Kurz: Ob ein „geschätzter“ oder „wahrer“ Korrelationskoeffizient kausal interpretiert werden kann, ist im Kern unabhängig davon, ob man weiß oder sich für die Annahme entscheidet, er sei Null oder eben nicht.

Die Vermischung der Problematik der kausalen Interpretierbarkeit von Korrelationen mit der inferenzstatistischen Thematik des Nullhypothesentestens bei Matthews verwirrt – wie auch sein eigenes inferenzstatistisches Vorgehen: Sind – im Ernst – die 17 erfassten europäischen Länder in irgendeinem Sinne als „Zufallsstichprobe“ aus einer (welcher?) Population zu verstehen, wie der durchgeführte Signifikanztest doch wohl voraussetzt? Gibt es irgendeinen Grund für die Annahme, die Residuen bei der Regression der Variablen „Geburtenrate“ auf die Variable „Störche“ seien als unabhängige Realisationen einer normalverteilten Zufallsgröße zu verstehen, wie der angewandte t-Test impliziert? Behauptet die Alternativhypothese des t-Testes etwa eine kausale (also nicht nur korrelative) Beziehung zwischen den Variablen, die Nullhypothese aber das Fehlen einer solchen?

Nun denn: Handelt es sich – unbeschadet der inferenzstatistischen Ungereimtheiten – dabei wenigstens um ein „nicht-triviales Beispiel einer Korrelation, die ... sich nicht unmittelbar erklären lässt“? Zunächst fällt auf, dass die präsentierte Korrelation für sich nicht sonderlich überzeugend ist, insofern sie wesentlich von zwei Ausreißern erzeugt scheint – Polen und Türkei. (Der Lehrer könnte Matthews Beispiel gerade zur Demonstration der Ausreißerempfindlichkeit der Produkt-Moment-Korrelation nutzen, ergänzt durch Hinweise auf in dieser Hinsicht weniger problematische Zusammenhangsmaße wie etwa den Rang-Korrelationskoeffizient von Spearman, der hier dann mit 0.338 auch erheblich niedriger liegt.)

Vor allem aber dürften die meisten Schüler sofort sehen, dass die Korrelation zwischen der absoluten Anzahl in einem Land pro Jahr geborener Menschen und der absoluten Anzahl der in einem Land lebenden Storchenpaare wesentlich durch die Größe des Landes zu erklären ist; insofern ist auch dieses Beispiel eher trivial – wie jedes Beispiel, in dem mit der Geburtenanzahl eine Variable korreliert wird, die von der Größe eines Landes beeinflusst ist. Trashiges Beispiel: Je öfter es auf den Straßen eines Landes „bumst“, desto mehr Kinder werden in dem Land geboren. „Bumsen bringt Babys. (p=0,0...).“

4 Kritik (konstruktiv)

Weniger trivial wäre das Beispiel erst dann, wenn man es um den allzu offensichtlichen Einfluss der Drittvariablen Größe des Landes – an Fläche und/oder Bevölkerung – bereinigte. Man korreliere etwa die Storchdichte (Quotient aus den oben genannten Variablen „Störche“ und „Fläche“) mit der Geburtenrate im Sinne der Geburten pro Jahr pro Kopf der Bevölkerung (Quotient aus den oben genannten Variablen „Geburtenrate“ und „Menschen“). Eine dabei verbleibende Korrelation – mein Rechner nennt für die Daten von Matthews allerdings nur den Schätzwert 0.146 – wäre dann nicht mehr so einfach zu erklären, und deshalb für die Schüler vielleicht verblüffender.

Als mögliche erklärende Drittvariable böte sich hier – nach einigem inhaltlichen Räsonieren – so etwas wie „Ländlichkeit“ an, etwa im Sinne der Quadratkilometer pro Einwohner (also des Quotienten aus den oben genannten Variablen „Fläche“ und „Menschen“). Nicht-triviale hypothetische Erklärungslogik: Ländlichkeit als geographische Variable begünstigt die Storchdichte, Ländlichkeit als soziologische Variable begünstigt die Geburtenrate.

