



Technische Universität Dortmund
Fakultät für Mathematik
Lehrstuhl IV - Stochastik



Masterarbeit im Fachbereich
Statistik in den Ingenieurwissenschaften

Asymptotische Eigenschaften eines Minimum-Distanz-Schätzers für selbstanregende Punktprozesse

vorgelegt von	Mirko A. Jakubzik
Matrikelnummer	147717
Datum der Ausgabe	16. September 2016
Datum der Abgabe	10. Februar 2017
Betreuer/in der Masterarbeit	Prof. Dr. Christine Müller
Gutachter/in der Masterarbeit	Prof. Dr. Jeannette Woerner

Danksagung

An dieser Stelle möchte ich mich bei all denjenigen bedanken, die mich bei der Anfertigung der vorliegenden Masterarbeit unterstützend begleitet haben.

Ganz besonders gilt mein Dank Frau Prof. Dr. Christine Müller, die in zahllosen fachlichen Diskussionen stets mit wertvoller Kritik und inspirierenden Anregungen aufwarten konnte und durch ihre moralische Unterstützung wesentlich zur Vollendung dieser Arbeit beigetragen hat.

Darüber hinaus möchte ich mich bei Frau Prof. Dr. Jeannette Woerner und Herrn Prof. Dr. Heribert Blum bedanken, die durch ihre Zustimmung die statistisch-mathematische Kooperation überhaupt erst möglich gemacht haben.

Auch möchte ich meinen Eltern, Wilfried und Martina Jakubzik, dafür danken, dass sie mich zeit meines Lebens dazu motivieren, meinen eigenen Weg zu beschreiten, und mir dabei stets den Rücken freihalten.

Schließlich gilt mein Dank Katharina Bata, die tapfer so manchen Anflug von Frustration ertragen und ins Gegenteil verkehrt hat.

Inhaltsverzeichnis

1	Einleitung	4
2	Mathematische Grundlagen	8
2.1	Grundlagen stochastischer Prozesse	8
2.1.1	Allgemeine stochastische Prozesse	8
2.1.2	Punkt- und Zählprozesse	9
2.1.3	Spezialfall: Verschobener Geburtsprozess	12
2.1.4	Spezialfall: Reiner verschobener Geburtsprozess mit Abbruch	20
2.1.5	Filtrationen und Martingale	32
2.1.6	Doob-Meyer-Zerlegung für nicht-negative Submartingale	35
2.1.7	Itô-Isometrie für quadratisch integrierbare Martingale	42
2.1.8	Selbstanregende Punkt- und Zählprozesse	44
2.2	U-Statistiken	48
2.2.1	Definition und das starke Gesetz der großen Zahlen für U-Statistiken	48
2.2.2	Das 0-1-Gesetz von Hewitt-Savage	50
2.2.3	Rückwärtsmartingale und der Doob'sche Konvergenzsatz	59
2.2.4	U-Statistiken als Rückwärtsmartingal	61
2.2.5	Der Beweis des starken Gesetzes der großen Zahlen für U-Statistiken	62
2.3	Straffheit im Raum der stetigen Funktionen	64
2.3.1	Das Straffheitskriterium von Kolmogorov	70
3	Der Minimum-Distanz-Schätzer ϑ_n	77
3.1	Definition des Schätzers ϑ_n	78
3.2	Folgerungen aus dem starken Gesetz der großen Zahlen für U-Statistiken . .	83
3.3	Konsistenz des Schätzers ϑ_n	90
3.4	Asymptotische Verteilung des Schätzers ϑ_n	100
3.4.1	Die Hilfsprozesse $\{\alpha_n\}$, $\{\beta_n\}$ und $\{\gamma_n\}$	102
	Darstellungssatz für die parametrischen Prozesse $\{\alpha_n\}$	103
	Darstellungssatz für die parametrischen Prozesse $\{\beta_n\}$	123
	Darstellungssatz für die parametrischen Prozesse $\{\gamma_n\}$	127
3.4.2	Die asymptotische Normalverteiltheit von ϑ_n	129
3.4.3	Asymptotik der approximativen Standardisierungsmatrix Φ_n	131
3.4.4	Beweis der asymptotischen Normalverteiltheit	137
4	Diskussion und Ausblick	142

A Anhang	144
A.1 Elementare Hilfssätze	144
A.2 Eine Abzählung des Gitters $[0, 1]^d \cap \delta\mathbb{Z}^d$	148
A.3 Das Lipschitz-Straffheitskriterium	151

1 Einleitung

Die vorliegende interdisziplinäre Arbeit trägt den Titel „Asymptotische Eigenschaften eines Minimum-Distanz-Schätzers für selbstanregende Punktprozesse“ und vereint die Statistik in den Ingenieurwissenschaften mit der Stochastik aus der mathematischen Domäne. Infolgedessen liegt es nahe, die Motivation für die nachfolgend diskutierte mathematische Theorie aus dem Anwendungskontext des Bauingenieurwesens zu beziehen.

Hierzu verweisen wir auf ein im Rahmen des SFB 823 (*“Statistical modelling of nonlinear dynamic processes”*) an der Technischen Universität Dortmund durchgeführtes Belastungsexperiment (siehe Szugat u. a. 2016):

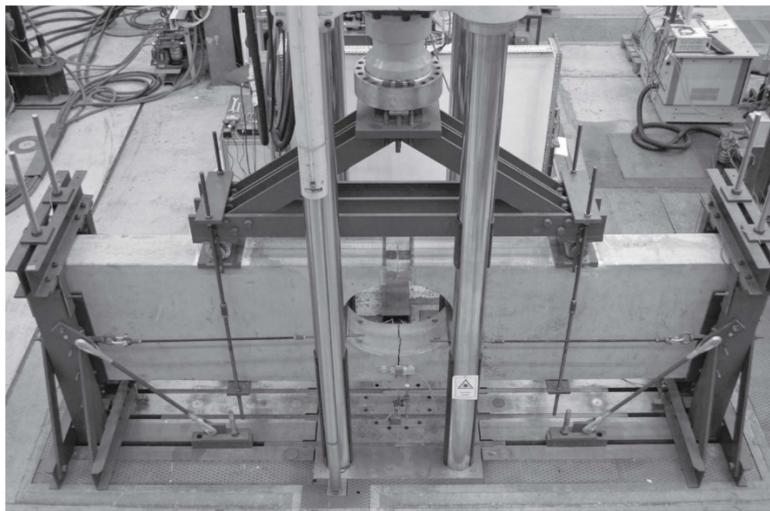


Abbildung 1: Fotografie des Versuchsaufbaus; siehe Szugat u. a. 2016.

In diesem Versuch wurden fünf Spannbetonträger (bezeichnet mit SB01 bis SB05) zyklischen Belastungen im Bereich von 80 MPa bis 200 MPa ausgesetzt, siehe Abbildung 1. Nachdem ein Initialriss im Spannbeton festgestellt werden konnte, wurde im weiteren Verlauf des Experiments eine stetige Überprüfung der Rissweite vorgenommen. Die dabei beobachteten sprunghaften Anstiege der Rissweite konnten mit dem Bruch der im Spannbeton verbauten Spanndrähte in Verbindung gebracht werden, siehe Abbildung 2. Dies ermöglichte es, den exakten Zeitpunkt eines Drahtbruches anhand des beobachtbaren Risswachstums zu observieren.

Die Versuchsergebnisse werden in Abbildung 3 grafisch dargestellt, wobei die Kennziffern TR01 bis TR05 weitere Spannbetonträger aus einem ebenfalls an der TU Dortmund durchgeführten Experiment indizieren.

In der Praxis werden derartige Spannbetonträger, wie sie im oben dargelegten Versuch verwendet wurden, beispielsweise zum Brückenbau verwendet. Konkret stammen die mit TR01 bis TR05 bezeichneten Träger aus einer im Jahre 2007 abgerissenen Brücke (siehe

wiederum Szugat u. a. 2016). Da der Bruch sämtlicher im Beton verbauter Spanndrähte mit dem Versagen des Bauwerks einhergeht, möchten wir eine möglichst präzise Vorhersage der Drahtbrüche treffen.



Abbildung 2: Im Zuge der Belastung gebrochene Spanndrähte; siehe Müller, Szugat und Maurer 2016.

Zu diesem Zwecke betrachten wir den mit den Zeitpunkten der Drahtbrüche assoziierten Zählprozess und interpretieren folglich die in Abbildung 3 präsentierten Versuchsergebnisse als Realisierungen unabhängiger Zählprozesse.

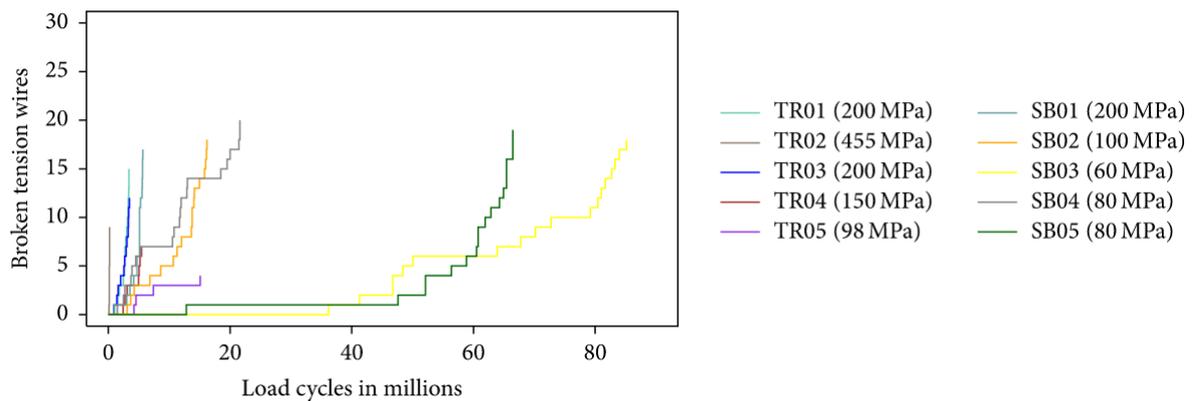


Abbildung 3: Plot der Anzahl der gebrochenen Spanndrähte in Abhängigkeit der Lastzyklen; siehe Szugat u. a. 2016.

Wie wir insbesondere dem Pfad des mit dem Spannbetonträger SB05 assoziierten Zählpro-

zesses entnehmen, folgen die Realisierungen einer selbstanregenden Dynamik, die mit der heuristischen Interpretation konform geht:

Je mehr Spanndrähte brechen, desto mehr Last wird von den verbleibenden Spanndrähten getragen, die somit wiederum zum schnelleren Brechen neigen.

Dies motiviert schließlich die Einführung der namensgebenden *selbstanregenden Punkt- oder Zählprozesse*. Zur Vorhersage eines solchen Prozesses erweist sich der Kompensator aus der Doob-Meyer-Zerlegung für nicht-negative Submartingale, den wir als Verallgemeinerung der kumulierten Intensitätsfunktion eines Zählprozesses kennenlernen, als geeignet. Infolgedessen setzen wir zur Beschreibung des qualitativen Verhaltens eines selbstanregenden Zählprozesses ein semiparametrisches Modell für den zugehörigen Kompensator an. Sodann führen wir einen Minimum-Distanz-Schätzer für den Parameter des modellierten Kompensators ein.

Den Kern dieser Arbeit bildet die Analyse der asymptotischen Eigenschaften dieses Parameterschätzers:

Der Minimum-Distanz-Schätzer ist stark konsistent und asymptotisch normalverteilt (siehe Stute und Kopperschmidt 2013).

Perspektivisch kann daher die Kenntnis der asymptotischen Verteilung des Schätzers genutzt werden, um ein approximatives Prognoseintervall für die Anzahl der Drahtbrüche zu einem gegebenen Zeitpunkt anzugeben.

Die Arbeit besteht neben der Einleitung aus drei weiteren Kapiteln.

Das zweite Kapitel befasst sich mit den *mathematischen Grundlagen* dieser Arbeit und ist in drei Abschnitte unterteilt:

- Zunächst behandeln wir die Theorie der *stochastischen Prozesse* und lernen insbesondere die Doob-Meyer-Zerlegung für nicht-negative Submartingale und die selbstanregenden Punkt- und Zählprozesse kennen, die wir in ihrer allgemeinsten Form definieren werden. Ferner betrachten wir eine Reduktion des obigen Experimentes und diskutieren den Spezialfall des verschobenen Geburtsprozesses. Dieses Beispiel eines selbstanregenden Zählprozesses überspannt die gesamte Arbeit. Die Idee und die mathematische Umsetzung derselbigen wurden vollständig vom Autor dieser Arbeit beigesteuert. Dies ermöglicht es in besonderem Maße, eine Brücke zwischen dem praktischen Kontext und der mathematischen Theorie zu schlagen.
- Der zweite Abschnitt thematisiert die sogenannten *U-Statistiken*. Das zentrale Resultat stellt eine Verallgemeinerung des starken Gesetzes der großen Zahlen dar, vermöge dessen wir die starke Konsistenz des Minimum-Distanz-Schätzers zeigen werden. Wir prä-

sentieren dem Leser nebst einem detaillierten Beweis eine Vielzahl technischer Grundlagen, zu denen auch das 0-1-Gesetz für permutierbare Mengen von Hewitt-Savage gehört.

- Der letzte Abschnitt befasst sich mit der *Straffheit von Wahrscheinlichkeitsmaßen über dem Raum der stetigen Funktionen*. Um die asymptotische Normalverteiltheit des Minimum-Distanz-Schätzers nachzuweisen, greifen wir später auf das Straffheitskriterium von Kolmogorov zurück. Dieses in seiner gewöhnlichen Form lediglich im eindimensionalen Fall gültige Kriterium wurde vom Autor dieser Arbeit eigenständig ins Mehrdimensionale transferiert. Außerdem wird im Anhang ein alternatives Kriterium diskutiert.

Das dritte Kapitel ist dem *Minimum-Distanz-Schätzer* gewidmet. Es stellt das zentrale Kapitel dieser Arbeit dar:

Nach einer kurzen Diskussion der zugrunde liegenden Situation definieren wir den besagten Schätzer, dessen asymptotische Eigenschaften nachfolgend im Vordergrund stehen.

Während wir uns grob an der Vorlage aus dem Artikel Stute und Kopperschmidt 2013 zum Nachweis der starken Konsistenz und asymptotischen Normalverteiltheit orientieren, nehmen wir zugleich eine kritische Haltung ein; wir begnügen uns nicht damit, die Beweise nachzuvollziehen, sondern fügen sinnvolle Erweiterungen hinzu, modifizieren sich als nicht hinreichend herausstellende Voraussetzungen und ergänzen fehlende Argumentationsschritte bis hin zu vollständigen Beweisen.

Das Ergebnis ist eine umfangreiche, mathematisch fundierte Analyse des Minimum-Distanz-Schätzers, in der auch technische Forderungen nicht unterschlagen werden. Gleichmaßen führen wir stets das zentrale Beispiel dieser Arbeit fort und zeigen die Praktikabilität der Modellierung mittels des im zweiten Kapitel thematisierten reinen verschobenen Geburtsprozesses mit Abbruch.

