

Verborgene Schätze heben: Die Offset-Option in PROC GENMOD

Olaf Kruse

VST Gesellschaft für Versicherungsstatistik mbH Hannover

Telefon: 0511 – 339 599-21

Olaf.Kruse@vst-gmbh.de

Zusammenfassung

Eine mächtige, aber wenig bekannte Option in PROC GENMOD ist die OFFSET-Option. Sie wird in der SAS-Dokumentation für PROC GENMOD eingeführt, um in Log-Linearen-Modellen den Einfluss der Anzahl der Versuche n_i auf die Anzahl der Erfolge y_i zu berücksichtigen.

Die OFFSET-Option eignet sich hervorragend, um bei Modellen, bei denen z.B. der Einfluss von Kovariablen durch externe Informationen vorgegeben ist, die freien Parameter in Abhängigkeit dieser Vorgaben zu optimieren. Aus diesem Ansatz leiten sich heuristische Methoden zur Modellauswahl und –Diagnose ab, auf die abschließend kurz eingegangen wird.

Keywords: Generalisiertes Lineares Modell, PROC GENMOD, Offset, a-priori-Information, Parameter, Restriktionen.

1 Einleitung

Die Familie der Generalisierten Linearen Modelle (GLMs) wurde erstmalig von Nelder & Wedderburn [1972] zusammenhängend dargestellt. Wie der Name andeutet, stellen GLMs eine Erweiterung des klassischen linearen Modells dar. Viele bekannte Modelle, wie das Logit-, Probit- oder loglineare Modell, gehören zu der Familie der GLMs. Auf eine Einführung der theoretischen Grundlagen von GLMs wird an dieser Stelle verzichtet. Hier sei auf die einschlägige Literatur, wie z.B. die SAS-Dokumentation für PROC GENMOD verwiesen.

Mit der PROC GENMOD bietet SAS ein mächtiges Tool nicht nur zur Modellierung einzelner GLMs, sondern vor allem zur Modellauswahl bzw. zum Modellvergleich. Eine i.d.R. wenig beachtete Modelloption ist die OFFSET-Option. Mit Hilfe dieser Option können extern vorgegebene Effekte („a-priori-Informationen“) bei der Modellierung bzw. Schätzung der verbleibenden freien Parameter explizit berücksichtigt werden.

2 Problemstellung

In der praktischen Anwendung von GLMs -z.B. für die Tarifikalkulation in der Versicherungswirtschaft- müssen häufig externe Vorgaben für einzelne Merkmale bzw. Modellvariablen berücksichtigt werden. Ein typisches Beispiel ist das Merkmal „Schadenfreiheitsrabatt“ in der Kraftfahrt-Haftpflicht-Versicherung. Die Indexwerte („Modell-Parameter“) für dieses Merkmal sind in den Versicherungsbedingungen fest verankert und werden i.d.R. nicht mit jeder neuen Tarifgeneration angepasst.

Diese vorgegebenen Modellparameter müssen natürlich bei der Ermittlung aller andern Parameter a priori berücksichtigt werden. Hierbei handelt es sich um eine Modellschätzung mit Restriktionen in den Parametern. Die prinzipielle Idee ist, a-priori-Informationen in geeigneter Weise vom Beobachtungsvektor y der zu erklärenden Variable zu subtrahieren und die Modellschätzung für die verbleibenden Parameter mit dem modifizierten Beobachtungsvektor y^* vorzunehmen. Auf diesem Ansatz basiert z.B. auch der verallgemeinerte F-Test zum Test auf eine Untermenge der Parameter.

3 A-priori-Informationen im klassischen linearen Modell

Einfach und anschaulich lässt sich der Ansatz zur Berücksichtigung von a-priori-Informationen in der Modellschätzung am Beispiel des klassischen linearen Modells der Form

$$y_t = \mathbf{x}_t' \boldsymbol{\beta} + e_t \quad \text{mit} \quad e_t \sim N(0, \sigma_t^2) \quad \text{und} \quad t = (1 \dots T) \quad (1)$$

darstellen. Hierbei bezeichnen y_t die Responsevariable bzw. \mathbf{x}_t den Vektor der erklärenden nicht stochastischen Variablen für die t-te der insgesamt T Beobachtungen und $\boldsymbol{\beta}$ den zu schätzenden Parametervektor mit insgesamt P

