

# Inferenz für geometrisch verteilte Verweildauern bei linkstrunkier- tem und rechtszensiertem Beob- achtungsdesign: Das AFiD-Panel

Dissertation zur  
Erlangung des akademischen Grades  
doctor rerum politicarum (Dr. rer. pol.)

der Wirtschafts- und Sozialwissenschaftlichen Fakultät  
der Universität Rostock

Name: Eric Scholz  
Matrikelnummer: 214203863  
Abgabedatum: 20.11.2024

Betreuer: Prof. Dr. Rafael Weißbach  
Universität Rostock  
Wirtschafts- und Sozialwissenschaftliche Fa-  
kultät

[https://doi.org/10.18453/rosdok\\_id00005072](https://doi.org/10.18453/rosdok_id00005072)

Dissertation am Institut für Volkswirtschaftslehre  
Wirtschafts- und Sozialwissenschaftliche Fakultät

---

Gutachter\_in 1: Vorname: Rafael  
Nachname: Weißbach  
Einrichtung: Universität Rostock

Gutachter\_in 2: Vorname: Thomas  
Nachname: Kneib  
Einrichtung: Georg-August-Universität Göttingen

Jahr der Einreichung: 2024  
Jahr der Verteidigung: 2025

---

## Lebenslauf

### Persönliche Daten:

Geburtstag: 30.11.1995  
Geburtsort: Hamburg  
Familienstand: ledig

### Berufserfahrung:

2021-2025: **Wissenschaftlicher Mitarbeiter**, *Universität Rostock, Wirtschafts- und Sozialwissenschaftliche Fakultät, Lehrstuhl für Statistik und Ökonometrie.*  
Forschung im Bereich Verweildaueranalyse;  
Durchführung von Übungen für die Fächer 'Grundlagen der Statistik', 'Statistische Modelle', 'Finanzstatistik' sowie 'Computergestützte Datenanalyse';  
Betreuung von Bachelor- und Masterarbeiten

### Studium:

2018-2021: **Master Studium**, *Universität Rostock,*  
Studium der Mathematik mit Zweitfach Physik.  
Abschlussnote: 1,6  
Masterarbeit: Intraklassenkorrelation zweiter Ordnung - Schätzung und Anwendung

2014-2018: **Bachelor Studium**, *Universität Rostock,*  
Studium der Mathematik mit Zweitfach Physik.  
Abschlussnote: 2,2  
Bachelorarbeit: Rekurrenzfreie Sequenz Labeling Ansätze

### Besondere Tätigkeiten:

2016 und 2018: **Tätigkeit als wissenschaftliche Hilfskraft**, *Universität Rostock*  
2020/2021: **Praktikum**, *Itzehoer Versicherung/Brandgilde von 1691 Versicherungsverein auf Gegenseitigkeit, Itzehoe.*

### Schulausbildung:

2006-2024: **Gymnasium**, *Sophie-Scholl-Gymnasium, Itzehoe,*  
Naturwissenschaftliches Profil.  
Abschlussnote: 2,7

2002-2006: **Grundschule**, *Grund- und Gemeinschaftsschule, Schenefeld.*

### Besondere Tätigkeiten:

2010-2011: **Propädeutikum**, *Young Business School, Heidelberg.*  
Förderprogramm für junge Menschen im Bereich Mathematik.

2011: **Praktikum**, *oPen Software GmbH, Elmshorn.*

2013: **Praktikum**, *inray Industriesoftware GmbH, Schenefeld.*

# Danksagung

An dieser Stelle möchte ich mich zunächst bei den Menschen bedanken, die mich in den letzten Jahren unterstützt und damit dazu beigetragen haben, dass meine Dissertation nun in dieser Form vorliegt.

Mein besonderer Dank gilt Professor Rafael Weißbach, der durch seinen hohen Einsatz bei der Betreuung eine sehr große Hilfe war. Er stand als Ansprechpartner immer zur Verfügung und hat sich viel Zeit für Diskussionen, anregende Gespräche und konstruktives sowie detailliertes Feedback genommen.

Außerdem möchte ich mich bei Anne-Marie Toparkus bedanken, die mich in den letzten Jahren mit vielen fachlichen und nicht-fachlichen Gesprächen sowie konstruktivem Feedback unterstützt hat.

Darüber hinaus danke ich Herrn Professor Dr. Thomas Kneib, der sich bereit erklärt hat für diese Arbeit als Zweitgutachter zu fungieren.

Ich danke außerdem Dr. Florian Köhler vom Landesamt für Statistik Niedersachsen für die Bereitstellung der Daten.

Ebenfalls bedanke ich mich bei Kristina Pieth, die in den letzten Jahren bei allen nicht fachlichen Anliegen stets ein offenes Ohr hatte und sich immer wieder hilfsbereit gezeigt hat.

Auch möchte ich mich bei all den anderen Kollegen bedanken, die durch Gespräche aller Art maßgeblich zu einem positiven Arbeitsalltag beigetragen haben: Florian, Lena, Nico, Philipp Budde und Philipp Lorscheider.

Meinen Eltern Uta und Mario sowie meiner Oma Gerda möchte ich für all den Rückhalt in den letzten Jahren danken. Ihr habt mich immer unterstützt und wart bei Problemen jeglicher Art immer für mich da. Danke für alles.

Abschließend sei all jenen Freunden gedankt, die in den letzten drei Jahren mit Gesprächen fern ab von fachlichen Belangen für reichlich Ablenkung gesorgt haben: Boffi, Bupka, Eike, Kelvin, Leon, Lukas, Marlon, Mattes, Sarah, Sebastian, Wiebke - danke für die schöne Zeit.

# Inhaltsverzeichnis

<b>1. Einleitung</b>	<b>1</b>
<b>2. Das Modell</b>	<b>4</b>
2.1. Die Lebensdauer als stochastischer Prozess . . . . .	4
2.2. Linkstrunkierung . . . . .	17
2.2.1. Die latente Filtration . . . . .	19
2.2.2. Linkstrunkierter Zählprozess und Kompensator . . . . .	26
2.2.3. Anpassung der Filtration . . . . .	32
2.2.4. Der Beitrag zur auf $\mathbb{1}_{\tilde{\mathbb{T}}}$ bedingten Likelihood-Funktion . . . . .	35
2.3. Rechtszensierte Historienteile . . . . .	39
<b>3. Inferenz für den Parameter der geometrischen Verteilung</b>	<b>52</b>
3.1. Der Punktschätzer . . . . .	52
3.2. Standardfehler im parametrischen Modell . . . . .	56
3.2.1. Konsistenz . . . . .	59
3.2.2. Asymptotische Normalität . . . . .	67
3.3. Standardfehler im semiparametrischen Modell . . . . .	74
3.3.1. Der vorhersagbare Variationsprozess . . . . .	79
3.3.2. Konsistenz und Asymptotische Normalität . . . . .	85
<b>4. Schätzung bei doppelt trunkiertem Design</b>	<b>88</b>
4.1. Das beobachtbare Modell . . . . .	89
4.2. Likelihood-Approximation . . . . .	92
<b>5. Das AFiD-Panel</b>	<b>100</b>
5.1. Beschreibung der Daten . . . . .	100
5.2. Punkt- und Intervallschätzung . . . . .	102
5.3. Vergleich mit anderen Quellen . . . . .	106
<b>6. Simulation</b>	<b>107</b>
<b>7. Fazit</b>	<b>112</b>
<b>Literatur</b>	<b>114</b>
<b>A. Appendix</b>	<b>117</b>
A.1. Elementares Beispiel zur Anwendung der Doob-Zerlegung . . . . .	117
A.2. Berechnung des Standardfehlers per Martingalgrenzwertsatz im parametrischen Modell . . . . .	121

A.3. Wichtige Definitionen, Eigenschaften und Sätze . . . . .	123
A.3.1. Bedingte Erwartungswerte . . . . .	123
A.3.2. Stochastische Konvergenzbegriffe . . . . .	124
A.3.3. Wichtige Sätze . . . . .	125
A.4. R-Code . . . . .	126
A.4.1. Analyse des AFiD-Panels . . . . .	126
A.4.2. Code zur Simulation . . . . .	130

# 1. Einleitung

Die vorliegende Arbeit behandelt die Schätzung der Wahrscheinlichkeit sowie die Berechnung des zugehörigen Standardfehlers in einem Bernoulli-Experiment. Allgemein werden Individuen untersucht, die innerhalb eines Jahres entstehen und innerhalb eines späteren Jahres wieder ausfallen. Das Experiment besteht darin zu bestimmen, ob ein Individuum im Laufe des nächsten Jahres ausfallen wird. Solche Überlebensexperimente werden dabei für mehrere Individuen und, für jedes einzelne Individuum, über mehrere, aufeinanderfolgende Jahre analysiert.

Zu diesem Zweck wird ein Überlebenszeitmodell entwickelt. Dieses wird, um im Einklang mit den Statistiken, die das Überleben eines Individuums jährlich dokumentieren, zu bleiben, das Überleben eines Individuums ebenso nicht kontinuierlich, sondern zeitdiskret in Form von Spielrunden messen. In der Literatur werden diese, unter anderem von Feller [1968], als Epochen bezeichnet. Damit eine begriffliche Übereinstimmung zwischen Theorie und Anwendung zumindest teilweise gewährleistet ist, werden die Spielrunden im Rahmen dieser Arbeit jedoch weiterhin Jahre genannt.

Angenommen, es wird eine Stichprobe bestehend aus  $n$  Individuen betrachtet, so bestünde ein intuitiver Ansatz zur Schätzung der Ausfallwahrscheinlichkeit darin, für das entsprechende Jahr die Anzahl der negativen Ereignisse, also der Ausfälle, in Relation zur Anzahl der Experimente zu setzen. Eben dieses Verhältnis soll im Rahmen dieser Arbeit ausführlich untersucht werden. Dabei wird angenommen, dass sich die Ausfallwahrscheinlichkeit im Laufe der Jahre nicht ändert. Die Lebensdauer eines Individuums unterliegt somit einer geometrischen Verteilung. Auch werden die Ausfälle der Individuen als paarweise unabhängig voneinander angenommen.

Zusätzlich gibt es für die Individuen einige Einschränkungen hinsichtlich der Beobachtbarkeit, da die Beobachtung der Individuen lediglich in einem fest definierten Zeitraum erfolgt.

Auf der einen Seite werden Individuen aus der Stichprobe vom Umfang  $n$ , deren Schließung bereits vor dem Beginn des Beobachtungszeitraums stattgefunden hat, gar nicht beobachtet und nicht einmal deren Existenz zu Kenntnis genommen. Folglich unterscheidet sich der Stichprobenumfang  $n$  von der Anzahl der Beobachtungen. Letztere ist dabei zufällig mit der Realisierung  $m$ . Literatur zu dieser Form der retrospektiven Stichprobenziehung und dem damit einhergehenden Informationsverlust wird unter dem Begriff 'Trunkierung' veröffentlicht. Da jedoch nur Individuen, deren Ausfall vor dem Beobachtungszeitraum stattgefunden hat, nicht beobachtet werden, wird im Rahmen dieser Arbeit von einer linksseitigen Trunkierung beziehungsweise Linkstrunkierung gesprochen.

Auf der anderen Seite unterliegen die Historien der Individuen, die den Beobachtungszeitraum überleben, einer Rechtszensur. Dementsprechend wird für solche Individuen die Existenz sowie das Überleben bis zum Ende des Beobachtungszeitraums registriert,

das exakte Ausfallalter allerdings nicht.

Das primäre Ziel dieser Arbeit ist die Schätzung des geometrischen Parameters sowie die Berechnung des Standardfehlers unter Berücksichtigung des linkstrunkierten und rechtszensierten Beobachtungsdesigns. Im Rahmen dessen resultieren aus dem fest definierten Beobachtungszeitraum zudem zwei Fragestellungen.

Die erste Frage ergibt sich daraus, dass aufgrund des Designs jedes Individuum erst ab einem gewissen Alter beobachtet werden kann. Das Überleben der Individuen bis zu diesem Alter wird dabei für die Schätzung des geometrischen Parameters nicht berücksichtigt. Daraus resultiert die Fragestellung, inwiefern das Vergessen der positiven Ereignisse vor dem Beginn des Beobachtungszeitraums Auswirkungen auf den Schätzer der Ausfallwahrscheinlichkeit hat.

Die zweite Frage resultiert daraus, dass nicht alle Individuen im selben Jahr entstehen und folglich das Alter zum Beginn des Beobachtungszeitraums nicht für jedes Individuum gleich ist. Es werden somit nicht für jedes Individuum identisch viele Überlebensexperimente gezählt. Im späteren Verlauf der Arbeit wird diesbezüglich untersucht, wie sich trotz der daraus resultierenden unterschiedlichen Varianzen ein asymptotischer Standardfehler ableiten lässt.

Die gezählte Anzahl der Jahre seit der Entstehung wird dabei im Rahmen dieser Arbeit als 'Alter' bezeichnet. Die Entstehung ist dabei im Überlebenszeitmodell gleichbedeutend mit der Geburt, da eine Messung am Individuum vor diesem Ereignis nicht definiert ist. Auch ist der Ausfall gleichbedeutend mit dem Tod, da nach diesem Ereignis keine Messung am Individuum mehr stattfindet und folglich dessen Registrierung beendet wird.

Zunächst wird in Kapitel 2.1 der Likelihood-Anteil eines Individuums zur Schätzung der Ausfallwahrscheinlichkeit hergeleitet, ohne dass Linkstrunkierung und Rechtszensur berücksichtigt werden. Dieses Modell wird dann sukzessiv in Kapitel 2.2 um die Linkstrunkierung und in Kapitel 2.3 um die Rechtszensur erweitert. Anschließend wird in Kapitel 3.1 der Likelihood-Anteil eines Individuums unter Berücksichtigung von Linkstrunkierung und Rechtszensur auf eine Stichprobe vom Umfang  $n$  ausgeweitet und ein Punktschätzer für den geometrischen Parameter ermittelt. Die Berechnung der entsprechenden Standardfehler erfolgt dann zum einen mithilfe der M-Schätzer-Theorie über Konsistenz und asymptotische Normalität in Kapitel 3.2 und zum anderen per Martingalgrenzwertsatz in Kapitel 3.3. Ergänzend wird noch ein weiterer Ansatz zur Konstruktion eines Punktschätzers der Ausfallwahrscheinlichkeit bei einem doppelt trunkierten Beobachtungsdesign in Kapitel 4 skizziert.

Eine anschließende konkrete Schätzung der Ausfallwahrscheinlichkeit sowie die Berechnung der Standardfehler findet im Rahmen dieser Arbeit auf der Grundlage des AFiD-Panels statt. Dieses umfasst deutsche Unternehmen, die in den Jahren 2013 bis 2017 gegründet wurden beziehungsweise entstanden sind. Für diese konnte ein potentieller Ausfall respektive Schließung in den Jahren 2018 und 2019 beobachtet werden. Das AFiD-Panel, dessen Zugang vom Forschungsdatenzentrum des statistischen Bundesamts für unter anderem diese Arbeit zur Verfügung gestellt wurde, wird in Kapitel 5 genauer erläutert.

Ergänzend wird in Kapitel 6 zur Evaluierung der Methoden eine Schätzung des geometrischen Parameters sowie die Berechnung der Standardfehler anhand von am Com-

puter generierten Daten simuliert. Abschließend wird in Kapitel 7 ein Fazit gezogen.

## Literaturüberblick

In Bezug auf das Anwendungsgebiet findet sich in der Literatur beispielsweise in Cochran [1981] eine ältere Referenz zu Statistiken über das Überleben von Unternehmen. Untersuchungen bezüglich der Lebenserwartung neuerer Unternehmen, jedoch ohne retrospektive Paneldaten, wurden von Mata und Portugal [1994] für Portugal, von Audretsch und Mahmood [1995] für die USA, von Pittiglio [2023] für Italien und von Honjo [2000] beziehungsweise Kato et al. [2022] für Japan durchgeführt. Ein früher retrospektiver Versuch stammt hingegen von Brüderl et al. [1992], die die Unternehmenslebensdauern von den jeweiligen Gründern erfragt und dabei vermehrt die Gründer von bereits geschlossenen Unternehmen kontaktiert haben. Auch wurde die Lebenserwartung auf der Grundlage portugiesischer Daten von Reis und Augusto [2015] untersucht, wohingegen eine Analyse der Demografie deutscher Unternehmen von Rink und Seiwert [2021] betrieben wurde.

Hinsichtlich methodischer Ähnlichkeiten wurden Ergebnisse zur beidseitigen Trunkierung, die auch als doppelte Trunkierung bezeichnet wird, mit der Anwendung auf Unternehmenslebensdaten von de Uña-Álvarez et al. [2024] für Spanien und von Toparkus und Weißbach [2025] für Deutschland veröffentlicht. Die Entwicklung eines Regressionsmodells für rechtstrunkierte und zeitdiskrete Überlebensdaten wurde von Gross und Huber-Carol [1992] durchgeführt. Die Effekte der linksseitigen Trunkierung auf die Untersuchung soziologischer Fragestellungen werden beispielsweise in Guo [1993] am Beispiel von Ehescheidungen in den Vereinigten Staaten diskutiert.

Das in dieser Arbeit entwickelte Modell knüpft an die zeitstetige biometrische Studie im Anhang A von Weißbach et al. [2024] an und kann als zeitdiskrete Version des Modells, das für die Studie 'Mortality of Diabetics in the County of Fyn' in Andersen et al. [1993, Beispiel III.3.6] verwendet wurde, interpretiert werden. Ebenfalls wird das Phänomen der linksseitigen Trunkierung als typischer Designfehler in Krankheitsmodellen angesehen [siehe beispielsweise Putter et al., 2006]. Das Ignorieren der Trunkierung führt dabei zum so genannten 'immortal time bias' [siehe beispielsweise Hernán et al., 2016 sowie Yadav und Lewis, 2021].

## 2. Das Modell

### 2.1. Die Lebensdauer als stochastischer Prozess

Bevor im späteren Verlauf der Arbeit die erwartete Lebensdauer deutscher Unternehmen geschätzt wird, erfolgt zunächst eine allgemeine Herleitung des verwendeten mathematischen Modells. Auf der sprachlichen Ebene werden dabei Individuen betrachtet, die eine Entstehung und einen Ausfall haben. In der späteren Anwendung handelt es sich bei den Individuen um Unternehmen mit einer Gründung (Entstehung) und einer Schließung (Ausfall). Die im späteren Verlauf untersuchten Daten beinhalten für die Unternehmen jeweils das Jahr der Gründung sowie das Jahr der Schließung. Da keine feinere zeitliche Untergliederung innerhalb der Daten vorliegt, ist es sinnvoll, ein zeitdiskretes mathematisches Modell zur Untersuchung dieser herzuleiten. Die Zeitspanne zwischen Entstehung und Ausfall wird dabei allgemein als Lebensdauer bezeichnet und aufgrund der Datenlage ebenfalls ganzzahlig in Jahren angegeben. Die Bezeichnung der Epochen als Jahre wird zwar im Rahmen dieser Arbeit beibehalten, jedoch können analog auch andere Zeiteinheiten wie Sekunden, Stunden oder Quartale verwendet werden.

Die Erstellung des Modells selbst erfolgt dabei in drei Schritten. In diesem Kapitel werden zunächst grundlegende Konzepte vorgestellt und ein Modell ohne Restriktionen bezüglich der Beobachtbarkeit beschrieben. Der Beobachtungszeitraum hat somit weder ein Anfang noch ein Ende, sodass jedes Individuum inklusive Lebensdauer unabhängig vom Jahr des Ausfalls beobachtet werden kann. In den folgenden Abschnitten wird dieses Modell sukzessiv um einen Beobachtungsbeginn sowie die damit verbundene linksseitige Trunkierung (Kapitel 2.2) und anschließend um ein Ende des Beobachtungszeitraums sowie die damit einhergehende rechtsseitige Zensur (Kapitel 2.3) ergänzt.

Als Grundgesamtheit dienen dabei alle Individuen, deren Entstehung frühestens im Jahr  $a$  und spätestens im Jahr  $b$  erfolgt ist, sodass der Entstehungszeitraum insgesamt  $G := |\{a, a + 1, \dots, b\}|$  Jahre umfasst. Die gesamte Modellentwicklung wird dabei zunächst für lediglich ein, aus der Grundgesamtheit zufällig entnommenes, Individuum erfolgen. Ein Übergang zu einer Stichprobe erfolgt zu einem späteren Zeitpunkt für das vollständige Modell in Kapitel 3.1.

Erfolgt die zufällige Ziehung eines Individuums aus der Grundgesamtheit, so werden die Elemente der Menge aller möglichen Ergebnisse  $\Omega := \{\omega_1, \omega_2, \dots\}$  folgendermaßen definiert:

$$\begin{aligned} \omega_{(x-1)G+t+1} := & \{ \text{Individuum hat eine Lebensdauer von } x \text{ Jahren} \\ & \text{und ist im letzten Jahr des Entstehungszeitraums} \\ & t \text{ Jahre alt} \}, \quad x \in \mathbb{N} \text{ und } t \in \{0, \dots, G - 1\} \end{aligned} \quad (2.1)$$

Diese Darstellung der Elemente der Ergebnismenge hat den Vorteil, dass die Elemente eigentlich von zwei Komponenten abhängen, allerdings nur mithilfe eines Indexes durchgezählt werden können.

Da die Komponente  $t$  auch das Alter des Individuums im Jahr  $b$  des Entstehungszeitraums angibt, kann mithilfe der  $t$ -Komponente das Entstehungsjahr des Individuums durch  $b - t$  ermittelt werden. Außerdem ergibt sich unter der Verwendung der  $x$ -Komponente das Jahr des Ausfalls für ein Individuum durch  $b - t + x$ .

Zur Visualisierung einer beispielhaften Ergebnismenge für einen Entstehungszeitraum, der aus  $G = 5$  Jahren besteht, dient Abbildung 2.1.

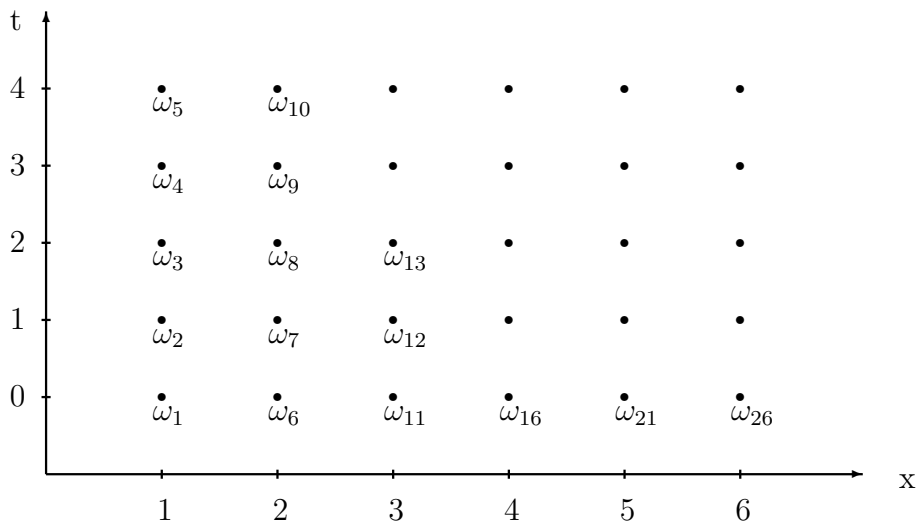


Abbildung 2.1.: Zuordnung der Elementarereignisse zu einem Entstehungs- und Ausfalljahr bei einem 5-jährigen Entstehungszeitraum. Beispielsweise gilt  $\omega_{12} = \omega_{(3-1) \cdot 5 + (1+1)}$ , sodass  $\omega_{12}$  bedeutet, dass das Individuum im letzten Jahr des Entstehungszeitraums ein Jahr alt war und insgesamt 3 Jahre gelebt hat. Das Individuum ist dann im Jahr  $b - 1$  entstanden und im Jahr  $b + 2$  ausgefallen.

Des Weiteren wird die Menge der Ereignisse mithilfe der  $\sigma$ -Algebra  $\mathcal{F} = \mathcal{P}(\Omega)$  dargestellt. Zusammen mit dem Wahrscheinlichkeitsmaß  $P_\theta$ , das von einem Parameter  $\theta \in \Theta$  abhängt und in den folgenden Abschnitten weiter spezifiziert wird, ergibt sich der zugrundeliegende Wahrscheinlichkeitsraum  $(\Omega, \mathcal{F}, P_\theta)$ .

Darüber hinaus wird die Lebensdauer eines Individuums in Jahren durch die univariate Zufallsvariable  $X$  mit

$$X : \Omega \rightarrow \mathbb{N} : \omega_{(x-1)G+t+1} \mapsto x \quad (2.2)$$

beschrieben, wobei  $x$  allgemein das  $x$ -te Jahr bezeichnet und  $\mathbf{x}$  für die konkrete Realisierung der Zufallsvariablen  $X$  steht.

Analog dazu lässt sich das Alter im letzten Jahr  $b$  des Entstehungszeitraums mithilfe

der univariaten Zufallsvariablen

$$T : \Omega \rightarrow \mathbb{N}_0 : \omega_{(x-1)G+t+1} \mapsto t \quad (2.3)$$

definieren, wobei  $t$  allgemein das Alter  $t$  bezeichnet und  $\mathbf{t}$  für die konkrete Realisierung der Zufallsvariablen  $T$  steht.

Da  $P_\theta$  jedem  $\omega \in \Omega$  eine Wahrscheinlichkeit zuordnet und sich die Elemente der Ergebnismenge sowohl in der  $x$ - als auch in der  $t$ -Komponente unterscheiden können, kann  $P_\theta$  als das gemeinsame Wahrscheinlichkeitsmaß der Zufallsvariablen  $X$  und  $T$  interpretiert werden. Des Weiteren wird angenommen, dass für ein zufällig aus der Grundgesamtheit entnommenes Individuum das Jahr der Entstehung unabhängig von der Lebensdauer ist.

Daraus resultiert die stochastische Unabhängigkeit von  $X$  und  $T$ . Auch wird davon ausgegangen, dass  $X$  von  $\theta$  abhängt und somit  $T$  von  $\theta$  unabhängig ist. Es folgt:

$$P_\theta(\omega) = P_X(\omega; \theta)P_T(\omega), \forall \omega \in \Omega, \theta \in \Theta \quad (2.4)$$

Es sei angemerkt, dass die Verteilung von  $T$  keinen Einfluss auf die folgende Modellentwicklung hat und  $P_T(\omega)$  somit beliebig gewählt werden kann.

Die Lebensdauer selbst wird hingegen als eine Art Rundenspiel formuliert. Das bedeutet, dass jedes Jahr als eine neue Spielrunde interpretiert wird, in der das Individuum mit einer bestimmten Wahrscheinlichkeit entweder ausfallen (Erfolg) oder nicht ausfallen (Misserfolg) kann. Auch wird angenommen, dass sich die Wahrscheinlichkeit, innerhalb eines Jahres auszufallen, mit den Jahren nicht ändert. Die Zufallsvariable  $X$  aus (2.2) kann somit als die Anzahl der Spielrunden bis zum ersten Erfolg bei konstanter Erfolgswahrscheinlichkeit interpretiert werden und ist folglich geometrisch verteilt [siehe beispielsweise Johnson et al., 2005, Kapitel 5.2]. Die verwendete Version der geometrischen Verteilung hat somit die Wahrscheinlichkeitsfunktion nach Johnson et al. [2005, Formel 10.71] und kann folgendermaßen definiert werden.

**Definition 1** (Geometrische Verteilung). *Sei  $X$  eine diskrete Zufallsvariable mit dem Bild  $\mathbf{x} \in \mathbb{N}$ . Dann heißt  $X$  geometrisch verteilt mit dem Parameter  $\theta \in \Theta = [\xi, 1 - \xi]$  mit  $\xi \in (0, 1/2)$  genau dann, wenn die zugehörige Wahrscheinlichkeitsfunktion gegeben ist durch:*

$$f_G(\mathbf{x}; \theta) := P_\theta\{X = \mathbf{x}\} = P_X\{X = \mathbf{x}; \theta\} = \begin{cases} \theta(1 - \theta)^{\mathbf{x}-1}, & \text{falls } \mathbf{x} = 1, 2, \dots \\ 0, & \text{sonst} \end{cases} \quad (2.5)$$

Die Verteilungsfunktion von  $X$  ist entsprechend definiert durch:

$$F_G(\mathbf{x}; \theta) := P_\theta\{X \leq \mathbf{x}\} = P_X\{X \leq \mathbf{x}; \theta\} = \begin{cases} 1 - (1 - \theta)^{\mathbf{x}}, & \text{falls } \mathbf{x} = 1, 2, \dots \\ 0, & \text{sonst} \end{cases} \quad (2.6)$$

Für den Erwartungswert einer geometrisch verteilten Zufallsvariable gilt:

$$E_{\theta}\{X\} = \frac{1}{\theta} \quad (2.7)$$

Entsprechend der Definition der geometrischen Verteilung bezeichnet  $\theta$  die Wahrscheinlichkeit dafür, dass das Individuum im Jahr  $x$  ausfällt, unter der Voraussetzung, dass ein Ausfall vorher noch nicht stattgefunden hat:

$$\begin{aligned} & \frac{P_{\theta}\{\omega_{(x-1)G+(t+1)} : t \in \{0, \dots, G-1\}\}}{P_{\theta}\{\omega_{(k-1)G+(t+1)} : k \geq x, t \in \{0, \dots, G-1\}\}} \\ &= \frac{P_{\theta}\{X = x\}}{P_{\theta}\{X \geq x\}} = \frac{P_{\theta}\{X = x\}}{1 - P_{\theta}\{X \leq x-1\}} = \frac{f_G(x; \theta)}{1 - F_G(x-1; \theta)} = \frac{\theta(1-\theta)^{x-1}}{1 - (1 - (1-\theta)^{x-1})} = \theta \end{aligned}$$

Dem gegenüber gibt der Erwartungswert (2.7) die erwartete Lebensdauer des Individuums in Jahren an.

Das wesentliche Ziel dieses Kapitels ist die Konstruktion einer Likelihood-Funktion zur Schätzung von  $\theta$ . Diese lässt sich auch direkt aus (2.5) durch das Vertauschen von Argument und Parameter gewinnen. Allerdings wird im Folgenden eine Herleitung beschrieben, die zwar umständlicher ist, jedoch viele Konzepte beinhaltet, die für spätere Kapitel notwendig sind. Es sei angemerkt, dass zunächst keine Einschränkung der Beobachtbarkeit vorliegt und lediglich der Parameter der geometrischen Verteilung, der der Zufallsvariablen  $X$  zugrunde liegt, geschätzt werden soll. Informationen bezüglich des Entstehungsjahres und somit über die  $t$ -Komponente von  $\omega \in \Omega$  sind somit zunächst irrelevant und werden erst in späteren Kapiteln weiter thematisiert.

Zunächst wird die Zufallsvariable  $N(1)$  eingeführt. Diese gibt an, ob ein Ausfall des Individuums bereits im ersten Jahr nach der Entstehung stattgefunden hat oder nicht:

$$N(1) := \mathbb{1}_{\{X \leq 1\}} \quad (2.8)$$

Die binäre Zufallsgröße  $N(1)$  genügt dabei einer Bernoulli-Verteilung mit dem Parameter  $\theta$  [siehe Johnson et al., 2005, Kapitel 3.1].

**Definition 2** (Bernoulli-Verteilung). Sei  $N(1)$  eine Zufallsvariable mit dem Bild  $\nu(1) \in \{0, 1\}$ . Dann heißt  $N(1)$  Bernoulli-verteilt mit dem Parameter  $\theta \in [\xi, 1 - \xi]$  mit  $\xi \in (0, 1/2)$  genau dann, wenn die zugehörige Wahrscheinlichkeitsfunktion gegeben ist durch:

$$f_B(\nu(1); \theta) = P_{\theta}\{N(1) = \nu(1)\} \begin{cases} \theta^{\nu(1)}(1-\theta)^{1-\nu(1)} & , \text{ falls } \nu(1) \in \{0, 1\} \\ 0 & , \text{ sonst} \end{cases} \quad (2.9)$$

Für eine Bernoulli-verteilte Zufallsvariable gilt:

$$E_{\theta}\{N(1)\} = \theta \text{ sowie } \text{Var}_{\theta}\{N(1)\} = \theta(1-\theta) \quad (2.10)$$

In der Statistik ist es gängig, die Zerlegung einer Zufallsvariablen  $N(1)$  in ein deterministisches Signal  $E_\theta\{N(1)\} = \theta$  und ein zufälliges Rauschen  $R$  vorzunehmen, sodass:

$$N(1) - E_\theta\{N(1)\} = N(1) - \theta = R \quad (2.11)$$

Es sei angemerkt, dass das Signal  $E_\theta\{N(1)\}$ , unabhängig von der zugrundeliegenden Ergebnismenge  $\Omega$ , genau einen Wert  $\theta$  hat.

Die Betrachtung solch einer Zerlegung für lediglich das erste Jahr ist im gegebenen Kontext allerdings wenig zielführend. Stattdessen sollen im Folgenden alle Zeitpunkte  $x \in \mathbb{N}$  betrachtet werden können. Dazu wird im ersten Schritt der Wahrscheinlichkeitsraum  $(\Omega, \mathcal{F}, P_\theta)$  mit der Filtration  $\{\mathcal{F}_x : x \in \mathbb{N}_0\}$  zu einer stochastischen Basis  $(\Omega, \mathcal{F}, \{\mathcal{F}_x : x \in \mathbb{N}_0\}, P_\theta)$  erweitert. Die Filtrationselemente  $\mathcal{F}_x$  geben dabei an, welche Informationen zum jeweiligen Zeitpunkt  $x$  zur Verfügung stehen. Dazu sei angemerkt, dass Informationen bezüglich des Entstehungsjahres derzeit irrelevant sind und folglich nur Informationen über die Lebensdauer mithilfe von  $\{\mathcal{F}_x : x \in \mathbb{N}_0\}$  modelliert werden sollen. Eine Erweiterung der Filtration um die Entstehungsinformationen erfolgt dann in Kapitel 2.2.

Da für alle  $x \in \mathbb{N}_0$  gelten muss, dass  $\mathcal{F}_x \subseteq \mathcal{F}_{x+1}$ , können zwar neue Informationen mit voranschreitender Zeit hinzugefügt, alte Informationen jedoch nicht wieder vergessen werden. Somit wird mithilfe einer Filtration eine Art von Gedächtnis simuliert. Die konkrete Darstellung von  $\mathcal{F}_x$  basiert für alle  $x \in \mathbb{N}_0$  auf folgenden Überlegungen:

$x = 0$  : Ein Ausfall des Individuums nach 0 Jahren ist im Rahmen des betrachteten Modells nicht möglich. Es folgt  $\mathcal{F}_0 := \sigma\{\emptyset\}$ .

$x = 1$  : Nach einem Jahr ist lediglich bekannt, ob das Individuum im ersten Jahr nach der Entstehung ausgefallen ist oder nicht. Der Ausfall im ersten Jahr wird dabei mithilfe des Ereignisses  $\mathbb{X}_1 := X^{-1}(1) = \{\omega_1, \dots, \omega_G\}$  beschrieben. Die  $\sigma$ -Algebra  $\mathcal{F}_1$ , die alle Informationen nach einem Jahr zusammenfasst, ist somit definiert durch:

$$\mathcal{F}_1 := \sigma\{\{\omega_1, \dots, \omega_G\}\} = \sigma\{\mathbb{X}_1\}$$

Im Rahmen von Abbildung 2.1 entspricht  $\mathbb{X}_1$  der Menge  $\{\omega_1, \dots, \omega_5\}$ . Generell bezeichnet  $\mathbb{X}_x$  die  $x$ -te Spalte, wenn die Elemente aus  $\Omega$  nach dem Vorbild von Abbildung 2.1 angeordnet werden.

$x = 2$  : Nach zwei Jahren ist bekannt, ob ein Ausfall im ersten Jahr, im zweiten Jahr, oder nach zwei Jahren noch nicht stattgefunden hat. Der Wissenszuwachs besteht entsprechend in der Information über einen Ausfall im zweiten Jahr, dargestellt durch das Ereignis  $\mathbb{X}_2 := X^{-1}(2) = \{\omega_{G+1}, \dots, \omega_{2G}\}$ . Die resultierende  $\sigma$ -Algebra kann folglich durch

$$\mathcal{F}_2 := \sigma\{\{\omega_1, \dots, \omega_G\}, \{\omega_{G+1}, \dots, \omega_{2G}\}\} = \sigma\{\mathbb{X}_1, \mathbb{X}_2\}$$

beschrieben werden.

$x \in \mathbb{N}$ : Allgemein ist im  $x$ -ten Jahr bekannt, ob das Individuum innerhalb der ersten  $x$  Jahre nach der Entstehung ausgefallen ist. Im Falle eines solchen Ausfalls ist ebenso bekannt, in welchem Jahr genau der Ausfall stattgefunden hat. Ein Ausfall innerhalb des  $x$ -ten Jahres lässt sich dabei mithilfe des Ereignisses

$$\mathbb{X}_x := X^{-1}(x) = \left\{ \omega_l \in \Omega : \left\lfloor \frac{l}{G} \right\rfloor = x \right\} = \{ \omega_{(x-1)G+1}, \dots, \omega_{xG} \} \quad (2.12)$$

modellieren. Somit gilt für die Filtrationselemente von  $\{\mathcal{F}_x : x \in \mathbb{N}_0\}$ :

$$\begin{aligned} \mathcal{F}_0 &= \sigma\{\emptyset\} \\ \mathcal{F}_x &:= \sigma\{ \{\omega_1, \dots, \omega_G\}, \dots, \{\omega_{(x-1)G+1}, \dots, \omega_{xG}\} \} \\ &= \sigma\{ \mathbb{X}_k : k \in \{1, \dots, x\} \} \text{ wenn } x \in \mathbb{N} \end{aligned}$$

Im nächsten Schritt wird unter Zuhilfenahme der Filtration  $\{\mathcal{F}_x : x \in \mathbb{N}_0\}$  die univariate Zufallsvariable  $N(1)$  (vgl. (2.8)) zu einem stochastischen Prozess  $N$  mit  $N(0) = 0$  sowie

$$\begin{aligned} N(x) &:= \mathbb{1}_{\{X \leq x\}} \\ &= \begin{cases} 1, & \text{Individuum ist in den ersten } x \text{ Jahren nach der Entstehung ausgefallen} \\ 0, & \text{sonst} \end{cases} \end{aligned} \quad (2.13)$$

erweitert. Es sei angemerkt, dass es sich bei dem Prozess  $N$  um die zeitdiskrete Version eines so genannten Zählprozesses handelt. Diese werden, basierend auf der Definition von Zählprozessen nach Fleming und Harrington [1991, Definition 1.2.6], folgendermaßen definiert:

**Definition 3** (Diskreter Zählprozess). *Ein stochastischer Prozess  $N := \{N(x) : x \in \mathbb{N}_0\}$  heißt diskreter Zählprozess bezüglich einer Filtration  $\{\mathcal{F}_x : x \in \mathbb{N}_0\}$  genau dann, wenn folgende Eigenschaften erfüllt sind:*

- (i) Für alle  $x \in \mathbb{N}_0$  ist  $N(x)$  messbar bezüglich  $\mathcal{F}_x$ . Der Prozess  $N$  heißt dann  $\mathcal{F}_x$ -adaptiert.
- (ii) Es gilt  $N(0) = 0$  und  $N(x) < \infty$ , fast sicher für alle  $x \in \mathbb{N}$ .
- (iii) Es gilt:

$$\Delta N(x) := N(x) - N(x-1) \in \{0, 1\}, \forall x \in \mathbb{N}$$

Es sei darauf hingewiesen, dass diskrete Zählprozesse im Rahmen dieser Arbeit als Zählprozesse bezeichnet werden. Auch sei angemerkt, dass Gleichungen sowie Ungleichungen, die Zufallsvariablen beinhalten, immer nur fast sicher gelten. Aus Gründen der Übersichtlichkeit wird besagter Zusatz im Rahmen dieser Arbeit allerdings weggelassen.

Mithilfe von Zählprozessen kann eingesehen werden, wie oft bestimmte Ereignisse innerhalb einer gewissen Zeit eingetreten sind. Im vorliegenden Fall (2.13) lässt sich durch die Differenz  $N(x_2) - N(x_1)$  mit  $x_1 \in \mathbb{N}_0$ ,  $x_2 \in \mathbb{N}$  sowie  $x_1 + 1 \leq x_2$  einsehen, ob der Ausfall des Individuums innerhalb der ersten  $\{x_1 + 1, \dots, x_2\}$  Jahre nach der Entstehung eingetreten ist.

**Bemerkung 1.** Für den Zählprozess  $N(x) = \mathbb{1}_{\{X \leq x\}}$  gilt, dass dieser genau dann gleich eins ist, wenn eines der in (2.12) beschriebenen Ereignisse  $\mathbb{X}_1, \dots, \mathbb{X}_x$  eingetreten ist. Andererseits gilt, dass  $\mathbb{X}_x$  genau dann eingetreten ist, wenn  $N(x) = 1$  und  $N(x - 1) = 0$ . Somit lässt sich jedem Zählprozessverlauf ein eindeutiges Ereignis zuweisen und umgekehrt. Für die Filtration  $\{\mathcal{F}_x : x \in \mathbb{N}_0\}$  folgt somit:

$$\mathcal{F}_x = \sigma\{\mathbb{X}_1, \dots, \mathbb{X}_x\} = \sigma\{N(k) : 0 \leq k \leq x\} \quad (2.14)$$

Für einen Zählprozess  $N$  wäre eine zu (2.11) analoge Zerlegung  $N(x) - E_\theta\{N(x)\}$  ebenso möglich. Das Signal  $E_\theta\{N(x)\}$  hätte dann genauso viele Werte, wie es mögliche Zeitpunkte gibt. Allerdings wird sich ein anderer Dekompositionsansatz als hilfreicher herausstellen. Dafür wird das Signal nicht mehr als der Erwartungswert  $E_\theta\{N(x)\}$ , sondern als ein  $\mathcal{F}_x$ -vorhersagbarer, stochastischer Prozess  $A := \{A(x, \theta) : x \in \mathbb{N}_0, \theta \in \Theta\}$  definiert. Intuitiv formuliert heißt ein Prozess vorhersagbar, wenn sein Zustand im Zeitpunkt  $x$  durch die Informationen aus dem Zeitraum  $\{0, \dots, x - 1\}$  eindeutig determiniert ist. Analog zu Fleming und Harrington [1991, Definitionen 1.4.1 und 1.4.2] wird dieses Konzept im Folgenden anhand von vorhersagbaren  $\sigma$ -Algebren definiert.

**Definition 4** (Vorhersagbarkeit). Sei  $\{\Omega, \mathcal{F}, \{\mathcal{F}_x : x \in \mathbb{N}_0\}, P_\theta\}$  eine stochastische Basis.

(i) Die  $\sigma$ -Algebra  $\mathcal{F}^V$  auf  $\mathbb{N}_0 \times \Omega$ , die von allen Mengen der Form

$$\{0\} \times \mathbb{V}, \mathbb{V} \in \mathcal{F}_0$$

sowie

$$\{v_1 + 1, \dots, v_2\} \times \mathbb{V}, 0 \leq v_1 < v_2 < \infty, v_1 \in \mathbb{N}_0, v_2 \in \mathbb{N}, \mathbb{V} \in \mathcal{F}_{v_1}$$

erzeugt wird, heißt vorhersagbare  $\sigma$ -Algebra bezüglich der Filtration  $\{\mathcal{F}_x : x \in \mathbb{N}_0\}$ .

(ii) Ein Prozess  $A$  heißt vorhersagbar bezüglich einer Filtration  $\{\mathcal{F}_x : x \in \mathbb{N}_0\}$ , wenn er messbar bezüglich  $\mathcal{F}^V$  ist. Dann wird  $A$  als  $\mathcal{F}_x$ -vorhersagbarer Prozess bezeichnet.

**Bemerkung 2.** Ein Prozess  $A$  ist genau dann  $\mathcal{F}^V$ -messbar und somit  $\mathcal{F}_x$ -vorhersagbar, wenn er  $\mathcal{F}_{x-1}$ -adaptiert ist.

Wird beispielsweise  $A(0, \theta) = 0$  und  $A(x, \theta) = N(x - 1)$  für alle  $x \in \mathbb{N}$  gewählt, so ergibt sich durch

$$N(x) = A(x, \theta) + \Delta N(x),$$

wobei  $\Delta N(x) := N(x) - N(x - 1)$  gilt, eine mögliche Zerlegung des Prozesses  $N$ , da  $A(x, \theta)$  offensichtlich  $\mathcal{F}_{x-1}$ -messbar ist und somit die  $\mathcal{F}_x$ -Vorhersagbarkeit von  $A$  folgt. Jedoch ist die Zerlegung eines Zählprozesses in einen  $\mathcal{F}_x$ -vorhersagbaren Prozess und einen Rest ohne weitere Bedingungen nicht eindeutig. Um dieses Problem zu lösen, wird zunächst der Startpunkt  $A(0, \theta) = 0$  festgelegt. Des Weiteren sollen auch Anforderungen an den Rest der Zerlegung gestellt werden. Bei diesem soll es sich um einen im Mittel fairen stochastischen Prozess  $M$  handeln. Diese werden als Martingale bezeichnet und folgendermaßen definiert [siehe Fleming und Harrington, 1991, Definition 1.2.8]:

**Definition 5** (Martingal). *Sei  $M := \{M(x) : x \in \mathbb{N}_0\}$  ein stochastischer Prozess. Dann heißt  $M$  ein  $(\mathcal{F}_x, P_\theta)$ -Martingal genau dann, wenn die folgenden Bedingungen erfüllt sind:*

- (i)  $M$  ist  $\mathcal{F}_x$ -adaptiert.
- (ii)  $E_\theta |M(x)| < \infty$ , für alle  $x \in \mathbb{N}_0$
- (iii)  $E_\theta \{M(x + \kappa) | \mathcal{F}_x\} = M(x)$  für alle  $x \in \mathbb{N}_0$  und  $\kappa \in \mathbb{N}$

**Bemerkung 3.** *Wird ein Zählprozess mit der Martingaldefinition abgeglichen, fällt auf, dass die Eigenschaften (i) und (ii) auch von Zählprozessen erfüllt werden. Lediglich Eigenschaft (iii) wird durch die Ungleichung  $E_\theta \{N(x + \kappa) | \mathcal{F}_x\} \geq N(x)$  ersetzt. Damit sind Zählprozesse eine spezielle Variante so genannter Submartingale.*

Die Dekomposition eines stochastischen Prozesses in ein Martingal und einen vorher-sagbaren Prozess kann mithilfe des folgenden Satzes, der Doob-Zerlegung, formalisiert werden [siehe Doob, 1953, Kapitel 7, Formel 1.5, beziehungsweise Chung, 2001, Theorem 9.3.2]:

**Satz 1** (Doob-Zerlegung). *Auf der Grundlage einer stochastischen Basis  $\{\Omega, \mathcal{F}, \{\mathcal{F}_x : x \in \mathbb{N}_0\}, P_\theta\}$  sei  $N$  ein  $\mathcal{F}_x$ -adaptierter Prozess, sodass  $E_\theta |N(x)| < \infty$ , für alle  $x \in \mathbb{N}_0$  erfüllt ist. Dann existieren ein  $(\mathcal{F}_x, P_\theta)$ -Martingal  $M$  und ein  $\mathcal{F}_x$ -vorhersagbarer Prozess  $A$  mit  $E_\theta \{A(x, \theta)\} < \infty$  für alle  $x$ , sodass:*

$$N(x) = M(x) + A(x, \theta) \tag{2.15}$$

*Diese Zerlegung ist eindeutig, vorausgesetzt es gilt  $A(0, \theta) = 0$ . Wenn es sich bei  $N$  um ein Submartingal handelt, ist  $A$  zusätzlich monoton wachsend.*

*Der Prozess  $A$  wird auch als  $(\mathcal{F}_x, P_\theta)$ -Kompensator von  $N$  bezeichnet und folgendermaßen berechnet:*

$$\begin{aligned} A(0, \theta) &= 0 \\ A(x, \theta) &= \sum_{k=1}^x E_\theta \{\Delta N(k) | \mathcal{F}_{k-1}\} \end{aligned} \tag{2.16}$$

Die Beschreibung bedingter Erwartungswerte kann Definition 16 im Appendix, Kapitel A.3.1, entnommen werden. Außerdem ist eine elementare Anwendung der Doob-Zerlegung inklusive der Funktionsweise von vorhersagbaren  $\sigma$ -Algebren im Appendix, Kapitel A.1, gegeben.

Die Dekomposition eines Zählprozesses  $N$  in einen Kompensator  $A$  und ein Martingal  $M$  mithilfe von Satz 1 hat den wesentlichen Vorteil, dass sich alle  $\Delta A(x, \theta)$  als die bedingten Wahrscheinlichkeiten der sukzessiven Bernoulli-Experimente  $\Delta N(x)$ , gegeben der in  $\mathcal{F}_{x-1}$  enthaltenen Informationen, interpretieren lassen:

$$\begin{aligned} P_\theta\{\Delta N(x) = 1|\mathcal{F}_{x-1}\} &= E_\theta\{\Delta N(x)|\mathcal{F}_{x-1}\} \\ &= E_\theta\{\Delta M(x)|\mathcal{F}_{x-1}\} + E_\theta\{\Delta A(x, \theta)|\mathcal{F}_{x-1}\} \\ &= E_\theta\{\Delta A(x, \theta)|\mathcal{F}_{x-1}\} \\ &= \Delta A(x, \theta) \end{aligned}$$

Dies hat zur Folge, dass sich die Likelihood-Funktion zur Schätzung des geometrischen Parameters  $\theta$  mithilfe der Prozesse  $\Delta N$  und  $\Delta A$  darstellen lässt. Das folgenden Lemma liefert das entsprechende Resultat.

**Lemma 1.** *Der Beitrag zur Likelihood-Funktion eines Individuums mit der Lebensdauer  $X$ , das bis zu einem beliebigen, aber festen Zeitpunkt  $\chi \in \mathbb{N}$  beobachtet wird, ist gegeben durch:*

$$L(\theta) = \prod_{x=1}^{\chi} \left\{ (1 - \Delta A(x, \theta))^{1-\Delta N(x)} (\Delta A(x, \theta))^{\Delta N(x)} \right\} \quad (2.17)$$

Dabei gilt  $\Delta A(x, \theta) = E_\theta\{\Delta N(x)|\mathcal{F}_{x-1}\} = P_\theta\{\Delta N(x) = 1|\mathcal{F}_{x-1}\}$ .

**Bemerkung 4.** *Die Darstellung des Likelihood-Beitrags (2.17) lässt sich als diskrete Version der so genannten Formel von Jacod [siehe Andersen et al., 1993, Kapitel II.7.1] interpretieren.*

*Beweis Lemma 1.* Zunächst wird, basierend auf dem bis  $\chi$  beobachteten Zählprozess  $N$ , der Vektor

$$(\Delta N(1), \Delta N(2) \dots, \Delta N(\chi))' \quad (2.18)$$

mit der zugehörigen Wahrscheinlichkeitsfunktion  $f^{\Delta N}$  aufgestellt, wobei  $\Delta N(x) = N(x) - N(x-1)$  für alle  $x \in \{1, \dots, \chi\}$  gilt. Die Likelihood-Funktion (2.17) ist dann zunächst definiert als:

$$L(\theta) = f^{\Delta N}(\Delta N) = f^{\Delta N(1), \dots, \Delta N(\chi)}(\Delta N(1), \dots, \Delta N(\chi))$$

Die Zufallsvariablen  $\Delta N(x)$  mit

$$\Delta N(x) = \mathbb{1}_{\{X=x\}} \quad (2.19)$$

sind dabei Bernoulli-verteilt (vgl. Definition 2), jedoch nicht unabhängig voneinander. Da ein Individuum nämlich nur ein Mal ausfallen kann, folgt aus  $\Delta N(x) = 1$  direkt,

dass  $\Delta N(k) = 0$  für alle  $k \in \{x+1, \dots, \chi\}$ . Wird diese longitudinale Abhängigkeit innerhalb des Vektors (2.18) berücksichtigt, ergibt sich folgende weitere Darstellung der Likelihood-Funktion:

$$\begin{aligned} L(\theta) &= f^{\Delta N(\chi)|\Delta N(\chi-1), \dots, \Delta N(1)}(\Delta N(\chi) | \Delta N(\chi-1), \dots, \Delta N(1)) \\ &\quad \cdot f^{\Delta N(\chi-1)|\Delta N(\chi-2), \dots, \Delta N(1)}(\Delta N(\chi-1) | \Delta N(\chi-2), \dots, \Delta N(1)) \\ &\quad \cdot \dots \cdot f^{\Delta N(1)}(\Delta N(1)) \end{aligned} \quad (2.20)$$

Im nächsten Schritt wird ein Faktor von (2.20) für einen beliebigen Zeitpunkt  $x \in \{1, \dots, \chi\}$  betrachtet. Dazu sei angemerkt, dass es sich bei den Informationen über  $\Delta N(x-1), \dots, \Delta N(1)$  genau um die Informationen handelt, die in der  $\sigma$ -Algebra  $\mathcal{F}_{x-1}$  enthalten sind. Mithilfe der Konvention  $0^0 = 1$  folgt:

$$\begin{aligned} &f^{\Delta N(x)|\Delta N(x-1), \dots, \Delta N(1)}(\Delta N(x) | \Delta N(x-1), \dots, \Delta N(1)) \\ &= P_\theta\{\Delta N(x) = 1 | \mathcal{F}_{x-1}\}^{\Delta N(x)} \cdot P_\theta\{\Delta N(x) = 0 | \mathcal{F}_{x-1}\}^{1-\Delta N(x)} \\ &= P_\theta\{\Delta N(x) = 1 | \mathcal{F}_{x-1}\}^{\Delta N(x)} \cdot (1 - P_\theta\{\Delta N(x) = 1 | \mathcal{F}_{x-1}\})^{1-\Delta N(x)} \end{aligned}$$

Des Weiteren folgt aus der Eigenschaft (A.9) von bedingten Erwartungswerten sowie aus  $\Delta N(x) = \mathbb{1}_{\{\Delta N(x)=1\}}$ :

$$\begin{aligned} &f^{\Delta N(x)|\Delta N(x-1), \dots, \Delta N(1)}(\Delta N(x) | \Delta N(x-1), \dots, \Delta N(1)) \\ &= \left( \mathbb{E}\{\mathbb{1}_{\{\Delta N(x)=1\}} | \mathcal{F}_{x-1}\} \right)^{\Delta N(x)} \left( 1 - \mathbb{E}\{\mathbb{1}_{\{\Delta N(x)=1\}} | \mathcal{F}_{x-1}\} \right)^{1-\Delta N(x)} \\ &= \left( \mathbb{E}\{\Delta N(x) | \mathcal{F}_{x-1}\} \right)^{\Delta N(x)} \left( 1 - \mathbb{E}\{\Delta N(x) | \mathcal{F}_{x-1}\} \right)^{1-\Delta N(x)} \end{aligned}$$

Unter der Verwendung von (2.16) folgt weiter:

$$\begin{aligned} &f^{\Delta N(x)|\Delta N(x-1), \dots, \Delta N(1)}(\Delta N(x) | \Delta N(x-1), \dots, \Delta N(1)) \\ &= \Delta A(x, \theta)^{\Delta N(x)} (1 - \Delta A(x, \theta))^{1-\Delta N(x)} \end{aligned} \quad (2.21)$$

Abschließend folgt durch Einsetzen von (2.21) in (2.20):

$$L(\theta) = \prod_{x=1}^{\chi} \left\{ (1 - \Delta A(x, \theta))^{1-\Delta N(x)} (\Delta A(x, \theta))^{\Delta N(x)} \right\}$$

□

Wird  $\theta$  als der Parameter der geometrischen Verteilung interpretiert, lässt sich die Darstellung (2.17) weiter vereinfachen. Dazu liefert das folgende Lemma eine alternative Darstellung für die Kompensatordifferenz  $\Delta A$ .

**Lemma 2.** *Sei  $X$  eine mit dem Parameter  $\theta$  geometrisch verteilte Zufallsvariable und  $N = \mathbb{1}_{\{X \leq x\}}$  ein Zählprozess, der bezüglich der Filtration  $\{\mathcal{F}_x : x \in \mathbb{N}_0\}$  mit  $\mathcal{F}_x =$*

$\sigma\{\mathbb{X}_1, \dots, \mathbb{X}_x\} = \sigma\{\{\omega_1, \dots, \omega_G\}, \dots, \{\omega_{(x-1)G+1}, \dots, \omega_{xG}\}\}$  adaptiert ist. Dann gilt:

$$\Delta A(x, \theta) = \mathbb{E}_\theta\{\Delta N(x) | \mathcal{F}_{x-1}\} = \theta Y(x-1) \quad (2.22)$$

Der Prozess  $Y$  wird als der Unter-Risiko-Prozess bezeichnet. Dabei gibt  $Y(x)$  an, ob das Individuum zum Zeitpunkt  $x$  noch lebt, somit in der Zukunft von  $x$  noch ausfallen kann, und folglich unter Risiko steht, oder nicht [siehe Fleming und Harrington, 1991, Beispiel 1.4.2]:

$$Y(x) := \mathbb{1}_{\{X_{-1} \geq x\}} = \begin{cases} 1, & \text{Individuum kann in der Zukunft von } x \text{ noch ausfallen} \\ 0, & \text{sonst} \end{cases} \quad (2.23)$$

**Bemerkung 5.** Da aus den Definitionen von  $N$  (vgl. 2.13) und  $Y$  (vgl. 2.23) folgt, dass

$$N(x) = 1 \Leftrightarrow Y(x) = 0$$

sowie

$$N(x) = 0 \Leftrightarrow Y(x) = 1$$

gelten, ergibt sich folgende alternative Darstellung der Filtrationselemente von  $\{\mathcal{F}_x : x \in \mathbb{N}_0\}$ :

$$\mathcal{F}_x = \sigma\{Y(k) : 0 \leq k \leq x\}$$

Somit ist auch  $Y$  ein  $\mathcal{F}_x$ -adaptierte Prozess.

*Beweis Lemma 2.* Nach Definition 16 (vgl. Appendix, Kapitel A.3.1) ist  $\theta Y(x-1)$  genau dann der bedingte Erwartungswert von  $\Delta N(x)$  bezüglich  $(\mathcal{F}_{x-1}, P_\theta)$ , wenn die folgenden Bedingungen erfüllt sind:

(i)  $\theta Y(x-1)$  ist messbar bezüglich  $\mathcal{F}_{x-1}$

(ii) Für alle  $\mathbb{F} \in \mathcal{F}_{x-1}$  gilt:

$$\int_{\mathbb{F}} \Delta N(x) dP_\theta = \int_{\mathbb{F}} \theta Y(x-1) dP_\theta$$

Um Bedingung (i) zu überprüfen, sei zunächst angemerkt, das  $\theta$  deterministisch ist. Auch ist  $Y$  nach Bemerkung 5 ein  $\mathcal{F}_x$ -adaptierter Prozess. Es folgt die Messbarkeit von  $Y(x-1)$  bezüglich  $\mathcal{F}_{x-1}$  und somit Bedingung (i).

Für den Nachweis von Bedingung (ii) folgt zunächst die Integrierbarkeit von Zähl- und Unter-Risiko-Prozess aus der Beschränktheit  $|N(x)| \leq 1$  beziehungsweise  $|Y(x)| \leq 1$  für alle  $x \in \mathbb{N}_0$

Aufgrund der Additivität der Integration genügt es normalerweise Bedingung (ii) nur für die erzeugenden Elemente nachzuweisen. Die  $\sigma$ -Algebra  $\mathcal{F}_{x-1}$  wird dabei von den Ereignissen  $\mathbb{X}_1, \dots, \mathbb{X}_{x-1}$  erzeugt, wobei gilt, dass:

$$X(\omega) \leq x - 1, \forall \omega \in \bigcup_{1 \leq k \leq x-1} \mathbb{X}_k$$

Dem folgend, ergibt sich für alle erzeugenden Ereignisse  $\mathbb{X}_k, k \in \{1, \dots, x - 1\}$ , dass  $N(x - 1) = 1$  sowie  $N(x) = 1$  und somit:

$$\int_{\mathbb{X}_k} \Delta N(x) dP_\theta = \int_{\mathbb{X}_k} N(x) - N(x - 1) dP_\theta = \int_{\mathbb{X}_k} 1 - 1 dP_\theta = \int_{\mathbb{X}_k} 0 dP_\theta = 0$$

Auch gilt für alle  $\omega \in \bigcup_{1 \leq k \leq x-1} \mathbb{X}_k$ , dass  $Y(x - 1) = 0$  und somit auch:

$$\int_{\mathbb{X}_k} \theta Y(x - 1) dP_\theta = \int_{\mathbb{X}_k} 0 dP_\theta = 0$$

Da für jedes erzeugende Element gilt, dass

$$\int_{\mathbb{F}} \Delta N(x) dP_\theta = \int_{\mathbb{X}_k} \theta Y(x - 1) dP_\theta = 0,$$

, könnte anstelle von  $\theta$  jede beliebige Konstante gewählt werden, ohne Bedingung (ii) zu verletzen. Folglich ist der Beweis an dieser Stelle noch nicht abgeschlossen. Ergänzend wird ein beliebiges Komplement  $\Omega \setminus \mathbb{X}_{\tilde{x}}$  mit  $\tilde{x} \in \{1, \dots, x - 1\}$  betrachtet. Für ein Element  $\omega \in \Omega \setminus \mathbb{X}_{\tilde{x}}$  gilt:

$$\Delta N(x)(\omega) = 1 \Leftrightarrow \omega \in \mathbb{X}_x$$

Somit ergibt sich für ein beliebiges Komplement  $\Omega \setminus \mathbb{X}_{\tilde{x}}$ :

$$\int_{\Omega \setminus \mathbb{X}_{\tilde{x}}} \Delta N(x) dP_\theta = \int_{\mathbb{X}_x} dP_\theta = P_\theta(\mathbb{X}_x) = \theta(1 - \theta)^{x-1}$$

Außerdem gilt für  $\omega \in \Omega \setminus \mathbb{X}_{\tilde{x}}$ :

$$Y(x - 1)(\omega) = 1 \Leftrightarrow \omega \in \Omega \setminus \left\{ \bigcup_{1 \leq k \leq x-1} \mathbb{X}_k \right\}$$

Mithilfe von

$$\begin{aligned} P_\theta \left\{ \Omega \setminus \left\{ \bigcup_{1 \leq k \leq x-1} \mathbb{X}_k \right\} \right\} &= 1 - P_\theta \left\{ \left\{ \bigcup_{1 \leq k \leq x-1} \mathbb{X}_k \right\} \right\} = 1 - F_G(x - 1; \theta) \\ &= 1 - (1 - (1 - \theta)^{x-1}) = (1 - \theta)^{1-x} \end{aligned}$$

folgt weiter:

$$\begin{aligned} \int_{\Omega \setminus \mathbb{X}_k} \theta Y(x-1) dP_\theta &= \theta \int_{\Omega \setminus \{\bigcup_{1 \leq k \leq x-1} \mathbb{X}_k\}} dP_\theta \\ &= \theta P_\theta \left\{ \Omega \setminus \left\{ \bigcup_{1 \leq k \leq x-1} \mathbb{X}_k \right\} \right\} = \theta(1-\theta)^{x-1} \end{aligned}$$

Bedingung (ii) ist damit auch erfüllt. Es folgt die Behauptung

$$\Delta A(x, \theta) = \theta Y(x-1), \quad \forall x \in \mathbb{N}$$

□

Unter der Verwendung von Lemma 2 und Gleichung (2.16) folgt die alternative Darstellung des Kompensators:

$$A(x, \theta) = \theta \sum_{k=1}^x Y(k-1) \quad (2.24)$$

Im letzten Schritt wird, basierend auf der allgemeinen Likelihood-Funktion (2.17), eine Likelihood-Funktion zur Schätzung des geometrischen Parameters  $\theta$  aufgestellt. Dazu sei angemerkt, dass  $\Delta N(x) = \mathbb{1}_{\{X=x\}}$  gilt (vgl. 2.19). Außerdem ist die Kompensator-differenz  $\Delta A$  für eine geometrisch verteilte Lebensdauer nach Lemma 2 gegeben durch:

$$\Delta A(x, \theta) = \theta Y(x-1) = \theta \mathbb{1}_{\{X \geq x\}} \quad (2.25)$$

Einsetzen von (2.19) und (2.25) in (2.17) liefert abschließend die Likelihood-Funktion zur Schätzung des Parameters der geometrischen Verteilung  $\theta$ :

$$\begin{aligned} L(\theta) &= \prod_{x=1}^X \left\{ (1 - \mathbb{1}_{\{X \geq x\}} \theta)^{1 - \mathbb{1}_{\{X=x\}}} (\mathbb{1}_{\{X \geq x\}} \theta)^{\mathbb{1}_{\{X=x\}}} \right\} \\ &= \left[ \prod_{x=1}^{\min(X-1, X)} (1 - \theta)^1 \theta^0 \right] \cdot \left[ (1 - \theta)^0 \theta^1 \right]^{\mathbb{1}_{\{X \leq X\}}} \\ &= \theta^{\mathbb{1}_{\{X \leq X\}}} (1 - \theta)^{\min(X-1, X)} \end{aligned} \quad (2.26)$$

Die Likelihood-Funktion (2.26) basiert dabei auf der Annahme, dass für ein zufällig aus der Grundgesamtheit gezogenes Individuum die Lebensdauer in jedem Fall beobachtet werden kann. Diese Annahme ist in der Praxis allerdings nicht haltbar. Entsprechend wird das Modell in den folgenden Kapiteln um Restriktionen bezüglich der Beobachtbarkeit, namentlich die linksseitige Trunkierung sowie die rechtsseitige Zensur, erweitert und die Likelihood-Funktion zur Schätzung der Ausfallwahrscheinlichkeit  $\theta$  entsprechend angepasst. Dazu wird zunächst in Kapitel 2.2 eine bedingte Likelihood-Funktion zur Schätzung von  $\theta$  unter Berücksichtigung der linksseitigen Trunkierung hergeleitet.

