

12 Parametrische und nichtparametrische Methoden der Einzelfallstatistik

von C. Möbus, G. Göricke und P. Kröh

12.1 EINLEITUNG

Die Auswertung von zeitbezogenen Daten macht i. allg. die Anwendung spezieller statistischer Verfahren notwendig. Die Begründung hierfür liegt in der sog. "Abhängigkeit" der Daten: Man nimmt im Gegensatz zur Querschnittuntersuchung an, daß die Daten eine gewisse "Trägheit" oder ein "Beharrungsvermögen" aufweisen. Diesen Effekt kann man auf physiologische Prozesse oder psychologische Ursachen wie "Erinnerungs"- und "Gewöhnungseffekte" oder auch regelmäßig wiederkehrende äußere Einflüsse (wie z.B. regelmäßige Berufstätigkeit) zurückführen. Diese zeitliche Abhängigkeit der Daten wird auch Autokorreliertheit genannt. Die Verwendung "normaler" statistischer Verfahren zur Analyse autokorrelierter Daten liefert verzerrte Prüfgrößen, die zu falschen Interpretationen führen.

Abhilfe schaffen die von uns vorgestellten Methoden. Zur besonderen Veranschaulichung haben wir sie mit einem aus der Literatur bekannten Beispiel durchgerechnet. Dabei erweisen sich die verteilungsfreien Randomisierungstests für den Praktiker wegen des geringeren Aufwands als geeigneter. Die verteilungsgebundenen Verfahren setzen in jedem Fall den Zugang zu einem Computer voraus.

12.2 VERTEILUNGSFREIE VERFAHREN: RANDOMISIERUNGS- BZW. PERMUTATIONS-TESTS

Einfache Auswertungsverfahren, die speziell für Zeitreihenexperimente geeignet sind, bieten sich mit dem Randomisierungs- oder Permutationstest an (Edgington 1967, 1969a, b, 1971, 1973, 1975a, b, 1980). Dabei geht die prinzipielle Konstruktionsidee auf Fisher (1951) zurück. Während die Prüfgrößen z.T. völlig äquivalent zu denen der klassischen statistischen Verfahren (t- oder F-Bruch, aber auch einfache Mittelwert- oder Mediandifferenzen) sein können, weichen Nullhypothesenformulierung und Inferenzmodell doch wesentlich vom klassischen Verfahren ab.

In einem $N=1$ -Zeitreihenexperiment gliedert man den Versuchsplan in mehrere A- und B-Phasen (z.B. $A_1B_1A_2B_2$), wobei A_1 meistens eine prä-experimentelle Beobachtungsphase darstellt. Diese wird dann von einer Treatment- oder B-Phase abgelöst. Dann folgt wieder eine reine Beobachtungs- oder Lösungsphase A_2 usw. Die Nullhypothese H_0 ("Kein Unterschied zwischen A- und B-Phasen") wird geprüft im Vergleich mit einer bedingten Wahrscheinlichkeitsverteilung, die sich aus allen theoretisch möglichen Aufteilungen dieser gerade erhobenen Daten auf die Treatments A und B ergibt. Unter der Gültigkeit der Nullhypothese liegen keine Treatmenteffekte vor. Die Daten spiegeln nur Zufallsein-

flüsse wider. Damit sind die Aufteilungen der Daten gerechtfertigt. Die so bestimmten Mittelwertunterschiede weisen ebenfalls nur Zufallseinflüsse auf. Wir erhalten unter H_0 eine Verteilung mit Mittelwertunterschieden. Diese Verteilung spielt die Rolle einer "Prüfverteilung", an der wir die empirisch vorgefundene Mittelwertdifferenz $\bar{Y}_B - \bar{Y}_A$ prüfen. Statt des Mittelwertes können auch andere Zentralitätsmaße (wie Mediane, Proportionen etc.) Verwendung finden.

Die Reduktion der Daten zu Phasenmittelwerten (oder anderen Summationsmaßen) bringt, wie Levin et al. (1978) gezeigt haben, beträchtliche Vorteile bei der Analyse autokorrelierter Daten. Die Mittelwerte weisen eine stark reduzierte Autokorrelation gegenüber der Originalzeitreihe auf.

Die Nullhypothese H_0 wird verworfen, wenn es unwahrscheinlich ist ($p < 0,05$), daß man durch einfache Umbenennungen der Phasen eine gleichgroße oder größere Mittelwertdifferenz als die beobachtete erhält. H_0 wird beibehalten, wenn einfache Umbenennungen der experimentellen Phasen gleiche oder größere Unterschiede als die beobachteten häufig auftreten lassen. Somit beschäftigt sich Inferenz nicht mehr, wie bei klassischen Experimenten, mit der Generalisation über Personen und Situationen, sondern mit der Generalisation über andere A- vs. B-Zuweisungen. Lehman (1975) bezeichnete dieses Inferenzmodell als "Randomisierungsmodell", das sich logisch vom üblichen "Stichproben-Populations-Modell" abhebt. Es werden keine Populationsannahmen gemacht, da die beobachteten Daten die Rolle der Population übernehmen.

Zur Illustration betrachten wir das folgende Beispiel (vgl. Möbus u. Nagl 1983) für die Auswertung des häufig verwendeten $A_1B_1A_2B_2$ -Designs. Im Experiment wurden die in Tabelle 1 dargestellten Daten beobachtet.

Tabelle 1. Daten aus einem Experiment, dem ein $A_1B_1A_2B_2$ -Design zugrunde liegt. (Aus Möbus u. Nagl, 1983, S.289; Nachdruck mit Genehmigung des Verlags J.C.Hogrefe)

1.	2.	3.	4. Phase	
A_1	B_1	A_2	B_2	
2.0	7.0	3.0	8.0	Mittelwerte \bar{Y}_i
(7.5 - 2.5) = ($\bar{Y}_B - \bar{Y}_A$) = 5.0 Differenz				

Da das Experiment $n_A = 2$ A-Phasen und $n_B = 2$ B-Phasen, d.h. insgesamt $n_A + n_B = 4$ Phasen umfaßt, lassen sich insgesamt $(n_A + n_B)! / (n_A! \cdot n_B!) = 4! / (2! \cdot 2!) = 6$ Anordnungen (Permutationen) der A- und B-Phasen finden (Tabelle 2).

Da die Wahrscheinlichkeit, durch eine andere Aufteilung eine mindestens ebenso große Mittelwertdifferenz wie die beobachtete zu finden, größer als 0,05 ist (nämlich exakt 1/6), wird die Nullhypothese beibehalten. Da diese $p = 1/6 = 0,16$ nach den üblichen Standards zu groß ist, kann das $A_1B_1A_2B_2$ -Design nicht über Mittelwertvergleiche innerhalb des Randomisierungsinferenzmodells ausgewertet werden!

