Garczorz, Ingo

Anwendung der Hazard-Analyse im Marketing: Einführung und Literaturüberblick

Manuskripte aus den Instituten für Betriebswirtschaftslehre der Universität Kiel, No. 548

Provided in Cooperation with:
Christian-Albrechts-University of Kiel, Institute of Business Administration

Suggested Citation: Garczorz, Ingo (2001) : Anwendung der Hazard-Analyse im Marketing: Einführung und Literaturüberblick, Manuskripte aus den Instituten für Betriebswirtschaftslehre der Universität Kiel, No. 548, Universität Kiel, Institut für Betriebswirtschaftslehre, Kiel

Nr. 548

Anwendung der Hazard-Analyse im Marketing

Einführung und Literaturüberblick

Ingo Garczorz

November 2001

Dipl.-Kfm. Ingo Garczorz
Institut für Betriebswirtschaftslehre
Lehrstuhl für Innovation, Neue Medien und Marketing (Prof. Dr. Sönke Albers)
Christian-Albrechts-Universität zu Kiel
Westring 425

Gliederung

Verzeichnis der Abbildungen

Verzeichnis der Tabellen

Verzeichnis der Abkürzungen

bzw.	beziehungsweise
d.h.	das heißt
et al.	et alii (und andere)
f.	und folgende Seite
ff.	und folgende Seiten
Hrsg.	Herausgeber
S.	Seite
Tab.	Tabelle
u.U.	unter Umständen
vgl.	vergleiche

Symbolverzeichnis

α	Parameter einer Verteilungsfunktion
A_i	Zensierungsindikator (Dummy-Variable) für das i-te Individuum
β_0	Konstante
β_k	Koeffizient der k-ten unabhängigen Variablen x_k ($k \in K$)
$\hat{\beta}$	geschätzter Koeffizient
C	Informationskriterium
C_i	Zensierungsindikator (Dummy-Variable) für das i-te Individuum
δ_i	Splitparameter für das i-te Individuum
d	Strafkomponente bei den Informationskriterien
Φ	Verteilungsparameter einer normalverteilten Zufallsvariable
$E(\cdot)$	Erwartungswert
ε	Heterogenitätskomponente
$f(\cdot)$	Dichtefunktion
$F(\cdot)$	Verteilungsfunktion
$g(\cdot)$	Transformationsfunktion
$G(\cdot)$	Verteilungsfunktion des Störterms ε
$\Gamma(\cdot)$	Gammafunktion
$h(\cdot)$	Hazardfunktion
$H(\cdot)$	kumulierte Hazardrate
$\hat{H}(\cdot)$	geschätzte kumulierte Hazardrate
γ	Split-Koeffizient
I	Indexmenge der Individuen bzw. Größe des Stichprobenumfanges
η	Parameter einer Verteilungsfunktion
λ_i	Parameter einer Verteilungsfunktion für das i-te Individuum

$\tilde{p}$	Anzahl der Parameter bei den Informationskriterien
$P_i = Prob(y_i = 1)$	Ereigniswahrscheinlichkeit des i-ten Individuums
π	Parameter einer Verteilungsfunktion
$\hat{r}_i$	Cox-Snell-Residuen des i-ten Individuums
$\hat{\sigma}^2$	Varianz
t_i	Episodendauer des i-ten Individuums
T	Endzeitpunkt der Beobachtungsperiode
θ	Parameter einer Verteilungsfunktion
$R(\cdot)$	Risikomenge bzw. Risk Set
$S(\cdot)$	Survivorfunktion
$\hat{S}(\cdot)$	geschätzte Survivorfunktion
$Var(\cdot)$	Varianz
ω	Proportionalitätskonstante
x_{ik}	Ausprägung der k-ten unabhängigen Variablen beim i-ten Individuum
X_i	Vektor der unabhängigen Variablen beim i-ten Individuum bzw. Vektor der Indikatorenvariablen beim i-ten Individuum
χ^2	Chi-Quadrat-Wert
z_1, z_2	Dummy-Variable

1 Einleitung

Regressionsanalytische Verfahren wie die lineare oder logistische Regression gehören zu den am häufigsten verwendeten Werkzeugen bei der Untersuchung multivariater Zusammenhänge in der Betriebswirtschaftslehre und hier insbesondere im Bereich des Marketing. Einhergehend mit dem zunehmenden Interesse an dynamischen Prozessen (Blossfeld, Hamerle und Mayer 1989, S. 235), deren Analyse aufgrund erhöhter Datenverfügbarkeit z.B. im Bereich der Neuen Medien ermöglicht wird, zeigen sich aber auch sehr schnell die Grenzen dieser Verfahren. Probleme wie die Behandlung zensierter Beobachtungen oder die Berücksichtigung zeitvariabler Einflussfaktoren können nicht adäquat berücksichtigt und abgebildet werden bzw. führen zu unvermeidlichen Verzerrungen der Untersuchungsergebnisse (vgl. dazu ausführlich Allison 1995, S. 4 Helsen und Schmittlein 1993, S. 399 f., Hutchinson 1988, S. 207 f., Litfin 2000, S. 61 ff. und Peterson 1991, S. 273 f.). Vor diesem Hintergrund rückt ein weiteres regressionsanalytisches Verfahren in den Blickpunkt des Interesses, das bis jetzt nur eine vergleichsweise geringe Aufmerksamkeit im Kontext der Betriebswirtschaftslehre wohl aber in anderen Wissenschaftsdisziplinen wie der Medizin und dem Ingenieurwesen erfahren hat (Hutchinson 1988, 205): die Hazard-Analyse. Aufgrund der vergleichsweise geringen Anzahl von Untersuchungen, die sich bisher dieses Instrumentes bedient haben, ist es das Ziel der nachfolgenden Ausführungen, dem Leser einen komprimierten Überblick über Idee und Vorgehensweise bei der Hazard-Analyse zu vermitteln. Damit werden die Hinweise von Mahajan/Muller/Bass (Mahajan, Muller und Bass 1990, S. 19 f.) und Rangaswamy/Gupta (Rangaswamy und Gupta 1999, S. 24 ff.) aufgegriffen, die dieses Verfahren für eines der am erfolgsversprechensten Instrumente bei der Analyse von Adoptions- bzw. Diffusionsprozessen, die exemplarisch für das Interesse an dynamischen Prozessen an sich stehen, halten.

2 Die Hazard-Analyse als Instrument zur Untersuchung dynamischer Prozesse

Die Hazard-Analyse[1] ist auf die Untersuchung des Auftretens von Ereignissen, als deren Folge sich ein Zustandswechsel beim Untersuchungsobjekt ergibt, im Zeitab-

[1] In der Literatur existieren zahlreiche Bezeichnungen für dieses Verfahren, u.a. die Bezeichnungen Ereignisanalyse, Event-History-Ansatz bzw. Übergangsratenmodelle, Überlebensmodelle und Survivormodelle. Für eine Einordnung dieser Modelle in die Ökonometrie vgl. Hansen 1991, S. 390 ff.

lauf auf Basis von Längsschnitts- und Paneldaten zugeschnitten (Blossfeld, Hamerle und Mayer 1986, S. 27). Im Gegensatz zu statischen Modellen wie der logistischen Regression oder der Diskriminanzanalyse ist aber nicht nur der Zustand zu einem bestimmten Zeitpunkt bedeutsam, sondern es interessiert zusätzlich noch die Zeitdauer bis zum Zustandswechsel. Die Verwendung dieser Information ist von Bedeutung, da die Wahrscheinlichkeit, dass es zu einem Zustandswechsel kommt, bei einer Vielzahl von Fragestellungen abhängig ist von der bereits vergangenen Verweildauer des Individuums in einem bestimmten Anfangszustand. Dies gilt z.B. für Adoptions- bzw. Diffusionsprozesse, bei denen die Übernahme allein aufgrund der Tatsache, dass sich ein Produkt über einen längeren Zeitraum zunehmend am Markt etabliert, wahrscheinlicher wird. Ähnlich gelagerte Zusammenhänge kann man auch für das (Wieder-)Kaufverhalten oder Fragen der Kundenbindung vermuten. Mit Hilfe der Hazard-Analyse kann diese funktionale Abhängigkeit des untersuchten Prozesses von der Zeit explizit berücksichtigt werden. Darüber hinaus können zeitveränderliche Einflussgrößen in die Analyse integriert werden. Dies ist ein weiterer entscheidender Vorteil gegenüber der Anwendung einer linearen bzw. logistischen Regression. Ferner erlauben Hazard-Modelle eine adäquate Erfassung der bereits erwähnten Zensierungsproblematik, ohne dass es zu Verzerrungen der geschätzten Parameter und den daraus möglicherweise resultierenden fehlerhaften Implikationen kommt. Diese Verbesserungen bzw. Vorteile der Modellierung der Zeitdauerabhängigkeit mit Hilfe der Hazard-Analyse erlauben umfassendere Aussagen bezüglich der Vorteilhaftigkeit von Promotionsmaßnahmen, der Bewertung von Kunden, des Timings von Marketingaktivitäten, des Managements eines existierenden Kundenbestandes (Helsen und Schmittlein 1993, S. 412). Dennoch gibt es erst vergleichsweise wenige Arbeiten im Bereich des Marketing wie die nachfolgende Bestandsaufnahme in Tabelle 2-1 der neueren empirischen Forschung zeigt.

Tabelle 2-1: Hazard Modelle im Marketing

Verfasser	Stichprobe/ Branchenfokus	Zu untersuchendes Phänomen	Modelltyp	Inhaltliche und/oder methodische Hauptbefunde	Besonderheiten
Kaufverhalten bei Konsumgütern: Wirkung von Marketing Mix Variablen					
Chintagunta 1998	Scanner-Paneldaten auf Haushaltsebene (400 Haushalte)/Konsumgüter	simultane Untersuchung von Kaufzeitpunkt und Markenwahl	Proportional Hazard (Box-Cox-Formulierung der Baseline) und Parametrische Formulierung: Log-Logistische Formulierung	Haushalte kaufen in relativ regelmäßigen Abständen und sind zu diesen Zeitpunkten relativ preissensibel.	Positive Log Likelihood Werte
Chintagunta und Haldar 1998	Haushaltsdaten/Konsumgüter	Abhängigkeit von Kaufintervallen in komplementären oder substitutiven Produktkategorien	Bivariate Hazardformulierung basierend auf Farlie-Gumble-Morgenstern-Verteilungsfamilie	Signifikanter und erwartungskonformer Zusammenhang zwischen Käufen in verwandten Produkt-Kategorien	Erste Studie im Marketing, in der ein bivariater Hazard-Ansatz verwendet wird
Gönül und Srinivasan 1993	84 Haushalte / Konsumgüter	Kaufverhalten insbesondere Markenwechsel sowie methodisch: Modellvergleich	Proportional Hazard und Parametrische Formulierung: Quadratische Gompertz Verteilung	Random-Effekt Modell mit Gompertz Formulierung der Hazard-Rate erweist sich als überlegen	Competing Risk Modell mit Heterogenität
Helsen und Schmittlein 1993	Paneldaten/Konsumgüter	Kaufintervalle	Parametrische Formulierung: Weibull und Quadratisch	Sonderpreise haben einen erwartungskonformen aber nicht signifikanten Einfluss.	Vergleich mit Probit- und Regressionsanalyse

Verfasser	Stichprobe/ Branchenfokus	Zu untersuchendes Phänomen	Modelltyp	Inhaltliche und/oder methodische Hauptbefunde	Besonderheiten
Hruschka, Stoiber und Hamerle 1998	236 Haushalte/Konsumgüter	Kaufverhalten insbesondere Markenwahl	Parametrische Formulierung: Weibull, Erlang-2, Log-Logistisch	Bezogen auf Käufe innerhalb einer Produktkategorie haben Promotionmaßnahmen keine intervall-verkürzende Wirkung wohl aber auf die Wiederkaufintervalle einzelner Marken	Mehrstufige Modellierung auf Produktkategorie- und Markenebene
Jain 1991	427 bzw. 166 Haushalte[2] / Konsumgüter	Kaufintervalle	Proportional Hazard (Box-Cox-Formulierung der Baseline)	Nicht-monotoner Verlauf der Hazard-Rate, d.h. die Zeitdauer bis zum nächsten Kauf steigt zunächst an und nimmt im weiteren Verlauf ab. Folglich können Kaufintervalle nicht mit den im allgemeinen verwendeten Verteilungen abgebildet werden.	Vergleich der Modellierungsmöglichkeit für unbeobachtete Heterogenität, S. 16 f.

[2] Es fand eine Segmentierung nach Art des untersuchten Produktes, hier Kaffee, statt. Wobei zwischen gemahlenem und löslichem Kaffee unterschieden wurde (Jain 1991, S. 9)

Verfasser	Stichprobe/ Branchenfokus	Zu untersuchendes Phänomen	Modelltyp	Inhaltliche und/oder methodische Hauptbefunde	Besonderheiten
Neuproduktentwicklung und -einführung					
Bähr-Seppelfricke 1999	Privathaushalte / Haushaltsgeräte, Unterhaltungs- und Kommunikationselektronik sowie Telekommunikation	Modellierung von Wiederholungskäufen aufgrund von Ersatzbeschaffung im Rahmen einer Diffusionsstudie	Parametrische Formulierung: Vergleich Weibull- und Gammaverteilungsfamilie	Wahl der Verteilungsfamilie hat keinen entscheidenden Einfluss auf die Ergebnisse	
Chandrashekaran und Sinha 1995	3236 Firmen unterschiedlicher Branchen	Focus auf die Diffusion von PCs	SPOT Modell, das sich – vereinfacht ausgedrückt- als Kombination aus Probit-/Tobit-/Hazard- bzw. Split-Hazard Modell darstellt	Signifikanter Einfluss von Firmencharakteristika auf das Adoptionsverhalten. Modellvergleich belegt Überlegenheit des SPOT-Modells	Modifiziertes Split-Hazard Modell
Litfin 2000	1005 Haushalte / Telekommunikation	Adoption eines innovativen Telekommunikationsdienstes	Parametrische Formulierung: Log-Normal, Log-Logistisch, Weibull und Exponential im Vergleich	Hypothesenkonforme, aber nur teilweise signifikante Einflüsse der auf den Rogerskriterien fußenden Adoptionsfaktoren, Überlegenheit des Split-Ansatzes	Split-Hazard Modell

Verfasser	Stichprobe/ Branchenfokus	Zu untersuchendes Phänomen	Modelltyp	Inhaltliche und/oder methodische Hauptbefunde	Besonderheiten
Sinha und Chandrashekaran 1992	3689 Banken	Diffusion von Geldausgabeautomaten	Parametrische Formulierung: Log-Normal, Weibull und Exponential sowie Vergleich der dominanten Log-Normalen- und Weibull-Spezifikation mit einem Proportional Hazard-Ansatz	Unterscheidung in Wahrscheinlichkeit und Zeitpunkt der Adoption sowie Zuordnung der Wirkung der Kovariate auf diese Aspekte von „Innovativeness" wird ermöglicht	Split-Hazard Modell
Kundenbeziehungsmanagement					
Li 1995	Kundenclub-Kundendaten / Telekommunikation	Beendigung der Mitgliedschaft	Proportional Hazard	Identifikation wechselfreudiger Kunden, Kundensegmentierung	Vergleich mit logistischer Regression
Bolton 1998	Kundenzufriedenheits- und Nutzungsdaten / Telekommunikation	Zusammenhang Kundenzufriedenheit und –bindung	Proportional Hazard	Kumulierte Zufriedenheit hat direkt und indirekt Einfluss auf die Bindung	Hinweis auf Bias bei OLS (S.56)

Verfasser	Stichprobe/ Branchenfokus	Zu untersuchendes Phänomen	Modelltyp	Inhaltliche und/oder methodische Hauptbefunde	Besonderheiten
Personalwirtschaft					
Darden, Hampton und Boatwright 1987	495 Angestellte / Einzelhandel	Dauer des Beschäftigungsverhältnisses	Proportional Hazard	Dauer des Beschäftigungsverhältnisse hängt vom Alter und der Position der Angestellten ab	Berücksichtigung zeitvariabler Kovariate
Hoverstad, Moncrief und Lucas 1990	3 Beschäftigungskohorten mit insgesamt 8923 Angestellten / Versicherungen	Dauer des Beschäftigungsverhältnisses	Nicht näher spezifizierte nicht-parametrische Sterbetafel Methode	Teilzeitangestellte weisen längere Beschäftigungs-verhältnisse auf als Vollzeitangestellte	
Moncrief III, Hoverstad und Lucas Jr. 1989	2411 Verkaufsaußendienstmitarbieter / Versicherungen	Dauer des Beschäftigungsverhältnisses	Nicht näher spezifizierte nicht-parametrische Sterbetafel Methode	Mitarbeiter, die höhere Umsätze erzielen, sind im Durchschnitt länger beschäftigt	

3 Statistisches Grundkonzept für den Ein-Episoden-Fall

Mit Hilfe der Hazard-Analyse wird der Übergang eines Untersuchungsobjektes von einem Anfangs- in einen oder mehrere Endzustände unter Berücksichtigung der Zeitdauer zwischen den Zustandswechseln untersucht. Dabei gilt, dass die Menge der potentiellen (End-) Zustände abzählbar ist. Diese Zustandswechsel bzw. Ereignisse können zu jedem beliebigen Zeitpunkt eintreten. Es werden also stochastische Prozesse mit stetiger Zeit und einer endlichen Anzahl von möglichen (End-) Zuständen untersucht,[3] wobei die Zeitdauer zwischen aufeinanderfolgenden Ereignissen als „Episode" bezeichnet wird.

Zur formalen Beschreibung dieser stochastischen Prozesse werden im folgenden die nachstehend aufgeführten Größen verwendet:

I: Menge der zu untersuchenden Individuen,

t_i: Episodenlänge des i-ten Individuums ($i \in I$, $t_i \geq 0$),

X_i: Vektor der erklärenden Kovariablen des i-ten Individuums ($i \in I$),

$F(t_i)$: Verteilungsfunktion der Episodendauer des i-ten Individuums ($i \in I$),

$F'(t_i)$: abgeleitete Verteilungsfunktion der Episodendauer des i-ten Individuums ($i \in I$),

$f(t_i)$: Dichtefunktion der Episodendauer des i-ten Individuums ($i \in I$),

$P(T \leq t_i)$: Adoptionswahrscheinlichkeit des i-ten Individuums ($i \in I$) während der Beobachtungsperiode,

$(0, T]$: Länge des Beobachtungszeitraumes.

Im einfachsten Fall ergibt sich ein Ein-Episoden-Modell mit einem Anfangs- und einem Endzustand. Die Episodendauer t_i ist je nach untersuchtem Individuum unterschiedlich und folgt in der Stichprobe einer spezifischen Verteilung, die mit Hilfe einer Verteilungsfunktion $F(t_i)$ oder auch der zugehörigen Dichtefunktion $f(t_i)$ beschrieben werden kann. Da der Bebachtungszeitraum die Länge $(0, T]$ hat, ergibt

---

[3] Auf eine Beschreibung zeitdiskreter Modelle wird an dieser Stelle verzichtet. Diese ergeben sich bei Modifikation des hier vorgestellten Grundmodells (vgl. dazu unter anderem Hamerle und Tutz 1989, S. 18 ff.). Zudem führen beide Varianten zu sehr ähnlichen Ergebnissen, wenn die Stichprobe groß genug ist und die Zeitintervalle entsprechend klein sind (vgl. die Ergebnisse der Monte-Carlo-Simulation von Galler 1985, S. 24 f.).

sich folgender Zusammenhang zwischen der Dichtefunktion und der kumulativen Verteilungsfunktion (vgl. dazu ausführlich insbesondere Allison 1995, S. 14 ff.; Andreß 1985, S. 45 ff.; Blossfeld, Hamerle und Mayer 1986, S. 31 ff.; Kalbfleisch und Prentice 1980, S. 21 ff.; Ronning 1991, S. 171 ff.):

$$(3\text{-}1) \qquad F(t_i) = P(T \leq t_i) = \int_0^{t_i} f(v_i)dv\,.$$

Die kumulierte Verteilungsfunktion gibt die Wahrscheinlichkeit an, mit der bis zum Zeitpunkt t_i das interessierende Ereignis beim i-ten Probanden eingetreten ist.

An allen Stellen, an denen F(t) differenzierbar ist, gilt

$$(3\text{-}2) \qquad f(t_i) = F'(t_i)\,.$$

Die Dichtefunktion f(t) gibt die Wahrscheinlichkeit an, dass das Ereignis in einem marginal kleinen Zeitintervall eintritt.

