Die Berechnung eines Gütemaßes für lineare Struktur–gleichungsmodelle mit SAS

Matthias Hunscher
Friedrich–Schiller–Universität Jena
Fachbereich Wirtschaftswissenschaften
Lehrstuhl für Wirtschafts- und Sozialstatistik
Carl–Zeiss Str. 3
07743 Jena
Tel.: 03641/943306, Fax: 03641/943302
e–mail: m.hunscher@wiwi.uni-jena.de

Zusammenfassung

In diesem Beitrag wird ein Verfahren, insbesondere dessen Umsetzung in ein SAS–Programm, zur Bestimmung der Güte linearer Strukturgleichungsmodelle vorgestellt. Hierbei handelt es sich um den sogenannten Tetrad–Score (Spirtes, Glymour, Scheines, 1993). Der Tetrad–Score entspricht einer Maßzahl, die eine Information über die Güte des aus einem Datensatz generierten Strukturgleichungsmodells liefert, wobei die Berechnung des Tetrad–Scores über die Durchführung eines multiplen Testverfahrens erfolgt. Die Elementartests des Verfahrens werden mit dem Programm Tetrad II (Scheines et al.,1994) durchgeführt. Der Programm–Output wird in SAS eingelesen und mit dem Programm IML wird der multiple Test nach Shaffer durchgeführt. Es soll insbesondere gezeigt werden, wie die das Testverfahren nach Shaffer bestimmenden charakteristischen Eigenschaften von Tetrad–Gleichungen in das SAS–Programm implementiert werden können.

1 Einführung und Problemstellung

In diesem Kapitel wird ein Instrument vorgestellt, das zur Beurteilung der Anpassung von rekursiven linearen Strukturgleichungsmodellen an vorliegende Daten genutzt werden kann. Dieses Instrument ist der *Tetrad–Score*. Zunächst ist die Angabe der Definition eines *linearen Strukturgleichungsmodells* notwendig. Eine entsprechende Definition findet sich in Scheines et al. (1994):

Definition:
Ein lineares Strukturgleichungsmodell ist dadurch gegeben, daß für jede im Modell enthaltene Variable Y eine Strukturgleichung aufgestellt wird. In dieser Gleichung wird Y als Linearkombination der direkten Ursachen von Y und eines Fehlerterms ε_y dargestellt. Für die Fehlerterme wird vorausgesetzt, daß diese mit Erwartungswert Null unabhängig und identisch verteilt sind. □

Bemerkung:
Für die im weiteren Verlauf dieses Artikels anzusprechenden Testverfahren ist zusätzlich die Normalverteilungsannahme notwendig. □

Enthält ein lineares Strukturgleichungsmodell keine zyklischen Strukturen, so wird es als *rekursives lineares Strukturgleichungsmodell* bezeichnet. Für jedes rekursive lineare Strukturgleichungsmodell mit mindestens vier Variablen lassen sich *Tetrad–Differenzen* bilden. Für jeweils vier aus dem vorliegenden Modell ausgewählte Variablen I, J, K und L können drei verschiedene Tetrad–Differenzen gebildet werden (ρ_{IJ} gibt den Korrelationskoeffizienten zwischen den Variablen I und J an):

$$\begin{aligned} \rho_{IJ}\rho_{KL} &- \rho_{IL}\rho_{JK} \\ \rho_{IK}\rho_{JL} &- \rho_{IL}\rho_{JK} \\ \rho_{IJ}\rho_{KL} &- \rho_{IK}\rho_{JL} \end{aligned}$$

Besitzt eine Tetrad–Differenz für jede Parametrisierung eines theoretischen Modells den Wert Null, so nennt man diese Tetrad–Differenz *Tetrad–Gleichung* und man sagt, daß die Tetrad–Gleichung vom Modell impliziert wird. Besitzen zwei der drei Tetrad-Differenzen, die aus vier Variablen eines gegebenen Models gebildet werden können, den Wert Null, so ist der Wert der dritten Tetrad–Differenz ebenfalls Null.

