

Schätzer und Test für den Schadenparameter in Krankenversicherungstarifen mit Selbstbeteiligung

Autor(en): **Bruchlos, Kai**

Objektyp: **Article**

Zeitschrift: **Mitteilungen / Schweizerische Aktuarvereinigung = Bulletin / Association Suisse des Actuaires = Bulletin / Swiss Association of Actuaries**

Band (Jahr): - **(2005)**

Heft 1

PDF erstellt am: **22.07.2024**

Persistenter Link: <https://doi.org/10.5169/seals-967315>

Nutzungsbedingungen

Die ETH-Bibliothek ist Anbieterin der digitalisierten Zeitschriften. Sie besitzt keine Urheberrechte an den Inhalten der Zeitschriften. Die Rechte liegen in der Regel bei den Herausgebern.

Die auf der Plattform e-periodica veröffentlichten Dokumente stehen für nicht-kommerzielle Zwecke in Lehre und Forschung sowie für die private Nutzung frei zur Verfügung. Einzelne Dateien oder Ausdrucke aus diesem Angebot können zusammen mit diesen Nutzungsbedingungen und den korrekten Herkunftsbezeichnungen weitergegeben werden.

Das Veröffentlichen von Bildern in Print- und Online-Publikationen ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. Die systematische Speicherung von Teilen des elektronischen Angebots auf anderen Servern bedarf ebenfalls des schriftlichen Einverständnisses der Rechteinhaber.

Haftungsausschluss

Alle Angaben erfolgen ohne Gewähr für Vollständigkeit oder Richtigkeit. Es wird keine Haftung übernommen für Schäden durch die Verwendung von Informationen aus diesem Online-Angebot oder durch das Fehlen von Informationen. Dies gilt auch für Inhalte Dritter, die über dieses Angebot zugänglich sind.

B. Wissenschaftliche Mitteilungen

KAI BRUCHLOS, Hamburg

Schätzer und Test für den Schadenparameter in Krankenversicherungstarifen mit Selbstbeteiligung

1 Begriff des Schadenparameters

Innerhalb eines Tarifes der Krankenversicherung, der auch eine Selbstbeteiligung – proportionale wie nicht proportionale Risikoteilung – des Versicherungsnehmers vorsieht, wird immer wieder beobachtet, daß sich der Kopfschaden nicht entsprechend der Selbstbeteiligung verändert. Beispielsweise gilt in der Regel für die 50%-Erstattungsstufe eines ambulanten Krankheitskostenvollversicherungstarifes nicht, daß der Kopfschaden die Hälfte der 100%-Erstattungsstufe ist, sondern weniger.

Eine derartige Änderung der Proportionalität der Kopfschäden liegt nicht nur bei Selbstbeteiligungen vor, sondern auch beim sogenannten *Pelzmanteleffekt* der Krankenhaustagegeldversicherung: Hat ein Versicherungsnehmer ein Tagegeld in Höhe von 5 Euro abgeschlossen, so wird er nicht länger als nötig im Krankenhaus bleiben. Bei einem Tagegeld von 100 Euro sieht dies schon ganz anders aus.

Derartige Effekte, die sich auf das Verhältnis von Kopfschäden auswirken, werden bei der Tarifikalkulation durch den sogenannten *Schadenparameter* berücksichtigt (Tosberg (1940), S. 21, Brünjes (1985), S. 180, Farny u.a. (1988), S. 402). Dabei stellt sich die Frage, wann und in welcher Größenordnung der Schadenparameter vorliegt, also ab welcher Selbstbeteiligung oder ab welcher Tagegeldhöhe in welcher Höhe er zu veranschlagen ist. Diese Frage soll im folgenden mit Hilfe der Regressionsanalyse beispielhaft für die Krankenhaustagegeldversicherung beantwortet werden.

Verhalten sich die Kopfschäden zweier verschiedener Leistungsstufen eines Tarifes entsprechend dem Verhältnis der Leistungen und werden die Kopfschäden der jeweiligen Tarifkombinationen paarweise zusammengefaßt, so bilden die Paare eine Gerade. Werden zum Beispiel die rohen Kopfschäden für Männer der 5 Euro Stufe und der 50 Euro Stufe eines Krankenhaustagegeldtarifes paarweise nach Altersgruppen zusammengefaßt, so sollten sie um eine Ursprungsgerade mit der Steigung 10 herum streuen. Um dies zu testen, bietet sich der t-Test der einfachen normalen linearen Regression mit Kleinsten-Quadrate-Schätzern an. Bei genauerer Betrachtung der Voraussetzungen dieser Regression stellt sich allerdings heraus, daß die Kopfschäden der 5 Euro Stufe fest vorzugeben sind und fehlerfrei

gemessen sein sollen, wohingegen die der 50 Euro Stufe stochastisch sind und mit Zufallsfehlern behaftet. Diese Asymmetrie des Modells kann nicht bestehen bleiben. Des Weiteren können die Kopfschäden der 5 Euro Stufe und ebenso der 50 Euro Stufe nicht fest vorgegeben und nur mit Zufallsfehlern beobachtet werden. Außerdem sollte bei der Vertauschung der 5 Euro Stufe mit der 50 Euro Stufe der t-Test für die Steigung $1/10$ dasselbe Ergebnis liefern wie bei der Steigung 10. Dem ist im allgemeinen nicht so (Schach und Schäfer, S.160f).

Dieser Fehler der 3. Art – die Asymmetrie des Modelles der einfachen normalen linearen Regression liegt in der betrachteten Anwendung nicht vor – ist im Fehler-in-den-Variablen Modell behoben, da beide Variablen symmetrisch behandelt werden. Da bei Krankenhaustagegeldtarifen das Verhältnis der Fehlervarianzen der rohen Kopfschäden, λ , bekannt ist und eine Ursprungsgerade vorliegt, kann zumindest für den Maximum-Likelihood-Schätzer nicht das oft verwendete strukturelle Modell herangezogen werden (Creasy (1956), S. 69). Für den Schadenparameter wird deshalb auf das funktionale Modell zurückgegriffen, welches in Abschnitt 2.2 dargestellt wird, nachdem in Abschnitt 2.1 das wahrscheinlichkeitstheoretische Modell der Krankenhaustagegeldversicherung dargelegt worden ist. Daß es einen konsistenten Maximum-Likelihood-Schätzer im funktionalen Modell für den Schadenparameter gibt, wird in Abschnitt 3 dargelegt. Bevor in Abschnitt 5 die Ergebnisse auf zwei Stufen eines Krankenhaustagegeldtarifes angewendet werden, wird in Abschnitt 4 ein Test für den Schadenparameter konstruiert.

2 Stochastisches Modell

Die Krankenhaustagegeldversicherung wird kurz mit *KHT-Versicherung* bezeichnet, *ML-Schätzer* ist die Abkürzung für Maximum-Likelihood-Schätzer. Für eine Teststatistik $T(X)$ wird die Konvention $T(x) := T(X)(x)$ für den Funktionswert der Teststatistik bei vorliegender Realisation x der Zufallsvariable X verwendet.

Grundsätzlich sei hier jede Zufallsvariable eine nicht konstante Funktion, damit bei Existenz der Varianz diese größer Null ist. Außerdem existiere zu jeder Zufallsvariable der Erwartungswert.

