

# Der Storch bringt die Babys zur Welt ( $p = 0.008$ )<sup>1</sup>

ROBERT MATTHEWS, BIRMINGHAM

Übersetzung: JOACHIM ENGEL, LUDWIGSBURG

**Zusammenfassung:** Dieser Aufsatz zeigt, dass eine statistisch hoch signifikante Korrelation zwischen der Anzahl der Störche und der Geburtenrate in den Ländern Europas besteht. Während Störche aber keine Babys zur Welt bringen können, kann eine unbeachtete Interpretation von Korrelation und  $p$ -Werten sehr wohl zu unzulässigen Schlüssen führen.

## 1 Einleitung

Einführende Lehrbücher in die Statistik warnen routinemäßig davor, Korrelation nicht mit Verursachung zu verwechseln. Während eine hohe Korrelation ein Anzeichen für einen linearen Zusammenhang ist, weisen die Bücher darauf hin, dass dies nicht als ein Maß für Kausalität angesehen werden darf. Diese Warnungen werden gewöhnlich von illustrativen Beispielen begleitet, wie z.B. der Korrelation zwischen Lesefähigkeit und Schuhgröße von Kindern oder der Beziehung zwischen Bildungsniveau und Arbeitslosigkeit (siehe z.B. Freedman *et al.*, 1998). Jedoch sind solche Beispiele oft entweder trivial aufgrund offensichtlicher erklärender Drittvariabler (z.B. Alter im Fall von Lesefähigkeit und Schuhgröße) oder sie sind keine offensichtlichen Fälle eines nicht-kausalen Zusammenhanges, d.h. sie haben zumindest offensichtliche kausale Anteile (z.B. kann das Bildungsniveau durchaus teilweise verantwortlich sein für die Länge der Arbeitslosigkeit). Im folgenden gebe ich ein Beispiel eines Zusammenhanges, der auf realen Daten basiert, bei dem Kausalität völlig absurd ist, aber auch nicht so einfach mit offensichtlichen Drittvariablen erklärt werden kann.

Mein Ausgangspunkt ist die bekannte Geschichte, dass der Klapperstorch die Babys zur Welt bringt. Die Ursprünge dieser Geschichte gehen auf die Verbindung zwischen dem Bild von Frauen als Lebensspenderinnen und dem Ernährungsverhalten von Störchen zurück, das einst als Suche nach embryonischen Leben im Wasser angesehen wurden (Cooper, 1992). Die Legende lebt bis zum heutigen Tag, wie man z.B. an Baby-bringenden Störchen sieht, die ein regelmäßiges Motiv auf Geburtstagsgrußkarten sind.

Während es offensichtlich ist, dass die Legende völliger Unfug ist, so ist es legitim, genauer nachzufra-

gen, wie man den wissenschaftlichen Widerspruch hier begründet. Geht man die Frage genauso an wie andere Zusammenhänge (z.B. zwischen bestimmten Diäten und Krebsrisiko), dann mag man sich durchaus dafür entscheiden, eine Korrelationsstudie durchzuführen, um zu sehen, ob es eine einfache Beziehung zwischen der Anzahl der Störche in einem Land und der Geburtenrate gibt. Obwohl das Vorhandensein einer statistisch signifikanten Korrelation nicht eine kausale Beziehung impliziert, würde seine Abwesenheit ein klares Argument gegen einen einfachen Zusammenhang der beiden Variablen ausdrücken. Diese Möglichkeit kann im vorliegenden Fall schnell mit Hilfe des Hypothesentestens untersucht werden, wobei die Nullhypothese lautet: "die Korrelation zwischen der Anzahl der Störche und der Anzahl der Lebendgeburten ist Null". Das werde ich nun untersuchen.

## 2 Ein Test der Storch-Geburtenrate Beziehung

Der weiße Storch (*Ciconia ciconia*) ist ein überraschend häufig vorkommender Vogel in vielen Teilen Europas und Daten bezüglich der Zahl der Brutpaare sind für 17 europäische Länder verfügbar (Harbard, 1999, persönliche Mitteilung); die jüngsten Zahlen, die sich auf die Zeit von 1980 bis 1990 erstrecken, sind zusammen mit Daten aus dem Britannica Jahrbuch von 1990 in Tabelle 1 wiedergegeben. Wenn man die Zahl der Störche in einem Streudiagramm gegen die Geburtenzahlen in jedem der 17 Länder einzeichnet, dann lassen sich deutlich Zeichen einer positiven Korrelation erkennen (siehe Abbildung 1).

