

Technische Universität München

ZENTRUM MATHEMATIK

**Anwendungen der
Überlebenszeitanalyse
in der Pflegeversicherung**

Diplomarbeit

von

Florian Rudolph

Themenstellerin: Prof. Dr. Czado

Betreuerin: Prof. Dr. Czado

Abgabetermin: 25. Januar 2000

Hiermit erkläre ich, daß ich die Diplomarbeit selbständig angefertigt und nur die angegebenen Quellen verwendet habe.

München, 25. Januar 2000

Danksagung

An dieser Stelle möchte ich allen meinen Dank aussprechen, die ihren Teil zur Entstehung dieser Arbeit beigetragen haben, in erster Linie Frau Professor Czado für die hervorragende Betreuung, die konstruktive Kritik und die Geduld, mit der sie mir über manche Hürde hinweghalf.

Desweiteren gebührt mein Dank Herrn Gerd Holzner für die Durchsicht der Arbeit, Herrn Zacharias und den Kollegen aus der Abteilung Leben- Mathematischer Service. Zuletzt möchte ich noch bei allen anderen Mitarbeitern am Lehrstuhl für Mathematische Statistik, die mir mit Rat und Tat zur Seite standen, danken.

Inhaltsverzeichnis

1	Einleitung	1
2	Theoretische Grundlagen	5
2.1	Stochastische Prozesse	5
2.2	Martingaltheorie	9
2.3	Zählprozeßtheorie	17
2.4	Grenzwerttheorie	22
2.5	Markovketten	24
3	Mathematische Grundlagen der Verweildaueranalyse	27
3.1	Die Überlebensfunktion	27
3.2	Die Hazardfunktion	28
3.3	Funktionale Darstellung der Hazardfunktion	29
3.4	Aufbau einer Studie	31
3.5	Likelihood-Techniken für Überlebenszeitmodelle	32
3.5.1	Likelihood-Konstruktion für Modelle mit unabhängiger Zensierung	32
3.5.2	Partial Likelihood	34
3.6	Schätzen der Überlebensfunktion	36
3.6.1	Der Produkt-Limit-Schätzer (Kaplan und Meier)	36
3.6.2	Der Nelson-Aalen-Schätzer	39
3.7	Das Proportional-Hazard-Modell	45
3.7.1	Modellierung	45
3.7.2	Partial Likelihood mit “ties”	47
3.7.3	Schätzung des Basis-Hazards (Breslow)	48
3.7.4	Asymptotische Eigenschaften	48
3.7.5	Hypothesentests im Proportional-Hazard-Modell	59
3.7.6	Lokale Tests	60
3.7.7	Modellbildung mit dem AIC	62
3.7.8	Residuenanalyse	62
3.7.9	Graphische Methoden zur Überprüfung des Proportional-Hazard	66
3.8	Parametrische Modelle	67

4	Datenanalyse zur gesetzlichen Pflegeversicherung	69
4.1	Überblick über die gesetzliche Pflegeversicherung	69
4.2	Beschreibung der Daten	71
4.3	Proportional-Hazard-Modell	73
4.3.1	Modellbildung (ohne Diagnosen)	73
4.3.2	Proportional Hazard (mit Diagnosen)	79
4.4	Test auf Proportional Hazard	80
4.4.1	Residuenanalyse	80
4.4.2	Graphische Tests des Propotional Hazard	85
4.5	Modellierung der Zustandsübergänge	94
4.5.1	Modellierung der Zustandsübergänge zwischen ambulanter und stationärer Pflege	94
4.5.2	Modellierung der Zustandsübergänge zwischen den verschiedenen Pflege- stufen	99
5	Entwicklung eines Versicherungsmodells	103
5.1	Versicherung als Markov-Prozeß	103
5.1.1	Zahlungsfunktionen	104
5.1.2	Berechnung von Barwerten und Erwartungswerten	105
5.1.3	Aktuarielle Werte	106
5.1.4	Beiträge und Reserven	107
5.2	Berechnung von Beiträgen für Pflegeversicherungs-Modelle	108
5.2.1	Versicherungsmodell mit Zustandsübergängen zwischen ambulanter und stationärer Pflege	109
5.2.2	Versicherungsmodell mit Zustandsübergängen zwischen Pflegestufen . . .	115
5.3	Beschreibung des Programms zur Beitragsberechnung	118
5.3.1	Beschreibung des Tarifs PET	118
5.3.2	Beschreibung des Programmes und Ergebnis der Beitragsberechnung . . .	119
6	Zusammenfassung und Ausblicke	121
A	Routinen zur graphischen Analyse	123
B	Funktionen zur Maximierung des Log-Likelihood	126
C	Tabellen	128
C.1	Pflegefalleintrittswahrscheinlichkeiten der Custodial Insurance, Japan	128
C.2	Einjährige Sterbewahrscheinlichkeiten Bayern für x -jährige Frauen und Männer (1986 bis 1988)	130

Kapitel 1

Einleitung

Die Überlebenszeit-, oder Verweildaueranalyse (engl.: Survival Analysis) ist ein Teilgebiet der Statistik, das in vielen Gebieten der Wissenschaft auf großes Interesse stößt. Zum Beispiel können in der Medizin mit Hilfe verweildaueranalytischer Methoden Aussagen über die Wirksamkeit neuer Medikamente, bzw. Behandlungsmethoden getroffen werden. In der Geologie werden anhand der Zerfallseigenschaften (“Überleben”) bestimmter Moleküle die Alter von Gesteinsschichten bestimmt.

Auch in der Personenversicherung ist die Verweildaueranalyse ein wichtiger Bestandteil. Die Anfänge gehen zurück auf das Jahr 1693, in dem die erste Sterbetafel (Absterbeordnung der Stadt Breslau) entwickelt wurde. De Moivre (1724), Gompertz (1827) und Weibull (1936) und andere versuchten, eine analytische Approximation für die Überlebensfunktion eines Menschen zu geben. Heute können mit Hilfe verweildaueranalytischer Methoden sichere Beiträge kalkuliert werden, Risiken abgeschätzt und neue Produkte entwickelt werden. Auf Volkszählungen basierende Sterbetafeln in Kombination mit medizinischem Know-how bilden dabei die Basis für eine Kalkulation, die es heute ermöglicht, sogar für sogenannte medizinische Risikogruppen (d.h. Personen, die aufgrund von Krankheiten, bzw. in Folge ihrer Krankengeschichte einem höheren Sterberisiko ausgesetzt sind) Lebensversicherungspolicen anzubieten. Die mathematische Grundlage zur Bestimmung eines Lebensversicherungsbeitrages sind einjährige Überlebenswahrscheinlichkeiten q_x

$$q_x = P(x \leq T < x + 1 | T \geq x),$$

wobei hier mit x das Lebensalter in Jahren und mit T die Zufallsvariable Lebenszeit bezeichnet ist. Mit q_x ist also die Wahrscheinlichkeit bezeichnet, daß eine Person das Lebensalter $x + 1$ nicht erreicht, unter der Voraussetzung, daß das Alter x erreicht wurde. In Abbildung 1.1 sehen wir den typischen Verlauf dieser einjährigen Sterbewahrscheinlichkeiten am Beispiel bayerischer Männer und Frauen (ermittelt im Zeitraum zwischen 1986 und 1988). Hier fallen einige Besonderheiten auf: Zunächst erkennt man, daß für Männer die Sterbewahrscheinlichkeit in allen Altersbereichen höher ist als für Frauen. Zudem bemerkt man bei Männern im Alter um ca. 20 Jahre eine “Überhöhung” der Sterblichkeit, die aus Motorrad- und Autounfällen resultiert

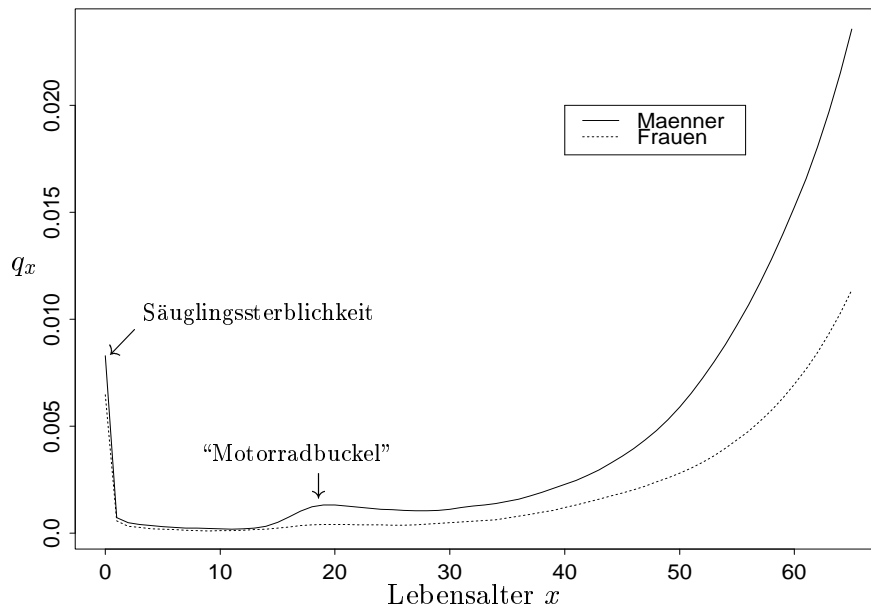


Abbildung 1.1: Einjährige Sterbewahrscheinlichkeiten für x -jährige Frauen und Männer in Bayern 1986-1988

und daher auch “Motorradbuckel” genannt wird. Die außergewöhnlich hohe Sterbewahrscheinlichkeit bei Neugeborenen ($q_0 \approx 10 q_1$), sowie das leichte (in der Abbildung kaum erkennbare) Abnehmen der einjährigen Sterbewahrscheinlichkeiten bis zu einem Alter von ca. 10 Jahren, führt zu einem Effekt, der in englischsprachiger Literatur oft mit “bathtub-shape” (Badewannen-Form) des Kurvenverlaufs bezeichnet wird.

Ein relativ junges Produkt in der Personenversicherung ist die Pflegeversicherung, die dem Versicherten Leistungen bietet, falls dieser nicht mehr in der Lage ist die grundlegenden Verrichtungen des täglichen Lebens, wie z.B. Körperwäsche, Nahrungsaufnahme, etc. selbständig zu bewerkstelligen. Diese Art der Versicherung ist im vergangenen Jahrzehnt ein immer aktuelleres Thema geworden. Zum einen führte der Wandel in der Sozialstruktur (hohe Geburtenrate direkt nach dem Krieg, sowie eine niedrige Geburtenrate in den letzten Jahren) zu einem Anstieg des sogenannten Abhängigkeitskoeffizienten (im wesentlichen das Verhältnis von arbeitender zu nicht-arbeitender Bevölkerung, siehe Abbildung 1.2), zum anderen führt auch die wachsende Lebenserwartung und die Änderung familiärer Strukturen in den Industrieländern dazu, daß es immer mehr Pflegebedürftige gibt, für deren Pflege teure Fachkräfte engagiert werden müssen, da in immer mehr Familien die Möglichkeit oder Bereitschaft für eine Pflege durch Angehörige fehlt. Das finanzielle Risiko dafür trägt in diesem Fall die Pflegeversicherung.

So hat sich diese Arbeit zum Ziel gesetzt, mit Hilfe von Daten aus der gesetzlichen Pflegeversicherung, die seit 1.4.1995 in Kraft ist, mittels moderner verweildaueranalytischer Methoden das

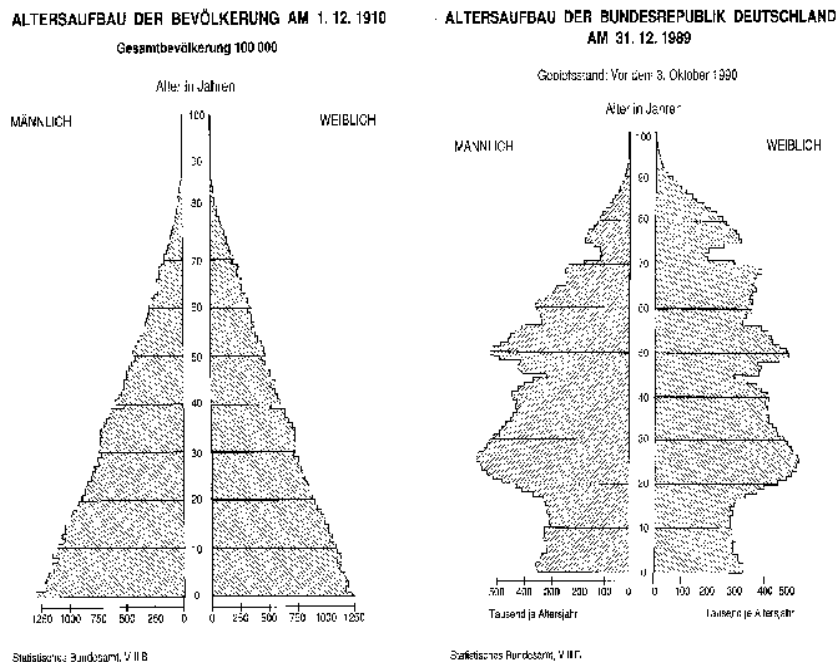


Abbildung 1.2: Verteilung der Bevölkerung in Deutschland auf Altersgruppen in den Jahren 1910 und 1989

Überleben von Pflegebedürftigen zu untersuchen und die Ergebnisse auf ein Versicherungsmodell anzuwenden.

Einen zentralen Punkt der Arbeit stellt das von David Cox 1972 [9] vorgeschlagene semiparametrische Regressionsmodell zur Bestimmung der Abhängigkeit der Überlebensfunktion von verschiedenen Kovariablen (das Überleben beeinflussende Faktoren, in unserem Beispiel wären das Alter, Geschlecht, pflegeverursachende Krankheiten, Pflegeart). Die Grundidee dieses Verfahrens ist, daß sich die sogenannte Hazardfunktion (Sterbeintensität), die für die Zufallsvariable Lebenszeit T folgendermaßen als Dichte von T , gegeben Überleben bis t , definiert ist

$$\lambda(t) = \frac{P(t \leq T < T + dt | T \geq t)}{dt},$$

formulieren läßt als Produkt aus einer, für alle Beobachtungen gleichen, "Basis-Hazardfunktion" und einem für jede Kovariable Z unterschiedlichen Multiplikator

$$\lambda(t|Z) = \lambda_0(t) \exp(\beta^t Z),$$

wobei β ein endlich-dimensionaler Regressionskoeffizient und Z der Kovariablenvektor ist. Dieses Verfahren heißt semiparametrisch, da die resultierende Überlebensfunktion

$$S(t) = P(T > t)$$

einerseits von dem endlichen Parameter β , zum anderen aber auch von einer "Basis-Überlebensfunktion" $S_0(t)$ abhängt, die in einem unendlich-dimensionalen Raum geschätzt wird.

Die Arbeit ist wie folgt untergliedert: In Kapitel 2 werden wesentliche Resultate aus der Theorie stochastischer Prozesse dargestellt, die zum einen dazu benötigt werden, asymptotische Konvergenzaussagen für das Cox-Regressions-Modell zu treffen, zum anderen, um ein auf Markov-Prozessen basiertes Versicherungsmodell zu entwickeln. In diesem Zusammenhang soll das Konzept lokaler Submartingale und deren Kompensatoren, insbesondere deren Anwendung auf Zählprozesse, vorgestellt werden. Ein wichtiges Resultat ist hier der zentrale Grenzwertsatz für Martingale von Rebolledo [33], der ein Hilfsmittel für den Beweis der asymptotischen Normalität des Proportional-Hazard-Modell auf Zählprozeß-Basis ist.

Eine Darstellung wichtiger Konzepte der Überlebenszeitanalyse ist die Zielsetzung von Kapitel 3. Dabei sollen der später in der Arbeit benötigte nicht-parametrische Nelson-Aalen-Schätzer für die Überlebensfunktion, das Regressionsmodell von Cox, sowie verschiedene Verfahren zur Analyse dieses Modells detailliert herausgearbeitet werden. Andere wichtige Resultate der Überlebenszeitanalyse, wie zum Beispiel das nicht-parametrische Schätzverfahren von Kaplan und Meier (1958)[22], sowie parametrische Regressionsansätze werden in knapper Form dargestellt.

In Kapitel 4 werden wir die vorher erarbeiteten Konzepte systematisch auf einen Datensatz anwenden. Der Datensatz beinhaltet Überlebensdaten, sowie zusätzliche Informationen (Geschlecht, Alter, Schwere und Art der Pflege, sowie diagnostizierte pflegeverursachende Krankheiten) von Pflegebedürftigen in der gesetzlichen Pflegeversicherung, die im Zeitraum vom 1.4.1995 bis zum 31.12.1998 beobachtet wurden. Hier analysieren wir zum einen natürlich das Überleben von Pflegebedürftigen, untersuchen aber auch Veränderungen (bzgl. Pflegeart und Pflegestufe) im Pflegeverlauf und deren Auswirkungen auf das Überleben von Pflegebedürftigen.

Schließlich werden wir in Kapitel 5 die Resultate aus Kapitel 4 im Rahmen eines Versicherungsmodells anwenden. Die theoretische Basis hierfür bilden Markov-Prozesse mit endlichem Zustandsraum, die die Entwicklung eines versicherten Risikos beschreiben. Mit Hilfe von zustandsabhängigen (zufälligen) Funktionen, die den Zahlungsstrom zwischen Versicherer und Versichertem beschreiben modellieren wir dann das Versicherungsprinzip. Dieses besagt im wesentlichen, daß jeder Versicherte für seinen erwarteten Schaden selbst aufkommt. Schließlich werden wir für zwei verschiedene Versicherungsmodelle die Basis für eine praktische Anwendung in der Form von Beitragsberechnungen schaffen.

Kapitel 2

Theoretische Grundlagen

Um Aussagen über asymptotische Konvergenz von Schätzern in der Verweildaueranalyse treffen zu können, bedient man sich häufig Mitteln aus der Theorie der Zählprozesse. In diesem Abschnitt wird ein knapper Abriss über stochastische Prozesse und die wesentlichen Konzepte von Martingalen und Zählprozessen gegeben. Näheres hierzu findet sich in Andersen u.a. (1993)[1]. Die wesentlichen Konzepte der stochastischen Integration sind orientiert an Rogers und Williams (1979 und 1987)[34], [35]. Einen guten Überblick über Zählprozesse gibt Brémaud (1981) [7]. Letztendlich benötigen wir noch einige elementare Aussagen über zeitstetige Markov-Prozesse.

2.1 Stochastische Prozesse

Satz und Definition 2.1 (bedingte Erwartung) Sei (Ω, \mathcal{F}, P) ein Wahrscheinlichkeitsraum und X eine Zufallsvariable mit $E(|X|) < \infty$. Sei nun $\mathcal{G} \subset \mathcal{F}$ eine Sub- σ -Algebra von \mathcal{F} . Dann gibt es eine Zufallsvariable Y mit

(i) Y ist \mathcal{G} -meßbar,

(ii) $E(|Y|) < \infty$,

(iii) $\forall G \in \mathcal{G}$ gilt $E(Y \cdot I(G)) = E(X \cdot I(G))$, ($I(G)(x) = 1$, falls $x \in G$ und 0 sonst).

Falls \tilde{Y} eine weitere Zufallsvariable ist, die die Eigenschaften (i)-(iii) erfüllt, so gilt: $\tilde{Y} = Y$ fast sicher. Die so definierte Zufallsvariable $Y = E(X|\mathcal{G})$ heißt **bedingte Erwartung** von X , gegeben \mathcal{G} .

Beweis:

Der Beweis findet sich zum Beispiel in Rogers und Williams (1979) [34].

Der folgende Satz enthält eine Auflistung einiger fundamentaler Eigenschaften bedingter Erwartungen.

Satz 2.2 (Eigenschaften der bedingten Erwartung) Seien (Ω, \mathcal{F}, P) ein Wahrscheinlichkeitsraum, \mathcal{G} und \mathcal{H} Sub- σ -Algebren von \mathcal{F} . Zudem sei für alle auf (Ω, \mathcal{F}, P) definierten Zufallsvariablen X der Erwartungswert $E(|X|) < \infty$. Dann gelten folgende Aussagen

A. Falls X \mathcal{G} -meßbar ist, dann gilt: $E(X|\mathcal{G}) = X$ fast sicher.

B. (Linearität) $E(aX_1 + bX_2|\mathcal{G}) = aE(X_1|\mathcal{G}) + bE(X_2|\mathcal{G})$

C. Falls $\mathcal{H} \subset \mathcal{G} \subset \mathcal{F}$, dann gilt:

$$E[E(X|\mathcal{G})|\mathcal{H}] = E[X|\mathcal{H}] \quad \text{fast sicher.}$$

D. Falls Z \mathcal{G} -meßbar und beschränkt ist, dann gilt:

$$E[ZX|\mathcal{G}] = ZE[X|\mathcal{G}] \quad \text{fast sicher.}$$

Zum Beweis siehe zum Beispiel Stirzaker (1994) [38] Seite 143.

Definition 2.3 (Filtration) Sei \mathcal{T} ein zeitstetiges Intervall

$$\mathcal{T} = [0, \tau) \quad \text{oder} \quad \mathcal{T} = [0, \tau] \quad \text{mit} \quad 0 < \tau \leq \infty$$

und (Ω, \mathcal{F}, P) ein Wahrscheinlichkeitsraum. Dann heißt eine Familie von σ -Algebren $\mathcal{F}_t : t \in \mathcal{T}$ **Filtration** wenn für alle $s \leq t$ gilt:

$$\mathcal{F}_s \subseteq \mathcal{F}_t. \quad (2.1)$$

Eine Filtration heißt **rechtsstetig** wenn für alle $t \in \mathcal{T}$ gilt:

$$\lim_{s \downarrow t} \mathcal{F}_s = \mathcal{F}_t \quad (2.2)$$

und **vollständig** falls

$$A \subset B \in \mathcal{F}, P(B) = 0, \Rightarrow A \in \mathcal{F}_0. \quad (2.3)$$

Bemerkung: Die Bedingungen (2.1) - (2.3) heißen auch die “gewöhnlichen Bedingungen”.

Definition 2.4 (stochastischer Prozeß) Ein **stochastischer Prozeß** ist eine zeitindizierte Menge von Zufallsvariablen $(X(t) : t \in \mathcal{T})$.

Mit $X(t, \omega)$ $\omega \in \Omega$ wird die **Realisierung** eines stochastischen Prozesses zum Zeitpunkt t bezeichnet.

Die Abbildung $t \mapsto X(t, \omega) \omega \in \Omega$ heißt **Pfad** eines stochastischen Prozesses.

Ein stochastischer Prozeß heißt **adaptiert** (gegenüber der Filtration \mathcal{F}_t), wenn $X(t)$ \mathcal{F}_t meßbar ist für alle $t \in \mathcal{T}$.

Ein stochastischer Prozeß heißt **càdlàg** (continu à droite, limité à gauche), wenn für fast alle $\omega \in \Omega$ die Pfade $X(t, \omega) : t \in \mathcal{T}$ rechtsstetig sind und der linksseitige Grenzwert $\lim_{s \uparrow t} X(s, \omega)$ existiert.

Bemerkungen:

- Normalerweise wird bei der Definition eines stochastischen Prozesses zwischen zeitstetigen und zeitdiskreten stochastischen Prozessen unterschieden. Die Definition 2.4 entspricht der Definition eines zeitstetigen stochastischen Prozesses, da hier der Index $t \in \mathbb{R}$ und damit überabzählbar ist. Im zeitdiskreten Fall betrachtet man eine Folge von Zufallsvariablen mit abzählbarer Indizierung (z.B. X_n $n \in \mathbb{N}$).
- Die Filtration \mathcal{F}_t wird auch Historie genannt. Man kann \mathcal{F}_t als Informationsstand zum Zeitpunkt t betrachten. Die σ - Algebra \mathcal{F} wird mit wachsendem t immer feiner.
- Häufig wird $\mathcal{F}_t := \sigma\{X(s) : s \leq t\}$ gesetzt, man spricht dann von der natürlichen Filtration. Mit $\mathcal{F}_{t-} := \sigma\{X(s) : s < t\}$ wird die Historie vor dem Zeitpunkt t bezeichnet.

Definition 2.5 (Version, Ununterscheidbarkeit) Seien X und Y zwei stochastische Prozesse auf (Ω, \mathcal{F}, P) . Y nennt man eine **Version** von X falls

$$P(\{\omega | X(t, \omega) \neq Y(t, \omega)\}) = 0 \quad \forall t.$$

X und Y heißen **ununterscheidbar** falls

$$P(\{\omega | X(t, \omega) \neq Y(t, \omega) \forall t\}) = 0,$$

was bedeutet, daß die beiden Prozesse fast sicher (bis auf P -Nullmengen) die selben Pfade haben.

Definition 2.6 (Stopzeit) Eine **Stopzeit** T ist eine Zufallsvariable, die Werte in \mathcal{T} annimmt, so daß

$$\{T \leq t\} \in \mathcal{F}_t \quad \text{für alle } t \in \mathcal{T}.$$

Beispiel 2.7 Wenn X càdlàg und adaptiert ist, dann ist $\inf\{t : |X(t)| \leq c\}$ eine Stopzeit.

Bemerkung: Intuitiv kann die Stopzeit als die zufällige Zeit bis zu einem bestimmten Ereignis interpretiert werden. Dabei kann zu jedem Zeitpunkt t festgestellt werden, ob das Ereignis schon eingetreten ist oder noch nicht.

Definition 2.8 (gestoppter Prozeß) Gegeben ist ein stochastischer Prozeß X und eine Stopzeit T . Dann ist der **gestoppte** Prozeß X^T definiert durch:

$$X^T(t) := X(t \wedge T) \quad \text{mit } t \wedge T = \min(t, T).$$

Definition 2.9 (lokale Eigenschaften) Eine monoton wachsende Folge von Stopzeiten T_n heißt **lokalisierend**, wenn gilt:

$$P(T_n \geq t) \rightarrow 1 \quad \text{für } n \rightarrow \infty.$$

Ein stochastischer Prozeß besitzt eine bestimmte Eigenschaft **lokal**, wenn eine lokalisierende Folge von Stopzeiten T_n existiert, so daß der Prozeß

$$I(T_n > 0)X^{T_n}$$

die entsprechende Eigenschaft besitzt. Ein Prozeß heißt **lokal beschränkt**, wenn eine lokalisierende Folge von Stopzeiten T_n , sowie Konstanten c_n existieren, so daß für alle n gilt:

$$\sup_{t \leq T_n} |X(t)| \leq c_n \quad \text{fast sicher für } T_n > 0.$$

Definition 2.10 (Variation) Ist $f : [a, b] \mapsto \mathbb{R}$ eine Funktion, so heißt

$$V(f; [a, b]) = \sup \left\{ \sum_{k=1}^n |f(x_k) - f(x_{k-1})| : a = x_0 < x_1 < \dots < x_n = b, n \in \mathbb{N} \right\}$$

die **Totalvariation** von f über $[a, b]$. Falls gilt

$$V(f; [a, b]) < \infty,$$

dann nennt man f von **beschränkter Variation**. Eine Funktion $f : \mathcal{T} \mapsto \mathbb{R}$ (mit $\mathcal{T} = [0, \tau)$, oder $\mathcal{T} \in [0, \tau]$, $0 < \tau \leq \infty$), heißt von **lokal beschränkter Variation**, wenn f für alle $t \in \mathcal{T}$ von beschränkter Variation ist.

Mit diesen Definitionen können wir nun ein stochastisches Integral einführen, das den späteren Anforderungen der Arbeit genügt.

Definition 2.11 (stochastische Integration) Für 2 stochastische Prozesse X, Y definieren wir **stochastische Integration** als pfadweises Lebesgue-Stiltjes-Integral. $\int X dY$ steht hier für die Abbildung

$$t \mapsto \int_0^t X(s) dY(s),$$

definiert für alle $(t, \omega) \in (\mathcal{T}, \Omega)$, für die gilt

$$\int_0^t |X(s)| |dY(s)| < \infty,$$

für die also das pfadweise Lebesgue-Stiltjes-Integral existiert.

Bemerkung: Da wir uns im folgenden Text auf stochastische Prozesse Y mit lokal beschränkter Variation als Integratoren beschränken, können wir Y als maßerzeugende Funktion verwenden, so daß die Existenz des Lebesgue-Stiltjes-Integrals gesichert ist und wir nicht auf den verallgemeinerten stochastischen Integrationsbegriff von Itô (siehe zum Beispiel in Rogers und Williams [35] (Kapitel 4)) angewiesen sind.

Definition 2.12 (gleichmäßige Integrierbarkeit) Ein auf dem Wahrscheinlichkeitsraum (Ω, \mathcal{F}, P) definierter stochastischer Prozeß $X(t)$ heißt **gleichmäßig integrierbar**, falls der Grenzwert

$$\lim_{c \rightarrow \infty} \int_{|X(t)| > c} |X(t)| dP = 0 \quad \text{gleichmäßig in } t \in \mathcal{T}.$$

Ein Kriterium für gleichmäßige Integrierbarkeit liefert folgender Satz.

Satz 2.13 Der auf dem Wahrscheinlichkeitsraum (Ω, \mathcal{F}, P) definierte stochastische Prozeß $X(t)$ $t \in \mathcal{T}$ mit $E(|X(t)|) < \infty \forall t \in \mathcal{T}$ ist genau dann gleichmäßig integrierbar, wenn gilt

$$\sup_{t \in \mathcal{T}} E(|X(t)|) < \infty$$

und es für alle $\delta > 0$ ein $\varepsilon > 0$ gibt, so daß

$$A \in \mathcal{F}, P[A] \leq \delta \Rightarrow \int_A |X(t)| dP < \varepsilon \quad \forall t \in \mathcal{T}.$$

Den Beweis dieses Satzes kann man in Bauer [4] Seite 106 nachlesen.

Definition 2.14 (Linksstetige Modifikation und Sprungprozeß) Für einen càdlàg Prozeß X ist die **linksstetige Modifikation** X_- definiert durch

$$X_-(t) = \lim_{s \uparrow t} X(s)$$

und sein **Sprungprozeß** $\Delta X(t)$ durch

$$\Delta X(t) = X(t) - X_-(t).$$

Bemerkung: Die linksstetige Modifikation $X_-(t)$ wird im Text auch mit $X(t-)$ bezeichnet.

2.2 Martingaltheorie

Definition 2.15 (Martingal) Ein adaptierter cadlag- Prozeß heißt **Martingal**, wenn er folgende Eigenschaften erfüllt:

- (i) $E(|M(t)|) < \infty$ für alle $t \in \mathcal{T}$,
- (ii) $E(M(t)|\mathcal{F}_s) = M(s)$ für alle $s \leq t$ (Martingaleigenschaft),

Gilt in (ii) " \leq " anstatt "=", so spricht man von einem **Supermartingal**, gilt " \geq ", so handelt es sich um ein **Submartingal**.

Ein Martingal heißt **quadratintegrierbar**, falls für alle $t \in \mathcal{T}$

$$E(M(t)^2) < \infty.$$

Da der für quadratintegrierbare Martingale M der Grenzwert

$$\lim_{t \rightarrow \tau} M(t)$$

existiert, können diese durch Grenzwertbildung auf $\overline{\mathcal{T}}$ erweitert werden.

Beispiel 2.16 (Standard Brownsche Bewegung) Die Standard Brown'sche Bewegung ist ein stochastischer Prozeß X mit stetigen Pfaden und $X(0) = 0$ für alle $\omega \in \Omega$. Die Zuwächse sind unabhängig normalverteilt, d.h.

$$X(s) - X(t) \sim N(0, s - t) \quad t \leq s.$$

Der so definierte stochastische Prozeß ist ein Martingal (Abbildung 2.1)

Beweis:

Zunächst muß man zeigen, daß der Erwartungswert $E(|X(t)|) < \infty \forall t$ existiert. $X(t)$ ist jedoch verteilt nach $N(0, t)$, so daß gilt:

$$\begin{aligned} E(|X(t)|) &= 2 \int_0^\infty s \frac{1}{\sqrt{2\pi t}} e^{-\frac{s^2}{2t}} ds = \\ &= \left[-2 \frac{t}{\sqrt{2\pi t}} e^{-\frac{s^2}{2t}} \right]_0^\infty = 2 \frac{t}{\sqrt{2\pi t}} < \infty \quad \forall t. \end{aligned}$$

Nun muß man noch zeigen, daß $\forall s \leq t$ gilt:

$$E[X(t)|\mathcal{F}_s] = X_s \iff E[X(t)|\mathcal{F}_s] - X(s) \stackrel{\text{Satz 2.2 A,B}}{=} E[X(t) - X(s)|\mathcal{F}_s].$$

Da jedoch $X(t) - X(s)$ normalverteilt mit Erwartungswert 0 sind, erhält man, daß

$$E[X(t) - X(s)|\mathcal{F}_s] = 0.$$

Somit ist die Aussage gezeigt. □

Definition 2.17 (Vorhersehbarkeit) Ein stochastischer Prozeß H heißt **vorhersehbar** (engl.: *predictable*), falls $H(\omega, t)$ hinsichtlich der durch die Klasse der linksstetigen, adaptierten Prozesse erzeugten σ -Algebra auf $\Omega \times \mathcal{T}$ meßbar ist.

Bemerkung: Nach dieser Definition sind alle linksstetigen, adaptierten Prozesse vorhersehbar. Jede meßbare Funktion auf \mathcal{T} ist, als stochastischer Prozeß betrachtet, natürlich auch vorhersehbar.

Definition 2.18 (Klasse D) Ein Submartingal X heißt von der Klasse **D** (Doob), wenn die Zufallsvariablen $X(T)$ (T beliebige Stopzeit) gleichmäßig integrierbar sind.

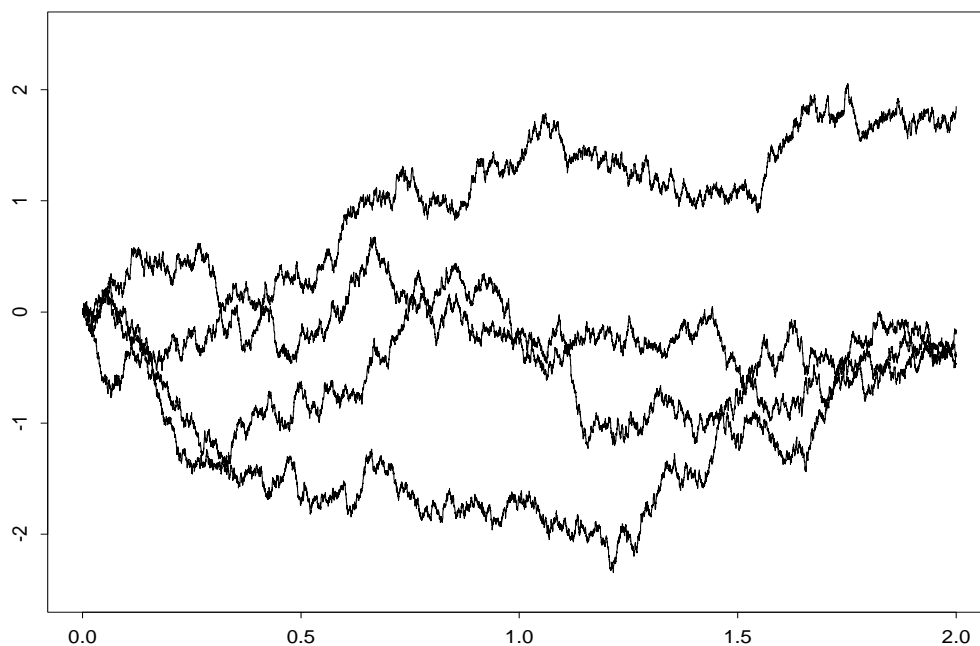


Abbildung 2.1: Simulierte Pfade einer Brown'schen Bewegung

Satz 2.19 (Doob-Meyer-Zerlegung) Sei X ein Submartingal der Klasse D . Dann existiert ein (bis auf Ununterscheidbarkeit) eindeutiger vorhersehbarer càdlàg Prozeß \tilde{X} , so daß

$$M := X - \tilde{X}$$

ein gleichmäßig integrierbares Martingal mit $M(0) = 0$ ist.

Beweis: Einen Beweis des Satzes findet man in Rogers und Williams [34] Seiten 153-154.

Bemerkung: Der Prozeß \tilde{X} heißt **Kompensator-Prozeß**

Beispiel 2.20 (Poisson-Prozeß) Sei N ein stochastischer Prozeß, dessen Pfade Werte in \mathbb{N}_0 annehmen, d.h.

$$N : \Omega \times \mathcal{T} \mapsto \mathbb{N}_0,$$

\mathcal{F}_t die von N erzeugte σ -Algebra und $\lambda > 0$. Für den Prozeß N gelten folgende Eigenschaften

(i)

$$P(N(t) = k) = \begin{cases} \frac{e^{-\lambda t} (\lambda t)^k}{k!} & \text{für } t < \infty \\ 0 & \text{sonst} \end{cases},$$

(ii) Die Zuwächse sind unabhängig Poisson-verteilt auf T , d.h. für $s < t < \infty$ gilt

$$P(N(t) - N(s) = k) = \frac{e^{-\lambda(t-s)} (\lambda(t-s))^k}{k!},$$

dann nennt man N einen **homogenen Poisson Prozeß**. Für $\lambda = 1$ heißt N **Standard Poisson Prozeß**. Der homogene Poisson Prozeß besitzt folgenden Erwartungswert

$$E(N(t)) = \sum_{k=0}^{\infty} k \frac{e^{-\lambda t} (\lambda t)^k}{k!} = \lambda t e^{-\lambda t} \sum_{k=1}^{\infty} \frac{(\lambda t)^{k-1}}{(k-1)!} = \lambda t e^{-\lambda t} e^{\lambda t} = \lambda t.$$

Analog gilt für den erwarteten Zuwachs für $s < t < \infty$

$$E(N(t) - N(s)) = \lambda(t - s).$$

Mit diesen Vorbereitungen können wir nun zeigen, daß λt der vorhersehbare Kompensatorprozeß des homogenen Poisson Prozeß ist. Die Vorhersehbarkeit ist klar, da λt eine deterministische Funktion ist. Jetzt bleibt noch zu zeigen, daß der Prozeß $M(t) = N(t) - \lambda t$ ein Martingal ist. Dazu müssen wir die Martingaleigenschaft zeigen (die Existenz des Erwartungswertes haben wir bereits gezeigt). Für $s < t < \infty$ gilt:

$$E(M(t)|\mathcal{F}_s) = E(N(t) - \lambda t|\mathcal{F}_s) = E(N(t)|\mathcal{F}_s) - E(\lambda t|\mathcal{F}_s) \stackrel{\text{Satz 2.2}}{=} N(s) + \lambda(t - s) - \lambda t = M(s).$$

□

Kommen wir nun zum Konzept der optionalen und vorhersehbaren Variation, letztere stellt ein Maß für das systematische Wachsen des Prozesses M^2 (M lokal quadratintegrierbares Martingal) dar (siehe auch Rogers und Williams [35] Seiten 42-51).

Definition 2.21 (vorhersehbare und optionale Variation) Seien M und M' lokale quadratintegrierbare Martingale, dann ist M^2 ein lokales Submartingal und MM' die Differenz zweier Submartingale, denn

$$MM' = \frac{1}{4}(M + M')^2 - \frac{1}{4}(M - M')^2.$$

Die zugehörigen Kompensatoren für M^2 und MM' , die mit $\langle M, M \rangle$ bzw. $\langle M, M' \rangle$ bezeichnet werden, heißen **vorhersehbarer Variationsprozeß**, bzw. **vorhersehbarer Kovariationsprozeß**. Falls

$$\langle M, M' \rangle = 0,$$

heißen M und M' **orthogonal**. Den vorhersehbaren Variationsprozeß $\langle M, M \rangle(t)$ erhält man, indem man den Grenzwert

$$\langle M \rangle = \langle M, M \rangle(t) = \lim_{n \rightarrow \infty} \sum_{i=1}^n \text{Var}(M(t_i) - M(t_{i-1})|\mathcal{F}_{t_{i-1}})$$

über die Partitionen $0 = t_1 < t_2 < \dots < t_n = t$ bildet. Die Prozesse $[M, M]$ und $[M, M']$, die durch

$$[M](t) = [M, M](t) = \sum_{s \leq t} \Delta M(s)^2 \quad \text{bzw.} \quad (2.4)$$

$$[M, M'](t) = \sum_{s \leq t} \Delta M(s) \Delta M'(s) \quad (2.5)$$

definiert sind, heißen **optionaler (Ko)variationsprozeß**.

Da in der späteren Anwendung oft stochastische Integrale mit vorhersehbaren Prozessen als Integranden und lokalen (bzw. lokal quadratintegrierbaren) Martingalen als Integratoren Verwendung finden, werden im folgenden Abschnitt einige Ergebnisse zur stochastischen Integration solcher Prozesse sowie deren vorhersehbare und optionale (Ko)variationsprozesse behandelt, zunächst wollen wir aber noch einen kurzen Blick auf die stochastische Integration vorhersehbarer Prozesse hinsichtlich zeitdiskreter bzw. zeitstetiger Sub- bzw. Supermartingale (im folgenden mit Semimartingal bezeichnet) werfen.

Definition 2.22 (stochastische Integration bzgl. zeitdiskretem Semimartingal)

Betrachten wir den mit einer Filtration \mathcal{F}_n $n \in \mathbb{N}$ versehenen Wahrscheinlichkeitsraum (Ω, \mathcal{F}, P) . Sei M_n $n \in \mathbb{N}$ ein adaptierter Prozeß, der die (Sub-, oder Super-) Martingal Eigenschaften erfüllt, sowie H_n ein vorhersehbarer Prozeß (im diskreten Fall bedeutet dies H_n ist \mathcal{F}_{n-1} -meßbar). Dann ist das stochastische Integral $(\int H dM)_n$ definiert als

$$\left(\int H dM \right)_n = \sum_{k=1}^n H_k (M_k - M_{k-1}).$$

Um eine analoge Definition des stochastischen Integrals bezüglich zeitstetige Semimartingale zu erhalten, müssen wir zunächst den Begriff des elementaren vorhersehbaren Prozesses definieren.

Definition 2.23 (elementarer vorhersehbarer Prozeß) Sei (Ω, \mathcal{F}, P) mit der Filtration \mathcal{F}_t ein filtrierter Wahrscheinlichkeitsraum. Seien $a, b \in \mathcal{T}$, $C \in \mathcal{F}_a$. Dann nennt man den stochastischen Prozeß

$$H(s, \omega) := I(a < s \leq b)C(\omega)$$

einen **elementaren vorhersehbaren Prozeß**.

Für elementare Prozesse ist das stochastische Integral wie folgt definiert

Definition 2.24 (Integration eines elementaren vorhersehbaren Prozesses) Sei $H(s, \omega) = I(a < s \leq b)C(\omega)$ ein elementarer stochastischer Prozeß und M ein lokales càdlàg Martingal, so definieren wir mit

$$\int H(s) dM(s) = C(\omega)(M(b, \omega) - M(a, \omega)),$$

sowie mit

$$\int_0^t H(s) dM(s) = \int_0^t H(s) I(0 \leq s \leq t) dM(s)$$

das stochastische Integral des elementaren Prozesses hinsichtlich des Martingals.

Das folgende Korollar zeigt, daß bei der Integration beschränkter elementarer Prozesse hinsichtlich lokaler Martingale die Martingaleigenschaft erhalten bleibt.

Korollar 2.25 Sei $H(s, \omega) = C(\omega)I(a < s \leq b)$ ein elementarer stochastischer Prozeß mit

$$\sup |H(s, \omega)| < \infty$$

und M ein Martingal, dann ist

$$\int_0^t H(s) dM(s)$$

ebenfalls ein Martingal.

Beweis:

$$\int_0^t H(s) dM(s) = \begin{cases} 0 & 0 \leq t < a \\ C(M(t) - M(a)) & a \leq t < b \\ C(M(b) - M(a)) & b \leq t < \infty \end{cases}.$$

Nachdem $M(t) \in \mathcal{F}_t$ ist und $C \in \mathcal{F}_a$, sieht man sofort, daß $\int_0^t H(s) dM(s) \in \mathcal{F}_t$. Aus der Beschränktheit von H folgt, daß

$$E \left| \int_0^t H(s) dM(s) \right| < \infty.$$

Nun müssen wir nur noch die Martingal-Eigenschaft $E(M(t)|\mathcal{F}_s) - M(s) = 0$ zu beweisen. Betrachten wir zunächst den Fall $a \leq s < t \leq b$

$$\begin{aligned} & E \left(\int_0^t H(u) dM(u) \middle| \mathcal{F}_s \right) - \int_0^s H(u) dM(u) \\ \stackrel{\text{Satz 2.2 A}}{=} & E \left(\int_0^t H(u) dM(u) - \int_0^s H(u) dM(u) \middle| \mathcal{F}_s \right) \\ = & E(C(M(t) - M(s)) | \mathcal{F}_s) \\ \stackrel{\text{Satz 2.2 C}}{=} & CE(M(t) - M(s) | \mathcal{F}_s) = 0. \end{aligned}$$

Definition 2.26 (einfacher vorhersehbarer Prozeß) Sei (Ω, \mathcal{F}, P) mit der Filtration \mathcal{F}_t ein filtrierter Wahrscheinlichkeitsraum, dann nennen wir die endliche Summe elementarer Prozesse

$$H(s, \omega) = \sum_{i=1}^n H_i(s, \omega) \quad n \in \mathbb{N}$$

einen **einfachen vorhersehbaren Prozess**.

Definition 2.27 (Integration eines einfachen vorhersehbaren Prozesse)

Für $H = H_1 + \dots + H_n$ ein einfacher stochastischer Prozess und M ein lokales càdlàg Martingal, so definieren wir mit

$$\int H(s) dM(s) = \sum_{i=1}^n \int H_i(s) dM(s),$$

sowie mit

$$\int_0^t H(s) dM(s) = \sum_{i=1}^n \int_0^t H_i(s) dM(s)$$

das stochastische Integral für den Prozess H .

Bemerkung: Als endlich Summe lokaler Martingale ist das stochastische Integral

$$\int H(s) dM(s) = \sum_{i=1}^n \int H_i(s) dM(s),$$

wieder ein Martingal.

Die Klasse der vorhersehbaren Prozesse erhält man nun als Grenzwerte (für $n \rightarrow \infty$) einfacher Prozesse, d.h. jeder vorhersehbare Prozess H ist durch

$$H = \lim_{n \rightarrow \infty} \sum_{i=1}^n H_i$$

darstellbar (siehe Durrett [11] Seiten 57 - 58).

Falls wir an dieser Stelle zusätzlich voraussetzen, daß der Integrator M von lokal beschränkter Variation ist, können wir das stochastische Integral, analog zum Lebesgue-Stiltjes-Integral (vergleiche Elstrodt [13]), durch Grenzwertbildung auf die Klasse der vorhersehbaren Prozesse erweitern.

Mit Hilfe der stochastischen Integration können wir nun einige wichtigen Zusammenhänge zwischen Semimartingalen und deren optionalen-, bzw. vorhersehbaren Variationsprozessen herausarbeiten.

Korollar 2.28 *Für den optionalen (Ko)variationsprozeß gilt:*

(i)

$$[M](t) = M(t)^2 - 2 \int_0^t M(s-) dM(s), \quad (2.6)$$

(ii)

$$[M, M'](t) = M(t)M'(t) - \int_0^t M(s-)dM'(s) - \int_0^t M'(s-)dM(s). \quad (2.7)$$

Den Beweis hierzu findet man in Rogers und Williams [35] Seite 59.

Satz 2.29 (Fundamentaltheorem) *Wenn H ein lokal beschränkter, vorhersehbarer Prozeß ist und M ein lokales Martingal, dann existiert $\int H dM$ und ist ein lokales Martingal.*

Beweis:

Der folgende Beweis ist für das zeitdiskrete Analogon durchgeführt, für den zeitstetigen Fall findet sich der Beweis in Rogers und Williams (1987) [35].

Sei $(\Omega, \mathcal{F}_n, P)$ ein filtrierter Wahrscheinlichkeitsraum und H ein beschränkter, vorhersehbarer Prozeß ($H_n \in \mathcal{F}_{n-1}$ für $n \in \mathbb{N}$). Dann ist:

$$\left(\int H dM \right)_n = \sum_{k=1}^n H_k (M_k - M_{k-1})$$

mit $(\int H dM)_0 = 0$ ein lokales Martingal. Dazu muß man zeigen, daß

$$E \left[\left(\int H dM \right)_n \middle| \mathcal{F}_{n-1} \right] = \left(\int H dM \right)_{n-1} \iff E \left[\left(\int H dM \right)_n - \left(\int H dM \right)_{n-1} \middle| \mathcal{F}_{n-1} \right] = 0.$$

Da aber $H_n \mathcal{F}_{n-1}$ meßbar ist gilt:

$$\begin{aligned} E \left[\left(\int H dM \right)_n - \left(\int H dM \right)_{n-1} \middle| \mathcal{F}_{n-1} \right] &= E[H_n(M_n - M_{n-1}) | \mathcal{F}_{n-1}] \\ &= H_n E[M_n - M_{n-1} | \mathcal{F}_{n-1}] = 0, \end{aligned}$$

und damit ist die Aussage gezeigt. □

Mit Hilfe von Korollar 2.28 und Satz 2.29 sehen wir, daß der stochastische Prozeß

$$M^2 - [M] \tag{2.8}$$

ein lokales Martingal ist, denn es gilt für alle $s < t$.

$$\begin{aligned} E[M^2(t) - [M](t) | \mathcal{F}_s] &= E[M^2(t) - \sum_{s \leq t} \Delta M(s)^2 | \mathcal{F}_s] \\ &\stackrel{\text{Kor 2.28}}{=} E[M^2(t) - (M^2(t) - 2 \int_0^t M(u-) dM(u)) | \mathcal{F}_s] \\ &= \underbrace{E[M^2(t) - M^2(t) | \mathcal{F}_s]}_0 + 2E[\int_0^t \underbrace{M(u-)}_{\text{vorhersehbar}} dM(u) | \mathcal{F}_s] \\ &\stackrel{\text{Satz 2.29}}{=} 2 \int_0^s M(u-) dM(u) \\ &\stackrel{\text{Kor 2.28}}{=} M^2(s) - [M](s). \end{aligned}$$

Der folgende Satz enthält ein wichtiges Resultat zum Verhalten des vorhersehbaren, bzw. optionalen (Ko)variationsprozesses bei stochastischer Integration.

Satz 2.30 *Sei M ein lokal integrierbares Martingal von endlicher Variation, H ein vorhersehbarer Prozeß und $\int H^2 d[M]$ lokal integrierbar, oder $\int H^2 d\langle M \rangle$ lokal endlich (automatisch gegeben, wenn H lokal beschränkt). Dann ist $\int H dM$ ein quadratintegrierbares Martingal und*

$$\begin{aligned} \left[\int H dM \right] &= \int H^2 d[M] \\ \left\langle \int H dM \right\rangle &= \int H^2 d\langle M \rangle. \end{aligned}$$

Dieser Satz wird zum Beispiel in Durett (1984) [11] Seiten 57-59 bewiesen.

2.3 Zählprozeßtheorie

In diesem Kapitel werden die wesentlichen Eigenschaften von Zählprozessen und deren Kompensatoren beschrieben. Die Darstellung ist an Brémaud (1981) [7] Seite 18-48 orientiert.