Tabelle 2. Mögliche Anordnungen der A- und B-Phasen des $A_1 B_1 A_2 B_2$ -Designs. (Aus Möbus u. Nagl, 1983, S.290; Nachdruck mit Genehmigung des Verlags J.C.Hogrefe)

Aufteilung	1.	2.	3.	4. Phase	\bar{Y}_A	\bar{Y}_B	$\bar{Y}_B - \bar{Y}_A$
1	A_1'	A_2'	B_1'	B_2'	4.5	5.5	1.0
2	A_1'	B_1'	A_2'	B_2'	2.5	7.5	5.0
3	A_1'	B_1'	B_2'	A_2'	5.0	5.0	0.0
4	B_1'	A_1'	A_2'	B_2'	5.0	5.0	0.0
5	B_1'	A_1'	B_2'	A_2'	7.5	2.5	-5.0
6	B_1'	B_2'	A_1'	A_2'	5.5	4.5	-1.0

Um ein kleineres p zu erhalten, müssen gewichtete Alternativhypothesen, wie z.B. $\bar{Y}_{A1} < \bar{Y}_{A2} < \bar{Y}_{B1} < \bar{Y}_{B2}$, geprüft werden. Hieraus leitet sich das ordinale Gewichtungsschema ab: Die \bar{Y}_t werden entsprechend der prognostizierten Größe mit ihren Rängen gewichtet: $g_{A1} = 1$, $g_{A2} = 2$, $g_{B1} = 3$, $g_{B2} = 4$. Mit diesen Gewichten werden die empirisch erhaltenen Maße (Mittelwerte, Mediane, Proportionen) multipliziert und aufsummiert. Damit ergibt sich für das obige Beispiel:

$$\Gamma_{\text{emp}} = \sum_{t=1}^4 g_t \bar{y}_t = 1 \cdot 2,0 + 2 \cdot 3,0 + 3 \cdot 7,0 + 4 \cdot 8,0 = 61,0. \quad (1)$$

Es gibt insgesamt $4! = 24$ Möglichkeiten, für 4 Phasen eine gewichtete Alternativhypothese aufzustellen. Die Verteilung dieser gewichteten Summen Γ_k ist in Tabelle 3 dargestellt.

Da die beobachtete Reihe von Mittelwerten in einer ordinal gewichteten Summe Γ von 61 resultiert, deren Auftretenswahrscheinlichkeit unter der Nullhypothese $1/24 = 0,042$ ist (s. auch rechte Spalte von Tabelle 3), wird H_0 zugunsten von H_1 verworfen: Die Phasen unterscheiden sich signifikant.

Levin et al. (1978) haben gezeigt, daß die durch ein ordinale Gewichtungsschema bestimmte Summe Γ_k annähernd normalverteilt ist, wenn der Versuchsplan mindestens 4 A- und 4 B-Phasen umfaßt. Somit läßt sich Γ_{emp} mit Hilfe der folgenden Prüfstatistik auf Signifikanz überprüfen:

$$z_{\text{emp}} = \frac{\Gamma_{\text{emp}} - E(\Gamma)}{\sqrt{\text{Var}(\Gamma)}} \quad (2)$$

mit

$$E(\Gamma) \approx \frac{1}{n} \cdot \left\{ \left(\sum_{t=1}^n g_t \right) \left(\sum_{t=1}^n \bar{y}_t \right) \right\} \quad (2a)$$

und

$$\text{Var}(\Gamma) \cong \frac{1}{n-1} \cdot \left\{ \sum_{t=1}^n (g_t - \bar{g})^2 \right\} \cdot \left\{ \sum_{t=1}^n (\bar{y}_t - \bar{y})^2 \right\}. \quad (2b)$$

Statt eines ordinalen Gewichtungsschemas ist auch ein Intervall- oder Verhältnisschema (nichtordinale Gewichtungsschemata) denkbar (vgl. Möbus u. Nagl 1983, S. 291 ff.).

Tabelle 3. Verteilung der gewichteten Summen Γ_k unter der Nullhypothese in einem $A_1 B_1 A_2 B_2$ -Design. (Aus Möbus u. Nagl, 1983, S.291; Nachdruck mit Genehmigung des Verlags J.C.Hogrefe)

Permutation	1.	2.	3.	4. Experimentelle Phase	Γ
K	2.0	7.0	3.0	8.0 Mittelwerte	
1	A ₁ '	A ₂ '	B ₁ '	B ₂ '	57
2	A ₁ '	A ₂ '	B ₂ '	B ₁ '	52
3	A ₁ '	B ₁ '	A ₂ '	B ₂ '	61
4	A ₁ '	B ₁ '	B ₂ '	A ₂ '	51
5	A ₁ '	B ₂ '	A ₂ '	B ₁ '	60
6	A ₁ '	B ₂ '	B ₁ '	A ₂ '	55
7	A ₂ '	A ₁ '	B ₁ '	B ₂ '	52
8	A ₂ '	A ₁ '	B ₂ '	B ₁ '	47
9	A ₂ '	B ₁ '	A ₁ '	B ₂ '	60
10	A ₂ '	B ₁ '	B ₂ '	A ₁ '	45
11	A ₂ '	B ₂ '	A ₁ '	B ₁ '	59
12	A ₂ '	B ₂ '	B ₁ '	A ₁ '	49
13	B ₁ '	A ₁ '	A ₂ '	B ₂ '	51
14	B ₁ '	A ₁ '	B ₂ '	A ₂ '	41
15	B ₁ '	A ₂ '	A ₁ '	B ₂ '	55
16	B ₁ '	A ₂ '	B ₂ '	A ₁ '	40
17	B ₁ '	B ₂ '	A ₁ '	A ₂ '	53
18	B ₁ '	B ₂ '	A ₂ '	A ₁ '	48
19	B ₂ '	A ₁ '	A ₂ '	B ₁ '	45
20	B ₂ '	A ₁ '	B ₁ '	A ₂ '	40
21	B ₂ '	A ₂ '	A ₁ '	B ₁ '	49
22	B ₂ '	A ₂ '	B ₁ '	A ₁ '	39
23	B ₂ '	B ₁ '	A ₁ '	A ₂ '	48
24	B ₂ '	B ₁ '	A ₂ '	A ₁ '	43

12.3 VERTEILUNGSGBUNDENE VERFAHREN: DAS TRANSFERMODELL VON BOX & TIAO

Bei der Überprüfung eines Interventionseffektes in Zeitreihen mit Hilfe klassischer statistischer Verfahren besteht ein grundsätzliches Hindernis in der Abhängigkeit der Daten bzw. der Residuen, so daß sich z.B. eine varianzanalytische Auswertung verbietet. Ein Verfahren, das die Abhängigkeit der seriell anfallenden Beobachtungen berücksichtigt, ist das Transfermodell von Box und Tiao. Hierbei werden die Zeitreihendaten in einen Interventionseffekt und einen nichtkontrollierbaren Effekt N_t zerlegt (s. auch Gl. (16)). Während der In-

terventionseffekt deterministisch ist, bildet N_t wiederum eine Zeitreihe, die seriell abhängig ist. Der nichtkontrollierbare Prozeß N_t wird nach dem Vorschlag von Box und Jenkins (1976) mit einem ARIMA-Prozeß so dargestellt, daß die Fehler a_t des ARIMA-Prozesses jetzt im Gegensatz zu den N_t voneinander unabhängig sind. Erst die unabhängigen a_t lassen den Einsatz hypothesenprüfender Verfahren zu.