Elementen. Es wird angenommen, dass die Störterme e_t unabhängig normalverteilt sind. Das Modell wird in der Form

$$y_t = \mathbf{x}_{1t}' \boldsymbol{\beta}_1 + \mathbf{x}_{2t}' \boldsymbol{\beta}_2 + e_t \quad \text{mit} \quad \boldsymbol{\beta}_{1[M \times 1]} \quad \text{und} \quad \boldsymbol{\beta}_{2[N \times 1]} \quad (2)$$

aufgeteilt, so dass die a-priori-Informationen in dem Vektor $\boldsymbol{\beta}_2$ zusammengefasst werden. Alle verbleibenden und noch zu schätzenden Parameter sind in dem Vektor $\boldsymbol{\beta}_1$ zusammengefasst. Nun werden die a-priori-Informationen auf beiden Seiten der Modellgleichung subtrahiert.

$$y_t^* = \mathbf{x}_{1t}' \boldsymbol{\beta}_1 + e_t^* \quad \text{mit} \quad y_t^* = [y_t - \mathbf{x}_{2t}' \boldsymbol{\beta}_2] \quad (3)$$

Der Parametervektor $\boldsymbol{\beta}_1$ des modifizierten Modells kann z.B. mit dem Kleinste-Quadrate-Schätzer

$$\hat{\boldsymbol{\beta}}_1 = (\mathbf{X}_1' \boldsymbol{\Omega} \mathbf{X}_1)^{-1} \mathbf{X}_1' \boldsymbol{\Omega} \mathbf{y}^* \quad (4)$$

geschätzt werden. In der Diagonalmatrix $\boldsymbol{\Omega}$ sind die Beobachtungsgewichte zusammengefasst. Im klassischen Fall mit identischen Beobachtungsgewichten hat die Matrix $\boldsymbol{\Omega}$ keinen Einfluss auf die Modellschätzung.

$$\frac{(\mathbf{e}^* \mathbf{e}^* - \mathbf{e}' \mathbf{e}) / M}{\mathbf{e}' \mathbf{e} / (T - M - N)} \sim F_{M, T - M - N}^M \quad (5)$$

Auf der Logik aus (2) und (3) basiert auch der allg. F-Test (5) zum Test auf eine Untermenge der Parameter. Im Vektor $\boldsymbol{\beta}_2$ sind Parameter, über die Hypothesen aufgestellt werden, zusammengefasst.

4 A-priori-Informationen im generalisierten linearen Modell

Die Verallgemeinerung dieses Ansatzes zur Berücksichtigung von a-priori-Informationen in der Modellschätzung für GLMs ist etwas aufwendiger. Das GLM in der Form

$$y_t = g^{-1}(\mathbf{x}' \boldsymbol{\beta}) + e_t \quad \text{mit} \quad \text{Var}(y_t) = \frac{\phi}{n_t} V(\mu_t), \quad (6)$$

wobei $g^{-1}(\cdot)$ die Linkfunktion und $V(\cdot)$ die Varianzfunktion des jeweiligen GLMs bezeichnen, wird wiederum so aufgeteilt, dass die a-priori-Informationen in dem Vektor $\boldsymbol{\beta}_2$ zusammengefasst werden:

O. Kruse

$$y_t = g^{-1}(\mathbf{x}_{1t}'\boldsymbol{\beta}_1 + \mathbf{x}_{2t}'\boldsymbol{\beta}_2) + e_t \quad \text{mit} \quad \text{Var}(y_t) = \frac{\phi}{n_t} V(\mu_t) \quad (7)$$

Alle verbleibenden und noch zu schätzenden Parameter sind in dem Vektor $\boldsymbol{\beta}_1$ zusammengefasst. Hierbei bezeichnen n_t die Beobachtungsgewichte und ϕ den i.d.R. unbekanntem Dispersionsparameter. Bei der Bildung des modifizierten Modells müssen die a-priori-Informationen über die Linkfunktion auf das Skalenniveau von y_t transformiert werden:

$$y_t^* = g^{-1}(\mathbf{x}_{1t}'\boldsymbol{\beta}_1) + e_t^* \quad \text{mit} \quad (8)$$
$$y_t^* = g^{-1}[g(y_t) - \mathbf{x}_{2t}'\boldsymbol{\beta}_2] \quad \text{und} \quad \text{Var}(y_t) = \frac{\phi}{n_t} V(\mu_t)$$

Zu beachten ist, dass die Varianzfunktion sich nicht ändert. Sie bezieht sich weiterhin auf die untransformierte Responsevariable y_t . Diese nur marginale Änderung in der Schätzung verhindert aber eine direkte Schätzung des modifizierten Modells mit PROC GENMOD.