Definition 2.31 (univariater Zählprozeß) Gegeben sei ein mit einer Filtration \mathcal{F}_t , die die gewöhnlichen Bedingungen bis auf (2.3) (\mathcal{F}_t muß nicht unbedingt vollständig sein) versehener Wahrscheinlichkeitsraum (Ω, \mathcal{F}, P) . Ein auf diesem filtrierten Wahrscheinlichkeitsraum definierter stochastischer Prozeß N heißt univariater Zählprozeß, falls er folgende Eigenschaften besitzt

- (i) $N(0) = 0$,
- (ii) N ist càdlàg und adaptiert bezüglich der Filtration \mathcal{F}_t ,
- (iii) N ist stückweise konstant und an den Sprüngen gilt

$$\Delta N(t) = +1.$$

Definition 2.32 (multivariater Zählprozeß) Sei (Ω, \mathcal{F}, P) ein Wahrscheinlichkeitsraum und \mathcal{F}_t eine Filtration die (2.1) und (2.2) erfüllt. Dann heißt der Vektor von adaptierter càdlàg Prozessen $N = (N_1, \dots, N_k)$ **k-variater Zählprozeß**, falls jede Komponente von N_k ein univariater Zählprozeß ist, wobei keine 2 Komponenten zur gleichen Zeit springen, d.h.

$$P(\Delta N_i \Delta N_j (1 - \delta_{ij}) = 1) = 0 \quad i, j \in \{1, \dots, k\},$$

mit

$$\delta_{ij} = \begin{cases} 1 & \text{falls } i = j \\ 0 & \text{sonst} \end{cases}.$$

Aufgrund der Tatsache, daß es keine gleichzeitigen Sprünge gibt, ist der Prozeß

$$N. = \sum_{i=1}^k N_i$$

auch ein Zählprozeß. Nun kann man für N . eine Folge von Zufallsvariablen

$$0 < T_1 \leq T_2 \leq T_3 \leq \dots$$

mit Werten aus \mathcal{T} , sowie

$$J_1, J_2, \dots,$$

die Werte aus $1, 2, \dots, k \cup 0$ annehmen kann, so daß für $n \leq N.(\tau)$ gilt:

$$\begin{aligned} T_n &\in \mathcal{T} = [0, \tau), \\ J_n &\neq 0, \\ T_n &> T_{n-1}, \end{aligned}$$

sowie

$$N.(T_n) = n \text{ und } \Delta N_{J_n}(T_n) = 1.$$

Für $n > N.(\tau)$ sei $T_n = \tau$ und $J_n = 0$. Da $\{T_n \leq t\}$ \mathcal{F}_{T_n} -meßbar ist, ist T_n eine Folge von Stopzeiten. Auch J_n ist \mathcal{F}_{T_n} -meßbar. Man kann die Zufallsvariablen T_n als Zeitpunkt, zu dem der Zählprozeß $N.$ auf den Wert n springt und die Zufallsvariable J_n als Index der zum Zeitpunkt n springenden Komponente interpretieren. Da für jede Komponente N_h , $h \in \{1, \dots, k\}$ des Zählprozesses $N.$ gilt: $0 \leq N_h^{T_n} \leq n$, ist T_n eine lokalisierende Folge von Stopzeiten für N_h . Da alle Pfade von N_h monoton wachsend sind gilt

$$\begin{aligned} E(N_h(s)|\mathcal{F}_t) &= E(N_h(t) + N_h(s) - N_h(t)|\mathcal{F}_t) \\ &= N_h(t) + E(\underbrace{N_h(s) - N_h(t)}_{\geq 0}|\mathcal{F}_t) \geq N_h(t) \quad \forall t \leq s. \end{aligned}$$

N_h ist also ein lokales Submartingal und für alle Stopzeiten T gleichmäßig integrierbar, d.h. N_h ein Submartingal der Klasse D . Nach Satz 2.19 besitzt N_h einen (bis auf Ununterscheidbarkeit) eindeutigen vorhersehbaren Kompensatorprozeß Λ_h , so daß der Prozeß

$$M_h = N_h - \Lambda_h$$

ein lokales Martingal ist.

Einige wichtige Eigenschaften dieses Kompensatorprozesses Λ_h , wollen wir nun genauer betrachten. Dazu sind jedoch noch einige vorbereitende Definitionen notwendig.

Satz und Definition 2.33 (Intensität eines Zählprozesses) Sei $N(t)$ ein Zählprozeß. Ein vorhersehbarer nichtnegativer Prozeß $\lambda(t)$, für den gilt

(i)

$$\int_0^t \lambda(s) ds < \infty \quad \forall t \text{ } P\text{-fast sicher.}$$

(ii) Für alle vorhersehbaren Prozesse $H(t)$ ist die Gleichung

$$E \left[\int_0^\infty H(s) dN(s) \right] = E \left[\int_0^\infty H(s) \lambda(s) ds \right]$$

erfüllt.

heißt **Intensitätsprozeß** und ist, bis auf Ununterscheidbarkeit, eindeutig definiert.

Bemerkungen:

- Brémaud verzichtet in seiner Definition für den Intensitätsprozeß auf die Vorhersehbarkeit und zeigt dann, daß man zu jedem so definierten Intensitätsprozeß $\lambda(t)$ eine bis auf Ununterscheidbarkeit eindeutige vorhersehbare Version $\tilde{\lambda}(t)$ finden kann.

- Eine äquivalente Definition kann gegeben werden durch

$$\lambda(t) = \lim_{dt \rightarrow 0} \frac{E[N(t+dt) - N(t) | \mathcal{F}_t]}{dt}.$$

Aven (1982) [3] beweist, daß unter bestimmten Voraussetzungen mit diesem so definierten $\lambda(t)$ der Prozess $N(t) - \int_0^t \lambda(s) ds$ ein Martingal ist, d.h. wir können den vorhersehbaren Kompensatorprozeß eines Zählprozesses angeben durch

$$\Lambda(t) = \int_0^t \lambda(s) ds.$$

- Aus der Existenz eines Intensitätsprozesses folgt, daß die Zuwächse in einem Zählprozeß unabhängig sind, der Beweis findet sich in Brémaud [7](Seite 25). Im folgenden Text werden wir immer Zählprozesse mit existierendem vorhersehbarem Intensitätsprozeß, d.h. mit unabhängigen Zuwächsen betrachten.

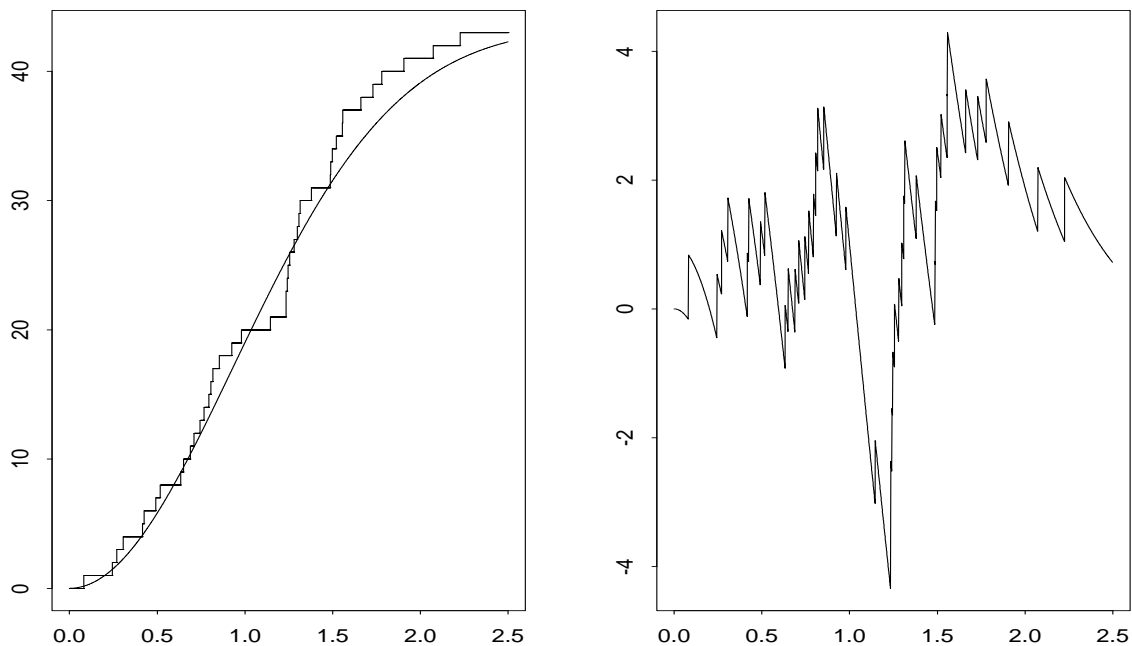


Abbildung 2.2: Simulierter Pfad für Zählprozeß von 50 Objekten mit Hazard $\lambda(t) = t$ und Zensierungsintensität $\lambda_Z(t) = 0.15 t$, sowie zugehöriger Martingalpfad

Beispiel 2.34 (simulierter Zählprozeß) *In diesem Beispiel soll der Zusammenhang zwischen Zählprozessen und deren Kompensatoren veranschaulicht werden. In einer Simulation betrachten wir 50 Individuen, die mit einer Intensität von $\lambda(t) = t$ ein terminierendes Ereignis erleben und dann aus der Risikomenge $R(t)$ (= Anzahl der Personen unter Beobachtung) ausscheiden. Zudem sei in dem Bestand eine Zensierungsintensität von $\lambda_Z(t) = 0.15 t$ gegeben, das bedeutet, daß*

mit einer Intensität von $0.15 t$ Individuen aus dem Bestand ausscheiden, ohne daß ein Ereignis stattfindet. Der Zählprozeß $N(t)$ ist nun folgendermaßen definiert.

$$N(t) = \text{Anzahl der Ereignisse bis zum Zeitpunkt } t.$$

Der so definierte Prozeß hat folgenden Kompensatorprozeß $\Lambda(t)$:

$$\begin{aligned} \Lambda(t) &= E \left[\int_0^t \lambda(s) R(s) ds \right] \\ &= \int_0^t \lambda(s) E[R(s)] ds \\ &= \int_0^t \lambda(s) 50 \exp \left(- \int_0^s (\lambda(u) + \lambda_Z(u)) du \right) ds \\ &= \int_0^t 50 s \exp \left(- \int_0^s 1.15u du \right) ds \\ &= 50 \int_0^t s \exp \left(- \frac{1.15}{2} s^2 \right) ds \\ &= \frac{50}{1.15} \left(1 - e^{-\frac{1.15}{2} t^2} \right). \end{aligned}$$

Bemerkung: In diesem Beispiel wird ein Ergebnis aus Kapitel 3 verwendet, für den Erwartungswert $E[R(t)]$ gilt

$$\begin{aligned} E[R(t)] &= 50 P(\text{Person zum Zeitpunkt } t \text{ unter Beobachtung}) \\ &= 50 S(t) \\ &\stackrel{\text{Satz 3.3}}{=} 50 \exp \left(- \int_0^s (\lambda(u) + \lambda_Z(u)) du \right). \end{aligned}$$

Da $\Lambda(t)$ der Kompensatorprozeß von $N(t)$ ist, ist $M(t) = N(t) - \Lambda(t)$ ein Martingal. In Abbildung 2.2 ist ein simulierter Pfad von $N(t)$ und dessen Kompensatorprozeß $\Lambda(t)$ (links), sowie das zugehörige Martingal $M(t)$ (rechts) dargestellt.

Der folgende Satz zeigt einen wichtigen Zusammenhang zwischen Zählprozessen und deren Intensitätsprozessen.

Satz 2.35 Gegeben sei der Zählprozeß $N(t)$ mit Intensitätsprozeß $\lambda(t)$. Sei $M(t) = N(t) - \int_0^t \lambda(s) ds$ und H lokal beschränkt und vorhersehbar. Dann gilt:

$$\langle M \rangle(t) = \int_0^t \lambda(s) ds, \quad (2.9)$$

$$[M](t) = N(t), \quad (2.10)$$

$$\left\langle \int H dM \right\rangle(t) = \int_0^t H^2(s) \lambda(s) ds, \quad (2.11)$$

$$\left[\int H dM \right](t) = \int_0^t H^2(s) dN(s). \quad (2.12)$$

Beweis:

Um die Aussage zu beweisen, zeigen wir zunächst die Gültigkeit von (2.10). Es gilt :

$$[M(t)] = \sum_{0 < s \leq t} (\Delta M(t))^2 = \sum_{0 < s \leq t} \Delta(N(s) - \int_0^s \lambda(u) du)^2.$$

Da $\Lambda(t) = \int_0^t \lambda(s) ds$ ein stetiger Prozess ist, hat $M(t)$ Sprünge der Höhe +1 an den Stellen, an denen $N(t)$ springt. Aus diesem Zusammenhang erhalten wir

$$\begin{aligned} [M(t)] &= \sum_{0 < s \leq t} \Delta(N(s) - \int_0^s \lambda(u) du)^2 \\ &= \sum_{0 < s \leq t} (\Delta N(s))^2 \\ &= \sum_{0 < s \leq t} \Delta N(s) = N(t). \end{aligned}$$

Um die Gültigkeit von (2.10) zu zeigen, benutzen wir die Tatsache, daß

$$M^2(t) - [M](t)$$

ein Martingal ist (siehe (2.8)). Der vorhersehbare Kompensatorprozeß von $[M](t) = N(t)$ ist gegeben durch $\Lambda(t) = \int_0^t \lambda(s) ds$, so daß

$$M^2(t) - [M](t) + N(t) - \int_0^t \lambda(s) ds$$

als Summe zweier Martingale wiederum ein Martingal ist und es gilt

$$\begin{aligned} &M^2(t) - [M](t) + N(t) - \int_0^t \lambda(s) ds \\ &= M^2(t) - [M](t) + [M](t) - \int_0^t \lambda(s) ds \\ &= M^2(t) - \int_0^t \lambda(s) ds. \end{aligned}$$

Da $\int_0^t \lambda(s) ds$ vorhersehbar ist, ist dies somit der (bis auf Ununterscheidbarkeit) eindeutige vorhersehbare Kompensatorprozeß von M^2 und damit gilt

$$\langle M \rangle = \int_0^t \lambda(s) ds.$$

(2.11) und (2.12) erhalten wir nun direkt aus Satz 2.30. Danach gilt:

$$\begin{aligned} \left\langle \int H^2 dM \right\rangle (t) &= \int_0^t H^2(s) d\langle M \rangle(s) \\ &= \int_0^t H^2(s) d\left(\int_0^s \lambda(u) du\right) \\ &= \int_0^t H^2(s) \lambda(s) ds, \end{aligned}$$

sowie

$$\begin{aligned} \left[\int H^2 dM \right] (t) &= \int_0^t H^2(s) d[M(s)] \\ &= \int_0^t H^2(s) dN(s). \end{aligned}$$

Damit ist der Satz bewiesen. \square

Bemerkung: Für einen k -variater Zählprozeß $N = (N_1, \dots, N_k)$ mit Intensitätsprozeß $\lambda = (\lambda_1, \dots, \lambda_k)$ und eine $p \times k$ -Matrix lokal beschränkter vorhersehbarer Prozesse H gelten analog zu 2.9 bis 2.12 die entsprechenden Zusammenhänge für die vorhersehbare bzw. optionale Kovariationsmatrix

$$\begin{aligned} \langle M \rangle (t) &= \text{diag} \int_0^t \lambda(s) ds, \\ [M] (t) &= \text{diag}(N(t)), \\ \left\langle \int H dM \right\rangle (t) &= \int_0^t H(s) \text{diag}(\lambda(s)) H^t(s) ds, \\ \left[\int H dM \right] (t) &= \int_0^t H(s) H(s)^t d(\text{diag} N(s)), \end{aligned}$$

wobei mit $\text{diag}(v)$, $v \in \mathbb{R}^p$ eine Matrix $V \in \mathbb{R}^{p \times p}$ definiert ist mit dem Vektor v als Eintrag auf der Diagonalen, d.h. $V_{ij} = \delta_{ij} v_i$. Der vorhersehbare, bzw. optionale Kovariationsprozeß eines k -variater Zählprozesses ist also eine $k \times k$ -Matrix mit folgenden Einträgen

$$\begin{aligned} \langle M_h, M_l \rangle (t) &= \delta_{hl} \int_0^t \lambda_h(s) ds, \\ [M]_{hl}(t) &= \delta_{ij} N_h(t), \\ \left\langle \sum_h \int H_{jh} dM_h, \sum_l \int H_{j'l} dM_l \right\rangle (t) &= \sum_h \int_0^t H_{jh}(s) H_{j'h}(s) \lambda_h(s) ds, \\ \left[\sum_h \int H_{jh} dM_h, \sum_l \int H_{j'l} dM_l \right] (t) &= \sum_h \int_0^t H_{jh}(s) H_{j'h}(s) dN_h(s). \end{aligned} \quad (2.13)$$

2.4 Grenzwerttheorie

Im diesem Abschnitt werden 2 wichtige Resultate über asymptotische Eigenschaften von Martingalen und deren Kompensatoren vorgestellt. Dazu sind aber zunächst noch einige vorbereitende Erklärungen nötig. Sei $M^{(n)} = (M_1^{(n)}, \dots, M_k^{(n)})$ $n \in \mathbb{N}$ eine Folge quadratintegrierbarer Martingale (mit jeweils k Komponenten) und $M_\varepsilon^{(n)} = (M_{\varepsilon_1}^{(n)}, \dots, M_{\varepsilon_k}^{(n)})$ eine Folge quadratintegrierbarer Martingale, die alle Sprünge von $M^{(n)}$ enthält, deren Absolutbetrag größer als ε ist. Mit $M^{(n)}$ und $M_\varepsilon^{(n)}$ ist auch die Differenz $M^{(n)} - M_\varepsilon^{(n)}$ ein lokal quadratintegrierbares Martingal und nach der Definition von $M_\varepsilon^{(n)}$ gilt, daß $|\Delta M_h^{(n)} - \Delta M_{\varepsilon_h}^{(n)}| \leq \varepsilon \quad \forall h \in \{1, \dots, k\}$.

Sei nun $M^{(\infty)}$ ein Martingal mit vorhersehbarem und optionalem (Ko)variationsprozeß $[M^{(\infty)}] = \langle M^{(\infty)} \rangle = V$ ($V : \mathcal{T} \mapsto \mathbb{R}^{p \times p}$), zudem seien die Zuwächse von M p -variat normalverteilt d.h.

$M^{(\infty)}(t) - M^{(\infty)}(s) \sim N_p(0, V(t) - V(s))$. Wir bezeichnen mit $\rightarrow_{\mathcal{P}}$ die Konvergenz in Wahrscheinlichkeit, sowie mit $\rightarrow_{\mathcal{D}}$ die Konvergenz in Verteilung. Mit diesen Vorbereitungen kann man nun einen Grenzwertsatz für Martingale formulieren.

Satz 2.36 (Satz von Rebolledo) *Sei $\mathcal{T}_0 \subset \mathcal{T}$ und gelte eine der beiden folgenden Bedingungen:*

$$\begin{aligned} \langle M^{(n)} \rangle(t) &\rightarrow_{\mathcal{P}} V(t) \text{ für alle } t \in \mathcal{T}_0 \text{ für } n \rightarrow \infty, \\ [M^{(n)}](t) &\rightarrow_{\mathcal{P}} V(t) \text{ für alle } t \in \mathcal{T}_0 \text{ für } n \rightarrow \infty, \end{aligned} \quad (2.14)$$

sowie

$$\langle M_{\varepsilon h}^{(n)} \rangle(t) \rightarrow_{\mathcal{P}} 0 \text{ für alle } t \in \mathcal{T}_0, h \text{ und } \varepsilon > 0 \text{ für } n \rightarrow \infty. \quad (2.15)$$

Dann gilt:

$$(M^{(n)}(t_1), \dots, M^{(n)}(t_l)) \rightarrow_{\mathcal{D}} (M^{(\infty)}(t_1), \dots, M^{(\infty)}(t_l)).$$

Falls zusätzlich \mathcal{T}_0 dicht in \mathcal{T} liegt und τ enthält falls $\tau \in \mathcal{T}$, so gilt mit denselben Voraussetzungen:

$$M^{(n)} \rightarrow_{\mathcal{D}} M^{(\infty)} \text{ für } n \rightarrow \infty$$

und $\langle M^{(n)} \rangle$ sowie $[M^{(n)}]$ konvergieren gleichmäßig auf kompakten Teilmengen von \mathcal{T} gegen V .

Dieser Satz wurde von Rebolledo (1980) [33] bewiesen.

Formulieren wir nun die Bedingungen (2.14) und (2.15) für stochastische Integrale, deren Integrand ein Zählprozeß ist. Für jedes $n \in \mathbb{N}$ sei $N^{(n)}$ ein k_n -variater Zählprozeß mit Intensitätsprozeß $\lambda^{(n)}$. Sei $H^{(n)}$ eine $k \times k_n$ Matrix lokal beschränkter vorhersehbarer Prozesse, definieren wir nun für $j = \{1, \dots, k\}$ die lokalen Martingale

$$M_j^{(n)}(t) = \sum_{h=1}^{k_n} \int_0^t H_{jh}^{(n)}(s) (dN_h^{(n)}(s) - \lambda_h^{(n)}(s) ds),$$

sowie mit

$$M_{j\varepsilon}^{(n)}(t) = \sum_{h=1}^{k_n} \int_0^t H_{jh}^{(n)}(s) I(|H_{jh}^{(n)}(s)| > \varepsilon) (dN_h^{(n)}(s) - \lambda_h^{(n)}(s) ds) \quad (2.16)$$

ein lokales Martingal, das alle Sprünge von M_j enthält mit $|\Delta M_j| > \varepsilon$. Dann erhalten wir mit Hilfe von Satz 2.35 für die Bedingungen (2.14) und (2.15) aus Satz 2.36

$$\begin{aligned} \langle M_j^{(n)}, M_{j'}^{(n)} \rangle(t) &= \sum_{h=1}^{k_n} \int_0^t H_{jh}^{(n)}(s) H_{j'h}^{(n)}(s) \lambda_h^{(n)}(s) ds, \\ [M_j^{(n)}, M_{j'}^{(n)}](t) &= \sum_{h=1}^{k_n} \int_0^t H_{jh}^{(n)}(s) H_{j'h}^{(n)}(s) dN_h^{(n)}(s), \end{aligned} \quad (2.17)$$

sowie

$$\langle M_{j\varepsilon}^{(n)}, M_{j'\varepsilon}^{(n)} \rangle(t) = \sum_{h=1}^{k_n} \int_0^t (H_{jh}^{(n)}(s))^2 I(|H_{jh}^{(n)}(s)| > \varepsilon) \lambda_h^{(n)}(s) ds. \quad (2.18)$$

Eine weitere wichtige Aussage ist die Ungleichung von Lengart. Mit ihrer Hilfe ist es möglich, einen stochastischen Prozeß mit Hilfe seines Kompensatorprozesses abzuschätzen.

Satz 2.37 (Ungleichung von Lenglart) Sei X ein lokales Martingal auf $\mathcal{T} = [0, \tau]$ oder $\mathcal{T} = [0, \tau]$ und \tilde{X} der zugehörige Kompensatorprozess nach Satz 2.19. Dann gilt $\forall \delta, \eta > 0$

$$P(\sup_{\mathcal{T}} X > \eta) \leq \frac{\delta}{\eta} + P(\tilde{X}(\tau) > \delta). \quad (2.19)$$

Der Satz wurde von Lenglart (1977) [27] bewiesen. Im folgenden Text findet jedoch meist eine leicht modifizierte Version von (2.19) Anwendung. Für ein lokal quadratintegrierbares Martingal M (d.h. M^2 hat den Kompensator $\langle M \rangle$) gilt nämlich nach Satz 2.37 für $\delta, \eta^2 > 0$

$$P(\sup_{\mathcal{T}} M^2 > \eta^2) \leq \frac{\delta}{\eta^2} + P(\langle M(\tau) \rangle > \delta),$$

was wiederum äquivalent ist zu

$$P(\sup_{\mathcal{T}} |M| > \eta) \leq \frac{\delta}{\eta^2} + P(\langle M(\tau) \rangle > \delta). \quad (2.20)$$

2.5 Markovketten

Definition 2.38 (Markov-Kette) Ein zeitstetiger Prozeß $X(t), t \geq 0$ mit abzählbarem Zustandsraum \mathcal{S} heißt zeitstetige **Markovkette**, wenn für alle n und jede endliche Menge von Zeitpunkten $0 \leq t_0 < \dots < t_n < u$ mit

$$P(X(t_0) = i_0 \wedge \dots \wedge X(t_n) = i_n \wedge X(u) = j) > 0,$$

die folgende Eigenschaft (Markov-Eigenschaft)

$$P(X(u) = j | X(t_0) = i_0 \wedge \dots \wedge X(t_n) = i_n) = P(X(u) = j | X(t_n) = i_n) \quad (2.21)$$

erfüllt ist.

Definition 2.39 (Übergangswahrscheinlichkeit) Die bedingten Wahrscheinlichkeiten

$$P_{ij}(t, u) := P(X(u) = j | X(t) = i)$$

heißen **Übergangswahrscheinlichkeiten**, mit

$$P_{\underline{ii}}(t, u) := P(X(z) = i \text{ für alle } z \in [t, u])$$

ist die Wahrscheinlichkeit bezeichnet im gesamten Intervall $[t, u]$ im Zustand i zu verbleiben. Die Zustände kann man in folgende 3 Gruppen unterteilen:

- Ein Zustand i heißt **absorbierend**, falls

$$P_{\underline{ii}}(t, u) = 1 \quad (0 \leq t \leq u).$$

- Ein Zustand i heißt **transient**, falls

$$P_{ii}(t, \infty) = 0 \quad (t \geq 0).$$

- Ein Zustand heißt **strikt transient**, falls

$$P_{ii}(t, u) = P_{\underline{ii}}(t, u) < 1 \quad (0 \leq t \leq u).$$

Ein absorbierender Zustand ist also ein Zustand, der nicht mehr verlassen wird, falls er einmal eingetreten ist, ein transienter Zustand ist ein Zustand, der auch wieder verlassen werden kann und ein strikt transienter Zustand ist ein Zustand, der genau einmal erreicht werden kann (das heißt, nachdem der Zustand einmal verlassen wurde, ist es unmöglich wieder einzutreten).

Definition 2.40 (Übergangintensitäten) Die Übergangintensitäten sind definiert als:

$$\lambda_{ij}(t) = \lim_{dt \rightarrow 0} \frac{P_{ij}(t, t + dt)}{dt}.$$

Zudem definieren als Gesamtintensität für das Verlassen des Zustands i

$$\lambda_i(t) = \sum_{j:j \neq i} \lambda_j(t).$$

Lemma 2.41 (Chapman-Kolmogorov) Für zeitstetige Markovketten erfüllen die Übergangswahrscheinlichkeiten die sogenannten Chapman-Kolmogorov-Gleichungen:

$$P_{ij}(t, u) = \sum_{k \in \mathcal{S}} P_{ik}(t, w) P_{kj}(w, u) \quad (t \leq w \leq u). \quad (2.22)$$

Beweis :

$$\begin{aligned} P_{ij}(t, u) &= P(X(u) = j | X(t) = i) \\ &= \sum_{k \in \mathcal{S}} P(X(u) = j \wedge X(w) = k | X(t) = i) \quad (t \leq w \leq u) \\ &= \sum_{k \in \mathcal{S}} P(X(u) = j | X(t) = i \wedge X(w) = k) P(X(w) = k | X(t) = i) \quad (t \leq w \leq u) \\ &\stackrel{(2.21)}{=} \sum_{k \in \mathcal{S}} P(X(u) = j | X(w) = k) P(X(w) = k | X(t) = i) \quad (t \leq w \leq u) \\ &= \sum_{k \in \mathcal{S}} P_{ik}(t, w) P_{kj}(w, u) \quad (t \leq w \leq u). \end{aligned}$$

Damit ist das Lemma gezeigt. □

Die Gleichungen von Chapman-Kolmogorov sind ein wichtiges Hilfsmittel bei der Herleitung der Kolmogorov-Differentialgleichungen.

Satz 2.42 (Kolmogorov-Differentialgleichungen) Für zeitstetige Markovketten gilt:

$$(i) \quad \frac{d}{dt}P_{ij}(z, t) = \sum_{k:k \neq j} P_{ik}(z, t)\lambda_{kj}(t) - P_{ij}(z, t)\lambda_j(t), \quad (2.23)$$

$$(ii) \quad \frac{d}{dz}P_{ij}(z, t) = P_{ij}(z, t)\lambda_i(z) - \sum_{k:k \neq i} P_{kj}(z, t)\lambda_{ik}(z). \quad (2.24)$$

Beweis:

(i) Nach (2.22) gilt:

$$P_{ij}(z, t + dt) = \sum_{k:k \neq j} P_{ik}(z, t)P_{kj}(t + dt) + P_{ij}(z, t)P_{jj}(t + dt).$$

Damit erhält man:

$$\frac{P_{ij}(z, t + dt) - P_{ij}(z, t)}{dt} = \sum_{k:k \neq j} P_{ik}(z, t) \frac{P_{kj}(t + dt)}{dt} + P_{ij}(z, t) \frac{P_{jj}(t, t + dt) - 1}{dt}.$$

Aus dem Zusammenhang, daß:

$$P_{jj}(t, t + dt) = 1 - \sum_{k \neq j} P_{jk}(t, t + dt)$$

folgt daß :

$$\frac{P_{ij}(z, t + dt) - P_{ij}(z, t)}{dt} = \sum_{k:k \neq j} P_{ik}(z, t) \frac{P_{kj}(t + dt)}{dt} + P_{ij}(z, t) \sum_{k:k \neq j} \frac{P_{jk}(t + dt)}{dt}.$$

Durch Grenzwertbildung $dt \rightarrow 0$ folgt schließlich die Aussage.

(ii) analog

□

Kapitel 3

Mathematische Grundlagen der Verweildaueranalyse

In diesem Kapitel werden die grundlegenden Konzepte der Verweildaueranalyse vorgestellt. Dies beinhaltet unter anderem verschiedene Ansätze zur Schätzung der Überlebensfunktion (nichtparametrische, semiparametrische und parametrische Methoden), die Modellierung von Maximum-Likelihood-Methoden für Überlebensdaten. Wichtige Standardwerke hierzu sind die Bücher von Kalbfleisch und Prentice (1980) [21], sowie von Klein und Moeschberger (1997) [25]. Eine sehr detailliertere Darstellung mit ausführlichen Beweisen von Konvergenzaussagen auf Basis von Zählprozessen findet man in Andersen u.a. (1993) [1]. Eine anwenderorientierte Beschreibung verweildaueranalytischer Methoden ist in dem Buch von Le (1997) [26] enthalten.

3.1 Die Überlebensfunktion

Die der Überlebenszeit zugrunde liegende Größe ist die zufällige Lebensdauer, oder die zufällige Dauer bis zu einem genau spezifizierten terminierenden Ereignis.

Beispiele:

- Dauer des Zeitraums von Geburt bis zum Tod,
- Dauer von Beginn einer Krankheit bis zur Genesung,
- ...

Definition 3.1 (Überlebensfunktion) *Gegeben sei die Zufallsvariable T mit mit Verteilungsfunktion $F(t)$ und Dichte $f(t)$. Dann ist die **Überlebensfunktion** (engl.: *Survival Function*) definiert als*

$$S(t) := P(T > t). \quad (3.1)$$

Bemerkung: Die Überlebensfunktion kann als Anteil der Überlebenden zum Zeitpunkt t gedeutet werden. Die Überlebensfunktion $S(t)$ läßt sich wie folgt als Funktion der Dichte und Verteilungsfunktion von T darstellen:

$$S(t) = P(T > t) = 1 - P(T \leq t) = 1 - F(t) = 1 - \int_0^t f(x) dx, \quad (3.2)$$

$$f(t) = -\frac{dS(t)}{dt}. \quad (3.3)$$

3.2 Die Hazardfunktion

Neben der Überlebensfunktion $S(t)$ ist die Hazardfunktion $\lambda(t)$ eine wichtige Größe in der Überlebenszeitanalyse. Die Hazardfunktion, in deutschsprachiger Literatur oft als Sterbeintensität bezeichnet, ist die bedingte Dichte von T , gegeben Überleben bis t .

Definition 3.2 Die Hazard-, oder Intensitätsfunktion ist definiert als

$$\lambda(t) = \lim_{dt \rightarrow 0} \frac{P(t \leq T < t + dt | T \geq t)}{dt}. \quad (3.4)$$

Die kumulative Hazardfunktion definiert man als

$$\Lambda(t) = \int_0^t \lambda(x) dx. \quad (3.5)$$

Zwischen Überlebens- und Hazardfunktion besteht folgender Zusammenhang:

Satz 3.3 Seien $S(t)$ Überlebensfunktion und $\lambda(t)$ Hazardfunktion zu Lebensdauer T . Dann gilt

(i)

$$S(t) = \exp\left(-\int_0^t \lambda(x) dx\right), \quad (3.6)$$

(ii)

$$\mu := E(T) = \int_0^\infty S(t) dt. \quad (3.7)$$

Beweis:

(i)

$$\begin{aligned} \lambda(t) &= \lim_{dt \rightarrow 0} \frac{P(t \leq T < t + dt | T \geq t)}{dt} \\ &= \lim_{dt \rightarrow 0} \frac{P(t \leq T < t + dt)}{dt P(T \geq t)} \\ &= \frac{f(t)}{S(t)} = -\frac{\frac{d}{dt} S(t)}{S(t)} = -\frac{d}{dt} \log(S(t)). \end{aligned}$$

Mittels Integration erhält man:

$$\ln(S(t)) = - \int_0^t \lambda(x) dx \iff S(t) = \exp\left(- \int_0^t \lambda(x) dx\right).$$

(ii)

$$\mu = E(T) = \int_0^\infty t dF(t) = \int_0^\infty \left[\int_0^t dx \right] dF(t) = \int_0^\infty \left[\int_x^\infty dF(t) \right] dx = \int_0^\infty S(x) dx.$$

□

Bemerkung: Nach Satz 3.3 ist also die Verteilung der Zufallsvariable T eindeutig durch $\lambda(t)$ bestimmt. Falls T eine diskrete Zufallsvariable ist mit Werten $x_1 < x_2 < \dots$ und

$$f(x_i) = P(T = x_i).$$

Dann ist die Überlebensfunktion eine Stufenfunktion mit

$$S(t) = 1 - \sum_{x_i < t} f(x_i) = \sum_{x_i \geq t} f(x_i).$$

Mit $\lambda_i := f(x_i)/S(x_i)$ erhält man einfache Ausdrücke für $f(x_i)$

$$f(x_i) = \lambda_i \prod_{j=1}^{i-1} (1 - \lambda_j), \quad (3.8)$$

$$S(x_i) = \prod_{j=1}^{i-1} (1 - \lambda_j). \quad (3.9)$$

3.3 Funktionale Darstellung der Hazardfunktion

Obwohl in der Arbeit semiparametrische Modelle im Vordergrund stehen, sollen hier einige wichtige funktionale Modelle in der Überlebenszeitanalyse vorgestellt werden.

Exponentielles Modell

Das exponentielle Modell ist wohl das bekannteste Überlebenszeitmodell. Die Hazardfunktion für das exponentielle Modell ist konstant für alle t

$$\lambda(t) = \lambda \quad \lambda > 0.$$

Die zugehörige Überlebensfunktion $S(t)$ ist demnach

$$S(t) = \exp\left[- \int_0^t \lambda ds\right] = 1 - e^{-\lambda t}.$$

Das bedeutet für die Verteilungsfunktion $F(t)$ der Zufallsvariable T

$$F(t) = 1 - S(t) = e^{-\lambda t}.$$

Die Überlebenszeit T ist also exponentialverteilt mit Parameter λ . Aus dem konstanten Hazard λ folgt sofort folgende wichtige Eigenschaft

$$P(T \geq t + s | T \geq t) = \frac{S(t + s)}{S(t)} = \exp(-\lambda(t + s - t)) = \exp(-\lambda s) = P(T \geq s).$$

Weibull-Modelle

Eine Verallgemeinerung der exponentiellen Modelle sind Weibull-verteilte Lebensdauern T mit Hazardfunktion

$$\lambda(t) = \lambda \alpha t^{\alpha-1} \quad \lambda, \alpha > 0.$$

Mit dieser funktionalen Darstellung können sowohl monoton wachsende ($\alpha \geq 1$) als auch monoton fallende ($\alpha \leq 1$) Hazardfunktionen modelliert werden. Für die Überlebensfunktion $S(t)$ gilt

$$S(t) = \exp \left[- \int_0^t \lambda(u) du \right] = \exp[-\lambda t^\alpha].$$

Die Verteilungsfunktion $F(t)$ und die Dichtefunktion $f(t)$ für Weibull-verteilte Lebensdauern sind demnach gegeben durch

$$\begin{aligned} F(t) &= 1 - S(t) = 1 - \exp[-\lambda t^\alpha], \\ f(t) &= \frac{d}{dt} F(t) = \alpha \lambda t^{\alpha-1} \exp(-\lambda t^\alpha). \end{aligned}$$

Log-Logistische Modelle

Die Zufallsvariable Lebensdauer T heißt log-logistisch verteilt, falls die Transformation:

$$Y = \log(T)$$

logistisch verteilt ist mit Dichtefunktion

$$f_Y(y) = \frac{\exp\left(-\frac{y-\mu}{\sigma}\right)}{\sigma \left[1 + \exp\left(\frac{y-\mu}{\sigma}\right)\right]^2}$$

und Verteilungsfunktion

$$F_Y(y) = \frac{1}{1 + \exp\left(-\frac{y-\mu}{\sigma}\right)} \quad \text{für } \mu \in \mathbb{R}, \sigma > 0.$$

Die Überlebens- und Hazardfunktion log-logistisch verteilter Überlebenszeiten haben die Form

$$S(t) = 1 - F_Y(\log(t)) = \frac{1}{1 + \beta t^\alpha},$$

sowie

$$\lambda(t) = \frac{\alpha \beta t^{\alpha-1}}{1 + \beta t^\alpha},$$

mit $\alpha = 1/\sigma > 0$ und $\beta = \exp(-\mu/\sigma)$. Mit log-logistischer Verteilung kann man mit $\alpha \leq 1$ monoton fallende Hazardfunktionen modellieren, für $\alpha > 1$ erhält man Hazardfunktionen, die bis zum Zeitpunkt $t = [(\alpha-1)/(\beta)]^{1/\alpha}$ monoton wachsen und dann monoton fallen mit Grenzwert in 0.

3.4 Aufbau einer Studie

Um das Überlebensverhalten eines Bestandes zu untersuchen, betrachtet man eine Gruppe von Probanden über einen längeren Zeitraum. Dazu werden im wesentlichen 2 Typen von Studien unterschieden.

- Retrospektive Studien
- Prospektive Studien

Retrospektive Studien betrachten Daten aus der Vergangenheit einer Kohorte (Menge beobachteter Objekte) und versuchen anhand dieser Daten Aussagen über das Überlebensverhalten der entsprechenden Kohorte zu treffen. Als Beispiel wäre hier zum Beispiel die Erstellung einer Sterbetafel für eine bestimmte Population zu nennen. Hier werden Daten, die meist aus einer Volkszählung resultieren, benutzt, um Aussagen über die Sterblichkeit dieser Population treffen zu können. Ein Vorteil dieser retrospektiven Studien ist, daß sie relativ kostengünstig sind, die Daten sind ja bereits erhoben, und zu schnellen Resultaten führen.

Im Gegensatz dazu wird für prospektive Studien eine Kohorte von Probanden über einen bestimmten Zeitraum $(0, \pi_1)$ rekrutiert und dann über einen weiteren Zeitraum (π_1, π_2) weiter beobachtet (siehe Abbildung). Natürlich ist auch $\pi_1 = \pi_2$ möglich (was bedeutet, daß über die gesamte Studiendauer neue Probanden rekrutiert werden können. Ein typisches Beispiel für prospektive Studien sind Testreihen für ein neues Medikament, in denen Patienten über einen bestimmten Zeitraum ein neu entwickeltes Medikament, bzw. ein Plazebo einnehmen. In diesem Fall wäre z.B. der Zeitraum von Beginn der Medikation bis zur Genesung als zufällige Überlebenszeit anzusetzen.

Ein besonderes Problem in der Verweildaueranalyse stellen sogenannte **zensierte** Beobachtungen dar. Eine zensierte Beobachtung ist ein Objekt, das das terminierende Ereignis nicht erlebt, da es den gesamten Beobachtungszeitraum $(0, \pi_2)$ überlebt, oder sich aus anderen Gründen der Beobachtung entzieht. In Abbildung 3.1 sind die rechtszensierten Beobachtungen durch den nicht ausgefüllten Punkt angedeutet.

Man kann verschiedene Arten der Zensierung und Trunkierung unterscheiden:

- Rechtszensierung: Eine Beobachtung heißt rechtszensiert, wenn aufgrund Beendigung des Beobachtungszeitraums keine Aussage darüber gemacht werden kann, ob das interessierende Ereignis bereits stattgefunden hat oder nicht. Dies ist wohl die bekannteste Art der Zensierung.

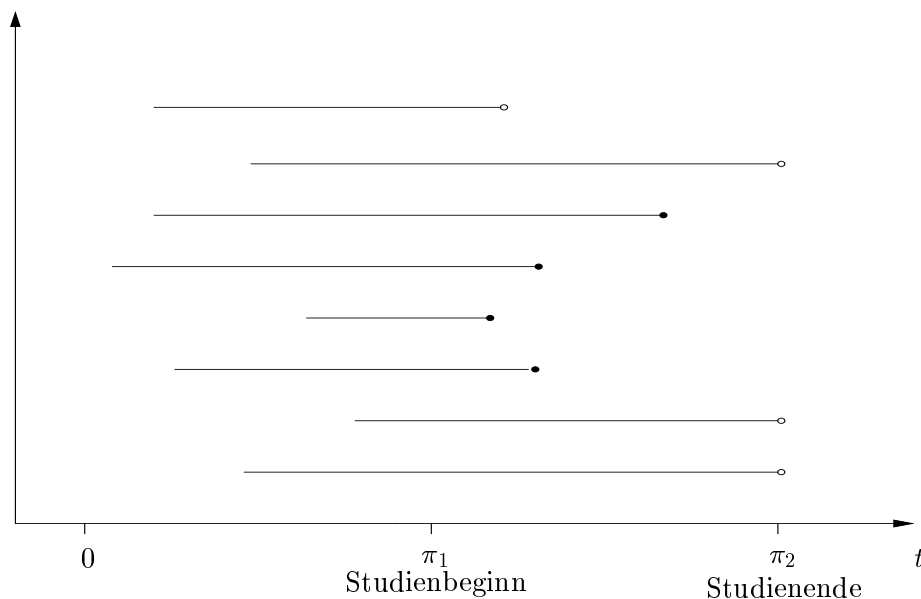


Abbildung 3.1: Schematische Darstellung einer Überlebensstudie

- **Linkszensierung:** Eine linkszensierte Beobachtung ist eine Beobachtung, deren Beginnzeitpunkt unbekannt ist. Als typisches Beispiel hierfür könnte man zum Beispiel eine medizinische Studie anführen, die den Zeitraum von HIV-Infektion bis zum Ausbruch von AIDS untersuchen möchte. In den wenigsten Fällen werden hier die Probanden den genauen Infektionszeitpunkt wissen.
- **Linkstrunkierung:** Eine linkstrunkierte Beobachtung ist eine Beobachtung, deren Beginnzeitpunkt vor Studienbeginn π_1 liegt, aber, im Gegensatz zu linkszensierten Beobachtungen, bekannt ist. Ein Beispiel für Linkstrunkierung ist die Erstellung einer Sterbetafel auf Basis einer Volkszählung. Hier ist der Beginnzeitpunkt (Geburt) in allen Fällen bekannt, aber der Studienbeginn ist der Beginn der Volkszählung.

3.5 Likelihood-Techniken für Überlebenszeitmodelle

In diesem Abschnitt sollen einige immer wieder in der Arbeit verwendete Techniken zur Herleitung von Likelihood-Schätzern in Überlebenszeitmodellen dargestellt werden (siehe Kalbfleisch und Prentice (1980) [21] Seiten 119 bis 142). Dabei stehen Maximum-Likelihood-Ansätze zur Bestimmung der Überlebensfunktion, sowie der sogenannte Partial-Likelihood im Zentrum der Betrachtung.

3.5.1 Likelihood-Konstruktion für Modelle mit unabhängiger Zensierung

Wir gehen aus von einer Studie mit insgesamt n Individuen, die vom Zeitpunkt 0 an beobachtet werden und deren Überleben durch die Hazardfunktion $\lambda(t, Z, \beta)$ bzw. durch die Dichtefunktion $f(t, Z, \beta)$ beschrieben wird, wobei hier $Z \in \mathbb{R}^p$ ein das Überleben genauer spezifizierender

Kovariablenvektor ist und $\beta \in \mathbb{R}^p$ ein Parametervektor. Die beobachteten Daten für das i -te Individuum sind t_i, δ_i, Z_i , hier ist t_i der Zeitpunkt des Beobachtungsendes, δ_i der sogenannte Zensierungsindikator, d.h.

$$\delta_i = \begin{cases} 1 & \text{Ereignis zum Zeitpunkt } t_i \\ 0 & \text{Zensierung zum Zeitpunkt } t_i \end{cases}$$

und Z_i der beobachtete Kovariablenvektor. Im Folgenden werden wir immer davon ausgehen, daß ein unabhängiger Zensierungsmechanismus vorliegt (nach Definition von Kalbfleisch und Prentice [21] Seite 120), der insbesondere dann gegeben ist, wenn die Zensierung zu einem festen Zeitpunkt (z.B. Studienende) stattfindet. Wir erhalten die Likelihoodfunktion allgemein als

$$L = \prod_{i=1}^n [f(t_i, Z_i, \beta)^{\delta_i} S(t_i, Z_i, \beta)^{1-\delta_i}]. \quad (3.10)$$

(3.10) kann man folgendermaßen erklären: Falls $\delta_i = 1$ wissen wir, daß ein Ereignis passiert ($T = t_i$), die Dichtefunktion dafür ist $f(t_i, Z_i, \beta)$, falls $\delta_i = 0$ erhalten wir die Information, daß das beobachtete Individuum mindestens bis zum Zeitpunkt t_i gelebt hat (d.h. $T \geq t_i$), die Wahrscheinlichkeit dafür ist $S(t_i, Z_i, \beta)$. Um nun Likelihood-Techniken für die Schätzung der Hazardfunktion $\lambda(t, Z, \beta)$ (bei gegebenem β), bzw. für den Parameter β (bei gegebenem λ) zu erhalten, stellen wir die Überlebensfunktion in Abhängigkeit von der Hazardfunktion dar. Zunächst haben wir in Satz (3.3) gezeigt, daß für die Überlebensfunktion der Zusammenhang

$$S(t) = \exp\left(-\int_0^t \lambda(s, Z, \beta) ds\right)$$

gilt. Zum anderen kann man eine absolut stetige Überlebensfunktion $S(t)$ in Abhängigkeit von $\lambda(t)$ durch folgenden Grenzwert darstellen:

$$S(t) = \lim_{r \rightarrow \infty} \prod_{k=0}^{r-1} (1 - \lambda(u_r) \Delta u_r), \quad (3.11)$$

wobei hier der Grenzwert über alle Partitionen $0 = u_0 < u_1 < \dots < u_r = t$ zu bilden ist und mit $\Delta u_r = u_{r+1} - u_r$ und

$$\lambda(u_r) \Delta u_r = P(u_r \leq T < u_r + \Delta u_r | T \geq u_r)$$

die Wahrscheinlichkeit, daß im Intervall von u_r bis u_{r+1} das terminierende Ereignis stattfindet, gegeben Überleben bis u_r . Der Beweis der Gleichung (3.11) findet sich in Andersen u.a. (1993) [1] Seiten 91 - 92. (3.11) nennt man auch Produkt-Integral-Darstellung der Überlebensfunktion.

Dementsprechend erhält man 2 verschiedene Darstellungen für die Likelihood-Funktion

$$\begin{aligned} L &= \prod_{i=1}^n [f(t_i, Z_i, \beta)^{\delta_i} S(t_i, Z_i, \beta)^{1-\delta_i}] \\ &\stackrel{f(t)=\lambda(t)S(t)}{=} \prod_{i=1}^n [\lambda(t_i, Z_i, \beta)^{\delta_i} S(t_i, Z_i, \beta)^{\delta_i} S(t_i, Z_i, \beta)^{1-\delta_i}] \end{aligned}$$

$$\begin{aligned}
&= \prod_{i=1}^n \left[\lambda(t_i, Z_i, \beta)^{\delta_i} \exp \left(- \int_0^{t_i} \lambda(u; Z_i, \beta) du \right) \right] \\
&= \left[\prod_{i=1}^n \lambda(t_i, Z_i, \beta)^{\delta_i} \right] \exp \left(- \int_0^{\infty} \sum_{k \in R(u)} \lambda(u, Z_k, \beta) du \right).
\end{aligned}$$

Hier ist $R(t)$ die sogenannte Risikomenge zum Zeitpunkt t

$$R(t) = \{\text{Beobachtungen mit } t_i \geq t\}.$$

In der Risikomenge $R(t)$ sind also auch alle Objekte enthalten, die zum Zeitpunkt t ein Ereignis erleben. Diese Feststellung ist sehr wichtig, da die Risikomenge so als linksstetiger stochastischer Prozeß mit rechtsseitigem Grenzwert betrachtet werden kann (R ist also vorhersehbar). Analog gilt für die Produkt-Integral-Darstellung

$$L = \prod_{i=1}^n [\lambda(t_i, Z_i, \beta)^{\delta_i} \lim_{r \rightarrow \infty} \sum_{k=1}^{r-1} (1 - \lambda(u_r) \Delta u_r)].$$

Im Abschnitt 3.6 werden wir je einen nichtparametrischen Schätzer auf Basis einer Darstellung der Überlebensfunktion wie in (3.11) (Kaplan-Meier-Schätzer) sowie nach Satz 3.3 (Nelson-Aalen-Schätzer) herleiten.

3.5.2 Partial Likelihood

An dieser Stelle wollen wir das Prinzip des Partial Likelihood erläutern, das später bei der Schätzung von Koeffizienten im Rahmen des Proportional-Hazard-Modells eine wesentliche Rolle spielen wird.

Nehmen wir an, die Verteilungsdichte einer Zufallsvariable X sei gegeben durch $f(x, \beta, \theta)$, wobei hier β der interessierende Parameter und θ ein unbekannter Störparameter, dessen Dimension meist sehr hoch, oder sogar, wie später im Proportional-Hazard-Modell, unendlich ist. Auf Basis einer Beobachtung Y wollen wir nun Aussagen über die Verteilung des Parameters treffen. Nehmen wir nun an, die beobachteten Daten Y werden nun transformiert in zwei Variablenmengen $A_1, B_1, A_2, B_2, \dots, A_m, B_m$ und sei $A^{(j)} = (A_1, \dots, A_j)$, sowie $B^{(j)} = (B_1, \dots, B_j)$.

Angenommen, die gemeinsame Dichte von $A^{(m)}$ und $B^{(m)}$ läßt sich darstellen als

$$\prod_{j=1}^m f(b_j | b^{(j-1)}, a^{(j-1)}, \theta, \beta) \prod_{j=1}^m f(a_j | b^{(j)}, a^{(j-1)}, \beta).$$

Dann nennt man den (von θ unabhängigen) zweiten Term **Partial Likelihood** basierend auf A . Dieser Term ist keine Likelihoodfunktion im herkömmlichen Sinn, d.h. er kann nicht direkt als Wahrscheinlichkeit interpretiert werden. Die Funktionsweise des Partial Likelihood soll mit Hilfe des folgenden Beispiels erläutert werden.

Beispiel 3.4 In diesem Beispiel betrachten wir 2 Stichproben, $SP1$ und $SP2$, mit $|SP1| = n$ und $|SP2| = m$ Elementen. Wir beobachten (nach Stichproben unterteilt) wie viele Ereignisse in je einem der 3 disjunkten Intervallen $I_1 = [c_0 = 0, c_1)$, $I_2 = [c_1, c_2)$, sowie $I_3 = [c_2, c_3 = \infty)$ passieren. Diese Beobachtungen können wir mit dem Datensatz $A_1, B_1, A_2, B_2, A_3, B_3$ beschreiben (siehe Abbildung 3.1), wobei

$$\begin{aligned} A_k &= \# \text{ Ereignisse in Intervall } I_k \text{ Stichprobe 1} \\ B_k &= \# \text{ Ereignisse in Intervall } I_k \text{ gesamt} \end{aligned}$$

Gehen wir davon aus, daß für in $SP1$ für die Wahrscheinlichkeit, daß in Intervall I_k $k = \{1, 2, 3\}$ das Ereignis passiert, gegeben Überleben bis c_k ,

$$p_1 = P(c_k \leq T < c_{k+1} | T \geq c_k, SP1), \quad k \in \{1, 2, 3\},$$

gilt

$$\frac{p_1}{1 - p_1} = \theta_k e^\beta$$

und in $SP2$ für

$$p_2 = P(c_k \leq T < c_{k+1} | T \geq c_k, SP2) \quad k = \{1, 2, 3\}$$

der Zusammenhang

$$\frac{p_2}{1 - p_2} = \theta_k$$

gilt.