2.1 Wahrscheinlichkeitstheoretisches Modell der Krankenhaustagegeldversicherung

Definition 2.1: Die Zufallsvariable $X: \Omega_1 \rightarrow \mathbf{W}$, wobei \mathbf{W} eine Teilmenge von \mathbf{R}_0^+ ist, heißt *Risiko*.

Gegeben sei ein bestimmter Versicherungsvertrag j innerhalb eines Beobachtungszeitraumes (Versicherungsperiode).

Definition 2.2: Die Zufallsvariable $N: \Omega_1 \rightarrow \mathbf{N}_0$ heißt *Schadenzahl* und deren zugehörige Verteilung *Schadenzahlverteilung*. Die Zufallsvariablen $\tilde{S}_i: \Omega_1 \rightarrow \mathbf{R}^+$, $i = 1, 2, \dots, N$ heißen *Schadenhöhen* und deren zugehörige Verteilungen *Schadenhöhenverteilung*.

Die Zufallsvariable N ist zu verstehen als die Anzahl der Schäden innerhalb einer Versicherungsperiode, die sich auf einen Versicherungsvertrag beziehen, die Zufallsvariable \tilde{S}_i als die Versicherungsleistung, die beim i -ten Schaden zu zahlen ist. Die Zufallsvariablen N und \tilde{S}_i sind Risiken. Es soll nun die Gesamtversicherungsleistung in einer Periode bezüglich eines Versicherungsvertrages j angegeben werden, also das *Einzelrisiko*

$$X_j := \begin{cases} 0, & \text{falls } N = 0 \\ \tilde{S}_1 + \tilde{S}_2 + \dots + \tilde{S}_N, & \text{falls } N > 0 \end{cases}.$$

Eingeführt sei noch die Bezeichnung $\tilde{S}_{j,i} := \tilde{S}_i$ für die i -te Schadenhöhe des j -ten Einzelrisikos und die Zufallsvariable N_j als Schadenzahl des j -ten Einzelrisikos X_j .

Sei ein Gesamtbestand von L KHT-Versicherungsverträgen (Einzelrisiken) eines bestimmten Geschlechts und eines Beobachtungszeitraumes mit der Tagegeldhöhe τ gegeben. Sei weiter $Y_{x,j} := X_j$ für $j = 1, 2, \dots, L_x$ die gesamte Versicherungsleistung (Schadenssumme) eines x -jährigen Versicherungsnehmers aus diesem Gesamtbestand, wobei L_x die Anzahl des Teilbestandes der x -jährigen Versicherungsnehmer vom Gesamtbestandes ist, d.h. $L = \sum_x L_x$, und

$$Y_x := \sum_{i=1}^{L_x} Y_{x,i} = \sum_{i=1}^{L_x} X_i.$$

Definition 2.3: Y_x heißt *Gesamtleistung einer Periode der x -Jährigen*, $Y_{x,j} \cdot \tau^{-1}$ *Kopfschaden*, $E(Y_{x,j} \cdot \tau^{-1})$ *erwarteter Kopfschaden* und $\sigma(Y_{x,j} \cdot \tau^{-1})$ *Kopfschadenstreuung des Einzelrisikos j* für $j = 1, 2, \dots, L_x$.

Bei der Kalkulation von Prämien für Krankenversicherungstarife wird zwischen Einzelaltergruppe, das sind alle Einzelrisiken eines Alters, und Altersgruppe, das ist die Zusammenfassung von in der Regel fünf Einzelaltergruppen, unterschieden (Bohn (1980), S. 134). Hier soll unter dem Begriff „Teilbestand der x -jährigen Versicherungsnehmer“ beides verstanden werden können. Dies bedeutet, daß

die nachfolgenden Ergebnisse sowohl für Einzelaltergruppen gelten als auch für Altersgruppen. Um beide Fälle zu erfassen, wird im folgenden von der Altersklasse x gesprochen.

Für den Erwartungswert der Gesamtleistung gilt:

$$E(Y_x) = \sum_{i=1}^{L_x} E(Y_{x,i}).$$

Es werden folgende Annahmen über die Bestandsstruktur getroffen:

1. Homogenität: Es gilt $E(Y_{x,i}) = E(Y_{x,j})$ und $\text{VAR}(Y_{x,i}) = \text{VAR}(Y_{x,j})$ für alle $i, j = 1, 2, \dots, L_x$
2. Stochastische Unabhängigkeit der einzelnen Versicherungsleistungen:
 $\tilde{S}_{j,i}$ sind paarweise stochastisch unabhängig für $j = 1, 2, \dots, L$,
 $i = 1, 2, \dots, N_j$. Daraus folgt, daß $Y_{x,1}, Y_{x,2}, \dots, Y_{x,L_x}$ stochastisch unabhängig sind.

Annahme 1 ermöglicht die Bezeichnung $K_x^\tau := E(Y_{x,j} \cdot \tau^{-1})$ für alle $j = 1, 2, \dots, L_x$.

Mit den beiden Annahmen gilt

$$\text{VAR}(Y_x) = \sum_{i=1}^{L_x} \text{VAR}(Y_{x,i}) = L_x \cdot \text{VAR}(Y_{x,j})$$

für alle $j \in \{1, 2, \dots, L_x\}$, woraus

$$\sigma(Y_x \cdot \tau^{-1}) = \sqrt{\text{VAR}(Y_x \cdot \tau^{-1})} = \sqrt{L_x} \cdot \sigma(Y_{x,j} \cdot \tau^{-1}) = \sqrt{L_x} \cdot \sigma_{x,j}^\tau$$

folgt. Mit Annahme 1 gilt für alle $j \in \{1, 2, \dots, L_x\}$:

$$E(Y_x) = \sum_{i=1}^{L_x} E(Y_{x,i}) = L_x \cdot E(Y_{x,j}).$$

Als Punktschätzer für den erwarteten Kopfschaden $K_x^\tau = E(Y_{x,j} \cdot \tau^{-1})$ bietet sich somit das folgende arithmetische Mittel an (Bohn (1980), S. 13, S. 23, Brünjes (1985), S. 180):

$$\hat{K}_x^\tau := \frac{T_x}{L_x}. \quad (2.4)$$

Hierbei ist T_x die Zufallsvariable der beobachteten gesamten Anzahl an Krankenhaustagen des Teilbestandes der x -jährigen Versicherungsnehmer. \widehat{K}_x^τ heißt *roher Kopfschaden der Altersklasse x* und ist die durchschnittliche Anzahl an Krankenhaustagen des Teilbestandes.

In dem Gesamtbestand gebe es z Altersklassen x_1, x_2, \dots, x_z . Demgemäß bezeichne $\widehat{K}_{x_i}^\tau, K_{x_i}^\tau$ den rohen Kopfschaden der Altersklasse x_i und der Tagegeldhöhe τ bzw. den erwarteten Kopfschaden des j -ten Einzelrisikos der Altersklasse x_i und der Tagegeldhöhe τ . Sei für ein fest gewähltes $l \in \{1, 2, \dots, z\}$

$$k_{x_i}^\tau := \frac{K_{x_i}^\tau}{K_{x_l}^\tau}$$

der *Profilwert* für $i = 1, 2, \dots, z$ und $(k_{x_1}^\tau, k_{x_2}^\tau, \dots, k_{x_z}^\tau)$ das *Profil*.