5 Modellierung von GLMs mit PROC GENMOD

Die Umsetzung in SAS wird an Hand des Datensatzes von Baxter, Coutts & Ross, (1980) demonstriert. Dieser Datensatz enthält 6.761 Schadenfälle mit Schadenzahlungen, die durch 3 Tarifmerkmale mit 5, 4 bzw. 3 Ausprägungen in 60 Zellen zusammengefasst werden. Der Schadenbedarf „SB“, der sich aus dem Quotient der Schadenaufwendungen „c“ und der Jahreseinheiten „je“ ergibt, wird durch drei Klassifikationsmerkmale „SFR“, „KW“ und „REGIO“ beschrieben. Ausgehend von dem Poisson-Modell mit logarithmischer Linkfunktion wird in mehreren Schritten gezeigt, wie die Berücksichtigung von a-priori-Informationen in PROC GENMOD implementiert werden kann.

Bei der Modellierung mit PROC GENMOD (Beispiel 1) müssen neben der aus PROC GLM bekannten MODEL-Anweisung mit Angabe der Modellvariablen lediglich die Verteilung der Responsevariablen und die Linkfunktion (DIST- und LINK-Option in der MODEL-Anweisung) definiert werden. Mit der CLASS-Anweisung werden nominalskalierte Klassifikationsmerkmale definiert. Mit der WEIGHT- oder SCWGT-Anweisung wird die GewichtungsvARIABLE n berücksichtigt. Mit der DSCALE-Option wird der Disper-

sionsparameter ϕ , der im Poisson-Modell eigentlich per Definition eins ist, zusätzlich aus den Daten geschätzt.

```

/*****
/**   Bsp. 1 Poisson-Modell "Einfach"   **/
/*****
proc genmod data=KSFE_daten;
  class SFR KW REGIO ;
  model SB = SFR KW REGIO
           / dist=poisson link=log dscale;
  weight JE;
run;

```

Beispiel 1: PROC GENMOD-Code

In Tab. 1 ist der SAS-Output für Beispiel 1, das für alle folgenden Beispiele als Referenzmodell dient, ausschnittsweise wiedergegeben. In den folgenden Beispielen werden a-priori-Informationen, die i.d.R. den SFR-Parametern aus Beispiel 1 entsprechen, für das Merkmal „SFR“ vorgegeben. Bei korrekter Spezifikation ist zu erwarten, dass die Parameterschätzer für die andern beiden Merkmale mit denen aus dem Referenzmodell aus Beispiel 1 zusammenfallen.

Analysis Of Parameter Estimates			
Parameter	DF	Estimate	Std. Error
Intercept	1	5.1714	0.0894
SFR S1	1	0.2230	0.0564
SFR S2	1	0.1319	0.0339
SFR S3	1	-0.1404	0.0442
SFR S4	1	-0.0168	0.0344
SFR S5	0	0.0000	0.0000
KW K1	1	-0.4025	0.0370
KW K2	1	-0.2586	0.0385
KW K3	0	0.0000	0.0000
REGIO R1	1	0.6760	0.0834
REGIO R2	1	0.5871	0.0837
REGIO R3	1	0.3255	0.0949
REGIO R4	0	0.0000	0.0000
Scale	0	18.8606	0.0000

Tabelle 1: Parameterschätzer für Beispiel 1

6 Manueller Offset in PROC GENMOD

In Beispiel 2 wird gezeigt, wie a-priori-Informationen –hier die Parameter für das Merkmal SFR aus Beispiel 1- manuell in PROC GENMOD berücksichtigt werden können. Analog der in (6) - (8) skizzierten Vorgehensweise müssen diese vorgegebenen Werte entsprechend transformiert von der zu erklärenden Variable „SB“ abgezogen werden. Neben dieser modifizierten erklärenden Variablen „SB_off“ wird auch die modifizierte GewichtungsvARIABLE „JE_off“ benötigt. Mit ihr wird sichergestellt, dass in PROC GENMOD die sich aus der unmodifizierten Variable „SB“ ergebene Varianzfunktion berücksichtigt wird.