## 2.2. Linkstrunkierung

Soll die Wahrscheinlichkeit für einen Ausfall innerhalb eines Jahres und, damit einhergehend, die erwartete Lebensdauer von Individuen untersucht werden, geschieht dies mithilfe von Daten, die oft im Rahmen einer Studie erhoben werden. Dabei gibt es in der Regel einen Zeitpunkt, an dem diese Studie beginnt. Auch wenn sich in der Regel ebenso ein Zeitpunkt für das Studienende definieren lässt, basiert das in diesem Kapitel entwickelte Modell auf der Annahme, dass die Studie ab dem Zeitpunkt des Beginns unbegrenzt andauert. Eine Ergänzung des Modells um ein Studienende und die damit einhergehende Rechtszensur der Historie erfolgt dann in Kapitel 2.3.

Es wird davon ausgegangen, dass die Studie ein Jahr nach dem Ende des Entstehungszeitraums beginnt. Umfasst der Entstehungszeitraum allgemein die  $G$  Jahre  $\{a, a + 1, \dots, b\}$ , so beginnt die Studie im Jahr  $b + 1$  (vgl. Abbildung 2.2). Folglich wird angenommen, dass die Grundgesamtheit nur Individuen umfasst, deren Entstehung zum Zeitpunkt des Studienbeginns höchstens  $G$  Jahre zurückliegt. Da ab dem Studienbeginn im Jahr  $b + 1$  der Ausfall eines Individuums beobachtet werden kann, wird der Zeitraum der Studie  $\{b + 1, b + 2, \dots\}$  auch als Beobachtungszeitraum bezeichnet.

Wird jedoch ein in der Grundgesamtheit enthaltendes Individuum, das vor dem Beginn der Studie ausfällt, zufällig gezogen, werden keine Informationen bezüglich der Lebensdauer beobachtet und nicht einmal die Existenz des Individuums zur Kenntnis genommen.

Dieses als linksseitige Trunkierung beziehungsweise Linkstrunkierung bezeichnete Phänomen soll im nächsten Schritt bei der Schätzung der Ausfallwahrscheinlichkeit  $\theta$  berücksichtigt werden.

Folglich gibt es zunächst zwei verschiedene Szenarien, nämlich den Ausfall vor und den Ausfall während der Studie, deren Illustrierung sich beispielhaft für  $G = 5$  der Abbildung 2.2 entnehmen lässt.

Das Ziel dieses Kapitels ist die Herleitung des Anteils eines Individuums zu einer bedingten Likelihood-Funktion, mithilfe derer der geometrische Parameter  $\theta$ , unter Berücksichtigung der Linkstrunkierung, geschätzt werden kann. Dazu muss im ersten Schritt der zugrundeliegende Wahrscheinlichkeitsraum diskutiert werden.

Die linksseitige Trunkierung kann nämlich auf zwei verschiedene Arten erfolgen, die in Toparkus und Weißbach [2025, Abbildung 2] visualisiert werden. Dabei besteht die eine Variante darin, erst aus der Grundgesamtheit alle Individuen, die vor dem Studienbeginn ausfallen, zu entfernen und anschließend ein Individuum aus der linkstrunkierten Teilgesamtheit zu ziehen. Dieser Ansatz wird jedoch im Rahmen dieser Arbeit nicht weiter verfolgt.

Stattdessen wird ein Individuum direkt aus der Grundgesamtheit gezogen und erst dann die Fallunterscheidung getroffen, ob dieses Individuum beobachtet werden kann oder nicht. Folglich beschreibt die Ziehung eines Individuums aus der Grundgesamtheit dasselbe Zufallsexperiment wie im vorherigen Kapitel 2.1, sodass erneut der Wahrscheinlichkeitsraum  $(\Omega, \mathcal{F}, P_\theta)$  als Grundlage verwendet werden kann. Die Elemente der Ergebnismenge  $\Omega = \{\omega_1, \omega_2, \dots\}$  sind dabei bereits in (2.1) definiert worden. Ebenso ändern sich  $\mathcal{F} = \mathcal{P}(\Omega)$  sowie das Wahrscheinlichkeitsmaß  $P_\theta$  (vgl. (2.4) und Definition 1) im Vergleich zum Modell ohne Restriktionen bezüglich der Beobachtbarkeit nicht.

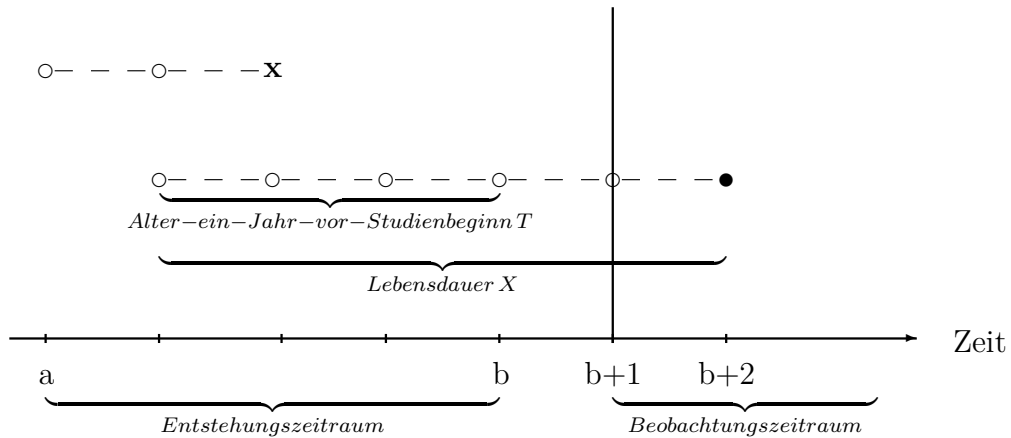


Abbildung 2.2.: Mögliche Pfade hinsichtlich der Beobachtbarkeit für einen Entstehungszeitraum der Länge  $G = |\{a, \dots, b\}| = 5$ ; die vertikale Linie zum Zeitpunkt  $b + 1$  markiert den Beginn der Studie; (i) Der obere Pfad beschreibt ein Individuum, das in  $a$  entsteht und nach zwei Jahren, also vor dem Studienbeginn, ausfällt. Weder das Ausfallalter noch das Individuum selbst werden beobachtet. (ii) Der untere Pfad beschreibt ein Individuum, dass in  $a + 1$  entstanden und in  $b + 2$ , also nach dem Studienbeginn, ausgefallen ist. Das Ausfallalter wird beobachtet.

Zusätzlich sind auf dem Wahrscheinlichkeitsraum  $(\Omega, \mathcal{F}, P_\theta)$  die Zufallsvariable  $X$ , die die Lebensdauer eines Individuums beschreibt (vgl. (2.2)), sowie die Zufallsvariable  $T$ , die das Alter des Individuums im letzten Jahr des Entstehungszeitraums angibt (vgl. (2.3)), definiert. Es sei angemerkt, dass  $T$  ebenso als das Alter ein Jahr vor dem Beginn des Beobachtungszeitraums interpretiert werden kann.

Wie sich Abbildung 2.2 entnehmen lässt, kann der Ausfall eines Individuums genau dann beobachtet werden, wenn dieser ein Jahr vor dem Beginn des Beobachtungszeitraums noch nicht stattgefunden hat. Somit ist die Beobachtbarkeit des Individuums und folglich dessen Ausfallalter lediglich für das Ereignis

$$\tilde{\mathbb{T}} := \{T + 1 \leq X\} \tag{2.27}$$

gewährleistet.

Welche Elemente der Ergebnismenge  $\omega \in \Omega$  mit der Beobachtbarkeit des Individuums einhergehen, kann für  $G = 5$  beispielhaft Abbildung 2.3 entnommen werden.

Bei der Herleitung einer bedingten Likelihood-Funktion zur Schätzung des geometrischen Parameters  $\theta$  unter Berücksichtigung der linksseitigen Trunkierung wird im Folgenden der Ansatz verfolgt, dass diese sich, analog zu (2.17), erneut mithilfe eines Zählprozesses und des zugehörigen Kompensators bezüglich einer Filtration darstellen lässt. Zunächst wird zu diesem Zweck, basierend auf Andersen et al. [1988, Kapitel 4] eine geeignete Filtration konstruiert.

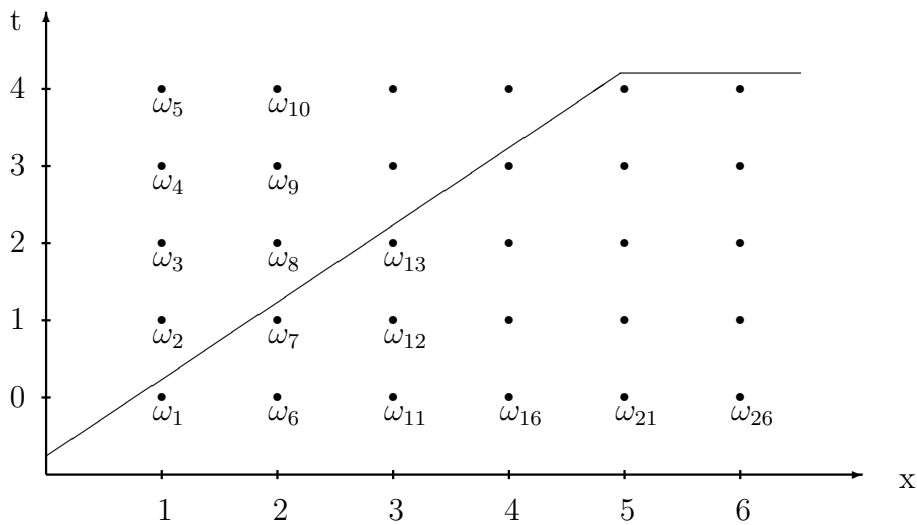


Abbildung 2.3.: Zuordnung der Elementarereignisse zu einem Entstehungs- und Ausfalljahr bei einem 5 jährigen Entstehungszeitraum inklusive Visualisierung der Beobachtbarkeit: Für alle  $\omega$  unterhalb der Linie ist das Individuum inklusive dessen Ausfall beobachtbar, alle  $\omega$  oberhalb der Linie repräsentieren ein nicht beobachtbares Individuum. Beispielsweise ist im Fall  $\omega_6 = \omega_{(2-1)5+(0+1)}$  das Ereignis  $\{T + 1 \leq X\}$  eingetreten und somit die Beobachtbarkeit gegeben, da  $X = 2$  und  $T = 0$ . Im Gegensatz dazu ist beispielsweise  $\omega_3 = \omega_{(1-1)5+(2+1)}$  mit  $X = 1$  und  $T = 2$  verbunden. Das Individuum fällt somit nach einem Jahr aus, tritt aber erst nach  $T + 1 = 3$  Jahren in die Studie ein und wird folglich nicht beobachtet.

### 2.2.1. Die latente Filtration

Um die Linkstrunkierung berücksichtigen zu können, sollte eine geeignete Filtration neben den Informationen bezüglich der Lebensdauer  $X$  auch Informationen hinsichtlich des Entstehungsjahres und somit über  $T$  beinhalten. Ebenso sollte die Frage, ob ein Individuum beobachtbar ist oder nicht, zu jedem Zeitpunkt beantwortet werden können. Das Ereignis  $\tilde{T}$  (vgl. (2.27)) sollte folglich in jedem Element der Filtration enthalten sein.

Als Ausgangslage wird die in (2.14) beschriebene Filtration  $\{\mathcal{F}_x : x \in \mathbb{N}_0\}$  verwendet. Diese enthält allerdings lediglich Informationen hinsichtlich der Lebensdauer eines Individuums und ist somit für weitere Betrachtungen zu grob. Im Folgenden wird daher eine Erweiterung von  $\{\mathcal{F}_x : x \in \mathbb{N}_0\}$  um Informationen über das in  $T$  codierte Entstehungsjahr und  $\tilde{T}$  konstruiert.

Zunächst wird eine neue Filtration  $\{\mathcal{G}_x : x \in \mathbb{N}_0\}$  eingeführt, deren Elemente  $\mathcal{G}_x$  zusätzlich zu den Informationen über die Lebensdauer, die in  $\mathcal{F}_x$  enthalten sind, auch Informationen bezüglich des Entstehungsjahres und somit über  $T$  beinhalten. Eine zusätzliche Erweiterung um die Messbarkeit von  $\tilde{T}$  erfolgt an einer späteren Stelle.

Analog zum, der Lebensdauer  $X$  zugeordneten, Zählprozess  $N$  (vgl. (2.13)) sei dazu  $\{\mathbb{1}_{\{T \leq t\}} : t \in \{0, \dots, G - 1\}\}$  der entsprechende Zählprozess für  $T$ . Zunächst wird

angenommen, dass für ein Individuum zu jedem Zeitpunkt  $x \in \mathbb{N}_0$  das Entstehungsjahr und somit der Wert von  $\mathbb{1}_{\{T \leq t\}}$  für alle  $t \in \{0, \dots, G-1\}$  bekannt sein soll. Es folgt:

$$\begin{aligned} \mathcal{G}_x &= \mathcal{F}_x \vee \sigma\{\mathbb{1}_{\{T \leq t\}} : t \in \{0, \dots, G-1\}\} \\ &= \sigma\{N(k), \mathbb{1}_{\{T \leq t\}} : k \in \{0, \dots, x\}, t \in \{0, \dots, G-1\}\} \end{aligned} \quad (2.28)$$

Dabei bedeutet das Symbol ' $\vee$ ', angewendet auf zwei  $\sigma$ -Algebren  $\mathcal{F}_x$  und  $\sigma\{\mathbb{1}_{\{T \leq t\}} : t \in \{0, \dots, G-1\}\}$ , dass die Verknüpfung  $\mathcal{G}_x$  sowohl von allen Mengen des Erzeugendensystems von  $\mathcal{F}_x$  als auch von allen Mengen des Erzeugendensystems von  $\sigma\{\mathbb{1}_{\{T \leq t\}} : t \in \{0, \dots, G-1\}\}$  erzeugt wird [siehe Chung, 2001, Seite 20].

Des Weiteren sei

$$\mathbb{T}_t := T^{-1}(t) = \{\omega_\iota \in \Omega : (\iota \bmod G = t + 1)\} \text{ mit } t \in \{0, \dots, G-1\} \quad (2.29)$$

die Menge aller Ereignisse, in denen das Individuum ein Jahr vor Studienbeginn  $t$  Jahre alt ist. Für den Fall  $G = 5$  (vgl. Abbildung 2.3) gilt beispielsweise  $\mathbb{T}_0 = \{\omega_1, \omega_6, \omega_{11}, \dots\}$  oder  $\mathbb{T}_1 = \{\omega_2, \omega_7, \omega_{12}, \dots\}$ . Allgemein beschreibt  $\mathbb{T}_t$  für eine Anordnung der Elemente des Ergebnisraums nach dem Vorbild von Abbildungen 2.3 die  $t$ -te Reihe.

Außerdem sei angemerkt, dass einerseits  $\mathbb{T}_t$  mit  $t \in \{0, \dots, G-1\}$  genau dann eingetreten ist, wenn  $\mathbb{1}_{\{T \leq t\}} = 1$  und  $\mathbb{1}_{\{T \leq (t-1)\}} = 0$ . Andererseits gilt  $\mathbb{1}_{\{T \leq t\}} = 1$  genau dann, wenn eines der Ereignisse  $\mathbb{T}_0, \dots, \mathbb{T}_t$  eingetreten ist. Jeder Verlauf des Prozesses  $\{\mathbb{1}_{\{T \leq t\}} : t \in \{0, \dots, G-1\}\}$  kann somit eindeutig mithilfe der Mengen  $\mathbb{T}_0, \dots, \mathbb{T}_t$  dargestellt werden und umgekehrt. Zusammen mit (2.14) sowie (2.28) resultiert die folgende, alternative Darstellung von  $\mathcal{G}_x$ :

$$\begin{aligned} \mathcal{G}_x &= \mathcal{F}_x \vee \sigma\{\mathbb{1}_{\{T \leq t\}}, t \in \{0, \dots, G-1\}\} \\ &= \sigma\{\mathbb{X}_1, \dots, \mathbb{X}_x\} \vee \sigma\{\mathbb{T}_t, t \in \{0, \dots, G-1\}\} \end{aligned} \quad (2.30)$$

Des Weiteren sei  $\mathcal{G}_0 := \sigma\{\mathbb{T}_t, t \in \{0, \dots, G-1\}\}$ , sodass

$$\mathcal{G}_x = \mathcal{F}_x \vee \mathcal{G}_0 \quad (2.31)$$

folgt.

**Satz 2.** *Der Zählprozess  $N$ , der in (2.13) definiert wurde, hat auch bezüglich  $\{\mathcal{G}_x : x \in \mathbb{N}_0\}$  und  $P_\theta$  den Kompensator  $A$  (vgl. (2.24)).*

*Beweis.* Zunächst folgt aus der  $\mathcal{F}_x$ -Vorhersagbarkeit von  $A$  und der Teilmengenbeziehung  $\mathcal{F}_x \subset \mathcal{G}_x$ , dass  $A$  auch bezüglich  $\{\mathcal{G}_x : x \in \mathbb{N}_0\}$  ein vorhersagbarer Prozess ist. Da  $N$  außerdem adaptiert bezüglich  $\{\mathcal{G}_x : x \in \mathbb{N}_0\}$  ist und alle weiteren Eigenschaften eines Zählprozesses erfüllt (vgl. Definition 3), ist  $N$  auch ein  $\mathcal{G}_x$ -Zählprozess.

Es bleibt zu zeigen, dass es sich bei  $M = N - A$  um ein  $(\mathcal{G}_x, P_\theta)$ -Martingal handelt.

Da  $M$  adaptiert bezüglich  $\{\mathcal{F}_x : x \in \mathbb{N}_0\}$  ist, ist  $M$  auch  $\mathcal{G}_x$ -adaptiert. Um die Integrierbarkeit von  $M$  zu zeigen, gilt zunächst mithilfe der Darstellung des Kompensators

(2.24):

$$\mathbb{E}_\theta |M(x)| = \mathbb{E}_\theta |N(x) - A(x, \theta)| \leq \mathbb{E}_\theta |N(x)| + \theta \mathbb{E}_\theta \left| \sum_{k=1}^x Y(k-1) \right|$$

Unter der Verwendung von  $\mathbb{E}_\theta |N(x)| \leq 1$  sowie  $Y(k-1) \in \{0, 1\}$  für  $k \in \mathbb{N}$  gilt weiter:

$$\mathbb{E}_\theta |M(x)| \leq 1 + \theta \sum_{k=1}^x \mathbb{E}_\theta |Y(k-1)| = 1 + \theta \sum_{k=1}^x \mathbb{E}_\theta \{Y(k-1)\}$$

Des Weiteren ergibt sich für  $\mathbb{E}_\theta \{Y(k-1)\}$ :

$$\mathbb{E}_\theta \{Y(k-1)\} = \mathbb{E}_\theta \{\mathbb{1}_{\{X \geq k\}}\} = P_\theta \{X \geq k\} = 1 - F_G(k-1; \theta) = (1 - \theta)^{k-1}$$

Abschließend folgt:

$$\mathbb{E}_\theta |M(x)| \leq 1 + \sum_{k=1}^x \theta (1 - \theta)^{k-1} = 1 + F_G(x; \theta) \leq 2 < \infty, \forall x \in \mathbb{N}_0$$

Zum Nachweis der Martingaleigenschaft  $\mathbb{E}_\theta \{M(x + \kappa) | \mathcal{G}_x\} = M(x)$  für  $x \in \mathbb{N}_0$  und  $\kappa \in \mathbb{N}$  wird stellvertretend

$$\mathbb{E}_\theta \{M(x) | \mathcal{G}_{x-1}\} = M(x-1), \forall x \in \mathbb{N}$$

nachgewiesen, da dann aufgrund der Turmregel (vgl. Satz 13, (ii))

$$\mathbb{E}_\theta \{M(x + \kappa) | \mathcal{G}_x\} = \mathbb{E}_\theta \left\{ \mathbb{E}_\theta \left\{ \dots \mathbb{E}_\theta \{M(x + \kappa) | \mathcal{G}_{x+\kappa-1}\} \dots \middle| \mathcal{G}_{x+1} \right\} \middle| \mathcal{G}_x \right\} = M(x)$$

folgt und der Beweis damit abgeschlossen ist.

Zunächst gilt:

$$\begin{aligned} & \mathbb{E}_\theta \{M(x) | \mathcal{G}_{x-1}\} = M(x-1) \\ \Leftrightarrow & \mathbb{E}_\theta \{\Delta M(x) | \mathcal{G}_{x-1}\} = 0 \\ \Leftrightarrow & \mathbb{E}_\theta \{\Delta N(x) - \Delta A(x, \theta) | \mathcal{G}_{x-1}\} = 0 \\ \Leftrightarrow & \mathbb{E}_\theta \{\Delta N(x) | \mathcal{G}_{x-1}\} = \mathbb{E}_\theta \{\Delta A(x, \theta) | \mathcal{G}_{x-1}\} \\ \Leftrightarrow & \mathbb{E}_\theta \{\Delta N(x) | \mathcal{G}_{x-1}\} = \theta Y(x-1) \end{aligned} \tag{2.32}$$

Somit bleibt zu zeigen, dass  $\theta Y(x-1)$  der bedingte Erwartungswert von  $\Delta N$  bezüglich  $\mathcal{G}_{x-1}$  und  $P_\theta$  ist. Dies ist genau dann der Fall, wenn die folgenden Bedingungen erfüllt sind (vgl. Definition 16):

(i)  $\theta Y(x-1)$  ist messbar bezüglich  $\mathcal{G}_{x-1}$

(ii) Für alle  $\mathbb{G} \in \mathcal{G}_{x-1}$  gilt:

$$\int_{\mathbb{G}} \Delta N(x) dP_\theta = \int_{\mathbb{G}} \theta Y(x-1) dP_\theta$$

Eigenschaft (i) folgt dabei direkt aus der Messbarkeit von  $\theta Y(x-1)$  bezüglich  $\mathcal{F}_{x-1}$ .

Für den Nachweis von Bedingung (ii) folgt zunächst die Integrierbarkeit von Zähl- und Unter-Risiko-Prozess aus der Beschränktheit  $|N(x)| \leq 1$  beziehungsweise  $|Y(x)| \leq 1$  für alle  $x \in \mathbb{N}_0$

Aufgrund der Additivität der Integration genügt es, Bedingung (ii) nur für die erzeugenden Elemente, die nicht auf die 0 abgebildet werden, nachzuweisen. Dabei wird  $\mathcal{G}_{x-1}$  von den Elementen  $\mathbb{X}_1, \dots, \mathbb{X}_{x-1}$  sowie  $\mathbb{T}_t$  mit  $t \in \{0, \dots, G-1\}$  erzeugt. Für alle  $\mathbb{X}_k$  mit  $k \in \{1, \dots, x-1\}$  respektive deren Komplemente wurde Eigenschaft (ii) bereits im Beweis von Lemma 2 nachgewiesen. Es bleibt somit

$$\int_{\mathbb{T}_t} N(x) - N(x-1) dP_\theta = \int_{\mathbb{T}_t} \theta Y(x-1) dP_\theta, \forall t \in \{0, \dots, G-1\} \quad (2.33)$$

zu zeigen.

Für die linke Seite von (2.33) gilt wegen

$$P_T(\omega_{(k-1)G+t+1}) = P_T(\omega_{(x-1)G+t+1})$$

für alle  $k \in \mathbb{N}$  und unter Zuhilfenahme der Unabhängigkeit von Lebensdauer und Entstehungsjahr (vgl. (2.4)) sowie der Wahrscheinlichkeitsfunktion der Lebensdauern (2.5):

$$\begin{aligned} \int_{\mathbb{T}_t} \Delta N(x) dP_\theta &= \sum_{\omega \in \mathbb{T}_t} \Delta N(x)(\omega) P_\theta(\omega) \\ &= \sum_{k=1}^{\infty} \Delta N(x)(\omega_{(k-1)G+t+1}) P_X(\omega_{(k-1)G+t+1}) P_T(\omega_{(k-1)G+t+1}) \\ &= P_T(\omega_{(x-1)G+t+1}) \sum_{k=1}^{\infty} \Delta N(x)(\omega_{(k-1)G+t+1}) \theta (1-\theta)^{k-1} \end{aligned} \quad (2.34)$$

Für alle  $k \leq x-1$  ist dabei:

$$\Delta N(x)(\omega_{(k-1)G+t+1}) = N(x)(\omega_{(k-1)G+t+1}) - N(x-1)(\omega_{(k-1)G+t+1}) = 1 - 1 = 0$$

Des Weiteren ist für alle  $k \geq x+1$  ebenfalls:

$$\Delta N(x)(\omega_{(k-1)G+t+1}) = N(x)(\omega_{(k-1)G+t+1}) - N(x-1)(\omega_{(k-1)G+t+1}) = 0 - 0 = 0$$

Lediglich für den Fall  $k = x$  ist:

$$\Delta N(x)(\omega_{(k-1)G+t+1}) = N(x)(\omega_{(k-1)G+t+1}) - N(x-1)(\omega_{(k-1)G+t+1}) = 1 - 0 = 1$$

Somit ergibt sich für (2.34):

$$\begin{aligned} \int_{\mathbb{T}_t} \Delta N(x) dP_\theta &= P_T(\omega_{(x-1)G+t+1}) \theta (1-\theta)^{x-1} \\ &= P_T(\omega_{(x-1)G+t+1}) P_X(\omega_{(x-1)G+t+1}) = P_\theta(\omega_{(x-1)G+t+1}) \end{aligned}$$

Für die rechte Seite von (2.33) ergibt sich aufgrund der Unabhängigkeit von Lebens-

dauer und Entstehungsjahr:

$$\begin{aligned} \int_{\mathbb{T}_t} \theta Y(x-1) dP_\theta &= \sum_{\omega \in \mathbb{T}_t} \theta Y(x-1)(\omega) P_\theta(\omega) \\ &= P_T(\omega_{(x-1)G+t+1}) \theta \sum_{k=1}^{\infty} Y(x-1)(\omega_{(k-1)G+t+1}) \theta (1-\theta)^{k-1} \end{aligned} \quad (2.35)$$

Dabei gilt

$$Y(x-1)(\omega_{(k-1)G+t+1}) = \begin{cases} 0, & \text{wenn } k \leq x-1 \\ 1, & \text{wenn } k \geq x \end{cases} \quad (2.36)$$

Abschließend liefert Einsetzen von (2.36) in (2.35):

$$\begin{aligned} \int_{\mathbb{T}_t} \theta Y(x-1) dP_\theta &= P_T(\omega_{(x-1)G+t+1}) \theta \sum_{k=x}^{\infty} \theta (1-\theta)^{k-1} \\ &= P_T(\omega_{(x-1)G+t+1}) \theta (1 - F_G(x-1; \theta)) \\ &= P_T(\omega_{(x-1)G+t+1}) \theta (1-\theta)^{x-1} = P_\theta(\omega_{(x-1)G+t+1}) \end{aligned}$$

Damit ist (2.33) erfüllt und der Beweis abgeschlossen.  $\square$

**Bemerkung 6.** Die Elemente der Filtration  $\{\mathcal{G}_x : x \in \mathbb{N}_0\}$  lassen sich alternativ durch

$$\mathcal{G}_x = \sigma\{\omega_1, \dots, \omega_{Gx}\} \vee \mathcal{G}_0$$

darstellen.

Zum Nachweis der Gleichheit genügt es zu zeigen, dass alle Elemente des Erzeugendensystems von  $\mathcal{G}_x$  in  $\sigma\{\omega_1, \dots, \omega_{Gx}\} \vee \mathcal{G}_0$  enthalten sind und umgekehrt. Die Ereignisse  $\mathbb{T}_t$  mit  $t \in \{0, \dots, G-1\}$  sind per Definition in  $\mathcal{G}_0$  und somit in  $\mathcal{G}_x$  sowie  $\sigma\{\omega_1, \dots, \omega_{Gx}\} \vee \mathcal{G}_0$  enthalten.

Für den Nachweis von  $\mathcal{G}_x \subseteq \sigma\{\omega_1, \dots, \omega_{Gx}\} \vee \mathcal{G}_0$  sei angemerkt, dass  $\sigma$ -Algebren stabil bezüglich abzählbaren Vereinigungen sind und die erzeugenden Mengen  $\mathbb{X}_k$  mit  $k \in \{1, \dots, x\}$  von  $\mathcal{G}_x$  darstellbar sind als:

$$\mathbb{X}_k = \bigcup_{Gk-(G-1) \leq \iota \leq Gk} \{\omega_\iota\}$$

Es folgt für alle  $k \in \{1, \dots, x\}$ , dass  $\mathbb{X}_k \in \sigma\{\omega_1, \dots, \omega_{Gx}\}$  und somit auch  $\mathcal{G}_x \subseteq \sigma\{\omega_1, \dots, \omega_{Gx}\} \vee \mathcal{G}_0$ .

Abschließend muss somit noch  $\omega_\iota \in \mathcal{G}_x$  für alle  $\iota \in \{1, \dots, Gx\}$  nachgewiesen werden. Dazu sei angemerkt, dass sich jede natürliche Zahl  $\iota \in \{1, \dots, Gx\}$  eindeutig schreiben lässt als  $\iota = (k-1)G + t + 1$  mit  $k \in \{1, \dots, x\}$  und  $t \in \{0, \dots, G-1\}$ . Da eine  $\sigma$ -Algebra durchschnittsstabil ist, gilt für alle Mengen  $\mathbb{X}_k \in \mathcal{G}_x$  und  $\mathbb{T}_t \in \mathcal{G}_x$ , dass ebenso

$\mathbb{X}_k \cap \mathbb{T}_t \in \mathcal{G}_x$  gelten muss. Mit (2.12) und (2.29) folgt:

$$\begin{aligned} \mathbb{X}_k \cap \mathbb{T}_t &= \left\{ \omega_\iota : \left\lceil \frac{\iota}{G} \right\rceil = k \right\} \cap \{ \omega_\iota : (\iota \bmod G) = t + 1 \} \\ &= \left\{ \omega_\iota : \left\lceil \frac{\iota}{G} \right\rceil = k \right\} \cap \{ \omega_{(\bar{k}-1)G+t+1}, \bar{k} \in \mathbb{N} \} \\ &= \left\{ \omega_{(\bar{k}-1)G+t+1} : \left\lceil \frac{(\bar{k}-1)G+t+1}{G} \right\rceil = k \right\} \\ &= \left\{ \omega_{(\bar{k}-1)G+t+1} : (\bar{k}-1) + \left\lceil \frac{t+1}{G} \right\rceil = k \right\} \end{aligned}$$

Da  $t + 1 \leq G$  gilt, folgt  $\lceil (t + 1)/G \rceil = 1$  und somit:

$$\mathbb{X}_k \cap \mathbb{T}_t = \left\{ \omega_{(\bar{k}-1)G+t+1} : \bar{k} = k \right\} = \{ \omega_{(k-1)G+t+1} \} \quad (2.37)$$

Somit kann jedes Element  $\omega_\iota$  mit  $\iota \in \{1, \dots, Gx\}$  als Schnitt einer Menge  $\mathbb{X}_k$  mit einer Menge  $\mathbb{T}_t$  realisiert werden. Es folgt  $\omega_\iota \in \mathcal{G}_x$  und somit  $\sigma\{\omega_1, \dots, \omega_{Gx}\} \vee \mathcal{G}_0 \subseteq \mathcal{G}_x$ . Insgesamt ergibt sich  $\sigma\{\omega_1, \dots, \omega_{Gx}\} \vee \mathcal{G}_0 = \mathcal{G}_x$ .

Da die Konstruktion einer Filtration  $\{\mathcal{G}_x : x \in \mathbb{N}_0\}$ , die für ein Individuum die benötigten Informationen hinsichtlich der Lebensdauer  $X$  sowie des in  $T$  codierten Entstehungsjahres enthält, abgeschlossen ist, erfolgt im nächsten Schritt eine zusätzliche Erweiterung.

Aufgrund der Tatsache, dass die Beobachtbarkeit des Individuums lediglich für das Ereignis  $\tilde{\mathbb{T}} = \{T + 1 \leq X\}$  (vgl. (2.27)) gegeben ist, ist es sinnvoll, die Filtration  $\{\mathcal{G}_x : x \in \mathbb{N}_0\}$  dahingehend zu modifizieren, dass zu jedem Zeitpunkt erkennbar ist, ob das Ereignis  $\tilde{\mathbb{T}}$  eingetreten ist oder nicht. Bezüglich der erweiterten Filtration sollte somit die Messbarkeit von  $\tilde{\mathbb{T}}$  für jeden Zeitpunkt  $x \in \mathbb{N}_0$  gegeben sein. Um so eine Erweiterung von  $\{\mathcal{G}_x : x \in \mathbb{N}_0\}$  zu konstruieren kann sich der Tatsache, dass es sich bei der Zufallsvariablen  $T$  um eine so genannte  $\mathcal{G}_x$ -Stopppzeit handelt, bedient werden. Diese werden folgendermaßen definiert [siehe beispielsweise Bauer, 1974, Kapitel XI., §58, Formel (58.2)]:

**Definition 6** (Stopppzeit). Sei  $\{\mathcal{G}_x : x \in \mathbb{N}_0\}$  eine Filtration auf einem Wahrscheinlichkeitsraum  $(\Omega, \mathcal{F}, P_\theta)$ . Eine nichtnegative Zufallsvariable  $T$  heißt eine  $\mathcal{G}_x$ -Stopppzeit genau dann, wenn

$$\{T = x\} \in \mathcal{G}_x, \forall x \in \mathbb{N}_0.$$

Intuitiv ist klar, dass für ein Individuum die Informationen zum Entstehungsjahr in  $\mathcal{G}_x$  enthalten sind und somit zu jedem Zeitpunkt  $x$  beantwortet werden kann, ob  $T = x$  gilt. Aus mathematischer Perspektive folgt aus

$$\{T = x\} = \{ \omega_\iota \in \Omega : \iota \bmod G = x + 1 \} = \begin{cases} \mathbb{T}_x, & \text{wenn } 0 \leq x \leq G - 1 \\ \emptyset, & \text{sonst,} \end{cases} \quad (2.38)$$

dass es sich bei  $T$  um eine  $\mathcal{G}_x$ -Stoppzeit handelt, da sowohl  $\mathbb{T}_x \in \mathcal{G}_x$  für alle  $x \in \{0, \dots, G-1\}$  als auch  $\emptyset \in \mathcal{G}_x$  erfüllt ist.

**Bemerkung 7.** Mithilfe der Teilmengenbeziehung  $\mathcal{G}_1 \subseteq \mathcal{G}_2 \subseteq \dots$  lässt sich einsehen, dass

$$\{T = x\} \in \mathcal{G}_x$$

auch direkt

$$\left(\{T = 0\} \cup \{T = 1\} \cup \dots \cup \{T = x\}\right) \in \mathcal{G}_x$$

und somit auch

$$\{T \leq x\} \in \mathcal{G}_x \tag{2.39}$$

impliziert.

Die gestoppte  $\sigma$ -Algebra, die alle bis  $T$  eingetretenen Ereignisse enthält, ist dann definiert als

$$\mathcal{G}_T := \sigma\{\mathbb{T} \in \mathcal{F} : \mathbb{T} \cap \{T = x\} \in \mathcal{G}_x, \forall x \in \mathbb{N}_0\}$$

Mithilfe von Bemerkung 7 kann diese folgendermaßen umformuliert werden [siehe Billingsley, 2012, Formel 35.20]:

$$\mathcal{G}_T = \sigma\{\mathbb{T} \in \mathcal{F} : \mathbb{T} \cap \{T \leq x\} \in \mathcal{G}_x, \forall x \in \mathbb{N}_0\} \tag{2.40}$$

Mithilfe von (2.31) und (2.40) kann dann die modifizierte Filtration  $\{{}_T\mathcal{G}_x : x \in \mathbb{N}_0\}$  mit

$${}_T\mathcal{G}_x := \mathcal{G}_x \vee \mathcal{G}_T, \tag{2.41}$$

die die Informationen zum Zeitpunkt  $x \vee T := \max(x, T)$  enthält, definiert werden. Der folgende Satz liefert nun die geforderte  ${}_T\mathcal{G}_x$ -Messbarkeit von  $\tilde{\mathbb{T}} = \{T + 1 \leq X\}$  für alle  $x \in \mathbb{N}_0$ .

**Satz 3.** Gegeben sei die Filtration  $\{{}_T\mathcal{G}_x : x \in \mathbb{N}_0\}$  mit  ${}_T\mathcal{G}_x := \mathcal{G}_x \vee \mathcal{G}_T$ , die auf dem Wahrscheinlichkeitsraum  $(\Omega, \mathcal{F}, P_\theta)$  (vgl. (2.1) und (2.4)) definiert ist. Dabei wurden  $\mathcal{G}_x$  in (2.30) und  $\mathcal{G}_T$  in (2.40) eingeführt. Das Ereignis  $\tilde{\mathbb{T}} = \{T + 1 \leq X\}$  mit  $X$  aus (2.2) und  $T$  aus (2.3) ist dann in der gestoppten  $\sigma$ -Algebra  $\mathcal{G}_T$  enthalten und somit messbar bezüglich  ${}_T\mathcal{G}_x$  für alle  $x \in \mathbb{N}_0$ .

*Beweis.* Um nachzuweisen, dass  $\tilde{\mathbb{T}}$  messbar bezüglich  ${}_T\mathcal{G}_x$  für alle  $x \in \mathbb{N}_0$  ist, genügt es zu zeigen, dass  $\tilde{\mathbb{T}}$  messbar bezüglich  $\mathcal{G}_T$  ist. Nach (2.40) muss folglich überprüft werden, ob

$$\tilde{\mathbb{T}} \cap \{T \leq x\} \in \mathcal{G}_x, \forall x \in \mathbb{N}_0 \tag{2.42}$$

erfüllt ist. Zunächst folgt aus  $\{T = x\} = \mathbb{T}_x$  für  $0 \leq x \leq G-1$  und  $\{T = x\} = \emptyset$  für

$x \geq G$  (vgl. (2.38)), dass:

$$\{T \leq x\} = \bigcup_{t=0}^{x \wedge (G-1)} \mathbb{T}_t$$

Außerdem gilt für  $\tilde{\mathbb{T}}$

$$\begin{aligned} \tilde{\mathbb{T}} = \{T + 1 \leq X\} &= \{\omega_{(k-1)G+t+1} : t \in \{0, \dots, G-1\}, k \in \mathbb{N}, t+1 \leq k\} \\ &= \bigcup_{t=0}^{G-1} \{\omega_{(k-1)G+t+1} : k \in \mathbb{N}, t+1 \leq k\} \\ &= \bigcup_{t=0}^{G-1} \left( \mathbb{T}_t \cap \left( \bigcup_{k=t+1}^{\infty} \mathbb{X}_k \right) \right) \\ &= \bigcup_{t=0}^{G-1} \left( \mathbb{T}_t \cap \left( \Omega \setminus \bigcup_{k=1}^t \mathbb{X}_k \right) \right) \\ &= \bigcup_{t=0}^{G-1} \left( \mathbb{T}_t \setminus \left( \bigcup_{k=1}^t \mathbb{X}_k \right) \right), \end{aligned}$$

wobei  $\bigcup_{k=1}^0 \mathbb{X}_k = \emptyset$  verwendet wird. Abschließend folgt:

$$\tilde{\mathbb{T}} \cap \{T \leq x\} = \left[ \bigcup_{t=0}^{G-1} \left( \mathbb{T}_t \setminus \left( \bigcup_{k=1}^t \mathbb{X}_k \right) \right) \right] \cap \left[ \bigcup_{t=0}^{x \wedge (G-1)} \mathbb{T}_t \right] = \bigcup_{t=0}^{x \wedge (G-1)} \left( \mathbb{T}_t \setminus \left( \bigcup_{k=1}^t \mathbb{X}_k \right) \right)$$

Da  $\mathbb{T}_t \in \mathcal{G}_x$  für  $t \in \{0, \dots, x \wedge (G-1)\}$  und  $\mathbb{X}_k \in \mathcal{G}_x$  für  $k \in \{1, \dots, t \leq x\}$  erfüllt ist (vgl. (2.30)), folgt (2.42) und damit die Behauptung.  $\square$

### 2.2.2. Linkstrunkierter Zählprozess und Kompensator

Nachdem mit  $\{ {}_T \mathcal{G}_x : x \in \mathbb{N}_0 \}$  eine zunächst geeignete Filtration konstruiert wurde, muss im nächsten Schritt die Frage nach dem Zählprozess, der bei der Berücksichtigung der linksseitigen Trunkierung beobachtbar ist, geklärt werden. Für  $X \leq T$  und somit  $N(T) = \mathbb{1}_{\{X \leq T\}} = 1$  wird nämlich die Lebensdauer des Individuums trunkiert (oberer Pfad in Abbildung 2.2) und somit der Ausfall nicht beobachtet. Außerdem findet keine beobachtbare Entwicklung nach  $T$  mehr statt. Somit lässt sich mithilfe des linkstrunkierten Zählprozesses  ${}_T N$  [siehe Andersen et al., 1988, Formel (4.1)] mit (vgl. Definition 3)

$${}_T N(x) := N(x) - N(x \wedge T) = \mathbb{1}_{\{T+1 \leq X \leq x\}} \quad (2.43)$$

einsehen, ob zu einem Zeitpunkt  $x \in \mathbb{N}_0$  ein sichtbarer Ausfall des Individuums bereits stattgefunden hat. Dabei ist  $x \wedge T$  als das Minimum von  $x$  und  $T$  definiert. Des Weiteren wird, basierend auf (2.23), der linkstrunkierte Unter-Risiko-Prozess [siehe Andersen et

al., 1988, Kapitel 4] definiert durch:

$${}_T Y(x) := \mathbb{1}_{\{T \leq x\}} Y(x) = \mathbb{1}_{\{T \leq x\}} \mathbb{1}_{\{X-1 \geq x\}} = \mathbb{1}_{\{T+1 \leq x+1 \leq X\}} \quad (2.44)$$

Es sei angemerkt, dass die Entscheidung, die individualspezifischen Experimente erst ab  $T$  zu beobachten, als Marginalisierung interpretiert werden kann. Auf diese wird jedoch im Verlauf der Arbeit nicht weiter eingegangen. Stattdessen werden  ${}_T N$  sowie  ${}_T Y$  als die gegebenen Daten interpretiert.

Für die Erstellung einer bedingten Likelihood-Funktion zur Schätzung des geometrischen Parameters muss zunächst der linkstrunkierte Zählprozess  ${}_T N$  in ein Martingal und einen Kompensator zerlegt werden. Da dieser Kompensator von  ${}_T Y$  abhängen wird, ist es sinnvoll, zunächst zu zeigen, dass es sich zum einen  ${}_T N$  tatsächlich einen bei  ${}_T \mathcal{G}_x$ -Zählprozess handelt und zum anderen  ${}_T Y$  ein  ${}_T \mathcal{G}_x$ -adaptierter Prozess ist. Der folgende Satz liefert die entsprechenden Resultate.

**Satz 4.** *Der Prozess  ${}_T N$  ist ein Zählprozess bezüglich der Filtration  $\{{}_T \mathcal{G}_x : x \in \mathbb{N}_0\}$ . Außerdem ist  ${}_T Y$  ein  ${}_T \mathcal{G}_x$ -adaptierter Prozess.*

*Beweis.* Da  $N$  und  $Y$  adaptiert bezüglich der Filtration  $\{\mathcal{F}_x : x \in \mathbb{N}_0\}$  sind (vgl. (2.14) und Bemerkung 5), folgt aus der Teilmengenbeziehung  $\mathcal{F}_x \subset \mathcal{G}_x \subseteq {}_T \mathcal{G}_x$  für alle  $x \in \mathbb{N}_0$ , dass  $N$  und  $Y$  auch  ${}_T \mathcal{G}_x$ -adaptierte Prozesse sind. Auch folgt aus der Definition von  ${}_T N$  (2.43), dass:

$${}_T N(x) = \mathbb{1}_{\{T+1 \leq X\}} \mathbb{1}_{\{X \leq x\}} = \mathbb{1}_{\tilde{\mathbb{T}}} N(x)$$

Die Indikatorfunktion  $\mathbb{1}_{\tilde{\mathbb{T}}}$  kann dabei folgendermaßen interpretiert werden:

$$\mathbb{1}_{\tilde{\mathbb{T}}} = \begin{cases} 1, & \text{wenn das Ereignis } \tilde{\mathbb{T}} \text{ eingetreten ist} \\ 0, & \text{sonst} \end{cases}$$

Das Ereignis  $\tilde{\mathbb{T}} = \{T+1 \leq X\}$  ist dabei per Konstruktion messbar bezüglich  ${}_T \mathcal{G}_x$  für alle  $x \in \mathbb{N}_0$ . Folglich ist der linkstrunkierte Zählprozess  ${}_T N$  als Komposition zweier  ${}_T \mathcal{G}_x$ -adaptierter Prozesse ebenfalls  ${}_T \mathcal{G}_x$ -adaptiert.

Außerdem gilt  ${}_T N(0) = N(0) - N(0) = 0$  sowie:

$${}_T N(x) = N(x) - N(x \wedge T) \leq N(x) < \infty$$

Darüber hinaus besitzt  ${}_T N$  nur maximal einen Sprung der Höhe 1. Damit ist  ${}_T N$  ein Zählprozess bezüglich  $\{{}_T \mathcal{G}_x : x \in \mathbb{N}_0\}$ .

Da es sich bei  $T$  um eine  ${}_T \mathcal{G}_x$ -Stoppzeit handelt, gilt außerdem  $\{T \leq x\} \in \mathcal{G}_x \subseteq {}_T \mathcal{G}_x$  (vgl. Bemerkung 7 und insbesondere (2.39)). Somit ist der Prozess  $\mathbb{1}_{\{T \leq x\}}$  messbar bezüglich  ${}_T \mathcal{G}_x$ . Folglich ist der linkstrunkierte Unter-Risiko-Prozess  ${}_T Y$  als Produkt (vgl. (2.44))

$${}_T Y(x) = \mathbb{1}_{\{T+1 \leq x+1 \leq X\}} = \mathbb{1}_{\{T+1 \leq x+1\}} \mathbb{1}_{\{X-1 \geq x\}} = \mathbb{1}_{\{T \leq x\}} Y(x)$$

zweier  ${}_T \mathcal{G}_x$ -adaptierter Prozesse selbst auch ein  ${}_T \mathcal{G}_x$ -adaptierter Prozess.  $\square$

Da  ${}_T N$  ein  ${}_T \mathcal{G}_x$ -Zählprozess ist, kann im nächsten Schritt eine zur Doob-Zerlegung von  $N$  (vgl. 2.15) analoge Dekomposition von  ${}_T N$  in ein Martingal und einen Kompensator erfolgen. Es scheint naheliegend, dass der zu  ${}_T N$  analog konstruierte Prozess  ${}_T A$  mit  ${}_T A(0, \theta) = 0$  und

$$\begin{aligned} {}_T A(x, \theta) &= A(x, \theta) - A(x \wedge T, \theta) \\ &= \sum_{k=1}^x \mathbb{E}_\theta \{ \Delta N(k) | \mathcal{G}_{k-1} \} - \sum_{k=1}^{x \wedge T} \mathbb{E}_\theta \{ \Delta N(k) | \mathcal{G}_{k-1} \} \\ &= \sum_{k=1}^x \mathbb{1}_{\{T+1 \leq k\}} \mathbb{E}_\theta \{ \Delta N(k) | \mathcal{G}_{k-1} \} \\ &= \sum_{k=1}^x \mathbb{1}_{\{T+1 \leq k\}} \theta Y(k-1) \\ &= \theta \sum_{k=1}^x {}_T Y(k-1) \end{aligned}$$

für  $x \in \mathbb{N}$  einen Kompensator von  ${}_T N$  darstellt.

Diese Vermutung kann durch das folgende Theorem, das einen Spezialfall von Proposition 4.1 aus Andersen et al. [1988] darstellt, bestätigt werden:

**Satz 5.** *Ist ein Individuum lediglich für das Ereignis  $\tilde{\mathbb{T}} = \{T + 1 \leq X\}$  mit*

$$\alpha_\theta := P_\theta \{ \tilde{\mathbb{T}} \} > 0$$

*beobachtbar, so hat der zugehörige linkstrunkierte Zählprozess  ${}_T N$  aus (2.43) den Kompensator  ${}_T A$  mit  ${}_T A(0, \theta) = 0$  sowie*

$${}_T A(x, \theta) = \theta \sum_{k=1}^x {}_T Y(k-1) \tag{2.45}$$

*Dieser ist definiert bezüglich der Filtration  $\{ {}_T \mathcal{G}_x : x \in \mathbb{N}_0 \}$  (vgl. 2.41) sowie dem bedingten Wahrscheinlichkeitsmaß  $P_\theta^{\tilde{\mathbb{T}}}$  mit*

$$P_\theta^{\tilde{\mathbb{T}}} \{ \mathbb{G} \} := \frac{P_\theta \{ \mathbb{G} \cap \tilde{\mathbb{T}} \}}{P_\theta \{ \tilde{\mathbb{T}} \}} = \frac{P_\theta \{ \mathbb{G} \cap \tilde{\mathbb{T}} \}}{\alpha_\theta}. \tag{2.46}$$

**Bemerkung 8.** (i) *Tatsächlich ist  ${}_T A$  auch der  ${}_T \mathcal{G}_x$ -Kompensator von  ${}_T N$  bezüglich des unbedingten Wahrscheinlichkeitsmaßes  $P_\theta$ . Jedoch liefert die Bedingung auf Beobachtbarkeit im späteren Verlauf einige Vorteile bei der Konstruktion der bedingten Likelihood-Funktion zur Schätzung der geometrischen Parameter  $\theta$ .*

(ii) *Erwartungswerte bezüglich  $P_\theta^{\tilde{\mathbb{T}}}$  können für beliebige Zufallsvariablen, hier beispielsweise dargestellt für Komponenten des Zählprozesses  $N(x)$  (vgl. (2.13)) mit der Realisierung  $\nu(x)$ , folgendermaßen formuliert werden [siehe Klenke 2020, Defini-*

tion 8.9]:

$$\mathbb{E}_{\theta}^{\tilde{\mathbb{T}}}\{N(x)\} = \sum_{\nu(x)} \nu(x) P_{\theta}\{N(x) = \nu(x) | \tilde{\mathbb{T}}\} = \mathbb{E}_{\theta}\{N(x) | \tilde{\mathbb{T}}\} = \frac{\mathbb{E}_{\theta}\{\mathbb{1}_{\tilde{\mathbb{T}}} N(x)\}}{\alpha_{\theta}} \quad (2.47)$$

Insbesondere gilt somit  $\mathbb{E}_{\theta}^{\tilde{\mathbb{T}}}\{N(x)\} = \mathbb{E}_{\theta}^{\tilde{\mathbb{T}}}\{ {}_T N(x) \} = \mathbb{E}_{\theta}\{ {}_T N(x) \} / \alpha_{\theta}$ .

*Beweis.* Es ist zu zeigen: Aus der Tatsache, dass es sich bei  $M = N - A$  um ein Martingal bezüglich der Filtration  $\{\mathcal{G}_x : x \in \mathbb{N}_0\}$  sowie dem Wahrscheinlichkeitsmaß  $P_{\theta}$  handelt (vgl. Satz 2), folgt, dass  ${}_T M$  mit

$${}_T M(x) = M(x) - M(x \wedge T) = {}_T N(x) - {}_T A(x, \theta)$$

ein Martingal bezüglich der Filtration  $\{ {}_T \mathcal{G}_x : x \in \mathbb{N}_0 \}$  sowie dem bedingten Wahrscheinlichkeitsmaß  $P_{\theta}^{\tilde{\mathbb{T}}}$  ist. Somit ist zu zeigen:

- (i) Der Prozess  ${}_T M$  ist  ${}_T \mathcal{G}_x$ -adaptiert.
- (ii) Für alle  $x \in \mathbb{N}_0$  gilt  $\mathbb{E}_{\theta}^{\tilde{\mathbb{T}}}| {}_T M(x) | < \infty$ .
- (iii) Es gilt die Martingaleigenschaft  $\mathbb{E}_{\theta}^{\tilde{\mathbb{T}}}\{ {}_T M(x + \kappa) | {}_T \mathcal{G}_x \} = {}_T M(x)$  für alle  $x \in \mathbb{N}_0$  und  $\kappa \in \mathbb{N}_0$ .

Zunächst wird, zum Nachweis von (i), die Doob-Zerlegung

$${}_T M(x) = {}_T N(x) - {}_T A(x, \theta) = {}_T N(x) - \theta \sum_{k=1}^x {}_T Y(k-1)$$

betrachtet.

Da  $\theta$  deterministisch ist und, nach Satz 4, die Prozesse  ${}_T N$  und  ${}_T Y$  adaptiert bezüglich  $\{ {}_T \mathcal{G}_x : x \in \mathbb{N}_0 \}$  sind, handelt es sich bei  ${}_T M$  als Komposition  ${}_T \mathcal{G}_x$ -adaptierter Prozesse ebenfalls um einen  ${}_T \mathcal{G}_x$ -adaptierten Prozess.

Um Eigenschaft (ii), also  $\mathbb{E}_{\theta}^{\tilde{\mathbb{T}}}| {}_T M(x) | < \infty, \forall x \in \mathbb{N}_0$ , nachzuweisen, wird zunächst unter der Verwendung von (2.47) die Zerlegung

$$\begin{aligned} \mathbb{E}_{\theta}^{\tilde{\mathbb{T}}}| {}_T M(x) | &= \frac{1}{\alpha_{\theta}} \mathbb{E}_{\theta}\{ \mathbb{1}_{\tilde{\mathbb{T}}}|M(x) - M(x \wedge T)| \} \\ &\leq \frac{1}{\alpha_{\theta}} \mathbb{E}_{\theta}\{ |M(x)| + |M(x \wedge T)| \} = \frac{1}{\alpha_{\theta}} \left( \mathbb{E}_{\theta}|M(x)| + \mathbb{E}_{\theta}|M(x \wedge T)| \right) \end{aligned}$$

betrachtet. Der Term  $\mathbb{E}_{\theta}|M(x)|$  ist endlich, da  $M$  ein  $(\mathcal{G}_x, P_{\theta})$ -Martingal ist. Der Prozess  $M(\cdot \wedge T)$  ist nach dem Theorem des optionalen Stoppens (vgl. Appendix, Satz 14) ebenfalls ein  $(\mathcal{G}_x, P_{\theta})$ -Martingal und folglich gilt auch  $\mathbb{E}_{\theta}|M(x \wedge T)| < \infty$  für alle  $x \in \mathbb{N}_0$ . Eigenschaft (ii) ist somit nachgewiesen.

Abschließend muss (iii) und somit die Martingaleigenschaft  $E_{\theta}^{\tilde{\mathbb{T}}}\{ {}_T M(x + \kappa) \mid {}_T \mathcal{G}_x \} = {}_T M(x)$  für alle  $x \in \mathbb{N}_0$  und  $\kappa \in \mathbb{N}$  nachgewiesen werden. Zunächst gilt

$$\begin{aligned} & E_{\theta}^{\tilde{\mathbb{T}}}\{ {}_T M(x + \kappa) \mid {}_T \mathcal{G}_x \} = {}_T M(x) \\ \Leftrightarrow & E_{\theta}^{\tilde{\mathbb{T}}}\{ {}_T M(x + \kappa) - {}_T M(x) \mid {}_T \mathcal{G}_x \} = 0 \\ \Leftrightarrow & E_{\theta}^{\tilde{\mathbb{T}}}\{ \mathbf{1}_{\mathbb{G}} [ {}_T M(x + \kappa) - {}_T M(x) ] \} = 0, \forall \mathbb{G} \in {}_T \mathcal{G}_x, \end{aligned} \quad (2.48)$$

Des Weiteren lässt sich mithilfe von (2.47) einsehen, dass (2.48) äquivalent ist zu:

$$\begin{aligned} & \frac{1}{\alpha_{\theta}} E_{\theta} \left\{ \mathbf{1}_{\mathbb{G} \cap \tilde{\mathbb{T}}} [ {}_T M(x + \kappa) - {}_T M(x) ] \right\} = 0 \\ \Leftrightarrow & E_{\theta} \left\{ \mathbf{1}_{\mathbb{G} \cap \tilde{\mathbb{T}}} [ {}_T M(x + \kappa) - {}_T M(x) ] \right\} = 0, \forall \mathbb{G} \in {}_T \mathcal{G}_x \end{aligned} \quad (2.49)$$

Da  $\{T \leq x\} \in \mathcal{G}_x \subseteq {}_T \mathcal{G}_x$  erfüllt ist (vgl. (2.39)) und somit auch für das Komplement  $\{T \geq x + 1\} \in {}_T \mathcal{G}_x$  folgt, bilden diese beiden Mengen eine Partition von  ${}_T \mathcal{G}_x$ . Folglich ist es ausreichend, Gleichung (2.49) für die beiden Fälle (a)  $\mathbb{G} \subseteq \{T \leq x\}$  und (b)  $\mathbb{G} \subseteq \{T \geq x + 1\}$  zu zeigen. Mengen, die teilweise in (a) und teilweise in (b) enthalten sind, können in zwei Teilmengen, die sich jeweils einem der Fälle eindeutig zuordnen lassen, zerlegt werden.

(a)  $\mathbb{G} \subseteq \{T \leq x\}$ : Für diesen Fall gilt:

$$\begin{aligned} & E_{\theta} \left\{ \mathbf{1}_{\mathbb{G} \cap \tilde{\mathbb{T}}} [ {}_T M(x + \kappa) - {}_T M(x) ] \right\} = 0, \forall \mathbb{G} \in {}_T \mathcal{G}_x \\ \Leftrightarrow & E_{\theta} \left\{ \mathbf{1}_{\{T \leq x\}} \mathbf{1}_{\mathbb{G} \cap \tilde{\mathbb{T}}} [ {}_T M(x + \kappa) - {}_T M(x) ] \right\} = 0, \forall \mathbb{G} \in {}_T \mathcal{G}_x \end{aligned} \quad (2.50)$$

Zunächst ergibt sich:

$$\tilde{\mathbb{T}} \cap \mathbb{G} = (\tilde{\mathbb{T}} \cap \{T \leq x\}) \cap (\mathbb{G} \cap \{T \leq x\})$$

Da  $\tilde{\mathbb{T}} \in \mathcal{G}_T$  (vgl. Satz 3), ist nach (2.40) der Mengenschnitt  $\tilde{\mathbb{T}} \cap \{T \leq x\} \in \mathcal{G}_x$ . Außerdem gilt  $\mathbb{G} \in {}_T \mathcal{G}_x$  und somit entweder  $\mathbb{G} \in \mathcal{G}_T$  oder  $\mathbb{G} \in \mathcal{G}_x$ . Falls  $\mathbb{G} \in \mathcal{G}_T$ , so ist erneut wegen (2.40) der Mengenschnitt  $\mathbb{G} \cap \{T \leq x\} \in \mathcal{G}_x$ . Falls  $\mathbb{G} \in \mathcal{G}_x$ , so ist ebenfalls  $\mathbb{G} \cap \{T \leq x\} = \mathbb{G} \in \mathcal{G}_x$ . Es folgt  $\tilde{\mathbb{T}} \cap \mathbb{G} \in \mathcal{G}_x$  und somit die Gültigkeit von (2.50), wenn

$$E_{\theta} \left\{ \mathbf{1}_{\{T \leq x\}} [ {}_T M(x + \kappa) - {}_T M(x) ] \mid \mathcal{G}_x \right\} = 0$$

erfüllt ist.

Darüber hinaus gilt für das Ereignis  $\{T \leq x\}$ :

$${}_T M(x + \kappa) - {}_T M(x) = M(x + \kappa) - M(T) - M(x) + M(T) = M(x + \kappa) - M(x)$$

Entsprechend folgt mithilfe der Martingaleigenschaft von  $M$  bezüglich  $\{\mathcal{G}_x : x \in \mathbb{N}_0\}$  und  $P_\theta$ :

$$\mathbb{E}_\theta \left\{ \mathbf{1}_{\{T \leq x\}} \left[ {}_T M(x + \kappa) - {}_T M(x) \right] \middle| \mathcal{G}_x \right\} = \mathbb{E}_\theta \left\{ M(x + \kappa) - M(x) \middle| \mathcal{G}_x \right\} = 0$$

Gleichung (2.49) ist damit für den Fall  $\mathbb{G} \subseteq \{T \leq x\}$  bewiesen.

(b)  $\mathbb{G} \subseteq \{T \geq x + 1\}$ :

Zunächst wird die Menge  $\mathbb{K} := \{T \leq x + \kappa\}$  definiert. Da für das Ereignis  $\mathbb{G}$  gilt, dass  $M(x \wedge T) \equiv M(x)$  ist, kann (2.49) folgendermaßen umgeformt werden:

$$\begin{aligned} & \mathbb{E}_\theta \left\{ \mathbf{1}_{\mathbb{G} \cap \tilde{\mathbb{T}}} \left[ {}_T M(x + \kappa) - {}_T M(x) \right] \right\} = 0 \\ \Leftrightarrow & \mathbb{E}_\theta \left\{ \mathbf{1}_{\mathbb{G} \cap \tilde{\mathbb{T}}} \left[ M(x + \kappa) - M((x + \kappa) \wedge T) - M(x) + M(x \wedge T) \right] \right\} = 0 \\ \Leftrightarrow & \mathbb{E}_\theta \left\{ \mathbf{1}_{\mathbb{G} \cap \tilde{\mathbb{T}}} \left[ M(x + \kappa) - M((x + \kappa) \wedge T) \right] \right\} = 0 \end{aligned}$$

Für das Ereignis  $\Omega \setminus \mathbb{K} = \{T \geq x + \kappa + 1\}$  ist  $M((x + \kappa) \wedge T) = M(x + \kappa)$  und Gleichung (2.49) somit erfüllt. Entsprechend muss lediglich noch die Gleichung

$$\mathbb{E}_\theta \left\{ \mathbf{1}_{\mathbb{K} \cap \mathbb{G} \cap \tilde{\mathbb{T}}} \left[ M(x + \kappa) - M((x + \kappa) \wedge T) \right] \right\} = 0 \quad (2.51)$$

nachgewiesen werden.