Intuitiv verständlicher ist die Interpretation der sogenannten Survivorfunktion. Diese gibt die Wahrscheinlichkeit an, dass bis zum Zeitpunkt t_i beim i-ten Individuum noch kein Ereignis eingetreten ist, die Episode also noch andauert. Es wird also die zu Gleichung (3-1) komplementäre Aussage formuliert:

$$(3\text{-}3) \qquad S(t_i) = 1 - F(t_i) = P(T > t_i)$$

Survivorfunktionen haben einen in Abhängigkeit von der Zeit monton fallenden Verlauf. Dies ist einsichtig, da mit zunehmender Zeitdauer bei immer mehr Probanden das interessierende Ereignis und damit ein Zustandswechsel eintreten wird.

Neben den Formulierungen in (3-1) bis (3-3) stellt die Hazard-Rate[4] eine weitere Möglichkeit zur Beschreibung dieses Zufallsprozesses das. Die Hazard-Rate ist das zentrale Element der Ereignisanalyse. Diekmann und Mitter (Diekmann und Mitter 1984, S. 42) veranschaulichen die Hazard-Rate als „Häufigkeit von Zustandswechseln in einem sehr kleinen Zeitintervall dividiert durch alle ‚Überlebenden', d.h. Kandidaten für einen Zustandswechsel [...]", so dass der in den Gleichungen (3-4) bzw. (3-5) formulierte Zusammenhang auch intuitiv verständlich ist:

---

[4] Für die Hazard-Rate existieren analog zum Grundmodell eine Fülle von Bezeichnungen, wie z. B. Intensitäts- oder Risikofunktion, Übergangs- oder Mortalitätsrate (Blossfeld, Hamerle und Mayer 1986, S. 31).

$$(3\text{-}4) \qquad h(t) = \frac{f(t)}{S(t)} = \frac{f(t)}{1-F(t)}.$$

Eine zu Gleichung (3-4) äquivalente Darstellung der Hazard-Rate ist gegeben durch:[5]

$$(3\text{-}5) \qquad h(t_i) = \lim_{\substack{\Delta t_i \to 0 \\ \Delta t_i > 0}} \frac{1}{\Delta t_i} P(t_i \le T < t_i + \Delta t_i \mid T \ge t_i).$$

Mit Hilfe der Hazard-Rate kann die Wahrscheinlichkeit bzw. das Risiko, dass das interessierende Ereignis, z.B. die Adoption einer Innovation oder der Wiederkauf eines Produktes, zum Zeitpunkt t_i beim Probanden i eintritt, quantifiziert werden. Die Darstellung in Gleichung (3-5) lässt sich wie folgt motivieren: Da T eine kontinuierliche Variable darstellt, ist die Wahrscheinlichkeit, dass ein Ereignis zu einem ganz bestimmten Zeitpunkt t_i eintritt, infinitesimal klein (Allison 1984, S. 23). Aus diesem Grund wird nicht ein Zeitpunkt, sondern ein sehr kleines Zeitintervall $[t_i;\ t_i + \Delta t_i)$ bzw. $(t_i \le T \le t_i + \Delta t_i)$ betrachtet. Aufgrund dieser Formulierung wäre es nun möglich durch die Wahl eines genügend großen Δt_i die Nutzungswahrscheinlichkeit zu inflationieren, indem man das Intervall genügend groß formuliert. Dies wird durch die Grenzwertbildung ($\lim \Delta t_i \to 0$) und durch Division der Wahrscheinlichkeit mit der Größe des Zeitraums ($1/\Delta t_i$) verhindert (Allison 1995, S. 15 f.; Kleinbaum 1995, S. 10 ff.).[6] Darüber hinaus muss sichergestellt sein, dass nur solche Individuen betrachtet werden, bei denen das interessierende Ereignis noch nicht eingetreten ist, da es sinnlos ist, das Risiko des Eintretens eines Ereignisses für das Zeitintervall $[t_i, t_i + \Delta t_i)$ anzugeben, wenn der Zustandswechsel bereits stattgefunden hat. Die Hazard-Rate kann folglich aufgefasst werden als der Grenzwert der bedingten Wahrscheinlichkeit, dass das Ereignis „Nutzung" im Zeitintervall $[t_i, t_i + \Delta t_i)$ stattfindet unter der Voraussetzung, dass bis zum Beginn dieses Intervalls noch keine Nutzung stattgefunden hat $(T \ge t_i)$. Wenn das i-te Individuum den Zeitpunkt t_i „überlebt", informiert die Hazard-Rate näherungsweise über den weiteren Verlauf der Eintrittswahrscheinlichkeit des Ereignisses. Die Hazard-Rate kann dabei sehr unterschiedliche Verläufe aufweisen. Die einzige Restriktion ist die Annahme nicht-negativer Ha-

---

[5] Hamerle (Hamerle 1987, S. 250) zeigt anhand dieser Formel, dass es sich bei dem Hazard-Ansatz um eine alternative Darstellung des Diffusionsmodells von Bass handelt (vgl. hierzu Bass 1969 und Mahajan, Muller und Srivastava 1990, S. 3 f.).

[6] Anzumerken ist, dass die Hazard-Rate keine bedingte Wahrscheinlichkeit, sondern lediglich eine bedingte Dichte darstellt (Arminger 1988, S. 79). Dies wird bei Annahme stetiger Zufallsvariablen deutlich, da die Hazard-Rate dann Werte größer als 1 annehmen kann. Daher sollte sie nur bei sehr kleinen Änderungen der Zufallsvariablen als Übergangswahrscheinlichkeit interpretiert werden (Blossfeld, Hamerle und Mayer 1986, S. 32).

zard-Raten. Abschließend sei noch darauf hingewiesen, dass das Integral der Hazard-Rate

$$(3\text{-}6) \qquad H(t) = \int_0^t h(v)dv$$

als kumulierte Hazard-Rate bezeichnet wird. Diese Größe spielt eine Rolle bei der Konstruktion grafischer Modelltests, die in Abschnitt 5.5 erläutert werden.

Der Zusammenhang zwischen den einzelnen Funktionen kann mit Hilfe der Gleichungen (3-1) bis (3-4) hergeleitet werden (Allison 1995, S. 16, Blossfeld, Hamerle und Mayer 1986, S. 33; Klein und Moeschberger 1997, S. 35), so dass die verschiedenen Funktionsformen letztendlich äquivalente Darstellungsweisen sind, die man ineinander überführen kann (Allison 1995, S. 16; Andreß 1985, S. 45; Kleinbaum 1995, S. 11). Tabelle 3-1 verdeutlicht diesen Zusammenhang.

Tabelle 3-1: Zusammenhänge zwischen den einzelnen Funktionen

gesucht / gegeben	Survivorfunktion	Dichtefunktion	Hazard-Rate
Survivor-funktion	--	$-S'(t_i)$	$\frac{-S'(t_i)}{S(t_i)}$
Dichte-funktion	$\int_{t_i}^{\infty} f(v_i)dv$	--	$\frac{f(t_i)}{\int_{t_i}^{\infty} f(v_i)dv}$
Hazard-rate	$\exp(-\int_0^{t_i} h(v_i)dv)$	$h(t_i) \times \exp(-\int_0^{t_i} h(v_i)dv)$	--

Quelle: In Anlehnung an: Wangler 1997, S. 15.

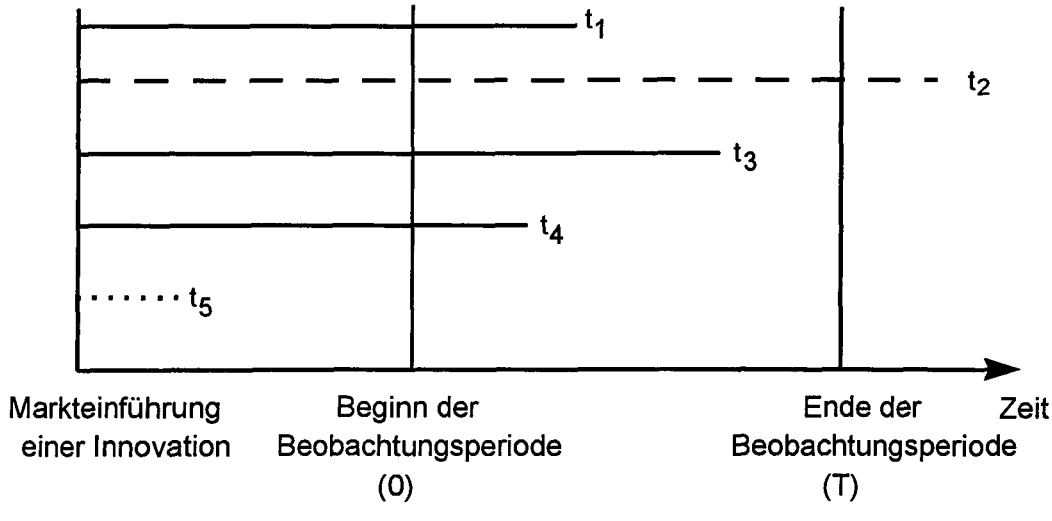
Die Hazard-Rate, Dichtefunktion und die Survivorfunktion sind also äquivalente Formen, um die kontinuierliche Wahrscheinlichkeitsverteilung der Episodendauer zu beschreiben. Wenn zum Beispiel die Survivorfunktion bekannt ist, ergibt sich die zugehörige Dichtefunktion aus der negativen Ableitung dieser Funktion. Obwohl der Prozessverlauf bereits durch eine dieser Funktionen vollständig beschrieben wird, ist eine Unterscheidung sinnvoll, da jeweils andere Aspekte im Zentrum der Analyse bzw. der Interpretation stehen. Zum Beispiel kann die Hazard-Rate wie beschrieben

als „Risiko" des Eintrittes eines Ereignisse interpretiert werden, während die Survivorfunktion Auskunft über die Wahrscheinlichkeit gibt, dass das Individuum diese Zeitperiode „überlebt", d.h., dass auch in der betrachteten Zeitspanne kein Ereignis stattfindet. Im allgemeinen gilt aber, dass die Hazard-Rate der mathematisch einfacher zu handhabende (Blossfeld, Hamerle und Mayer 1989, S. 220) und inhaltlich der intuitiv verständlichere Ansatz ist (Allison 1995, S. 17).

4 Die Behandlung zensierter Beobachtungen

Im vorgegangen Abschnitt ist dargestellt worden, dass es mit der Hazard-Analyse möglich ist, den Eintritt eines interessierenden Ereignisses als stochastischen Prozess unter Berücksichtigung der vergangenen Zeit bis zu seinem Eintritt zu modellieren und mit Hilfe verschiedener, ineinander überführbarer Funktionen zu beschreiben. Die Frage ist nun, wie dabei Beobachtungen berücksichtigt werden, bei denen das interessierende Ereignis erst nach Ablauf der Beobachtungsperiode aufgetreten ist. Abbildung 4-1 verdeutlicht diese Problematik.

Abbildung 4-1 Problematik zensierter Beobachtungen



Quelle: Litfin 2000, S. 70.

Die Bestimmung der Verweildauern von t_1 sowie t_3 und t_4 ist vollkommen unproblematisch, da das Ereignis, im Schaubild z.B. die Übernahme einer Innovation, innerhalb der Beobachtungsperiode eingetreten ist. Diese Länge der Verweildauern ergibt sich als zeitliche Differenz zwischen Markteinführung und Adoption durch das i-te Individuum. Die Verweildauer von t_5 kann in der Regel durch eine retrospektive

Erfragung des Übernahmezeitpunktes erhoben werden.[7] Einzig die Verweildauer des zweiten Untersuchungsobjektes t_2 kann nicht bestimmt werden, da die Übernahme nach Ende der Beobachtungsperiode stattfindet. Im Gegensatz zur linearen oder auch logistischen Regression bietet die Hazard-Analyse eine Möglichkeit, diese Beobachtung in der Untersuchung zu berücksichtigen. Dabei bedient man sich der Survivorfunktion, die ja gerade die Wahrscheinlichkeit angibt, dass das interessierende Ereignis zu einem bestimmten Zeitpunkt noch nicht eingetreten ist. Genau dies ist bei zensierten Beobachtung t_2 zum Ende des Beobachtungszeitraumes T der Fall, so dass diese Beobachtung mit Hilfe der zugehörigen Survivorfunktion zum Zeitpunkt T beschrieben werden kann.[8] Um nun vollständige und zensierte Beobachtungen zusammenfassen zu können, wird der Zensierungsindikator C_i eingeführt. Dieser nimmt den Wert 1 annimmt, wenn bei der i-ten Person das interessierende Ereignis innerhalb der Beobachtungsperiode eintritt. Anderenfalls ist die Beobachtung zensiert und C_i nimmt den Wert 0 an (Vgl. zu den verschiedenen Zensierungsmodellen Allison 1984, S. 28 f.; Blossfeld, Hamerle und Mayer 1986, S. 72; Kalbfleisch und Prentice 1980, S. 119 ff.; Maller und Zhou 1996, S. 29 ff.) Insgesamt erfordert dieses einfache Zensierungsmodell folgende Daten je Individuum:

- Die Länge der Beobachtungsperiode: $(0, T]$,
- Den Zustand der Beobachtung: Vollständig oder unvollständig, d.h. die Beobachtung ist rechtszensiert. Formal wird dies mit Hilfe des Zensierungsindikator C_i ausgedrückt: $C_i = \begin{cases} 1 & \text{wenn } t_i \leq T \\ 0 & \text{wenn } t_i > T \end{cases}$.

Auf der Basis dieser kann dann eine zu maximierende Likelihood-Funktion konstruiert werden. Für die vollständige Beobachtungen ergibt sich die Nutzungswahrscheinlichkeit wie folgt:[9]

---

[7] Für die Fälle, in denen dies nicht möglich sein sollte, sei auf Verfahren zur Behandlung dann linkszensierter Daten bei Klein und Moeschberger 1997, S. 62 ff. verwiesen bei, deren Darstellung im Zusammenhang der vorliegenden Untersuchung zu weit führen würde.

[8] Dabei ist sicherzustellen, dass der Zensierungsmechanismus unabhängig von den Kovariablen sowie anderen die Verteilung der Episode beeinflussenden Faktoren ist, es sich also um einen nicht informativen Zensierungsprozess handelt (Blossfeld, Hamerle und Mayer 1986, S. 74). Im folgenden wird angenommen, dass diese Unabhängigkeit gegeben ist.

[9] Aufgrund der kontinuierlichen Darstellung der Zeitdauer ergibt sich für die Wahrscheinlichkeit des Nutzungsbeginnes in einem Zeitpunkt stets 0. Folgerichtig wäre eine Grenzwertbetrachtung vorzunehmen, um die Wahrscheinlichkeit in einem infinitesimal kleinen Zeitintervall bestimmen zu können. Hierauf wird aus Vereinfachungsgründen verzichtet (Vgl. dazu. Sinha und Chandrashekaran 1992, S. 118).

(4-1) $$P(C_i = 1, t = t_i) = f(t_i)$$

Für die zensierten Beobachtungen ist nur bekannt, dass die Nutzung bis zum Ende des Beobachtungszeitraumes T nicht erfolgt ist. Dies lässt sich mit der Survivorfunktion beschreiben:

(4-2) $$P(C_i = 0) = P(t_i > T) = 1 - F(T) = S(T).$$

Damit ergibt sich die folgende Likelihood- Funktion:[10]

(4-3) $$L = \prod_{i=1}^{I} P(t_i, C_i) = \prod_{i=1}^{I} [f_i(t_i)^{C_i}] \cdot [(S_i(t_i)^{1-C_i}]$$

bzw. Log-Likelihood-Funktion

(4-4) $$\ln L = \sum_{i=1}^{I} (C_i \cdot \ln[f_i(t_i)] + (1 - C_i) \cdot \ln[S_i(t_i)]).$$

Unter Berücksichtigung der in Tabelle 3-1 dargestellten Zusammenhänge ergeben sich noch die folgenden Formulierungsmöglichkeiten, auf die im weiteren Verlauf der Darstellung zurückgegriffen wird:

(4-5) $$\ln L = \sum_{i=1}^{I} (C_i \cdot \ln[h_i(t_i)] + \ln[S_i(t_i)]) = \sum_{i=1}^{I} C_i \cdot \ln[h_i(t_i)] - \int_{o}^{t} h_i(u)du.$$

Damit ist eine Log-Likelihood-Funktion unter Verwendung der verschiedenen Funktionen, die im Rahmen des statistischen Grundkonzeptes in Abschnitt 3 zur Beschreibung der Episodendauer t_i verwendet worden sind, konstruiert worden. Es können nun sowohl vollständige als auch zensierte Beobachtung berücksichtigt werden, so dass die bei der linearen und logistischen Regression ungelöste Problematik zensierter Beobachtungen bewältigt werden kann. Auf Basis von Gleichung (4-4) bzw. (4-5) werden nun unabhängige Variablen, die in der Hazard-Analyse als Kovariablen bezeichnet werden, berücksichtigt und die Funktion $S(\cdot)$, $f(\cdot)$ bzw. $h(\cdot)$ näher spezifiziert, so dass sich ein vollständiges Regressionsmodell ergibt. Ist dies erfolgt, kann eine Maximierung der resultierenden Log-Likelihood-Funktionen in Abhängigkeit der zu schätzenden Parameter mittels eines iterativen Verfahrens, z.B. des Newton-Verfahrens, vorgenommen werden.[11]

---

[10] Eine ausführlichere Herleitung der Likelihood-Funktion findet sich unter Blossfeld, Hamerle und Mayer 1986, S. 72 ff. sowie Klein und Moeschberger 1997, S.65 ff.

[11] Eine ausführliche Darstellung der Maximum-Likelihood-Methode im Kontext der Ereignisanalyse findet sich unter anderem Allison 1995, S. 81 ff.; Blossfeld und Rohwer 1995, S. 82 ff., Blossfeld, Hamerle und Mayer 1986, S. 74 ff.; Diekmann und Mitter 1984, S. 52 ff.; Kalbfleisch und Prentice

5 Regressionsmodelle

Neben der Modellierung der Zeitdauerabhängigkeit des untersuchten Prozesse liegt ein Hauptaugenmerk der Hazard-Analyse auf der Ermittlung des Einflusses quantitativer und qualitativer erklärender Variablen, sogenannter Kovariablen, auf die Dauer der Episode. (Blossfeld, Hamerle und Mayer 1986, S. 48 f.). Dazu dienen regressionsanalytische Verfahren, die in diesem Abschnitt vorgestellt werden. Analog zur klassischen multiplen Regression geht man davon aus, dass der Einfluss der Kovariablen linear in den Parametern erfolgt. Allerdings beeinflussen diese nicht die Verweil- beziehungsweise Episodendauern T direkt, sondern in der Regel eine Funktion von T, etwa ln T (Blossfeld, Hamerle und Mayer 1986, S.50).

5.1.1 Parametrische Regressionsmodelle

Bei parametrischen Regressionsmodellen wird sowohl der Einfluss der Kovariablen als auch der Effekt der Episodendauer auf die Hazard-Rate parametrisiert, d.h. es wird eine spezielle Verteilungsannahme für die Episodendauer getroffen. Dies bietet sich an, wenn a priori Informationen über die Art der Zeitdauerabhängigkeit des beobachteten Prozesses bestehen. Auf Basis dieser Kenntnis wird dann die Verteilung für t_i zugrunde gelegt, die den unterstellten Zusammenhang zwischen Eintritt des Ereignisses und Episodenlänge adäquat beschreibt. Die Wahl einer bestimmten Verteilung determiniert dann den Verlauf der Hazard-Rate und damit natürlich auch den der zugehörigen Dichte- bzw. Survivorfunktion (Blossfeld, Hamerle und Mayer 1989, S. 220).

In der Literatur finden eine Vielzahl verschiedenster Verteilungen Anwendung bei der Modellierung dynamischer Prozesse mit Hilfe von Hazard-Modellen. Eine Diskussion aller prinzipiell möglichen Verteilungen würde den Rahmen dieser Arbeit sprengen und ist nicht zielführend. Die folgenden Ausführungen beschränken sich daher auf einige ausgewählte Verteilungen, anhand derer die verschiedenen Eigenschaften und Besonderheiten parametrischer Modelle exemplarisch aufgezeigt werden können.

1980, S. 54 ff. Grundlagen der Maximum-Likelihood Schätzung werden ausführlich in Mood, Graybill und Boes 1974, S. 276 ff. beschrieben.