Betrachten wir nun die Tagegeldhöhen τ_1, τ_2 mit den zugehörigen Beständen L^{τ_1}, L^{τ_2} , wobei $\tau_1 < \tau_2$ ist. Sei die Notation um τ_1, τ_2 entsprechend erweitert. Gibt es eine möglicherweise geschlechtsabhängige, aber altersunabhängige Konstante $P_{x(y)}^{\tau_1} \geq 1$ mit

$$K_{x_i}^{\tau_2} = P_{x(y)}^{\tau_1} \cdot K_{x_i}^{\tau_1} \tag{2.5}$$

für alle $i = 1, 2, \dots, z$, dann heißt die Konstante $P_{x(y)}^{\tau_1}$ *Schadenparameter* (Bohn (1980), S. 23, Brünjes (1985), S. 180). Die Existenz des Schadenparameters hat zur Folge, daß das Profil des Tagegeldes τ_1 gleich dem Profil des Tagegeldes τ_2 ist, also $k_{x_i}^{\tau_1} = k_{x_i}^{\tau_2}$ für $i = 1, 2, \dots, z$. Geometrisch betrachtet ist der Schadenparameter die Steigung einer Ursprungsgeraden in der $(K_{x_i}^{\tau_1}, K_{x_i}^{\tau_2})$ -Ebene.

2.2 Statistisches Modell für den Schadenparameter

Sei $(\Omega_2, \mathcal{A}, P)$ ein Wahrscheinlichkeitsraum, \mathbf{B}^n die Borel σ -Algebra über \mathbf{R}^n , Θ eine Menge von Verteilungsparametern, $\mathcal{P} := \{P_\vartheta := (P_{\vartheta,1}, P_{\vartheta,2}, \dots, P_{\vartheta,n}) \mid \vartheta \in \Theta\}$ eine Familie von Wahrscheinlichkeitsmaßen auf dem Meßraum $(\mathcal{X}, \mathcal{B}) := (\mathbf{R}^n, \mathbf{B}^n)$ mit der Eigenschaft $P_{\vartheta,1} = P_{\vartheta,2} = \dots = P_{\vartheta,n}$ und $\mathcal{Z} := (Z_1, Z_2, \dots, Z_n) : \Omega_2 \rightarrow \mathbf{R}^n$ eine Zufallsvariable, deren Verteilung unter P in \mathcal{P} liegt. Die Zufallsvariablen Z_1, Z_2, \dots, Z_n seien unabhängig und identisch verteilt. Für den Test betrachten wir somit das *statistische Modell* $(\mathbf{R}^n, \mathbf{B}^n, \mathcal{P})$.

Sei z_1, z_2, \dots, z_n eine Zufallsstichprobe aus der zu betrachtenden Grundgesamtheit, wobei die z_i Realisationen der Zufallsvariable $Z : \Omega_2 \rightarrow \mathbf{R}$ sind mit dem Wahrscheinlichkeitsmaß $P_{\tilde{\vartheta}} \in \mathcal{P}$. Z_1, Z_2, \dots, Z_n sollen dieselbe Verteilung wie Z haben.

In \mathcal{P} wird eine nicht leere, echte Teilmenge H_0 festgelegt und das Komplement von H_0 mit $H_1 := \mathcal{P} \setminus H_0$ bezeichnet. Ein statistischer Test ist nun eine Entscheidungsregel, die zu jeder Stichprobe $(z_1, z_2, \dots, z_n) \in \mathbf{R}^n$ festlegt, ob man sich für die *Nullhypothese* $P_{\tilde{\delta}} \in H_0$ oder die *Alternativhypothese* $P_{\tilde{\delta}} \in H_1$ entscheiden soll, formal:

Definition 2.6: Die Abbildung

$$\delta: \mathcal{X} \rightarrow \{0, 1\},$$

wobei $\delta(v) = 1$ die Ablehnung der Nullhypothese bedeutet und $\delta(v) = 0$ die Nicht-Ablehnung der Nullhypothese, heißt *statistischer Test*.

Seien zwei metrisch skalierte Merkmale K und Ξ aus einer Grundgesamtheit gegeben, deren Merkmalswerte $\tilde{\kappa}_i$ und $\tilde{\xi}_i$ Bilder der Funktionen κ_i und ξ_i für $i = 1, 2, \dots, n$ sind. Da gemäß der allgemeinen Voraussetzung in Abschnitt 2 eine Zufallsvariable keine konstante Funktion sein soll, wird hier festgelegt: Die Funktionen κ_i und ξ_i sind entweder konstante Funktionen oder Zufallsvariablen und es gelte der Zusammenhang

$$\xi_i = \alpha + \beta \cdot \kappa_i$$

für $i = 1, 2, \dots, n$ mit zwei Konstanten $\alpha, \beta \in \mathbf{R}$. Die Merkmalswerte $\tilde{\kappa}_i$ und $\tilde{\xi}_i$ können aber nicht direkt beobachtet werden, sondern nur

$$v_i := \tilde{\kappa}_i + d_i \quad \text{und} \quad w_i := \tilde{\xi}_i + e_i$$

für $i = 1, 2, \dots, n$, d.h. $\tilde{\kappa}_i$ und $\tilde{\xi}_i$ unterliegen jeweils einem zufälligen Beobachtungsfehler d_i und e_i . Sei e_i eine Realisation der Zufallsvariablen E_i und d_i eine der Zufallsvariablen D_i für $i = 1, 2, \dots, n$. $D_1, D_2, \dots, D_n, E_1, E_2, \dots, E_n$ seien stochastisch unabhängig, D_1, D_2, \dots, D_n sowie E_1, E_2, \dots, E_n jeweils identisch verteilt mit Erwartungswert Null und endlicher Varianz. Insbesondere gilt also:

$$\begin{aligned} E(D_i) &= E(E_i) = 0 && \text{für } i = 1, 2, \dots, n \\ \text{Var}(E_i) &= \text{Var}(E_j) && \text{für } i, j \in \{1, 2, \dots, n\}, \sigma_E^2 := \text{Var}(E_i) \\ \text{Var}(D_i) &= \text{Var}(D_j) && \text{für } i, j \in \{1, 2, \dots, n\}, \sigma_D^2 := \text{Var}(D_i) \\ \text{Cov}(D_i, D_j) &= \text{Cov}(E_i, E_j) = 0 && \text{für } i \neq j, \\ \text{Cov}(D_i, E_j) &= 0 && \text{für } i, j \in \{1, 2, \dots, n\} \end{aligned}$$

Führen wir jetzt noch zu den Realisationen v_i und w_i für $i = 1, 2, \dots, n$ die Zufallsvariablen $V_i = \kappa_i + D_i$ und $W_i = \xi_i + E_i$ ein, dann heißt dieses Modell (Schach und Schäfer (1978), S. 152, Kendall und Stuart (1979), S. 400, Miller (1996), S. 202f)

Fehler-in-den-Variablen Modell. (2.7)

Üblich ist die folgende empirische Darstellung (Schach und Schäfer (1978), S. 152, Fuller (1987), S. 30, Miller (1996), S. 202f):

$$\begin{aligned}
 v_i &= \tilde{\kappa}_i + d_i, & w_i &= \tilde{\xi}_i + e_i && \text{Eigenschaft der beobachteten Werte} \\
 w_i &= \alpha + \beta \cdot \tilde{\kappa}_i + e_i && && \text{Eigenschaft der Stichprobenvariablen} \\
 &= \alpha + \beta \cdot v_i - \beta \cdot d_i + e_i && && \\
 \hat{w}_i &= \hat{\alpha} + \hat{\beta} \cdot v_i && && \text{Regressionsgerade} \\
 \xi &= \alpha + \beta \cdot \kappa && && \text{zugrunde liegende lineare Beziehung} \\
 \kappa &= \alpha' + \beta' \cdot \xi \text{ mit } \alpha' := -\alpha/\beta, \text{ Symmetrie} && && \\
 &&&&&& \beta' := 1/\beta
 \end{aligned}$$