Stichprobe	I_1			I_2			I_3	
	Ges.	Ereig.	Überl. bis c_1	Ges.	Ereig.	Überl. bis c_2	Ges.	Ereig.
1	n	a_1	$n - a_1$	$n - a_2$	a_2	$n - a_1 - a_2$	$n - a_1 - a_2$	a_3
2	m	r_1	$m - r_1$	$m - r_1$	r_2	$m - r_1 - r_2$	$m - r_1 - r_2$	r_3
		b_1			b_2			b_3

Tabelle 3.1: Beobachtete Überlebenszeiten in Beispiel 3.4

Hier ist β der interessierende Parameter und θ_k $k = 1, 2, 3$ der Störparameter. Den ersten Faktor des Partial-Likelihood erhält man dann als bedingte Wahrscheinlichkeit, daß im Intervall I_1 a_1 Ereignisse in $SP 1$ beobachtet werden können, unter der Bedingung, daß insgesamt b_1 Ereignisse passieren

$$\begin{aligned} P(A_1 = a_1 | B_1 = b_1) &= \frac{P(A_1 = a_1, B_1 = b_1)}{P(B_1 = b_1)} \\ &= \frac{\binom{n}{a_1} p_{11}^{a_1} (1 - p_{11})^{n - a_1} \binom{m}{b_1 - a_1} p_{12}^{b_1 - a_1} (1 - p_{12})^{m - (b_1 - a_1)}}{\sum_l \binom{n}{l} p_{11}^l (1 - p_{11})^{n - l} \binom{m}{b_1 - l} p_{12}^{b_1 - l} (1 - p_{12})^{m - (b_1 - l)}} \\ &= \frac{\binom{n}{a_1} \left(\frac{p_{11}}{1 - p_{11}}\right)^{a_1} (1 - p_{11})^n \binom{m}{b_1 - a_1} \left(\frac{p_{12}}{1 - p_{12}}\right)^{b_1 - a_1} (1 - p_{12})^m}{\sum_l \binom{n}{l} \left(\frac{p_{11}}{1 - p_{11}}\right)^l (1 - p_{11})^n \binom{m}{b_1 - l} \left(\frac{p_{12}}{1 - p_{12}}\right)^{b_1 - l} (1 - p_{12})^m} \end{aligned}$$

$$\begin{aligned}
 &= \frac{\binom{n}{a_1} \theta_1^{a_1} e^{\beta a_1} \binom{m}{b_1 - a_1} \theta_1^{b_1 - a_1}}{\sum_l \binom{n}{l} \theta_k^l e^{l\beta} \binom{m}{b_1 - l} \theta_k^{b_1 - l}} \\
 &= \frac{\binom{n}{a_1} \binom{m}{b_1 - a_1} e^{a_1 \beta}}{\sum \binom{n}{l} \binom{m}{b_1 - l} e^{l\beta}}.
 \end{aligned}$$

Man sieht, daß der Störparameter θ_1 in diesem Term gekürzt werden kann. Für den zweiten Faktor erhalten wir

$$P(A_2 = a_2 | B_1 = b_1, A_1 = a_1, B_2 = b_2) = \frac{\binom{n - a_1}{a_2} \binom{m - r_1}{b_2 - a_2} e^{a_2 \beta}}{\sum \binom{n - a_1}{l} \binom{m - r_1}{b_2 - l} e^{l\beta}}$$

und schließlich

$$P(A_3 = n - a_1 - a_2 | B_1 = b_1, A_1 = a_1, B_2 = b_2, A_2 = a_2, B_3 = b_3) = 1.$$

Mit einer ähnlichen Konstruktion wie in Beispiel 3.4 werden wir später im semiparametrischen Proportional-Hazard-Modell Regressionskoeffizienten für die Hazardfunktion schätzen.

3.6 Schätzen der Überlebensfunktion

In diesem Abschnitt betrachten wir 2 unterschiedliche, nichtparametrische Ansätze für die Schätzung der Überlebensfunktion. Die Schätzmethoden leiten wir mit den im vorherigen Abschnitt dargestellten Likelihood-Techniken her. Da wir im Folgenden oft die Menge der geordneten Todeszeitpunkte, sowie die Menge der geordneten Endzeitpunkte der Beobachtungen (mit möglichen Zensierungen) benötigen, treffen wir nun folgende Konvention: Mit $t_1 < t_2 < \dots < t_k$ bezeichnen wir die Menge der geordneten Zeitpunkte, an denen ein Ereignis passiert, mit $t_1 \leq \dots \leq t_n$ die Menge der Endzeitpunkte der Beobachtungen in einer geordneten Stichprobe der Länge n , d.h. im ersten Fall indizieren wir die Zeitpunkte bis zum Index k , im zweiten bis zum Index n .

3.6.1 Der Produkt-Limit-Schätzer (Kaplan und Meier)

Der Produkt-Limit-Schätzer (Kaplan und Meier, 1958)[22] ist wohl der bekannteste und am häufigsten eingesetzte Schätzer für die Überlebensfunktion. Mit diesem Schätzer werden an den Zeitpunkten $t_i, i = 1, \dots, k$, an denen ein Ereignis stattfindet, die zugehörigen Hazards $\lambda_i = \lambda(t_i), i = 1, \dots, k$, geschätzt. Aus der diskreten Approximation

$$S(t) = \lim_{r \rightarrow \infty} \prod_{k=0}^{r-1} (1 - \lambda(u_r) \Delta u_r) \approx \prod_{t_i \leq t} (1 - \lambda_i)$$

des Produkt-Integrals (3.11) resultiert dann die Produkt-Limit-Schätzung für die Überlebensfunktion. Nun müssen wir noch $\lambda_i, i = 1, \dots, k$, schätzen.

Seien nun $t_1 < t_2 < \dots < t_k$ die Zeitpunkte, an denen ein terminierendes Ereignis stattfindet und

$$R_i := \{\text{beobachtete Objekte mit } T \geq t_i\} \quad i = 1, \dots, k,$$

sowie

$$D_i := \{\text{beobachtete Objekte mit } T = t_i\} \quad i = 1, \dots, k,$$

mit $d_i = |D_i|$ und $n_i = |R_i|$.

Als Likelihoodfunktion (in Abhängigkeit von dem diskreten Hazard λ_i zum Zeitpunkt t_i) für die Wahrscheinlichkeit, daß d_i von n_i Objekten das terminierende Ereignis am Zeitpunkt t_i erleben, erhält man

$$L_i(\lambda_i) = \frac{[f(t_i)]^{d_i} [S(t_i)]^{n_i - d_i}}{[S(t_i)]^{n_i}} = \lambda_i^{d_i} (1 - \lambda_i)^{n_i - d_i}.$$

Sei nun

$$L(\lambda) = \prod_{i=1}^k L_i(\lambda_i) = \prod_{i=1}^k \lambda_i^{d_i} (1 - \lambda_i)^{n_i - d_i}$$

und

$$LL(\lambda) = \log(L(\lambda)) = \sum_{i=1}^k (d_i \log(\lambda_i) + (n_i - d_i) \log(1 - \lambda_i))$$

der zugehörige Log-Likelihood mit Ableitung nach λ_i

$$\frac{\partial}{\partial \lambda_i} LL(\lambda) = \frac{d_i}{\lambda_i} + \frac{d_i - n_i}{1 - \lambda_i}.$$

Dieser Term wird 0 für $\hat{\lambda}_i = d_i/n_i$, so daß die Produkt-Limit-Schätzung für die Überlebensfunktion gegeben ist durch

$$\hat{S}_{KM}(t_k) = \prod_{i=1}^k (1 - \hat{\lambda}_i) \quad (3.12)$$

Im folgenden betrachten wir noch kurz einige Eigenschaften des Kaplan-Meier-Schätzer. Die Varianz von \hat{S}_{KM} kann man mit der Formel von Greenwood (1926) [18]

$$\widehat{\text{Var}}[\hat{S}(t)_{KM}] = \hat{S}_{KM}(t) \sum_{t_i \leq t} \frac{d_i}{R_i(R_i - d_i)}$$

bestimmen. Auf Basis asymptotischer Normalität, die wir Abschnitt 3.6.2 für den Nelson-Aalen-Schätzer noch genauer untersuchen werden, kann man nun punktweise $100(1 - \alpha)\%$ -Konfidenzintervalle für $\hat{S}(t)$ angeben. Dazu benötigen wir noch folgende Notation

$$\sigma_s^2(t) = \frac{\widehat{\text{Var}}[\hat{S}(t)]}{\hat{S}^2(t)}.$$

Damit erhalten wir als $100(1 - \alpha)\%$ Konfidenzintervall zum Zeitpunkt t_0

$$[\text{untere Schranke}, \text{obere Schranke}] = [\hat{S}(t_0) - c_{1-\alpha/2} \sigma_S(t_0) \hat{S}(t_0) + c_{1-\alpha/2} \sigma_S(t_0) \hat{S}(t_0)], \quad (3.13)$$

wobei $c_{1-\alpha/2}$ das $1-\alpha/2$ -Quantil der Standardnormalverteilung ist.

Bemerkung: Punktweises Konfidenzintervall bedeutet im diesem Zusammenhang, daß das in (3.13) definierte zufällige Intervall zu jedem Zeitpunkt t_0 mit $(1 - \alpha)100\%$ -iger Wahrscheinlichkeit den (wahren) Wert der Überlebensfunktion S an der Stelle t_0 beinhaltet.

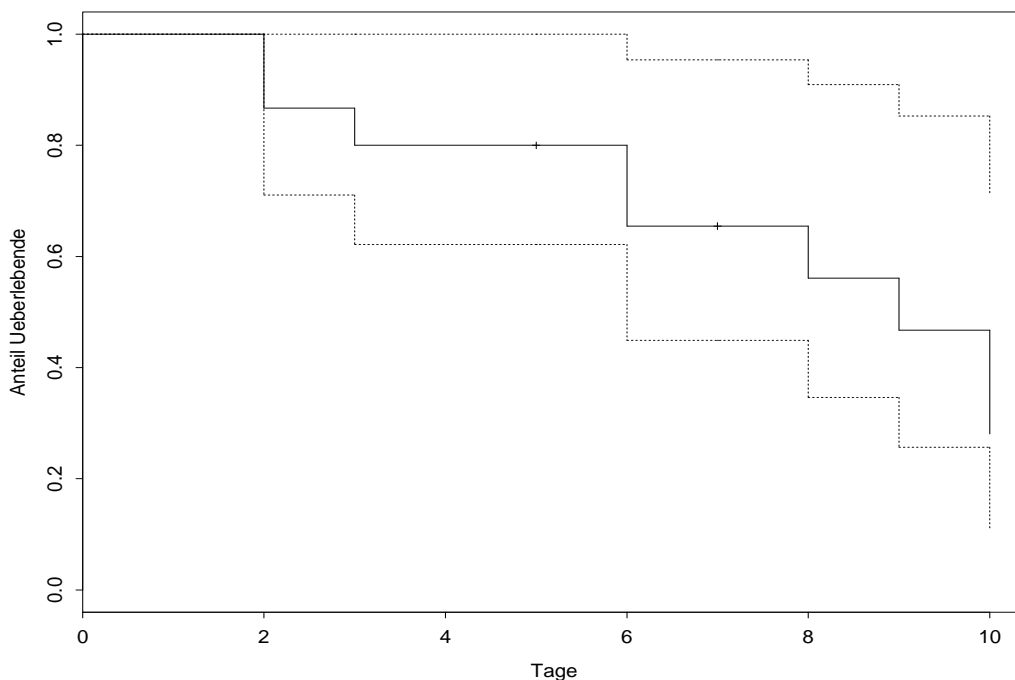


Abbildung 3.2: Kaplan-Meier-Schätzer mit zugehörigen 95%- Konfidenzintervallen

Beispiel 3.5 Mit folgendem fiktiven Zahlenbeispiel soll die Funktionsweise des Kaplan- Meier- Schätzers erläutert werden (+ steht für Zensierung).

Beobachtete Lebensdauern: 2, 2, 3, 5+, 6, 6, 7+, 7+, 8, 9, 10, 10, 10+, 10+, 10+ (Anzahl der Beobachtungen $n = 15$).

In Tabelle 3.2 sind die geschätzten Koeffizienten, sowie die geschätzten 95 % Konfidenzintervalle aufgeführt.

t_i	n_i	d_i	$\hat{S}_{KM}(t_i)$	unt.Schr.(95 %)	ob.Schr.(95 %)
2	15	2	$1 - \frac{2}{15} = 0.867$	0.711	1.000
3	13	1	$(1 - \frac{2}{15})(1 - \frac{1}{13}) = 0.800$	0.621	1.000
6	11	2	$(1 - \frac{2}{15}) \cdots (1 - \frac{2}{11}) = 0.655$	0.449	0.954
8	7	1	$(1 - \frac{2}{15}) \cdots (1 - \frac{1}{7}) = 0.561$	0.346	0.909
9	6	1	$(1 - \frac{2}{15}) \cdots (1 - \frac{1}{6}) = 0.468$	0.256	0.853
10	5	2	$(1 - \frac{2}{15}) \cdots (1 - \frac{2}{5}) = 0.281$	0.110	0.714

Tabelle 3.2: Geschätzte Werte der Überlebensfunktion (Kaplan-Meier-Schätzer)

In Abbildung 3.2 erkennt man, daß der Kaplan-Meier-Schätzer eine Stufenfunktion mit Sprüngen an den beobachteten Todeszeitpunkten ist. Der Kaplan-Meier-Schätzer ist nach unten verzerrt.

3.6.2 Der Nelson-Aalen-Schätzer

Der Nelson-Aalen-Schätzer ist ein alternativer Ansatz für die Schätzung der Überlebensfunktion $S(t)$. Es gibt verschiedene Möglichkeiten, diesen Schätzer herzuleiten, im folgenden soll hier, analog zum Produkt-Limit-Schätzer, die nichtparametrische Maximum-Likelihood-Herleitung vorgestellt werden. Nelson (1972) [31] schlug diesen Schätzer im Rahmen eines Artikels über graphische Methoden zur Darstellung der Hazardfunktion vor. Seien $t_1 \leq \dots \leq t_n$ die geordneten Endzeitpunkte der Beobachtungen und δ_i die zugehörigen Zensierungsindikatoren (d.h. $\delta_i = 0$ falls Beobachtung zensiert, $\delta_i = 1$ sonst).

Die Wahrscheinlichkeit dafür, daß ein Individuum zum Zeitpunkt t_i das terminierende Ereignis erlebt, ist

$$f(t_i) = \lambda(t_i) S(t_i) = \lambda(t_i) \exp[-\Lambda(t_i)].$$

Damit erhält man als Likelihoodfunktion

$$L(\lambda(t)) = \prod_{i=1}^n \lambda(t_i)^{\delta_i} \exp(-\Lambda(t_i)).$$

Um diesen Likelihood zu maximieren, betrachten wir die diskreten Hazardraten λ_i $i = 1, \dots, k$ zu den Zeitpunkten $t_1 < \dots < t_k$, an denen ein Ereignis passiert. Sei mit R_i wiederum die Risikomenge zum Zeitpunkt t_i bezeichnet und sei $|R_i| = n_i$. Dann reduziert sich der Likelihood zu

$$L(\lambda(t)) = \prod_{i=1}^k \lambda_i \exp(-\Lambda(t_i)) = \prod_{i=1}^k [\lambda_i \exp(-\sum_{t_j \leq t_i} \lambda_j)] = \prod_{i=1}^k \lambda_i \exp(-\sum_{j \in R_i} \lambda_j),$$

mit zugehörigem Log-Likelihood

$$LL(\lambda) = \sum_{i=1}^k \log(\lambda_i) - \sum_{j \in R_i} \lambda_j$$

und Ableitung

$$\frac{\partial}{\partial \lambda_i} LL(\lambda_i) = \frac{1}{\lambda_i} - n_i.$$

Die Ableitung hat den Wert 0 an der Stelle

$$\hat{\lambda}_i = \frac{1}{n_i},$$

so daß der Nelson-Aalen-Schätzer für die Überlebensfunktion definiert ist als

$$\hat{S}_{NA}(t) = \exp(-\sum_{t_i \leq t} \hat{\lambda}_i). \quad (3.14)$$

Bemerkung: Falls an einem Zeitpunkt mehrere Ereignisse stattfinden, wird der Schätzer für die Hazardfunktion modifiziert zu: $\hat{\lambda}_i = d_i/n_i$, wobei d_i die Anzahl der Ereignisse zum Zeitpunkt t_i

zählt.

Asymptotische Eigenschaften des Nelson-Aalen-Schätzers

Um asymptotische Eigenschaften des Nelson-Aalen Schätzers zu beweisen, bedient man sich einer anderen Herleitung. Betrachtet man nämlich den Zählprozeß

$$N(t) = \sum_{i=1}^n I(T_i \leq t)$$

(Summe der Toten zum Zeitpunkt t), wobei jedes Objekt mit der Intensität $\lambda(s)$ ausscheidet. Sei nun $R(s)$ die Anzahl der beobachteten Objekte zum Zeitpunkt s , so ist

$$M(t) = N(t) - \int_0^t \lambda(s)R(s) ds \quad (3.15)$$

ein Martingal bzgl. der kanonischen, d.h. der von $M(t)$ erzeugten Filtration \mathcal{F}_t . Dies sieht man sofort, da man $N(t)$ als Summe n univariater Zählprozesse darstellen kann

$$N(t) = \sum_{i=1}^n Y_i(t)N_i(t),$$

mit

$$Y_i(t) = \begin{cases} 1 & \text{falls Objekt zum Zeitpunkt } t \text{ unter Beobachtung} \\ 0 & \text{sonst} \end{cases}$$

und

$$N_i(t) = \# \text{ Ereignisse des Objektes } i.$$

N_i nimmt also die Werte 0 oder 1 an. Nach der Definition des vorhersehbaren Intensitätsprozesses für Zählprozesse ist

$$E \left[\int_0^\infty H(u)\lambda(u) du \right] = E \left[\int_0^\infty H(u)dN_i(u) \right]$$

für alle vorhersehbaren Prozesse H und alle $i \in 1, \dots, n$ erfüllt. Insbesondere also auch für den vorhersehbaren Prozeß

$$I(s \leq u \leq t) \sum_{i=1}^n Y_i(u) = R(u)I(s \leq u \leq t) \quad (3.16)$$

und wir erhalten für $s < t$

$$\begin{aligned} & E \left[N(t) - \int_0^t \lambda(u)R(u) du \middle| \mathcal{F}_s \right] \\ &= E \left[N(t) - N(s) + N(s) - \left(\int_0^s \lambda(u)R(u) du + \int_s^t \lambda(u)R(u) du \right) \middle| \mathcal{F}_s \right] \\ \stackrel{\text{Satz 2.2}}{=} & E \left[N(t) - N(s) - \int_s^t \lambda(u)R(u) du \middle| \mathcal{F}_s \right] + N(s) - \int_0^s R(u)\lambda(u) du \\ &= E \left[\sum_{i=1}^n \left(\int_s^t Y_i(u) dN_i(u) - \int_s^t Y_i(u)\lambda(u) du \right) \middle| \mathcal{F}_s \right] + N(s) - \int_0^s R(u)\lambda(u) du \\ \stackrel{(3.16)}{=} & N(s) - \int_0^s R(u)\lambda(u) du, \end{aligned}$$

womit die Martingaleigenschaft gezeigt wäre.

Mit

$$dM(t) = dN(t) - \lambda(s)R(s),$$

wobei $dM(t)$ als zufälliges Rauschen zu betrachten ist, erhält man als Schätzer für $\lambda(t)$

$$\hat{\lambda}(t) = \frac{dN(t)}{R(t)},$$

so daß man für den kumulativen Hazard $\Lambda(t)$ den Schätzer

$$\hat{\Lambda}(t) = \int_0^t \frac{1}{R(s)} dN(s) \quad (3.17)$$

angeben kann. Seien nun wiederum $t_1 < \dots < t_k$ die geordneten Todeszeitpunkte, dann gilt

$$dN(t) = \begin{cases} 1 & \text{für } t = t_i \quad i = 1, \dots, k \\ 0 & \text{sonst} \end{cases},$$

so daß sich (3.17) reduziert zu

$$\hat{\Lambda}(t) = \sum_{i:t_i \leq t} \frac{1}{R(t_i)}. \quad (3.18)$$

Schließlich sei noch die Zufallsvariable $J(t) = I(R(t) > 0)$ eingeführt, sowie

$$\Lambda^*(t) = \int_0^t \lambda(s)J(s) ds.$$

Da der Zählprozeß $N(t)$ nur dann springt, wenn $R(t) > 0$ kann man für (3.17) folgende äquivalente Notation benutzen:

$$\hat{\Lambda}(t) = \int_0^t \frac{J(s)}{R(s)} dN(s).$$

Aus der Tatsache, daß $R(t)$ ein vorhersehbarer Prozeß ist, folgt mit Satz 2.29, daß

$$\begin{aligned} \hat{\Lambda}(t) - \Lambda^*(t) &= \int_0^t \frac{J(s)}{R(s)} dN(s) - \int_0^t \lambda(s)J(s) ds \\ &= \int_0^t \frac{J(s)}{R(s)} d(N(s) - \int_0^s \lambda(u)R(u) du) = \int_0^t \frac{J(s)}{R(s)} dM(s) \end{aligned}$$

ein lokales Martingal ist. Nach Satz 2.35 hat $\hat{\Lambda} - \Lambda^*$ den vorhersehbaren Variationsprozeß

$$\langle \hat{\Lambda}(t) - \Lambda^*(t) \rangle = \int_0^t \left(\frac{J(s)}{R(s)} \right)^2 \lambda(s)R(s) ds \stackrel{J^2(s) \equiv J(s)}{=} \int_0^t \frac{J(s)}{R(s)} \lambda(s) ds.$$

Da Λ^* der vorhersehbare Kompensatorprozeß von $\hat{\Lambda}$ ist, gilt für den Erwartungswert

$$E(\hat{\Lambda}(t)) = E(\Lambda^*(t)) = \int_0^t \lambda(s)P(R(s) > 0) ds.$$

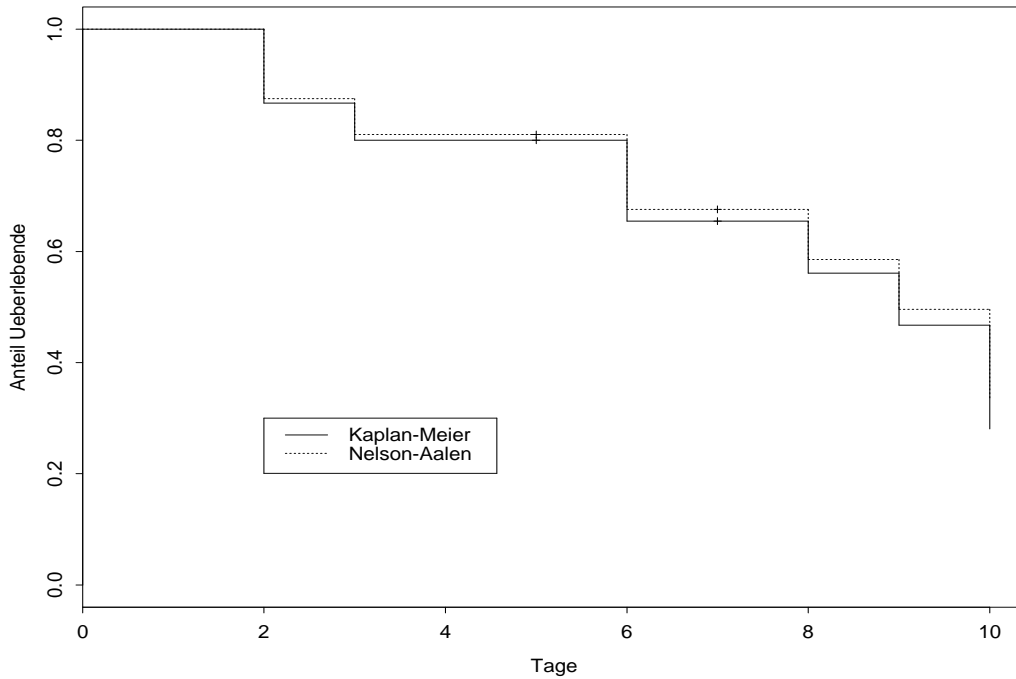


Abbildung 3.3: Vergleich von Kaplan-Meier und Nelson-Aalen-Schätzer

Daraus folgt, daß die Verzerrung $b(\hat{\Lambda}(t))$ des Schätzer gegeben ist durch

$$\begin{aligned}
 b(\hat{\Lambda}(t)) &= E(\hat{\Lambda}(t)) - \Lambda(t) = \int_0^t \lambda(s)P(R(s) > 0) ds - \int_0^t \lambda(s) ds = \\
 &= \int_0^t \lambda(s) \underbrace{[P(R(s) > 0) - 1]}_{=-P(R(s)=0)} ds = - \int_0^t \lambda(s)P(R(s) = 0) ds \leq 0
 \end{aligned}$$

Da die Überlebensfunktion eine monoton fallende Funktion in Abhängigkeit von $\Lambda(t)$ ist, folgt daß der Nelson-Aalen-Schätzer $\hat{S}_{NA}(t)$ für die Überlebensfunktion im Gegensatz zum Kaplan-Meier-Schätzer $\hat{S}_{KM}(t)$ (3.12) nach oben verzerrt ist und es gilt

$$\hat{S}_{KM} \leq \hat{S}_{NA} \text{ für alle } t$$

denn

$$\begin{aligned}
 \hat{S}_{NA}(t) &= \exp\left(-\sum_{t_i \leq t} \hat{\lambda}_i\right) = \prod_{t_i \leq t} e^{-\hat{\lambda}_i} \\
 &= \prod_{t_i \leq t} \left(1 - \hat{\lambda}_i + \underbrace{\hat{\lambda}_i^2 - \hat{\lambda}_i^3 + \dots}_{\geq 0}\right) \\
 &\geq \prod_{t_i \leq t} (1 - \hat{\lambda}_i) = \hat{S}_{KM}(t).
 \end{aligned}$$

In der Praxis erweist sich jedoch für große Datensätze die Verzerrung als vernachlässigbar gering, in Abbildung 3.3 wird erkenntlich, daß schon für das fiktive Beispiel zum Kaplan-Meier-Schätzer

die Abweichungen relativ gering sind.

Als Schätzer für die Varianz der kumulativen Hazardfunktion $\Lambda(t)$ dient der optionale Variationsprozeß von $\hat{\Lambda} - \Lambda^*$. Mit Hilfe von Satz 2.35 erhält man:

$$\hat{\sigma}^2(t) = [\hat{\Lambda}(t) - \Lambda^*(t)] = \left[\int_0^t \frac{J(s)}{R(s)} dM(s) \right] = \int_0^t \frac{J(s)}{(R(s))^2} dN(s), \quad (3.19)$$

mit

$$M(s) = N(s) - \int_0^s \lambda(u) R(u) du.$$

Mit Hilfe dieser Vorbereitungen kann man nun die Konvergenz des Nelson-Aalen Schätzers zeigen. Dazu betrachten wir eine Folge von Zählprozessen $N^{(n)}$. Desweiteren bezeichnen wir mit $R^{(n)}$ die zugehörige Folge der Risikomengen, mit $J^{(n)}(t)$ die Folge der Prozesse $I(R^{(n)}(t) > 0)$ und mit $\hat{\Lambda}^{(n)}$ und $\Lambda^{*(n)}$ die Prozesse

$$\begin{aligned} \hat{\Lambda}^{(n)}(t) &= \int_0^t \frac{1}{R^{(n)}(s)} dN^{(n)}(s), \\ \Lambda^{*(n)}(t) &= \int_0^t \lambda(s) J^{(n)}(s) ds. \end{aligned}$$

Satz 3.6 (Konsistenz von $\hat{\Lambda}$) Sei $t \in \mathcal{T}$ und gelte für $n \rightarrow \infty$

$$\int_0^t \frac{J^{(n)}(s)}{R^{(n)}(s)} \lambda(s) ds \rightarrow_{\mathcal{P}} 0 \text{ für } n \rightarrow \infty \quad (3.20)$$

sowie

$$\int_0^t (1 - J^{(n)}(s)) \lambda(s) ds \rightarrow_{\mathcal{P}} 0 \text{ für } n \rightarrow \infty \quad (3.21)$$

dann gilt für $n \rightarrow \infty$

$$\sup_{s \in [0, t]} |\hat{\Lambda}^{(n)}(s) - \Lambda(s)| \rightarrow_{\mathcal{P}} 0.$$

Beweis:

Mit Hilfe der Zerlegung von $\sup_{s \in [0, t]} |\hat{\Lambda}^{(n)}(s) - \Lambda(s)|$ in

$$\begin{aligned} \sup_{s \in [0, t]} |\hat{\Lambda}^{(n)}(s) - \Lambda(s)| &= \sup_{s \in [0, t]} |\hat{\Lambda}^{(n)}(s) - \Lambda^{*(n)}(s) + \Lambda^{*(n)}(s) - \Lambda(s)| \\ &\leq \sup_{s \in [0, t]} |\hat{\Lambda}^{(n)}(s) - \Lambda^{*(n)}(s)| + \sup_{s \in [0, t]} |\Lambda^{*(n)}(s) - \Lambda(s)|, \end{aligned}$$

zeigt man zunächst mit Hilfe der Ungleichung von Lenglart, daß gilt

$$\begin{aligned} P \left\{ \sup_{s \in [0, t]} |\hat{\Lambda}^{(n)}(s) - \Lambda^{*(n)}(s)| > \eta \right\} &\leq \frac{\delta}{\eta^2} + P \left\{ \underbrace{\int_0^t \frac{J^{(n)}(s)}{R^{(n)}(s)} \lambda(s) ds}_{\rightarrow 0 \text{ nach Vor. (3.20)}} > \delta \right\} \\ &\implies \sup_{s \in [0, t]} |\hat{\Lambda}^{(n)}(s) - \Lambda^{(n)*}(s)| \rightarrow_{\mathcal{P}} 0. \end{aligned}$$

Mit (3.21) erhält man direkt

$$\sup_{s \in [0, t]} |\Lambda^{*(n)}(s) - \Lambda(s)| = \int_0^t (1 - J^{(n)}(s)) \lambda(s) ds \rightarrow_{\mathcal{P}} 0,$$

so daß die Aussage gezeigt ist. \square

Bemerkung: Die Bedingungen (3.20) und (3.21) sind offensichtlich erfüllt, wenn $R^{(n)}(t) \rightarrow \infty$ für $n \rightarrow \infty$.

Mit folgendem Satz wird die asymptotische Normalität des Nelson-Aalen-Schätzers gezeigt

Satz 3.7 (Asymptotische Normalität) *Sei $t \in \mathcal{T}$ und angenommen, es gibt eine monoton wachsende Folge positiver Konstanten a_n , mit $a_n \rightarrow \infty$ für $n \rightarrow \infty$ und nicht-negative Funktionen $y(\cdot)$, so daß $\frac{\lambda(\cdot)}{y(\cdot)}$ in ganz $[0, t]$ integrierbar ist. Sei*

$$\sigma^2(s) = \int_0^s \frac{\lambda(u)}{y(u)} du$$

und seien folgende Bedingungen erfüllt

A. Für jedes $s \in [0, t]$ gilt

$$a_n^2 \int_0^s \frac{J^{(n)}(u)}{R^{(n)}(u)} \lambda(u) du \rightarrow_{\mathcal{P}} \sigma^2(s) \text{ für } n \rightarrow \infty.$$

B. Für alle $\varepsilon > 0$ gilt

$$a_n^2 \int_0^t \frac{J^{(n)}(u)}{R^{(n)}(u)} \lambda(u) du \rightarrow_{\mathcal{P}} 0 \text{ für } n \rightarrow \infty.$$

C.

$$a_n \int_0^t (1 - J^{(n)}(u)) \lambda(u) du \rightarrow_{\mathcal{P}} 0 \text{ für } n \rightarrow \infty.$$

Dann gilt

$$a_n(\hat{\Lambda}^{(n)} - \Lambda) \rightarrow_{\mathcal{D}} U \text{ für } n \rightarrow \infty,$$

wobei U ein Martingal ist, dessen Zuwächse unabhängig $N(0, \sigma^2(t-s))$ verteilt sind für $s < t$. Zudem gilt

$$\sup_{s \in [0, t]} |a_n^2 \hat{\sigma}^2(s) - \sigma^2(s)| \rightarrow_{\mathcal{P}} 0 \text{ für } n \rightarrow \infty.$$

Beweis: Die asymptotische Normalität wird bewiesen in Andersen u.a. (1993) [1] Seiten 191 und 192.

Konfidenzintervalle für den Nelson-Aalen-Schätzer

Aus der asymptotischen Normalität des Nelson-Aalen-Schätzers $\hat{\Lambda}(t)$ kann man nun Konfidenzintervalle für die kumulative Hazardfunktion angeben (Bie und Borgan (1987) [5]). Zunächst

soll ein punktweises Konfidenzintervall für Λ zu einem festen Zeitpunkt t_0 bestimmt werden.

Aus der asymptotischen Normalität des Nelson-Aalen-Schätzers erhält man als Konfidenzintervall für den kumulativen Hazard $\Lambda(t_0)$

$$\hat{\Lambda}(t_0) \pm c_{\alpha/2} \hat{\sigma}(t_0),$$

wobei hier $c_{\alpha/2}$ das $1-\alpha/2$ -Quantil der Standardnormalverteilung und $\hat{\sigma}(t_0)$ die nach (3.19) geschätzte Varianz des Nelson-Aalen-Schätzers. Es zeigt sich jedoch, daß für kleine Stichprobengrößen der so definierte Schätzer eine relativ ungenaue Approximation des Konfidenzintervalles liefert. Daher wurden mit Hilfe der funktionalen δ -Methode (siehe z.B. Gill(1989) [17]) Transformationen für des Konfidenzintervall entwickelt, die bessere asymptotische Eigenschaften für kleine Stichprobengrößen haben. Nach der funktionalen δ -Methode gilt für eine in einer Umgebung von $\Lambda(t_0)$ differenzierbare Funktion $g(x)$ mit $g(x) \neq 0$ für $x = \Lambda(t_0)$ daß

$$\frac{g(\hat{\Lambda}(t_0)) - g(\Lambda(t_0))}{|g'(\hat{\Lambda}(t_0))| \hat{\sigma}(t_0)} \rightarrow_{\mathcal{D}} N(0, 1).$$

Als $100(1 - \alpha)\%$ Konfidenzintervall für die transformierte Hazardfunktion $g(\Lambda(t_0))$ erhält man daher

$$g(\hat{\Lambda}(t_0)) \pm c_{\alpha/2} |g'(\hat{\Lambda}(t_0))| \sigma(t_0).$$

Für die logarithmische Transformation mit $g(x) = \log(x)$ erhält man als Konfidenzintervall für den transformierten kumulativen Hazard $\log(\Lambda(t_0))$

$$\log(\hat{\Lambda}(t_0)) \pm c_{\alpha/2} \frac{\hat{\sigma}(t_0)}{\hat{\Lambda}(t_0)},$$

so daß das Konfidenzintervall für die kumulative Hazardfunktion $\Lambda(t_0)$ gegeben ist durch

$$\hat{\Lambda}(t_0) \exp \left(\pm c_{\alpha/2} \frac{\hat{\sigma}(t_0)}{\hat{\Lambda}(t_0)} \right).$$

Eine weitere wichtige Transformation ist die arcsin-Transformation. Genaueres hierzu kann man in [5] nachlesen. Simulationen haben gezeigt, daß sowohl die logarithmische, als auch die arcsin-Transformation deutlich bessere Eigenschaften für kleine Stichprobengrößen haben (Borgan und Liestøl (1990) [6]).

3.7 Das Proportional-Hazard-Modell

3.7.1 Modellierung

Bis jetzt wurden nur nichtparametrische Schätzer der Überlebensfunktion betrachtet. Im folgenden soll ein semiparametrisches Regressionsmodell vorgestellt werden, das sogenannte Proportional-Hazard-Modell. Das Proportional-Hazard-Modell geht davon aus, daß die Hazardfunktion

eines Individuums abhängig ist von einer allen Individuen zugrundeliegenden Basis-Hazard-Funktion $\lambda_0(t)$ und einem eventuell zeitabhängigen Kovariablenvektor

$$Z(t) = (Z_1(t), \dots, Z_p(t))^t \in \mathbb{R}^p.$$

Das Proportional-Hazard-Modell ist dann gegeben durch:

$$\lambda(t|Z(t)) = \lambda_0(t) \exp\left(\sum_{i=1}^p \beta_i Z_i(t)\right),$$

wobei $\beta = (\beta_1, \dots, \beta_p) \in \mathbb{R}^p$ den Vektor der zu schätzenden Regressionskoeffizienten darstellt.

Die Schätzung der Koeffizienten erfolgt durch den sogenannten Partial-Likelihood. Diese Technik wurde von Cox (1975) [10] für das Proportional-Hazard-Modell vorgeschlagen und ermöglicht es den Regressionskoeffizienten zu schätzen, ohne daß man dazu Informationen über den zugrundeliegenden Basis-Hazard benötigt.

Die Herleitung des Partial-Likelihood erfolgt nun ähnlich wie in Beispiel 3.4. Dazu betrachten wir an jedem Zeitpunkt, an dem ein Ereignis passiert, die bedingte Wahrscheinlichkeit, daß die beobachtete Person stirbt unter der Bedingung, daß ein Ereignis an diesem Zeitpunkt passiert.

Seien $t_1 < \dots < t_k$ die geordneten Todeszeitpunkte und R_i die Risikomenge zum Zeitpunkt t_i , sowie $Z_i(t_i) = (Z_{i1}(t_i), \dots, Z_{ip}(t_i))$ der zugehörige Kovariablenvektor. Aufgrund der zeitstetigen Modellierung beobachten wir fast sicher genau ein Ereignis pro Zeitpunkt t_i $i = 1, \dots, k$.

Für jeden beobachteten Todeszeitpunkt ist die Wahrscheinlichkeit, daß die ausgeschiedene Person zum Zeitpunkt t_i stirbt, gegeben ein Toter zum Zeitpunkt t_i :

$$\begin{aligned} P(\text{Person stirbt zum Zeitpunkt } t_i | \text{ein Toter zum Zeitpunkt } t_i) &= \\ \frac{P(\text{Person stirbt zum Zeitpunkt } t_i | \text{überlebt bis } t_i^-)}{P(\text{Ein Toter zum Zeitpunkt } t_i | \text{überlebt bis } t_i^-)} &= \\ \frac{\lambda_0(t_i) \exp[\beta^t Z_i(t_i)]}{\sum_{j \in R_i} \lambda_0(t_i) \exp[\beta^t Z_j(t_i)]} &= \frac{\exp[\beta^t Z_i(t_i)]}{\sum_{j \in R_i} \exp[\beta^t Z_j(t_i)]}. \end{aligned}$$

Der Partial Likelihood ist nun als Produkt dieser Wahrscheinlichkeiten definiert:

$$L(\beta) = \prod_{i=1}^k \frac{\exp[\beta^t Z_i(t_i)]}{\sum_{j \in R_i} \exp[\beta^t Z_j(t_i)]}$$

und

$$LL(\beta) = \log(L(\beta)) = \sum_{i=1}^k \beta^t Z_i(t_i) - \sum_{i=1}^k \left[\log\left(\sum_{j \in R_i} \exp[\beta^t Z_j(t_i)] \right) \right] \quad (3.22)$$

der zugehörige Log-Likelihood. Um die Likelihood-Funktion zu maximieren benötigt man noch den sogenannten Score-Vektor:

$$U(\beta) = \left(\frac{\partial}{\partial \beta_1} LL(\beta), \dots, \frac{\partial}{\partial \beta_p} LL(\beta) \right)^t \in \mathbb{R}^p, \quad (3.23)$$

sowie die Matrix der zweiten Ableitungen:

$$\frac{\partial}{\partial \beta} U(\beta) = -I(\beta) = \begin{pmatrix} \frac{\partial^2}{\partial \beta_1 \partial \beta_1} LL(\beta) & \cdots & \frac{\partial^2}{\partial \beta_1 \partial \beta_p} LL(\beta) \\ \vdots & \ddots & \vdots \\ \frac{\partial^2}{\partial \beta_p \partial \beta_1} LL(\beta) & \cdots & \frac{\partial^2}{\partial \beta_p \partial \beta_p} LL(\beta) \end{pmatrix} \in \mathbb{R}^{p \times p}.$$

Bemerkung: Die Matrix I heißt Informationsmatrix.

Nun kann mittels numerischer Verfahren (zum Beispiel Verfahren von Newton-Raphson) die Nullstelle von $U(\beta)$ bestimmt werden, d.h. der Partial-Likelihood maximiert werden. Eine Beschreibung dieses Verfahrens kann man in Klein (1997) [25] Anhang A nachlesen. Auf wichtige Eigenschaften des Schätzers eingehen werden wir genauer in Abschnitt 3.7.4 eingehen.

3.7.2 Partial Likelihood mit “ties”

Obwohl die Wahrscheinlichkeit dafür, daß zwei oder mehr Ereignisse (“ties”) zum selben Zeitpunkt passieren, Null ist, wird man in der Praxis, aufgrund begrenzter Genauigkeit in der Zeitmessung, oft mit Datensätzen konfrontiert, in denen mehrere Ereignisse an ein und demselben Zeitpunkt passieren. Während in Kapitel 3.7.1 die Herleitung des Partial Likelihood für den Fall, daß genau ein Ereignis pro Zeitpunkt passiert, vorgenommen wurde, beschäftigt sich dieser Teil mit verschiedenen Techniken zur Bestimmung des Partial Likelihood für mehrere Ereignisse pro Zeitpunkt.

Seien nun wiederum t_1, \dots, t_k die geordneten Zeitpunkte, an denen ein Ereignis passiert, D_i die Menge der Toten zum Zeitpunkt t_i mit $d_i = |D_i|$ und R_i die Risikomenge zum Zeitpunkt t_i . Sei nun zusätzlich:

$$Q_i = \{M \subset R_i \mid |M| = d_i\}$$

die Menge aller Teilmengen der Risikomenge mit genau d_i Elementen. Für jedes Element $q = (q_1, \dots, q_{d_i})^t$ aus Q_i sei nun

$$s_q^*(t_i) := \sum_{j=1}^{d_i} Z_{q_j}(t_i)$$

die Summe der Kovariablenvektoren über alle Elemente aus q . Für die Menge D_i definieren wir mit

$$s(t_i) := \sum_{j \in D_i} Z_j(t_i)$$

die zugehörige Summe der Kovariablen aller Objekte (insgesamt d_i), die zum Zeitpunkt t_i sterben. Man kann analog zu (3.22) den diskreten Partial-Likelihood, d.h. die Wahrscheinlichkeit, daß zu jedem Zeitpunkt t_i genau die beobachteten Personen sterben, gegeben daß d_i Personen sterben, herleiten

$$L_1(\beta) = \prod_{i=1}^k \frac{\exp(\beta^t s(t_i))}{\sum_{q \in Q_i} \exp(\beta^t s_q^*(t_i))}. \quad (3.24)$$

Für große “ties” erweist sich die Maximierung von 3.24 als numerisch zu aufwendig, da Q_i $\binom{|R_i|}{|D_i|}$ Elemente enthält. Um den Partial Likelihood zumindest näherungsweise bestimmen zu können, wurden verschiedene Approximationen entwickelt. Die erste, hier mit $L_2(\beta)$ bezeichnet, wurde von Breslow (1974) [8] vorgeschlagen

$$L_2(\beta) = \prod_{i=1}^k \frac{\exp(\beta^t s(t_i))}{[\sum_{j \in R_i} \exp(\beta^t Z_j(t_i))]^{d_i}}. \quad (3.25)$$

Die Approximation von Efron (1977)[12] für den diskreten Partial Likelihood ($L_3(\beta)$) ist gegeben durch

$$L_3(\beta) = \prod_{i=1}^k \frac{\exp(\beta^t s(t_i))}{\prod_{j=1}^{d_i} [\sum_{r \in R_i} \exp(\beta^t Z_r(t_i)) - \frac{i-1}{d_i} \sum_{r \in D_i} \exp(\beta^t Z_r(t_i))]} \quad (3.26)$$

3.7.3 Schätzung des Basis-Hazards (Breslow)

In Abschnitt 3.7.1 wurden Techniken gezeigt, wie man ohne Kenntnis des Basis-Hazards den Koeffizienten β schätzt. Nun soll die zugrundeliegende Hazard-Funktion geschätzt werden. Die Grundlage hierfür ist ein Log-Likelihood-Ansatz, ähnlich dem zur Bestimmung des Nelson-Aalen-Schätzers (3.14).

Um den Schätzer zu erhalten, maximiert man den Likelihood in Abhängigkeit von den beobachteten Daten $(t_1, \delta_1), \dots, (t_n, \delta_n)$, wobei hier wiederum t_i das Beobachtungsende und δ_i der Zensierungsindikator ist

$$L(\lambda_0(t)) = \prod_{i=1}^n [\lambda_0(t_i) \exp(\beta^t Z_i(t_i))]^{\delta_i} \exp[-\Lambda_0(t_i) \exp(\beta^t Z_i)].$$

Gehen wir nun davon aus, daß die Schätzung $\hat{\beta}$ des Koeffizientenvektors β bereits mit dem Partial-Likelihood-Ansatz durchgeführt wurde. Analog zur Herleitung von (3.14) erhält man durch Maximierung des Log-Likelihood (unter Verwendung des schon geschätzten Parameters $\hat{\beta}$) den sogenannten Breslow-Schätzer für den kumulativen Basis-Hazard $\Lambda_0(t) = \int_0^t \lambda_0(s) ds$

$$\hat{\Lambda}_0(t) = \sum_{t_i \leq t} \frac{d_i}{\sum_{j \in R_i} \exp(\hat{\beta}^t Z_j(t_i))}. \quad (3.27)$$

Hier sind $t_1 < \dots < t_k$ die geordneten Zeitpunkte, an denen ein Ereignis passiert, d_i die Anzahl der Ereignisse zum Zeitpunkt t_i , sowie R_i die Risikomenge zum Zeitpunkt t_i .

3.7.4 Asymptotische Eigenschaften

In diesem Abschnitt beweisen wir die Konvergenz des geschätzten Regressionsparameter $\hat{\beta}$ gegen den wahren Wert β_0 , sowie asymptotische Aussagen über die Verteilung von $\hat{\beta}$. Schließlich zeigen wir noch die Konvergenz des Breslow-Schätzers für den Basis-Hazard. Der Beweis folgt der Darstellung von Andersen und Gill (1982) [2] Dabei bedient man sich erneut einiger Aussagen aus der Martingaltheorie, wobei wiederum der Grenzwertsatz für Martingale von Rebolledo (Satz

2.36) und die Ungleichung von Lengart (Satz 2.37) eine zentrale Rolle spielen. Um asymptotische Eigenschaften zu beweisen, betrachten wir eine Folge multivariater Zählprozesse $N_i^{(n)}(t)$. Der Zeitparameter t nimmt Werte aus einem kompakten Intervall an, d.h. o.E. $t \in \mathcal{T} = [0, 1]$, wobei $N_i^{(n)}(t)$ die beobachteten Ereignisse für Individuum i zum Zeitpunkt t zählt. Jede Komponente des so definierten Zählprozesses hat den Intensitätsprozeß $\lambda^{(n)}(t) = (\lambda_1^{(n)}(t), \lambda_2^{(n)}(t), \dots, \lambda_n^{(n)}(t))$, mit

$$\lambda_i^{(n)}(t) = Y_i^{(n)}(t) \lambda_0(t) \exp(\beta_0^t Z_i^{(n)}(t)), \quad (3.28)$$

wobei $Z_i^{(n)}(t)$ der Kovariablenvektor für Individuum i ist, $\beta_0 \in \mathbb{R}^p$ der zu schätzende Regressionskoeffizient und $Y_i^{(n)}(t)$ ein vorhersehbarer Indikatorprozeß mit

$$Y_i^{(n)}(t) = \begin{cases} 1 & \text{i-tes Objekt zum Zeitpunkt } t \text{ unter Beobachtung} \\ 0 & \text{sonst} \end{cases}.$$

Insbesondere springt der Zählprozeß $N_i^{(n)}(t)$ nur dann, wenn $Y_i^{(n)}(t) = 1$. Diese Formulierung ist eine Erweiterung des ursprünglichen Modells von Cox, da hier mehrere Ereignisse pro Objekt beobachtet werden können. Für die Modellierung des Übergangs "lebendig-tot" nimmt der Zählprozeß nur Werte aus $\{0, 1\}$ an.

Da $\int_0^t \lambda_i^{(n)}(s) ds$ der vorhersehbare Kompensator des Zählprozesses $N_i^{(n)}(t)$ ist, kann man nun eine Folge lokal quadratintegrierbarer Martingale $M^{(n)}(t) = (M_1^{(n)}(t), M_2^{(n)}(t), \dots, M_n^{(n)}(t))$ auf dem Intervall $[0, 1]$ definieren

$$M_i^{(n)}(t) = N_i^{(n)}(t) - \int_0^t \lambda_i^{(n)}(u) du, \quad (3.29)$$

die nach Satz 2.35 die vorhersehbaren (Ko)variationsprozesse

$$\langle M_i^{(n)}, M_j^{(n)} \rangle(t) = \delta_{ij} \int_0^t \lambda_i^{(n)}(u) du$$

besitzen, d.h. die Martingale $M_i^{(n)}(t)$ und $M_j^{(n)}(t)$ sind orthogonal für $i \neq j$.

Im Folgenden werden die hochgestellten Indizes (n) weggelassen. Mit diesem Zählprozeßmodell läßt sich jetzt der Partial-Likelihood (3.22) verallgemeinern

$$LL(\beta, t) = \sum_{i=1}^n \int_0^t \beta^t Z_i(s) dN_i(s) - \int_0^t \log \left\{ \sum_{i=1}^n Y_i(s) \exp(\beta^t Z_i(s)) \right\} d\bar{N}(s), \quad (3.30)$$

wobei

$$\bar{N}(s) = \sum_{i=1}^n N_i(s).$$

Man erkennt, daß das Modell (3.22) $LL(\beta, 1)$ nach dem Zählprozeßmodell entspricht. Der geschätzte Parameter $\hat{\beta}$ ist dann die Lösung der Gleichung $U(\beta, 1) = (\partial/\partial\beta)LL(\beta, 1) = 0$, wobei

$$U(\beta, t) = \sum_{i=1}^n \int_0^t Z_i(s) dN_i(s) - \int_0^t \frac{\sum_{i=1}^n Y_i(s) Z_i(s) \exp(\beta^t Z_i(s))}{\sum_{i=1}^n Y_i(s) \exp(\beta^t Z_i(s))} d\bar{N}(s).$$

Die Gleichung $U(\beta, 1) = (\partial/\partial\beta)LL(\beta, 1) = 0$ kann man mit dem Verfahren von Newton-Raphson unter Zuhilfenahme der negativ semidefiniten Matrix der zweiten Ableitungen

$$\begin{aligned} \frac{\partial}{\partial\beta}U(\beta, t) &= -I(\beta, t) \\ &= \int_0^t \left[\left(\frac{\sum_{i=1}^n Y_i(s) Z_i(s) \exp(\beta^t Z_i(s))}{\sum_{i=1}^n Y_i(s) \exp(\beta^t Z_i(s))} \right) \left(\frac{\sum_{i=1}^n Y_i(s) Z_i(s) \exp(\beta^t Z_i(s))}{\sum_{i=1}^n Y_i(s) \exp(\beta^t Z_i(s))} \right)^t \right. \\ &\quad \left. - \frac{\sum_{i=1}^n Y_i(s) Z_i(s) Z_i(s)^t \exp(\beta^t Z_i(s))}{\sum_{i=1}^n Y_i(s) \exp(\beta^t Z_i(s))} \right] d\bar{N}(s) \end{aligned} \quad (3.31)$$

numerisch lösen kann.

Im folgenden Korollar wird für den Parameter β_0 $U(\beta_0, 1)$ als stochastisches Integral mit lokalem Martingal als Integrator hergeleitet.

