Zeitschrift für experimentelle und angewandte Psychologie
1980, Band XXVII, Heft 3, S. 460—467

Eine exakte nonparametrische Prüfung auf Kovariation zweier autokorrelierter Zeitreihen

Hans-Peter Krüger und Gustav A. Lienert

Fachbereich Erziehungs- und Kulturwissenschaften der Universität Erlangen-Nürnberg

Häufig tritt in der psychologischen und psychiatrischen Forschung die Frage nach dem Zusammenhang zwischen zwei oder mehreren Zeitreihen auf. Übliche Kovariationsmaße versagen dabei vor allem deshalb, weil sie nicht inferentiell ausgewertet werden können. An einem Beispiel aus der Pharmakopsychologie wird ein auf Pfanzagl (1963) zurückgehendes Verfahren vorgeschlagen, das es erlaubt, voraussetzungsfrei den Zusammenhang zwischen autokorrelierten Meßwertreihen zu prüfen. Als Unabhängigkeit zweier Zeitreihen wird definiert, daß zwischen ihnen bei Erhaltung der Autokorrelation kein höherer Zusammenhang besteht als nach dem Zufall zu erwarten.

Problemstellung

Häufig werden in der psychologischen und psychiatrischen Forschung wiederholte Messungen zu festen oder variierenden Zeitintervallen vorgenommen. Üblicherweise wird unterstellt, daß die einzelnen Beobachtungen unabhängig sind, um die Daten z.B. über eine Profilanalyse (siehe Cronbach & Gleser, 1953) auswerten zu können. Diese Annahme wird besonders dann fraglich, wenn die Zeitintervalle sehr kurz gewählt werden, da hier grundsätzlich mit autokorrelierten Meßwertreihen gerechnet werden muß. Will man bei der Überprüfung der Kovarianz zweier Zeitreihen nicht die Voraussetzungen einer parametrischen Zeitreihenanalyse (siehe z.B. Revenstorff, 1979) in Kauf nehmen, bietet sich eine Auswertung nach einem von Pfanzagl (1963) vorgeschlagenen Rationale an, das völlig verteilungsfrei arbeitet.

Das Rationale an einem Beispiel

Bei einem Pbn wurde nach Verabreichung eines Tranquilizers in $t = 15$ Halbstundenintervallen jeweils die Zahl der gesprochenen Worte

(X_t) und die mittlere Herzfrequenz (Y_t) erhoben. Die Meßwerte wurden Stanine-transformiert. Die daraus resultierenden x_t- und y_t-Werte sind der Tabelle 1 zu entnehmen.

Tabelle 1

Stanine-Werte der Worthäufigkeit und Herzfrequenz über $t = 15$ Zeitabschnitte, Signum ($0 \triangleq$ „—", $1 \triangleq$ „+") der ersten Differenzen der Meßwerte und das Kreuzprodukt T

t	Worte x_t	Herz y_t	$d_t = \operatorname{sgn}(x_{t+1} - x_t)$	$e_t = \operatorname{sgn}(y_{t+1} - y_t)$	$T_t = d_t \cdot e_t$
1	3	1			
2	5	4	1	1	1
3	6	7	1	1	1
4	7	8	1	1	1
5	5	5	0	0	0
6	4	2	0	0	0
7	1	3	0	1	0
8	2	4	1	1	1
9	3	5	1	1	1
10	4	6	1	1	1
11	6	7	1	1	1
12	9	9	1	1	1
13	5	6	0	0	0
14	7	4	1	0	0
15	8	3	1	0	0
					$\Sigma T_t = 8$

Im ersten Schritt werden die t-1 Vorzeichen der ersten Differenzen in den beiden Meßreihen nach folgender Regel bestimmt:

$$d_t = \operatorname{sgn}(x_{t+1} - x_t) \rightarrow \text{„0" für „—", „1" für „+"}$$
$$e_t = \operatorname{sgn}(y_{t+1} - y_t)$$

Treten in der Folge der Meßwerte gleiche Beobachtungen auf, resultieren bei der Bildung der ersten Differenzen Bindungen (ties). Für diesen Fall sind d_t und e_t nicht definiert. Konservativ entscheidet man, wenn die ties so aufgelöst werden, daß die Zahl der Nullen und Einsen möglichst gleich wird.

