

„Trendanalysen“ mit Fangzahlen

VON

Erwin Rudolf S c h e r n e r

"Trendanalysen" mit Fangzahlen

von

Erwin Rudolf S c h e r n e r

Die Notwendigkeit einer speziellen "Einführung in die Statistik für Feldornithologen" begründete FLIEGE (1976) insbesondere mit der Existenz von Lehrbüchern und ähnlichen Publikationen, die "manchem Anfänger den Mut zur Lektüre nehmen", dem "Amateur das Eindringen in ein neues Gebiet" erschweren oder so verfaßt sind, daß "der Laie Verständnisschwierigkeiten hat". Fehlerhafte oder gar unzulässige Anwendungen statistischer Methoden in der Vogelkunde wären demnach schon deshalb kaum verwunderlich, weil dieser Wissenschaftszweig hauptsächlich von Personen getragen wird, die weder einer Forschungsinstitution angehören noch über eine geeignete akademische Ausbildung verfügen.

Bei der Durchsicht renommierter Periodika wie "Ecology", "American Midland Naturalist" und "Journal of Animal Ecology" entdeckte HURLBERT (1984) 101 Freilandstudien mit Signifikanzprüfungen, von denen immerhin 48 % unzulässigerweise erfolgt waren. Eine noch höhere Quote falscher Anwendungen, nämlich 78 %, ermittelte UNDERWOOD (1981) für 143 von Meeresbiologen veröffentlichte Varianzanalysen. Der Mißbrauch induktiver Statistik ist demnach keine Besonderheit der Ornithologie, sondern ein generelles, offenbar weit verbreitetes Problem in den Biowissenschaften.

Falscher Gebrauch induktiver Statistik ist meist darauf zurückzuführen, daß die verwendeten Daten keine Stichproben darstellen oder "statistisch signifikante" Befunde automatisch als biologisch bedeutsam angesehen werden. Aktuelle Beispiele dafür bieten die von FLIEGE (1986) beschriebenen, im "Mettnau-Reit-Illmitz-Programm" der Vogelwarte Radolfzell schon mehrmals vorgenommenen "Trendanalysen" mit Fangzahlen (s. BERTHOLD 1977, BERTHOLD & QUERNER 1978, BERTHOLD et al. 1986).

An einem Fangplatz bei Illmitz während der Wegzugperiode erfaßte Drosselrohrsänger (Tab. 1) waren für FLIEGE Anlaß zur Untersuchung folgender Frage: "Hat sich der Bestand signifikant verändert?" Die Antwort fehlt jedoch. Die Berechnung des SPEARMAN-Rangkorrelationskoeffizienten ($r_s = -0,7091$) und die anschließende Signifikanzprüfung ($p < 0,05$) ergaben stattdessen: "Die Fangzahlen des Drosselrohrsängers nehmen ab". Mit dieser nicht ganz korrekt formulierten Aussage wird die Problematik derartiger "Trendanalysen" sehr deutlich.

Tab. 1: Fangzahlen des Drosselrohrsängers (*Acrocephalus arundinaceus*) bei Illmitz am Neusiedler See während der Wegzugperiode (aus FLIEGE 1976).

Jahr	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983
Fänglinge	565	400	718	580	409	201	298	205	213	212

Daß die Fangzahl des Drosselrohrsängers oder, genauer, die Anzahl der bei Illmitz während der Wegzugperiode gefangenen Drosselrohrsänger im Zeitraum 1974-1983 abgenommen hat, ist aus den vorhandenen Daten unmittelbar ersichtlich (Tab. 1). Informationen über r_s und p sind also entbehrlich. Rückschlüsse auf den "Bestand" wären ohnehin zweifelhaft, weil FLIEGE diesen nicht definiert hat. Der hier zu beurteilende "Bestand" könnte beispielsweise die bei Illmitz, am gesamten Neusiedler See, im Burgenland oder in Mitteleuropa während der Wegzugperiode erscheinenden Drosselrohrsänger umfassen. Eine klare Definition ist jedenfalls unerlässlich, da sie maßgeblich die Aussagekraft der vorliegenden Befunde bestimmt. Zudem bleibt zu prüfen, ob die Fangzahlen einen alljährlich gleichen Anteil der Grundgesamtheit repräsentieren. Dabei sind zwei Fälle zu unterscheiden:

- (a) Die Fangquote ist konstant, so daß Veränderungen der Grundgesamtheit unmittelbar ersichtlich und Signifikanzprüfungen überflüssig sind. Die Beurteilung von Ausmaß (z.B. geringfügig) und Ursache (evtl. "zufällig") der aufgezeigten Fluktuationen obliegt ohnehin dem Sachverstand des Biologen (SCHERNER 1989).
- (b) Die Fangquote ist variabel, weshalb hinsichtlich der Fluktuationen des "Bestandes" Ungewißheiten existieren, die mit induktiver Statistik nicht zu beseitigen sind und den Gebrauch letzterer eigentlich ausschließen.