Korollar 3.8 *Für den wahren Regressionskoeffizienten β_0 hat $(\partial/\partial\beta)LL(\beta, t)$ folgende Darstellung:*

$$U(\beta_0, t) = \sum_{i=1}^n \int_0^t Z_i(s) dM_i(s) - \int_0^t \frac{\sum_{i=1}^n Y_i(s) Z_i(s) \exp(\beta_0^t Z_i(s))}{\sum_{i=1}^n Y_i(s) \exp(\beta_0^t Z_i(s))} d\bar{M}(s), \quad (3.32)$$

mit

$$M_i(t) = N_i(t) - \int_0^t \lambda_i(s) ds$$

und

$$\bar{M}(t) = \sum_{i=1}^n M_i(t).$$

Beweis:

Um zu zeigen, daß (3.32) gilt, muß man $U(\beta_0, t)$ folgendermaßen darstellen können:

$$\begin{aligned} U(\beta_0, t) &= \sum_{i=1}^n \int_0^t Z_i(s) d(N_i(s) - \int_0^s \lambda_i(u) du) \\ &\quad - \int_0^t \frac{\sum_{i=1}^n Y_i(s) Z_i(s) \exp(\beta_0^t Z_i(s))}{\sum_{i=1}^n Y_i(s) \exp(\beta_0^t Z_i(s))} d(\bar{N}(s) - \sum_{i=1}^n \int_0^s \lambda_i(u) du). \end{aligned}$$

Mit den Zusammenhang aus (3.28) zeigt man nun, daß

$$\begin{aligned} \sum_{i=1}^n \int_0^t Z_i(s) d\lambda_i(s) &= \int_0^t \frac{\sum_{i=1}^n Y_i(s) Z_i(s) \exp(\beta_0^t Z_i(s))}{\sum_{i=1}^n Y_i(s) \exp(\beta_0^t Z_i(s))} d\left(\sum_{i=1}^n \int_0^s \lambda_i(u) du\right) \\ &\iff \\ \int_0^t \sum_{i=1}^n \lambda_0(s) Y_i(s) \exp(\beta_0^t Z_i(s)) ds &= \int_0^t \frac{\sum_{i=1}^n Y_i(s) Z_i(s) \exp(\beta_0^t Z_i(s))}{\sum_{i=1}^n Y_i(s) \exp(\beta_0^t Z_i(s))} \sum_{i=1}^n \lambda_0(s) \exp(\beta_0^t Z_i(s)) Y_i(s) ds. \end{aligned}$$

Die Integrale auf der rechten und linken Seite sind genau dann gleich, wenn der Integrand für alle t gleich ist, so daß zu zeigen bleibt, daß

$$\sum_{i=1}^n \lambda_0(s) Y_i(s) Z_i(s) \exp(\beta_0^t Z_i(s)) = \frac{\sum_{i=1}^n Y_i(s) Z_i(s) \exp(\beta^t Z_i(s))}{\sum_{i=1}^n Y_i(s) \exp(\beta_0^t Z_i(s))} \sum_{i=1}^n \lambda_0(s) \exp(\beta_0^t Z_i(s)) Y_i(s).$$

Dies ist offensichtlich der Fall, so daß die Aussage gezeigt ist. \square

Im folgenden seien nun noch einige wichtige Definitionen und Bedingungen definiert, die bei der Herleitung der asymptotischen Aussagen oft benötigt werden.

Definition 3.9 Für eine Matrix A mit Elementen $a_{i,j}$ sowie für einen Vektor $a = a_1, \dots, a_n$ sind die Normen $\|A\| = \sup_{i,j} |a_{i,j}|$, $\|a\| = \sup_i |a_i|$, sowie $|a| = \sqrt{a^t a}$ definiert. Weitere wichtige Definitionen sind:

$$\begin{aligned} S^{(0)}(\beta, t) &= \frac{1}{n} \sum_{i=1}^n Y_i(t) \exp(\beta^t Z_i(t)) \in \Omega \times \mathcal{T} \mapsto \mathbb{R}, \\ S^{(1)}(\beta, t) &= \frac{1}{n} \sum_{i=1}^n Z_i(t) Y_i(t) \exp(\beta^t Z_i(t)) \in \Omega \times \mathcal{T} \mapsto \mathbb{R}^p, \\ S^{(2)}(\beta, t) &= \frac{1}{n} \sum_{i=1}^n Z_i(t) Z_i(t)^t Y_i(t) \exp(\beta^t Z_i(t)) \in \Omega \times \mathcal{T} \mapsto \mathbb{R}^{p \times p}, \\ E(\beta, t) &= \frac{S^{(1)}(\beta, t)}{S^{(0)}(\beta, t)} \in \Omega \times \mathcal{T} \mapsto \mathbb{R}^p, \\ V(\beta, t) &= \frac{S^{(2)}(\beta, t)}{S^{(0)}(\beta, t)} - E(\beta, t) E(\beta, t)^t \in \Omega \times \mathcal{T} \mapsto \mathbb{R}^{p \times p}. \end{aligned}$$

Bemerkung: $S^{(0)}$ ist also ein stochastischer Prozeß, der, da $Y_i(t)$ und $Z_i(t)$ vorhersehbar sind, ebenfalls vorhersehbar ist. $S^{(1)}$ ist ein p -variater vorhersehbarer Prozeß und $S^{(2)}$ eine $p \times p$ -Matrix mit vorhersehbaren Prozessen als Einträgen.

Bedingungen:

- A. (Endliches Intervall). $\int_0^1 \lambda_0(t) dt < \infty$.
- B. (Asymptotische Stabilität). Es gibt eine Nachbarschaft \mathcal{B} von β_0 und Folgen skalar-, matrix- und vektorwertiger Funktionen $s^{(0)}, s^{(1)}, s^{(2)}$, definiert auf $\mathcal{B} \times [0, 1]$, so daß für $j = 0, 1, 2$ gilt

$$\sup_{t \in [0, 1], \beta \in \mathcal{B}} \|S^{(j)}(\beta, t) - s^{(j)}(\beta, t)\| \rightarrow_{\mathcal{P}} 0.$$

- C. (Lindeberg-Bedingung). Es gibt ein $\delta > 0$, so daß

$$\frac{1}{\sqrt{n}} \sup_{i,t} |Z_i(t) Y_i(t) I\{\beta_0^t Z_i(t) > -\delta |Z_i(t)|\}| \rightarrow_{\mathcal{P}} 0.$$

D. (Asymptotische Regularitäts-Bedingungen). Seien \mathcal{B} , $s^{(0)}$, $s^{(1)}$ und $s^{(2)}$ gegeben wie in Bedingung B und $e = s^{(1)}/s^{(0)}$, sowie $e = s^{(1)}/s^{(0)}$ und $v = s^{(2)}/s^{(0)} - ee^t$. Für alle $\beta \in \mathcal{B}$ und $t \in [0, 1]$ gelte:

$$s^{(1)}(\beta, t) = \frac{\partial}{\partial \beta} s^{(0)}(\beta, t), \quad s^{(2)} = \frac{\partial^2}{\partial \beta^2} s^{(0)}(\beta, t).$$

Die Funktionen $s^{(0)}(\cdot, t)$, $s^{(1)}(\cdot, t)$ und $s^{(2)}(\cdot, t)$ seien stetig in $\beta \in \mathcal{B}$, gleichmäßig in $t \in [0, 1]$, $s^{(0)}$, $s^{(1)}$, $s^{(2)}$ seien beschränkt in $\mathcal{B} \times [0, 1]$; $s^{(0)} \neq 0$ auf $\mathcal{B} \times [0, 1]$ und die Matrix

$$\Sigma = \int_0^1 v(\beta_0, t) s^{(0)}(\beta_0, t) \lambda_0(t) dt$$

sei positiv definit

Zuletzt wird noch ein Korollar aus der konvexen Analysis benötigt:

Korollar 3.10 *Sei B eine offene konvexe Menge im \mathbb{R}^p and F_1, F_2, \dots , eine Folge zufälliger konkaver Funktionen auf B , so daß $\forall x \in B, F_n(x) \rightarrow_{\mathcal{P}} f(x)$ für $n \rightarrow \infty$, wobei $f(x)$ eine beliebige reellwertige Funktion ist, dann ist f auch konkav und für alle kompakten Mengen $A \subset B$ gilt:*

$$\sup_{x \in A} |F_n(x) - f(x)| \rightarrow_{\mathcal{P}} 0 \quad \text{für } n \rightarrow \infty.$$

Falls f genau ein Maximum in $\hat{x} \in B$ hat und die Zufallsvariable \hat{X}_n die Funktion F_n maximiert, dann gilt:

$$\hat{X}_n \rightarrow_{\mathcal{P}} \hat{x} \quad \text{für } n \rightarrow \infty.$$

Beweis:

Der Beweis findet sich zum Beispiel in Andersen und Gill (1982) [2] Seite 1116. □

Satz 3.11 (Konsistenz von $\hat{\beta}$) *Falls die Bedingungen A, B und D gelten, dann konvergiert $\hat{\beta}$ in Wahrscheinlichkeit gegen β_0 für $n \rightarrow \infty$.*

Beweis:

Betrachte die folgendermaßen definierten Prozesse X und A :

$$\begin{aligned} X(\beta, t) &:= \frac{1}{n} (LL(\beta, t) - LL(\beta_0, t)) \\ &= \frac{1}{n} \left[\sum_{i=1}^n \int_0^t (\beta - \beta_0)^t Z_i(s) dN_i(s) - \int_0^t \log \left(\frac{S^{(0)}(\beta, u)}{S^{(0)}(\beta_0, u)} \right) d\bar{N}(s) \right], \end{aligned}$$

$$\begin{aligned} A(\beta, t) &:= \frac{1}{n} \left[\sum_{i=1}^n \int_0^t (\beta - \beta_0)^t Z_i(u) \lambda_i(u) du - \int_0^t \log \left\{ \frac{S^{(0)}(\beta, u)}{S^{(0)}(\beta_0, u)} \right\} \bar{\lambda}(u) du \right] \\ &= \int_0^t \left[(\beta - \beta_0)^t S^{(1)}(\beta_0, u) - \log \left\{ \frac{S^{(0)}(\beta, u)}{S^{(0)}(\beta_0, u)} \right\} S^{(0)}(\beta_0, u) \right] \lambda_0(u) du, \end{aligned}$$

mit

$$\begin{aligned}\bar{\lambda}(u) &= \sum_{i=1}^n \lambda_i(u) = \lambda_0(u) \sum_{i=1}^n Y_i(u) \exp(\beta_0^t Z_i(t)) \\ &= \lambda_0(u) n S^{(0)}(\beta_0, u).\end{aligned}$$

Für jedes β ist nach Definition

$$X(\beta, t) - A(\beta, t) = \frac{1}{n} \sum_{i=1}^n \left[\int_0^t \left[(\beta - \beta_0)^t Z_i(u) - \log \left\{ \frac{S^{(0)}(\beta, u)}{S^{(0)}(\beta_0, u)} \right\} \right] d(N_i(u) - \int_0^u \lambda_i(s) ds) \right]$$

ein stochastisches Integral mit vorhersehbarem Prozeß als Integrand und dem lokalen Martingal $N_i(\cdot) - \int_0^\cdot \lambda_i(s) ds$ als Integrator. Nach Satz 2.29 ist $X - A$ demnach selbst ein lokales Martingal und hat nach Satz 2.35 folgenden vorhersehbaren Variationsprozeß:

$$\begin{aligned}\langle X(\beta, t) - A(\beta, t) \rangle &= \frac{1}{n^2} \sum_{i=1}^n \int_0^t \left[(\beta - \beta_0)^t Z_i(u) - \log \left\{ \frac{S^{(0)}(\beta, u)}{S^{(0)}(\beta_0, u)} \right\} \right]^2 \underbrace{d\left(\int_0^u \lambda_i(s) ds \right)}_{\lambda_i(u)} du \\ &= \frac{1}{n} \int_0^t \left((\beta - \beta_0)^t S^{(2)}(\beta_0, u) (\beta - \beta_0) \right. \\ &\quad \left. - 2(\beta - \beta_0)^t S^{(1)}(\beta_0, u) \log \left(\frac{S^{(0)}(\beta, u)}{S^{(0)}(\beta_0, u)} \right) \right. \\ &\quad \left. + \left[\log \left(\frac{S^{(0)}(\beta, u)}{S^{(0)}(\beta_0, u)} \right) \right] S^{(0)}(\beta_0, u) \lambda_0(u) du \right) \quad (3.33)\end{aligned}$$

Nach den Voraussetzungen A, B und D gilt für alle $\beta \in \mathcal{B}$, daß

$$A(\beta, 1) \rightarrow_{\mathcal{P}} \int_0^1 \left[(\beta - \beta_0)^t s^{(1)}(\beta_0, u) - \log \left\{ \frac{s^{(0)}(\beta, u)}{s^{(0)}(\beta_0, u)} \right\} s^{(0)}(\beta_0, u) \right] \lambda_0(u) du$$

Außerdem gilt, daß der Prozeß $n\langle X - A \rangle$ auf \mathcal{B} gegen eine von β abhängige Konstante $c(\beta)$ konvergiert

$$n\langle X - A \rangle \rightarrow_{\mathcal{P}} c(\beta) < \infty \quad \forall t \in [0, 1], \beta \in \mathcal{B}.$$

Dies sieht man aus (3.33), denn $S^{(0)}(\beta, u)$, $S^{(1)}(\beta, u)$ und $S^{(2)}(\beta, u)$ konvergieren für $\beta \in \mathcal{B}$ nach Voraussetzung B gegen $s^{(0)}(\beta, u)$, $s^{(1)}(\beta, u)$, sowie $s^{(2)}(\beta, u)$. Zudem ist der Wert $s^{(0)}(\beta_0, u) \neq 0$ nach Bedingung D. Demzufolge konvergiert der Integrand von (3.33) gegen einen von β abhängigen Wert. Nachdem das Integral $\int_0^1 \lambda(t) dt < \infty$ konvergiert der gesamte Term.

Aus Version (2.20) der Ungleichung von Lenglart und $n\langle X - A \rangle \rightarrow_{\mathcal{P}} c(\beta)$ folgt:

$$P\left(\sup_{t \in [0, 1]} |X(t, \beta) - A(t, \beta)| > \eta \right) \leq \frac{\delta}{\eta^2} + P\left(\underbrace{\langle X - A \rangle(1, \beta)}_{\rightarrow_{\mathcal{P}} 0 \text{ für } n \rightarrow \infty} > \delta \right) \quad \forall \delta, \eta > 0,$$

woraus mit $\eta = \delta \rightarrow 0$ folgt, daß

$$\|X(\beta, t) - A(\beta, t)\| \rightarrow_{\mathcal{P}} 0 \quad \text{für } n \rightarrow \infty \quad \text{und } \beta \in \mathcal{B},$$

d.h. X und A konvergieren in Wahrscheinlichkeit gegen den selben Grenzwert. Nach Voraussetzung D kann man nun die erste und zweite Ableitung des Prozesses A berechnen und erhält:

$$\frac{\partial}{\partial \beta} A(\beta, 1) = \int_0^1 \left[s^{(1)}(\beta_0, u) - s^{(1)}(\beta, u) \frac{s^{(0)}(\beta_0, u)}{s^{(0)}(\beta, u)} \right] \lambda_0(u) du,$$

sowie

$$\frac{\partial^2}{\partial \beta^2} A(\beta, 1) = - \underbrace{\int_0^1 v(\beta, u) s^{(0)}(\beta, u) \lambda_0(u) du}_{\text{neg. definit für } \beta_0 \text{ nach Vor. D}}.$$

Die erste Ableitung wird 0 in β_0 und die zweite ist kleiner 0 für β_0 . Daher konvergiert $X(\beta, 1)$ in Wahrscheinlichkeit $\forall \beta \in \mathcal{B}$ gegen eine konkave Zufallsfunktion mit genau einem Maximum in $\beta = \beta_0$. Da $\hat{\beta}$ die konkave Zufallsfunktion $X(\cdot, 1)$ maximiert, folgt aus Korollar 3.10, daß $\hat{\beta} \rightarrow_{\mathcal{P}} \beta_0$. \square

Kommen wir nun zur asymptotischen Normalverteilung von β_0 . Sie bildet die Grundlage zur Herleitung asymptotischer Tests und zur Bestimmung von Konfidenzintervallen.

Satz 3.12 (Asymptotische Normalität von $\hat{\beta}$) Falls die Voraussetzungen A bis D erfüllt sind, gilt für die Folge $\hat{\beta}^{(n)}$ (mit $\hat{\beta}^{(n)} \rightarrow_{\mathcal{P}} \beta_0$ für $n \rightarrow \infty$)

$$\sqrt{n}(\hat{\beta} - \beta_0) \rightarrow_{\mathcal{D}} N(0, \Sigma^{-1}).$$

Bemerkung: Da nach dem vorhergehenden Satz der Schätzer $\hat{\beta}$ für β konsistent ist, ist demnach $\hat{\beta}$ asymptotisch normalverteilt.

Beweis:

Zunächst stellen wir fest, daß wir in einer Umgebung um β_0 , in der $\frac{d}{d\beta} U(\beta, 1)$ existiert für $U(\beta_0, 1)$ folgende Taylor-Entwicklung erhalten

$$U(\beta_0, 1) = U(\beta, 1) - \frac{d}{d\beta} U(\beta, 1)(\beta - \beta_0) + \text{Terme höherer Ordnung},$$

was äquivalent ist zu

$$U(\beta, 1) - U(\beta_0, 1) = -I(\beta^*, 1)(\beta - \beta_0), \quad (3.34)$$

wobei β^* auf dem Liniensegment zwischen β und β_0 liegt und $-I(\beta^*, 1)$ die Matrix der zweiten Ableitungen (hinsichtlich β) des Log-Likelihood an der Stelle $t = 1$ und $\beta = \beta^*$. einsetzen von $\hat{\beta}$ in (3.34) ergibt

$$\frac{1}{\sqrt{n}} U(\beta_0, 1) = \left[\frac{1}{n} I(\beta^*, 1) \right] \sqrt{n}(\hat{\beta} - \beta_0).$$

Um nun die asymptotische Normalität von $\sqrt{n}(\hat{\beta} - \beta_0)$ zu beweisen, reicht es zu zeigen, daß der Prozeß $(1/\sqrt{n})U(\beta_0, \cdot)$ in Verteilung gegen einen stochastischen Prozeß konvergiert, dessen Zuwächse p -variat normalverteilt sind mit Kovarianzmatrix Σ^{-1} und dann, daß $(1/n)I(\hat{\beta}, 1)$ in Wahrscheinlichkeit gegen Σ konvergiert.

Das weitere Vorgehen ist demnach in 2 Abschnitte untergliedert. Zunächst zeigen wir, daß

$$\frac{1}{\sqrt{n}}U(\beta_0, 1) \rightarrow_{\mathcal{D}} N(0, \Sigma)$$

und dann, daß

$$\frac{1}{n}I\beta^*, 1) \rightarrow_{\mathcal{P}} \Sigma$$

Für irgendeine zufälliges $\beta^* = \beta^{*(n)}$ mit $\beta^{*(n)} \rightarrow_{\mathcal{P}} \infty$ für $n \rightarrow \infty$.

Für den ersten Teil nutzen wir die Tatsache, daß wir nach (3.32) gilt

$$\begin{aligned} \frac{1}{\sqrt{n}}U(\beta_0, t) &= \frac{1}{\sqrt{n}} \sum_{i=1}^n \int_0^t Z_i(u) dM_i(u) - \frac{1}{\sqrt{n}} \int_0^t E(\beta_0, u) d\bar{M}(u) \\ &= \sum_{i=1}^n \int_0^t \underbrace{\frac{1}{\sqrt{n}}[Z_i(u) - E(\beta_0, u)]}_{H_i(u)} dM_i(u). \end{aligned}$$

$(1/\sqrt{n})U(\beta_0, t)$ läßt sich also darstellen als Summe stochastischer Integrale mit den vorhersehbaren Prozessen $H_i(t), i \in \{1, \dots, n\}$ als Integranden und den lokal quadratintegrierbaren Martingalen $M_i(t), i \in \{1, \dots, n\}$ als Integratoren. Somit hat $(1/\sqrt{n})U(\beta_0, 1)$ einen vorhersehbaren Variationsprozeß, auf den wir nun den Satz von Rebolledo (Satz 2.36) anwenden.

Um die Bedingung (2.14) des Satzes 2.36 zu verifizieren, stellen wir zunächst fest, daß der vorhersehbare Variationsprozeß $\langle (1/\sqrt{n})U(\beta_0, t) \rangle$ nach Satz 2.35 folgende Darstellung besitzt:

$$\left\langle \frac{1}{\sqrt{n}}U(\beta_0, t) \right\rangle = \sum_{i=1}^n \int_0^t H_i(u) H_i(u)^t d\langle M_i(u) \rangle.$$

Bemerkung: $\langle \frac{1}{\sqrt{n}}U(\beta_0, t) \rangle$ ist also eine $p \times p$ -Matrix mit lokales Submartingalen als Einträgen.

$\langle M_i(u) \rangle$ ist nach Satz 2.35 gegeben durch

$$\Lambda_i(u) = \int_0^u \lambda_0(s) Y_i(s) \exp(\beta_0^t Z_i(s)) ds,$$

so daß wir

$$\begin{aligned} \left\langle \frac{1}{\sqrt{n}}U(\beta_0, 1) \right\rangle &= \int_0^1 \sum_{i=1}^n \frac{1}{n} [Z_i(u) - E(\beta_0)] [Z_i(u) - E(\beta_0)]^t \exp(\beta_0^t Z_i(u)) \lambda_0(u) du \\ &= \int_0^1 \left[S^{(2)}(\beta_0, u) - \frac{S^{(1)}(\beta_0, u) S^{(1)}(\beta_0, u)^t}{S^{(0)}(\beta_0, u)} \right] \lambda_0(u) du \end{aligned}$$

erhalten. Nach Bedingungen A, B und D konvergiert dieser Term in Wahrscheinlichkeit

$$\int_0^1 \left[S^{(2)}(\beta_0, u) - \frac{S^{(1)}(\beta_0, u)S^{(1)}(\beta_0, u)^t}{S^{(0)}(\beta_0, u)} \right] \lambda_0(u) du \rightarrow_{\mathcal{P}} \int_0^t v(\beta_0, u) s^{(0)}(\beta_0, u) \lambda_0(u) du.$$

Um die Bedingung (2.15) zu verifizieren, zeigen wir zunächst die Gültigkeit der folgenden Ungleichung, die für zwei reelle Zahlen a, b gilt

$$|a - b|^2 I(|a - b| > \varepsilon) \leq 4|a|^2 I(|a| > \frac{\varepsilon}{2}) + 4|b|^2 I(|b| > \frac{\varepsilon}{2}) \quad (3.35)$$

Denn damit $|a - b| > \varepsilon$ muß entweder $|a| > \frac{\varepsilon}{2}$ oder $|b| > \frac{\varepsilon}{2}$ sein. Wären nämlich sowohl $|a| \leq \frac{\varepsilon}{2}$, als auch $|b| \leq \frac{\varepsilon}{2}$, dann wäre

$$|a - b| \leq |a| + |b| \leq 2 \frac{\varepsilon}{2} = \varepsilon,$$

was einen Widerspruch zu $|a - b| > \varepsilon$ darstellt. Das führt dann zu

$$\begin{aligned} |a - b|^2 I(|a - b| > \varepsilon) &\leq |a - b|^2 \max[I(|a| > \frac{\varepsilon}{2}), I(|b| > \frac{\varepsilon}{2})] \\ &\leq |a - b|^2 [I(|a| > \frac{\varepsilon}{2}) + I(|b| > \frac{\varepsilon}{2})] \\ &\leq |a|^2 [I(|a| > \frac{\varepsilon}{2}) + I(|b| > \frac{\varepsilon}{2})] + |b|^2 [I(|a| > \frac{\varepsilon}{2}) + I(|b| > \frac{\varepsilon}{2})] \\ &\leq 4|a|^2 I(|a| > \frac{\varepsilon}{2}) + 4|b|^2 I(|b| > \frac{\varepsilon}{2}). \end{aligned}$$

Um die Bedingung (2.18) zu zeigen, bilden wir mit Hilfe der komponentenweisen Darstellung von H

$$H_{jh} = \frac{1}{\sqrt{n}} (Z_h(t) - E_h(\beta_0, t))_j \quad j \in \{1, \dots, p\} \quad h \in \{1, \dots, n\}.$$

Analog zu (2.16) konstruieren wir uns nun ein lokales Martingal $M_{j\varepsilon}$, das alle Sprünge von $(1/\sqrt{n})U(\beta_0, t)$ enthält, deren Absolutbetrag größer ist als ε :

$$M_{j\varepsilon}(t) = \sum_{h=1}^n \int_0^t H_{jh}(s) I(|H_{jh}(s)| > \varepsilon) (dN_h - \lambda_h(s) ds).$$

Die Zuhilfenahme von (3.35) führt zu folgender Abschätzung des vorhersehbaren Kompensatorprozesses $\langle M_{j\varepsilon}, M_{j\varepsilon} \rangle$ an der Stelle 1:

$$\begin{aligned} \langle M_{j\varepsilon}, M_{j\varepsilon} \rangle(1) &= \int_0^1 \sum_{h=1}^n [H_{jh}(s)]^2 I(|H_{jh}(s)| > \varepsilon) \lambda_h(s) ds \\ &= \int_0^1 \sum_{h=1}^n \frac{1}{n} [(Z_h(s) - E(\beta_0, s))_j]^2 I(n^{-1/2} |(Z_h(s) - E(\beta_0, s))_j| > \varepsilon) \lambda_h(s) ds \\ &\stackrel{(3.35)}{\leq} 4 \left\{ \int_0^1 \sum_{h=1}^n \frac{1}{n} [(E(\beta_0, s))_j]^2 I(n^{-1/2} |(E(\beta_0, s))_j| > \frac{\varepsilon}{2}) \lambda_h(s) ds \right. \\ &\quad \left. + \int_0^1 \sum_{h=1}^n \frac{1}{n} [(Z_h(s))_j]^2 I(n^{-1/2} |(Z_h(s))_j| > \frac{\varepsilon}{2}) \lambda_h(s) ds \right\} \end{aligned}$$

$$= 4 \left\{ \int_0^1 \sum_{h=1}^n \frac{1}{n} [(E(s))_j]^2 I(n^{-1/2} |(E(s))_j| > \frac{\varepsilon}{2}) \exp(\beta_0^t Z_h(s)) \lambda_0(s) ds \right. \quad (3.36)$$

$$+ \int_0^1 \sum_{h=1}^n \frac{1}{n} [(Z_h(s))_j]^2 I(n^{-1/2} |(Z_h(s))_j| > \frac{\varepsilon}{2}, \\ \beta_0^t Z_l(s) \leq -\delta |(Z_h(s))_j|) \exp(\beta_0^t Z_h(s)) \lambda_0(s) ds \quad (3.37)$$

$$+ \int_0^1 \sum_{h=1}^n \frac{1}{n} [(Z_h(s))_j]^2 I(n^{-1/2} |(Z_h(s))_j| > \frac{\varepsilon}{2}, \\ \beta_0^t Z_h(s) > -\delta |(Z_h(s))_j|) \exp(\beta_0^t Z_h(s)) \lambda_0(s) ds \left. \right\}. \quad (3.38)$$

Die Konvergenz von (3.36) sieht man sofort, denn mit den Bedingungen A, B und D gilt

$$\int_0^1 \sum_{h=1}^n \frac{1}{n} [(E(\beta_0, s))_j]^2 I(n^{-1/2} |(E(\beta_0, s))_j| > \frac{\varepsilon}{2}) \lambda_h(s) ds \\ = \int_0^1 \underbrace{[(E(\beta_0, s))_j]^2}_{\rightarrow (e(\beta_0, s))_j^2 < \infty} I(\underbrace{n^{-1/2} |(E(\beta_0, s))_j| > \frac{\varepsilon}{2}}_{\rightarrow 0 \text{ für } n \rightarrow \infty \text{ da } e(\beta_0, s) \text{ beschränkt}}) \\ \underbrace{\sum_{h=1}^n n^{-1} Y_h(s) \exp(\beta_0^t Z_h(s)) \lambda_0(s) ds}_{\rightarrow s^{(0)}(\beta, s)} \rightarrow_{\mathcal{P}} 0 \text{ für } n \rightarrow \infty.$$

Nach Bedingung C gilt für $I\{n^{-1/2} |(Z_h(s))_j| > \frac{\varepsilon}{2}, \beta_0^t Z_h(s) \leq -\delta |(Z_h(s))_j|\}$ in (3.38), daß

$$P[n^{-1/2} |(Z_h(s))_j| > \frac{\varepsilon}{2}, \beta_0^t Z_l(s) \leq -\delta |(Z_h(s))_j|] \rightarrow_{\mathcal{P}} 0 \text{ für } n \rightarrow \infty,$$

sowie nach Bedingungen B und D daß

$$\frac{1}{n} \sum_{l=1}^{\infty} [(Z_h(s))_j]^2 Y_l(s) \exp(\beta_0^t Z_h(s)) = S_j^{(2)}(\beta_0, s) \rightarrow_{\mathcal{P}} s_j^{(2)}(\beta_0, s) < \infty \text{ für } n \rightarrow \infty,$$

so daß zusammen mit $\int_0^1 \lambda(s) ds < \infty$ nach Bedingung A auch Term (3.38) gegen 0 konvergiert.

Schließlich können wir Term (3.37) abschätzen mit

$$\int_0^1 \sum_{h=1}^n \frac{1}{n} [(Z_h(s))_j]^2 I(n^{-1/2} |(Z_h(s))_j| > \frac{\varepsilon}{2}, \beta_0^t Z_h(s) \leq -\delta |(Z_h(s))_j|) \exp(\beta_0^t Z_h(s)) \lambda_0(s) ds \\ \leq \int_0^1 \frac{1}{n} \sum_{h=1}^n \underbrace{[(Z_h(s))_j]^2 \exp(-\delta |(Z_h(s))_j|)}_{< \infty} \lambda_0(s) \underbrace{I(n^{-1/2} |(Z_h(s))_j| > \frac{\varepsilon}{2})}_{\rightarrow 0 \text{ für } \rightarrow \infty} ds.$$

Da $\lim_{x \rightarrow \infty} x^2 e^{-x} = 0$ folgt, daß $[(Z_h(s))_j]^2 \exp(-\delta |Z_h|)$ beschränkt durch einen Wert η , so daß auch dieser Term in Wahrscheinlichkeit gegen 0 konvergiert für $n \rightarrow \infty$. Damit ist der erste Teil bewiesen.

Für den zweiten Teil des Beweises $\frac{I(\beta^*, 1)}{n} \rightarrow_{\mathcal{P}} \Sigma$ für irgendein zufälliges β^* mit bilden wir mit

Hilfe von

$$\frac{1}{n}I(\beta^*, 1) = \frac{1}{n} \int_0^1 V(\tilde{\beta}, t) d\bar{N}(t) \quad (\text{vergl. (3.31)})$$

und

$$\Sigma = \int_0^1 v(\beta_0, t) s^{(0)}(\beta_0, t) \lambda_0(t) dt$$

folgende Abschätzung

$$\begin{aligned} \left\| \frac{1}{n}I(\beta^*, 1) - \Sigma \right\| &\leq \left\| \int_0^1 \frac{1}{n} [V(\beta^*, t) - v(\beta^*, t)] d\bar{N}(t) \right\| \\ &+ \left\| \int_0^1 \frac{1}{n} [v(\beta^*, t) - v(\beta_0, t)] d\bar{N}(t) \right\| + \left\| \int_0^1 \frac{1}{n} v(\beta_0, t) (d\bar{N}(t) - \bar{\lambda}(t) dt) \right\| \\ &+ \left\| \int_0^1 v(\beta_0, t) [S^{(0)}(\beta_0, t) - s^{(0)}(\beta_0, t)] \lambda_0(t) dt \right\|. \end{aligned} \quad (3.39)$$

Zunächst zeigen wir, daß

$$\lim_{\eta \rightarrow \infty} \lim_{n \rightarrow \infty} P \left[\frac{\bar{N}(1)}{n} > \eta \right] = 0. \quad (3.40)$$

Da $\bar{N}(t)$ den vorhersehbaren Kompensatorprozeß

$$\int_0^t \sum_{k=1}^n Y_i(s) \lambda_0(s) \exp(\beta_0^t Z_i(s)) ds = n \int_0^t S(\beta_0, t) \lambda_0(s) ds$$

hat, gilt nach Satz 2.37 (Ungleichung von Lengart)

$$P \left[\sup_{t \in [0,1]} \frac{\bar{N}(t)}{n} > \eta \right] = P \left[\frac{\bar{N}(1)}{n} > \eta \right] \leq \frac{\delta}{\eta} + P \left[\int_0^1 \underbrace{S^{(0)}(\beta_0, t)}_{\rightarrow s^{(0)}(\beta_0, t)} \lambda_0(t) dt > \eta \right].$$

Mit $\eta \rightarrow \infty$ und $\delta = \sqrt{\eta}$, sowie $\int_0^1 \lambda_0(t) < \infty$ erhält man schließlich (3.40). Aus der Bedingung B und der Beschränktheit von $s^{(0)}$, $s^{(1)}$ und $s^{(2)}$ in einer Umgebung \mathcal{B} von β_0 , sowie aus $s^{(0)} \neq 0$ in \mathcal{B} folgt, daß

$$\sup_{t \in [0,1], \beta \in \mathcal{B}} \|V(\beta, t) - v(\beta, t)\| \rightarrow_P 0.$$

Da $\beta^* \rightarrow_P \beta_0$, konvergiert der erste Term der Abschätzung (3.39) in Wahrscheinlichkeit gegen 0. Den dritten Term aus (3.39) können wir direkt mit der Ungleichung (2.20) abschätzen

$$P \left[\left| \int_0^1 v_{ij}(\beta_0, t) d\frac{\bar{M}}{n}(t) \right| > \delta \right] \leq \eta/\delta^2 + P \left[\frac{1}{n} \int_0^1 \underbrace{[v_{ij}(\beta_0, t)]^2 S^{(0)}(\beta_0, t) \lambda_0(t)}_{< \infty \text{ nach Bed. A, B, D}} dt > \eta \right]. \quad (3.41)$$

Die Bedingung B, sowie $\int_0^1 \lambda_0(t) dt < \infty$ (aus Bedingung A), sowie die Beschränkungsbedingungen aus D haben zur Folge, daß der rechte Term aus (3.41) gegen 0 konvergiert für $n \rightarrow \infty$. Damit ist gezeigt, daß auch der dritte Summand aus (3.39) gegen 0 läuft. Es bleibt noch zu zeigen, daß der vierte Term gegen 0 konvergiert. Dies ist jedoch ein direktes Resultat aus Bedingungen A, B und D (v ist beschränkt, $\int_0^1 \lambda_0(t) dt$ ist beschränkt und $\|S^{(0)}(\beta_0, t) - s^{(0)}(\beta_0, t)\| \rightarrow_P 0$). Damit ist die asymptotische Normalität von $\hat{\beta}$ bewiesen. \square

3.7.5 Hypothesentests im Proportional-Hazard-Modell

Für den Test, ob der Kovariablenvektor $\beta \in \mathbb{R}^p$ einen bestimmten Wert $\beta_0 \in \mathbb{R}^p$ annimmt, kann man verschiedene Statistiken anführen. Im folgenden sollen 3 verschiedene Teststatistiken für Hypothesentests für die Nullhypothese:

$$H_0 : \beta = \beta_0$$

gegen

$$H_1 : \beta \neq \beta_0$$

vorgestellt werden. Zum Beispiel ist man oft interessiert an der Frage, ob ein Proportional-Hazard-Modell vorliegt, die Nullhypothese wäre in diesem Fall $H_0 : \beta_0 = 0$. Eine ausführlichere Darstellung dieser Thematik findet man in Fahrmeier (1994) [15] Seiten 45 bis 48.

Likelihood-Ratio-Test

Der Likelihood-Ratio-Teststatistik ist im wesentlichen ein Maß dafür, wie sich für den partiellen Likelihood der Quotient $L(\hat{\beta})/L(\beta_0)$ verhält. In diesem Fall lehnt man die Nullhypothese für große Werte ab, d.h. wenn der Zähler wesentlich größer ist als der Nenner. Die Likelihood-Ratio-Statistik betrachtet die logarithmische Transformation dieses Quotienten:

$$\chi_{LR}^2 = 2 \left[\log \frac{L(\hat{\beta})}{L(\beta_0)} \right] = 2 [LL(\hat{\beta}) - LL(\beta_0)]. \quad (3.42)$$

Man kann zeigen, daß (3.42) asymptotisch χ^2 verteilt ist mit p Freiheitsgraden.

Wald-Test

Der Wald-Test basiert auf der asymptotischen Normalverteilung des Parameters $\hat{\beta}$. Für große n gilt, daß:

$$\hat{\beta} \sim N_p(\beta, \mathbf{i}^{-1}(\beta)).$$

$\mathbf{i}(\beta)$ ist die Fisher-Informations-Matrix:

$$\mathbf{i}(\beta) = E_{\beta} \left[-\frac{\partial}{\partial \beta} U(\beta) \right] = E_{\beta} \left[\begin{pmatrix} \frac{\partial^2}{\partial \beta_1 \partial \beta_1} LL(\beta) & \cdots & \frac{\partial^2}{\partial \beta_1 \partial \beta_p} LL(\beta) \\ \vdots & \ddots & \vdots \\ \frac{\partial^2}{\partial \beta_p \partial \beta_1} LL(\beta) & \cdots & \frac{\partial^2}{\partial \beta_p \partial \beta_p} LL(\beta) \end{pmatrix} \right].$$

Da die Informationsmatrix $\mathbf{i}(\beta)$ oft schwierig zu berechnen ist, benutzt man die beobachtete Informationsmatrix $I(\hat{\beta})$ als Schätzung. Damit kann man zeigen, daß die quadratische Form:

$$\chi_W^2 = (\hat{\beta} - \beta_0)^t I(\hat{\beta}) (\hat{\beta} - \beta_0) \quad (3.43)$$

asymptotisch χ^2 verteilt mit p Freiheitsgraden. (3.43) heißt Wald-Test.

Score-Test

Der Score- oder Rao-Test basiert auf der effizienten Score-Statistik. Für den Fall, daß $\beta = \beta_0$ ist der Score-Vektor $U(\beta_0)$ asymptotisch verteilt nach $N_p(0, \mathbf{i}(\beta_0))$. Mit der Näherung I^{-1} für \mathbf{i}^{-1} erhält man den Score-Test

$$\chi_S^2 = U(\beta_0)^t I^{-1}(\beta_0) U(\beta_0). \quad (3.44)$$

Der Score-Test ist auch asymptotisch χ^2 -verteilt mit n Freiheitsgraden.

In Rao (1965) [32] finden sich in Kapitel 6 die Beweise der asymptotischen Normalität der Teststatistiken (3.42) - (3.44). In Abbildung 3.4 ist der Zusammenhang zwischen Likelihood-

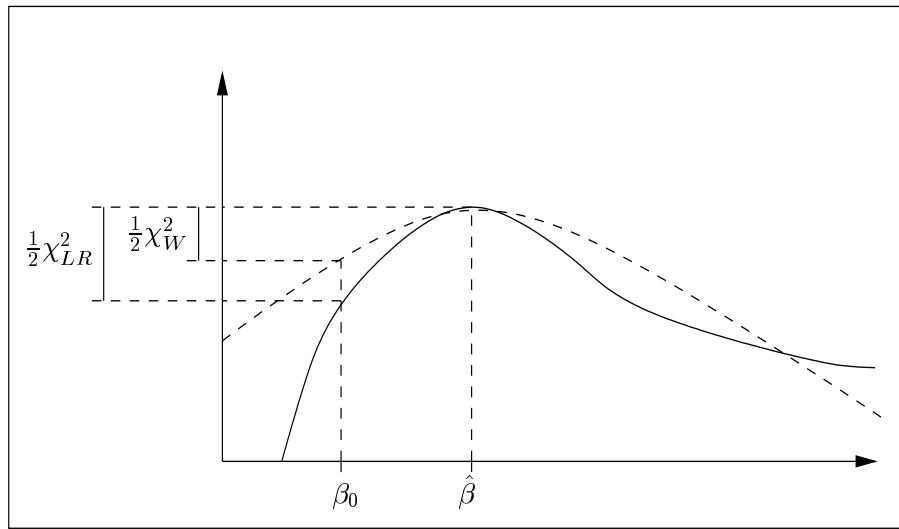


Abbildung 3.4: Zusammenhang zwischen Likelihood-Ratio- und Wald-Test (— Likelihood-Funktion, - - - quadratische Approximation)

Ratio und Wald-Statistik erläutert. Um die Wald-Statistik zu erhalten wird die Likelihood-Funktion an der Stelle $\hat{\beta}$ mit Hilfe der Matrix der zweiten Ableitungen quadratisch approximiert. Im eindimensionalen Fall erhält man als Approximation eine Parabel, die im Bild mit einer gestrichelten Linie dargestellt ist.

3.7.6 Lokale Tests

Oft ist man nicht nur an Hypothesentests für alle β_i , $i = 1, \dots, p$ sondern an Teilmengen (man ist zum Beispiel oft interessiert daran, ob eine einzelne Kovariable Einfluß auf den Hazard hat). Das heißt: Wir unterteilen den Vektor β in $\beta^t = (\beta_1^t, \beta_2^t)$ mit $\beta_1 \in \mathbb{R}^q$ und $\beta_2 \in \mathbb{R}^{p-q}$. Die Nullhypothese ist in diesem Fall:

$$H_0 : \beta_1 = \beta_{10},$$

womit β_2 nur noch als Störparameter in das Modell eingeht. Die im Kapitel 3.7.5 beschriebenen Tests lassen sich wie folgt für lokale Tests modifizieren:

Bestimme $\hat{\beta}_2(\beta_{10})$, so daß der partielle Log-Likelihood in Abhängigkeit von β_{10}

$$LL(\beta_{10}, \beta_2)$$

für $\hat{\beta}_2(\beta_{10})$ maximal wird. Zerlege die Informationsmatrix und deren Inverse in

$$I = \begin{pmatrix} I_{\beta_1\beta_1} & I_{\beta_1\beta_2} \\ I_{\beta_2\beta_1} & I_{\beta_2\beta_2} \end{pmatrix} \text{ und } I^{-1} = \begin{pmatrix} \tilde{I}_{\beta_1\beta_1} & \tilde{I}_{\beta_1\beta_2} \\ \tilde{I}_{\beta_2\beta_1} & \tilde{I}_{\beta_2\beta_2} \end{pmatrix}, \quad (3.45)$$

mit $I_{\beta_1\beta_1} \in \mathbb{R}^{q \times q}$, $I_{\beta_2\beta_2} \in \mathbb{R}^{(p-q) \times (p-q)}$. Die Teststatistiken aus 3.7.5 für die Nullhypothese $H_0 : \beta_1 = \beta_{10}$ sind dann:

Likelihood-Ratio:

$$\chi_{LR}^2(\beta_{10}) = 2(LL(\hat{\beta}) - LL(\beta_{10}, \hat{\beta}_2(\beta_{10}))). \quad (3.46)$$

Wald-Test:

$$\chi_W^2(\beta_{10}) = (\hat{\beta}_1 - \beta_{10})^t [\tilde{I}_{\beta_1\beta_1}(\hat{\beta})]^{-1} (\hat{\beta}_1 - \beta_{10}). \quad (3.47)$$

Score-Test:

$$\chi_S^2(\beta_{10}) = U(\beta_{10}, \hat{\beta}_2(\beta_{10}))^t \tilde{I}_{\beta_1\beta_1}(\beta_{10}, \hat{\beta}_2(\beta_{10})) U(\beta_{10}, \hat{\beta}_2(\beta_{10})). \quad (3.48)$$

Sämtliche Teststatistiken sind asymptotisch χ^2 -verteilt mit q Freiheitsgraden. In der Anwendung spielt vor allem der Wald-Test eine Rolle, da für diese Teststatistik kein neues Proportional Hazard Modell gefittet werden muß.

Die z-Statistik

Mit Hilfe der **z-Statistik** kann man die Proportional-Hazard-Annahme für einen einzelnen Parameter testen. Die Nullhypothese lautet hier für einen Koeffizienten β_k $k \in 1, \dots, p$:

$$H_0 : \beta_k = 0.$$

Diese Teststatistik ist für β_k wie folgt definiert

$$z = \frac{\hat{\beta}_k}{\sqrt{\hat{\sigma}_k^2}},$$

wobei $\hat{\sigma}_k^2$ das k -te Diagonalelement der Matrix I ist. Wie man leicht sieht, ist

$$z^2 = \chi_W^2(\beta_{k0}) \quad \beta_{k0} = 0$$

ein Spezialfall des Wald-Tests für die Nullhypothese $H_0 : \beta_k = 0$ und somit χ^2 verteilt mit einem Freiheitsgrad.

3.7.7 Modellbildung mit dem AIC

Nachdem im vorherigen Abschnitt Verfahren zum Testen von Hypothesen vorgestellt wurden, soll hier nun eine Methode gezeigt werden, wie man die Kovariablen im Proportional-Hazard modelliert und wie man entscheidet, ob eine Kovariable starken Einfluß auf den betrachteten Überlebensprozeß hat oder nicht.

Eine Möglichkeit zur Examinierung, welche Kovariablen wesentlichen Einfluß im Modell haben ist das sogenannte AIC (Akaike's Information Criterion)

$$\text{AIC} = -2LL(\hat{\beta}) + np.$$

Hier ist $LL(\hat{\beta})$ der Partial Likelihood für den geschätzten Koeffizienten $\hat{\beta}$, p die Anzahl der Freiheitsgrade im Modell und n eine Konstante, auf die später noch näher eingegangen wird.

Der Grundgedanke ist nun Folgender: Falls eine Kovariable wesentlichen Einfluß auf das Modell hat, wird sich dies im Likelihood ausdrücken (der Likelihood wird grösser, da mit der neuen Kovariablen die Wahrscheinlichkeit dafür, daß das gewählte Modell der Realität entspricht, wächst). Da der Log-Likelihood stets negative Werte annimmt bedeutet das jedoch, daß der AIC-Wert für einen großen Likelihood sinkt. Im Gegensatz dazu nimmt der Summand $2np$ zu, für jede Kovariable die neu hinzugefügt wird. Das bedeutet, daß der zweite Term eine Art Bestrafung für jede neu hinzugefügte Kovariable ist. Das Ziel ist nun die Minimierung des AIC um das optimale Modell zu erhalten. Die Konstante n ist demzufolge dafür verantwortlich, wie stark die neue Hinzunahme von Kovariablen bestraft werden soll. Ein niedriges n bestraft die Hinzunahme von neuen Variablen weniger, so daß das endgültige Modell in diesem Fall mehr Kovariablen beinhalten wird als ein Modell, das mit hohem n gewählt wurde.

3.7.8 Residuenanalyse

Mit Hilfe von Residuen kann die Proportional-Hazard-Annahme überprüft werden, bzw. die funktionale Form einer Kovariablen bestimmt werden. Die hier vorgestellten Residuen werden mit Hilfe der Zählprozeßtheorie hergeleitet. Eine Herleitung verschiedener Residuen auf Martingal-Basis kann man in dem Artikel von Therneau u.a. (1990) [39] nachlesen.

Martingal-Residuen

Sei $N_i(t)$ $i = 1, \dots, n$ der Zählprozeß der beobachteten Ereignisse für Person i (d.h. $N_i(t) = 1$ falls Person i gestorben, $N_i(t) = 0$ sonst) und $\hat{\Lambda}_0(t)$ der mit dem Schätzer von Breslow bestimmte kumulative Basishazard, sowie $Y_i(t)$ die Indikatorfunktion:

$$Y_i(t) = \begin{cases} 1 & \text{Person } i \text{ wird zum Zeitpunkt } t \text{ beobachtet} \\ 0 & \text{sonst} \end{cases}.$$

Dann sind die **Martingal-Residuen** $\hat{M}_i(t)$ wie folgt definiert:

$$\hat{M}_i(t) := N_i(t) - \int_0^t Y_i(s) \exp(\hat{\beta}^t Z_i(s)) d\hat{\Lambda}_0(s) \text{ und } \hat{M}_i = \hat{M}_i(\infty). \quad (3.49)$$

Bemerkungen:

- Die Martingal-Residuen lassen sich als Differenz aus beobachteten Toten und den nach dem Proportional Hazard-Modell erwarteten Toten interpretieren.
- Der Name der Martingal-Residuen basiert auf der Tatsache, daß der Prozeß $N(t) - \Lambda(t)$ ein Martingal ist. Für den Fall, daß das Proportional-Hazard-Modell gilt, ist

$$\Lambda(t) = \int_0^t \exp(\beta^t Z_i(s)) d\Lambda_0(s),$$

d.h. genau dann ist der Prozeß:

$$M_i(t) := N_i(t) - \int_0^t Y_i(s) \exp(\beta^t Z_i(s)) d\Lambda_0(s)$$

ein Martingal.

Durch Plotten der Martingalresiduen gegen eine qualitative Kovariable läßt sich die Proportional-Hazard-Annahme überprüfen. Falls nämlich die Proportional-Hazard-Annahme gilt, ist für jeden Wert der Kovariablen der Erwartungswert $E(\hat{M}|\mathcal{F}_0) = 0$, d.h. falls man eine Schätzung des Erwartungswertes $E(\hat{M}|\mathcal{F}_0)$ (z.B. durch Glättung) durch den Plot legt, sollte diese ungefähr gleich 0 sein.

Ein weiterer wichtiger Anwendungsbereich von Martingal-Residuen ist die Bestimmung der funktionalen Form einer Kovariablen. Für den Kovariablenvektor $Z = (Z_1, Z_{-1}) \in \mathbb{R}^p$ (mit $Z_{-1} = (Z_2, \dots, Z_p)$) ist man daran interessiert, ob die (quantitative) Komponente Z_1 in einer bestimmten funktionalen Form (z.B. $Z_1^2, \log(Z_1)$) in das Proportional-Hazard-Modell eingeht. D.h. man möchte die Hazardfunktion folgendermaßen darstellen

$$\lambda(t|Z) = h(Z_1) \exp(\beta^t Z_{-1}) \lambda_0(t) = \exp(f(Z_1) + \beta^t Z_{-1}) \lambda_0(t),$$

wobei hier $h(Z_1)$ die gesuchte funktionale Form ist und

$$f(Z_1) = \log(h(Z_1)).$$

Therneau, Grambsch und Fleming (1990) [39] zeigen für \bar{h} , einer Funktion, die mit Hilfe gewichteter Mittel über Zeit und erwartete Zusammensetzung der Risikomenge definiert wird (siehe [39]) und für unabhängige Z_1, Z_{-1} ungefähr konstant ist, daß

$$E[\hat{M}(t)|Z_1] \approx \left(1 - \frac{\bar{h}}{h(Z_1)}\right) E[N(t)|Z_1]$$

und man für $t \rightarrow \infty$ folgende Approximationen annehmen kann

$$E(\hat{M}|Z_1) \approx c(h(Z_1) - \bar{h})$$

oder

$$E(\hat{M}|Z_1) \approx c(f(Z_1) - \bar{f})$$

mit $\bar{f} = \log(\bar{h})$. Durch den Plot der Kovariablen gegen die geglätteten Martingal Residuen erhält man die funktionale Darstellung der Kovariablen.

Devianz Residuen

Eine Schwäche der Martingal-Residuen ist die weite Streuung (die Residuen können Werte zwischen $[1, -\infty)$ annehmen). Daher sind sie nicht dazu geeignet einzelne Beobachtungen, die nicht in das Modell passen, herauszufiltern. Um solche Ausreißer festzustellen, skaliert man die Martingal-Residuen folgendermaßen:

$$\hat{d}_i = \text{sign}(\hat{M}_i)[-2\{\hat{M}_i + N_i(\infty) \log(N_i(\infty) - \hat{M}_i)\}]^{1/2}. \quad (3.50)$$

Die Devianz-Residuen stammen aus der Theorie der generalisierten linearen Modelle (siehe z.B. McCullagh und Nelder (1983) [28]). Die Devianz eines Modells ist hiernach definiert als $2[LL(\text{gesättigtes Modell}) - LL(\text{berechnetes Modell})]$. Hier ist das “gesättigte Modell” ein Modell, in den für jede Beobachtung ein einzelner Parameter geschätzt wird. In Therneau u.a. (1990) [39] ist die Herleitung dieser Residuen ausführlich beschrieben.

Für Überlebenszeitmodelle (d.h. für Zählprozesse, in denen maximal ein Sprung pro Beobachtung möglich ist), erhält man für die Devianz-Residuen

$$\hat{d}_i = \text{sign}(\hat{M}_i)[-2\{\hat{M}_i + \delta_i \log(\delta_i - \hat{M}_i)\}]^{1/2}.$$

Hier ist δ_i wie üblich der Zensierungsindikator ist. Werfen wir nun noch einmal einen Blick auf die Skalierung: Für positive Martingal Residuen gilt $\delta_i = 1$. Für Martingal Residuen, die in der Nähe von 1 liegen, bewirkt der Logarithmus

$$\log(\underbrace{\delta_i - \hat{M}_i}_{\text{nahe 0}})$$

eine weitere Streuung, zum anderen sorgt die Wurzel dafür, daß stark negative Martingal-Residuen “gestaucht” werden.