5.1.1.1 Das Exponential-Modell

Im Rahmen der Hazard-Analyse wird häufig unterstellt, dass die Episodendauer t_i einer Exponentialverteilung folgt. Dieses Modell wird vielfach auch als Basis- bzw. Referenzmodell zum Vergleich alternativer Spezifikationen der Verweildauerabhängigkeit herangezogen (Sinha und Chandrashekaran 1992, S. 120). Der Grund hierfür ist, dass es sich bei den in der empirischen Anwendung beliebten auf einer Weibull-, Exponential- oder einer Log-Normalverteilung beruhenden Modellen um sogenannte genestete Modelle handelt. Ein Modell wird als genestet bezeichnet, wenn es sich aufgrund von Parameterrestriktionen als ein spezieller Fall eines zweiten, allgemeineren Modells ergibt. In diesem Fall stellt die sogenannte Verallgemeinerte-Gamma-Verteilung (Kalbfleisch und Prentice 1980 , S. 27) ein solches allgemeineres Modell dar:

$$(5\text{-}1) \qquad f(t) = \frac{\lambda \cdot \eta \cdot (\lambda t)^{\eta\alpha - 1} \cdot \exp\left[-(\lambda \cdot t)^{\eta}\right]}{\Gamma(\cdot)} \qquad t > 0,\ \Gamma(\cdot)\text{: Gammafunktion.}$$

Die aufgeführten Parameterrestriktion führen dann zu den angegebenen Verteilungen bzw. Modellen (Allison 1995, S. 89):

- $\eta = 1$ Weibull-Verteilung,
- $\eta = 1,\ \alpha = 1$ Exponential-Verteilung,
- $\eta = 0$ Log-Normalverteilung.

Allerdings gilt nicht, dass die erwähnten Modelle auch untereinander genestet sind. Nur das Exponential-Modell ergibt sich als Spezialfall eines auf der Weibull- bzw. der Standard-Gamma-Verteilung beruhenden allgemeineren Modells. Dennoch ist diese Eigenschaft bei der Konstruktion von Gütekritierien, wie sie in Abschnitt 6 beschrieben werden, von Bedeutung.

Zur weiteren Charakterisierung des Exponential-Modells dienen wie bereits bei der Darstellung des statistischen Grundkonzeptes die Dichte- und Survivorfunktion bzw. die Hazard-Rate:

$$(5\text{-}2) \qquad f(t_i|X_i) = \lambda_i \cdot e^{-\lambda_i \cdot t_i},$$

$$(5\text{-}3) \qquad S(t_i|X_i) = \exp(-\lambda_i \cdot t_i),$$

$$(5\text{-}4) \qquad h(t_i|X_i) = \lambda_i,$$

mit

λ_i: Parameter mit $\lambda_i = e^{-\beta_0 - x_i \cdot \beta_i}$,

X_i: Vektor der Erklärungsvariablen des i-ten Individuums ($i \in I$),

β: Vektor der Koeffizienten.

Am Beispiel des Exponential-Modells soll die Vorgehensweise bei der Konstruktion eines Regressionsmodells kurz erläutert werden. Exponential-Modelle werden vollständig durch den Parameter λ determiniert. Da die durchschnittliche Verweildauer eines Individuums $1/\lambda$ ist, bietet es sich an, den Kovariableneinfluss auf die Verweildauer z.B. in der Form $1/\lambda = g(x;\beta)$ zu modellieren, wobei β den unbekannten Parametervektor darstellt. Die Funktion g ist dann so zu spezifizieren, dass sie nur positive Werte annehmen kann, da die Hazard-Rate wie o.a. nur positive Werte annehmen kann:[12]

(5-5) $$g(X'\beta) = \exp(X'\beta)$$

bzw.

(5-6) $$g(X'\beta) = \lambda_0 \exp(X'\beta) \qquad (\lambda_0 > 0)$$

$$g(X'\beta) = \exp(\beta_0 + x'\beta) \qquad \text{mit } \beta_0 = \ln\lambda_0 .$$

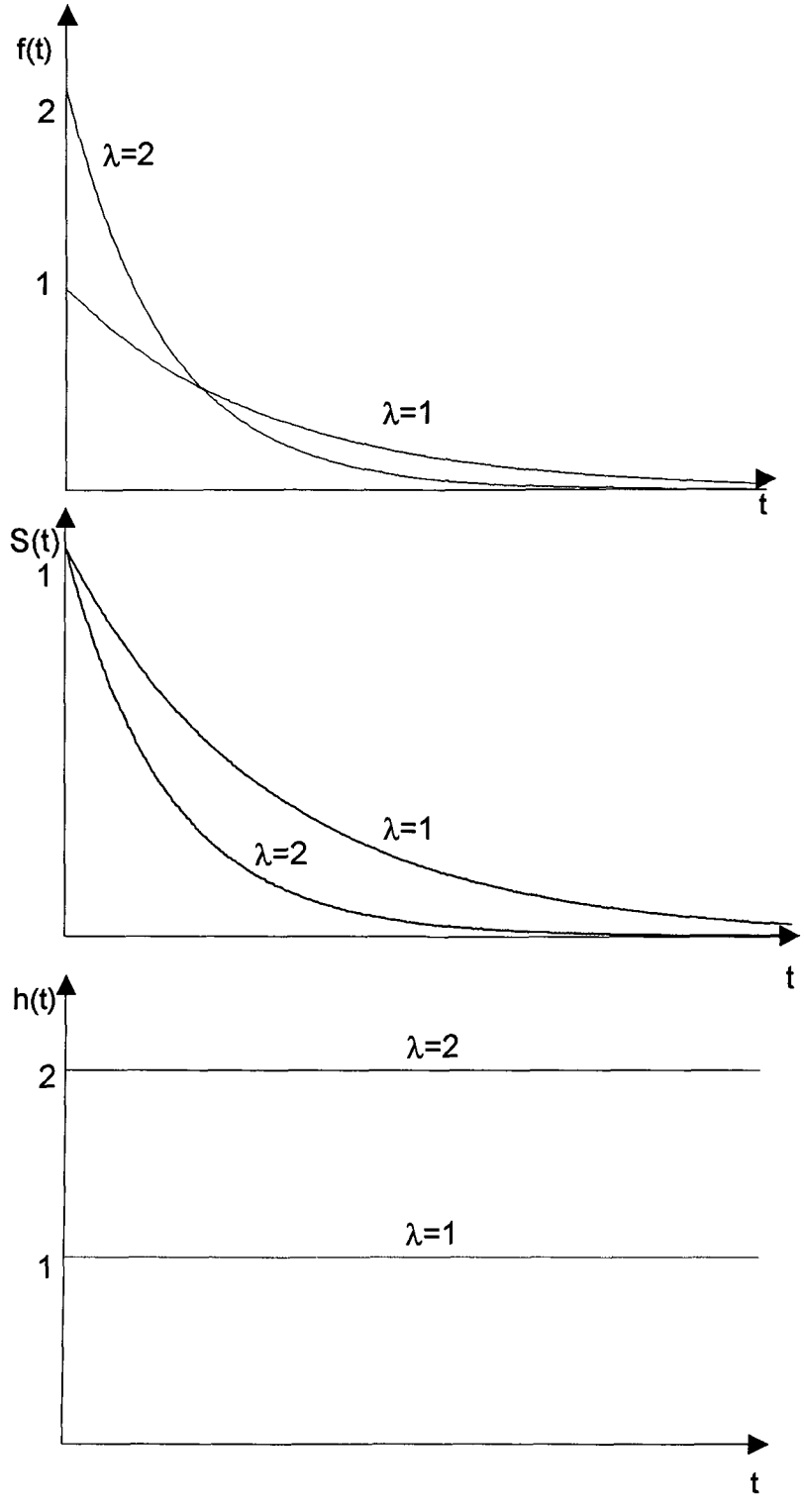
β_0 ist dabei das konstante Glied des Regressionsansatzes. Die Episodenlänge T besitzt nun eine Exponentialverteilung mit dem in Abhängigkeit der Kovariablen spezifizierten Parameter $\lambda = \exp(-\beta_0 - x'\beta)$ (Blossfeld, Hamerle und Mayer 1986, S. 51 f.).

Wie aus der Parametrisierung der Hazard-Rate deutlich wird, ist diese beim Exponential-Modell über die Zeit konstant. D.h. die Eintrittswahrscheinlichkeit des Ereignisses ist unabhängig von der betrachteten Verweildauer in dem Anfangszustand. Es handelt sich also um eine „gedächtnislose" Verteilungsannahme (Diekmann und Mitter 1984, S. 146; Klein und Moeschberger 1997 S. 38). Diese Annahme erscheint ausgesprochen restriktiv. Litfin (Litfin 2000, S. 73) argumentiert im Fall der Adoption eines innovativen Telekommunikationsdienstes, dass die Annahme einer zeitunabhängiger Eintrittswahrscheinlichkeit unplausibel ist, da bestimmte adoptionsfördernde Effekte wie die Verbreitung von Informationen über die Neuerung im Sozialen System und der Aufbau von derivativem Nutzen bei dem von ihm untersuchten Netzeffekt-Gut sich im Zeitablauf verstärken und damit zu einer steigenden Adoptionswahrscheinlichkeit führen.

---

[12] Würde die identische Funktion g(z) = z gewählt, träten unkontrollierbare Restriktionen für den Parameter β auf, da λ der Restriktion $\lambda > 0$ unterliegt, vgl. z.B. Mantel und Myers 1971, S. 489.

Abbildung 5-1: Funktionsverläufe des Exponential-Modells



Quelle: In Anlehnung an: Blossfeld, Hamerle und Mayer 1986, S. 35.

Eine von der Verweildauer unabhängige Übernahmewahrscheinlichkeit bedeutet jedoch nicht, dass die Nutzungswahrscheinlichkeit ebenfalls konstant ist über alle Individuen. Die individuellen Charakteristika der Untersuchungsobjekte werden, wie

bereits in Gleichung (5-6) beschrieben, über die log-lineare Parametrisierung von λ_i berücksichtigt. Folglich resultieren individuenspezifische Werte für diesen Parameter, so dass unterschiedliche Hazard-Raten resultieren. Abbildung 5-1 veranschaulicht dies exemplarisch, wobei für λ_i die Werte 1 bzw. 2 gewählt werden (Blossfeld, Hamerle und Mayer 1986, S. 52 f.).

Eine weitere Eigenschaft des Exponential-Modells besteht darin, dass die Hazard-Raten zweier Individuen sich um einen konstanten Faktor unterscheiden, also proportional zueinander sind. Diese Eigenschaft wird bei der Quotientenbildung der jeweils betrachteten Hazard-Raten deutlich:

$$\text{(5-7)} \qquad \frac{h(t|X_i)}{h(t|X_j)} = \frac{\exp(\beta_{0j} + X_j'\beta_i)}{\exp(\beta_{0i} + X_i'\beta_j)} = c \text{ für } i \neq j .$$

5.1.1.2 Das Weibull-Modell

Eine erste Möglichkeit, die restriktive Annahme zeitkonstanter Hazard-Raten zu relaxieren, stellt das Weibull-Modell dar, das formal wie folgt beschrieben werden kann:

$$\text{(5-8)} \qquad f(t_i|X_i) = \alpha \cdot \lambda_i \cdot (\lambda_i \cdot t_i)^{\alpha-1} \cdot \exp(-(\lambda_i \cdot t_i)^{\alpha}),$$

$$\text{(5-9)} \qquad S(t_i|X_i) = \exp(-(\lambda_i \cdot t_i)^{\alpha}),$$

$$\text{(5-10)} \qquad h(t_i|X_i) = \lambda_i \cdot \alpha(\lambda_i \cdot t_i)^{\alpha-1},$$

wobei

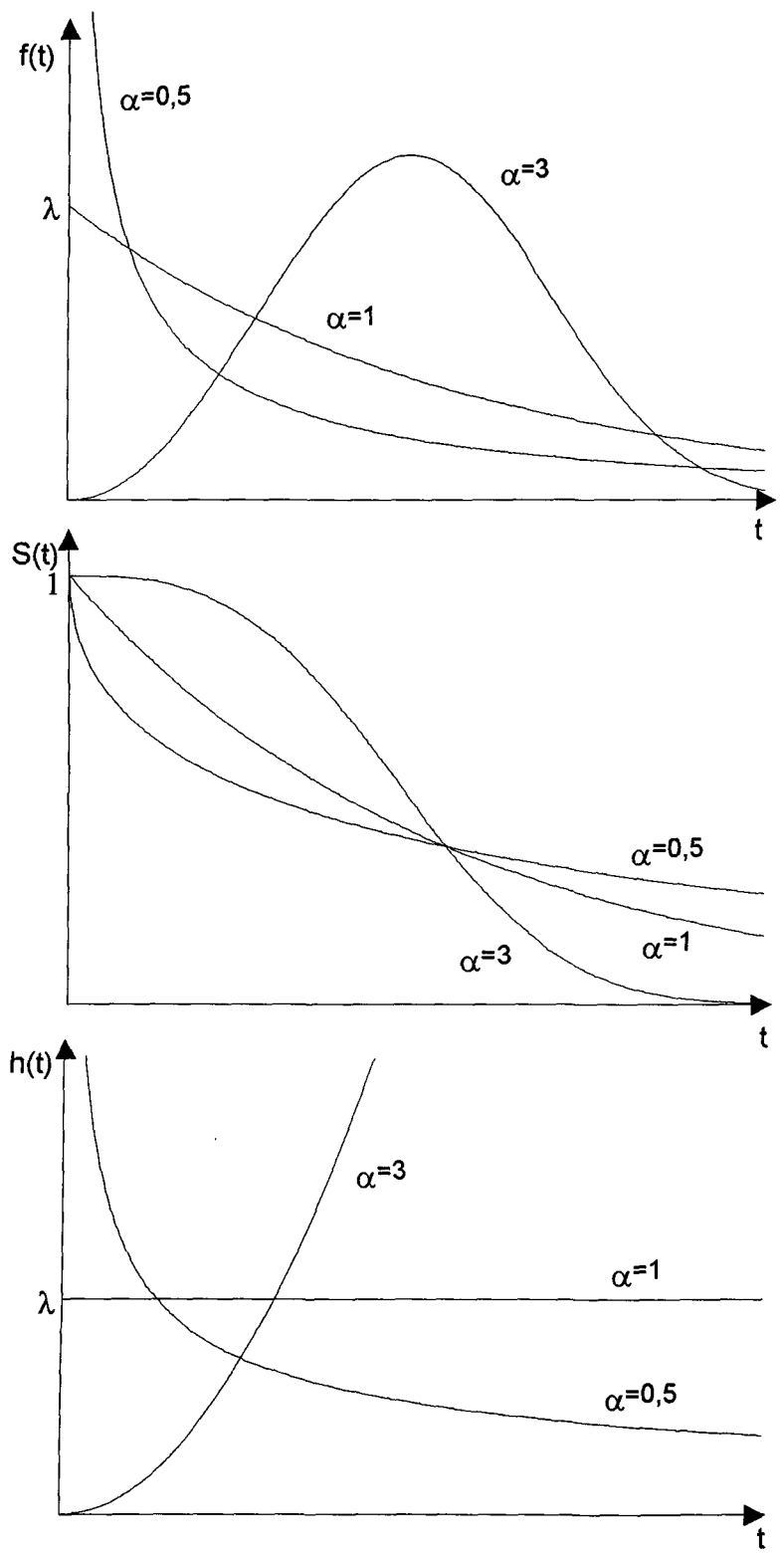
λ_i: Parameter mit $\lambda_i = e^{-\beta_0 - x_i\beta_i}$ und

α: (Shape-)Parameter.

Wie beim Exponential-Modell stehen auch hier die Kovariablen in einem log-linearen Zusammenhang zum Parameter λ_i. Die gegenüber dem Exponential-Modell erhöhte Flexibilität dieses Modells wird durch den Shape-Parameter α erreicht, der für alle Personen den gleichen Wert annimmt. Bei isolierter Betrachtung ergibt sich für Werte $\alpha < 1$ eine im Zeitablauf abnehmende Hazard-Rate. Eine mit der Zeit zunehmende Hazard-Rate ergibt sich für α-Werte größer 1. Wie bereits eingangs beschrieben, reduziert sich das Weibull- zum Exponential-Modell, wenn α den Wert 1 annimmt.

Um die Übersichtlichkeit in den Funktionsverläufen der Weibull-Verteilung in Abbildung 5-2 zu gewährleisten, sind nur die Werte des Shape-Parameters α variiert worden.

Abbildung 5-2: Funktionsverläufe des Weibull-Modells



Quelle: Blossfeld, Hamerle und Mayer 1986, S. 37.

Durch Quotientenbildung kann zudem gezeigt werden, dass auch das Weibull-Modell zur Klasse der Proportional-Hazard-Modelle gehört, da der im Vergleich zum Exponential-Modell zusätzlich auftretende Parameter α und damit auch der Quotient der Hazard-Raten zweier beliebiger (Populations-)Subgruppen konstant ist.

5.1.1.3 Das Log-Normal-Modell

Das Weibull-Modell ist zwar nicht auf zeitkonstante Hazard-Raten wie das Exponential-Modell beschränkt, jedoch resultieren für Werte $\alpha \neq 1$ entweder monton steigende oder fallende Hazard-Raten. Damit wird also nicht unterstellt, dass die Zeitdauer gar keinen Einfluss hat, aber es wird implizit angenommen, dass dieser in Bezug auf die Richtung des Einflusses konstant ist. Eine Alternative stellt das Log-Normal-Modell dar. Dieses erhält man, wenn für die logarithmierten Verweildauern eine Normalverteilung angenommen wird. Im Gegensatz zum Weibull-Modell können hier nicht nur sinkende oder steigende Hazard-Raten abgebildet werden. So steigt die Hazard-Rate für Werte von $\alpha < 1$ zunächst von Null bis zu einem maximalen Wert an und fällt dann wieder. Bei unendlichen Zeitdauern nähert sie sich asymptotisch wieder der Null. Der Bezug zur Normal-Verteilung resultiert aus dem Umstand, dass auch die Log-Normal-Verteilung durch Mittelwert und Varianz vollständig beschrieben wird. Kritisch ist anzumerken, dass eine im Zeitablauf sinkende und im Extremfall auf Null reduzierte Hazard-Rate teilweise wie zum Beispiel bei der Modellierung von Lebensdauern nicht plausibel ist. Werden keine (extrem) langen Verweildauern betrachtet, hat sich dieses Hazard-Modell in einer Vielzahl empirischer Untersuchung als sehr geeignet erwiesen (Klein und Moeschberger 1997, S. 39 f.). Die Funktionen für diese Verteilung lauten:

$$f(t_i|X_i) = \frac{\alpha \cdot \exp\left(\frac{-\alpha^2 (\ln(\lambda_i \cdot t_i))^2}{2}\right)}{(2\pi)^{1/2} \cdot t_i}, \tag{5-11}$$

$$S(t_i|X_i) = 1 - \Phi(\alpha \cdot \ln(\lambda_i \cdot t_i)), \tag{5-12}$$

$$h(t_i|X_i) = \frac{(2 \cdot \pi)^{-1/2} \cdot \alpha \cdot t^{-1} \cdot \exp\left(\frac{-\alpha^2 \cdot (\ln(\lambda_i \cdot t_i))^2}{2}\right)}{1 - \Phi(\ln(\lambda_i \cdot t_i) \cdot \alpha)}, \tag{5-13}$$

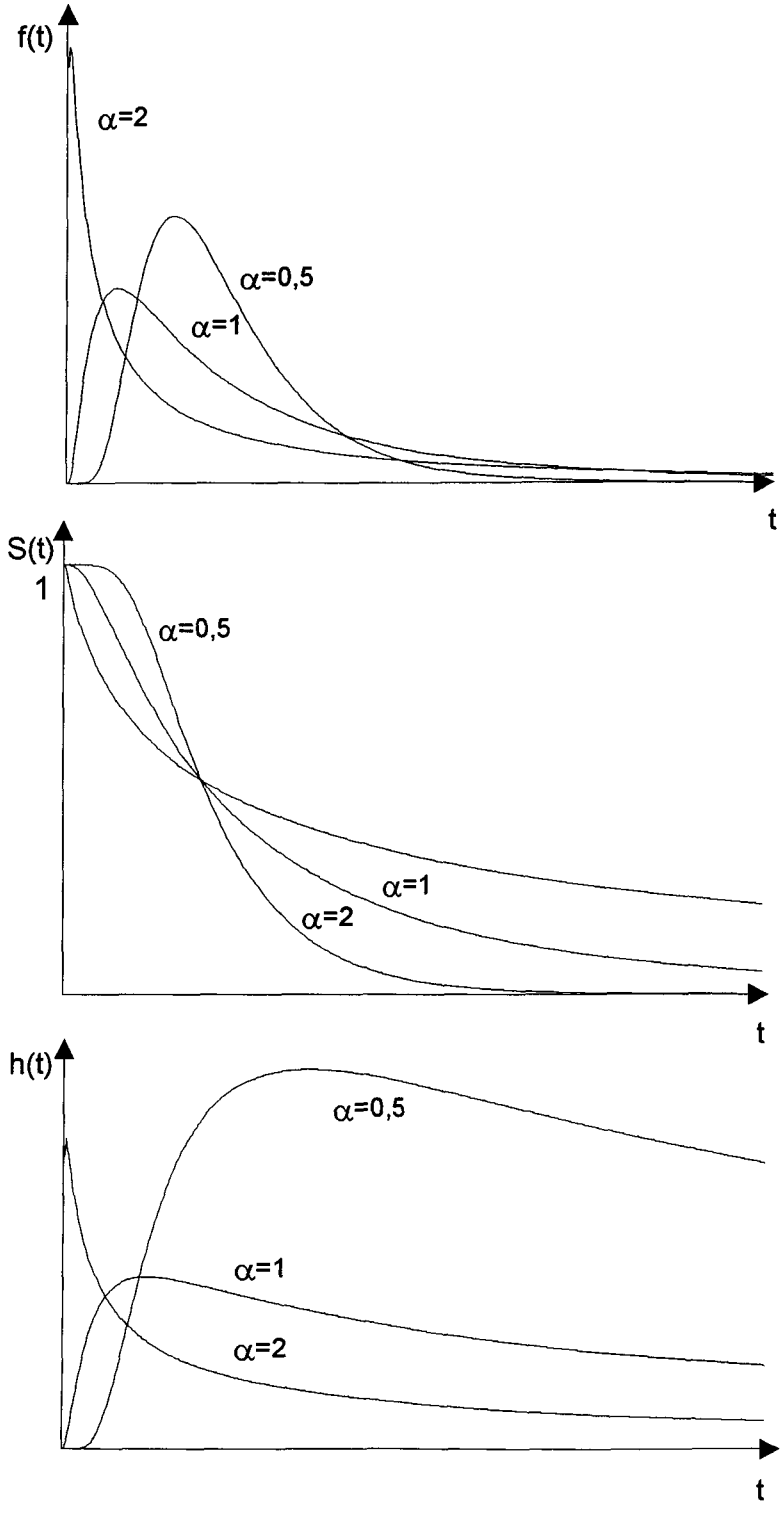
wobei:

λ_i: Parameter mit $\lambda_i = e^{-\beta_0 - x_i' \cdot \beta_i}$,

α: Parameter,

Φ: Verteilungsfunktion einer normalverteilten Zufallsvariable.