Wie könnte anknüpfend an diese Problemlage ein Unterricht aussehen?

5 Kausalität und Handeln

Moderne philosophisch-wissenschaftstheoretische Reflexionen zum Kausalitätsbegriff betonen – ganz im Unterschied etwa zu platonischen Vorstellungen der Antike – eine pragmatische Komponente: A ist Ursache für B, wenn ich durch das Herstellen oder Tun von A auch B herstellen oder tun kann. Ein Unterricht zum Thema Korrelation und Kausalität sollte daher diesen Handlungskontext darstellen. Erst aus diesem Kontext wird ja überhaupt das Problem dringend, dass korrelative Studien zumeist auch auf eine erfolgreiches Handeln ermöglichende Erkenntnis zielen, diese Erkenntnis aber nicht wie experimentelle Studien auf ein Probehandeln stützen können, in dem man – wiederholt – A herstellt oder tut und dann beobachtet, ob sich – regelmäßig – B ergibt oder nicht.

So sollte denn auch das Storch-Baby-Beispiel möglichst in einen Handlungskontext eingebettet werden. Beispiel: Ein sich um die Rentenversicherung und damit um den ausbleibenden Nachwuchs sorgender Politiker liest von einer wissenschaftlichen Studie, in der sich eine po4

sitive Korrelation zwischen Storchdichte und Geburtenrate ergeben habe. Sollte er sich nun dafür einsetzen, die Storchdichte zu erhöhen (A), um den Nachwuchs zu mehren (B)?

Das durch diesen Einstieg erzeugte Erlebnis einer offensichtlich kausal nicht interpretierbaren, weil zu erfolgreichem Handeln nicht führenden Korrelation dürfte dann das Bedürfnis wecken nach Methoden, Korrelationen – und mehr ist in vielen nichtexperimentellen wissenschaftlichen Feldern nicht zu haben – durch zusätzliche Annahmen auf ihre kausale Interpretierbarkeit hin zu analysieren. Solche Methoden sind – zumindest auf der gymnasialen Oberstufe – auch schon auf der Schule diskutierbar.

6 Kausale Modellierung

Etwa so: Es seien $Z_1, Z_2, ..., Z_p$ p der Unabhängigkeit von speziellen Skaleneigenschaften wegen auf den Erwartungs- bzw. Mittelwert 0 und die Varianz 1 standardisierte Variablen mit den – ggf. empirisch geschätzten – Interkorrelationen $r_{ij} = \mathrm{E}(Z_i \cdot Z_j)$, i,j=1, ..., p. Es sei nun angenommen, dass die Variablen $Z_1, Z_2, ..., Z_p$ so angeordnet sind, dass eine Variable jeweils auf alle nachfolgenden Variablen einen kausalen Einfluss haben kann, nicht aber auf vorangehende Variablen („kausale Reihung“). Dabei sei „kausaler Einfluss“ hier im Sinne einer direkten linearen Abhängigkeit zu verstehen: Erhöhe ich die Variable Z_i um eine (Standard-)Einheit, so erhöht sich allein dadurch „im Schnitt“ – abgesehen also von zusätzlichen Zufallseinflüssen – die Variable Z_j um b_{ij} (Standard-)Einheiten.

Präziser: Es gibt zu jeder abhängigen Modellvariablen Z_j, j>1, eine spezifische Fehler- oder Residualvariable E_j mit folgenden Eigenschaften („kausale Abgeschlossenheit“): Erstens hat sie den Erwartungs- oder Mittelwert 0, weil sich in ihr alle im Modell nicht enthaltenen Zufallseinflüsse „ausmitteln“. Zweitens korreliert sie mit allen Modellvariablen Z_i mit i≠j nicht – das heißt zu null; dadurch wird ausgeschlossen, dass eine Korrelation zwischen zwei „endogenen“ Modellvariablen durch eine „exogene“ Drittvariable außerhalb des Modells bedingt ist.