Die d_t und e_t sind Spalte 4 und 5 der Tabelle 1 zu entnehmen[1]). Als

1) Um dem Leser den Vergleich mit der Veröffentlichung von Pfanzagl zu ermöglichen, wurde das Beispiel so gewählt, daß die gleiche Nullverteilung von T resultiert.

Maß für die Kovariation beider Zeitreihen wird die Teststatistik T eingeführt, die wie folgt definiert ist:

$$T = \sum_{k=1}^{t-1} d_k \cdot e_k$$

T stellt damit das Kreuzprodukt beider Meßwertreihen dar. Zur inferentiellen Prüfung dieser Teststatistik muß die Nullverteilung (= T-Werte bei Gültigkeit von H_0, daß keine Kovariation vorliegt) *unter der Bedingung* erstellt werden, daß die Autokorrelation beider Meßwertreihen erhalten bleibt. Diese Autokorrelation drückt sich bei Alternativdaten (wie sie bei der vorliegenden Prüfung durch die Signum-Vorschrift entstehen) in der Sequenz von Einsen und Nullen aus (Iterationen). Die Nullverteilung muß dem Rechnung tragen, indem sie weder die Anzahl noch die Länge der Iterationen verändert, sondern lediglich vertauscht. Das Vorgehen ist deshalb wie folgt:

1. Stelle aus d_t alle Permutationen her, für die Zahl und Länge der Iterationen gleich sind wie bei der Ausgangsreihe. Diese i Permutationen ergeben sich am Beispiel daraus, daß in

$$d_t\colon\ 1\quad 1\quad 1\quad 0\quad 0\quad 0\quad 1\quad 1\quad 1\quad 1\quad 1\quad 0\quad 1\quad 1$$

$i_1 = 3$ Iterationen von Einsen und $i_0 = 2$ Iterationen von Nullen vorliegen. Diese können zu $i_1!\ i_0! = 12$ Möglichkeiten permutiert werden ($i = i_1!\ i_0!$).

Anschaulich wird das, wenn man die Einseriterationen mit Großbuchstaben, die Nulleriterationen mit Kleinbuchstaben bezeichnet. Im Beispiel heißen die Einseriterationen dann A, B, C, die Nulleriterationen a, b. Folgende Permutationen sind dann möglich:

1. Permutation von A, B, C	A		B		C		
1. Permutation von a, b		a		b		→	AaBbC
2. Permutation von a, b		b		a		→	AbBaC
2. Permutation von A, B, C	A		C		B		
1. Permutation von a, b		a		b		→	AaCbB
2. Permutation von a, b		b		a		→	AbCaB
3. Permutation von A, B, C	B		A		C		
usf.							

Im linken oberen Teil der Tabelle 2 sind alle $3!\ 2! = 12$ Permutationen von d_t zu finden.

2. Stelle nach dem gleichen Verfahren wie unter 1. aus e_t alle Permutationen her. Diese j Permutationen ergeben sich analog zu $j = j_1!\ j_0!$. Die Beispielreihe besitzt $j_1 = 2$ Einseriterationen und $j_0 = 2$ Nulleriterationen,

Tabelle 2

Permutationen von d_t und e_t, Berechnung der Prüfgröße T für alle Kombinationen und Erstellen der Nullverteilung