Im vierten und letzten Kapitel schließlich möchten wir einerseits retrospektiv die Komplikationen der vorherigen Kapitel beleuchten, während wir andererseits die weiterführenden Arbeitsschritte des Autors skizzieren.

Den Abschluss dieser Arbeit bildet der Anhang, der aufgrund der vom Autor eigenständig bewiesenen Resultate zur Abzählung eines mehrdimensionalen Gitters sowie dem Lipschitz-Straffheitskriterium nicht ungesehen übersprungen werden sollte. Hier finden sich auch die im dritten Kapitel utilisierten Hilfssätze, die ebenfalls im Artikel Stute und Kopperschmidt 2013 formuliert werden.

2 Mathematische Grundlagen

Dieses Kapitel der vorliegenden Arbeit dient der Herstellung eines Konsens bezüglich der mathematischen Vorkenntnisse. Außerdem wird eine Vielzahl von Schreibweisen eingeführt, die in den nachfolgenden Abschnitten übernommen werden sollen. Dem kenntnisreichen Leser sei es also nahegelegt, von einem unnötig detaillierten Studium der hier dargelegten Definitionen und Resultate abzusehen und lediglich die verwendeten Notationen rückwirkend nachzuschlagen. Hierbei mag lediglich der *verschobene Geburtsprozess* aus den Unterabschnitten 2.1.3 und 2.1.4 eine Ausnahme darstellen, fungiert er doch als motivierendes Beispiel eines *selbstanregenden Punktprozesses*.

Als hinreichend für das Verständnis der zentralen Resultate späterer Abschnitte werden erweiterte Grundlagen aus der Maß- und Wahrscheinlichkeitstheorie sowie der Topologie und Funktionalanalysis erachtet.

Um auch dem interessierten Leser mit weniger ausgeprägtem Fundus mathematischer Grundlagen die Möglichkeit zu bieten, die komplexeren Beweisschritte im Laufe dieser Arbeit nachzuvollziehen, soll auf den nächsten Seiten nichtsdestoweniger eine knappe Einführung in die oben genannten mathematischen Felder erfolgen.

Da auf elementare Konzepte und Begrifflichkeiten wie σ -Algebren, Zufallsvariablen, Wahrscheinlichkeitsmaße und ähnliches nicht eingegangen werden soll, sei an dieser Stelle auf die ausführlichen Lehrbücher Bauer 1991 und Bauer 1992 verwiesen, sollte der Leser nachfolgend auf ihm unbekannte Begriffe stoßen.

2.1 Grundlagen stochastischer Prozesse

2.1.1 Allgemeine stochastische Prozesse

Bevor wir uns im weiteren Verlauf dieses Abschnittes auch speziellen stochastischen Prozessen wie den namensgebenden *selbstanregenden Punktprozessen* zuwenden, führen wir zunächst allgemeine *stochastische Prozesse* ein.

Definition 2.1.1. (*Stochastischer Prozess*, siehe Ethier und Kurtz 1986, S. 49)

Es bezeichne (Ω, \mathcal{F}, P) einen Wahrscheinlichkeitsraum und (E, \mathcal{E}) einen Messraum.

Ein *stochastischer Prozess* X mit *Indexmenge* \mathcal{I} und *Zustandsraum* (E, \mathcal{E}) über dem Wahrscheinlichkeitsraum (Ω, \mathcal{F}, P) ist eine auf $\mathcal{I} \times \Omega$ definierte Funktion mit Werten in E dergestalt, dass für alle $t \in \mathcal{I}$ die Funktion

$$X(t, \cdot) : \Omega \longrightarrow E$$

eine E -wertige Zufallsvariable ist, das heißt es gilt:

$$\{\omega \in \Omega : X(t, \omega) \in A\} \in \mathcal{F}, \quad \text{für alle } A \in \mathcal{E}.$$

Häufig schreiben wir dann $X = (X_t)_{t \in I}$.

Als Indexmenge \mathcal{I} werden wir beispielsweise die natürlichen Zahlen \mathbb{N} , ein kompaktes Intervall $I \subset \mathbb{R}$ oder auch eine Teilmenge des d -dimensionalen Raumes \mathbb{R}^d für $d \in \mathbb{N}$ betrachten. Ist die Indexmenge eine Teilmenge der reellen Zahlen, so werden die Indizes $t \in \mathcal{I}$ häufig - aber nicht ausschließlich - mit *Zeitpunkten* assoziiert. In diesem Falle spricht man von einem stochastischen Prozess in diskreter Zeit, falls \mathcal{I} diskret ist (etwa $\mathcal{I} = \mathbb{N}$), beziehungsweise in stetiger Zeit, falls \mathcal{I} ein Kontinuum ist (etwa $\mathcal{I} = [0, 1]$).

Definition 2.1.2. (*Messbarkeit und Stetigkeit eines Prozesses in stetiger Zeit*, siehe Ethier und Kurtz 1986, S. 50)

Es sei $I \subset \mathbb{R}$ ein Intervall und $X = (X_t)_{t \in I}$ ein stochastischer Prozess in stetiger Zeit. Der Prozess X heißt dann *messbar*, falls $X : I \times \Omega \rightarrow E$ als Abbildung $(\mathcal{B}(I) \times \mathcal{F})$ - \mathcal{E} - messbar ist, wobei $\mathcal{B}(I)$ die *Borel'sche σ -Algebra* über dem Intervall I bezeichne¹. Ferner heißt X (*linksseitig/rechtsseitig*) *stetig*, falls für fast alle $\omega \in \Omega$ der sogenannte *Pfad* $X(\cdot, \omega)$ (*linksseitig/rechtsseitig*) stetig ist.

Bevor uns diese Eigenschaften im Rahmen der Martingalthorie ab Unterabschnitt 2.1.5 erneut gewahr werden, möchten wir uns nun ganz speziellen stochastischen Prozessen zuwenden: Den *Punkt-* und *Zählprozessen*.

2.1.2 Punkt- und Zählprozesse

Obgleich wir im Rahmen dieser Arbeit eine Vielzahl verschiedener stochastischer Prozesse kennenlernen werden, liegt unser Hauptaugenmerk doch auf den *Punkt-* und *Zählprozessen*, die wir in diesem Abschnitt einführen möchten. Wir werden eine zwischen *Punkt-* und *Zählprozessen* bestehende Dualität erkennen; infolgedessen mag es den Leser nicht verwundern, dass in der Fachliteratur häufig nur der Begriff des *Punktprozesses* auftaucht, dem wir in der nachfolgenden Definition Inhalt verleihen möchten.

Definition 2.1.3. (*Punktprozess*, vgl. Jacobsen 2006, S. 9)

Es sei $t_0 \in \mathbb{R}$. Ferner sei $T = (T_n)_{n \in \mathbb{N}}$ ein stochastischer Prozess mit Indexmenge \mathbb{N} und Zustandsraum $([t_0, \infty], \mathcal{B}([t_0, \infty]))$ über dem Wahrscheinlichkeitsraum (Ω, \mathcal{F}, P) . Dann heißt T ein *Punktprozess*, falls die nachfolgenden Bedingungen erfüllt sind:

- (i) $P(\{t_0 < T_1 \leq T_2 \leq \dots\}) = 1,$
- (ii) $P\left(\left\{\lim_{n \rightarrow \infty} T_n = \infty\right\}\right) = 1.$

Ist die Bedingung (ii) nicht erfüllt, so heißt T ein *Punktprozess mit Explosion*.

¹Messbarkeit kann analog beispielsweise für diskrete Indexmengen definiert werden.

Bemerkung 2.1.4.

Die Bedingungen (i) und (ii) aus der Definition eines Punktprozesses bedeuten, dass ein Punktprozess $T = (T_n)_{n \in \mathbb{N}}$ fast sicher eine monoton wachsende Folge von Zufallsvariablen mit Grenzwert ∞ ist, wobei der Wert ∞ tatsächlich angenommen werden kann.

Anders als in Jacobsen 2006 fordern wir dabei in Definition 2.1.3 *nicht* zusätzlich:

$$(iii) \quad P(\{T_n < T_{n+1}, T_n < \infty\}) = P(\{T_n < \infty\}), \quad n \in \mathbb{N}.$$

Diese Bedingung würde sicherstellen, dass die Folge der Zufallsvariablen *streng* monoton wachsend ist, solange sie endlich ist.

Sowohl der Verzicht auf diese Bedingung, als auch der Begriff der *Explosion* aus Definition 2.1.3 werden nachvollziehbar, sobald wir unseren Blick den *Zählprozessen* aus Definition 2.1.5 zuwenden.

Definition 2.1.5. (Zählprozess, vgl. Jacobsen 2006, S. 11)

Es sei $t_0 \in \mathbb{R}$ und T ein Punktprozess gemäß Definition 2.1.3.

Der mit T assoziierte *Zählprozess* ist der stochastische Prozess $N = (N_t)_{t \geq t_0}$ mit Indexmenge $[t_0, \infty)$ und Zustandsraum $(\mathbb{N}_0, \mathcal{P}(\mathbb{N}_0))$ über dem Wahrscheinlichkeitsraum (Ω, \mathcal{F}, P) , der definiert ist durch

$$N_t = \sum_{n=1}^{\infty} \mathbb{1}_{\{T_n \leq t\}}, \quad t \geq t_0.$$

Für eine Menge $A \subset \Omega$ bezeichne dabei $\mathbb{1}_A$ die *Indikatorfunktion* oder *Charakteristische Funktion* der Menge A , das heißt die Abbildung

$$\mathbb{1}_A : \Omega \longrightarrow \{0, 1\} : \omega \longmapsto \mathbb{1}_A(\omega) = \begin{cases} 1, & \omega \in A, \\ 0, & \omega \notin A. \end{cases}$$

Außerdem bezeichne $\mathcal{P}(\mathbb{N}_0)$ die Potenzmenge von \mathbb{N}_0 .

Ein mit einem Punktprozess T assoziierter Zählprozess N ist also ein rechtsseitig stetiger, stochastischer Prozess, der die Anzahl der Punkte in den Intervallen $[t_0, t]$ für $t \geq t_0$ zählt. Insbesondere ist ein solcher Zählprozess fast sicher monoton wachsend.

Aus einem gegebenen Zählprozess $N = (N_t)_{t \geq t_0}$ lässt sich umgekehrt leicht der assoziierte Punktprozess $T = (T_n)_{n \in \mathbb{N}}$ gewinnen, da fast sicher gilt (siehe Jacobsen 2006, S. 12):

$$T_n = \inf\{t \geq t_0 : N_t \geq n\}$$

Insbesondere gilt schließlich die angesprochene Dualität zwischen Punkt- und Zählprozes-

sen, die durch die Äquivalenz

$$T_n \leq t \iff N_t \geq n \quad (2.1.1)$$

ausgedrückt werden kann.

Bemerkung 2.1.6.

Wir möchten die Bedingungen (i) bis (iii) aus Definition 2.1.3 und Bemerkung 2.1.4 im Hinblick auf den assoziierten Zählprozess interpretieren:

- (i) Die erste Bedingung stellt sicher, dass der Zählprozess fast sicher monoton wachsend ist und dass $N_{t_0} = 0$ gilt. Sie stellt außerdem die Grundlage für die Dualität aus (2.1.1) dar.
- (ii) Die zweite Bedingung verhindert, dass $N_t = \infty$ für ein $t \in [t_0, \infty)$ ist. Der Begriff der *Explosion* des Punktprozesses ist also dahingehend begründet, dass der assoziierte Zählprozess in *endlicher* Zeit den Wert ∞ erreicht.
- (iii) Die dritte Bedingung schließlich verhindert, dass *Mehrfachsprünge* auftreten können. In der in Jacobsen 2006 vorliegenden Definition kann also zu jedem Zeitpunkt fast sicher nur ein einfacher *Sprung* des assoziierten Zählprozesses erfolgen, das heißt für alle $t > t_0$ gilt fast sicher

$$N_t - \lim_{s \uparrow t} N_s \in \{0, 1\}.$$

Um mit der Definition aus Snyder und Miller 1991 konform zu gehen, verzichten wir auf diese Bedingung, obgleich wir zu Beginn des kommenden Abschnittes eine vergleichbare Forderung stellen möchten.

Auf Grundlage der Dualität aus (2.1.1) werden wir uns zukünftig lediglich mit Zählprozessen auseinandersetzen. In diesem Sinne überrascht es nicht, dass wir später auf die Definition eines *selbstanregenden Zählprozesses* anstelle eines *selbstanregenden Punktprozesses* stoßen werden. Die Definition findet sich am Ende von Abschnitt 2.1 im Unterabschnitt 2.1.8 und erfordert die Einführung weiterer Konzepte, weswegen wir den Leser an dieser Stelle zunächst trösten möchten.

Der nächste Abschnitt sei nun einem speziellen Zählprozess gewidmet, der als Motivation für die nachfolgende Theorie der *selbstanregenden Zählprozesse* dienen soll:

Dem *verschobenen Geburtsprozess*.

2.1.3 Spezialfall: Vershobener Geburtsprozess

Befasst man sich mit der Theorie der Zählprozesse, so wird man unweigerlich mit dem wohl bekanntesten Vertreter dieser Klasse stochastischer Prozesse konfrontiert, dem *Poisson-Prozess*.

Nichtsdestoweniger werden wir erkennen, dass dieser Prozess aufgrund seiner inhärenten Eigenschaften für die Beschreibung diverser Zusammenhänge etwa aus den Ingenieurwissenschaften ungeeignet erscheint.

Sodann lernen wir den *verschobenen Geburtsprozess* kennen, der uns durch seine *selbstanregende* Natur Zugang zu einem breiteren Spektrum der modellierbaren Zusammenhänge verschafft. Dieser Gedankengang soll im nachfolgenden Beispiel ausführlicher illustriert werden.

Beispiel 2.1.7. (*Homogener Poisson-Prozess & Lastumverteilung*, vgl. Snyder und Miller 1991, S. 41)

Es bezeichne $N = (N_t)_{t \geq 0}$ einen Zählprozess, der die nachfolgenden Bedingungen erfülle:

- (i) Für $0 \leq s < t$ sind die Zuwächse $N_t - N_s$ Poisson-verteilt mit Parameter $\lambda(t - s)$, wobei $\lambda \in (0, \infty)$ eine positive Konstante sei.
- (ii) Der Zählprozess hat unabhängige Zuwächse, das heißt für jede endliche Menge von Zeitpunkten $0 = t_0 < t_1 < t_2 < \dots < t_k$, $k \in \mathbb{N}$, sind die Zuwächse

$$N_{t_j} - N_{t_{j-1}}, \quad 1 \leq j \leq k$$

stochastisch unabhängig.

Dann heißt der Prozess N ein *homogener Poisson-Prozess mit Parameter λ* .

Wir betrachten nun ein simplifiziertes Beispiel aus den Ingenieurwissenschaften, für dessen Modellierung sich dieser Prozess als unzureichend erweist; die Motivation für dieses Beispiel entstammt dem in der Einleitung vorgestellten Experiment.