Beim Aufruf von PROC GEMOD sind entsprechend die modifizierten Variablen „SB_off“ und „JE_off“ zu berücksichtigen und die Variable „SFR“ aus der MODEL-Anweisung zu streichen.

```
/******  
/**      Offset für das Merkmal SFR      **/  
/******  
data KSFE_neu;  
set KSFE_daten;  
  if SFR = 'S1' then off= 0.2230;  
  if SFR = 'S2' then off= 0.1319;  
  if SFR = 'S3' then off= -0.1404;  
  if SFR = 'S4' then off= -0.0168;  
  if SFR = 'S5' then off= 0;  
  SB_off = sb / exp(off);  
  JE_off = je * exp(off);  
run;  
proc genmod data=KSFE_neu;  
  class SFR KW REGIO ;  
  model SB_off = KW REGIO  
        /dist=p link=log dscale;  
  weight JE_off;  
run;
```

Beispiel 2: Manuelle Berücksichtigung von a-priori-Informationen

Wie erwartet entsprechen die in Tabelle 2 zusammengefassten Schätzer für die verbleibenden Merkmale „REGIO“ und „KW“ exakt denen aus dem Referenzmodell. Die Unterschiede in den Varianz-Schätzern sind ausschließlich auf den durch den Gewinn von 4 Freiheitsgraden durch die Ein-

sparung des Merkmals „SFR“ geänderten Schätzer für den Dispersionsparameter ϕ zurückzuführen.

Analysis Of Parameter Estimates				
Parameter		DF	Estimate	StdErr
Intercept		1	5.17142	0.083993
KW	K1	1	-0.40249	0.035301
KW	K2	1	-0.25859	0.036867
KW	K3	0	0	0
REGIO	R1	1	0.675963	0.08012
REGIO	R2	1	0.587116	0.08033
REGIO	R3	1	0.325549	0.09113
REGIO	R4	0	0	0
Scale		0	18.12065	

Tabelle 2: Parameterschätzer für Beispiel 2

Diese Vorgehensweise ist umständlich, da die Modell-Modifikationen entsprechend der jeweils gewählten Link- und Varianzfunktion angepasst werden müssen. Mit der OFFSET-Option in der MODEL-Anweisung von PROC GENMOD ist hingegen eine elegante und bequeme Vorgabe von a-priori-Informationen möglich.

7 Die OFFSET-Option in PROC GENMOD

In der Dokumentation für PROC GENMOD wird die OFFSET-Option in einem Zusammenhang eingeführt, der nicht automatisch den Rückschluss auf unser Problem zulässt. Das entsprechende Code-Segment im linken Block von Beispiel 3 wurde lediglich auf unsere Nomenklatur angepasst. Er produziert identischen Output wie der rechte Code-Block, der unserem Beispiel 1 entspricht.

<pre> Proc genmod data=KSFE_daten; Class SFR KW REGIO Model c = SFR KW REGIO / dist = poisson link = log offset = ln; run; </pre>	<pre> Proc genmod data=KSFE_daten; Class SFR KW REGIO Model c = SFR KW REGIO / dist = poisson link = log; Weight je; run; </pre>
---	--

Beispiel 3: OFFSET-Option in der SAS-Dokumentation

In dem Beispiel aus der SAS-Dokumentation wird nicht der Schadenbedarf SB , sondern die Summe der Schadenaufwendungen c modelliert, wobei die Anzahl der Jahreseinheiten je als Skalierungsgröße berücksichtigt werden muss. Das mit (6) bzw. Beispiel 1 vergleichbare Referenzmodell

$$SB_t = \left[\frac{c_t}{je_t} \right] = \exp(\mathbf{x}_t' \boldsymbol{\beta}) \quad (9)$$

wird über folgendem Zwischenschritt

$$\ln \left[\frac{c_t}{je_t} \right] = (\mathbf{x}_t' \boldsymbol{\beta}) \quad (10)$$

so umcodiert, dass die Gewichtungvariable je_t aus dem Nenner der Response-Variable auf die Seite der erklärenden Variable verschoben wird:

$$\ln[c_t] = (\mathbf{x}_t' \boldsymbol{\beta}) + \ln(je_t) \quad (11)$$

Diese Darstellung (11) entspricht exakt unserem modifizierten Modell

$$g(\mu_t) = \mathbf{x}_{1t}' \boldsymbol{\beta}_1 + \mathbf{x}_{2t}' \boldsymbol{\beta}_2 \quad (12)$$

aus (7). Der zweite Term auf der Seite der erklärenden Variablen aus (11) wird mit der OFFSET-Option vorgegeben. Mit der OFFSET-Option können also a-priori-Informationen auf dem Skalenniveau des linearen Prädiktors berücksichtigt werden.