Im Rahmen der Tetrad-Analyse kann der sogenannte *Tetrad-Score* (Spirtes et al., 1993) berechnet werden. Bei der Konstruktion des Tetrad–Scores werden die hier genannten Eigenschaften der Tetrad–Gleichungen ausgenutzt. Von einem gegebenen theoretischen Strukturgleichungsmodell wird eine bestimmte Zahl von Tetrad–Gleichungen impliziert und dies unabhängig von einer konkreten Parametrisierung. Prüft man nun für einen Datensatz ob diese Tetrad–Gleichungen von den Daten empirisch bestätigt werden, so erhält man einen Anhaltspunkt ob Daten und theoretisches Modell "zusammenpassen". Auf diesem Prinzip beruhend liefert der Tetrad-Score eine Maßzahl, die zur Bewertung der Güte der Anpassung eines theoretisches Modells an vorliegende Daten verwendet werden kann. Der Tetrad-Score ist definiert als

$$T = \sum_{t \in \mathrm{Implied}_{\mathrm{H}}} P(t) - \sum_{t \in \mathrm{Implied}_{\sim\mathrm{H}}} w\,(1 - P(t)).$$

$\mathrm{Implied}_{\mathrm{H}}$: die Menge der theoretischen Tetrad-Gleichungen, die von den Daten bestätigt werden.

$\mathrm{Implied}_{\sim\mathrm{H}}$: die Menge der theoretischen Tetrad-Gleichungen, die nicht von den Daten bestätigt werden.

$P(t)$: p-Wert des Tests $\mathrm{H}_0 : t = 0$ vs $\mathrm{H}_1 : t \neq 0$

w: Gewichtskonstante, $0 \leq w \leq 1$ □

Der Tetrad-Score beruht auf folgendem theoretischen Ansatz: Es wird für jede vom theoretischen Modell implizierte Tetrad-Gleichung mittels eines statistischen Tests überprüft, ob diese von den zugrundeliegenden Daten bestätigt wird, d.h. für die entsprechende empirische Tetrad–Differenz wird obiger Test durchgeführt. Eine vom theoretischen Modell implizierte Tetrad-Gleichung geht mit einem positiven Wert in den Tetrad-Score ein, wenn diese von den Daten per Test bestätigt wird. Im gegenteiligen Fall, d.h. Nichtbestätigung durch die Daten, geht die Tetrad-Gleichung mit einem negativen Wert in den Tetrad-Score ein. Aus Gründen der Interpretierbarkeit wird der oben definierte Tetrad-Score T

so normiert, daß er als minimalen Wert 0 und als maximalen Wert 100 annehmen kann. Je größer der Wert des Tetrad–Scores ist, um so besser ist die Anpassung des theoretischen Modells an die Daten, denn bei einem hohen Score wird ein großer Anteil der vom theoretischen Modell implizierten Tetrad–Gleichungen von den Daten bestätigt.

Die Qualität des Tetrad–Scores hängt in erster Linie von der Güte des zu verwendenden Tests auf verschwindende Tetrad–Differenzen ab. In den nachfolgenden Kapiteln werden verschiedene Testverfahren für die Berechnung des Tetrad–Scores vorgestellt.

2 Die Berechnung des Tetrad–Scores als multiples Testproblem

Werden von einem theoretischen Modell zwei oder mehr Tetrad–Gleichungen impliziert, d.h. es gilt $k \geq 2$ ($k \cong$ Anzahl der vom Modell implizierten Tetrad–Gleichungen), so handelt es sich um ein multiples Testproblem. Ein multiples Testproblem liegt vor, wenn gleichzeitig (parallel) mehrere Hypothesen, sogenannte Elementarhypothesen, zu prüfen sind (Horn, Vollandt 1995). Im vorliegenden Fall entsprechen die Hypothesen $\mathrm{H}_{0i} : t_i = 0$, $i = 1, ..., k$, bzgl. der theoretisch implizierten Tetrad–Gleichungen den Elementarhypothesen. Diese Hypothesen werden mittels der vorliegenden Daten statistisch überprüft. Eine formale Definition des multiplen Testproblems findet sich in Schlittgen (1996):