Zunächst sei angemerkt, dass  $\mathbb{G} \cap \{T \leq x\} = \emptyset \in \mathcal{G}_x$  und somit  $\mathbb{G} \in \mathcal{G}_T$  erfüllt ist (vgl. (2.40)). Da auch  $\tilde{\mathbb{T}} \in \mathcal{G}_T$ , folgt  $\mathbb{G} \cap \tilde{\mathbb{T}} \in \mathcal{G}_T$  und nach (2.40) weiter  $\mathbb{G} \cap \tilde{\mathbb{T}} \cap \{T \leq x\} \in \mathcal{G}_x$ . Des Weiteren ergibt sich mithilfe der Teilmengenbeziehung

$$\{T \leq x\} \subseteq \{T \leq x + \kappa\} = \mathbb{K}$$

direkt

$$\mathbb{K} \cap \mathbb{G} \cap \tilde{\mathbb{T}} \cap \{T \leq x\} = \mathbb{G} \cap \tilde{\mathbb{T}} \cap \{T \leq x\} \in \mathcal{G}_x$$

und somit  $\mathbb{K} \cap \mathbb{G} \cap \tilde{\mathbb{T}} \in \mathcal{G}_T$ . Außerdem folgt aus  $\mathbb{G} \cap \tilde{\mathbb{T}} \in \mathcal{G}_T$  ebenfalls

$$\mathbb{K} \cap \mathbb{G} \cap \tilde{\mathbb{T}} = (\mathbb{G} \cap \tilde{\mathbb{T}}) \cap \{T \leq x + \kappa\} \in \mathcal{G}_{x+\kappa}$$

Zusammengefasst gilt somit  $\mathbb{K} \cap \mathbb{G} \cap \tilde{\mathbb{T}} \in \mathcal{G}_{x+\kappa} \cap \mathcal{G}_T = \mathcal{G}_{(x+\kappa) \wedge T}$ . Folglich reicht es

$$\mathbb{E}_\theta \left\{ M(x + \kappa) - M((x + \kappa) \wedge T) \middle| \mathcal{G}_{(x+\kappa) \wedge T} \right\} = 0$$

nachzuweisen, um zu zeigen, dass 2.51 gilt.

Nach dem Theorem des optionalen Stoppens (vgl. Appendix, Satz 14) ist  $M(\cdot \wedge T)$

ein  $\mathcal{G}_{x \wedge T}$ -Martingal und es gilt:

$$\begin{aligned} & \mathbb{E}_\theta \left\{ M(x + \kappa) - M((x + \kappa) \wedge T) \middle| \mathcal{G}_{(x + \kappa) \wedge T} \right\} \\ &= \mathbb{E}_\theta \left\{ M(x + \kappa) \middle| \mathcal{G}_{(x + \kappa) \wedge T} \right\} - \mathbb{E}_\theta \left\{ M((x + \kappa) \wedge T) \middle| \mathcal{G}_{(x + \kappa) \wedge T} \right\} \\ &= M((x + \kappa) \wedge T) - M((x + \kappa) \wedge T) = 0 \end{aligned}$$

Gleichung (2.49) ist somit auch für  $\mathbb{G} \subseteq \{T \geq x + 1\}$  bewiesen. Das Prozess  ${}_T A$  ist folglich der  $({}_T \mathcal{G}_x, P_\theta^{\tilde{\mathbb{T}}})$ -Kompensator von  ${}_T N$ . □

Es sei angemerkt, dass aus Satz 5 folgt, dass sich alle  $\Delta_{{}_T A}(x, \theta)$  als die bedingten Wahrscheinlichkeiten der sukzessiven Bernoulli-Experimente  $\Delta_{{}_T N}(x)$ , gegeben  $\tilde{\mathbb{T}}$  sowie der in  ${}_T \mathcal{G}_{x-1}$  enthaltenen Informationen, interpretieren lassen:

$$P_\theta^{\tilde{\mathbb{T}}} \{\Delta_{{}_T N}(x) = 1 \mid {}_T \mathcal{G}_{x-1}\} = \mathbb{E}_\theta^{\tilde{\mathbb{T}}} \{\Delta_{{}_T N}(x) \mid {}_T \mathcal{G}_{x-1}\} = \Delta_{{}_T A}(x, \theta) = \theta {}_T Y(x - 1)$$

Zur Erinnerung, besteht das Ziel dieses Kapitels in der Herleitung einer bedingten Likelihood-Funktion, mithilfe derer eine Schätzung des geometrischen Parameters  $\theta$  unter Berücksichtigung der linksseitigen Trunkierung erfolgen kann. Diese wird sich analog zu  $L(\theta)$  (vgl. (2.17)) mithilfe des Zählprozesses  ${}_T N$  sowie des zugehörigen Kompensators  ${}_T A$  bezüglich einer Filtration und bedingt auf  $\mathbb{1}_{\tilde{\mathbb{T}}} \in \{0, 1\}$  darstellen lassen.

Vorher muss allerdings noch die zugrundeliegende Filtration  $\{{}_T \mathcal{G}_x : x \in \mathbb{N}_0\}$  in Frage gestellt und gegebenenfalls angepasst werden.

### 2.2.3. Anpassung der Filtration

Im Folgenden wird die Filtration  $\{{}_T \mathcal{G}_x : x \in \mathbb{N}_0\}$  so angepasst, dass diese eine auch für praktische Anwendungen geeignete adäquate Beschreibung der Individuumhistorie liefert. Es wird sich nämlich herausstellen, dass in der Praxis nicht alle in  $\{{}_T \mathcal{G}_x : x \in \mathbb{N}_0\}$  enthaltenen Informationen beobachtbar sind. Dies kann mithilfe des folgenden Lemmas eingesehen werden.

**Lemma 3.** *Sei  $\{\mathcal{G}_x : x \in \mathbb{N}_0\}$  eine Filtration, deren Elemente in (2.30) und (2.31) definiert wurden. Sei  $T$  eine  $\mathcal{G}_x$ -Stoppzeit und  $\mathcal{G}_T$  die entsprechende gestoppte  $\sigma$ -Algebra nach (2.40). Sei weiter  $\mathcal{G}_0 = \sigma\{\mathbb{T}_t : t \in \{0, \dots, G - 1\}\} \subseteq \mathcal{G}_x$ . Dann gilt  $\mathcal{G}_0 \subseteq \mathcal{G}_T$ .*

*Beweis.* Es muss gezeigt werden, dass alle erzeugenden Elemente  $\mathbb{T}_t$  mit  $t \in \{0, \dots, G - 1\}$  von  $\mathcal{G}_0$  in  $\mathcal{G}_T$  enthalten sind. Nach (2.40) muss dafür gezeigt werden, dass

$$\mathbb{T}_t \cap \{T \leq x\} \in \mathcal{G}_x, \forall x \in \mathbb{N}_0$$

Dazu sei angemerkt, dass  $\mathbb{T}_t = \{T = t\}$  gilt und wegen (2.30) folglich

$$\mathbb{T}_t \cap \{T \leq x\} = \begin{cases} \mathbb{T}_t \in \mathcal{G}_x, & \text{falls } t \leq x \leq G - 1 \\ \emptyset \in \mathcal{G}_x, & \text{sonst} \end{cases}$$

Somit ist  $\mathcal{G}_0 \subseteq \mathcal{G}_T$  nachgewiesen.  $\square$

Mithilfe von Lemma 3 sowie den Darstellungen (2.41) von  ${}_T\mathcal{G}_x$  und (2.31) von  $\mathcal{G}_x$  folgt:

$${}_T\mathcal{G}_x = \mathcal{G}_x \vee \mathcal{G}_T = \mathcal{F}_x \vee \mathcal{G}_0 \vee \mathcal{G}_T = \mathcal{F}_x \vee \mathcal{G}_T \quad (2.52)$$

Ausgehend von (2.52) kann nun eingesehen werden, dass  $\{{}_T\mathcal{G}_x : x \in \mathbb{N}_0\}$  für praktische Anwendungen ungeeignet ist. Für ein Individuum, das bereits vor dem Beginn der Studie ausfällt und somit linksseitig trunkiert wird (vgl. oberer Pfad in Abbildung 2.2), werden in der Praxis keine Informationen beobachtet und nicht einmal dessen Existenz zu Kenntnis genommen. Jedoch sind auch für ein trunkiertes Individuum alle Informationen bis zum Alter  $T$  in  $\mathcal{G}_T$  und somit in  $\{{}_T\mathcal{G}_x : x \in \mathbb{N}_0\}$  enthalten. Ebenso sind Informationen bezüglich eines Ausfalls (oder Nicht-Ausfalls) vor dem Beginn der Studie in  $\{\mathcal{F}_x : x \in \mathbb{N}_0\}$  und folglich in  $\{{}_T\mathcal{G}_x : x \in \mathbb{N}_0\}$  enthalten, in der Praxis aber nicht beobachtbar. Die Filtration  $\{{}_T\mathcal{G}_x : x \in \mathbb{N}_0\}$  enthält somit in der Praxis nicht beobachtbare Informationen. Zur adäquaten Beschreibung der Historien wird folglich eine gröbere Filtration benötigt. Dazu sei angemerkt, dass die in  $T$  startenden Informationen über die Prozesse  ${}_TN$  und  ${}_TY$  bis zum Zeitpunkt  $x$  gegeben sind durch:

$$\sigma\{{}_TN(k), {}_TY(k) : k \in \{T, \dots, x\}\}$$

Darauf aufbauend, liefert das folgende Lemma eine beobachtbare Filtration zur adäquaten Beschreibung der Historien [siehe Andersen et al., 1988, Kapitel 4, oder Weißbach et al., 2024, Kapitel A.1.2].

**Lemma 4.** *Für die Filtration  $\{{}_T\mathcal{F}_x : T \leq x\}$  mit*

$${}_T\mathcal{F}_x := \sigma\{{}_TN(k), {}_TY(k) : k \in \{T, \dots, x\}\} \quad (2.53)$$

*gilt:*

- (i)  ${}_T\mathcal{F}_x \subseteq {}_T\mathcal{G}_x$  für alle  $x \in \mathbb{N}_0$ .
- (ii) Es gilt  $\tilde{\mathbb{T}} = \{T + 1 \leq X\} \in {}_T\mathcal{F}_x$  für alle  $x \geq T$ .
- (iii) Der Prozess  ${}_TM = {}_TN - {}_TA$  ist ein  $({}_T\mathcal{F}_x, P_{\theta}^{\tilde{\mathbb{T}}})$ -Martingal.
- (iv)  $\{{}_T\mathcal{F}_x : T \leq x\}$  enthält nur beobachtbare Informationen.

*Beweis.* (i) Nach Satz 4 sind die Prozesse  ${}_TN$  und  ${}_TY$  adaptiert bezüglich der Filtration  $\{{}_T\mathcal{G}_x : x \in \mathbb{N}_0\}$ . Somit folgt direkt  ${}_T\mathcal{F}_x \subseteq {}_T\mathcal{G}_x$  für alle  $x \in \mathbb{N}_0$ .

- (ii) Zum Nachweis der Messbarkeit von  $\tilde{\mathbb{T}} = \{T + 1 \leq X\}$  bezüglich  ${}_T\mathcal{F}_x$  für alle  $x \geq T$  genügt es, das kleinste Filtrationselement  ${}_T\mathcal{F}_T$  zu betrachten. Da  ${}_TN(T) = 0$  mit Wahrscheinlichkeit 1 gilt, folgt

$${}_T\mathcal{F}_T = \sigma\{{}_TN(T), {}_TY(T)\} = \sigma\{{}_TY(T)\}$$

Des Weiteren gilt nach der Definition von  ${}_T Y$  (2.43):

$${}_T Y(T) = \mathbb{1}_{\{T+1 \leq T+1 \leq X\}} = \mathbb{1}_{\{T+1 \leq X\}} = \mathbb{1}_{\tilde{\mathbb{T}}}$$

Somit ergibt sich

$${}_T \mathcal{F}_T = \sigma\{\tilde{\mathbb{T}}\} \subset {}_T \mathcal{F}_{T+1} \subset \dots$$

und folglich die Messbarkeit von  $\tilde{\mathbb{T}}$  bezüglich  ${}_T \mathcal{F}_x$  für alle  $x \geq T$ .

(iii) Es wird gezeigt, dass  ${}_T M$  ein  $({}_T \mathcal{F}_x, P_\theta^{\tilde{\mathbb{T}}})$ -Martingal ist. Da

$${}_T M(T) = M(T) - M(T \wedge T) = 0$$

erfüllt ist und für alle  $x \geq T + 1$  gilt, dass sowohl  ${}_T N(x)$  als auch  ${}_T Y(x - 1)$  messbar bezüglich  ${}_T \mathcal{F}_x$  sind, ist wegen (2.45)

$$\begin{aligned} {}_T M(x) &= {}_T N(x) - {}_T A(x, \theta) \\ &= {}_T N(x) - \theta \sum_{k=1}^x {}_T Y(k - 1) \end{aligned}$$

adaptiert bezüglich  ${}_T \mathcal{F}_x$ .

Die Eigenschaft  $\mathbb{E}_\theta^{\tilde{\mathbb{T}}} |{}_T M(x)| < \infty$  für alle  $x \in \mathbb{N}_0$  wurde bereits im Beweis von Satz 5 diskutiert und ist folglich auch erfüllt. Da  ${}_T M$  ein  $({}_T \mathcal{G}_x, P_\theta^{\tilde{\mathbb{T}}})$ -Martingal ist, folgt abschließend die Martingaleigenschaft  $\mathbb{E}_\theta^{\tilde{\mathbb{T}}} \{ {}_T M(x + \kappa) | {}_T \mathcal{F}_x \} = {}_T M(x)$  für alle  $x \geq T$  und  $\kappa \in \mathbb{N}$  mithilfe des in (i) bewiesenen Resultats  ${}_T \mathcal{F}_x \subseteq {}_T \mathcal{G}_x$  sowie unter Zuhilfenahme der Turmregel für bedingte Erwartungswerte (vgl. Satz 13, (ii)):

$$\begin{aligned} \mathbb{E}_\theta^{\tilde{\mathbb{T}}} \{ {}_T M(x + \kappa) | {}_T \mathcal{F}_x \} &= \mathbb{E}_\theta^{\tilde{\mathbb{T}}} \left\{ \mathbb{E}_\theta^{\tilde{\mathbb{T}}} \{ {}_T M(x + \kappa) | {}_T \mathcal{G}_x \} \middle| {}_T \mathcal{F}_x \right\} \\ &= \mathbb{E}_\theta^{\tilde{\mathbb{T}}} \{ {}_T M(x) | {}_T \mathcal{F}_x \} = {}_T M(x) \end{aligned}$$

(iv) Die Eigenschaft der Beobachtbarkeit folgt direkt da  ${}_T N$  sowie  ${}_T Y$  immer beobachtbar sind. □

**Bemerkung 9.** Da es sich bei  ${}_T Y$  um einen  ${}_T \mathcal{F}_x$ -adaptierten Prozess handelt, folgt, dass  ${}_T A$  nach (2.45) ein  ${}_T \mathcal{F}_x$ -vorhersagbarer Prozess ist.

Lemma 4 stellt eine Anwendung des Innovationstheorems [siehe Andersen et al. 1993, Kapitel II.4.2] dar.

Aussage (i) ist dabei in erster Linie eine Bedingung für Aussage (iii). Zusätzlich lässt sich mithilfe von Aussage (ii) zu jedem Zeitpunkt  $x \geq T$  einsehen, ob das Individuum beobachtbar ist oder nicht.

Dabei ist die Teilmengenbeziehung  ${}_T\mathcal{F}_x \subseteq {}_T\mathcal{G}_x$  intuitiv klar, da für ein Individuum, das  $x \vee T$  Jahre alt ist, alle Informationen über die Vergangenheit und die Gegenwart in  ${}_T\mathcal{G}_x$  enthalten sind. Im Gegensatz dazu enthält  ${}_T\mathcal{F}_x$  lediglich einige, aber nicht zwangsläufig alle, Informationen über Vergangenheit und Gegenwart eines  $x \geq T$  Jahre alten Individuums.

Aus Aussage (iii) und Bemerkung 9 folgt wiederum, dass der Prozess  ${}_TA$  der  $({}_T\mathcal{F}_x, P_\theta^{\tilde{\mathbb{T}}})$ -Kompensator von  ${}_TN$  ist.

Somit resultiert:

$$P_\theta^{\tilde{\mathbb{T}}}\{\Delta {}_TN(x) = 1 | {}_T\mathcal{F}_{x-1}\} = E_\theta^{\tilde{\mathbb{T}}}\{\Delta {}_TN(x) | {}_T\mathcal{F}_{x-1}\} = \Delta {}_TA(x, \theta) \quad (2.54)$$

Mithilfe von Aussage (iv) können die  $\Delta {}_TA(x, \theta)$  somit als die Wahrscheinlichkeiten der beobachteten sukzessiven Bernoulli-Experimente  $\Delta {}_TN(x)$  für  $x \geq T$ , gegeben  $\tilde{\mathbb{T}}$  sowie der in  ${}_T\mathcal{F}_{x-1}$  enthaltenen Informationen, interpretiert werden. Darauf aufbauend wird im Folgenden eine entsprechende Beitrag zu einer bedingten Likelihood-Funktion formuliert.

#### 2.2.4. Der Beitrag zur auf $\mathbb{1}_{\tilde{\mathbb{T}}}$ bedingten Likelihood-Funktion

Das folgende Lemma liefert eine Darstellung des Beitrags eines Individuums zur auf  $\mathbb{1}_{\tilde{\mathbb{T}}}$  bedingten Likelihood-Funktion, mithilfe derer eine Schätzung des geometrischen Parameters  $\theta$  unter Berücksichtigung der Linkstrunkierung vollzogen werden kann. Dabei werden beide Fälle  $\mathbb{1}_{\tilde{\mathbb{T}}} \in \{0, 1\}$  mit  $\tilde{\mathbb{T}} = \{T + 1 \leq X\}$  untersucht, sodass zwischen sichtbaren und unsichtbaren Individuen unterschieden werden kann. Das Resultat kann erneut als eine zeitdiskrete Version einer Formel von Jacod [siehe Andersen et al., 1993, Kapitel II.7.1] interpretiert werden.

**Lemma 5.** *Der bedingte Likelihood-Beitrag eines Individuums, das zum Zeitpunkt des Ausfalls  $X$  Jahre alt ist, erst ab dem Alter von  $T + 1$  sowie bedingt auf dem Ereignis  $\tilde{\mathbb{T}} = \{T + 1 \leq X\}$  beobachtet werden kann und bis zu einem beliebigen, aber festen Zeitpunkt  $\chi \geq G$  beobachtet wird, ist:*

$$\begin{aligned} {}_TL(\theta) &= 1, \text{ wenn } \mathbb{1}_{\tilde{\mathbb{T}}} = 0 \\ {}_TL(\theta) &= \prod_{x=T+1}^{\chi} \left\{ (1 - \Delta {}_TA(x, \theta))^{1 - \Delta {}_TN(x)} (\Delta {}_TA(x, \theta))^{\Delta {}_TN(x)} \right\}, \text{ wenn } \mathbb{1}_{\tilde{\mathbb{T}}} = 1 \end{aligned} \quad (2.55)$$

Dabei ist

$$\begin{aligned} \Delta {}_TN(x) &= N(x) - N(x \wedge T) - N(x - 1) + N((x - 1) \wedge T) \\ &= \mathbb{1}_{\{T+1 \leq X\}} \mathbb{1}_{\{X=x\}} \\ &= \mathbb{1}_{\tilde{\mathbb{T}}} \mathbb{1}_{\{X=x\}} \end{aligned}$$

sowie

$$\Delta {}_TA(x, \theta) = \theta {}_TY(x - 1) = \theta \mathbb{1}_{\{T+1 \leq x \leq X\}}.$$

*Beweis.* Es soll der auf  $\mathbb{1}_{\{\tilde{T}\}} \in \{0, 1\}$  bedingte Beitrag zu einer Likelihood-Funktion ermittelt werden, mithilfe derer die Wahrscheinlichkeit für das Eintreten der beobachteten Ergebnisse maximiert werden kann.

Wird das Leben eines Individuums unter Berücksichtigung der Linkstrunkierung bis zu einem Zeitpunkt  $\chi \in \mathbb{N}$  mit  $\chi \geq G$  untersucht, lassen sich die Realisierungen der Prozesse  ${}_T N$  und  ${}_T Y$  ab  $T$  als eben diese beobachteten Ergebnisse interpretieren.

Gilt für ein  $x \in \{T + 1, \dots, \chi\}$ , dass  ${}_T N(x) = 1$ , so folgt  $\Delta {}_T N(k) = 1$  für ein  $k \in \{T + 1, \dots, x\}$ . Außerdem impliziert  $\Delta {}_T N(x) = 1$ , dass  ${}_T N(x) = 1$  und  ${}_T N(x - 1) = 0$  gelten muss. Folglich lässt sich jedem Verlauf von  ${}_T N$  eindeutig ein Verlauf von  $\Delta {}_T N$  zuordnen und umgekehrt. Der Prozess  $\Delta {}_T N$  enthält somit dieselben Informationen wie  ${}_T N$ .

Gilt zudem  ${}_T Y(x - 1) = 0$ , so folgt direkt  ${}_T Y(x) = 0$ . Darüber hinaus ergibt sich aus  ${}_T Y(x - 1) = 1$  und  ${}_T N(x) = 0$  direkt  ${}_T Y(x) = 1$ . Ebenso impliziert  ${}_T Y(x - 1) = 1$  und  ${}_T N(x) = 1$ , dass  ${}_T Y(x) = 0$  gelten muss. Der Wert von  ${}_T Y(x)$  lässt sich somit eindeutig über die Werte von  ${}_T Y(x - 1)$  und  ${}_T N(x)$  bestimmen.

Zusammenfassend sind die beobachtbaren Informationen in der Matrix

$${}_T J := \begin{pmatrix} \Delta {}_T N(T + 1) & \Delta {}_T N(T + 2) & \dots & \Delta {}_T N(\chi) \\ {}_T Y(T) & {}_T Y(T + 1) & \dots & {}_T Y(\chi - 1) \end{pmatrix}' \quad (2.56)$$

mit der zugehörigen bedingten Wahrscheinlichkeitsfunktion  $f^{TJ|\mathbb{1}_{\{\tilde{T}\}}}$  codiert [siehe Gouriéroux und Monfort, 1995, Definition 7.2 (ii)].

Die Beitrag zur bedingten Likelihood-Funktion  ${}_T L$  ergibt sich dann als:

$$\begin{aligned} {}_T L(\theta) &= f^{TJ|\mathbb{1}_{\{\tilde{T}\}}}(TJ|\mathbb{1}_{\{\tilde{T}\}}) \\ &= f^{\Delta {}_T N(T+1), \dots, \Delta {}_T N(\chi), {}_T Y(T), \dots, {}_T Y(\chi-1)|\mathbb{1}_{\{\tilde{T}\}}} \\ &\quad \left( \Delta {}_T N(T + 1), \dots, \Delta {}_T N(\chi), {}_T Y(T), \dots, {}_T Y(\chi - 1) \middle| \mathbb{1}_{\{\tilde{T}\}} \right) \end{aligned} \quad (2.57)$$

Zunächst gilt für den Fall  $\mathbb{1}_{\{\tilde{T}\}} = 0$ , dass für alle  $x \in \{T + 1, \dots, \chi\}$  mit Wahrscheinlichkeit 1 gilt, dass:

$$\Delta {}_T N(x) = {}_T Y(x - 1) = 0$$

Somit folgt für diesen Fall direkt  ${}_T L \equiv 1$ .

Entsprechend muss lediglich der Fall  $\mathbb{1}_{\{\tilde{T}\}} = 1$  weiter untersucht werden. Dabei ist für ein  $x \in \{T + 1, \dots, \chi\}$  die Zufallsvariable  $\Delta {}_T N(x)$  abhängig von  $\Delta {}_T N(T + 1), \dots, \Delta {}_T N(x - 1)$  sowie  ${}_T Y(T), \dots, {}_T Y(x - 1)$ . Unter der Berücksichtigung dieser Abhängigkeit lässt sich die bedingte Likelihood-Funktion nach (2.57) folgendermaßen umformulieren:

$$\begin{aligned}
{}_T L(\theta) &= f^{T\Delta N(x) | T\Delta N(x-1), \dots, T\Delta N(T+1), T Y(x-1), \dots, T Y(T), \mathbb{1}_{\tilde{T}}=1} \\
&\quad \left( \Delta {}_T N(x) \middle| \Delta {}_T N(x-1), \dots, \Delta {}_T N(T+1), T Y(x-1), \dots, T Y(T), \mathbb{1}_{\tilde{T}}=1 \right) \\
&\quad \cdot f^{T Y(x-1) | T\Delta N(x-1), \dots, T\Delta N(T+1), T Y(x-2), \dots, T Y(T), \mathbb{1}_{\tilde{T}}=1} \\
&\quad \left( T Y(x-1) \middle| \Delta {}_T N(x-1), \dots, \Delta {}_T N(T+1), T Y(x-2), \dots, T Y(T), \mathbb{1}_{\tilde{T}}=1 \right) \\
&\quad \cdot f^{T\Delta N(x-1) | T\Delta N(x-2), \dots, T\Delta N(T+1), T Y(x-2), \dots, T Y(T), \mathbb{1}_{\tilde{T}}=1} \\
&\quad \left( \Delta {}_T N(x-1) \middle| \Delta {}_T N(x-2), \dots, \Delta {}_T N(T+1), T Y(x-2), \dots, T Y(T), \mathbb{1}_{\tilde{T}}=1 \right) \\
&\quad \cdot \dots \cdot f^{\Delta {}_T N(T+1) | T Y(T), \mathbb{1}_{\tilde{T}}=1} \left( \Delta {}_T N(T+1) \middle| T Y(T), \mathbb{1}_{\tilde{T}}=1 \right) \\
&\quad \cdot f^{T Y(T) | \mathbb{1}_{\tilde{T}}=1} \left( T Y(T) \middle| \mathbb{1}_{\tilde{T}}=1 \right) \tag{2.58}
\end{aligned}$$

Außerdem gilt für alle  $x \in \{T, \dots, \chi-1\}$ , dass  ${}_T Y(x)$  eindeutig durch die Ausprägungen von  $\Delta {}_T N(x)$  und  ${}_T Y(x-1)$  determiniert ist, sodass:

$$\begin{aligned}
& f^{T Y(x) | T\Delta N(x), \dots, T\Delta N(T+1), T Y(x-1), \dots, T Y(T), \mathbb{1}_{\tilde{T}}=1} \\
& \left( T Y(x) \middle| \Delta {}_T N(x), \dots, \Delta {}_T N(T+1), T Y(x-1), \dots, T Y(T), \mathbb{1}_{\tilde{T}}=1 \right) \\
& = f^{T Y(x) | T\Delta N(x), T Y(x-1)} \left( T Y(x) \middle| \Delta {}_T N(x), T Y(x-1) \right) = 1
\end{aligned}$$

Außerdem gilt

$$f^{T Y(T) | \mathbb{1}_{\tilde{T}}=1} \left( T Y(T) \middle| \mathbb{1}_{\tilde{T}}=1 \right) = f^{T Y(T) | \mathbb{1}_{\tilde{T}}=1} \left( T Y(T) \middle| T Y(T) = 1 \right) = 1,$$

sodass (2.58) folgendermaßen vereinfacht werden kann:

$$\begin{aligned}
{}_T L(\theta) &= f^{T\Delta N(x) | T\Delta N(x-1), \dots, T\Delta N(T+1), T Y(x-1), \dots, T Y(T), \mathbb{1}_{\tilde{T}}=1} \\
&\quad \left( \Delta {}_T N(x) \middle| \Delta {}_T N(x-1), \dots, \Delta {}_T N(T+1), T Y(x-1), \dots, T Y(T), \mathbb{1}_{\tilde{T}}=1 \right) \\
&\quad \cdot f^{T\Delta N(x-1) | T\Delta N(x-2), \dots, T\Delta N(T+1), T Y(x-2), \dots, T Y(T), \mathbb{1}_{\tilde{T}}=1} \\
&\quad \left( \Delta {}_T N(x-1) \middle| \Delta {}_T N(x-2), \dots, \Delta {}_T N(T+1), T Y(x-2), \dots, T Y(T), \mathbb{1}_{\tilde{T}}=1 \right) \\
&\quad \cdot \dots \cdot f^{\Delta {}_T N(T+1) | T Y(T), \mathbb{1}_{\tilde{T}}=1} \left( \Delta {}_T N(T+1) \middle| T Y(T), \mathbb{1}_{\tilde{T}}=1 \right) \tag{2.59}
\end{aligned}$$

Im nächsten Schritt wird ein Faktor von (2.59) für einen beliebigen Zeitpunkt  $x \in \{T+1, \dots, \chi\}$  betrachtet:

$$\begin{aligned}
& f^{T\Delta N(x) | T\Delta N(x-1), \dots, T\Delta N(T+1), T Y(x-1), \dots, T Y(T), \mathbb{1}_{\tilde{T}}=1} \\
& \left( \Delta {}_T N(x) \middle| \Delta {}_T N(x-1), \dots, \Delta {}_T N(T+1), T Y(x-1), \dots, T Y(T), \mathbb{1}_{\tilde{T}}=1 \right) \tag{2.60}
\end{aligned}$$

Es sei angemerkt, dass es sich bei den Informationen über  $\Delta {}_T N(x-1), \dots, T\Delta N(1)$  sowie über  ${}_T Y(x-1), \dots, T Y(1)$  genau um die Informationen handelt, die in der  $\sigma$ -Algebra  ${}_T \mathcal{F}_{x-1}$  enthalten sind (vgl. (2.53)). Außerdem ermöglicht die zusätzliche Bedingung auf

$\mathbb{1}_{\tilde{\tau}} = 1$  den Übergang zum bedingten Wahrscheinlichkeitsmaß  $P_{\theta}^{\tilde{\tau}}$ . Somit gilt weiter:

$$\begin{aligned} & f_{T\Delta N(x) | T\Delta N(x-1), \dots, T\Delta N(T+1), TY(x-1), \dots, TY(T), \mathbb{1}_{\tilde{\tau}}=1} \\ & \left( \Delta_{TN}(x) \mid \Delta_{TN}(x-1), \dots, \Delta_{TN}(T+1), TY(x-1), \dots, TY(T), \mathbb{1}_{\tilde{\tau}} = 1 \right) \\ & = P_{\theta}^{\tilde{\tau}} \{ \Delta_{TN}(x) = 1 \mid T\mathcal{F}_{x-1} \}^{\Delta_{TN}(x)} \cdot P_{\theta}^{\tilde{\tau}} \{ \Delta_{TN}(x) = 0 \mid T\mathcal{F}_{x-1} \}^{1-\Delta_{TN}(x)} \\ & = P_{\theta}^{\tilde{\tau}} \{ \Delta_{TN}(x) = 1 \mid T\mathcal{F}_{x-1} \}^{\Delta_{TN}(x)} \cdot \left( 1 - P_{\theta}^{\tilde{\tau}} \{ \Delta_{TN}(x) = 1 \mid T\mathcal{F}_{x-1} \} \right)^{1-\Delta_{TN}(x)} \end{aligned}$$

Darüber hinaus ergibt sich aus der Eigenschaft (A.9) von bedingten Erwartungswerten sowie aus  $\Delta_{TN}(x) = \mathbb{1}_{\{\Delta_{TN}(x)=1\}}$ :

$$\begin{aligned} & P_{\theta}^{\tilde{\tau}} \{ \Delta_{TN}(x) = 1 \mid T\mathcal{F}_{x-1} \}^{\Delta_{TN}(x)} \cdot \left( 1 - P_{\theta}^{\tilde{\tau}} \{ \Delta_{TN}(x) = 1 \mid T\mathcal{F}_{x-1} \} \right)^{1-\Delta_{TN}(x)} \\ & = \left( E_{\theta}^{\tilde{\tau}} \{ \mathbb{1}_{\{\Delta_{TN}(x)=1\}} \mid T\mathcal{F}_{x-1} \} \right)^{\Delta_{TN}(x)} \left( 1 - E_{\theta}^{\tilde{\tau}} \{ \mathbb{1}_{\{\Delta_{TN}(x)=1\}} \mid T\mathcal{F}_{x-1} \} \right)^{1-\Delta_{TN}(x)} \\ & = \left( E_{\theta}^{\tilde{\tau}} \{ \Delta_{TN}(x) \mid T\mathcal{F}_{x-1} \} \right)^{\Delta_{TN}(x)} \left( 1 - E_{\theta}^{\tilde{\tau}} \{ \Delta_{TN}(x) \mid T\mathcal{F}_{x-1} \} \right)^{1-\Delta_{TN}(x)} \end{aligned}$$

Da der Prozess  ${}_TA$  der Kompensator von  ${}_TN$  bezüglich der Filtration  $\{T\mathcal{F}_x : T \leq x\}$  sowie dem bedingten Wahrscheinlichkeitsmaß  $P_{\theta}^{\tilde{\tau}}$  ist, folgt weiter unter der Verwendung von (2.54):

$$\begin{aligned} & \left( E_{\theta}^{\tilde{\tau}} \{ \Delta_{TN}(x) \mid T\mathcal{F}_{x-1} \} \right)^{\Delta_{TN}(x)} \left( 1 - E_{\theta}^{\tilde{\tau}} \{ \Delta_{TN}(x) \mid T\mathcal{F}_{x-1} \} \right)^{1-\Delta_{TN}(x)} \\ & = \Delta_{TA}(x, \theta)^{\Delta_{TN}(x)} \left( 1 - \Delta_{TA}(x, \theta) \right)^{1-\Delta_{TN}(x)} \end{aligned} \quad (2.61)$$

Abschließend folgt durch Einsetzen von (2.61) in (2.58):

$${}_TL(\theta) = \prod_{x=T+1}^{\chi} \left\{ (1 - \Delta_{TA}(x, \theta))^{1-\Delta_{TN}(x)} (\Delta_{TA}(x, \theta))^{\Delta_{TN}(x)} \right\}$$

Der Beweis ist somit abgeschlossen.  $\square$

Im letzten Schritt wird die allgemeine Darstellung des Beitrag zur bedingten Likelihood-Funktion (2.55) konkretisiert. Folgt die Zufallsvariable  $X$  einer geometrischen Verteilung mit dem Parameter  $\theta$ , so lässt sich der in Lemma 5 hergeleitete bedingte Likelihood-Beitrag weiter vereinfachen zu:

$$\begin{aligned} {}_TL(\theta) & = \prod_{x=T+1}^{\chi} \left\{ (1 - \mathbb{1}_{\{T+1 \leq x \leq X\}} \theta)^{1 - \mathbb{1}_{\{T+1 \leq x\}} \mathbb{1}_{\{X=x\}}} (\mathbb{1}_{\{T+1 \leq x \leq X\}} \theta)^{\mathbb{1}_{\{T+1 \leq x\}} \mathbb{1}_{\{X=x\}}} \right\} \\ & = \left[ \prod_{x=T+1}^{\min(X-1, \chi)} (1 - \theta)^1 \theta^0 \right] \cdot \left[ (1 - \theta)^0 \theta^1 \right]^{\mathbb{1}_{\{X \leq \chi\}}} \\ & = \theta^{\mathbb{1}_{\{X \leq \chi\}}} (1 - \theta)^{\min(X-1, \chi) - T} \end{aligned}$$

Damit ist mit der linksseitigen Trunkierung die erste Einschränkung der Beobachtbarkeit in das Modell implementiert worden. Der bedingte Likelihood-Beitrag ähnelt zwar

der Wahrscheinlichkeitsfunktion der geometrischen Verteilung (2.5). Jedoch werden für ein Individuum, dessen Ausfall beobachtbar und dessen gesamte Lebensdauer sowie dessen Anzahl an Jahren ohne Ausfall folglich bekannt ist, alle Jahre vor dem Beginn der Studie und somit die ersten  $T$  Jahre ohne Ausfall "vergessen".

Im folgenden Kapitel wird das Design um ein Studienende ergänzt. Damit einhergehend erfolgt eine zusätzliche Erweiterung des bisherigen Modells um die so genannte Rechtszensur.

### 2.3. Rechtszensierte Historienteile

Im folgenden Kapitel soll das in den Kapiteln 2.1 und 2.2 erstellte Modell um eine weitere Einschränkung hinsichtlich der Beobachtbarkeit ergänzt werden.

In Kapitel 2.1 wurde zunächst davon ausgegangen, dass der Beobachtungszeitraum weder einen Anfang noch ein Ende besitzt und dieser somit den gesamten Zeitstrahl umfasst. Entsprechend konnte für jedes im Entstehungszeitraum  $\{a, \dots, b\}$  mit  $G = |\{a, \dots, b\}|$  entstandene Individuum sowohl die Existenz als auch die genaue Lebensdauer in Jahren  $X$  beobachtet werden.

Im nächsten Schritt wurde in Kapitel 2.2 für den Beobachtungszeitraum ein eindeutiger Anfang im Jahr  $b+1$  definiert. Fiel ein Individuum in den Jahren  $\{b+1, b+2, \dots\}$  aus, so konnten sowohl Existenz als auch die Lebensdauer  $X$  sowie das Alter ein Jahr vor dem Beginn der Studie  $T$  beobachtet werden. Bei einem Ausfall im Entstehungszeitraum und somit vor dem Beginn des Beobachtungszeitraums war die Beobachtung von  $X$  sowie  $T$  und sogar der Existenz des Individuums allerdings nicht mehr möglich. Folglich lag der linkstrunkierte Fall vor.

Als letzter Schritt im Rahmen der Modellentwicklung soll in diesem Kapitel der Beobachtungszeitraum zusätzlich um ein eindeutiges Ende erweitert werden. Dazu bezeichne allgemein  $s \in \mathbb{N}$  die Länge der Studie, sodass der Beobachtungszeitraum die Jahre  $\{b+1, \dots, b+s\}$  umfasst.

Entsprechend gibt es nun drei mögliche Fälle hinsichtlich der Beobachtbarkeit, die in Abbildung 2.4 illustriert werden.

Fällt ein Individuum vor beziehungsweise während des Beobachtungszeitraums aus, so ist es linksseitig trunkiert (oberer Pfad in Abbildung 2.4) beziehungsweise beobachtbar (mittlerer Pfad in Abbildung 2.4). Diese beiden Fälle wurden bereits in den Kapiteln 2.1 und 2.2 ausführlich diskutiert.

Findet der Ausfall des Individuums allerdings nach dem Jahr  $b+s$  und somit nach dem Ende der Studie statt, so wird zwar die Existenz des Individuums zur Kenntnis genommen, die genaue Lebensdauer kann allerdings nicht mehr beobachtet werden (unterer Pfad in Abbildung 2.4). Bekannt sind dann lediglich das Alter des Individuums ein Jahr vor dem Beginn der Studie  $T$  und somit auch das Alter  $T+s$  im Jahr  $b+s$  sowie die Information, dass der Ausfall am Ende des Beobachtungszeitraums noch nicht stattgefunden hat. Dieses Phänomen wird im Folgenden als rechtsseitige Zensur beziehungsweise Rechtszensur bezeichnet.

Das Ziel dieses Kapitels wird, analog zu den vorherigen Kapiteln 2.1 und 2.2, die Herleitung des Beitrags eines Individuums zu einer marginalen, bedingten Likelihood-

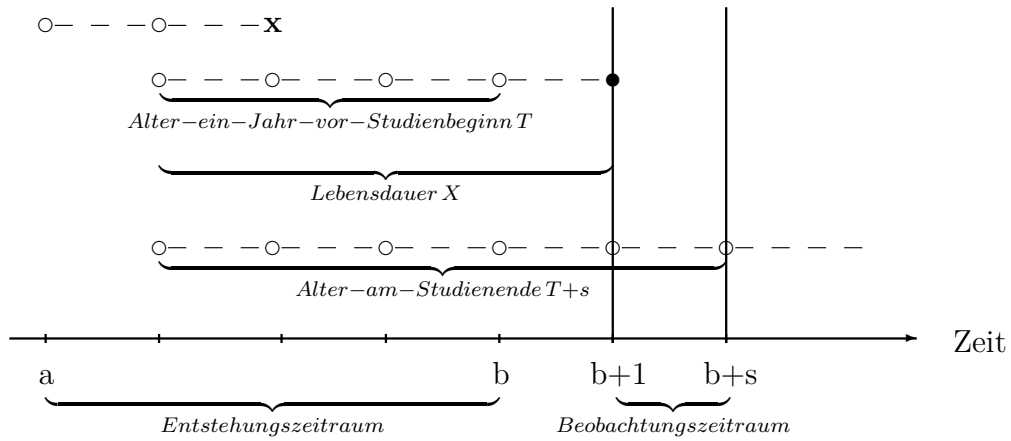


Abbildung 2.4.: Mögliche Pfade hinsichtlich der Beobachtbarkeit für einen beispielhaften Entstehungszeitraum der Länge  $G = |\{a, \dots, b\}| = 5$  sowie einen Beobachtungszeitraum der Länge  $s = 2$ ; die Vertikalen an den Zeitpunkten  $b + 1$  und  $b + s$  begrenzen den Beobachtungszeitraum; der obere und der mittlere Pfad repräsentieren ein trunkiertes beziehungsweise ein vollständig beobachtetes Individuum (vgl. Abb 2.2); (iii) der untere Pfad beschreibt ein Individuum, das in  $a + 1$  entstanden, zum Zeitpunkt  $b + s$  jedoch noch nicht ausgefallen ist. Beobachtet wird das Entstehungsjahr und das Überleben bis einschließlich  $b + s$ , jedoch nicht die genaue Lebensdauer.

Funktion zur Schätzung des geometrischen Parameters  $\theta$  unter Berücksichtigung von Linkstrunkierung sowie Rechtszensur sein. Das weitere Vorgehen basiert dabei auf Andersen et al. [1988, Kapitel 3] sowie Weißbach et al. [2024, Kapitel A.1.2].

Als Grundlage der mathematischen Betrachtungen kann erneut der Wahrscheinlichkeitsraum  $(\Omega, \mathcal{F}, P_\theta)$  verwendet werden, wobei die Elemente der Ergebnismenge  $\omega \in \Omega$  in (2.1) definiert wurden,  $\mathcal{F} = \mathcal{P}(\Omega)$  gilt und  $P_\theta$  im Vergleich zu (2.4) ebenfalls erhalten bleibt. Des Weiteren seien erneut  $X$  das Alter zum Zeitpunkt des Ausfalls respektive die Lebensdauer (vgl. (2.2)) und  $T$  das Alter ein Jahr vor dem Beginn des Beobachtungszeitraums (vgl. (2.3)).

Zusammenfassend unterliegt ein Individuum der Linkstrunkierung und wird folglich nicht beobachtet, wenn  $X \leq T$ . Im Falle von  $T + 1 \leq X \leq T + s$  liegt hingegen die Beobachtbarkeit des in  $T$  codierten Entstehungsjahres sowie der Lebensdauer  $X$  vor. Für  $T + s + 1 \leq X$  wird das Entstehungsjahr sowie das Überleben für die ersten  $T + s$  Jahre dokumentiert. Nach den ersten  $T + s$  Jahren wird die Historie des Individuums jedoch zensuriert und keine weiteren Informationen, inklusive der über die exakte Lebensdauer, mehr beobachtet. Eine beispielhafte Zuordnung der  $\omega \in \Omega$  zu den verschiedenen Beobachtungsszenarien für den Fall  $G = 5$  und  $s = 2$  lässt sich Abbildung 2.5 entnehmen.

Analog zu vorherigen Kapiteln wird sich der marginale, bedingte Likelihood-Beitrag mithilfe eines Zählprozesses und des zugehörigen Kompensators bezüglich einer Filtration darstellen lassen.

Die Frage nach einer geeigneten Filtration sollte dabei einleitend geklärt werden.

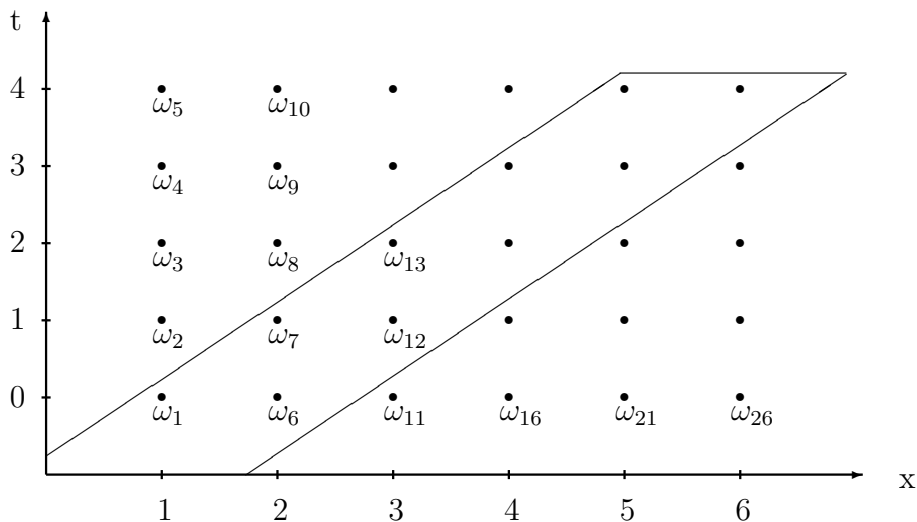


Abbildung 2.5.: Zuordnung der Elementarereignisse zu einem Entstehungs- und Ausfalljahr bei einem  $G = 5$  jährigen Entstehungszeitraum sowie  $s = 2$  jährigen Beobachtungszeitraum inklusive Visualisierung der Beobachtbarkeit: (i) Alle  $\omega$  oberhalb des Parallelogramms repräsentieren ein Ereignis mit  $X \leq T$ , in dem folglich ein Individuum trunziert und entsprechend nicht beobachtet wird (vgl. Abb (2.3)); (ii) Für alle  $\omega$  innerhalb des Parallelogramms liegt die Beobachtbarkeit von Entstehungs- und Ausfalljahr vor. Für  $\omega_6 = \omega_{(2-1)5+(0+1)}$  ist beispielsweise  $X = 2$ ,  $T = 0$  und  $T + s = 2$ , sodass  $T + 1 \leq X \leq T + s$  erfüllt ist; (iii) Alle  $\omega$  unterhalb des Parallelogramms repräsentieren ein Ereignis, in dem die Historie eines Individuums zensiert und folglich nur für die ersten  $T + s$  Jahre beobachtet wird. Für  $\omega_{11} = \omega_{(3-1)5+(0+1)}$  ist beispielsweise  $X = 3$ ,  $T = 0$  und  $T + s = 2$  und somit  $T + s + 1 \leq X$  erfüllt.

In Kapitel 2.2 wurde zunächst in (2.41) die Filtration  $\{ {}_T\mathcal{G}_x : x \in \mathbb{N}_0 \}$  mit  ${}_T\mathcal{G}_x = \mathcal{G}_T \vee \mathcal{G}_x$  eingeführt, wobei  $\mathcal{G}_T$  in (2.40) und  $\{ \mathcal{G}_x : x \in \mathbb{N}_0 \}$  in (2.30) definiert wurden. Diese enthielt sowohl die Information bezüglich der Lebensdauer  $X$  als auch hinsichtlich des Alters ein Jahr vor dem Beginn der Studie  $T$ . Außerdem war zu jedem Zeitpunkt erkennbar, ob das Ereignis  $\tilde{T} = \{ T + 1 \leq X \}$  eingetreten und somit die Beobachtbarkeit des Individuums gegeben ist oder nicht.

Für folgende Untersuchungen muss zusätzlich für jeden Zeitpunkt  $x \in \mathbb{N}_0$  unterschieden werden können, ob das Ende des Beobachtungszeitraums bereits erreicht wurde oder nicht. Folglich muss für alle  $x \in \mathbb{N}_0$  das Ereignis  $\{ x \leq T + s \}$  und somit das entsprechende Komplement  $\{ T + s + 1 \leq x \}$  messbar bezüglich des jeweiligen Filtrationselements sein.

Jedoch handelt es sich bei  $T$  um eine  $\mathcal{G}_x$ -Stoppzeit und wegen  $\mathcal{G}_x \subseteq {}_T\mathcal{G}_x$  für alle  $x \in \mathbb{N}_0$  auch um eine  ${}_T\mathcal{G}_x$ -Stoppzeit (vgl. Definition 6), sodass nach Bemerkung 7

$$\{ T \leq x \} \subseteq {}_T\mathcal{G}_x$$

erfüllt ist. Da die Studiendauer  $s \in \mathbb{N}$  und somit auch  $s + 1$  deterministisch ist, ist die Summe  $T + s + 1$  nach Bauer [1974, Kapitel XI., §58.] ebenfalls eine  ${}_T\mathcal{G}_x$ -Stopzeit. Folglich gilt:

$$\{T + s + 1 \leq x\} \in {}_T\mathcal{G}_x \text{ sowie } \{x \leq T + s\} \in {}_T\mathcal{G}_x, \forall x \in \mathbb{N}_0$$

Die Filtration  $\{{}_T\mathcal{G}_x : x \in \mathbb{N}_0\}$  enthält somit alle benötigten Informationen und kann als Ausgangslage für weitere Betrachtungen verwendet werden.

In den nächsten Schritten werden bezüglich  $\{{}_T\mathcal{G}_x : x \in \mathbb{N}_0\}$  ein geeigneter Zählprozess sowie der zugehörige Kompensator definiert und anschließend die Filtration so angepasst, dass diese nur noch beobachtbare Informationen enthält.

Zu diesem Zweck wird zunächst der Prozess  $C$  mit

$$C(x) := \mathbb{1}_{\{x \leq T+s\}} \quad (2.62)$$

eingeführt, um festzustellen, ob zum Zeitpunkt  $x$  das Ende des Beobachtungszeitraums bereits überschritten wurde.

**Bemerkung 10.** *Es gilt genau dann  $C(x) = 1$ , wenn für das zugrundeliegende Ereignis  $\omega_{(x-1)G+t+1} \in \Omega$  gilt, dass  $x-s \leq t$ . Somit ist mithilfe der Definition von  $\mathbb{T}_t$  (vgl. (2.29)):*

$$C(x) = 1, \forall \omega \in \bigcup_{\max(0, x-s) \leq t \leq G-1} \mathbb{T}_t \quad (2.63)$$

Da nach Lemma 3 für alle  $x \in \mathbb{N}_0$  gilt, dass

$$\mathcal{G}_0 = \sigma\{\mathbb{T}_t, t \in \{0, \dots, G-1\}\} \subset \mathcal{G}_T \subseteq {}_T\mathcal{G}_x,$$

ist der Prozess  $C$  adaptiert bezüglich  $\{{}_T\mathcal{G}_x : x \in \mathbb{N}_0\}$ .

Folglich kann unter Hinzunahme der Rechtszensur für ein Individuum zum Zeitpunkt  $x$  der Wert des linksseitig trunkierten Zählprozesses  ${}_TN(x)$  genau dann beobachtet werden, wenn  $C(x) = 1$ . Somit lässt sich, unabhängig davon, ob ein Individuum trunkiert, zensiert oder beobachtet wird, der Prozess  ${}_TN^c$  [siehe Weißbach et al., 2024, Kapitel A.1.2] mit

$${}_TN^c(x) := \sum_{k=1}^x C(k) \Delta {}_TN(k) \quad (2.64)$$

als beobachtbar interpretieren.

Da  $C$  nach Bemerkung 10 und  ${}_TN$  nach Satz 4 adaptiert bezüglich der Filtration  $\{{}_T\mathcal{G}_x : x \in \mathbb{N}_0\}$  sind, handelt es sich folglich bei  ${}_TN^c$  ebenfalls um einen  $\{{}_T\mathcal{G}_x : x \in \mathbb{N}_0\}$ -

adaptierten Prozess. Außerdem gilt:

$$\begin{aligned}
 {}_T N^c(x) &= \sum_{k=1}^x \mathbb{1}_{\{k \leq T+s\}} \mathbb{1}_{\{T+1 \leq X\}} \mathbb{1}_{\{X=k\}} \\
 &= \mathbb{1}_{\{T+1 \leq X\}} \sum_{k=1}^x \mathbb{1}_{\{k \leq T+s\}} \mathbb{1}_{\{X=k\}} \\
 &= \mathbb{1}_{\{T+1 \leq X\}} \mathbb{1}_{\{X \leq \min(x, T+s)\}} \\
 &= \mathbb{1}_{\{T+1 \leq X \leq \min(x, T+s)\}}
 \end{aligned}$$

Folglich ist  ${}_T N^c(0) = 0$  sowie  ${}_T N^c(x) \leq 1 < \infty$  für alle  $x \in \mathbb{N}$ . Außerdem besitzt  ${}_T N^c$  maximal einen Sprung der Höhe 1. Damit ist  ${}_T N^c$  nach Definition 3 ein diskreter  ${}_T \mathcal{G}_x$ -Zählprozess.

Somit kann im nächsten Schritt bezüglich  $\{{}_T \mathcal{G}_x : x \in \mathbb{N}_0\}$  und  $P_\theta^{\tilde{\mathbb{T}}}$  eine Dekomposition von  ${}_T N^c$  in ein Martingal  ${}_T M^c$  und einen Kompensator  ${}_T A^c$  erfolgen.

Zu diesem Zweck liefert das folgende Lemma den  ${}_T \mathcal{G}_x$ -Kompensator  ${}_T A^c$  von  ${}_T N^c$  [siehe Andersen et al., 1993, Kapitel III.2.2].

**Lemma 6.** *Für den Prozess  ${}_T A^c$  mit*

$${}_T A^c(x, \theta) = \sum_{k=1}^x C(k) \Delta {}_T A(k, \theta) \tag{2.65}$$

*gilt:*

- (i)  ${}_T A^c$  ist  ${}_T \mathcal{G}_x$ -vorhersagbar.
- (ii) Der Prozess  ${}_T M^c = {}_T N^c - {}_T A^c$  ist ein Martingal bezüglich der Filtration  $\{{}_T \mathcal{G}_x : x \in \mathbb{N}_0\}$  sowie des bedingten Wahrscheinlichkeitsmaßes  $P_\theta^{\tilde{\mathbb{T}}}$  (vgl. (2.46)).

Somit ist  ${}_T A^c$  der  $({}_T \mathcal{G}_x, P_\theta^{\tilde{\mathbb{T}}})$ -Kompensator von  ${}_T N^c$ .

*Beweis.* Für den Nachweis von (i) sei angemerkt, dass  ${}_T A$  der  $({}_T \mathcal{G}_x, P_\theta^{\tilde{\mathbb{T}}})$ -Kompensator von  ${}_T N$  und somit  ${}_T \mathcal{G}_x$ -vorhersagbar ist. Folglich muss nur noch die  ${}_T \mathcal{G}_x$ -Vorhersagbarkeit von  $C$  nachgewiesen werden.

Zunächst ist bekannt, dass es sich bei der Zufallsvariablen  $T$  um eine  ${}_T \mathcal{G}_x$ -Stopppzeit handelt. Folglich ist dann auch  $T + s$  eine  ${}_T \mathcal{G}_x$ -Stopppzeit, da  $s$  konstant ist. Somit kann zu jedem Zeitpunkt  $x$  eingesehen werden, ob das Ereignis  $\{x = T + s\}$  eingetreten ist oder nicht. Darauf aufbauend kann nun eine Fallunterscheidung vorgenommen werden.

- $x \leq T + s - 1$ : Dann ist zunächst  $C(x) = 1$ . Da zusätzlich bekannt ist, dass das Ereignis  $\{x = T + s\}$  nicht eingetreten ist, folgt direkt  $C(x + 1) = 1$ .
- $x = T + s$ : Dann ist zunächst  $C(x) = 1$ . Da zusätzlich bekannt ist, dass das Ereignis  $\{x = T + s\}$  eingetreten ist, folgt direkt  $C(x + 1) = 0$ .
- $x \geq T + s + 1$ : Dann ist  $C(x) = 0$  und somit auch  $C(x + 1) = 0$ .

Zusammenfassend kann in jedem Fall der Wert von  $C(x+1)$  eindeutig zum Zeitpunkt  $x$  ermittelt werden. Folglich ist  $C$  und somit auch  ${}_T A^c$  ein  ${}_T \mathcal{G}_x$ -vorhersagbarer Prozess.

Für den Nachweis von (ii) folgt zunächst die  ${}_T \mathcal{G}_x$ -Adaption von  ${}_T M^c$  direkt, da es sich bei  ${}_T N$ ,  ${}_T A$  und  $C$  um  ${}_T \mathcal{G}_x$ -adaptierte Prozesse handelt.

Außerdem ist der Prozess  ${}_T M^c$  integrierbar, da für alle  $x \in \mathbb{N}_0$  gilt, dass

$$\begin{aligned} \mathbb{E}_{\tilde{\theta}}^{\mathbb{T}} |{}_T M^c(x)| &= \frac{1}{\alpha_{\theta}} \mathbb{E}_{\theta} |1_{\tilde{\mathbb{T}}} {}_T M^c(x)| \\ &\leq \frac{1}{\alpha_{\theta}} \mathbb{E}_{\theta} \left\{ \sum_{k=1}^x |C(k) [\Delta {}_T N(k) - \Delta {}_T A(k, \theta)]| \right\} \\ &< \frac{1}{\alpha_{\theta}} \mathbb{E}_{\theta} \left\{ \sum_{k=1}^x |C(k)| \right\} \leq \frac{1}{\alpha_{\theta}} \min(x, T+s) < \infty \end{aligned} \quad (2.66)$$

Abschließend wird noch die Martingaleigenschaft  $\mathbb{E}_{\tilde{\theta}}^{\mathbb{T}} \{ {}_T M^c(x+\kappa) | {}_T \mathcal{G}_x \} = {}_T M^c(x)$  für alle  $x \in \mathbb{N}_0$  und  $\kappa \in \mathbb{N}$  gezeigt:

$$\begin{aligned} \mathbb{E}_{\tilde{\theta}}^{\mathbb{T}} \{ {}_T M^c(x+\kappa) | {}_T \mathcal{G}_x \} &= \mathbb{E}_{\tilde{\theta}}^{\mathbb{T}} \left\{ \sum_{k=1}^{x+\kappa} C(k) [\Delta {}_T N(k) - \Delta {}_T A(k, \theta)] \middle| {}_T \mathcal{G}_x \right\} \\ &= \mathbb{E}_{\tilde{\theta}}^{\mathbb{T}} \left\{ \sum_{k=1}^x C(k) [\Delta {}_T N(k) - \Delta {}_T A(k, \theta)] \middle| {}_T \mathcal{G}_x \right\} \\ &\quad + \mathbb{E}_{\tilde{\theta}}^{\mathbb{T}} \left\{ \sum_{k=x+1}^{x+\kappa} C(k) [\Delta {}_T N(k) - \Delta {}_T A(k, \theta)] \middle| {}_T \mathcal{G}_x \right\} \\ &= \sum_{k=1}^x C(k) [\Delta {}_T N(k) - \Delta {}_T A(k, \theta)] \\ &\quad + \mathbb{E}_{\tilde{\theta}}^{\mathbb{T}} \left\{ \sum_{k=x+1}^{x+\kappa} C(k) [\Delta {}_T N(k) - \Delta {}_T A(k, \theta)] \middle| {}_T \mathcal{G}_x \right\} \\ &= {}_T M^c(x) + \sum_{k=x+1}^{x+\kappa} \mathbb{E}_{\tilde{\theta}}^{\mathbb{T}} \left\{ C(k) [\Delta {}_T N(k) - \Delta {}_T A(k, \theta)] \middle| {}_T \mathcal{G}_x \right\} \end{aligned}$$

Folglich reicht es zu zeigen, dass

$$\begin{aligned} \mathbb{E}_{\tilde{\theta}}^{\mathbb{T}} \left\{ C(k) [\Delta {}_T N(k) - \Delta {}_T A(k, \theta)] \middle| {}_T \mathcal{G}_x \right\} &= 0, \forall k \in \{x+1, \dots, x+\kappa\} \\ \Leftrightarrow \mathbb{E}_{\tilde{\theta}}^{\mathbb{T}} \left\{ 1_{\mathbb{G}} C(k) [\Delta {}_T N(k) - \Delta {}_T A(k, \theta)] \right\} &= 0, \forall k \in \{x+1, \dots, x+\kappa\}, \forall \mathbb{G} \in {}_T \mathcal{G}_x \end{aligned} \quad (2.67)$$

erfüllt ist.

Des Weiteren sei angemerkt, dass  $T$  eine  $\mathcal{G}_x$ -Stopppzeit ist und somit für alle  $k \in \{x+1, \dots, x+\kappa\}$  für das Ereignis

$$\Omega \setminus \{k \leq T+s\} = \{T+s+1 \leq k\} = \{T \leq k-s-1\}$$

gilt, dass:

$$\{T \leq k - s - 1\} \cap \{T \leq x\} = \{T \leq \min(k - s - 1, x)\} \in \mathcal{G}_{(k-s-1) \wedge x} \subseteq \mathcal{G}_x, \forall x \in \mathbb{N}_0$$

Nach der Definition der gestoppten  $\sigma$ -Algebra  $\mathcal{G}_T$  (vgl. (2.40)) ist somit:

$$\{T \leq k - s - 1\} \in \mathcal{G}_T \subseteq {}_T\mathcal{G}_x, \forall x \in \mathbb{N}_0$$

Aufgrund der Stabilität von  $\sigma$ -Algebren bezüglich Komplementen ist somit auch  $\{k \leq T + s\} \in {}_T\mathcal{G}_x$ . Entsprechend kann jedes  $\mathbb{G} \in {}_T\mathcal{G}_x$  folgendermaßen dargestellt werden:

$$\mathbb{G} = \mathbb{G} \cap (\{T \leq k - s - 1\} \cup \{k \leq T + s\}) = (\mathbb{G} \cap \{T \leq k - s - 1\}) \cup (\mathbb{G} \cap \{k \leq T + s\})$$

Folglich ist (2.67) äquivalent zu:

$$\begin{aligned} & \mathbb{E}_{\theta}^{\tilde{\mathbb{T}}} \left\{ \mathbb{1}_{\mathbb{G} \cap \{T+s+1 \leq k\}} C(k) [\Delta {}_T N(k) - \Delta {}_T A(k, \theta)] \right\} \\ & + \mathbb{E}_{\theta}^{\tilde{\mathbb{T}}} \left\{ \mathbb{1}_{\mathbb{G} \cap \{k \leq T+s\}} C(k) [\Delta {}_T N(k) - \Delta {}_T A(k, \theta)] \right\} = 0 \end{aligned} \quad (2.68)$$

Für den Fall  $T + s + 1 \leq k$  ist  $C(k) = 0$ . Außerdem gilt  $C(k) = 1$  für  $k \leq T + s$ . Folglich ist (2.68) wegen  $\mathbb{G} \in {}_T\mathcal{G}_x$  äquivalent zu:

$$\begin{aligned} & \mathbb{E}_{\theta}^{\tilde{\mathbb{T}}} \left\{ \mathbb{1}_{\{k \leq T+s\}} C(k) [\Delta {}_T N(k) - \Delta {}_T A(k, \theta)] \Big| {}_T\mathcal{G}_x \right\} \\ & = \mathbb{E}_{\theta}^{\tilde{\mathbb{T}}} \left\{ \Delta {}_T N(k) - \Delta {}_T A(k, \theta) \Big| {}_T\mathcal{G}_x \right\} \\ & = \mathbb{E}_{\theta}^{\tilde{\mathbb{T}}} \left\{ \Delta {}_T M(k) \Big| {}_T\mathcal{G}_x \right\} = 0, \forall k \in \{x + 1, \dots, x + \kappa\} \end{aligned} \quad (2.69)$$

Da es sich bei dem Prozess  ${}_T M$  nach Satz 5 um ein Martingal bezüglich der Filtration  $\{{}_T\mathcal{G}_x : x \in \mathbb{N}_0\}$  sowie des bedingten Wahrscheinlichkeitsmaßes  $P_{\theta}^{\tilde{\mathbb{T}}}$  handelt, ist (2.69) erfüllt. Folglich ist die Gültigkeit von (2.67) gezeigt und somit das Lemma bewiesen.  $\square$

**Bemerkung 11.** *Unabhängig davon, ob ein Individuum trunkiert, zensiert oder beobachtet wird, kann der beobachtbare Unter-Risiko-Prozess [siehe Andersen et al., 1993, Kapitel III.2.2] definiert werden als*

$${}_T Y^c(x) := C(x + 1) {}_T Y(x) \quad (2.70)$$

$$= \mathbb{1}_{\{x+1 \leq T+s\}} \mathbb{1}_{\{T+1 \leq x+1 \leq X\}} = \mathbb{1}_{\{T+1 \leq x+1 \leq \min(X, T+s)\}} \quad (2.71)$$

Mithilfe von (2.70) kann der Kompensator  ${}_T A^c$  (vgl. (2.65)) folgendermaßen dargestellt werden:

$${}_T A^c(x, \theta) = \sum_{k=1}^x C(k) \Delta {}_T A(k, \theta) = \theta \sum_{k=1}^x C(k) {}_T Y(k - 1) = \theta \sum_{k=1}^x {}_T Y^c(k - 1) \quad (2.72)$$

Wie bereits im vorherigen Kapitel 2.2 diskutiert, kann auf der Grundlage von  ${}_T\mathcal{G}_x =$

$\mathcal{F}_x \vee \mathcal{G}_T$  (vgl. (2.52)) eingesehen werden, dass die Filtration  $\{\mathcal{G}_x, x \in \mathbb{N}_0\}$  Informationen enthält, die in der Praxis nicht beobachtet werden können. So sind auch für  $x \leq T$  alle Informationen bis zum Alter  $T$  in  $\mathcal{G}_T$  und somit in  $\mathcal{G}_x$  enthalten. Auch lassen sich Informationen bezüglich eines Ausfalls vor dem Studienbeginn aus der Filtration  $\{\mathcal{F}_x, x \in \mathbb{N}_0\}$  und folglich auch aus  $\{\mathcal{G}_x, x \in \mathbb{N}_0\}$  entnehmen.