Die Devianz-Residuen lassen sich nun folgendermaßen für die Herausfilterung von schlecht vorhergesagten Daten einsetzen. Man plottet die Devianz-Residuen gegen den Wert einer (quantitativen) Kovariablen. Falls alle Beobachtungen durch das Modell erklärbar sind, sollte der Plot aussehen, wie die zufällige Realisation einer $N(0, 1)$ verteilten Zufallsvariable.

Partielle- oder Schoenfeld Residuen

Oft kann man in einem Proportional-Hazard-Modell erkennen, daß der Regressionskoeffizient β nicht für alle t konstant ist, sondern sich mit zunehmender Lebensdauer verändert. Im Regressions-Modell bedeutet das, daß wir für die Hazardfunktion einen zeitabhängigen Regressionskoeffizienten erhalten

$$\lambda(t|Z(t)) = \lambda_0(t) \exp(\beta(t)Z(t)),$$

wobei

$$\beta : \mathcal{T} \mapsto \mathbb{R}$$

nun eine reelwertige Funktion ist. Mit Hilfe von partiellen Residuen (auch Schoenfeld-Residuen genannt) läßt sich zeitliche Variation im Koeffizienten β feststellen. Schoenfeld (1982) [37] schlug diese Residuen vor, die folgendermaßen definiert sind:

Betrachten wir eine nach Endzeitpunkten geordnete Beobachtung der Größe n mit $t_1 \leq \dots \leq t_n$ und zugehörigen Zensierungsindikatoren δ_i $i \in \{1, \dots, n\}$, sowie

$$\hat{r}_{ij} = \delta_i \left[Z_{ij}(t_i) - \frac{\sum_{k \in R(t_i)} \exp(\hat{\beta}^t Z_k(t_i)) Z_{kj}(t_i)}{\sum_{k \in R(t_i)} \exp(\hat{\beta}^t Z_k(t_i))} \right] \quad i \in \{1, \dots, n\}, \quad j \in \{1, \dots, p\}. \quad (3.51)$$

Dann ist für alle Beobachtungen i , an denen ein Ereignis passiert ($\delta_i = 1$), das partielle Residuum für die Komponente j durch \hat{r}_{ij} gegeben. Mit folgendem Vorgehen können wir nun Aussagen zur funktionalen Form treffen. Wiederum schätzen wir mit einem Glättungsschätzer den Erwartungswert der partiellen Residuen (pro Komponente j) in Abhängigkeit von der Zeit, d.h. wir glätten $\{t_i, \hat{r}_{ij}\}$, $i = 1, \dots, n$, $j = 1, \dots, p$. Durch einen Plot der geglätteten partiellen Residuen gegen die Zeit erhalten wir einen Indikator für die zeitliche Abhängigkeit des jeweiligen Koeffizienten (pro Komponente).

Bemerkung: Die partiellen Residuen bilden also für jede Beobachtung, an der ein Ereignis passiert, die Differenz aus der beobachteten Komponente j des Kovariablenvektor und einer Art gewichtetem Mittel über die j -te Komponente aller Kovariablen von Individuen, die sich zu diesem Zeitpunkt unter Risiko befinden.

Score Residuen

Die sogenannten **Score Residuen** werden verwendet, um den Einfluß einer einzelnen Beobachtung i auf den geschätzten Koeffizienten $\hat{\beta}$ zu ermitteln. Eine optimale Methode dies zu erreichen, wäre die Beobachtung i aus dem Datensatz zu entfernen und für die verbleibenden Beobachtungen den Koeffizienten $\hat{\beta}_{(i)}$ mit einem Cox-Regressionsmodell zu schätzen. Falls die Beobachtung nur geringen Einfluß hat, ist

$$\hat{\beta} - \hat{\beta}_{(i)} \approx 0.$$

Um dies zu realisieren müssen also für ein Modell mit n Beobachtungen insgesamt $n + 1$ Proportional-Hazard-Modelle berechnet werden, ein Vorgehen, das für große Werte von n nicht mehr praktikierbar ist.

Als Näherung von $\hat{\beta} - \hat{\beta}_{(i)}$ (siehe z.B. Klein und Moeschberger (1997) [25]) benutzt man hier die Score-Residuen \hat{S}_{ij} , die für die Beobachtung $i \in \{1, \dots, n\}$ und die Komponente $j \in \{1, \dots, p\}$ des Kovariablenvektors wie folgt definiert werden

$$\hat{S}_{ij}(t) = \int_0^t [Z_{ij}(s) - \bar{Z}_j(s)] d\hat{M}_i(s),$$

mit

$\hat{M}_j(s)$ = Martingalresiduum für Beobachtung j zum Zeitpunkt s

und

$$\bar{Z}_j(s) = \frac{\sum_{k=1}^n Y_k(s) \exp(\hat{\beta}^t Z_k(s)) Z_{kj}(s)}{\sum_{k=1}^n Y_k(s) \exp(\hat{\beta}^t Z_j(s))} \in \mathbb{R}.$$

$Y_j(s)$ ist hier wiederum der Indikator, ob sich ein Objekt zum Zeitpunkt s unter Beobachtung steht, oder nicht. Man erkennt, daß zu jedem Zeitpunkt, an dem ein Ereignis stattfindet, der Integrator von $\hat{S}_{ij}(t)$ den Wert des partiellen Residuums \hat{r}_{ij} annimmt. Das Score-Residuum \hat{S}_{ij} ist schließlich definiert als

$$\hat{S}_{ij} = S_{ij}(\infty) = \int_0^\infty [Z_{ij}(s) - \bar{Z}_j(s)] d\hat{M}_j(s).$$

Hiermit kann man nun direkt den Einfluß einer Beobachtung i auf die Komponente j des Kovariablenvektor feststellen. Je näher \hat{S}_{ij} an 0, desto geringer ist der Einfluß der Beobachtung i auf die Komponente j .

3.7.9 Graphische Methoden zur Überprüfung des Proportional-Hazard

In diesem Teil werden einige Techniken zur graphischen Überprüfung der Proportional-Hazard-Annahme vorgestellt. Dabei ist man zunächst daran interessiert, ob man für eine bestimmte Komponente des Kovariablenvektors von einem Proportional-Hazard-Modell ausgehen kann. Sei nun ohne Einschränkung Z_1 die interessierende Komponente des Kovariablenvektors $Z = (Z_1, Z_2^t)^t$, wobei hier Z_2 der Vektor der verbleibenden $p - 1$ Komponenten ist. Außerdem geht man davon aus, daß die Kovariable Z_1 nur endlich viele Werte annimmt (die Werte seien ohne Einschränkung $\{0, 1, \dots, K\}$).

Um nun die Proportional-Hazard-Annahme zu testen berechnet man ein nach der Kovariablen Z_1 stratifiziertes Modell und schätzt mit Hilfe des Breslow-Schätzers den zugrundeliegenden kumulativen Basis-Hazard für jedes Stratum $\Lambda_{0i}(t)$; $i = 1, \dots, K$. Falls ein Proportional-Hazard-Modell vorliegt, gilt für jedes $i \neq j$:

$$\begin{aligned} \frac{\Lambda_{0i}(t)}{\exp(i \beta_1)} &= \frac{\Lambda_{0j}(t)}{\exp(j \beta_1)} \\ \iff \frac{\Lambda_{0i}(t)}{\Lambda_{0j}(t)} &= \exp((i - j)\beta_1), \end{aligned}$$

das bedeutet, daß

$$\log(\Lambda_{0i}(t)) - \log(\Lambda_{0j}(t)) = \beta_1(i - j) = \text{konstant über } t$$

Um nun zu überprüfen, ob Proportional-Hazard vorliegt plottet man t gegen $\log(\hat{\Lambda}_{0i}(t)) - \log(\hat{\Lambda}_{0j}(t))$. Dieser Graph sollte für den Fall, daß Proportional Hazard gilt ungefähr eine Parallele zur Zeitachse darstellen. Für binäre Kovariablen, wie z.B. Geschlecht, wurden die Plots in

der S-Plus-Routine `cumloghazard` (siehe Anhang A.1) realisiert.

Ein etwas modifizierter Ansatz sind die sogenannten “**Andersen-Plots**” (1982) Mit den Bezeichnungen wie vorher plottet man hier zu jedem Zeitpunkt $t_1 < \dots < t_k$, an dem ein Ereignis passiert, die Hazardraten $\Lambda_{0i}(t)$ (x -Wert) und $\Lambda_{0j}(t)$ (y -Wert) gegeneinander. Falls man von Proportional Hazard ausgehen kann, sollte der Plot ungefähr einer Gerade durch den Ursprung mit Steigung

$$\frac{\exp(\beta_1 j)\lambda_0(t)}{\exp(\beta_1 i)\lambda_0(t)} = \frac{\exp(\beta_1 j)}{\exp(\beta_1 i)}$$

entsprechen. Für binäre Kovariablen wurde dieses graphische Hilfsmittel in der Routine `andplot` (siehe Anhang) programmiert.

3.8 Parametrische Modelle

Am Beispiel Weibull-verteilter Zufallsvariablen, wollen wir hier auf Schätzmethoden in parametrischen Modellen eingehen. Eine ausführliche Beschreibung parametrischer Schätzmethoden enthält das Buch von Kalbfleisch und Prentice [21].

Die Zufallsvariable Überlebenszeit sei im folgenden nach dem Weibull-Modell verteilt, d.h.

$$S_T(t) = \exp(-\lambda t^\alpha),$$

mit $\alpha, \lambda > 0$ und zugehöriger Hazardfunktion

$$\lambda_T(t) = \lambda \alpha t^{\alpha-1}$$

und Dichte

$$f_T(t) = \alpha \lambda t^{\alpha-1} \exp(-\lambda t^\alpha).$$

Um nun Schätzer für die Parameter λ und α zu erhalten, transformieren wir das Modell folgendermaßen:

$$Y = \log T$$

und

$$W = \frac{Y + \mu}{\sigma},$$

mit $\sigma = 1/\alpha$ und $\lambda = \exp(\mu/\sigma)$.

Die Zufallsvariable W ist verteilt nach

$$F_W(w) = 1 - \exp(-e^w), \tag{3.52}$$

mit Dichte

$$f_W(w) = \exp(w - e^w), \tag{3.53}$$

denn

$$\begin{aligned}
 F_W(w) &= P(W \leq w) = P\left(\frac{Y - \mu}{\sigma} \leq w\right) = P(T \leq e^{\mu + \sigma w}) \\
 &= 1 - S_T(e^{\mu + \sigma w}) = 1 - \exp\left(-\lambda e^{\alpha(\mu + \sigma w)}\right) \\
 &= 1 - \exp\left(-\exp\left(-\frac{\mu}{\sigma}\right) \exp\left(\frac{\mu + \sigma w}{\sigma}\right)\right) \\
 &= 1 - \exp(-e^w).
 \end{aligned} \tag{3.54}$$

Y beschreibt demnach ein log-lineares Modell

$$Y = \log T = \mu + \sigma W.$$

Die Überlebensfunktion von Y erhält man mit Hilfe von (3.52)

$$\begin{aligned}
 S_Y(y) &= 1 - F_Y(y) = 1 - P(Y \leq y) \\
 &= 1 - P\left(W \leq \frac{y - \mu}{\sigma}\right) = \exp\left(-\exp\left(\frac{y - \mu}{\sigma}\right)\right).
 \end{aligned}$$

Die Dichtefunktion von Y ist demnach

$$f_Y(y) = \frac{1}{\sigma} \exp\left(\frac{y - \mu}{\sigma} - \exp\left(\frac{y - \mu}{\sigma}\right)\right).$$

Seien nun wiederum t_1, \dots, t_n die Beobachtungszeitpunkte und

$$\delta_j = \begin{cases} 1 & \text{falls } T_j = t_j \\ 0 & \text{sonst} \end{cases}$$

der zugehörige Zensierungsindikator, dann definieren wir für $y = \log(t_i)$ die folgende Likelihood-Funktion

$$L(\mu, \sigma) = \prod_{i=1}^n [f_Y(y_i)]^{\delta_i} [S_Y(y_i)]^{(1-\delta_i)},$$

die mit Hilfe numerischer Methoden maximiert werden kann (siehe Klein und Moeschberger [25] Seiten 375-377).

Kapitel 4

Datenanalyse zur gesetzlichen Pflegeversicherung

In diesem Kapitel werden mit Hilfe der vorher beschriebenen verweildaueranalytischen Methoden Überlebensdaten für Leistungsempfänger aus der gesetzlichen Pflegeversicherung analysiert. Alle Berechnungen und graphischen Analysen wurden mit dem Programm S-Plus realisiert (eine detaillierte Einführung in S-Plus geben Venables und Ripley (1994) [40]). Insbesondere fanden für Schätzungen des Proportional Hazard die S-Plus-Funktionen `coxph` und `coxph.detail` Anwendung, für Schätzungen der Überlebensfunktion wurden die Funktion `survfit` verwendet, zur Residuen-Analyse besitzt die S-Plus-Funktion `residuals` eigene Standardmethode für `coxph`-Objekte. Für einige graphische Analysen mußten noch eigene S-Plus-Routinen entwickelt werden. Mit diesem Paket können in Kapitel 3 vorgestellten Techniken in S-Plus angewendet werden.

4.1 Überblick über die gesetzliche Pflegeversicherung

Am 26.Mai 1994 beschloß der deutsche Bundestag die Einführung einer allgemeinen gesetzlichen Pflegeversicherung in der Sozialversicherung. Als Träger wurden die gesetzlichen Krankenversicherer bestimmt. Die Krankenversicherer wurden verpflichtet, sämtliche Krankenversicherte ohne erneute Gesundheitsüberprüfung in der gesetzlichen Pflegeversicherung zu übernehmen.

Am 01.April 1995 trat die Pflegeversicherung in Kraft, aber zunächst wurden nur Leistungen im Fall stationärer Pflege gewährt, ab 1.Juli 1996 wurde das Leistungsspektrum dann auch auf stationäre Pflege ausgedehnt.

Die Pflegebedürftigen werden nach Schwere des Pflegefalls in 3 Kategorien (sogenannte Pflegestufen) unterteilt. Die Stufenunterteilung wird nach folgender Definition unternommen:

- Stufe 1: Der Pflegebedürftige benötigt mindestens 90 Minuten Hilfe pro Tag bei Verrichtungen des täglichen Lebens (erheblich Pflegebedürftig).

- Stufe 2: Der Pflegebedürftige benötigt mindestens 180 Minuten Hilfe pro Tag bei Verrichtungen des täglichen Lebens (schwerpflegebedürftig).
- Stufe 3: Der Pflegebedürftige benötigt mindestens 300 Minuten Hilfe pro Tag bei Verrichtungen des täglichen Lebens (schwerstpflegebedürftig).

Diese Verrichtungen des täglichen Lebens beinhalten zum Beispiel Aufstehen/ Zubettgehen, An- und Auskleiden, Waschen, Kämmen, Rasieren, usw. Zu beachten ist hierbei, daß Hilfe bei hauswirtschaftlicher Versorgung in dieser Definition nicht enthalten ist. Die Stufeneinteilung wird von einem unabhängigen medizinischen Gutachter vorgenommen. Leistungshöhe und -art sind abhängig von der jeweiligen Pflegestufe und Art der Pflege. In der gesetzlichen Pflegeversicherung werden dem Pflegebedürftigen Kosten für häusliche Pflegehilfe, teil- und vollstationäre Pflege, technische Hilfsmittel, wie zum Beispiel Rollstuhl, sowie Kosten zur Verbesserung des Wohnumfeldes (Einbau eines Treppenlifts, Verbreiterung des Wohnungseingangs, etc.) erstattet. Zusätzlich zur gesetzlichen Pflegeversicherung wird in Deutschland von Lebens- und Kranken-

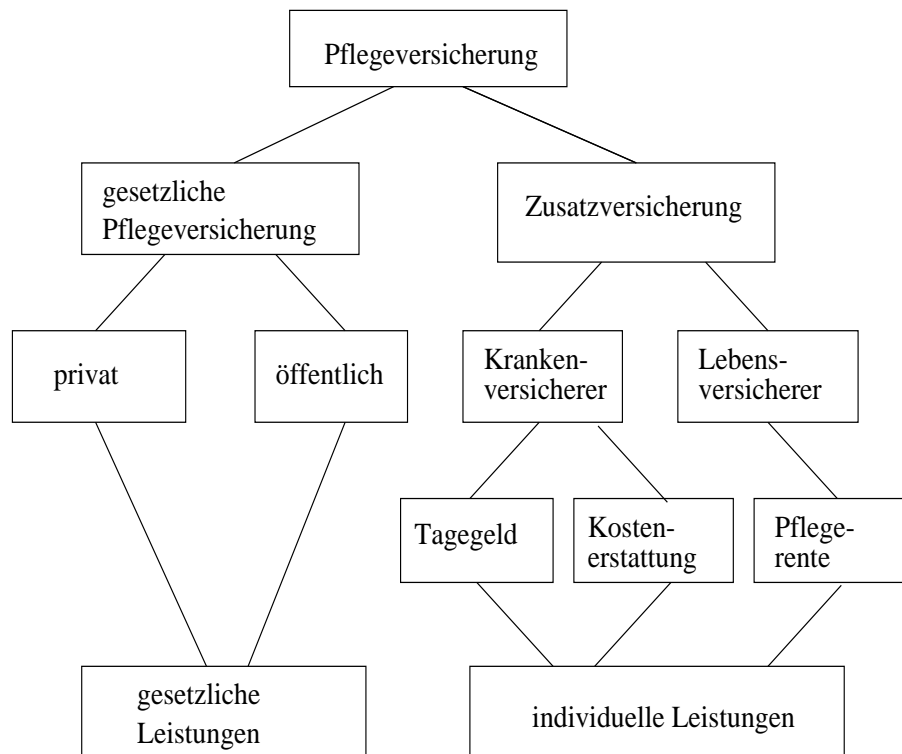


Abbildung 4.1: Übersicht über private und gesetzliche Pflegeversicherung in Deutschland

versicherern eine private Pflegeversicherung angeboten. Dabei unterscheiden sich die Tarife der Krankenversicherer von denen, die von Lebensversicherern angeboten werden in der Art der Leistung. Während Krankenversicherer meist Ergänzungsprodukte zur gesetzlichen Pflegeversicherung anbieten, die meist das Leistungsspektrum der gesetzlichen Versicherung erweitern und somit auch die Definition des Pflegefalls aus der gesetzlichen Versicherung übernehmen, bieten Lebensversicherer meist monatliche Pflegerenten an. Auch für die Definition des Pflegefalls

benutzen Lebensversicherer ein anderes System, das sogenannte ADL-System (ADL= Activities of Daily Living). Hier wird anhand eines fünf bis sechs Punkte beinhaltenden ADL-Katalogs bestimmt, wieviele dieser Verrichtungen des täglichen Lebens, der Pflegebedürftige nicht mehr selbständig verrichten kann. In Abbildung 4.1 ist die Untergliederung der privaten und gesetzlichen Pflegeversicherung schematisch dargestellt.

4.2 Beschreibung der Daten

Die Daten wurden zwischen dem 1. April 1995 und dem 31.12.1998 bei insgesamt 5603 Leistungsempfängern aus der privaten gesetzlichen Pflegeversicherung erhoben. Diese unterteilten sich in 3511 Frauen und 2092 Männer. In Abbildung 4.2 soll die Struktur der beobachteten Daten anhand von 2 Beobachtungen verdeutlicht werden. Mit den senkrechten Linien ist der Zeitraum markiert, in dem wir die Pflegefälle beobachten konnten. Betrachten wir nun den mit der durchgezogenen Linie dargestellten Pflegefall genauer. Der Pflegebeginn fällt in den Beobachtungszeitraum, zunächst befindet sich der Pflegebedürftige in ambulanter Pflege der Stufe 1. Von dort wechselte er in ambulante Pflege der Stufe 2, d.h. der Zustand verschlechterte sich. Der nächste zu beobachtende Wechsel war von ambulanter zu stationärer Pflege (jeweils Stufe 2). Hier wechselte der Pflegefall von Stufe 2 zu Stufe 3, wo er sich am Ende des Beobachtungszeitraums befand. Es handelt sich hierbei also offenbar um eine zensierte Beobachtung, da wir das terminierende Ereignis ("Tod oder Genesung des Pflegebedürftigen") nicht observieren konnten. Der mit der gestrichelten Linie dargestellte Pflegefall stellt eine linkstrunkierte Beobachtung

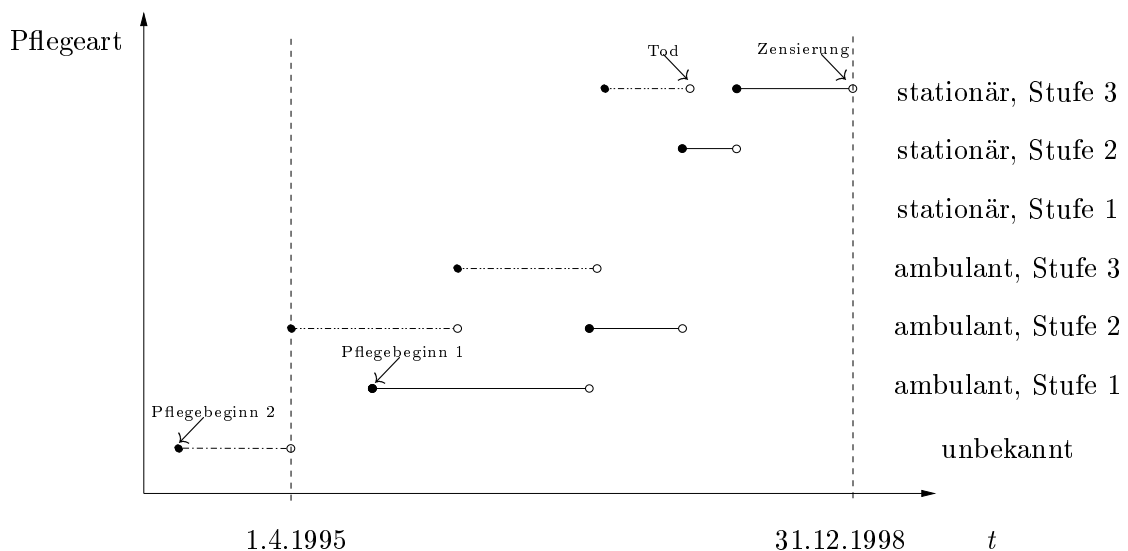


Abbildung 4.2: Zustandsübergänge im Pflegeprozess am Beispiel von 2 Pflegefällen

dar. Das bedeutet, der Beginn der Pflegebedürftigkeit ist bekannt (im medizinischen Gutachten wird nach dem Pflegebeginn gefragt), fällt aber in den Zeitraum vor Beginn der gesetzlichen Pflegeversicherung und konnte demzufolge nicht beobachtet werden.

	zensiert	tot	reakt.	Amb1	Amb2	Amb3	Stat1	Stat2	Stat3
Amb1	1012	279	28	–	444	75	118	71	34
Amb2	877	598	1	47	–	296	9	208	59
Amb3	308	632	2	2	20	–	0	4	87
Stat1	248	85	3	11	1	0	–	108	26
Stat2	451	263	2	1	6	1	7	–	116
Stat3	376	437	0	1	1	6	2	9	–

Tabelle 4.1: Anzahl der Zustandsübergänge im Datensatz

Insgesamt wurden 7372 Stufenübergänge beobachtet, deren empirische Häufigkeit in Tabelle 4.1 dargestellt ist.

Hierbei stehen "Amb i " und "Stat i " jeweils für ambulante, bzw. stationäre Pflege in Stufe i . Der Übergang zu "zensiert" bedeutet, daß der jeweilige vorherige Zustand der letzte beobachtete Zustand am 31. Dezember 1998 war. Die Abkürzung "reakt." steht für Pflegebedürftige, die wieder Nicht- Pflegebedürftig (=reaktiviert) werden. Die Abkürzung "gek." steht für Pflegebedürftige deren Verträge gekündigt wurden (von Seiten des Versicherers oder von Seiten des Versicherten). Diese Übergänge werden technisch wie zensierte Beobachtungen gehandhabt.

Auf den ersten Blick erkennt man in den Daten sofort, daß Übergänge zu einer "schlechteren" Pflegestufe weitaus häufiger sind als zu einer besseren. So wechselten zum Beispiel 444 Pflegebedürftige von ambulanter Pflege, Stufe 1, in ambulante Pflege Stufe 2, während nur 47 Pflegebedürftige von Stufe 2 in Stufe 1 wechselten. Auch die niedrige Anzahl von Reaktivierungen (insgesamt 36) ist ein Anzeichen dafür, daß Verbesserungen im Pflegestatus relativ unwahrscheinlich sind.

Um einen genaueren Überblick über den Datensatz zu bekommen, wurden nach Geschlechtern getrennt, die Anzahl der Personen in den verschiedenen Pflegestufen bzw. in ambulanter oder stationärer Pflege gezählt. Da ein Individuum aufgrund Wechsel der Pflegestufe, beziehungsweise der Pflegeart in mehreren Kategorien gezählt werden kann, entspricht die Summe über die Kategorien nicht der Gesamtanzahl der Individuen. Die Zahlen sind in Tabelle 4.2 dargestellt. Wir sehen hier, daß 40.6% aller beobachteten Frauen sich im Laufe ihrer Pflegehistorie einmal in stationärer Pflege befunden hatten, aber nur 22.2% aller beobachteten Männer. Die Werte summieren sich zu mehr als 100%, da sich einige beobachtete Personen sowohl in stationärer, als auch in ambulanter Pflege befunden hatten. Zudem wurden noch bei 5044 Pflegefällen (davon 3176 Frauen und 1868 Männer) die pflegeauslösenden Diagnosen erfaßt. Die Häufigkeiten der einzelnen Diagnosen sind in Tabelle 4.3 zusammengefaßt.

	Frauen	Männer	Frauen in %	Männer in %
Ambulant	2526	1798	72.0	86.0
Stationär	1427	464	40.6	22.2
Stufe 1	1581	827	40.0	39.5
Stufe 2	1711	1011	48.7	48.3
Stufe 3	1074	719	30.6	34.4

Tabelle 4.2: Häufigkeiten für Pflegeart und Pflegestufe, unterteilt nach Geschlecht

Diagnose	Frauen	Männer	Frauen in %	Männer in %
Demenzerkrankungen	1469	694	46.3	37.2
Schlaganfälle	497	549	15.6	29.4
Psychosen	679	577	21.4	30.9
Tumore	350	346	11.0	18.5
Blindheit	117	75	3.7	4.0
Knochenkrankheiten	865	152	27.2	8.1
Arthrosen	433	103	13.6	5.5
Herzerkrankungen	1352	572	42.6	30.6
Lungenerkrankungen	50	45	1.6	2.4
Nierenerkrankungen	2	3	0.1	0.2
Geburtsschäden	17	20	0.5	1.1

Tabelle 4.3: Häufigkeiten der verschiedenen Diagnosen im Datensatz

4.3 Proportional-Hazard-Modell

Zunächst soll hier ein Modell entwickelt werden, das die Überlebenszeit in Abhängigkeit von den Kovariablen Alter, Geschlecht, Art der Pflege und Pflegestufe schätzt. Die pflegeverursachenden Diagnosen sind hier noch nicht berücksichtigt, da das hier ermittelte Modell später auch bei der Entwicklung eines Versicherungsmodells Anwendung finden soll.

4.3.1 Modellbildung (ohne Diagnosen)

Um einen ersten Überblick über die Struktur der Daten zu erhalten wurde ein erstes Proportional-Hazard-Modell gefittet, bei dem folgende Kovariablen berücksichtigt wurden:

$Z_{\text{Alter}}(t)$ als zeitabhängige Kovariable, die das Alter eines Pflegebedürftigen bei Eintritt in eine neue Pflegestufe enthält (d.h. Z_{Alter} kann als Sprungfunktion mit Werten in \mathbb{R} und Sprüngen an allen Stufenübergängen interpretiert werden). $Z_{\text{Geschlecht}}$, $Z_{\text{Stufe2}}(t)$ und $Z_{\text{Stufe3}}(t)$ sind als binäre Kovariablen definiert, wobei $Z_{\text{Geschlecht}} = 1$ falls die pflegebedürftige Person weiblich ist und 0 sonst, $Z_{\text{Stufe2}}(t) = 1$ bzw. $Z_{\text{Stufe3}}(t) = 1$ falls sich die pflegebedürftige Person zum Zeit-

punkt t in der Pflegestufe 1 bzw. 2 befindet und $Z_{\text{Stat}}(t) = 1$ falls sich die beobachtete Person zum Zeitpunkt t in stationärer Pflege befindet.

Für die Hazardfunktion wurde ein additives Modell gewählt:

$$\lambda_Z(t) = \lambda_0(t) \exp[\beta_1 Z_{\text{Alter}}(t) + \beta_2 Z_{\text{Geschlecht}} + \beta_3 Z_{\text{Stat}}(t) + \beta_4 Z_{\text{Stufe2}}(t) + \beta_5 Z_{\text{Stufe3}}(t)]. \quad (4.1)$$

Das so gewählte Modell ist sicherlich noch nicht geeignet, die Realität ausreichend zu modellieren, da es überhaupt keine Interaktionen zwischen Kovariablen berücksichtigt, es soll an dieser Stelle trotzdem erwähnt werden, da es relativ einfach zu interpretieren ist. Mit der Efron-Approximation für den partiellen Likelihood führt das Modell 4.1 zu den in Tabelle 4.4 dargestellten Resultaten.

Variable	Koeffizient	exp(Koeff.)	SE(Koeff)	z-Statistik	p-Wert
Z_{Alter}	0.0244	1.025	0.00179	13.6056	$< 10^{-15}$
$Z_{\text{Geschlecht}}$	-0.3495	0.705	0.04580	-7.6320	$2.3 \cdot 10^{-14}$
Z_{Art}	-0.0022	0.998	0.04778	-0.0459	0.96
Z_{Stufe2}	0.7610	2.140	0.06313	12.0536	$< 10^{-15}$
Z_{Stufe3}	1.5577	4.748	0.06358	24.4998	$< 10^{-15}$

Tabelle 4.4: Geschätzte Koeffizienten im Modell 4.1

Im ersten Fit wird für die Kovariablen Z_{Alter} , Z_{Stufe2} , Z_{Stufe3} und $Z_{\text{Geschlecht}}$ ein signifikanter Einfluß der Kovariablen auf den Hazard festgestellt. Die Werte können folgendermaßen interpretiert werden: Der positive Koeffizient für die Kovariable Alter belegt, daß mit zunehmendem Alter auch die Sterbeintensität zunimmt, was zu erwarten war. Auch die signifikante Abhängigkeit des Hazards vom Geschlecht überrascht hier nicht sonderlich, denn nicht nur in der Gesamtbevölkerung, sondern auch in der Pflegeversicherung kann man eine höhere Sterblichkeit von Männern beobachten (der ersten Schätzung zufolge liegt der Hazard für Frauen bei ca. 70 % des Hazards für Männer). Die Kovariablen für Stufe 2 und Stufe 3 stellen ein erstes Maß dafür dar, inwieweit sich die Schwere der Pflege auf die Sterbeintensität auswirkt. Interessanterweise hat Kovariable für stationäre Pflege keinen signifikanten Einfluß auf die Pflegebedürftigkeit (der P-Wert liegt bei 96 %), das heißt daß die Nullhypothese “die Kovariable hat keinen Einfluß auf den Hazard” nicht abgelehnt werden kann. Da dieses Modell jedoch noch sehr grob gewählt ist, kann man die Kovariable “Pflegeart” nicht einfach unberücksichtigt lassen, da sie in Interaktion mit anderen Variablen trotzdem einen Einfluß auf das Modell haben kann.

Inwieweit man wirklich von einem proportionalen Einfluß auf den Hazard ausgehen kann (d.h. ob $\lambda_Z(t) = \lambda_0(t) \exp(\beta^t Z(t))$ für alle t) muß durch weitere Tests begründet werden.

Um den Zusammenhang zwischen stationärer und ambulanter Pflegebedürftigkeit genauer zu untersuchen, wurde ein nach stationärer Pflege stratifiziertes (d.h. für stationäre und ambulante Pflege separat berechnetes) Proportional-Hazard-Modell gefittet und die mittels des Breslow-Schätzers für den Basis-Hazard berechneten Kurven gegeneinander geplottet. In Abbildung 4.3

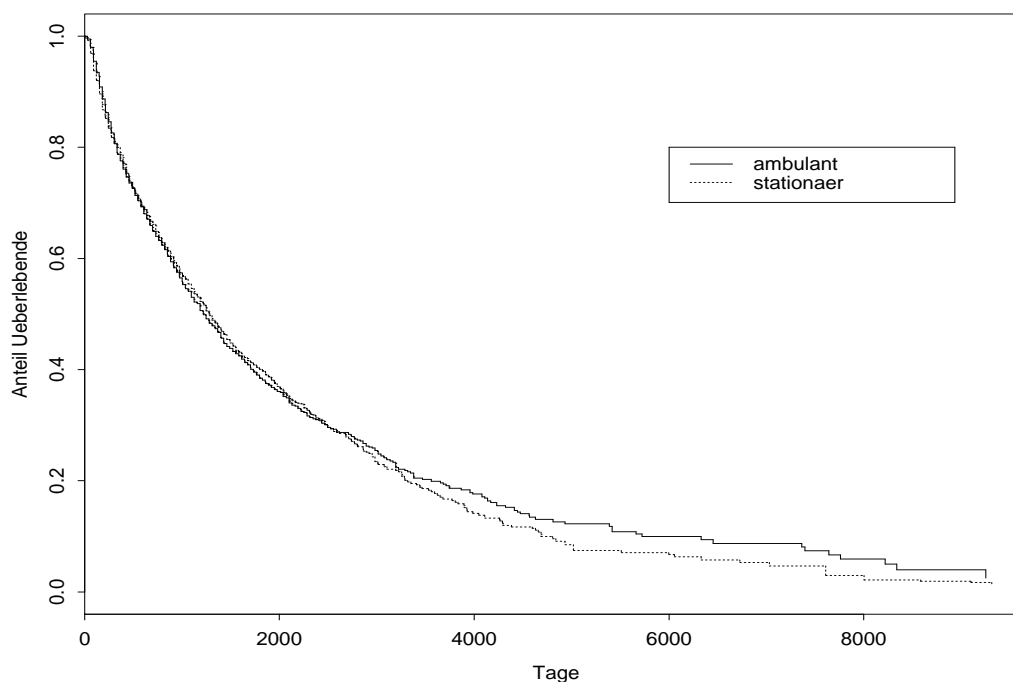


Abbildung 4.3: Mit Breslow-Schätzer geschätzte Überlebensfunktionen für ambulante und stationäre Pflege

erkennt man, daß bei langen Pflegedauern (größer als 3000 Tage), die Überlebensfunktion für Pflege in einem Pflegeheim deutlich unter der entsprechenden Funktion für ambulante Pflege liegt, was dafür spricht, daß Interaktionen zwischen den Kovariablen bestehen, die bislang noch nicht berücksichtigt wurden.

Modellierung der Interaktionen

Die Erkennung und Modellierung eventuell bestehender Interaktionen soll nun mit Hilfe des AIC durchgeführt werden. Das Vorgehen wird dabei in 2 Schritte unterteilt:

- Bestimmung zwischen wie vielen Faktoren Interaktion modelliert werden kann (2-Faktor, 3-Faktor, 4 Faktor-Interaktion, oder überhaupt keine).
- Bestimmung, welche Interaktionen schließlich signifikanten Einfluß auf das Modell haben.

Für den ersten Schritt wird folgendes Vorgehen gewählt: Für den Datensatz werden 4 verschiedene Modelle gefittet und zwar jeweils eines ohne Interaktion, mit allen 2-Faktor-Interaktionen, mit allen Interaktionen, die bis zu 3 Faktoren enthalten und letztendlich ein Modell mit sämtlichen Interaktionen. Solange Interaktionen Einfluß auf das Modell haben, wird sich dieser Einfluß durch ein Wachsen der Likelihoodfunktion und damit auch des Log-Likelihood bemerkbar machen. Falls Interaktionen hinzugefügt werden, die keinen, oder kaum Einfluß haben, so wird sich

Modell	Log-Likelihood	df	AIC ($n = 2$)
ohne Interaktion	-15909.2	5	31828.4
2 Faktoren	-15878.2	14	31784.4
≤ 3 Faktoren	-15875.8	21	31793.6
alle Interaktionen	-15874.6	23	31795.2

Tabelle 4.5: AIC-Kriterium für Modelle mit Interaktionen

der Log-Likelihood nur unwesentlich ändern, der Wert des AIC wird jedoch größer, da das neue Modell mehr Freiheitsgrade beinhaltet. Die Ergebnisse sind in Tabelle 4.3.1 dargestellt. Man erkennt, daß das Modell ohne Interaktionen (im Vergleich zum Modell mit 2-Faktor-Interaktionen) einen relativ hohen Likelihood hat, der sich auch im Informations-Kriterium bemerkbar macht. Außerdem erkennt man, daß sich der Log-Likelihood nur unwesentlich ändert, wenn man 3-Faktor-, bzw. alle Interaktionen betrachtet (Der Likelihood für 2-Faktor-Interaktionen unterscheidet sich vom Likelihood für alle Interaktionen nur um 3.6, das volle Modell enthält aber 9 (23 - 14) Freiheitsgrade mehr!).

Variable	Koeffizient	exp(Koeff.)	SE(Koeff)	z-Statistik	p-Wert
Z_{Alter}	0.0272	1.028	0.0050	5.483	$4.2 \cdot 10^{-8}$
$Z_{\text{Geschlecht}}$	0.1401	1.150	0.3013	0.465	$6.4 \cdot 10^{-1}$
Z_{Stat}	-0.5948	0.552	0.4114	-1.446	$1.5 \cdot 10^{-1}$
Z_{Stufe2}	1.0679	2.909	0.4326	2.469	$1.4 \cdot 10^{-2}$
Z_{Stufe3}	1.8610	6.430	0.4219	4.411	$1.0 \cdot 10^{-5}$
$Z_{\text{Alter}} \times Z_{\text{Geschlecht}}$	-0.0064	0.994	0.0035	-1.829	$6.7 \cdot 10^{-2}$
$Z_{\text{Alter}} \times Z_{\text{Stat}}$	0.0161	1.016	0.0047	3.407	$6.6 \cdot 10^{-4}$
$Z_{\text{Alter}} \times Z_{\text{Stufe2}}$	-0.0048	0.995	0.0054	-0.894	$3.7 \cdot 10^{-1}$
$Z_{\text{Alter}} \times Z_{\text{Stufe3}}$	-0.0027	0.997	0.0053	-0.510	$6.1 \cdot 10^{-1}$
$Z_{\text{Geschlecht}} \times Z_{\text{Stat}}$	-0.4515	0.637	0.0995	-4.537	$5.7 \cdot 10^{-6}$
$Z_{\text{Geschlecht}} \times Z_{\text{Stufe2}}$	0.2275	1.255	0.1321	1.721	$8.5 \cdot 10^{-2}$
$Z_{\text{Geschlecht}} \times Z_{\text{Stufe3}}$	0.1842	1.202	0.1292	1.425	$1.5 \cdot 10^{-1}$
$Z_{\text{Stat}} \times Z_{\text{Stufe2}}$	-0.3210	0.725	0.1470	-2.184	$2.9 \cdot 10^{-2}$
$Z_{\text{Stat}} \times Z_{\text{Stufe3}}$	-0.6737	0.510	0.1416	-4.758	$2.0 \cdot 10^{-6}$

Tabelle 4.6: Geschätzte Koeffizienten im Modell mit allen 2-Faktor-Interaktionen

Daher deutet alles darauf hin, daß im "optimalen" Modell nur 2-Faktor-Interaktionen vorkommen. Daher soll nun in einem zweiten Schritt das Modell mit allen Interaktionen zwischen 2 Faktoren genauer betrachtet werden und durch Entfernen von eventuell nicht signifikanten Interaktionen noch verbessert werden. Das Modell mit allen 2-Faktor-Interaktionen ist in Tabelle 4.6 dargestellt

Tabelle 4.6 legt die Vermutung nahe, daß immer noch zu viele Kovariablen, die keinen signifikanten Einfluß auf das Modell haben, vorhanden sind. Im Falle der Interaktion zwischen Alter und Stufe, sprechen der p-Wert von 0.37 für die Interaktion $Z_{\text{Alter}} \times Z_{\text{Stufe2}}$, bzw. von 0.61 für die Interaktion $Z_{\text{Alter}} \times Z_{\text{Stufe3}}$ dafür, daß noch zu viele Kovariablen berücksichtigt werden. Der lokale Likelihood-Ratio-Test für eine Interaktion zwischen Alter und Pflegestufe ergibt einen p-Wert von 0.85, d.h. man kann nicht davon ausgehen, daß diese Interaktion signifikanten Einfluß im Modell hat. Um diese Vermutung zu belegen soll nun durch iteratives Entfernen von Interaktionen eine weitere Verbesserung des AIC erzielt werden. In Tabelle 4.7 sind die Ergebnisse für das Entfernen von jeweils einer Interaktion (aus dem Modell mit allen Interaktionen) dargestellt.

Entfernen von	Log-Likelihood	df	AIC ($n = 2$)
Alter \times Geschlecht	-15879.8	13	31785.6
Alter \times Pflegeart	-15884.3	13	31794.6
Alter \times Pflegestufe	-15878.7	12	31781.4
Geschlecht \times Pflegeart	-15888.2	13	31802.4
Geschlecht \times Pflegestufe	-15879.7	12	31783.4
Pflegeart \times Pflegestufe	-15891.3	12	31806.6

Tabelle 4.7: Entwicklung des AIC bei Entfernen von Interaktionen

In Tabelle 4.7 findet sich die Vermutung bestätigt, daß sich durch Entfernen der Interaktion zwischen Alter und Pflegestufe eine Verbesserung des Informationskriteriums erreichen läßt (der Wert beträgt 31781.4 im Vergleich zu 31784.4 für Modell mit allen Interaktionen). Man erkennt jedoch auch, daß man für das Modell ohne Interaktion Geschlecht \times Pflegestufe eine leichte Verbesserung des AIC erzielt. Daher sollen durch analoges Vorgehen für das Modell ohne Interaktion zwischen Alter und Pflegestufe getestet werden, ob weitere Interaktionen, die nicht zur Erklärung des Modells beitragen, bestehen (Tabelle 4.8).

Nach diesem Schritt sieht man eine Verbesserung des AIC bei Herausnahme der Interaktion

Entfernen von	Log-Likelihood	df	AIC ($n = 2$)
Alter \times Geschlecht	-15880.5	11	31783.0
Alter \times Pflegeart	-15885.1	11	31792.2
Geschlecht \times Pflegeart	-15888.6	11	31799.2
Geschlecht \times Pflegestufe	-15879.9	10	31779.8
Pflegeart \times Pflegestufe	-15892.2	10	31804.4

Tabelle 4.8: Entwicklung des AIC bei Entfernen weiterer Interaktionen (ausgehend von Modell ohne Interaktion zwischen Alter und Pflegestufe)

zwischen Alter und Pflegestufe. Falls man diese Vorgehensweise nun nochmals anwendet, findet keine Verbesserung mehr statt (siehe Tabelle 4.9), so daß das optimale Modell alle Interaktionen bis Alter \times Pflegestufe und Geschlecht \times Pflegeart beinhaltet.

Entfernen von	Log-Likelihood	df	AIC ($n = 2$)
Alter \times Geschlecht	-15881.7	9	31781.4
Alter \times Pflegeart	-15886.2	9	31792.2
Geschlecht \times Pflegeart	-15889.6	9	31797.2
Pflegeart \times Pflegestufe	-15893.6	8	31803.2

Tabelle 4.9: Entwicklung des AIC bei Entfernen weiterer Interaktionen (ausgehend von Modell ohne Interaktionen: Alter \times Pflegestufe, Geschlecht \times Pflegestufe)

Die Hazardfunktion besitzt nach diesem Modell folgende Darstellung:

$$\begin{aligned}
 \lambda(t) = & \lambda_0(t) \exp[\beta_1 Z_{\text{Alter}}(t) + \beta_2 Z_{\text{Geschlecht}} + \beta_3 Z_{\text{Stat}}(t) + \beta_4 Z_{\text{Stufe2}}(t) \\
 & + \beta_5 Z_{\text{Stufe3}}(t) + \beta_6 Z_{\text{Alter}}(t) \times Z_{\text{Geschlecht}} + \beta_7 Z_{\text{Alter}}(t) \times Z_{\text{Stat}}(t) \\
 & + \beta_8 Z_{\text{Geschlecht}} \times Z_{\text{Stat}}(t) + \beta_9 Z_{\text{Stat}}(t) \times Z_{\text{Stufe2}}(t) \\
 & + \beta_{10} Z_{\text{Stat}}(t) \times Z_{\text{Stufe3}}(t)]. \tag{4.2}
 \end{aligned}$$

Mit der Efron-Approximation für den Partial-Likelihood erhält man für das Modell 4.2 die in Tabelle 4.10 dargestellten Koeffizienten. Durch die vielen Interaktionen ist eine Interpretation

Variable	Koeffizient	exp(Koeff.)	SE(Koeff)	z-Statistik	p-Wert
Z_{Alter}	0.02421	1.025	0.0025	9.78	$< 10^{-15}$
$Z_{\text{Geschlecht}}$	0.3202	1.377	0.2811	-1.47	$2.5 \cdot 10^{-1}$
Z_{Stat}	-0.5969	0.551	0.4063	1.14	$1.4 \cdot 10^{-1}$
Z_{Stufe2}	0.8180	2.266	0.0729	11.22	$< 10^{-15}$
Z_{Stufe3}	1.7553	5.785	0.0734	23.91	$< 10^{-15}$
$Z_{\text{Alter}} \times Z_{\text{Geschlecht}}$	-0.0065	0.993	0.0035	-1.89	$5.9 \cdot 10^{-2}$
$Z_{\text{Alter}} \times Z_{\text{Stat}}$	0.0162	1.016	0.0047	3.45	$5.7 \cdot 10^{-4}$
$Z_{\text{Geschlecht}} \times Z_{\text{Stat}}$	-0.4405	0.644	0.0989	-4.46	$8.4 \cdot 10^{-6}$
$Z_{\text{Stat}} \times Z_{\text{Stufe2}}$	-0.3320	0.717	0.1446	-2.30	$2.2 \cdot 10^{-2}$
$Z_{\text{Stat}} \times Z_{\text{Stufe3}}$	-0.6753	0.509	0.1391	-4.86	$1.2 \cdot 10^{-6}$

Tabelle 4.10: Geschätzte Koeffizienten im Modell (4.2)

der Koeffizienten kaum mehr möglich. Zum Beispiel erstaunt, der positive Koeffizient für die Kovariable $Z_{\text{Geschlecht}}$ (in Modell (4.1) hatte der Koeffizient den Wert -0.3495). Da jedoch die Koeffizienten für die Interaktion zwischen Alter und Geschlecht negativ sind, wird dieser Effekt gerade für hohe Alter wieder kompensiert. In Tabelle 4.11 ist daher für das durchschnittliche

Alter im Datensatz von 78.72 Jahren dargestellt, wie sich der Multiplikator für den Basis-Hazard $\exp(\beta^t Z)$ für verschiedene Kovariablenwerte Z entwickelt.

	Frauen	Männer
ambulant, Stufe 1	5.548	6.720
ambulant, Stufe 2	12.573	15.227
ambulant, Stufe 3	32.100	38.874
stationär, Stufe 1	12.786	13.242
stationär, Stufe 2	20.787	21.529
stationär, Stufe 3	37.649	38.993

Tabelle 4.11: Multiplikator $\exp(\beta^t Z)$ für verschiedene Kovariablenkombinationen (bei fixiertem Alter v. 78.72 Jahren)

4.3.2 Proportional Hazard (mit Diagnosen)

In diesem Abschnitt soll versucht werden, mit Hilfe der pflegeverursachenden Diagnosen das Proportional-Hazard-Modell eventuell noch zu verbessern. Natürlich liegt die Vermutung nahe, daß die festgestellten pflegeauslösenden Krankheiten einen massiven Einfluß auf die Überlebensfunktion haben. Auch hier sollen mit Informationskriterium von Akaike Diagnosen, die einen signifikanten Einfluß auf das Modell haben, herausgefiltert werden. Da nicht für alle Pflegefälle Diagnosen erfaßt wurden, berechnen wir zunächst noch das Informationskriterium für Modell (4.2) auf Basis aller Daten, für die Diagnosen erfaßt wurden, um so ein Vergleichskriterium zu erhalten. In einem zweiten Fit betrachten wir dann ein Modell, das zusätzlich zu den in (4.2) spezifizierten Kovariablen für jede Diagnose eine weitere Kovariable berücksichtigt. In einzelnen bedeutet dies

$$Z_{\text{Diagnose}} = \begin{cases} 1 & \text{entsprechende Diagnose wurde bei der Person festgestellt} \\ 0 & \text{sonst} \end{cases}$$

Der Log-Likelihood und das AIC für diese Modelle sind in Tabelle 4.12 dargestellt. Wie erwartet

Modell	Log-Likelihood	df	AIC ($n = 2$)
ohne Diagnosen	-14467.1	10	28954.2
mit allen Diagnosen	-14270.1	21	28582.0

Tabelle 4.12: AIC-Kriterium für Modelle mit, bzw. ohne Berücksichtigung der pflegeverursachenden Diagnosen

haben also die pflegeverursachenden Krankheiten einen entscheidenden Einfluß auf das Überleben in der gesetzlichen Pflegeversicherung. In Tabelle 4.13 sind die mit der Efron-Approximation für den proportionalen Hazard geschätzten Koeffizienten dargestellt.

Variable	Koeffizient	exp(Koeff.)	SE(Koeff)	z-Statistik	p-Wert
Z_{Demenz}	-0.1351	0.874	0.0483	-2.80	$5.1 \cdot 10^{-3}$
Z_{Schlag}	-0.2067	0.813	0.0573	-3.61	$3.1 \cdot 10^{-4}$
Z_{Psychose}	-0.2056	0.814	0.0549	-3.74	$1.8 \cdot 10^{-4}$
Z_{Tumore}	1.0504	2.859	0.0570	18.43	$< 10^{-15}$
$Z_{\text{Blindheit}}$	-0.1071	0.898	0.1133	-0.95	$3.4 \cdot 10^{-1}$
Z_{Knochenk}	-0.0818	0.921	0.0583	-1.40	$1.6 \cdot 10^{-1}$
$Z_{\text{Arthrosen}}$	-0.0798	0.923	0.0746	-1.07	$2.9 \cdot 10^{-1}$
Z_{Herz}	-0.0784	1.082	0.0460	1.70	$8.8 \cdot 10^{-2}$
Z_{Lungen}	0.3391	1.404	0.1393	2.43	$1.5 \cdot 10^{-2}$
Z_{Nieren}	0.4271	1.533	0.7152	0.60	$5.5 \cdot 10^{-1}$
Z_{Geburt}	-0.9219	0.398	0.8846	-1.04	$3.0 \cdot 10^{-1}$

Tabelle 4.13: Geschätzte Koeffizienten im Modell (4.2)

Bei genauerer Betrachtung der geschätzten Koeffizienten fällt vor allem der hohe Wert des Koeffizienten und die hohe Signifikanz (p-Wert $< 10^{-15}$) der Kovariable Z_{Tumore} auf. Hier handelt es sich wohl meist um Krebspatienten im Endstadium, deren Heilungschancen gering sind. Zudem erkennt man positive Koeffizienten bei Erkrankungen lebenswichtiger Organe (Herz, Lunge und Nieren), alle weiteren Koeffizienten sind negativ. Auf eine detailliertere Betrachtung der Diagnosen hinsichtlich Interaktion mit anderen Kovariablen, bzw. Interaktion mit anderen Diagnosen möchte ich in der Arbeit verzichten, da in das später im Text formulierte Versicherungsmodell nur die Kovariablen für Alter, Geschlecht, Pflegeart und Pflegestufe eingehen. Der Grund dafür liegt darin, daß in den meisten Versicherungsverträgen Art und Höhe der Leistungen abhängig von der Schwere der Pflege, nicht aber von pflegeverursachenden Diagnosen sind.