12.3.1 Die Modellierung des Fehlerprozesses N_t

Die ARIMA-Modelle setzen sich aus a) dem autoregressiven Prozeß (AR-Prozeß), b) dem integrierenden Prozeß (I-Prozeß) und c) dem Moving Average-Prozeß (MA-Prozeß) zusammen.

Der ARIMA(p,d,q)-Prozeß läßt sich in Rohwerten Y_t (Hibbs 1977, S. 140) darstellen als

$$(1 - \phi_1 B - \dots - \phi_p B^p)(1 - B)^d Y_t = \theta_0 + (1 - \theta_1 B - \dots - \theta_q B^q) a_t \quad (3a)$$

oder

$$\phi(B)(1 - B)^d Y_t = \theta_0 + \theta(B) a_t \quad (3b)$$

bzw.

$$Y_t = \frac{\theta_0 + \theta(B) a_t}{\phi(B)(1 - B)^d}, \quad (3c)$$

wobei B der Backshiftoperator ist: $B^k Y_t = Y_{t-k}$. Alternative Formulierungen für das allgemeine ARIMA-Modell findet man u.a. bei Box u. Jenkins (1976), Glass et al. (1975), Jenkins (1979) und Gottman (1981). Der Zufallsprozeß a_t in Gl. (3) ist so definiert, daß er unabhängig und normalverteilt ist mit

$$E(a_t) = 0 \quad \text{und} \quad \text{Var}(a_t) = \sigma_a^2. \quad (4)$$

Ein solcher Zufallsprozeß wird von einigen Autoren mit "white noise" (weißem Rauschen) und von anderen als Innovationsprozeß bezeichnet.

In Gl. (3) ist $\phi(B) = 1 - \phi_1 B - \dots - \phi_p B^p$ ein stationärer AR-Operator der Ordnung p und $\theta(B) = 1 - \theta_1 B - \dots - \theta_q B^q$ ein invertierbarer MA-Operator der Ordnung q . Eine notwendige und hinreichende Bedingung für die Stationarität des AR-Operators bzw. der Invertierbarkeit des MA-Operators ist, daß alle Lösungen der charakteristischen Gleichung

$$\phi(\lambda) = 1 - \phi_1\lambda - \dots - \phi_p\lambda^p = 0$$

bzw.

$$\theta(\lambda) = 1 - \theta_1\lambda - \dots - \theta_q\lambda^q = 0$$

in ihrem Absolutbetrag größer als eins sind. Nur wenn die Stationaritätsbedingung (Invertierbarkeitsbedingung) erfüllt ist, läßt sich der ARIMA(p,d,q)-Prozeß als unendlicher MA-(AR-)Prozeß formulieren. Der Hauptgrund für die Kombination der 3 Prozesse in einem Modell liegt im Bestreben, mit möglichst wenigen Parametern möglichst komplizierte Zeitreihen beschreiben zu können (vgl. Möbus und Nagl 1983).

Zur Illustration betrachten wir den unendlichen MA-Prozeß

$$Y_t = \frac{\theta_0}{1 - 0,5B} + a_t + a_{t-1} + 0,5a_{t-2} + 0,25a_{t-3} + \dots \quad (5a)$$

bzw.

$$Y_t = \frac{\theta_0}{1 - 0,5B} + \frac{(1 + 0,5B)}{(1 - 0,5B)} \cdot a_t \quad (5b)$$

Der Prozeß Gl. (5) ist aber dem ARIMA(1,0,1)-Prozeß mit $\phi_1 = 0,5$ und $\theta_1 = -0,5$ äquivalent:

$$(1 - 0,5B)Y_t = \theta_0 + (1 + 0,5B)a_t \quad (6)$$

Bei der Formulierung eines gemischten Prozesses ($p \neq 0$ und $q \neq 0$) sollte man die Gefahr der Überparametrisierung einkalkulieren (Box u. Jenkins 1976, Kap. 7.3.5).

12.3.1.1 Modellidentifikation:

Unter Identifikation von ARIMA-Modellen versteht man die Festlegung der Prozeßparameter p , d und q vor der Parameterschätzung. Dabei stützt sich die Identifikation auf die Autokorrelationsfunktion (PACF). Die ACF ist definiert als:

$$\rho(k) = \frac{\gamma(k)}{\gamma(0)}, \quad k = 0, 1, 2, \dots \quad (7)$$

$$\text{mit } \gamma(k) = E [(Y_t - \mu_y)(Y_{t+k} - \mu_y)].$$

Die Autokorrelation mißt die Größe des Zusammenhangs der Zeitreihe mit der um k Zeitpunkte verschobenen Zeitreihe. Ein gebräuchlicher Schätzer für die ACF ist:

$$\text{ACF}(k) = r_k = \frac{C(k)}{C(0)}, \quad 0 \leq k \leq T-1 \quad (8)$$

mit der Kovarianz

$$C(k) = \frac{1}{T^*} \cdot \sum_{t=1}^{T-k} (Y_t - \bar{Y})(Y_{t+k} - \bar{Y}) \quad (9)$$

mit $T^* = T$ (Box u. Jenkins 1976)
bzw. $T^* = T - k$ (McCleary u. Hay 1980).

Mit zunehmend größerem Lag k wird dabei die Schätzung unsicherer, da die Zahl der Kreuzprodukte in Gl. (9) mit wachsendem k ständig abnimmt.

Die PACF läßt sich am besten durch eine Analogie zur schrittweisen multiplen Regression erklären, wobei der partielle Autokorrelationskoeffizient jeweils dem letzten Gewicht der Regressionsgleichung entspricht. Unter der Annahme der Stationarität ($E(Y_t) = \mu_t$ für alle t) folgt:

$$y_t = \phi_{11}y_{t-1} + a_t$$

$$y_t = \phi_{21}y_{t-1} + \phi_{22}y_{t-2} + a_t$$

$$y_t = \phi_{31}y_{t-1} + \phi_{32}y_{t-2} + \phi_{33}y_{t-3} + a_t$$

...

$$\text{mit } y_t = Y_t - \mu_y. \quad (10)$$

ϕ_{kk} spiegelt den direkten Einfluß des Zeitpunktes $t-k$ auf t wider. Die Folge $\phi_{11}, \phi_{22}, \phi_{33}, \dots$ stellt die PACF dar. Die Schätzung der ϕ_{kk} erfolgt über die Yule-Walker-Gleichungen (Box u. Jenkins 1976; Nelson 1973):

$$\begin{bmatrix} \phi_{k1} \\ \phi_{k2} \\ \vdots \\ \phi_{kk} \end{bmatrix} = \begin{bmatrix} 1 & r_1 & r_2 & \cdots & r_{k-1} \\ r_1 & 1 & r_1 & \cdots & r_{k-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ r_{k-1} & r_{k-2} & r_{k-3} & \cdots & 1 \end{bmatrix}^{-1} \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_k \end{bmatrix} \quad (11)$$

Diese Gleichungen besitzen eine Parallele bei der Bestimmung von Regressionskoeffizienten in der multiplen Regression (s. auch Cooley u. Lohnes 1971, S. 53).