Sind $\kappa_1, \kappa_2, \dots, \kappa_n$ konstante Funktionen und D_i, E_i normalverteilt für $i = 1, 2, \dots, n$, und habe die Verteilung der Grundgesamtheit, aus der die $\tilde{\kappa}_i$ stammen, eine endliche Varianz größer Null und existieren die vierten Momente dieser Verteilung, dann heißt das Modell (Schach und Schäfer (1978), S. 163)

funktionales (Fehler-in-den-Variablen) Modell. (2.8)

V_1, V_2, \dots, V_n sowie W_1, W_2, \dots, W_n sind jeweils stochastisch unabhängig und es gilt für $i = 1, 2, \dots, n$

$$V_i \sim N(\tilde{\kappa}_i, \sigma_D^2) \quad \text{und} \quad W_i \sim N(\alpha + \beta \cdot \tilde{\kappa}_i, \sigma_E^2), \quad (2.9)$$

insbesondere

$$E(W_i) = \alpha + \beta \cdot E(V_i).$$

Der Zufallsvektor $((V_1, W_1), (V_2, W_2), \dots, (V_n, W_n))^T$ ist multivariat normalverteilt. Der Erwartungsvektor besteht aus den $2 \cdot n$ Parametern $E(V_1), E(W_1), E(V_2), E(W_2), \dots, E(V_n), E(W_n)$.

Sei im funktionalen Modell (2.8) $V_i := \widehat{K}_{x_i}^{\tau_1}$, $W_i := \widehat{K}_{x_i}^{\tau_2}$, $\widetilde{\kappa}_i := K_{x_i}^{\tau_1}$, $\widetilde{\xi}_i := K_{x_i}^{\tau_2}$, $G_i := E_i - \beta \cdot D_i$ und $\widetilde{H}_i := a_1 \cdot W_i + a_2 \cdot V_i$ mit $a_1, a_2 \in \mathbf{R} \setminus \{0\}$ für $i = 1, 2, \dots, z$. Es gelte $\sigma_E^2 = \sigma_D^2$ und $K_{x_i}^{\tau_2} = P_{x(y)}^{\tau_1} \cdot K_{x_i}^{\tau_1}$ mit $P_{x(y)}^{\tau_1} \geq 1$ – siehe (2.5).

$\widehat{K}_{x_i}^{\tau_1}$ bzw. $\widehat{K}_{x_i}^{\tau_2}$ sind erwartungstreue Schätzer der erwarteten Kopfschäden $K_{x_i}^{\tau_1}$ bzw. $K_{x_i}^{\tau_2}$ und es gilt

$$E(W_i) = \beta \cdot E(V_i) = P_{x(y)}^{\tau_1} \cdot K_{x_i}^{\tau_1} = K_{x_i}^{\tau_2}.$$

Sei $\lambda := \sigma_E^2 / \sigma_D^2$ und

$$\widehat{\beta}^* := \frac{\sum_{i=1}^z W_i^2 - \lambda \cdot \sum_{i=1}^z V_i^2 + \sqrt{\left(\sum_{i=1}^z W_i^2 - \lambda \cdot \sum_{i=1}^z V_i^2\right)^2 + 4 \cdot \lambda \cdot \left(\sum_{i=1}^z V_i \cdot W_i\right)^2}}{2 \cdot \sum_{i=1}^z V_i \cdot W_i} \quad (2.10)$$

für $\sum_{i=1}^z V_i \cdot W_i \neq 0$ und sonst 0. $\widehat{\beta}^*$ ist der konsistente ML-Schätzer für $P_{x(y)}^{\tau_1}$

(Lemma 3.1, (iii) und Lemma 3.2).

Sei F^* die Verteilungsfunktion von $\widehat{\beta}^*$ mit den Parametern

$$\vartheta := (\beta, \vartheta_2, \vartheta_3, \dots, \vartheta_m),$$

$$F_{\vartheta;1}^* := F_{\vartheta;2}^* := \dots := F_{\vartheta;2;z}^* := F^*,$$

$$\mathcal{P}^{(1)} := \{F_{\vartheta}^* := (F_{\vartheta;1}^*, F_{\vartheta;2}^*, \dots, F_{\vartheta;2;z}^*) \mid \vartheta = (\beta, \vartheta_2, \dots, \vartheta_m) \in \Theta\}$$

und

$$H_0^{(1)} := \{F_{\vartheta}^* \mid \vartheta = (\beta_0, \vartheta_2, \vartheta_3, \dots, \vartheta_m)\}.$$

Kurz formuliert testen wir $H_0^{(1)} : \beta = \beta_0$ gegen $H_1^{(1)} : \beta \neq \beta_0$.

Mit der Teststatistik

$$T_1(V_1, V_2, \dots, V_z, W_1, W_2, \dots, W_z) := \frac{\mathcal{R} \cdot \sqrt{z-2}}{\sqrt{1-\mathcal{R}^2}}, \quad (2.11)$$

wobei

$$\mathcal{R} := \frac{\sum_{i=1}^z \mathcal{M}_i \cdot \mathcal{N}_i}{\sqrt{\sum_{i=1}^z \mathcal{M}_i^2 \cdot \sum_{i=1}^z \mathcal{N}_i^2}}$$

mit

$$\mathcal{M}_i := \widehat{K}_{x_i}^{\tau_2} - \left(\frac{1}{z} \cdot \sum_{j=1}^z \widehat{K}_{x_j}^{\tau_2} \right) - \beta_0 \cdot \left(\widehat{K}_{x_i}^{\tau_1} - \frac{1}{z} \cdot \sum_{j=1}^z \widehat{K}_{x_j}^{\tau_1} \right)$$

und

$$\mathcal{N}_i := \beta_0 \cdot \widehat{K}_{x_i}^{\tau_2} + \widehat{K}_{x_i}^{\tau_1} - \frac{1}{z} \cdot \sum_{j=1}^z (\beta_0 \cdot \widehat{K}_{x_j}^{\tau_2} + \widehat{K}_{x_j}^{\tau_1})$$

ist, erhalten wir für $H_0^{(1)} : \beta = \beta_0$ gegen $H_1^{(1)} : \beta \neq \beta_0$ folgenden statistischen Test (Satz 4.1, Satz 4.5, Lemma 4.6):

$$\begin{aligned} \delta_1 : \mathbf{R}^{2 \cdot z} &\rightarrow \{0, 1\}, \\ (v_1, v_2, \dots, v_z, w_1, w_2, \dots, w_z) &\mapsto \begin{cases} 1, & \text{falls } |T_1(v_1, \dots, v_z, w_1, \dots, w_z)| \\ & > t_{z-2; 1-\alpha/2} \\ 0, & \text{falls } |T_1(v_1, \dots, v_z, w_1, \dots, w_z)| \\ & \leq t_{z-2; 1-\alpha/2} \end{cases} \end{aligned} \quad (2.12)$$