Formal:

$$(2)\, Z_2 = b_{12}Z_1 + E_2$$
$$(3)\, Z_3 = b_{13}Z_1 + b_{23}Z_2 + E_3$$
$$\vdots$$
$$(p)\, Z_p = b_{1p}Z_1 + b_{2p}Z_2 + \cdots + b_{(p-1)p}Z_{p-1} + E_p$$

Oder allgemein:

$$(j)\, Z_j = \sum_{i=1}^{j-1} b_{ij}Z_i + E_j \text{ mit j} = 2, ..., \text{p.}$$

Dabei sei $\mathrm{E}(E_j) = 0$ für alle j>1 sowie

$r(Z_i, E_j) = 0$ für alle $i \neq j$, $j > 1$.

Offensichtlich sind die p(p-1)/2 Koeffizienten b_{ij} in dem eben skizzierten Sinne als kausale Effekte interpretierbar: Erhöhe ich Z_i um eine Einheit, erhöht sich „im Schnitt" Z_j (mit i<j) um b_{ij} Einheiten – und zwar unabhängig davon, ob sich über den kausalen Umweg einer gleichzeitig bewirkten Änderung anderer „zwischengelagerter" Modellvariablen Z_k (mit i<k<j) weitere Änderungen in Z_j ergeben.

Wäre etwa, um an die inhaltliche Frage oben anzuknüpfen, Z_1 die Ländlichkeit, Z_2 die Storchdichte und Z_3 die Geburtenrate, so sollten – sofern wir oben Recht haben – die Koeffizienten b_{12} und b_{13} deutlich positiv, der Koeffizient b_{23} indessen (zumindest ungefähr) null sein.

Wie lassen sich nun – dies ist die Leitfrage für den weiteren Unterricht – unter den genannten Modellannahmen aus den p(p-1)/2 nichttrivialen Korrelationskoeffizienten r_{ij} mit i,j=1, ..., p, i<j, die ebenfalls p(p-1)/2 Kausaleffekte b_{ij} ermitteln? Hier hilft folgende Ableitung (i<j):

$$r_{ij} = \mathrm{E}(Z_i, Z_j)$$
$$= \mathrm{E}\left(Z_i \cdot \left(\sum_{k=1}^{j-1} b_{kj}Z_k + E_j \right) \right) \quad \text{(nach Gleichung (j))}$$
$$= \mathrm{E}\left(\sum_{k=1}^{j-1} b_{kj}Z_iZ_k + Z_iE_j \right) \quad \text{(Distributivgesetz)}$$
$$= \sum_{k=1}^{j-1} b_{kj}\mathrm{E}(Z_iZ_k) + \mathrm{E}(Z_iE_j) \quad \text{(Erwartungswert als lineares Funktional)}$$
$$= \sum_{k=1}^{j-1} b_{kj}r_{ik} \quad (\text{Def. von } r_{ik};\ \mathrm{E}(Z_iE_j) = \mathrm{cov}(Z_i, E_j) = 0, \text{ da nach Vor. } r(Z_i, E_j) = 0 \text{ und } \mathrm{E}(E_j) = 0)$$

Es ergibt sich also ein lineares Gleichungssystem aus p(p-1)/2 Gleichungen

$$r_{ij} = \sum_{k=1}^{j-1} b_{kj}r_{ik} \text{ für alle i, j} = 1, ..., \text{p mit i} < \text{j}$$

mit p(p-1)/2 „Unbekannten", nämlich den zu bestimmenden Kausaleffekten b_{kj}, und p(p-1)/2 „Bekannten", nämlich den Korrelationskoeffizienten r_{ik}.

Der Rest ist reine Rechenarbeit.