Revenstorff u. Keeser (1979, S. 195) geben die in Abb. 1 dargestellte Zusammenfassung für die Identifikation von ARIMA-Modellen.

Zunächst wird die ACF inspiziert, um d festzulegen. In der Literatur (Box u. Jenkins 1976; Glass et al. 1975) wird als Kriterium für Nichtstationarität im Level angegeben, daß die ACF einen besonders flachen Verlauf nimmt. Für die Bestimmung von p und q werden dann die ACF und PACF der d -fach differenzierten Zeitreihe untersucht. Die ACF eines ARIMA($p, d, 0$)-Prozesses setzt sich u.a. aus einem gedämpften exponentiellen Abfall und/oder gedämpften Sinusschwingungen zusammen, während für die theoretische PACF $\phi_{kk} \neq 0$ für $k \leq p$ und $\phi_{kk} = 0$ für $k > p$ gilt.

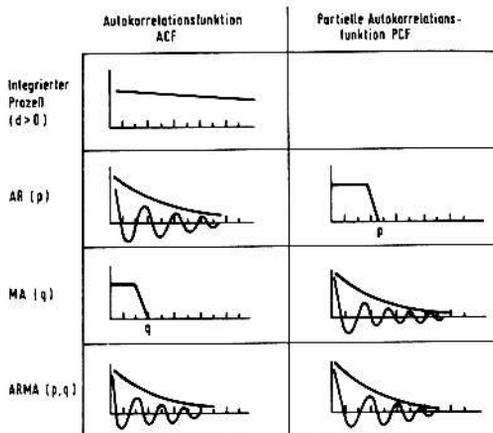


Abb. 1. Grobschema zur Identifikation von ARIMA-Prozessen nach Box u. Jenkins. (Aus Revenstorff u. Keeser 1979, S. 195; Nachdruck mit Genehmigung des Verlags Urban & Schwarzenberg)

Die PACF eines ARIMA($0, d, q$)-Prozesses verhält sich wie die ACF beim ARIMA($p, d, 0$)-Prozeß, während für die ACF $ACF(k) \neq 0$ für $k \leq q$ und $ACF(k) = 0$ für $k > q$ gilt.

Bei gemischten Prozessen ($p \neq 0$ und $q \neq 0$) haben sowohl die ACF als auch die PACF eine unendliche Ausdehnung.

Da die ACF und PACF mit Daten geschätzt werden, versucht man, die empirisch erhaltenen Muster mit den theoretischen zu vergleichen, um so die Grade p, d, q des ARIMA-Modells festzulegen. Daher gehen in die Identifikationsphase stark subjektive Momente ein (Chatfield 1975, S. 25). Einen Hinweis, inwieweit sich die empirische ACF und PACF von den theoretischen Funktionen unterscheiden, geben die Verteilungen der r_k und ϕ_{kk} . Nach Bartlett (1946) gilt:

$$\hat{\sigma} [r_k] = \frac{1}{\sqrt{T}} \cdot \left\{ 1 + 2 \cdot \sum_{i=1}^q r_i^2 \right\}^{1/2}, \quad k > q, \quad (12)$$

sowie nach Quenouille (1949):

$$\hat{\sigma} [\phi_{kk}] = \frac{1}{\sqrt{T}} \quad \text{für } k > p. \quad (13)$$

Mit Gl. (12) und (13) lassen sich Konfidenzintervalle konstruieren. Gilt $|r_k| > 2 \cdot \hat{\sigma} [r_k]$ bzw. $|\phi_{kk}| > 2 \cdot \hat{\sigma} [\phi_{kk}]$, so sind die r_k für $k > q$ bzw. die ϕ_{kk} für $k > p$ auf dem 5%-Niveau signifikant verschieden von Null.

12.3.1.2 Parameterschätzung:

Der vorläufigen Identifikation schließt sich die Schätzung der Parameter an. Für die Bestimmung der Maximumlikelihood-Schätzer gibt es u.a.:

1. die "Conditional-least-squares-(CLS-)Methode", die auf Aström und Bohlin (1966) zurückgeht und von Anderson (1975, 1978) weiter untersucht wurde;
2. die "Unconditional-least-squares-Methode", auch "Backforecasting-Methode" genannt (Box u. Jenkins 1976).

Beide Methoden stellt z.B. das Programmpaket BMDP für die Parameterschätzung zur Verfügung. Die geschätzten Parameter müssen innerhalb der Stationaritäts- und Invertierbarkeitsgrenzen liegen. In einem überparametrisierten Modell müssen die überflüssigen Parameter nicht signifikant sein.

12.3.1.5 Modelldiagnostik:

Nach der Identifikation und der Parameterschätzung wird die Adäquatheit des stochastischen Modells getestet. Für die "Richtigkeit" eines Modells spricht v.a., daß die Residuen a_t des Modells einem White-noise-Prozeß folgen. Um zu überprüfen, ob die gesamte ACF nicht signifikant von der ACF eines White-noise-Prozesses abweicht, wurde von Box und Pierce (1970) der sog. Portemanteau-Test vorgelegt:

$$Q = (T - d) \cdot \sum_{i=1}^k r_i^2 \sim \chi^2_{df=k-p-q}. \quad (14)$$

Ist Q signifikant, muß die Nullhypothese verworfen werden. Verschiedene Autoren empfehlen, für k den Wert 20-30 zu wählen, da Q sehr sensitiv von k abhängt (s. auch Box u. Jenkins 1976, S. 291).

Weitere Prüfungen des vorläufigen Modells schließen die Berechnung des multiplen Wertes R^2 als Maß für die Datenanpassung ein:

$$R^2 = 1 - \sum_{t=1}^T \left(\frac{a_t^2}{y_t^2} \right) \quad (15)$$

$$\text{mit } y_t = \left\{ \begin{array}{l} Y_t - \frac{\theta_0}{1 - \phi_1 - \dots - \phi_p} \\ (1 - B)^d Y_t - \frac{\theta_0}{1 - \phi_1 - \dots - \phi_p} \end{array} \right.$$

Weitere praktische Hinweise zur Anpassung von ARIMA(p,d,q)-Modellen finden sich bei Glass et al. (1975), Anderson (1975), Gottman u. Glass (1978), Makridakis u. Wheelwright (1978a, b), Revenstorf und Keeser (1979), Revenstorf (1979), McCain u. McCleary (1979) und Gottman (1981).

12.3.2 Die Modellierung des Interventionsprozesses

Während der Ansatz von Box und Tiao (1965) und Glass et al. (1975) stark am allgemeinen linearen Modell orientiert ist, haben Box und Tiao (1975) die Interventionskomponente konsequent in das Zeitreihenmodell integriert. Einen Vergleich sowie Querverbindungen zwischen den Modellansätzen von Glass et al. (1975) und Box und Tiao (1975) geben Möbus et al. (1983). In dieser Arbeit wird auch gezeigt, daß sich der Modellansatz von Glass et al. (1975) aus dem Box-Tiao-Modell herleiten läßt.