Dazu werde in einem Belastungsexperiment ein Verbund von Drahtseilen zyklischen Belastungen ausgesetzt. Die Zeitpunkte der Drahtbrüche² bilden die Realisierung eines Punktprozesses $(T_n)_{n \in \mathbb{N}}$. Wir sind nun daran interessiert, die Anzahl der Drahtbrüche zu einem gegebenen Zeitpunkt $t > 0$ nach Beginn des Experiments *vorherzusagen*. Ein mögliche Herangehensweise ist nun der Ansatz eines Poisson-Modells für den assoziierten Zählprozess $N = (N_t)_{t \geq 0}$:

Hierbei nehmen wir an, N sei ein homogener Poisson-Prozess mit unbekanntem Parameter

²In der Praxis wird zumeist nicht die *Zeit*, sondern die *Anzahl der Lastzyklen* gemessen. Des einfacheren Verständnisses halber werden wir nachfolgend dennoch von einer *zeitlichen* Komponente sprechen.

$\lambda_0 > 0$. Diesen Parameter schätzen wir aus den beobachteten Daten (einige Verfahren hierfür werden ausführlich in Snyder und Miller 1991, Abschnitt 2.4, diskutiert); die Kenntnis der Verteilung von N erlaubt uns dann das Aufstellen von Prognoseintervallen unter Berücksichtigung der Unsicherheit beim Schätzen von λ_0 .

Die Naivität dieses Vorgehens liegt in der Annahme, die mit den beobachteten Drahtbrüchen assoziierte Realisierung des Zählprozesses N entstamme einem Poisson-Prozess. Diese Annahme ist in der Praxis kaum haltbar:

Der Parameter λ eines Poisson-Prozesses N kann im obigen Kontext als (einheitenloses) *Intensitätsniveau* der Belastung interpretiert werden, gilt doch bekanntermaßen (siehe etwa Snyder und Miller 1991, S. 43):

$$E[N_t] = \lambda t, \quad t \geq 0, \quad (2.1.2)$$

das heißt die Anzahl der erwarteten Drahtbrüche zum Zeitpunkt t nimmt mit wachsendem λ zu. Hierbei fallen uns sogleich *zwei* wesentliche Kritikpunkte ins Auge:

1. Obgleich die Anzahl der möglichen Drahtbrüche durch die Gesamtanzahl der Drahtseile im Verbund beschränkt ist, ist der Erwartungswert in Gleichung (2.1.2) für $t \rightarrow \infty$ unbeschränkt.
2. Bricht ein Drahtseil aus dem Verbund, so wird die Gesamtlast auf die verbleibenden Drahtseile umverteilt. Je mehr Drahtbrüche also erfolgen, desto mehr Belastung wird auf die verbleibenden Drahtseile ausgeübt, die somit erwartungsgemäß schneller brechen. Dieser Umstand wird in Gleichung (2.1.2) nicht berücksichtigt und ist mit der Annahme *unabhängiger Zuwächse* nicht vereinbar.

Das Phänomen, das im zweiten Kritikpunkt angesprochen wurde, bezeichnen wir suggestiv als *Lastumverteilung*.

Wir möchten eine Klasse von Zählprozessen einführen, für die das *Intensitätsniveau* λ nicht konstant ist, sondern von der Realisierung des Zählprozesses selbst abhängen darf, um die Lastumverteilung modellieren zu können. Heuristisch gesprochen ist ein solcher Zählprozess *selbstanregend* im Sinne des Wortlauts, da vergangene Sprünge das Auftreten weiterer Sprünge beeinflussen können. Der sogenannte *verschobene Geburtsprozess* aus Definition 2.1.9 genügt dieser Interpretation.

Bevor wir jedoch den *verschobenen Geburtsprozess* kennenlernen, bedarf es einer Definition, deren Notwendigkeit aus dem Verzicht auf die Bedingung (iii) in der Definition eines Punktprozesses erwächst (siehe Bemerkung 2.1.4). Wie in Bemerkung 2.1.6 erwähnt, erlauben wir definitorisch zunächst Zählprozesse mit Mehrfachsprüngen. Nun jedoch führen wir den Begriff des *bedingt geordneten Zählprozesses* ein, der die Beachtung von Mehrfachsprüngen obsolet macht.

Definition 2.1.8. (*Bedingt geordnete Zählprozesse*, vgl. Snyder und Miller 1991, S. 45)

Es sei $t_0 \in \mathbb{R}$, $(N_t)_{t \geq t_0}$ bezeichne einen Zählprozess und $\sigma(\{N_s : t_0 \leq s < t\})$ die von den Zufallsvariablen $\{N_s : t_0 \leq s < t\}$ erzeugte σ -Algebra der *Vergangenheit des Zählprozesses*. Der Zählprozess $(N_t)_{t \geq t_0}$ heißt dann *bedingt geordnet zum Zeitpunkt $t \geq t_0$* , falls für alle Ereignisse $A \in \sigma(\{N_s : t_0 \leq s < t\})$ und $\varepsilon > 0$ ein $\widetilde{\Delta t} = \widetilde{\Delta t}(A, \varepsilon) > 0$ existiert, so dass für alle $\Delta t < \widetilde{\Delta t}$ gilt:

$$P(\{N_{t+\Delta t} - N_t > 1\} | A) \leq \varepsilon P(\{N_{t+\Delta t} - N_t = 1\} | A). \quad (2.1.3)$$

Ist die Bedingung aus Gleichung (2.1.3) für jeden Zeitpunkt $t \geq t_0$ erfüllt, so heißt der Zählprozess $(N_t)_{t \geq t_0}$ *bedingt geordnet*.

Ist ein Zählprozess $(N_t)_{t \geq t_0}$ bedingt geordnet, so bedeutet dies anschaulich, dass gegeben irgendein Ereignis A , das nur von der Vergangenheit des Prozesses abhängt (z.B. dass $N_s > 1$ ist für ein $t_0 \leq s < t$, und dass der erste Sprung zum Zeitpunkt $s_0 \leq s$ eingetreten ist), die Wahrscheinlichkeit für mehr als einen Sprung in einem - hinreichend klein gewählten - Intervall ein beliebig kleiner Bruchteil der Wahrscheinlichkeit für genau einen Sprung in diesem Intervall ist.

Qualitativ bedeutet dies, dass Sprünge nicht simultan auftreten, sondern gewissermaßen *geordnet* (s. Snyder und Miller 1991, S. 45).

Wir sind nun in der Lage, die Definition eines *verschobenen Geburtsprozesses* anzugeben.

Definition 2.1.9. (*Verschobener Geburtsprozess & Intensitätsfunktion*, vgl. Snyder und Miller 1991, S. 95)

Es sei wiederum $t_0 \in \mathbb{R}$. $(N_t)_{t \geq t_0}$ bezeichne einen Zählprozess, der die nachfolgenden Annahmen erfülle:

1. $(N_t)_{t \geq t_0}$ sei bedingt geordnet (siehe Definition 2.1.8).
2. Für alle $t \geq t_0$ und $A \in \sigma(\{N_s : t_0 \leq s < t\})$ existiere der Grenzwert von $\frac{1}{\Delta t} P(\{N(t, t + \Delta t) = 1\} | A)$ für $\Delta t \rightarrow 0$, der nur von t und N_t abhängt und mit $\lambda(t, N_t)$ bezeichnet werde. Für alle $n \in \mathbb{N}$ sei die Funktion λ_n mit $\lambda_n(t) = \lambda(t, n)$ beschränkt und bezüglich t für alle $t^0 > t_0$ über dem Intervall $[t_0, t^0]$ integrierbar.
3. $P(\{N_{t_0} = n\}) = \delta_{0n}$.

Dabei gelte $N(s, t) = N_t - N_s$, δ_{ij} bezeichne das Kronecker-Delta und wir vereinbaren die Notation $P_n(t) = P(\{N_t = n\})$. Wir nennen einen Zählprozess, der die obigen Bedingungen erfüllt, einen *verschobenen Geburtsprozess*. Die Funktion $\lambda(t, N_t)$ bezeichnen wir fortan als *Intensitätsfunktion*.

Bemerkung 2.1.10.

Um Verwirrung vorzubeugen, weisen wir an dieser Stelle auf zwei Aspekte gesondert hin:

1. Die Definition des verschobenen Geburtsprozesses aus Definition 2.1.9 stimmt nicht mit der üblichen Definition eines Geburtsprozesses überein, für den

$$P(\{N_{t_0} = n\}) = \delta_{1n}$$

anstelle der obigen dritten Bedingung gefordert wird. Dies ist durch den Umstand motiviert, dass der Geburtsprozess der Namensgebung entsprechend das Wachstum einer Population beschreibe, die jedoch in der Situation von Definition 2.1.9 gar nicht existieren würde.

2. In der Definition der Intensitätsfunktion λ bedingen wir auf Ereignisse aus der σ -Algebra $\sigma(\{N_s : t_0 \leq s < t\})$, weshalb λ bisweilen als *bedingte Intensitätsfunktion* bezeichnet wird. Dass λ nicht von diesem konkreten Ereignis abhängen soll und somit *unabhängig* von der Vergangenheit des Zählprozesses N ist, geht ebenfalls mit der Idee eines Geburtsprozesses konform:

Die Veränderung einer Population hängt *nur* von der aktuellen Population, nicht jedoch ihrer Historie ab.

Wir können den in Beispiel 2.1.7 definierten homogenen Poisson-Prozess als Spezialfall eines verschobenen Geburtsprozesses auffassen:

Ist die Intensitätsfunktion λ des Zählprozesses $(N_t)_{t \geq t_0}$ aus Definition 2.1.9 eine deterministische Funktion der Zeit t und hängt folglich *nicht* von der Realisierung von N ab, so nennt man den Zählprozess einen *inhomogenen Poisson-Prozess* (siehe Snyder und Miller 1991, S. 46). Ist λ zudem auch bezüglich t konstant, so erhält man gerade den oben eingeführten homogenen Poisson-Prozess mit Parameter λ .

Beispiel 2.1.7. (1. Fortsetzung)

Wie aus der zweiten Bedingung aus 2.1.9 und dem zweiten Punkt der Bemerkung 2.1.10 ersichtlich wird, eignet sich ein verschobener Geburtsprozess anschaulich zur Modellierung der Lastumverteilung. Dass wir zugleich auch das Problem des *unbeschränkten Erwartungswertes*, welches im gleichen Rahmen thematisiert wurde, beseitigen können, werden wir in der nächsten Fortsetzung des Beispiels erkennen.

Bevor wir uns genauer mit dem verschobenen Geburtsprozess auseinandersetzen werden, formulieren wir zunächst ein Lemma, dessen Nutzen uns erst während des Beweises des nachfolgenden Satzes gewahr werden wird.

Lemma 2.1.11.

Es sei $(N_t)_{t \geq t_0}$ ein verschobener Geburtsprozess. Dann gilt für alle $A \in \sigma(N_s : t_0 \leq s < t)$:

$$\lim_{\Delta t \rightarrow 0} P(\{N(t, t + \Delta t) \geq 1\} | A) = 0. \quad (2.1.4)$$

Beweis³:

Da der Zählprozess nach der ersten Bedingung aus Definition 2.1.9 bedingt geordnet ist, existiert zu $\varepsilon > 0$ ein $\widetilde{\Delta t} = \widetilde{\Delta t}(\varepsilon) > 0$ so, dass für alle $\Delta t < \widetilde{\Delta t}$ gilt:

$$\begin{aligned} P(\{N(t, t + \Delta t) \geq 1\} | A) &= P(\{N(t, t + \Delta t) = 1\} | A) + P(\{N(t, t + \Delta t) > 1\} | A) \\ &\leq (1 + \varepsilon) P(\{N(t, t + \Delta t) = 1\} | A) \end{aligned} \quad (2.1.5)$$

Nach der zweiten Bedingung aus Definition 2.1.9 ist nun die rechte Seite von Gleichung (2.1.5) ein $\mathcal{O}(\Delta t)$ für $\Delta t \rightarrow 0$. Damit folgt für $\Delta t \rightarrow 0$ sofort die Behauptung. \square

Unter Verwendung dieses Lemmas können wir nun den nachfolgenden Satz für verschobene Geburtsprozesse beweisen.

Satz 2.1.12.

Es sei $(N_t)_{t \geq t_0}$ ein verschobener Geburtsprozess. Dann erfüllt $P_n(t) = P(\{N_t = n\})$ für alle $n \in \mathbb{N}$ die Differentialgleichung

$$\begin{cases} \frac{d}{dt} P_n(t) = -\lambda_n(t) P_n(t) + \lambda_{n-1}(t) P_{n-1}(t), \\ P_n(t_0) = \delta_{0n}. \end{cases} \quad (2.1.6)$$

Die Gleichung bleibt auch für $n = 0$ gültig, wenn wir $\lambda_{-1} \equiv 0$ und $P_{-1} \equiv 0$ setzen.

Beweis⁴:

Wir beachten zunächst die Gültigkeit der nachstehenden Gleichung unter Verwendung der Definition der bedingten Wahrscheinlichkeit:

$$\begin{aligned} P_n(t + \Delta t) &= P(\{N_{t+\Delta t} = n\}) = \sum_{k=0}^n P(\{N_t = n - k, N(t, t + \Delta t) = k\}) \\ &= P(\{N(t, t + \Delta t) = 0\} | \{N_t = n\}) P(\{N_t = n\}) \\ &\quad + P(\{N(t, t + \Delta t) = 1\} | \{N_t = n - 1\}) P(\{N_t = n - 1\}) \\ &\quad + \sum_{k=2}^n P(\{N(t, t + \Delta t) = k\} | \{N_t = n - k\}) P(\{N_t = n - k\}). \end{aligned} \quad (2.1.7)$$

³Dieser Beweis wurde vom Autor dieser Arbeit selbst geführt.

⁴Der Beweis geht im Wesentlichen zurück auf *Snyder und Miller*, siehe Snyder und Miller 1991, S. 47ff. wird dort jedoch nur für den Spezialfall, dass die Intensitätsfunktion *nicht* von n abhängt, gezeigt.

Dabei fallen im Falle $n = 0$ die letzten beiden und im Falle $n = 1$ der letzte Term weg. Wir betrachten den Fall $n \geq 2$.

Im Hinblick auf Lemma 2.1.11 ist es uns fortan ohne Einschränkung möglich, auf Ereignisse aus $\sigma(\{N_s : t_0 \leq s \leq t\})$ anstelle der größeren σ -Algebra $\sigma(\{N_s : t_0 \leq s < t\})$ zu bedingen, da ein Sprung zum Zeitpunkt t mit Wahrscheinlichkeit 0 auftritt. Man beachte, dass sich hierfür das Zusammenwirken der ersten *beiden* Bedingungen aus der Definition des verschobenen Geburtsprozesses als notwendig erweist.