4.4 Test auf Proportional Hazard

4.4.1 Residuenanalyse

Zunächst wird hier nun mit Hilfe der Martingal-Residuen untersucht, ob im Modell (4.2) die quantitative Kovariable Alter bereits gut modelliert ist, oder ob eventuell eine andere funktionale Form für die Kovariable gefunden werden kann. Dazu fitten wir ein Modell, mit den Kovariablen und Interaktionen aus (4.2), die nicht vom Alter abhängen

$$\begin{aligned}
 \lambda(t) &= \lambda_0(t) \exp[\beta_1 Z_{\text{Geschlecht}} + \beta_2 Z_{\text{Stat}}(t) + \beta_3 Z_{\text{Stufe2}}(t) + \beta_4 Z_{\text{Stufe3}}(t) \\
 &+ \beta_5 Z_{\text{Geschlecht}} \times Z_{\text{Stat}}(t) + \beta_6 Z_{\text{Stat}}(t) \times Z_{\text{Stufe2}}(t) \\
 &+ \beta_7 Z_{\text{Stat}}(t) \times Z_{\text{Stufe3}}(t)].
 \end{aligned} \tag{4.3}$$

Für das nach (4.3) berechnete Modell plottet man dann die Martingal-Residuen gegen die Kovariable Alter und schätzt mit einem Glättungsoperator den Erwartungswert dieser Residuen $E(\hat{M}|Z_{\text{Alter}})$ (Abbildung 4.4, linke Seite). Der Plot sollte dann ungefähr die funktionale Form darstellen, in der die Kovariable in das Modell eingeht (siehe Abschnitt 3.7.8).

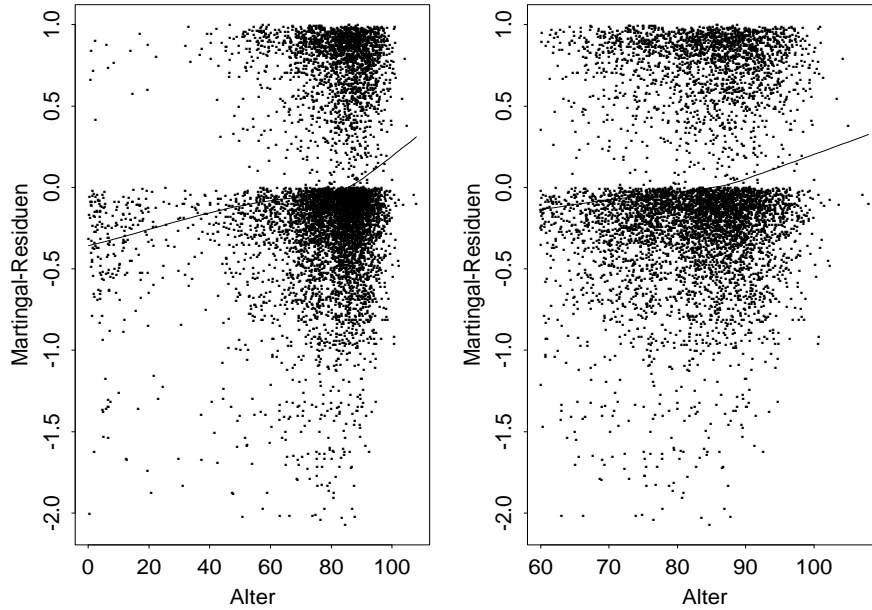


Abbildung 4.4: Martingal-Residuen zur Bestimmung der funktionalen Form der Kovariablen Z_{Alter} in Modell (4.2)

Man erkennt hier, daß sich die Funktion stückweise linear verhält, mit einem Knick im Altersbereich von ca. 80 Jahren. Um diesen Knick genauer zu lokalisieren wurde die Residuen nochmals für den Alterbereich zwischen 60 und 100 geplottet (Abbildung 4.4, rechte Seite), in dieser Graphik sehen wir, daß sich für ein Alter von ca. 85 Jahren die Steigung der Kurve erhöht. Der Plot erlaubt nun folgende Interpretation: Die Kovariable Alter geht linear in das Modell ein, ab einem bestimmten Alter θ (im Bereich von ca. 85) ändert sich jedoch der Regressionskoeffizient für die Kovariable Alter.

Um dieses Alter θ zu finden kann man nun folgendermaßen vorgehen (vgl. Klein und Moeschberger [25], Seiten 334 - 336): Man erweitert das Modell (4.2) um eine weitere Kovariable

$$Z_{\theta}(t) = \begin{cases} 1 & \text{falls } Z_{\text{Alter}}(t) \geq \theta \\ 0 & \text{sonst} \end{cases},$$

sowie um die Interaktion $Z_{\text{Alter}}(t) \times Z_{\theta}(t)$ zu

$$\begin{aligned} \lambda(t) &= \lambda_0(t) \exp[\beta_1 Z_{\text{Alter}}(t) + \beta_2 Z_{\text{Geschlecht}} + \beta_3 Z_{\text{Stat}}(t) + \beta_4 Z_{\text{Stufe2}}(t) \\ &+ \beta_5 Z_{\text{Stufe3}}(t) + \beta_6 Z_{\text{Alter}}(t) \times Z_{\text{Geschlecht}} + \beta_7 Z_{\text{Alter}}(t) \times Z_{\text{Stat}}(t) \end{aligned}$$

$$\begin{aligned}
& + \beta_8 Z_{\text{Geschlecht}} \times Z_{\text{Stat}}(t) + \beta_9 Z_{\text{Stat}}(t) \times Z_{\text{Stufe2}}(t) \\
& + \beta_{10} Z_{\text{Stat}}(t) \times Z_{\text{Stufe3}}(t)] + \beta_{11} Z_{\theta}(t) + \beta_{12} Z_{\text{Alter}}(t) \times Z_{\theta}(t). \quad (4.4)
\end{aligned}$$

Ziel ist nun die Maximierung des Partial-Likelihood für verschiedene Alter θ . Dafür wurden für den Altersbereich $[80, 92.5]$ für die Werte $\theta = 80 + 0.5i$ $i \in \{1, \dots, 25\}$ mit der Efron-Approximation für den Partial-Likelihood nach (4.4) berechnet und die Auswirkungen auf den Log-Likelihood betrachtet (Abbildung 4.5).

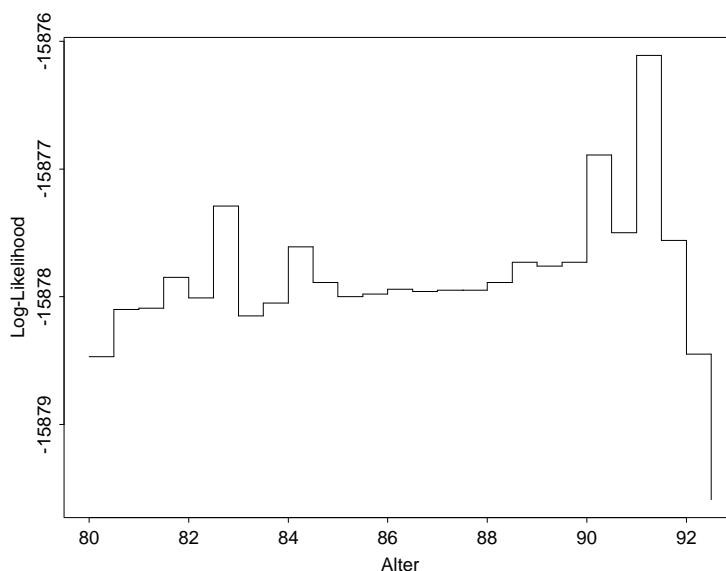


Abbildung 4.5: Entwicklung des Log-Likelihood für verschiedene Werte von θ in Modell (4.4)

Hier können wir folgendes erkennen: Die Funktion hat ein globales Maximum für $\theta = 91,5$, zudem fällt auf, daß an der Stelle 82.5 ein lokales Maximum liegt. Sowohl für $\theta = 82.5$, als auch für $\theta = 91.5$ verbessert sich das AIC im Vergleich zu Modell (4.2) (31778.6 bzw. 31776.2 im Vergleich zu 31779.8). Der partielle Log-Likelihood verbessert sich um ca. 3.8. Mit dem lokalen Likelihood-Ratio-Test (3.46) sieht man, daß diese Änderung im partiellen Log-Likelihood wirklich signifikant ist. Das 0.95-Quantil der χ^2 -Verteilung mit 2 Freiheitsgraden liegt bei 5.99, die doppelte Differenz im partiellen Log-Likelihood liegt bei ca. $2 \cdot 3.8 = 7.6$.

Die 2 Maxima, die sich im Plot beobachten lassen, legen die Vermutung nahe, daß der “Knick” für Männer und Frauen zu verschiedenen Zeitpunkten vorliegt. Um dies zu untersuchen wurde jeweils ein Modell für Frauen und ein Modell für Männer berechnet, dessen Hazardfunktion nach folgendem Proportional-Hazard-Modell konstruiert wurde

$$\begin{aligned}
\lambda(t) & = \lambda_0(t) \exp[\beta_1 Z_{\text{Stat}}(t) + \beta_2 Z_{\text{Stufe2}}(t) + \beta_3 Z_{\text{Stufe3}}(t) \\
& + \beta_4 Z_{\text{Stat}}(t) \times Z_{\text{Stufe2}}(t) + \beta_5 Z_{\text{Stat}}(t) \times Z_{\text{Stufe3}}(t)] \quad (4.5)
\end{aligned}$$

und die korrespondierenden Martingal-Residuen für Männer und Frauen getrennt gegen die Kovariable Alter geplottet (Abbildung 4.6). Wir erkennen hier für Frauen einen relativ stark

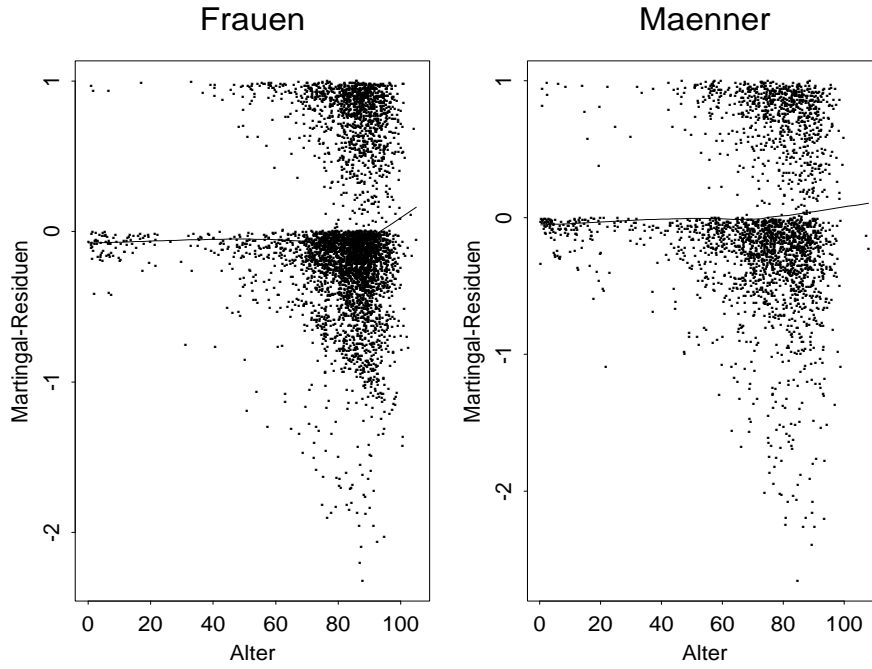


Abbildung 4.6: Martingal-Residuen (nach Geschlecht getrennt) zur Bestimmung der funktionalen Form der Kovariablen Z_{Alter} in Modell (4.2)

ausgeprägten Knick im Alter von ca. 90 Jahren und für Männer einen leichten Knick im Altersbereich zwischen 70 und 80 Jahren. Daher betrachten wir nun ein Modell mit folgenden zusätzlichen Kovariablen:

$$Z_{\theta_w} = \begin{cases} 1 & \text{falls } Z_{\text{Alter}} \geq \theta_w \text{ und } Z_{\text{Geschlecht}} = 1 \\ 0 & \text{sonst} \end{cases},$$

sowie

$$Z_{\theta_m} = \begin{cases} 1 & \text{falls } Z_{\text{Alter}} \geq \theta_m \text{ und } Z_{\text{Geschlecht}} = 0 \\ 0 & \text{sonst} \end{cases}$$

und maximieren den Log-Likelihood für folgendes Proportional-Hazard-Modell

$$\begin{aligned} \lambda(t) &= \lambda_0(t) \exp[\beta_1 Z_{\text{Alter}}(t) + \beta_2 Z_{\text{Geschlecht}} + \beta_3 Z_{\text{Stat}}(t) + \beta_4 Z_{\text{Stufe2}}(t) \\ &+ \beta_5 Z_{\text{Stufe3}}(t) + \beta_6 Z_{\text{Alter}}(t) \times Z_{\text{Geschlecht}} + \beta_7 Z_{\text{Alter}}(t) \times Z_{\text{Stat}}(t) \\ &+ \beta_8 Z_{\text{Geschlecht}} \times Z_{\text{Stat}}(t) + \beta_9 Z_{\text{Stat}}(t) \times Z_{\text{Stufe2}}(t) \\ &+ \beta_{10} Z_{\text{Stat}}(t) \times Z_{\text{Stufe3}}(t) + \beta_{11} Z_{\theta_w}(t) + \beta_{12} Z_{\theta_m}(t) \\ &+ \beta_{13} Z_{\text{Alter}}(t) \times Z_{\theta_w}(t) + \beta_{14} Z_{\text{Alter}}(t) \times Z_{\theta_m}(t)]. \end{aligned} \quad (4.6)$$

Um zu erkennen, ob sich mit Modell (4.6) eine Verbesserung im AIC-Kriterium ergibt wurden wiederum für verschiedene Werte von $\theta_w = 80 + 0.25i$ $i \in \{0, \dots, 80\}$ und $\theta_m = 70 + 0.25i$ $i \in \{0, \dots, 80\}$ der Log-Likelihood berechnet. Dieser wurde maximiert für $\theta_w = 91$ und $\theta_m = 75.25$,

der Log-Likelihood an dieser Stelle hat den Wert 15774.4. Man erkennt eine viel stärkere Ausprägung des Maximums für Frauen. Bei genauerer Analyse des Modells fällt auf, daß die Interaktionen und $Z_{\text{Alter}}(t) \times Z_{\theta_w}(t)$ und $Z_{\text{Alter}}(t) \times Z_{\theta_m}(t)$ kaum signifikant sind (die lokalen Wald-Tests gegen die Nullhypothesen: $\beta_{13} = 0$ bzw. $\beta_{14} = 0$ führen zu p-Werten von 0.92 bzw. 0.37). Auch im AIC erkennt führt die hohe Anzahl an Freiheitsgraden zu einer leichten Verschlechterung im Vergleich zu Modell (4.4) .

Variable	Koeffizient	exp(Koeff.)	SE(Koeff)	z-Statistik	p-Wert
Z_{Alter}	0.0279	1.028	0.0036	7.68	$1.6 \cdot 10^{-14}$
$Z_{\text{Geschlecht}}$	0.7637	2.146	0.3202	2.39.	$1.7 \cdot 10^{-2}$
Z_{Stat}	-0.3724	0.689	0.4035	-0.92	$3.6 \cdot 10^{-1}$
Z_{Stufe2}	0.8101	2.248	0.0730	11.10	$< 10^{-15}$
Z_{Stufe3}	1.7521	5.767	0.0734	23.87	$< 10^{-15}$
Z_{θ_w}	0.2146	1.239	0.0737	2.91	$3.6 \cdot 10^{-3}$
Z_{θ_m}	-0.1401	0.869	0.1041	-1.34	$1.8 \cdot 10^{-1}$
$Z_{\text{Alter}} \times Z_{\text{Geschlecht}}$	-0.0138	0.986	0.0045	-3.08	$2.1 \cdot 10^{-3}$
$Z_{\text{Alter}} \times Z_{\text{Stat}}$	0.0134	1.014	0.0047	2.87	$4.1 \cdot 10^{-3}$
$Z_{\text{Geschlecht}} \times Z_{\text{Stat}}$	-0.4374	0.646	0.0993	-4.40	$4.4 \cdot 10^{-5}$
$Z_{\text{Stat}} \times Z_{\text{Stufe2}}$	-0.3254	0.722	0.1447	-2.25	$2.5 \cdot 10^{-2}$
$Z_{\text{Stat}} \times Z_{\text{Stufe3}}$	-0.6733	0.510	0.1391	-4.84	$1.3 \cdot 10^{-6}$

Tabelle 4.14: Geschätzte Koeffizienten im Modell (4.7)

Um eine Optimierung in unserem Modell zu erreichen, sollten wir daher die Interaktionen zwischen θ_m , θ_w und Alter unberücksichtigt lassen und dann wiederum die optimalen Werte für θ_m und θ_w mittels Maximierung des Log-Likelihood bestimmen, das bedeutet, daß wir den Log-Likelihood für Modell

$$\begin{aligned}
 \lambda(t) &= \lambda_0(t) \exp[\beta_1 Z_{\text{Alter}}(t) + \beta_2 Z_{\text{Geschlecht}} + \beta_3 Z_{\text{Stat}}(t) + \beta_4 Z_{\text{Stufe2}}(t) \\
 &+ \beta_5 Z_{\text{Stufe3}}(t) + \beta_6 Z_{\text{Alter}}(t) \times Z_{\text{Geschlecht}} + \beta_7 Z_{\text{Alter}}(t) \times Z_{\text{Stat}}(t) \\
 &+ \beta_8 Z_{\text{Geschlecht}} \times Z_{\text{Stat}}(t) + \beta_9 Z_{\text{Stat}}(t) \times Z_{\text{Stufe2}}(t) \\
 &+ \beta_{10} Z_{\text{Stat}}(t) \times Z_{\text{Stufe3}}(t) + \beta_{11} Z_{\theta_w}(t) + \beta_{12} Z_{\theta_m}(t)] \\
 &= \lambda_0(t) \exp(\beta^t Z(t))
 \end{aligned} \tag{4.7}$$

in Abhängigkeit von θ_w und θ_m maximieren.

Wie auch in (4.6) wird der Log-Likelihood maximal für $\theta_w = 91$ und $\theta_m = 75.25$. Mit einem Wert von 15874.79 erhalten wir so einen minimal höheren Wert als für Modell (4.6), obwohl wir die Anzahl der Freiheitsgrade von 14 auf 12 reduziert haben. Daraus resultiert eine Verbesserung des AIC-Wertes für dieses Modell (31773.58), so daß wir uns bei weiteren Analysen auf dieses Modell konzentrieren wollen.

4.4.2 Graphische Tests des Propotional Hazard

Um für Modell (4.7) die Abhängigkeit des Hazard von Kovariablen zu analysieren, unterteilen wir den Datensatz in 12 verschiedene Gruppen (Frauen ambulant Stufe 1, Frauen ambulant Stufe 2,..., Männer stationär Stufe 3) und schätzen dann mit des Nelson-Aalen-Schätzer für jede einzelne Gruppe die kumulative Hazard- oder Intensitätsfunktion. Mit diesen Intensitätsfunktionen plotten wir jeweils zwei interessierende Untergruppen gegeneinander (Andersen-Plots, siehe Abschnitt 3.7.9) . Falls die Verteilung der Pflegedauern nach einem Proportional-Hazard-Modell gegeben ist, sollte der Plots ungefähr auf einer Geraden liegen, deren Steigung dem relativen Risiko unter der Proportional-Hazard-Annahme

$$RR_{PH} = \frac{\lambda_0(t) \exp(\hat{\beta}^t Z_{\text{Gruppe 1}})}{\lambda_0(t) \exp(\hat{\beta}^t Z_{\text{Gruppe 2}})} = \frac{\exp(\hat{\beta}^t Z_{\text{Gruppe 1}})}{\exp(\hat{\beta}^t Z_{\text{Gruppe 2}})}$$

entspricht. Um das relative Risiko zu erhalten berechnen wir deshalb mit Modell (4.7) die zugehörigen Multiplikatoren für die Basis-Hazardfunktion (als Werte für die Kovariable Alter verwenden wir das Durchschnittalter in der jeweiligen Untergruppe). Die Verwendung des Mittelwertes für die zeitabhängige Kovariable Z_{Alter} führt approximativ zu einem linearen Zusammenhang der jeweiligen relativen Risiken in den einzelnen Untergruppen.

Wenden wir uns zunächst der Analyse der Kovariablen $Z_{\text{Geschlecht}}$ zu (Abbildung 4.7). Für die 6 Kategorien (ambulant Stufe 1, ... , stationär Stufe 3) wurden hier jeweils die geschätzten kumulativen Hazardfunktionen für Frauen $\hat{\Lambda}_w(t)$ (x-Achse) und Männer $\hat{\Lambda}_m(t)$ (y-Achse) gegeneinander geplottet. Die Gerade im Plot entspricht dem nach dem Proportional-Hazard-Modell (4.7) erwarteten Verhältnis der kumulativen Hazardfunktionen. Man erkennt zunächst folgendes: In Stufe 2 und Stufe 3 entsprechen die für Männer und Frauen mit dem nichtparametrischen Nelson-Aalen-Schätzer ermittelten kumulativen Hazardfunktionen ungefähr den mit dem Proportional-Hazard-Modell geschätzten kumulativen Hazards. In Stufe 1 fällt jedoch auf, daß sowohl im ambulanten, als auch im stationären Bereich die Hazardfunktion nicht besonders gut durch ein Proportional-Hazard-Modell approximiert wird. Bei genauerer Betrachtung des Plot für ambulante Pflege fällt auf, daß bis zu einem bestimmten Zeitpunkt t_0 das relative Risiko für Frauen zu sterben niedriger ist, als nach dem Proportional Hazard-Modell erwartet (der Plot entfernt sich von der Geraden, die das nach dem Proportional-Hazard-Modell erwartete relative Risiko darstellt) und ab t_0 dann höher. Zu welchen Zeitpunkt t_0 sich dieser Trend umkehrt, kann man mit den Andersen-Plots nicht feststellen. Im Bereich der stationären Pflege erkennt man einen umgekehrten Trend, d.h. zunächst ist für Männer das relative Risiko niedriger als für Frauen.

Um diesen Zeitpunkt t_0 festzustellen, an dem sich der Trend umkehrt plotten wir die Differenz der mit dem Nelson-Aalen-Schätzer geschätzten kumulativen Hazards von Männern und Frauen gegeneinander, d.h. wir plotten die Zeit t gegen $\log(\hat{\Lambda}_w(t)) - \log(\hat{\Lambda}_m(t))$ (Abbildung 4.8). Da gilt

$$\lambda(t) = \frac{d}{dt} \Lambda(t)$$

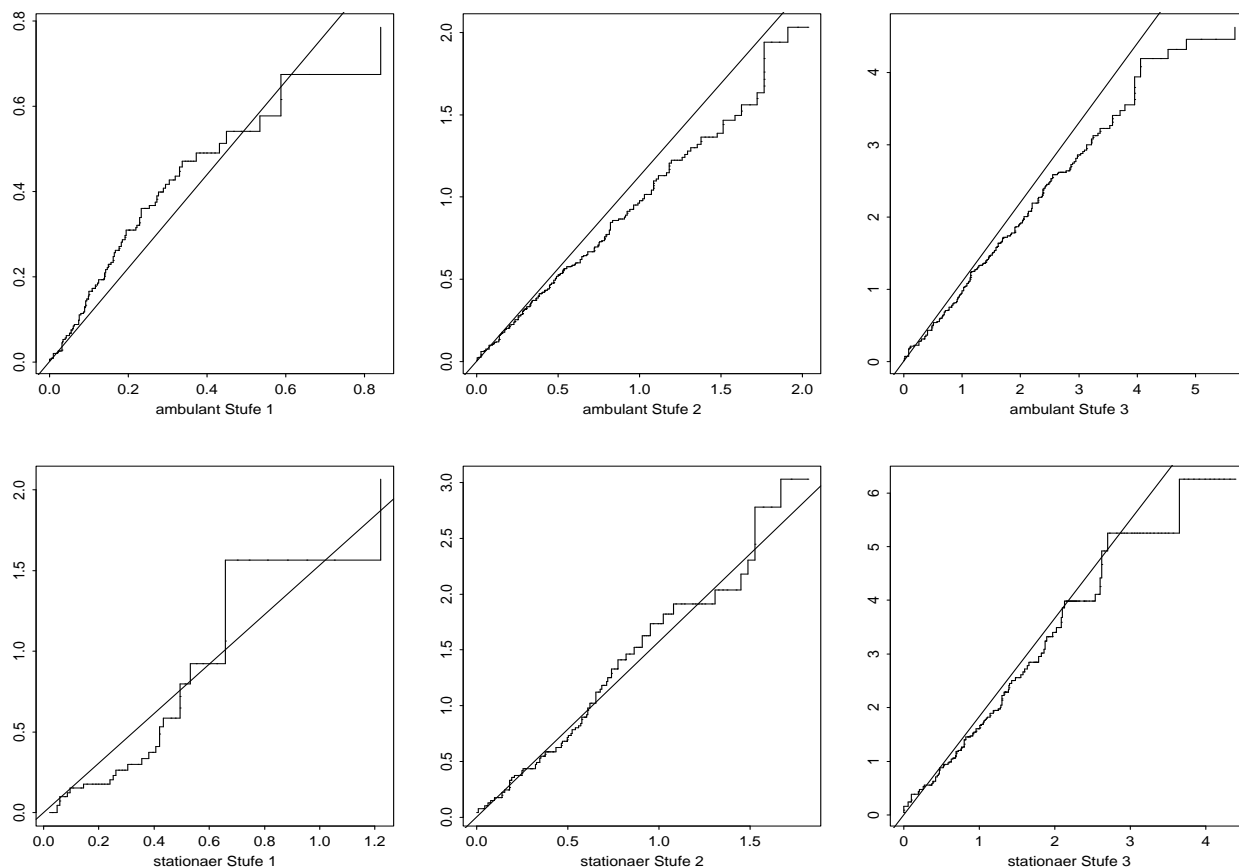


Abbildung 4.7: Andersen-Plots unterteilt nach Geschlecht, x-Achse: Frauen, y-Achse: Männer

und die log-Funktion echt monoton wachsend ist, kann man den Plot folgendermaßen interpretieren. In den monoton fallenden Bereichen ist das im nichtparametrischen Modell beobachtete relative Risiko zwischen Frauen und Männer kleiner, als das nach dem Proportional-Hazard-Modell erwartete. Mit der horizontalen Linie ist die nach dem Proportional-Hazard-Modell geschätzte Differenz $\log \Lambda_w - \log \Lambda_m$ dargestellt. Seien also mit $\hat{\lambda}_w(t)$, bzw. $\hat{\lambda}_m(t)$ die mit dem Nelson-Aalen-Verfahren (nichtparametrisch) geschätzten Hazardfunktionen, dann gilt

$$\frac{\hat{\lambda}_w(t)}{\hat{\lambda}_m(t)} < \frac{\exp(\hat{\beta}^t Z_w)}{\exp(\hat{\beta}^t Z_m)}.$$

Z_w und Z_m sind die zugehörigen Kovariablen-Vektoren einzelnen Untergruppe, zum Beispiel hätte der Kovariablenvektor Z_w für die stationär Pflegebedürftigen der Stufe 1 die Werte

$$\begin{aligned} Z_{\text{Geschlecht}} &= 1, \\ Z_{\text{Stat}} &= 1, \\ Z_{\text{Stufe 1}} &= 0, \\ Z_{\text{Alter}} &= \text{Durchschnittsalter stationär pflegebed. Frauen.} \end{aligned}$$

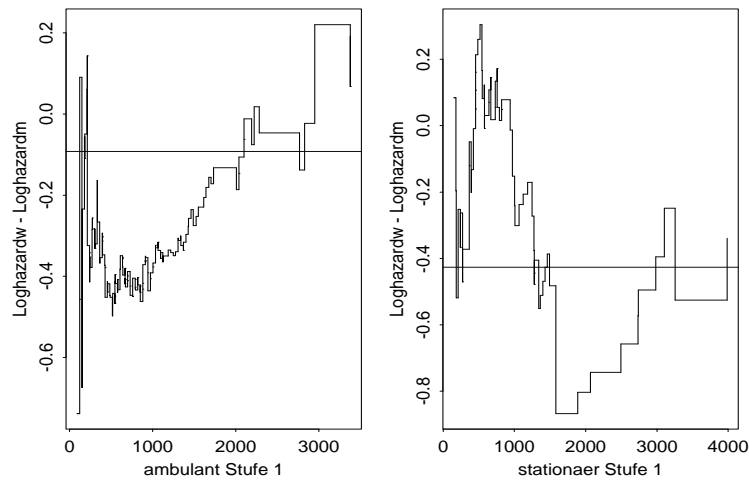


Abbildung 4.8: Differenz des Kumulativen Hazard für Männer, bzw. Frauen

Auch bei den Plots für die Kovariable “Pflegeart” Abbildung 4.9 erkennen wir, daß in Pflegestufe 1 der Andersen-Plot nicht sehr gut der Geraden des nach Modell (4.7) erwarteten relativen Risikos folgt. Hier sieht man, daß für Frauen in Stufe 1 das relative Risiko in ambulanter Pflege zu sterben, zunächst höher ist und sich dann einpendelt, bei Männern sieht man den umgekehrten Trend. In den Stufen 2 und 3 folgen die Andersen-Plots relativ gut den Geraden, die das nach dem Proportional-Hazard-Modell (4.7) erwartete Verhältnis ausdrücken.

Zuletzt betrachten wir noch die Andersen-Plots für die Kovariable Pflegestufe 4.10. Hier wurden für die 4 Untergruppen (Männer ambulant, ..., Frauen stationär) die jeweiligen kumulativen Hazardfunktionen für Stufe 1 (x-Achse) gegen Stufe 2 (y, Achse, gestrichelte Linie), bzw. Stufe 3 (y-Achse, durchgezogene Linie) geplottet. Wir erkennen, daß das Verhältnis des kumulativen Hazards für Stufe 3 zum kumulativen Hazard für Stufe 1 nur sehr unzureichend durch das Cox-Regressionsmodell beschrieben wird. Dies erkennt man, in unterschiedlicher Ausprägung bei allen 4 Plots. Für d.h. gerade zu Beginn der Pflegezeit ist zu beobachten

$$\frac{\hat{\lambda}_{\text{Stufe3}}(t)}{\hat{\lambda}_{\text{Stufe1}}(t)} > \frac{\exp(\hat{\beta}^t Z_{\text{Stufe 3}})}{\exp(\hat{\beta}^t Z_{\text{Stufe 1}})} \quad \text{für kleine } t$$

mit $\hat{\lambda}_{\text{Stufe3}}$ und $\hat{\lambda}_{\text{Stufe1}}$ als den jeweiligen nichtparametrischen Nelson-Aalen-Schätzern für Stufe 3, bzw. Stufe 1. das heißt also gerade in der Zeit kurz nach Pflegebeginn ist in Stufe 3 die Sterbeintensität niedriger als nach dem Modell erwartet. Nach der graphischen Analyse kann man zusammenfassend folgende Feststellung treffen:

Für einige Kovariablen ist das relative Risiko mit dem Proportional-Hazard Modell sehr gut modelliert. Diese sind im einzelnen

- Relatives Risiko Frauen zu Männern (in Stufe 2 und Stufe 3)

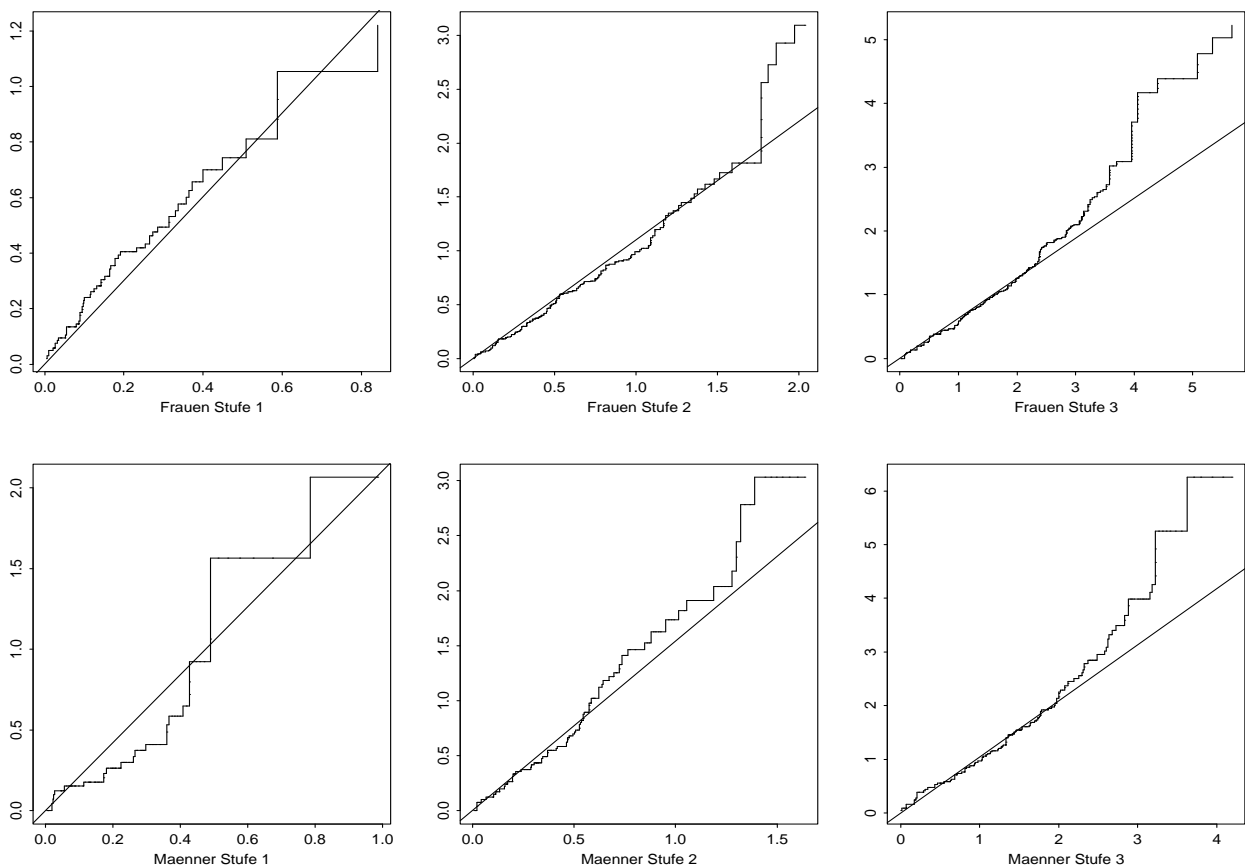


Abbildung 4.9: Andersen-Plots unterteilt nach Pflegeart, x-Achse: ambulant, y-Achse: stationär

- Relatives Risiko ambulante Pflege zu stationärer Pflege (in Stufe 2 und Stufe 3)
- Relatives Risiko Stufe 2 zu Stufe 1

Eine schlechte Proportional-Hazard-Modellierung erhält man für

- Relatives Risiko Frauen zu Männern (in Stufe 1)
- Relatives Risiko ambulante Pflege zu stationärer Pflege (in Stufe 1)
- Relatives Risiko Stufe 3 zu Stufe 1

Man sieht, daß hier noch eine unzureichende Modellierung der Realität vorliegt, allem Anschein nach verändern sich die einige Koeffizienten mit der Pflegedauer (vor allem für die Kovariable Stufe). Diese zeitliche Veränderung von Koeffizienten wollen wir nun mit der Hilfe von Schoenfeld-Residuen (siehe Abschnitt 3.7.9) genauer analysieren

Untersuchung der zeitlichen Abhängigkeit der Regressionskoeffizienten

Um ein Gefühl für eventuelle zeitliche Abhängigkeit von Regressionskoeffizienten von der Pflegedauer t zu bekommen, plotten wir den durch Glättung (mit der S-Plus Funktion `coxph.zph`,

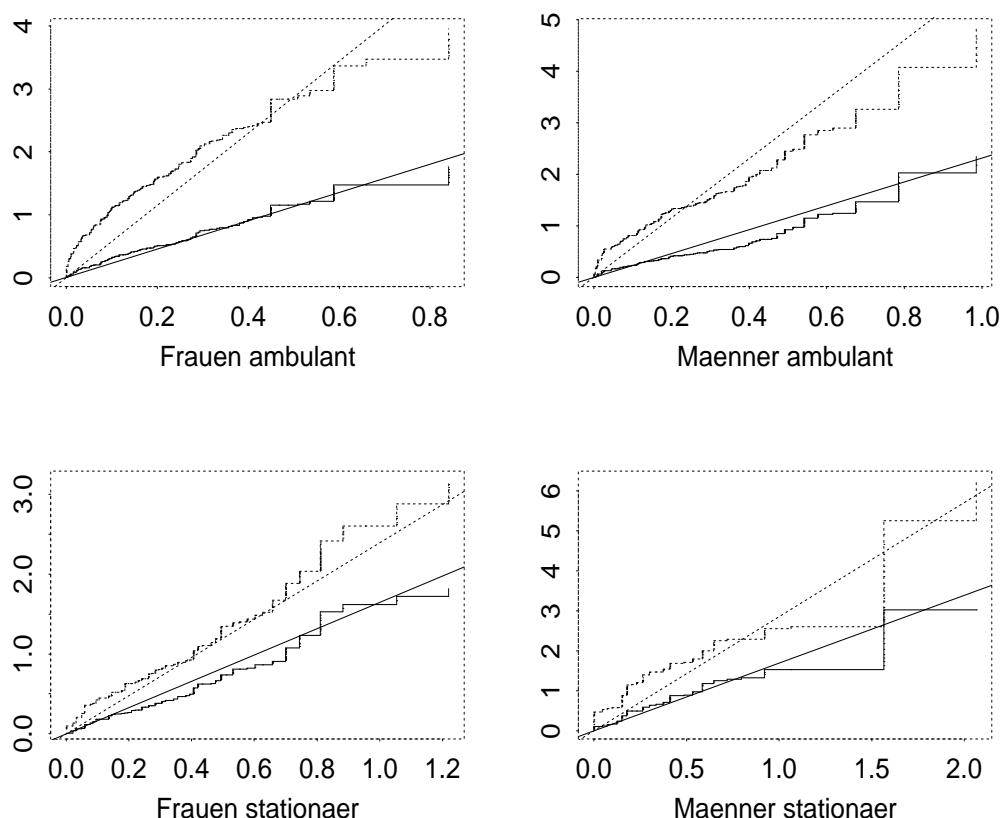


Abbildung 4.10: Andersen-Plots unterteilt nach Pflegestufe, x-Achse: Stufe 1, y-Achse: Stufe 2 (durchgezogene Linie) bzw. Stufe 3 (gestrichelte Linie)

als Glätter verwenden wir einen kubischen Spline mit 40 äquidistanten Stützstellen) geschätzten Erwartungswert von skalierten Schoenfeld-Residuen (3.51) gegen die Pflegedauer t (Abbildung 4.11). Die gestrichelten Linien sind 95%-Konfidenzbänder für den Erwartungswert. Der so erhaltene Plot gibt ein Indiz für die qualitativen Entwicklung des Regressionskoeffizienten β in Abhängigkeit von der Pflegedauer.

Man erkennt hier, daß für fast alle Kovariablen im Koeffizienten β eine starke Zeitabhängigkeit festzustellen ist (vor allem in den Zeitbereichen zwischen 0 und 2000 Tagen, also in den Bereichen, in denen viele Beobachtungen vorliegen). Insbesondere erkennt man in diesem Zeitintervall eine “Trendwende”, d.h. viele der Kurven haben im Zeitbereich von ca. 1000 Tagen ein lokales Maximum, bzw. Minimum. Betrachtet man zum Beispiel die Plots für die Kovariable für Stufe 2 (rechts, oben) und Stufe 3 (links, 2. von oben), sieht man, daß das relative Risiko zu sterben (im Vergleich zu Stufe 1) am Anfang viel höher ist und dann fällt, bis bei einem Zeitpunkt von ca. 1000 Tagen ein lokales Minimum erreicht ist. Danach erkennt man nur noch eine leichte Veränderung des Koeffizienten. Mit folgenden 2 Schritten versuchen wir nun das Modell noch

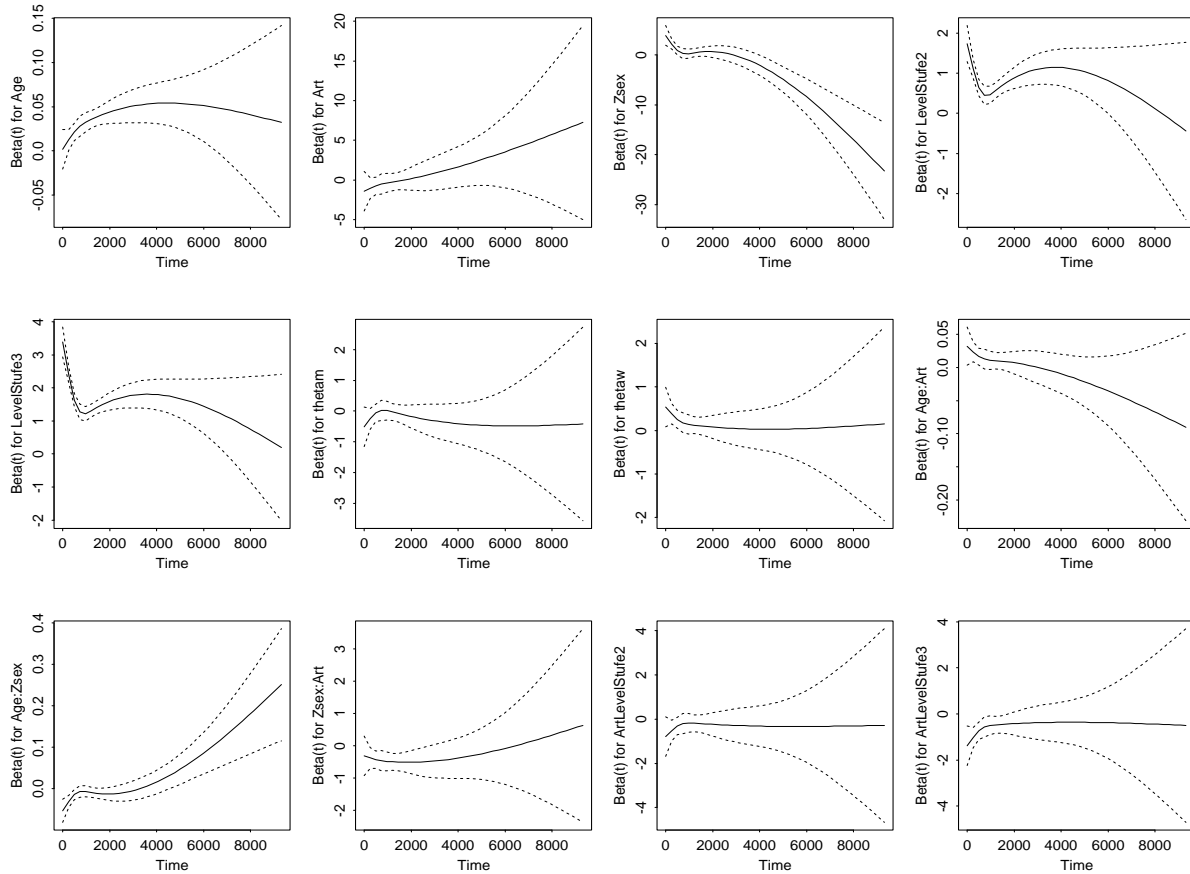


Abbildung 4.11: Geglättete Schoenfeld-Residuen zur Bestimmung der funktionalen Abhängigkeit der Regressionskoeffizienten

zu verbessern:

- Um den Zeitpunkt einer “globalen Trendwende” t_0 zu finden, betrachten wir zunächst ein Modell, in dem für Pflegedauern $T \leq t_0$ und $T > t_0$ separat Regressionskoeffizienten geschätzt werden. Dieses Modell kann folgendermaßen dargestellt werden:

$$\lambda(t) = \lambda_0(t) \exp[\beta^t Z(t) + Z_{\leq t_0}(t)(\alpha^t Z(t))].$$

Hier sind

$$Z_{\leq t_0} = \begin{cases} 1 & t \leq t_0 \\ 0 & \text{sonst} \end{cases},$$

$$Z(t) \in \mathbb{R}^{12}$$

der Kovariablenvektor aus Modell (4.7) und

$$\alpha, \beta \in \mathbb{R}^{12}$$

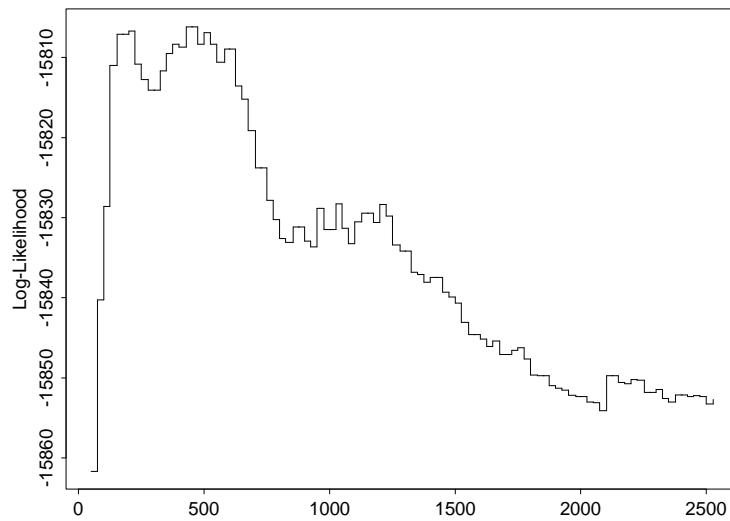


Abbildung 4.12: Entwicklung des Log-Likelihood für verschiedene Werte von t_0 bei Schätzung von separaten Koeffizienten für Pflegedauern $T \leq t_0$ und $T > t_0$

die zu schätzenden Parameter. Für dieses Modell maximieren wir den partiellen Log-Likelihood in Abhängigkeit von t_0 .

- In einem zweiten Schritt schätzen wir differenziert für jede einzelne Kovariablen aus Modell (4.7) separate Koeffizienten für Pflegedauern $T \leq t_0$ und $T > t_0$ um damit die Koeffizienten zu bestimmen, bei denen die größte Abhängigkeit von der Pflegedauer vorliegt. Zum Beispiel erhalten wir für die Kovariable Alter folgende Darstellung der Hazardfunktion:

$$\lambda(t) = \lambda_0(t) \exp[\gamma Z_{\text{Alter}}(t) \times Z_{\leq t_0}(t) + \beta^t Z(t)].$$

Wiederum sind $Z(t)$ der Vektor aus Modell (4.7) und

$$\gamma \in \mathbb{R}, \beta \in \mathbb{R}^{12}$$

die zu schätzenden Parameter.

Bemerkung: Eine zeitstetige Modellierung von Regressionskoeffizienten (z.B. in Form einer linearen Abhängigkeit $\beta(t) = \beta_0 + \beta_1 t$) ist mit dem Proportional-Hazard-Schätzer generell möglich, mit dem Paket `coxph` aus S-Plus nicht zu realisieren. Um dies zu ermöglichen müsste die Designmatrix an jedem Zeitpunkt, an dem ein Ereignis beobachtet wird, neu berechnet werden. Für kleine Datensätze kann man dies noch erreichen, indem man den Datensatz aufsplittet und an jedem Beobachtungspunkt einen neuen Kovariablenvektor in Abhängigkeit von der Beobachtungszeit definiert (mit dieser Technik wurden auch die zeitabhängigen Kovariablen für Pflegestufe und Pflegeart modelliert, hier ist die Zeitabhängigkeit jedoch in Form einer stückweise konstanten Sprungfunktion vorgelegen).

In Abbildung 4.12 sehen wir zunächst, wie sich der Log-Likelihood für verschiedene Werte von t_0 verhält. Man erkennt ein globales Maximum für $t_0 = 450$ Tage. Der Wert des Log-Likelihood beträgt an dieser Stelle -15806.16, was eine wesentliche Verbesserung im Vergleich zu Modell (4.7) (mit einem Wert von 15874.79) ist.

Um zu erkennen, in welcher Kovariablen die stärkste zeitliche Variation im Koeffizienten vorliegt, wurde für jeden einzelnen Koeffizienten aus Modell (4.7) ein separater Koeffizient für Pflegedauern $T \leq t_0$, sowie $T > t_0$ berechnet und gegen t_0 geplottet (Abbildung 4.13)

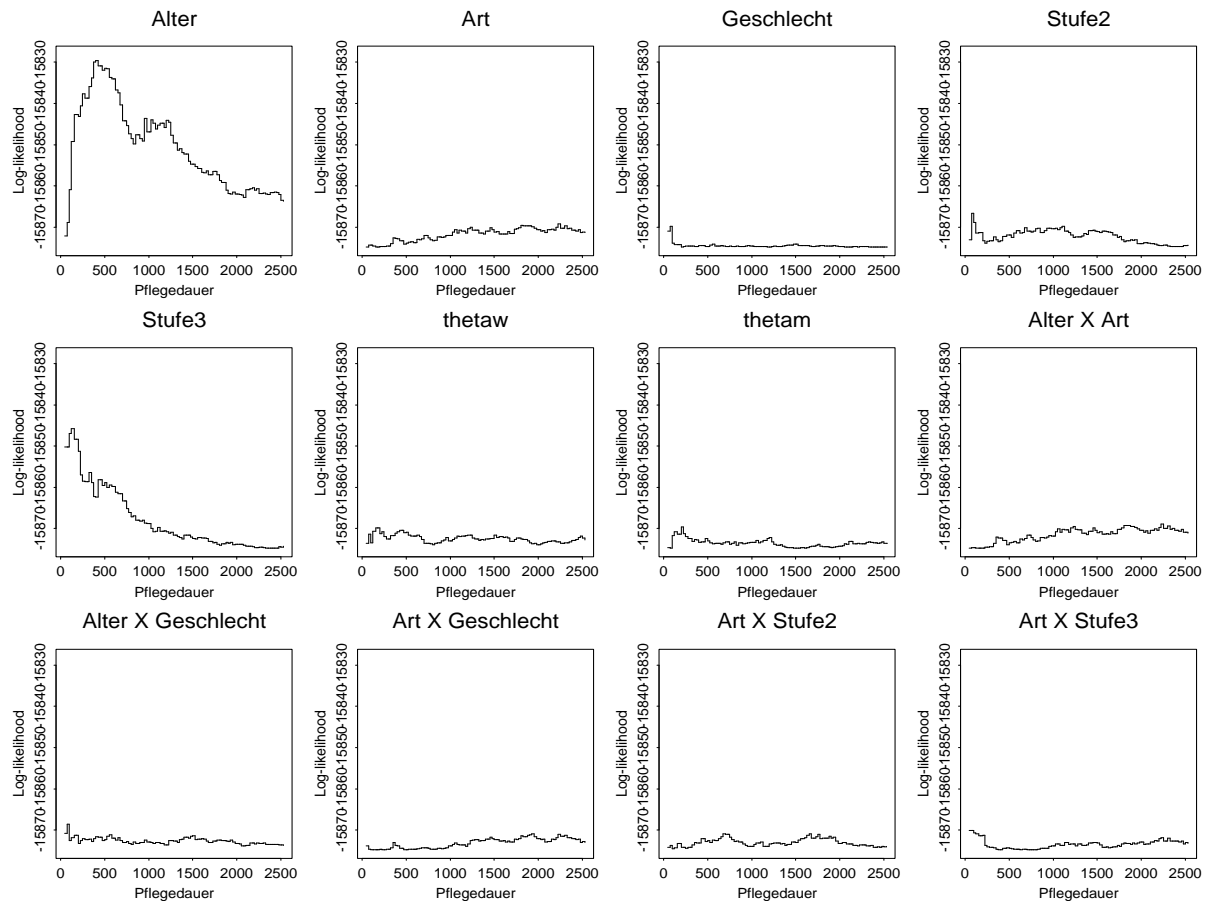


Abbildung 4.13: Entwicklung des Log-Likelihood für verschiedene Werte von t_0 bei Schätzung von separaten Koeffizienten für Pflegedauern $T \leq t_0$ und $T > t_0$

In Abbildung 4.13 erkennen wir vor allem für die Kovariable Alter, sowie für Pflegestufe 2 und 3 eine starke Verbesserung des Log-Likelihood bei der Schätzung der jeweiligen Koeffizienten für $T \leq t_0$ und $T > t_0$. In Tabelle 4.15 erkennt man alle 3 Kovariablen eine signifikante Verbesserung im Log-Likelihood (hinsichtlich des p-Wert für den lokalen Likelihood-Ratio-Test)

Daher modifizieren wir das Modell (4.7) und schätzen für die Kovariablen Alter und Pflegestufe Koeffizienten separat für $T \leq t_0$ und $T > t_0$. Technisch wurde dies erreicht, indem wir zu den

Kovariablen	t_0	LL (Modell für $T \leq t_0, T > t_0$)	LL (Modell(4.7))	p-Wert
Z_{Alter}	400	-15829.59	-15874.79	$< 10^{-15}$
Z_{Stufe2}	75	-15866.69	-15874.79	$5.6 \cdot 10^{-5}$
Z_{Stufe3}	125	-15845.82	-15874.79	$2.7 \cdot 10^{-14}$

Tabelle 4.15: Maximum des Log-Likelihood für separate Schätzung der Kovariablen für Alter, Stufe 1 und Stufe 2 (unterteilt nach Pflegedauern $T \leq t_0$ und $T > t_0$)

Koeffizienten in (4.7) drei weitere Kovariablen $Z_{\leq 75}$, $Z_{\leq 125}$, sowie $Z_{\leq 400}$ mit

$$Z_{\leq t_0}(t) = \begin{cases} 1 & t \leq t_0 \\ 0 & \text{sonst} \end{cases}$$

eingeführt und deren Interaktion mit den jeweiligen Variablen für Alter, Stufe 2 und Stufe 3 betrachtet. Für dieses Modell verbessert sich der Log-Likelihood auf 15801.4. Man erkennt

Variable	Koeffizient	exp(Koeff.)	SE(Koeff)	z-Statistik	p-Wert
Z_{Alter}	0.0401	1.041	0.0041	9.65	$< 10^{-15}$
$Z_{\text{Alter}} \times Z_{\leq 400}$	-0.0327	0.968	0.0035	-9.47	$< 10^{-15}$
$Z_{\text{Geschlecht}}$	0.7986	2.222	0.3214	2.49	0.01
Z_{Stat}	-0.0531	0.948	0.4004	-0.13	0.89
Z_{Stufe2}	0.8011	2.228	0.0733	10.93	$< 10^{-15}$
$Z_{\text{Stufe2}} \times Z_{\leq 75}$	-0.8080	0.446	0.4982	-1.62	0.10
Z_{Stufe3}	1.6420	5.165	0.0752	21.83	$< 10^{-15}$
$Z_{\text{Stufe3}} \times Z_{\leq 125}$	1.135	3.113	0.1828	6.21	$5.2 \cdot 10^{-10}$
Z_{θ_w}	0.1815	1.199	0.0743	2.443	0.02
Z_{θ_m}	-0.1206	0.886	0.1052	-1.15	0.25
$Z_{\text{Alter}} \times Z_{\text{Geschlecht}}$	-0.0141	0.986	0.0045	-3.13	$1.8 \cdot 10^{-3}$
$Z_{\text{Alter}} \times Z_{\text{Stat}}$	0.0093	1.009	0.0047	1.99	0.05
$Z_{\text{Geschlecht}} \times Z_{\text{Stat}}$	-0.4257	0.653	0.0996	-4.27	$1.9 \cdot 10^{-5}$
$Z_{\text{Stat}} \times Z_{\text{Stufe2}}$	-0.3420	0.710	0.1448	-2.36	0.02
$Z_{\text{Stat}} \times Z_{\text{Stufe3}}$	-0.6533	0.520	0.1393	-4.69	$2.7 \cdot 10^{-6}$

Tabelle 4.16: Geschätzte Koeffizienten im Modell mit separaten Koeffizienten für $T \leq t_0$ und $T > t_0$ für Alter und Stufe

speziell für Pflegestufe 3, daß gerade die ersten 400 Tage (also ungefähr das erste Jahr) das Risiko zu sterben im Vergleich zu Pflegestufe 1 extrem hoch ist. Wie schon vorher erwähnt, würde die Modellierung stetig von der Zeit abhängiger Koeffizienten eine Verbesserung des Modells ergeben. Aus vorher genannten Gründen war eine Schätzung von Koeffizienten für diese Modellierung aufgrund der Größe des Datensatzes nicht zu realisieren. Daher möchte ich an die

ser Stelle mit der Analyse der Sterbeintensitäten enden und später, bei der Entwicklung eines Versicherungsmodells, die in Tabelle 4.16 dargestellten Ergebnisse verwenden.