Das Interventionsmodell von Box und Tiao (1975) kennt 3 Variablen-sätze:

1. die Interventionsvariable I_t , die zu bestimmten a priori festgelegten Zeitpunkten die Werte 0 oder 1 annimmt, je nachdem ob ein Interventionseffekt "ein-" oder "ausgeschaltet" wird;
2. die Effektvariable Y_t^* , die eine Folge dieser Intervention ist, $Y_t^* = f(I_t)$, und
3. die beobachtbare Zeitreihe Y_t , die einerseits durch die Effektvariable Y_t^* und andererseits durch anderweitige, nichtkontrollierte, Einflüsse N_t geprägt wird:

$$Y_t = \left\{ \begin{array}{l} \text{Interventions-} \\ \text{effekte} \end{array} \right\} + \left\{ \begin{array}{l} \text{nichtkontrollierte} \\ \text{Effekte} \end{array} \right\}. \quad (16)$$

$$Y_t^* \qquad \qquad \qquad N_t$$

Die nichtkontrollierten Einflüsse N_t , die die Interventionseffekte überlagern, lassen sich durch ein ARIMA-Modell repräsentieren. Box und Tiao (1975) geben die folgende formale Darstellung von Gl. (16):

$$Y_t = \underbrace{\frac{\Omega(B)}{\delta(B)}}_{Y_t^*} \cdot I_t + \underbrace{\frac{\theta_0 + \theta(B)a_t}{(1-B)^d \phi(B)}}_{N_t} \quad (17)$$

Aus Gl. (17) folgt:

$$\delta(B)Y_t^* = \Omega(B)I_t \quad (18a)$$

bzw. mit $\Omega(B) = \omega(B)B^b$

$$\delta(B)Y_t^* = \omega(B)I_{t-b} \quad (18b)$$

bzw. mit Hilfe des Backshiftoperators

$$(1 - \delta_1 B - \dots - \delta_r B^r)Y_t^* = (\omega_0 - \omega_1 B - \dots - \omega_s B^s)I_{t-b} \quad (18c)$$

bzw.

$$Y_t^* = \delta_1 Y_{t-1}^* + \dots + \delta_r Y_{t-r}^* + \omega_0 I_{t-b} - \dots - \omega_s I_{t-b-s} \quad (18d)$$

Dabei spiegeln die Indizes r und s das "Gedächtnis" der Interventionskomponenten wider. Der Index b gibt an, daß die Intervention sich mit einer Zeitverzögerung von b Zeiteinheiten auswirkt.

Das Transfermodell Gl. (18) ist stabil (d.h. explodiert nicht), wenn für die Lösungen der charakteristischen Gleichung

$$1 - \delta_1 \lambda - \dots - \delta_r \lambda^r = 0 \quad (19)$$

$$|\lambda_i| > 1 \text{ für } i = 1, 2, \dots, r \text{ gilt.}$$

Die Stabilitätsgrenzen entsprechen numerisch den Grenzen der Stationarität von autoregressiven Prozessen.

In der Literatur unterscheidet man bei der Indikatorvariablen I_t zwischen einem Stepinput:

$$I_t = \begin{cases} 0 & \text{für } t < t_1 \\ 1 & \text{für } t \geq t_1 \end{cases} \quad (20)$$

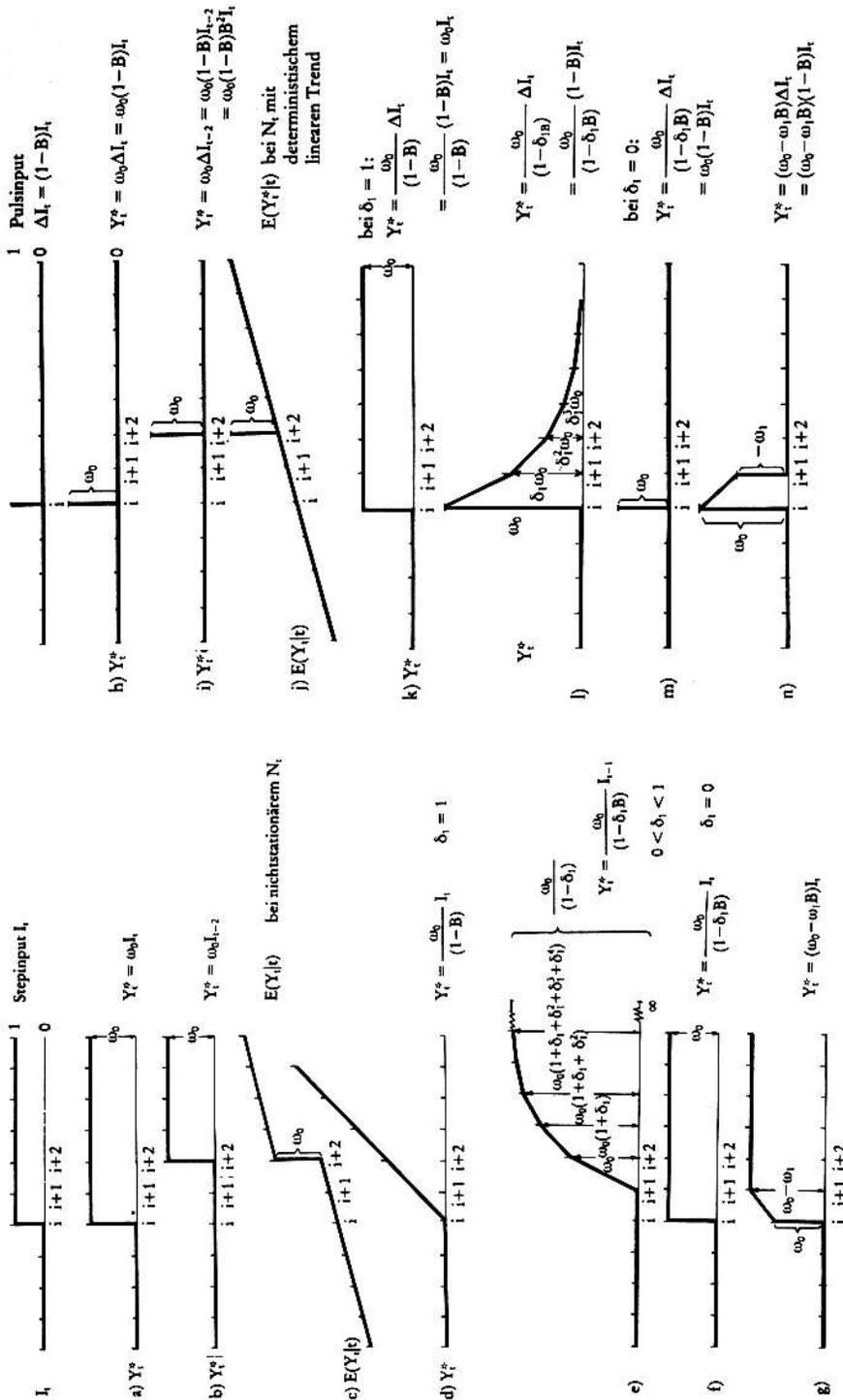


Abb. 2. Anregungen für die Formulierung einiger grundlegender Interventionseffekte Y_t mit Hilfe des Stepinputs (Funktionen a - g) bzw. des Pulsinputs (Funktionen h - n). (Aus Möbus u. Nagl 1983, S. 306-307; Nachdruck mit Genehmigung des Verlags J.C.Hogrefe)

und einem Pulsinput:

$$\Delta I_t = (1 - B)^d I_t = \begin{cases} 0 & \text{für } t \neq t_1 \\ 1 & \text{für } t = t_1 \end{cases} \quad (21)$$

In Gl. (20) und (21) bezeichnet t_1 den ersten Meßzeitpunkt nach Durchführung der Intervention.