Definition:
Ein multipler Test ist eine Abbildung

$$\varphi : \mathbb{R}^n \longrightarrow \{0,1\}^k, \quad k \geq 2$$

mit den Komponenten $\varphi_i : \mathbb{R}^n \longrightarrow \{0,1\}, i = 1, \ldots, k$. Gilt $\varphi_i(x) = 0$, so wird die Elementarhypothese H_{oi} beibehalten, im Falle von $\varphi_i(x) = 1$ wird für H_{1i} entschieden. □

Im Rahmen der multiplen Testverfahren gibt es drei verschiedene Ansätze ein Signifikanzniveau α festzulegen. Tests können zu einem lokalen, zu einem globalen und zu einem multiplen Signifikanzniveau durchgeführt werden. Für das Tetrad–Score–Problem ist das multiple Signifikanzniveau zu verwenden.

Ein multipler Test hält das multiple Signifikanzniveau α ein, wenn die Wahrscheinlichkeit, mindestens eine Elementarhypothese irrtümlich abzulehnen, höchstens α beträgt, unabhängig davon, wieviele und welche Elementarhypothesen wahr sind. Die formale Bedingung für die Einhaltung eines multiplen Niveaus α bei gegebenen Hypothesen $\mathrm{H}_{01}, \ldots, \mathrm{H}_{0k}$ lautet demgemäß (I entspricht der Menge der Indizes der wahren Hypothesen):

$$P(\mathrm{H}_{1i_1}, ..., \mathrm{H}_{1i_j} | \mathrm{H}_{0i} \; \forall i \in I) \leq \alpha \quad \forall \, \{i_1, ..., i_j\} \subseteq I, \; I \subseteq \{1, ..., k\}$$

Der Tetrad-Score nimmt einen falschen Wert an, wenn schon einer der k Elementartests zu einer falschen Entscheidung führt, denn: Wird eine Elementarhypothese irrtümlich abgelehnt, so wird ein zu niedriger Tetrad-Score berechnet. Im umgekehrten Fall, der irrtümlichen Nichtablehnung einer Elementarhypothese, wird ein zu großer Tetrad-Score berechnet. Die Kontrolle des Fehlers, mindestens eine Elementarhypothese fälschlicherweise abzulehnen, gelingt mit der Durchführung eines Tests zu einem multiplen Niveau, denn laut Definition hält ein multipler Test das multiple Signifikanzniveau α ein, wenn die Wahrscheinlichkeit, mindestens eine Elementarhypothese irrtümlich abzulehnen, höchstens α beträgt (Horn, Vollandt 1995). Für einen multiplen Test zum multiplen Niveau α gilt also, daß mit einer Sicherheitswahrscheinlichkeit $(1-\alpha)$ alle wahren Hypothesen nicht verworfen werden.

In den Tetrad-Score gehen Ergebnisse von Elementartests ein, die nicht verworfen werden können. Für einen auf das Tetrad-Score-Problem anzuwendenden multiplen Test bedeutet dies, daß dieser eine möglichst hohe Güte besitzen sollte, so daß falsche Nullhypothesen mit einer hohen Wahrscheinlichkeit abgelehnt werden, d.h. es muß bei jedem Elementartest Wert auf die Kontrolle des Fehlers 2. Art gelegt werden. Ziel der folgenden Betrachtungen ist es, für die Berechnung des Tetrad-Scores einen multiplen Test mit möglichst hoher Güte zu finden.

Multiple Testverfahren werden als Einschritt- oder als Mehrschrittprozeduren realisiert. Einschritt-Verfahren (*single-step procedures*) sind Methoden, bei denen man alle Elementarhypothesen gleichzeitig und auf demselben (lokalen) Signifikanzniveau prüft. Im Rahmen von Mehrschritt-Verfahren (*multi-step procedures*) werden die Elementarhypothesen in Abhängigkeit von den Realisationen der Teststatistiken der Elementartests in einer bestimmten Reihenfolge und im allgemeinen auf unterschiedlichen (lokalen) Signifikanzniveaus geprüft.