Nun existiert folglich zu gegebenem $\varepsilon > 0$ ein $\widetilde{\Delta t} > 0$, so dass für alle $\Delta t < \widetilde{\Delta t}$ und $k = 0, 1, \dots, n$ gilt⁵:

$$P(\{N(t, t + \Delta t) > 1\} | \{N_t = n - k\}) \leq \varepsilon P(\{N(t, t + \Delta t) = 1\} | \{N_t = n - k\}). \quad (2.1.8)$$

Die Definition der Intensitätsfunktion $\lambda_n(t)$ liefert für Δt hinreichend klein bei erneuter Berücksichtigung der Aussage von Lemma 2.1.11 zudem⁶:

$$\left| \frac{1}{\Delta t} P(\{N(t, t + \Delta t) = 1\} | \{N_t = n\}) - \lambda_n(t) \right| \leq \varepsilon. \quad (2.1.9)$$

Unter Verwendung der Abschätzungen (2.1.8) und (2.1.9) erhalten wir somit für den letzten Term aus Gleichung (2.1.7):

$$\begin{aligned} & \sum_{k=2}^n P(\{N(t, t + \Delta t) = k\} | \{N_t = n - k\}) P(\{N_t = n - k\}) \\ & \leq \sum_{k=2}^n P(\{N(t, t + \Delta t) > 1\} | \{N_t = n - k\}) P(\{N_t = n - k\}) \\ & \leq \varepsilon \sum_{k=2}^n P(\{N(t, t + \Delta t) = 1\} | \{N_t = n - k\}) P(\{N_t = n - k\}) \\ & \leq \varepsilon \sum_{k=2}^n \Delta t (\lambda_{n-k}(t) + \varepsilon) P(\{N_t = n - k\}) \\ & \leq \varepsilon \Delta t \left(\max_{k=0,1,\dots,n-2} \lambda_k(t) + \varepsilon \right). \end{aligned} \quad (2.1.10)$$

Unter Beachtung der Gleichung

$$\begin{aligned} P(\{N(t, t + \Delta t) = 0\} | \{N_t = n\}) &= 1 - P(\{N(t, t + \Delta t) = 1\} | \{N_t = n\}) \\ &\quad - P(\{N(t, t + \Delta t) > 1\} | \{N_t = n\}), \end{aligned}$$

⁵Man wähle dazu für jedes $k = 0, 1, \dots, n$ ein solches $\widetilde{\Delta t}$ und gehe anschließend zum Minimum über.

⁶Man beachte, dass mit $N_t = n$ definitionsgemäß $\lambda(t, N_t) = \lambda(t, n) = \lambda_n(t)$ gilt.

erhalten wir ferner bei erneuter Verwendung der Gleichungen (2.1.8) und (2.1.9) für alle $n \in \mathbb{N}_0$:

$$\begin{aligned}
& P(\{N(t, t + \Delta t) = 0\} | \{N_t = n\}) \\
& \leq 1 - P(\{N(t, t + \Delta t) = 1\} | \{N_t = n\}) \\
& \leq 1 - \Delta t (\lambda_n(t) - \varepsilon) ,
\end{aligned} \tag{2.1.11}$$

$$\begin{aligned}
& P(\{N(t, t + \Delta t) = 0\} | \{N_t = n\}) \\
& \geq 1 - (1 + \varepsilon) P(\{N(t, t + \Delta t) = 1\} | \{N_t = n\}) \\
& \geq 1 - (1 + \varepsilon) \Delta t (\lambda_n(t) + \varepsilon) .
\end{aligned} \tag{2.1.12}$$

Einsetzen der Ungleichungen (2.1.9) bis (2.1.11) in (2.1.7) liefert schließlich für $n \geq 2$:

$$\begin{aligned}
& \frac{1}{\Delta t} (P(\{N_{t+\Delta t} = n\}) - P(\{N_t = n\})) \\
& \quad + \lambda_n(t) P(\{N_t = n\}) - \lambda_{n-1}(t) P(\{N_t = n-1\}) \\
& = \frac{1}{\Delta t} \left((P(\{N(t, t + \Delta t) = 0\} | \{N_t = n\}) - 1) P(\{N_t = n\}) \right. \\
& \quad + P(\{N(t, t + \Delta t) = 1\} | \{N_t = n-1\}) P(\{N_t = n-1\}) \\
& \quad + \sum_{k=2}^n P(\{N(t, t + \Delta t) = k\} | \{N_t = n-k\}) P(\{N_t = n-k\}) \left. \right) \\
& \quad + \lambda_n(t) P(\{N_t = n\}) - \lambda_{n-1}(t) P(\{N_t = n-1\}) \\
& \leq \frac{1}{\Delta t} \left(-\Delta t (\lambda_n(t) - \varepsilon) P(\{N_t = n\}) + \Delta t (\varepsilon + \lambda_{n-1}(t)) P(\{N_t = n-1\}) \right. \\
& \quad \left. + \varepsilon \Delta t \left(\max_{k=0,1,\dots,n-2} \lambda_k(t) + \varepsilon \right) \right) + \lambda_n(t) P(\{N_t = n\}) - \lambda_{n-1}(t) P(\{N_t = n-1\}) \\
& = (\varepsilon - \lambda_n(t)) P(\{N_t = n\}) + (\varepsilon + \lambda_{n-1}(t)) P(\{N_t = n-1\}) \\
& \quad + \varepsilon \left(\max_{k=0,1,\dots,n-2} \lambda_k(t) + \varepsilon \right) + \lambda_n(t) P(\{N_t = n\}) - \lambda_{n-1}(t) P(\{N_t = n-1\}) \\
& = \varepsilon \left(P(\{N_t = n\}) + P(\{N_t = n-1\}) + \varepsilon + \max_{k=0,1,\dots,n-2} \lambda_k(t) \right) \\
& \leq \varepsilon \left(2 + \varepsilon + \max_{k=0,1,\dots,n-2} \lambda_k(t) \right) ,
\end{aligned} \tag{2.1.13}$$

Völlig analog dazu erhält man mit Ungleichung (2.1.12) anstelle von (2.1.11):

$$\begin{aligned}
& \frac{1}{\Delta t} (P(\{N_{t+\Delta t} = n\}) - P(\{N_t = n\})) \\
& \quad + \lambda_n(t)P(\{N_t = n\}) - \lambda_{n-1}(t)P(\{N_t = n-1\}) \\
& \geq (1 + \varepsilon)(-\varepsilon - \lambda_n(t))P(\{N_t = n\}) + (-\varepsilon + \lambda_{n-1}(t))P(\{N_t = n-1\}) \\
& \quad + \lambda_n(t)P(\{N_t = n\}) - \lambda_{n-1}(t)P(\{N_t = n-1\}) \\
& = -\varepsilon((1 + \varepsilon + \lambda_n(t))P(\{N_t = n\}) + P(\{N_t = n-1\})) \\
& \geq -\varepsilon(2 + \varepsilon + \lambda_n(t)) . \tag{2.1.14}
\end{aligned}$$

Durch Kombination der Abschätzungen aus (2.1.13) und (2.1.14) erhalten wir damit:

$$\begin{aligned}
& \left| \frac{1}{\Delta t} (P(\{N_{t+\Delta t} = n\}) - P(\{N_t = n\})) + \lambda_n(t)P(\{N_t = n\}) - \lambda_{n-1}(t)P(\{N_t = n-1\}) \right| \\
& \leq \varepsilon \left(2 + \varepsilon + \max_{k=0,1,\dots,n} \lambda_k(t) \right) . \tag{2.1.15}
\end{aligned}$$

Aus der Beschränktheit der λ_k bezüglich t und der Beliebigkeit von ε folgt somit die Gültigkeit der zu zeigenden Differentialgleichung für alle $n \geq 2$, da Gleichung (2.1.15) für alle $\Delta t < \widetilde{\Delta t}$ (also insbesondere für $\Delta t \rightarrow 0$) gilt.

Die Abschätzungen in (2.1.13) und (2.1.14) bleiben auch im Falle $n = 1$ gültig (es fällt lediglich die Summe weg, die zuvor gegen $\varepsilon \left(\max_{k=0,1,\dots,n-2} \lambda_k(t) + \varepsilon \right)$ abgeschätzt wurde; dieser Summand fehlt folglich in der Abschätzung), so dass die Differentialgleichung für alle $n \in \mathbb{N}$ gilt.

Für $n = 0$ schließlich folgt die Behauptung bereits unmittelbar aus (2.1.11) und (2.1.12) mit $k = n$:

$$\begin{aligned}
& \frac{1}{\Delta t} (P(\{N_{t+\Delta t} = 0\}) - P(\{N_t = 0\})) \\
& \quad + \lambda_0(t)P(\{N_t = 0\}) \\
& \leq (\varepsilon - \lambda_0(t))P(\{N_t = 0\}) + \lambda_0(t)P(\{N_t = 0\}) \leq \varepsilon ,
\end{aligned}$$

$$\begin{aligned}
& \frac{1}{\Delta t} (P(\{N_{t+\Delta t} = 0\}) - P(\{N_t = 0\})) \\
& \geq -(1 + \varepsilon)(\lambda_0(t) + \varepsilon)P(\{N_t = 0\}) + \lambda_0(t)P(\{N_t = 0\}) \\
& \geq -\varepsilon(1 + \varepsilon + \lambda_0(t))P(\{N_t = 0\}) \\
& \geq -\varepsilon(1 + \varepsilon + \lambda_0(t)) .
\end{aligned}$$

Die Behauptung folgt dann wie zuvor leicht durch Kombination dieser Abschätzungen.

Die Anfangswerte letztlich folgen unmittelbar aus der Definition des verschobenen Geburtsprozesses. Dies vervollständigt den Beweis. \square

2.1.4 Spezialfall: Reiner verschobener Geburtsprozess mit Abbruch

Wir betrachten nun einen speziellen verschobenen Geburtsprozess, den wir nachfolgend definieren möchten:

Definition 2.1.13. (*Reiner verschobener Geburtsprozess mit Abbruch*)

Es bezeichne $N = (N_t)_{t \geq t_0}$ einen verschobenen Geburtsprozess. Ferner betrachten wir die nachfolgenden Bedingungen:

Zeithomogenität: Die Intensitätsfunktion $\lambda_n(t)$ ist nur eine Funktion von $n = N_t$ und hängt darüber hinaus nicht von t ab.

Abbruch: Der Zählprozess N ist durch ein $M \in \mathbb{N}_0$ beschränkt.

Erfüllt N die Bedingung der Zeithomogenität, so nennen wir N einen *reinen verschobenen Geburtsprozess*.

Ist ferner auch die zweite obige Bedingung erfüllt, so heißt N ein *reiner verschobener Geburtsprozess mit Abbruch bei M* .

Um die Notation zu vereinfachen, sei fortan $t_0 = 0$ (den Fall $t_0 \neq 0$ erhält man dann durch eine einfache Verschiebung).

Betrachten wir einen reinen verschobenen Geburtsprozess mit Abbruch bei M , so gilt für die zugehörige Intensitätsfunktion, dass

$$\lambda_n(t) = g(n) \cdot \mathbb{1}_{\{n < M\}}$$

mit einer Funktion $g : \mathbb{N}_0 \rightarrow \mathbb{R}$ ist.

Beispiel 2.1.7. (*2. Fortsetzung*)

Besteht der Drahtseilverbund aus M einzelnen Drahtseilen und bezeichne N wiederum den mit den Zeitpunkten der Drahtbrüche assoziierten Zählprozess, so ist N durch M beschränkt.

Für einen Zählprozess aus der Klasse der *reinen verschobenen Geburtsprozesse mit Abbruch* gilt tatsächlich $P(\{N_t = M + k\}) \equiv 0$ für alle $k \in \mathbb{N}$, wie wir in Satz 2.1.14 zeigen werden. Derartige Zählprozesse eignen sich folglich besser als ein homogener Poisson-Prozess, um den Zusammenhang aus dem Belastungsexperiment zu beschreiben.

Mittels der Differentialgleichung für P_n aus Satz 2.1.12 ist es nun sogar möglich, die Verteilung von N_t für alle $t \geq 0$ zu bestimmen. Das Ergebnis fasst der nächste Satz zusammen.

Satz 2.1.14.

Es sei $(N_t)_{t \geq 0}$ ein reiner verschobener Geburtsprozess mit Abbruch bei $M \in \mathbb{N}$. Ferner sei die Intensitätsfunktion λ_n streng monoton steigend in n . Dann gilt:

$$\begin{aligned}
P_0(t) &= e^{-\lambda_0 t}, \\
P_n(t) &= \lambda_0 \cdots \lambda_{n-1} \sum_{k=0}^n \left[\prod_{\substack{j=0, \\ j \neq k}}^n (\lambda_j - \lambda_k) \right]^{-1} e^{-\lambda_k t}, \quad n = 1, \dots, M-1, \\
P_M(t) &= \lambda_0 \cdots \lambda_{M-1} \sum_{k=0}^{M-1} \frac{1}{\lambda_k} \left[\prod_{\substack{j=0, \\ j \neq k}}^{M-1} (\lambda_j - \lambda_k) \right]^{-1} (1 - e^{-\lambda_k t}), \\
P_{M+k} &\equiv 0, \quad k \in \mathbb{N}.
\end{aligned}$$

Beweis⁷:

Wir zeigen für $n = 0, 1, \dots, M$, dass die so definierten P_n die Differentialgleichung (2.1.6) aus Satz 2.1.12 erfüllen. Aus den Voraussetzungen folgt, dass $\lambda_i \neq \lambda_j$ für $i \neq j$ mit Wahrscheinlichkeit 1 gilt.

Für $n = 0$ liefert Differentiation und Einsetzen von $t = 0$ sofort die Behauptung.

Für $n = 1$ erhalten wir mittels einer Nullergänzung die Behauptung:

$$\begin{aligned}
\frac{d}{dt} P_1(t) &= \frac{d}{dt} \left[\lambda_0 \left(\frac{1}{\lambda_1 - \lambda_0} e^{-\lambda_0 t} - \frac{1}{\lambda_1 - \lambda_0} e^{-\lambda_1 t} \right) \right] = \frac{d}{dt} \left[\frac{\lambda_0}{\lambda_1 - \lambda_0} (e^{-\lambda_0 t} - e^{-\lambda_1 t}) \right] \\
&= \frac{\lambda_0}{\lambda_1 - \lambda_0} (\lambda_1 e^{-\lambda_1 t} - \lambda_0 e^{-\lambda_0 t}) = \frac{\lambda_0}{\lambda_1 - \lambda_0} (\lambda_1 (-e^{-\lambda_0 t} + e^{-\lambda_1 t} + e^{-\lambda_0 t}) - \lambda_0 e^{-\lambda_0 t}) \\
&= \lambda_0 e^{-\lambda_0 t} - \frac{\lambda_0 \lambda_1}{\lambda_1 - \lambda_0} (e^{-\lambda_0 t} - e^{-\lambda_1 t}) = -\lambda_1 P_1(t) + \lambda_0 P_0(t). \tag{2.1.16}
\end{aligned}$$

Aus den Umformungen innerhalb der Gleichungskette (2.1.16) wird zudem unmittelbar klar, dass $P_1(0) = 0$ gilt.