Die gröbere, in (2.53) eingeführte Filtration  $\{\mathcal{F}_x, x \in \mathbb{N}_0\}$  ist unter der Berücksichtigung der Rechtszensur zur adäquaten Beschreibung der Historien ebenfalls ungeeignet. So sind Informationen bezüglich eines Ausfalls nach dem Ende des Beobachtungszeitraums zwar in  $\{\mathcal{F}_x, x \in \mathbb{N}_0\}$  enthalten, in der Praxis jedoch nicht mehr beobachtbar.

Das folgende Lemma liefert eine geeignete Filtration zur Beschreibung der Historie des Individuums [siehe Weißbach et al., 2024, Kapitel A.1.2].

**Lemma 7.** *Für die Filtration  $\{\mathcal{F}_x^c, T \leq x\}$  mit*

$${}_T\mathcal{F}_x^c := \sigma\{{}_TN^c(k), {}TY^c(k), k \in \{T, \dots, x\}\} \quad (2.73)$$

*gilt.*

- (i)  ${}_T\mathcal{F}_x^c \subseteq {}_T\mathcal{G}_x$  für alle  $x \in \mathbb{N}_0$ .
- (ii) Das Ereignis  $\tilde{\mathbb{T}}$  ist für alle  $x \geq T$  messbar bezüglich  ${}_T\mathcal{F}_x^c$ .
- (iii) Der Prozess  ${}_TM = {}_TN - {}_TA$  ist bezüglich dem bedingten Wahrscheinlichkeitsmaß  $P_\theta^{\tilde{\mathbb{T}}}$  ein  ${}_T\mathcal{F}_x^c$ -Martingal.
- (iv)  $\{\mathcal{F}_x^c, T \leq x\}$  enthält nur beobachtbare Informationen.

*Beweis Lemma 7.* (i) Nach Satz 4 sind die Prozesse  ${}_TN$  und  ${}_TY$  adaptiert bezüglich der Filtration  $\{\mathcal{G}_x : x \in \mathbb{N}_0\}$ . Außerdem ist nach Bemerkung 10 der Prozess  $C$  ebenfalls  ${}_T\mathcal{G}_x$ -adaptiert. Entsprechend sind  ${}_TN^c$  sowie  ${}_TY^c$  adaptiert bezüglich  $\{\mathcal{G}_x : x \in \mathbb{N}_0\}$  und es folgt direkt  ${}_T\mathcal{F}_x^c \subseteq {}_T\mathcal{G}_x$  für alle  $x \geq T$ .

- (ii) Zum Nachweis der Messbarkeit von  $\tilde{\mathbb{T}} = \{T+1 \leq X\}$  bezüglich  ${}_T\mathcal{F}_x^c$  für alle  $x \geq T$  genügt es, das kleinste Filtrationselement  ${}_T\mathcal{F}_T^c$  zu betrachten. Da  ${}_TN(T)^c = 0$  mit Wahrscheinlichkeit 1 gilt, folgt

$${}_T\mathcal{F}_T^c = \sigma\{{}_TN^c(T), {}TY(T)\} = \sigma\{{}_TY^c(T)\}$$

Des Weiteren gilt nach der Definition von  ${}_TY^c$  (vgl. 2.71):

$${}_TY^c(T) = \mathbb{1}_{\{T+1 \leq T+1 \leq \min(X, T+s)\}} = \mathbb{1}_{\{T+1 \leq \min(X, T+s)\}} = \mathbb{1}_{\tilde{\mathbb{T}}}$$

Da  $s \in \mathbb{N}$  folgt  $\mathbb{1}_{\{T+1 \leq T+s\}} = 1$ , sodass sich

$${}_TY^c(T) = \mathbb{1}_{\tilde{\mathbb{T}}}$$

und entsprechend

$${}_T\mathcal{F}_T^c = \sigma\{\tilde{\mathbb{T}}\} \subset {}_T\mathcal{F}_{T+1}^c \subset \dots$$

ergibt. Somit resultiert die Messbarkeit von  $\tilde{\mathbb{T}}$  bezüglich  ${}_T\mathcal{F}_x^c$  für alle  $x \geq T$ .

(iii) Zunächst ist der Prozess  ${}_TM^c$  mit

$${}_TM^c(x) = {}_TN^c(x) - {}_TA^c(x, \theta) = {}_TN^c(x) - \theta \sum_{k=1}^x {}_TY^c(k-1)$$

offensichtlich  ${}_T\mathcal{F}_x^c$ -adaptiert. Außerdem ist nach (2.66) auch  $E_{\tilde{\theta}}^{\tilde{\mathbb{T}}}|{}_TM^c(x)| < \infty$  für alle  $x \in \mathbb{N}_0$  und somit auch für alle  $x \geq T$  erfüllt. Da  ${}_TM^c$  nach Lemma 6 ein  ${}_T\mathcal{G}_x$ -Martingal ist und  ${}_T\mathcal{F}_x^c \subseteq {}_T\mathcal{G}_x$  gilt, folgt abschließend die Martingaleigenschaft

$$\begin{aligned} E_{\tilde{\theta}}^{\tilde{\mathbb{T}}}\{{}_TM^c(x+\kappa)|{}_T\mathcal{F}_x^c\} &= E_{\tilde{\theta}}^{\tilde{\mathbb{T}}}\{E_{\tilde{\theta}}^{\tilde{\mathbb{T}}}\{{}_TM^c(x+\kappa)|{}_T\mathcal{G}_x\}|{}_T\mathcal{F}_x^c\} \\ &= E_{\tilde{\theta}}^{\tilde{\mathbb{T}}}\{{}_TM^c(x)|{}_T\mathcal{F}_x^c\} = {}_TM^c(x) \end{aligned}$$

für alle  $x \geq T$  und  $\kappa \in \mathbb{N}$ .

(iv) Da die Prozesse  ${}_TN^c$  und  ${}_TY^c$  immer beobachtbar sind, enthält  $\{{}_T\mathcal{F}_x^c, T \leq x\}$  nur beobachtbare Informationen. □

**Bemerkung 12.** Da  ${}_TY^c$  messbar bezüglich  $\{{}_T\mathcal{F}_x^c, T \leq x\}$  ist, folgt direkt die  ${}_T\mathcal{F}_{x-1}^c$ -Messbarkeit von  ${}_TY^c(x-1)$ . Der Prozess  ${}_TA^c$  ist folglich  ${}_T\mathcal{F}_x^c$ -vorhersagbar und kann somit als  $({}_T\mathcal{F}_x^c, P_{\tilde{\theta}}^{\tilde{\mathbb{T}}})$ -Kompensator von  ${}_TN^c$  bezeichnet werden.

Lemma 7 stellt erneut eine Anwendung des Innovationstheorems [siehe Andersen et al. 1993, Kapitel II.4.2] dar. Aussage (i) ist dabei eine notwendige Bedingung für Aussage (iii). Außerdem impliziert Aussage (ii), dass für jeden Zeitpunkt  $x \geq T$  bekannt ist, ob das Individuum beobachtbar ist.

Da wegen Aussage (iii) und Bemerkung 12 der Prozess  ${}_TA^c$  der  $({}_T\mathcal{F}_x^c, P_{\tilde{\theta}}^{\tilde{\mathbb{T}}})$ -Kompensator von  ${}_TN^c$  ist, folgt:

$$P_{\tilde{\theta}}^{\tilde{\mathbb{T}}}\{\Delta {}_TN^c(x) = 1 | {}_T\mathcal{F}_{x-1}^c\} = E_{\tilde{\theta}}^{\tilde{\mathbb{T}}}\{\Delta {}_TN^c(x) | {}_T\mathcal{F}_{x-1}^c\} = \Delta {}_TA^c(x, \theta) \quad (2.74)$$

Aussage (iv) impliziert, dass die  $\Delta {}_TA^c(x, \theta)$  somit als die Wahrscheinlichkeiten der beobachteten sukzessiven Bernoulli-Experimente  $\Delta {}_TN^c(x)$  für  $x \geq T$ , gegeben  $\tilde{\mathbb{T}}$  sowie der in  ${}_T\mathcal{F}_{x-1}^c$  enthaltenen Informationen, interpretiert werden können. Darauf aufbauend wird im Folgenden der Beitrag eines Individuums zu einer marginalen, bedingten Likelihood-Funktion nach dem Vorbild von Jacod's Formeln [siehe Andersen et al., 1993, Kapitel II.7.1] aufgestellt.

**Lemma 8.** Gegeben sei ein Individuum, das zum Zeitpunkt des Ausfalls  $X$  Jahre alt ist, erst ab dem Alter von  $T+1$  sowie bedingt auf dem Ereignis  $\tilde{\mathbb{T}} = \{T+1 \leq X\}$  beobachtet werden kann und dessen Historie nach  $T+s$  Jahren zensiert wird, wobei  $s \in \mathbb{N}$ . Dann ist der Beitrag dieses Individuums zur marginalen, bedingten Likelihood-Funktion bis zu

einem beliebigen, aber festen, Zeitpunkt  $\chi \geq G$  gegeben durch:

$$\begin{aligned} {}_T L^c(\theta) &= 1, \text{ wenn } \mathbb{1}_{\tilde{\tau}} = 0 \\ {}_T L^c(\theta) &= \prod_{x=T+1}^{\min(\chi, T+s)} \left\{ (1 - \Delta {}_T A^c(x, \theta))^{1 - \Delta {}_T N^c(x)} (\Delta {}_T A^c(x, \theta))^{\Delta {}_T N^c(x)} \right\} \text{ wenn } \mathbb{1}_{\tilde{\tau}} = 1 \end{aligned} \quad (2.75)$$

Dabei ist

$$\Delta {}_T N^c(x) = C(x) \Delta {}_T N(x) = \mathbb{1}_{\{x \leq T+s\}} \mathbb{1}_{\{T+1 \leq x\}} \mathbb{1}_{\{X=x\}} = \mathbb{1}_{\{T+1 \leq x \leq T+s\}} \mathbb{1}_{\{X=x\}} \quad (2.76)$$

sowie

$$\Delta {}_T A^c(x, \theta) = \theta C(x) {}_T Y(x-1) = \theta \mathbb{1}_{\{T+1 \leq x \leq \min(X, T+s)\}}$$

*Beweis.* Analog zum Beweis von Lemma 5 soll der Beitrag eines Individuum zu einer marginalen, bedingten Likelihood-Funktion ermittelt werden, mithilfe derer die Wahrscheinlichkeit für die beobachteten Ergebnisse maximiert werden kann.

Dabei sind die beobachtbaren Informationen in der Matrix

$${}_T \tilde{J}^c := \begin{pmatrix} \Delta {}_T N^c(T+1) & \Delta {}_T N^c(T+2) & \dots & \Delta {}_T N^c(\chi) \\ {}_T Y^c(T) & {}_T Y^c(T+1) & \dots & {}_T Y^c(\chi-1) \end{pmatrix}'$$

mit der zugehörigen bedingten Wahrscheinlichkeitsfunktion  $f^{T \tilde{J}^c | \mathbb{1}_{\tilde{\tau}}}$  codiert. Da jedoch für alle  $x \geq T+s+1$  gilt, dass sowohl  $\Delta {}_T N^c(x) = 0$  als auch  ${}_T Y^c(x-1) = 0$  mit Wahrscheinlichkeit 1 erfüllt sind, genügt es, die Matrix

$${}_T J^c := \begin{pmatrix} \Delta {}_T N^c(T+1) & \Delta {}_T N^c(T+2) & \dots & \Delta {}_T N^c(\min(\chi, T+s)) \\ {}_T Y^c(T) & {}_T Y^c(T+1) & \dots & {}_T Y^c(\min(\chi, T+s)-1) \end{pmatrix}' \quad (2.77)$$

mit der entsprechenden marginalen, bedingten Wahrscheinlichkeitsfunktion  $f^{T J^c | \mathbb{1}_{\tilde{\tau}}}$  für weitere Untersuchungen zu verwenden [siehe Gouriéroux und Monfort, 1995, Definition 7.2 (i) und (ii)]. Für den marginalen, bedingten Likelihood-Beitrag  ${}_T L^c$  ergibt sich zunächst:

$$\begin{aligned} {}_T L^c(\theta) &= f^{T J^c | \mathbb{1}_{\tilde{\tau}}}({}_T J | \mathbb{1}_{\tilde{\tau}}) \\ &= f^{\Delta {}_T N^c(T+1), \dots, \Delta {}_T N^c(\min(\chi, T+s)), {}_T Y^c(T), \dots, {}_T Y^c(\min(\chi, T+s)-1) | \mathbb{1}_{\tilde{\tau}}} \\ &\quad \left( \Delta {}_T N^c(T+1), \dots, \Delta {}_T N^c(\min(\chi, T+s)), {}_T Y^c(T), \dots, {}_T Y^c(\min(\chi, T+s)-1) \middle| \mathbb{1}_{\tilde{\tau}} \right) \end{aligned} \quad (2.78)$$

Dabei gilt, im Falle von  $\mathbb{1}_{\tilde{\tau}} = 0$ , dass  ${}_T N^c \equiv {}_T Y^c \equiv 0$  mit Wahrscheinlichkeit 1 und somit  ${}_T L^c \equiv 1$ . Folglich genügt es, lediglich den Fall  $\mathbb{1}_{\tilde{\tau}} = 1$  weiter zu betrachten.

Zunächst ergibt sich aufgrund der Abhängigkeit von  $\Delta {}_T N^c(x)$  bezüglich  $\Delta {}_T N^c(T+$

$1), \dots, \Delta_T N^c(x-1)$  sowie  ${}_T Y^c(T), \dots, {}_T Y^c(x-1)$  für alle  $x \in \{T+1, \dots, \min(\chi, T+s)\}$  folgende Umformung von (2.78):

$$\begin{aligned}
{}_T L^c(\theta) &= f^{T\Delta N^c(\min(\chi, T+s)) | T\Delta N^c(\min(\chi, T+s)-1), \dots, T\Delta N^c(T+1), {}_T Y^c(\min(\chi, T+s)-1), \dots, {}_T Y^c(T), \mathbb{1}_{\tilde{\mathbb{T}}} = 1} \\
&\quad \left( \Delta_T N^c(\min(\chi, T+s)) \middle| \Delta_T N^c(\min(\chi, T+s)-1), \dots, {}_T Y^c(T), \mathbb{1}_{\tilde{\mathbb{T}}} = 1 \right) \\
&\quad \cdot f^{T Y^c(\min(\chi, T+s)-1) | T\Delta N^c(\min(\chi, T+s)-1), \dots, T\Delta N^c(T+1), {}_T Y^c(\min(\chi, T+s)-2), \dots, {}_T Y^c(T), \mathbb{1}_{\tilde{\mathbb{T}}} = 1} \\
&\quad \left( {}_T Y^c(\min(\chi, T+s)-1) \middle| \Delta_T N^c(\min(\chi, T+s)-1), \dots, {}_T Y^c(T), \mathbb{1}_{\tilde{\mathbb{T}}} = 1 \right) \\
&\quad \cdot \dots \cdot f^{\Delta_T N^c(T+1) | {}_T Y^c(T), \mathbb{1}_{\tilde{\mathbb{T}}} = 1} \left( \Delta_T N^c(T+1) \middle| {}_T Y(T), \mathbb{1}_{\tilde{\mathbb{T}}} = 1 \right) \\
&\quad \cdot f^{T Y^c(T) | \mathbb{1}_{\tilde{\mathbb{T}}} = 1} \left( {}_T Y^c(T) \middle| \mathbb{1}_{\tilde{\mathbb{T}}} = 1 \right) \tag{2.79}
\end{aligned}$$

Außerdem gilt für alle  $x \in \{T, \dots, \min(\chi, T+s)-1\}$ , dass  ${}_T Y^c(x)$  eindeutig durch die Ausprägungen von  $\Delta_T N^c(x)$  und  ${}_T Y^c(x-1)$  determiniert ist, sodass:

$$\begin{aligned}
& f^{T Y^c(x) | T\Delta N^c(x), \dots, T\Delta N^c(T+1), {}_T Y^c(x-1), \dots, {}_T Y^c(T), \mathbb{1}_{\tilde{\mathbb{T}}} = 1} \\
& \left( {}_T Y^c(x) \middle| \Delta_T N^c(x), \dots, \Delta_T N^c(T+1), {}_T Y^c(x-1), \dots, {}_T Y^c(T), \mathbb{1}_{\tilde{\mathbb{T}}} = 1 \right) \\
& = f^{T Y^c(x) | T\Delta N^c(x), {}_T Y^c(x-1)} \left( {}_T Y^c(x) \middle| \Delta_T N^c(x), {}_T Y^c(x-1) \right) = 1
\end{aligned}$$

Insbesondere gilt

$$f^{T Y(T) | \mathbb{1}_{\tilde{\mathbb{T}}} = 1} \left( {}_T Y(T) \middle| \mathbb{1}_{\tilde{\mathbb{T}}} = 1 \right) = f^{T Y(T) | \mathbb{1}_{\tilde{\mathbb{T}}} = 1} \left( {}_T Y(T) \middle| {}_T Y(T) = 1 \right) = 1$$

sodass (2.79) folgendermaßen vereinfacht werden kann:

$$\begin{aligned}
{}_T L^c(\theta) &= f^{T\Delta N^c(\min(\chi, T+s)) | T\Delta N^c(\min(\chi, T+s)-1), \dots, T\Delta N^c(T+1), {}_T Y^c(\min(\chi, T+s)-1), \dots, {}_T Y^c(T), \mathbb{1}_{\tilde{\mathbb{T}}} = 1} \\
&\quad \left( \Delta_T N^c(\min(\chi, T+s)) \middle| \Delta_T N^c(\min(\chi, T+s)-1), \dots, {}_T Y^c(T), \mathbb{1}_{\tilde{\mathbb{T}}} = 1 \right) \\
&\quad \cdot f^{T\Delta N^c(\min(\chi, T+s)-1) | T\Delta N^c(\min(\chi, T+s)-2), \dots, T\Delta N^c(T+1), {}_T Y^c(\min(\chi, T+s)-2), \dots, {}_T Y^c(T), \mathbb{1}_{\tilde{\mathbb{T}}} = 1} \\
&\quad \left( \Delta_T N^c(\min(\chi, T+s)-1) \middle| \Delta_T N^c(\min(\chi, T+s)-2), \dots, {}_T Y^c(T), \mathbb{1}_{\tilde{\mathbb{T}}} = 1 \right) \\
&\quad \cdot \dots \cdot f^{\Delta N^c(T+1) | {}_T Y^c(T), \mathbb{1}_{\tilde{\mathbb{T}}} = 1} \left( \Delta N^c(T+1) \middle| {}_T Y^c(T), \mathbb{1}_{\tilde{\mathbb{T}}} = 1 \right) \tag{2.80}
\end{aligned}$$

Wird ein Faktor von (2.80) für einen beliebigen Zeitpunkt  $x \in \{T+1, \dots, \min(\chi, T+s)\}$  betrachtet, ergibt sich:

$$\begin{aligned}
& f^{T\Delta N^c(x) | T\Delta N^c(x-1), \dots, T\Delta N^c(T+1), {}_T Y^c(x-1), \dots, {}_T Y^c(T), \mathbb{1}_{\tilde{\mathbb{T}}} = 1} \\
& \left( \Delta_T N^c(x) \middle| \Delta_T N^c(x-1), \dots, \Delta_T N^c(T+1), {}_T Y^c(x-1), \dots, {}_T Y^c(T), \mathbb{1}_{\tilde{\mathbb{T}}} = 1 \right) \tag{2.81}
\end{aligned}$$

Es sei angemerkt, dass es sich bei den Informationen über  $\Delta_T N^c(x-1), \dots, T\Delta N^c(1)$  sowie über  $\Delta_T Y^c(x-1), \dots, T\Delta Y^c(1)$  genau um die Informationen handelt, die in der  $\sigma$ -Algebra  ${}_T \mathcal{F}_{x-1}^c$  enthalten sind (vgl. (2.73)). Außerdem ermöglicht die zusätzliche Bedingung auf  $\mathbb{1}_{\tilde{\mathbb{T}}} = 1$  den Übergang zum bedingten Wahrscheinlichkeitsmaß  $P_{\theta}^{\tilde{\mathbb{T}}}$ . Somit

folgt:

$$\begin{aligned}
& f_{T\Delta N^c(x) | T\Delta N^c(x-1), \dots, T\Delta N^c(T+1), TY^c(x-1), \dots, TY^c(T), \mathbb{1}_{\tilde{T}=1}} \\
& \left( \Delta_T N^c(x) \middle| \Delta_T N^c(x-1), \dots, \Delta_T N^c(T+1), TY^c(x-1), \dots, TY^c(T), \mathbb{1}_{\tilde{T}=1} \right) \\
& = P_{\theta}^{\tilde{T}} \{ \Delta_T N^c(x) = 1 | T\mathcal{F}_{x-1}^c \}^{\Delta_T N^c(x)} \cdot P_{\theta}^{\tilde{T}} \{ \Delta_T N^c(x) = 0 | T\mathcal{F}_{x-1}^c \}^{1-\Delta_T N^c(x)} \\
& = P_{\theta}^{\tilde{T}} \{ \Delta_T N^c(x) = 1 | T\mathcal{F}_{x-1}^c \}^{\Delta_T N^c(x)} \cdot \left( 1 - P_{\theta}^{\tilde{T}} \{ \Delta_T N^c(x) = 1 | T\mathcal{F}_{x-1}^c \} \right)^{1-\Delta_T N^c(x)}
\end{aligned}$$

Darüber hinaus ergibt sich aus der Eigenschaft (A.9) von bedingten Erwartungswerten sowie aus  $\Delta_T N^c(x) = \mathbb{1}_{\{\Delta_T N^c(x)=1\}}$ :

$$\begin{aligned}
& P_{\theta}^{\tilde{T}} \{ \Delta_T N^c(x) = 1 | T\mathcal{F}_{x-1}^c \}^{\Delta_T N^c(x)} \cdot \left( 1 - P_{\theta}^{\tilde{T}} \{ \Delta_T N^c(x) = 1 | T\mathcal{F}_{x-1}^c \} \right)^{1-\Delta_T N^c(x)} \\
& = \left( E_{\theta}^{\tilde{T}} \{ \mathbb{1}_{\{\Delta_T N^c(x)=1\}} | T\mathcal{F}_{x-1}^c \} \right)^{\Delta_T N^c(x)} \left( 1 - E_{\theta}^{\tilde{T}} \{ \mathbb{1}_{\{\Delta_T N^c(x)=1\}} | T\mathcal{F}_{x-1}^c \} \right)^{1-\Delta_T N^c(x)} \\
& = \left( E_{\theta}^{\tilde{T}} \{ \Delta_T N^c(x) | T\mathcal{F}_{x-1}^c \} \right)^{\Delta_T N^c(x)} \left( 1 - E_{\theta}^{\tilde{T}} \{ \Delta_T N^c(x) | T\mathcal{F}_{x-1}^c \} \right)^{1-\Delta_T N^c(x)}
\end{aligned}$$

Da der Prozess  $T A^c$  der Kompensator von  $T N^c$  bezüglich der Filtration  $\{T\mathcal{F}_x^c : T \leq x\}$  sowie dem bedingten Wahrscheinlichkeitsmaß  $P_{\theta}^{\tilde{T}}$  ist, folgt weiter unter der Verwendung von (2.74):

$$\begin{aligned}
& \left( E_{\theta}^{\tilde{T}} \{ \Delta_T N^c(x) | T\mathcal{F}_{x-1}^c \} \right)^{\Delta_T N^c(x)} \left( 1 - E_{\theta}^{\tilde{T}} \{ \Delta_T N^c(x) | T\mathcal{F}_{x-1}^c \} \right)^{1-\Delta_T N^c(x)} \\
& = \Delta_T A^c(x, \theta)^{\Delta_T N^c(x)} \left( 1 - \Delta_T A^c(x, \theta) \right)^{1-\Delta_T N^c(x)} \tag{2.82}
\end{aligned}$$

Abschließend folgt durch Einsetzen von (2.82) in (2.79):

$${}_T L^c(\theta) = \prod_{x=T+1}^{\min(\chi, T+s)} \left\{ \left( 1 - \Delta_T A^c(x, \theta) \right)^{1-\Delta_T N^c(x)} \left( \Delta_T A^c(x, \theta) \right)^{\Delta_T N^c(x)} \right\}$$

Der Beweis ist somit abgeschlossen.  $\square$

Unter der Bedingung, dass die Zufallsvariable  $X$  geometrisch verteilt mit dem Parameter  $\theta$  ist, gilt für den marginalen, bedingten Likelihood-Beitrag (2.75):

$$\begin{aligned}
{}_T L^c(\theta) &= \prod_{x=T+1}^{\min(\chi, T+s)} \left\{ \left( 1 - \mathbb{1}_{\{T+1 \leq x \leq \min(X, T+s)\}} \theta \right)^{1-\mathbb{1}_{\{T+1 \leq x \leq T+s\}}} \mathbb{1}_{\{X=x\}} \right. \\
&\quad \left. \left( \mathbb{1}_{\{T+1 \leq x \leq \min(X, T+s)\}} \theta \right)^{\mathbb{1}_{\{T+1 \leq x \leq T+s\}}} \mathbb{1}_{\{X=x\}} \right\}
\end{aligned}$$

Abschließend ergibt sich:

$$\begin{aligned}
{}_T L^c(\theta) &= \left[ \prod_{x=T+1}^{\min(X-1, T+s, \chi)} (1-\theta)^1 \theta^0 \right] \cdot \left[ (1-\theta)^0 \theta^1 \right]^{\mathbb{1}_{\{X \leq \min(T+s, \chi)\}}} \\
&= \theta^{\mathbb{1}_{\{X \leq \min(T+s, \chi)\}}} (1-\theta)^{\min(X-1, T+s, \chi) - T} \tag{2.83}
\end{aligned}$$

Da die Herleitung des Beitrags eines Individuums zu einer marginalen, bedingten Likelihood-Funktion zur Schätzung des geometrischen Parameters  $\theta$  unter der Berücksichtigung von linksseitiger Trunkierung sowie rechtsseitiger Zensur abgeschlossen ist, kann im nächsten Schritt der Übergang von einem zufällig aus der Grundgesamtheit gezogenen Individuum zu einer Stichprobe erfolgen. Darauf aufbauend kann dann ein Punktschätzer für  $\theta$  sowie ein entsprechender Standardfehler ermittelt werden.

# 3. Inferenz für den Parameter der geometrischen Verteilung

## 3.1. Der Punktschätzer

Nachdem der marginale, bedingte Likelihood-Beitrag eines zufällig aus der Grundgesamtheit entnommenen Individuums unter Berücksichtigung von linksseitiger Trunkierung und rechtsseitiger Zensur hergeleitet wurde, erfolgt in diesem Kapitel eine Punktschätzung für die Wahrscheinlichkeit, innerhalb eines Jahres auszufallen, sowie die Berechnung zugehöriger Standardfehler. Die Berechnung der Standardfehler erfolgt dabei zunächst unter der Annahme, dass das Alter im letzten Jahr des Entstehungszeitraums  $T$  diskret gleichverteilt ist. Später wird diese Annahme wieder fallen gelassen und der Standardfehler in einem semiparametrischen Modell ermittelt. Vorher werden jedoch drei Darstellungen des Punktschätzers vorgestellt. Die Erste davon wird sich intuitiv als Nullstelle der abgeleiteten, logarithmierten, marginalen und bedingten Likelihood-Funktion ergeben. Die zweite Darstellungsform dient hingegen im späteren Verlauf als Grundlage zur Berechnung des Standardfehlers im semiparametrischen Modell mithilfe eines Martingalgrenzwertsatzes von Rebolledo [1979]. Zusätzlich dient die dritte Formel für den Punktschätzer zur Minimierung des Berechnungsaufwandes bei praktischen Anwendungen.

Dazu erfolgt ein Übergang zum Stichprobenmodell. Somit wird im Folgenden nicht mehr ein Individuum, sondern eine Stichprobe, bestehend aus  $n$  Individuen, untersucht. Dazu sei angemerkt, dass  $n$  auch nicht beobachtbare Individuen enthält und folglich unbekannt ist. Des Weiteren seien  $X_i$  die geometrisch verteilte (vgl. Definition 1) Lebensdauer und  $T_i$  das Alter ein Jahr vor dem Beginn der Studie für das  $i$ -te Individuum, wobei  $i \in \{1, \dots, n\}$ . Da zur Untersuchung der verschiedenen Schätzaspekte eine Unterscheidung zwischen wahren Wert, Parameter und Schätzer getroffen werden muss, erfolgt im ersten Schritt eine Einführung der Notation für eben diese Begriffe:

- $\theta_0$ : Für ein Individuum aus der Grundgesamtheit wird die wahre Wahrscheinlichkeit, innerhalb eines Jahres auszufallen, mit  $\theta_0$  bezeichnet. Somit sind alle  $X_i$  mit  $i \in \{1, \dots, n\}$  identisch geometrisch verteilt mit dem wahren Parameterwert  $\theta_0$ . Auch ist  $\theta_0$  unbekannt und soll im Folgenden geschätzt werden.
- $\theta$ : Der allgemeine Parameter der geometrischen Verteilung und somit das Argument der marginalen, bedingten Likelihood-Funktion wird mit  $\theta$  bezeichnet.
- $\hat{\theta}_n$ : Die Stelle des Maximums der marginalen, bedingten Likelihood-Funktion für eine Stichprobe vom Umfang  $n$  und somit der Punktschätzer von  $\theta_0$  wird mit  $\hat{\theta}_n$  bezeichnet.

Dabei werden die Daten als der linksseitig trunkeerte Teil einer einfachen Zufallsstichprobe, die aus der Grundgesamtheit gezogen wurde, interpretiert [siehe beispielsweise Andersen et al., 1988, Efron und Petrosian, 1999, Weißbach und Wied, 2022, Toparkus und Weißbach, 2025].

**Annahmen 1.** Es werden zusammenfassend folgende Annahmen getroffen:

- (A1) Für alle  $i \in \{1, \dots, n\}$  sind die bivariaten Zufallsvariablen  $(X_i, T_i)'$  auf dem gemeinsamen Wahrscheinlichkeitsraum  $(\Omega, \mathcal{F}, P_{\theta_0})$ , mit  $\Omega$  aus (2.1),  $\mathcal{F} = \mathcal{P}\{\Omega\}$  und  $P_{\theta_0}$  aus (2.4), definiert. Dabei entspricht die marginale Wahrscheinlichkeitsfunktion von  $X_i$  der Wahrscheinlichkeitsfunktion der geometrischen Verteilung (vgl. (2.5)). Außerdem ist  $X_i$  unabhängig von  $T_i$ .
- (A2) Das Tupel  $(X_i, T_i)'$  ist genau dann beobachtbar, wenn  $T_i + 1 \leq X_i$  erfüllt ist. Auch wird davon ausgegangen, dass für mindestens ein  $i \in \{1, \dots, n\}$  gilt, dass  $T_i + 1 \leq X_i$  und folglich mindestens ein Individuum beobachtbar ist.
- (A3) Für den wahren Parameterwert  $\theta_0$  gilt  $\theta_0 \in \Theta := [\xi, 1 - \xi]$  für ein kleines  $\xi \in (0, \frac{1}{2})$ .
- (A4)  $(X_i, T_i)', i \in \{1, \dots, n\}, n \in \mathbb{N}$  beschreibt eine einfache Zufallsstichprobe. Somit sind  $(X_i, T_i)'$  unabhängige und identisch verteilte Zufallsvariablen, die der Verteilungsannahme (A1) mit dem wahren Parameterwert  $\theta_0$  aus Annahme (A3) genügen.

Des Weiteren sei  ${}_T L_i^c$ ,  $i = 1, \dots, n$  der Beitrag (vgl. (2.83)) des  $i$ -ten Individuums zur marginalen, bedingten Likelihood-Funktion der Stichprobe  ${}_T L^c$ .

Es sei angemerkt, dass das  $i$ -te Individuum nur genau dann beobachtet werden kann, wenn  $T_i + 1 \leq X_i$ . Die marginale, bedingte Likelihood-Funktion für eine Stichprobe vom Umfang  $n$ , die über einen Zeitraum von  $\chi \geq G$  Jahren beobachtet wird, kann folglich mithilfe von (2.83) und Annahme (A4) folgendermaßen formuliert werden:

$${}_T L^c(\theta) = \prod_{i=1}^n [{}_T L_i^c(\theta)]^{\mathbb{1}_{\{T_i+1 \leq X_i\}}} = \prod_{i=1}^n [\theta^{\mathbb{1}_{\{X_i \leq \min(T_i+s, \chi)\}}} (1 - \theta)^{\min(X_i-1, T_i+s, \chi) - T_i}]^{\mathbb{1}_{\{T_i+1 \leq X_i\}}} \quad (3.1)$$

Folglich ergibt sich für die logarithmierte Likelihood-Funktion:

$$\begin{aligned} \log {}_T L^c(\theta) &= \log \theta \sum_{i=1}^n \mathbb{1}_{\{T_i+1 \leq X_i\}} \mathbb{1}_{\{X_i \leq \min(T_i+s, \chi)\}} \\ &\quad + \log(1 - \theta) \sum_{i=1}^n \mathbb{1}_{\{T_i+1 \leq X_i\}} [\min(X_i - 1, T_i + s, \chi) - T_i] \end{aligned} \quad (3.2)$$

Ableiten liefert

$$\begin{aligned} \frac{\partial \log {}_T L^c(\theta)}{\partial \theta} &= \frac{\sum_{i=1}^n \mathbb{1}_{\{T_i+1 \leq X_i\}} \mathbb{1}_{\{X_i \leq \min(T_i+s, \chi)\}}}{\theta} \\ &\quad - \frac{\sum_{i=1}^n \mathbb{1}_{\{T_i+1 \leq X_i\}} [\min(X_i - 1, T_i + s, \chi) - T_i]}{1 - \theta} \end{aligned} \quad (3.3)$$

und somit den eindeutigen Punktschätzer:

$$\begin{aligned}\hat{\theta}_n &= \left[ \sum_{i=1}^n \mathbb{1}_{\{T_i+1 \leq X_i\}} \mathbb{1}_{\{X_i \leq \min(T_i+s, \chi)\}} \right] \left[ \sum_{i=1}^n \mathbb{1}_{\{T_i+1 \leq X_i\}} \mathbb{1}_{\{X_i \leq \min(T_i+s, \chi)\}} \right. \\ &\quad \left. + \sum_{i=1}^n \mathbb{1}_{\{T_i+1 \leq \min(X_i, \chi)\}} \left[ \min(X_i - 1, T_i + s, \chi) - T_i \right] \right]^{-1} \\ &= \frac{\sum_{i=1}^n \mathbb{1}_{\{T_i+1 \leq X_i \leq \min(T_i+s, \chi)\}}}{\sum_{i=1}^n \mathbb{1}_{\{T_i+1 \leq X_i\}} \left[ \mathbb{1}_{\{X_i \leq \min(T_i+s, \chi)\}} + \min(X_i - 1, T_i + s, \chi) - T_i \right]} \quad (3.4)\end{aligned}$$

Es sei angemerkt, dass wegen Annahme 1, (A2) der Nenner von (3.4) nicht Null werden kann.

Im nächsten Schritt soll der Faktor  $\mathbb{1}_{\{X_i \leq \min(T_i+s, \chi)\}} + \min(X_i - 1, T_i + s, \chi) - T_i$  im Nenner von (3.4) vereinfacht werden. Dazu wird folgende Fallunterscheidung durchgeführt:

- Falls  $X_i \leq \min(T_i + s, \chi)$  erfüllt ist, folgt  $\min(X_i - 1, T_i + s, \chi) = X_i - 1$  und somit:

$$\mathbb{1}_{\{X_i \leq \min(T_i+s, \chi)\}} + \min(X_i - 1, T_i + s, \chi) - T_i = X_i - T_i$$

- Falls  $X_i > \min(T_i + s, \chi)$  erfüllt ist, folgt  $\min(X_i - 1, T_i + s, \chi) = \min(T_i + s, \chi)$  und somit:

$$\mathbb{1}_{\{X_i \leq \min(T_i+s, \chi)\}} + \min(X_i - 1, T_i + s, \chi) - T_i = \min(T_i + s, \chi) - T_i$$

Zusammenfassend ergibt sich

$$\mathbb{1}_{\{X_i \leq \min(T_i+s, \chi)\}} + \min(X_i - 1, T_i + s, \chi) - T_i = \min(X_i, T_i + s, \chi) - T_i \quad (3.5)$$

und somit die erste Darstellung des Punktschätzers:

$$\hat{\theta}_n = \frac{\sum_{i=1}^n \mathbb{1}_{\{T_i+1 \leq X_i \leq \min(T_i+s, \chi)\}}}{\sum_{i=1}^n \mathbb{1}_{\{T_i+1 \leq X_i\}} \left[ \min(X_i, T_i + s, \chi) - T_i \right]} \quad (3.6)$$

Zur weiteren Vereinfachung von (3.6) seien im Folgenden  ${}_T N_i^c$  (vgl. 2.64) und  ${}_T Y_i^c$  (vgl. 2.70) der linksseitig trunkeerte und rechtsseitig zensierte Zähl- beziehungsweise Unter-Risiko-Prozess des  $i$ -ten Individuums.

Mithilfe von (2.76) folgt direkt:

$$\sum_{i=1}^n \sum_{x=1}^{\chi} \Delta {}_T N_i^c(x) = \sum_{i=1}^n \sum_{x=1}^{\chi} \mathbb{1}_{\{T_i+1 \leq x \leq T_i+s\}} \mathbb{1}_{\{X_i=x\}} = \sum_{i=1}^n \mathbb{1}_{\{T_i+1 \leq X_i \leq \min(T_i+s, \chi)\}} \quad (3.7)$$

Des Weiteren lässt sich aus (2.71) die Gleichheit

$$\begin{aligned} \sum_{i=1}^n \sum_{x=1}^{\chi} {}_T Y_i^c(x-1) &= \sum_{i=1}^n \sum_{x=1}^{\chi} \mathbb{1}_{\{T_i+1 \leq x \leq \min(X_i, T_i+s)\}} \\ &= \sum_{i=1}^n \mathbb{1}_{\{T_i+1 \leq X_i\}} \left[ \min(X_i, T_i+s, \chi) - T_i \right] \end{aligned} \quad (3.8)$$

ableiten.

Durch Einsetzen von (3.7) und (3.8) in (3.6) ergibt sich somit die folgende Darstellung des Punktschätzers:

$$\hat{\theta}_n = \frac{\sum_{i=1}^n \sum_{x=1}^{\chi} \Delta_T N_i^c(x)}{\sum_{i=1}^n \sum_{x=1}^{\chi} {}_T Y_i^c(x-1)} \quad (3.9)$$

Abschließend erfolgt die Herleitung einer Version des Punktschätzers zur Minimierung des Berechnungsaufwands. Dazu sei angenommen, dass  $\chi \geq (G-1) + s$  für alle  $i \in \{1, \dots, n\}$  erfüllt ist, sodass im Punktschätzer alle Informationen bis zum Ende des Beobachtungszeitraums verarbeitet werden. Auch sei angemerkt, dass  $n$  den Stichprobenumfang ohne Berücksichtigung der Trunkierung bezeichnet und folglich in der Praxis nicht beobachtbar ist. Zwar gibt es Beispiele für die Analyse aus trunkierten Daten, die nicht zwischen der Anzahl der Beobachtungen und dem Stichprobenumfang unterscheiden [siehe beispielsweise Heckman, 1976], im Rahmen dieser Arbeit werden Beobachtungsanzahl und Stichprobenumfang jedoch klar voneinander abgegrenzt. Dazu bezeichne im Gegensatz zum Stichprobenumfang  $n$  das Symbol  $m$  im Folgenden die Anzahl der Beobachtungen. Sei weiter

$$m = m_{uncens} + m_{cens}$$

eine Zerlegung von  $m$  in die Individuen, deren Ausfall innerhalb des Beobachtungszeitraum stattgefunden hat ( $m_{uncens}$ ), und die Individuen, die am Ende des Beobachtungszeitraums noch nicht ausgefallen sind und folglich zensiert wurden ( $m_{cens}$ ). Zunächst gilt mit  $\chi \geq (G-1) + s$  für den Zähler von (3.6):

$$\sum_{i=1}^n \mathbb{1}_{\{T_i+1 \leq X_i \leq \min(T_i+s, \chi)\}} = \sum_{i=1}^n \mathbb{1}_{\{T_i+1 \leq X_i \leq T_i+s\}} = m_{uncens}$$

Um den Nenner von (3.6) weiter zu vereinfachen, werden die Summanden

$$\sum_{i=1}^n \mathbb{1}_{\{T_i+1 \leq X_i\}} \left[ \min(X_i, T_i+s) - T_i \right] \quad (3.10)$$

in zwei Kategorien eingeteilt:

- Für Individuen, deren Ausfall beobachtet wird, gilt  $\min(X_i, T_i+s) - T_i = X_i - T_i$ . Seien dazu  $x_j^{obs}$  und  $t_j^{obs}$  mit  $j \in \{1, \dots, m_{uncens}\}$  die beobachteten Realisierungen

der Alter bei Ausfall und am Ende des Entstehungszeitraums. Dann ist

$$d^{obs} := \sum_{j=1}^{m_{uncens}} x_j^{obs} - t_j^{obs}$$

die Summe über alle realisierten, beobachtbaren Lebensspannen innerhalb der Studie und somit der Anteil von (3.10), der alle Individuen mit beobachtbarem Ausfall repräsentiert.

- Für alle  $m_{cens}$  Individuen, deren Historie rechtsseitig zensiert und deren Ausfall entsprechend nicht innerhalb der Studie beobachtet wird, gilt  $\min(X_i, T_i + s) - T_i = (T_i + s) - T_i = s$ . Dieser Anteil kann somit durch  $sm_{cens}$  beschrieben werden.

Abschließend ergibt sich die dritte und letzte Darstellung des Punktschätzers für die Wahrscheinlichkeit, dass ein Individuum innerhalb eines Jahres ausfällt, folgendermaßen:

$$\hat{\theta}_n = \frac{m_{uncens}}{d^{obs} + sm_{cens}} \quad (3.11)$$

**Bemerkung 13.** Aus dem Vergleich von (3.9) mit (3.11) lässt sich das Verhältnis

$$\sum_{i=1}^n \sum_{x=1}^{(G-1)+s} {}_T Y_i^c(x-1) = d^{obs} + sm_{cens} \quad (3.12)$$

ableiten.

Nachdem der Punktschätzer  $\hat{\theta}_n$  für die Wahrscheinlichkeit, innerhalb eines Jahres auszufallen, berechnet wurde, besteht das nächste Ziel in der Angabe eines Konfidenzintervalls zum Niveau  $1 - \alpha$  für den wahren Parameterwert  $\theta_0$ . Dazu werden im Folgenden einige Methoden zur Berechnung des Standardfehlers von  $\hat{\theta}_n$  vorgestellt.

## 3.2. Standardfehler im parametrischen Modell

Mithilfe von (3.9) wurde ein Punktschätzer  $\hat{\theta}_n$  für die wahre Wahrscheinlichkeit, innerhalb eines Jahres auszufallen,  $\theta_0$ , basierend auf einer Stichprobe, die aus  $n$  geometrisch verteilten Lebensdauern  $X_1, \dots, X_n$  besteht, ermittelt. Dabei wurden sowohl die linksseitige Trunkierung als auch die rechtsseitige Zensur berücksichtigt.

Meistens liegen in der Praxis jedoch Abweichungen zwischen dem wahren Parameterwert und der entsprechenden Punktschätzung vor. Aus diesem Grund werden bevorzugt Bereichsschätzungen in Form von Konfidenzintervallen zum Niveau  $1 - \alpha$  mit  $\alpha \in (0, 1)$  angegeben. Dabei handelt es sich um ein Intervall, das mit einer zuvor festgelegten Wahrscheinlichkeit  $1 - \alpha$  den wahren Parameterwert  $\theta_0$  enthält. Um eben so einen Bereich angeben zu können, wird neben dem Punktschätzer noch der Standardfehler SE

benötigt. Dieser ist definiert als der Schätzer der Standardabweichung von  $\hat{\theta}_n$  mit:

$$\sqrt{\text{Var}_{\theta_0}\{\hat{\theta}_n\}} = \frac{\sqrt{\text{Var}_{\theta_0}\{\sqrt{n}\hat{\theta}_n\}}}{\sqrt{n}} = \frac{\sigma}{\sqrt{n}} \quad (3.13)$$

Bei

$$\sigma := \sqrt{\text{Var}_{\theta_0}\{\sqrt{n}\hat{\theta}_n\}}$$

handelt es sich um die Standardabweichung der asymptotischen Normalverteilung der Folge  $(\sqrt{n}(\hat{\theta}_n - \theta_0))_n$  für  $n \rightarrow \infty$ . Als Grundvoraussetzung für diese asymptotische Normalität muss zunächst die Konsistenz der Schätzfolge  $(\hat{\theta}_n)_n$  nachgewiesen werden.

Die dafür verwendeten Sätze stammen aus van der Vaart [1998, Lemma 5.10 und Theorem 5.41] und basieren auf dem Konzept der M-Schätzer. Zunächst werden diese inklusive einer entsprechenden Notation sowie weitere Modellannahmen eingeführt. Anschließend können dann die Konsistenz von  $(\hat{\theta}_n)_n$  sowie die asymptotische Normalität von  $(\sqrt{n}(\hat{\theta}_n - \theta_0))_n$  nachgewiesen sowie der Standardfehler

$$\text{SE} := \sqrt{\widehat{\text{Var}_{\theta_0}\{\hat{\theta}_n\}}}$$

ermittelt werden.

An dieser Stelle sei hervorgehoben, dass die Berechnung des Standardfehlers mithilfe von Konsistenz und asymptotischer Normalität nach van der Vaart [1998] ein parametrisches Modell voraussetzt. Während für  $X_i, i \in \{1, \dots, n\}$  eine geometrische Verteilung mit dem wahren Parameterwert  $\theta_0$  angenommen wurde, musste bisher für  $T_i, i \in \{1, \dots, n\}$  keine Verteilungsannahme getroffen werden.

Da in der später erfolgenden Anwendung (siehe Kapitel 5) die Lebensdauer von Unternehmen untersucht werden soll und die Analyse von Unternehmen, die vor hunderten von Jahren gegründet wurden, wenig sinnvoll erscheint, wird für die Verteilung von  $T_i$  ein endlicher Träger vorausgesetzt. Werden die Entstehungen der Individuen zusätzlich als homogener Poisson-Prozess modelliert, ergibt sich für  $T_i$  eine Gleichverteilung [siehe Dörre, 2020], sodass die Gleichverteilung als Referenz für die Verteilung von  $T_i$  interpretiert werden kann.

Entsprechend wird für den Nachweis von Konsistenz und asymptotischer Normalität vorausgesetzt, dass  $T_1, \dots, T_n$  unabhängig und identisch diskret gleichverteilte Zufallsvariablen auf dem Träger  $\{0, \dots, G-1\}$  sind.

**Definition 7** (Diskrete Gleichverteilung). *Sei  $T$  eine diskrete Zufallsvariable mit dem Bild  $\mathbf{t}$ . Dann heißt  $T$  diskret gleichverteilt auf dem Intervall  $\{0, \dots, G-1\}$  genau dann, wenn die zugehörige Wahrscheinlichkeitsfunktion gegeben ist durch:*

$$f_U(\mathbf{t}) = P_\theta\{T = \mathbf{t}\} = P_T\{T = \mathbf{t}\} = \begin{cases} \frac{1}{G-1}, & \text{falls } \mathbf{t} = \{0, \dots, G-1\} \\ 0, & \text{sonst} \end{cases} \quad (3.14)$$

Die Verteilungsfunktion von  $T$  ist entsprechend definiert durch:

$$F_U(\mathbf{t}) = P_\theta\{T \leq \mathbf{t}\} = P_T\{T \leq \mathbf{t}\} = \begin{cases} 0, & \text{falls } \mathbf{t} < 0 \\ \frac{\mathbf{t}}{G-1}, & \text{falls } \mathbf{t} = \{0, \dots, G-1\} \\ 1, & \text{falls } \mathbf{t} \geq G \end{cases} \quad (3.15)$$

Außerdem wird davon ausgegangen, dass für die aus der Grundgesamtheit gezogene Stichprobe vom Umfang  $n \in \mathbb{N}$  die beiden folgenden Szenarien nicht eintreten:

- (S1) Alle Individuen fallen vor dem Beginn der Studie aus oder werden zensiert, sodass  $T_i \geq X_i$  oder  $X_i > T_i + s$  für alle  $i \in \{1, \dots, n\}$  erfüllt ist. In diesem Szenario würde  $\hat{\theta}_n = 0$  gelten.
- (S2) Alle Individuen fallen im ersten Jahr des Beobachtungszeitraums aus, sodass  $X_i = T_i + 1$  für alle  $i \in \{1, \dots, n\}$  erfüllt ist. In diesem Szenario würde  $\hat{\theta}_n = 1$  gelten.

Das Vermeiden der Szenarien (S1) und (S2) garantiert, dass  $\Theta = [\xi, 1 - \xi]$  mit  $\xi \in (0, 1/2)$  als Parameterraum gewählt werden kann. Für dieses Modell können im Folgenden M-Schätzer definiert werden.

Als M-Schätzer wird ein Schätzer bezeichnet, der Maximum-Likelihood-artig ist [siehe van der Vaart, 1998, Kapitel 5.1]. Soll der Verteilungsparameter  $\theta_0 \in \Theta$  der Zufallsvariablen  $X_1, \dots, X_n$ , die auf einem gemeinsamen Wahrscheinlichkeitsraum  $(\Omega, \mathcal{F}, P_{\theta_0})$  definiert sind, geschätzt werden, geschieht dies über die Maximierung einer Funktion der Form

$$\theta \mapsto M_n(\theta) := \frac{1}{n} \log {}_T L^c(\theta) = \frac{1}{n} \sum_{i=1}^n m_\theta(X_i, T_i) \quad (3.16)$$

Dabei ist  $m_\theta : \mathbb{N}^2 \rightarrow \mathbb{R}$  eine bekannte Funktion, die entsprechend (3.16) und der in (3.2) aufgestellten, logarithmierten, marginalen und bedingten Likelihood-Funktion im Rahmen dieser Arbeit durch

$$m_\theta(X_i, T_i) := \mathbb{1}_{\{T_i+1 \leq X_i\}} \left[ \log(\theta) \mathbb{1}_{\{X_i \leq \min(T_i+s, \chi)\}} + \log(1 - \theta) \left( \min(X_i - 1, T_i + s, \chi) - T_i \right) \right]$$

definiert wird. Zur Gewährleistung einer besseren Übersichtlichkeit wird zusätzlich angenommen, dass  $\chi \geq G + s$ , sodass sich

$$m_\theta(X_i, T_i) := \mathbb{1}_{\{T_i+1 \leq X_i\}} \left[ \log(\theta) \mathbb{1}_{\{X_i \leq T_i+s\}} + \log(1 - \theta) \left( \min(X_i - 1, T_i + s) - T_i \right) \right]$$

ergibt.

Allgemein werden Schätzer, die  $M_n(\theta)$  über  $\Theta$  maximieren, als M-Schätzer bezeichnet. Die Stelle des Maximums lässt sich dann als Nullstelle der Ableitung bestimmen. M-

Schätzer können folglich auch als Lösung der Gleichung

$$\Psi_n(\theta) := \frac{1}{n} \sum_{i=1}^n \psi_\theta(X_i, T_i) = 0 \quad (3.17)$$

definiert werden. Da im vorliegenden Fall der M-Schätzer  $\hat{\theta}_n$  als das Maximum der logarithmierten, marginalen und bedingten Likelihood-Funktion (3.2) verstanden wird, folgt:

$$\psi_\theta(X_i, T_i) := \frac{\partial m_\theta(X_i, T_i)}{\partial \theta} = \mathbb{1}_{\{T_i+1 \leq X_i\}} \left[ \frac{\mathbb{1}_{\{X_i \leq T_i+s\}}}{\theta} - \frac{\min(X_i - 1, T_i + s) - T_i}{1 - \theta} \right] \quad (3.18)$$

### 3.2.1. Konsistenz

Soll der wahre Parameter  $\theta_0$  der Grundgesamtheit anhand einer Stichprobe, bestehend aus  $n$  Merkmalsträgern, geschätzt werden, sollte bei steigendem Stichprobenumfang auch die Genauigkeit der Schätzung zunehmen. Der wahre Parameterwert  $\theta_0$  sollte folglich bei steigendem Stichprobenumfang mit zunehmender Genauigkeit durch die Schätzfolge  $(\hat{\theta}_n)_n$  approximiert werden können. Dieses Qualitätsmerkmal einer Schätzfolge lässt sich anhand des Konsistenzbegriffes mathematisch formulieren [siehe van der Vaart, 1998, Kapitel 5.2]:

**Definition 8** (Konsistenz). *Eine Schätzfolge  $(\hat{\theta}_n)_n$  heißt (schwach) konsistent, wenn die Folge für alle möglichen Werte von  $\theta$  in Wahrscheinlichkeit gegen  $\theta$  konvergiert:*

$$\hat{\theta}_n \xrightarrow{p} \theta, \forall \theta \in \Theta \quad (3.19)$$

Die für den Konsistenznachweis notwendige Konvergenz in Wahrscheinlichkeit wird dabei in Definition 18 eingeführt:

Das nächste große Ziel ist der Nachweis der Konvergenz in Wahrscheinlichkeit von  $(\hat{\theta}_n)_n$  gegen den wahren Parameterwert  $\theta_0$ . Dazu wird eine Grenzfunktion  $\theta \mapsto \Psi(\theta)$  ermittelt, sodass  $\Psi_n(\theta) \xrightarrow{p} \Psi(\theta)$  für alle  $\theta \in \Theta$  gilt und  $\theta_0$  eine Nullstelle von  $\Psi$  ist. Dann konvergiert unter gewissen Voraussetzungen die Folge der Nullstellen  $(\hat{\theta}_n)_n$  von  $(\Psi_n)_n$  in Wahrscheinlichkeit gegen die Nullstelle  $\theta_0$  von  $\Psi$ .

Zunächst wird zu diesem Zweck gezeigt, dass die Funktionenfolge  $(\Psi_n)_n$  gleichmäßig in Wahrscheinlichkeit gegen die Funktion

$$\Psi(\theta) := E_\theta\{\psi_\theta(X_i, T_i)\} \quad (3.20)$$

konvergiert, wobei  $\psi$  in (3.18) definiert wurde.

In Newey und McFadden [1994, Kapitel 2.3, Lemma 2.4] wird eine Variante des gleichmäßigen Gesetzes der großen Zahlen angegeben, die für diesen Zweck verwendet werden kann.

**Satz 6** (Gleichmäßiges Gesetz der großen Zahlen). *Seien  $(X_1, T_1)', \dots, (X_n, T_n)'$  unabhängige und identisch verteilte bivariate Zufallsvariablen,  $\Theta$  kompakt und  $\theta \rightarrow \psi_\theta(x, t)$  stetig für fast alle  $(x, t)'$ . Außerdem existiert eine Abbildung  $h(x, t)$  mit  $\|\psi_\theta(x, t)\| \leq h(x, t)$  für alle  $\theta \in \Theta$  sowie  $E_\theta\{h(X_i, T_i)\} < \infty$ .*

*Dann ist  $E_\theta\{\psi_\theta(X_i, T_i)\} = \Psi(\theta)$  stetig und es gilt:*

$$\sup_{\theta \in \Theta} \left\| \frac{1}{n} \sum_{i=1}^n (\psi_\theta(X_i, T_i) - E_\theta\{\psi_\theta(X_i, T_i)\}) \right\| \xrightarrow{p} 0 \quad (3.21)$$

Im Folgenden werden die Eigenschaften von Satz 6 überprüft. Die bivariaten Zufallsvariablen  $(X_i, T_i)', i \in 1, \dots, n$  sind nach der Annahme 1,(A4), unabhängig und identisch verteilt. Der Parameterraum  $\Theta = [\xi, 1 - \xi]$  mit  $\xi \in (0, 1/2)$  ist darüber hinaus als abgeschlossene und beschränkte Teilmenge des eindimensionalen euklidischen Raumes kompakt. Da  $\theta \in \Theta$  und somit  $0 < \theta < 1$  gilt, ist  $\theta \mapsto \psi_\theta(x, t)$  nach (3.18) als Komposition stetiger Abbildungen selbst auch stetig.

Zur Bestimmung einer Funktion  $h(x, t)$  mit  $\|\psi_\theta(x, t)\| \leq h(x, t)$  sei zunächst angemerkt, dass

$$\|\psi_\theta(x, t)\| = \sqrt{\psi_\theta^2(x, t)} = |\psi_\theta(x, t)|$$

Auf Grundlage dessen und mithilfe von  $\xi \leq \theta \leq 1 - \xi$  kann basierend auf (3.18) folgende Abschätzung vorgenommen werden:

$$\begin{aligned} \|\psi_\theta(x, t)\| &= \left| \mathbb{1}_{\{t+1 \leq x\}} \left[ \frac{\mathbb{1}_{\{x \leq t+s\}}}{\theta} - \frac{\min(x-1, t+s) - t}{1-\theta} \right] \right| \\ &\leq \left| \frac{\mathbb{1}_{\{t+1 \leq x\}} \mathbb{1}_{\{x \leq t+s\}}}{\theta} \right| + \left| \frac{\mathbb{1}_{\{t+1 \leq x\}} \min(x-1, t+s)}{1-\theta} \right| + \left| \frac{\mathbb{1}_{\{t+1 \leq x\}} t}{1-\theta} \right| \\ &< \frac{1}{\theta} + \frac{x+s}{1-\theta} + \frac{t}{1-\theta} < \frac{x+t+s+1}{\xi} =: h(x, t) \end{aligned}$$

Es bleibt zu zeigen, dass  $E_\theta\{h(X_i, T_i)\} < \infty$ :

$$E_\theta\{h(X_i, T_i)\} = E_\theta\left\{ \frac{X_i + T_i + s + 1}{\xi} \right\} = \frac{1}{\xi G} \sum_{t=0}^{G-1} \sum_{x=1}^{\infty} \theta(1-\theta)^{x-1} (x+t+s+1) \quad (3.22)$$

Da es sich bei  $\theta$  um den Parameter einer geometrischen Verteilung handelt, gilt:

$$\sum_{x=1}^{\infty} \theta(1-\theta)^{x-1} x = E_\theta\{X_i\} = \frac{1}{\theta} \quad (3.23)$$

$$\sum_{x=1}^{\infty} \theta(1-\theta)^{x-1} t = t \quad (3.24)$$

$$\sum_{x=1}^{\infty} \theta(1-\theta)^{x-1} (s+1) = s+1 \quad (3.25)$$

Da der rechte Rand  $G - 1$  des Trägers der Verteilung von  $T_i$  endlich ist, folgt durch einsetzen von (3.23)-(3.25) in (3.22):

$$\mathbb{E}_\theta\{h(X_i, T_i)\} = \frac{1}{\xi G} \sum_{t=0}^{G-1} \left( \frac{1}{\theta} + t + (s+1) \right) < \infty$$

Somit sind alle Bedingungen von Satz 6 erfüllt. Mit

$$\frac{1}{n} \sum_{i=1}^n \mathbb{E}_\theta(\psi_\theta(X_i, T_i)) = \frac{1}{n} n \Psi(\theta) = \Psi(\theta)$$

folgt:

$$\sup_{\theta \in \Theta} \|\Psi_n(\theta) - \Psi(\theta)\| \xrightarrow{p} 0$$

Somit konvergiert  $(\Psi_n)_n$  gleichmäßig in Wahrscheinlichkeit gegen  $\Psi$ .

Im nächsten Schritt ist nachzuweisen, dass dann auch die Folge der Nullstellen  $(\hat{\theta}_n)_n$  von  $(\Psi_n)_n$  in Wahrscheinlichkeit gegen die Nullstelle  $\theta_0$  von  $\Psi$  konvergiert. Dazu kann der folgende Satz verwendet werden [siehe van der Vaart, 1998, Lemma 5.10]:

**Satz 7.** Sei  $\Theta \subseteq \mathbb{R}$  und  $(\Psi_n(\theta))_n$  eine Folge von Funktionen, sodass für alle  $n$  gilt, dass die Funktion  $\theta \mapsto \Psi_n(\theta)$  stetig ist und genau eine Nullstelle  $\hat{\theta}_n$  hat. Außerdem konvergiere die Folge  $(\Psi_n(\theta))_n$  für alle  $\theta$  in Wahrscheinlichkeit gegen eine feste Funktion  $\theta \mapsto \Psi(\theta)$ . Des Weiteren erfülle  $\theta_0$  die Bedingung  $\Psi(\theta_0 - \epsilon) > 0 > \Psi(\theta_0 + \epsilon)$  für alle  $\epsilon > 0$ . Dann gilt  $\hat{\theta}_n \xrightarrow{p} \theta_0$ .

Im Folgenden wird überprüft, ob im Rahmen des besprochenen Modells die Bedingungen von Satz 7 erfüllt werden. Zunächst folgt mithilfe von (3.17) und (3.18), dass:

$$\Psi_n(\theta) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{\{T_i+1 \leq X_i\}} \left[ \frac{\mathbb{1}_{\{X_i \leq T_i+s\}}}{\theta} - \frac{\min(X_i - 1, T_i + s) - T_i}{1 - \theta} \right] \quad (3.26)$$

Dabei folgt direkt aus  $\theta \in [\xi, 1 - \xi]$  mit  $\xi \in (0, 1/2)$ , dass  $\Psi_n(\theta)$  als Komposition stetiger Funktionen stetig ist. Des Weiteren ist  $\hat{\theta}_n$  als Nullstelle von  $\partial \log_T L^c(\theta) / \partial \theta$  (vgl. (3.3)) auch eine Lösung der Gleichung

$$\Psi_n(\theta) = \frac{1}{n} \frac{\partial \log_T L^c(\theta)}{\partial \theta} = 0$$

Um die Eindeutigkeit der Lösung nachzuweisen, wird gezeigt, dass die Funktion  $\Psi_n$  (vgl. 3.26) auf dem gesamten Definitionsbereich  $[\xi, 1 - \xi]$  streng monoton fallend ist. Für die Ableitung  $\dot{\Psi}_n$  gilt:

$$\dot{\Psi}_n(\theta) = -\frac{1}{n} \sum_{i=1}^n \mathbb{1}_{\{T_i+1 \leq X_i\}} \left[ \frac{\mathbb{1}_{\{X_i \leq T_i+s\}}}{\theta^2} + \frac{\min(X_i - 1, T_i + s) - T_i}{(1 - \theta)^2} \right]$$

Da nach Voraussetzung die Szenarien (S1) und (S2) nicht eintreten, folgt für die Sum-

manden:

$$\sum_{i=1}^n \frac{\mathbb{1}_{\{T_i+1 \leq X_i\}} \mathbb{1}_{\{X_i \leq T_i+s\}}}{\theta^2} > \frac{0}{\theta^2} = 0$$

$$\sum_{i=1}^n \mathbb{1}_{\{T_i+1 \leq X_i\}} \frac{\min(X_i - 1, T_i + s) - T_i}{(1 - \theta)^2} > \frac{0}{(1 - \theta)^2} = 0$$

und somit:

$$\dot{\Psi}_n(\theta) < 0, \forall \theta \in \Theta = [\xi, 1 - \xi]$$

Da  $\Psi_n$  eine stetige und streng monotone Funktion ist, ist die Nullstelle  $\hat{\theta}_n$  eindeutig.