Abbildung 5-3: Funktionsverläufe des Log-Normal-Modells



Quelle: Blossfeld, Hamerle und Mayer 1986, S.

Neben dem Umstand, dass das Exponential-Modell anders als das Weibull-Modell nicht mit dem Log-Normal-Modell genestet ist, zeigt sich auch, dass das Log-Normal-Modell nicht in die Klasse der Proportional-Hazard-Modelle fällt (Blossfeld, Hamerle und Mayer 1986, S. 54 f.). Der variable Term t lässt nicht mehr kürzen, da

bei der Betrachtung zweier Individuen in der Regel $t_i \neq t_j$ gilt, so dass der Quotient der Hazard-Raten zweier Subpopulationen nicht mehr konstant ist und somit die Eigenschaft des Proportional-Hazard nicht erfüllt werden kann. Mögliche Verläufe der Log-Normal-Verteilung werden in Abbildung 5-3 veranschaulicht.

5.1.1.4 Das Log-Logistische-Modell

Eine Alternative zum insbesondere in Bezug auf die resultierende Hazardfunktion komplexen Log-Normal-Modell stellt das Log-Logistische-Modell dar. Hier folgen die logarithmierten Zeitdauern einer logistischen Verteilung, die der Normalverteilung vor allem im mittleren Wertebereich sehr ähnlich ist mit dem Unterschied, dass die resultierende Funktion für die Hazard-Rate mathematischer leichter zu handhaben ist (Klein und Moeschberger 1997, S. 41). Die zugehörigen Funktionen lauten:

(5-14) $$f(t_i|X_i) = \frac{\lambda_i \cdot \alpha(\lambda_i t_i)^{\alpha-1}}{(1+(\lambda_i t_i)^{\alpha})^2},$$

(5-15) $$S(t_i|X_i) = \frac{1}{1+(\lambda_i \cdot t_i)^{\alpha}},$$

(5-16) $$h(t_i|X_i) = \frac{\lambda_i \cdot \alpha(\lambda_i \cdot t_i)^{\alpha-1}}{1+(\lambda_i \cdot t_i)^{\alpha}},$$

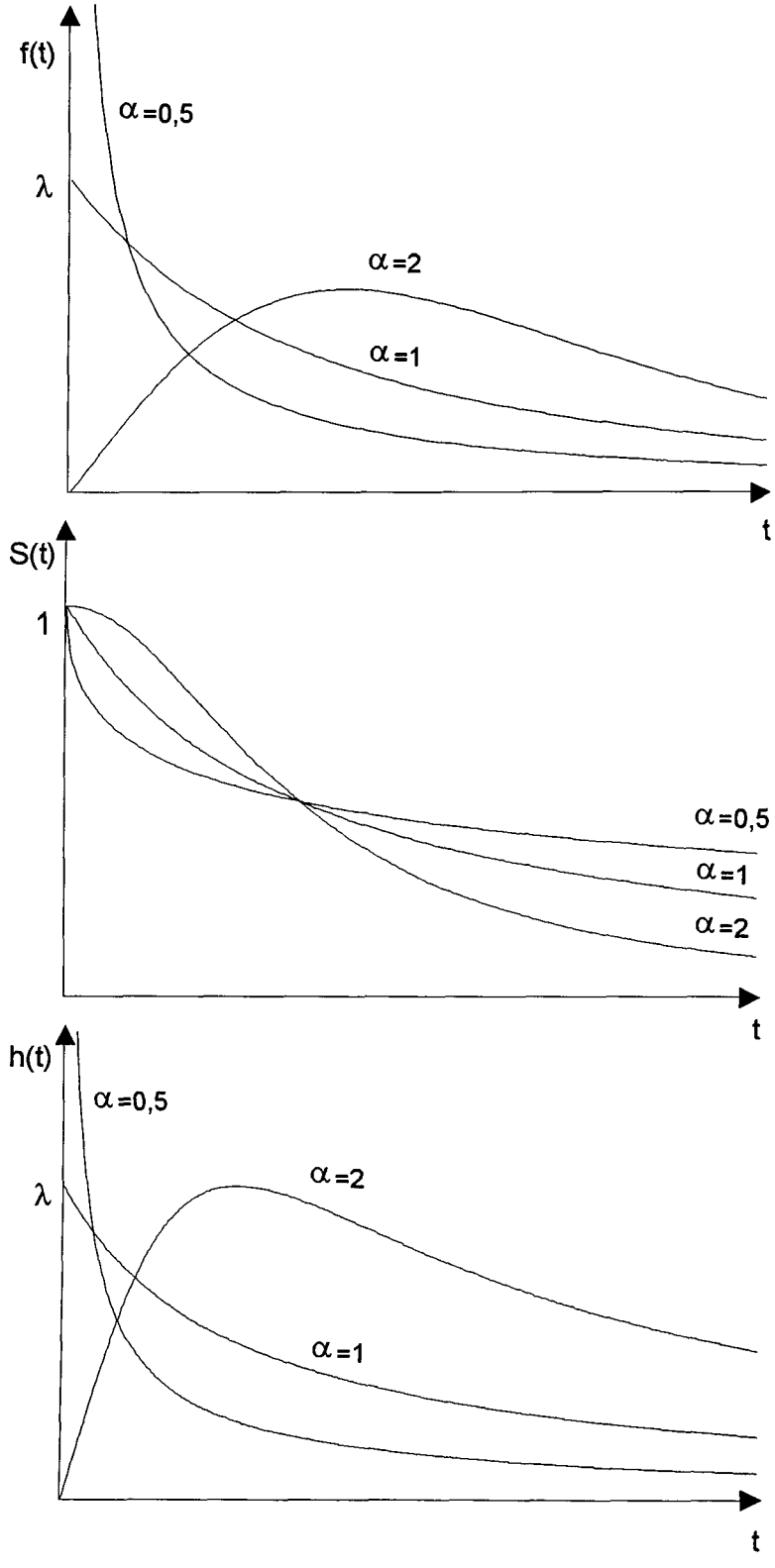
wobei:

λ_i: Parameter mit $\lambda_i = e^{-\beta_0 - x_i\beta_i}$,

α : Parameter.

Es fällt auf, dass der Term im Zähler der Hazard-Rate des Log-Logistischen-Modells exakt der Hazard-Rate des Weibull-Modells entspricht. Der flexiblere Verlauf der Hazard-Rate resultiert aus dem zusätzlichen Term im Nenner der Hazard-Rate. Dabei gilt, dass abnehmende Risiken für $\alpha < 1$ und zunächst steigende und erst im späteren Verlauf abnehmende Risiken für $\alpha > 1$ abgebildet werden können. Ebenso wie beim Log-Normal-Modell gehört das Log-Logistische-Modell nicht zur Klasse der Proportional-Hazard, da auch hier der Zeitparameter nicht aus dem Quotienten der Hazard-Raten der beiden Untergruppen i und j ($i \neq j$) herausgekürzt werden kann.

Abbildung 5-4: Funktionsverläufe des Log-Logistischen-Modells



Quelle: Blossfeld, Hamerle und Mayer 1986, S. 40.

5.1.2 Zusammenfassung und Beurteilung parametrischer Hazard-Modelle

Damit endet die Darstellung ausgewählter Verteilungsannahmen für parametrische Hazard-Modelle. Es konnte gezeigt werden, dass in Abhängigkeit von der gewählten Verteilungsannahme eine flexible Modellierung der jeweils betrachteten Zeitdauerabhängigkeit des zu untersuchenden Prozesses möglich ist. Darüber hinaus konnten auch Eigenschaften wie die des Proportional-Hazard oder auch die genesteter Modelle verdeutlicht werden, die für den weiteren Umgang mit Hazard-Modellen von Nutzen sein werden. Eine noch deutlich über die hier vorgenommene Beschreibung einzelner Verteilungsannahmen hinausgehende Darstellung findet sich bei Klein und Moeschberger (Klein und Moeschberger 1997, S. 36 ff.).

Bereits das hier vorgestellte Grundmodell der Hazard-Analyse stellt ein im Vergleich zu statischen Ansätzen wie der logistischen Regression oder einem dynamischen Ansatz auf Grundlage der linearen Regression deutlich besseres Instrumentarium zur Untersuchung von dynamischen Prozessen dar. Insbesondere die Möglichkeit zur adäquaten Berücksichtigung zensierter Beobachtungen und die Modellierung der Zeitdauerabhängigkeit des untersuchten Prozesses durch Wahl einer geeigneten Verteilungsannahme sind hier zu nennen. Dennoch ergeben sich aus der bisherigen Darstellung auch Einschränkungen bzw. mögliche Erweiterungen:

1. Bei dem bisher vorgestelltem Grundmodell der Hazard-Analyse ergibt sich zwangsläufig (Vgl. zur Anschauung die Graphen der Survivorfunktionen in den Abbildung 5-1 bis Abbildung 5-4), dass letztendlich bei allen Individuen das interessierende Ereignis eintritt, wenn der betrachtete Zeitraum nur groß genug ist. Dies ist eine unberechtigt restriktive Annahme. Es ist durchaus vorstellbar, dass erst weit nach Ablauf der Beobachtungsperiode ein Ereignis z.B. in Form einer Adoptions- oder Nutzungsentscheidung, die dann positiv oder negativ ausfallen kann, stattfindet. Eine Berücksichtigung der Heterogenität der Individuen in Bezug auf dieses (u.U. nicht beobachtbare) Ereignis ist mit Hilfe des sogenannten Split-Hazard- Ansatzes möglich.
2. Neben der Heterogenität in Bezug darauf, ob das interessierende Ereignis tatsächlich eintritt, kann unmöglich davon ausgegangen werden, dass die erhobenen unabhängigen Variablen ausreichend sind, um die Heterogenität zwischen den Individuen vollständig zu erfassen. Diese Problematik erstreckt sich nicht exklusiv auf Hazard-Modelle, sie bedarf aber wegen ihrer besonderen Auswirkungen bezüglich dieser Modelle einer intensiveren Diskussion.

3. Die Wahl einer geeigneten Verteilungsannahme für die Zeit- bzw. Verweildauern kann sich unter Umständen als sehr schwierig darstellen, auch wenn geeignete Tests zur Überprüfung einer gewählten Verteilung vorliegen. Hier stellt der Proportional-Hazard Ansatz von Cox eine Alternative dar, der ohne die (ex-ante) Annahme einer Verteilung die Durchführung einer Hazard-Analyse ermöglicht.
4. In der bisherigen Darstellung wurde in den Regressionsmodellen von der Konstanz der berücksichtigten Kovariablen ausgegangen. Allerdings ist es durchaus denkbar, dass sich die Werte von Kovariablen im Beobachtungszeitraum ändern. Dies ist insbesondere bei typischen Marketing-Mix Variablen wie Preis oder auch Werbung vorstellbar. Die Berücksichtigung zeitveränderlicher Kovariable ist ein komplexes und schwieriges Unterfangen (Allison 1995, S. 138), das im Rahmen dieser Arbeit einführend diskutiert werden soll.

Im folgenden werden die dargestellten Modellerweiterungen bzw. Modellalternativen in der Reihenfolge ihrer Nennung näher vorgestellt. Darüber hinaus existieren noch eine Vielzahl weiterer Modellerweiterungen und Varianten, deren Diskussion aber den Rahmen dieser Arbeit sprengen würde. Stattdessen sei auf die bereits vielfach zitierte Literatur und hier insbesondere auf Allison 1995, Blossfeld, Hamerle und Mayer 1989, Heckmann und Singer 1984a, Kalbfleisch und Prentice 1980 sowie Kiefer 1988 hingewiesen.

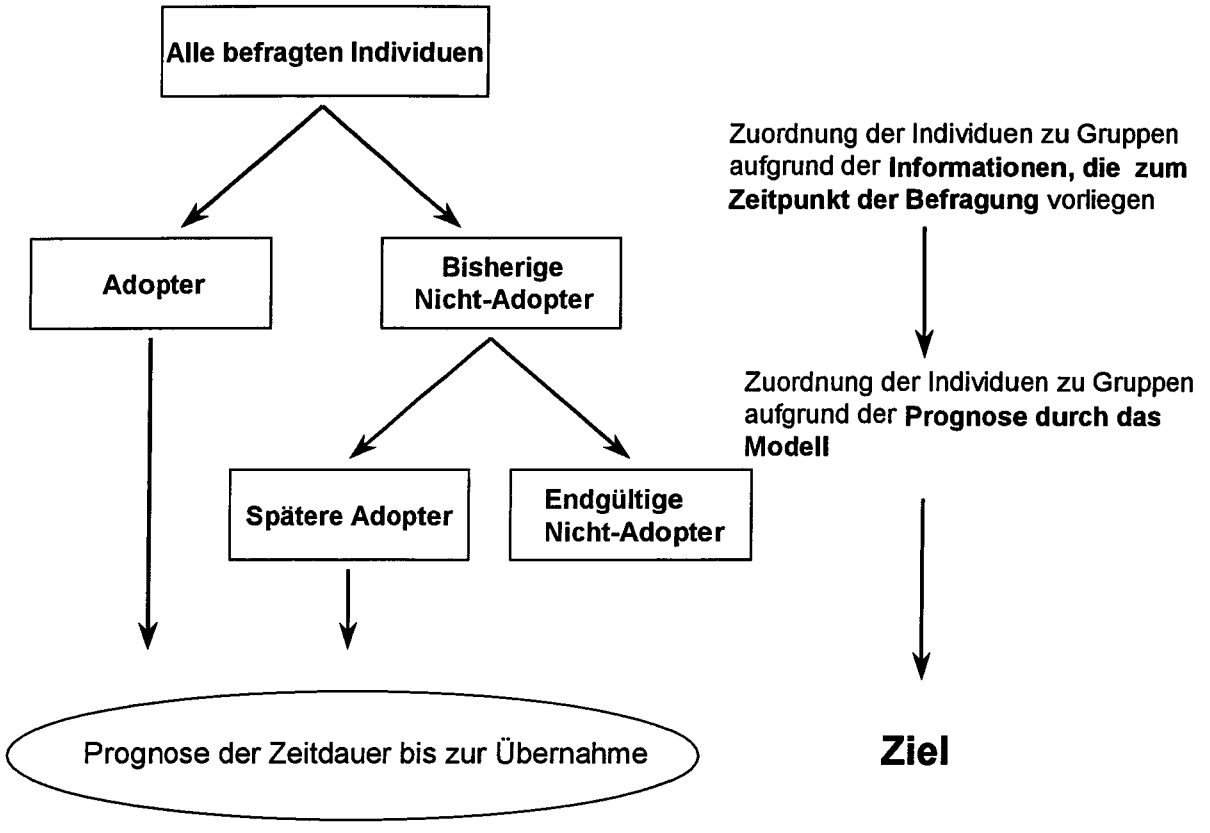
5.2 Der Split-Hazard- Ansatz

Der Split-Hazard- Ansatz wurde das erste Mal 1992 von Sinha und Chandrashekaran in der betriebswirtschaftlichen Literatur vorgestellt.[13] Bezeichnenderweise belegten sie die Überlegenheit dieses Ansatzes gegenüber den in Abschnitt 1 vorgestellten Modellen der Hazard-Analyse am Beispiel der Adoption von Geldautomaten durch Kreditinstitute. Die im folgenden verwendete Notation orientiert sich weitestgehend an dieser Arbeit, zumal auch Litfin (Litfin 2000, S, 84 ff.), der ebenfalls im Kontext der Adoptionsforschung diesen Ansatz verwendete, sich an dieser Notation orientiert.

---

[13] Grundlage der Arbeit von Sinah und Chandrashekaran war allerdings die Arbeit von Schmidt und Witte aus dem Jahr 1989 (Schmidt und Witte 1989), die mit Hilfe des Split-Hazard- Ansatzes zu deutlich verbesserten Aussagen und Prognosen über die Rückfälligkeit von Straftätern in den USA gelangten. Gemäß dem Grundmodell der Hazard-Analyse hätten über kurz oder lang alle Straftäter erneut mit dem Gesetz in Konflikt kommen müssen, was sicher eine ebenso traurige wie unrealistische Vorstellung ist.

Abbildung 5-5: Grundidee des Split-Hazard-Modells am Beispiel „Adoption"



Quelle: In Anlehnung an Litfin 2000, S. 85.

Abbildung 5-5 verdeutlicht die Grundidee des Split-Hazard-Ansatzes. Der Unterschied zum Grundmodell der Hazard-Analyse besteht in der Behandlung zensierter Beobachtungen. Anstatt aufgrund der Konstruktion des Grundmodells zu unterstellen, dass es bei sämtlichen Individuen, bei denen das interessierende Ereignis im Beobachtungszeitraum nicht registriert worden ist, früher oder später zu einem Eintritt dieses Ereignisses kommt, wird nun explizit Heterogenität - in diesem Beispiel bezüglich der Adoptionsentscheidung - zugelassen.

Für die formale Darstellung wird ein unbeobachtbarer Vektor $A = \{A_1,...,A_i,...,A_I\}$ eingeführt, in dem A_i anzeigt, ob die i-te Person zur Gruppe der Adopter bzw. der Nicht-Adopter gehören wird. Es gelte folgende Definition:

$$A_i = \begin{cases} 1 & \text{wenn das i-te Individuum ein potentieller Adopter ist,} \\ 0 & \text{wenn das i-te Individuum die Innovation nie adoptiert} \end{cases}$$

Die Wahrscheinlichkeit, dass $A_i = 1$ ist, kann dann in Abhängigkeit der individuenspezifischen Charakteristika modelliert werden:

(5-17) $$P(A_i = 1) = \delta_i = \delta(X_i)$$

bzw.