Das Vorhandensein dieser Korrelation wird bestätigt, wenn man eine lineare Regression der jährlichen Geburtenrate in jedem Land (die letzte Spalte in Tabelle 1) gegen die Zahl der Brutpaare des weißen Storches (Spalte 3) durchführt. Dies führt zu einem Korrelationskoeffizienten von  $r = 0.62$ , dessen statistische Signifikanz mit dem  $t$ -Test beurteilt werden kann, wobei  $t = r \cdot \sqrt{(n-2)/(1-r^2)}$  und  $n$  der Stichprobenumfang ist. In unserem Fall ist  $n = 17$ , so dass  $t = 3.06$ , was bei  $n - 2 = 15$  Freiheitsgraden zu einem  $p$ -Wert von 0.008 führt.

<sup>1</sup>Übersetzung aus *Teaching Statistics*, 2000 (2), 36-38

Land	Fläche (km <sup>2</sup> )	Störche (Paare)	Menschen (10 <sup>6</sup> )	Geburtenrate (10 <sup>3</sup> /Jahr)
Albanien	28 750	100	3.2	83
Belgien	30 520	1	9.9	87
Bulgarien	111 000	5000	9.0	117
Dänemark	43 100	9	5.1	59
Deutschland	357 000	3300	78	901
Frankreich	544 000	140	56	774
Griechenland	132 000	2500	10	106
Holland	41 900	4	15	188
Italien	301 280	5	57	551
Österreich	83 860	300	7.6	87
Polen	312 680	30 000	38	610
Portugal	92 390	1500	10	120
Rumänien	237 500	5000	23	23
Spanien	504 750	8000	39	439
Schweiz	41 290	150	6.7	82
Türkei	779 450	25 000	56	1 576
Ungarn	93 000	5000	11	124

Tab. 1: Geografische Daten sowie Geburtenrate und Störche für 17 Länder Europas

### 3 Analyse

Was sollen wir mit diesem Result anfangen, das auf eine hoch-signifikante statistische Korrelation zwischen der Zahl der Storchenpaare und den Geburtenraten hindeutet? Der Korrelationskoeffizient ist nicht übermäßig hoch, aber gemäß seinem  $p$ -Wert gibt es nur eine 1 zu 125 Chance einen mindestens so außergewöhnlichen Wert zu erhalten, *vorausgesetzt* die Nullhypothese "Keine Korrelation" ist wahr. Dennoch impliziert dies *nicht* - wie mit jedem  $p$ -Wert (und im Gegensatz zu dem, was unachtsame Nutzer glauben), dass die Wahrscheinlichkeit, reiner Zufall sei *tatsächlich* die korrekte Erklärung, 1 zu 125 sei; noch weniger impliziert das Ergebnis eine  $124/125 = 99.2\%$  Wahrscheinlichkeit dafür, dass der Klapperstorch die Babys bringt.

Solche scheinbar pedantischen Unterscheidungen werden oft von Konsumenten von  $p$ -Werten übersehen. Im Fall der Korrelation zwischen den Störchen und den menschlichen Geburten scheinen sie aber nicht mehr länger pedantisch: tatsächlich weisen diese Überlegungen den Weg, wie sich hier eine offenkundig absurde Schlussfolgerung vermeiden lässt. Die aller-plausibelste Erklärung für die beobachtete Korrelation ist natürlich die Existenz einer beide Phänomene erklärenden Drittvariable: irgendein Faktor, den beide Variable, Geburtenrate und Anzahl der Storchenpaare, gemeinsam haben, der - wie das Alter in der Korrelation zwischen Lesefähigkeit und

Schuhgröße - zu einer statistischen Korrelation zwischen den beiden Variablen führt, die nicht direkt miteinander verbunden sind. Ein möglicher Kandidat für eine explanative Drittvariable ist die Landfläche: die Leser sind aufgefordert, diese Möglichkeit mit Hilfe der Daten in Tabelle 1 zu untersuchen.