3 ML-Schätzer für die Steigung

Zu welchen Ergebnissen führt die ML-Methode zur Schätzung von β im funktionalen Modell (2.8)? Schach und Schäfer (1978), S. 165 sowie Creasy (1956), S. 67 geben zur Existenz folgende Auskunft:

Lemma 3.1:

- (i) *Im funktionalen Modell existiert im allgemeinen kein ML-Schätzer für β .*
- (ii) *Ist im funktionalen Modell zusätzlich nur bekannt, daß $\alpha = 0$ ist, dann existiert im allgemeinen kein ML-Schätzer.*
- (iii) *Ist im funktionalen Modell λ gegeben und des weiteren nur noch $\alpha = 0$, so ist $\widehat{\beta}^*$ der ML-Schätzer für β .*

Mit Blick auf die Likelihood-Funktion (Schach und Schäfer (1978), S. 166) gilt folgende asymptotische Eigenschaft (Cramér (1946), S. 500):

Lemma 3.2: *Ist im funktionalen Modell λ gegeben und des weiteren nur noch $\alpha = 0$, dann ist $\widehat{\beta}^*$ ein konsistenter Schätzer für β .*

4 Test für die Steigung

Ziel dieses Abschnittes ist die Konstruktion des statistischen Tests δ_1 im funktionalen Modell, welche auf der Formulierung einer äquivalenten Nullhypothese zu $H_0^{(1)}$ beruht. Hinsichtlich des strukturellen Modells vergleiche Fuller (1987), S. 45. Ausgangspunkt ist für $H_0^{(1)}$ gegen $H_1^{(1)}$ der Test (2.12)

$$\delta_1 : \mathbf{R}^{2 \cdot z} \rightarrow \{0, 1\},$$

$$(v_1, v_2, \dots, v_z, w_1, w_2, \dots, w_z) \mapsto \delta_1(v_1, v_2, \dots, v_z, w_1, w_2, \dots, w_z)$$

Sei $G : \Omega_3 \rightarrow \mathbf{R}$ bzw. $\tilde{H} : \Omega_3 \rightarrow \mathbf{R}$ eine Zufallsvariable mit den stochastisch unabhängigen und identisch wie G bzw. \tilde{H} verteilten Stichprobenvariablen G_1, G_2, \dots, G_z bzw. $\tilde{H}_1, \tilde{H}_2, \dots, \tilde{H}_z$. Sei

$$R := \frac{\sum_{i=1}^z (G_i - \bar{G}) \cdot (\tilde{H}_i - \bar{H})}{\sqrt{\sum_{i=1}^z (G_i - \bar{G})^2 \cdot \sum_{i=1}^z (\tilde{H}_i - \bar{H})^2}}$$

mit $\bar{G} := \frac{1}{z} \cdot \sum_{i=1}^z G_i$ und $\bar{H} := \frac{1}{z} \cdot \sum_{i=1}^z \tilde{H}_i$ der sogenannte *Stichprobenkorrelationskoeffizient*. Gemäß Fisz (1976), S. 422f gilt der

Satz 4.1: Seien G und \tilde{H} normalverteilt und $f : \mathbf{R} \rightarrow \mathbf{R}$, $r \mapsto f(r)$ die Dichtefunktion der Zufallsvariablen R . Für festes $z \in \mathbf{N}$ gilt dann

$$f(r) = \frac{z-2}{\pi} \cdot (1-\varrho^2)^{\frac{z-1}{2}} \cdot (1-r^2)^{\frac{z-4}{2}} \cdot \int_0^1 \frac{x^{z-2}}{(1-\varrho^{r \cdot x})^{z-1} \cdot \sqrt{1-x^2}} dx,$$

wobei $\varrho := \frac{\text{Cov}(G, \tilde{H})}{\sqrt{\text{Var}(G) \cdot \text{Var}(\tilde{H})}}$ die Korrelation von G und \tilde{H} ist. Ist $\varrho = 0$, so gilt

$$\frac{R \cdot \sqrt{z-2}}{\sqrt{1-R^2}} \sim t_{z-2}.$$

Sei F_ϱ die Verteilungsfunktion von R , $\mathcal{P}^{(2)} := \{(F_\varrho, F_\varrho, \dots, F_\varrho) \mid \varrho \in [-1; 1]\}$ und $H_0^{(2)} := \{(F_\varrho, F_\varrho, \dots, F_\varrho) \mid \varrho = 0\}$. Kurz formuliert testen wir $H_0^{(2)} : \varrho = 0$ gegen $H_1^{(2)} : \varrho \neq 0$.

Als Teststatistik wählen wir

$$T_2(G_1, G_2, \dots, G_z, \tilde{H}_1, \tilde{H}_2, \dots, \tilde{H}_z) := \frac{R \cdot \sqrt{z-2}}{\sqrt{1-R^2}}, \quad (4.2)$$

welche gemäß Satz 4.1 für $\varrho = 0$ t-verteilt ist mit $z-2$ Freiheitsgraden. Wenn G und \tilde{H} normalverteilt sind, haben wir für $H_0^{(2)}: \varrho = 0$ gegen $H_1^{(2)}: \varrho \neq 0$ den statistischen Test (Hartung u.a. (1993), S. 547)

$$\delta_2: \mathbf{R}^{2 \cdot z} \rightarrow \{0, 1\},$$

$$(g_1, g_2, \dots, g_z, \tilde{h}_1, \tilde{h}_2, \dots, \tilde{h}_z) \mapsto \begin{cases} 1, & \text{falls } \left| T_2(g_1, \dots, g_z, \tilde{h}_1, \dots, \tilde{h}_z) \right| \\ & > t_{z-2; 1-\alpha/2} \\ 0, & \text{falls } \left| T_2(g_1, \dots, g_z, \tilde{h}_1, \dots, \tilde{h}_z) \right| \\ & \leq t_{z-2; 1-\alpha/2} \end{cases}. \quad (4.3)$$

Sei nun das Fehler-in-den-Variablen Modell (2.7) gegeben und $G_i := E_i - \beta \cdot D_i$ sowie $\tilde{H}_i := a_1 \cdot W_i + a_2 \cdot V_i$ für $i = 1, 2, \dots, z$ mit $a_1, a_2 \in \mathbf{R} \setminus \{0\}$. Entsprechend der allgemeinen Voraussetzung in Abschnitt 2 existiert $E(\tilde{H}_i)$.

Lemma 4.4:

- (i) *Es ist $\text{Cov}(G_i, \tilde{H}_i) = (a_1 \cdot \lambda - a_2 \cdot \beta) \cdot \sigma_D^2$ für $i = 1, 2, \dots, z$.*
- (ii) *Sei $\mathcal{C}: \mathbf{R} \times \mathbf{R} \rightarrow \mathbf{R}$, $(a_1, \beta) \mapsto a_1 \cdot \lambda - \beta$ eine Abbildung. Ist $(\zeta^{-1} \cdot \eta, \eta)$ eine Nullstelle von \mathcal{C} , dann ist auch $(-\eta^{-1}, -\zeta \cdot \eta^{-1})$ eine. (Inverse Eigenschaft von $\text{Cov}(G_i, \tilde{H}_i)$)*

Beweis: (i): Es ist $E(G_i) = 0$ und

$$\begin{aligned} \tilde{H}_i &= a_1 \cdot W_i + a_2 \cdot V_i \\ &= a_1 \cdot (\alpha + \beta \cdot \kappa_i) + a_2 \cdot \kappa_i + a_1 \cdot E_i + a_2 \cdot D_i. \end{aligned}$$