8 Anwendungsbeispiele für die OFFSET-Option

Die a-priori-Informationen aus Beispiel 2 können über die OFFSET-Option im MODEL-Befehl wesentlich eleganter berücksichtigt werden. Die Offset-

Informationen in der Variable „off“ werden nur mit der OFFSET-Option eingeführt. Der Output entspricht exakt dem aus Beispiel 2 (vgl. Tab. 2).

```

/*****
/**      Offset für das Merkmal SFR      **/
/*****

data KSFE_neu;
set KSFE_daten;
  if SFR = 'S1' then off= 0.2230;
  if SFR = 'S2' then off= 0.1319;
  if SFR = 'S3' then off= -0.1404;
  if SFR = 'S4' then off= -0.0168;
  if SFR = 'S5' then off= 0;
run;

proc genmod data=KSFE_neu;
  class SFR KW REGIO ;
  model SB = KW REGIO
    /dist=p link=log dscale offset=off;
  weight JE;
run;

```

Beispiel 4: Beispiel 2 mit OFFSET-Option

Das Merkmal „SFR“ darf nicht mehr als erklärende Variable mitgeführt werden. Da die a-priori-Informationen auf dem Skalenniveau des linearen Prädiktors vorgegeben werden, zieht –entgegen Beispiel 2– eine Änderung der Varianz- und Linkfunktion keine weiteren Anpassung der orgaben nach sich.

Alternativ können auch a-priori-Informationen nicht für alle, sondern nur für eine Ausprägung des Klassifikationsmerkmals „SFR“ vorgegeben werden. Zu berücksichtigen ist, dass für diese Ausprägung (hier „S1“) kein Parameter geschätzt werden darf. Deshalb wird die Ausprägung von „S1“ auf die durch die Cornerpoint-Parametrisierung gegebene Referenzkategorie „S5“ umcodiert.

```

/*****
/*      Offset für SFR S1      */
*****/
data KSFE_daten_neu;
set KSFE_daten;
  if SFR = 'S1' then do
    off = 0.123;
    SFR = 'S5';
  end;
run;
proc genmod data=KSFE_neu;
  class SFR KW REGIO ;
  model SB = SFR KW REGIO
    /dist=p link=log dscale offset=off;
  weight JE;
run;

```

Beispiel 5: OFFSET für nur eine Ausprägung eines Merkmals

Die geschätzten Parameter aus dem Beispiel 5 sind in Tab. 3 zusammengefasst. Wie erwartet wird für die Ausprägung „S1“ kein Parameter geschätzt. Da der Offset-Wert für „S1“ mit 0,123 nicht dem entsprechenden Schätzer aus dem Referenzmodell entspricht, verschieben sich die Parameter für die andern Merkmale leicht.

Analysis Of Parameter Estimates				
Parameter		DF	Estimate	StdErr
Intercept		1	5.1831	0.0909
SFR	S2	1	0.1166	0.0334
SFR	S3	1	-0.1556	0.0442
SFR	S4	1	-0.0320	0.0339
SFR	S5	0	0	0
KW	K1	1	-0.4037	0.0377
KW	K2	1	-0.2606	0.0392
KW	K3	0	0	0
REGIO	R1	1	0.6793	0.0852
REGIO	R2	1	0.5935	0.0854
REGIO	R3	1	0.3288	0.0969
REGIO	R4	0	0	0
Scale		0	19.2568	0

Tabelle 3: Parameterschätzer für Beispiel 5

Ohne die Umcodierung von „S1“ auf die Referenzkategorie „S5“ würde für den Parameter S1 ein Wert von 0,100 –also die Differenz zwischen Offset-Wert 0,1230 und Schätzwert ohne Offset von 0,2230– geschätzt werden. Alle andern Parameter würden wiederum exakt mit denen aus dem Referenzmodell zusammenfallen. In diesem Fall wäre nur in die Parametrisierung und nicht –wie eigentlich beabsichtigt- in die Modellstruktur eingegriffen worden.