Im nächsten Schritt soll mithilfe von (3.18) und (3.20) die Grenzfunktion  $\Psi$  der Folge  $(\Psi_n)_n$  explizit berechnet werden. Es gilt:

$$\begin{aligned} \Psi(\theta) &= \mathbb{E}_{\theta_0} \left\{ \mathbb{1}_{\{T_i+1 \leq X_i\}} \left[ \frac{\mathbb{1}_{\{X_i \leq T_i+s\}}}{\theta} - \frac{\min(X_i - 1, T_i + s) - T_i}{1 - \theta} \right] \right\} \\ &= \frac{1}{G} \sum_{t=0}^{G-1} \mathbb{E}_{\theta_0} \left\{ \mathbb{1}_{\{t+1 \leq X_i\}} \left[ \frac{\mathbb{1}_{\{X_i \leq t+s\}}}{\theta} - \frac{\min(X_i - 1, t + s) - t}{1 - \theta} \right] \right\} \\ &= \frac{1}{G} \sum_{t=0}^{G-1} \left[ \frac{1}{\theta} \mathbb{E}_{\theta_0} \left\{ \mathbb{1}_{\{t+1 \leq X_i \leq t+s\}} \right\} - \frac{1}{1 - \theta} \mathbb{E}_{\theta_0} \left\{ \mathbb{1}_{\{t+1 \leq X_i\}} (\min(X_i - 1, t + s) - t) \right\} \right] \end{aligned}$$

wobei

$$\mathbb{E}_{\theta_0} \left\{ \mathbb{1}_{\{t+1 \leq X_i \leq t+s\}} \right\} = \sum_{x=t+1}^{t+s} f_G(x; \theta_0) = \theta_0 \left( \sum_{x=t}^{t+s-1} (1 - \theta_0)^x \right)$$

und

$$\begin{aligned} &\mathbb{E}_{\theta_0} \left\{ \mathbb{1}_{\{t+1 \leq X_i\}} (\min(X_i - 1, t + s) - t) \right\} \\ &= \sum_{x=1}^{\infty} f_G(x; \theta_0) \mathbb{1}_{\{t+1 \leq x\}} (\min(x - 1, t + s) - t) = \sum_{x=t+1}^{\infty} f_G(x; \theta_0) (\min(x - 1, t + s) - t) \\ &= \sum_{x=t+1}^{t+s} (x - t - 1) f_G(x; \theta_0) + \sum_{x=t+s+1}^{\infty} s f_G(x; \theta_0) \\ &= \sum_{x=t+1}^{t+s} (x - t - 1) f_G(x; \theta_0) + s \left( 1 - \sum_{x=1}^{t+s} f_G(x; \theta_0) \right) \\ &= \sum_{x=t+1}^{t+s} (x - t - 1) \theta_0 (1 - \theta_0)^{x-1} + s \left( 1 - \sum_{x=1}^{t+s} \theta_0 (1 - \theta_0)^{x-1} \right) \end{aligned}$$

Insgesamt folgt:

$$\Psi(\theta) = \frac{1}{G} \sum_{t=0}^{G-1} \left[ \frac{\theta_0 \left( \sum_{x=t}^{t+s-1} (1-\theta_0)^x \right)}{\theta} - \frac{\sum_{x=t+1}^{t+s} (x-t-1)\theta_0(1-\theta_0)^{x-1} + s \left( 1 - \sum_{x=1}^{t+s} \theta_0(1-\theta_0)^{x-1} \right)}{1-\theta} \right]$$

Abschließend muss  $\Psi(\theta_0 - \epsilon) > 0 > \Psi(\theta_0 + \epsilon)$  für ein  $\epsilon > 0$  überprüft werden. Dies geschieht anhand der folgenden zwei Schritte:

- (1) Es wird gezeigt, dass  $\theta_0$  eine Nullstelle von  $\Psi$  ist.
- (2) Es wird gezeigt, dass die Funktion  $\Psi$  in einer Umgebung von  $\theta_0$  streng monoton fallend ist.

Um (1) und somit

$$\Psi(\theta_0) = \frac{1}{G} \sum_{t=0}^{G-1} \left[ \frac{\theta_0 \left( \sum_{x=t}^{t+s-1} (1-\theta_0)^x \right)}{\theta_0} - \frac{\sum_{x=t+1}^{t+s} (x-t-1)\theta_0(1-\theta_0)^{x-1} + s \left( 1 - \sum_{x=1}^{t+s} \theta_0(1-\theta_0)^{x-1} \right)}{1-\theta_0} \right] = 0 \quad (3.27)$$

zu zeigen, sei angemerkt, dass unter der Verwendung von

$$\begin{aligned} \frac{s \left( 1 - \sum_{x=1}^{t+s} \theta_0(1-\theta_0)^{x-1} \right)}{1-\theta_0} &= \frac{s \sum_{x=t+s+1}^{\infty} \theta_0(1-\theta_0)^{x-1}}{1-\theta_0} \\ &= s \sum_{x=t+s}^{\infty} \theta_0(1-\theta_0)^{x-1} = s \left( 1 - \mathbb{1}_{\{t+s \leq 2\}} \sum_{x=1}^{t+s-1} \theta_0(1-\theta_0)^{x-1} \right) \end{aligned}$$

gilt, dass:

$$\begin{aligned} \Psi(\theta_0) &= \frac{1}{G} \sum_{t=0}^{G-1} \left[ \sum_{x=t}^{t+s-1} (1-\theta_0)^x - \sum_{x=t+1}^{t+s} (x-t-1)\theta_0(1-\theta_0)^{x-2} \right. \\ &\quad \left. - s \left( 1 - \mathbb{1}_{\{t+s \geq 2\}} \sum_{x=1}^{t+s-1} \theta_0(1-\theta_0)^{x-1} \right) \right] \end{aligned}$$

Es wird nun gezeigt, dass für alle  $t \in \{0, \dots, G-1\}$  und alle  $s \in \mathbb{N}$

$$\sum_{x=t}^{t+s-1} (1-\theta_0)^x - \sum_{x=t+1}^{t+s} (x-t-1)\theta_0(1-\theta_0)^{x-2} - s \left( 1 - \mathbb{1}_{\{t+s \geq 2\}} \sum_{x=1}^{t+s-1} \theta_0(1-\theta_0)^{x-1} \right) = 0 \quad (3.28)$$

erfüllt ist. Zunächst gilt:

$$\mathbb{1}_{\{t+s \geq 2\}} = \begin{cases} 0, & \text{wenn } t = 0 \text{ und } s = 1 \\ 1, & \text{sonst} \end{cases}$$

Zur Überprüfung von (3.28) erfolgt für den Fall  $t = s-1 = 0$  eine separate Betrachtung. Es ergibt sich:

$$\sum_{x=0}^0 (1-\theta_0)^x - \sum_{x=1}^1 (x-1)\theta_0(1-\theta_0)^{x-2} - 1 = 1 - 0 - 1 = 0$$

Für die weitere Betrachtung gilt  $\mathbb{1}_{\{t+s \geq 2\}} = 1$ , sodass dieser Term entsprechend weggelassen werden kann. Es erfolgt ein Nachweis von (3.28) per vollständiger Induktion, sowohl über  $t$  als auch über  $s$ . Dazu wird zunächst  $s = 1$  gesetzt und eine Induktion in  $t$ -Richtung durchgeführt. Sei für den Induktionsanfang  $t = 1$ . Dann gilt für (3.28):

$$\begin{aligned} & \sum_{x=1}^1 (1-\theta_0)^x - \sum_{x=2}^2 (x-1-1)\theta_0(1-\theta_0)^{x-2} - \left( 1 - \sum_{x=1}^1 \theta_0(1-\theta_0)^{x-1} \right) \\ & = (1-\theta_0) - 0 - (1-\theta_0) = 0 \end{aligned}$$

Es folgt der Induktionsschritt  $t \mapsto t+1$ :

Die Induktionsvoraussetzung lautet:

$$\begin{aligned} & (1-\theta_0)^t - ((t+1) - t - 1)\theta_0(1-\theta_0)^{t-1} - \left( 1 - \sum_{x=1}^t \theta_0(1-\theta_0)^{x-1} \right) \\ & = (1-\theta_0)^t - \left( 1 - \sum_{x=1}^t \theta_0(1-\theta_0)^{x-1} \right) = 0 \end{aligned}$$

Darauf aufbauend soll nun

$$\begin{aligned} & (1-\theta_0)^{t+1} - ((t+2) - (t+1) - 1)\theta_0(1-\theta_0)^{x-2} - \left( 1 - \sum_{x=1}^{t+1} \theta_0(1-\theta_0)^{x-1} \right) \\ & = (1-\theta_0)^{t+1} - \left( 1 - \sum_{x=1}^{t+1} \theta_0(1-\theta_0)^{x-1} \right) = 0 \end{aligned} \quad (3.29)$$

gezeigt werden.

Es gilt:

(i)

$$(1 - \theta_0)^{t+1} = (1 - \theta_0)(1 - \theta_0)^t$$

(ii)

$$1 - \sum_{x=1}^{t+1} \theta_0(1 - \theta_0)^{x-1} = 1 - \sum_{x=1}^t \theta_0(1 - \theta_0)^{x-1} - \theta_0(1 - \theta_0)^t$$

Einsetzen von (i) und (ii) in (3.29) liefert mithilfe der Induktionsvoraussetzung:

$$\begin{aligned} & (1 - \theta_0)(1 - \theta_0)^t - \left(1 - \sum_{x=1}^t \theta_0(1 - \theta_0)^{x-1} - \theta_0(1 - \theta_0)^t\right) \\ &= (1 - \theta_0)^t - \left(1 - \sum_{x=1}^t \theta_0(1 - \theta_0)^{x-1}\right) - \theta_0(1 - \theta_0)^t + \theta_0(1 - \theta_0)^t \\ &= (1 - \theta_0)^t - \left(1 - \sum_{x=1}^t \theta_0(1 - \theta_0)^{x-1}\right) = 0 \end{aligned}$$

Die Gültigkeit von (3.28) ist somit für  $s = 1$  und alle  $t \in \{0, \dots, G - 1\}$  nachgewiesen. Darauf aufbauend wird im Folgenden eine vollständige Induktion über  $s$  durchgeführt. Der Induktionsanfang mit  $s = 1$  ist bereits abgeschlossen. Für den Induktionsschritt  $s \mapsto s + 1$  gilt als Voraussetzung:

$$\sum_{x=t}^{t+s-1} (1 - \theta_0)^x - \sum_{x=t+1}^{t+s} (x - t - 1)\theta_0(1 - \theta_0)^{x-2} - s \left(1 - \sum_{x=1}^{t+s-1} \theta_0(1 - \theta_0)^{x-1}\right) = 0$$

Darauf basierend muss folglich

$$\sum_{x=t}^{t+s} (1 - \theta_0)^x - \sum_{x=t+1}^{t+s+1} (x - t - 1)\theta_0(1 - \theta_0)^{x-2} - (s + 1) \left(1 - \sum_{x=1}^{t+s} \theta_0(1 - \theta_0)^{x-1}\right) = 0 \quad (3.30)$$

gezeigt werden. Es gilt:

(i)

$$\sum_{x=t}^{t+s} (1 - \theta_0)^x = \sum_{x=t}^{t+s-1} (1 - \theta_0)^x + (1 - \theta_0)^{t+s}$$

(ii)

$$\sum_{x=t+1}^{t+s+1} (x - t - 1)\theta_0(1 - \theta_0)^{x-2} = \sum_{x=t+1}^{t+s} (x - t - 1)\theta_0(1 - \theta_0)^{x-2} + s\theta_0(1 - \theta_0)^{t+s-1}$$

(iii)

$$\begin{aligned} & (s+1) \left( 1 - \sum_{x=1}^{t+s} \theta_0 (1-\theta_0)^{x-1} \right) \\ &= s \left( 1 - \sum_{x=1}^{t+s-1} \theta_0 (1-\theta_0)^{x-1} \right) - s \theta_0 (1-\theta_0)^{t+s-1} + \left( 1 - \sum_{x=1}^{t+s} \theta_0 (1-\theta_0)^{x-1} \right) \end{aligned}$$

Einsetzen von (i), (ii) und (iii) in (3.30) liefert mithilfe der Voraussetzung

$$\begin{aligned} & \sum_{x=t}^{t+s-1} (1-\theta_0)^x - \sum_{x=t+1}^{t+s} (x-t-1) \theta_0 (1-\theta_0)^{x-2} - s \left( 1 - \sum_{x=1}^{t+s-1} \theta_0 (1-\theta_0)^{x-1} \right) \\ &+ s \theta_0 (1-\theta_0)^{t+s-1} - \left( 1 - \sum_{x=1}^{t+s} \theta_0 (1-\theta_0)^{x-1} \right) + (1-\theta_0)^{t+s} - s \theta_0 (1-\theta_0)^{t+s-1} \\ &= s \theta_0 (1-\theta_0)^{t+s-1} - \left( 1 - \sum_{x=1}^{t+s} \theta_0 (1-\theta_0)^{x-1} \right) + (1-\theta_0)^{t+s} - s \theta_0 (1-\theta_0)^{t+s-1} \\ &= (1-\theta_0)^{t+s} - \left( 1 - \sum_{x=1}^{t+s} \theta_0 (1-\theta_0)^{x-1} \right) = 0 \end{aligned}$$

Die letzte Gleichheit folgt dabei direkt aus der Gültigkeit von (3.29). Somit ist gezeigt, dass es sich für alle  $t \in \{1, \dots, G-1\}$  und  $s \in \mathbb{N}$  bei  $\theta_0$  um die Nullstelle von  $\Psi$  handelt.

Für den Beweis von (2) wird zunächst  $\dot{\Psi}$  berechnet:

$$\begin{aligned} \dot{\Psi}(\theta) = & -\frac{1}{G} \sum_{t=0}^{G-1} \left[ \frac{\theta_0 \left( \sum_{x=t}^{t+s-1} (1-\theta_0)^x \right)}{\theta^2} \right. \\ & \left. + \frac{\sum_{x=t+1}^{t+s} (x-t-1) \theta_0 (1-\theta_0)^{x-1} + s \left( 1 - \sum_{x=1}^{t+s} \theta_0 (1-\theta_0)^{x-1} \right)}{(1-\theta)^2} \right] \end{aligned}$$

Da sowohl  $\theta$  als auch  $\theta_0$  in dem Intervall  $[\xi, 1-\xi]$  enthalten sind, lassen sich die Summanden folgendermaßen abschätzen:

$$\begin{aligned} & \frac{\theta_0 \left( \sum_{x=t}^{t+s-1} (1-\theta_0)^x \right)}{\theta^2} > \frac{\xi \left( \sum_{x=t}^{t+s-1} \xi^x \right)}{1-\xi} > 0 \\ & \frac{\sum_{x=t+1}^{t+s} (x-t-1) \theta_0 (1-\theta_0)^{x-1}}{(1-\theta)^2} \geq \frac{\sum_{x=t+1}^{t+s} (x-t-1) \xi (1-\xi)^{x-1}}{(1-\xi)^2} \geq 0 \\ & \frac{s \left( 1 - \sum_{x=1}^{t+s} \theta_0 (1-\theta_0)^{x-1} \right)}{(1-\theta)^2} > \frac{s(1-1)}{(1-\xi)^2} = 0 \end{aligned}$$

Dies impliziert  $\dot{\Psi}(\theta) < 0$  für alle  $\theta \in \Theta$ . Die Funktion  $\Psi(\theta)$  ist folglich auf dem ganzen

Intervall  $[\xi, 1 - \xi]$  streng monoton fallend.

Zusammenfassend folgt aus (1) und (2), dass für alle  $\epsilon > 0$  die Eigenschaft  $\Psi(\theta_0 - \epsilon) > 0 > \Psi(\theta_0 + \epsilon)$  erfüllt ist.

Satz 7 ist folglich anwendbar und die Schätzfolge  $(\hat{\theta}_n)_n$  konvergiert in Wahrscheinlichkeit gegen den wahren Parameterwert  $\theta_0$ . Der Konsistenznachweis ist somit erbracht.

### 3.2.2. Asymptotische Normalität

In diesem Kapitel werden die asymptotische Normalität sowie die klassischen Bedingungen für diese formuliert [siehe van der Vaart, 1998, Kapitel 5.3 und 5.6], da über diesen Weg direkt ein Konfidenzintervall für den wahren Parameterwert der geometrischen Verteilung  $\theta_0$  hergeleitet werden kann.

Im vergangenen Kapitel wurde mithilfe des Konsistenzbegriffs die Frage beantwortet, ob die Schätzfolge  $(\hat{\theta}_n)_n$  gegen den wahren Parameterwert  $\theta_0$  in Wahrscheinlichkeit konvergiert. Im Folgenden wird sich der Frage nach der Konvergenzrate von  $(\hat{\theta}_n - \theta_0)_n$  gegen 0 angenommen. Die Antwort auf diese Frage ist situationsabhängig, jedoch liegt die Ordnung sehr oft, wenn die Schätzer auf der  $n$ -fachen Wiederholung eines Experiments basieren, bei  $n^{-1/2}$ . Wird  $(\hat{\theta}_n - \theta_0)_n$  mit der inversen Rate  $\sqrt{n}$  multipliziert, ergibt sich ein Gleichgewicht, das in den meisten Fällen in Verteilung gegen eine Normalverteilung konvergiert. Dieses Phänomen wird asymptotische Normalität genannt.

Analog zu vorherigen Kapiteln werden die auf dem gemeinsamen Wahrscheinlichkeitsraum  $(\Omega, \mathcal{F}, P_{\theta_0})$  (vgl. (2.1) und folgende Seiten) definierten bivariaten Zufallsvariablen  $(X_1, T_1)', \dots, (X_n, T_n)'$  (vgl. 2.2 und 2.3) betrachtet.

Zunächst beschreibt der folgende Satz die klassischen Bedingungen für die asymptotische Normalität [siehe van der Vaart, 1998, Theorem 5.41].

**Satz 8.** *Für alle  $\theta \in \Theta$ , wobei  $\Theta$  eine Teilmenge des eindimensionalen euklidischen Raums darstellt, sei die Abbildung  $\theta \mapsto \psi_\theta(x, t)$  für alle  $x$  und für alle  $t$  zweimal stetig differenzierbar. Es wird angenommen, dass  $E_{\theta_0}\{\psi_{\theta_0}\} = 0$  sowie  $E_{\theta_0}\|\psi_{\theta_0}\|^2 < \infty$ . Der Erwartungswert  $E_{\theta_0}\{\dot{\psi}_{\theta_0}\}$  existiere mit  $E_{\theta_0}\{\dot{\psi}_{\theta_0}\} \neq 0$ . Außerdem erfüllt die Ableitung zweiter Ordnung die Ungleichung*

$$\left| \frac{\partial^2 \psi_\theta(x, t)}{\partial \theta^2} \right| \leq \ddot{\psi}(x, t) \quad (3.31)$$

für alle  $\theta$  in der Nähe von  $\theta_0$  und für eine integrierbare und messbare Funktion  $\ddot{\psi}$ . Des Weiteren sei  $(\hat{\theta}_n)_n$  eine konsistente Schätzfolge, die  $\Psi_n(\hat{\theta}_n) = 0$  für alle  $n$  erfüllt. Dann ist die Sequenz  $(\sqrt{n}(\hat{\theta}_n - \theta_0))_n$  asymptotisch normalverteilt mit dem Erwartungswert 0 und der Varianz:

$$\sigma^2 = \frac{E_{\theta_0}\{\psi_{\theta_0}^2\}}{(E_{\theta_0}\{\dot{\psi}_{\theta_0}\})^2} \quad (3.32)$$

Im Folgenden wird der Punktschätzer des geometrischen Parameters bei linksseitig trunziertem und rechtsseitig zensiertem Beobachtungsdesign hinsichtlich der asympto-

tischen Normalität mithilfe von Satz 8 untersucht und gegebenenfalls die Varianz der resultierenden Normalverteilung berechnet. Dazu werden zunächst alle Bedingungen von Satz 8 nachgewiesen, beginnend mit der zweifachen stetigen Differenzierbarkeit von  $\theta \mapsto \psi_\theta(x, t)$  für alle Tupel  $(x, t)$ . Die Funktion  $\theta \mapsto \psi_\theta(x, t)$  wurde bereits in Kapitel 3.2.1 hergeleitet, sodass für diesen Abschnitt auf die Gleichung (3.18) zurückgegriffen werden kann:

$$\psi_\theta(X_i, T_i) = \frac{\partial m_\theta(X_i, T_i)}{\partial \theta} = \mathbf{1}_{\{T_i+1 \leq X_i\}} \left[ \frac{\mathbf{1}_{\{X_i \leq T_i+s\}}}{\theta} - \frac{\min(X_i - 1, T_i + s) - T_i}{1 - \theta} \right]$$

Die partiellen Ableitungen erster und zweiter Ordnung in Richtung des Wahrscheinlichkeitsparameters  $\theta$  haben entsprechend folgende Gestalt:

$$\dot{\psi}_\theta(X_i, T_i) = \frac{\partial \psi_\theta(X_i, T_i)}{\partial \theta} = -\mathbf{1}_{\{T_i+1 \leq X_i\}} \left[ \frac{\mathbf{1}_{\{X_i \leq T_i+s\}}}{\theta^2} + \frac{\min(X_i - 1, T_i + s) - T_i}{(1 - \theta)^2} \right] \quad (3.33)$$

$$\ddot{\psi}_\theta(X_i, T_i) = \frac{\partial^2 \psi_\theta(X_i, T_i)}{\partial \theta^2} = \mathbf{1}_{\{T_i+1 \leq X_i\}} \left[ \frac{2\mathbf{1}_{\{X_i \leq T_i+s\}}}{\theta^3} - \frac{2(\min(X_i - 1, T_i + s) - T_i)}{(1 - \theta)^3} \right] \quad (3.34)$$

Die Funktionen  $\dot{\psi}_\theta(X_i, T_i)$  und  $\ddot{\psi}_\theta(X_i, T_i)$  sind für alle  $\theta \in \Theta = [\xi, 1 - \xi]$  als Kompositionen von stetigen Funktionen selbst auch wieder stetig.

Um  $E_{\theta_0}\{\psi_{\theta_0}\} = 0$  einzusehen, sei angemerkt, dass  $E_{\theta_0}\{\psi_{\theta_0}\} = \Psi(\theta_0)$  gilt und  $\Psi(\theta_0) = 0$  bereits gezeigt wurde (vgl. (3.27) und die folgenden Seiten).

Die nächste zu überprüfende Bedingung von Satz 8 ist  $E_{\theta_0}\|\psi_{\theta_0}\|^2 < \infty$ . Da im betrachteten Modell die wahre Ausfallwahrscheinlichkeit  $\theta_0$  einen eindimensionalen Parameterwert darstellt, ergibt sich  $E_{\theta_0}\|\psi_{\theta_0}\|^2 = E_{\theta_0}\{\psi_{\theta_0}^2\}$ . Weiter gilt:

$$\begin{aligned} E_{\theta_0}\{\psi_{\theta_0}^2\} &= E_{\theta_0} \left\{ \left[ \mathbf{1}_{\{T_i+1 \leq X_i\}} \left[ \frac{\mathbf{1}_{\{X_i \leq T_i+s\}}}{\theta_0} - \frac{\min(X_i - 1, T_i + s) - T_i}{1 - \theta_0} \right] \right]^2 \right\} \\ &= \frac{1}{G} \sum_{t=0}^{G-1} E_{\theta_0} \left\{ \mathbf{1}_{\{t+1 \leq X_i\}} \left[ \left[ \frac{\mathbf{1}_{\{X_i \leq t+s\}}}{\theta_0} - \frac{\min(X_i - 1, t + s) - t}{1 - \theta_0} \right] \right]^2 \right\} \\ &= \frac{1}{G} \sum_{t=0}^{G-1} \left[ \sum_{x=1}^{\infty} \theta_0(1 - \theta_0)^{x-1} \mathbf{1}_{\{t+1 \leq x\}} \left[ \frac{(1 - \theta_0)\mathbf{1}_{\{x \leq t+s\}} - \theta_0(\min(x - 1, t + s) - t)}{\theta_0(1 - \theta_0)} \right]^2 \right] \\ &= \frac{1}{G} \sum_{t=0}^{G-1} \left[ \sum_{x=t+1}^{\infty} \theta_0^{-1}(1 - \theta_0)^{x-3} \left( (1 - \theta_0)^2 \mathbf{1}_{\{x \leq t+s\}} \right. \right. \\ &\quad \left. \left. - 2(1 - \theta_0)\mathbf{1}_{\{x \leq t+s\}}\theta_0(\min(x - 1, t + s) - t) + \theta_0^2(\min(x - 1, t + s) - t)^2 \right) \right] \end{aligned}$$

Die einzelnen Summanden lassen sich dabei folgendermaßen vereinfachen:

(i)

$$\sum_{x=t+1}^{\infty} \theta_0^{-1}(1-\theta_0)^{x-3}(1-\theta_0)^2 \mathbf{1}_{\{x \leq t+s\}} = \sum_{x=t+1}^{t+s} \theta_0^{-1}(1-\theta_0)^{x-1}$$

(ii)

$$\begin{aligned} & \sum_{x=t+1}^{\infty} \theta_0^{-1}(1-\theta_0)^{x-3} 2(1-\theta_0) \mathbf{1}_{\{x \leq t+s\}} \theta_0 (\min(x-1, t+s) - t) \\ &= \sum_{x=t+1}^{t+s} 2(1-\theta_0)^{x-2} (\min(x-1, t+s) - t) = \sum_{x=t+1}^{t+s} 2(1-\theta_0)^{x-2} (x-t-1) \end{aligned}$$

(iii)

$$\begin{aligned} & \sum_{x=t+1}^{\infty} \theta_0^{-1}(1-\theta_0)^{x-3} \theta_0^2 (\min(x-1, t+s) - t)^2 \\ &= \sum_{x=t+2}^{\infty} \theta_0 (1-\theta_0)^{x-3} (\min(x-1, t+s) - t)^2 \\ &= \sum_{x=t+2}^{t+s+1} \theta_0 (1-\theta_0)^{x-3} (x-t-1)^2 + s^2 \sum_{x=t+s+2}^{\infty} \theta_0 (1-\theta_0)^{x-3} \\ &= \sum_{x=t+2}^{t+s+1} \theta_0 (1-\theta_0)^{x-3} (x-t-1)^2 + s^2 \left( 1 - \mathbf{1}_{\{t+s \geq 2\}} \sum_{x=1}^{t+s-1} \theta_0 (1-\theta_0)^{x-1} \right) \end{aligned}$$

Einsetzen liefert:

$$\begin{aligned} \mathbb{E}_{\theta_0} \{ \psi_{\theta_0}^2 \} &= \frac{1}{G} \sum_{t=0}^{G-1} \left[ \sum_{x=t+1}^{t+s} \theta_0^{-1}(1-\theta_0)^{x-1} \right. \\ & \quad - 2 \sum_{x=t+1}^{t+s} (1-\theta_0)^{x-2} (x-t-1) \\ & \quad + \sum_{x=t+2}^{t+s+1} \theta_0 (1-\theta_0)^{x-3} (x-t-1)^2 \\ & \quad \left. + s^2 \left( 1 - \mathbf{1}_{\{t+s \geq 2\}} \sum_{x=1}^{t+s-1} \theta_0 (1-\theta_0)^{x-1} \right) \right] \quad (3.35) \end{aligned}$$

Auf der Grundlage von (3.35) kann  $\mathbb{E}_{\theta_0} \|\psi_{\theta_0}\|^2$  abgeschätzt werden. Unter der Verwen-

dung von  $0 < \xi \leq \theta_0 \leq 1 - \xi < 1$  folgt für die einzelnen Summanden:

$$\sum_{x=t+1}^{t+s} \theta_0^{-1} (1 - \theta_0)^{x-1} < \frac{\sum_{x=t+1}^{t+s} 1}{\xi} = \frac{s}{\xi} \quad (3.36)$$

$$-2 \sum_{x=t+1}^{t+s} (1 - \theta_0)^{x-2} (x - t - 1) \leq 0 \quad (3.37)$$

$$\sum_{x=t+2}^{t+s+1} \theta_0 (1 - \theta_0)^{x-3} (x - t - 1)^2 \leq s^2 \frac{\sum_{x=t}^{t+s-1} (1 - \theta_0)^x}{1 - \theta_0} \leq \frac{s^2 (s - 1)}{\xi} \quad (3.38)$$

$$s^2 \left(1 - \mathbb{1}_{\{t+s \geq 2\}} \sum_{x=1}^{t+s-1} \theta_0 (1 - \theta_0)^{x-1}\right) \leq s^2 (1 - 0) = s^2 \quad (3.39)$$

Durch Einsetzen von (3.36)-(3.39) in (3.35) folgt mit  $s < \infty$ :

$$\mathbb{E}_{\theta_0} \|\psi_{\theta_0}\|^2 < \frac{1}{G} \sum_{t=0}^{G-1} \left[ \frac{s}{\xi} - 0 + \frac{s^2 (s - 1)}{\xi} + s^2 \right] = \frac{s}{\xi} + \frac{s^2 (s - 1)}{\xi} + s^2 < \infty$$

Im nächsten Schritt sollen Existenz und Invertierbarkeit von  $\mathbb{E}_{\theta} \{\dot{\psi}_{\theta_0}\}$  nachgewiesen werden. Unter der Verwendung von (3.33) folgt:

$$\begin{aligned} \mathbb{E}_{\theta_0} \{\dot{\psi}_{\theta_0}\} &= -\mathbb{E}_{\theta_0} \left\{ \mathbb{1}_{\{T_i+1 \leq X_i\}} \left[ \frac{\mathbb{1}_{\{X_i \leq T_i+s\}}}{\theta_0^2} + \frac{\min(X_i - 1, T_i + s) - T_i}{(1 - \theta_0)^2} \right] \right\} \\ &= \frac{-1}{G} \sum_{t=0}^{G-1} \mathbb{E}_{\theta_0} \left\{ \mathbb{1}_{\{t+1 \leq X_i\}} \left[ \frac{\mathbb{1}_{\{X_i \leq t+s\}}}{\theta_0^2} + \frac{\min(X_i - 1, t + s) - t}{(1 - \theta_0)^2} \right] \right\} \\ &= \frac{-1}{G} \sum_{t=0}^{G-1} \sum_{x=t+1}^{\infty} \theta_0 (1 - \theta_0)^{x-1} \left[ \frac{\mathbb{1}_{\{x \leq t+s\}}}{\theta_0^2} + \frac{\min(x - 1, t + s) - t}{(1 - \theta_0)^2} \right] \end{aligned} \quad (3.40)$$

wobei

$$\sum_{x=t+1}^{\infty} \theta_0 (1 - \theta_0)^{x-1} \frac{\mathbb{1}_{\{x \leq t+s\}}}{\theta_0^2} = \sum_{x=t+1}^{t+s} \theta_0^{-1} (1 - \theta_0)^{x-1} \quad (3.41)$$

und

$$\begin{aligned} &\sum_{x=t+1}^{\infty} \theta_0 (1 - \theta_0)^{x-1} \frac{\min(x - 1, t + s) - t}{(1 - \theta_0)^2} \\ &= \sum_{x=t+2}^{\infty} \theta_0 (1 - \theta_0)^{x-3} (\min(x - 1, t + s) - t) \\ &= \sum_{x=t+2}^{t+s+1} \theta_0 (1 - \theta_0)^{x-3} (x - t - 1) + s \sum_{x=t+s+2}^{\infty} \theta_0 (1 - \theta_0)^{x-3} \\ &= \sum_{x=t+2}^{t+s+1} \theta_0 (1 - \theta_0)^{x-3} (x - t - 1) + s \left(1 - \mathbb{1}_{\{t+s \geq 2\}} \sum_{x=1}^{t+s-1} \theta_0 (1 - \theta_0)^{x-1}\right) \end{aligned} \quad (3.42)$$

Einsetzen von (3.41) und (3.42) in (3.40) liefert:

$$\begin{aligned} \mathbb{E}_{\theta_0}\{\dot{\psi}_{\theta_0}\} = \frac{-1}{G} \sum_{t=0}^{G-1} \left[ \sum_{x=t+1}^{t+s} \theta_0^{-1}(1-\theta_0)^{x-1} \right. \\ \left. + \sum_{x=t+2}^{t+s+1} \theta_0(1-\theta_0)^{x-3}(x-t-1) + s \left(1 - \mathbb{1}_{\{t+s \geq 2\}} \sum_{x=1}^{t+s-1} \theta_0(1-\theta_0)^{x-1}\right) \right] \end{aligned} \quad (3.43)$$

Damit ist die Existenz von  $\mathbb{E}_{\theta_0}\{\dot{\psi}_{\theta_0}\}$  gezeigt. Zum Nachweis der Invertierbarkeit von (3.43) genügt es  $\mathbb{E}_{\theta_0}\{\dot{\psi}_{\theta_0}\} \neq 0$  einzusehen. Da einerseits  $0 < \theta_0 < 1$  und somit auch  $0 < (1 - \theta_0)$  gilt und andererseits

$$\mathbb{1}_{\{t+s \geq 2\}} \sum_{x=1}^{t+s-1} \theta_0(1-\theta_0)^{x-1} < 1$$

und somit

$$s \left(1 - \mathbb{1}_{\{t+s \geq 2\}} \sum_{x=1}^{t+s-1} \theta_0(1-\theta_0)^{x-1}\right) > 0$$

erfüllt ist, folgt direkt  $\mathbb{E}_{\theta_0}\{\dot{\psi}_{\theta_0}\} < 0$  und somit die Invertierbarkeit von (3.43).

Abschließend ist die Existenz einer integrierbaren und messbaren Funktion  $\ddot{\psi}(x, t)$  nachzuweisen, die in der Nähe des wahren Parameters  $\theta_0$  eine obere Schranke von  $|\ddot{\psi}_{\theta}(x, t)|$  darstellt. Mithilfe von (3.34) sowie  $0 < \xi \leq \theta \leq 1 - \xi < 1$  gilt:

$$\begin{aligned} |\ddot{\psi}_{\theta}(x, t)| &= \left| \mathbb{1}_{\{t+1 \leq x\}} \left[ \frac{2\mathbb{1}_{\{x \leq t+s\}}}{\theta^3} - \frac{2(\min(x-1, t+s) - t)}{(1-\theta)^3} \right] \right| \\ &\leq \left| \frac{2\mathbb{1}_{\{t+1 \leq x\}}\mathbb{1}_{\{x \leq t+s\}}}{\theta^3} \right| + \left| \frac{2\mathbb{1}_{\{t+1 \leq x\}}(\min(x-1, t+s) - t)}{(1-\theta)^3} \right| \\ &< \frac{2}{\xi^3} + \frac{2|\min(x-1, t+s) - t|}{\xi^3} = \frac{2(|\min(x-1, t+s) - t| + 1)}{\xi^3} =: \ddot{\psi}(x, t) \end{aligned}$$

Da sowohl das Minimum als auch die Betragsfunktion stetige Abbildungen sind, folgt die Stetigkeit von  $\ddot{\psi}(x, t)$ . Damit ist die Funktion  $\ddot{\psi}(x, t)$  messbar und integrierbar.

Somit sind alle Bedingungen von Satz 8 nachgewiesen. Für die konsistente Schätzfolge  $(\hat{\theta}_n)_n$  nach (3.6) gilt somit, dass die Sequenz  $(\sqrt{n}(\hat{\theta} - \theta_0))_n$  asymptotisch normalverteilt mit Erwartungswert 0 und Varianz

$$\sigma^2 = \frac{\mathbb{E}_{\theta_0}\{\dot{\psi}_{\theta_0}^2\}}{(\mathbb{E}_{\theta_0}\{\dot{\psi}_{\theta_0}\})^2}$$

ist. Mithilfe von (3.35) und (3.43) folgt abschließend:

$$\begin{aligned} \sigma^2 = & G \left[ \sum_{t=0}^{G-1} \left[ \sum_{x=t+1}^{t+s} \theta_0^{-1} (1 - \theta_0)^{x-1} + \sum_{x=t+2}^{t+s+1} \theta_0 (1 - \theta_0)^{x-3} (x - t - 1) \right. \right. \\ & \left. \left. + s \left( 1 - \mathbb{1}_{\{t+s \geq 2\}} \sum_{x=1}^{t+s-1} \theta_0 (1 - \theta_0)^{x-1} \right) \right] \right]^{-2} \\ & \left[ \sum_{t=0}^{G-1} \left[ \sum_{x=t+1}^{t+s} \theta_0^{-1} (1 - \theta_0)^{x-1} - 2 \sum_{x=t+1}^{t+s} (1 - \theta_0)^{x-2} (x - t - 1) \right. \right. \\ & \left. \left. + \sum_{x=t+2}^{t+s+1} \theta_0 (1 - \theta_0)^{x-3} (x - t - 1)^2 + s^2 \left( 1 - \mathbb{1}_{\{t+s \geq 2\}} \sum_{x=1}^{t+s-1} \theta_0 (1 - \theta_0)^{x-1} \right) \right] \right] \quad (3.44) \end{aligned}$$

Davon ausgehend kann im Folgenden der Standardfehler des Punktschätzers  $\hat{\theta}_n$  berechnet werden. Dieser wird zunächst auf zwei verschiedene Arten ermittelt.

Für den ersten Berechnungsansatz gilt zunächst, basierend auf (3.13):

$$SE_1^2 = \frac{\hat{\sigma}^2}{n}$$

Dabei ergibt sich  $\hat{\sigma}^2$  aus (3.44), wobei der unbekannte wahre Parameterwert  $\theta_0$  durch die entsprechende Punktschätzung  $\hat{\theta}_n$  (vgl. (3.9)) ersetzt wurde.

Das Problem, das dieser Variante zur Standardfehlerberechnung zugrunde liegt, ist, dass von Individuen, die vor dem Beginn des Beobachtungszeitraums ausfallen, nicht einmal die Existenz beobachtet wird, sodass  $n$  unbekannt ist und dementsprechend geschätzt werden muss. Somit gilt  $SE_1^2 = \hat{\sigma}^2 / \hat{n}$ . Dazu sei  $\alpha_{\theta_0} := P_{\theta_0} \{T_i + 1 \leq X_i\}$  die Wahrscheinlichkeit für das Ereignis, dass ein zufälliges Individuum aus der Stichprobe auch sichtbar ist.

Die Anzahl der Beobachtungen ist folglich zufällig mit der Realisierung  $m$ . Somit resultiert das Verhältnis

$$m = n\alpha_{\theta_0}.$$

beziehungsweise

$$n = \frac{m}{\alpha_{\theta_0}}.$$

Folglich beschreibt  $m$  die Anzahl aller Individuen, die während oder nach der Studie ausfallen.

Mithilfe von (2.5) und (3.14) kann  $\alpha_{\theta_0}$  folgendermaßen berechnet werden:

$$\begin{aligned}
 \alpha_{\theta_0} = P_{\theta_0}\{T_i + 1 \leq X_i\} &= \sum_{t=0}^{G-1} \sum_{x=t+1}^{\infty} f_U(t) f_G(x; \theta_0) \\
 &= \frac{1}{G} \sum_{t=0}^{G-1} \sum_{x=t+1}^{\infty} \theta_0 (1 - \theta_0)^{x-1} \\
 &= \frac{\theta_0}{G} \sum_{t=0}^{G-1} \left[ \sum_{x=0}^{\infty} (1 - \theta_0)^x - \sum_{x=0}^{t-1} (1 - \theta_0)^x \right] \\
 &= \frac{\theta_0}{G} \sum_{t=0}^{G-1} \left[ \frac{1}{\theta_0} - \frac{1 - (1 - \theta_0)^t}{\theta_0} \right] \\
 &= \frac{1}{G} \sum_{t=0}^{G-1} (1 - \theta_0)^t = \frac{1 - (1 - \theta_0)^G}{\theta_0 G} \tag{3.45}
 \end{aligned}$$

Der geschätzte Wert von  $n$  wird dann mithilfe von

$$\hat{n} = \frac{m}{\alpha_{\hat{\theta}_n}}$$

berechnet. Zusammenfassend folgt:

$$\begin{aligned}
 \text{SE}_1^2 &= \frac{\hat{\sigma}^2}{\hat{n}} = \frac{\alpha_{\hat{\theta}_n} \hat{\sigma}^2}{m} = \frac{1}{m} \frac{1 - (1 - \hat{\theta}_n)^G}{\hat{\theta}_n} \\
 &\quad \left[ \sum_{t=0}^{G-1} \left[ \sum_{x=t+1}^{t+s} \hat{\theta}_n^{-1} (1 - \hat{\theta}_n)^{x-1} + \sum_{x=t+2}^{t+s+1} \hat{\theta}_n (1 - \hat{\theta}_n)^{x-3} (x - t - 1) \right. \right. \\
 &\quad \left. \left. + s \left( 1 - \mathbf{1}_{\{t+s \geq 2\}} \sum_{x=1}^{t+s-1} \hat{\theta}_n (1 - \hat{\theta}_n)^{x-1} \right) \right] \right]^{-2} \\
 &\quad \left[ \sum_{t=0}^{G-1} \left[ \sum_{x=t+1}^{t+s} \hat{\theta}_n^{-1} (1 - \hat{\theta}_n)^{x-1} - 2 \sum_{x=t+1}^{t+s} (1 - \hat{\theta}_n)^{x-2} (x - t - 1) \right. \right. \\
 &\quad \left. \left. + \sum_{x=t+2}^{t+s+1} \hat{\theta}_n (1 - \hat{\theta}_n)^{x-3} (x - t - 1)^2 + s^2 \left( 1 - \mathbf{1}_{\{t+s \geq 2\}} \sum_{x=1}^{t+s-1} \hat{\theta}_n (1 - \hat{\theta}_n)^{x-1} \right) \right] \right] \tag{3.46}
 \end{aligned}$$

Alternativ kann die Schätzung von  $n$  zur Berechnung des Standardfehlers umgangen werden, wenn im Gegenzug auf die Darstellung (3.44) der asymptotischen Varianz  $\sigma^2$  verzichtet wird. Stattdessen wird sich, basierend auf (3.13) und (3.32), dem Verhältnis

$$\text{Var}_{\theta_0}\{\hat{\theta}_n\} = \frac{1}{n} \text{Var}_{\theta_0}\{\sqrt{n}\hat{\theta}_n\} = \frac{\sigma^2}{n} = \frac{1}{n} \frac{\text{E}_{\theta_0}\{\psi_{\theta_0}^2\}}{[\text{E}_{\theta_0}\{\dot{\psi}_{\theta_0}\}]^2}$$

bedient. Da  $\theta_0$  als wahrer Parameterwert der Grundgesamtheit unbekannt und folglich eine exakte Berechnung der Erwartungswerte  $\text{E}_{\theta_0}\{\psi_{\theta_0}^2\}$  und  $\text{E}_{\theta_0}\{\dot{\psi}_{\theta_0}\}$  nicht möglich ist, werden zunächst entsprechende Schätzwerte ermittelt. Zu diesem Zweck wird erneut  $\theta_0$

durch die Punktschätzung  $\hat{\theta}_n$  ersetzt (vgl. (3.6)). Außerdem erfolgt eine Approximation der Erwartungswerte mithilfe der zugehörigen arithmetischen Mittel, sodass sich der quadrierte Standardfehler folgendermaßen berechnen lässt:

$$\begin{aligned} \text{SE}_2^2 &= \frac{1}{n} \frac{\hat{\mathbb{E}}_{\theta_0} \{\psi_{\theta_0}^2\}}{[\hat{\mathbb{E}}_{\theta_0} \{\dot{\psi}_{\theta_0}\}]^2} = \frac{1}{n} \frac{\frac{1}{n} \sum_{i=1}^n \psi_{\hat{\theta}_n}^2(X_i, T_i)}{\left[\frac{1}{n^2} \left[\sum_{i=1}^n \dot{\psi}_{\hat{\theta}_n}(X_i, T_i)\right]^2\right]} = \frac{\sum_{i=1}^n \psi_{\hat{\theta}_n}^2(X_i, T_i)}{\left[\sum_{i=1}^n \dot{\psi}_{\hat{\theta}_n}(X_i, T_i)\right]^2} \\ &= \left[ \sum_{i=1}^n \mathbb{1}_{\{T_i+1 \leq X_i\}} \left[ \frac{\mathbb{1}_{\{X_i \leq T_i+s\}}}{\hat{\theta}_n} - \frac{\min(X_i-1, T_i+s) - T_i}{1 - \hat{\theta}_n} \right]^2 \right] \\ &\quad \left[ \sum_{i=1}^n \mathbb{1}_{\{T_i+1 \leq X_i\}} \left[ \frac{\mathbb{1}_{\{X_i \leq T_i+s\}}}{\hat{\theta}_n^2} + \frac{\min(X_i-1, T_i+s) - T_i}{(1 - \hat{\theta}_n)^2} \right] \right]^{-2} \end{aligned}$$

Diese Art der Berechnung geht mit dem Vorteil einher, dass der unbekannte Stichprobenumfang  $n$  entfällt und folglich nicht geschätzt werden muss. Sei dazu  $m$  die Realisierung der zufälligen Beobachtungsanzahl und somit die Anzahl aller Individuen mit  $T_i + 1 \leq X_i, i \in 1, \dots, n$ . Außerdem werden die sichtbaren Individuen mithilfe der bivariaten Zufallsvariablen  $(\tilde{X}_j, \tilde{T}_j)'$  mit  $j \in \{1, \dots, m\}$  beschrieben. Dann gilt:

$$\begin{aligned} \text{SE}_2^2 &= \left[ \sum_{j=1}^m \left[ \frac{\mathbb{1}_{\{\tilde{X}_j \leq \tilde{T}_j+s\}}}{\hat{\theta}_n} - \frac{\min(\tilde{X}_j-1, \tilde{T}_j+s) - \tilde{T}_j}{1 - \hat{\theta}_n} \right]^2 \right] \\ &\quad \left[ \sum_{j=1}^m \left[ \frac{\mathbb{1}_{\{\tilde{X}_j \leq \tilde{T}_j+s\}}}{\hat{\theta}_n^2} + \frac{\min(\tilde{X}_j-1, \tilde{T}_j+s) - \tilde{T}_j}{(1 - \hat{\theta}_n)^2} \right] \right]^{-2} \end{aligned} \quad (3.47)$$

Der Nachteil dieser Methode besteht darin, dass die Erwartungswerte mithilfe des arithmetischen Mittels geschätzt werden müssen und sich die Darstellung der asymptotischen Varianz (3.44) nicht in die Berechnung integrieren lässt. Auf der anderen Seite besteht der Vorteil darin, dass zumindest im Rahmen der Formel (3.47) keine Verteilungsannahme für  $T_i, i \in \{1, \dots, n\}$ , getroffen werden muss. Jedoch lassen sich einige Bedingungen für Konsistenz und Asymptotische Normalität anhand der Sätze 7 und 8 nicht nachweisen, wenn  $T_i, i \in \{1, \dots, n\}$ , keiner Verteilungsannahme unterliegt. So kann beispielsweise die Grenzfunktion  $\Psi$  nach (3.20) nicht explizit berechnet werden, wenn die Verteilung von  $T_i, i \in \{1, \dots, n\}$ , unbekannt ist.

Demgegenüber liefert das folgende Kapitel eine Variante, den Standardfehler von  $\hat{\theta}_n$  gänzlich ohne eine Verteilungsannahme für  $T_i, i \in \{1, \dots, n\}$ , zu berechnen.

### 3.3. Standardfehler im semiparametrischen Modell

Im Folgenden soll unter Zuhilfenahme eines Martingalgrenzwertsatzes von Rebolledo [1979] [siehe Kalbfleisch und Prentice, 2002] eine Darstellung des Standardfehlers von  $\hat{\theta}_n$  (vgl. (3.9)) hergeleitet und dabei auf eine Verteilungsannahme von  $T_i, i \in \{1, \dots, n\}$ , (vgl. (2.3)) verzichtet werden.

Als Ausgangslage dient dabei die bereits ermittelte, marginale und bedingte Likelihood-Funktion (3.1). Diese wird jedoch als stochastischer Prozess formuliert, wobei die bis-

her beliebig gewählte, aber feste, Beobachtungsgrenze  $\chi$  als weiteres Argument in die marginale, bedingte Likelihood-Funktion aufgenommen wird und somit für weitere Betrachtungen die Zeit-Richtung repräsentiert. Somit wird für das folgende Kapitel  $\chi \in \mathbb{N}_0$  angenommen. Zusammenfassend gilt unter der Annahme, dass die Lebensdauer  $X_i$ ,  $i \in \{1, \dots, n\}$ , weiterhin geometrisch verteilt ist (vgl. Definition 1):

$${}_T L^c(\chi, \theta) = \prod_{i=1}^n \left[ \theta^{\mathbb{1}_{\{X_i \leq \min(T_i + s, \chi)\}}} (1 - \theta)^{\min(X_i - 1, T_i + s, \chi) - T_i} \right]^{\mathbb{1}_{\{T_i + 1 \leq \chi\}}}$$

Darauf aufbauend wird der Prozess  ${}_T V^c$  folgendermaßen definiert:

$$\begin{aligned} {}_T V^c(\chi, \theta) &:= \frac{1}{\sqrt{n}} \frac{\partial \log {}_T L^c(\chi, \theta)}{\partial \theta} = \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{\partial \log {}_T L_i^c(\chi, \theta)}{\partial \theta} \\ &= \frac{1}{\sqrt{n}} \sum_{i=1}^n \mathbb{1}_{\{T_i + 1 \leq \chi\}} \left[ \frac{\mathbb{1}_{\{T_i + 1 \leq X_i \leq \min(T_i + s, \chi)\}}}{\theta} - \frac{\mathbb{1}_{\{T_i + 1 \leq X_i\}} (\min(X_i - 1, T_i + s, \chi) - T_i)}{1 - \theta} \right] \end{aligned}$$

Mithilfe von (vgl. (3.5))

$$\begin{aligned} &\frac{\mathbb{1}_{\{T_i + 1 \leq X_i\}} (\min(X_i - 1, T_i + s, \chi) - T_i)}{1 - \theta} + \frac{\mathbb{1}_{\{T_i + 1 \leq X_i \leq \min(T_i + s, \chi)\}}}{1 - \theta} \\ &= \frac{\mathbb{1}_{\{T_i + 1 \leq X_i\}} (\min(X_i, T_i + s, \chi) - T_i)}{1 - \theta} \end{aligned}$$

folgt weiter:

$$\begin{aligned} {}_T V^c(\chi, \theta) &= \frac{1}{\sqrt{n}} \sum_{i=1}^n \mathbb{1}_{\{T_i + 1 \leq \chi\}} \left[ \frac{\mathbb{1}_{\{T_i + 1 \leq X_i \leq \min(T_i + s, \chi)\}}}{\theta} \right. \\ &\quad \left. - \frac{\mathbb{1}_{\{T_i + 1 \leq X_i\}} (\min(X_i, T_i + s, \chi) - T_i)}{1 - \theta} + \frac{\mathbb{1}_{\{T_i + 1 \leq X_i \leq \min(T_i + s, \chi)\}}}{1 - \theta} \right] \end{aligned} \quad (3.48)$$

Des Weiteren lässt sich unter der Verwendung von (3.7) und (3.8) einsehen, dass:

$$\mathbb{1}_{\{T_i + 1 \leq \chi\}} \mathbb{1}_{\{T_i + 1 \leq X_i \leq \min(T_i + s, \chi)\}} = \sum_{x=1}^{\chi} \Delta {}_T N_i^c(x) \quad (3.49)$$

$$\mathbb{1}_{\{T_i + 1 \leq \chi\}} \mathbb{1}_{\{T_i + 1 \leq X_i\}} (\min(X_i, T_i + s, \chi) - T_i) = \sum_{x=1}^{\chi} {}_T Y_i^c(x - 1) \quad (3.50)$$

Einsetzen von (3.49) und (3.50) in (3.48) liefert

$$\begin{aligned} {}_T V^c(\chi, \theta) &= \frac{1}{\sqrt{n}} \sum_{i=1}^n \left[ \sum_{x=1}^{\chi} \frac{\Delta {}_T N_i^c(x)}{\theta} - \sum_{x=1}^{\chi} \frac{{}_T Y_i^c(x-1)}{1-\theta} + \sum_{x=1}^{\chi} \frac{\Delta {}_T N_i^c(x)}{1-\theta} \right] \\ &= \frac{1}{\sqrt{n}} \sum_{i=1}^n \left[ \frac{1}{\theta(1-\theta)} \sum_{x=1}^{\chi} \left( \Delta {}_T N_i^c(x) - \theta {}_T Y_i^c(x-1) \right) \right] \end{aligned}$$

und somit:

$${}_T V^c(\chi, \theta) = \sum_{i=1}^n \sum_{x=1}^{\chi} \frac{1}{\sqrt{n}\theta(1-\theta)} \Delta {}_T M_i^c(x) \quad (3.51)$$

Dies motiviert die Definition:

$$\begin{aligned} {}_T V_i^c(0, \theta) &:= 0 \\ {}_T V_i^c(\chi, \theta) &:= \sum_{x=1}^{\chi} \frac{1}{\sqrt{n}\theta(1-\theta)} \Delta {}_T M_i^c(x) \text{ für } \chi \in \mathbb{N}_0 \end{aligned} \quad (3.52)$$

Im späteren Verlauf soll ein Martingalgrenzwertsatz auf den Prozess  ${}_T V^c$  angewendet werden, da sich aus den damit erzielten Ergebnissen der Standardfehler des Punktschätzers  $\hat{\theta}_n$  berechnen lässt. Dazu ist es notwendig, zunächst eine Filtration zu konstruieren, bezüglich der durch  ${}_T V^c$  ein Martingal definiert ist.

Dazu sei zunächst die Filtration  $\{ {}_T \mathcal{F}_{i,\chi}^c : \chi \in \mathbb{N}_0 \}$  gegeben mit

$$\begin{aligned} {}_T \mathcal{F}_{i,\chi}^c &= \sigma\{\emptyset\} \text{ für } \chi \in \{0, \dots, T_i - 1\} \text{ sowie} \\ {}_T \mathcal{F}_{i,\chi}^c &= \sigma\{ {}_T N_i^c(k), {}_T Y_i^c(k) : k \in \{T_i, \dots, \chi\} \} \end{aligned}$$

Da  ${}_T M_i^c$  nach Lemma 7 ein  $({}_T \mathcal{F}_{i,x}^c, P_{\theta}^{\mathbb{T}})$ -Martingal und  $(\sqrt{n}\theta(1-\theta))^{-1}$  deterministisch und somit vorhersagbar und beschränkt ist, ist der Prozess  ${}_T V_i^c$  ein  $({}_T \mathcal{F}_{i,\chi}^c, P_{\theta}^{\mathbb{T}})$ -Martingal [siehe Fleming und Harrington, 1991, Theorem 1.5.1]. Abschließend liefert das folgende Lemma eine Filtration, bezüglich der die Summe  ${}_T V^c$  ebenfalls ein Martingal ist.

**Lemma 9.** *Sei*

$${}_T V^c(\chi, \theta) = \sum_{i=1}^n {}_T V_i^c(\chi, \theta)$$

, wobei  ${}_T V_i^c(\chi, \theta)$  nach (3.52) für alle  $i \in \{1, \dots, n\}$  ein  $({}_T \mathcal{F}_{i,\chi}^c, P_{\theta}^{\mathbb{T}})$ -Martingal ist. Des Weiteren sei  $\{ {}_T \mathcal{F}_{:, \chi}^c : \chi \in \mathbb{N}_0 \}$  eine Filtration mit

$${}_T \mathcal{F}_{:, \chi}^c := \bigvee_{i=1}^n {}_T \mathcal{F}_{i,\chi}^c = \bigvee_{i=1}^n \sigma\{ {}_T N_i^c(x), {}_T Y_i^c(x) : T_i \leq x \leq \chi \} \quad (3.53)$$

Dann ist der Prozess  ${}_T V^c$  ein  $({}_T \mathcal{F}_{:, \chi}^c, P_{\theta}^{\mathbb{T}})$ -Martingal mit  ${}_T V^c(0, \theta) = 0$ .

*Beweis.* Da für  $\chi = 0$  automatisch  ${}_T V_i^c(0, \theta) = 0$  für alle  $i \in \{1, \dots, n\}$  folgt, ist

${}_T V^c(0, \theta) = 0$  ebenfalls erfüllt.

Des Weiteren folgt die  ${}_T \mathcal{F}_{:, \chi}^c$ -Adaption von  ${}_T V^c$  direkt aus der Tatsache, dass es sich bei  ${}_T V_i^c$  für alle  $i \in \{1, \dots, n\}$  um  ${}_T \mathcal{F}_{i, \chi}^c$ -adaptierte und somit auch um  ${}_T \mathcal{F}_{:, \chi}^c$ -adaptierte Prozesse handelt.

Da  ${}_T V_i^c$  Martingale bezüglich  $({}_T \mathcal{F}_{i, \chi}^c, P_{\theta}^{\tilde{\mathbb{T}}})$  sind, folgt zusätzlich  $E_{\theta}^{\tilde{\mathbb{T}}} |{}_T V_i^c(\chi, \theta)| < \infty$  für alle  $\chi \in \mathbb{N}_0$  und somit auch

$$E_{\theta}^{\tilde{\mathbb{T}}} |{}_T V^c(\chi, \theta)| \leq \sum_{i=1}^n E_{\theta}^{\tilde{\mathbb{T}}} |{}_T V_i^c(\chi, \theta)| < \infty$$

Folglich bleibt zu zeigen, dass

$$E_{\theta}^{\tilde{\mathbb{T}}} \{\Delta {}_T V^c(\chi, \theta) | {}_T \mathcal{F}_{:, \chi-1}^c\} = 0, \forall \chi \in \mathbb{N} \quad (3.54)$$

erfüllt ist.

Zunächst ist (3.54) äquivalent zu:

$$E_{\theta}^{\tilde{\mathbb{T}}} \{\Delta {}_T V^c(\chi, \theta) | \mathbb{F}\} = 0, \forall \mathbb{F} \in {}_T \mathcal{F}_{:, \chi-1}^c$$

Dabei gilt für alle  $\mathbb{F} \in {}_T \mathcal{F}_{:, \chi-1}^c$ , dass:

$$\mathbb{F} = \bigcup_{i=1}^n \mathbb{F}_i, \text{ mit } \mathbb{F}_i \in {}_T \mathcal{F}_{i, \chi-1}^c$$

Dabei gilt  $\mathbb{F}_i = \emptyset$ , wenn  $\chi \leq T_i$ . Entsprechend folgt die Äquivalenz von (3.54) zu:

$$\begin{aligned} E_{\theta}^{\tilde{\mathbb{T}}} \left\{ \Delta {}_T V^c(\chi, \theta) \middle| \bigcup_{i=1}^n \mathbb{F}_i \right\} &= 0, \forall \mathbb{F}_i \in {}_T \mathcal{F}_{i, \chi-1}^c \\ \Leftrightarrow \sum_{i=1}^n E_{\theta}^{\tilde{\mathbb{T}}} \left\{ \Delta {}_T V_i^c(\chi, \theta) \middle| \bigcup_{i=1}^n \mathbb{F}_i \right\} &= 0, \forall \mathbb{F}_i \in {}_T \mathcal{F}_{i, \chi-1}^c \end{aligned} \quad (3.55)$$

Im Folgenden werden die einzelnen Summanden von (3.55) betrachtet. Sei dazu  $\Delta {}_T v_i^c(\chi, \theta)$  die Realisierung von  $\Delta {}_T V_i^c(\chi, \theta)$  für ein  $i \in \{1, \dots, n\}$ . Dann gilt

$$\begin{aligned} E_{\theta}^{\tilde{\mathbb{T}}} \left\{ \Delta {}_T V_i^c(\chi, \theta) \middle| \bigcup_{i=1}^n \mathbb{F}_i \right\} & \\ = \sum_{\Delta {}_T v_i^c(\chi, \theta)} \Delta {}_T v_i^c(\chi, \theta) \frac{P_{\theta}^{\tilde{\mathbb{T}}} \left\{ \left\{ \Delta {}_T V_i^c(\chi, \theta) = \Delta {}_T v_i^c(\chi, \theta) \right\} \cap \left\{ \bigcup_{i=1}^n \mathbb{F}_i \right\} \right\}}{P_{\theta}^{\tilde{\mathbb{T}}} \left\{ \bigcup_{i=1}^n \mathbb{F}_i \right\}} & \end{aligned} \quad (3.56)$$

und somit:

$$\begin{aligned} \mathbb{E}_{\theta}^{\tilde{\mathbb{T}}}\left\{\Delta_{TV_i^c}(\chi, \theta) \middle| \bigcup_{i=1}^n \mathbb{F}_i\right\} &= \sum_{\Delta_{Tv_i^c}(\chi, \theta)} \Delta_{TV_i^c}(\chi, \theta) \left[ \frac{P_{\theta}^{\tilde{\mathbb{T}}}\left\{\left\{\Delta_{TV_i^c}(\chi, \theta) = \Delta_{Tv_i^c}(\chi, \theta)\right\} \cap \mathbb{F}_i\right\}}{P_{\theta}^{\tilde{\mathbb{T}}}\left\{\bigcup_{i=1}^n \mathbb{F}_i\right\}} \right. \\ &\quad \left. + \frac{P_{\theta}^{\tilde{\mathbb{T}}}\left\{\left\{\Delta_{TV_i^c}(\chi, \theta) = \Delta_{Tv_i^c}(\chi, \theta)\right\} \cap \left\{\bigcup_{i=1}^n \mathbb{F}_i \setminus \mathbb{F}_i\right\}\right\}}{P_{\theta}^{\tilde{\mathbb{T}}}\left\{\bigcup_{i=1}^n \mathbb{F}_i\right\}} \right] \end{aligned}$$

Weiter folgt:

$$\begin{aligned} &\mathbb{E}_{\theta}^{\tilde{\mathbb{T}}}\left\{\Delta_{TV_i^c}(\chi, \theta) \middle| \bigcup_{i=1}^n \mathbb{F}_i\right\} \\ &= \sum_{\Delta_{Tv_i^c}(\chi, \theta)} \Delta_{TV_i^c}(\chi, \theta) \frac{P_{\theta}^{\tilde{\mathbb{T}}}\{\mathbb{F}_i\}}{P_{\theta}^{\tilde{\mathbb{T}}}\left\{\bigcup_{i=1}^n \mathbb{F}_i\right\}} \frac{P_{\theta}^{\tilde{\mathbb{T}}}\left\{\left\{\Delta_{TV_i^c}(\chi, \theta) = \Delta_{Tv_i^c}(\chi, \theta)\right\} \cap \mathbb{F}_i\right\}}{P_{\theta}^{\tilde{\mathbb{T}}}\{\mathbb{F}_i\}} \\ &+ \sum_{\Delta_{Tv_i^c}(\chi, \theta)} \Delta_{TV_i^c}(\chi, \theta) \frac{P_{\theta}^{\tilde{\mathbb{T}}}\left\{\left\{\Delta_{TV_i^c}(\chi, \theta) = \Delta_{Tv_i^c}(\chi, \theta)\right\} \cap \left\{\bigcup_{i=1}^n \mathbb{F}_i \setminus \mathbb{F}_i\right\}\right\}}{P_{\theta}^{\tilde{\mathbb{T}}}\left\{\bigcup_{i=1}^n \mathbb{F}_i\right\}} \end{aligned} \quad (3.57)$$

Es bleibt zu zeigen, dass beide Summen von (3.57) gleich 0 sind. Für die erste Summe gilt:

$$\begin{aligned} &\frac{P_{\theta}^{\tilde{\mathbb{T}}}\{\mathbb{F}_i\}}{P_{\theta}^{\tilde{\mathbb{T}}}\left\{\bigcup_{i=1}^n \mathbb{F}_i\right\}} \sum_{\Delta_{Tv_i^c}(\chi, \theta)} \Delta_{TV_i^c}(\chi, \theta) \frac{P_{\theta}^{\tilde{\mathbb{T}}}\left\{\left\{\Delta_{TV_i^c}(\chi, \theta) = \Delta_{Tv_i^c}(\chi, \theta)\right\} \cap \mathbb{F}_i\right\}}{P_{\theta}^{\tilde{\mathbb{T}}}\{\mathbb{F}_i\}} \\ &= \frac{P_{\theta}^{\tilde{\mathbb{T}}}\{\mathbb{F}_i\}}{P_{\theta}^{\tilde{\mathbb{T}}}\left\{\bigcup_{i=1}^n \mathbb{F}_i\right\}} \mathbb{E}_{\theta}^{\tilde{\mathbb{T}}}\left\{\Delta_{TV_i^c}(\chi, \theta) \middle| \mathbb{F}_i\right\} \end{aligned}$$

Da es sich bei  $TV_i^c$  um ein  $(\mathcal{F}_{i,\chi}^c, P_{\theta}^{\tilde{\mathbb{T}}})$ -Martingal handelt, gilt  $\mathbb{E}_{\theta}^{\tilde{\mathbb{T}}}\left\{\Delta_{TV_i^c}(\chi, \theta) \middle| \mathbb{F}_i\right\} = 0$  für alle  $\mathbb{F}_i \in \mathcal{F}_{i,\chi-1}^c$ .