Somit gilt:

$$\begin{aligned}
\text{Cov}(G_i, \tilde{H}_i) &= E \left[\left(\tilde{H}_i - E(\tilde{H}_i) \right) \cdot \left(G_i - E(G_i) \right) \right] \\
&= E \left[a_1 \cdot (\alpha + \beta \cdot \kappa_i) \cdot E_i + a_2 \cdot \kappa_i \cdot E_i \right. \\
&\quad + a_1 \cdot E_i^2 + a_2 \cdot D_i \cdot E_i \\
&\quad - a_1 \cdot (\alpha + \beta \cdot \kappa_i) \cdot \beta \cdot D_i - a_2 \cdot \kappa_i \cdot \beta \cdot D_i \\
&\quad - a_1 \cdot E_i \cdot \beta \cdot D_i \\
&\quad \left. - a_2 \cdot D_i^2 \cdot \beta - E(\tilde{H}_i) \cdot G_i \right] \\
&= E \left[a_1 \cdot E_i^2 - a_2 \cdot \beta \cdot D_i^2 \right] \\
&= a_1 \cdot E(E_i^2) - a_1 \cdot E(E_i)^2 \\
&\quad - a_2 \cdot \beta \cdot E(D_i^2) + a_2 \cdot \beta \cdot E(D_i)^2 \\
&= a_1 \sigma_E^2 - a_2 \cdot \beta \cdot \sigma_D^2 \\
&= (a_1 \lambda - a_2 \cdot \beta) \cdot \sigma_D^2
\end{aligned}$$

(ii): $0 = \zeta^{-1} \cdot \eta \cdot \lambda - \eta \Rightarrow$

$$0 = (\zeta^{-1} \cdot \eta \cdot \lambda - \eta) \cdot (-\zeta \cdot \eta^{-2}) = -\eta^{-1} \cdot \lambda + \zeta \cdot \eta^{-1} \quad \blacksquare$$

Es folgt jetzt die alternative Darstellung von $H_0^{(1)}$:

Satz 4.5: *Es existiere die Varianz von \tilde{H} . Ist $\beta \geq 0$, $a_1 := \lambda^{-1} \cdot \beta_0$ und $a_2 := 1$, dann ist die Aussage „ $\beta = \beta_0$ “ äquivalent zur Aussage „ $\varrho = 0$ “, also $H_0^{(1)} \iff H_0^{(2)}$.*

Beweis: Aus $\beta = \beta_0$ folgt $\text{Cov}(G_i, \tilde{H}_i) = 0$ für $i = 1, 2, \dots, z$ mit Lemma 4.4, (i). Somit ist $\varrho = 0$.

Aus $\varrho = 0$ folgt $\text{Cov}(G_i, \tilde{H}_i) = 0$ für alle $i = 1, 2, \dots, z$, da die Varianzen gemäß der allgemeinen Voraussetzung in Abschnitt 2 größer Null sind. Hieraus folgt $\beta = \beta_0$ mit Lemma 4.4, (i).

Es könnte aus $\text{Cov}(G_i, \tilde{H}_i) = 0$ mit der Inversen Eigenschaft von $\text{Cov}(G_i, \tilde{H}_i)$ – Lemma 4.4, (ii) – zwar auch $\beta = -\lambda \cdot \beta_0^{-1}$ folgen, wobei dann $a_1 = \lambda^{-1} \cdot (-\lambda \cdot \beta_0^{-1})$ wäre. Dies ist wegen der Voraussetzung $\beta \geq 0$ nicht möglich (vgl. Fuller (1987), S. 45). \blacksquare

Vermerkt sei, daß der Test δ_2 (4.3) ohne die Berechnung eines Schätzers für β auskommt.

Da im funktionalen Modell (2.8) G_i und \tilde{H}_i normalverteilt für $i = 1, 2, \dots, z$ sind – beachte (2.9) –, sind die Voraussetzungen von Satz 4.1 und Satz 4.5 erfüllt. Das folgende Lemma ermöglicht die explizite Berechnung der Teststatistik $T_2(G_1, G_2, \dots, G_z, \tilde{H}_1, \tilde{H}_2, \dots, \tilde{H}_z)$:

Lemma 4.6: Sei λ gegeben, $a_1 := \lambda^{-1} \cdot \beta$ und $a_2 := 1$. Dann gilt

$$R = \frac{\sum_{i=1}^z \mathcal{A}_i \cdot \mathcal{B}_i}{\sqrt{\sum_{i=1}^z \mathcal{A}_i^2 \cdot \sum_{i=1}^z \mathcal{B}_i^2}}$$

mit $\mathcal{A}_i := W_i - \left(\frac{1}{z} \cdot \sum_{j=1}^z W_j\right) - \beta \cdot \left(V_i - \frac{1}{z} \cdot \sum_{j=1}^z V_j\right)$ und $\mathcal{B}_i := \lambda^{-1} \cdot \beta \cdot W_i + V_i - \frac{1}{z} \cdot \sum_{j=1}^z (\lambda^{-1} \cdot \beta \cdot W_j + V_j)$.

Beweis: Es gilt für $i = 1, 2, \dots, z$:

$$\begin{aligned} G_i - \bar{G} &= E_i - \beta \cdot D_i - \frac{1}{z} \cdot \sum_{j=1}^z (E_j - \beta \cdot D_j) \\ &= \beta \cdot \kappa_i + E_i - \frac{1}{z} \cdot \sum_{j=1}^z E_j - \beta \cdot \kappa_i - \beta \cdot D_i + \beta \cdot \frac{1}{z} \cdot \sum_{j=1}^z D_j \\ &= \alpha + \beta \cdot \kappa_i + E_i - \frac{1}{z} \cdot \sum_{j=1}^z (\alpha + \beta \cdot \kappa_j + E_j) \\ &\quad - \beta \cdot (\kappa_i + D_i) + \beta \cdot \frac{1}{z} \cdot \sum_{j=1}^z (\kappa_j + D_j) \\ &= \xi_i + E_i - \frac{1}{z} \cdot \sum_{j=1}^z (\xi_j + E_j) - \beta \cdot (\kappa_i + D_i) \\ &\quad + \beta \cdot \frac{1}{z} \cdot \sum_{j=1}^z (\kappa_j + D_j) \\ &= W_i - \left(\frac{1}{z} \cdot \sum_{j=1}^z W_j\right) - \beta \cdot \left(V_i - \frac{1}{z} \cdot \sum_{j=1}^z V_j\right). \end{aligned}$$

Weiter ist $\tilde{H}_i = a_1 \cdot W_i + a_2 \cdot V_i = \lambda^{-1} \cdot \beta \cdot W_i + V_i$ für $i = 1, 2, \dots, z$.
Es folgt die Behauptung. ■

5 Beispiel

Folgende Realisationen roher Kopfschäden (2.4) für die Tagegeldhöhen $\tau_1 = 10$ Euro und $\tau_2 = 25$ Euro eines KHT-Tarifes des Deutscher Ring Krankenversicherungsverein a.G. sollen hinsichtlich eines Schadenparameters analysiert werden:

Alters- gruppe	Männer		Frauen	
	$\widehat{K}_{x_i}^{10}$	$\widehat{K}_{x_i}^{25}$	$\widehat{K}_{y_i}^{10}$	$\widehat{K}_{y_i}^{25}$
22	0,58	0,74	0,88	1,36
27	0,61	0,81	1,37	1,64
32	0,43	0,84	1,34	1,64
37	0,79	1,06	1,11	1,27
42	0,77	1,31	1,08	1,62
47	1,30	1,80	1,57	2,02
52	1,39	2,91	1,75	2,50
57	3,10	3,36	2,48	3,09
62	3,60	4,00	3,42	3,38
67	4,97	5,41	3,72	4,71
72	5,88	7,72	5,25	6,01
77	7,53	9,95	7,99	7,95

Als erstes soll festgestellt werden, ob ein linearer Zusammenhang zwischen den beiden Tagegeldstufen besteht. Dafür verwenden wir den Test δ_2 (4.3) mit der Teststatistik (4.2)

$$T_2(V_1, V_2, \dots, V_{12}, W_1, W_2, \dots, W_{12}) = \frac{R \cdot \sqrt{10}}{\sqrt{1 - R^2}},$$

wobei

$$R := \frac{\sum_{i=1}^{12} (V_i - \bar{V}) \cdot (W_i - \bar{W})}{\sqrt{\sum_{i=1}^{12} (V_i - \bar{V})^2 \cdot \sum_{i=1}^{12} (W_i - \bar{W})^2}}$$

ist. Voraussetzung hierfür ist, daß $V_i = \widehat{K}_{x_i}^{10}$ und $W_i = \widehat{K}_{x_i}^{25}$ normalverteilt sind. Diese Bedingung ist im funktionalen Modell erfüllt (vgl. (2.9)). Wird die Nullhypothese $H_0^{(2)}$ verworfen, so liegt ein linearer Zusammenhang zwischen V und W vor beziehungsweise es ist $\beta \neq 0$ (Kendall und Stuart (1979), S. 413, Hartung u.a. (1993), S. 545ff). Dies wiederum ist der statistische Nachweis der Existenz des Schadenparameters.

Für Männer ergibt sich $R = 0,985$ und $T_2(v_1, \dots, v_{12}, w_1, \dots, w_{12}) = 17,8$. Da bei einem Sicherheitsniveau von $\alpha = 0,01$ $t_{10;1-0,01/2} = 3,58$ ist, wird die Nullhypothese verworfen. Bei den Männern sind die rohen Kopfschäden also hoch korreliert, d.h. es liegt ein linearer Zusammenhang zwischen den Kopfschäden der Tagegeldhöhen 10 Euro und 25 Euro vor.

Bei Frauen ist $R = 0,989$ und $T_2(v_1, \dots, v_{12}, w_1, \dots, w_{12}) = 20,9$. Mit dem Sicherheitsniveau $\alpha = 0,01$ wird auch bei Frauen die Nullhypothese abgelehnt, womit auch hier ein linearer Zusammenhang zwischen den Kopfschäden der Tagegeldhöhen 10 Euro und 25 Euro vorliegt.

Grundsätzlich kann auch der Test δ_1 (2.12), der ja auf dem eben eingesetzten Test δ_2 (4.3) beruht, zur Prüfung des Zusammenhanges zwischen V und W herangezogen werden, wenn G und \tilde{H} normalverteilt sind, λ und $\beta \geq 0$ beziehungsweise $\beta \leq 0$ bekannt sind. – Bei der Steigung β ist folgendes zu beachten. Ist das Vorzeichen von β bekannt, dann kann der Test δ_1 angewendet werden, wobei für $\beta \leq 0$ der Test entsprechend gilt (Fuller (1987), S. 45). Entscheidend ist nur, daß die Inverse Eigenschaft von $\text{Cov}(G_i, \tilde{H}_i)$, Lemma 4.4, nicht zum Tragen kommt. Aber auch im Falle, daß das Vorzeichen von β nicht bekannt ist, gibt es die grundsätzliche Möglichkeit, exakt zu testen (Fuller (1987), S. 45ff). Es besteht allerdings die Schwierigkeit, daß zwei Tests im Hinblick auf eine bestimmte Anwendung geeignet zu kombinieren sind.

Als nächstes soll die Frage beantwortet werden, ob der Schadenparameter (2.5) $P_{x(y)}^\tau$ größer als Eins ist. Um Satz 4.5 und Lemma 4.6 anwenden zu können, sind die „zufälligen Fehler“ $D_i = \hat{K}_{x_i}^{\tau_1} - K_{x_i}^{\tau_1}$ und $E_i = \hat{K}_{x_i}^{\tau_2} - K_{x_i}^{\tau_2}$ näher zu betrachten.

D_i und E_i sollen gemäß den getroffenen Annahmen stochastisch unabhängig und identisch normalverteilt mit Erwartungswert Null sein. Diese Eigenschaften können als zutreffend angesehen werden, wenn es bei der Leistungsabrechnung der Tagegelder zu keinen systematischen Fehlern kommt, die einzelnen Versicherungsnehmer getrennt abgerechnet werden und die Rechnungsprüfung grundsätzlich bei jeder Rechnungseinreichung dieselbe ist. Natürlich wird eine Rechnung mit einer Tagegeldhöhe von 250 Euro in der Regel genauer geprüft als mit einer Tagegeldhöhe von 10 Euro. Dies sind aber nicht die – wie sich gleich zeigen wird – für den Test relevanten Tagegeldhöhen, sondern 25 Euro und 10 Euro. Und hier kann von einer gleichen Abrechnungsart ausgegangen werden, insbesondere von $\sigma_D^2 = \sigma_E^2$, mithin $\lambda = 1$.

Es ist noch zu begründen, warum $\beta \geq 0$ ist. Bei einer Ursprungsgeraden würde eine negative Steigung bedeuten, daß bei der einen Tagegeldstufe der Versicherungsnehmer Tagegeld erhält, bei der anderen Tagegeld an das Versicherungsunternehmen zu zahlen hat. Ist der Abszissenabschnitt ungleich Null, so heißt dies,

daß bei einer Tagegeldstufe geleistet wird und bei der anderen nicht. Dies ist nicht möglich, da völlig unabhängig von der Tagegeldstufe genau dann Tagegeld an den Versicherungsnehmer gezahlt wird, wenn ein Krankenhausaufenthalt vorliegt.

Somit liegen die Voraussetzungen von Satz 4.1 und Satz 4.5 vor, d.h. es kann $H_0^{(1)} : \beta = \beta_0$ gegen $H_1^{(1)} : \beta \neq \beta_0$ getestet werden. Mit $a_1 := \beta_0$ und $a_2 := 1$ gilt mit Lemma 4.6

$$\mathcal{R} := R = \frac{\sum_{i=1}^z \mathcal{M}_i \cdot \mathcal{N}_i}{\sqrt{\sum_{i=1}^z \mathcal{M}_i^2 \cdot \sum_{i=1}^z \mathcal{N}_i^2}}$$

mit

$$\mathcal{M}_i := \widehat{K}_{x_i}^{\tau_2} - \left(\frac{1}{z} \cdot \sum_{j=1}^z \widehat{K}_{x_j}^{\tau_2} \right) - \beta_0 \cdot \left(\widehat{K}_{x_i}^{\tau_1} - \frac{1}{z} \cdot \sum_{j=1}^z \widehat{K}_{x_j}^{\tau_1} \right)$$

und

$$\mathcal{N}_i := \beta_0 \cdot \widehat{K}_{x_i}^{\tau_2} + \widehat{K}_{x_i}^{\tau_1} - \frac{1}{z} \cdot \sum_{j=1}^z (\beta_0 \cdot \widehat{K}_{x_j}^{\tau_2} + \widehat{K}_{x_j}^{\tau_1}).$$

Wir lehnen die Nullhypothese $H_0^{(1)} : \beta = \beta_0$ ab, wenn – siehe (2.11), (2.12) –

$$|T_1(v_1, \dots, v_z, w_1, \dots, w_z)| = \frac{|r| \cdot \sqrt{z-2}}{\sqrt{1-r^2}} > t_{z-2; 1-\alpha/2}$$

ist, wobei r mit den vorliegenden Realisationen der Funktionswert von \mathcal{R} ist. Für $\beta_0 = 1$ bedeutet dies, daß der Schadenparameter (2.5) $P_{x(y)}^{\tau_1} \neq 1$ ist.