(5-18) $$P(A_i = 0) = 1 - \delta_i = 1 - \delta(X_i)$$

Der Parameter δ_i kann z.B. als Anteil der potentiellen Adopter an der gesamten Stichprobe aufgefasst werden. Nach Einführung dieses zusätzlichen Parameters lässt sich nun analog zur Vorgehensweise und bei gleicher Notation wie in Abschnitt 0 eine zu maximierende Likelihood-Funktion unter Berücksichtigung nicht-zensierter Beobachtungen ($C_i = 1$)

(5-19) $$P(C_i = 1, t = t_i) = P(A_i = 1) \cdot f(t_i \mid X_n, A_i = 1) = \delta_i \cdot f(t_i \mid X_i, A_i = 1)$$

sowie zensierter Beobachtungen ($C_i = 0$)

(5-20) $$\begin{aligned} P(C_i = 0) &= P(A_i = 0) + P(A_i = 1) \cdot P(t_i > T \mid A_i = 1) \\ &= 1 - \delta_i + \delta_i \cdot [1 - F(T \mid X_i, A_i = 1)] \\ &= 1 - \delta_i + \delta_i \cdot S(T|X_i). \end{aligned}$$

konstruieren. Es wird deutlich, dass bei den zensierten Beobachtungen sowohl die Möglichkeit einer späteren Adoption mit Hilfe von $P(A_i = 1) \cdot P(t_i > T \mid A_i = 1)$ als auch die endgültige Nicht- Adoption über $P(A_i = 0)$ erfasst wird. Es ist ebenfalls darauf hinzuweisen, dass natürlich nur für potentielle Nutzer der mögliche Zeitpunkt der Adoption über $P(t_i > T \mid A_i = 1)$ spezifiziert wird. Die resultierende Likelihood-Funktion ist dann mit Hilfe der Maximum-Likelihood-Methode zu maximieren ist. Es ergibt sich:

(5-21) $$L = \prod_{i=1}^{I} [(\delta_i \cdot f(t_i|X_i, A_i = 1))^{C_i} \cdot [1 - \delta_i + \delta_i \cdot S(T|X_i)]^{1-C_i}]$$

bzw. $$\ln L = \sum_{I=1}^{I} [C_i \cdot \ln[\delta_i \cdot f(t_i|X_i, A_i = 1)] + (1 - C_i) \cdot \ln[1 - \delta_i + \delta_i \cdot S(T|X_i)].$$

Die Funktionen $f(\cdot)$ und $S(\cdot)$ beschreiben analog zum Grundmodell die zeitliche Verteilung der Zeitpunkte des Nutzungsbeginns und können z.B. einer der in 5.1.1 dargestellten Verteilungen und Parametrisierung folgen. Noch ungeklärt ist, welcher Verteilung der Split-Parameter δ_i (Schmidt und Witte 1989, S. 151) folgt. Soll dieser Parameter die Wahrscheinlichkeit, zur Gruppe der Adopter bzw. Nicht-Adopter zu gehören, abbilden, erscheint die Annahme eines logistischen Modells in Abhängigkeit der individuenspezifischen Charakteristka analog zur logistischen Regression plausibel:

$$\delta_i = \frac{1}{1+\exp(-\gamma \cdot X_i)}. \quad (5\text{-}22)$$

Es ist darauf hinzuweisen, dass es nicht notwendig ist, die einzelnen Modellkomponenten in Abhängigkeit der gleichen Variablen zu spezifizieren (Greene 1998, S. 744). Darüber hinaus gilt, da δ_i - wie hier beispielsweise angenommen - die Rate oder den Anteil der potentiellen Nutzer darstellt, dass sich dieses verallgemeinerte Modell für $\delta_i = \delta = 1$ auf das Grundmodell der Hazard-Analyse reduziert. Für $\delta_i = \delta \leq 1$ ergibt sich als weitere Möglichkeit ein Split-Modell, bei dem die Wahrscheinlichkeit der Zugehörigkeit zu einer Gruppe konstant, d.h. nicht abhängig von den individuenspezifischen Charakteristika ist (Schmidt und Witte 1989, S. 152). Da δ_i bzw. δ im Kontext der Diffusionsforschung den Anteil potentieller Adopter darstellt, bedeutet dies nichts anderes, als dass das Marktpotential einer Innovation exogen vorgegeben wird (Sinha und Chandrashekaran 1992, S. 122).

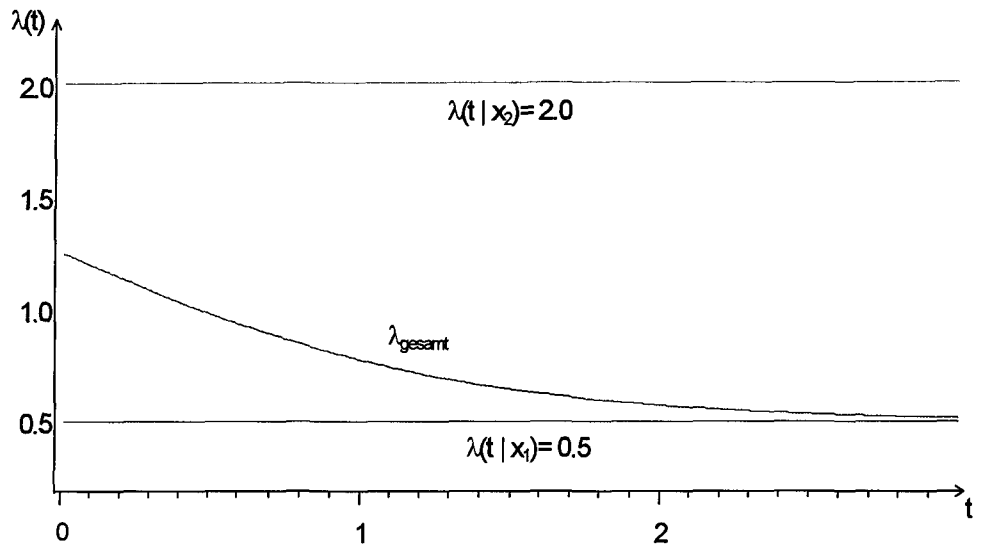
Zusammenfassend lässt sich festhalten, dass der Split-Hazard-Ansatz sowohl die Heterogenität der Individuen bezüglich der hier beispielhaft angenommenen Adoptionsentscheidung als auch bezüglich des Zeitpunkt der Adoption erfasst. Dabei wird Heterogenität bezüglich der Nutzungsentscheidung mittels des Split-Parameters δ_i, der sich in Abhängigkeit des bereits aus der logistischen Regression bekannten Zusammenhanges zwischen Nutzungsentscheidung und Kovaribalen ergibt, berücksichtigt. Die unterschiedlichen Zeitpunkte des Nutzungsbeginns werden analog zum Grundmodell der Hazard-Analyse über die Funktionen f(·) und S(·) erfasst.

5.3 Berücksichtigung unbeobachteter Heterogenität

Mit der Einführung der Split-Komponente wurden die Basis-Hazard-Modelle um einen wesentlichen Faktor zur Erklärung der Heterogenität unter den betrachteten Individuen erweitert. Implizit wird aber auch im Split-Hazard- Modell angenommen, dass zwei Individuen, die durch gleiche Werte für die Kovariablen gekennzeichnet sind, auch gleiche Hazard-Funktionen besitzen. Es ist aber unrealistisch anzunehmen, dass mit Hilfe der erfassten Kovariablen und bei Einbeziehung der Split-Komponente alle Unterschiede zwischen den Individuen berücksichtigt werden (Klein und Moeschberger 1997, S. 405). Die Auswirkungen unbeobachteter Heterogenität auf die Aussagekraft empirischer Modelle und Möglichkeiten zu ihrer Berücksichtigung werden insbesondere in Bezug auf Logit-Modelle intensiv im Marketing diskutiert (Vgl. für eine aktuelle Übersicht und Diskussion Ailwadi, Gedenk und Neslin 1999). Die Folgen unbeobachtbarer Heterogenität im Rahmen der Hazard-

Analyse lassen sich einfach illustrativ darstellen und machen deutlich, dass dieses Phänomen auch im Rahmen der Hazard-Analyse einer ausführlichen Erörterung bedarf:

Abbildung 5-6: Folgen der Mischung zweier konstanter Hazard-Raten



Quelle: In Anlehnung an Vaupel und Yashin 1985, S. 177.

Abbildung 5-6 veranschaulicht die schwerwiegendste Folge von unbeobachteter Heterogenität in Hazard-Modellen: Es ergibt sich unter Umständen eine im Zeitablauf fallende Hazard-Rate, auch wenn die tatsächliche Hazard-Rate für alle Individuen im Sample über die Zeit konstant ist. Dieses Phänomen lässt sich folgendermaßen erklären: Im Sample liegen zwei (oder mehr) Subgruppen vor, die jeweils eine konstante aber voneinander verschiedene Hazard-Rate aufweisen, ohne dass die für die Erklärung dieses Unterschiedes relevanten unabhängigen Variablen erfasst worden sind. Dann resultiert aufgrund der Mischung der beiden Hazard-Raten über das gesamte Sample eine scheinbar im Zeitablauf fallende Hazard-Rate, die nur dadurch begründet ist, dass die Individuen mit dem höheren Risiko (λ_2=2) über den gesamten Zeitraum relativ mehr Ereignisse erfahren. Als Folge besteht das verbleibende Risk Set zunehmend aus Individuen mit relativ geringem Risiko (λ_1=0.5). Daraus folgt, dass die für das gesamte Sample geschätzte Hazard-Rate anfänglich einen Wert von 1,25 (wenn man gleiche Gruppengrößen unterstellt) aufweist und sich dann asymptotisch dem Wert 0,5 nähert und somit den Anschein einer negativen Zeitabhängigkeit erweckt (Vgl. für eine formalere Darstellung Blossfeld und Rohwer 1995, S. 239 ff.). Ebenso möglich ist die Verursachung einer zunächst

scheinbar positiven Zeitdauerabhängigkeit aufgrund von unbeobachteter Heterogenität (Blossfeld und Hamerle 1992, 159 f.). In der Literatur spricht man in diesem Zusammenhang auch von mischenden Verteilungen (So z.B. Allison 1995, S. 235 oder auch Blossfeld und Rohwer 1995, S. 243). Dabei kann es auch zu Mischungen von Hazard-Raten kommen, die unterschiedlichen Verteilungen folgen (Vgl. für eine Vielzahl von derartigen Beispielen Blossfeld und Rohwer 1995, S. 241 ff.). Neben einer resultierenden scheinbaren Verweildauerabhängigkeit der Hazard-Rate können Verzerrungen der Koeffizientenschätzer auftreten, wenn die nicht berücksichtigten bzw. unbeobachtbaren Variablen mit den im Modell explizit berücksichtigten Variablen korreliert sind (Chamberlain 1985, S. 21).

In der Literatur werden verschiedene Ansätze zur Berücksichtigung von unbeobachteter Heterogenität diskutiert (Vgl. u.a. Allison 1995, S. 235 ff., Blossfeld und Rohwer 1995, S. 244 ff., Blossfeld, Hamerle und Mayer 1986, S. 251 ff., Horowitz 1999, Hosmer (Jr.) und Lemeshow 1999, S. 317 ff., S. 1002 ff., Klein und Moeschberger 1997, S. 406). Allerdings stellen Blossfeld und Rohwer 1995, S. 256 fest, dass die einzig erfolgversprechende Strategie zur Überwindung von unbeobachteter Heterogenität die Erhebung besserer bzw. vollständigerer Daten ist. Da dies in der Regel nicht möglich ist, soll im folgenden ein Einblick in Techniken zur Berücksichtigung unbeobachteter Heterogenität in Hazard-Modellen gegeben werden.

Den beliebtesten Ansatz stellt die Berücksichtigung von Heterogenität mittels eines unbeobachtbaren Zufallseffekts dar (Klein und Moeschberger 1997, S. 405), der auf alle Objekte der unbeobachtbaren Teilgruppen des Samples in gleicher Art und Weise einwirkt. In der Literatur werden diese Modelle als "Frailty Models" (so z.B. bei Hosmer (Jr.) und Lemeshow 1999, S. 318) oder auch als "Mixture Models" (Blossfeld und Rohwer 1995, S. 246) bezeichnet.[14] Dieser Zufallseffekt kann in Analogie zur linearen Regression als Störterm ε in die Formulierung der Hazard-Rate eingeführt werden (Wangler 1997, S. 65). Damit erweitert sich die bereits aus (3-4) bekannte Formulierung der Hazard-Rate ganz allgemein zu:

$$h(t_i|X_i, \varepsilon) = h(t_i|X_i) \cdot \varepsilon \qquad (5\text{-}23)$$

[14] Der Begriff "Frailty" stammt aus dem Englischen und bedeutet so viel wie "zerbrechlich". Damit wird der Umstand beschrieben, dass einige Objekte des Samples trotz gleicher Ausprägung der beobachtbaren Kovariablen "zerbrechlicher" sind als andere, woraus die bereits diskutierte scheinbare Zeitabhängigkeit der Hazard-Rate resultiert (Klein und Moeschberger 1997, S. 405).

Bei dieser Formulierung werden in Bezug auf ε folgende Annahmen getroffen (Blossfeld und Rohwer 1995, S. 247 sowie Hosmer (Jr.) und Lemeshow 1999, S. 319):

1. Aufgrund der in Abschnitt 3 unterstellten Positivität der Hazard-Rate gilt, dass $\varepsilon > 0$ sein muss.
2. ε ist eine zeitinvariante, unbeobachtbare Konstante für jedes Objekt des Samples.
3. ε lässt sich als Realisierung eines Zufallsprozesses ausdrücken, der einer über alle Objekte des Samples gleichen Verteilung folgt und unabhängig von den beobachtbaren Kovariablen sowie von eventuell auftretender Zensierung ist.

Wird Gleichung (5-23) in log-linearer Form dargestellt, wird unmittelbar die Ähnlichkeit zum linearen Modell der Regressionsanalyse deutlich. Diese kommt auch und vor allem in Annahme 3 zum Ausdruck. Die modifizierte Dichtefunktion stellt sich dann nach Berücksichtigung einer noch nicht näher spezifizierten Verteilungsannahme für die Heterogenitätskomponente und bei gegebenem Kovariablenvektor sowie nach Umformung gemäß Gleichung (4-5) wie folgt dar:

$$f(t|x) = \int_0^\infty f(t|x, \varepsilon) \cdot dG(\varepsilon) = \int_0^\infty h(t|x, \varepsilon) \cdot S(t|x, \varepsilon) \cdot dG(\varepsilon) \tag{5-24}$$

Dabei wird die von ε nach Annahme 3 unbedingte Verteilung $f(t|x)$ als Mischverteilung bezeichnet, mit G(ε) als mischende Verteilung. Um in Gleichung (5-23) über G(ε) zu integrieren, muss G(ε) vollständig spezifiziert sein, d.h. es muss eine Verteilungsannahme für ε getroffen werden. Ist diese gegeben, können die unbekannten Parameter von G(ε) zusammen mit den Parametern von $f(t|x)$ bzw. $h(t|x)$ und $S(t|x)$ auf Basis der ML-Methode geschätzt werden. Die Maximierung der Log-Likelihood-Funktion kann allerdings mit erheblichen numerischen Problemen einhergehen, da bei jedem Iterationsschritt eine Integration bezüglich der Verteilung von ε nötig ist. Dies wird aus der Darstellung der resultierenden, sogenannten „marginalen" Log-Likelihood-Funktion in (5-25) deutlich, wobei C_i wieder den Zensierungsindikator bezeichnet (Blossfeld, Hamerle und Mayer 1986, S. 97 f.):

$$\ln L_M = \sum_{i=1}^{I} \ln \int_0^\infty h(t_i|x_i, \varepsilon)^{c_i} \cdot S(t_i|x_i, \varepsilon)\, dG(\varepsilon) \tag{5-25}$$

Wird die Heterogenitätskomponente – wie im folgenden – als stetig mit der Dichtefunktion g(ε) angenommen ergibt sich:

$$(5\text{-}26) \qquad \ln L_M = \sum_{i=1}^{I} \ln \int_0^{\infty} h(t_i|x_i, \varepsilon)^{c_i} \cdot S(t_i|x_i, \varepsilon) \, g(\varepsilon) d\varepsilon$$

Eine weitere, praktische Problematik bei der Berücksichtigung unbeobachteter Heterogenität bei Hazard Modellen besteht in dem Umstand, dass es so gut wie keine Software zur Berücksichtigung der Heterogenitätskomponente gibt. Zwar existieren individuelle Lösungen, von deren Verwendung aber aufgrund einer fehlenden Testung abgeraten wird. Stattdessen wird die Benutzung der wenigen, bereits heute in Standard-Softwarepaketen implementierten Routinen empfohlen (Hosmer (Jr.) und Lemeshow 1999, S. 320). Im folgenden wird daher die Berücksichtung unbeobachteter Heterogenität anhand eines Weibull-Regressionsmodell mit einer gammaverteilten Heterogenitätskomponente, wie in dem Softwarepaket LIMDEP realisiert, konkretisiert. In der Literatur wird auf die (einparametrische) Gamma-Verteilung zudem als beliebte bzw. am häufigsten verwendete Verteilungsannahme verwiesen (Klein und Moeschberger 1997, S. 406 bzw. Hosmer (Jr.) und Lemeshow 1999, S. 319), die auch schon im Marketing Anwendung gefunden hat (Gönül und Srinivasan 1993, S.1222).

Legt man eine Gamma-Verteilung zugrunde und berücksichtigt die Annahme, dass die Hazard-Rate im Durchschnitt tatsächlich $h(t_i|X_i)$ entspricht, ergibt sich für ε die folgende Dichtefunktion (Vgl. zu diesem Abschnitt Greene 2000, S. 946 f., Klein und Moeschberger 1997, S. 410 ff. sowie ausführlicher Blossfeld, Hamerle und Mayer 1986, S. 97 ff. oder auch Blossfeld und Rohwer 1995, S. 252 ff.):

$$(5\text{-}27) \qquad f(\varepsilon) = \frac{\varepsilon^{(1/\theta - 1)} \cdot \exp(-\varepsilon/\theta)}{\Gamma[1/\theta] \cdot \theta^{1/\theta}} \qquad \theta > 0$$

$\Gamma(1/\theta)$ bezeichnet die Gamma-Funktion. Die Annahme, dass die Hazard-Rate im Durchschnitt tatsächlich $h(t_i|X_i)$ entspricht, bedeutet, dass sich als Erwartungswert der Heterogenitätskomponente $E(\varepsilon) = 1$ ergibt. Dies resultiert in einer Identität der die Gammaverteilung in ihrer allgemeineren Form determinierenden zwei Parameter. Daher wird diese Verteilung in ihrer Darstellung gemäß (5-27) auch als einparametrische Gammaverteilung (Klein und Moeschberger 1997, S. 406) bezeichnet mit einem Erwartungswert $E(\varepsilon) = 1$ und der Varianz $Var(\varepsilon) = \theta$.

Berücksichtigt man diese Formulierung, ergibt sich folgende Survivorfunktion für ein Weibull-Regressionsmodell mit Gamma-Heterogenität:

$$(5\text{-}28) \qquad S(t) = \int_0^\infty S(t|\varepsilon) \cdot f(\varepsilon) d\varepsilon = [1+\theta \cdot (\lambda t)^\alpha]^{-1/\theta}.$$

Die zugehörige Hazard-Rate ergibt sich als

$$(5\text{-}29) \qquad h(t_i|X_i) = \lambda_i \cdot \alpha \cdot (\lambda_i t_i)^{\alpha-1} \cdot [S(t)]^\theta.$$

Hier wird unmittelbar die Verwandtschaft zum Weibull-Modell ohne Heterogenität deutlich (Vgl. Abschnitt 5.1.1.2), da sich die beiden Hazard-Raten nur um den Faktor $[S(t)]^\theta$ unterscheiden. Der Effekt der Heterogenität auf die Hazard-Rate ist um so größer ist, je ausgeprägter die Varianz der Störgröße ε ausfällt. Dies ist auch intuitiv verständlich: Sind die Untersuchungsobjekte hinsichtlich nicht erhobener Merkmale wenig heterogen oder wurden sogar alle Merkmale vollständig erhoben, kann nur ein geringer bzw. gar kein Effekt der nicht beobachteten Größen auf die Hazard-Rate resultieren, da die unbeobachteten Größen nur relativ wenig zur Unterscheidung zwischen den untersuchten Objekten beitragen. Folgerichtig fällt dann die mittlere quadratische Abweichung von ε entsprechend klein aus bzw. nimmt sogar den Wert Null an.

Die wesentlichen Vorteile des hier vorgestellten Ansatzes bestehen in der Möglichkeit einer konsistenten Schätzung der Parameter des resultierenden Regressionsmodells und der Vielfalt möglicher Modellierungen der Heterogenitätskomponente auf Basis unterschiedlicher Verteilungsannahmen. Zugleich existiert aber auch wesentliche Kritik an diesem Ansatz, die im folgenden kurz aufgezeigt werden soll.[15]

Zunächst muss Annahme 3, die Unabhängigkeit von Heterogenitätskomponente und beobachtbaren Kovariablen, in Frage gestellt werden. Die Annahme, dass die nichtbeobachtbaren individuenspezifischen Merkmale völlig unabhängig von den beobachtbaren Merkmalen seien, erscheint unrealistisch, so dass trotz Berücksichtigung einer Heterogenitätskomponente mit einer Verzerrung der Resultate gerechnet werden muss (Blossfeld, Hamerle und Mayer 1986, S. 100). Dies ist aber keine unerwartete Erkenntnis. Überraschender ist der Hinweis von Allison (Allison 1995, S.

---

[15] Diesen Problemen wird in der (statistischen) Literatur viel Aufmerksamkeit geschenkt. Eine ausführlichere Diskussion würde jedoch den Rahmen und den Anspruch dieser Arbeit sprengen, so dass auf die entsprechende Literatur verwiesen wird. Einen gut verständlichen Überblick gibt hier z.B. Wangler 1997, S. 73 ff.