Es sei also $n \in \{2, \dots, M-1\}$. Dann gilt erneut unter Berücksichtigung der Linearität des Differentialoperators $\frac{d}{dt}$:

⁷Dieser Satz wurde vom Autor dieser Arbeit gefunden und bewiesen.

$$\begin{aligned}
\frac{d}{dt}P_n(t) &= \lambda_0 \cdots \lambda_{n-1} \sum_{k=0}^n -\lambda_k \left[\prod_{\substack{j=0, \\ j \neq k}}^n (\lambda_j - \lambda_k) \right]^{-1} e^{-\lambda_k t} \\
&= \lambda_0 \cdots \lambda_{n-1} \sum_{k=0}^n (\lambda_n - \lambda_k - \lambda_n) \left[\prod_{\substack{j=0, \\ j \neq k}}^n (\lambda_j - \lambda_k) \right]^{-1} e^{-\lambda_k t} \\
&= -\lambda_n P_n(t) + \lambda_0 \cdots \lambda_{n-1} \sum_{k=0}^n (\lambda_n - \lambda_k) \left[\prod_{\substack{j=0, \\ j \neq k}}^n (\lambda_j - \lambda_k) \right]^{-1} e^{-\lambda_k t} \\
&= -\lambda_n P_n(t) + \lambda_0 \cdots \lambda_{n-1} \sum_{k=0}^{n-1} \left[\prod_{\substack{j=0, \\ j \neq k}}^{n-1} (\lambda_j - \lambda_k) \right]^{-1} e^{-\lambda_k t} \\
&= -\lambda_n P_n(t) + \lambda_{n-1} P_{n-1}(t).
\end{aligned}$$

Es bleibt zu zeigen, dass zudem die Bedingung $P_n(0) = 0$ erfüllt ist.

Können wir dazu zeigen, dass die nachfolgende Gleichheit gilt, so folgt dies bereits durch Einsetzen von $t = 0$.

$$\sum_{k=0}^{n-1} \left[\prod_{\substack{j=0, \\ j \neq k}}^n (\lambda_j - \lambda_k) \right]^{-1} = - \left[\prod_{j=0}^{n-1} (\lambda_j - \lambda_n) \right]^{-1}. \quad (2.1.17)$$

Im Folgenden beweisen wir also Gleichung (2.1.17).

Dazu beobachten wir zunächst, dass mit $\lambda_i \neq \lambda_j$ für $i \neq j$ gilt:

$$\begin{aligned}
\sum_{k=0}^{n-1} \left[\prod_{\substack{j=0, \\ j \neq k}}^n (\lambda_j - \lambda_k) \right]^{-1} &= - \left[\prod_{j=0}^{n-1} (\lambda_j - \lambda_n) \right]^{-1} \\
\iff \sum_{k=0}^{n-1} \prod_{\substack{j=0, \\ j \neq k}}^{n-1} \frac{\lambda_j - \lambda_n}{\lambda_j - \lambda_k} &= 1 \\
\iff -1 + \sum_{k=0}^{n-1} \prod_{\substack{j=0, \\ j \neq k}}^{n-1} \frac{\lambda_j - \lambda_n}{\lambda_j - \lambda_k} &= 0. \quad (2.1.18)
\end{aligned}$$

Motiviert durch Gleichung (2.1.18) definieren wir das Polynom $Q(x)$, indem wir λ_n durch

x ersetzen:

$$Q(x) = -1 + \sum_{k=0}^{n-1} \prod_{\substack{j=0, \\ j \neq k}}^{n-1} \frac{\lambda_j - x}{\lambda_j - \lambda_k}.$$

Offensichtlich beträgt der Grad von Q maximal $n-1$, da in jedem Summanden ein Produkt von $n-1$ Linearfaktoren vorliegt. Andererseits überzeugt man sich leicht, dass $Q(\lambda_l) = 0$ für $l = 0, 1, \dots, n-1$ gilt:

- Ist $k \neq l$, so ist

$$\prod_{\substack{j=0, \\ j \neq k}}^{n-1} \frac{\lambda_j - \lambda_l}{\lambda_j - \lambda_k} = 0,$$

da schließlich für $j = l$ der Faktor 0 auftritt.

- Ist hingegen $k=l$, so folgt

$$\prod_{\substack{j=0, \\ j \neq k}}^{n-1} \frac{\lambda_j - \lambda_l}{\lambda_j - \lambda_k} = \prod_{\substack{j=0, \\ j \neq k}}^{n-1} 1 = 1.$$

- Insgesamt erhält man folglich:

$$Q(\lambda_l) = -1 + \sum_{k=0}^{n-1} \delta_{kl} = -1 + 1 = 0.$$

Da Q somit mindestens n Nullstellen aufweist, muss nach dem Fundamentalsatz der Algebra bereits $Q \equiv 0$ gelten. Mit (2.1.18) folgt ergo die Gültigkeit von Gleichung (2.1.17) und deshalb $P_n(0) = 0$ für $n = 2, \dots, M-1$.

Abschließend betrachten wir noch den Fall $n = M$; die Fälle $n = M + k$ für $k \in \mathbb{N}$ sind dann trivial.

$P_M(0) = 0$ ist offensichtlich erfüllt. Differentiation nach t liefert wegen $\lambda_M = 0$ erneut:

$$\begin{aligned} \frac{d}{dt} P_M(t) &= \lambda_0 \cdot \dots \cdot \lambda_{M-1} \sum_{k=0}^{M-1} \left[\prod_{\substack{j=0, \\ j \neq k}}^{M-1} (\lambda_j - \lambda_k) \right]^{-1} e^{-\lambda_k t} \\ &= \lambda_{M-1} P_{M-1}(t) = -\lambda_M P_M(t) + \lambda_{M-1} P_{M-1}(t). \end{aligned}$$

Dies vervollständigt den Beweis von Satz 2.1.14. □

Bemerkung 2.1.15.

1. In Satz 2.1.14 wird gefordert, dass die Intensitätsfunktion λ_n streng monoton steigend in n ist. Der Beweis des Satzes zeigt, dass wir tatsächlich nur fordern müssen, dass $\lambda_i \neq \lambda_j$ für $i \neq j$ gilt.
2. Das Resultat aus Satz 2.1.14 zeigt insbesondere, dass die Verteilung von N_t für ein festes $t \geq 0$ vollständig durch die Intensitätsfunktion λ_n charakterisiert werden kann. Infolgedessen erscheint es natürlich, die Modellierung eines solchen Zählprozesses auf Ebene der Intensitätsfunktion vorzunehmen.
3. Da das $M \in \mathbb{N}$ aus Satz 2.1.14 prinzipiell beliebig gewählt werden kann, liefert Satz 2.1.14 auch die Verteilung für N_t im Falle eines reinen verschobenen Geburtsprozesses *ohne* Abbruch.

Unter Verwendung von Satz 2.1.14 können wir zudem den gewöhnlichen *homogenen Poisson-Prozess* gewissermaßen als Grenzfall eines reinen verschobenen Geburtsprozesses auffassen. Diese Aussage soll im nachfolgenden Satz formuliert werden.

Satz 2.1.16.

Es sei $(N_t^{(h)})_{t \geq 0}$ ein reiner verschobener Geburtsprozess, der die Intensitätsfunktion $\lambda_n^{(h)} = \lambda_0 + nh$ für ein $\lambda_0 > 0$ aufweist. Dann konvergiert die Verteilung von $N_t^{(h)}$ für $h \rightarrow 0$ gegen eine Poisson-Verteilung mit Parameter $\lambda_0 t$.

Beweis⁸:

Wir bemerken zunächst, dass folgende Identität gilt:

$$\prod_{\substack{j=0, \\ j \neq k}}^n \frac{1}{j-k} = \frac{(-1)^k}{k!(n-k)!}, \quad k = 0, 1, \dots, n. \quad (2.1.19)$$

Dies sieht man wie folgt ein:

- Für $k = 0$ ist die Behauptung unmittelbar klar, da gilt:

$$\prod_{\substack{j=0, \\ j \neq 0}}^n \frac{1}{j-0} = \prod_{j=1}^n \frac{1}{j} = \frac{1}{1 \cdot \dots \cdot n} = \frac{1}{n!}.$$

- Für $k = n$ folgt die Behauptung wie für $k = 0$, indem jeder Faktor mit -1 multipliziert wird:

$$\prod_{\substack{j=0, \\ j \neq n}}^n \frac{1}{j-n} = (-1)^n \prod_{j=0}^{n-1} \frac{1}{n-j} = \frac{(-1)^n}{n \cdot \dots \cdot 1} = \frac{(-1)^n}{n!}.$$

⁸Dieser Satz wurde vom Autor dieser Arbeit gefunden und bewiesen.

- Für $k = 1, \dots, n-1$ kombinieren wir die obigen Ansätze und erhalten:

$$\begin{aligned} \prod_{\substack{j=0, \\ j \neq k}}^n \frac{1}{j-k} &= \prod_{j=0}^{k-1} \frac{1}{j-k} \cdot \prod_{j=k+1}^n \frac{1}{j-k} = (-1)^k \prod_{j=0}^{k-1} \frac{1}{k-j} \cdot \prod_{j=k+1}^n \frac{1}{j-k} \\ &= (-1)^k \frac{1}{k \cdot \dots \cdot 1} \cdot \frac{1}{1 \cdot \dots \cdot (n-k)} = \frac{(-1)^k}{k! (n-k)!}. \end{aligned}$$

Setzen wir nun die obige Intensitätsfunktion in die Darstellung der P_n für $n \geq 1$ aus Satz 2.1.14 ein, so erhalten wir mit Gleichung (2.1.19) und der verallgemeinerten binomischen Formel:

$$\begin{aligned} P\left(\left\{N_t^{(h)} = n\right\}\right) &= (\lambda_0 (\lambda_0 + h) \cdot \dots \cdot (\lambda_0 + (n-1)h)) \sum_{k=0}^n \left[\prod_{\substack{j=0, \\ j \neq k}}^n \frac{1}{(j-k)h} \right] e^{-(\lambda_0 + kh)t} \\ &= (\lambda_0^n + \mathcal{O}(h)) \sum_{k=0}^n \frac{(-1)^k}{k! (n-k)!} \frac{1}{h^n} e^{-(\lambda_0 + kh)t} \\ &= \frac{(-1)^n (\lambda_0^n + \mathcal{O}(h))}{n!} e^{-\lambda_0 t} \sum_{k=0}^n \binom{n}{k} (-1)^{n-k} \frac{e^{-kht}}{h^n} \\ &= \frac{(-1)^n (\lambda_0^n + \mathcal{O}(h))}{n!} e^{-\lambda_0 t} \sum_{k=0}^n \binom{n}{k} \left(\frac{e^{-ht}}{h}\right)^k \left(\frac{-1}{h}\right)^{n-k} \\ &= \frac{(-1)^n (\lambda_0^n + \mathcal{O}(h))}{n!} e^{-\lambda_0 t} \left(\frac{e^{-ht} - 1}{h}\right)^n \\ &= \frac{(-1)^n (\lambda_0^n + \mathcal{O}(h))}{n!} e^{-\lambda_0 t} \left(\frac{d}{d\lambda} e^{-\lambda t} \Big|_{\lambda=0} + \mathcal{O}(h)\right)^n \\ &= \frac{(-1)^n (\lambda_0^n + \mathcal{O}(h))}{n!} e^{-\lambda_0 t} (-t + \mathcal{O}(h))^n \\ &= \frac{(-1)^n (\lambda_0^n + \mathcal{O}(h))}{n!} e^{-\lambda_0 t} ((-t)^n + \mathcal{O}(h)) \\ &= \frac{(\lambda_0 t)^n}{n!} e^{-\lambda_0 t} + \mathcal{O}(h) \\ &\rightarrow \frac{(\lambda_0 t)^n}{n!} e^{-\lambda_0 t} \quad (h \rightarrow 0). \end{aligned} \tag{2.1.20}$$

Gleichung (2.1.20) können wir nun entnehmen, dass die Verteilung von $N_t = \lim_{h \rightarrow 0} N_t^{(h)}$ für $t \geq t_0$ gerade die Poisson-Verteilung mit Parameter $\lambda_0 t$ ist, und dies zeigt die Behauptung. \square

Wir möchten fortan die sogenannte *kumulierte Intensitätsfunktion*, die im Falle eines verschobenen Geburtsprozesses durch

$$\Lambda(t) = \int_0^t \lambda_{N_s}(s) \, ds \quad (2.1.21)$$

gegeben ist, betrachten. Wir werden erkennen, dass sie wesentliche Informationen über den zugehörigen Zählprozess beinhaltet und im Rahmen der sogenannten *Doob-Meyer-Zerlegung für nicht-negative Submartingale* eine Verallgemeinerung kennenlernen.

Speziell interessiert uns nun der Erwartungswert der kumulierten Intensitätsfunktion und damit auch von N_t für $t \geq 0$, wie wir später sehen werden. Unter Verwendung eines Fubini-Arguments stellen wir dabei fest (vgl. auch Snyder und Miller 1991, S. 332):

$$\begin{aligned} \mathbb{E}[\Lambda(t)] &= \int_0^t \mathbb{E}[\lambda_{N_s}(s)] \, ds = \int_0^t \sum_{n \in \mathbb{N}_0} P(\{N_s = n\}) \lambda_n(s) \, ds \\ &= \sum_{n \in \mathbb{N}_0} \int_0^t P_n(s) \lambda_n(s) \, ds. \end{aligned} \quad (2.1.22)$$

Die Gültigkeit von Gleichung (2.1.22) wird klar, wenn wir den Ausdruck für einen reinen verschobenen Geburtsprozess und Abbruch bei $M \in \mathbb{N}$ betrachten. In diesem Falle erhalten wir wegen $\Lambda_{M+k} = 0$ für $k \geq 0$:

$$\mathbb{E}[\Lambda(t)] = \sum_{n=0}^{M-1} \lambda_n \int_0^t P_n(s) \, ds. \quad (2.1.23)$$

Unter Verwendung von Satz 2.1.14 können wir folglich aus (2.1.23) eine explizite Darstellung von $\mathbb{E}[\Lambda(t)]$ gewinnen. Diese wird in Satz 2.1.17 formuliert.

Satz 2.1.17.

Es sei erneut $(N_t)_{t \geq 0}$ ein reiner verschobener Geburtsprozess mit Abbruch bei $M \in \mathbb{N}$ dergestalt, dass Satz 2.1.14 gültig ist.