Bezogen auf die Daten des KHT-Tarifes wird die Hypothese $H_0^{(2)} : \beta = 1$ abgelehnt, wenn mit einem Sicherheitsniveau von $\alpha = 0,01$

$$\frac{|r| \cdot \sqrt{10}}{\sqrt{1-r^2}} > t_{10; 1-0,01/2} = 3,58$$

ist. Bei Männern ergibt sich $r = 0,777$ und $T_1(v_1, \dots, v_{12}, w_1, \dots, w_{12}) = 3,90$, bei Frauen $r = -0,089$ und $T_1(v_1, \dots, v_{12}, w_1, \dots, w_{12}) = 0,28$. Dies bedeutet, daß bei Männern ein „echter“ Schadenparameter (größer Eins) vorliegt, wohingegen bei Frauen vom statistischen Test her nicht von einem Schadenparameter

(ungleich Eins) ausgegangen werden kann. Der Fall $0 < P_{x(y)}^{\tau_1} < 1$ liegt bei den Männern nicht vor, da alle Werte der rohen Kopfschäden der Tagegeldhöhe 25 Euro größer als die der Tagegeldhöhe 10 Euro sind.

Zur Berechnung des Schätzers $\hat{\beta}^*$ (2.10) für die Steigung β ist noch die Voraussetzung zu prüfen, ob der Achsenabschnitt Null ist. In der Literatur konnten dazu nur zwei asymptotische Tests im Rahmen eines speziellen strukturellen Modells gefunden werden (Fuller (1987), S. 43f). Die Annahme, daß hier eine Ursprungsgerade vorliegt, muss aber nicht getestet, sondern kann wie oben begründet werden: Ein Abszissenabschnitt ungleich Null würde bedeuten, daß bei der einen Tagegeldstufe geleistet wird, bei der anderen nicht.

Somit ergibt sich mit $\lambda = 1$ und $\alpha = 0$ für Männer $\hat{\beta}^* = 1,27$ und für Frauen $\hat{\beta}^* = 1,10$.

Bibliographie

- BOHN, K. (1980) *Die Mathematik der deutschen Privaten Krankenversicherung*. Hg. Deutsche Gesellschaft für Versicherungsmathematik e.V. 1. Aufl. Schriftreihe angewandte Versicherungsmathematik 11. Versicherungswirtschaft, Karlsruhe.
- BRÜNJES, C. (1985) Spezifische Rechnungsgrundlagen der Krankentagegeldversicherung. *Blätter der Deutsche Gesellschaft für Versicherungsmathematik e.V.* **17.2**, 179–195.
- CRAMÉR, H. (1946) *Mathematical Methods of Statistics*. 1. Aufl. Princeton University Press, Princeton.
- CREASY, M. A. (1956) Confidence limits for the gradient in the linear functional relationship. *Journal of the Royal Statistical Society, Series B.* **18**, 65–69.
- FARNY, D. u.a., Hg. (1988) *Handwörterbuch der Versicherung*. 1. Aufl. Versicherungswirtschaft, Karlsruhe.
- FISZ, M. (1976) *Wahrscheinlichkeitsrechnung und mathematische Statistik*. 8. Aufl. VEB Deutscher Verlag der Wissenschaften, Berlin.
- FULLER, W.A. (1987) *Measurement Error Models*. 1. Aufl. John Wiley & Sons, New York.
- HARTUNG, J., ELPELT, B. und KLÖSENER, K.-H. (1993) *Statistik: Lehr- und Handbuch der angewandten Statistik*. 9., durchges. Aufl. Oldenbourg, München u.a.
- KENDALL, M. und Stuart, A. (1979) *The Advanced Theory of Statistics. Volume 2: Inference and Relationship*. 4. Aufl. Charles Griffin & Company, London.
- MILLER, R.G. (1996) *Grundlagen der angewandten Statistik*. 1. Aufl. Oldenbourg, München u.a.
- SCHACH, S. und SCHÄFER, T. (1978) *Regressions- und Varianzanalyse*. 1. Aufl. Springer, Berlin u.a.
- TOSBERG, A. (1940) *Rechnungsgrundlagen und Schadentafeln der Krankheitskostenversicherung*. 1. Aufl. E.S. Mittler & Sohn, Berlin.

Kai Bruchlos
 Deutscher Ring
 Krankenversicherungsverein a.G.
 Ludwig-Erhard-Str. 22
 D-20459 Hamburg

Zusammenfassung

Schätzer und Test für den Schadenparameter
in Krankenversicherungstarifen mit Selbstbeteiligung

In der Krankenversicherung gibt es bei Tarifen mit Selbstbeteiligung den sogenannten Schadenparameter, welcher das Verhältnis von Kopfschäden wiedergibt, die aus zwei Tarifvarianten stammen. Zur Bestimmung des Schadenparameters soll ein statistischer Test herangezogen werden. Der t-Test der einfachen normalen linearen Regression ist hierfür wenig geeignet. Deshalb wird die Theorie des Fehler-in-den-Variablen Modells verwendet.

Beispielhaft wird der Schadenparameter in der Krankenhaustagegeldversicherung betrachtet, für den im funktionalen Modell ein Test konstruiert wird. Die theoretischen Ergebnisse werden anhand der Daten zweier Tagegeldstufen durchgerechnet.

Summary

Estimator and test for the Schadenparameter
in special tariffs in health insurance

In special tariffs in health insurance there is a linear interrelation between average losses which belong to two variants of the tariff. Such a linear interrelation is the so-called „Schadenparameter“, a special kind of parameter for calculation. For the calculation of the Schadenparameter a statistical test should be consulted. The t-test of the ordinary linear regression analysis is not a suitable test. Therefore the theory of the errors in variables model is considered.

The Schadenparameter is considered in the health insurance of daily benefits during hospitalisation, for which a test is constructed in the functional model. The theoretical results are checked on the basis of data from two levels of daily benefit.

Résumé

Estimateur et teste pour le paramètre de dégât
dans le système des assurances maladies

Dans le système des assurances maladies il y a deux tarifs avec ticket modérateur, le paramètre de dégâts qui reproduit la relation des dégâts individuel qui parviennent de deux variantes tarifaires. Pour déterminer le paramètre de dégât, on se sert d'un test statistique. Le teste T de la régression linéaire ne convient que peu pour ce propos. C'est pour cela que la théorie du modèle „de la faute dans les variables“ est peu utilisé.

On regarde exemplairement le paramètre de dégât d'allocation forfaitaire en cas d'hospitalisation, pour lequel dans le modèle fonctionnel un teste a été développé. Les résultats théoriques sont calculés en s'appuyant sur les dates des deux degrés du montant quotidien.