236) auf die Ergebnisse von Gail, Wieand und Piantadosi 1984, die zu dem Ergebnis gelangen, dass selbst im Falle der Unabhängigkeit die Berücksichtigung unbeobachteter Heterogenität die Parameterschätzer der Kovariablen in Richtung Null abschwächt. Diese Ergebnisse können allerdings nicht ohne weiteres auf alle Modellklassen und –typen übertragen werden. Insbesondere weisen die Autoren darauf hin, dass u.a. für das Weibull-Regressionsmodell noch Forschungsbedarf bestehe (Gail, Wieand und Piantadosi 1984, S. 443).

Ein weiterer Kritikpunkt besteht in dem Problem der Identifizierbarkeit der Modelle (Allison 1995, S. 236). So ist es prinzipiell möglich, eine vorhandene Datensituation durch zwei oder sogar mehr grundsätzliche verschiedene strukturelle Modelle zu erklären (Heckmann und Singer 1984a, S. 80, Heckmann und Singer 1984b, S. 273 f. und S. 276). Dieses Problem lässt sich nur mittels Parameterrestriktionen lösen z.B. hinsichtlich der funktionalen Form der Hazard-Rate (Heckmann und Singer 1984b, S. 275). Derartige Restriktionen lassen sich aber in der Regel nicht ökonomisch begründen.

Dies gilt in der Regel auch für die Auswahl der Mischverteilung. Hier wird ebenfalls das Fehlen einer ökonomischen Begründung für die Wahl einer bestimmten Verteilung beklagt und darauf hingewiesen, dass eher technische als inhaltliche Überlegungen die Entscheidung für oder gegen eine bestimmte Verteilung determinieren (Heckmann und Singer 1984b, S. 274). Darüber hinaus ist die praktische Auswahl einer Mischverteilung noch aus einem anderen Grund problematisch. Die Darstellung der marginalen Log-Likelihood-Funktionen in Gleichung (5-25) bzw. (5-26) zeigt, dass die Wahl der Verteilungsfunktion von ε die Form der marginalen Log-Likelihood-Funktion und damit auch die Parameterschätzer beeinflussen kann. Insbesondere Heckmann und Singer 1984b setzen sich intensiv mit dieser Problematik auseinander und demonstrieren beispielhaft, wie Schätzergebnisse in Abhängigkeit der zugrundegelegten Verteilungsannahme für die Heterogenitätskomponente z.T. dramatisch von einander abweichen können (Heckmann und Singer 1984b, S. 276). Derartige Abweichungen zeigen sich jedoch nicht in allen empirischen Analysen, die sich mit der Frage der Sensitivität der Schätzergebnisse in Bezug auf getroffene Verteilungsannahmen beschäftigen. So können Newman und McCulloch 1984 die Ergebnisse von Heckmann/Singer nicht in dem geschilderten Umfang bestätigen. Sie stellen vielmehr fest, dass die Parameterschätzer relativ stabil waren unabhängig von der Verteilung von ε. Allerdings zeigte sich, dass bei diskreter Approximation der Verteilung von ε die Schätzer insofern sensibel waren, als dass in Abhängigkeit

der in der diskreten Approximation modellierten Varianz eine Art Kompensationseffekt bei den Schätzwerten der erklärenden Variablen des strukturellen Modells zu beobachten war. Trotz dieses Effektes können sie keine qualitative Änderung ihrer Ergebnisse feststellen (Newman und McCulloch 1984, S. 957 f.).

Aufgrund dieser Probleme schlagen Heckmann/Singer eine simultane Schätzung der strukturellen Modellparameter und der Verteilung der Heterogenitätskomponente vor. Dabei handelt es sich um eine nicht-parametrische Vorgehensweise, bei der nun keine ex-ante Spezifikation der Verteilung von ε mehr notwendig ist (Vgl. ausführlich Heckmann und Singer 1984b, S. 300 ff.). Stattdessen wird nun die Verteilung von ε durch eine diskrete Verteilung mit einer empirisch zu bestimmenden Anzahl an Stützstellen spezifiziert. Auf Basis der vorhandenen Daten werden dann in einem iterativen Prozess die Position und Wahrscheinlichkeitsmasse je Stützstelle bestimmt. Vilcassim/Jain zeigen eine Anwendung dieser Vorgehensweise bei der empirischen Untersuchung des Markenwechselverhaltens. Sie leiten dabei aus nur drei Stützstellen die Verteilung der Heterogenitätskomponente ab und weisen darauf hin, dass eine analoge Gestalt dieser Verteilung von ε auch mit Hilfe des beschriebenen parametrischen Ansatzes zu erzielen sei (Vilcassim und Jain 1991 , S. 34 und 37 f.).

Allerdings ist auch dieser Ansatz nicht frei von Kritik. So weisen Blossfeld/Rohwer auf eine Untersuchung von Trussel/Richards (Trussell und Richards 1985) hin, die zeigen, dass nun die Ergebnisse bezüglich der Verteilung der Heterogenitätskomponente abhängig sind von der parametrischen Spezifizierung der Hazard-Rate (Blossfeld und Rohwer 1995, S. 255). Dies ist unter Umständen ein weiterer Beleg für den bereits beschriebenen, von Newman und McCulloch festgestellten Kompensationseffekt. Aber auch eine vollständige nicht-parametrische Modellierung sowohl der Verweildauerabhängigkeit als auch der Heterogenität ist mit schwerwiegenden Problemen hinsichtlich Identifizierbarkeit und Sensitivität der Schätzergebnisse belastet und stellt damit ebenfalls keine Lösung dar (Blossfeld und Rohwer 1995, S. 256). Darüber hinaus zeigt die Diskussion bei Bearse, Canals und Rilstone 1998, dass bezüglich der Eigenschaften der resultierenden Schätzer noch Forschungsbedarf besteht, so dass Vorgehensweisen und Schlussfolgerungen bei Verwendung dieses Verfahrens z.T. auf Ad hoc Annahmen beruhen, denen dann eine theoretische Fundierung fehlt (Bearse, Canals und Rilstone 1998, S. 153).

Die Diskussion des Phänomens unbeobachteter Heterogenität in Hazard-Modellen hat gezeigt, dass es weder eine einfache noch eine eindeutige Lösung für dieses

Problem gibt (Blossfeld und Rohwer 1995, S. 256). Darüber hinaus muss man berücksichtigen, dass mit Hilfe der hier aufgezeigten Verfahren zwar unter Umständen Verbesserungen der Schätzergebnisse in technischer Hinsicht möglich sind, inhaltlich aber keine Vorteile bei der Interpretation erzielt werden können. So sind Jain/Vilcassim in ihrer o.a. Arbeit auch nur in der Lage nachzuweisen, dass ein großer Teil der beobachten Varianz bezüglich des Markenwechsel auf unbeobachtete Faktoren zurückzuführen ist, über deren inhaltliche Bedeutung sie dann allerdings nur spekulieren können (Vilcassim und Jain 1991, S. 38). Dieser Eindruck wird durch einen von Wangler durchgeführten Vergleich von Hazard-Modellen mit und ohne Berücksichtigung der Heterogenitätskomponente bestätigt (Wangler 1997, S. 93 ff.). Eine weitere, umfassende Diskussion dieses Themas findet sich bei Heckmann und Taber 1994, wobei deutlich wird, dass noch keine abschließenden Lösungen gefunden worden sind, sondern noch bedeutender Forschungsbedarf in methodischer Hinsicht besteht.

5.4 Das Proportional-Hazard-Modell von Cox

Stehen keine ausreichenden Vorabinformationen z.B. in Form von theoriegeleiteten Überlegungen für eine genaue Spezifikation der zeitlichen Abhängigkeit der Hazard-Rate zur Verfügung, ist eine vollständige Parametrisierung, wie sie im Abschnitt 5.1.1 vorgestellt worden ist, nicht möglich bzw. schwer zu rechtfertigen (Allison 1984, S. 33). In diesen Fällen stellt das 1972 von Cox (Cox 1972) erstmals vorgeschlagene Proportional-Hazard-Modell eine angemessene Modellvariante dar (Hutchinson 1988, S. 213). Hierbei handelt es sich um einen semiparametrischen Ansatz mit einer unspezifizierten Baseline-Hazard-Rate, die auch Grund-Hazard-Rate genannt wird. Die Hazard-Rate des Proportional-Hazard-Modells von Cox ist unter Berücksichtigung von k Kovariablen gegeben durch:

$$\lambda(t|x) = \lambda_0(t) \cdot \exp(\beta_1 \cdot x_1 + \ldots + \beta_k \cdot x_k) \tag{5-30}$$

Damit stellt sich die Hazard-Rate als Produkt zweier Funktionen dar: $\lambda_0(t)$ beschreibt, wie sich die Hazard-Rate in Abhängigkeit der Zeit- bzw. Überlebensdauer verhält, während $\exp(x'\beta)$ die individuenspezifischen Einflüsse auf die Hazard-Rate abbildet (Hosmer (Jr.) und Lemeshow 1999, S. 90). Dabei ist $\lambda_0(t)$ eine unspezifizierte Baseline-Hazard-Rate. Sie bildet die Hazard-Rate unter der Annahme ab, dass die Kovariablen alle ohne Einfluss sind. Es wird also nur der Einfluss der Kovariablen auf die Hazard-Rate parametrisiert, ohne dass eine parametrische Spezifikation der zeitlichen Abhängigkeit der Hazard-Rate selbst ex ante vorgegeben wird. In den meisten Fällen wird eine exponentielle Parametrisierung des Einflusses der Kovariablen vorgenommen wie in Gleichung (5-30). Dies führt zu einer linearen Ab-

hängigkeit der logarithmierten Hazard-Rate von den exogenen Variablen, was eine besonders einfache Interpretation des Einflusses dieser Variablen ermöglicht. Allerdings soll an dieser Stelle der Vollständigkeit halber darauf hingewiesen werden, dass auch alternative Modellierungen des Kovariablen-Einflusses möglich sind (Koop und Ruhm 1993, S. 421 ff.). Das Modell von Cox bietet sich also insbesondere in Situationen an, in denen das Hauptaugenmerk auf der Ermittlung des Einflusses der Kovariablen z.B. bei der Analyse von Marketingmaßnahmen liegt (Li 1995, S. 20). Dadurch, dass keine spezifische Zeitabhängigkeit der Hazard-Rate unterstellt wird, kann in diesem Zusammenhang ein höheres Maß an Flexibilität in der Modellierung erreicht werden.

Eine weitere wesentliche Eigenschaft des Cox-Modells ist die Proportionalität individuen-spezifischer Hazard-Raten, die man analog zu der Vorgehensweise in den Abschnitten 5.1.1.1 und 5.1.1.2 erkennen kann, wenn man den Quotienten der Hazard-Raten zweier beliebiger Untersuchungsobjekte bildet:

$$\frac{\lambda_i(t|x_i)}{\lambda_j(t|x_j)} = \exp\left[\beta_1(x_{i1} - x_{j1}) + \ldots + \beta_k(x_{ik} - x_{jk})\right] \tag{5-31}$$

Dieser Ausdruck hängt nicht mehr von der Baseline-Hazard-Rate, sondern nur von den interessierenden Kovariablen ab. Genau diese Eigenschaft des Proportional-Hazard Modells kann man sich für die Ermittlung der Schätzer für β zu nutze machen, ohne die Baseline-Hazard-Rate bestimmen zu müssen. Dazu wird die resultierende Likelihood-Funktion

$$L = \prod_{i=1}^{I}\left[\lambda_0(t_i)\cdot\exp(x_i'\beta)\right]^{c_i}\cdot\exp\left[-\int_0^{t_i}\lambda_0(u)\cdot\exp(x_i'\beta)du\right] \tag{5-32}$$

um den Ausdruck

$$\sum_{j\in R(t_i)}\lambda_0(t_i)\cdot e^{x_j'\beta} \tag{5-33}$$

erweitert (Blossfeld, Hamerle und Mayer 1986, S. 76), wobei $R(t_i)$ die Risikomenge bezeichnet, also alle Objekte mit Überlebens- oder Zensierungszeiten, die größer oder gleich der interessierenden Zeit t_i sind (Hosmer (Jr.) und Lemeshow 1999, S. 94). Auf der rechten Seite von Gleichung (5-32) ergibt sich dann nach Kürzen der Baseline-Hazard-Rate als erster Faktor der Ausdruck

(5-34) $$\prod_{i=1}^{I}\left[\frac{e^{x_i'\beta}}{\sum_{j\in R(t_i)} e^{x_j'\beta}}\right]^{c_i}.$$

Cox schlug vor, diesen Faktor, der nur von den interessierenden Koeffizienten β abhängt und als „Partial Likelihood" bezeichnet wird, wie eine gewöhnliche LikelihoodFunktion zu behandeln und in Abhängigkeit von β zu maximieren (Blossfeld, Hamerle und Mayer 1986, 77). Bei dieser Vorgehensweise wird der zweite nach Erweiterung von Gleichung (5-32) mit dem Ausdruck nach (5-33) resultierende Term vernachlässigt. Dieser enthält aber ebenfalls noch den bzw. die Parameter β. Cox spekulierte trotz des bei dieser Vorgehensweise auftretenden Informationsverlustes darauf, dass die resultierenden „Partial Likelihood" Schätzer die gleichen (Verteilungs-) Eigenschaften wie die Schätzer auf Basis der vollständigen Likelihood-Funktion aufweisen (Hosmer (Jr.) und Lemeshow 1999, S. 95). Der Nachweis, dass die resultierenden Schätzer tatsächlich wünschenswerte Eigenschaften wie asymptotische Konsistenz und asymptotische Normalität besitzen, erfolgte dann später (Blossfeld, Hamerle und Mayer 1986, S. 77). Gleichung (5-34) lässt sich intuitiv als (bedingte) Wahrscheinlichkeit, dass in einem bestimmten Zeitpunkt nur für ein bestimmtes Individuum gerade ein Ereignis stattfindet, interpretieren. Der Partial Likelihood ergibt sich dann als Produkt dieser individuellen Wahrscheinlichkeiten über alle i Zeitpunkte. Daraus resultiert, dass in Unkenntnis der Baseline nur die beobachtbare Reihenfolge der Ereignisse Informationen über die unbekannten Parameter β liefert (Kiefer 1988, S. 668). Dementsprechend ist bei der Anwendung zu beachten, dass die Zeitdauern t_i ausreichend genau erfasst werden, so dass zumindest keine messtechnisch bedingten identischen Werte für die Zeitdauern, sogenannte Ties, vorliegen. Ansonsten ist eine Korrektur des Partial Likelihood notwendig oder bei zu großer Anzahl von Ties der Übergang zu einem diskreten Hazard-Modell (Blossfeld, Hamerle und Mayer 1986, S. 77).

Abschließend soll auf den Zusammenhang zwischen dem Proportional-Hazard Modell von Cox und den parametrischen Modellen hingewiesen werden. Dieser wird deutlich, wenn man für die Baseline-Hazard-Rate einen bestimmten zeitlichen Verlauf spezifizieren würde. Als Ergebnis erhielte man ein voll parametrisiertes Proportional-Hazard-Modell. Die im vorangegangen Abschnitt vorgestellten Weibull- bzw. Exponential-Regressionsmodelle sind zwei Beispiele hierfür (Vgl. u.a. Allison 1995, S. 113 und Hamerle 1987, S. 253).

5.5 Zeitvariable Kovariable

Üblicherweise werden in der Literatur Kovariablen als zeitvariabel bezeichnet, deren Werte sich im Zeitablauf ändern. Yamaguchi weist jedoch darauf hin, dass auch solche Kovariablen als zeitvariabel bezeichnet werden können, deren Einfluss, d.h. deren Koeffizienten im Zeitablauf variiert (Yamaguchi 1993, 280 f.). Dieser Aspekt von Zeitvariabilität soll im Folgenden nicht weiter betrachtet werden. Stattdessen sei auf Yamaguchi 1993 verwiesen. Nach diesen einleitenden Bemerkungen erfolgt die Betrachtung zeitvariabler Kovariable in zwei Schritten. In einem ersten Schritt muss zwischen verschiedenen Arten zeitvariabler Kovariabler unterschieden werden. Im zweiten Schritt wird geklärt, wie diese Einflussgrößen in Abhängigkeit der Art der Zeitvariabilität, aber auch in Abhängigkeit des zugrundeliegenden Hazard-Modells berücksichtigt werden können.

Eine gängige Einteilung von zeitvariablen Kovariablen geht auf Kalbfleisch/Prentice Kalbfleisch und Prentice 1980, S. 122 ff.) zurück. Diese unterscheiden zwischen determinstischen und stochastischen zeitvariablen Kovariablen. Deterministische Kovariable werden auch als „definiert" bezeichnet. Mit diesem Begriff werden Einflussgrößen beschrieben, deren Abhängigkeit von der Zeit sich in eine vorher festgelegte funktionale Form fassen lässt (Blossfeld, Hamerle und Mayer 1986, S. 90). Stochastische zeitvariable Einflussgrößen werden in einem zweiten Schritt noch in interne und externe zeitvariable Kovariable eingeteilt. Bei nach dieser Einteilung externen Kovariablen wird der Wert dieser Variablen nicht von dem eigentlich interessierendem und zu beobachtendem Prozess beeinflusst. Zu diesem Typ von Variablen zählen insbesondere wirtschaftliche Rahmenbedingungen, die auf alle Untersuchungsobjekte gleichermaßen einwirken, wie z.B. sich im Zeitablauf ändernde Preise für bestimmte im Rahmen der Untersuchung relevante Güter und Dienstleistungen. Im Gegensatz dazu handelt es sich bei internen zeitvariablen Kovariablen um individuenspezifische Variablen, deren Wert durch die Untersuchungsobjekte während der Untersuchung selbst generiert wird. Ein Beispiel sind die Erfahrungen eines Individuums im Umgang mit dem Untersuchungsgegenstand z.B. der Umfang der Berufserfahrung bei der Analyse des Arbeitsplatzwechsels. Darüber hinaus können diese Variablen danach unterschieden werden, ob es sich um diskrete oder kontinuierliche zeitveränderliche Größen handelt (Peterson 1986a, S. 281 f.), also um Größen, die ihren Wert ständig oder nur zu bestimmten Zeitpunkten ändern. Letztere haben grafisch die Gestalt einer Treppenfunktion. Problematisch in der Behandlung sind insbesondere interne zeitvariable Kovariable. Diese stellen zum einen Einflussgrößen des zu untersuchenden Prozesses dar, werden aber ihrerseits durch genau diesen beeinflusst. Es liegt also ein Rückkopplungseffekt vor, so dass diese Prozesse auch als interdependent oder als dynamisches System bezeichnet werden (Blossfeld und

(Blossfeld und Rohwer 1995, S. 122. Die detailierte Darstellung der Behandlung interner zeitvariabler Kovaribale würde den Rahmen und Anspruch dieser Arbeit sprengen. Daher wird an dieser Stelle bezüglich der weiteren Behandlung dieser Prozesse auf die zitierte Literatur, insbesondere auf Blossfeld und Rohwer 1995, S. 123 ff. verwiesen.

Zur Berücksichtigung zeitvariabler Kovariablen in Hazard-Modellen können drei Standard-Vorgehensweisen unterschieden werden (Blossfeld und Rohwer 1995, S. 120):

1. Ist die funktionale Abhängigkeit bei „definierten" Kovariablen nicht zu komplex, kann diese direkt in der für die Schätzung resultierenden Log-Likelihood-Funktion erfasst und die Optimierung vorgenommen werden (Blossfeld, Hamerle und Mayer 1986, S. 91).
2. Bei Kovariablen, deren Wert sich nur zu bestimmten Zeitpunkten ändert bietet sich sowohl bei parametrischen wie auch bei dem semiparametrischen Cox-Modell die Methode des Episoden-Splitting an (Blossfeld und Rohwer 1995, S. 120).
3. Werden sich kontinuierlich ändernde Einflussfaktoren berücksichtigt, kann versucht werden, dies durch die Approximation mittels einer Treppenfunktion bzw. einer stückweise linearen Funktion abzubilden (Peterson 1991, S. 292 ff.).

Dabei bedürfen insbesondere die unter den Punkten 2 und 3 angegebenen Verfahren noch einer weitergehenden Erläuterung.