4.5 Modellierung der Zustandsübergänge

Bis jetzt haben wir uns nur intensiv mit dem Ereignis “Tod des Pflegebedürftigen” und dessen Abhängigkeit von verschiedenen Einflußfaktoren beschäftigt. In einem weiteren Schritt wollen wir nun die Übergänge zwischen den verschiedenen Pflegearten und -stufen genauer untersuchen. Mit Hilfe zweier verschiedener Modelle sollen der Datensatz hinsichtlich folgende Fragestellungen untersucht werden:

- Fluktuation im Bestand hinsichtlich Pflegeart (ambulant, stationär)
- Fluktuation im Bestand hinsichtlich Pflegestufe

Mit Hilfe der in Abschnitt 2.5 entwickelten Theorie sollen für diese Modelle Übergangintensitäten, bzw. Wahrscheinlichkeiten hergeleitet werden.

4.5.1 Modellierung der Zustandsübergänge zwischen ambulanter und stationärer Pflege

Im folgenden soll nun zunächst Markov-Modell mit den 3 verschiedenen Zuständen: ambulante Pflege, stationäre Pflege und Tod, betrachtet werden. Die Markoveigenschaft spiegelt sich in der Annahme wieder, daß Übergangintensitäten $\lambda(t)$ nur von der Pflegedauer t und vom Zustand, in dem sich die pflegebedürftige Person zum Zeitpunkt t befindet und einem Kovariablenvektor $Z(t)$ (Alter, Pflegestufe, Geschlecht und eventuell Interaktionen) nicht jedoch von der bisherigen Pflegehistorie abhängen. Ein Beispiel für die Modellierung eines Markov-Modells für Knochen-transplantationsdaten findet man zum Beispiel in dem dem Artikel von Klein und Keiding (1994) [24].

In einem ersten Schritt sollen die Übergangintensitäten $\lambda_{ij}(t)$, $i \in \{1, 2, 3\}$, $j \in \{1, 2, 3\}$ zu den in Abbildung 4.14 schematisch dargestellten Zustandsübergänge geschätzt werden. In einem zweiten Schritt werden dann mit Hilfe der geschätzten Intensitäten, die Übergangswahrscheinlichkeiten $P_{ij}(z, t)$ $z \leq t$ berechnet.

Bei der Schätzung der Intensitäten konzentrieren wir uns vor allem auf die Übergangintensitäten zwischen ambulanter und stationärer Pflege, da wir für die Sterbeintensitäten $\lambda_{13}(t)$, bzw. $\lambda_{23}(t)$ eine ausführliche Untersuchung im vorherigen Abschnitt durchgeführt haben. Insgesamt wurden 590 Übergänge von ambulanter zu stationärer Pflege und 28 Übergänge von stationärer zu ambulanter Pflege beobachtet.

Mit Hilfe der Zählprozeßnotation für den Proportional-Hazard können wir zur Schätzung dieser

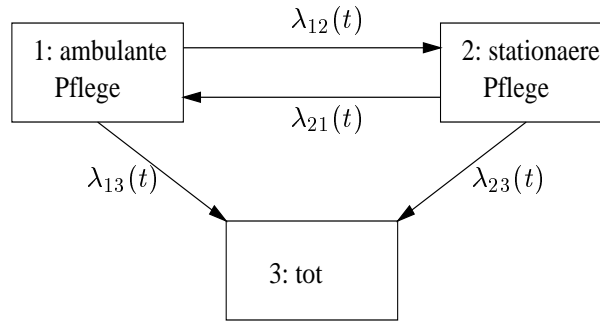


Abbildung 4.14: Schematische Übersicht der Zustandsübergänge ambulant stationär

Intensitäten ein Proportional-Hazard-Modell fitten. Dabei interessieren uns folgende Zählprozesse

$$N_{12}(t) = \# \text{ Übergänge von ambulanter zu stationärer Pflege bis zum Zeitpunkt } t,$$

$$N_{21}(t) = \# \text{ Übergänge von stationärer zu ambulanter Pflege bis zum Zeitpunkt } t,$$

für die wir die Kompensator-, oder kumulative Hazardfunktion

$$\Lambda_{ij}(t) = \int_0^t \lambda_{ij}(u) du \quad i, j \in \{1, 2\} \quad i \neq j$$

in einem Proportional-Hazard-Modell schätzen. Zunächst gehen wir davon aus, daß λ_{ij} folgende Darstellung hat

$$\begin{aligned} \lambda_{ij}(t) = & \lambda_{ij0}(t) \exp[\beta_1 Z_{\text{Alter}}(t) + \beta_2 Z_{\text{Geschlecht}} + \beta_3 Z_{\text{Stufe } 2}(t) \\ & + \beta_4 Z_{\text{Stufe } 3}(t)] \quad i, j \in 1, 2 \quad i \neq j, \end{aligned} \quad (4.8)$$

mit einem unspezifizierten Basis-Hazard $\lambda_{ij0}(t)$. Für die Übergangintensität $\lambda_{12}(t)$ von stationärer zu ambulanter Pflege erhält man die in Tabelle 4.17 dargestellten Koeffizienten. Bei einem Signifikanzniveau von 5% haben alle Kovariablen bis auf $Z_{\text{Stufe } 3}$ (der p-Wert liegt hier bei 0.19) signifikanten Einfluß auf das Modell. Dieses Modell werden wir später noch genauer hinsichtlich Interaktionen zwischen Kovariablen untersuchen. Für die Übergänge von stationärer zu ambu-

Variable	Koeffizient	exp(Koeff.)	SE(Koeff.)	z-Statistik	p-Wert
Z_{Alter}	0.0305	1.03	0.0038	7.99	$1.2 \cdot 10^{-15}$
$Z_{\text{Geschlecht}}$	0.3842	1.47	0.0956	4.02	$5.9 \cdot 10^{-5}$
$Z_{\text{Stufe } 2}$	0.3191	1.38	0.0912	3.50	$4.6 \cdot 10^{-4}$
$Z_{\text{Stufe } 3}$	0.1661	1.18	0.1279	1.30	$1.9 \cdot 10^{-1}$

Tabelle 4.17: Geschätzte Koeffizienten für die Übergangintensität λ_{12}

lanter Pflege kann man für keine Kovariable einen signifikanten Einfluß feststellen. Auch der p-Wert der Likelihood-Ratio-Statistik für den Gesamteinfluß der Kovariablen von 0.192 (Wert

der Teststatistik ist 6.09 bei χ^2 Verteilung mit 4 Freiheitsgraden) ist relativ hoch. Ein Grund dafür, daß kein signifikanter Proportional-Hazard erkennbar ist, ist wohl die extrem niedrige Anzahl an Übergängen von stationärer zu ambulanter Pflege.

Mit Hilfe der Kolmogorov-Differentialgleichungen können wir nun die Übergangswahrscheinlichkeiten herleiten. Die Anwendung von (2.23) führt zu folgendem Differentialgleichungssystem:

$$\begin{aligned}
\frac{d}{dt}P_{11}(z, t) &= P_{12}(z, t)\lambda_{21}(t) - P_{11}(z, t)(\lambda_{12}(t) + \lambda_{13}(t)), \\
\frac{d}{dt}P_{12}(z, t) &= P_{11}(z, t)\lambda_{12}(t) - P_{12}(z, t)(\lambda_{21}(t) + \lambda_{23}(t)), \\
\frac{d}{dt}P_{13}(z, t) &= P_{11}(z, t)\lambda_{13}(t) + P_{12}(z, t)\lambda_{23}(t), \\
\frac{d}{dt}P_{21}(z, t) &= P_{22}(z, t)\lambda_{21}(t) - P_{21}(z, t)(\lambda_{12}(t) + \lambda_{13}(t)), \\
\frac{d}{dt}P_{22}(z, t) &= P_{21}(z, t)\lambda_{12}(t) - P_{22}(z, t)(\lambda_{21}(t) + \lambda_{23}(t)), \\
\frac{d}{dt}P_{23}(z, t) &= P_{22}(z, t)\lambda_{23}(t) + P_{21}(z, t)\lambda_{13}(t).
\end{aligned} \tag{4.9}$$

Die Differentialgleichungen für die Verweilwahrscheinlichkeiten lauten:

$$\begin{aligned}
\frac{d}{dt}P_{\underline{11}}(z, t) &= -P_{\underline{11}}(z, t)(\lambda_{12}(t) + \lambda_{13}(t)), \\
\frac{d}{dt}P_{\underline{22}}(z, t) &= -P_{\underline{22}}(z, t)(\lambda_{21}(t) + \lambda_{23}(t)).
\end{aligned} \tag{4.10}$$

Eine explizite Lösung des Differentialgleichungssystems (4.9) ist nicht möglich. Da wir jedoch festgestellt haben, daß es kaum Stufenübergänge von stationärer zu ambulanter Pflege gibt, wollen wir im Folgenden die Übergänge von stationärer zu ambulanter Pflege vernachlässigen (d.h. wir gehen davon aus, daß $\lambda_{21}(t) = 0$ für alle t). Die Differentialgleichungen (4.9) vereinfachen sich somit zu

$$\begin{aligned}
\frac{d}{dt}P_{11}(z, t) &= -P_{11}(z, t)(\lambda_{12}(t) + \lambda_{13}(t)), \\
\frac{d}{dt}P_{12}(z, t) &= P_{11}(z, t)\lambda_{12}(t) - P_{12}(z, t)\lambda_{23}(t), \\
\frac{d}{dt}P_{13}(z, t) &= P_{11}(z, t)\lambda_{13}(t) + P_{12}(z, t)\lambda_{23}(t), \\
\frac{d}{dt}P_{22}(z, t) &= -P_{22}(z, t)(\lambda_{21}(t) + \lambda_{23}(t)), \\
\frac{d}{dt}P_{23}(z, t) &= P_{22}(z, t)\lambda_{23}(t).
\end{aligned} \tag{4.11}$$

Insbesondere werden die Zustände 2 und 3 zu strikt transienten Zuständen, d.h.

$$\begin{aligned}
P_{\underline{11}}(z, t) &= P_{11}(z, t), \\
P_{\underline{22}}(z, t) &= P_{22}(z, t).
\end{aligned}$$

Diese Differentialgleichungen lassen sich einfach lösen. Für die Verweilwahrscheinlichkeiten in Stufe 1 erhält man eine Lösung durch Integration

$$\int_z^t \frac{d}{ds} \frac{P_{11}(z, s)}{P_{11}(z, s)} ds = - \int_z^t [\lambda_{12}(s) + \lambda_{13}(s)] ds,$$

was äquivalent ist zu

$$\log(P_{11}(z, t)) - \underbrace{\log(P_{11}(z, z))}_{=0} = - \int_z^t [\lambda_{12}(s) + \lambda_{13}(s)] ds,$$

so daß man mit

$$P_{11}(z, t) = \exp \left(- \int_z^t [\lambda_{12}(s) + \lambda_{13}(s)] ds \right)$$

eine Lösung erhält. Analoges Vorgehen ergibt für die Verweilwahrscheinlichkeit in Stufe 2 $P_{22}(z, t)$

$$P_{22}(z, t) = \exp \left(- \int_z^t \lambda_{23}(s) ds \right).$$

Um die Übergangswahrscheinlichkeit $P_{12}(z, t)$ zu erhalten, müssen wir die Differentialgleichung

$$\frac{d}{dt} P_{12}(z, t) = -\lambda_{23} P_{12}(z, t) + P_{11}(z, t) \lambda_{12}(t)$$

mittels Variation der Konstanten (z.B. in Königsberger (1992)[20] Seite 270) lösen. Zuerst bestimmen wir die Lösungen $y(z, t)$ des homogenen Systems

$$\frac{d}{dt} y(z, t) = -\lambda_{23}(t) y(z, t),$$

die gegeben sind durch

$$y(z, t) = c \exp \left(- \int_z^t \lambda_{23}(s) ds \right) \quad c \in \mathbb{R}.$$

Eine partikuläre Lösung der inhomogenen Differentialgleichung ist

$$P_{12}(z, t) = \int_z^t P_{11}(z, s) \lambda_{12}(s) \exp \left(\int_z^t \lambda_{23}(u) du \right) ds$$

$$\stackrel{\text{Fubini}}{=} \left[\int_z^t P_{11}(z, s) \lambda_{12}(s) ds \right] \left[\exp \left(\int_z^t \lambda_{23}(u) du \right) \right],$$

so daß wir nun alle Lösungen der Differentialgleichung angeben können

$$P_{12}(z, t) = \left[\left\{ \int_z^t P_{11}(z, s) \lambda_{12}(s) ds \right\} \left\{ \exp \left(\int_z^t \lambda_{23}(u) du \right) \right\} + c \right] \exp \left(- \int_z^t \lambda_{23}(s) ds \right)$$

Unter Berücksichtigung der Randbedingung

$$\lim_{dz \rightarrow 0} \frac{d}{dt} P_{12}(z, z + dz) = \lim_{dz \rightarrow 0} \underbrace{P_{11}(z, z + dz)}_{\rightarrow 1} \lambda_{12}(z + dz) - \underbrace{P_{12}(z, z + dz)}_{\rightarrow 0} \lambda_{23}(z + dz) = \lambda_{12}(z),$$

die für $c = 0$ erfüllt ist, erhält man als Lösung

$$P_{12}(z, t) = \int_z^t P_{11}(z, u) \lambda_{12}(u) P_{22}(u, t) du.$$

Die Übergangswahrscheinlichkeiten für P_{23} erhalten wir durch Integration

$$\begin{aligned}
 P_{23}(z, t) &= \int_z^t P_{22}(z, u) \lambda_{23}(u) du \\
 &= \int_z^t \underbrace{\exp\left(-\int_z^u \lambda_{23}(s) ds\right)}_{\text{Stammfunktion: } \exp\left(-\int_z^u \lambda_{23}(s) ds\right)} \lambda_{23}(u) \\
 &= -\left[\exp\left(-\int_z^u \lambda_{23}(s) ds\right) - 1\right] = 1 - P_{22}.
 \end{aligned}$$

Zuletzt erhalten die verbleibende Übergangswahrscheinlichkeit $P_{13}(z, t)$ als Differenz

$$P_{13}(z, t) = 1 - P_{11}(z, t) - P_{12}(z, t).$$

Kommen wir nun noch einmal auf die Intensitäten für den Übergang von ambulanter zu stationärer Pflege zurück. Auch hier wollen durch weitere Analysen unsere Modellwahl optimieren.

Modellierung eventuell vorhandener Interaktionen

Wie auch bei der Herleitung der Sterbeintensitäten sind wir auch bei den Übergängen von ambulanter zu stationärer Pflege daran interessiert, ob signifikante Interaktionen zwischen den verbleibenden Kovariablen Z_{Alter} , $Z_{\text{Geschlecht}}$ und Z_{Stufe} bestehen. Dazu schauen wir uns die Entwicklung des Log-Likelihood (bzw. das AIC) für ein Modell mit allen Interaktionen im Vergleich zum Modell ohne Interaktionen an. (Tabelle 4.5.1)

Modell	Log-Likelihood	df	AIC ($n = 2$)
ohne Interaktion	-4080.10	4	8168.20
alle Interaktionen	-4078.44	11	8178.88

Tabelle 4.18: AIC-Kriterium für Modell ohne Interaktion und Modell mit allen Interaktionen

Hier sieht man, daß die Hinzunahme von Interaktionen den Log-Likelihood nur minimal (um 1.74) erhöht, so daß es wohl nicht sinnvoll ist ein um Interaktionen erweitertes Modell zu betrachten.

4.5.2 Modellierung der Zustandsübergänge zwischen den verschiedenen Pflegestufen

Analog zum Vorgehen in Abschnitt 4.5.1 betrachten wir hier ein Markovmodell, das die Zustandsübergänge für die verschiedenen Pflegestufen beschreibt.

Insgesamt wurden 1334 Stufenübergänge beobachtet, deren empirische Häufigkeit in Tabelle 4.19

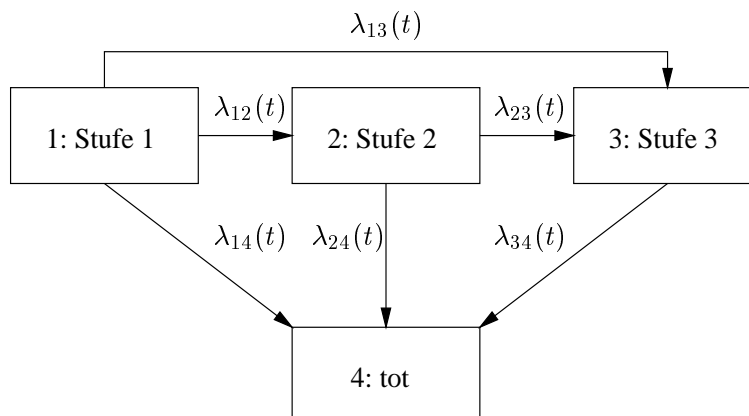


Abbildung 4.15: Schematische Übersicht der Zustandsübergänge zwischen den verschiedenen Pflegestufen

dargestellt ist.

	Stufe 1	Stufe 2	Stufe 3
Stufe 1	–	624	135
Stufe 2	64	–	472
Stufe 3	5	34	–

Tabelle 4.19: Anzahl der Stufenübergänge im Datensatz

Insgesamt konnten wir bei 1231 Übergängen eine Verschlechterung (d.h. eine Erhöhung des Grades an Pflegebedürftigkeit), nur 103 Pflegebedürftige wechselten in eine bessere Pflegestufe. Man erkennt also auch hier einen wesentlichen Trend zur Verschlechterung. Daher sollen bei unserem Modell wieder die vereinfachende Annahme getroffen werden, daß die Übergangintensitäten $\lambda_{31}(t) = \lambda_{21}(t) = \lambda_{32}(t) = 0$ sind. Wiederum schätzen wir die kumulativen Übergangintensitäten $\int_0^t \lambda_{12}(s) ds$, $\int_0^t \lambda_{13}(s) ds$ und $\int_0^t \lambda_{23}(s) ds$ als Kompensatoren der Zählprozesse:

$$\begin{aligned}
 N_{12}(t) &= \# \text{ Übergänge von Pflegestufe 1 zu Pflegestufe 2 bis zum Zeitpunkt } t, \\
 N_{13}(t) &= \# \text{ Übergänge von Pflegestufe 1 zu Pflegestufe 3 bis zum Zeitpunkt } t, \\
 N_{23}(t) &= \# \text{ Übergänge von Pflegestufe 2 zu Pflegestufe 3 bis zum Zeitpunkt } t
 \end{aligned}$$

und nehmen wiederum eine Abhängigkeit in den Kovariablen $Z_{\text{Geschlecht}}$, Z_{Alter} , Z_{Stat} nach einem Proportional-Hazard-Modell an, d.h.

$$\lambda_{ij}(t) = \lambda_{ij0} \exp(\beta^t Z(t)) \quad i, j \in \{1, 2, 3\} \quad i < j, \quad (4.12)$$

mit

$$\beta^t Z(t) = \beta_1 Z_{\text{Alter}}(t) + \beta_2 Z_{\text{Geschlecht}} + \beta_3 Z_{\text{Stat}}(t).$$

Mit dem Ansatz (4.12) erhalten wir die in den Tabellen 4.20 - 4.22 aufgeführten Koeffizienten.

Variable	Koeffizient	exp(Koeff.)	SE(Koeff)	z-Statistik	p-Wert
Z_{Alter}	0.0206	1.02	0.0032	6.041	$1.5 \cdot 10^{-9}$
$Z_{\text{Geschlecht}}$	0.0007	1.00	0.0909	0.007	0.99
Z_{Stat}	0.0363	1.04	0.1076	0.338	0.74

Tabelle 4.20: Geschätzte Koeffizienten für die Übergangsintensität λ_{12}

Variable	Koeffizient	exp(Koeff.)	SE(Koeff)	z-Statistik	p-Wert
Z_{Alter}	0.035	1.04	0.0087	3.98	$7 \cdot 10^{-5}$
$Z_{\text{Geschlecht}}$	-0.368	0.70	0.1873	-1.97	0.05
Z_{Stat}	0.1452	1.16	0.2234	0.65	0.52

Tabelle 4.21: Geschätzte Koeffizienten für die Übergangsintensität λ_{13}

Variable	Koeffizient	exp(Koeff.)	SE(Koeff)	z-Statistik	p-Wert
Z_{Alter}	0.0106	1.01	0.0032	3.33	$8.8 \cdot 10^{-4}$
$Z_{\text{Geschlecht}}$	0.1735	0.84	0.0996	-1.74	0.08
Z_{Stat}	0.1271	0.88	0.1109	-1.15	0.25

Tabelle 4.22: Geschätzte Koeffizienten für die Übergangsintensität λ_{23}

Modellierung von Interaktionen

Um zu erkennen, ob Interaktionen signifikanten Einfluß auf das Modell haben, betrachten wir jeweils einzeln für die Übergangsintensitäten $\lambda_{12}(t)$, $\lambda_{13}(t)$, $\lambda_{23}(t)$ die Entwicklung des Log-Likelihood im Modell. Die Ergebnisse sind in Tabelle 4.23 aufgeführt.

Wir erkennen, daß sich nur bei der Schätzung der Übergangsintensitäten $\lambda_{12}(t)$ eine signifikante Verbesserung des Log-Likelihoods und damit auch des AIC-Kriterium beobachtet werden kann. Daher fitten wir nun ein Cox-Regressionsmodell für $\lambda_{12}(t)$ in dem alle Interaktionen

	Modell	Log-Likelihood	df	AIC ($n = 2$)
λ_{12}	ohne Interaktion	-3965.2	3	7936.4
	alle Interaktionen	-3954.7	7	7923.4
λ_{13}	ohne Interaktion	849.9	3	1705.8
	alle Interaktionen	-847.0	7	1708.0
λ_{23}	ohne Interaktion	-2975.1	3	5956.2
	alle Interaktionen	-2972.9	7	5959.8

Tabelle 4.23: AIC-Kriterium für Modell ohne Interaktion und Modell mit allen Interaktionen für die Übergangintensitäten $\lambda_{12}(t)$, $\lambda_{13}(t)$ und $\lambda_{23}(t)$

berücksichtigt sind, d.h.

$$\begin{aligned}
\lambda_{12}(t) &= \lambda_{120}(t) \exp(\beta_1 Z_{\text{Alter}}(t) + \beta_2 Z_{\text{Geschlecht}} + \beta_3 Z_{\text{Stat}}(t)) \\
&+ \beta_4 Z_{\text{Alter}}(t) \times Z_{\text{Geschlecht}} + \beta_5 Z_{\text{Alter}}(t) \times Z_{\text{Stat}}(t) \\
&+ \beta_6 Z_{\text{Geschlecht}} \times Z_{\text{Stat}}(t) + \beta_7 Z_{\text{Alter}}(t) \times Z_{\text{Geschlecht}} \times Z_{\text{Stat}}(t) \quad (4.13)
\end{aligned}$$

Hier sehen wir eine starke 3-Faktor-Interaktion zwischen Alter, Geschlecht und Pflegeart (siehe Tabelle 4.24), so daß wir später bei den Kalkulationen im Versicherungsmodell auf (4.13) zurückgreifen werden.

Variable	Koeffizient	exp(Koeff.)	SE(Koeff)	z-Statistik	p-Wert
Z_{Alter}	0.0143	1.01	0.0047	3.06	0.002
$Z_{\text{Geschlecht}}$	-0.8544	0.43	0.5588	-1.53	0.130
Z_{Stat}	-10.7437	$2.6 \cdot 10^{-5}$	3.5743	-3.01	0.03
$Z_{\text{Alter}} \times Z_{\text{Geschlecht}}$	0.0099	1.01	0.0069	1.42	0.150
$Z_{\text{Alter}} \times Z_{\text{Stat}}$	0.1221	1.13	0.0407	3.00	0.003
$Z_{\text{Geschlecht}} \times Z_{\text{Art}}$	11.7833	$1.31 \cdot 10^5$	3.7654	3.13	0.002
$Z_{\text{Alter}} \times Z_{\text{Geschlecht}} \times Z_{\text{Stat}}$	-0.1323	0.88	0.0429	-3.08	0.002

Tabelle 4.24: Geschätzte Koeffizienten für die Übergangintensität λ_{12} bei Berücksichtigung aller Interaktionen

Aus Tabelle 4.23 sieht man, daß keine signifikanten Interaktionseffekte für $\lambda_{13}(t)$ und $\lambda_{23}(t)$ vorhanden sind. Bei Betrachtung der Zustandsübergänge von Zustand 1 in Zustand 3 ($\lambda_{13}(t)$) sowie von Zustand 2 in Zustand 3 ($\lambda_{23}(t)$) erkennt man keine signifikante Abhängigkeit der Übergangintensität von der Kovariable Z_{Stat} (die korrespondierenden p-Werte betragen 0.52, bzw. 0.25), so daß wir diese Kovariable im folgenden nicht berücksichtigen und die Intensitätsfunktionen

Variable	Koeffizient	exp(Koeff.)	SE(Koeff)	z-Statistik	p-Wert
Z_{Alter}	0.035	1.04	0.0087	3.98	$7 \cdot 10^{-5}$
$Z_{\text{Geschlecht}}$	-0.368	0.70	0.1873	-1.97	0.05

Tabelle 4.25: Geschätzte Koeffizienten für die Übergangintensität λ_{13}

modellieren mit

$$\lambda_{13}(t) = \lambda_{130}(t) \exp[\beta_1 Z_{\text{Alter}}(t) + \beta_2 Z_{\text{Geschlecht}}] \quad (4.14)$$

$$\lambda_{23}(t) = \lambda_{230}(t) \exp[\beta_1 Z_{\text{Alter}}(t) + \beta_2 Z_{\text{Geschlecht}}]. \quad (4.15)$$

Mit den Modellen (4.14) und (4.15) erhalten wir die Tabellen 4.25 und 4.26 aufgeführten Koeffizienten.

Variable	Koeffizient	exp(Koeff.)	SE(Koeff)	z-Statistik	p-Wert
Z_{Alter}	0.035	1.04	0.0087	4.06	$4.8 \cdot 10^{-5}$
$Z_{\text{Geschlecht}}$	-0.359	0.70	0.1865	-1.92	0.06

Tabelle 4.26: Geschätzte Koeffizienten für die Übergangintensität λ_{23}

Die Übergangswahrscheinlichkeiten kann man wiederum mit Hilfe der Differentialgleichungen bestimmen. Man erhält

$$\begin{aligned}
 P_{11}(z, t) &= \exp\left(-\int_z^t [\lambda_{12}(u) + \lambda_{13}(u) + \lambda_{14}(u)] du\right), \\
 P_{22}(z, t) &= \exp\left(-\int_z^t [\lambda_{23}(u) + \lambda_{24}(u)] du\right), \\
 P_{33}(z, t) &= \exp\left(-\int_z^t \lambda_{34}(u) du\right), \\
 P_{34}(z, t) &= 1 - P_{33}(z, t), \\
 P_{23}(z, t) &= \int_z^t P_{22}(z, u) \lambda_{23}(u) P_{33}(u, t) du, \\
 P_{24}(z, t) &= 1 - P_{22}(z, t) - P_{23}(z, t), \\
 P_{12}(z, t) &= \int_z^t P_{11}(z, u) \lambda_{12}(u) P_{22}(u, t) du, \\
 P_{13}(z, t) &= \int_z^t (P_{11}(z, u) \lambda_{13} + P_{12}(z, u) \lambda_{23}(u)) P_{33}(u, t) du, \\
 P_{14}(z, t) &= 1 - P_{11}(z, t) - P_{12}(z, t) - P_{13}(z, t).
 \end{aligned} \quad (4.16)$$

Mit diesen Vorbereitungen können wir uns nun der Konstruktion eines Versicherungsmodells zuwenden.

Kapitel 5

Entwicklung eines Versicherungsmodells

Mit Hilfe der in 2.5 hergeleiteten Theorie zu Markov-Prozessen und der in 4.5 hergeleiteten Übergangintensitäten soll nun ein Versicherungsmodell zur Pflegeversicherung entwickelt werden. Dazu sind aber noch einige vorbereitende Definitionen nötig, die im folgenden Abschnitt kurz dargestellt sind. Eine sehr ausführliche Darstellung zur Anwendung von Multi-State-Modellen in der Personenversicherung findet man in Haberman (1999) [19]. Auf Grundlagen der klassischen Lebensversicherung werden wir in diesem Abschnitt nur so weit wie nötig eingehen, eine Darstellung dieses Themas ist in Gerber (1997)[16] gegeben. In dem Handbuch zur Pflegeversicherung der Münchener Rück (1992) [36] findet man ausführliche Modelle zur Produktgestaltung und -kalkulation in der Pflegeversicherung, sowie statistisches Rohmaterial zur Gewinnung von Rechnungsgrundlagen.

5.1 Versicherung als Markov-Prozeß

Die Entwicklung eines versicherten Risikos kann als zeitstetiger Markov-Prozeß S mit endlichem Zustandsraum $Z = \{1, \dots, n\}$ beschrieben werden, das heißt

$$S : \mathcal{T} \times \Omega \mapsto \{1, \dots, n\}.$$

Hier ist $\mathcal{T} = [0, \tau)$, $\tau < \infty$ das Intervall von Versicherungsbeginn bis Versicherungsende. Der Versicherungsprozeß wird im Markovprozeß als ein (zustandsabhängiger) Zahlungsstrom zwischen dem Versicherten, oder Versicherungsnehmer (kurz: VN) und dem Versicherer (kurz: VR). Die einmaligen, oder kontinuierlichen Zahlungen des Versicherten an den Versicherer heißen Beiträge oder Prämien, die des Versicherers an den Versicherungsnehmer Leistungen. Für die Übergangswahrscheinlichkeiten und -intensitäten verwenden wir die Bezeichnungen aus Abschnitt 2.5.

5.1.1 Zahlungsfunktionen

Um den Zahlungsstrom beschreiben zu können, definiert man nun folgende von Zeitpunkt und Zustand abhängige Zahlungsfunktionen:

Definition 5.1 (Zahlungsfunktionen) Für den Versicherungsprozeß $S(t)$ sind folgende **Zahlungsfunktionen** definiert

- (i) $p_i(t)$: Eine stetige Prämie, die der VN bezahlt, solange er sich im Zustand i befindet.
- (ii) $b_i(t)$: Eine stetige Rente, die vom Versicherer bezahlt wird, solange sich der Versicherungsnehmer im Zustand i befindet.
- (iii) $c_{ij}(t)$: Die Zahlung einer bestimmten Summe falls der Versicherte vom Zustand i in den Zustand j wechselt.
- (iv) $d_i(t_0)$: Die Zahlung einer bestimmten Summe, falls sich der Versicherungsnehmer zum Zeitpunkt t_0 im Zustand i befindet.

Beispiel 5.2 (gemischte Lebensversicherung)

Die gemischte Lebensversicherung ist in Deutschland die häufigste Art der Lebensversicherung. Der Versicherungsnehmer zahlt hier bis zu einem bestimmten Zeitpunkt t_0 (zum Beispiel bis zum Alter 65 Jahre) einen stetigen Beitrag der Höhe p . Falls der Versicherungsnehmer stirbt, spätestens aber zum Zeitpunkt t_0 zahlt der Versicherer eine einmalige Leistung der Höhe c .

Bei der Modellierung als Markov-Modell interessieren bei diesem Modell nur die Zustände tot und lebendig (Abbildung 5.1). Der Zahlungsstrom für diese Versicherung sieht wie folgt aus

$$\begin{aligned} p_1(t) &= \begin{cases} p & t < t_0 \\ 0 & \text{sonst} \end{cases}, \\ c_{12}(t) &= \begin{cases} c & t < t_0 \\ 0 & \text{sonst} \end{cases}, \\ d_1(t_0) &= c, \\ b_i(t) &= 0 \quad i = 1, 2. \end{aligned}$$

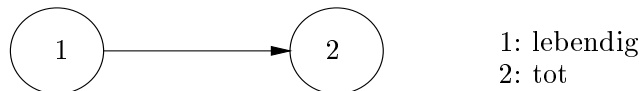


Abbildung 5.1: Zustandsübergänge in Lebensversicherungsprozeß

Beispiel 5.3 (private Altersrente) Der Versicherungsnehmer zahlt solange er lebt, jedoch maximal bis zu einem bestimmten Alter t_0 eine stetige Prämie und erhält ab einem Alter

$t_1 \geq t_0$ eine stetige Rente der Höhe b . Auch für dieses Beispiel sind die uns interessierenden Zustandsübergänge in Abbildung 5.1 skizziert. Als Zahlungsstrom für dieses Modell erhalten wir

$$\begin{aligned} p_1(t) &= \begin{cases} p & t < t_0 \\ 0 & \text{sonst} \end{cases}, \\ b_1(t) &= \begin{cases} b & t \geq t_1 \\ 0 & \text{sonst} \end{cases}, \\ c_{12}(t) &= 0. \end{aligned}$$

5.1.2 Berechnung von Barwerten und Erwartungswerten

Um Barwerte berechnen zu können benötigt man eine finanzielle Struktur, das bedeutet, zu jedem Zeitpunkt t muß der Wert einer finanziellen Leistung ermittelt werden können. Dies erreicht man mittels einer sogenannten Zinsstruktur. Im unserem Fall soll die einfachste finanzielle Struktur angenommen werden, das heißt wir betrachten ein Modell mit stetiger Zinsstruktur und konstanter Zinsintensität $\delta(t) = \delta$. Die Diskontierungsfunktion hat damit den Wert:

$$\exp\left(-\int_0^t \delta ds\right) = e^{-\delta t}.$$

Sei mit v der jährliche Diskontierungsfaktor bezeichnet, so gilt:

$$v = e^{-\delta}.$$

Bemerkung: Die Zinsintensität läßt sich folgendermaßen interpretieren: Ein Anleger möchte zum Zeitpunkt $t < u$ genau so viel Geld anlegen, daß er zum Zeitpunkt u durch stetige Verzinsung mit Zinsintensität δ den Wert 1 erhält. Falls t und u in die Zeit in Jahren angibt, benötigt der Anleger genau:

$$\exp\left(-\int_t^u \delta ds\right) = e^{-\delta(u-t)} = v^{u-t}.$$

Nun können wir die (zufälligen) Barwerte einer Leistung für die in Definition 5.1 definierten Zahlungsfunktionen angeben. Diese Barwerte sind sozusagen der Wert einer Leistung zum Zeitpunkt t , falls die Leistung zum Zeitpunkt u , bzw. im Zeitintervall $[u_1, u_2]$ erbracht werden.

Definition 5.4 (Barwert) Für die in Definition 5.1 (i) - (iv) beschriebenen Zahlungsfunktionen werden folgende (zufälligen) **Barwerte** definiert:

(i) Der Barwert einer stetigen Prämie p_j zum Zeitpunkt t ist folgendermaßen definiert

$$Y_t^{p_j}(u, u + du) := v^{u-t} I(S(u) = j) p_j(u) du.$$

für den Prämienbarwert über das Intervall $[u_1, u_2]$ mit $t \leq u_1 < u_2$ erhält man

$$Y_t^{p_j}(u_1, u_2) := \int_{u_1}^{u_2} v^{u-t} I(S(u) = j) p_j(u) du.$$

(ii) Analog ist der Barwert für eine stetige Rente b_j zum Zeitpunkt t für das Intervall $[u_1, u_2)$

$$Y_t^{b_j}(u_1, u_2) = \int_{u_1}^{u_2} v^{u-t} I(S(u) = j) b_j(u) du.$$

(iii) Der Barwert einer einmaligen Zahlung c_{ij} bei Übergang von Zustand i nach Zustand j ist

$$Y_t^{c_{ij}}(u) = v^{u-t} I(S(u-) = i \wedge S(u) = j).$$

(iv) Zuletzt erhalten wir noch für eine einmalige Zahlung $d_j(t_0)$, die der Versicherungsnehmer erhält, falls er sich zum Zeitpunkt t_0 im Zustand j befindet

$$Y_t^{d_j}(t_0) = v^{u-t} I(S(u) = j) d_j(t_0).$$

5.1.3 Aktuarielle Werte

Nachdem wir in Abschnitt 5.1.2 mit den Barwerten den Wert verschiedener in der Zukunft liegender Zahlungsströme zwischen dem Versicherungsnehmer und dem Versicherer modelliert haben, interessieren wir uns nun für die sogenannten aktuariellen Werte, die den Erwartungswert (hinsichtlich des Markovprozesses S) dieser Zahlungsströme beschreiben. Mit Hilfe dieser aktuariellen Werte kann später das Versicherungsmodell formuliert werden.

Definition 5.5 (aktuarielle Werte) Für einen Versicherten, der sich zum Zeitpunkt t im Zustand i des Versicherungsprozesses S befindet ist der **aktuarielle Wert** einer Zahlungsfunktion für einen Zeitpunkt u bzw. für ein Zeitintervall $[u_1, u_2)$ als der für diesen Zeitpunkt, bzw. -raum erwartete Barwert (gegeben $S(t) = i$) definiert. Für die in Definition 5.1 beschriebenen Zahlungsfunktionen bedeutet dies im Einzelnen

(i) Der aktuarielle Wert einer kontinuierlichen Prämie p_j für das infinitesimale Intervall $[u, u + du)$ ist

$$E[Y_t^{p_j}(u, u + du) | S(t) = i] = v^{u-t} P_{ij}(t, u) p_j(u) du.$$

Für das Intervall $[u_1, u_2)$ erhält man den aktuariellen Wert

$$E[Y_t^{p_j}(u_1, u_2) | S(t) = i] = \int_{u_1}^{u_2} v^{u-t} P_{ij}(t, u) p_j(u) du.$$

(ii) Wiederum kann man analog für eine stetige Rente $b_j(t)$ den aktuariellen Wert angeben mit

$$E[Y_t^{b_j}(u_1, u_2) | S(t) = i] = \int_{u_1}^{u_2} v^{u-t} P_{ij}(t, u) b_j(u) du.$$

(iii) Für einmalige Zahlungen an den Versicherungsnehmer, falls dieser von Zustand j in Zustand k wechselt, $c_{jk}(u)$ bedeutet dies

$$E[Y_t^{c_{jk}}(u) | S(t) = i] = v^{u-t} P_{ij}(t, u) \lambda_{jk}(u) c_{jk}(u) du$$

für das Intervall $[u_1, u_2)$ ist der aktuarielle Wert

$$E[Y_t^{c_{jk}}(u_1, u_2) | S(t) = i] = \int_{u_1}^{u_2} v^{u-t} P_{ij}(t, u) c_{jk}(u) \lambda_{jk}(u) du$$

(iv) Schließlich erhält für eine einmalige Zahlung des Versicherers zum Zeitpunkt t_0 den Barwert

$$E[Y^{d_j}(t_0)|S(t) = i] = v^{t_0-t} P_{ij}(t, t_0) d_j(t_0)$$

5.1.4 Beiträge und Reserven

Das jeder Versicherung zugrundeliegende Prinzip ist das versicherungsmathematische Äquivalenzprinzip. Salopp ausgedrückt besagt dieses Prinzip, daß jeder Versicherte für seinen erwarteten Schaden selbst aufkommt. Für den Zahlungsstrom bedeutet dies, daß zu Policenbeginn ($t=0$) der aktuarielle Wert aller Zahlungsströme vom Versicherer zum Versicherungsnehmer dem aktuariellen Wert der Zahlungsströme vom Versicherungsnehmer zum Versicherer entspricht. In der bisher entwickelten Terminologie läßt sich dieses Äquivalenzprinzip wie folgt beschreiben

Definition 5.6 Für einen Versicherungsprozeß $S(t)$ mit Policenende τ definieren wir mit

$$\begin{aligned} \mathcal{B}_i(t, \tau) &= \int_t^\tau v^{u-t} \left[\sum_{j \in \mathcal{S}} P_{ij}(t, u) b_j(u) \right] du \\ &+ \int_t^\tau v^{u-t} \left[\sum_{j \in \mathcal{S}} \sum_{k \neq j} P_{ij}(t, u) \lambda_{jk}(u) c_{jk}(u) \right] du \\ &+ \sum_{u: u \geq t} v^{u-t} \left[\sum_{j \in \mathcal{S}} P_{ij}(t, u) d_j(u) \right] \end{aligned}$$

die Summe aller erwarteten Leistungen zum Zeitpunkt t , gegeben $S(t) = i$ und analog mit

$$\mathcal{P}_i(t, \tau) = \int_t^\tau v^{u-t} \left[\sum_{j \in \mathcal{S}} P_{ij}(t, u) p_j(u) \right] du$$

die Summe aller erwarteten Prämien zum Zeitpunkt t , gegeben $S(t) = i$.

Mit dieser Definition läßt sich das Äquivalenzprinzip nun folgendermaßen formulieren:

Definition 5.7 (Versicherungsmathematisches Äquivalenzprinzip) Für eine Versicherungspolice mit Endzeitpunkt τ und Anfangszustand des versicherten Risikos $S(0) = 1$ ist das **versicherungsmathematische Äquivalenzprinzip** genau dann erfüllt wenn gilt:

$$\mathcal{P}_1(0, \tau) = \mathcal{B}_1(0, \tau). \quad (5.1)$$

Gerade in der Personenversicherung ist das Äquivalenzprinzip meist nur zu Beginn der Versicherungspolice erfüllt. Der Grund dafür liegt unter anderem an der Tatsache, daß das versicherte Risiko mit wachsendem Alter immer größer wird, der gezahlte Versicherungsbeitrag aber gleich bleibt. Wichtig in diesem Zusammenhang ist, daß während der gesamten Vertragslaufzeit gelten muß

$$\mathcal{P}_1(t, \tau) \leq \mathcal{B}_1(t, \tau). \quad (5.2)$$

was bedeutet, daß der erwartete Zahlungsstrom von Versicherungsnehmer zum Versicherer immer mindestens genauso groß ist, wie der umgekehrte Strom. Der Grund dafür ist einleuchtend, auf diese Weise möchte man verhindern, daß der Versicherer zum “Gläubiger” des Versicherungsnehmer wird, die Position des Versicherers wird also gestärkt.

Definition 5.8 Für einen Versicherten im Zustand i nennt man die Differenz aus aktuariellem Wert aller Beiträge und aktuariellem Wert aller Leistungen **prospektive Reserve** zum Zeitpunkt t

$$\bar{V}_i(t, \tau) = \mathcal{B}_i(t, \tau) - \mathcal{P}_i(t, \tau).$$

(5.1) ist keine eindeutige Bedingung für die Gestaltung der Versicherungsprämie, vielmehr sind (5.1) und (5.2) Randbedingungen, innerhalb deren sich die Versicherungsprämie bewegen kann. Wichtige Beispiele für die Gestaltung der Prämienstruktur sind

- **Einmalprämien:** Der Versicherte zahlt die gesamte Prämie zu Policenbeginn
- **gleichbleibende Jahresprämien:** Der Versicherungsnehmer zahlt in jährlichen Abständen einen Beitrag gleichbleibender Höhe.

5.2 Berechnung von Beiträgen für Pflegeversicherungs-Modelle

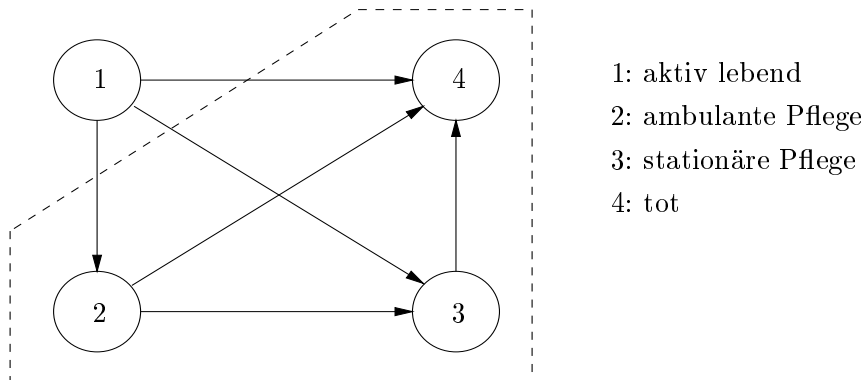


Abbildung 5.2: Zustandsübergänge in Pflegeversicherungsprozess

In diesen Abschnitt wollen wir darstellen, wie mit Hilfe der bisherigen Ergebnisse Pflegeversicherungsbeiträge kalkuliert werden können. Um Berechnungen realisieren zu können müssen zunächst die Markov-Modelle um einen weiteren Zustand, den wir im folgenden mit “aktiv-lebend” bezeichnen erweitern. In den Abbildungen 5.2 bzw. 5.3 sehen wir die Graphen der so erweiterten Markovmodelle. Mit der gestrichelten Linie sind die Zustände, die im Datensatz beobachtbar waren, umrandet. Da dem Datensatz nur Informationen über pflegebedürftige Versicherte entnehmbar waren, werden wir die Übergangswahrscheinlichkeiten für “aktiv lebende” einer anderen Quelle entnehmen (einjährige Pflegefall-Eintrittswahrscheinlichkeiten Custodial-Care-Insurance, aus [36] Anhang 6, für die Übergänge von aktiv lebend zu pflegebedürftig).

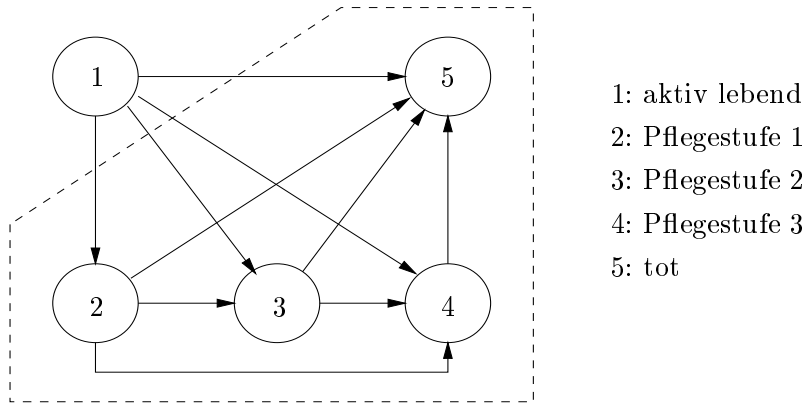


Abbildung 5.3: Zustandsübergänge in Pflegeversicherungsprozess

5.2.1 Versicherungsmodell mit Zustandsübergängen zwischen ambulanter und stationärer Pflege

Betrachten wir nun das in Abbildung 5.2 skizzierte Versicherungsmodell. Zunächst erkennt man, daß die Übergangintensitäten für die Zustände 2, 3 und 4 schon als Lösung des Differentialgleichungssystems (4.11) vorliegen. Für den Zustand 1 erhalten wir 4 weitere Differentialgleichungen.

$$\begin{aligned}
 \frac{d}{dt}P_{11}(z, t) &= -P_{11}(z, t)(\lambda_{12}(t) + \lambda_{13}(t) + \lambda_{14}(t)), \\
 \frac{d}{dt}P_{12}(z, t) &= P_{11}(z, t)\lambda_{12}(t) - P_{12}(z, t)(\lambda_{23}(t) + \lambda_{24}(t)), \\
 \frac{d}{dt}P_{13}(z, t) &= P_{11}(z, t)\lambda_{13}(t) + P_{12}(z, t)\lambda_{23}(t) - P_{13}(z, t)\lambda_{34}(t), \\
 \frac{d}{dt}P_{14}(z, t) &= P_{11}(z, t)\lambda_{14}(t) + P_{12}(z, t)\lambda_{24}(t) + P_{13}(z, t)\lambda_{34}(t).
 \end{aligned} \tag{5.3}$$

Für das Differentialgleichungssystem erhalten wir folgende Lösung

$$\begin{aligned}
 P_{11}(z, t) &= \exp\left(-\int_z^t [\lambda_{12}(u) + \lambda_{13}(u) + \lambda_{14}(u)] du\right), \\
 P_{22}(z, t) &= \exp\left(-\int_z^t [\lambda_{23}(u) + \lambda_{24}(u)] du\right), \\
 P_{33}(z, t) &= \exp\left(-\int_z^t \lambda_{34}(u) du\right), \\
 P_{34}(z, t) &= 1 - P_{33}(z, t), \\
 P_{23}(z, t) &= \int_z^t P_{22}(z, u)\lambda_{23}(u)P_{33}(u, t) du, \\
 P_{24}(z, t) &= 1 - P_{22}(z, t) - P_{23}(z, t), \\
 P_{12}(z, t) &= \int_z^t P_{11}(z, u)\lambda_{12}(u)P_{22}(u, t) du, \\
 P_{13}(z, t) &= \int_z^t [P_{11}(z, u)\lambda_{13}(u) + P_{12}(z, u)\lambda_{23}(u)P_{33}(u, t)] du, \\
 P_{14}(z, t) &= 1 - P_{11}(z, t) - P_{12}(z, t) - P_{13}(z, t).
 \end{aligned} \tag{5.4}$$

Diskretisierung des Modells

Die Rechnungsgrundlagen in der klassischen Lebensversicherungsmathematik sind normalerweise Übergangswahrscheinlichkeiten für einen diskreten Zeitraum. So sind im Bereich der Pflegeversicherung für Rechnungsgrundlagen in der folgenden Notation üblich (mit T ist hier die Zufallsvariable Lebensalter bezeichnet, z ist das erreichte Lebensalter, falls nach Geschlechtern differenzierte Rechnungsgrundlagen Anwendung finden, wir als Index für das erreichte Lebensalter für Männer ein x und für Frauen ein y verwendet).