i	Permutationen von $d_t(i)$	$d_t(i) \times e_t(1)$	$d_t(i) \times e_t(2)$	$d_t(i) \times e_t(3)$	$d_t(i) \times e_t(4)$
$d_t(1)$	1 1 1 0 0 0 1 1 1 1 1 0 1 1	8	8	6	5
$d_t(2)$	1 1 1 0 1 1 1 1 1 0 0 0 1 1	7	6	6	5
$d_t(3)$	1 1 1 0 0 0 1 1 0 1 1 1 1 1	7	8	5	6
$d_t(4)$	1 1 1 0 1 1 0 0 0 1 1 1 1 1	6	6	7	8
$d_t(5)$	1 1 0 0 0 1 1 1 0 1 1 1 1 1	7	7	5	6
$d_t(6)$	1 1 0 1 1 1 0 0 0 1 1 1 1 1	5	5	7	8
$d_t(7)$	1 1 0 0 0 1 1 1 0 1 1 1 1 1	7	7	5	6
$d_t(8)$	1 1 0 1 1 1 1 1 0 0 0 1 1 1	5	5	5	6
$d_t(9)$	1 1 1 1 1 0 0 0 1 1 0 1 1 1	5	6	7	7
$d_t(10)$	1 1 1 1 1 0 1 1 0 0 0 1 1 1	5	6	5	6
$d_t(11)$	1 1 1 1 1 0 0 0 1 1 1 0 1 1	6	6	8	7
$d_t(12)$	1 1 1 1 1 0 1 1 1 0 0 0 1 1	6	6	6	5

j	Permutationen von $e_t(j)$
$e_t(1)$	1 1 1 0 0 1 1 1 1 1 1 0 0 0
$e_t(2)$	1 1 1 0 0 0 1 1 1 1 1 1 0 0
$e_t(3)$	1 1 1 1 1 1 0 0 1 1 1 0 0 0
$e_t(4)$	1 1 1 1 1 1 0 0 0 1 1 1 0 0

Nullverteilung von $T = d_t(i) \times e_t(j)$

T	5	6	7	8	
f(T)	14	17	11	6	$\Sigma f(T) = 48$

so daß $j = 2!\,2! = 4$ Möglichkeiten resultieren. Diese Permutationen sind im mittleren Teil der Tabelle 2 zu finden.

3. Berechne für jede Kombination von $d_t(i)$ und $e_t(j)$ die Teststatistik T_{ij}. Insgesamt resultieren daraus $j \cdot j$ T-Werte (im Beispiel $12 \cdot 4 = 48$). In Tabelle 2, rechter oberer Teil sind diese T-Werte aufgeführt.

4. Bilde die Nullverteilung von T, in der die Häufigkeit der möglichen T's verzeichnet sind. Dieser Schritt wurde im unteren Teil der Tabelle 2 durchgeführt. Wie man sieht, stellt das im Beispiel beobachtete $T = 8$ zwar einen extremen Wert dar, aber immerhin sechs der möglichen T's sind von gleicher Höhe.

Alle $i \cdot j = 48$ T-Werte sind unter H_0 gleichwahrscheinlich. In den Ablehnbereich von H_0 müssen all jene T-Werte, die gleich oder „extremer“ sind als der beobachtete $T°$-Wert. Bezeichnen wir mit f_T die Häufigkeit des Auftretens eines bestimmten T-Werts, ergibt sich die Überschreitungswahrscheinlichkeit P zu

$$P = \Sigma\, f_T \,/\, (i \cdot j)$$

Dabei ist diese Summation entsprechend der Alternativhypothese H_1 durchzuführen, und zwar:

a. unterstellt H_1 eine positive Kovariation, müssen alle f_T für $T \geq T°$ aufsummiert werden.

b. unterstellt H_1 eine negative Kovariation, müssen alle f_T für $T \leq T°$ aufsummiert werden.

c. ist H_1 zweiseitig angelegt, müssen alle T's in den Ablehnbereich, deren Überschreitungswahrscheinlichkeit kleiner oder gleich $\alpha/2$ ist, wenn von jedem Ende der Nullverteilung her aufsummiert wird.

In unserem Beispiel würde H_1 eine positive Kovariation zwischen Sprechhäufigkeit und Herzfrequenz unterstellen, so daß nach obiger Vorschrift a. summiert werden muß. Wie aus Tabelle 2 hervorgeht, haben $f_T = 6$ Zeitreihenpaare das gleiche $T = 8$, so daß sich die Überschreitungswahrscheinlichkeit P zu $P = 6/48 = 0{,}125$ ergibt. Wäre vorher ein Signifikanzniveau von $\alpha = 5\%$ vereinbart worden, müßte die Nullhypothese keiner Kovariation beider Zeitreihen beibehalten werden.