3 Multiple Testverfahren zur Bestimmung des Tetrad–Scores

Im folgenden werden zwei multiple Testverfahren für die Berechnung des Tetrad-Scores vorgestellt. Hierbei handelt es sich um die Mehrschrittverfahren nach Holm (1979) und nach Shaffer (1986). Das Verfahren nach Shaffer entspricht einer Weiterentwicklung des Verfahrens nach Holm, so daß es als nützlich für das Verständnis erscheint, das Verfahren nach Holm kurz vorzustellen.

3.1 Der multiple Test nach Holm

Der multiple Test nach Holm (1979), der das multiple Niveau α einhält, besitzt bezogen auf das Tetrad-Score-Problem folgenden Ablauf: Bei k vom theoretischen Modell implizierten Tetrad-Gleichungen sind k Elementarhypothesen zu prüfen. Für diese Hypothesen berechnet man zunächst die zugehörigen p-Werte. Die nach ihrer Größe geordneten p-Werte werden mit $p_{(1)}, ..., p_{(k)}$ bezeichnet. Im ersten Schritt vergleicht man den kleinsten p-Wert $p_{(1)}$ mit α/k, wobei α das multiple Niveau bezeichnet. Falls $p_{(1)} > \alpha/k$ gilt, bricht das Verfahren ab und es wird keine Elementarhypothese abgelehnt. Für den Tetrad-Scores bedeutet dies, daß jede theoretisch implizierte Tetrad-Gleichung per Test durch die Daten bestätigt wird und somit der Tetrad-Score den maximalen Wert 100 annimmt. Gilt $p_{(1)} \leq \alpha/k$, so wird die zum p-Wert $p_{(1)}$ gehörige Hypothese abgelehnt. Im nächsten Verfahrensschritt wird $p_{(2)}$ mit $\alpha/(k-1)$ verglichen. Falls $p_{(2)} > \alpha/(k-1)$ gilt, bricht das Verfahren ab und es wird außer der bereits im ersten Schritt abgelehnten Hypothese keine weitere abgelehnt. Gilt $p_{(2)} \leq \alpha/(k-1)$, so wird auch die zu $p_{(2)}$ gehörige Hypothese abgelehnt. Allgemein wird $p_{(i)}$ mit $\alpha/(k-i+1)$ bis zu dem i verglichen, bei dem erstmalig

$p(i) > \alpha/(k-i+1)$ gilt. Ist ein derartiges i gefunden bricht das Verfahren ab und die verbleibenden Elementarhypothesen werden nicht abgelehnt. Führt das Verfahren nicht zu einem Abbruch, so werden alle Elementarhypothesen abgelehnt. Das bedeutet, daß keine theoretisch implizierte Tetrad-Gleichung durch die Daten per Test bestätigt wird. In diesem Fall besitzt der Tetrad-Score den Wert Null.

3.2 Der multiple Test nach Shaffer

Shaffer (1986) hat das multiple Testverfahren von Holm weiterentwickelt. Das resultierende Testverfahren besitzt eine höhere Güte als das Verfahren von Holm.

Dem Testverfahren von Shaffer, welches das multiple Niveau α einhält, liegt folgender Ablauf zugrunde: Die zu den k Elementartests gehörigen p-Werte werden der Größe nach geordnet. Bei Durchführung der Testprozedur nach Holm wird unter der Voraussetzung der Ablehnung der $(i-1)$ vorhergehenden Elementarhypothesen H_{0i} abgelehnt, wenn gilt: $p_{(i)} \leq \alpha/(k-i+1)$. Diese Vorschrift ist von Shaffer so abgeändert worden, daß die Hypothese H_{0i} abgelehnt wird, wenn gilt: $p_{(i)} \leq \alpha/r_i$. r_i gibt die maximal mögliche Anzahl an wahren Hypothesen gegeben, daß die $(i-1)$ vorhergehenden Elementarhypothesen falsch sind, an. Befindet sich das Verfahren in Stufe i verbleiben $(k-i+1)$ noch zu testende Hypothesen, d.h. es gilt $r_i \leq k-i+1$. Dies bedeutet, daß die Elementarhypothesen auf der jeweiligen Stufe zu dem gleichen oder zu einem höheren Signifikanzniveau getestet werden. In dieser Aussage ist die Begründung enthalten, warum das Verfahren von Shaffer eine höhere Güte als das Verfahren von Holm besitzt: Hypothesen, die nach dem Verfahren von Holm abgelehnt werden, werden auch nach dem Verfahren von Shaffer abgelehnt, da das Signifikanzniveau auf jeder Stufe mindestens ebenso hoch ist. Dies gilt insbesondere für falsche Hypothesen.