Was sollten die Schüler zum Verständnis der obigen Ableitung wissen?

Erstens: Der Produkt-Moment-Korrelationskoeffizient zweier Variablen lässt sich (auch) als Quotient aus der Kovarianz beider Variablen und dem Produkt der Standardabweichungen beider Variablen darstellen. Dies dürfte vermutlich jeder Unterricht über die Korrelation erwähnen, auch dann, wenn er vernünftigerweise nicht auf der Kovarianz, sondern auf dem Bestimmtheitsmaß aufbaut.

Zweitens: Im Falle z-standardisierter Variablen, also Variablen mit Erwartungs- oder Mittelwert 0 und Varianz bzw. Standardabweichung 1, ist die Korrelation ebenso wie die Kovarianz zweier Variablen schlicht der Erwartungs- bzw. Mittelwert des Produkts beider Variablen. Dies lässt sich unmittelbar zeigen.

Drittens: Im Falle zentrierter Variablen, also Variablen mit Erwartungs- oder Mittelwert 0, ist die Kovarianz zweier Variablen schlicht der Erwartungs- bzw. Mittelwert des Produkts beider Variablen. Trivial.

Viertens: Der Erwartungs- oder Mittelwert ist ein lineares Funktional: Der Erwartungs- oder Mittelwert einer Linearkombination von Variablen entspricht derselben Linearkombination der Erwartungs- bzw. Mittelwerte der Variablen. Auch diese häufig benutzte Eigenschaft lässt sich unmittelbar „nachrechnen".

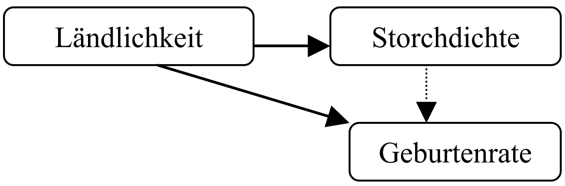
Die gedankliche Herausforderung des Unterrichts dürfte kaum in dem Aufweis dieser vier formalen Tatsachen liegen, sondern in der Analyse, welche kausalen Beziehungen einer Korrelation zwischen zwei Variablen *X* und *Y* zu Grunde liegen können. Nämlich erstens ein direkter oder indirekter kausaler Einfluss von *X* auf *Y*, zweitens ein direkter oder indirekter kausaler Einfluss von *Y* auf *X*, drittens aber ein kausaler Einfluss einer Drittvariablen auf die beiden – sich selbst gegenseitig nicht kausal beeinflussenden – Variablen *X* und *Y*, und natürlich viertens alle möglichen Kombinationen dieser drei

Möglichkeiten.

Denn nur auf dem Hintergrund dieser Verhältnisse werden die obigen Modellannahmen plausibel: Die Annahme der kausalen Reihung schließt kausale Einflüsse in umgekehrter Richtung aus, und die Annahme der kausalen Abgeschlossenheit schließt das kausale Wirken von nicht im Modell enthaltenen Drittvariablen aus. Dass dies alles unter der zusätzlichen starken Annahme steht, dass alle thematisierten kausalen Effekte – auch die der Größe null – linear sind, sollte der Unterricht nicht vergessen zu betonen.

7 Beispiel

Zurück zu unserem Beispiel mit den Variablen Ländlichkeit Z_1, Storchdichte Z_2 und Geburtenrate Z_3: Unterstellt man fiktiverweise im Unterricht für die Korrelation r_{12} zwischen Ländlichkeit und Storchdichte den Wert 0.7, für die „scheinbare" Korrelation r_{23} zwischen Storchdichte und Geburtenrate den Wert 0.5 und schließlich für die Korrelation r_{13} zwischen Ländlichkeit und Geburtenrate wieder den Wert 0.7, so ergibt sich in dem entsprechenden linearen Gleichungssystem für den fraglichen kausalen Effekt b_{23} von der Storchdichte auf die Geburtenrate mit rund 0.02 ein Wert in der Nähe der erwarteten Null. Die Störche brächten also nicht die Babys.