Für die zweite Summe von (3.57) ergibt sich außerdem mithilfe der Rechenregel (i) aus Satz 13 für bedingte Erwartungswerte:

$$\begin{aligned} &\frac{P_{\theta}^{\tilde{\mathbb{T}}}\left\{\bigcup_{i=1}^n \mathbb{F}_i \setminus \mathbb{F}_i\right\}}{P_{\theta}^{\tilde{\mathbb{T}}}\left\{\bigcup_{i=1}^n \mathbb{F}_i\right\}} \sum_{\Delta_{Tv_i^c}(\chi, \theta)} \Delta_{TV_i^c}(\chi, \theta) P_{\theta}^{\tilde{\mathbb{T}}}\left\{\left\{\Delta_{TV_i^c}(\chi, \theta) = \Delta_{Tv_i^c}(\chi, \theta)\right\}\right\} \\ &= \frac{P_{\theta}^{\tilde{\mathbb{T}}}\left\{\bigcup_{i=1}^n \mathbb{F}_i \setminus \mathbb{F}_i\right\}}{P_{\theta}^{\tilde{\mathbb{T}}}\left\{\bigcup_{i=1}^n \mathbb{F}_i\right\}} \mathbb{E}_{\theta}^{\tilde{\mathbb{T}}}\left\{\Delta_{TV_i^c}(\chi, \theta)\right\} \\ &= \frac{P_{\theta}^{\tilde{\mathbb{T}}}\left\{\bigcup_{i=1}^n \mathbb{F}_i \setminus \mathbb{F}_i\right\}}{P_{\theta}^{\tilde{\mathbb{T}}}\left\{\bigcup_{i=1}^n \mathbb{F}_i\right\}} \mathbb{E}_{\theta}^{\tilde{\mathbb{T}}}\left\{\mathbb{E}_{\theta}^{\tilde{\mathbb{T}}}\left\{\Delta_{TV_i^c}(\chi, \theta) \middle| \mathcal{F}_{i,\chi-1}^c\right\}\right\} = 0 \end{aligned}$$

Somit ist auch die Gültigkeit von (3.55) nachgewiesen. Der Prozess  $TV^c$  ist folglich

ein  $({}_T\mathcal{F}_{\cdot,\chi}^c, P_{\theta}^{\tilde{\Pi}})$ -Martingal.  $\square$

**Bemerkung 14.** Aus dem Beweis von Lemma 9 lässt sich ebenso entnehmen, dass es sich bei  ${}_TV_i^c$  nach (3.52) für alle  $i \in \{1, \dots, n\}$  um ein  $({}_T\mathcal{F}_{\cdot,\chi}^c, P_{\theta}^{\tilde{\Pi}})$ -Martingal handelt. Außerdem ist dann auch  ${}_TM_i^c$  nach (3.52) für alle  $i \in \{1, \dots, \}$  ein  $({}_T\mathcal{F}_{\cdot,x}^c, P_{\theta}^{\tilde{\Pi}})$ -Martingal.

Insbesondere ist dann der Prozess  ${}_TA^c$  der  $({}_T\mathcal{F}_{\cdot,x}^c, P_{\theta}^{\tilde{\Pi}})$ -Kompensator von  ${}_TM^c$ , sodass sich

$$P_{\theta}^{\tilde{\Pi}}\{\Delta {}_TN^c(x) = 1 \mid {}_T\mathcal{F}_{\cdot,x-1}^c\} = E_{\theta}^{\tilde{\Pi}}\{\Delta {}_TN^c(x) \mid {}_T\mathcal{F}_{\cdot,x-1}^c\} = \Delta {}_TA^c(x, \theta) \quad (3.58)$$

ergibt.

Da gezeigt wurde, dass es sich bei  ${}_TV^c$  um ein  $({}_T\mathcal{F}_{\cdot,\chi}^c, P_{\theta}^{\tilde{\Pi}})$ -Martingal handelt, kann im nächsten Schritt der vorhersagbare Variationsprozess für  ${}_TV^c$  eingeführt und ermittelt werden. Dieser wird zur Überprüfung einiger Voraussetzungen des Martingalgrenzwertsatzes benötigt.

### 3.3.1. Der vorhersagbare Variationsprozess

In den nächsten Schritten soll mithilfe des Martingalgrenzwertsatzes von Rebolledo [1979] zunächst die asymptotische Normalverteilung von  ${}_TV^c$  und darauf aufbauend die asymptotische Normalverteilung von  $(\sqrt{n}(\hat{\theta}_n - \theta_0))_n$  für  $n \rightarrow \infty$  ermittelt werden. Zur Überprüfung der Grenzwertsatzvoraussetzungen muss zunächst das Konzept des vorhersagbaren Variationsprozesses eingeführt werden. Es gilt [siehe Fleming und Harrington, 1991, Korollar 1.4.2, beziehungsweise Kalbfleisch und Prentice, 2002, Kapitel 5.3.1]:

**Lemma 10.** Sei  ${}_TV^c$  ein Martingal bezüglich der Filtration  $\{{}_T\mathcal{F}_{\cdot,\chi}^c : \chi \in \mathbb{N}_0\}$ , wobei  $E_{\theta}^{\tilde{\Pi}}\{({}_TV^c(\chi, \theta))^2\} < \infty$  für alle  $\chi \in \{0, 1, 2, \dots\}$  und alle  $\theta \in \Theta$  gilt. Dann existiert ein eindeutiger, vorhersagbarer Prozess  $\langle {}_TV^c \rangle$ , der der vorhersagbare Variationsprozess von  ${}_TV^c$  genannt wird und folgende Eigenschaften erfüllt:

- (i) Es gilt  $\langle {}_TV^c \rangle(0, \theta) = 0$  und  $\langle {}_TV^c \rangle(\chi, \theta) < \infty$  für alle  $\chi \in \{1, 2, \dots\}$  und alle  $\theta \in \Theta$ .
- (ii) Der Prozess  $({}_TV^c)^2 - \langle {}_TV^c \rangle$  ist ein  ${}_T\mathcal{F}_{\cdot,\chi}^c$ -Martingal.

Der vorhersagbare Variationsprozess kann folgendermaßen berechnet werden:

$$\langle {}_TV^c \rangle(\chi, \theta) = \sum_{x=1}^{\chi} \text{Var}_{\theta}^{\tilde{\Pi}}\{\Delta {}_TV^c(x, \theta) \mid {}_T\mathcal{F}_{\cdot,x-1}^c\} \quad (3.59)$$

#### Nachweis der Existenz

Bevor  $\langle {}_TV^c \rangle(\chi, \theta)$  berechnet werden kann, wird vorher dessen Existenz nachgewiesen. Dazu wird  $E_{\theta}^{\tilde{\Pi}}\{({}_TV^c(\chi, \theta))^2\} < \infty$  für alle  $\chi \in \{1, 2, \dots\}$  und alle  $\theta \in \Theta$  überprüft.

Zur Vereinfachung der Notation werden zunächst die Summen der Zählprozesse, Unter-Risiko-Prozesse, Kompensatoren und Martingale über die Stichprobe folgendermaßen definiert:

$${}_T N^c(x) := \sum_{i=1}^n {}_T N_i^c(x) = \sum_{i=1}^n \mathbf{1}_{\{T_i+1 \leq X_i \leq \min(x, T_i+s)\}} \quad (3.60)$$

$${}_T Y^c(x) := \sum_{i=1}^n {}_T Y_i^c(x) = \sum_{i=1}^n \mathbf{1}_{\{T_i+1 \leq x+1 \leq \min(X_i, T_i+s)\}} \quad (3.61)$$

$${}_T A^c(x, \theta) := \theta \sum_{k=1}^x {}_T Y^c(k-1) \quad (3.62)$$

$${}_T M^c(x) := {}_T N^c(x) - {}_T A^c(x, \theta) \quad (3.63)$$

Dann gilt nach (3.51):

$$\begin{aligned} \mathbb{E}_{\tilde{\theta}}^{\tilde{\Pi}} \{({}_T V^c(\chi, \theta))^2\} &= \frac{1}{n\theta^2(1-\theta)^2} \mathbb{E}_{\tilde{\theta}}^{\tilde{\Pi}} \left\{ \left( \sum_{x=1}^{\chi} \Delta {}_T M^c(x) \right)^2 \right\} \\ &= \frac{1}{n\theta^2(1-\theta)^2} \left[ \sum_{x=1}^{\chi} \mathbb{E}_{\tilde{\theta}}^{\tilde{\Pi}} \left\{ \left( \Delta {}_T M^c(x) \right)^2 \right\} + 2 \sum_{\substack{x_1, x_2=1 \\ x_1 < x_2}}^{\chi} \mathbb{E}_{\tilde{\theta}}^{\tilde{\Pi}} \left\{ \left( \Delta {}_T M^c(x_1) \right) \left( \Delta {}_T M^c(x_2) \right) \right\} \right] \end{aligned}$$

Des Weiteren ergibt sich mithilfe der Eigenschaften bedingter Erwartungswerte (vgl. Satz 13, (iii)) für den hinteren Term

$$\begin{aligned} &\sum_{\substack{x_1, x_2=1 \\ x_1 < x_2}}^{\chi} \mathbb{E}_{\tilde{\theta}}^{\tilde{\Pi}} \left\{ \left( \Delta {}_T M^c(x_1) \right) \left( \Delta {}_T M^c(x_2) \right) \right\} \\ &= \sum_{\substack{x_1, x_2=1 \\ x_1 < x_2}}^{\chi} \mathbb{E}_{\tilde{\theta}}^{\tilde{\Pi}} \left\{ \mathbb{E}_{\tilde{\theta}}^{\tilde{\Pi}} \left\{ \left( \Delta {}_T M^c(x_1) \right) \left( \Delta {}_T M^c(x_2) \right) \middle| {}_T \mathcal{F}_{\cdot, x_1}^c \right\} \right\} \\ &= \sum_{\substack{x_1, x_2=1 \\ x_1 < x_2}}^{\chi} \mathbb{E}_{\tilde{\theta}}^{\tilde{\Pi}} \left\{ \left( \Delta {}_T M^c(x_1) \right) \mathbb{E}_{\tilde{\theta}}^{\tilde{\Pi}} \left\{ \left( \Delta {}_T M^c(x_2) \right) \middle| {}_T \mathcal{F}_{\cdot, x_1}^c \right\} \right\} \\ &= \sum_{\substack{x_1, x_2=1 \\ x_1 < x_2}}^{\chi} \mathbb{E}_{\tilde{\theta}}^{\tilde{\Pi}} \left\{ \left( \Delta {}_T M^c(x_1) \right) \cdot 0 \right\} = 0, \end{aligned}$$

sodass

$$\mathbb{E}_{\tilde{\theta}}^{\tilde{\Pi}} \{({}_T V^c(\chi, \theta))^2\} = \frac{1}{n\theta^2(1-\theta)^2} \left[ \sum_{x=1}^{\chi} \mathbb{E}_{\tilde{\theta}}^{\tilde{\Pi}} \left\{ \mathbb{E}_{\tilde{\theta}}^{\tilde{\Pi}} \left\{ \left( \Delta {}_T M^c(x) \right)^2 \middle| {}_T \mathcal{F}_{\cdot, x-1}^c \right\} \right\} \right] \quad (3.64)$$

Im nächsten Schritt wird der innere, bedingte Erwartungswert von (3.64) berechnet. Es gilt:

$$\begin{aligned} \mathbb{E}_{\theta}^{\tilde{\mathbb{T}}}\left\{\left(\Delta_{TM^c}(x)\right)^2\middle|T\mathcal{F}_{\cdot,x-1}^c\right\} &= \mathbb{E}_{\theta}^{\tilde{\mathbb{T}}}\left\{\left(\Delta_{TN^c}(x)\right)^2\middle|T\mathcal{F}_{\cdot,x-1}^c\right\} + \mathbb{E}_{\theta}^{\tilde{\mathbb{T}}}\left\{\left(\Delta_{TA^c}(x,\theta)\right)^2\middle|T\mathcal{F}_{\cdot,x-1}^c\right\} \\ &\quad - 2\mathbb{E}_{\theta}^{\tilde{\mathbb{T}}}\left\{\Delta_{TN^c}(x)\Delta_{TA^c}(x,\theta)\middle|T\mathcal{F}_{\cdot,x-1}^c\right\} \end{aligned} \quad (3.65)$$

Aufgrund der  $T\mathcal{F}_{x-1}^c$ -Messbarkeit von  $\Delta_{TA^c}(x,\theta)$  folgt weiter:

$$\begin{aligned} &\mathbb{E}_{\theta}^{\tilde{\mathbb{T}}}\left\{\left(\Delta_{TM^c}(x)\right)^2\middle|T\mathcal{F}_{\cdot,x-1}^c\right\} \\ &= \mathbb{E}_{\theta}^{\tilde{\mathbb{T}}}\left\{\left(\Delta_{TN^c}(x)\right)^2\middle|T\mathcal{F}_{\cdot,x-1}^c\right\} - 2\Delta_{TA^c}(x,\theta)\mathbb{E}_{\theta}^{\tilde{\mathbb{T}}}\left\{\Delta_{TN^c}(x)\middle|T\mathcal{F}_{\cdot,x-1}^c\right\} + \left(\Delta_{TA^c}(x,\theta)\right)^2 \end{aligned} \quad (3.66)$$

Im Folgenden werden die einzelnen Terme von (3.66) betrachtet. Zunächst gilt

$$\left(\Delta_{TA^c}(x,\theta)\right)^2 = \theta^2\left(TY^c(x-1)\right)^2 \quad (3.67)$$

sowie

$$\begin{aligned} 2\Delta_{TA^c}(x,\theta)\mathbb{E}_{\theta}^{\tilde{\mathbb{T}}}\left\{\Delta_{TN^c}(x)\middle|T\mathcal{F}_{\cdot,x-1}^c\right\} &= 2\theta TY^c(x-1)\mathbb{E}_{\theta}^{\tilde{\mathbb{T}}}\left\{\Delta_{TN^c}(x)\middle|T\mathcal{F}_{\cdot,x-1}^c\right\} \\ &= 2\theta^2\left(TY^c(x-1)\right)^2, \end{aligned} \quad (3.68)$$

wobei der letzte Umformungsschritt intuitiv auf der Überlegung basiert, dass zum Zeitpunkt  $x$  insgesamt  $TY^c(x-1)$  Individuen mit Wahrscheinlichkeit  $\theta$  ausfallen können und somit  $\mathbb{E}_{\theta}^{\tilde{\mathbb{T}}}\left\{\Delta_{TN^c}(x)\middle|T\mathcal{F}_{\cdot,x-1}^c\right\} = \theta TY^c(x-1)$  folgt. Mathematisch ergibt sich das Resultat aus der Tatsache, dass es sich bei dem Prozess  $TA^c$  um den  $(P_{\theta}^{\tilde{\mathbb{T}}}, T\mathcal{F}_{\cdot,x})$ -Kompensator von  $TN^c$  handelt.

Abschließend muss der Term  $\mathbb{E}_{\theta}^{\tilde{\mathbb{T}}}\left\{\left(\Delta_{TN^c}(x)\right)^2\middle|T\mathcal{F}_{\cdot,x-1}^c\right\}$  berechnet werden. Es gilt:

$$\begin{aligned} \mathbb{E}_{\theta}^{\tilde{\mathbb{T}}}\left\{\left(\Delta_{TN^c}(x)\right)^2\middle|T\mathcal{F}_{\cdot,x-1}^c\right\} &= \mathbb{E}_{\theta}^{\tilde{\mathbb{T}}}\left\{\left(\sum_{i=1}^n \Delta_{TN_i^c}(x)\right)^2\middle|T\mathcal{F}_{\cdot,x-1}^c\right\} \\ &= \mathbb{E}_{\theta}^{\tilde{\mathbb{T}}}\left\{\sum_{i=1}^n \left(\Delta_{TN_i^c}(x)\right)^2\middle|T\mathcal{F}_{\cdot,x-1}^c\right\} + 2\mathbb{E}_{\theta}^{\tilde{\mathbb{T}}}\left\{\sum_{\substack{i_1, i_2 \in \{1, \dots, n\} \\ i_1 < i_2}} \left(\Delta_{TN_{i_1}^c}(x)\Delta_{TN_{i_2}^c}(x)\right)\middle|T\mathcal{F}_{\cdot,x-1}^c\right\} \end{aligned}$$

Da  $\Delta_{TN_i^c}(x) \in \{0, 1\}$  gilt und die Ausfälle unabhängig voneinander erfolgen (vgl. Annahme 1, (A4)), ergibt sich

$$\begin{aligned} \mathbb{E}_{\theta}^{\tilde{\mathbb{T}}}\left\{\left(\Delta_{TN^c}(x)\right)^2\middle|T\mathcal{F}_{\cdot,x-1}^c\right\} &= \mathbb{E}_{\theta}^{\tilde{\mathbb{T}}}\left\{\sum_{i=1}^n \Delta_{TN_i^c}(x)\middle|T\mathcal{F}_{\cdot,x-1}^c\right\} \\ &\quad + 2 \sum_{\substack{i_1, i_2 \in \{1, \dots, n\} \\ i_1 < i_2}} \mathbb{E}_{\theta}^{\tilde{\mathbb{T}}}\left\{\Delta_{TN_{i_1}^c}(x)\middle|T\mathcal{F}_{\cdot,x-1}^c\right\} \mathbb{E}_{\theta}^{\tilde{\mathbb{T}}}\left\{\Delta_{TN_{i_2}^c}(x)\middle|T\mathcal{F}_{\cdot,x-1}^c\right\} \end{aligned}$$

und somit:

$$\begin{aligned} \mathbb{E}_{\theta}^{\tilde{\Pi}}\left\{(\Delta_{TN^c}(x))^2 \middle| \mathcal{F}_{:,x-1}^c\right\} &= \mathbb{E}_{\theta}^{\tilde{\Pi}}\left\{\Delta_{TN^c}(x) \middle| \mathcal{F}_{:,x-1}^c\right\} \\ &\quad + 2\theta^2 \sum_{\substack{i_1, i_2 \in \{1, \dots, n\} \\ i_1 < i_2}} {}_TY_{i_1}^c(x-1) {}_TY_{i_2}^c(x-1) \end{aligned}$$

Die hintere Summe entspricht dabei der Anzahl aller Kombinationen verschiedener Individuen, deren potentieller Ausfall in  $x$  beobachtbar ist. Zum Zeitpunkt  $x-1$  gibt es insgesamt  ${}_TY^c(x-1)$  solcher Individuen. Somit folgt:

$$\begin{aligned} \mathbb{E}_{\theta}^{\tilde{\Pi}}\left\{(\Delta_{TN^c}(x))^2 \middle| \mathcal{F}_{:,x-1}^c\right\} &= \mathbb{E}_{\theta}^{\tilde{\Pi}}\left\{\Delta_{TN^c}(x) \middle| \mathcal{F}_{:,x-1}^c\right\} + 2\theta^2 \binom{{}_TY^c(x-1)}{2} \\ &= \theta {}_TY^c(x-1) + 2\theta^2 \frac{{}_TY^c(x-1)!}{2!({}_TY^c(x-1) - 2)!} \\ &= \theta {}_TY^c(x-1) + \theta^2 {}_TY^c(x-1)({}_TY^c(x-1) - 1) \\ &= \theta {}_TY^c(x-1)(1 + \theta {}_TY^c(x-1) - \theta) \end{aligned} \quad (3.69)$$

Als Nächstes folgt durch Einsetzen von (3.67), (3.68) und (3.69) in (3.65):

$$\begin{aligned} \mathbb{E}_{\theta}^{\tilde{\Pi}}\left\{(\Delta_{TM^c}(x))^2 \middle| \mathcal{F}_{:,x-1}^c\right\} &= \theta {}_TY^c(x-1)(1 + \theta {}_TY^c(x-1) - \theta) \\ &\quad - 2\theta^2 ({}_TY^c(x-1))^2 + \theta^2 ({}_TY^c(x-1))^2 \\ &= \theta(1 - \theta) {}_TY^c(x-1) \end{aligned} \quad (3.70)$$

und somit:

$$\begin{aligned} \mathbb{E}_{\theta}^{\tilde{\Pi}}\left\{({}_TV^c(\chi, \theta))^2\right\} &= \frac{1}{n\theta^2(1 - \theta_0)^2} \left[ \sum_{x=1}^{\chi} \mathbb{E}_{\theta}^{\tilde{\Pi}}\left\{\mathbb{E}_{\theta}^{\tilde{\Pi}}\left\{(\Delta_{TM^c}(x))^2 \middle| \mathcal{F}_{:,x-1}^c\right\}\right\} \right] \\ &= \frac{1}{n\theta(1 - \theta)} \sum_{x=1}^{\chi} \mathbb{E}_{\theta}^{\tilde{\Pi}}\left\{{}_TY^c(x-1)\right\} = \frac{1}{\theta(1 - \theta)} \sum_{x=1}^{\chi} \mathbb{E}_{\theta}^{\tilde{\Pi}}\left\{{}_TY_i^c(x-1)\right\} \\ &< \frac{\min(T + s, \chi)}{\theta(1 - \theta)} \leq \frac{\min(G - 1 + s, \chi)}{\theta(1 - \theta)} < \infty \end{aligned}$$

Damit ist  $\mathbb{E}_{\theta}^{\tilde{\Pi}}\left\{({}_TV^c(\chi, \theta))^2\right\} < \infty$  für alle  $\chi \in \{1, 2, \dots\}$  und somit die Existenz eines vorhersagbaren Variationsprozesses  $\langle {}_TV^c \rangle$  nachgewiesen.

### Berechnung

Im Folgenden wird  $\langle {}_TV^c \rangle(\chi, \theta)$  berechnet. Zunächst gilt

$$\langle {}_TV^c \rangle(\chi, \theta) = \sum_{x=1}^{\chi} \text{Var}_{\theta}^{\tilde{\Pi}}\left\{\Delta\left(\sum_{i=1}^n {}_TV_i^c(x, \theta)\right) \middle| \mathcal{F}_{:,x-1}^c\right\}$$

und somit:

$$\begin{aligned} \langle {}_T V^c \rangle(\chi, \theta) &= \sum_{x=1}^{\chi} \left[ \sum_{i=1}^n \text{Var}_{\tilde{\theta}} \left\{ \Delta {}_T V_i^c(x, \theta) \middle| {}_T \mathcal{F}_{\cdot, x-1}^c \right\} \right. \\ &\quad \left. + 2 \sum_{\substack{i_1, i_2=1 \\ i_1 < i_2}}^n \text{Cov}_{\tilde{\theta}} \left\{ \Delta {}_T V_{i_1}^c(x, \theta), \Delta {}_T V_{i_2}^c(x, \theta) \middle| {}_T \mathcal{F}_{\cdot, x-1}^c \right\} \right] \end{aligned} \quad (3.71)$$

Für den ersten Term ergibt sich unter der Verwendung von (3.52):

$$\begin{aligned} &\sum_{i=1}^n \text{Var}_{\tilde{\theta}} \left\{ \Delta {}_T V_i^c(x, \theta) \middle| {}_T \mathcal{F}_{\cdot, x-1}^c \right\} \\ &= \frac{1}{n\theta^2(1-\theta)^2} \sum_{i=1}^n \text{Var}_{\tilde{\theta}} \left\{ \Delta {}_T M_i^c(x) \middle| {}_T \mathcal{F}_{\cdot, x-1}^c \right\} \\ &= \frac{1}{n\theta^2(1-\theta)^2} \sum_{i=1}^n \text{E}_{\tilde{\theta}} \left\{ \left( \Delta {}_T M_i^c(x) \right)^2 \middle| {}_T \mathcal{F}_{\cdot, x-1}^c \right\} \\ &= \frac{1}{n\theta^2(1-\theta)^2} \sum_{i=1}^n \left[ \text{E}_{\tilde{\theta}} \left\{ \Delta {}_T N_i^c(x) \middle| {}_T \mathcal{F}_{\cdot, x-1}^c \right\} \right. \\ &\quad \left. - 2\theta {}_T Y_i^c(x-1) \text{E}_{\tilde{\theta}} \left\{ \Delta {}_T N_i^c(x) \middle| {}_T \mathcal{F}_{\cdot, x-1}^c \right\} + \theta^2 \left( {}_T Y_i^c(x-1) \right)^2 \right] \end{aligned}$$

Unter der Verwendung von  $\text{E}_{\tilde{\theta}} \left\{ \Delta {}_T N_i^c(x) \middle| {}_T \mathcal{F}_{\cdot, x-1}^c \right\} = \theta {}_T Y_i^c(x-1)$  und  ${}_T Y_i^c(x-1) \in \{0, 1\}$  folgt weiter:

$$\begin{aligned} &\sum_{i=1}^n \text{Var}_{\tilde{\theta}} \left\{ \Delta {}_T V_i^c(x, \theta) \middle| {}_T \mathcal{F}_{\cdot, x-1}^c \right\} \\ &= \frac{\sum_{i=1}^n \left[ \theta {}_T Y_i^c(x-1) - 2\theta^2 {}_T Y_i^c(x-1) + \theta^2 {}_T Y_i^c(x-1) \right]}{n\theta^2(1-\theta)^2} \\ &= \frac{1}{n\theta(1-\theta)} \sum_{i=1}^n {}_T Y_i^c(x-1) \end{aligned}$$

Für den hinteren Term von (3.71) folgt zunächst:

$$\begin{aligned} &2 \sum_{\substack{i_1, i_2=1 \\ i_1 < i_2}}^n \text{Cov}_{\tilde{\theta}} \left\{ \Delta {}_T V_{i_1}^c(x, \theta), \Delta {}_T V_{i_2}^c(x, \theta) \middle| {}_T \mathcal{F}_{\cdot, x-1}^c \right\} \\ &= \frac{2}{n\theta^2(1-\theta)^2} \sum_{\substack{i_1, i_2=1 \\ i_1 < i_2}}^n \text{Cov}_{\tilde{\theta}} \left\{ \Delta {}_T M_{i_1}^c(x), \Delta {}_T M_{i_2}^c(x) \middle| {}_T \mathcal{F}_{\cdot, x-1}^c \right\} \end{aligned}$$

Weiter gilt wegen  $E_{\tilde{\theta}}^{\mathbb{T}}\{\Delta_{TM_i^c}(x) | T\mathcal{F}_{\cdot, x-1}^c\} = 0$ :

$$\begin{aligned}
 & 2 \sum_{\substack{i_1, i_2=1 \\ i_1 < i_2}}^n \text{Cov}_{\tilde{\theta}}^{\mathbb{T}}\left\{\Delta_{TV_{i_1}^c}(x, \theta), \Delta_{TV_{i_2}^c}(x, \theta) \mid T\mathcal{F}_{\cdot, x-1}^c\right\} \\
 &= \frac{2}{n\theta^2(1-\theta)^2} \sum_{\substack{i_1, i_2=1 \\ i_1 < i_2}}^n E_{\tilde{\theta}}^{\mathbb{T}}\left\{\Delta_{TM_{i_1}^c}(x)\Delta_{TM_{i_2}^c}(x) \mid T\mathcal{F}_{\cdot, x-1}^c\right\} \\
 &= \frac{2}{n\theta^2(1-\theta)^2} \\
 & \quad \sum_{\substack{i_1, i_2=1 \\ i_1 < i_2}}^n E_{\tilde{\theta}}^{\mathbb{T}}\left\{(\Delta_{TN_{i_1}^c}(x) - \theta_{TY_{i_1}^c}(x-1))(\Delta_{TN_{i_2}^c}(x) - \theta_{TY_{i_2}^c}(x-1)) \mid T\mathcal{F}_{\cdot, x-1}^c\right\} \quad (3.72)
 \end{aligned}$$

**Bemerkung 15.** Im Rahmen von zeitstetigen Ausfallzeitanalysen wird häufig [siehe beispielsweise Borgan, 1984, Fleming und Harrington, 1991, Andersen et al., 1993] die Annahme getroffen, dass es sich bei dem Prozess  $TN^c$  um einen multivariaten Zählprozess handelt. Das bedeutet, dass zum einen jede Komponente  $TN_i^c$  einen Zählprozess darstellt und zum anderen niemals zwei Komponenten gleichzeitig ausfallen können, sodass durch  $\Delta_{TN_{i_1}^c}(x) = 1$  direkt  $\Delta_{TN_{i_2}^c}(x) = 0$  impliziert wird. In diesem Fall ließe sich nämlich ohne Beschränkung der Allgemeinheit in (3.72)  $\Delta_{TN_{i_2}^c}(x) = 0$  setzen, sodass wegen

$$E_{\tilde{\theta}}^{\mathbb{T}}\{(\Delta_{TN_{i_1}^c}(x) | T\mathcal{F}_{\cdot, x-1}^c) = \theta_{TY_{i_1}^c}(x-1)$$

sowie der  $T\mathcal{F}_{\cdot, x-1}^c$ -Messbarkeit von  $TY_{i_2}^c(x-1)$  direkt

$$2 \sum_{\substack{i_1, i_2=1 \\ i_1 < i_2}}^n \text{Cov}_{\tilde{\theta}}^{\mathbb{T}}\left\{\Delta_{TV_{i_1}^c}(x), \Delta_{TV_{i_2}^c}(x) \mid T\mathcal{F}_{\cdot, x-1}^c\right\} = 0$$

folgen würde.

Für das in dieser Arbeit verwendete zeitdiskrete Modell kann die Annahme eines multivariaten Zählprozesses jedoch nicht getroffen werden. Stattdessen wird ersatzweise die Unabhängigkeit der Individuumsausfälle gefordert, sodass ebenfalls

$$\begin{aligned}
 & 2 \sum_{\substack{i_1, i_2=1 \\ i_1 < i_2}}^n \text{Cov}_{\tilde{\theta}}^{\mathbb{T}}\left\{\Delta_{TV_{i_1}^c}(x), \Delta_{TV_{i_2}^c}(x) \mid T\mathcal{F}_{\cdot, x-1}^c\right\} \\
 &= \frac{2}{n\theta_0^2(1-\theta_0)^2} \sum_{\substack{i_1, i_2=1 \\ i_1 < i_2}}^n E_{\tilde{\theta}}^{\mathbb{T}}\left\{\Delta_{TM_{i_1}^c}(x) \mid T\mathcal{F}_{x-1}^c\right\} E_{\tilde{\theta}}^{\mathbb{T}}\left\{\Delta_{TM_{i_2}^c}(x) \mid T\mathcal{F}_{\cdot, x-1}^c\right\} = 0
 \end{aligned}$$

resultiert

Zusammenfassend lässt sich der vorhersagbare Variationsprozess von  ${}_T V^c$  darstellen als:

$$\langle {}_T V^c \rangle(\chi, \theta) = \frac{1}{n\theta(1-\theta)} \sum_{i=1}^n \sum_{x=1}^{\chi} {}_T Y_i^c(x-1)$$

Somit handelt es sich bei  ${}_T V^c$  um ein  $({}_T \mathcal{F}_{\cdot, \chi}^c, P_{\theta}^{\mathbb{T}})$ -Martingal mit dem zugehörigen vorhersagbaren Variationsprozess  $\langle {}_T V^c \rangle$ . Darauf aufbauend kann anschließend der Martingalgrenzwertsatz nach dem Vorbild von Rebolledo [1979] eingeführt und auf  ${}_T V^c$  angewendet werden.

### 3.3.2. Konsistenz und Asymptotische Normalität

Nachdem die Existenz eines vorhersagbaren Variationsprozesses von  ${}_T V^c$  nachgewiesen und dieser berechnet wurde, kann darauf aufbauend im Folgenden die Konsistenz und asymptotische Normalität von  ${}_T V^c$  gezeigt und aus den Ergebnissen die asymptotische Normalverteilung von  $(\sqrt{n}(\hat{\theta}_n - \theta_0))_n$  für  $n \rightarrow \infty$  abgeleitet werden. Zu diesem Zweck wird ein Martingalgrenzwertsatz von Rebolledo [1979] verwendet [siehe Kalbfleisch und Prentice, 2002, Theorem 5.1]:

**Satz 9** (Martingalgrenzwertsatz). *Sei  $\tilde{x}$  ein beliebiger, fester Zeitpunkt mit  $\tilde{x} \in \{1, \dots, \chi\}$ . Dann wird durch die Bedingungen*

$$(i) \quad \langle {}_T V^c \rangle(\tilde{x}, \theta) \xrightarrow{P} \sigma_V^2(\tilde{x}, \theta) \text{ für } n \rightarrow \infty$$

$$(ii) \quad \langle {}_T V_{\cdot, \epsilon}^c \rangle(\tilde{x}, \theta) \xrightarrow{P} 0 \text{ für } n \rightarrow \infty$$

*impliziert, dass  ${}_T V^c(\tilde{x}, \theta)$  asymptotisch normalverteilt mit Erwartungswert 0 und Varianz  $\sigma_V^2(\tilde{x}, \theta)$  ist. Der Prozess  $\langle {}_T V_{\cdot, \epsilon}^c \rangle$  ist dabei definiert durch:*

$$\langle {}_T V_{\cdot, \epsilon}^c \rangle(\tilde{x}, \theta) := \frac{1}{n\theta(1-\theta)} \sum_{i=1}^n \sum_{x=1}^{\tilde{x}} \mathbf{1}_{\{ |(\sqrt{n}\theta(1-\theta))^{-1}| > \epsilon \}} {}_T Y_i^c(x-1)$$

Zunächst werden die beiden Eigenschaften von Satz 9 überprüft. Da die Unter-Risiko-Prozesse  ${}_T Y_i^c$  für jeden Zeitpunkt  $x-1$  mit  $x \in \mathbb{N}$  unabhängig und identisch verteilt sind und sich somit das schwache Gesetz der großen Zahlen (vgl. Satz 15) anwenden lässt, konvergiert das arithmetische Mittel der Unter-Risiko-Prozesse gegen den Erwartungswert von  ${}_T Y_i^c$  bezüglich des unbedingten Wahrscheinlichkeitsmaßes  $P_{\theta}$  (vgl. (2.4)). Es folgt:

$$\langle {}_T V^c \rangle(\tilde{x}, \theta) = \frac{1}{\theta(1-\theta)} \sum_{x=1}^{\tilde{x}} \frac{1}{n} \sum_{i=1}^n {}_T Y_i^c(x-1) \xrightarrow{P} \frac{1}{\theta(1-\theta)} \sum_{x=1}^{\tilde{x}} E_{\theta} \{ {}_T Y_i^c(x-1) \} := \sigma_V^2(\tilde{x}, \theta)$$

Zur Überprüfung der zweiten Bedingung sei angemerkt, dass

$$\frac{1}{\sqrt{n}\theta(1-\theta)} \xrightarrow{n \rightarrow \infty} 0$$

gilt und somit direkt

$$\begin{aligned} \langle {}_T V_{\cdot, \epsilon}^c \rangle(\tilde{x}, \theta) &= \frac{1}{\theta(1-\theta)} \sum_{x=1}^{\tilde{x}} \mathbb{1}_{\{ |(\sqrt{n}\theta(1-\theta))^{-1}| > \epsilon \}} \frac{1}{n} \sum_{i=1}^n {}_T Y_i^c(x-1) \\ &\xrightarrow{p} \frac{1}{\theta(1-\theta)} \sum_{x=1}^{\tilde{x}} 0 \cdot \mathbb{E}_\theta \{ {}_T Y_i^c(x-1) \} = 0 \end{aligned}$$

folgt. Der Prozess  ${}_T V^c$  ist somit asymptotisch normalverteilt mit Erwartungswert 0 und Varianz  $\sigma_V^2(\tilde{x}, \theta)$ .

Im letzten Schritt kann nun die asymptotische Normalverteilung von  $(\sqrt{n}(\hat{\theta}_n - \theta_0))_n$  berechnet werden. Für eine standardnormalverteilte Zufallsvariable  $Z \sim N(0, 1)$  gilt:

$$\begin{aligned} {}_T V^c(\chi, \theta_0) &\rightsquigarrow \sigma_V(\chi, \theta_0) \cdot Z \\ &\Leftrightarrow \frac{1}{\sqrt{n}} \sum_{x=1}^{\chi} \frac{1}{\theta_0(1-\theta_0)} \Delta {}_T M^c(x) \rightsquigarrow \sqrt{\frac{1}{\theta_0(1-\theta_0)} \sum_{x=1}^{\chi} \mathbb{E}_{\theta_0} \{ {}_T Y_i^c(x-1) \}} \cdot Z \\ &\Leftrightarrow \frac{1}{\sqrt{n}} \sum_{x=1}^{\chi} (\Delta {}_T N^c(x) - \theta_0 {}_T Y^c(x-1)) \rightsquigarrow \sqrt{\theta_0(1-\theta_0) \sum_{x=1}^{\chi} \mathbb{E}_{\theta_0} \{ {}_T Y_i^c(x-1) \}} \cdot Z \end{aligned}$$

Die durch das Symbol ' $\rightsquigarrow$ ' gekennzeichnete Konvergenz in Verteilung wird dabei im Appendix, Definition 17, beschrieben.

Da  $\mathbb{E}_{\theta_0} \{ {}_T Y_i^c(x-1) \}$  deterministisch ist, folgt mit dem Satz von Slutsky (vgl. Appendix, Satz 16, (iii)):

$$\frac{1}{\sqrt{n}} \frac{\sum_{x=1}^{\chi} (\Delta {}_T N^c(x) - \theta_0 {}_T Y^c(x-1))}{\frac{1}{n} \sum_{x=1}^{\chi} {}_T Y^c(x-1)} \rightsquigarrow \frac{\sqrt{\theta_0(1-\theta_0) \sum_{x=1}^{\chi} \mathbb{E}_{\theta_0} \{ {}_T Y_i^c(x-1) \}}}{\sum_{x=1}^{\chi} \mathbb{E}_{\theta_0} \{ {}_T Y_i^c(x-1) \}} \cdot Z$$

Dabei ist die linke Seite wegen der Formel für den Punktschätzer (3.9) sowie den Darstellungen (3.60) und (3.61) mit  $\chi \geq s + G - 1$  gleich  $\sqrt{n}(\hat{\theta}_n - \theta_0)$  und es gilt:

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \rightsquigarrow \frac{\sqrt{\theta_0(1-\theta_0)}}{\sqrt{\sum_{x=1}^{s+G-1} \mathbb{E}_{\theta_0} \{ {}_T Y_i^c(x-1) \}}} \cdot Z = \frac{1}{\sigma_V(s+G-1, \theta_0)} \cdot Z \quad (3.73)$$

Somit ist die Sequenz  $\sqrt{n}(\hat{\theta}_n - \theta_0)$  asymptotisch normalverteilt mit Erwartungswert 0 und Standardabweichung  $(\sigma_V(s+G-1, \theta_0))^{-1}$ .

Abschließend ergibt sich der quadrierte Standardfehler mithilfe von (3.13) sowie (3.73) als Schätzer von:

$$\frac{\text{Var}_{\theta_0} \{ \sqrt{n} \hat{\theta}_n \}}{n} = \frac{1}{n \sigma_V(s+G-1, \theta_0)^2} = \frac{1}{n} \frac{\theta_0(1-\theta_0)}{\sum_{x=1}^{s+G-1} \mathbb{E}_{\theta_0} \{ {}_T Y_i^c(x-1) \}} \quad (3.74)$$

Zur Standardfehlerberechnung wird zunächst in (3.74) der unbekannte, wahre Parameterwert  $\theta_0$  durch den zugehörigen Punktschätzer  $\hat{\theta}_n$  (vgl. 3.9) ersetzt. Des Weiteren werden, um keine Verteilungsannahme für  $T_i$  mit  $i \in \{1, \dots, n\}$  treffen zu müssen, die

Erwartungswerte  $E_{\theta_0}\{TY_i^c(x-1)\}$  für  $x \in \{1, \dots, s+G-1\}$  mithilfe der entsprechenden arithmetischen Mittel geschätzt. Es ergibt sich:

$$SE_3^2 = \frac{1}{n} \frac{\hat{\theta}_n(1 - \hat{\theta}_n)}{\sum_{x=1}^{s+G-1} \frac{1}{n} \sum_{i=1}^n TY_i^c(x-1)} = \frac{\hat{\theta}_n(1 - \hat{\theta}_n)}{\sum_{x=1}^{s+G-1} \sum_{i=1}^n TY_i^c(x-1)}$$

Es sei angemerkt, dass sich  $E_{\theta_0}\{TY_i^c(x-1)\}$  auch explizit darstellen lässt, sofern eine Verteilungsannahme für  $T_i$  mit  $i \in \{1, \dots, n\}$  getroffen wurde. Eine beispielhafte Berechnung des Standardfehlers unter der Annahme, dass alle  $T_i$  unabhängig und identisch diskret gleichverteilt auf dem Träger  $\{0, \dots, G-1\}$  sind, lässt sich im Appendix, Kapitel A.2, finden.

Abschließend lässt sich die Formel für den Standardfehler  $SE_3^2$  mithilfe von Bemerkung 13 beziehungsweise Formel (3.12) sowie (3.11) folgendermaßen umschreiben:

$$\begin{aligned} SE_3^2 &= \frac{1}{d^{obs} + sm_{cens}} \frac{m_{uncens}}{d^{obs} + sm_{cens}} \left( 1 - \frac{m_{uncens}}{d^{obs} + sm_{cens}} \right) \\ &= \frac{m_{uncens}(d^{obs} + sm_{cens}) - m_{uncens}^2}{(d^{obs} + sm_{cens})^3} \end{aligned} \quad (3.75)$$

Nachdem nun ein Punktschätzer für die Ausfallwahrscheinlichkeit  $\hat{\theta}_n$  sowie Formeln für zugehörige Standardfehler hergeleitet wurden, können mithilfe dieser Resultate in Kapitel 5 konkrete Daten analysiert und insbesondere die verschiedenen Standardfehler miteinander verglichen werden. Vorher erfolgt jedoch ein Exkurs über eine alternative Methode zur Herleitung eines Punktschätzers bei einem doppelt trunkierten Beobachtungsdesign.

## 4. Schätzung bei doppelt trunkiertem Design

Im Folgenden wird ein alternativer Ansatz zur Erstellung eines Punktschätzers von  $\theta_0$  bei einem beidseitig trunkierten Beobachtungsdesign hergeleitet. Die vorgestellte Herleitung kann dabei als diskrete Version von Weißbach und Wied [2022, Kapitel 2.1 und Kapitel 3] interpretiert werden. Diese Methode liefert zwar für simulierte Daten (siehe Kapitel 6) die plausible Ergebnisse, jedoch konnten im Rahmen der Anwendung auf das AFiD-Panel (siehe Kapitel 5) keine glaubhaften Ergebnisse erzielt werden. Folglich wird dieser Ansatz nach der Berechnung des Punktschätzers nicht weiter verfolgt.

Analog zu vorherigen Kapiteln sei eine Stichprobe, bestehend aus  $n$  Individuen, gegeben. Für diese werden, definiert auf einem gemeinsamen Wahrscheinlichkeitsraum  $(\Omega, \mathcal{F}, P_{\theta_0})$  (vgl. (2.1) und folgende Seiten), jeweils sowohl die Lebensdauer  $X_i$ ,  $i = 1, \dots, n$  (vgl. (2.2)) als auch das Alter ein Jahr vor dem Studieneintritt  $T_i$ ,  $i = 1, \dots, n$  (vgl. (2.3)) mit  $T_i \in \{0, \dots, G-1\}$  in Jahren gemessen. Entstehung und Ausfall werden dabei für jedes Individuum als unabhängig vorausgesetzt. Es wird angenommen, dass alle  $X_i$  unabhängig und identisch geometrisch verteilt (vgl. Definition 1) mit dem wahren Parameterwert  $\theta_0 \in \Theta = [\xi, 1-\xi]$  mit  $\xi \in (0, 1/2)$  sind. Dahingegen sind alle  $T_i$  unabhängig diskret gleichverteilt (vgl. Definition 7) auf dem Träger  $\{0, \dots, G-1\}$ .

Der Raum der möglichen Ergebnisse von  $(X_i, T_i)$  kann entsprechend als  $\mathbb{S} = \mathbb{N} \times \{0, \dots, G-1\}$  definiert werden. Des Weiteren sei  $\mathcal{U}$  das von  $\mathbb{S}$  erzeugte  $\sigma$ -Feld.

**Annahmen 2.** *Zusammenfassend sollen für das latente Modell, in dem für jedes Individuum sowohl die Lebensdauer als auch das Alter ein Jahr vor dem Studieneintritt beobachtbar sind, folgende Annahmen gelten:*

- (B1) *Für alle  $i \in \{1, \dots, n\}$  sind die bivariaten Zufallsvariablen  $(X_i, T_i)'$  auf dem gemeinsamen Wahrscheinlichkeitsraum  $(\Omega, \mathcal{F}, P_{\theta_0})$ , mit  $\Omega$  aus (2.1),  $\mathcal{F} = \mathcal{P}\{\Omega\}$  und  $P_{\theta_0}$  aus (2.4), definiert.*
- (B2) *Die marginale Wahrscheinlichkeitsfunktion von  $X_i$  entspricht der Wahrscheinlichkeitsfunktion der geometrischen Verteilung (vgl. (2.5)). Die marginale Wahrscheinlichkeitsfunktion von  $T_i$  entspricht der Wahrscheinlichkeitsfunktion der diskreten Gleichverteilung mit dem Träger  $\{0, \dots, G-1\}$  (vgl. (3.14)). Außerdem ist  $X_i$  unabhängig von  $T_i$ .*
- (B3) *Für den wahren Parameterwert  $\theta_0$  gilt  $\theta_0 \in \Theta := [\xi, 1-\xi]$  für ein kleines  $\xi \in (0, \frac{1}{2})$ .*
- (B4)  *$(X_i, T_i)'$ ,  $i \in \{1, \dots, n\}$ ,  $n \in \mathbb{N}$  beschreibt eine einfache Zufallsstichprobe. Somit sind  $(X_i, T_i)'$  unabhängige und identisch verteilte Zufallsvariablen, die der Vertei-*

lungsannahme (B2) mit dem wahren Parameterwert  $\theta_0$  aus Annahme (B3) genügen.

## 4.1. Das beobachtbare Modell

Für das beobachtbare Modell wird zusätzlich zu den Voraussetzungen (B1)-(B4) die Annahme getroffen, dass ein Individuum genau dann beobachtet werden kann, wenn sein Ausfallzeitpunkt innerhalb des Beobachtungszeitraums liegt. Sei dafür  $s \in \mathbb{N}$  die Studiendauer. Fällt das  $i$ -te Individuum vor dem Studienbeginn oder nach dem Ende der Studie aus, gilt also  $X_i \leq T_i$  oder  $X_i \geq T_i + s + 1$ , so werden keine Informationen beobachtet und nicht einmal die Existenz dieses Individuums zur Kenntnis genommen. Somit liegt ein doppelt trunkiertes Beobachtungsdesign vor. Dies führt zu folgender Annahme:

(B5) Für eine bekannte, ganzzahlige Konstante  $s \in \mathbb{N}$  wird der Spaltenvektor  $(X_i, T_i)'$  genau dann beobachtet, wenn:

$$(X_i, T_i)' \in \mathbb{D} := \left\{ (x, t)' \mid 1 \leq t + 1 \leq x \leq t + s, x \in \mathbb{N}, t \in \{0, \dots, G - 1\} \right\}$$

Um die gesamte Stichprobe von der Anzahl der sichtbaren Individuen zu separieren, wird ein beobachtetes Individuum im Folgenden über den Spaltenvektor  $(\tilde{X}_j, \tilde{T}_j)'$ ,  $j = 1, \dots, m \leq n$ , beschrieben.

Zu beachten ist, dass zwar die Anzahl der Beobachtungen  $m$  eingesehen werden kann, wohingegen der Umfang  $n$  der Stichprobe allerdings unbekannt ist. Folglich fungiert  $n$  nach der Hinzunahme von (B5) als zweiter Parameter neben  $\theta_0$  im beobachtbaren Modell.

Im nächsten Schritt wird die Wahrscheinlichkeit  $\alpha_{\theta_0}^{DT}$  für das Ereignis, dass ein Individuum bei doppelt trunkiertem Design beobachtet wird, berechnet:

$$\begin{aligned} \alpha_{\theta_0}^{DT} &= P_{\theta_0} \{ (X_i, T_i)' \in \mathbb{D} \} = P_{\theta_0} \{ T_i + 1 \leq X_i \leq T_i + s \} \\ &= \sum_{t=0}^{G-1} \sum_{x=t+1}^{t+s} f_U(t) f_G(x; \theta_0) dx dt \\ &= \sum_{t=0}^{G-1} \frac{1}{G} \sum_{x=t+1}^{t+s} \theta_0 (1 - \theta_0)^{x-1} dx dt \\ &= \sum_{t=0}^{G-1} \frac{1}{G} \sum_{x=t+1}^{t+s} \theta_0 (1 - \theta_0)^{x-1} \\ &= \sum_{t=0}^{G-1} \frac{\theta_0}{G} \sum_{x=t}^{t+s-1} (1 - \theta_0)^x \\ &= \sum_{t=0}^{G-1} \frac{\theta_0}{G} \left[ \sum_{x=0}^{t+s-1} (1 - \theta_0)^x - \sum_{x=0}^{t-1} (1 - \theta_0)^x \right] \end{aligned}$$

Unter Berücksichtigung des Verhältnisses

$$\sum_{x=0}^{t-1} (1 - \theta_0)^x = \frac{1 - (1 - \theta_0)^t}{\theta_0}$$

folgt weiter:

$$\begin{aligned} \alpha_{\theta_0}^{DT} &= \frac{\theta_0}{G} \sum_{t=0}^{G-1} \left[ \frac{1 - (1 - \theta_0)^{t+s}}{\theta_0} - \frac{1 - (1 - \theta_0)^t}{\theta_0} \right] \\ &= \frac{1}{G} \left[ \sum_{t=0}^{G-1} (1 - \theta_0)^{t-1} - \sum_{t=0}^{G-1} (1 - \theta_0)^{t+s} \right] \\ &= \frac{1}{G} \left[ \sum_{t=0}^{G-1} (1 - \theta_0)^{t-1} - (1 - \theta_0)^s \sum_{t=0}^{G-1} (1 - \theta_0)^t \right] \\ &= \frac{1}{G} \left[ \frac{1 - (1 - \theta_0)^G}{\theta_0} - (1 - \theta_0)^s \frac{1 - (1 - \theta_0)^G}{\theta_0} \right] \\ &= \frac{1}{\theta_0 G} \left[ (1 - (1 - \theta_0)^s) (1 - (1 - \theta_0)^G) \right] \end{aligned}$$

Abschließend liefert ausmultiplizieren:

$$\alpha_{\theta_0}^{DT} = \frac{(1 - \theta_0)^{G+s} - (1 - \theta_0)^G - (1 - \theta_0)^s + 1}{\theta_0 G} \quad (4.1)$$

**Bemerkung 16.** *Im späteren Verlauf wird zur Maximierung einer logarithmierten Likelihood-Funktion die partielle Ableitung  $\partial \alpha_{\theta_0} / \partial \theta_0$  benötigt. Diese ist gegeben durch:*

$$\begin{aligned} \frac{\partial \alpha_{\theta_0}^{DT}}{\partial \theta_0} &= \\ &= \frac{-1}{\theta_0^2 G} \left[ ((G + s - 1)\theta_0 + 1)(1 - \theta_0)^{G+s-1} - ((G - 1)\theta_0 + 1)(1 - \theta_0)^{G-1} \right. \\ &\quad \left. - ((s - 1)\theta_0 + 1)(1 - \theta_0)^{s-1} + 1 \right] \end{aligned} \quad (4.2)$$

Als Nächstes muss die gemeinsame Verteilungsfunktion, bedingt auf der Beobachtbarkeit der Individuen, definiert werden:

**Definition 9.** *Die Spaltenvektoren  $(\tilde{X}_1, \tilde{T}_1)', (\tilde{X}_2, \tilde{T}_2)', \dots$  seien unabhängig identisch verteilt mit der gemeinsamen Verteilungsfunktion:*

$$F^{\tilde{X}, \tilde{T}}(x, t) = P\{X_i \leq x, T_i \leq t | T_i + 1 \leq X_i \leq T_i + s\} \quad (4.3)$$

Mithilfe der Erkenntnisse (4.1) und (4.3) lässt sich auch die gemeinsame Wahrschein-



Des Weiteren sei  $\mathcal{L}(X_i, T_i)$  die unbedingte Verteilung von  $(X_i, T_i)'$ . Somit gilt:

$$\begin{aligned} \alpha_{\theta_0}^{DT} F^{\tilde{X}, \tilde{T}}(x, t) &= \mathcal{L}(X_i, T_i)(\mathbb{E}_1) = F_G(x; \theta_0) F_U(t) - \mathcal{L}(X_i, T_i)(\mathbb{E}_2 \cup \mathbb{E}_3) \\ &= (1 - (1 - \theta_0)^x) \frac{t + 1}{G} - \mathcal{L}(X_i, T_i)(\mathbb{E}_2) - \mathcal{L}(X_i, T_i)(\mathbb{E}_3) \end{aligned}$$

Zur Berechnung von  $\mathcal{L}(X_i, T_i)(\mathbb{E}_2)$  und  $\mathcal{L}(X_i, T_i)(\mathbb{E}_3)$  werden für alle Punkte, die in den Mengen enthalten sind, die entsprechenden Wahrscheinlichkeiten aufsummiert. Für  $(x, t)' \in \mathbb{D}$  haben  $\mathbb{E}_2$  und  $\mathbb{E}_3$  eine Dreiecksstruktur und es ergibt sich:

$$\begin{aligned} R(x, t) &= \mathcal{L}(X_i, T_i)(\mathbb{E}_2) + \mathcal{L}(X_i, T_i)(\mathbb{E}_3) \\ &= \mathbf{1}_{\{1 \leq t\}} \sum_{l=1}^t \frac{1}{G} \sum_{k=1}^l \theta_0 (1 - \theta_0)^{k-1} + \mathbf{1}_{\{s+1 \leq x\}} \sum_{k=s+1}^x \theta_0 (1 - \theta_0)^{k-1} \sum_{l=0}^{k-s-1} \frac{1}{G} \quad (4.5) \end{aligned}$$

Außerdem gilt dann für  $(x, t)' \in \mathbb{D}$ :

$$\begin{aligned} &R(x, t) - R(x-1, t) - R(x, t-1) + R(x-1, t-1) \\ &= \frac{\theta_0}{G} \left[ \mathbf{1}_{\{1 \leq t\}} \sum_{l=1}^t \sum_{k=1}^l (1 - \theta_0)^{k-1} + \mathbf{1}_{\{s+1 \leq x\}} \sum_{k=s+1}^x \sum_{l=0}^{k-s-1} (1 - \theta_0)^{k-1} \right. \\ &\quad - \mathbf{1}_{\{1 \leq t\}} \sum_{l=1}^t \sum_{k=1}^l (1 - \theta_0)^{k-1} - \mathbf{1}_{\{s+1 \leq x-1\}} \sum_{k=s+1}^{x-1} \sum_{l=0}^{k-s-1} (1 - \theta_0)^{k-1} \\ &\quad - \mathbf{1}_{\{1 \leq t-1\}} \sum_{l=1}^{t-1} \sum_{k=1}^l (1 - \theta_0)^{k-1} - \mathbf{1}_{\{s+1 \leq x\}} \sum_{k=s+1}^x \sum_{l=0}^{k-s-1} (1 - \theta_0)^{k-1} \\ &\quad \left. + \mathbf{1}_{\{1 \leq t-1\}} \sum_{l=1}^{t-1} \sum_{k=1}^l (1 - \theta_0)^{k-1} + \mathbf{1}_{\{s+1 \leq x-1\}} \sum_{k=s+1}^{x-1} \sum_{l=0}^{k-s-1} (1 - \theta_0)^{k-1} \right] = 0 \quad (4.6) \end{aligned}$$

□

Nachdem nun die grundlegenden Zusammenhänge hergeleitet wurden, kann darauf aufbauend im folgenden Kapitel eine Likelihood-Funktion zur Schätzung der Ausfallwahrscheinlichkeit  $\theta_0$  bei einem beidseitig trunkiertem Beobachtungsdesign approximiert werden.

## 4.2. Likelihood-Approximation

Im Folgenden soll die Likelihood-Funktion zur Bestimmung einer Schätzung von  $\theta_0$  und  $n$  hergeleitet werden. Dazu wird die nach der Trunkierung beobachtbare Stichprobe als stochastischer Prozess, der mithilfe eines gemischten empirischen Prozesses approximiert werden kann, beschrieben [siehe Weißbach und Wied, 2022, Kapitel 3].

Für einen Punkt  $(x, t)' \in \mathbb{S}$  und eine Menge  $\mathbb{U} \in \mathcal{U}$  sei das zugehörige Dirac-Maß

definiert durch [siehe Reiss, 1993, Kapitel 1.1]:

$$\epsilon_{(x,t)'}(\mathbb{U}) := \begin{cases} 1, & \text{wenn } (x,t)' \in \mathbb{U} \\ 0, & \text{sonst} \end{cases} \quad (4.7)$$

Zunächst erfolgt eine Betrachtung der Stichprobe vor der Trunkierung, bestehend aus den Individuen  $(X_i, T_i)$ ,  $i \in 1, \dots, n$ , sowie den Annahmen 2, (B1)-(B4). Mithilfe des Dirac-Maßes (4.7) lässt sich dann die Anzahl aller sich in  $\mathbb{U}$  befindenden Punkte berechnen durch:

$$\left| \left\{ (X_i, T_i) \in \mathbb{S} : i \in \{1, \dots, n\} \right\} \cap \mathbb{U} \right| = \sum_{i=1}^n \epsilon_{(x_i, t_i)'}(\mathbb{U})$$

Die Anzahl der Punkte, die sich in  $\mathbb{U}$  befinden, kann folglich durch:

$$\mu(\mathbb{U}) := \sum_{i=1}^n \epsilon_{(x_i, t_i)'}(\mathbb{U}) \quad (4.8)$$

repräsentiert werden. Die Funktion  $\mu : \mathcal{U} \mapsto \mathbb{N}_0$  beschreibt dann ein Maß auf dem  $\sigma$ -Feld  $\mathcal{U}$ . So ein diskretes Maß  $\mu$  wird auch als Punktmaß bezeichnet.

Weiter sei der Raum der Punktmaße auf  $\mathcal{U}$  gegeben durch  $\mathbb{M} := \mathbb{M}(\mathbb{S}, \mathcal{U})$  [siehe Reiss, 1993, Formel (1.1)], versehen mit der  $\sigma$ -Algebra  $\mathcal{M}$ .

Mithilfe von (4.8) lassen sich Punktprozesse definieren.

**Definition 10** (Punktprozess). *Sei  $(\Omega, \mathcal{F}, P_{\theta_0})$  ein Wahrscheinlichkeitsraum, auf dem die Zufallsvektoren  $(X_i, T_i)'$  mit  $i \in \{1, \dots, n\}$  definiert sind. Eine Abbildung*

$$H : \Omega \mapsto \mathbb{M}(\mathbb{S}, \mathcal{U})$$

*heißt Punktprozess auf  $(\mathbb{S}, \mathcal{U})$ , wenn  $H$  messbar bezüglich  $\mathcal{F}$  und  $\mathcal{M}$  ist. Folglich handelt es sich bei  $H$  um ein zufälliges Punktmaß auf  $(\mathbb{S}, \mathcal{U})$ .*

**Bemerkung 17.** *Besitzt ein Punktprozess die Summendarstellung*

$$H_n = \sum_{i=1}^n \epsilon_{(X_i, T_i)'} \quad (4.9)$$

*so kann dieser auch als empirischer Prozess bezeichnet werden.*

Basierend auf (4.9) können die Daten des beobachtbaren Modells als trunkierter empirischer Prozess dargestellt werden [siehe Reiss, 1993, Kapitel 1.1, Beispiel 1.1.1 (iii)]:

$$H_{n, \mathbb{D}} := H_n(\cdot \cap \mathbb{D}) = \sum_{i=1}^n \epsilon_{(X_i, T_i)'}(\cdot \cap \mathbb{D}) \quad (4.10)$$

Die Anzahl der sichtbaren Individuen ist dann gegeben durch  $H_{n, \mathbb{D}}(\mathbb{S}) = M$  mit der Realisierung  $m$ . Es sei darauf hingewiesen, dass die Bezeichnung  $M$  im restlichen Teil der

Arbeit das Martingal in der Doob-Zerlegung eines diskreten Zählprozesses bezeichnet und lediglich für dieses Kapitel 4.2 eine andere Bedeutung bekommt.

Um eine Likelihood-Funktion zur Schätzung von  $n$  und  $\theta_0$  herzuleiten, wird im nächsten Schritt das so genannte Intensitätsmaß von  $H_{n,\mathbb{D}}$  eingeführt [siehe Reiss, 1993, Formel (1.7)].

**Definition 11** (Intensitätsmaß). *Sei  $H_{n,\mathbb{D}}$  ein zufälliges Maß auf dem Messraum  $(\mathbb{S}, \mathcal{U})$ . Dann ist das Intensitätsmaß von  $H_{n,\mathbb{D}}$  gegeben durch:*

$$\eta_{n,\mathbb{D}}(\mathbb{U}) := E_{\theta_0}\{H_{n,\mathbb{D}}(\mathbb{U})\}, \quad \text{mit } \mathbb{U} \in \mathcal{U} \quad (4.11)$$

Es genügt, das Intensitätsmaß für Mengen der Form  $\{1, \dots, x\} \times \{0, \dots, t\}$  zu berechnen, da alle anderen Gebiete als Komposition von Mengen dieser Form dargestellt werden können. Beispielsweise gilt für  $\kappa \in \mathbb{N}$  mit  $\kappa \leq x$  sowie für  $\lambda \in \mathbb{N}_0$  mit  $\lambda \leq t$ , dass

$$\begin{aligned} & \{\kappa, \dots, x\} \times \{\lambda, \dots, t\} \\ &= (\{1, \dots, x\} \times \{0, \dots, t\}) \setminus (\{1, \dots, x\} \times \{0, \dots, \lambda\} \cup \{1, \dots, \kappa\} \times \{0, \dots, t\}) \end{aligned}$$

Mithilfe von (4.4) ergibt sich:

$$\begin{aligned} \eta_{n,\mathbb{D}}(\{1, \dots, x\} \times \{0, \dots, t\}) &:= E_{\theta_0}\left[H_{n,\mathbb{D}}(\{1, \dots, x\} \times \{0, \dots, t\})\right] \\ &= nP_{\theta_0}\left\{(X_i, T_i)' \in (\{1, \dots, x\} \times \{0, \dots, t\}) \cap \mathbb{D}\right\} \quad (4.12) \end{aligned}$$

$$\begin{aligned} &= n\alpha_{\theta_0}^{DT} F^{\tilde{X}, \tilde{T}}(x, t) \\ &= n\left[\left(1 - (1 - \theta_0)^x\right) \frac{t+1}{G} - R(x, t)\right] \quad (4.13) \end{aligned}$$

Es sei angemerkt, dass  $\eta_{n,\mathbb{D}}(\mathbb{S}) = n\alpha_{\theta_0}$  gilt. Des Weiteren kann mithilfe des folgenden Satzes [siehe Reiss, 1993, Theorem 1.4.1] gezeigt werden, dass  $H_{n,\mathbb{D}}$  verteilungsgleich zu einem gemischten empirischen Binomial-Prozess ist. Dabei wird Verteilungsgleichheit im Folgenden mit dem Symbol  $=^d$  gekennzeichnet.