Beiden Verfahren liegt eine grundsätzlich gleiche Vorgehensweise zugrunde. Dabei wird die betrachtete Gesamtepisode in mehrere Teil- bzw. Subepisoden unterteilt bzw. aufgesplittet. Innerhalb der gebildeten Subepisoden wird dann die Hazard-Rate auf Basis der aktuellen, innerhalb dieser Subepisode konstanten bzw. als konstant angenommen Werte der zeitvariablen Kovariable geschätzt. Die Hazard-Rate für die Gesamtepisode wird dann auf Basis der Ergebnisse aus den Teilepisoden durch Summation bestimmt (Newman und McCulloch 1984, S. 946; Peterson 1991, S. 292 f.). Der Unterschied zwischen den beiden Methoden besteht zum einen in der Bestimmung der Subepisoden. Das Episoden-Splitting wird bei zeitvariablen Kovariablen angewendet, deren Verlauf die Gestalt einer Treppenfunktion hat. Dabei wird zu jedem Zeitpunkt, an dem eine Wertänderung auftritt, eine neue Subepisode gebildet (Blossfeld und Rohwer 1995, S. 128). Die Vorgehensweise ist bei sich kontinuierlich ändernden Einflussgrößen ähnlich. Bei diesen Variablen werden jedoch mit

einer gewissen Willkür die Zeitpunkte, an denen Teilepisoden gebildet werden, und auch die Anzahl der zu bildenden Teilepisoden bestimmt, wobei mit einer feineren Unterteilung tendenziell auch eine Verbesserung der Genauigkeit der Schätzergebnisse einhergeht (Allison 1995, S. 107). Hier wird auch gleich der zweite wesentliche Unterschied deutlich. Bei sich kontinuierlich ändernden Kovariablen ist es nicht möglich für die Subepisoden einen exakten Wert anzugeben. Stattdessen kann man versuchen, den Verlauf dieser Variablen entweder durch Treppenfunktionen oder durch stückweise lineare Funktion zwischen den Messpunkten zu approximieren, so dass im Ergebnis wieder die Hazard-Raten der einzelnen Abschnitte variieren und so den Effekt der zeitvariablen Einflussgrößen abbilden (Allison 1995, S. 104 ff.). Im Marketing wird dieses Verfahren beispielsweise von Gupta im Rahmen seines Modellvergleich zur Abbildung des Wiederkaufverhaltens diskutiert (Gupta 1991, S. 6), allerdings ohne in der abschließenden Beurteilung einen Vergleich von Modellen mit und ohne zeitvariable Einflussgrößen vorzunehmen.

Während die Idee dieser Verfahren relativ einfach erscheint, stellt die Sammlung und Aufbereitung des notwendigen Datenmaterials unter Umständen eine Herausforderung dar (Hosmer (Jr.) und Lemeshow 1999, 251). Über die bisher benötigten Daten hinaus, wie Länge der Gesamtepisode, Anfangs- und Endzustand des Untersuchungsobjektes und Werte der Kovariablen müssen bei zeitabhängigen Einflussfaktoren zusätzlich auch die Zeitpunkte auftretender Wertänderungen sowie die daraus resultierende Dauer, die Anfangs- und Endzustände sowie der Zensierungsstatus je Subepisode erfasst werden. Dies setzt gegebenenfalls eine kontinuierliche Datenerhebung voraus, die vielfach nicht oder nur unter großen Umständen realisierbar sein wird. Insbesondere beim semiparametrischem Cox-Modell müssen zu jedem Zeitpunkt, an dem ein Ereignis stattfindet, die Werte der Kovariablen bekannt sein. Wurden die zeitvariablen Einflussfaktoren nicht permanent beobachtet, bietet es in diesen Fällen an, den beobachtbaren Wert für die entsprechende Kovariable zu verwenden, der zeitlich dem beobachteten Ereignis bei einem Untersuchungsobjekt am nächsten liegt (Blossfeld, Hamerle und Mayer 1986, S. 92).

Nachdem die Idee, Arten und das prinzipielle Vorgehen bei der Berücksichtigung zeitvariabler Kovariable dargestellt worden sind, ist abschließend unbedingt darauf hinzuweisen, dass besondere Vorsicht bei der Berücksichtigung dieser Variablen geboten ist (Allison 1995, S. 98; Heckmann und Singer 1984a, S. 81). Dieser Hinweis liegt in dem Umstand begründet, dass diese Variablen gemäß der o.a. Klassifizierung identifiziert werden müssen, um insbesondere bei der Berücksichtigung der hier nicht näher diskutierten internen Kovariablen die herrschende Interdependenz

abzubilden (Hosmer (Jr.) und Lemeshow 1999, S. 248). Darüber hinaus muss sichergestellt werden, dass die Effekte dieser zeitvariablen Einflussgrößen von der Verweildauerabhängigkeit des Prozesses insgesamt getrennt werden können, was hohe Anforderungen an das vorhandene Datenmaterial im Hinblick auf die Variation von x(t) stellt, da ansonsten ein Multikollinearitätsproblem auftreten kann (Heckmann und Singer 1984a, S. 81).

6 Überprüfung der Modellannahmen und Hypothesentests

Den in den vorangegangenen Abschnitten vorgestellten parametrischen Modellen und dem semiparametrischen Ansatz von Cox liegen Annahmen bezüglich der Verteilung bzw. bezüglich der Proportionalität der Hazard-Rate beliebiger Subpopulationen zugrunde. In diesem Abschnitt sollen Testverfahren vorgestellt werden, mit deren Hilfe die Überprüfung dieser Annahmen vorgenommen werden kann. Diese Vorgehensweise ist analog zur Vorgehensweise bei der linearen bzw. logistischen Regression, bei denen ebenfalls Tests der vielfältigen Modellannahmen notwendig sind, um die Aussagekraft der statistischen Schätzungen abzusichern (Diekmann und Mitter 1984, S. 131). Zunächst werden Tests zur Überprüfung der Modellannahmen bei parametrischen Regressionsmodellen, dann bei dem semiparametrischen Modell von Cox und abschließend zur Überprüfung der resultierenden Koeffizienten selbst vorgestellt. Dabei kann unmöglich auf alle existieren Verfahren eingegangen werden, so dass nur eine Auswahl dargestellt werden kann.

6.1 Überprüfung der Modellannahmen und Modellvergleiche für parametrische Regressionsmodelle

Die Überprüfung der Verteilungsannahmen bei den parametrischen Modellen erfolgt mit Hilfe grafischer Verfahren. Bei den grafischen Verfahren können prinzipiell zwei Vorgehensweisen unterschieden werden. Zum einen wird versucht, durch Transformation einen linearen Zusammenhang zwischen der Zeitdauer und der zugrundeliegenden Verteilungsannahme zu konstruieren (Klein und Moeschberger 1997, S. 389). Dabei ergeben sich die in Tabelle 6-1 nachfolgend dargestellten linearen Zusammenhänge. Unter Berücksichtigung der Beziehung zwischen Survivorfunktion und Hazard-Rate gemäß Tabelle 3-1 wird deutlich, dass diese Linearisierungen Funktionen der Survivorfunktion bzw. der kumulativen Hazard-Rate $H(t) = -\ln S(t)$ sind, was exemplarisch für die Exponential-Verteilung deutlich gemacht worden ist.

Tabelle 6-1: Linearisierungen für ausgewählte Verteilungsannahmen

Funktion / Verteilung	S(t)	Linearisierung
Exponential-Verteilung	$\exp(-\lambda \cdot t)$	$-\ln S(t) = H(t) = \lambda t$
Weibull-Verteilung	$\exp\left(-(\lambda \cdot t)^{\alpha}\right)$	$\ln[-\ln S(t)] = \alpha \ln t + \alpha \ln \lambda$
Log-Normalverteilung	$1 - \Phi(\alpha \ln(\lambda \cdot t))$	$\Phi^{-1}[1 - S(t)] = \alpha \ln t + \alpha \ln \lambda$ [16]
Log-Logistische-Verteilung	$\frac{1}{1+(\lambda \cdot t)^{\alpha}}$	$\ln\left[\frac{1-S(t)}{S(t)}\right] = \alpha \ln t + \alpha \ln \lambda$

Quelle: Eigene Erstellung in Anlehnung an Christensen 1999, S. 65 und Klein und Moeschberger 1997, S. 389

Trägt man nun die einzelnen Zusammenhänge gegeneinander ab, so muss sich bei Gültigkeit der Verteilungsannahme eine Gerade ergeben, so dass sich z.B. im Fall der Exponentialverteilung, bei der die Schätzung $-\ln\hat{S}(t)$ bzw. $\hat{H}(t)$ gegen t abgetragen wird, eine Ursprungsgerade mit der Steigung λ ergibt. Die Schätzung $\hat{S}(t)$ wird dabei mit Hilfe des nicht-parametrischen Kaplan-Meir Schätzers ermittelt, der sich einfach als Verhältnis von Individuen, deren Ereigniszeitpunkt größer als der betrachtete Zeitpunkt t, und der Gesamtzahl der betrachten Individuen ergibt (Vgl. hierzu ausführlicher Allison 1995, S. 30 f.). Bei diesem Verfahren macht man sich den Umstand zu Nutze, dass die Survivorfunktion bzw. die kumulative Hazard-Rate einen monoton fallenden bzw. steigenden Verlauf haben, der dann nur noch linear zu transformieren ist. Diese Vorgehensweise ist jedoch auf den univariaten Fall beschränkt, d.h. es werden keine Kovariablen berücksichtigt. Daher ist es durchaus möglich, dass Modelle, die nach diesem Test eine gute Anpassung erwarten lassen, unter Berücksichtigung von Kovariablen u.U. nur einen schlechten Fit erzielen und umgekehrt, so dass die Aussagekraft dieser Tests allenfalls eingeschränkt ist (Allison 1995, S. 94).

Sollen die Einflüsse der Kovariablen berücksichtigt werden, können analog zur linearen Regression die Residuen der Regressionsgleichungen der Hazard-Modelle für eine weitere grafische Analyse herangezogen werden (Kiefer 1988, 674). Im Unterschied zur linearen Regressionsanalyse kann bei Hazard-Modellen eine Vielzahl

[16] Dabei ist Φ^{-1} die Inverse der Verteilungsfunktion einer normalverteilten Zufallsvariable.

möglicher Residuen zur Untersuchung der Modellannahmen verwendet werden.[17] Am geeignetsten haben sich die sogenannten Cox-Snell Residuen erwiesen (Allison 1995, S. 94). Diese ergeben sich auf Basis des unter Berücksichtigung von Kovariablen geschätzten Modells und sind allgemein definiert als (Cox und Snell 1968 , S. 248 ff.):

$$\hat{r}_i = \hat{H}(t_i|X_i) = -\log\hat{S}(t_i|X_i) \tag{6-1}$$

Für die in dieser Arbeit vorgestellten Verteilungsannahmen lassen sich die formalen Ausdrücke für die Cox-Snell Residuen Tabelle 6-1 ableiten. Dazu müssen nur die entsprechenden Ausdrücke der Survivorfunktionen logarithmiert, um den Vektor X_i der Kovariablen erweitert und als Schätzungen gekennzeichnet werden. Eine entsprechende Übersicht findet sich bei Klein und Moeschberger 1997, S. 394. Bei der Konstruktion eines grafischen Tests kann der Umstand ausgenutzt werden, dass die Cox-Snell Residuen bei Gültigkeit der unterstellten Verteilungsannahme näherungsweise eine Exponentialverteilung mit einem auf den Wert 1 festgelegten Parameter λ haben. Diese Verteilung kann dann analog mittels der für den univariaten Fall beschriebenen grafischen Methode getestet werden. Hier muss nun ein Plot von $-\ln\hat{S}(\hat{r})$ gegen $\hat{r}$, eine bei Gültigkeit der unterstellten Verteilung eine Ursprungsgerade mit einer Steigung von 1 ergeben (Blossfeld, Hamerle und Mayer 1986, S. 85).

Mit Hilfe dieser grafischen Verfahren ist es zum einen möglich, die Richtigkeit der getroffenen Verteilungsannahme zu überprüfen, zum anderen ermöglichen die sich ergebenden Plots u.U. bereits die Identifikation des Modells mit der besten Anpassung an die Daten, sofern mehrere konkurrierende Modelle zur Auswahl stehen, bei denen die zugrunde gelegte Verteilungsannahme auf Basis der Plots nicht verworfen kann. Eine solche Situation ist z.B. beim Vergleich eines Weibull- mit einem Log-Logistischen Regressionsmodell vorstellbar. In diesem Fall wäre eine erkennbar bessere Anpassung der Cox-Snell Residuen ein Hinweis auf die Überlegenheit des einen oder anderen Modells. Allerdings ist eine Auswahlentscheidung allein aufgrund eines grafischen Tests nicht empfehlenswert. Hierzu sind statistische Tests heranzuziehen. Bei der Anwendung statistischer Tests muss zwischen genesteten und nicht-genesteten Modellen unterschieden werden. So konnte in Anschnitt 5.1.1 gezeigt werden, dass sich durch die Restringierung des Shape-Parameters α auf den Wert 1 das Exponential-Modell als Spezialfall des Weibull-Modells ergibt, die

[17] Eine Auswahl möglicher Residuen findet sich bei Klein und Moeschberger 1997, S. 328 ff.

sich wiederum beide als Spezialfall eines Modells, dem die verallgemeinerte Gamma-Verteilung zugrunde liegt, ergeben. Diese Eigenschaft genesteter Modelle macht man sich bei der Konstruktion eines Test zu Nutze. Es werden die doppelten Differenzen der Log-Likelihoods der speziellen und der verallgemeinerten Verteilung gebildet. Mit Hilfe des Likelihood-Ratio-Tests wird dann ein Test der mit der Wahl einer bestimmten Verteilung implizit angenommenen Parameterrestriktionen durchgeführt. Hierbei handelt es sich um eine χ^2-verteilte Teststatistik mit einem bzw. beim Vergleich zwischen der verallgemeinerten Gamma- und der Exponential-Verteilung 2 Freiheitsgraden. Dabei dient die in Abschnitt 5.1.1 vorgestellte Verallgemeinerte Gamma-Verteilung als Grundmodell. Die wesentliche Einschränkung dieses Verfahrens ist, dass nur genestete Modelle miteinander verglichen werden können. Außerdem kann keine Aussage über die Güte des Grundmodells der Verallgemeinerten Gamma-Verteilung gemacht werden (Vgl. für diesen Absatz Allison 1995, S. 88 f.).

Eine Alternative zur Beurteilung der globalen Modellgüte stellen sogenannte Informationskriterien dar. Auch diese basieren auf den Log-Likelihood-Werten der geschätzten Modelle und wägen den mit der Aufnahme weiterer Kovariablen einhergehenden Zuwachs der Modellanpassung im Vergleich zu der Anzahl der zu schätzenden Parameter ab. Sie weisen damit von ihrer Intention her eine Ähnlichkeit zum adjustierten bzw. korrigiertem R^2 der linearen Regression auf (Klein und Moeschberger 1997, S. 254). Informationskriterien C können in allgemeiner Form definiert werden als:

$$C = -2 \cdot \ln L + \tilde{p} \cdot d, \tag{6-2}$$

wobei $\tilde{p}$ die Anzahl der Parameter bezeichnet und d eine Strafkomponente für den mit zunehmender Zahl zu schätzender Parameter zu verbessernden Modellfit (Wedel und Kamakura 1998, S. 90). Wie aus Tabelle 2-1 ersichtlich, werden bei der Beurteilung von Hazard-Modellen im Marketing u.a. das Akaike Informationskriterium (AIC) und das Bayessche Informationskriterium (BIC) häufig verwendet. Diese beiden Kriterien unterscheiden sich in der Operationalisierung der Strafkomponente d. Während das AIC (d=2) eine hohe Anzahl von Parametern nur mit einer geringen Strafe belegt, fällt diese beim BIC (d=ln(I) mit I: Anzahl der Individuen) vor allem bei großen Populationen deutlich größer aus. Das BIC wird als konservativer eingeschätzt, da es sparsame Modelle bevorzugt.[18] Neben den Informationskriterien finden sich in der Literatur bei Balasubramanian und Jain 1994 alternative Ansätze

[18] Eine ausführliche Diskussion der Informationskriterien und ihrer Eignung zur Modellauswahl findet sich bei Wannhoff 1990, S. 29 ff.

zum Vergleich nicht-genesteter Modelle. Diese sollen an dieser Stelle allerdings nicht diskutiert werden, da sie bisher auch in neueren Studien und Lehrbüchern zur Hazard-Analyse keine Anwendung gefunden haben. Eine andere Alternative bzw. Ergänzung zeigt Litfin auf, der in seiner Studie als Modellauswahlkriterien zusätzlich noch die bereits aus der logistischen Regressionsanalyse bekannte Klassifikationsgüte und auch McFaddens Pseudo R^2 heranzieht. Darüber findet noch ein Vergleich der Übernahmeverläufe der einzelnen Regressionsmodelle statt (Litfin 2000, S. 242 ff.). Dazu ist anzumerken, dass McFaddens R^2 nicht geeignet ist, um einen Vergleich alternativer Modelle vorzunehmen, da es sich ja um ein Maß handelt, dass lediglich die Verbesserung der Anpassung gegenüber dem Nullmodell unter Annahme der gleichen Verteilung quantifiziert. Ein Vergleich über verschiedene Verteilungen hinweg wiederum wäre nur dann möglich, wenn die Modelle untereinander genestet wären, so dass letztendlich unterschiedliche Parametrisierungen eines gemeinsamen Grundmodells auf Basis der sich ergebenden Likelihood-Werte miteinander verglichen werden. Sollen nicht miteinander genestete Modelle verglichen werden, wäre als Grundmodell das semiparametrische Cox-Modell vorstellbar, das sich ja durch das Fehlen einer bestimmten Verteilungsannahme hinsichtlich der Verweildauern auszeichnet. Die Verbesserungen der Anpassung der vollparametrisierten parametrischen Regressionsmodelle könnten dann mit einem semiparametrischem Nullmodell verglichen werden, so dass die in Relation dazu erzielten Verbesserungen in der Anpassung als Vergleichsmaßstab der Modelle untereinander verwendet werden können.

6.2 Überprüfung der Modellannahmen im semiparametrischen Modell von Cox

Im semiparametrischen Cox-Modell ist die Proportionalitätsannahme zu überprüfen. Dies kann ebenfalls mittels grafischer und statistischer Tests erfolgen. Ein einfacher grafischer Test beruht auf der Überlegung, dass die Hazard-Raten der betrachteten Teilpopulation sich bei proportionalen Risiken nur um einen konstanten Faktor unterscheiden:

$$\lambda_1(t|x) = \omega\lambda_2(t|x). \tag{6-3}$$

Gemäß Tabelle 3-1 kann diese Beziehung auch mit Hilfe der entsprechenden Survivorfunktionen ausgedrückt werden. Nach zweifacher Logarithmierung ergibt sich:

$$\ln[-\ln S_1(t|x)] = \ln\omega + \ln[-\ln S_2(t|x)]. \tag{6-4}$$

Ein Plot der so transformierten Survivorfunktionen muss bei Gültigkeit der Proportionalitätsannahme zwei parallele Kurven im Abstand von ln ω ergeben (Blossfeld, Ha-

merle und Mayer 1986, S. 139). Eine Verfeinerung dieser Methode stellt der grafische Vergleich der gleitenden Durchschnitte der Hazard-Raten dar, die ebenfalls parallel zueinander verlaufen sollten (Allison 1995, 114).

Eine statistische Überprüfung der Proportionalitätsannahme findet immer im Hinblick auf ein Merkmal bzw. eine Kovariable statt, anhand dessen sich die Grundgesamtheit in nur zwei Teilgesamtheiten aufspalten lässt z.B. „Geschlecht". Bei Vorliegen proportionaler Risiken hinsichtlich dieses Merkmals können die Hazard-Raten geschrieben werden als

(6-5) $$\lambda_1(t|x) = \lambda_0(t) \cdot \exp(x'\beta) \quad \text{bzw.}$$

(6-6) $$\lambda_2(t|x) = \lambda_0(t) \cdot \exp(\alpha + x'\beta),$$

wobei x der Vektor der anderen im Modell enthaltenen Variablen ist. Diese beiden Hazard-Raten können dann mittels einer Dummy-Variable z_1, die für Elemente der zweiten Gruppe den Wert 1 annimmt, zusammengefasst werden zu

(6-7) $$\lambda(t|x, z_1) = \lambda_0(t) \cdot \exp(z_1\alpha_1 + x'\beta).$$

Für die Durchführung des Test wird nun eine zusätzliche Kovariable $z_2 = z_1 * \ln t$ definiert mit dem Parameter α_2 und in Gleichung (6-7) integriert, so dass sich diese ergibt als

(6-8) $$\lambda(t|x, z_1, z_2) = \lambda_0(t) \cdot \exp(z_1\alpha_1 + z_2\alpha_2 + x'\beta).$$

Mit Hilfe der Kovariablen z_2 wird eine Interaktion zwischen dem hinsichtlich der Proportionalitätsannahme zu untersuchenden Merkmal und der Zeit modelliert. Eine Überprüfung der Hypothese H_0: $\alpha_2=0$ ist damit ein Test auf Proportionalität hinsichtlich der ausgewählten Variablen (Allison 1995, S. 157, Blossfeld, Hamerle und Mayer 1986, S. 143). Sofern der Koeffizient α_2 ungleich Null ist, trifft die Annahme proportionaler Risiken nicht mehr zu, da dann der Quotient der Hazard-Raten der Teilpopulationen, wie in Abschnitt 5.1.1 gezeigt wurde, nicht über die Zeit konstant, sondern vielmehr abhängig von der Verweildauer ist. Diese Überprüfung muss für jede interessierende Variable in analoger Form durchgeführt werden.