Die Abb. 2 zeigt Anregungen für die Postulierung einiger grundlegender Interventionshypothesen. Weitere Erläuterungen hierzu finden sich bei Möbus und Nagl (1983, S. 305 ff.). Die Abb. 3 zeigt eine Abfolge zur Prüfung von einfachen Interventionshypothesen.

Zunächst gilt es, ein adäquates ARIMA-Modell für den Noiseprozeß zu bestimmen. Die Autoren Box u. Tiao (1975) und Hibbs (1977) schlagen hierzu vor, entweder die gesamte Zeitreihe oder nur die entsprechenden Beobachtungen vor der Intervention zu verwenden.

Nach Cook u. Campbell (1979) wird danach das formulierte Interventionsmodell zum identifizierten Noisemodell addiert und alle Parameter des Transfer-ARIMA-Modells Gl. (17) geschätzt. Schließlich wird dieses Modell analog der Modelldiagnostik in Abschnitt 12.3.1.3 auf Adäquatheit geprüft sowie ein Signifikanztest für die Interventionsparameter durchgeführt.

Liegen k Interventionen vor, so lautet das allgemeine Transfermodell von Box und Tiao (1975):

$$Y_t = \sum_{j=1}^k Y_{jt}^* + N_t = \sum_{j=1}^k \frac{\omega_j(B)}{\delta_j(B)} I_{jt} + N_t \quad (22)$$

12.4 EIN EMPIRISCHES BEISPIEL

Zu Illustrationszwecken betrachten wir die Zeitreihe für die täglichen Messungen der Wahrnehmungsgeschwindigkeit bei einem schizophrenen Patienten, die aus einer Studie von Meffert stammt (vgl. Holtzman 1963).

Die Zeitreihe in Abb. 4 enthält über den betrachteten Zeitraum von 180 Tagen eine exogene Treatmentvariable mit Langzeitintervallen von 60 Tagen zwischen den Treatmentänderungen.

Der Patient erhielt in den ersten 60 Tagen ein Placebo (Phase I). In den folgenden 60 Tagen wurde das Placebo durch Chlorpromazin ersetzt (Phase II). Vom 120. bis zum 179. Tag (Phase III) wurde die Chlorpromazinbehandlung fortgesetzt, allerdings überlagert von einer Elektroschocktherapie. Da die Phasen I - III "genügend viele" Daten umfassen, wird die Bestimmung eines adäquaten Noisemodells zunächst für jede Phase getrennt durchgeführt, um so zu erkennen, ob die Interventionen eine strukturelle Veränderung des der Zeitreihe zugrunde liegenden ARIMA-Prozesses bewirkt haben.

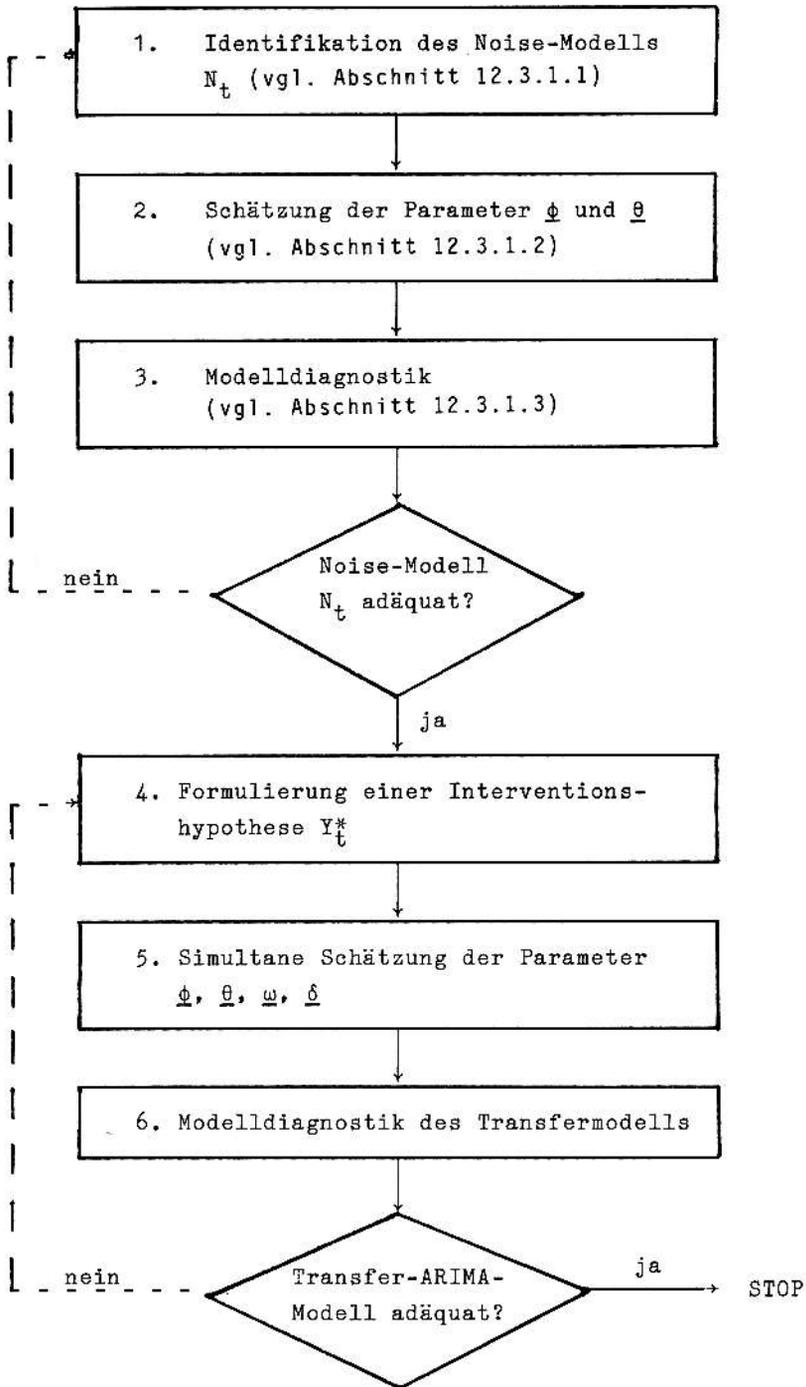


Abb. 3. Schematische Darstellung zur Überprüfung von Interventionshypothesen nach Box und Tiao