Die Koeffizienten $b_k^{(n)}$, c_k , c und π_n seien wie folgt definiert⁹:

$$\begin{aligned} b_k^{(n)} &= \left[\prod_{\substack{j=0, \\ j \neq k}}^n (\lambda_j - \lambda_k) \right]^{-1}, & n = 1, \dots, M-1, \quad k = 0, 1, \dots, n, \\ \pi_n &= \lambda_0 \cdot \dots \cdot \lambda_n, & n = 0, \dots, M, \end{aligned}$$

⁹Dies wird motiviert durch Satz 2.1.14 und den Beweis dieses Satzes.

$$c_k = \sum_{n=k}^{M-1} \frac{\pi_n b_k^{(n)}}{\lambda_k}, \quad k = 1, \dots, M-1,$$

$$c_0 = 1 + \sum_{n=1}^{M-1} \frac{\pi_n b_0^{(n)}}{\lambda_0}.$$

Dann gilt für den Erwartungswert der kumulierten Intensitätsfunktion:

$$E[\Lambda(t)] = \sum_{k=0}^{M-1} c_k \left(1 - e^{-\lambda_k t}\right).$$

Beweis¹⁰:

Mit den Notationen aus Satz 2.1.17 erhält man unter Verwendung von Satz 2.1.14 mittels einfacher Integration der einzelnen Summanden die nachfolgende Gleichungskette:

$$\begin{aligned} \sum_{n=0}^{M-1} \lambda_n \int_0^t P_n(s) ds &= \lambda_0 \int_0^t e^{-\lambda_0 s} ds + \sum_{n=1}^{M-1} \pi_n \sum_{k=0}^n b_k^{(n)} \int_0^t e^{-\lambda_k s} ds \\ &= \left(1 - e^{-\lambda_0 t}\right) + \sum_{n=1}^{M-1} \sum_{k=0}^n \frac{\pi_n b_k^{(n)}}{\lambda_k} \left(1 - e^{-\lambda_k t}\right). \end{aligned} \quad (2.1.24)$$

Umsortieren der Terme aus (2.1.24) liefert dann unmittelbar die retrospektive Begründung für die Definition der Koeffizienten c_k , da somit gilt:

$$E[\Lambda(t)] = \sum_{n=0}^{M-1} \lambda_n \int_0^t P_n(s) ds = \sum_{k=0}^{M-1} c_k \left(1 - e^{-\lambda_k t}\right).$$

Dies beweist die Darstellung des Erwartungswertes der kumulierten Intensitätsfunktion. \square

Beispiel 2.1.7. (3. Fortsetzung)

Wir nehmen nun an, dass der mit den Drahtbrüchen assoziierte Zählprozess $N = (N_t)_{t \geq 0}$ ein reiner Geburtsprozess mit Abbruch bei $M \in \mathbb{N}$ ist, dessen Intensitätsfunktion λ_n der parametrischen Familie

$$\left\{ \lambda : \lambda_n = \vartheta \frac{M}{M-n} \mathbb{1}_{\{n < M\}} \text{ mit } \vartheta > 0 \right\}$$

entstammt. Es existiert folglich eine Konstante $\vartheta_0 > 0$, so dass gilt:

$$\lambda(N_t) = \vartheta_0 \frac{M}{M - N_t} \mathbb{1}_{\{N_t < M\}}.$$

¹⁰Dieser Satz wurde vom Autor dieser Arbeit gefunden und bewiesen.

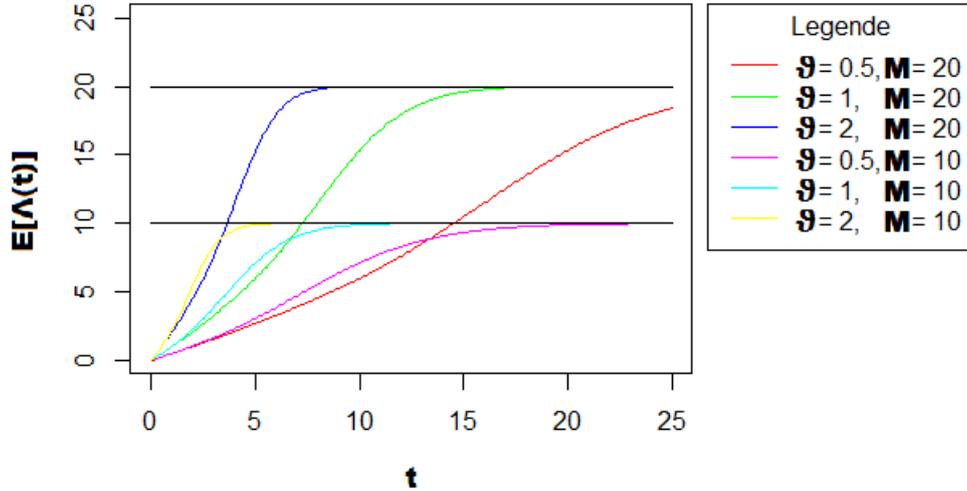


Abbildung 4: Plot der Erwartungswerte der kumulierten Intensitätsfunktion $E[\Lambda(t)]$ für verschiedene $\vartheta > 0$ und $M \in \mathbb{N}$. Diese Grafik wurde mit der statistischen Software *R* erstellt, siehe R Core Team 2015.

Der Parameter ϑ_0 kann als *Grundintensität* der Belastung interpretiert werden, die auf jedes Drahtseil wirkt. Die Familie von Intensitätsfunktionen ist dabei so konstruiert, dass sich die *Gesamtintensität* $\vartheta_0 M$ jederzeit gleichmäßig auf die intakten Drahtseile verteilt; mit jedem Drahtbruch steigt ergo die Belastung auf die verbleibenden Drahtseile, bis schließlich der gesamte Verbund gebrochen ist.

Um einen anschaulichen Zugang zum Resultat von Satz 2.1.17 zu gewährleisten, möchten wir $E[\Lambda(t)] = E[N_t]$ (wir zeigen diese Identität formal in Satz 2.1.19) als Funktion der Variable t für verschiedene ϑ und M betrachten. Für ausgewählte Werte finden sich Plots dieser Funktion in Abbildung 4.

Die genauere Betrachtung von Abbildung 4 legt nahe, dass sich die Funktion $E[\Lambda(t)]$ asymptotisch der Geraden $y = M$ annähert. Dies ist verträglich mit der heuristischen Interpretation, dass der Abbruch beim M -ten Sprung auftritt, insbesondere kann folglich auch kein weiterer Sprung mehr *erwartet* werden.

Diesen Sachverhalt möchten wir mathematisch nachweisen. Das Ergebnis halten wir im nächsten Satz fest.

Satz 2.1.18.

Es sei erneut die Situation aus Satz 2.1.17 gegeben. Dann gilt:

$$\lim_{t \rightarrow \infty} E[\Lambda(t)] = M.$$

Beweis¹¹:

Wir bemerken zunächst, dass für jedes $\lambda > 0$ gilt:

$$\lim_{t \rightarrow \infty} (1 - e^{-\lambda t}) = 1.$$

Gemäß Gleichung (2.1.24) genügt es also zu zeigen, dass folgende Identität gilt:

$$1 + \sum_{n=1}^{M-1} \sum_{k=0}^n \frac{\pi_n b_k^{(n)}}{\lambda_k} = M. \quad (2.1.25)$$

Die Gültigkeit von Gleichung (2.1.25) folgt dann unmittelbar aus der folgenden Behauptung:

$$\sum_{k=0}^n \frac{\pi_n b_k^{(n)}}{\lambda_k} \stackrel{!}{=} 1, \quad n = 1, \dots, M-1. \quad (2.1.26)$$

Dazu beachten wir, dass für $n > 1$ mit Gleichung (2.1.17) gilt:

$$\begin{aligned} \sum_{k=0}^n \frac{\pi_n b_k^{(n)}}{\lambda_k} &= \lambda_0 \cdot \dots \cdot \lambda_n \sum_{k=0}^n \frac{b_k^{(n)}}{\lambda_k} = \lambda_0 \cdot \dots \cdot \lambda_{n-1} \left[b_n^{(n)} + \sum_{k=0}^{n-1} \frac{\lambda_n}{\lambda_k} b_k^{(n)} \right] \\ &= \lambda_0 \cdot \dots \cdot \lambda_{n-1} \left[- \sum_{k=0}^{n-1} b_k^{(n)} + \sum_{k=0}^{n-1} \frac{\lambda_n}{\lambda_k} b_k^{(n)} \right] \\ &= \lambda_0 \cdot \dots \cdot \lambda_{n-1} \left[\sum_{k=0}^{n-1} \frac{\lambda_n - \lambda_k}{\lambda_k} b_k^{(n)} \right] \\ &= \lambda_0 \cdot \dots \cdot \lambda_{n-1} \sum_{k=0}^{n-1} \frac{b_k^{(n-1)}}{\lambda_k} = \sum_{k=0}^{n-1} \frac{\pi_{n-1} b_k^{(n-1)}}{\lambda_k}. \end{aligned} \quad (2.1.27)$$

Außerdem gilt offensichtlich:

$$\sum_{k=0}^1 \frac{\pi_1 b_k^{(1)}}{\lambda_k} = \lambda_1 b_0^{(1)} + \lambda_0 b_1^{(1)} = \frac{\lambda_1}{\lambda_1 - \lambda_0} + \frac{\lambda_0}{\lambda_0 - \lambda_1} = 1. \quad (2.1.28)$$

Kombination von (2.1.27) und (2.1.28) liefert den Beweis für die Behauptung aus (2.1.26). Das wiederum vollendet den Beweis. \square

¹¹Dieser Satz wurde vom Autor dieser Arbeit gefunden und bewiesen.

Wie bereits im Vorfeld von Gleichung (2.1.22) angekündigt, gilt in der vorliegenden Situation für alle $t \geq 0$ die Identität

$$\mathbb{E} [\Lambda(t)] = \mathbb{E} [N_t] .$$

Dies folgt einerseits aus einem Resultat, das wir im Kontext der *Doob-Meyer-Zerlegung für nicht-negative Submartingale* formulieren möchten, kann andererseits jedoch auch explizit gezeigt werden.

Der Beweis der obigen Identität im Falle eines reinen verschobenen Geburtsprozesses mit Abbruch, die Gegenstand des nachfolgenden Satzes ist, kann elementar unter Verwendung von Satz 2.1.14 geführt werden, ein deutlich eleganterer Beweis basiert jedoch auf der Differentialgleichung aus Satz 2.1.12. Wir möchten es dem Leser überlassen, die Formel aus Satz 2.1.17 ein weiteres mal herzuleiten, und zeigen daher hier die alternative Herangehensweise.

Satz 2.1.19.

Es sei $(N_t)_{t \geq 0}$ ein reiner verschobener Geburtsprozess mit Abbruch bei $M \in \mathbb{N}$ dergestalt, dass Satz 2.1.14 gültig ist . Dann gilt für alle $t \geq 0$ die Identität

$$\mathbb{E} [\Lambda(t)] = \mathbb{E} [N_t] . \tag{2.1.29}$$

Beweis¹²:

Aufgrund von Satz 2.1.14 gilt $P_{M+k}(t) = 0$ für alle $t \geq 0$ und $k \in \mathbb{N}$. Wir erhalten folglich für $t \geq 0$:

$$\mathbb{E} [N_t] = \sum_{n=1}^M n P_n(t) . \tag{2.1.30}$$

Andererseits liefert uns Gleichung (2.1.23):

$$\mathbb{E} [\Lambda(t)] = \sum_{n=0}^{M-1} \int_0^t \lambda_n(s) P_n(s) \, ds . \tag{2.1.31}$$

Wir erinnern die Differentialgleichung (2.1.6) aus Satz 2.1.12 und setzen die Identität

$$\lambda_n(s) P_n(s) = \left. \frac{d}{dt} P_{n+1}(t) \right|_{t=s} + \lambda_{n+1}(s) P_{n+1}(s) , \tag{2.1.32}$$

die für alle $n \in \mathbb{N}_0$ gültig ist, in Gleichung (2.1.31) ein. Sodann erhalten wir unter Verwendung des *Hauptsatzes der Differential- und Integralrechnung* wegen $P_n(0) = 0$ für alle $n \in \mathbb{N}$:

¹²Dieser Beweis wurde vom Autor dieser Arbeit selbst geführt.

$$\begin{aligned}
\mathbb{E} [\Lambda(t)] &= \sum_{n=0}^{M-1} \int_0^t \lambda_n(s) P_n(s) \, ds = \sum_{n=0}^{M-1} \int_0^t \frac{d}{dt} P_{n+1}(t) \Big|_{t=s} + \lambda_{n+1}(s) P_{n+1}(s) \, ds \\
&= \sum_{n=1}^M \left[P_n(t) - P_n(0) + \int_0^t \lambda_n(s) P_n(s) \, ds \right] = \sum_{n=1}^M P_n(t) + \sum_{n=1}^M \int_0^t \lambda_n(s) P_n(s) \, ds.
\end{aligned} \tag{2.1.33}$$

Setzen wir nun wiederum die Identität aus (2.1.32) für den Integranden aus (2.1.33) ein, so liefert sukzessives Wiederholen dieses Verfahrens wegen $P_{M+1} \equiv 0$ schließlich:

$$\begin{aligned}
\mathbb{E} [\Lambda(t)] &= \sum_{n=1}^M P_n(t) + \sum_{n=1}^M \int_0^t \lambda_n(s) P_n(s) \, ds \\
&= \sum_{n=1}^M P_n(t) + \sum_{n=2}^{M+1} \left[P_n(t) - P_n(0) + \int_0^t \lambda_n(s) P_n(s) \, ds \right] \\
&= \sum_{n=1}^M P_n(t) + \sum_{n=2}^M P_n(t) + \sum_{n=2}^M \int_0^t \lambda_n(s) P_n(s) \, ds \\
&= \dots = \sum_{k=1}^M \sum_{n=k}^M P_n(t) = \sum_{n=1}^M n P_n(t),
\end{aligned}$$

und dies ist gerade die Darstellung des Erwartungswertes von N_t aus Gleichung (2.1.30). \square

Bemerkung 2.1.20.

- Der Beweis zeigt insbesondere, dass die Aussage von Satz 2.1.19 gültig bleibt, wenn die Intensitätsfunktion $\lambda_n(t)$ *nicht* zeithomogen ist.
- In Satz 2.1.19 haben wir explizit gezeigt, dass sich der Erwartungswert eines reinen verschobenen Geburtsprozesses mit Abbruch zu einem gegebenen Zeitpunkt $t \geq 0$ über die kumulierte Intensitätsfunktion ermitteln lässt. Dies führt zu der Idee, dass die kumulierte Intensitätsfunktion das qualitative Verhalten des Zählprozesses widerspiegelt. Auf Grundlage dieser Erkenntnis werden wir später einen Schätzer für die *Verallgemeinerung* dieser Funktion einführen:

Dem *Kompensator* aus der *Doob-Meyer-Zerlegung für nicht-negative Submartingale*, den wir im Unterabschnitt 2.1.6 kennenlernen.

2.1.5 Filtrationen und Martingale

In den nächsten Unterabschnitten präsentieren wir nebst grundlegender Definitionen ein überschaubares Potpourri wichtiger Resultate aus der Martingalthorie. Hiermit wird es uns schlussendlich in Unterabschnitt 2.1.8 möglich sein, die in der vorliegenden Arbeit angewandte Definition eines *selbstanregenden Zählprozesses* anzugeben.

Die zentralen Sätze aus den Unterabschnitten 2.1.6 und 2.1.7, wie etwa die *Doob-Meyer-Zerlegung für nicht-negative Submartingale* und die *Itô-Isometrie für quadratisch integrierbare Martingale*, werden in Grundlagenvorlesungen zur Wahrscheinlichkeitstheorie behandelt und sollen hier deshalb ohne Beweis angegeben werden. Für den interessierten Leser findet sich dabei stets eine Literaturangabe, um die mitunter technisch anspruchsvollen Beweise im Selbststudium nachzuvollziehen.