Die Autokorrelation beider Zeitreihen verbietet, ein übliches Korrelationsmaß (z.B. den Vierfelderkoeffizienten) auf die Daten anzuwenden. Um dennoch zu einem deskriptiven Zusammenhangsmaß zu kommen, wird (über Pfanzagl hinausgehend) vorgeschlagen, die erhaltene Überschreitungswahrscheinlichkeit nach der Logit-Funktion zu transformieren (siehe dazu Fisher & Yates, 1963, 16). Dabei gilt

$$z = \frac{1}{2} \ln\,(p/q)$$

Diese Transformation ist äquivalent zu der r-z-Transformation ($r = \tanh z$) mit $p = P$ (= Überschreitungswahrscheinlichkeit)

$$r = |2\,p - 1|$$

Die Betragsstriche deuten an, daß das Vorzeichen von r aus der Richtung der Kovariation der beiden Reihen bestimmt werden muß. Im Beispiel wird

$$r_T = 1 - 2(.125) = +.750$$

Durch die Regel, daß in den Ablehnbereich alle T kleiner *oder gleich* einbezogen werden, reagiert der Test bei kleinen Stichprobenumfängen stark konservativ. Pfanzagl schlägt vor, unter der Stetigkeitsannahme die Einfallsgruppe des tatsächlich beobachteten T°-Werts zu halbieren, so daß (für den Fall einer positiven Kovariation)

$$P_{T^\circ} = \left[\sum_{T > T^\circ} f_T + f_{T^\circ}/2\right] / (i \cdot j)$$

Im Beispiel resultiert dann, da $T^\circ = \max T$

$$P = (6/2)/48 = .0625$$

Diskussion

Das Procedere nach Pfanzagl führt für zwei Zeitreihen eine neue Definition der Unabhängigkeit ein: zwei Zeitreihen sollen dann als unabhängig gelten, wenn (bei Erhaltung der Autokorrelation) zwischen ihnen kein höherer Zusammenhang besteht als nach dem Zufall zu erwarten. Gegenüber der theoretisch nicht gerechtfertigten Anwendung üblicher Korrelationsmaße auf Zeitreihen bietet die Statistik T vor allem vier Vorteile:

a) Es werden alle Formen der Parallelität von Zeitreihen miteinbezogen, auch solche, gegenüber denen ein monotoner Korrelationskoeffizient versagt.

b) Die Nullverteilung kann explizit angegeben werden. Damit ist eine Signifikanzbestimmung möglich. Werden übliche Korrelationskoeffizienten verwendet, versagen die zugehörigen Signifikanztests.

Das wird unmittelbar evident, wenn man auf das Beispiel der Tabelle 1 den phi-Koeffizienten anwendet. Tabelle 3a gibt die zugehörige Vierfeldertafel wieder, aus der ein phi-Koeffizient von .519 resultiert. Wollte man die exakte Überschreitungswahrscheinlichkeit nach Fisher & Yates bestimmen, müßte als extremere Tafel auch die Vierfeldertafel der

Tabelle 3

Darstellung der Kovariation der beiden Zeitreihen in einer Vierfeldertafel (Tabelle 3a). Extremste, bei der Enumeration entstehende Vierfeldertafel mit a = T = 9 (Tabelle 3b)

3a

		e_t 1	0	Σ
d_t	1	8	2	10
	0	1	3	4
	Σ	9	5	14

3b

		e_t 1	0	Σ
d_t	1	9	1	10
	0	0	4	4
	Σ	9	5	14

Tabelle 3b in den Ablehnbereich einbezogen werden. Die Zellfrequenz a = 9 entspricht der Teststatistik T. Ein Blick in Tabelle 2 zeigt, daß der Wert T = 9 überhaupt nicht auftreten kann, *wenn die Autokorrelation der beiden Zeitreihen erhalten bleibt.*

c) Die Methode ist völlig verteilungsfrei in bezug auf die Verteilung der Daten und die Höhe der Autokorrelationen in den beiden Zeitreihen.

d) Da lediglich die ersten Differenzen verwendet werden, genügen bereits Meßwerte, die lediglich nach „verstärkt — vermindert", „mehr — weniger" usw. skaliert sind.