<pre> /***** /* Schritt 1 */ /* Modell mit KW u REGIO */ /***** proc genmod data=KSFE_daten; class sfr kw regio; model sb = kw regio / dist=p link=log obstats; make 'obstats' out=obs_1; weight je; run; data KSFE_neu (rename=(xbeta=off)); merge KSFE_daten obs_1(keep=xbeta); run; </pre>	<pre> /***** /* Schritt 2 */ /* Nur SFR, Model l= offset */ /***** proc genmod data=KSFE_neu; class sfr kw regio; model sb = sfr / dist=p link=log obstats offset=off; make 'obstats' out=obs_2; weight je; run; </pre>
--	--

Beispiel 6: Beliebiges OFFSET für y

Die Offset-Werte müssen sich nicht immer auf einzelne Parameter (-gruppen) beziehen. Es können z.B. auch Offset-Werte direkt für die abhängige Variable y vorgegeben werden. In Beispiel 6 wird in Schritt 1 ein Modell nur mit den beiden Klassifikationsmerkmalen „KW“ und „REGIO“ geschätzt. Über die ODS-Optionen werden die Schätzer „XBETA“ für den linearen Prädiktor jeder einzelnen Beobachtung wieder an den Datensatz „KSFE_daten“ angespielt.

In Schritt 2 wird ein Modell nur mit dem dritten Klassifikationsmerkmal „SFR“ geschätzt, wobei der lineare Prädiktor aus Schritt 1 als Offset vorge-

geben wird. In Tab. 4 sind die Parameterschätzer für beide Modelle aus Beispiel 6 angegeben.

Parameter Estimates			Parameter Estimates		
Parameter	DF	Estimate	Parameter	DF	Estimate
Intercept	1	5.1867	Intercept	1	-0.0227
KW K1	1	-0.4053	SFR S1	1	0.2213
KW K2	1	-0.2619	SFR S2	1	0.1311
KW K3	0	0	SFR S3	1	-0.1409
REGIO R1	1	0.6826	SFR S4	1	-0.0173
REGIO R2	1	0.6035	SFR S5	0	0
REGIO R3	1	0.3335	Scale	0	17.9770
REGIO R4	0	0			
Scale	0	26.51660			

Tabelle 4: Parameterschätzer für Beispiel 6

Der zweite Schritt aus Beispiel 6 kann als GLM für die Residuen aus Schritt 1 interpretiert werden. Diese Art der Residuenanalyse eignet sich hervorragend, um über Indikatorvariablen –wie im Beispiel 6 das Merkmal „SFR“-systematische und noch nicht im Modell berücksichtigte Einflüsse anschaulich aufzuzeigen. Durch die Fixierung der Parameterschätzer aus Schritt 1 über die OFFSET-Option weisen die Parameter aus Schritt 2 nur den isolierten zusätzlichen Einfluss aus.

Im Gegensatz zu dem üblichen Ansatz, die Indikatorvariable ohne Offset zusätzlich in das Basismodell aufzunehmen, beeinträchtigen Korrelationen bzw. Wechselwirkungen zwischen dem Basismodell und der Indikatorvariable die Ergebnisse bzw. die Schätzer für die Indikatorvariable nicht. Weiterhin sind diese Ergebnisse kompakter -und auch anschaulicher zu interpretieren- als z.B. Residuen-Plots.

9 Schlussbemerkung

Im Bereich der generalisierten linearen Modelle stellt PROC GENMOD ein sehr vielseitiges und leistungsfähiges Werkzeug dar. Verschiedene Facetten zur Berücksichtigung von a-priori-Informationen wurden theoretisch und in der Umsetzung mit Hilfe der OFFSET-Option in PROC GENMOD aufgezeigt.

Über die Berücksichtigung von a-priori-Informationen kann auch der Einfluss einzelner Variablen in einem Modell isoliert werden. Aus dieser Überlegung leiten sich anschauliche heuristische Ansätze zur Modellauswahl und Residuenanalyse ab, auf die kurz eingegangen wurde.

Literatur

- [1] L.A. Baxter, S.M. Coutts & G.A.F. Ross, (1980): Applications of Linear Models, TICA 21, S. 11-29
- [2] O. Kruse (1997): Modelle zur Analyse und Prognose des Schadenbedarfs in der Kfz-Haftpflichtversicherung, Karlsruhe: Verlag Versicherungswirtschaft
- [3] P. McCullagh & J.A Nelder (1989): Generalized Linear Models (2. ed.), London New York: Chapman & Hall.
- [4] J.A. Nelder & R.M.W. Wedderburn (1972): Generalized Linear Models, Journal of the Royal Statistical Society A, 1972, S.370-384
- [5] SAS Institute Inc. (1993): SAS Technical Report P-243: The GENMOD Procedure, Cary: SAS Institute Inc.