Die Identifikation, Parameterschätzung und Modelldiagnostik (vgl. Abschn. 12.3.1) ergeben, daß für jede Phase das ARIMA(0,1,1)-Modell

$$(1 - B)Y_t = \theta_0 + (1 - \theta_1 B)a_t \quad (23)$$

mit $\theta_1 = 0,76$ und damit für die gesamte Zeitreihe adäquat ist. Diese Homogenitätsprüfung ähnelt dabei der Homogenitätsprüfung in der Kovarianzanalyse oder der Strukturbruchanalyse in der Regression (Schneeweiß 1974). Der Graph der Zeitreihe zeigt deutliche Niveauschwankungen, was ebenfalls auf Nichtstationarität bezüglich des Niveaus ($d=1$) schließen läßt. Aus Gl. (23) folgt:

$$Y_t = \frac{\theta_0 + (1 - \theta_1 B)a_t}{(1 - B)} = N_t \quad (24)$$

Aus Abb. 4 ist zu erkennen, daß mit der Verabreichung des Tranquiliziers ein Levelsprung einsetzt. Für die 2. Intervention wurde zunächst ein einfaches Interventionsmodell mit einem Levelsprung am 120. Tag angenommen. Die geschätzten Residuen wichen signifikant von White-noise-Residuen ab. Daher werden für die Intervention jetzt 2 Effekte angenommen:

1. eine vorübergehende Niveauerhöhung vom 120. bis zum 131. Tag und

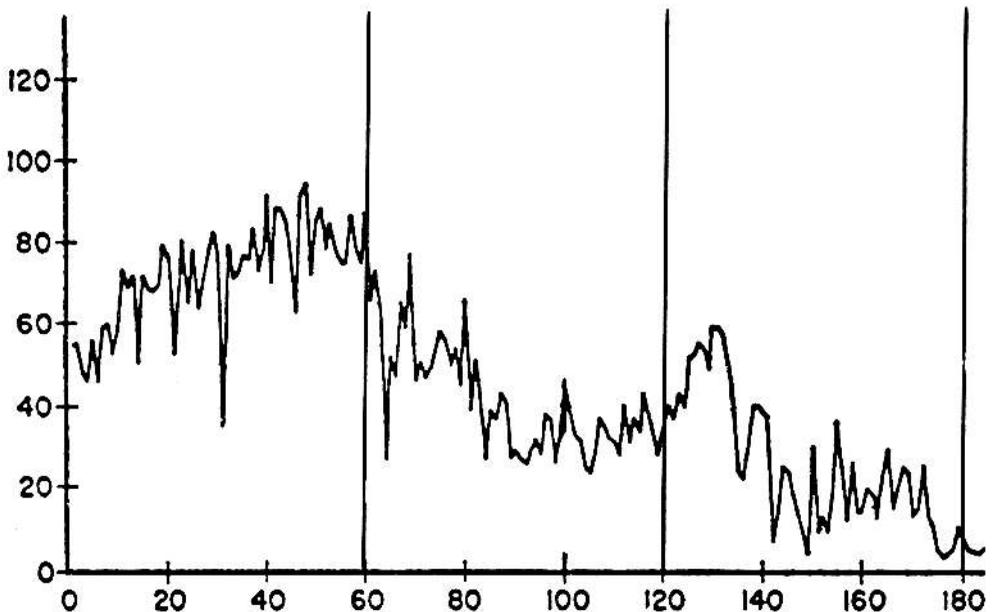


Abb. 4. Tägliche Messungen der Wahrnehmungsgeschwindigkeit bei einem schizophrenen Patienten. (Aus Holtzman 1963)

2. vom 132. bis zum 180. Beobachtungszeitpunkt ein exponentielles Abfallen der Wahrnehmungsgeschwindigkeit.

Die formale Darstellung der Interventionshypothese lautet:

$$Y_t^* = \omega_0 I_t^{(60)} + \omega_1 (I_t^{(120)} - I_t^{(131)}) + \frac{\omega_2}{(1 - \delta_1 B)} B^{12} I_t^{(120)}. \quad (25)$$

Das Interventionsmodell nach Gl. (17) ergibt sich zu:

$$\begin{aligned} Y_t &= Y_t^* + N_t \\ &= \omega_0 I_t^{(60)} + \omega_1 (I_t^{(120)} - I_t^{(131)}) + \frac{\omega_2}{(1 - \delta_1 B)} B^{12} I_t^{(120)} \\ &\quad + \frac{\theta_0 + (1 - \theta_1 B) a_t}{(1 - B)}. \end{aligned} \quad (26)$$

Die simultane Schätzung der Parameter des Interventionsmodells Gl. (26) liefert $\hat{\theta}_0 = -2,099$; $\hat{\theta}_1 = 0,76$; $\hat{\omega}_0 = -21,95$; $\hat{\omega}_1 = 9,09$; $\hat{\omega}_2 = -4,618$ so wie $\hat{\delta}_1 = 0,998$. Bis auf θ_0 sind alle Parameter signifikant verschieden von Null.

Im folgenden wollen wir die Auswertung des Zeitreihenexperiments von Meffert mit Hilfe des Randomisierungstests demonstrieren. Nach der Beobachtungsphase (Basislinie) A erfolgt die 1. Behandlungsphase B_1 , die direkt von der 2. Intervention B_2 abgelöst wird. Somit liegt dem Experiment ein A B_1 B_2 -Design zugrunde.

Um beim Randomisierungstest die approximative Prüfverteilung anwenden zu können, müssen jedoch mindestens 4 A- und 4 B-Phasen vorliegen (vgl. Möbus u. Nagl 1983, S. 292). Da jede Phase des A B_1 B_2 -Designs 60 Daten umfaßt, wird daher jede dieser Phasen in 4 Teilphasen zu jeweils 15 Beobachtungswerten eingeteilt, so daß wir ein

$$A_1 A_2 A_3 A_4 \mid B_1 B_2 B_3 B_4 \mid B_5 B_6 B_7 B_8 \text{ -Design erhalten.}$$

$I_1 \qquad I_2$

Die Zusammenfassung von jeweils 15 Tagen zu einer Einheit wird hier nicht inhaltlich begründet, sondern ist aus dem Wunsch zu erklären, prinzipiell die Nullhypothese verwerfen zu können. Dies wäre bei nur 3 Phasen nicht mehr möglich. Die Null- und die gerichtete Alternativhypothese lauten:

$$H_0 : \text{Kein Unterschied zwischen A- und B-Phase,}$$

$$\begin{aligned} H_1 : \bar{Y}_{B_8} = \bar{Y}_{B_7} = \bar{Y}_{B_6} = \bar{Y}_{B_5} < \bar{Y}_{B_4} = \bar{Y}_{B_3} = \bar{Y}_{B_2} = \bar{Y}_{B_1} \\ < \bar{Y}_{A_4} = \bar{Y}_{A_3} = \bar{Y}_{A_2} = \bar{Y}_{A_1}. \end{aligned}$$