Nun wird gezeigt, wie für das Tetrad-Score-Problem in Verfahrensstufe i die Anzahl der noch möglichen wahren Hypothesen bestimmt werden kann. Grundlage der folgenden Aussagen ist, daß in einem Dreierblock von Tetrad-Differenzen (Dreierblock $\cong$ die drei aus vier ausgewählten Variablen generierbaren Tetrad-Differenzen) keine, eine oder drei Tetrad-Gleichungen impliziert werden können, jedoch nicht zwei Tetrad-Gleichungen. Dies führt in Verfahrensstufe $i,\ i \in \{1, ..., k\}$, zu einer Fallunterscheidung, in der für jeden Dreierblock aus Tetrad-Differenzen separat bestimmt wird, wieviele Hypothesen noch wahr sein können. Eine Summation über die Einzelergebnisse ergibt die Gesamtzahl der noch möglichen Hypothesen. Die Fallunterscheidung besitzt die Gestalt:

1. Fall:
Keiner der drei Tests auf eine verschwindende Tetrad-Differenz ist bis zur Stufe i abgelehnt worden. Dies bedeutet, daß alle drei Hypothesen wahr sein können.

2. Fall:
Einer der zum Dreierblock gehörigen Tests ist bis zum Erreichen der Stufe i abgelehnt worden. Dies bedeutet, daß nur eine der beiden verbleibenden Hypothesen wahr sein kann.

3. Fall:
Zwei der zum Dreierblock gehörigen Tests sind bis zum Erreichen von Stufe i abgelehnt worden. Dies bedeutet, daß die verbleibende Hypothese wahr sein kann.

4. Fall:
Alle drei Tests sind bereits vor Erreichen der Stufe i abgelehnt worden. Dies bedeutet, daß es in dem betrachteten Dreierblock keine wahre Hypothese gibt.

Diese Fallunterscheidung enthält die Annahme, daß alle drei Tetrad–Gleichungen eines Blockes theoretisch impliziert werden. Auch wenn dies nicht zutrifft bleibt die Argumentation erhalten. Es muß zusätzlich überprüft werden, ob die betrachtete Tetrad–Differenz für das Testproblem relevant ist. Das Eintreten von Fall 2 führt im Vergleich zum Verfahren von Holm zu einer Erhöhung des Signifikanzniveaus in Stufe i. In den Fällen 1 und 3 entspricht die Anzahl der noch offenen Testentscheidungen der Anzahl der Hypothesen, die wahr sein können. Wenn Fall 2 für keinen der Dreier–Blöcke mehr eintreten kann, entspricht r_i der Anzahl der noch nicht betrachteten Hypothesen.

4 Aspekte der Implementierung in SAS

In diesem Kapitel wird die Umsetzung des multiplen Testverfahrens nach Shaffer für das Tetrad–Score–Problem in ein SAS–Programm dargestellt.

Ausgangspunkt für die Realisierung ist der Output des Tetrad–Moduls *Tetrads* (Scheines et al., 1994). Input in dieses Modul sind ein graphisches Modell und ein Datensatz. Im Output von *Tetrads* sind folgende relevante Informationen enthalten:

- Für jede Tetrad–Differenz des gegebenen graphischen Modells wird angegeben, ob die zugehörige Tetrad–Gleichung theoretisch impliziert wird.
- Für jede Tetrad–Differenz wird der p–Wert zum Test $H_0 : t_i = 0$ ausgegeben.
- Der Output der Prozedur *Tetrads* ist nach den genannten Dreierblöcken geordnet, so daß Abfragen gemäß der in Kapitel 3 geschilderten Fallunterscheidung im SAS–Programm innerhalb einer Schleife mit Schrittweite Drei getätigt werden können.