Bevor wir den Begriff des *Martingals* definieren können, müssen wir zunächst sogenannte *Filtrationen* einer σ -Algebra einführen. Anschließend lernen wir *filtrierte Wahrscheinlichkeitsräume*, *adaptierte stochastische Prozesse* und die *kanonische Filtration* eines Prozesses kennen.

Definition 2.1.21. (*Filtrationen & adaptierte Prozesse*, siehe Ethier und Kurtz 1986, S. 50 und Bauer 1991, S. 138)

Es sei \mathcal{F} eine σ -Algebra und die Indexmenge \mathcal{I} eine Teilmenge der reellen Zahlen.

Eine *Filtration der σ -Algebra \mathcal{F}* ist eine Menge $\{\mathcal{F}_t\}_{t \in \mathcal{I}}$ von Unter- σ -Algebren von \mathcal{F} mit

$$\mathcal{F}_s \subset \mathcal{F}_t, \quad \text{für alle } s, t \in \mathcal{I} \text{ mit } s \leq t.$$

Ist (Ω, \mathcal{F}, P) ein Wahrscheinlichkeitsraum und $\{\mathcal{F}_t\}_{t \in \mathcal{I}}$ eine Filtration der σ -Algebra \mathcal{F} , so heißt $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \in \mathcal{I}}, P)$ ein *filtrierter Wahrscheinlichkeitsraum*.

Ein messbarer stochastischer Prozess X mit Indexmenge \mathcal{I} über dem filtrierten Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \in \mathcal{I}}, P)$ heißt *adaptiert* bezüglich der Filtration $\{\mathcal{F}_t\}_{t \in \mathcal{I}}$, falls X_t \mathcal{F}_t -messbar für alle $t \in \mathcal{I}$ ist¹³.

Die *kanonische Filtration* des stochastischen Prozesses X ist definiert durch

$$\mathcal{F}_t^X = \sigma(\{X_s : s \in \mathcal{I}, s \leq t\}), \quad t \in \mathcal{I},$$

und ist die grösste Filtration, bezüglich welcher der Prozess X adaptiert ist.

¹³Wie in Jacobsen 2006, S. 302 bemerkt wird, ist ein rechtsseitig stetiger Prozess $X = (X_t)_{t \in \mathcal{I}}$ dergestalt, dass X_t für alle $t \in \mathcal{I}$ \mathcal{F}_t -messbar ist, automatisch *messbar*. Insbesondere muss also für einen Zählprozess *keine* Messbarkeitsannahme getroffen werden.

Ist X adaptiert bezüglich einer Filtration $\{\mathcal{F}_t\}_{t \in \mathcal{I}}$, so gilt folglich

$$\mathcal{F}_t^X \subset \mathcal{F}_t, \quad t \in \mathcal{I}.$$

Wir verwenden bisweilen die suggestive Bezeichnung eines *filtrierten Prozesses* für das Tupel $(X_t, \mathcal{F}_t)_{t \in \mathcal{I}}$, die ihre Bewandnis insbesondere mit der Definition eines *Martingals* hat.

Obleich wir den Begriff der Filtration erst jetzt in Gebrauch nehmen, sind wir in Abschnitt 2.1.3 bereits einer Filtration begegnet:

Die in den Definitionen 2.1.8 und 2.1.9 betrachte Filtration $\{\mathcal{F}_t\}_{t \geq t_0}$ mit

$$\mathcal{F}_t = \sigma(\{N_s : t_0 \leq s < t\}), \quad t \geq t_0,$$

die auch als *Historie des Zählprozesses* N bezeichnet wird (siehe Daley und Vere-Jones 2003, S. 424). So verwundert es also nicht, dass mit einer Filtration $\{\mathcal{F}_t\}_{t \in \mathcal{I}}$ häufig die zum Zeitpunkt t verfügbare *Information* assoziiert wird, wobei dieser Zusammenhang rein heuristisch anmutet.

Wir möchten nun den Begriff des *vorhersagbaren Prozesses* beleuchten. *Vorhersagbare Prozesse* besitzen für uns keine inhaltliche Relevanz über die mit ihnen verknüpften Resultate dieses Abschnitts hinaus, weshalb sich der genügsame Leser mit der bloßen *Kenntnis* des Begriffes, der hier der Vollständigkeit halber nicht fehlen soll, begnügen mag.

Definition 2.1.22. (*Vorhersagbare Prozesse*, siehe Kuo 2006, S. 79 sowie Daley und Vere-Jones 2003, S. 425)

Es sei $I \subset \mathbb{R}$ ein Intervall und $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \in I}, P)$ ein filtrierter Wahrscheinlichkeitsraum. Ferner bezeichne \mathbb{L} die Menge aller stochastischen Prozesse X mit Indexmenge I , die linksseitig stetig und adaptiert bezüglich der Filtration $\{\mathcal{F}_t\}_{t \in I}$ sind.

Dann heißt die von \mathbb{L} erzeugte σ -Algebra $\sigma(\mathbb{L})$ die *vorhersagbare σ -Algebra* und ist eine Sub- σ -Algebra von $\mathcal{B}(I) \times \mathcal{F}$.

Ein stochastischer Prozess heißt dann *vorhersagbar* bezüglich $\{\mathcal{F}_t\}_{t \in I}$, wenn er $\sigma(\mathbb{L})$ - \mathcal{E} -messbar ist.

Definitionsgemäß ist der archetypische vorhersagbare Prozess linksseitig stetig, und dies motiviert auch die Bezeichnung der *Vorhersagbarkeit*:

Für einen linksseitig stetigen stochastischen Prozess wird der Wert zum Zeitpunkt t durch die zu vorherigen Zeitpunkten gegebenen Informationen bestimmt. Dass sich diese Aussage für vorhersagbare Prozesse übertragen lässt, zeigt das nachfolgende Lemma, das wir ohne Beweis angeben möchten.

Lemma 2.1.23. (siehe Daley und Vere-Jones 2003, S. 425)

In der Situation aus Definition 2.1.22 definieren wir die Filtration $\{\mathcal{F}_t^-\}_{t \in I}$ durch

$$\mathcal{F}_{t_0}^- = \mathcal{F}_{t_0} \quad \text{und} \quad \mathcal{F}_t^- = \bigcup_{\substack{s \in I \\ s < t}} \mathcal{F}_s, \quad \text{für } t \in I \setminus \{t_0\}, \quad \text{mit } t_0 = \inf I.$$

Dann ist jeder bezüglich $\{\mathcal{F}_t\}_{t \in I}$ vorhersagbare Prozess adaptiert bezüglich $\{\mathcal{F}_t^-\}_{t \in I}$.

Beweis: Ein Beweis findet sich ebenfalls in Daley und Vere-Jones 2003. \square

Die Aussage von Lemma 2.1.23 geht nun konform mit der heuristischen Interpretation einer Filtration als Ansammlung von Informationen:

Während wir für einen adaptierten Prozess zumindest bei Vorhandensein aller Informationen bis inklusive dem Zeitpunkt t den Wert des Prozesses zum Zeitpunkt t kennen, benötigen wir für einen vorhersagbaren Prozess nur Kenntnis über die Vergangenheit des Prozesses.

Wir möchten nun den Begriff der Vorhersagbarkeit beiseite schieben und uns der zentralen Definition dieses Abschnittes zuwenden: den *Martingalen*.

Definition 2.1.24. (*Martingale in diskreter und stetiger Zeit*, siehe Bauer 1991, S. 138)

Es sei $\mathcal{I} = \mathbb{N}$ oder $\mathcal{I} = I$ für ein Intervall $I \subset \mathbb{R}$ und $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \in \mathcal{I}}, P)$ ein filtrierter Wahrscheinlichkeitsraum. Ferner sei $X = (X_t)_{t \in \mathcal{I}}$ ein adaptierter stochastischer Prozess. Dann heißt der filtrierte Prozess $(X_t, \mathcal{F}_t)_{t \in \mathcal{I}}$ ein *Martingal*, falls die nachfolgenden Bedingungen erfüllt sind:

- (i) Für jedes $t \in \mathcal{I}$ ist X_t integrierbar, d.h. es ist $E |X_t| < \infty$ für alle $t \in \mathcal{I}$.
- (ii) Es gilt fast sicher $E(X_t | \mathcal{F}_s) = X_s$ für alle $s, t \in \mathcal{I}$ mit $s \leq t$.

Gilt Bedingung (ii) mit \geq bzw. \leq anstelle von $=$, so heißt X ein *Supermartingal* bzw. *Submartingal*.

Wir folgern hieraus leicht ein Resultat für Zählprozesse, das uns später die Anwendung der *Doob-Meyer-Zerlegung für nicht-negative Submartingale* erlaubt.

Lemma 2.1.25. (*Zählprozesse als nicht-negative Submartingale*)

Es sei $N = (N_t)_{t \geq t_0}$ ein adaptierter Zählprozess über dem filtrierten Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq t_0}, P)$. Ist N_t für alle $t \geq t_0$ integrierbar, so ist $(N_t, \mathcal{F}_t)_{t \geq t_0}$ ein nicht-negatives Submartingal.

Beweis¹⁴:

Nach Konstruktion - man erinnere Definition 2.1.5 - ist $N_t \geq 0$ für alle $t \geq t_0$. Ebenfalls nach Konstruktion gilt fast sicher $N_t \geq N_s$ für $t \geq s \geq t_0$. Aus der Monotonie des bedingten Erwartungswertes folgern wir schließlich

$$E(N_t | \mathcal{F}_s) \geq E(N_s | \mathcal{F}_s) = N_s \quad \text{fast sicher, für } t \geq s \geq t_0,$$

da N_s nach Voraussetzung \mathcal{F}_s - messbar ist. Somit gilt die Bedingung (ii) aus Definition 2.1.24. Da die Bedingung (i) aus Definition 2.1.24 nach Voraussetzung erfüllt ist, ist $(N_t, \mathcal{F}_t)_{t \geq t_0}$ ein nicht-negatives Submartingal. \square

Beispiel 2.1.7. (4. Fortsetzung)

Ein reiner Geburtsprozess mit Abbruch ist stets ein nicht-negatives Submartingal bezüglich der kanonischen Filtration, da die Beschränkung des Zählprozesses durch ein $M \in \mathbb{N}$ die Integrierbarkeit impliziert.

Wir sind nun nahezu in der Lage, die *Doob-Meyer-Zerlegung für nicht-negative Submartingale* zu formulieren.

2.1.6 Doob-Meyer-Zerlegung für nicht-negative Submartingale

Das nachfolgend diskutierte Resultat stellt einen Spezialfall der allgemeineren *Doob-Meyer-Zerlegung* dar. Die im Rahmen dieses Theorems betrachtete Filtration muss jedoch die *üblichen Bedingungen* (engl. “*usual conditions*”) erfüllen, die wie folgt lauten (siehe Jacobsen 2006, S. 301):

Es sei $t_0 \in \mathbb{R}$. Ein filtrierter Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq t_0}, P)$ erfüllt die *üblichen Bedingungen*, falls die nachfolgenden Bedingungen erfüllt sind:

- (i) Der Wahrscheinlichkeitsraum (Ω, \mathcal{F}, P) ist *vollständig*:
Ist $A_0 \in \mathcal{F}$ eine P -Nullmenge, d.h. es ist $P(A_0) = 0$, so gilt für jede Teilmenge $A \subset A_0$ bereits $A \in \mathcal{F}$.
- (ii) Es gelte $\mathcal{N} \subset \mathcal{F}_{t_0}$, wobei \mathcal{N} die Menge der P -Nullmengen sei, d.h.

$$\mathcal{N} = \{A \in \mathcal{F} : P(A) = 0\}.$$

- (iii) Die Filtration $\{\mathcal{F}_t\}_{t \geq t_0}$ ist *rechtsseitig stetig*, d.h. es gilt

$$\mathcal{F}_t = \mathcal{F}_t^+ = \bigcap_{s > t} \mathcal{F}_s, \quad \text{für alle } t \geq t_0.$$

¹⁴Dieser leichte Beweis wurde vom Autor dieser Arbeit selbst geführt.

Häufig wird bereits vorab angenommen, dass die üblichen Bedingungen stets erfüllt seien (vgl. etwa Protter 2005, S. 3). Wenn wir die kanonische Filtration eines Zählprozesses betrachten, so ist hingegen nicht unmittelbar klar, ob dies ohne Einschränkung möglich ist. Die nachfolgenden Lemmata liefern jedoch eine positive Antwort auf diese Frage.

Lemma 2.1.26. (*Vervollständigung eines Wahrscheinlichkeitsraumes*, siehe Jacobsen 2006, S. 301)

Es sei (Ω, \mathcal{F}, P) ein Wahrscheinlichkeitsraum. Dann existiert ein vollständiger Wahrscheinlichkeitsraum $(\Omega, \overline{\mathcal{F}}, \overline{P})$ mit $\mathcal{F} \subset \overline{\mathcal{F}}$ und

$$\overline{P}(A) = P(A), \quad \text{für alle } A \in \mathcal{F}.$$

Beweis¹⁵:

Wir definieren die Menge

$$\mathcal{N} = \{A \subset \Omega : \exists A_0 \in \mathcal{F} \text{ mit } P(A_0) = 0 \text{ und } A \subset A_0\},$$

und setzen sodann

$$\overline{\mathcal{F}} = \{F \cup N : F \in \mathcal{F}, N \in \mathcal{N}\}. \quad (2.1.34)$$

Das Mengensystem $\overline{\mathcal{F}}$ ist eine σ -Algebra; wir zeigen hier nur die Stabilität von $\overline{\mathcal{F}}$ unter der Komplementbildung, da die verbleibenden Eigenschaften trivial sind:

Es seien also $F \in \mathcal{F}$ und $N \in \mathcal{N}$. Dann existiert nach Konstruktion ein $N_0 \in \mathcal{F}$ mit $N \subset N_0$ und wir definieren $\tilde{N} = N_0 \setminus N \in \mathcal{N}$. Damit gilt:

$$\begin{aligned} (F \cup N)^c &= \left(F \cup (N_0 \setminus \tilde{N}) \right)^c = F^c \cap (N_0 \setminus \tilde{N})^c \\ &= F^c \cap (N_0^c \cup \tilde{N}) = (F \cup N_0)^c \cup (F^c \cap \tilde{N}) \in \overline{\mathcal{F}}, \end{aligned}$$

da $F \cup N_0 \in \mathcal{F}$ gilt und $F^c \cap \tilde{N} \subset N_0$ ein Element von \mathcal{N} ist.

Auf $\overline{\mathcal{F}}$ definieren wir das Wahrscheinlichkeitsmaß \overline{P} via

$$\overline{P}(F \cup N) = P(F), \quad \text{für } F \in \mathcal{F}, N \in \mathcal{N}. \quad (2.1.35)$$

Nach Konstruktion ist zudem $(\Omega, \overline{\mathcal{F}}, \overline{P})$ ein vollständiger Wahrscheinlichkeitsraum. Dies liefert in Verbindung mit Gleichung (2.1.35) die Behauptung. \square

¹⁵Der Beweis findet sich ebenfalls in Jacobsen 2006, S. 301. Wir zeigen hier zusätzlich die Stabilität des in (2.1.34) definierten Mengensystems unter der Komplementbildung.