Sofern, wie im zweiten Fall, induktive Statistik eigentlich unzulässig ist, erscheint zweifelhaft, ob Signifikanzprüfungen dann wenigstens "als Orientierungshilfen zu werten" sind (BERTHOLD et al. 1986). Für den Rangkorrelationskoeffizienten¹⁾ läßt sich diese Problematik anhand einer Urne erläutern, die Kugeln unterschiedlicher Größe enthält (Grundgesamtheit).

Wenn n Kugeln der Urne zufällig entnommen, nach ihrer Größe (X) geordnet (Rangfolge 1, 2, 3, ..., n) und dann mit der Reihenfolge (Y) ihrer Ziehungen (Rangfolge 1, 2, 3, ..., n) verglichen werden, sind $n!$ verschiedene Anordnungen (Permutationen) möglich, die gleiche Wahrscheinlichkeiten haben, falls Kugelgröße und Ziehungszeitpunkt nicht korrelieren, tatsächlich also $R = 0$ gilt (Grundgesamtheit). Für $n = 3$ ergeben sich z.B. $3! = 3 \cdot 2 \cdot 1 = 6$ Permutationen mit Einzelwahrscheinlichkeiten von $1/6$ oder $0,17$ bzw. 17% (s. Tab. 2).

Tab. 2: Mögliche Anordnungen der Rangfolgen nach Größe (X) und Ziehungszeitpunkt (Y) für eine aus drei Kugeln bestehende Zufallsstichprobe mit den zugehörigen Korrelationskoeffizienten (r_s).

Permutationen	X	Y	X	Y	X	Y	X	Y	X	Y	X	Y
	1	3	1	3	1	2	1	2	1	1	1	1
	2	2	2	1	2	3	2	1	2	3	2	2
	3	1	3	2	3	1	3	3	3	2	3	3
r_s	- 1,0		- 0,5		- 0,5		+ 0,5		+ 0,5		+ 1,0	

Offensichtlich ist eine Auswahl von nur drei Kugeln kein überzeugender Beweis für einen in der Grundgesamtheit vorhandenen Zusammenhang zwischen X und Y ($R \neq 0$). Eine derartige Annahme wäre mit einem ziemlich hohen Irrtumsrisiko behaftet. Selbst der Extremwert $r_s = + 1,0$ böte

1) Für Stichproben r_s , in der Grundgesamtheit R ("wahrer" Wert des Korrelationskoeffizienten).

noch zu wenig Sicherheit, weil er ja auch bei Fehlen einer Korrelation ($R = 0$) in durchschnittlich jeder sechsten Zufallsstichprobe eintreffen würde.

Statistische Rückschlüsse auf Grundgesamtheiten sind stets mit einem Irrtumsrisiko verbunden. Dieses wird üblicherweise toleriert, wenn es höchstens 5 % beträgt ($p \leq 0,05$). Der in Stichproben vorgefundene Sachverhalt gilt dann als "signifikant" (mit $p \leq 0,01$ oft auch "hoch signifikant"). Klarheit darüber erbringt der "Signifikanztest", dem das folgende Beispiel einer zufälligen Auswahl von fünf Kugeln gewidmet ist. Den Hintergrund bilden $5! = 120$ Permutationen, denen 21 verschiedene r_s -Werte mit Wahrscheinlichkeiten von 0,8 bis 8,3 % zukommen, sofern zwischen X und Y keine Korrelation besteht (Tab. 3).

Tab. 3: Mögliche r_s -Werte und ihre Häufigkeiten für sämtliche, gleichermaßen wahrscheinliche Permutationen der Rangfolgen nach Größe und Ziehungszeitpunkt von fünf Kugeln.

r_s	Permutationen	relative Häufigkeit
+ 1,0	1	0,008
+ 0,9	4	0,033
+ 0,8	3	0,025
+ 0,7	6	0,050
+ 0,6	7	0,058
+ 0,5	6	0,050
+ 0,4	4	0,033
+ 0,3	10	0,083
+ 0,2	6	0,050
+ 0,1	10	0,083
0,0	6	0,050
- 0,1	10	0,083
- 0,2	6	0,050
- 0,3	10	0,083
- 0,4	4	0,033
- 0,5	6	0,050
- 0,6	7	0,058
- 0,7	6	0,050
- 0,8	3	0,025
- 0,9	4	0,033
- 1,0	1	0,008
Summe	120	1,000 $\hat{=} 100,0$ %