Der Nachteil des Verfahrens ist darin zu sehen, daß bislang keine Asymptotik existiert. Zwar zeigen die T-Werte relativ schnell eine gute Annäherung an die Normalverteilung, doch sind weder Erwartungswert noch Varianz der Verteilungen bekannt. Deshalb muß vorläufig die aufwendige Auswertung über den exakten Test in Kauf genommen werden.

Anwendungen für das Verfahren ergeben sich vor allem in der Einzelfallanalyse bei Fragestellungen wie:

1. Intraindividuelle Kovariation zweier wiederholt gemessener Variablen (analog dem Beispiel).
2. Intraindividuelle Kovariation einer Variablen, die über zwei Indikatoren gemessen wird (z.B. verbale und motorische Aktivität als Maß für die Inhibition bei Depressiven).
3. Paarweise intraindividuelle Kovariation von k Variablen in der Zeit (das entspricht der P-Technik in der Faktoranalyse von Ratingskalen).

Weitere Anwendungen sind auch beim Vergleich mehrerer (= N) Pbn denkbar:

4. Interindividuelle Kovariation einer mehrfach an N Pbn gemessenen Variablen (z.B. Schlafstörungen von N Ptn über k Nächte) als Analogon zur S-Technik der Faktoranalyse.
5. Intraindividuelle Kovariation zweier wiederholt an N Pbn gemessenen Variablen und Überprüfung der Homogenität der Kovariation (z.B. Kovariation von Sprechhäufigkeit und Herzfrequenz bei N Neurotikern).
6. Intraindividuelle Kovariation von k Variablen, die wiederholt an N Pbn gemessen wurden zur Überprüfung der Homogenität der Kovariationen.

Summary

The problem of examining the relation between two or more series of successive observations occurs often in the psychological and psychiatric research. The usual methods of measuring covariation cannot be employed here, particularly because they cannot be used to draw inferences. The present paper describes a new method to examine the relation between autocorrelated series of successive observations, without using any presuppositions. Two successive series are considered to be independent, when the relation between them is not higher than that expected by chance. The method described here is based upon a paper by Pfanzagl (1963), and is illustrated with the help of an example from the field of Pharmacopsychology.

Résumé

En recherche psychologique et psychiatrique, on pose fréquemment le problème de la liaison entre deux ou plusieurs séries temporelles. Les mesures usuelles de la covariation sont inadéquates, parce qu'elles ne permettent pas d'inférences. A l'aide d'un exemple tiré de la pharmakopsychologie, l'auteur propose une méthode fondée sur les travaux de Pfanzagl (1963) qui permet de tester la liaison entre des séries temporelles autocorrélées. Deux séries temporelles sont jugées indépendantes, lorsque, tout en maintenant l'autocorrélation, leur liaison ne varie pas plus qu'on ne pourrait s'y attendre par hasard.

Literatur

Cronbach, L. J. & Gleser, Goldine C.: Assessing similarity between profiles. Psychological Bulletin, **50,** 1953, 456—473.

Fisher, R. A. & Yates, F.: Statistical tables for biological, agricultural and medical research. 6th ed. Edinburgh: Oliver u. Boyd, 1963.

Pfanzagl, J.: Über die Parallelität von Zeitreihen. Metrika, **6,** 1963, 100—113.

Revenstorff, D.: Zeitreihenanalyse für klinische Daten. Weinheim: Beltz, 1979.

Anschrift der Verfasser:
Prof. Dr. Dr. G. A. Lienert
PD Dr. Hans-Peter Krüger
Lehrstuhl für Psychologie
Universität Erlangen–Nürnberg
Regensburger Straße 160
8500 Nürnberg