Im folgenden werden wichtige Aspekte der einzelnen Ablaufschritte des in SAS implementierten Programms zur Durchführung des multiplen Testverfahrens nach Shaffer dargestellt.

1. Schritt:
Ein Beispiel für den Output der Prozedur *Tetrads* besitzt folgende Gestalt:

```
x1 x2, x3 x4 - x1 x3, x2 x4              -0.1488          0.0000
x1 x3, x2 x4 - x1 x4, x2 x3               0.1475          0.0000
x1 x4, x2 x3 - x1 x2, x3 x4               0.0013   0   0  0.7875

x1 x2, x3 x5 - x1 x3, x2 x5              -0.1587          0.0000
x1 x3, x2 x5 - x1 x5, x2 x3               0.1504          0.0000
x1 x5, x2 x3 - x1 x2, x3 x5               0.0083   0   0  0.0717
                                             ⋮
```

Dieser Output wird im Data–Step mit folgender Input–Zeile eingelesen:

```
input t_gl $ 1-30 t_wert +1 i $ 1. +2 h $ 1. +3 prob 6.4;
```

Für die Durchführung des multiplen Tests sind die Variablen i und *prob* von Relevanz. Wenn die entsprechende Tetrad–Gleichung theoretisch impliziert wird, so besitzt die Variable i entweder die Ausprägung '0' oder 'T'. Ansonsten ist die Variable i nicht belegt. In der Variable *prob* sind die p–Werte der innerhalb von *Tetrads* durchgeführten Elementartests abgespeichert.

2. Schritt:
Innerhalb des Data–Steps wird durch Generierung von zwei Variablen das multiple Signifikanzniveau und die in der Berechnungsvorschrift des Tetrad–Scores enthaltene Gewichtskonstante w festgelegt:

```
alpha = ...;
weight = ...;
```

3. Schritt:
Die für die Durchführung des multiplen Testverfahrens nach Shaffer relevanten Variablen werden in SAS–IML mit dem IML–Befehl *read* eingelesen:

- Das multiple Testniveau.
- Die p–Werte der Elementartests.
- Die Anzahl der Tetrad–Differenzen des gegebenen Modells.
- Die Anzahl der vom Modell theoretisch implizierten Tetrad–Gleichungen.
- Eine Indikatorvariable, die anzeigt, ob die entsprechende Tetrad–Gleichung vom Modell impliziert wird.

4. Schritt
In diesem Schritt wird der multiple Test nach Shaffer durchgeführt. Der Testablauf ist in eine *Do–Until*–Schleife eingebunden. Der Durchlauf dieser Schleife wird abgebrochen, wenn keine Elementarhypothese mehr wahr sein kann. In dieser Schleife werden die Berechnungen der zu der in Kapitel 3 beschriebenen Fallunterscheidung durchgeführt, d.h. für jeden Dreierblock wird unter Beachtung der Anzahl der im Block bereits abgelehnten Hypothesen bestimmt, wieviele Hypothesen noch wahr sein können. Dies geschieht unter Beachtung der Tatsache, daß die Anzahl der noch möglichen Hypothesen von der Zahl der relevanten Tetrad–Differenzen abhängt. Relevant sind in diesem Sinne die Tetrad–Differenzen bzw. Tetrad–Gleichungen, die vom Modell theoretisch impliziert werden. Der entsprechende Programmteil zur Bestimmung der noch möglichen wahren Hypothesen besitzt folgende Gestalt

```
do j = 1 to MAXI by 3;
  HILF3 = H[j] + H[j+1] + H[j+2];
  M = (j + 2) / 3;
  S = X[j] + X[j+1] + X[j+2];
  if HILF3 = 0 then do;
   if S = 3 then L[M] = 3;
```

```
    if S = 2 then L[M] = 2;
    if S = 1 then L[M] = 1;
    if S = 0 then L[M] = 0;
  end;
  if HILF3 = 1 then do;
    if (S - HILF3) = 2 then L[M] = 1;
    if (S - HILF3) = 1 then L[M] = 1;
    if (S - HILF3) = 0 then L[M] = 0;
  end;
  if HILF3 = 2 then do;
    if (S - HILF3) = 1 then L[M] = 1;
    if (S - HILF3) = 0 then L[M] = 0;
  end;
  if HILF3 = 3 then L[M] = 0;
end;
```