Dabei gilt im einzelnen:

$$\begin{array}{llll} \bar{Y}_{B_8} = 13,4 & \bar{Y}_{B_7} = 19,33 & \bar{Y}_{B_6} = 23,933 & \bar{Y}_{B_5} = 50,0, \\ \bar{Y}_{B_4} = 34,067 & \bar{Y}_{B_3} = 31,4 & \bar{Y}_{B_2} = 42,867 & \bar{Y}_{B_1} = 55,6, \\ \bar{Y}_{A_4} = 80,53 & \bar{Y}_{A_3} = 76,267 & \bar{Y}_{A_2} = 71,533 & \bar{Y}_{A_1} = 58,26. \end{array}$$

Aufgrund der Struktur von H_1 erhalten die künstlichen Phasenmittellwerte in einem ordinalen Gewichtungsschema die Gewichte:

$$g_{B_i} = 1 \quad \text{für } i = 1, 2, 3, 4,$$

$$g_{B_j} = 2 \quad \text{für } j = 5, 6, 7, 8,$$

$$g_{A_k} = 3 \quad \text{für } k = 1, 2, 3, 4.$$

Gemäß Gl. (2) erhalten wir:

$$z_{\text{emp}} = \frac{\Gamma_{\text{emp}} - E(\Gamma)}{\sqrt{\text{Var}(\Gamma)}} = \frac{1244,331 - 1114,344}{64,064} = 2,029. \quad (27)$$

Bei einseitiger Fragestellung auf dem 5%-Niveau ist dieses Ergebnis signifikant.

LITERATUR

- Anderson OD (1975) Time series analysis and forecasting: The Box-Jenkins approach. Butterworths, London
- Anderson TW (1978) Repeated measurements in autoregressive processes. J Am Statist Assoc 73:371-378
- Aström KJ, Bohlin T (1966) Numerical identification of linear dynamic systems from normal operating records. In: Hammond PH (ed) Theory of self-adaptive control systems. Plenum, New York
- Bartlett MS (1946) On the theoretical specifications of the sampling properties of autocorrelated time series. J R Statist Soc B 8:27-41
- Box GEP, Jenkins GM (1976) Time series analysis: Forecasting and control. Holden-Day, San Francisco

- Box GEP, Pierce DA (1970) Distribution of residual autocorrelations in ARIMA time series models. *J Am Statist Assoc* 64:1509-1526
- Box GEP, Tiao GC (1965) A change in level of a non-stationary time-series. *Biometrika* 52:181-192
- Box GEP, Tiao GC (1975) Intervention analysis with applications to economic and environmental problems. *J Am Statist Assoc* 70:70-79
- Chatfield C (1975) *The analysis of time-series: Theory and practice*. Chapman & Hall, London
- Cook TD, Campbell DT (1979) Quasi-experiments: Nonequivalent control groups designs. In: Cook TD, Campbell TD (eds) *Quasi-experimentation: Design and analysis issues for field settings*. McNally, Chicago, pp 95-146
- Cooley WW, Lohnes PR (1971) *Multivariate data analysis*. Wiley & Sons, New York
- Edgington ES (1967) Statistical inference from $N = 1$ experiments. *J Psychol* 65:195-199
- Edgington ES (1969a) *Statistical inference: The distribution free approach*. McGraw-Hill, New York
- Edgington ES (1969b) Approximate randomization tests. *J Psychol* 72:143-149
- Edgington ES (1971) Randomization tests with statistical control over concomitant variables. *J Psychol* 79:13-19
- Edgington ES (1973) Randomization tests: Computer time requirements. *J Psychol* 85:89-95
- Edgington ES (1975a) Randomization tests for one-subject operant experiments. *J Psychol* 90:57-68
- Edgington ES (1975b) Randomization for predicted trends. *Can Psychol Rev* 16:49-53
- Edgington ES (1980) *Randomization tests*. Dekker, New York
- Fisher RA (1951) *The design of experiments*. Hafner, London
- Glass GV, Willson UL, Gottman JM (1975) *Design and analysis of time-series experiments*. Colorado Associated University Press, Boulder
- Gottman JM (1981) *Time-series analysis*. Cambridge University Press, Cambridge
- Gottman JM, Glass GV (1978) Analysis of interrupted time-series experiments. In: Kratochwill TR (ed) *Single subject research: Strategies for evaluating change*. Academic Press, New York, pp 197-235
- Hibbs DA (1977) On analyzing the effects of policy interventions: BOX-JENKINS and BOX-TIAO vs. structural equation models. In: Heise DR (ed) *Sociological methodology*. Jossey-Bass, New York, pp 137-179

- Holtzman WH (1963) Statistical models for the study of change in the single case. In : Harris C (ed) Problems in measuring change. University of Wisconsin Press, Madison Wisconsin
- Jenkins GM (1979) Practical experiences with modelling and forecasting time series. GJP Ltd., Channel Island
- Lehman EL (1975) Nonparametrics: Statistical methods based on ranks. Holden-Day, San Francisco
- Levin JR, Marascuilo LA, Hubert LJ (1978) $N = 1$ - Nonparametric randomization tests. In: Kratochwill TR (ed) Single subject research. Academic Press, New York, pp 167-196
- Makridakis S, Wheelwright SC (1978a) Forecasting - methods and applications. Wiley & Sons, New York
- Makridakis S, Wheelwright SC (1978b) Interactive forecasting. Holden-Day, San Francisco
- McCain LJ, McCleary R (1979) The statistical analysis of the simple interrupted time-series quasi-experiment. In: Cook TD, Campbell DT (eds) Quasi-experimentation: Design and analysis for field settings. McNally, Chicago, pp 233-293
- McCleary R, Hay RA (1980) Applied time-series analysis for the social sciences. Sage, Beverly Hills
- Möbus C, Nagl W (1983) Messung, Analyse und Prognose von Veränderungen. In: Bredenkamp J, Feger H (Hrsg) Hypothesenprüfung, Bd 5 der Serie Forschungsmethoden der Psychologie der Enzyklopädie der Psychologie. Verlag für Psychologie, Göttingen, S 239-470
- Möbus C, Göricke G, Kröh PA (1983) Statistical analysis of single-case experimental designs: Conditional equivalence of the general-linear-model approach of Glass, Willson and Gottman with the intervention model of Box and Tiao. EDV in Med Biol 14:98-108
- Nelson CR (1973) Applied time-series analysis. Holden-Day, San Francisco
- Quenouille MH (1949) Approximate tests of correlation in time series. J R Statist Soc B 11:68-74
- Revenstorf D (1979) Zeitreihenanalyse für klinische Daten: Methodik und Anwendungen. Beltz, Weinheim
- Revenstorf D, Keeser W (1979) Zeitreihenanalyse von Therapieverläufen. In: Petermann F, Hehl FJ (Hrsg) Einzelfallanalyse. Urban & Schwarzenberg, München, S 183-228
- Schneeweiß H (1974) Ökonometrie, 2. Aufl. Physica, Würzburg Wien