Bei "zweiseitiger" Fragestellung darf ein Zusammenhang zwischen X und Y in der Grundgesamtheit ($R \neq 0$) angenommen werden, wenn die Stichprobe $r_s = - 1,0$ oder $r_s = + 1,0$ ergibt. Das Irrtumsrisiko ist tolerierbar klein ($p = 0,016$). Die Wahrscheinlichkeit, daß Zufallsstichproben (mit $n = 5$) auch dann, wenn tatsächlich keine Korrelation besteht, zu Werten führen, die außerhalb des Intervalls $- 0,9 < r_s \leq + 0,9$ liegen, beträgt nur 1,6 %. (Der etwas kürzere Bereich $- 0,8 \leq r_s \leq + 0,8$ wäre hier ungeeignet, da er immerhin 8,2 % der in Tab. 3 genannten r_s -Werte nicht enthält.)

Unter gewissen Voraussetzungen kann "einseitig" geprüft werden, ob in der Grundgesamtheit eine positive Korrelation vorliegt ($R > 0$). Eine solche ist anzunehmen, wenn die Stichprobe $r_s = + 0,9$ oder $r_s = + 1,0$ ergibt. Die Wahrscheinlichkeit, daß Zufallsstichproben (mit $n = 5$) Resultate liefern, welche den "kritischen" Wert $r_s = + 0,8$ überschreiten,

obwohl X und Y tatsächlich voneinander unabhängig sind, beträgt lediglich 4,1 % ($p = 0,041$). Entsprechend bildet $r_s = -0,8$ den Zufalls-höchstwert, falls für die Grundgesamtheit $R < 0$ erwartet wird; s. Tab. 3).

Wie das vorstehende Beispiel zeigt, lassen sich Irrtumswahrscheinlichkeiten exakt berechnen, wenn die Stichproben zufällig, gleichsam im Lotterieverfahren erfaßte Teilmengen ihrer Grundgesamtheiten repräsentieren²⁾. Wird diese Bedingung ignoriert, entstehen systematische Fehler, deren Konsequenzen kaum abschätzbar sind, und die wahren p-Werte bleiben unbestimmt. Ob die von BERTHOLD et al. mit "allerdings nicht zufällig gewonnenen" Fangzahlen vorgenommenen Signifikanzprüfungen die erhofften "Orientierungshilfen" bieten können, ist demnach ungewiß.

Ebenfalls Beachtung verdient die Tatsache, daß p-Werte die Zuverlässigkeit statistischer Folgerungen, nicht aber deren Inhalt betreffen³⁾. Die von BERTHOLD & QUERNER (1978) sowie BERTHOLD et al. (1986) mit Irrtumswahrscheinlichkeiten vorgenommene Unterscheidung zwischen "Trend" (bei $p < 0,05$) und "Tendenz" (bei $p > 0,05$) ist deshalb mehr als nur ein semantisches Problem. Es handelt sich um den Versuch einer generalisierenden Aussage, nachdem der Signifikanztest ergeben hat, daß die verfügbaren Daten keine ausreichende Grundlage dafür bieten (vgl. SCHERNER 1986).

Werden die im Modell verwendeten Kugeln durch Fangzahlen ersetzt, so tritt das entscheidende Argument gegen die Zulässigkeit induktiver Statistik zutage: Eine bei Illmitz oder anderenorts bestehende Grundgesamtheit von Fangzahlen, der man im Rahmen eines Forschungsprogrammes - zufällig oder auf andere Weise - eine Stichprobe entnehmen könnte, ist wohl unvorstellbar.

Für "Trendanalysen" mit Fangzahlen ist abschließend festzustellen, daß der SPEARMAN-Rangkorrelationskoeffizient lediglich eine Möglichkeit der vereinfachenden Beschreibung von Datenreihen darstellt, nicht anders als Mittelwert oder Variationsbreite bei Meßserien (vgl. JAMES & MCCULLOCH 1985). Signifikanzprüfungen jedoch sind hier unlogisch und daher irreführend, was freilich selten so offenkundig wird wie im folgenden, von BERTHOLD (1977) formulierten Ergebnis fünfjähriger Berin-gungsarbeit am Bodensee: "Die Anzahl gefangener Gartenrotschwänze nahm möglicherweise ab, die des Schilfrohrsängers möglicherweise zu ($p < 0,10$)."