Die Bedeutung der Variablen:

MAXI	:	Die Anzahl der Tetrad–Differenzen.
HILF3	:	Die Anzahl der in einem Dreierblock bereits abgelehnten Hypothesen.
S	:	Die Anzahl der relevanten Tetrad–Differenzen.
L	:	Die Anzahl der noch möglichen Hypothesen

5. Schritt
Das Ergebnis des IML–Programms wird in der 1/0–Variable h abgespeichert. In dieser Variable ist vermerkt, ob die jeweilige Tetrad–Gleichung von den Daten im Rahmen des multiplen Tests bestätigt wurde, d.h. daß der zugehörige Elementartest nicht abgelehnt wurde. Die genannte Variable wird mit den Befehlen *create* und *append* in den Data–Step zur weiteren Verarbeitung übergeben.

6. Schritt:
Im Data–Step erfolgt die eigentliche Berechnung des Tetrad–Scores. Hierfür liegen die benötigten Informationen vor: Die Variable i (Indikator für die theoretische Implizierung), die Variable h (Indikator für die Implizierung durch die Daten), die Variable *prob* (die p–Werte der Elementartests) und die Gewichtungsvariable *weight*. Die Berechnungsformel des Tetrad–Scores wird auf jede Beobachtung des Datensatzes, d.h. auf jede relevante Tetrad–Differenz angewendet. Der Tetrad–Score wird dann durch die Bildung kumulierter Summen berechnet, d.h. das gesuchte Ergebnis erscheint in der letzten Zeile der generierten SAS–Datei.

Alternativ ist es natürlich auch möglich, den Tetrad–Score mit einem IML–Programm zu berechnen.

5 Zusammenfassung

In diesem Artikel ist eine Verbindung zwischen der Theorie der multiplen Tests und der Berechnung des Tetrad–Scores hergestellt worden. Es wurde begründet, daß als Signifikanzniveau für die Berechnung des Tetrad–Scores mittels eines multiplen Tests ein multiples Niveau anzusetzen ist. Desweiteren wurde für die klassischen multiplen Testverfahren nach Holm und Shaffer beschrieben, wie diese konkret auf das Tetrad–Score–Problem anzuwenden sind. Dies geschah unter dem Aspekt, für die Berechnung des Tetrad–Scores einen Test mit hoher Güte zu verwenden. In einem weiteren Kapitel wurde darauf eingegangen, wie der Tetrad–Score konkret mit SAS berechnet werden kann.

Diese Arbeit entstand im Rahmen eines Projekts des Landes Thüringen.

Literatur:

Holm, S. (1979): *A Simple Sequentially Rejective Multiple Test Procedure*. Scand. J. Statist. 6, 65–70.

Horn, M., Vollandt, R. (1995): Multiple Tests und Auswahlverfahren. Stuttgart, Jena, New York: G. Fischer.

SAS/IML Software: Usage and Reference, Version 6, First Edition. SAS Institute Inc.

Scheines, R., Spirtes, P., Glymour, C., Meek, C. (1994): Tetrad II – tools for causal modeling. New York: Lawrence Erlbaum Associates.

Schlittgen, R. (1996): Statistische Inferenz. München: Oldenbourg.

Shaffer, J. P. (1986): *Modified sequentially rejective multiple test procedures*. Journal of the American Statistical Association 81, 826–831.

Spirtes, P., Glymour, C., Scheines, R. (1993): Causation, Prediction and Search. New York: Springer–Verlag.