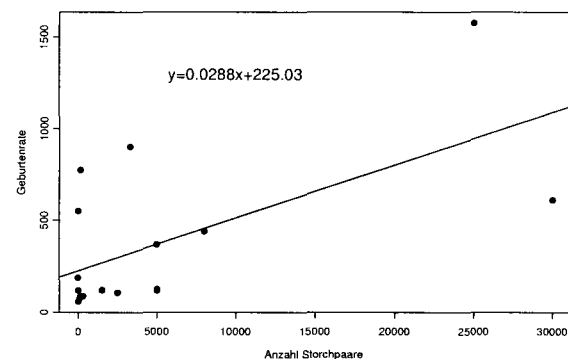


Abb. 1: Menschliche Geburtenrate und Anzahl brütender Storchenpaare in 17 Ländern Europas ändert

### 4 Schlussfolgerungen

Statistische Standardlehrbücher warnen vor dem Irrtum, Korrelation mit Verursachung zu verwechseln, aber ihre Beispiele sind entweder trivial, weil erklärende Drittvariable vom Kontext her offensicht-

lich sind, oder die Beziehungen zwischen den beiden untersuchten Variablen erscheinen eben doch zumindest partiell kausal. Die empirische Beziehung zwischen der Zahl der brütenden Storchpaare und der menschlichen Geburtenrate in 17 europäischen Ländern liefert hingegen ein nicht-triviales Beispiel einer Korrelation, die statistisch hoch signifikant ist, sich nicht unmittelbar erklären lässt und kausal doch unsinnig ist. Tatsächlich hat die klare Absurdität didaktischen Wert über den Korrelation/Verursachungs-Irrtum hinaus, da sie zu vermehrter Aufmerksamkeit gegenüber dem  $p$ -Wert zwingt und eine größere Anerkennung der Tatsache verlangt, dass die Ablehnung der Nullhypothese nicht die Richtigkeit der Ausgangshypothese impliziert.

### Danksagung

Der Verfasser dankt Chris Harbard von der Royal Society for the Protection of Birds für die Storch-Daten und Professor Dennis Lindley für wertvolle Diskussionen.

### Literatur

Cooper, J.C. (Hrsg.) (1992). *Brewer's Myth and Legend*. London: Cassell.

Freedman, D., Pisani, R. und Purves, R. (1998). *Statistics* (3. Auflage). New York: W.W. Norton.

Anschrift des Verfassers

Robert Matthews

Aston University

Birmingham, England [rajm@compuserve.com](mailto:rajm@compuserve.com)

## TAGUNGSANKÜNDIGUNG

Tagung des Arbeitskreises *Stochastik in der Schule* der Gesellschaft für Didaktik der Mathematik:

### Stochastisches Denken und Perspektiven für das Stochastik-Curriculum 1-13

Zeit: 9. bis 11. November 2001  
Ort: Reinhardswaldschule bei Kassel  
Anmeldung: bis 15. September 2001 über  
[ak-stochastik@ph-ludwigsburg.de](mailto:ak-stochastik@ph-ludwigsburg.de) oder  
AK Stochastik, Institut für Mathematik und Informatik,  
PH Ludwigsburg, Postfach 220, 71602 Ludwigsburg

#### Programm:

1. Vorträge zu Begriff und Strukturen stochastischer Denkweisen
  - Manfred Borovcnik (Klagenfurt): Vom stochastischen Denken zur aktuellen Diskussion um statistisches Denken
  - Rolf Biehler (Kassel): Konzepte zum statistischen Denken
  - Hans-Dieter Sill (Rostock): Zur Prozessbeobachtung zufälliger Erscheinungen und ihren Konsequenzen für den Unterricht
2. Diskussion zum Stochastik-Curriculum und Erarbeitung eines Positionspapiers mit folgenden Impulsen:
  - Ergebnisse einer Lehrplananalyse zur Stochastik (Arbeitsgruppe Borovcnik, Neubert, Sill, Warmuth)
  - Positionen zu einem nationalen Stochastikcurriculum für die Klassen 1-13 (Arbeitsgruppe Borovcnik, Neubert, Sill, Warmuth)
  - Überblick über die NCTM-Standards (Engel)
3. Vortrag von Siegfried Zseby (Berlin): Die Rolle von Simulationen in der Finanzmathematik

Ziel des Seminars ist es, ein Positionspapier des Arbeitskreises zum Stellenwert der Stochastik im Schulcurriculum der Klassen 1 - 13 zu diskutieren und zu verabschieden. An alle angemeldeten Teilnehmer werden weitere Papiere zur Vorbereitung dieser Stellungnahme verschickt.