Schrifttum

B e r t h o l d , P. (1977): Über die Bestandsentwicklung von Kleinvogelpopulationen: Fünfjährige Untersuchungen in SW-Deutschland. Vogelwelt 98: 193-197. - B e r t h o l d , P., G. F l i e g e , U. Q u e r n e r & H. W i n k l e r (1986): Die Bestandsentwicklung von Kleinvögeln in Mitteleuropa: Analyse von Fangzahlen. J. Orn. 127: 397-437. - B e r t h o l d , P., & U. Q u e r n e r (1978): Über Bestandsentwicklung und Fluktuationsrate von Kleinvogelpopulationen: Fünfjährige Untersuchungen in Mitteleuropa. Orn. Fennica 56: 110-123. - F l i e g e , G. (1986): Einführung in die Sta-

2) In der Praxis sind solche Berechnungen meist nicht erforderlich. Bei Signifikanzprüfungen werden gewöhnlich die in Lehrbüchern für verschiedene Irrtumswahrscheinlichkeiten und Stichproben tabellierten Zufallshöchstwerte benutzt, denen freilich die hier am Modell erläuterten Gedankengänge zugrunde liegen.

3) Die in manchen Veröffentlichungen erkennbare Annahme, daß ein "hoch signifikanter" Unterschied zwangsläufig besonders groß oder wichtig sei, beruht auf einem Mißverständnis induktiver Statistik.

tistik für Feldornithologen. Vogelwarte 33: 257-280. - H u r l b e r t , S.H. (1984): Pseudoreplication and the Design of Ecological Field Experiments. Ecol. Monographs 54: 187-211. - J a m e s , F.C., & C.E. M c C u l l o c h (1985): Data Analysis and the Design of Experiments in Ornithology. In: R.F. J o h n s t o n : Current Ornithology. Bd. 2. New York und London: 1-64. - S c h e r n e r , E.R. (1986): [Rezension.] Vogelk. Ber. Niedersachsen 18: 105. - S c h e r n e r , E.R. (1989): Welche Signifikanz haben Ergebnisse langfristiger Brutvogel-Bestandsaufnahmen? Limicola 3: 137-143. - U n d e r w o o d , A.J. (1981): Techniques of Analysis of Variance in Experimental Marine Biology and Ecology. Oceanography Marine Biol. 19: 513-605.

Anschrift des Verfassers: Dr. E.R. Scherner, Gesellschaft für biologische Landeserkundung, Strandallee 3, D-2893 Butjadingen-Tossens

Literaturbesprechungen

B a u r , H.W. (1987): Gewässergüte bestimmen und beurteilen. Praktische Anleitung für Gewässerwarte und alle an der Qualität unserer Gewässer interessierten Kreise. 141 S., 158 Einzeldarstellungen in 68 Abb., 20 Tab., 2., neubearb. Aufl. ISBN 3-490-04414-2. Parey. Hamburg, Berlin. -

Vorteile der alle wichtigen chemischen und biologischen Beurteilungskriterien zusammenfassenden Bearbeitung sind gründlichere Beschreibungen der Bio-Indikatoren (vgl. S. 22-53). Die Tabellen A-D bieten Anfängern eine einfach nachzuvollziehende Anleitung. Sie richtet sich vorzugsweise auf Fischereibelange aus, wozu auch die Empfehlung für Untersuchungsgeräte (Kompaktlabor) des Verbandes Deutscher Sportfischer dient.

B a y e r i s c h e s L a n d e s a m t f ü r U m w e l t s c h u t z (1988): Beiträge zum Artenschutz 7. Naturschutz in der Kulturlandschaft. 270 S., Abb., Photos, Tab. ISSN 0723-0028. München. Zu beziehen über: Bayerisches Landesamt für Umweltschutz, Rosenkavalierplatz 3, 8000 München 81. -

In 18, z.T. sehr detaillierten Beiträgen werden theoretisch, aber noch mehr praktisch-forschend Problembereiche des Naturschutzes in der Kulturlandschaft von Bayern beschrieben (s. Methoden zur Umweltverträglichkeitsprüfung - UVP; Streuobstwiesen, Rebflächen, Trockenhänge, Enzian-Schillergrasrasen, Uferstreifen an kleineren Fließgewässern, Teichgebiete, Acker- und Wiesenrandstreifenprogramme, extensive Ackerntzung, Schutzprogramme für Ackerwildkräuter). Die nützlichen Anregungen dieses Heftes werden sicherlich bei bestimmten Problembereichen auch außerhalb von Bayern größeres Interesse finden.

ZOBODAT - www.zobodat.at

Zoologisch-Botanische Datenbank/Zoological-Botanical Database

Digitale Literatur/Digital Literature

Zeitschrift/Journal: [Beiträge zur Naturkunde Niedersachsens](#)

Jahr/Year: 1989

Band/Volume: [42](#)

Autor(en)/Author(s): Scherner Erwin Rudolf

Artikel/Article: [„Trendanalysen“ mit Fangzahlen 100-104](#)