6.3 Überprüfung der resultierenden Koeffizienten

Neben der Überprüfung der allgemeinen Güte des Modells sind insbesondere Hypothesen-Tests bezüglich der erklärenden Kovariablen für den empirischen Forscher von Interesse. Zur simultanen Prüfung mehrerer Regressionskoeffizienten bezie-

hungsweise Parameter stehen bei den beschriebenen Modellen drei Testverfahren zur Verfügung: der bereits in Abschnitt 6.1 kurz erläuterte Likelihood-Ratio – sowie der Wald- und Score- bzw. Lagrange-Multiplikator-Test (Blossfeld, Hamerle und Mayer 1989, S. 89).[19] Alle drei Teststatistiken sind unter der Nullhypothese, dass alle zu überprüfenden Parameter gleich Null sind, asymptotisch χ^2-verteilt mit r Freiheitsgraden, wobei r die Anzahl der Parameterrestriktionen unter der Nullhypothese H_0 ist. Da die Verteilungseigenschaften der Tests nur asymptotisch gelten, ist bei der Anwendung in der Praxis auf einen ausreichend großen Stichprobenumfang zu achten. Darüber hinaus liefern die drei Verfahren asymptotisch gleiche Ergebnisse, so dass keine konkrete Empfehlung für das eine oder andere Verfahren ausgesprochen werden kann (Kiefer 1988, S. 674). Allison merkt allerdings an, dass es Anzeichen dafür gibt, dass die Likelihood-Ratio-Teststatistik in kleinen Stichproben eine bessere Anpassung an eine χ^2-Verteilung darstellt als die anderen beiden Teststatistiken, so dass eine Reihe von Autoren eine Präferenz für diese Methode hat (Allison 1995, S. 86). Neben der gemeinsamen Prüfung aller oder mehrerer Parameter kann auch eine Prüfung einzelner Parameter der Art H_0: $\beta_i = \xi$ erfolgen. Dies geschieht mit Hilfe der Teststatistik

$$\frac{\hat{\beta}_i - \xi}{\sqrt{\hat{\mathrm{Var}}(\hat{\beta}_i)}}, \tag{6-9}$$

die unter H_0 asymptotisch standardnormalverteilt ist. Setzt man $\xi = 0$, so erhält man einen Test auf Signifikanz einer einzelnen Variablen. Insgesamt weisen damit die hier vorgestellten Testverfahren eine prinzipielle Verwandtschaft bzw. Ähnlichkeit zu denen der linearen Regressionsanalyse auf (Kiefer 1988, S. 674).

7 *Interpretation der Ergebnisse*

Die Interpretation der resultierenden Koeffizienten kann auf verschiedene Arten vorgenommen werden. Da die Regressionsmodelle wie in Abschnitt 0 erläutert als loglineare Modelle konstruiert sind, zeigt ein Regressionskoeffizient ceteris paribus direkt Richtung und Stärke der Änderung der logarithmierten Hazard-Rate bei Änderung der betrachteten unabhängigen Variable um eine Einheit an (Allison 1984, S. 27). Dies ermöglicht zwar eine Aussage über die Richtung des Effektes, im Hinblick auf die Stärke des Effektes ist diese Interpretation wenig intuitiv, da diese sich auf die logarithmierte Hazard-Rate bezieht (Allison 1995, S. 65). Folgerichtig bedient

[19] Für eine ausführliche Darstellung dieser Testverfahren vgl. Johnston und DiNardo 1997, S. 147 ff.

man sich bei der Interpretation der Stärke des Einflusses der Antilogarithmen α der Regressionskoeffizienten $\alpha_i = e^{-\beta_i}$ bzw. $\alpha_i =_i \exp(-\beta_i)$.[20] Wendet man diese Transformation auf das zugrundeliegende log-lineare Regressionsmodell an, lassen sich Richtung und Stärke des Effektes auf die Hazard-Rate bei einer Veränderung der unabhängigen Variable um eine Einheit wie folgt ermitteln:

(7-1) $$\Delta h(t_i|X_i(t)) = (\exp(-\beta_i) - 1) \cdot 100.$$

Hierbei handelt es sich um eine relative Änderung, d.h. die Veränderung der Hazard-Rate wird in Prozent angegeben. Aufgrund der vorgenommen Transformation des ursprünglich log-linearen Modells ist zu beachten, dass bei gleichzeitiger Änderung mehrerer unabhängiger Variablen keine additive sondern eine multiplikative Verknüpfung der Effekte in der Form

(7-2) $$(\exp(-\beta_i) \cdot \exp(-\beta_j) - 1) \cdot 100$$

vorzunehmen ist (Blossfeld, Hamerle und Mayer 1986, S. 148). Ähnlich vorsichtig muss vorgegangen werden, wenn Variablen unterschiedlich skaliert sind, wenn z.B. die Variable „Alter" in Jahren, die Variable „Berufserfahrung" in Monaten gemessen wird. Will man dann die Einflüsse dieser beiden Variablen miteinander vergleichen, muss eine der Skalen transformiert werden, wobei es aufgrund des zugrundeliegenden log-linearen Modells folgender Transformation bedarf, wenn z.B. eine Betrachtung in Jahren statt in Monaten gewünscht wird (Blossfeld, Hamerle und Mayer 1986, S. 148):

(7-3) $$((\exp(-\hat{\beta}_i))^{12} - 1) \cdot 100.$$

Liegen Dummy- bzw. kategoriale Variable vor, ist mittels des Antilogarithmus des Regressionskoeffizienten e^{β_i} bzw. $\exp(\beta_i)$ eine Aussage in Relation zur Gruppe, bei der das betrachte Merkmal nicht aufgetreten ist bzw. zur gewählten Referenzkategorie möglich. Daher wird der Ausdruck e^{β_i} auch als Hazard- bzw. Risk-Ratio bezeichnet und wird insbesondere bei der Interpretation von Proportional Hazard Modellen verwendet. Die Anwendung dieser Interpretationshilfe ist auch nicht auf Dummy- bzw. kategoriale Variable beschränkt. Vielmehr kann durch Differenzbildung zweier logarithmierter Proportional Hazard Modelle, die log-linear in ihren Parametern sind, eine Hazardratio

---

[20] Üblicherweise finden sich in der Literatur Antilogarithmen $\alpha_i =_i \exp(\beta_i)$, die aus einem log-linearen Modell der Form $\lambda = e^{\beta_0 + \beta_1 \cdot x_1}$ resultieren. Hier wurde aber ein Modell $\lambda = e^{-\beta_0 - \beta_1 \cdot x_1}$ angenommen.

(7-4) $$HR(t,a,b,\beta) = \exp\left(h(t,x_i = a,\beta_i) - h(t,x_i = b,\beta_i)\right)$$

erzeugt werden, die sich vereinfacht zu

(7-5) $$\frac{h(t,a,\beta_i)}{h(t,b,\beta_i)} = e^{(a-b)\cdot\beta_i}.$$

Aufgrund der Eigenschaften der Proportional Hazard Modelle, zu denen ja auch das Exponential- bzw. das Weibull- Modell aus den Abschnitten 5.1.1.1 und 5.1.1.2 gehören, ist die Hazardratio unabhängig von der Verweildauer und kann damit direkt zur Interpretation der Einflusse von Kovariablen genutzt werden (Hosmer (Jr.) und Lemeshow 1999, S. 114 f.). Für den Fall, dass a=1 und b=0, also das Vorliegen einer Dummy-Variable, ergibt sich dann wieder e^{β_i}. An dieser Stelle wird auch deutlich, dass es keine prinzipiellen Unterschiede bei der Vorgehensweise der Interpretation von Koeffizienten zwischen parametrischen und dem semiparametrischen Modell von Cox gibt (Allison 1995, S. 117).

8 Zusammenfassung

Das zunehmende Interesse an der Analyse dynamischer Prozesse und die sich ständig verbessernde Verfügbarkeit der zu ihrer Untersuchung notwenigen Datensätze lassen die Frage nach geeigneten Verfahren aufkommen. Traditionelle Methoden wie logistische und lineare Regression sind entweder nicht in der Lage, den Prozesscharakter zu erfassen, oder resultieren in verzerrten Schätzungen. Vor diesem Hintergrund stellt die Hazard-Analyse aufgrund seiner im Großen und Ganzen recht einfachen analytischen Behandlung, der intuitiv eingängigen Idee des Verfahrens, der guten Interpretierbarkeit der Ergebnisse sowie der Anpassungsfähigkeit an verschiedenste Anforderungen, die aus dem jeweiligen Untersuchungsdesign resultieren können, eine beachtenswerte Alternative dar, die sich zudem großer Beliebtheit in anderen Wissenschaftsbereichen erfreut. Folglich soll dieser Beitrag als Anregung und erste Anleitung dienen, sich nach Maßgabe der Forschungsfrage und Datensituation dieses Instruments als ergänzende oder alternative Analysemethode zu bedienen.

9 Literaturverzeichnis

Ailwadi, K.L., K. Gedenk und S.A. Neslin (1999): Heterogeneity and purchase event feedback in choice models: An empirical analysis with implications for model building, *International Journal of Research in Marketing*, 16, 177-198.

Allison, P.D. (1984): *Event History Analysis: Regression for Longitudinal Data*, Newbury Park, London, New Delhi.

Allison, P.D. (1995): *Survival Analysis Using the SAS System: A Practical Guide*, Cary, NC.

Andreß, H.-J. (1985): *Multivariate Analyse von Verlaufsdaten: Statistische Grundlagen und Anwendungsbeispiele für die dynamische Analyse nicht-metrischer Merkmale*, Mannheim.

Arminger (1988): Modelle zur Analyse qualitativer Variablen in stetigem Zeitverlauf, in: Meier, F. (Hrsg.), *Prozeßforschung in den Sozialwissenschaften*, Stuttgart, New York, 77-91.

Bähr-Seppelfricke, U. (1999): *Der Einfluß von Produkteigenschaften auf die Diffusion neuer Produkte*, Wiesabden.

Balasubramanian, S.K. und D.C. Jain (1994): Simple approaches to evaluate competing non-nested models in marketing, *International Journal of Research in Marketing*, 11, 53-72.

Bass, F., M. (1969): A New Product Growth Model for Consumer Durables, *Management Science*, 15, 215-227.

Bearse, P., J. Canals und P. Rilstone (1998): Consistent standard errors for semiparametric duration models with unobserved heterogeneity, *Economics Letters*, 59, 153-156.

Blossfeld, H.-P. und A. Hamerle (1992): Unobserved Heterogeneity in event history models, *Quality & Quantity*, 26, 157-168.

Blossfeld, H.-P., A. Hamerle und K.U. Mayer (1986): *Ereignisanalyse: Statistische Theorie und Anwendungen in den Wirtschafts- und Sozialwissenschaften*, Frankfurt am Main, New York.

Blossfeld, H.-P., A. Hamerle und K.U. Mayer (1989): Hazard-Raten Modelle in den Wirtschafts und Sozialwissenschaften, *Allgemeines Statistisches Archiv*, 73, 213-238.

Blossfeld, H.-P. und G. Rohwer (1995): *Techniques of Event History Modeling: New Approaches to Causal Analysis*, Mahwah, New Jersey.

Bolton, R.N. (1998): A Dynamic Model of the Duration of the Customer´s Relationship with a Continuous Service Provider: The Role of Satisfaction, *Marketing Science*, 17, 45-65.

Chamberlain, G. (1985): Heterogeneity, omitted variable bias, and duration dependence, in: Heckmann, J. J. und B. Singer (Hrsg.), *Longitudinal analysis of labor market data*, Cambridge et al, 3-38.

Chandrashekaran, M. und R.K. Sinha (1995): Isolating the Determinants of Innovativeness: A Split-Population Tobit (SPOT) Duration Model of Timing and Volume of First and Repeat Purchase, *Journal of Marketing Research*, 32, 444-456.

Chintagunta, P. und S. Haldar (1998): Investigating Purchase Timing Behavior in Two Related Product Categories, *Journal of Marketing Research*, 35, 43 - 53.

Chintagunta, P.K.P., Alok R. (1998): An Empirical Investigation of the "Dynamic McFadden" Model of Purchase Timing and Brand Choice: Implications for the Market Structure, *Journal of Business & Economic Statistics*, 16, 2-12.

Christensen, B. (1999): Determinanten der beruflichen Mobilität - Eine ökonomische Untersuchung auf Basis des Sozioökonomischen Panels, *Institut für Statistik und Ökonometrie, Kiel*, Christian-Albrechts-Universität zu Kiel, 177.

Cox, D.R. (1972): Regression Models and Life-Tables, *Journal of the Royal Statistical Society*, 34, 187-220.

Cox, D.R. und E.J. Snell (1968): A General Definition of Residuals, *Journal of the Royal Statistical Society*, B 30, 248-275.

Darden, W.R., R.D. Hampton und E.W. Boatwright (1987): Investigating Retail Employee Turnover: An Application of Survival Analysis, *Journal of Retailing*, 63, 69-88.

Diekmann, A. und P. Mitter (1984): *Methoden zur Analyse von Zeitverläufen: Anwendungen stochastischer Prozesse bei der Untersuchung von Erreignisdaten*, Stuttgart.

Gail, M.H., S. Wieand und S. Piantadosi (1984): Biased estimates of treatment effect in randomized experiments with nonlinear regressions and omitted covariates, *Biometrika*, 71, 431-444.

Galler, H.P. (1985): *Übergangsratenmodelle bei intervalldatierten Ereignissen, Working Paper, Sonderforschungsbereich 3: Mikroanalytische Grundlagen der Gesellschaftspolitik, Vol. 164, S. 1-32,*

Gönül, F. und K. Srinivasan (1993): Consumer Purchase Behavior in a Frequently Bought Product Category: Estimating Issues and Managerial Insights from a Hazard-Function Model with Heterogeneity, *Journal of the American Statistical Association*, 88, 1219-1227.

Greene, W.H. (1998): *LIMDEP, Version 7.0: User's Manual (Revised Edition)*, Plainview (NY), Castle Hill (NSW).

Greene, W.H. (2000): *Econometric Analysis, Fourth Edition*, Englewood Cliffs, New Jersey.

Gupta, S. (1991): Stochastic Models of Interpurchase Time With Time-Dependent Covariates, *Journal of Marketing Research*, 28, 1-15.

Hamerle, A. (1987): Der "Event-History-Ansatz" zur Modellierung von Diffusions- und allgemeinen Kaufentscheidungsprozessen, *Marketing ZFP*, 10, 248-256.

Hamerle, A. und G. Tutz (1989): *Diskrete Modelle zur Analyse von Verweildauern und Lebenszeiten*, Frankfurt am Main; New York.

Hansen, G. (1991): Neuere Entwicklungen auf dem Gebiet der Ökonometrie, *Zeitschrift für Sozialwissenschaften (ZWS)*, 111, 337-399.

Heckmann, J. und B. Singer (1984a): Econometric Duration Analysis, *Journal of Econometrics*, 24, 63-132.

Heckmann, J. und B. Singer (1984b): A Method For Minimizing The Impact Of Distributional Assumptions In Econometric Models For Duration Data, *Econometrica*, 52, 271-320.

Heckmann, J.J. und C.R. Taber (1994): Econometric Mixture Models and More General Models for Unobservables in Duration Analysis, Cambridge, MA, National Bureau of Economic Research, Inc., 1-34.

Helsen, K. und D.C. Schmittlein (1993): Analyzing Duration Times in Marketing: Evidence for the Effectiveness of Hazard Rate Models, *Marketing Science*, 11, 395-414.

Horowitz, J.L. (1999): Semiparametric Estimation of a Proportional Hazard Model with unobserved Heterogeneity, *Econometrica*, 67, 1001-1028.

Hosmer (Jr.), D.W. und S. Lemeshow (1999): *Applied Survival Analysis*, New York et al.

Hoverstad, R., W.C.I. Moncrief und G.H.j. Lucas (1990): The Use of Survival Analysis to Examine Sales Force Turnover of Part-Time and Full-Time Sales Employees, *International Journal of Research in Marketing*, 7, 109-119.

Hruschka, H., H. Stoiber und A. Hamerle (1998): Analyzing purchase incidence and brand choice by hazard models, *OR Spektrum*, 20, 55-63.

Hutchinson, D. (1988): Event History and survival analysis in the social sciences, *Quality & Quantity*, 22, 203-229.

Jain, D.C.V., Naufel J. (1991): Investigating Household Purchase Timing Decisions: A Conditional Hazard Function Approach, *Marketing Sciences*, 10, 1-23.

Johnston, J. und J. DiNardo (1997): *Econometric Methods*, New York et al.

Kalbfleisch, J.D. und R.L. Prentice (1980): *The Statistical Analysis of Failure Time Data*, New York.

Kiefer, N.M. (1988): Economic Duration Data and Hazard Functions, *Journal of Economic Literature*, 26, 646-679.

Klein, J.P. und M.L. Moeschberger (1997): *Survival Analysis: Techniques for Censored and Truncated Data*, New York et al.

Kleinbaum, D.G. (1995): *Survival Analysis: A Self-Learning Text*, New York et al.

Koop, G. und C.J. Ruhm (1993): Econometric Estimation of Proportional Hazard Models, *Journal of Economics and Business*, 45, 421-430.

Li, S. (1995): Survival Analysis, *Marketing Research*, 7, 17-23.

Litfin, T. (2000): *Adoptionsfaktoren - Eine empirische Analyse am Beispiel eines innovativen Telekommunikationsdienstes*, Wiesbaden.

Mahajan, V., E. Muller und F.M. Bass (1990): New Product Diffusion Models in Marketing: A Review and Directions for Research, *Journal of Marketing*, 54, 1-26.

Mahajan, V., E. Muller und R.K. Srivastava (1990): Determination of Adopter Categories by Using Innovation Diffusion Models, *Journal of Marketing Research*, 27, 37-50.

Maller, R. und X. Zhou (1996): *Survival Analysis with Long-Term Survivors*, Chichester et al.

Mantel, N. und M. Myers (1971): Problems of Convergence of Maximum Likelihood Interative Procedures in Multiparameter Situations, *Journal of the American Statistical Association*, 66, 484-481.

Moncrief III, W.C., R. Hoverstad und G.H. Lucas Jr. (1989): Survival Analysis: A New Approach to Analyzing Sales Force Retention, *Journal of Personal Selling and Sales Management*, 9, 19-30.

Mood, A.M., F.A. Graybill und D.C. Boes (1974): *Introduction to the Theory of Statistics*, Auckland et al.

Newman, J.L. und C.E. McCulloch (1984): A Hazard Approach to the Timing of Births, *Econometrica*, 52, 939-961.

Peterson, T. (1986a): Fitting Parametric Survival Models with Time-Dependent Covariates, *Applied Statistics*, 35, 281-288.

Peterson, T. (1991): The Statistical Analysis of Event Histories, *Sociological Methods & Research*, 19, 270-323.

Rangaswamy, A. und S. Gupta (1999): Innovation Adoption And Diffusion In The Digital Environment: Some Research Opportunities, Pennsylvania, eBusiness Reserach Center - Pennsylvania State University, 1-35.

Ronning, G. (1991): *Mikroökonometrie*, Berlin et al.

Schmidt, P. und A.D. Witte (1989): Predicting Criminal Recidivism Using 'Split Population' Survival Time Models, *Journal of Econometrics*, 40, 141-159.

Sinha, R.K. und M. Chandrashekaran (1992): A Split Hazard Model for Analyzing the Diffusion of Innovations, *Journal of Marketing Research*, 29, 116-127.

Trussell, J. und T. Richards (1985): Correcting for unmeasured Heterogeneity in Hazard Models using the Heckman-Singer procedure, *Sociological Methodology*, 15, 242-276.

Vaupel, J.W. und A.I. Yashin (1985): Heterogeneity´s Ruses: Some Surprising Effects of Selection on Population Dynamics, *The American Statistican*, 39, 176--185.

Vilcassim, N.J. und D.C. Jain (1991): Modeling Purchase-Timing and Brand-Switching Behaviour Incorporating Explanatory Variables and Unobserved Heterogeneity, *Journal of Marketing Research*, 28, 29-41.

Wangler, A. (1997): *Heterogenitätsprobleme in der Verlaufsdatenanalyse*, Frankfurt am Main et al.

Wannhoff, J. (1990): *Zur Analyse von Mischverteilungen auf der Basis von Informationskriterien*, Bergisch Gladbach.

Wedel, M. und W.A. Kamakura (1998): *Market Segmentation: Conceptual and Methodological Foundations*, Boston, Dordrecht, London.

Yamaguchi, K. (1993): Modelling Time-Varying Effects of Covariates in Event-History Analysis using Statistics from the Saturated Hazard Rate Model, *Sociological Methodology*, 23, 279-317.