**Satz 10.** *Für  $\mathbb{D} \in \mathcal{U}$  sei der trunkierte empirische Prozess gegeben durch:*

$$H_{n,\mathbb{D}} = \sum_{i=1}^n \epsilon_{(X_i, T_i)'}(\cdot \cap \mathbb{D})$$

*Dabei sind  $(X_i, T_i)'$  unabhängig identisch verteilte Zufallsvektoren aus  $\mathbb{S}$  mit der unbedingten, gemeinsamen Verteilung  $\mathcal{L}(X_i, T_i)$ . Dann gilt*

$$H_{n,\mathbb{D}} =^d H_B = \sum_{i=1}^{\beta} \epsilon_{(Y_i)}$$

wobei  $\beta, Y_1, \dots, Y_n$  unabhängig mit

$$\mathcal{L}(Y_i) = \frac{\mathcal{L}(X_i, T_i)(\cdot \cap \mathbb{D})}{\mathcal{L}(X_i, T_i)(\mathbb{D})} = \frac{1}{\alpha_{\theta_0}^{DT}} \mathcal{L}(X_i, T_i)(\cdot \cap \mathbb{D})$$

und

$$\mathcal{L}(\beta) = B_{(n, \alpha_{\theta_0}^{DT})}$$

sind.  $B_{(n, \alpha_{\theta_0}^{DT})}$  bezeichnet dabei die Binomialverteilung mit den Parametern  $n$  und  $\alpha_{\theta_0}^{DT}$ .

Ist die Studiendauer  $s$  im Vergleich zur Länge des Beobachtungszeitraums  $G$  relativ klein, ist folglich auch  $\mathbb{D}$  relativ klein im Verhältnis zu  $\mathbb{S}$ . Dann kann  $H_{n, \mathbb{D}}$ , anstatt durch einen gemischten empirischen Binomial-Prozess, auch durch einen gemischten empirischen Poisson-Prozess  $H_n^*$  approximiert werden. Dabei ist die Poisson-Verteilung folgendermaßen definiert:

**Definition 12** (Poisson-Verteilung). Sei  $\tilde{Z}$  eine diskrete Zufallsvariable mit dem Bild  $\tilde{\mathbf{z}} \in \mathbb{N}_0$ . Dann heißt  $\tilde{Z}$  poissonverteilt mit dem Parameter  $n\alpha_{\theta_0}^{DT} > 0$  genau dann, wenn die zugehörige Wahrscheinlichkeitsfunktion gegeben ist durch:

$$f_P(\tilde{\mathbf{z}}; n\alpha_{\theta_0}^{DT}) = P_{n\alpha_{\theta_0}^{DT}}\{\tilde{Z} = \tilde{\mathbf{z}}\} = \begin{cases} \frac{(n\alpha_{\theta_0}^{DT})^{\tilde{\mathbf{z}}}}{\tilde{\mathbf{z}}!} \exp\{-n\alpha_{\theta_0}^{DT}\}, & \text{falls } \tilde{\mathbf{z}} \in \mathbb{N}_0 \\ 0, & \text{sonst} \end{cases} \quad (4.14)$$

Auch sei angemerkt, dass sich, basierend auf (4.12), die unbedingte Verteilung der beobachtbaren Individuen folgendermaßen berechnen lässt:

$$\begin{aligned} & \mathcal{L}(\tilde{X}_j, \tilde{T}_j)\left(\{1, \dots, x\} \times \{0, \dots, t\}\right) & (4.15) \\ & = P_{\theta_0}\left\{(\tilde{X}_j, \tilde{T}_j)' \in \left(\{1, \dots, x\} \times \{0, \dots, t\}\right)\right\} \\ & = P_{\theta_0}\left\{(X_i, T_i)' \in \left(\{1, \dots, x\} \times \{0, \dots, t\}\right) \mid (X_i, T_i)' \in \mathbb{D}\right\} \\ & = \frac{P_{\theta_0}\left\{\left\{(X_i, T_i)' \in \left(\{1, \dots, x\} \times \{0, \dots, t\}\right)\right\} \cap \left\{(X_i, T_i)' \in \mathbb{D}\right\}\right\}}{P_{\theta_0}\{(X_i, T_i)' \in \mathbb{D}\}} \\ & = \frac{\eta_{n, \mathbb{D}}\left(\{1, \dots, x\} \times \{0, \dots, t\}\right)}{n\alpha_{\theta_0}^{DT}} & (4.16) \end{aligned}$$

Im nächsten Schritt kann mithilfe des Existenz- und Eindeigkeitstheorems (vgl. Satz 17) ein gemischter empirischer Poisson-Prozess  $H_n^*$  definiert werden. Später wird sich herausstellen, dass dieser Prozess unter gewissen Bedingungen eine gute Approximation für den trunkierten empirischen Prozess  $H_{n, \mathbb{D}}$  darstellt [siehe Weißbach und Wied, 2022, Definition 2].

**Definition 13.** Angenommen (B1)-(B5) seien erfüllt und  $\tilde{Z}$  sei eine Poisson-verteilte Zufallsvariable mit dem Parameter  $n\alpha_{\theta_0}^{DT}$ . Des Weiteren seien  $(\tilde{X}_j, \tilde{T}_j)$ ,  $j \in \{1, \dots, m\}$

von  $\tilde{Z}$  unabhängige Zufallsvektoren aus Definition 9. Dann ist

$$H_n^* := \sum_{j=1}^{\tilde{Z}} \epsilon_{(\tilde{X}_j, \tilde{T}_j)'}.$$

wegen  $\eta_{n, \mathbb{D}}(\mathbb{S}) = n\alpha_{\theta_0}^{DT} < \infty$  und  $\mathcal{L}(\tilde{X}_j, \tilde{T}_j) = \eta_{n, \mathbb{D}}/(n\alpha_{\theta_0}^{DT})$  (siehe (4.16)) ein Poisson-Prozess mit dem Intensitätsmaß  $\eta_n^* \equiv \eta_{n, \mathbb{D}}$  sowie  $H_n^*(\mathbb{S}) = \tilde{Z}$ .

Damit der trunkierte empirische Prozess  $H_{n, \mathbb{D}}$  durch den gemischten empirischen Poisson-Prozess  $H_n^*$  approximiert werden kann, müssen die beiden Prozesse, intuitiv formuliert, nahe beieinander liegen. Um für diesen Zweck den Abstand zwischen den Verteilungen zweier empirischer Prozesse quantifizieren zu können, wird zunächst die so genannte Hellingermetrik eingeführt [siehe Reiss, 1993, Formel (1.21)].

**Definition 14** (Hellingermetrik). Gegeben seien zwei empirische Prozesse  $H_{n, \mathbb{D}}$  und  $H_n^*$ . Dann ist die Hellingermetrik, auch Hellingerabstand genannt, der Verteilungen  $\mathcal{L}(H_{n, \mathbb{D}})$  und  $\mathcal{L}(H_n^*)$  gegeben durch:

$$H(\mathcal{L}(H_{n, \mathbb{D}}), \mathcal{L}(H_n^*)) := \left( \int (\sqrt{f_{n, \mathbb{D}}} - \sqrt{f_n^*})^2 d\eta \right)^{1/2}$$

Dabei sind  $f_{n, \mathbb{D}}$  und  $f_n^*$  die Dichten von  $\mathcal{L}(H_{n, \mathbb{D}})$  respektive  $\mathcal{L}(H_n^*)$  bezüglich eines Maßes  $\eta$ .

Auf der Grundlage der Hellingermetrik kann im nächsten Schritt das Approximationstheorem eingeführt werden [siehe Reiss, 1993, Theorem 1.4.2]:

**Satz 11** (Approximationstheorem). Sei  $H_{n, \mathbb{D}}$  ein trunkierter empirischer Prozess mit der Darstellung (4.10) und  $\alpha_{\theta_0}^{DT} = P\{(X_i, T_i)' \in \mathbb{D}\}$  für  $i \in \{1, \dots, n\}$ . Des Weiteren sei  $H_n^*$  ein gemischter empirischer Poisson-Prozess mit demselben Intensitätsmaß  $\eta_{n, \mathbb{D}}$  wie  $H_{n, \mathbb{D}}$ . Dann gilt:

$$H(\mathcal{L}(H_{n, \mathbb{D}}), \mathcal{L}(H_n^*)) \leq 3^{1/2} \alpha_{\theta_0}^{DT} \quad (4.17)$$

Wenn folglich  $\mathbb{D}$  im Verhältnis zu  $\mathbb{S}$  klein ist, so nimmt  $\alpha_{\theta_0}^{DT}$  ebenfalls einen kleinen Wert an. Das Approximationstheorem besagt, dass in diesem Fall  $H_{n, \mathbb{D}}$  und  $H_n^*$  im Sinne der Hellingermetrik nah beieinander liegen. In diesem Fall kann stellvertretend für  $H_{n, \mathbb{D}}$  die Likelihood für  $H_n^*$  berechnet werden. Dazu wird zunächst die Radon-Nikodym-Dichte  $f_{\theta_0}$  von  $\eta_{n, \mathbb{D}}$  bezüglich eines dominierenden Maßes  $\eta_0$  berechnet. Dabei ist  $\eta_0$  das Intensitätsmaß des folgenden Poisson-Prozesses [siehe Weißbach und Wied, 2022, Definition 3].

**Definition 15.** Sei  $H_0$  ein empirischer Prozess der Form

$$H_0 = \sum_{j=1}^{\tilde{Z}_0} \epsilon_{(X_j^0, T_j^0)'} \quad (4.18)$$

wobei  $Z_0$  poissonverteilt mit Parameter  $((G-1)+s)G$  und  $(X_i^0, T_i^0)'$  diskret gleichverteilt auf dem Träger  $\mathbb{G} := \{1, \dots, (G-1)+s\} \times \{0, \dots, G-1\}$  ist. Die Zufallsvariable  $\tilde{Z}_0$  sei dabei stochastisch unabhängig von  $(X_j^0, T_j^0)'$ .

Für das zugehörige Intensitätsmaß  $\eta_0$  von  $H_0$  ergibt sich für  $\mathbb{U} \in \mathcal{U}$ :

$$\begin{aligned} \eta_0(\mathbb{U}) &= \mathbb{E}\{H_0(\mathbb{U})\} = \mathbb{E}\left\{\sum_{j=1}^{\tilde{Z}_0} \epsilon_{(X_j^0, T_j^0)'}(\mathbb{U})\right\} \\ &= \mathbb{E}\{\tilde{Z}_0\} \mathbb{E}\{\epsilon_{(X_j^0, T_j^0)'}(\mathbb{U})\} \\ &= ((G-1)+s)G \frac{\#\{(x, t)' : (x, t)' \in \mathbb{U} \cap (\{1, \dots, (G-1)+s\} \times \{0, \dots, G-1\})\}}{((G-1)+s)G} \\ &= \#\{(x, t)' : (x, t)' \in \mathbb{U} \cap \mathbb{G}\} \end{aligned}$$

Somit ist  $\eta_0$  das auf  $\mathbb{G}$  restringierte Zählmaß und es gilt  $\eta_0(\mathbb{S}) = ((G-1)+s)G < \infty$ . Außerdem ist  $\tilde{Z}_0$  poissonverteilt mit Parameter  $\eta_0(\mathbb{S})$  und es gilt:

$$P\{(X_j^0, T_j^0)' = (x_j^0, t_j^0)'\} = \frac{\eta_0((x_j^0, t_j^0)')}{\eta_0(\mathbb{S})}$$

Nach dem Existenz- und Eindeigkeitstheorem für Poisson-Prozesse (vgl. Appendix, Satz 17) ist  $H_0$  dann ein Poisson-Prozess mit dem endlichen Intensitätsmaß  $\eta_0$ .

Im nächsten Schritt kann mithilfe des folgenden Satzes [siehe Reiss, 1993, Theorem 3.1.1] unter Zuhilfenahme von  $H_0$  die  $\mathcal{L}(H_0)$ -Dichte von  $\mathcal{L}(H_n^*)$  bestimmt werden.

**Satz 12.** *Seien  $H_n^*$  und  $H_0$  Poisson-Prozesse mit den endlichen Intensitätsmaßen  $\eta_{n, \mathbb{D}}$  und  $\eta_0$ , die auf dem gemeinsamen Messraum  $(\mathbb{S}, \mathcal{U})$  definiert sind. Sei  $f_{\theta_0} : \mathbb{S} \mapsto \mathbb{R}_0^+$  die Radon-Nikodym-Dichte von  $\eta_{n, \mathbb{D}}$  bezüglich  $\eta_0$ . Dann ist die  $\mathcal{L}(H_0)$ -Dichte von  $\mathcal{L}(H_n^*)$ , die mit  $g$  bezeichnet wird, gegeben durch:*

$$g(\mu) = \left( \prod_{i=1}^{\mu(\mathbb{S})} f_{\theta_0}(X_i, T_i) \right) \exp(\eta_0(\mathbb{S}) - \eta_{n, \mathbb{D}}(\mathbb{S})) \quad (4.19)$$

Dabei ist  $\mu = \sum_{j=1}^{\mu(\mathbb{S})} \epsilon_{(\tilde{x}_j, \tilde{t}_j)'}$  und es gilt die Konvention  $\prod_{i=1}^0 f_{\theta_0}(X_i, T_i) = 1$

Die Radon-Nikodym Dichte von  $\eta_{n, \mathbb{D}}$  bezüglich  $\eta_0$  ist eine Abbildung  $f_{\theta_0} : \mathbb{S} \mapsto \mathbb{R}_0^+$ , sodass für alle  $\mathbb{U} \in \mathcal{U}$  gilt:

$$\eta_{n, \mathbb{D}}(\mathbb{U}) = \int_{\mathbb{U}} f_{\theta_0} d\eta_0 \quad (4.20)$$

Die Existenz einer solchen Abbildung ist genau dann gegeben, wenn  $\eta_{n, \mathbb{D}}$  absolut stetig bezüglich  $\eta_0$  ist, also  $\eta_0(\mathbb{U}) = 0 \Rightarrow \eta_{n, \mathbb{D}}(\mathbb{U}) = 0$  für alle  $\mathbb{U} \in \mathcal{U}$  erfüllt ist. Mit  $\mathbb{D} \subset$

$\{1, \dots, (G-1) + s\} \times \{0, \dots, G-1\}$  und (4.12) gilt:

$$\begin{aligned} \eta_0(\mathbb{U}) = 0 &\Rightarrow \#(x, t)' : (x, t)' \in \mathbb{U} \cap (\{1, \dots, (G-1) + s\} \times \{0, \dots, G-1\}) \\ &\Rightarrow \#(x, t)' : (x, t)' \in \mathbb{U} \cap \mathbb{D} \\ &\Rightarrow P\{(x, t)' \in \mathbb{U} \cap \mathbb{D}\} = 0 \\ &\Rightarrow \eta_{n, \mathbb{D}}(\mathbb{U}) = 0 \end{aligned}$$

Folglich existiert eine Abbildung  $f_{\theta_0}$  mit der Eigenschaft (4.20). Für Mengen der Form  $\mathbb{U} = \{1, \dots, x\} \times \{0, \dots, t\}$  ergibt sich:

$$\begin{aligned} \eta_{n, \mathbb{D}}(\{1, \dots, x\} \times \{0, \dots, t\}) &= \int_1^x \int_0^t f_{\theta_0}(a_1, a_2) \eta_0(da_1) \eta_0(da_2) \\ &= \sum_{k=1}^x \sum_{l=0}^t f_{\theta_0}(k, l) \end{aligned}$$

Umstellen nach  $f_{\theta_0}(x, t)$  liefert

$$\begin{aligned} f_{\theta_0}(x, t) &= \eta_{n, \mathbb{D}}(\{1, \dots, x\} \times \{0, \dots, t\}) - \eta_{n, \mathbb{D}}(\{1, \dots, x-1\} \times \{0, \dots, t\}) \\ &\quad - \eta_{n, \mathbb{D}}(\{1, \dots, x\} \times \{0, \dots, t-1\}) + \eta_{n, \mathbb{D}}(\{1, \dots, x-1\} \times \{0, \dots, t-1\}) \\ &= n \left[ (1 - (1 - \theta_0)^x) \frac{t+1}{G} - (1 - (1 - \theta_0)^{x-1}) \frac{t+1}{G} \right. \\ &\quad - (1 - (1 - \theta_0)^x) \frac{t}{G} + (1 - (1 - \theta_0)^{x-1}) \frac{t}{G} \\ &\quad \left. - R(x, t) + R(x-1, t) + R(x, t-1) - R(x-1, t-1) \right] \\ &= \frac{n}{G} \left[ (1 - (1 - \theta_0)^x)(t+1) - (1 - (1 - \theta_0)^{x-1})(t+1) \right. \\ &\quad \left. - (1 - (1 - \theta_0)^x)t + (1 - (1 - \theta_0)^{x-1})t \right] \\ &= \frac{n}{G} \left[ (1 - (1 - \theta_0)^x) - (1 - (1 - \theta_0)^{x-1}) \right] \\ &= \frac{n}{G} \left[ (1 - \theta_0)^{x-1} - (1 - \theta_0)^x \right] \end{aligned} \tag{4.21}$$

wobei die zweite Gleichheit durch einsetzen von (4.13) und die dritte Gleichheit durch Eigenschaft (4.6) erfolgt.

Abschließend ergibt sich die Likelihood-Funktion zur Schätzung der Ausfallwahrscheinlichkeit  $\theta_0$  sowie des Stichprobenumfangs  $n$  als die Dichte von  $H_n^*$ , die an der Realisierung  $h_n^*$  evaluiert wird. Mithilfe von (4.19) gilt:

$$L^{DT}(\theta, n; h_n^*) = \frac{n^{h_n^*(\mathbb{S})}}{G^{h_n^*(\mathbb{S})}} \prod_{j=1}^{h_n^*(\mathbb{S})} \left( (1 - \theta)^{\tilde{x}_j - 1} - (1 - \theta)^{\tilde{x}_j} \right) \exp \left( G((G-1) + s) - n\alpha_{\theta}^{DT} \right) \tag{4.22}$$

In dem zu untersuchenden Modell ist  $m$  als Realisierung von  $H_n^*(\mathbb{S})$  gegeben. Da  $n$  im Rahmen des beobachtbaren Modells durch Hinzunahme von (B5) einen Parameter darstellt, wird im Folgenden der ML-Schätzer  $\hat{n}$  von  $n$  ermittelt. Zunächst ergibt sich folgende logarithmierte Likelihood-Funktion:

$$\begin{aligned} \log L^{DT}(\theta, n; m, \tilde{\mathbf{x}}_1, \dots, \tilde{\mathbf{x}}_m) \\ = m(\log n - \log G) + \sum_{j=1}^m \log\left((1 - \theta)^{\tilde{\mathbf{x}}_{j-1}} - (1 - \theta)^{\tilde{\mathbf{x}}_j}\right) + G\left((G - 1) + s\right) - n\alpha_\theta^{DT} \end{aligned}$$

Ableiten nach  $n$  liefert

$$\frac{\partial}{\partial n} \log L^{DT}(\theta, n; m, \tilde{\mathbf{x}}_1, \dots, \tilde{\mathbf{x}}_m) = \frac{m}{n} - \alpha_\theta^{DT}$$

und somit den ML-Schätzer

$$\hat{n} = \frac{m}{\alpha_\theta^{DT}} \quad (4.23)$$

Abschließend soll (4.22) an der Stelle des ML-Schätzers  $\hat{n} = m/\alpha_\theta$  in  $\theta$ -Richtung maximiert werden. Es gilt:

$$\begin{aligned} & \frac{\partial \log L^{DT}(\theta, m/\alpha_\theta^{DT}; m, \tilde{\mathbf{x}}_1, \dots, \tilde{\mathbf{x}}_m)}{\partial \theta} \\ &= \frac{\partial}{\partial \theta} \left[ m \log\left(\frac{m}{\alpha_\theta^{DT}}\right) + \sum_{j=1}^m \log\left((1 - \theta)^{\tilde{\mathbf{x}}_{j-1}} - (1 - \theta)^{\tilde{\mathbf{x}}_j}\right) \right] \\ &= \frac{\partial}{\partial \theta} \left[ \sum_{j=1}^m \log\left((1 - \theta)^{\tilde{\mathbf{x}}_{j-1}} - (1 - \theta)^{\tilde{\mathbf{x}}_j}\right) \right] - \frac{m}{\alpha_\theta^{DT}} \frac{\partial \alpha_\theta^{DT}}{\partial \theta} \\ &= \sum_{j=1}^m \frac{1 - \tilde{\mathbf{x}}_j \theta}{\theta(1 - \theta)} + \frac{m}{\theta \left[ (1 - \theta)^{G+s} - (1 - \theta)^G - (1 - \theta)^s + 1 \right]} \\ & \quad \cdot \left[ [(G + s - 1)\theta + 1](1 - \theta)^{G+s-1} - [(G - 1)\theta + 1](1 - \theta)^{G-1} \right. \\ & \quad \left. - [(s - 1)\theta + 1](1 - \theta)^{s-1} + 1 \right] \end{aligned} \quad (4.24)$$

Der Punktschätzer  $\hat{\theta}_n^{DT}$  für den wahren Parameterwert  $\theta_0$  kann dann als Nullstelle von (4.24) ermittelt werden. Abschließend liefert Einsetzen der Punktschätzung der Ausfallwahrscheinlichkeit in (4.23) eine Punktschätzung für den Stichprobenumfang  $n$ .

Die in diesem Kapitel hergeleiteten Schätzer werden in den folgenden Kapiteln im Rahmen der Auswertung des AFiD-Panels (Kapitel 5) und einer Simulation (Kapitel 6) neben dem Schätzer des geometrischen Parameters inklusive der Standardfehler bei linkstrunkiertem und rechtszensiertem Beobachtungsdesign (siehe Kapitel 3.1 bis 3.3.2) ebenfalls evaluiert.

# 5. Das AFiD-Panel

## 5.1. Beschreibung der Daten

Bisher wurden auf der sprachlichen Ebene Individuen betrachtet, die in einem bestimmten Jahr aus  $\{a, \dots, b\}$  entstehen und in einem späteren Jahr wieder ausfallen. Diese Individuen sind im Folgenden deutsche Unternehmen, die innerhalb eines Jahres gegründet und innerhalb eines anderen Jahres wieder geschlossen werden. Im folgenden Kapitel soll mithilfe des entwickelten Punktschätzers  $\hat{\theta}_n$  für die Ausfallwahrscheinlichkeit (vgl.(3.11)) die durchschnittliche Lebenserwartung deutscher Unternehmen geschätzt werden. Anschließend erfolgt eine Berechnung der zugehörigen Standardfehler  $SE_1$  (vgl. (3.46)),  $SE_2$  (vgl. (3.47)),  $SE_3$  (vgl. (3.75)) sowie  $SE_4$  (vgl. (A.8)). Als Datengrundlage dienen dabei die vom Forschungsdatenzentrum des statistischen Bundesamts (Destatis) erhobenen 'Amtlichen Firmendaten für Deutschland' (AFiD).

Destatis hat mit dem Berichtsjahr 2018 begonnen, Berichte zur Geschäftstätigkeit so zu erstellen, dass die dort untersuchten statistischen Einheiten der Unternehmensdefinition der Europäischen Kommission entsprechen. Darauf basierend wurde von Destatis der Zugang zu Angaben bezüglich Unternehmensschließungen in den Jahren 2018 und 2019 sowie hinsichtlich Informationen über Unternehmen, die das Jahr 2019 überdauert haben, für unter Anderem diese Arbeit zur Verfügung gestellt. Da alle Unternehmen, zu denen Angaben vorliegen, in den Jahren  $\{2013, \dots, 2017\}$  gegründet wurden, ergibt sich ein retrospektives Stichprobendesign.

Die Angaben, wie viele Unternehmen in welchem Jahr gegründet beziehungsweise geschlossen wurden, kann dabei Tabelle 5.1 entnommen werden.

Gründung	Schließung			Summe
	2018	2019	3000	
2013	37.016	17.295	191.803	246.114
2014	42.272	20.305	200.411	262.988
2015	35.588	23.353	209.649	268.590
2016	34.579	27.464	278.223	340.266
2017	18.687	18.633	292.566	329.886
Summe	168.142	107.050	1.172.652	1.447.844

Tabelle 5.1.: Tabellarische Übersicht der Gründungs- und Schließungsjahre von insgesamt ungefähr 1.4 Millionen deutschen Unternehmen; die Spalte mit dem Schließungsjahr 3000 umfasst alle Unternehmen, die bis einschließlich 2019 nicht geschlossen wurden und somit rechtszensiert sind

Eine detaillierte Beschreibung der Daten liefert Rink und Seiwert [2021] sowie die darin enthaltenen Referenzen. Dabei wurden die Merkmale 'urs\_we\_beginn\_datum' als Gründungsjahr und 'urs\_we\_ende\_datum' als Jahr der Schließung verwendet.

Als Grundgesamtheit wurde die Menge aller deutschen Unternehmen, die in den Kalenderjahren  $\{a, \dots, b\} = \{2013, \dots, 2017\}$  gegründet wurden, definiert. Somit beschreibt diese Zeitspanne den Entstehungszeitraum, der im folgenden auch als Gründungszeitraum bezeichnet wird, mit einer Anzahl von  $G = 5$  Jahren. Es sei angemerkt, dass für das  $i$ -te Unternehmen, mit  $i \in \{1, \dots, n\}$ , die Zufallsvariable  $T_i$  (vgl. (2.3)) das Alter des Unternehmens im Jahr 2017 angibt. Des Weiteren bezeichne  $X_i$  die mit dem wahren Parameterwert  $\theta_0$  geometrisch verteilte Lebensdauer des  $i$ -ten Unternehmens in Jahren.

Es sei angemerkt, dass das AFiD-Panel alle in den Jahren  $\{2013, \dots, 2017\}$  gegründeten Unternehmen enthält. Der unbekannte Umfang  $n$  der Stichprobe gleicht somit dem Umfang der Grundgesamtheit.

Beobachtungen hinsichtlich der Unternehmensschließungen liegen jedoch lediglich für die Jahre 2018 sowie 2019 vor. Die Länge des Beobachtungszeitraums beträgt somit  $s = |\{2018, 2019\}| = 2$  Jahre. Unternehmen, deren Schließung bereits vor 2018 stattgefunden haben, werden nicht beobachtet und entsprechend in den Daten nicht aufgeführt. Es liegt somit ein linksseitig trunkiertes Design vor. Die Unternehmen, deren Existenz trotz der Linkstrunkierung beobachtet werden können, bilden folglich die Menge der Beobachtungen, bestehend aus  $m = 1.447.844$  Unternehmen.

Auf der anderen Seite werden Unternehmen, die das Jahr 2019 ohne Schließung überstehen, zwar beobachtet, jedoch bleibt der genaue Zeitpunkt der Schließung unbekannt. Das Beobachtungsdesign ist somit zusätzlich rechtsseitig zensiert. Unternehmen mit einer solchen rechtszensierten Historie werden in den Daten mit dem Schließungsjahr 3000 gekennzeichnet. Somit konnte bei 1.172.652 Unternehmen keine Schließung in den Jahren 2018 und 2019 beobachtet werden.

Eine Übersicht der verschiedenen Verläufe eines Unternehmenslebens hinsichtlich der Beobachtbarkeit am Beispiel der Anwendung kann Abbildung 5.1 entnommen werden.

### **Angemessenheit der Modellannahmen**

Bevor in Kapitel 5.2 eine Punkt- und Intervallschätzung der Schließungswahrscheinlichkeit sowie der durchschnittlichen Lebenserwartung deutscher Unternehmen basierend auf dem AFiD-Panel durchgeführt wird, werden zunächst die Annahmen 1 (vgl. Kapitel 3.1) mit Hinblick auf die Daten diskutiert.

In Bezug auf (A1) sei angemerkt, dass die Annahme von geometrisch verteilten Lebensdauern  $X_i$  mit  $i \in \{1, \dots, n\}$  impliziert, dass die Schließungswahrscheinlichkeit im Laufe der Zeit nicht variiert, sondern konstant bleibt. In der Praxis kann diese Annahme zwar nicht garantiert werden, allerdings ist auch kein empirischer Nachweis für das Gegenteil bekannt. Bezüglich der Unabhängigkeit von  $X_i$  und  $T_i$  existieren erste Indizien dafür, dass diese im Kontext von Unternehmensdemografien nicht gegeben ist [siehe de Uña-Álvarez et al., 2024, Toparkus und Weißbach, 2025].

Die Annahmen (A2) und (A3) sind hingegen klar erfüllt.

Bezüglich Annahme (A4) muss hinsichtlich der Unabhängigkeitsannahme zwischen

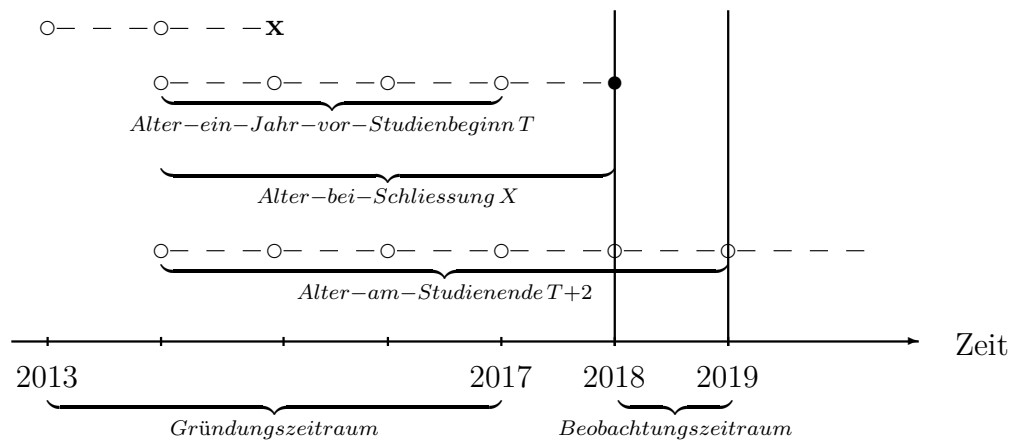


Abbildung 5.1.: Mögliche Pfade von Unternehmensleben, die sich hinsichtlich der Beobachtbarkeit unterscheiden; oberer Pfad: Gründung in 2013 und Schließung in 2015 und somit vor Beginn des Beobachtungszeitraums, weder Lebensdauer noch Existenz des Unternehmens werden beobachtet (Linkstrunkierung); mittlerer Pfad: Gründung in 2014 und Schließung in 2018 und somit im Beobachtungszeitraum, Lebensdauer sowie Gründungsjahr werden beobachtet; unterer Pfad: Gründung in 2014 und Schließung nach 2019 und somit nach Ende des Beobachtungszeitraums, die beobachtete Information ist das Gründungsjahr sowie das Überleben bis einschließlich 2019 (Rechtszensur)

den Unternehmen darauf hingewiesen werden, dass eine gewisse Abhängigkeit in der Praxis durchaus denkbar sein kann, da sich die Lebensdauern teilweise stark überschneiden. Ebenso bilden  $(X_i, T_i)'$ ,  $i \in \{1, \dots, n\}$  keine einfache Zufallsstichprobe, sondern die Grundgesamtheit, da die latenten Messungen alle deutschen Unternehmen, die in den Jahren  $\{2013, \dots, 2017\}$  gegründet wurden, abdecken.

## 5.2. Punkt- und Intervallschätzung

### Punkt- und Intervallschätzung bei linkstrunkiertem und rechtszensiertem Beobachtungsdesign

Zunächst wird basierend auf den Daten (siehe Tabelle 5.1) die logarithmierte, marginale und bedingte Likelihood-Funktion (3.2) berechnet. Das Ergebnis lässt sich Abbildung 5.2 entnehmen.

Zum Erhalt einer analytischen Lösung für die Punktschätzung des geometrischen Parameters kann sich der Formel (3.11) bedient werden. Dabei lässt sich  $m_{uncens}$  als Summe aller in 2018 und 2019 beobachteten Schließungen berechnen:

$$m_{uncens} = 168.142 + 107.050 = 275.192$$

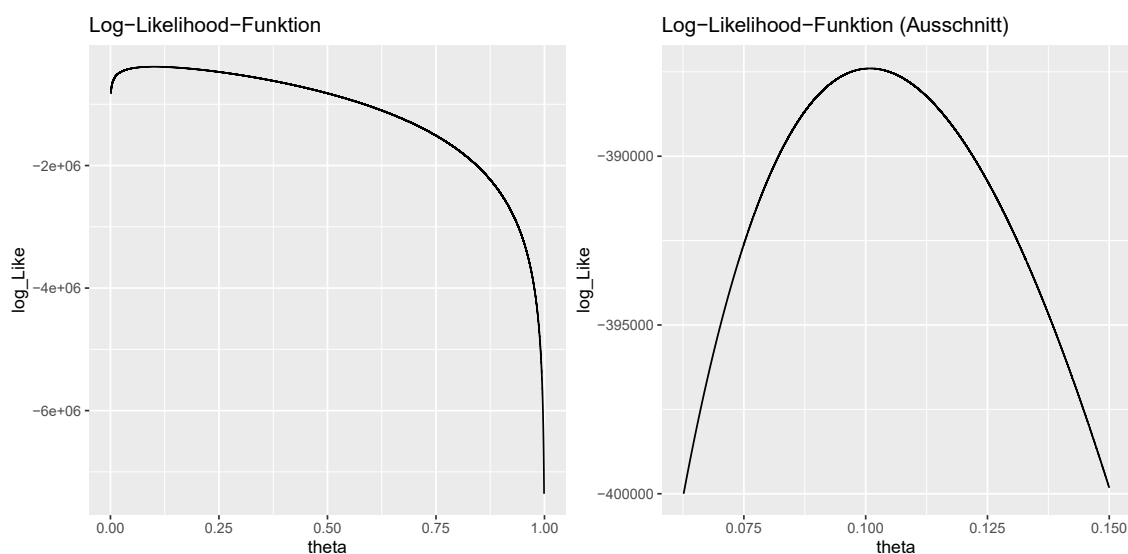


Abbildung 5.2.: Links: Graph der logarithmierten, marginalen und bedingten Likelihood-Funktion zur Schätzung der Schließungswahrscheinlichkeit deutscher Unternehmen auf dem gesamten Parameterraum  $\Theta = [0,001; 0,999]$ ; Rechts: Ausschnitt der Funktion auf dem Intervall  $[0,0625; 0,15]$ . Das Maximum ist bei ungefähr 0,1 zu erkennen.

Analog dazu ergibt sich die Anzahl der zensierten Historien durch  $m_{cens} = 1.172.652$ . Zu berechnen bleibt somit die Summe der innerhalb der Studie beobachteten Lebensdauern von bis einschließlich 2019 geschlossenen Unternehmen  $d^{obs}$ . Dabei wurden Unternehmen mit einer Schließung im Jahr 2018 ein Jahr beobachtet, wohingegen 2019 geschlossene Unternehmen zwei Jahre beobachtet wurden. Es ergibt sich:

$$d^{obs} = 1 \cdot 168.142 + 2 \cdot 107.050 = 382.242$$

Abschließend folgt für den Punktschätzer der Wahrscheinlichkeit für eine Schließung innerhalb eines Jahres:

$$\hat{\theta}_n = \frac{m_{uncens}}{d^{obs} + sm_{cens}} = \frac{275.192}{382.242 + 2 \cdot 1.172.652} \approx 0,10089362 \approx 0,1009$$

Deutsche Unternehmen haben folglich jedes Jahr eine ungefähr 10%-tige Wahrscheinlichkeit geschlossen zu werden. Somit resultiert für deutsche Unternehmen, die zwischen 2013 und 2017 gegründet wurden, eine geschätzte erwartete Lebensdauer von

$$\widehat{E_{\theta}\{X\}} = E_{\hat{\theta}_n}\{X\} = \frac{1}{\hat{\theta}_n} \approx \frac{1}{0,1009} \approx 9,91$$

Jahren.

Außerdem lassen sich mithilfe des Punktschätzers  $\hat{\theta}_n$  die Quantile der geometrischen

Verteilung berechnen. So gibt (vgl. (2.6))

$$F_G(x; 0, 1009) = 1 - (1 - 0, 1009)^x = 1 - 0, 8991^x$$

den Anteil der Unternehmen an, die innerhalb von  $x$  Jahren nach der Gründung wieder geschlossen werden. So werden beispielsweise 10, 09 Prozent der Unternehmen innerhalb des ersten Jahres geschlossen. Eine Übersicht über weitere Quantile ist in der folgenden Tabelle 5.2 gegeben:

Jahre nach Gründung	2	3	4	5	8	10	15
Anteil Schließungen in %	19, 16	27, 32	34, 65	41, 25	57, 30	65, 48	79, 72

Tabelle 5.2.: Auflistung einiger Quantile der geometrischen Verteilung mit dem Parameter  $\hat{\theta}_n \approx 0, 1009$  und somit der Anteile der Unternehmen, die innerhalb einer gewissen Anzahl an Jahren nach der Gründung geschlossen wurden. So schließen beispielsweise 19, 16% in den ersten zwei Jahren und 65, 48% in den ersten 10 Jahren.

Darüber hinaus resultiert ein Median der Lebensdauer [siehe Chattamvelli und Shanmugam, 2020, Kapitel 4.3] von

$$\text{ME}_{\hat{\theta}_n}\{X\} := \left\lceil \frac{-1}{\log_2(1 - \hat{\theta}_n)} \right\rceil = \left\lceil \frac{-1}{\log_2(0, 8991)} \right\rceil = [6, 52] = 7 \quad (5.1)$$

Jahren.

Im nächsten Schritt erfolgt eine Angabe der Standardfehler von  $\hat{\theta}_n$ . Der Standardfehler  $\text{SE}_3$  (vgl. (3.75)), der mithilfe eines Martingalgrenzwertsatzes (vgl. Satz 9) hergeleitet wurde, kann dabei direkt berechnet werden:

$$\begin{aligned} \text{SE}_3 &= \sqrt{\frac{m_{uncens}(d^{obs} + sm_{cens}) - m_{uncens}^2}{(d^{obs} + sm_{cens})^3}} \\ &= \sqrt{\frac{275.192 \cdot (382.242 + 2 \cdot 1.172.652) - 275.192^2}{(382.242 + 2 \cdot 1.172.652)^3}} \approx 1, 8237 \cdot 10^{-4} \end{aligned}$$

Zur Berechnung der anderen Standardfehler  $\text{SE}_1$  (vgl. 3.46),  $\text{SE}_2$  (vgl. 3.47) und  $\text{SE}_4$  (vgl. A.8) erfolgte eine entsprechende Programmierung in  $R$  (siehe Kapitel A.4.1). Die Ergebnisse sind in der Tabelle 5.3 aufgelistet.

Zusammenfassend liefern alle Standardfehler  $\text{SE}_1$  bis  $\text{SE}_4$  ähnliche Ergebnisse. In jedem Fall beläuft sich, gerundet auf zwei Nachkommastellen, das 95%-Konfidenzintervall auf [9, 88; 9, 95]. Insgesamt liefern die Standardfehler  $\text{SE}_1$  (vgl. (3.46)) und  $\text{SE}_4$  (vgl. (A.8)), die auf der Annahme, die Alter ein Jahr vor dem Studieneintritt  $T_i$  mit  $i \in \{1, \dots, n\}$  seien unabhängig und identisch diskret gleichverteilt auf dem Träger  $\{0, \dots, G-1\}$ , basieren, die schmalsten Konfidenzintervalle. Insbesondere ergeben  $\text{SE}_1$  und  $\text{SE}_4$  exakt denselben Wert. Außerdem liefert der Standardfehler im semiparametrischen Modell

	SE <sub>1</sub>	SE <sub>2</sub>	SE <sub>3</sub>	SE <sub>4</sub>
	$1,8164 \cdot 10^{-4}$	$1,84 \cdot 10^{-4}$	$1,8237 \cdot 10^{-4}$	$1,8164 \cdot 10^{-4}$
KI $\theta_0$	[0, 1005; 0, 1012]	[0, 1005; 0, 1013]	[0, 1005; 0, 1013]	[0, 1005; 0, 1012]
KI $E_{\theta}\{X\}$	[9, 8766; 9, 9465]	[9, 8761; 9, 9470]	[9, 8764; 9, 9467]	[9, 8766; 9, 9465]

Tabelle 5.3.: Tabellarische Übersicht der vier verschiedenen Berechnungsansätze für den Standardfehler inklusive der daraus resultierenden 95%-Prozent Konfidenzintervalle für die wahre Schließungswahrscheinlichkeit  $\theta_0$  und die erwartete Lebensdauer  $E_{\theta}\{X\}$  deutscher Unternehmen.

SE<sub>3</sub> (vgl. (3.75)), der mithilfe des Martingalgrenzwertsatzes von Rebolledo [1979] [siehe Kalbfleisch und Prentice, 2002, Theorem 5.1] hergeleitet wurde, ein etwas schmaleres Konfidenzintervall als der mithilfe der M-Schätzer-Theorie nach van der Vaart [1998] hergeleitete Schätzer SE<sub>2</sub> (vgl. (3.47)).

### Punktschätzung bei doppelt trunkiertem Beobachtungsdesign

Ein weiterer möglicher Ansatz zur Berechnung einer Punktschätzung der Ausfallwahrscheinlichkeit  $\theta_0$  wäre, ein beidseitig trunkiertes Beobachtungsdesign anzunehmen. Anschließend kann auf der Grundlage der doppelt trunkierten Daten die Nullstelle der abgeleiteten logarithmierten Likelihood-Funktion (4.24) berechnet werden.

Dazu werden zum Erhalt eines beidseitig trunkierten Beobachtungsdesigns die Überlebensdaten aus Tabelle 5.1 verwendet und dabei alle Unternehmen mit einer rechtsseitig zensurierten Historie entfernt. Die Nullstelle der abgeleiteten logarithmierten Likelihood-Funktion (4.24) wird somit basierend auf den in Tabelle 5.4 angegebenen Daten berechnet.

Gründung	Schließung		Summe
	2018	2019	
2013	37.016	17.295	54.311
2014	42.272	20.305	62.577
2015	35.588	23.353	58.941
2016	34.579	27.464	62.043
2017	18.687	18.633	37.320
Summe	168.142	107.050	275.192

Tabelle 5.4.: Tabellarische Übersicht der Gründungs- und Schließungsjahre von insgesamt 275.192 deutschen Unternehmen bei beidseitig trunkiertem Beobachtungsdesign

Obwohl das in Kapitel 4.2 hergeleitete Modell und die damit einhergehende Likelihood-Funktion auf der Grundlage von Simulationsdaten plausible Ergebnisse liefert (siehe Kapitel 6), ergibt sich basierend auf den in den Tabelle 5.4 gegebenen Daten eine

Punktschätzung der Wahrscheinlichkeit, dass ein Unternehmen binnen eines Jahres geschlossen wird, von

$$\hat{\theta}_n = 8 \cdot 10^{-6}$$

und somit ein geschätzter Stichprobenumfang von (vgl. (4.1) und (4.23)):

$$\hat{n} = \frac{m}{\alpha_{\hat{\theta}_n}^{DT}} \approx \frac{275.192}{1,6 \cdot 10^{-5}} \approx 17.199.844.686$$

Da diese Ergebnisse unplausibel sind, wurde dieser Ansatz in Bezug auf die Anwendung auf das AFiD-Panel als unbrauchbar erachtet und somit nicht weiter verfolgt.

### 5.3. Vergleich mit anderen Quellen

Zur Einordnung der Resultate aus anderen Arbeiten sei zunächst angemerkt, dass eine einheitliche und präzise Definition des Ausfalls der Unternehmen gegeben sein muss. Beispielsweise sind Insolvenzen nicht gleichzusetzen mit Schließungen, da ein Unternehmen auch aus anderen Gründen geschlossen werden kann [siehe Cochran, 1981, Abbildung 1]. Anderenfalls gestaltet sich ein Vergleich von Unternehmensdemografien, insbesondere wenn die Daten aus unterschiedlichen Zeiten und Ländern stammen, als schwierig.

In Bezug auf Deutschland berechneten Brüderl et al. [1992], basierend auf 1.621 Unternehmen, für die Schließungsverteilung ein 25%-Quantil von 2 Jahren und ein 37%-Quantil von 5 Jahren. Dieses Ergebnis liegt relativ nahe bei dem in dieser Arbeit errechneten 19%-Quantil von 2 Jahren sowie dem 41%-Quantil von 5 Jahren.

Eine frühere, in Japan durchgeführte Studie [siehe Honjo, 2000] ermittelte, dass lediglich 4,8% der Unternehmen innerhalb von 8 Jahren geschlossen werden. Dieses Resultat impliziert eine wesentlich geringe Schließungswahrscheinlichkeit als die, die im Rahmen dieser Arbeit errechnet wurde, da 8 Jahre hier als das 57%-Quantil zu interpretieren sind. Allerdings gab derselbe Autor für eine spätere japanische Kohorte (2003 bis 2013) an, die Schließungsrate hätte sich verdoppelt [siehe Kato et al., 2022].

Für portugiesische Unternehmen wurde eine mediane Lebensdauer von 4 Jahren sowie ein 20%-Quantil von einem Jahr ermittelt [siehe Mata und Portugal, 1994]. Dies deutet im Vergleich zu den in Tabelle 5.2 angegebenen Quantilen auf eine schnellere Schließungsrate hin.

Das für Pakistan von Ullah et al. [2016] ermittelte 80%-Quantil von 4 Jahren weist im Vergleich zum vorliegenden 80%-Quantil von 15 Jahren auf eine deutlich schnellere Schließungsrate hin.

Für frühere Kohorten wurde in den USA ein Median von 5 Jahren bis zu Schließung angegeben [siehe Cochran, 1981]. Dies deutet auf eine schnellere Schließungsrate als die, die in dieser Arbeit errechnet wurde, hin, da hier eine mediane Lebenserwartung von 7 Jahren ermittelt wurde.

## 6. Simulation

In diesem Kapitel soll eine Schätzung des wahren Parameterwerts  $\theta_0$  der geometrischen Verteilung (vgl. Definition 1) sowie eine Berechnung der entsprechenden Standardfehler  $SE_1$  (vgl. (3.46)),  $SE_2$  (vgl. (3.47)),  $SE_3$  (vgl. (3.75)) sowie  $SE_4$  (vgl. (A.8)) unter der Berücksichtigung von linksseitiger Trunkierung sowie rechtsseitiger Zensur simuliert werden. Dazu wird der wahre Parameterwert  $\theta_0$  sowie der Umfang  $n$ , den die Stichprobe ohne Einschränkungen der Beobachtbarkeit hat, als bekannt vorausgesetzt. Außerdem werden die Dauer des Entstehungszeitraums  $G$  sowie die Länge des Beobachtungszeitraums  $s$  im Vorfeld festgelegt. Zur Berechnung der wahren Varianz des Punktschätzers  $\text{Var}_{\theta_0}\{\hat{\theta}_n\}$  wird auf Formel (A.7) zurückgegriffen.

**Beispiel 1.** Wird  $\theta_0 = 0,1$  sowie  $G = 5$ ,  $s = 2$  und  $n = 50$  festgelegt, ergibt sich:

$$\begin{aligned} \text{Var}_{\theta_0}\{\hat{\theta}_n\} &= \sqrt{\frac{G\theta_0(1-\theta_0)}{n \sum_{x=1}^{s+G-1} \left[ \min(x, s, G, s+G-x) \left(1 - \mathbb{1}_{\{2 \leq x\}} \sum_{k=1}^{x-1} \theta_0(1-\theta_0)^{k-1}\right) \right]}} \\ &= \sqrt{\frac{5 \cdot 0,1 \cdot 0,9}{50 \sum_{x=1}^6 \left[ \min(x; 2; 7-x) \left(1 - \mathbb{1}_{\{2 \leq x\}} \sum_{k=1}^{x-1} 0,1 \cdot 0,9^{k-1}\right) \right]}} \approx 0,034 \end{aligned}$$

Anschließend werden mithilfe eines Computerprogramms die Realisierungen der  $n$  unabhängigen und identisch verteilten, bivariaten Zufallsvariablen

$$(X_i^{sim}, T_i^{sim})', \quad i \in \{1, \dots, n\}$$

gezogen. Dabei folgen für alle  $i \in \{1, \dots, n\}$  die  $X_i^{sim}$  einer geometrischen Verteilung mit dem Parameter  $\theta_0$ , wohingegen alle  $T_i^{sim}$  diskret gleichverteilt auf dem Träger  $\{0, \dots, G-1\}$  sind.

Anschließend werden, um die Linkstrunkierung zu berücksichtigen, aus der Stichprobe alle Realisierungen entfernt, für die  $\mathbf{x}_i^{sim} \leq \mathbf{t}_i^{sim}$  erfüllt ist, sodass lediglich die  $m \leq n$  Realisierungen

$$(\tilde{\mathbf{x}}_j^{sim}, \tilde{\mathbf{t}}_j^{sim})', \quad j \in \{1, \dots, m\}$$

übrig bleiben. Damit die Historien zusätzlich rechtsseitig zensiert sind, werden die Realisierungen  $\tilde{\mathbf{x}}_j^{sim}$  durch  $\bar{\mathbf{x}}_j^{sim}$  mit

$$\bar{\mathbf{x}}_j^{sim} := \min(\tilde{\mathbf{x}}_j^{sim}, \tilde{\mathbf{t}}_j^{sim} + s)$$

ersetzt. Um einsehen zu können, ob eine Zensur vorliegt, werden die Tupel  $(\bar{\mathbf{x}}_j^{sim}, \tilde{\mathbf{t}}_j^{sim})'$

abschließend zu den Tripeln  $(\bar{\mathbf{x}}_j^{sim}, \tilde{\mathbf{t}}_j^{sim}, \tilde{\mathbf{c}}_j^{sim})'$  erweitert. Dabei bezeichnet  $\tilde{\mathbf{c}}_j^{sim}$  die Realisierung der Zufallsvariablen:

$$\tilde{C}_j^{sim} := \mathbb{1}_{\{\tilde{X}_j^{sim} \leq \tilde{T}_j^{sim} + s\}}$$

Anhand der auf diese Weise generierten Daten können zunächst der Punktschätzer  $\hat{\theta}_n$  (vgl. (3.11)) und anschließend die verschiedenen Standardfehler  $SE_1$  (vgl. (3.46)),  $SE_2$  (vgl. (3.47)),  $SE_3$  (vgl. (3.74)) und  $SE_4$  (vgl. (A.8)) berechnet werden. Abschließend wird der quadratische Fehler ermittelt:

$$e^2 := \begin{pmatrix} e_\theta^2 \\ e_1^2 \\ e_2^2 \\ e_3^2 \\ e_4^2 \end{pmatrix} = \begin{pmatrix} (\hat{\theta}_n - \theta_0)^2 \\ (SE_1 - \sqrt{\text{Var}_{\theta_0}\{\hat{\theta}_n\}})^2 \\ (SE_2 - \sqrt{\text{Var}_{\theta_0}\{\hat{\theta}_n\}})^2 \\ (SE_3 - \sqrt{\text{Var}_{\theta_0}\{\hat{\theta}_n\}})^2 \\ (SE_4 - \sqrt{\text{Var}_{\theta_0}\{\hat{\theta}_n\}})^2 \end{pmatrix}$$

Diese Prozedur, bestehend aus der Generierung der Daten, anschließender Schätzung von  $\theta_0$  und Berechnung des Standardfehlers sowie abschließender Berechnung der quadratischen Fehler wird insgesamt  $W$  mal wiederholt. Die quadratischen Fehler werden dabei aufsummiert und durch  $W$  dividiert. Das Ergebnis ist der mittlere quadratische Fehler:

$$\text{MSE} := \begin{pmatrix} \text{MSE}_\theta \\ \text{MSE}_1 \\ \text{MSE}_2 \\ \text{MSE}_3 \\ \text{MSE}_4 \end{pmatrix} = \frac{1}{W} \sum_{l=1}^W \begin{pmatrix} e_{\theta,l}^2 \\ e_{1,l}^2 \\ e_{2,l}^2 \\ e_{3,l}^2 \\ e_{4,l}^2 \end{pmatrix} = \frac{1}{W} \sum_{l=1}^W \begin{pmatrix} (\hat{\theta}_{n,l} - \theta_0)^2 \\ (SE_{1,l} - \sqrt{\text{Var}_{\theta_0}\{\hat{\theta}_n\}})^2 \\ (SE_{2,l} - \sqrt{\text{Var}_{\theta_0}\{\hat{\theta}_n\}})^2 \\ (SE_{3,l} - \sqrt{\text{Var}_{\theta_0}\{\hat{\theta}_n\}})^2 \\ (SE_{4,l} - \sqrt{\text{Var}_{\theta_0}\{\hat{\theta}_n\}})^2 \end{pmatrix}$$

Das Subskript  $l \in \{1, \dots, W\}$  bezeichnet dabei das entsprechende Ergebnis in der  $l$ -ten Wiederholung des Experiments. Der zugehörige Quelltext zur Simulation kann Kapitel A.4.2 entnommen werden.

Zunächst wird, analog zu Beispiel 1,  $G = 5$ ,  $s = 2$  und  $\theta_0 = 0, 1$  sowie  $W = 1000$  gesetzt. Für dieses Szenario sind die mittleren quadratischen Fehler für unterschiedlich große  $n$  in der Tabelle 6.1 aufgeführt.

Anschließend erfolgt eine Darstellung der verschiedenen mittleren quadratischen Fehler in Abhängigkeit von der Anzahl der wiederholten Experimente. Dazu sei im folgenden erneut  $G = 5$ ,  $s = 2$  und  $\theta_0 = 0, 1$ . Außerdem sei  $n = 10.000$  gegeben und  $W$  variabel. Die Ergebnisse sind in der Tabelle 6.2 aufgelistet.

Abschließend erfolgen die Berechnungen der mittleren quadratischen Fehler bei variablem  $G$  beziehungsweise variablem  $s$ . Die entsprechenden Ergebnisse werden in Tabelle 6.4 dargestellt.

6. Simulation

$G = 5, s = 2, \theta_0 = 0, 1, W = 1000$					
n	100	1.000	10.000	100.000	1.000.000
$MSE_\theta$	$6,15 \cdot 10^{-4}$	$6,23 \cdot 10^{-5}$	$5,18 \cdot 10^{-6}$	$5,52 \cdot 10^{-7}$	$6,11 \cdot 10^{-8}$
$MSE_1$	$8,46 \cdot 10^{-6}$	$8,2 \cdot 10^{-8}$	$6,94 \cdot 10^{-10}$	$7,36 \cdot 10^{-12}$	$8,13 \cdot 10^{-14}$
$MSE_2$	$8,98 \cdot 10^{-6}$	$8,37 \cdot 10^{-8}$	$7,17 \cdot 10^{-10}$	$7,61 \cdot 10^{-12}$	$8,23 \cdot 10^{-14}$
$MSE_3$	$8,57 \cdot 10^{-6}$	$8,2 \cdot 10^{-8}$	$6,96 \cdot 10^{-10}$	$7,39 \cdot 10^{-12}$	$8,11 \cdot 10^{-14}$
$MSE_4$	$8,46 \cdot 10^{-6}$	$8,2 \cdot 10^{-8}$	$6,94 \cdot 10^{-10}$	$7,36 \cdot 10^{-12}$	$8,13 \cdot 10^{-14}$

Tabelle 6.1.: Tabellarische Übersicht der mittleren quadratischen Fehler bei  $G = 5$ ,  $s = 2$ ,  $\theta_0 = 0, 1$ ,  $W = 1000$  und variablem  $n$

$G = 5, s = 2, \theta_0 = 0, 1, n = 10.000$					
W	10	100	1.000	10.000	100.000
$MSE_\theta$	$6,36 \cdot 10^{-6}$	$6,01 \cdot 10^{-6}$	$5,86 \cdot 10^{-6}$	$5,82 \cdot 10^{-6}$	$5,74 \cdot 10^{-6}$
$MSE_1$	$7,89 \cdot 10^{-10}$	$7,99 \cdot 10^{-10}$	$7,88 \cdot 10^{-10}$	$7,78 \cdot 10^{-10}$	$7,68 \cdot 10^{-10}$
$MSE_2$	$8,12 \cdot 10^{-10}$	$8,17 \cdot 10^{-10}$	$8,06 \cdot 10^{-10}$	$7,95 \cdot 10^{-10}$	$7,9 \cdot 10^{-10}$
$MSE_3$	$8,93 \cdot 10^{-10}$	$7,99 \cdot 10^{-10}$	$7,89 \cdot 10^{-10}$	$7,79 \cdot 10^{-10}$	$7,7 \cdot 10^{-10}$
$MSE_4$	$7,89 \cdot 10^{-10}$	$7,99 \cdot 10^{-10}$	$7,88 \cdot 10^{-10}$	$7,78 \cdot 10^{-10}$	$7,68 \cdot 10^{-10}$

Tabelle 6.2.: Tabellarische Übersicht der mittleren quadratischen Fehler bei  $G = 5$ ,  $s = 2$ ,  $\theta_0 = 0, 1$ ,  $n = 10.000$  und variablem  $W$

$s = 2, \theta_0 = 0, 1, n = 10.000, W = 1.000$					
G	1	2	5	10	15
$MSE_\theta$	$4,69 \cdot 10^{-6}$	$5,06 \cdot 10^{-6}$	$5,54 \cdot 10^{-6}$	$7,76 \cdot 10^{-6}$	$1,01 \cdot 10^{-5}$
$MSE_1$	$4,91 \cdot 10^{-10}$	$5,69 \cdot 10^{-10}$	$7,46 \cdot 10^{-10}$	$1,33 \cdot 10^{-9}$	$2,18 \cdot 10^{-9}$
$MSE_2$	$5,06 \cdot 10^{-10}$	$5,81 \cdot 10^{-10}$	$7,65 \cdot 10^{-10}$	$1,37 \cdot 10^{-9}$	$2,23 \cdot 10^{-9}$
$MSE_3$	$4,93 \cdot 10^{-10}$	$5,69 \cdot 10^{-10}$	$7,47 \cdot 10^{-10}$	$1,34 \cdot 10^{-9}$	$2,19 \cdot 10^{-9}$
$MSE_4$	$4,91 \cdot 10^{-10}$	$5,69 \cdot 10^{-10}$	$7,46 \cdot 10^{-10}$	$1,33 \cdot 10^{-9}$	$2,18 \cdot 10^{-9}$

Tabelle 6.3.: Übersicht der mittleren quadratischen Fehler bei  $s = 2$ ,  $\theta_0 = 0, 1$ ,  $n = 10.000$ ,  $W = 1.000$  und variablem  $G$

$G = 5, \theta_0 = 0, 1, n = 10.000, W = 1.000$					
s	1	2	3	5	10
$MSE_\theta$	$1,11 \cdot 10^{-5}$	$5,68 \cdot 10^{-6}$	$4,26 \cdot 10^{-6}$	$2,61 \cdot 10^{-6}$	$1,62 \cdot 10^{-6}$
$MSE_1$	$2,48 \cdot 10^{-9}$	$7,64 \cdot 10^{-10}$	$4,58 \cdot 10^{-10}$	$2,14 \cdot 10^{-10}$	$1,23 \cdot 10^{-10}$
$MSE_2$	$2,48 \cdot 10^{-9}$	$8 \cdot 10^{-10}$	$4,92 \cdot 10^{-10}$	$2,52 \cdot 10^{-10}$	$1,7 \cdot 10^{-10}$
$MSE_3$	$2,48 \cdot 10^{-9}$	$7,7 \cdot 10^{-10}$	$4,63 \cdot 10^{-10}$	$2,17 \cdot 10^{-10}$	$1,26 \cdot 10^{-10}$
$MSE_4$	$2,48 \cdot 10^{-9}$	$7,64 \cdot 10^{-10}$	$4,58 \cdot 10^{-10}$	$2,14 \cdot 10^{-10}$	$1,23 \cdot 10^{-10}$

Tabelle 6.4.: Übersicht der mittleren quadratischen Fehler bei  $G = 5$ ,  $\theta_0 = 0, 1$ ,  $n = 10.000$ ,  $W = 1.000$  und variablem  $s$

Zusammenfassend lässt sich in allen Simulationen erkennen, dass die parametrischen Standardfehler  $SE_1$  sowie  $SE_4$  etwas kleinere mittlere quadratische Fehler aufweisen als der semiparametrische Standardfehler  $SE_3$ . Des Weiteren sind die mittleren quadratischen Fehler insgesamt für  $SE_3$  am größten. Diese Unterschiede sind allerdings sehr gering. Aus Tabelle 6.1 lässt sich zudem entnehmen, dass der MSE für den Schätzer des geometrischen Parameters sowie für alle Standardfehler für  $n \rightarrow \infty$  gegen 0 konvergiert.

Mit zunehmender Anzahl der Wiederholungen der Experimente  $W \rightarrow \infty$  lässt sich hingegen keine merkliche Änderung der mittleren quadratischen Fehler festmachen (vgl. Tabelle 6.2).

Darüber hinaus zeigt Tabelle 6.3, dass bei längerem Entstehungszeitraum und somit steigendem  $G$  die mittleren quadratischen Fehler größer werden, wenn  $s = 2$ ,  $\theta_0 = 0, 1$ ,  $n = 10.000$  und  $W = 1.000$  gilt.

Abschließend zeigt Tabelle 6.4, dass bei längerem Beobachtungszeitraum und folglich steigendem  $s$  die mittleren quadratischen Fehler kleiner werden. Dies liegt daran, dass bei steigendem  $s$  auch mehr Individuen beobachtet werden können und somit auch die Schätzungen genauer werden.

## Simulierte Schätzung des geometrischen Parameters bei doppelt trunziertem Design

Im Folgenden soll ergänzend zu den bisherigen Simulationen die in Kapitel 4.2 hergeleitete Methode zur Schätzung des geometrischen Parameters bei doppelt trunziertem Beobachtungsdesign evaluiert werden. Dabei werden erneut die Dauer des Entstehungszeitraums  $G$ , die Länge des Beobachtungszeitraums  $s$  sowie der wahre Parameterwert der geometrischen Verteilung  $\theta_0$  mit  $G = 5$ ,  $s = 2$  und  $\theta_0 = 0, 1$  als bekannt vorausgesetzt.

Anschließend werden erneut mithilfe eines Computerprogramms die  $n$  Realisierungen der unabhängigen und identisch verteilten, bivariaten Zufallsvariablen

$$(X_i^{sim}, T_i^{sim})', i \in \{1, \dots, n\}$$

gezogen. Dabei folgen für alle  $i \in \{1, \dots, n\}$  die  $X_i^{sim}$  einer geometrischen Verteilung mit dem Parameter  $\theta_0$ , wohingegen alle  $T_i^{sim}$  diskret gleichverteilt auf dem Träger  $\{0, \dots, G-1\}$  sind.

Im nächsten Schritt werden alle Realisierungen mit  $\mathbf{x}_i^{sim} \leq \mathbf{t}_i^{sim}$  oder  $\mathbf{t}_i^{sim} + s + 1 \leq \mathbf{x}_i^{sim}$  aus dem simulierten Datensatz entfernt. Das Ergebnis sind die Beobachtungen:

$$(\tilde{\mathbf{x}}_j^{sim}, \tilde{\mathbf{t}}_j^{sim})', j \in \{1, \dots, m\}$$

Auf der Grundlage der auf diese Weise generierten Daten kann anschließend approximativ die Nullstelle der abgeleiteten, logarithmierten Likelihood-Funktion (4.24) berechnet werden. Mithilfe des daraus resultierenden Punktschätzers  $\hat{\theta}_n^{DT}$  kann anschließend der quadratische Fehler

$$e_{DT}^2 := (\hat{\theta}_n^{DT} - \theta_0)^2$$

ermittelt werden. Dieses Experiment wird insgesamt  $W$  mal wiederholt und anschlie-

ßend der mittlere quadratische Fehler

$$\text{MSE}_\theta^{DT} := \frac{1}{W} \sum_{l=1}^W e_{DT,l}^2 = \frac{1}{W} \sum_{l=1}^W (\hat{\theta}_{n,l}^{DT} - \theta_0)^2$$

berechnet. Das Subskript  $l \in \{1, \dots, W\}$  bezeichnet dabei das entsprechende Ergebnis in der der  $l$ -ten Wiederholung des Experiments.