- Mit

$$q_z = P(z \leq T < z + 1 | T \geq z)$$

ist die Wahrscheinlichkeit für eine z -jährige Person, im Lebensjahr z zu sterben, gegeben Überleben bis z , bezeichnet.

- Mit

$$q_z^a = P(z \leq T < z + 1 | T \geq z, \text{ aktiv lebend})$$

ist die Wahrscheinlichkeit für einen z -jährigen aktiv-lebenden im Lebensjahr z zu sterben, gegeben Überleben bis z , bezeichnet.

- Analog dazu ist

$$q_z^i = P(z \leq T < z + 1 | T \geq z, \text{ pflegebedürftig})$$

die Wahrscheinlichkeit für einen z -jährigen Pflegebedürftigen im Lebensjahr z zu sterben.

- Zuletzt wird noch mit

$$i_z = P(\text{pflegebedürftig bei Erreichen des Alters } z + 1 | T \geq z, \text{ nicht pflegebedürftig})$$

die sogenannte einjährige Pflegefalleintrittswahrscheinlichkeit für eine z -jährige Person bezeichnet

Da wir aufgrund fehlender Beobachtungen für aktiv lebende, teilweise auch auf diese einjährigen Übergangswahrscheinlichkeiten angewiesen sind, betrachten wir im Folgenden ein zeitdiskretes Versicherungsmodell mit den einjährigen Übergangswahrscheinlichkeiten

$$p_{ij}(n) = P(S(n+1) = j | S(n) = i)$$

wobei hier mit $n \in \mathbb{N}_0$ die Versicherungsdauer in Jahren angibt. Mit Hilfe der Übergangsmatrix

$$\Pi(n) = (p_{ij}(n))_{i,j} = \begin{pmatrix} p_{11}(n) & p_{12}(n) & p_{13}(n) & p_{14}(n) \\ 0 & p_{22}(n) & p_{23}(n) & p_{24}(n) \\ 0 & 0 & p_{33}(n) & p_{34}(n) \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

können wir nun analog zu (5.4) die Übergangswahrscheinlichkeiten für das zeitdiskrete Modell angeben, für $n \leq m \in \mathbb{N}_0$ erhalten wir

$$P_{ij}(n, m) = P(S(m) = j | S(n) = i) = \delta_{ij} \text{ falls } n = m.$$

Da unsere Zustände 1, 2 und 3 strikt transient sind erhalten wir für die Wahrscheinlichkeit im Zustand i zu verbleiben:

$$P_{ii}(n, n+1) = 1 - \sum_{j>i} p_{ij}(n).$$

Aus dem rekursiven Zusammenhang

$$P_{ii}(n, m) = P_{ii}(n, m-1)P_{ii}(m-1, m)$$

bestimmen wir die Wahrscheinlichkeit, daß man sich zum Zeitpunkt m in Zustand i befindet, falls man sich schon zum Zeitpunkt n in Zustand i befunden hat. Es gilt

$$P_{ii}(n, m) = \prod_{i=0}^{m-n-1} [1 - \sum_{j>i} p_{ij}(n+i)].$$

Damit erhalten wir

$$\begin{aligned} P_{11}(n, m) &= \prod_{i=0}^{m-n-1} [1 - p_{12}(n+i) - p_{13}(n+i) - p_{14}(n+i)], \\ P_{22}(n, m) &= \prod_{i=0}^{m-n-1} [1 - p_{23}(n+i) - p_{24}(n+i)], \\ P_{33}(n, m) &= \prod_{i=0}^{m-n-1} [1 - p_{34}(n+i)], \\ P_{34}(n, m) &= 1 - P_{34}(n, m), \\ P_{23}(n, m) &= \sum_{i=0}^{m-n-1} P_{22}(n, n+i)p_{23}(n+i)P_{33}(n+i+1, m), \\ P_{24}(n, m) &= 1 - P_{22}(n, m) - P_{23}(n, m), \\ P_{12}(n, m) &= \sum_{i=0}^{m-n-1} P_{11}(n, n+i)p_{12}(n+i)P_{22}(n+i+1, m), \\ P_{13}(n, m) &= \sum_{i=0}^{m-n-1} (P_{11}(n, n+i)p_{13}(n+i) + P_{12}(n, n+i)p_{23}(n+i))P_{33}(n+i+1, m), \\ P_{14}(n, m) &= 1 - P_{11}(n, m) - P_{12}(n, m) - P_{13}(n, m). \end{aligned} \tag{5.5}$$

Modellierung des Versicherungsprozesses

Wir betrachten nun ein Versicherungsmodell in dem der Versicherungsnehmer gegen eine gleichbleibende Jahresprämie der Höhe π folgende Leistungen erhält:

- Jeweils einmalige Zahlungen der Höhe c_{1j} falls er von Zustand 1 in Zustand j wechselt $j \in \{2, 3\}$. In vielen Pflegeversicherungstarifen existieren solche Leistungen, damit werden meist zu Pflegebeginn entstehende, einmalige Kosten (wie z.B. Einbau eines Treppenliftes) gedeckt.
- Gleichbleibende, jährliche Rentenzahlungen der Höhe b_j , falls sich der Versicherte in Zustand j befindet ($j = \{2, 3\}$).

Wir gehen zudem davon aus, daß die Versicherung leistet, solange der Versicherte lebt. Um diese beliebig lange Versicherungsdauer aktuariell zu modellieren, setzt man für ein bestimmtes Alter ω (in der Praxis meist $\omega = 105$) die Wahrscheinlichkeit dafür, daß der Versicherte im Lebensjahr ω stirbt auf 1. ω heißt technisches Endalter. Mit diesen Modellvoraussetzungen und den Zusammenhängen aus (5.5) erhalten wir für die verschiedenen Leistungen folgende aktuarielle Werte zum Zeitpunkt 0 (Versicherungsbeginn) als Erwartungswert der jeweiligen Zahlungsströme hinsichtlich des Versicherungsprozesses $S(n)$, $n \in \mathbb{N}_0$.

Für die den aktuariellen Wert (zu Versicherungsbeginn) einmaliger Zahlungen bei Übergang von Zustand 1 (aktiv lebend) zu Zustand 2 (ambulant pflegebedürftig) c_{12} , den wir mit $\mathcal{B}_{1,c_{12}}(0)$ bezeichnen, erhalten wir

$$\begin{aligned} \mathcal{B}_{1,c_{12}}(0) &= \sum_{i=0}^{\omega-x-1} P_{11}(0, i) p_{12}(i) v^i c_{12} \\ &= \sum_{i=0}^{\omega-x-1} \prod_{j<i} [1 - p_{12}(j) - p_{13}(j) - p_{14}(j)] p_{12}(i) v^i c_{12}. \end{aligned}$$

Analog gilt natürlich für den aktuariellen Wert einer einmaligen Leistung, die bei Übergang von Zustand 1 (aktiv lebend) in Zustand 3 (stationär pflegebedürftig) gezahlt wird

$$\mathcal{B}_{1,c_{13}}(0) = \sum_{i=0}^{\omega-x-1} \prod_{j<i} [1 - p_{34}(j) - p_{24}(j)] p_{13}(i) v^i c_{13}.$$

Für den aktuariellen Wert einer jährlichen Rente b_2 , die an den Pflegebedürftige bezahlt wird, solange sich dieser in Zustand 2 befindet

$$\begin{aligned} \mathcal{B}_{1,b_2}(0) &= \sum_{i=0}^{\omega-x-1} P_{12}(0, i) v^i b_2 \\ &= \sum_{i=0}^{\omega-x-1} \left[\sum_{j<i} P_{11}(0, j) p_{12}(j) P_{22}(j+1, i) \right] v^i b_2 \\ &= \sum_{i=0}^{\omega-x-1} \left[\sum_{j<i} \prod_{k<j} [1 - p_{12}(k) - p_{13}(k) - p_{14}(k)] p_{12}(j) \right. \\ &\quad \left. \prod_{k<i-(j+1)} [1 - p_{23}(j+k+1) - p_{24}(j+k+1)] \right] v^i b_2. \end{aligned}$$

Der aktuarielle Wert einer jährlichen Rente b_3 für einen Pflegebedürftigen in Zustand 3 ist

$$\begin{aligned}
 \mathcal{B}_{1,b_3}(0) &= \sum_{i=0}^{\omega-x-1} P_{13}(0,i)v^i b_3 \\
 &= \sum_{i=1}^{\omega-x-1} \left\{ \sum_{j<i} \left[P_{11}(0,j)p_{13}(j)P_{33}(j+1,i) + P_{12}(0,j) \cdot p_{23}(j)P_{33}(j+1,i) \right] \right\} v^i b_3 \\
 &= \sum_{i=1}^{\omega-x-1} \left\{ \sum_{j<i} \left[\prod_{k<j} [1 - p_{12}(k) - p_{13}(k) - p_{14}(k)] p_{13}(j) \prod_{k<i-(j+1)} [1 - p_{34}(j+k+1)] \right. \right. \\
 &\quad \left. \left. + \sum_{k<j} \prod_{l<k} [1 - p_{12}(l) - p_{13}(l) - p_{14}(l)] p_{12}(k) \prod_{l<j-(k+1)} [1 - p_{23}(k+l+1)] \right. \right. \\
 &\quad \left. \left. - p_{24}(k+l+1) \right] p_{23}(j) \prod_{l<i-(j+1)} [1 - p_{34}(j+l+1)] \right\} v^i b_3.
 \end{aligned}$$

Schließlich erhalten wir als Barwert für die von Versicherungsnehmer zu zahlende Jahresprämie (solange er sich in Zustand 1 befindet)

$$\begin{aligned}
 \mathcal{P}_{1,\pi}(0) &= \sum_{i=0}^{\omega-x-1} P_{11}(0,i)v^i \pi \\
 &= \sum_{i=0}^{\omega-x-1} \prod_{j<i} [1 - p_{12}(j) - p_{13}(j) - p_{14}(j)] v^i \pi.
 \end{aligned}$$

Nun können wir den Jahresbeitrag π bestimmen, denn nach dem Äquivalenzprinzip muß gelten

$$\mathcal{B}_{1,b_2}(0) + \mathcal{B}_{1,b_3}(0) + \mathcal{B}_{1,c_{12}}(0) + \mathcal{B}_{1,c_{13}}(0) = \mathcal{P}_{1,\pi}(0).$$

Diesen Ausdruck kann man nach π auflösen und erhält

$$\pi = \frac{\mathcal{B}_{1,b_2}(0) + \mathcal{B}_{1,b_3}(0) + \mathcal{B}_{1,c_{12}}(0) + \mathcal{B}_{1,c_{13}}(0)}{\sum_{i=0}^{\omega-x-1} \prod_{j<i} [1 - p_{12}(j) - p_{13}(j) - p_{14}(j)] v^i}.$$

Modellierung von Rechnungsgrundlagen

Wie schon vorher erwähnt, sind in der Arbeit nur Übergangintensitäten für die Übergänge zwischen den Zuständen “pflegebedürftig” und “tot” untersucht, die Übergangintensitäten zwischen “aktiv lebend”-“Pflegebedürftig” und “aktiv lebend”-“tot” waren auf Basis der gegebenen Daten nicht ermittelbar. Zudem müssen wir beachten, daß die Übergangswahrscheinlichkeiten p_{ij} , $j \in \{2, 3, 4\}$ primär vom Lebensalter x abhängen, für die Übergangswahrscheinlichkeiten p_{24} und p_{34} und p_{23} mit der Pflegedauer (im folgenden mit d bezeichnet) eine weitere Zeitabhängigkeit modelliert wird. Die Markov-Eigenschaft geht dadurch verloren, da in diesem Modell die Übergangswahrscheinlichkeiten p_{24} , p_{34} und p_{23} von der Ersteintrittszeit in Zustand 2, bzw. 3 (Pflegebeginn) abhängig sind. Als Modellgrößen sollen für die Berechnungen folgende Daten verwendet werden:

- Für die Übergangswahrscheinlichkeiten p_{12} und p_{13} verwenden wir die von Alter und Geschlecht abhängigen Pflegefalleintrittswahrscheinlichkeiten der Custodial-Insurance, Japan (siehe [36], Anhang 6). Diese sind nach Alter und Geschlecht differenziert. Um eine Unterscheidung nach Pflegeart vornehmen zu können, betrachten wir die relativen Häufigkeiten der Neueintritte in die verschiedenen Pflegearten im Zeitraum zwischen 1.1.1997 und 31.12.1998 in unserem Bestand. Dieses Zeitintervall ist so gewählt, da erst ab 1.7.1996 Leistungen für stationäre Pflege gewährt werden. Bei Neueintritten, unterteilt nach ambulanten und stationärer Pflegebedürftigkeit, konnten wir folgende relativen Häufigkeiten feststellen:

	Frauen (in %)	Männer (in %)
ambulant	79.43	85.37
stationär	20.57	14.63

- In dem in Abbildung 5.2 dargestellten Versicherungsmodell sind keine Übergänge zwischen Pflegestufen vorgesehen. Bei der Beitragsberechnung berücksichtigen wir die Pflegestufen, indem wir für jede Stufe separat Barwerte und Beiträge berechnen und diese dann gewichten. Als Gewichte verwenden wir die Pflegedauer in den einzelnen Stufen. Wir konnten folgende Häufigkeiten beobachten

	Frauen (in %)	Männer (in %)
Stufe 1	36.41	41.50
Stufe 2	39.82	26.27
Stufe 3	23.77	33.73

- Als Modellgröße für die “Aktivensterblichkeit” p_{14} verwenden wir die einjährigen Sterbewahrscheinlichkeit für bayerische Männer und Frauen zwischen 1986 und 1988 (Quelle [14], Seiten 50-51), das technische Endalter ω liegt bei 101 Jahren. Hier ist zu beachten, daß diese einjährigen Pflegewahrscheinlichkeiten auf Basis der gesamten (also auch pflegebedürftigen) Bevölkerung ermittelt wurden, die einjährigen Sterbewahrscheinlichkeiten für “aktiv lebende” sind demnach etwas niedriger.
- Als Größe für die Übergangswahrscheinlichkeiten p_{23} , p_{24} und p_{34} diskretisieren wir die in Kapitel 4 hergeleiteten Sterbeintensitäten. Dazu betrachten wir die mit dem Breslow-Schätzer für den Basis-Hazard geschätzten Intensitäten an den Beobachtungszeitpunkten und bilden die einjährige Übergangswahrscheinlichkeiten. Für Lebensalter x und Pflegedauer d (in Jahren) erhalten wir

$$p_{ij}(x, d) = \sum_{d \leq t_k < d+1} [\hat{\lambda}_{ij}(t_k, Z_{\text{Alter}} = x) \prod_{d \leq t_l < t_k} (1 - \hat{\lambda}_{ij}(t_l, Z_{\text{Alter}} = x))] \quad \begin{array}{l} i \in \{2, 3\}, \\ j \in \{2, 3, 4\}, \end{array}$$

d.h. wir approximieren einjährige Überlebenswahrscheinlichkeiten mit dem Produkt-Limit Ansatz (siehe Abschnitt 3.6.1). Hier ist $\hat{\lambda}_{ij}(t_k, Z_{\text{Alter}} = x)$ das Produkt aus Basis-Hazard für den Zustandsübergang von i nach j und Kovariable nach dem Proportional-Hazard-Modell, wobei die Komponente Z_{Alter} den Wert x annimmt.

5.2.2 Versicherungsmodell mit Zustandsübergängen zwischen Pflegestufen

Betrachten wir nun das in 5.3 skizzierte Modell. Wir erkennen, daß die Struktur dem in Abschnitt 5.2.1 betrachteten Modell sehr ähnlich ist (die Zustände $\{1, \dots, 4\}$ sind strikt transient, der Zustand 5 ist absorbierend). Die Herleitung der Übergangswahrscheinlichkeiten für ein zeitdiskretes Markovmodell, sowie die Berechnungen von aktuariellen Werten wird wie im vorherigen Abschnitt durchgeführt.

Diskrete Übergangswahrscheinlichkeiten im Modell

Mit Hilfe der Übergangsmatrix

$$\Pi(n) = (p_{ij}(n))_{i,j} = \begin{pmatrix} p_{11}(n) & p_{12}(n) & p_{13}(n) & p_{14}(n) & p_{15}(n) \\ 0 & p_{22}(n) & p_{23}(n) & p_{24}(n) & p_{25}(n) \\ 0 & 0 & p_{33}(n) & p_{34}(n) & p_{35}(n) \\ 0 & 0 & 0 & p_{44}(n) & p_{45}(n) \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

erhalten folgende Übergangswahrscheinlichkeiten

$$\begin{aligned} P_{11}(n, m) &= \prod_{i=0}^{m-n-1} [1 - p_{12}(n+i) - p_{13}(n+i) - p_{14}(n+i) - p_{15}(n+i)], \\ P_{22}(n, m) &= \prod_{i=0}^{m-n-1} [1 - p_{23}(n+i) - p_{24}(n+i) - p_{25}(n+i)], \\ P_{33}(n, m) &= \prod_{i=0}^{n-m-1} [1 - p_{34}(n+i) - p_{35}(n+i)], \\ P_{44}(n, m) &= \prod_{i=1}^{n-m-1} [1 - p_{45}(n+i)], \\ P_{45}(n, m) &= 1 - P_{44}(n, m), \\ P_{34}(n, m) &= \sum_{i=0}^{n-m-1} P_{33}(n, n+i) p_{34}(n+i) P_{44}(n+i+1, m), \\ P_{23}(n, m) &= \sum_{i=0}^{n-m-1} P_{22}(n, n+i) p_{23}(n+i) P_{33}(n+i+1, m), \\ P_{12}(n, m) &= \sum_{i=0}^{n-m-1} P_{11}(n, n+i) p_{12}(n+i) P_{22}(n+i+1, m), \\ P_{35}(n, m) &= 1 - P_{34}(n, m) - P_{33}(n, m), \end{aligned}$$

$$\begin{aligned}
P_{24}(n, m) &= \sum_{i=1}^{n-m-1} [P_{22}(n, n+i)p_{24}(n+i) + P_{23}(n, n+i)p_{34}(n+i)]P_{44}(n+i+1, m), \\
P_{13}(n, m) &= \sum_{i=1}^{n-m-1} [P_{11}(n, n+i)p_{13}(n+i) + P_{12}(n, n+i)p_{23}(n+i)]P_{33}(n+i+1, m), \\
P_{25}(n, m) &= 1 - P_{22}(n, m) - P_{23}(n, m) - P_{24}(n, m), \\
P_{14}(n, m) &= \sum_{i=0}^{n-m-1} [p_{11}(n, n+i)p_{14}(n+i) + P_{12}(n, n+i)p_{24}(n+i), \\
&\quad + P_{13}(n, n+i)p_{34}(n+i)]P_{44}(n+i+1, m), \\
P_{15}(n, m) &= 1 - P_{11}(n, m) - P_{12}(n, m) - P_{13}(n, m) - P_{14}(n, m). \tag{5.6}
\end{aligned}$$

Modellierung des Versicherungsprozesses

Analog zu Abschnitt 5.2.1 soll das Versicherungsmodell gegen eine gleichbleibende Jahresprämie π folgende Leistungen gedeckt sein

- Jeweils einmalige Zahlungen der Höhe c_{ij} falls er von Zustand i in Zustand j wechselt $i, j \in \{1, 2, 3, 4\}$ $i < j$.
- Gleichbleibende, jährliche Rentenzahlungen der Höhe b_j , falls sich der Versicherte in Zustand j befindet ($j = \{2, 3, 4\}$).

Die aktuariellen Werte lassen sich nun ähnlich wie im vorherigen Modell bestimmen. Für die einmaligen Zahlungen bei Zustandsübergängen erhalten wir

$$\begin{aligned}
\mathcal{B}_{1,c_{12}}(0) &= \sum_{i=1}^{\omega-x-1} \prod_{j<i} [1 - p_{12}(j) - p_{13}(j) - p_{14}(j) - p_{15}(j)] p_{12}(i) v^i c_{12}, \\
\mathcal{B}_{1,c_{13}}(0) &= \sum_{i=1}^{\omega-x-1} \prod_{j<i} [1 - p_{12}(j) - p_{13}(j) - p_{14}(j) - p_{15}(j)] p_{13}(i) v^i c_{13}, \\
\mathcal{B}_{1,c_{14}}(0) &= \sum_{i=1}^{\omega-x-1} \prod_{j<i} [1 - p_{12}(j) - p_{13}(j) - p_{14}(j) - p_{15}(j)] p_{14}(i) v^i c_{14}.
\end{aligned}$$

Die aktuarielle Wert für jährliche Rentenzahlungen, die der Versicherungsnehmer erhält, falls er sich in den Zuständen 2 und 3 befindet berechnen wir analog zum Vorgehen in Abschnitt 5.2.1.

Wir erhalten für \mathcal{B}_{b_2}

$$\begin{aligned}
\mathcal{B}_{b_2} &= \sum_{i=0}^{\omega-x-1} \left[\sum_{j<i} \prod_{k<j} [1 - p_{12}(k) - p_{13}(k) - p_{14}(k) - p_{15}(k)] p_{12}(j) \right. \\
&\quad \left. \prod_{k<i-j} [1 - p_{23}(j+k) - p_{24}(j+k) - p_{25}(j+k)] \right] v^i b_2.
\end{aligned}$$

sowie für \mathcal{B}_{b_3}

$$\begin{aligned}
\mathcal{B}_{b_3} &= \sum_{i=1}^{\omega-x-1} \left\{ \sum_{j<i} \left[\prod_{k<j} [1 - p_{12}(k) - p_{13}(k) - p_{14}(k) - p_{15}(k)] p_{13}(j) \right. \right. \\
&\quad \prod_{k<i-(j+1)} [1 - p_{34}(j+k+1) - p_{35}(j+k+1)] \\
&\quad + \sum_{k<j} \prod_{l<k} [1 - p_{12}(l) - p_{13}(l) - p_{14}(l) - p_{15}(l)] p_{12}(k) \\
&\quad \prod_{l<j-(k+1)} [1 - p_{23}(k+l+1) - p_{24}(k+l+1) - p_{25}(k+l+1)] \\
&\quad \left. \left. p_{23}(j) \prod_{l<i-(j+1)} [1 - p_{34}(j+l+1) - p_{35}(j+l+1)] \right] \right\} v^i b_3.
\end{aligned}$$

Für die Pflegerente in Zustand 4 erhalten wir

$$\begin{aligned}
\mathcal{B}_{1,b_4}(0) &= \sum_{i=0}^{\omega-x-1} P_{14}(0, i) v^i b_4 \\
&= \sum_{i=0}^{\omega-x-1} \sum_{j<i} \left[P_{11}(0, j) p_{14}(j) + P_{12}(0, j) p_{24}(j) + P_{13}(0, j) p_{34}(j) \right] p_{44}(j+1, i) v^i b_4 \\
&= \sum_{i=0}^{\omega-x-1} \sum_{j<i} \left[\prod_{k<j} [1 - p_{12}(k) - p_{13}(k) - p_{14}(k) - p_{15}(k)] p_{14}(j) \right. \\
&\quad + \sum_{k<j} \left\{ P_{11}(0, k) p_{12}(k) P_{22}(k+1, j) \right\} p_{24}(j) \\
&\quad + \sum_{k<j} \left\{ P_{11}(0, k) P_{13}(k) + P_{12}(0, k) p_{23}(k) \right\} P_{33}(k+1, j) p_{34}(j) \left. \right] \\
&\quad \prod_{k<i-(j+1)} [1 - P_{45}(k+j+1)] v^i b_4 \\
&= \sum_{i=1}^{\omega} \sum_{j<i} \left[\prod_{k<j} [1 - p_{12}(k) - p_{13}(k) - p_{14}(k) - p_{15}(k)] p_{14}(j) \right. \\
&\quad + \sum_{k<j} \left\{ \prod_{l<k} [1 - p_{12}(l) - p_{13}(l) - p_{14}(l) - p_{15}(l)] p_{24}(k) \right. \\
&\quad \left. \prod_{l<j-(k+1)} [1 - p_{23}(k+1+l) - p_{24}(k+1+l) - p_{25}(k+1+l)] \right\} p_{24}(j) \\
&\quad + \sum_{k<j} \left\{ \prod_{l<k} [1 - p_{12}(l) - p_{13}(l) - p_{14}(l) - p_{15}(l)] p_{13}(k) \right. \\
&\quad + \sum_{l<k} \prod_{m<l} [1 - p_{12}(m) - p_{13}(m) - p_{14}(m) - p_{15}(m)] p_{12}(l) \\
&\quad \left. \prod_{m<k-(l+1)} [1 - p_{23}(l+k+1) - p_{24}(l+k+1) - p_{25}(l+k+1)] p_{23}(k) \right\} \\
&\quad \left. \prod_{l<j-(k+1)} [1 - p_{34}(k+1+l) - p_{35}(k+1+l)] p_{34}(j) \right]
\end{aligned}$$

$$\prod_{k < i - (j+1)} [1 - P_{45}(k + j + 1)] v^i b_4.$$

Zur Bestimmung des Jahresbeitrags erhalten wir somit

$$\mathcal{B}_{1,b_2}(0) + \mathcal{B}_{1,b_3}(0) + \mathcal{B}_{1,b_4}(0) \mathcal{B}_{1,c_{12}}(0) + \mathcal{B}_{1,c_{13}}(0) + \mathcal{B}_{1,c_{14}}(0) = \mathcal{P}_{1,\pi}(0)$$

nach dem versicherungsmathematischen Äquivalenzprinzip. Auflösen nach π ergibt

$$\pi = \frac{\mathcal{B}_{1,b_2}(0) + \mathcal{B}_{1,b_3}(0) + \mathcal{B}_{1,b_4}(0) + \mathcal{B}_{1,c_{12}}(0) + \mathcal{B}_{1,c_{13}}(0) + \mathcal{B}_{1,c_{14}}(0)}{\sum_{i=0}^{\omega-x-1} \prod_{j < i} [1 - p_{12}(j) - p_{13}(j) - p_{14}(j) - p_{15}(j)] v^i}.$$

5.3 Beschreibung des Programms zur Beitragsberechnung

Hier wollen wir nun die für das in Abschnitt 5.2.1 vorgestellte Modell ein C-Programm entwickeln. Dieses beschreiben wir am Beispiel des Pflegekostentagegeldtarifs PET der Deutschen Krankenversicherung (DKV).

```

Terminal
Window Edit Options Help
Jaehrliche Pflegerente, ambulant, Stufe 1 ?
912.5
Jaehrliche Pflegerente, ambulant, Stufe 2 ?
1825
Jaehrliche Pflegerente, ambulant, Stufe 3 ?
2737.5
Jaehrliche Pflegerente, stationaer, Stufe 1 ?
3650
Jaehrliche Pflegerente, stationaer, Stufe 2 ?
3650
Jaehrliche Pflegerente, stationaer, Stufe 3 ?
3650
Hoehe der einmaligen Zahlung bei Uebergang aktiv -> ambulante Pflege (Stufe 1) ?
0
Hoehe der einmaligen Zahlung bei Uebergang aktiv -> ambulante Pflege (Stufe 2) ?
0
Hoehe der einmaligen Zahlung bei Uebergang aktiv -> ambulante Pflege (Stufe 3) ?
0
Hoehe der einmaligen Zahlung bei Uebergang aktiv -> stationaere Pflege (Stufe 1)
0
Hoehe der einmaligen Zahlung bei Uebergang aktiv -> stationaere Pflege (Stufe 2)
0
Hoehe der einmaligen Zahlung bei Uebergang aktiv -> stationaere Pflege (Stufe 3)
0
Hoehe des rechnerischen Zinssatzes?
0.035

```

Abbildung 5.4: Eingabeabfrage des Programms

5.3.1 Beschreibung des Tarifs PET

Der Pflegekostentagegeldtarif der DKV stellt eine Zusatzversicherung zur gesetzlichen Pflegeversicherung dar. Im Rahmen dieses Tarifs wird ein vom Versicherer zu zahlendes Pflegegeld vereinbart.

Der Leistungsfall wird im Stufensystem, analog zur gesetzlichen Pflegeversicherung, definiert. Der Versicherer verpflichtet sich im Pflegefall die folgenden Leistungen zu erbringen:

Für ambulante Pflege werden

- in der Pflegestufe 1 25 %
- in der Pflegestufe 2 50 %
- in der Pflegestufe 3 75 %

des vereinbarten Pfl egetagegeldes gezahlt.

Für stationäre Pflege werden 100 % des versicherten Tagegeldes gezahlt, wenn ambulante Pflege nicht möglich ist, oder wegen der Besonderheiten des einzelnen Falles nicht in Betracht kommt.

Falls der Pflegebedürftige vollstationäre Pflege wählt, obwohl diese nicht erforderlich ist, so erhält er

- in der Pflegestufe 1 25 %
- in der Pflegestufe 2 50 %
- in der Pflegestufe 3 75 %

des vereinbarten Pfl egetagegeldes.

5.3.2 Beschreibung des Programmes und Ergebnis der Beitragsberechnung

Alter	Programm		Werte DKV	
	Frauen	Männer	Frauen	Männer
20	2.81	2.31	2.12	1.70
25	3.50	2.89	2.92	2.33
30	4.41	3.65	3.90	3.10
35	5.63	4.67	5.05	4.01
40	7.28	6.06	6.44	5.13
45	9.48	7.91	8.16	6.52
50	12.49	10.42	10.39	8.36
55	16.67	13.93	13.32	10.86
60	22.56	18.90	17.31	14.40
65	30.98	26.04	22.01	18.84
70	42.91	36.40	29.04	25.71

Tabelle 5.1: Vergleich der Beiträge für den Tarif PET der DKV. Linke Spalten: Beiträge nach Berechnung mit dem C-Programm, rechte Spalten: Beiträge der DKV.

Als Eingabe benötigt das Programm die Höhe der vereinbarten jährlichen Pflegerenten, unterteilt nach Pflegestufe und Pflegeart (siehe Abbildung 5.4). Wir wollen den Tarif der DKV für ein

vereinbartes Pflagegeld in der Höhe von 10 DM modellieren. Das Tagesgeld für ambulante Pflege in Stufe 1 entspricht einer jährlichen Rente von

$$365 \cdot 0.25 \cdot 10 \text{ DM} = 912.5 \text{ DM}.$$

Das Pflagegeld in ambulanter Pflege in Stufe 2 (50 % des vereinbarten Satzes) entspricht einer jährlichen Rente von 1825 DM, das Pflagegeld für Stufe 3 einer Rente von 2737.5 DM. Für stationäre Pflege rechnen wir mit dem vollen Tagesgeldsatz von 10 DM, was einer jährlichen Rente von 3650 DM entspricht.

Als Ergebnis der Berechnungen erhalten wir Einmalbeiträge und jährliche gleichbleibende Beiträge. Für den Tarif PET sind gleichbleibende Monatsbeiträge angegeben, so daß wir die mit dem Programm berechneten Beiträge ebenfalls in Monatsbeiträge umrechnen, um so eine Vergleichsbasis herzustellen. Zusätzlich erhalten wir die für 1 DM normierten aktuariellen Werte \mathcal{B}_{b_2} für eine ambulante Pflagerente, sowie \mathcal{B}_{b_3} für eine stationäre Pflagerente (in Abhängigkeit von Alter und Geschlecht).

Die Berechnung der Beiträge für den Tarif PET brachte das in Tabelle 5.1 dargestellte Ergebnis. Wir sehen, daß die Beiträge der DKV deutlich niedriger sind, als die nach unserem Modell berechneten Werte. Ein Vergleich ist jedoch schwierig, da wir Bruttobeiträge der DKV gegeben haben, die noch zusätzliche Kosten wie z.B. Provisionen beinhalten. Zudem verwendeten wir unterschiedliche Rechnungsgrundlagen (japanische Pflegefalleintrittswahrscheinlichkeiten und bayerische Sterbewahrscheinlichkeiten). Interessant wäre eine Berechnung auf Basis aller Daten (d.h. Daten, die Informationen über die Sterblichkeit des gesamten versicherten Portefeuilles, sowie über die Häufigkeit von Pflegefällen enthalten) des betreffenden Versicherers.

Kapitel 6

Zusammenfassung und Ausblicke

Mit dieser Arbeit wollten wir die Möglichkeiten einer Anwendung des Cox-Regressions-Modells und anderer verweildaueranalytischer Methoden im Bereich der Pflegeversicherung untersuchen. Fassen wir noch einmal die einzelnen Schritte, die zur Herleitung des Modells nötig waren, zusammen.

Zunächst modellierten wir die Sterbeintensitäten mit Hilfe des Proportional-Hazard-Ansatzes in Abhängigkeit verschiedener Kovariablen. Dabei stellten wir zum einen fest, daß einige Kovariablen stark korreliert waren. Für die Kovariablen, die den Einfluß von Alter und Pflegestufe auf die Pflegedauer modellierten, stellten wir zudem eine starke Zeitabhängigkeit im Koeffizienten fest. Durch separates Schätzen der Koeffizienten für je 2 verschiedene Zeitintervalle konnte eine Verbesserung des Modells erzielt werden. Ziel weiterer Untersuchungen könnte in diesem Zusammenhang eine zeitstetige Modellierung von Koeffizienten sein, die bis jetzt mit den in S-Plus zur Verfügung stehenden Werkzeugen für Datensätze dieser Größe nicht praktikierbar ist.

In einem zweiten Schritt konstruierten wir ein zeitstetiges Markov-Modell, mit dessen Hilfe wir die Veränderungen im Pflegeverlauf hinsichtlich Pflegestufe und Pflegeart analysieren konnten. Mit Hilfe der Verallgemeinerung des Cox-Regressionsmodells für Zählprozesse war es möglich für dieses Modell Übergangintensitäten zu schätzen. Um analytische Lösungen für die Kolmogorov-Differentialgleichungen zu erhalten, wurde die vereinfachende Annahme getroffen, daß im Modell nur eine Verschlechterung im Pflegeverlauf möglich ist. Als Hauptargument dafür, daß diese Näherung eine relativ gute Approximation der Wirklichkeit ist, konnten wir die extrem niedrige Anzahl an Personen, die Verbesserungen im Pflegeverlauf erlebten, anführen.

Zuletzt wurde die Basis dafür geschaffen, diese Ergebnisse der Datenanalyse für ein Versicherungsmodells anzuwenden. Mit Hilfe der Implementierung der geschätzten Basis-Häufigkeiten in ein diskretes Versicherungsmodell hatten wir nun Möglichkeit, Beiträge für Pflegeversicherungsmodelle zu kalkulieren. Als Problem stellte sich hierbei heraus, daß unser Datensatz nur Personen beinhaltete, die schon pflegebedürftig waren. Schlüsse bezüglich der Zustandsübergänge von “aktiv lebend” zu “pflegebedürftig” konnten nicht gezogen werden.

In einer Anwendung entwickelten wir schließlich noch ein C-Programm zur Beitragsberechnung. Damit kalkulierten wir Beiträge für den Tarif PET der DKV. Es stellte sich heraus, daß die nach dem Markov-Modell kalkulierten (Netto-) Beiträge höher waren, als die (Brutto-) Beiträge der DKV. Obwohl die Ergebnisse mit Vorsicht zu genießen sind, da nur ein Teil der Rechnungsgrundlagen wirklich unserem Datensatz entsprang, geben sie doch Anlaß, die Berechnungsbasis von Pflegeversicherungsverträgen zu überprüfen und gegebenenfalls anzupassen.

“Wird’s besser, wird’s schlimmer?”

fragt man alljährlich.

Seien wir ehrlich:

Leben ist immer

lebensgefährlich

Erich Kästner

Anhang A

Routinen zur graphischen Analyse

Funktion cumloghazard

```
function(time1, surv1, time2, surv2)
{
  i <- 1
  j <- 1
  time <- max(time1[i], time2[j])
  if(time == time1[i]) {
    while(time2[j] < time) {
      j <- j + 1
    }
  }
  else {
    while(time1[i] < time) {
      i <- i + 1
    }
  }
  loghazard <- log(- log(surv1[i])) - log(- log(surv2[j]))
  while((i < length(time1) + 1) && (j < length(time2) + 1)) {
    if(time1[i] < time2[j]) {
      loghazard <- c(loghazard, log(- log(surv1[i])) - log(-
        log(surv2[j])))
      time <- c(time, time1[i])
      i <- i + 1
    }
    else {
      if(time1[i] > time2[j]) {
        loghazard <- c(loghazard, log(- log(surv1[i])) -
          log(- log(surv2[j])))
        time <- c(time, time2[j])
        j <- j + 1
      }
      else {
        loghazard <- c(loghazard, log(- log(surv1[i])) -
          log(- log(surv2[j])))
        time <- c(time, time2[j])
      }
    }
  }
}
```

```

        i <- i + 1
        j <- j + 1
    }
}
}
frame <- data.frame(time, loghazard)
plot(frame$time, frame$loghazard, xlab="", ylab="")
frame
}

```

Funktion andplot

```

function(time1, surv1, time2, surv2)
{
# Plot der kum-Hazardrate 2 verschiedener surv-objekte gegeneinander
  i <- 1
  j <- 1
  totaltime <- NULL
  hazardges1 <- NULL
  hazardges2 <- NULL
  while((i < length(time1)) && (j < length(time2))) {
    time <- min(time1[i], time2[j])
    if((time == time1[i]) && (time == time2[j])) {
      totaltime <- c(totaltime, time1[i])
      hazardges1 <- c(hazardges1, - log(surv1[i]))
      hazardges2 <- c(hazardges2, - log(surv2[j]))
      if(j < length(time2)) {
        j <- j + 1
      }
      if(i < length(time1)) {
        i <- i + 1
      }
    }
  }
  else {
    if(time == time1[i]) {
      totaltime <- c(totaltime, time1[i])
      hazardges1 <- c(hazardges1, - log(surv1[i]))
      if(time2[1] > time1[i]) {
        hazardges2 <- c(hazardges2, 0)
      }
      else {
        hazardges2 <- c(hazardges2, - log(surv2[j]))
      }
      if(i < length(time1)) {
        i <- i + 1
      }
    }
    else {
      totaltime <- c(totaltime, time2[j])
      hazardges2 <- c(hazardges2, - log(surv2[j]))
    }
  }
}

```

```
    if(time1[1] > time2[j]) {
      hazardges1 <- c(hazardges1, 0)
    }
    else {
      hazardges1 <- c(hazardges1, - log(surv1[i]))
    }
    if(j < length(time2)) {
      j <- j + 1
    }
  }
}
frame <- data.frame(totaltime, hazardges1, hazardges2)
plot(frame$hazardges2, frame$hazardges1, xlab = "", ylab = "")
frame
}
```


Anhang B

Funktionen zur Maximierung des Log-Likelihood

Funktion agetheta

```
function(daten, n, steps, beginagem, beginagew)
{
  # Funktion zur Maximierung des log-Likelihood
  # in Abhaengigkeit von thetaw und thetam
  attach(daten)
  result <- matrix(1:(n + 1) * (n + 1), n + 1, n + 1)
  tempm <- 1:length(daten[, 1])
  tempw <- 1:length(daten[, 1])
  nullo <- rep(0, length(daten[, 1]))
  tempm <- nullo
  tempw <- nullo
  thetam <- nullo
  thetaw <- nullo
  tempm[daten$Zsex == "m"] <- 1
  tempw[daten$Zsex == "w"] <- 1
  for(i in 1:(n + 1)) {
    for(j in 1:(n + 1)) {
      print(i)
      print(j)
      thetam <- nullo
      thetaw <- nullo # print(thetam)
      thetaw[daten$Age >= beginagew + (i - 1) * steps] <- 1
      thetam[daten$Age >= beginagem + (j - 1) * steps] <- 1
      thetaw <- tempw * thetaw
      thetam <- tempm * thetam
      temp <- data.frame(ltcdata, thetam, thetaw)
      result[i, j] <- coxph(Surv(LTCbegin, LTCend, Death) ~
        Age * (Art + Zsex) + Zsex * Art + Art * Level +
        thetam + thetaw, data = temp)$loglik[2]
      print(result[i, j])
    }
  }
}
```

```
    }  
  }  
  result  
}
```

Anhang C

Tabellen

C.1 Pflegefalleintrittswahrscheinlichkeiten der Custodial Insurance, Japan

Alter	i_x (Maenner)	i_y (Frauen)	Alter	i_x (Maenner)	i_y (Frauen)
20	0.00010	0.00010	40	0.00027	0.00027
21	0.00010	0.00010	41	0.00030	0.00030
22	0.00010	0.00010	42	0.00034	0.00034
23	0.00010	0.00010	43	0.00039	0.00039
24	0.00010	0.00010	44	0.00044	0.00044
25	0.00011	0.00011	45	0.00049	0.00050
26	0.00011	0.00011	46	0.00056	0.00056
27	0.00011	0.00011	47	0.00064	0.00064
28	0.00011	0.00011	48	0.00073	0.00072
29	0.00011	0.00011	49	0.00082	0.00082
30	0.00012	0.00012	50	0.00093	0.00093
31	0.00012	0.00012	51	0.00105	0.00105
32	0.00012	0.00012	52	0.00119	0.00119
33	0.00013	0.00013	53	0.00135	0.00135
34	0.00014	0.00013	54	0.00154	0.00154
35	0.00014	0.00014	55	0.00174	0.00174
36	0.00015	0.00015	56	0.00198	0.00198
37	0.00016	0.00016	57	0.00225	0.00225
38	0.00017	0.00017	58	0.00255	0.00255
39	0.00018	0.00018	59	0.00290	0.00290

Alter	i_x (Maenner)	i_y (Frauen)	Alter	i_x (Maenner)	i_y (Frauen)
60	0.00329	0.00329	80	0.04360	0.04369
61	0.00374	0.00374	81	0.04967	0.04980
62	0.00425	0.00425	82	0.05657	0.05678
63	0.00483	0.00483	83	0.06445	0.06475
64	0.00549	0.00549	84	0.07343	0.07386
65	0.00625	0.00625	85	0.08366	0.08427
66	0.00711	0.00710	86	0.09262	0.09336
67	0.00808	0.00808	87	0.10039	0.10120
68	0.00919	0.00919	88	0.10841	0.10933
69	0.01046	0.01046	89	0.11671	0.11777
70	0.01190	0.01190	90	0.12529	0.12657
71	0.01355	0.01354	91	0.13421	0.13579
72	0.01542	0.01542	92	0.14351	0.14542
73	0.01755	0.01755	93	0.15323	0.15550
74	0.01999	0.01998	94	0.16345	0.16612
75	0.02276	0.02276	95	0.17346	0.17651
76	0.02591	0.02592	96	0.18055	0.18373
77	0.02951	0.02953	97	0.18696	0.19014
78	0.03361	0.03364	98	0.19302	0.19604
79	0.03828	0.03834	99	0.19870	0.20136
			100	0.20405	0.20605

C.2 Einjährige Sterbewahrscheinlichkeiten Bayern für x -jährige Frauen und Männer (1986 bis 1988)

Alter	q_x (Männer)	q_y (Frauen)	Alter	q_x (Männer)	q_y (Frauen)
0	0.00830	0.00646	31	0.00119	0.00052
1	0.00072	0.00058	32	0.00126	0.00056
2	0.00049	0.00032	33	0.00131	0.00059
3	0.00041	0.00026	734	0.00138	0.00063
4	0.00035	0.00020	35	0.00148	0.00070
5	0.00030	0.00018	36	0.00159	0.00079
6	0.00026	0.00016	37	0.00174	0.00088
7	0.00024	0.00013	38	0.00191	0.00098
8	0.00024	0.00012	39	0.00210	0.00107
9	0.00022	0.00010	40	0.00228	0.00119
10	0.00020	0.00011	41	0.00247	0.00131
11	0.00019	0.00012	42	0.00270	0.00145
12	0.00020	0.00014	43	0.00296	0.00159
13	0.00023	0.00016	44	0.00327	0.00173
14	0.00032	0.00019	45	0.00360	0.00188
15	0.00050	0.00024	46	0.00395	0.00203
16	0.00076	0.00029	47	0.00435	0.00220
17	0.00102	0.00035	48	0.00480	0.00239
18	0.00123	0.00039	49	0.00531	0.00259
19	0.00132	0.00040	50	0.00590	0.00280
20	0.00131	0.00041	51	0.00654	0.00304
21	0.00126	0.00041	52	0.00724	0.00331
22	0.00121	0.00039	53	0.00800	0.00362
23	0.00117	0.00038	54	0.00882	0.00396
24	0.00112	0.00038	55	0.00972	0.00435
25	0.00109	0.00037	56	0.01068	0.00477
26	0.00106	0.00037	57	0.01171	0.00523
27	0.00104	0.00039	58	0.01281	0.00574
28	0.00104	0.00042	59	0.01398	0.00631
29	0.00106	0.00045	60	0.01523	0.00695
30	0.00111	0.00048	61	0.01523	0.00767

Alter	q_x (Männer)	q_y (Frauen)	Alter	q_x (Männer)	q_y (Frauen)
62	0.01806	0.00845	82	0.12011	0.08630
63	0.01970	0.00933	83	0.13116	0.09668
64	0.02153	0.01029	84	0.14295	0.10796
65	0.02357	0.01136	85	0.15550	0.12017
66	0.02583	0.01256	86	0.16881	0.13332
67	0.02835	0.01392	87	0.18290	0.14743
68	0.03114	0.01547	88	0.19778	0.16248
69	0.03426	0.01728	89	0.21345	0.17846
70	0.03774	0.01940	90	0.22991	0.19536
71	0.04164	0.02188	91	0.24715	0.21313
72	0.04597	0.02476	92	0.26515	0.23174
73	0.05077	0.02809	93	0.28391	0.25112
74	0.05607	0.03193	94	0.30338	0.27121
75	0.06190	0.03632	95	0.32353	0.29193
76	0.06830	0.04131	96	0.34431	0.31318
77	0.07530	0.04694	97	0.36569	0.33488
78	0.08293	0.05327	98	0.38759	0.35690
79	0.09120	0.06033	99	0.40995	0.37914
80	0.10015	0.06816	100	0.43271	0.40147
81	0.10978	0.07681	101	1.00000	1.00000

Literaturverzeichnis

- [1] P. K. Andersen, O. Borgan, R. D. Gill, and N. Keiding. *Statistical Models Based on Counting Processes*. Springer, New York, 1993.
- [2] P. K. Andersen and R. D. Gill. Cox's regression model for counting processes. *The Annals of Statistics*, 10:1100–1120, 1982.
- [3] T. Aven. A theorem for determining the compensator of a counting process. *Scandinavian Journal of Statistics*, 12:69–72, 1982.
- [4] H. Bauer. *Wahrscheinlichkeitstheorie*. de Gruyter, Berlin, 3rd edition, 1978.
- [5] O. Bie and Ø. Borgan. Confidence intervals and confidence bands for the cumulative hazard rate function and their small sample properties. *Scandinavian Journal of Statistics*, 14:221–233, 1987.
- [6] Ø. Borgan and K. Listøl. A note on confidence intervals and bands for the survival curve based on transformations. *Scandinavian Journal of Statistics*, 17:35–41, 1990.
- [7] P. Brémaud. *Point Processes and Queues; Martingale Dynamics*. Springer, New York, 1981.
- [8] N. Breslow. Covariance analysis of censored survival data. *Biometrics*, 30:89–99, 1974.
- [9] D. R. Cox. Regression models and life-tables. *Journal of the Royal Stat. Society*, B 34:187–220, 1972.
- [10] D.R. Cox. Partial likelihood. *Biometrika*, 62:269–276, 1975.
- [11] R. Durrett. *Brownian Motion and Martingales in Analysis*. Wadsworth, Belmont, 1984.
- [12] B. Efron. The efficiency of cox's likelihood function for censored data. *Journal of the American Statistician Association*, 72:557–565, 1977.
- [13] J. Elstrodt. *Maß- und Integrationstheorie*. Springer, Heidelberg, 1996.
- [14] Bayerisches Landesamt für Statistik und Datenverarbeitung. *Statistisches Jahrbuch 1998*. 1998.

- [15] L. Fahrmeier and G. Tutz. *Multivariate Statistical Modelling Based On Generalized Linear Models*. Springer, New York, 1994.
- [16] H. U. Gerber. *Life Insurance Mathematics*. Springer, Heidelberg, 3rd edition, 1997.
- [17] R. D. Gill. Non- and semi-parametric maximum likelihood estimators and the von mises method. *Scandinavian Journal of Statistics*, 16:97–128, 1989.
- [18] M. Greenwood. The natural duration of cancer. *Reports on Public Health and Medical Subjects*, 33:1–26, 1926.
- [19] S. Haberman and E. Pitacco. *Actuarial Models for Disability Insurance*. Chapman & Hall/CRC, 1999.
- [20] K. Königsberger. *Analysis 1*. Springer, Heidelberg, 2nd edition, 1992.
- [21] J. D. Kalbfleisch and R. L. Prentice. *The Statistical Analysis of Failure Data*. Wiley, New York, 1980.
- [22] E. L. Kaplan and P. Meier. Non-parametric estimation from incomplete observations. *Journal of the Amer. Statist. Assoc.*, 53:457–481, 562 – 563, 1958.
- [23] B. W. Kerningham and D. M. Ritchie. *The C Programming Language*. Prentice Hall, New Jersey, 2nd edition, 1988.
- [24] J. P. Klein and N. Keiding. Plotting summary predictions in multistate survival models. *Statistics in Medicine*, 13:2315–2331, 1994.
- [25] J. P. Klein and M. L. Moeschberger. *Survival Analysis*. Springer, New York, 1997.
- [26] C. T. Le. *Applied Survival Analysis*. Wiley, New York, 1997.
- [27] E. Lengkart. Relation de domination entre deux processus. *Ann. Inst. Henri Poincaré*, 13:171–179, 1977.
- [28] P. McCullagh and J. A. Nelder. *Generalized Linear Models*. Chapman & Hall, London, 1983.
- [29] R. G. Jr. Miller. *Survival Analysis*. Wiley, New York, 1981.
- [30] R. H. Myers. *Classical And Modern Regression With Applications*. Duxbury, Boston, 2nd edition, 1990.
- [31] W. Nelson. Theory and applications of hazard plotting for censored failure data. *Technometrics*, 14:945–966, 1972.
- [32] C. R. Rao. *Linear Statistical Inference and Its Applications*. Wiley, New York, 1965.

- [33] R. Rebolledo. Central limit theorems for local martingales. *Z. f. Wahrscheinlichkeitstheorie u. verw. Gebiete*, 51:269–286, 1980.
- [34] L. C. G. Rogers and D. Williams. *Diffusions, Markov Processes and Martingales: Foundations*, volume 1. Wiley, Chichester, 2nd edition, 1979.
- [35] L. C. G. Rogers and D. Williams. *Diffusions, Markov Processes and Martingales: Itô-Calculus*, volume 2. Wiley, Chichester, 1987.
- [36] Münchener Rückversicherung. *Die Absicherung des Pflegefallrisikos*. Münchener Rückversicherungsgesellschaft, 1992.
- [37] D. Schoenfeld. Partial residuals for the proportional hazards regression model. *Biometrika*, 69:239–241, 1982.
- [38] D. Stirzaker. *Elementary Probability*. Cambridge University Press, 1994.
- [39] T. M. Therneau, P. M. Grambsch, and Thomas R. Fleming. Martingale-based residuals for survival models. *Biometrika*, 1:147–160, 1990.
- [40] W. N. Venables and B. D. Ripley. *Modern Applied Statistics With S-Plus*. Springer, New York, 1994.