Arbeitet man aber mit den empirischen Korrelationen aufgrund der erwähnten Daten von Matthews ($r_{13} = 0.292$, $r_{23} = 0.146$ und $r_{13} = 0.257$) , so ergibt sich für den kausalen Effekt b_{23} von der Storchdichte auf die Geburtenrate nach meinen Berechnungen immerhin noch der Wert 0.077. Dass man dies dann kaum so interpretieren kann, dass die Störche doch „ein bisschen" die Babys bringen, dies zu erarbeiten könnte den Unterricht abrunden:

Neben dem naheliegenden inferenzstatistischen Hinweis darauf, dass die empirischen Korrelationen ja nur ungefähre Schätzwerte für die zugrundeliegenden „wahren" Korrelationen sind – die möglicherweise alle den Wert null haben –, sollten auch die Schüler das Argument finden, dass dieser vermeintliche kausale Effekt auch daran liegen mag, dass die Annahmen des Modells – insbesondere kausale Reihung wie kausale Abgeschlossenheit – vielleicht nicht die Wirklichkeit widerspiegeln, das Modell also fehlspezifiziert sein mag.

8 Abrundung

Die hier in ihren Grundideen skizzierte kausale Interpretation korrelativer Beziehungen mithilfe linearer Modelle hat – übrigens unter verschiedensten Bezeichnungen – ihre Domäne in solchen Wissenschaften, die in ihrer Empirie auf die bloße „passive" Beobachtung des Kovariierens von Variablen angewiesen sind, ohne experimentell in das Geschehen eingreifen und „aktiv" Bedingungen, insbesondere „unabhängige Variablen", variieren zu können: Ökonomie, Soziologie, Politologie, in Teilen aber auch Medizin, Psychologie usw.

Insofern die in diesen nichtexperimentellen Bereichen einzig verfügbaren Korrelationen allein häufig keine hinreichende Basis für erfolgreiches Handeln bieten („Scheinkorrelation"), ist der Wunsch verständlich, diese dennoch wenigstens unter bestimmten mehr oder minder plausiblen Modellannahmen – etwa kausale Reihung und Abgeschlossenheit – kausal interpretieren und dann, sofern halt die Modellannahmen gelten, zu erfolgreichem Handeln nutzen zu können. Hier erweist sich also Mathematisierung als nützlich.

Den Unterricht könnten zwei Themen abrunden.

Erstens der Hinweis darauf, dass Korrelationen durchaus nicht notwendig kausal interpretiert werden können müssen, um sie für erfolgreiches Handeln zu nutzen. Gäbe es etwa eine positive Korrelation zwischen dem Reichtum von Männern und der Schönheit ihrer Frauen, so könnte dies dem Anlageberater auf der Suche nach reichen Männern ebenso nützen wie dem Liebhaber auf der Suche nach schönen Frauen – völlig unabhängig davon, ob der Reichtum eines Mannes die Schönheit seiner Frau beeinflusst und/oder umgekehrt.

Zweitens weckt die Problematik der auf Drittvariablen beruhenden „Scheinkorrelation" den Wunsch, das mögliche Wirken solch störender Drittvariablen auszuschalten, zu kontrollieren oder zu neutralisieren. Und von da aus ist es nicht mehr weit bis zur zentralen Idee der Randomisierung, wie sie einen Großteil experimenteller Forschung leitet. Erfolgt die Zuordnung von Versuchsobjekten auf die verschiedenen Experimentalbedingungen nach Zufall, so wird es mit zunehmender Stichprobengröße immer unwahrscheinlicher, dass sich die Experimentalgruppen in störenden Drittvariablen unterscheiden.

Literatur

Matthews, R. (2001): Der Storch bringt die Babys zur Welt (p=0.008). Stochastik in der Schule 21 (2), 21-23.

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