Für  $W = 1000$  Wiederholungen, einen Entstehungszeitraum von  $G = 5$ , einen Beobachtungszeitraum von  $s = 2$  Jahren sowie einen wahren Wert des geometrischen Parameters von  $\theta_0 = 0,1$  sind die mittleren quadratischen Fehler bei variablem  $n$  in der folgenden Tabelle 6.5 zusammengefasst.

$G = 5, s = 2, \theta_0 = 0,1, W = 1.000$					
n	100	1.000	10.000	100.000	1.000.000
MSE $_\theta^{DT}$	0,013	$2,26 \cdot 10^{-3}$	$2,33 \cdot 10^{-4}$	$2,37 \cdot 10^{-5}$	$2,37 \cdot 10^{-6}$

Tabelle 6.5.: Übersicht der mittleren quadratischen Fehler für  $G = 5, s = 2, \theta_0 = 0,1, W = 1.000$  und verschiedenen Stichprobenumfängen  $n$  bei doppelt trunziertem Beobachtungsdesign

Der Tabelle 6.5 kann entnommen werden, dass der mittlere quadratische Fehler des geometrischen Parameters für  $n \rightarrow \infty$  gegen 0 konvergiert. Im Vergleich zu zum mittleren quadratischen Fehler bei linkstrunziertem und rechtszensiertem Design MSE $_\theta$  (vgl. Tabelle 6.1) ist MSE $_\theta^{DT}$  allerdings für alle untersuchten Stichprobenumfänge etwas größer. Dies ist darauf zurückzuführen, dass das beidseitige Trunkieren zu einem größeren Informationsverlust führt als das linksseitige Trunkieren und rechtsseitige Zensieren.

## 7. Fazit

Häufig wird in der Statistik untersucht, inwieweit ein Merkmal  $Y$  ein anderes Merkmal  $X$  beeinflusst. Dabei werden in der Regel die Auswirkungen von  $Y$  auf den Erwartungswert von  $X$  untersucht. Im einfachsten Fall ist  $Y$  dabei dichotom, sodass lediglich zwei Erwartungswerte miteinander verglichen werden müssen. In der vorliegenden Arbeit beschreibt  $X$  dabei die Lebensdauer eines Individuums. Der Mehrwert der vorliegenden Arbeit besteht, vor einer Aufnahme von  $Y$  in die Analyse, in der Schätzung eines unbedingten Erwartungswertes von  $X$  bei einem linkstrunkierten und rechtszensierten Beobachtungsdesign. Wird das Modell um  $Y$  ergänzt, ergeben sich mehrere, nach den Ausprägungen von  $Y$  stratifizierte Erwartungswerte  $E\{X\}$ . In einer verlässlichen Datenanalyse muss das Mittel dieser stratifizierten Erwartungswerte dem in dieser Arbeit errechneten unbedingten Erwartungswert entsprechen.

Zur Schätzung des unbedingten Erwartungswertes wurde, um im Einklang mit dem in Kapitel 5 analysierten AFiD-Panel zu bleiben, ein zeitdiskretes Modell, das auf der Doob-Zerlegung basiert, erstellt. Als Vorbild dienten dabei einige zeitstetige Modelle [siehe Weißbach et al., 2024, Anhang A, sowie Andersen et al., 1993, Beispiel III.3.6]. Dabei bietet die Verwendung eines zeitdiskreten Modells im Vergleich zu einer zeitstetigen Variante [siehe beispielsweise Fleming und Harrington, 1991] einige Vorteile. So kann beispielsweise auf die Doob-Meyer-Zerlegung eines Zählprozesses in ein Martingal und einen Kompensator verzichtet und stattdessen auf die einfachere Doob-Zerlegung zurückgegriffen werden. Auch gelingt die Darstellung der marginalen, bedingten Likelihood-Funktion im Zeitstetigen nur mithilfe von Produktintegralen über die Zeit, die sich im Diskreten zu Produkten über die einzelnen Zeitpunkte vereinfachen (vgl. (2.75)). Ein wesentlicher Nachteil besteht jedoch darin, dass bei der zeitdiskreten Betrachtung einer Stichprobe durchaus der Fall eintreten kann, dass zwei Individuen zeitgleich ausfallen, sodass es sich bei der Menge der einzelnen Zählprozesse selbst um keinen multivariaten Zählprozess mehr handelt. Die Zählprozesssumme  $TN^c$  (vgl. (3.60)) ist dann selbst kein Zählprozess mehr, sondern nur noch allgemein ein Submartingal. Stattdessen wurde gefordert, dass sich die Ausfälle unabhängig voneinander ereignen (vgl. Bemerkung 15). Diese Annahme ist in der Anwendung auf das AFiD-Panel allerdings kritisch zu betrachten, da sich die Lebensdauern der Unternehmen teilweise stark überschneiden.

Für das theoretische Modell wurde zunächst eine marginale, bedingte Likelihood-Funktion hergeleitet, deren Maximum dann den Punktschätzer für den geometrischen Parameter darstellte. Davon ausgehend wurden sowohl mithilfe der M-Schätzer-Theorie als auch mithilfe eines Martingalgrenzwertsatzes entsprechende Standardfehler berechnet. Für beide Varianten wurden alle Voraussetzungen für die Konsistenz sowie asymptotische Normalität erfüllt. Auch konnte in einer Simulation gezeigt werden, dass mithilfe des vorgestellten Modells der geometrische Parameter sowie die Varianz des

Schätzers mit steigendem Stichprobenumfang zunehmend genauer approximiert werden können.

Eine alternative Schätzmethode des geometrischen Parameters mithilfe von Punktprozessen bei doppelt trunkiertem Beobachtungsdesign wurde ebenfalls vorgestellt. Eine durchgeführte Simulation lieferte zwar plausible Resultate, allerdings konnten von diesem Ansatz ausgehend im Rahmen der Anwendung auf das AFiD-Panel keine sinnvollen Ergebnisse erzielt werden.

Es sei außerdem angemerkt, dass im AFiD-Panel alle Unternehmensausfälle in Deutschland registriert wurden. Somit bilden die latenten Messungen keine einfache Zufallsstichprobe, sondern die Grundgesamtheit. In diesem Sinne liefert das Maximieren der Likelihood-Funktion den wahren Parameterwert  $\theta_0$ , eine Interpretation der Standardfehler ist allerdings schwierig, da die Zufälligkeit der Stichprobenziehung entfällt.

Eine weitere Anmerkung betrifft die im Rahmen der Trunkierung verwendete Marginalisierung der Verteilung von  $N_i$  für  $i \in \{1, \dots, n\}$ , die erst zum Zeitpunkt  $T_i$  startet. Da die marginale, bedingte Likelihood-Funktion (3.1) für jedes Individuum ein Produkt über die Zeitpunkte  $\{T_i + 1, T_i + 2, \dots, \chi\}$  darstellt, werden pro Individuum unterschiedlich viele Experimente berücksichtigt. Wären die Alter am Ende des Entstehungszeitraums  $T_i$  nicht zufällig, sondern wären verschiedene deterministische  $t_i$  gegeben, würde das Produkt in (3.1) nicht zu einer ordnungsgemäßen Wahrscheinlichkeitsfunktion führen. Das liegt daran, dass das dominierende Maß  $P_\theta^{\mathbb{T}}$  (vgl. (2.46)) von  $T_i$  abhängt beziehungsweise bei fester Trunkierung von  $t_i$  abhängen würde. Somit würden aus unterschiedlichen  $t_i$  auch unterschiedliche dominierende Maße resultieren. Für den Erhalt der Produktstruktur einer gemeinsamen Wahrscheinlichkeitsfunktion unabhängiger Zufallsvariablen scheinen jedoch übereinstimmende dominierende Maße der Randverteilungen ratsam. Zufällige Trunkierung behebt dabei diese Ungenauigkeit, auch wenn diese für die Analyse des vorliegenden AFiD-Panels kein optimales Design darstellt. Für zeitstetige Lebensdauern wurde dieses Argument in Weißbach et al. [2024, Kapitel A.1.1] ebenfalls gegeben.

Auch ist die Annahme einer geometrisch verteilten Lebensdauer nicht allgemein gültig und sollte, abhängig von den zu untersuchenden Daten entsprechend angepasst werden. Für die Verteilung der Gründungsjahre und somit von  $T_i, i \in \{1, \dots, n\}$ , musste glücklicherweise keine Verteilungsannahme getroffen werden. Jedoch kann das Wissen über eine parametrische Verteilung von  $T_i$  zu Verbesserungen in der Analyse führen. So sind die Standardfehler  $SE_1$  und  $SE_4$ , die unter der Annahme,  $T_i$  sei gleichverteilt, berechnet wurden, sowohl in der Anwendung als auch in der Simulation stets etwas kleiner als der semiparametrische Standardfehler  $SE_3$ . Eine andere Wahl der Verteilung von  $T_i$  ist im Rahmen des in dieser Arbeit vorgestellten Modells allerdings auch möglich.

Ebenso könnte das Modell um Kovariablen erweitert werden. Für deutsche Unternehmen ließe sich dann beispielsweise bestimmen, welche Auswirkung das Bundesland, in dem das Unternehmen sitzt, oder die Rechtsform des Unternehmens auf die durchschnittliche Lebenserwartung hat.

# Literatur

- [1] Andersen, P., Ø. Borgan, R. Gill und N. Keiding. 1988. Censoring, truncation and filtering in statistical models based on counting processes. *Statistical inference from stochastic processes*, ed. Prabhu, N.U., Volume 80, pages 19-60, Center for Mathematics and Computer Science, Amsterdam.
- [2] Andersen, P., Ø. Borgan, R. Gill und N. Keiding. 1993. *Statistical Models Based on Counting Processes*. Springer, New York.
- [3] Audretsch, D.B. und T. Mahmood. 1995. New firm survival: New results using a hazard function. *Review of Economics and Statistics* 77: 97-103 .
- [4] Bauer, H. 1974. *Wahrscheinlichkeitstheorie und Grundzüge der Maßtheorie*. De Gruyter, Berlin, Boston.
- [5] Billingsley, P. 2012. *Probability and Measure* (4th ed.). Wiley, New York.
- [6] Borgan, Ø. 1984. Maximum Likelihood Estimation in Parametric Counting Process Models, with Applications to Censored Failure Time Data. *Scandinavian Journal of Statistics* 11: 1–16 .
- [7] Brüderl, J., P. Preisendörfer, und R. Ziegler. 1992. Survival chances of newly founded business organizations. *American Sociological Review* 57: 227-242 .
- [8] Chattamvelli, R. und R. Shanmugam. 2020. *Discrete Distributions in Engineering and the Applied Sciences*. Springer Cham.
- [9] Chung, K. 2001. *A Course in Probability Theory* (3rd ed.). Academic Press, San Diego.
- [10] Cochran, A. 1981. Small business mortality rates: A review of the literature. *Journal of Small Business Economy* 19: 50-59 .
- [11] de Uña-Álvarez, J., A. Martínez-Senra, M. Otero-Giráldez, und M. Quintás. 2024. Cox regression with doubly truncated responses and time-dependent covariates: the impact of innovation on firm survival. *Journal of Applied Statistics* 51: 780-792 .
- [12] Doob, J. 1953. *Stochastic Processes*. John Wiley & Sons, New York.
- [13] Dörre, A. 2020. Bayesian estimation of a lifetime distribution under double truncation caused by time-restricted data collection. *Statistical Papers* 61: 945-965 .

- [14] Efron, B. und V. Petrosian. 1999. Nonparametric methods for doubly truncated data. *Journal of the American Statistical Association* 94: 824-834 .
- [15] Feller, W. 1968. *An Introduction to Probability Theory and Its Applications* (3rd ed.), Volume 1. Wiley, New York.
- [16] Fleming, T.R. und D. Harrington. 1991. *Counting Processes and Survival Analysis*. Wiley, Hoboken.
- [17] Gouriéroux, C. und A. Monfort. 1995. *Statistics and Econometric Models*, Volume 1. Cambridge University Press, Cambridge.
- [18] Gross, S.T. und C. Huber-Carol. 1992. Regression Models for Truncated Survival Data. *Scandinavian Journal of Statistics* 19: 193-213 .
- [19] Guo, G. 1993. Event-History Analysis for Left-Truncated Data. *Sociological Methodology* 23: 217-243 .
- [20] Heckman, J. 1976. The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models. *Annals of Economic and Social Measurement* 5: 475-492 .
- [21] Hernán, M., B. Sauer, S. Hernández-Díaz, R. Platt, und I. Shrier. 2016. Specifying a target trial prevents immortal time bias and other self-inflicted injuries in observational analyses. *Journal of Clinical Epidemiology* 79: 70-75 .
- [22] Honjo, Y. 2000. Business failure of new firms: An empirical analysis using a multiplicative hazards model. *International Journal of Industrial Organization* 18: 557-574 .
- [23] Johnson, N., A. Kemp, und S. Kotz. 2005. *Univariate Discrete Distributions* (3rd ed.). John Wiley & Sons, New York.
- [24] Kalbfleisch, J. und R. Prentice. 2002. *The Statistical Analysis of Failure Time Data* (2nd ed.). John Wiley & Sons, New York.
- [25] Kato, A., K. Onishi, und Y. Honjo. 2022. Does patenting always help new firm survival? understanding heterogeneity among exit routes. *Small Business Economy* 59: 449-475 .
- [26] Klenke, A. 2020. *Wahrscheinlichkeitstheorie* (4th ed.). Springer Spektrum Berlin, Heidelberg.
- [27] Mata, J. und P. Portugal. 1994. Life duration of new firms. *Journal of Industrial Economics* 42: 227-245 .
- [28] Newey, W.K. und D. McFadden. 1994. Large Sample Estimation and Hypothesis Testing. *Handbook of Econometrics* 4: 2111-2245, 2129

- [29] Pittiglio, R. 2023. Counterfeiting and firm survival. do international trade activities matter? *International Business Review* 32: 102145 .
- [30] Putter, H., M. Fiocco, und R. Geskus. 2006. Tutorial in biostatistics: Competing risks and multi-state models. *Statistics in Medicine* 26: 2389-2430 .
- [31] Rebolledo, R. 1979. La méthode des martingales appliquée à l'étude de la convergence en loi de processus. *Mémoires de la Société Mathématique de France* 62.
- [32] Reis, P.N. und M.G. Augusto. 2015. What is a firms life expectancy? empirical evidence in the context of portuguese companies. *Journal of Business Valuation and Economic Loss Analysis* 10: 45-75 .
- [33] Reiss, R.D. 1993. *A Course on Point Processes*. Springer, New York.
- [34] Rink, A. und I. Seiwert. 2021. Aktuelle Entwicklungen der Unternehmensdemografie. *Wirtschaft und Statistik* 2021/2: 41-58 .
- [35] Toparkus, A.M. und R. Weißbach. 2025. Testing truncation dependence: The Gumbel-Barnett copula. *Journal of Statistical Planning and Inference* 234: 106194 .
- [36] Ullah, M., N. Naimi, und R.B. Md Yusoff. 2016. Are small and medium enterprises (SMEs) in Lahore failing at the rate suggested in prior studies? an analysis of the degree of financial stress on small and medium enterprises and its impact on their life expectancy. *International Business Management* 10: 4258-4267 .
- [37] van der Vaart, A. 1998. *Asymptotic Statistics*. Cambridge University Press, Cambridge.
- [38] Weißbach, R., A. Dörre, D. Wied, G. Doblhammer, und A. Fink. 2024. Left-truncated health insurance claims data: Theoretical review and empirical application. *ASTA Advances in Statistical Analysis* 108: 31-68 .
- [39] Weißbach, R. und D. Wied. 2022. Truncating the exponential with a uniform distribution. *Statistical Papers* 63: 1247-1270 .
- [40] Yadav, K. und R. Lewis. 2021. Immortal time bias in observational studies. *Journal of the American Medical Association* 325: 686-687 .

# A. Appendix

## A.1. Elementares Beispiel zur Anwendung der Doob-Zerlegung

Betrachtet wird der Wahrscheinlichkeitsraum  $(\Omega, \mathcal{F} = \mathcal{P}(\Omega), P)$  mit der Ergebnismenge  $\Omega = \{\omega_1, \omega_2, \omega_3\}$ . Dabei beschreibt  $\omega \in \Omega$  einen Pfad über drei Zeitpunkte, wobei  $P\{\omega\} = 1/3, \forall \omega \in \Omega$ . Die genaue Darstellung der Pfade lässt sich der Abbildung A.1 entnehmen:

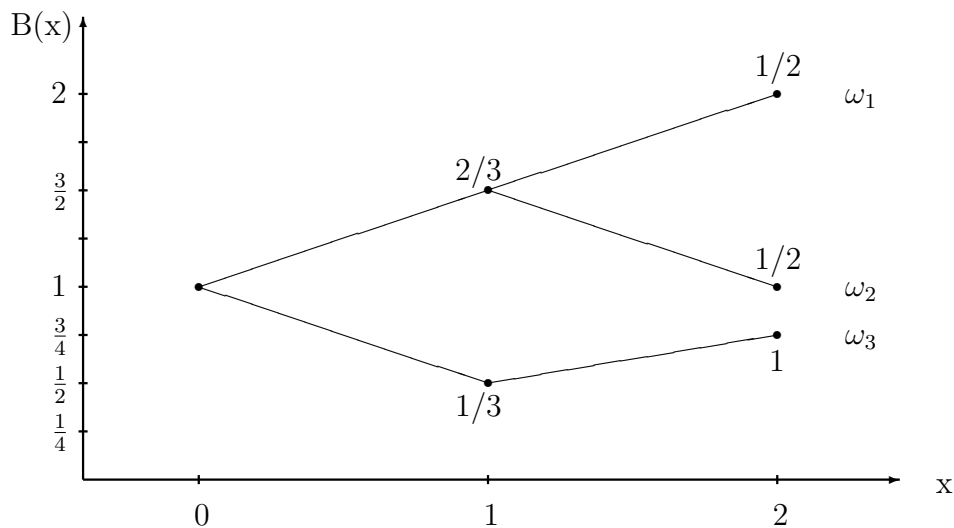


Abbildung A.1.: Darstellung aller möglichen Pfade des Prozesses  $B$

Betrachtet wird für die Zeitpunkte  $x \in \{0, 1, 2\}$  der Prozess  $B$  mit  $B(0, \omega) = 1$  für alle  $\omega \in \Omega$  sowie

$$P\{\{\omega : B(1, \omega) = 3/2\}\} = P\{\{\omega_1, \omega_2\}\} = P\{\omega_1 \cup \omega_2\} = \frac{2}{3}$$

und

$$P\{\{\omega : B(1, \omega) = 1/2\}\} = P\{\omega_3\} = \frac{1}{3}$$

Außerdem gilt

$$\begin{aligned} & P\left\{\{\omega : B(2, \omega) = 2\} \mid \{\omega : B(1, \omega) = 3/2\}\right\} \\ &= \frac{P\left\{\{\omega : B(2, \omega) = 2\} \cap \{\omega : B(1, \omega) = 3/2\}\right\}}{P\left\{\{\omega : B(1, \omega) = 3/2\}\right\}} = \frac{P\{\omega_1\}}{P\{\{\omega_1, \omega_2\}\}} = \frac{1/3}{2/3} = \frac{1}{2} \end{aligned}$$

sowie

$$\begin{aligned} & P\left\{\{\omega : B(2, \omega) = 1\} \mid \{\omega : B(1, \omega) = 3/2\}\right\} \\ &= \frac{P\left\{\{\omega : B(2, \omega) = 1\} \cap \{\omega : B(1, \omega) = 3/2\}\right\}}{P\left\{\{\omega : B(1, \omega) = 3/2\}\right\}} = \frac{P\{\omega_2\}}{P\{\{\omega_1, \omega_2\}\}} = \frac{1/3}{2/3} = \frac{1}{2} \end{aligned}$$

und

$$\begin{aligned} & P\left\{\{\omega : B(2, \omega) = 3/4\} \mid \{\omega : B(1, \omega) = 1/2\}\right\} \\ &= \frac{P\left\{\{\omega : B(2, \omega) = 3/4\} \cap \{\omega : B(1, \omega) = 1/2\}\right\}}{P\left\{\{\omega : B(1, \omega) = 1/2\}\right\}} = \frac{P\{\omega_3\}}{P\{\omega_3\}} = \frac{1/3}{1/3} = 1 \end{aligned}$$

Des Weiteren wird auf dem zugrundeliegenden Wahrscheinlichkeitsraum folgende Filtration  $\{\mathcal{F}_x : x \in \{0, 1, 2\}\}$  definiert:

$x = 0$  : Es gibt für alle  $\omega \in \Omega$  nur die Möglichkeit  $B(0, \omega) = 1$ , es folgt  $\mathcal{F}_0 = \{\emptyset, \Omega\}$ .

$x = 1$  : Es wird entschieden, ob der obere Ast, bestehend aus  $\omega_1$  und  $\omega_2$ , oder  $\omega_3$  gewählt wird. Es folgt  $\mathcal{F}_1 = \{\emptyset, \Omega, \{\omega_1, \omega_2\}, \omega_3\}$ .

$x = 2$  : Letztendlich wird entschieden, ob  $\omega_1$  oder  $\omega_2$  realisiert wird. Es ergibt sich die  $\sigma$ -Algebra  $\mathcal{F}_2 = \sigma\{\emptyset, \Omega, \{\omega_1, \omega_2\}, \omega_1, \omega_2, \omega_3\}$

Basierend auf Definition 4 lässt sich die vorhersagbare  $\sigma$ -Algebra  $\mathcal{F}^V$  folgendermaßen definieren:

$$\begin{aligned} \mathcal{F}^V = \sigma\{ & \{0\} \times \emptyset, \{0\} \times \Omega, \{1\} \times \emptyset, \{1\} \times \Omega, \{1, 2\} \times \emptyset, \{1, 2\} \times \Omega, \\ & \{2\} \times \emptyset, \{2\} \times \Omega, \{2\} \times \{\omega_1, \omega_2\}, \{2\} \times \{\omega_3\}\} \end{aligned}$$

Außerdem gilt

$$\begin{aligned} E\{B(1, \omega) \mid \mathcal{F}_0\} &= \frac{3}{2} \cdot P\left\{\{\omega : B(1, \omega) = 3/2\} \mid \mathcal{F}_0\right\} + \frac{1}{2} \cdot P\left\{\{\omega : B(1, \omega) = 1/2\} \mid \mathcal{F}_0\right\} \\ &= \frac{3}{2} \cdot \frac{2}{3} + \frac{1}{2} \cdot \frac{1}{3} = \frac{7}{6} > 1 = B(0) \end{aligned} \tag{A.1}$$

sowie

$$\begin{aligned}
 E\{B(2, \omega) | \mathcal{F}_1\} &= 2 \cdot P\{\{\omega : B(2, \omega) = 2\} | \mathcal{F}_1\} + 1 \cdot P\{\{\omega : B(2, \omega) = 1\} | \mathcal{F}_1\} \\
 &\quad + \frac{3}{4} \cdot P\{\{\omega : B(2, \omega) = 3/4\} | \mathcal{F}_1\} \\
 &= \mathbb{1}_{\{\omega : B(1, \omega) = 3/2\}}(\omega) \left( 2 \cdot P\{\{\omega : B(2, \omega) = 2\} | \{\omega : B(1, \omega) = 3/2\}\} \right. \\
 &\quad + 1 \cdot P\{\{\omega : B(2, \omega) = 1\} | \{\omega : B(1, \omega) = 3/2\}\} \\
 &\quad \left. + \frac{3}{4} \cdot P\{\{\omega : B(2, \omega) = 3/4\} | \{\omega : B(1, \omega) = 3/2\}\} \right) \\
 &+ \mathbb{1}_{\{\omega : B(1, \omega) = 1/2\}}(\omega) \left( 2 \cdot P\{\{\omega : B(2, \omega) = 2\} | \{\omega : B(1, \omega) = 1/2\}\} \right. \\
 &\quad + 1 \cdot P\{\{\omega : B(2, \omega) = 1\} | \{\omega : B(1, \omega) = 1/2\}\} \\
 &\quad \left. + \frac{3}{4} \cdot P\{\{\omega : B(2, \omega) = 3/4\} | \{\omega : B(1, \omega) = 1/2\}\} \right)
 \end{aligned}$$

Mithilfe der Definition bedingter Wahrscheinlichkeiten folgt weiter für alle  $\omega \in \Omega$ :

$$\begin{aligned}
 E\{B(2, \omega) | \mathcal{F}_1\} &= 2 \left( \frac{P\{\{\omega : B(2, \omega) = 2\} \cap \{\omega : B(1, \omega) = 3/2\}\}}{P\{\{\omega : B(1, \omega) = 3/2\}\}} \mathbb{1}_{\{\omega_1, \omega_2\}}(\omega) \right) \\
 &+ 1 \left( \frac{P\{\{\omega : B(2, \omega) = 1\} \cap \{\omega : B(1, \omega) = 3/2\}\}}{P\{\{\omega : B(1, \omega) = 3/2\}\}} \mathbb{1}_{\{\omega_1, \omega_2\}}(\omega) \right) \\
 &+ \frac{3}{4} \left( \frac{P\{\{\omega : B(2, \omega) = 3/4\} \cap \{\omega : B(1, \omega) = 1/2\}\}}{P\{\{\omega : B(1, \omega) = 1/2\}\}} \mathbb{1}_{\{\omega_3\}}(\omega) \right) \\
 &= \mathbb{1}_{\{\omega_1, \omega_2\}}(\omega) \left( 2 \frac{1/3}{2/3} + 1 \frac{1/3}{2/3} \right) + \mathbb{1}_{\{\omega_3\}}(\omega) \left( \frac{3}{4} \frac{1/3}{1/3} \right) \\
 &= \frac{3}{2} \cdot \mathbb{1}_{\{\omega_1, \omega_2\}}(\omega) + \frac{3}{4} \cdot \mathbb{1}_{\{\omega_3\}}(\omega) \geq B(1, \omega) \tag{A.2}
 \end{aligned}$$

Der Prozess  $B$  ist folglich ein Submartingal und kann in ein Martingal  $M$  und einen vorhersagbaren Kompensator  $A$  zerlegt werden.

Basierend auf (2.16) kann das Martingal für  $x \in \{0, 1, 2\}$  folgendermaßen berechnet werden:

$$M(x, \omega) := B(x, \omega) - A(x, \omega) = B(0, \omega) + \sum_{k=1}^x (B(k, \omega) - E\{B(k, \omega) | \mathcal{F}_{k-1}\})$$

Es lässt sich schnell einsehen, dass es sich bei  $M$  tatsächlich um ein  $\mathcal{F}_x$ -Martingal handelt, denn mithilfe einiger Rechenregeln für bedingte Erwartungswerte (vgl. Satz

13, (ii)-(iv)) folgt für  $x \in \{0, 1, 2\}$  und  $\kappa \in \mathbb{N}$  mit  $x + \kappa \leq 2$ :

$$\begin{aligned}
 \mathbb{E}\{M(x + \kappa, \omega) | \mathcal{F}_x\} &= \mathbb{E}\{B(0, \omega) | \mathcal{F}_x\} + \sum_{k=1}^{x+\kappa} \left( \mathbb{E}\{B(k, \omega) | \mathcal{F}_x\} - \mathbb{E}\{\mathbb{E}\{B(k, \omega) | \mathcal{F}_{k-1}\} | \mathcal{F}_x\} \right) \\
 &= B(0, \omega) + \sum_{k=1}^x \left( \mathbb{E}\{B(k, \omega) | \mathcal{F}_x\} - \mathbb{E}\{\mathbb{E}\{B(k, \omega) | \mathcal{F}_{k-1}\} | \mathcal{F}_x\} \right) \\
 &\quad + \sum_{k=x+1}^{x+\kappa} \left( \mathbb{E}\{B(k, \omega) | \mathcal{F}_x\} - \mathbb{E}\{\mathbb{E}\{B(k, \omega) | \mathcal{F}_{k-1}\} | \mathcal{F}_x\} \right) \\
 &= B(0, \omega) + \sum_{k=1}^x \left( B(k, \omega) - \mathbb{E}\{\mathbb{E}\{B(k, \omega) | \mathcal{F}_x\} | \mathcal{F}_{k-1}\} \right) \\
 &\quad + \sum_{k=x+1}^{x+\kappa} \left( \mathbb{E}\{B(k, \omega) | \mathcal{F}_x\} - \mathbb{E}\{B(k, \omega) | \mathcal{F}_x\} \right) \\
 &= B(0, \omega) + \sum_{k=1}^x \left( B(k, \omega) - \mathbb{E}\{B(k, \omega) | \mathcal{F}_{k-1}\} \right) = M(x, \omega)
 \end{aligned}$$

Abschließend soll die  $\mathcal{F}_x$ -Vorhersagbarkeit von  $A$  nachgewiesen werden. Wird  $A$  als eine Abbildung von  $(\Omega, \mathcal{F})$  nach  $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$  interpretiert, gilt unter der Verwendung von (A.1) und (A.2):

$$\begin{aligned}
 A(0, \omega) &= 0, \quad \text{für } \omega \in \{\omega_1, \omega_2, \omega_3\} \\
 A(1, \omega) &= \mathbb{E}\{B(1, \omega) | \mathcal{F}_0\} - B(0, \omega) = \frac{7}{6} - 1 = \frac{1}{6}, \quad \text{für } \omega \in \{\omega_1, \omega_2, \omega_3\} \\
 A(2, \omega) &= \frac{1}{6} + \mathbb{E}\{B(2, \omega) | \mathcal{F}_1\} - B(1, \omega) = \frac{1}{6} + \frac{3}{2} - \frac{3}{2} = \frac{1}{6} \quad \text{für } \omega \in \{\omega_1, \omega_2\} \\
 A(2, \omega) &= \frac{1}{6} + \mathbb{E}\{B(2, \omega) | \mathcal{F}_1\} - B(1, \omega) = \frac{1}{6} + \frac{3}{4} - \frac{1}{2} = \frac{5}{12} \quad \text{für } \omega = \omega_3
 \end{aligned}$$

Des Weiteren ist  $A$  messbar bezüglich der vorhersagbaren  $\sigma$ -Algebra  $\mathcal{F}^V$ , da für alle  $\mathbb{B} \in \mathcal{B}(\mathbb{R})$  gilt, dass  $A^{-1}(\mathbb{B}, \omega) \in \mathcal{F}^V$  ist. Beispielsweise gilt:

$$A^{-1}\left(\left[\frac{1}{4}, \frac{1}{2}\right]\right) = A^{-1}\left(\frac{5}{12}\right) = \{2\} \times \{\omega_3\} \in \mathcal{F}^V$$

Der Prozess  $A$  ist entsprechend  $\mathcal{F}_x$ -vorhersagbar. Im Gegensatz dazu gilt beispielsweise:

$$B^{-1}\left(\left[\frac{9}{10}, \frac{11}{10}\right]\right) = B^{-1}(1) = \{0\} \times \{\Omega\} \cup \{2\} \times \{\omega_2\} \notin \mathcal{F}^V$$

Der Prozess  $B$  ist somit nicht vorhersagbar.

## A.2. Berechnung des Standardfehlers per Martingalgrenzwertsatz im parametrischen Modell

In Kapitel 3.3.2 wurde mit (3.74) eine Formel für die Varianz von  $\hat{\theta}_n$  (vgl. (3.9)) hergeleitet. Darauf aufbauend soll im Folgenden eine explizite Darstellung des Standardfehlers unter der Annahme, dass alle  $T_i$  (vgl. (2.3)) unabhängig und diskret gleichverteilt auf dem Träger  $\{0, \dots, G-1\}$  (vgl. Definition 7) sind, ermittelt werden.

Dazu muss  $E_{\theta_0}\{TY_i^c(x-1)\}$  für alle  $x \in \{1, \dots, \chi\}$  berechnet werden, wobei  $TY_i^c$  in (2.71) definiert wurde. Es wird angenommen, dass  $\chi \geq (G-1) + s$  erfüllt ist.

Es sei angemerkt, dass sich zunächst für ein Individuum aufgrund der Unabhängigkeit von Entstehung und Ausfall (vgl. Annahme (A2))

$$E_{\theta_0}\{TY_i^c(x-1)\} = E_{\theta_0}\{\mathbb{1}_{\{T_i+1 \leq x \leq \min(X_i, T_i+s)\}}\} = E_{\theta_0}\{\mathbb{1}_{\{T_i+1 \leq x \leq T_i+s\}}\}E_{\theta_0}\{\mathbb{1}_{\{x \leq X_i\}}\} \quad (\text{A.3})$$

ergibt. Zur Berechnung des ersten Erwartungswertes  $E_{\theta_0}\{\mathbb{1}_{\{T_i+1 \leq x \leq T_i+s\}}\}$  gilt, dass  $T_i + 1 \leq x \leq T_i + s$  genau dann erfüllt ist, wenn

$$\textcircled{O} := \{x - s \leq T_i \leq x - 1\}$$

gilt, wobei  $T_i \in \{0, \dots, G-1\}$ .

Wir betrachten folgende Fälle:

- (i)  $1 \leq x < \min(s, G)$ : Dann tritt das Ereignis  $\textcircled{O}$  ein, wenn  $T_i \in \{0, \dots, x-1\}$  gilt. Es folgt:

$$E_{\theta_0}\{\mathbb{1}_{\{T_i+1 \leq x \leq T_i+s\}}\} = \frac{x}{G}$$

- (ii)  $\min(s, G) \leq x \leq \max(s, G)$ : Dieser Fall wird nochmals in zwei Fälle unterteilt:

- (a)  $G \leq s$ : Dann tritt das Ereignis  $\textcircled{O}$  ein, wenn  $T_i \in \{0, \dots, G-1\}$  gilt. Es folgt:

$$E_{\theta_0}\{\mathbb{1}_{\{T_i+1 \leq x \leq T_i+s\}}\} = \frac{G}{G} = 1$$

- (b)  $G > s$ : Dann tritt das Ereignis  $\textcircled{O}$  ein, wenn  $T_i \in \{x-s, \dots, x-1\}$  gilt. Es folgt:

$$E_{\theta_0}\{\mathbb{1}_{\{T_i+1 \leq x \leq T_i+s\}}\} = \frac{s}{G}$$

- (iii)  $\max(s, G) < x \leq s + G - 1$ : Dann tritt das Ereignis  $\textcircled{O}$  ein, wenn  $T_i \in \{x-s, \dots, G-1\}$  gilt. Es folgt:

$$E_{\theta_0}\{\mathbb{1}_{\{T_i+1 \leq x \leq T_i+s\}}\} = \frac{s + G - x}{G}$$

(iv)  $s + G \leq x$ : Dann kann das Ereignis  $\mathbb{O}$  nicht mehr eintreten. Es folgt:

$$E_{\theta_0} \{ \mathbb{1}_{\{T_{i+1} \leq x \leq T_i + s\}} \} = 0$$

Insgesamt ergibt sich aus den Fällen (i) – (iv) folgender Erwartungswert:

$$E_{\theta_0} \{ \mathbb{1}_{\{T_{i+1} \leq x \leq T_i + s\}} \} = \frac{\min(x, s, G, \max(s + G - x, 0))}{G} \quad (\text{A.4})$$

Für den zweiten Erwartungswert von (A.3) folgt unter Berücksichtigung der geometrischen Verteilung von  $X_i$ :

$$E_{\theta_0} \{ \mathbb{1}_{\{x \leq X_i\}} \} = \sum_{k=x}^{\infty} \theta_0 (1 - \theta_0)^{k-1} = 1 - \mathbb{1}_{\{2 \leq x\}} \sum_{k=1}^{x-1} \theta_0 (1 - \theta_0)^{k-1} \quad (\text{A.5})$$

Abschließend ergibt sich durch einsetzen von (A.4) und (A.5) in (A.3)

$$E_{\theta_0} \{ {}_T Y_i^c(x-1) \} = \frac{\min(x, s, G, \max(s + G - x, 0))}{G} \left( 1 - \mathbb{1}_{\{2 \leq x\}} \sum_{k=1}^{x-1} \theta_0 (1 - \theta_0)^{k-1} \right) \quad (\text{A.6})$$

Des Weiteren ergibt sich durch Einsetzen von (A.6) in (3.74) eine explizite Form der Varianz von  $\hat{\theta}_n$ :

$$\text{Var}_{\theta_0} \{ \hat{\theta}_n \} = \frac{G \theta_0 (1 - \theta_0)}{n \sum_{x=1}^{s+G-1} \left[ \min(x, s, G, s + G - x) \left( 1 - \mathbb{1}_{\{2 \leq x\}} \sum_{k=1}^{x-1} \theta_0 (1 - \theta_0)^{k-1} \right) \right]} \quad (\text{A.7})$$

Da der Stichprobenumfang  $n$  unbekannt ist, muss dieser mithilfe des Verhältnisses  $m = n \alpha_{\theta_0}$  ermittelt werden, wobei  $m$  die Anzahl der sichtbaren Individuen und  $\alpha_{\theta_0}$  die Wahrscheinlichkeit, dass ein Individuum beobachtet wird, darstellt. Die Beobachtungswahrscheinlichkeit  $\alpha_{\theta_0}$  ist dabei gegeben durch (vgl. 3.45)

$$\alpha_{\theta_0} = \frac{1}{G} \sum_{t=0}^{G-1} (1 - \theta_0)^t = \frac{1 - (1 - \theta_0)^G}{\theta_0 G}$$

sodass sich

$$\text{Var}_{\theta_0} \{ \hat{\theta}_n \} = \frac{(1 - (1 - \theta_0)^G)(1 - \theta_0)}{m \sum_{x=1}^{s+G-1} \left[ \min(x, s, G, s + G - x) \left( 1 - \mathbb{1}_{\{2 \leq x\}} \sum_{k=1}^{x-1} \theta_0 (1 - \theta_0)^{k-1} \right) \right]}$$

ergibt.

Abschließend wird der unbekannte wahre Parameterwert  $\theta_0$  durch den zugehörigen

Punktschätzer  $\hat{\theta}_n$  ersetzt, um einen Standardfehler zu erhalten:

$$SE_4 = \sqrt{\frac{(1 - (1 - \hat{\theta}_n)^G)(1 - \hat{\theta}_n)}{m \sum_{x=1}^{s+G-1} \left[ \min(x, s, G, s + G - x) \left(1 - \mathbb{1}_{\{2 \leq x\}} \sum_{k=1}^{x-1} \hat{\theta}_n (1 - \hat{\theta}_n)^{k-1}\right)\right]}} \quad (\text{A.8})$$

### A.3. Wichtige Definitionen, Eigenschaften und Sätze

Die Definitionen und Konzepte sowie daraus resultierende Eigenschaften und Sätze, die im Rahmen dieser Arbeit zwar nicht von zentraler Bedeutung sind, aber durchaus Verwendung finden, werden in diesem Kapitel aufgelistet.

#### A.3.1. Bedingte Erwartungswerte

Zunächst wird eine Definition des bedingten Erwartungswertes, basierend auf Fleming und Harrington [1991, Definition 1.2.7] angegeben:

**Definition 16** (Bedingter Erwartungswert). *Sei  $X$  eine Zufallsvariable auf dem Wahrscheinlichkeitsraum  $(\Omega, \mathcal{F}, P)$  und  $\mathcal{F}_1$  eine Unter- $\sigma$ -Algebra von  $\mathcal{F}$ . Sei  $E\{X|\mathcal{F}_1\}$  eine Zufallsvariable mit:*

- (i)  $E\{X|\mathcal{F}_1\}$  ist  $\mathcal{F}_1$ -messbar
- (ii)  $\int_{\mathbb{F}} X dP_\theta = \int_{\mathbb{F}} E\{X|\mathcal{F}_1\} dP_\theta, \forall \mathbb{F} \in \mathcal{F}_1$

Dann heißt  $E\{X|\mathcal{F}_1\}$  der bedingte Erwartungswert von  $X$  gegeben  $\mathcal{F}_1$ . Weiterhin gilt:  $E|X| < \infty \Rightarrow E\{X|\mathcal{F}_1\}$  existiert. Des Weiteren gilt:

$$\mathbb{F} \in \mathcal{F} \Rightarrow P\{\mathbb{F}|\mathcal{F}_1\} = E\{\mathbb{1}_{\{\mathbb{F}\}}|\mathcal{F}_1\} \quad (\text{A.9})$$

Der folgende Satz listet alle Rechenregeln für bedingte Erwartungswerte, die im Rahmen dieser Arbeit angewendet werden, auf [siehe Fleming und Harrington, 1991, Kapitel 1.2]

**Satz 13** (Eigenschaften bedingter Erwartungswerte). *Sei  $(\Omega, \mathcal{F}, P)$  ein Wahrscheinlichkeitsraum, auf dem die Zufallsvariablen  $X_1$  und  $X_2$  definiert sind. Außerdem seien  $\mathcal{F}_k$  sowie  $\mathcal{F}_x$  mit  $\mathcal{F}_k \subset \mathcal{F}_x$  Unter- $\sigma$ -Algebren von  $\mathcal{F}$ . Dann gilt:*

- (i)  $E\{E\{X_1|\mathcal{F}_x\}\} = E\{X_1\}$
- (ii) Turmregel:  $E\{E\{X_1|\mathcal{F}_k\}|\mathcal{F}_x\} = E\{E\{X_1|\mathcal{F}_x\}|\mathcal{F}_k\} = E\{X_1|\mathcal{F}_k\}$
- (iii) Sei  $X_2$  messbar bezüglich  $\mathcal{F}_x$ , also  $\sigma\{X_2\} \subset \mathcal{F}_x$ . Dann gilt

$$E\{X_1 X_2|\mathcal{F}_x\} = X_2 E\{X_1|\mathcal{F}_x\}$$

sowie  $E\{X_2|\mathcal{F}_x\} = X_2$

$$(iv) \ E\{c_1X_1 + c_2X_2|\mathcal{F}_x\} = c_1E\{X_1|\mathcal{F}_x\} + c_2E\{X_2|\mathcal{F}_x\}$$

### A.3.2. Stochastische Konvergenzbegriffe

Im Folgenden werden verschiedene Arten beschrieben, auf die eine Folge von Zufallsvariablen  $(X_n)_{n \in \mathbb{N}}$  gegen eine Zufallsvariable  $X$  konvergieren kann [siehe van der Vaart, 1998, Kapitel 2.1].

Die schwächste Form der stochastischen Konvergenz ist die so genannte Konvergenz in Verteilung oder auch schwache Konvergenz.

**Definition 17** (Konvergenz in Verteilung). *Sei  $X$  eine Zufallsvariable. Dann wird die Verteilungsfunktion  $F$  von  $X$  definiert durch die Abbildung*

$$F : x \mapsto P\{X \leq x\}$$

*Eine Folge von Zufallsvariablen  $(X_n)_n$  mit den entsprechenden Verteilungsfunktionen  $(F_n)_n$  konvergiert in Verteilung gegen  $X$ , wenn*

$$\lim_{n \rightarrow \infty} F_n(x) = F(x) \tag{A.10}$$

*für alle Stetigkeitsstellen von  $F$  erfüllt ist. Im Rahmen dieser Arbeit wird für Konvergenz in Verteilung die Notation  $X_n \rightsquigarrow X$  verwendet.*

Ein stärkerer stochastischer Konvergenzbegriff ist die Konvergenz in Wahrscheinlichkeit.

**Definition 18** (Konvergenz in Wahrscheinlichkeit). *Sei  $X$  eine Zufallsvariable und  $d : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}_+$  eine Metrik. Dann konvergiert eine Folge von Zufallsvariablen  $(X_n)_n$  in Wahrscheinlichkeit gegen  $X$ , wenn für alle  $\epsilon > 0$ :*

$$\lim_{n \rightarrow \infty} P[d(X_n, X) > \epsilon] = 0 \tag{A.11}$$

*Im Rahmen dieser Arbeit wird für Konvergenz in Verteilung die Notation  $X_n \xrightarrow{p} X$  verwendet.*

Alternativ kann für Funktionen eine stärkere Version der Konvergenz in Wahrscheinlichkeit, die so genannte Gleichmäßige Konvergenz in Wahrscheinlichkeit, definiert werden.

**Definition 19** (Gleichmäßige Konvergenz in Wahrscheinlichkeit). *Eine Folge von reellwertigen Funktionen  $(\Psi_n(\theta))_n$  konvergiert gleichmäßig in Wahrscheinlichkeit gegen die Grenzfunktion  $\Psi(\theta)$ , wenn*

$$\sup_{\theta \in \Theta} |\Psi_n(\theta) - \Psi(\theta)| \xrightarrow{p} 0. \tag{A.12}$$

*Die gleichmäßige Konvergenz in Wahrscheinlichkeit (A.12) impliziert die punktweise Konvergenz in Wahrscheinlichkeit (A.11).*

### A.3.3. Wichtige Sätze

Der folgende Satz liefert ein wichtiges Resultat über gestoppte stochastische Prozesse

**Satz 14** (Theorem des optionalen Stoppens). *Sei  $\{M(x) : x \in \mathbb{N}_0\}$  ein  $\mathcal{G}_x$ -(Sub-)Martingal und  $T$  eine  $\mathcal{G}_x$ -Stoppzeit. Dann ist der gestoppte Prozess  $\{M(x \wedge T) : x \in \mathbb{N}_0\}$  ein (Sub-)Martingal sowohl bezüglich  $\{\mathcal{G}_x : x \in \mathbb{N}_0\}$  als auch bezüglich der gestoppten Filtration  $\{\mathcal{G}_{x \wedge T} : x \in \mathbb{N}_0\}$ .*

Ein Beweis für das Theorem des optionalen Stoppens findet sich in Doob [1953, Kapitel VII, Theorem 2.1].

Als Nächstes folgen zwei Resultate aus der asymptotischen Statistik, beginnend mit dem schwachen Gesetz der großen Zahlen [siehe Billingsley, 2012, Seite 90/91], gefolgt von dem Satz von Slutsky [siehe van der Vaart, 1998, Kapitel 2.1, Lemma 2.8].

**Satz 15** (Schwaches Gesetz der großen Zahlen). *Seien  $X_1, X_2, \dots$  unabhängige, auf einem gemeinsamen Wahrscheinlichkeitsraum  $(\Omega, \mathcal{F}, P_\theta)$  definierte und identisch verteilte Zufallsvariablen. Sei außerdem  $S_n = X_1 + \dots + X_n$  mit  $n \in \mathbb{N}$ . Dann gilt*

$$\lim_{n \rightarrow \infty} P\{|n^{-1}S_n - E_\theta\{X_n\}| \geq \epsilon\} = 0$$

oder, alternativ formuliert:

$$\frac{1}{n} \sum_{i=1}^n X_i \xrightarrow{p} E_\theta\{X_n\}$$

**Satz 16** (Satz von Slutsky). *Seien  $(X_n)_n$  und  $(Y_n)_n$  Zufallsfolgen und  $X$  eine Zufallsvariable. Wenn  $X_n \rightsquigarrow X$  und  $Y_n \rightsquigarrow c$  für eine Konstante  $c$  erfüllt ist, dann gelten folgende Aussagen:*

- (i)  $X_n + Y_n \rightsquigarrow X + c$
- (ii)  $Y_n X_n \rightsquigarrow cX$
- (iii)  $Y_n^{-1} X_n \rightsquigarrow c^{-1}X$ , falls  $c \neq 0$ .

Abschließend folgt das Existenz- und Eindeigkeitstheorem, mithilfe dessen die Existenz sowie Eindeigkeit eines Poisson-Prozesses mit entsprechendem Intensitätsmaß unter gewissen Bedingungen abgeleitet werden kann [siehe Reiss, 1993, Theorem 1.2.1].

**Satz 17** (Existenz- und Eindeigkeitstheorem). *Seien  $(\tilde{X}_j, \tilde{T}_j)'$ ,  $j \in \{1, \dots, m\}$  unabhängig identisch verteilte Zufallsvektoren, die in den Messraum  $(\mathbb{S}, \mathcal{U})$  abbilden. Sei weiter  $H_n^*$  ein Punktprozess mit endlichem Intensitätsmaß  $\eta_n^*$ , der gegeben ist durch:*

$$H_n^* = \sum_{j=1}^{\tilde{Z}} \epsilon_{(\tilde{X}_j, \tilde{T}_j)'}$$

wobei  $\tilde{Z}$  eine von  $(\tilde{X}_j, \tilde{T}_j)'$  unabhängige, Poisson-verteilte Zufallsgröße mit dem Parameter  $\eta_n^*(\mathbb{S})$  ist. Gelten zusätzlich  $\eta_n^*(\mathbb{S}) > 0$  sowie  $\mathcal{L}[(\tilde{X}_j, \tilde{T}_j)] = \eta_n^*/\eta_n^*(\mathbb{S})$ , so ist  $H_n^*$  ein Poisson-Prozess mit dem Intensitätsmaß  $\eta_n^*$ . Außerdem sind andere Poisson-Prozesse, die dasselbe (endliche) Intensitätsmaß  $\eta_n^*$  haben, verteilungsgleich zu  $H_n^*$ .

## A.4. R-Code

### A.4.1. Analyse des AFiD-Panels

```

1 rm(list=ls())
2
3 install.packages("tidyverse")
4 install.packages("lubridate")
5
6 library(tidyverse)
7 library(lubridate)
8
9 #Defintion der Daten (N) inklusive Lebensdauern (X) und Alter bei Studieneintritt (T)
10 N <- matrix(c(37016,17295,191803,
11 42272,20305,200411,
12 35588,23353,209649,
13 34579,27464,278223,
14 18687,18633,292566
15 ), byrow=TRUE, nrow=5)
16
17 X <- matrix(c(5,6,7,
18 4,5,6,
19 3,4,5,
20 2,3,4,
21 1,2,3), byrow=TRUE, nrow=5)
22
23 T <- matrix(c(4,4,4,
24 3,3,3,
25 2,2,2,
26 1,1,1,
27 0,0,0), byrow=TRUE, nrow=5)
28
29 #Berechnung Punktschaetzer
30 #Zaehler(A)
31 A=0
32 for (i in 1:5){
33 for (j in 1:2){
34 A=A+N[i,j]
35 }
36 }
37
38 #Nenner (A+B)
39 B=sum(N*((X-1)-T))
40
41 #Punktschaetzer
42 theta=A/(A+B)

```

```

43
44
45 #log-Likelihood fuer Plot
46 theta_plot <- c(1000:999000)/1000000
47 log_Like <- c()
48 for (i in 1:998001){
49 log_Like[i]=A*log10(theta_plot[i])+B*log10((1-theta_plot[i]))
50 }
51
52 #Plot der log-Likelihood entlang des gesamten Parameterraums
53 Likeplot <- tibble(theta_plot, log_Like) %>%
54 rename(theta=theta_plot)
55 Likeplot %>%
56 ggplot(aes(x=theta, y=log_Like))+
57 geom_line()+
58 ggtitle ("Logarithmierte Likelihood-Funktion")
59
60
61
62 #Plot der log-Likelihood im Bereich des Maximums
63 Likeplot %>%
64 ggplot(aes(x=theta, y=log_Like))+
65 geom_line()+
66 scale_x_continuous(limits=c(0.0625, 0.15))+
67 scale_y_continuous(limits=c(-400000, -387350))+
68 ggtitle ("Log-Likelihood-Funktion")
69
70 #####Standardfehler#####
71 #Berechnung der asymptotischen Varianz nach Formel (3.44)
72 #Schätzung von n ueber alpha_theta -> SE1 (Formel (3.46))
73
74 #Berechnung sigma^2 (Formel (3.44))
75 G=5
76 s=2
77
78 Zaehler=0
79 for (t in 0:(G-1)){
80 Zaehler=Zaehler+s^2
81 for (x in (t+1):(t+s)){
82 Zaehler=Zaehler+(1/theta)*(1-theta)^(x-1)
83 }
84 for (x in (t+1):(t+s)){
85 Zaehler=Zaehler-2*(1-theta)^(x-2)*(x-t-1)
86 }
87 for (x in (t+2):(t+s+1)){
88 Zaehler=Zaehler+theta*(1-theta)^(x-3)*(x-t-1)^2
89 }
90 if (t+s >=2){
91 for (x in 1:(t+s-1)){
92 Zaehler=Zaehler-s^2*(theta*(1-theta)^(x-1))
93 }
94 }
95 }
96

```

```

97 Nenner=0
98 for (t in 0:(G-1)){
99   Nenner=Nenner+s
100  for (x in (t+1):(t+s)){
101    Nenner=Nenner+(1/theta)*(1-theta)^(x-1)
102  }
103  for (x in (t+2):(t+s+1)){
104    Nenner=Nenner+theta*(1-theta)^(x-3)*(x-t-1)
105  }
106  if (t+s >=2){
107    for (x in 1:(t+s-1)){
108      Nenner=Nenner-s*(theta*(1-theta)^(x-1))
109    }
110  }
111 }
112
113 Nenner=Nenner^2
114 sigma=Zaehler/Nenner
115
116 #Berechnung Standardfehler SE1
117 Faktor=1/sum(N)*(1-(1-theta)^G)/theta
118 SE1=sqrt(Faktor*Zaehler/Nenner)
119
120 #95 Prozent-Konfidenzintervalle fuer theta0 und Lebenserwartung
121 Q=1.96
122 KI1 <- c(theta-Q*SE1, theta+Q*SE1)
123 LE1=1/KI1
124
125
126
127 #Ansatz Van-der-Vaart, Erwartungswerte ueber Mittelwerte schaezen -> SE2(Formel (3.47))
128 #Berechnung Standardfehler
129 Zaehler=(1/theta)^2*sum(N[,1])+(1/theta - 1/(1-theta))^2*sum(N[,2])+(2/(1-theta))^2*sum(N
[,3])
130 Nenner=( (1/(theta^2))*sum(N[,1]) + (1/(theta^2)+1/((1-theta)^2))*sum(N[,2]) + (2/((1-theta)
^2))*sum(N[,3]) )^2
131 SE2=sqrt(Zaehler/Nenner)
132
133 #Berechnung Konfidenzintervalle
134 KI2 <- c(theta-Q*SE2, theta+Q*SE2)
135 LE2=1/KI2
136
137
138
139 #Ansatz Semiparametrisch per Martingalgrenzwertsatz -> SE3(Formel vor (3.75))
140 #Berechnung Standardfehler
141 X2 <- matrix(c(5,6,6,
142 4,5,5,
143 3,4,4,
144 2,3,3,
145 1,2,2), byrow=TRUE, nrow=5)
146
147 Y=sum(N*(X2-T))
148 SE3=sqrt((theta*(1-theta))/Y)

```

```

149
150 #Berechnung Konfidenzintervalle
151 KI3 <- c(theta-Q*SE3, theta+Q*SE3)
152 LE3=1/KI3
153
154
155
156 #Ansatz Parametrisch per Martingalgrenzwertsatz -> SE4(Formel (A.8))
157 #Berechnung Standardfehler
158 Zaehler4=(1-(1-theta)^G)*(1-theta)
159 Nenner4=0
160 for (x in 1:(s+G-1)){
161 Rest=0
162 if (x >=2){
163 for (k in 1:(x-1)){
164 Rest=Rest+theta*(1-theta)^(k-1)
165 }
166 }
167 Nenner4=Nenner4+min(x,s,G,s+G-x)*(1-Rest)
168 }
169 Nenner4=Nenner4*sum(N)
170 SE4=sqrt(Zaehler4/Nenner4)
171
172 #Berechnung Konfidenzintervalle
173 KI4 <- c(theta-Q*SE4, theta+Q*SE4)
174 LE4=1/KI4
175
176
177
178
179 #####Punktschaetzer bei doppelt trunkiertem Design (Kapitel 4)#####
180 #Umformen der Daten
181 G=5
182 s=2
183 NDT <- N[,1:2]
184 XDT <- X[,1:2]
185 m=sum(NDT)
186 vektor <- c()
187 count=1
188
189 for (i in 1:2){
190 for (j in 1:5){
191 for (k in 1:NDT[j,i]){
192 vektor[count] <- XDT[j,i]
193 count=count+1
194 }
195 }
196 }
197
198 DTdata <- tibble(vektor)
199 DTdata <- DTdata %>%
200 rename(X=vektor)
201
202

```

```

203 #Nullstelle der abgeleiteten log-Likelihood naeherungsweise bestimmen
204 scale=1000000
205 L <- c(1:scale)
206 thetav <- c(1:scale)/(scale+1)
207 for (i in 1:scale){
208   theta0=thetav[i]
209   L[i]=0
210   Z=m*(((G+s-1)*theta0+1)*(1-theta0)^(G+s-1)-((G-1)*theta0+1)*(1-theta0)^(G-1)
211     -((s-1)*theta0+1)*(1-theta0)^(s-1)+1)
212   N=theta0*((1-theta0)^(G+s)-(1-theta0)^G - (1-theta0)^(s+1))
213   Summand=sum((1-DTdata$X*theta0)/(theta0*(1-theta0)))
214   Bruch=Z/N
215   L[i]=Bruch+Summand
216 }
217 L=abs(L)
218 MLE=which.min(L)
219 MLE=MLE/scale
220
221 #Berechnung alpha_theta^DT sowie geschaetzter Stichprobenumfang
222 alphaDT=((1-MLE)^(G+s)-(1-MLE)^G-(1-MLE)^(s+1))/(MLE*G)
223 n_est=m/alphaDT

```

#### A.4.2. Code zur Simulation

```

1 rm(list=ls())
2 install.packages("tidyverse")
3 install.packages("lubridate")
4
5 library(tidyverse)
6 library(lubridate)
7
8 #Festlegen der Parameter fuer Simulationsdurchlauf
9 w=1000
10 n=10000
11 G=5
12 s=2
13 theta0=0.1
14
15 #Berechnung wahre Varianz des Punktschaetzers (SE0)
16 Zaehler0=theta0*G*(1-theta0)
17
18 Nenner0=0
19 for (x in 1:(s+G-1)){
20   Rest=0
21   if (x >=2){
22     for (k in 1:(x-1)){
23       Rest=Rest+theta0*(1-theta0)^(k-1)
24     }
25   }
26   Nenner0=Nenner0+min(x,s,G,s+G-x)*(1-Rest)
27 }
28 Nenner0=Nenner0*n

```

```

29
30 SE0=sqrt(Zaehler0/Nenner0)
31
32 #####Simulation#####
33 #Erstellung der Vektoren (z-te Komponente entspricht z-tem Simulationsdurchlauf)
34 theta <- c()
35 m <- c()
36 SE1 <- c()
37 SE2 <- c()
38 SE3 <- c()
39 SE4 <- c()
40
41 for (z in 1:(w)){
42   #Erstellung der Daten
43   T <- rdunif(n, 0, G-1)
44   Xu <- rgeom(n, theta0)+1
45   data <- tibble(Xu,T)
46   data <- data %>%
47     filter (T+1<=Xu) %>%
48     mutate(X=pmin(Xu, T+s+1),
49            X2=pmin(Xu, T+s),
50            C=ifelse(T+s+1-Xu>0, 1, 0),
51            Life=X-1-T) %>%
52     select (X, X2, Xu, T, C, Life)
53
54
55   #Punktschaetzung
56   Estimation <- data %>%
57     filter (X !=T+s+1) %>%
58     summarize(Deaths = n()) %>%
59     mutate(Lifes=sum(data$Life),
60            theta_est=Deaths/(Deaths+Lifes))
61   theta[z] <- Estimation$theta_est
62
63   #Standardfehler SE1
64   m[z]=length(data$X)
65   Zaehler=0
66   for (t in 0:(G-1)){
67     Zaehler=Zaehler+s^2
68     for (x in (t+1):(t+s)){
69       Zaehler=Zaehler+(1/theta[z])*(1-theta[z])^(x-1)
70     }
71     for (x in (t+1):(t+s)){
72       Zaehler=Zaehler-2*(1-theta[z])^(x-2)*(x-t-1)
73     }
74     for (x in (t+2):(t+s+1)){
75       Zaehler=Zaehler+theta[z]*(1-theta[z])^(x-3)*(x-t-1)^2
76     }
77     if (t+s >=2){
78       for ( x in 1:(t+s-1)){
79         Zaehler=Zaehler-s^2*(theta[z]*(1-theta[z])^(x-1))
80       }
81     }
82   }

```

```

83
84 Nenner=0
85 for (t in 0:(G-1)){
86   Nenner=Nenner+s
87   for (x in (t+1):(t+s)){
88     Nenner=Nenner+(1/theta[z]*(1-theta[z])^(x-1)
89   }
90   for (x in (t+2):(t+s+1)){
91     Nenner=Nenner+theta[z]*(1-theta[z])^(x-3)*(x-t-1)
92   }
93   if (t+s >=2){
94     for (x in 1:(t+s-1)){
95       Nenner=Nenner-s*(theta[z]*(1-theta[z])^(x-1))
96     }
97   }
98 }
99
100 Nenner=Nenner^2
101 Faktor=1/m[z]*(1-(1-theta[z])^G)/theta[z]
102
103 SE1[z]=sqrt(Faktor*Zaehler/Nenner)
104
105 #Standardfehler SE2
106 Zaehler2=sum((data$C/theta[z]-(data$Life)/(1-theta[z]))^2)
107 Nenner2=(sum((data$C/(theta[z]^2)+(data$Life)/((1-theta[z]^2))))^2)
108 SE2[z]=sqrt(Zaehler2/Nenner2)
109
110 #Standardfehler SE3
111 Y=sum(data$X2-data$T)
112 SE3[z]=sqrt((theta[z]*(1-theta[z]))/Y)
113
114 #Standardfehler SE4
115 Zaehler4=(1-(1-theta[z])^G)*(1-theta[z])
116 Nenner4=0
117 for (x in 1:(s+G-1)){
118   Rest=0
119   if (x >=2){
120     for (k in 1:(x-1)){
121       Rest=Rest+theta[z]*(1-theta[z])^(k-1)
122     }
123   }
124   Nenner4=Nenner4+min(x,s,G,s+G-x)*(1-Rest)
125 }
126
127 Nenner4=Nenner4*m[z]
128 SE4[z]=sqrt(Zaehler4/Nenner4)
129 }
130
131 #Berechnung der mittleren quadratischen Fehler
132 ME_theta = sum((theta-theta0)^2)/w
133 ME1 = sum((SE1-SE0)^2)/w
134 ME2 = sum((SE2-SE0)^2)/w
135 ME3 = sum((SE3-SE0)^2)/w
136 ME4 = sum((SE4-SE0)^2)/w

```

```

137
138
139 #####Simulation Punktschaetzung bei doppelt trunkiertem Design#####
140 #Paramterwahl
141 w=1000
142 n=100
143 G=5
144 s=2
145 theta=0.1
146
147 #Erstellung Ergebnis-Vektor
148 MLE <- c()
149
150 for (z in 1:w){
151   #Generierung der Daten
152   T <- rdunif(n, 0, G-1)
153   Xu <- rgeom(n, theta)+1
154
155   DTdata <- tibble(Xu, T)
156   DTdata <- DTdata %>%
157     filter (T+1 <= Xu)%>%
158     filter (Xu <= T+s)%>%
159     rename (X=Xu)
160
161   #Berechnung der Punktschaetzung (naerungsweise Nullstellenberechnung von Formel (4.24))
162   m = length(DTdata$X)
163   scale=100000
164   L <- c(1:scale)
165   for (i in 1:scale){
166     theta0=i/(scale+1)
167     L[i]=0
168     Y=m*(((G+s-1)*theta0+1)*(1-theta0)^(G+s-1)-((G-1)*theta0+1)*(1-theta0)^(G-1)
169     -((s-1)*theta0+1)*(1-theta0)^(s-1)+1)
170     N=theta0*((1-theta0)^(G+s)-(1-theta0)^G - (1-theta0)^(s)+1)
171     Summand=sum((1-DTdata$X*theta0)/(theta0*(1-theta0)))
172     Bruch=Y/N
173     L[i]=Bruch+Summand
174   }
175
176   L=abs(L)
177   MLE[z]=which.min(L)
178
179   MLE[z]=MLE[z]/scale
180 }
181 MSE=(MLE-theta)^2
182 MSE=sum(MSE)/w


```

# Eidesstattliche Versicherung

Ich erkläre hiermit, dass ich die vorliegende Arbeit ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe; die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sind als solche kenntlich gemacht.

Die Arbeit wurde bisher weder im Inland noch im Ausland in gleicher oder ähnlicher Form einer Prüfungsbehörde zur Erlangung eines akademischen Grades vorgelegt.

Rostock, den 20.11.2024



Eric Scholz