Aufgrund von Lemma 2.1.26 können wir also ohne Einschränkung von einem vollständigen Wahrscheinlichkeitsraum ausgehen; die erste übliche Bedingung ist somit erfüllt.

Ist nun $\{\mathcal{F}_t^X\}_{t \geq t_0}$ die kanonische Filtration eines stochastischen Prozesses, so können wir völlig analog zur *vervollständigten Filtration* $\{\overline{\mathcal{F}}_t^X\}_{t \geq t_0}$ übergehen, indem $\overline{\mathcal{F}}_t^X$ wie in Gleichung (2.1.34) definiert wird.

Diese Filtration erfüllt dann die zweite gewöhnliche Bedingung.

Da die vervollständigte Filtration nur durch Hinzunahme von P -Nullmengen konstruiert wurde, ist heuristisch gesprochen kein *Informationsgewinn* zu verzeichnen, so dass wir in der Praxis nicht zwischen der ursprünglichen und der vervollständigten Filtration unterscheiden müssen.

Den Schlüssel zur dritten gewöhnlichen Bedingung liefert das nachfolgende Lemma.

Lemma 2.1.27. (*Rechtsseitige Stetigkeit der kanonischen Filtration eines Zählprozesses*, siehe Protter 2005, S. 16)

Es sei $N = (N_t)_{t \geq t_0}$ ein Zählprozess. Dann ist die kanonische Filtration $\{\mathcal{F}_t^N\}_{t \geq t_0}$ des Zählprozesses rechtsseitig stetig.

Beweis¹⁶:

Es sei $I = [0, \infty]$. Für $t \geq t_0$ definieren wir die Abbildungen

$$\pi_t : \Omega \longrightarrow \Gamma : \omega \longmapsto [s \mapsto N_{\min(s,t)}(\omega)] ,$$

wobei Γ definiert sei als der Messraum

$$\Gamma = \left(\prod_{t \geq t_0} I, \bigotimes_{t \geq t_0} \mathcal{B}(I) \right) ,$$

das heißt der Menge aller Abbildung von $[t_0, \infty)$ nach I , versehen mit der Borel'schen Produkt- σ -Algebra (für eine ausführliche Definition und mathematische Hintergründe siehe etwa Bauer 1991, S. 58-67).

Da die Pfade im Bildbereich von π_t gemäß Konstruktion jenseits des Zeitpunktes t konstant sind, gilt $\sigma(\{\pi_t\}) = \sigma(\{N_t : t_0 \leq s \leq t\}) = \mathcal{F}_t^N$; wir haben folglich den Zählprozess N bis zum Zeitpunkt t durch eine einzige - dann jedoch funktionswertige - Zufallsvariable ausgedrückt¹⁷.

¹⁶Der Beweis findet sich ebenfalls in Protter 2005, S. 16 und wird um einige Erläuterungen ergänzt.

¹⁷Wir werden dieses Vorgehen später erneut utilisieren, indem wir Zählprozesse als Zufallsvariablen mit Werten im *Skorokhod-Raum* interpretieren.

Wir betrachten nun ein Ereignis

$$A \in \bigcap_{n \in \mathbb{N}} \mathcal{F}_{t+\frac{1}{n}}^N = \bigcap_{n \in \mathbb{N}} \sigma\left(\{\pi_{t+\frac{1}{n}}\}\right).$$

Definitionsgemäß existiert zu jedem $n \in \mathbb{N}$ eine Menge

$$B_n \in \bigotimes_{t \geq t_0} \mathcal{B}(I)$$

mit $A = \pi_{t+\frac{1}{n}}^{-1}(B_n) = \{\pi_{t+\frac{1}{n}} \in B_n\}$. Durch $W_n = \{\pi_t = \pi_{t+\frac{1}{n}}\}$ definieren wir eine monoton aufsteigende Mengenfolge.

Aufgrund der rechtsseitigen Stetigkeit von N und der Tatsache, dass der mit N assoziierte Punktprozess fast sicher gegen ∞ konvergiert, existiert zu jedem $\omega \in \Omega$ ein $n \in \mathbb{N}$ dergestalt, dass der Pfad $s \mapsto N_s(\omega)$ auf dem Intervall $[t, t + \frac{1}{n}]$ konstant ist, woraus wir folgern:

$$\Omega = \bigcup_{n \in \mathbb{N}} W_n.$$

Da die Mengenfolge $(W_n)_{n \in \mathbb{N}}$ monoton aufsteigend ist, gilt folglich

$$\lim_{n \rightarrow \infty} W_n = \Omega,$$

und somit erhalten wir schließlich:

$$\begin{aligned} A &= \lim_{n \rightarrow \infty} (W_n \cap A) = \lim_{n \rightarrow \infty} \left(W_n \cap \{\pi_{t+\frac{1}{n}} \in B_n\} \right) \\ &= \lim_{n \rightarrow \infty} (W_n \cap \{\pi_t \in B_n\}) = \lim_{n \rightarrow \infty} \{\pi_t \in B_n\} \in \sigma(\{\pi_t\}) = \mathcal{F}_t^N, \end{aligned}$$

und daher

$$\bigcap_{n \in \mathbb{N}} F_{t+\frac{1}{n}}^N \subset F_t^N. \quad (2.1.36)$$

Die umgekehrte Inklusion in (2.1.36) gilt nach Definition einer Filtration; insgesamt folgt damit die Behauptung. \square

Die Lemmata 2.1.26 und 2.1.27 erlauben nun ohne Einschränkung die Voraussetzung der üblichen Bedingungen an den filtrierte Wahrscheinlichkeitsraum im Kontext der Analyse von Zählprozessen.

Wir sind nun in der Lage, die *Doob-Meyer-Zerlegung für nicht-negative Submartingale* anzugeben.

Satz 2.1.28. (*Doob-Meyer-Zerlegung für nicht-negative Submartingale*, siehe Pang, Talreja und Whitt 2007, S. 208 und Ethier und Kurtz 1986, S. 74)

Es sei $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq t_0}, P)$ ein filtrierter Wahrscheinlichkeitsraum, der die üblichen Bedingungen erfülle. Es sei ferner $(X_t, \mathcal{F}_t)_{t \geq t_0}$ ein rechtsseitig stetiges, nicht-negatives Submartingal.

Dann existiert ein bezüglich $\{\mathcal{F}_t\}_{t \geq t_0}$ vorhersagbarer, integrierbarer und rechtsseitig stetiger Prozess $\Lambda = (\Lambda_t)_{t \geq t_0}$ mit $\Lambda_{t_0} = 0$ und monoton wachsenden Pfaden dergestalt, dass der Prozess $M = X - \Lambda$ ein $\{\mathcal{F}_t\}_{t \geq t_0}$ -Martingal ist.

Wir nennen Λ den *Kompensator* des stochastischen Prozesses X .

Der Kompensator ist eindeutig in dem Sinne, dass die Pfade zweier Versionen des Kompensators fast sicher übereinstimmen¹⁸.

Beweis:

Die Doob-Meyer-Zerlegung kann in allgemeinerer Form für rechtsseitig stetige Submartingale *der Klasse (DL)* bewiesen werden, worauf wir hier nicht näher eingehen möchten. Jedes nach unten beschränkte Submartingal gehört zu dieser Klasse (siehe Pang, Talreja und Whitt 2007, S. 208), weshalb wir an dieser Stelle lediglich auf den Beweis aus Ethier und Kurtz 1986 verweisen. \square

Wie wir in Lemma 2.1.25 gesehen haben, genügt ein adaptierter Zählprozess den Voraussetzungen der Doob-Meyer-Zerlegung aus Satz 2.1.28.

Für die Untersuchung der generellen Dynamik eines solchen Zählprozesses bietet es sich also an, den vorhersagbaren Kompensator des Zählprozesses zu studieren, siehe auch Daley und Vere-Jones 2008, S. 355.

Die Sinnhaftigkeit dieses Vorgehens wird durch die nachfolgende Bemerkung gestützt.

Bemerkung 2.1.29.

1. Wir betrachten den homogenen Poisson-Prozess N mit Parameter λ aus Beispiel 2.1.7. Als Filtration $\{\mathcal{F}_t\}_{t \geq 0}$ wählen wir die - gegebenenfalls vervollständigte - kanonische Filtration dieses Zählprozesses und definieren $\Lambda_t = \lambda t$. Dann gilt für $t \geq s \geq 0$:

$$\begin{aligned} \mathbb{E}[N_t - \Lambda_t | \mathcal{F}_s] &= \mathbb{E}[(N_t - N_s) - (\Lambda_t - \Lambda_s) | \mathcal{F}_s] + \mathbb{E}[N_s - \Lambda_s | \mathcal{F}_s] \\ &= \mathbb{E}[N_t - N_s | \mathcal{F}_s] - \lambda(t - s) + N_s - \Lambda_s \\ &= \mathbb{E}[N_t - N_s] - \lambda(t - s) + N_s - \Lambda_s \\ &= \lambda(t - s) - \lambda(t - s) + N_s - \Lambda_s = N_s - \Lambda_s. \end{aligned}$$

Hierbei haben wir ausgenutzt, dass definitionsgemäß die Zuwächse $N_t - N_s$ unabhängig von \mathcal{F}_s sind, da wir die kanonische Filtration des Zählprozesses N betrachten,

¹⁸Man nennt solche Prozesse *ununterscheidbar* (engl. “*indistinguishable*”), siehe Protter 2005, S. 4.

und aus dem selben Grund ist die Zufallsvariable $N_s - \Lambda_s$ offensichtlich \mathcal{F}_s -messbar. Wir erkennen also, dass $(N_t - \Lambda_t, \mathcal{F}_t)_{t \geq 0}$ ein Martingal ist und können schließlich folgern, dass der Prozess $\Lambda = (\Lambda_t)_{t \geq 0}$ der Kompensator des homogenen Poisson-Prozesses ist.

Der Kompensator ist hier also eine deterministische Funktion der Zeit t ; der homogene Poisson-Prozess lässt sich ergo als Summe eines deterministischen Anteils und eines Martingals darstellen.

2. Ist allgemein Λ der Kompensator eines stochastischen Prozesses X über dem filtrierten Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq t_0}, P)$, so gilt für $t \geq s \geq t_0$ mit den Notationen aus Satz 2.1.28:

$$\begin{aligned} \mathbb{E}[X_t - X_s | \mathcal{F}_s] &= \mathbb{E}[(X_t - \Lambda_t) - (X_s - \Lambda_s) | \mathcal{F}_s] + \mathbb{E}[\Lambda_t - \Lambda_s | \mathcal{F}_s] \\ &= \mathbb{E}[M_t - M_s | \mathcal{F}_s] + \mathbb{E}[\Lambda_t - \Lambda_s | \mathcal{F}_s] = \mathbb{E}[\Lambda_t - \Lambda_s | \mathcal{F}_s]. \end{aligned}$$

Der erwartete Zuwachs des Prozesses X im Intervall $[s, t]$, gegeben die *Information* zum Zeitpunkt s , entspricht also gerade dem erwarteten Zuwachs des vorhersagbaren Kompensators Λ im selben Zeitraum.

3. Der Kompensator aus der Doob-Meyer-Zerlegung für nicht-negative Submartingale ist definitionsgemäß ein vorhersagbarer und rechtsseitig stetiger stochastischer Prozess. Da Vorhersagbarkeit häufig mit linksseitiger Stetigkeit assoziiert wird (vergleiche die Ausführungen im Anschluss an Definition 2.1.22), erscheint es nicht unnatürlich, dass wir später die *Stetigkeit* des Kompensators fordern werden.

Der nachfolgende Satz zeigt, dass der in Satz 2.1.28 eingeführte Kompensator eines Zählprozesses als Verallgemeinerung der kumulierten Intensitätsfunktion aus Gleichung (2.1.21) aufgefasst werden kann. Wir setzen nachfolgend voraus, dass der zugrunde liegende filtrierte Wahrscheinlichkeitsraum die üblichen Bedingungen erfüllt.

Satz 2.1.30. (*Kompensator eines Zählprozesses*, siehe Daley und Vere-Jones 2003, S. 241)

Es sei $N = (N_t)_{t \geq t_0}$ ein Zählprozess, der bezüglich einer Filtration $(\mathcal{F}_t)_{t \geq t_0}$ adaptiert sei. Die *bedingte Intensitätsfunktion* $\lambda(t)$ sei definiert via

$$\lambda(t) = \lim_{\Delta t \rightarrow 0} \frac{P(\{N(t, t + \Delta t) = 1\} | \mathcal{F}_t^-)}{\Delta t}, \quad (2.1.37)$$

so der Grenzwert für alle $t \geq t_0$ existiert (vergleiche Lemma 2.1.23).

Ist die Funktion $\lambda(t)$ linksseitig stetig und definiert man die *kumulierte Intensitätsfunktion*

durch

$$\Lambda(t) = \int_{t_0}^t \lambda(s) \, ds,$$

dann ist Λ der Kompensator des filtrierten Zählprozesses $(N_t, \mathcal{F}_t)_{t \geq t_0}$ aus der Doob-Meyer-Zerlegung für nicht-negative Submartingale.

Beweis:

Eine Beweisidee findet sich ebenfalls in Daley und Vere-Jones 2003, S. 241. Für eine ausführlichere Studie der Kompensatoren von Zählprozessen siehe Daley und Vere-Jones 2008, Kapitel 14. □

Betrachtet man in Satz 2.1.30 die kanonische Filtration $\mathcal{F}_t = \mathcal{F}_t^N$ des Zählprozesses N , so stimmt die Definitionen der bedingten Intensitätsfunktionen gerade mit Definition 2.1.9 überein. Wir haben also in den Unterabschnitten 2.1.3 und 2.1.4 implizit stets die kanonische Filtration zugrunde gelegt.

Der obige Satz stellt nun die Verknüpfung zwischen diesem Abschnitt und der Doob-Meyer-Zerlegung für nicht-negative Submartingale her, ist doch die kumulierte Intensitätsfunktion nichts anderes als der Kompensator des verschobenen Geburtsprozesses. Dieser Umstand wurde bereits im Vorfeld von Satz 2.1.19 thematisiert und liefert zugleich einen impliziten Beweis jenes Satzes.

Abschließend möchten wir das Beispiel 2.1.7 im Kontext von Satz 2.1.30 diskutieren.

Beispiel 2.1.7. (5. Fortsetzung)

Wir erinnern an die in der dritten Fortsetzung dieses Beispiels eingeführte parametrische Familie bedingter Intensitätsfunktionen:

$$\left\{ \lambda : \lambda_n = \vartheta \frac{M}{M-n} \mathbb{1}_{\{n < M\}} \text{ mit } \vartheta > 0 \right\}$$

Wir haben sodann einen reinen verschobenen Geburtsprozess N mit Abbruch bei $M \in \mathbb{N}$ betrachtet, dessen bedingte Intensitätsfunktion λ dieser Familie entstammt, das heißt

$$\lambda(t) = \vartheta_0 \frac{M}{M - N_t} \mathbb{1}_{\{N_t < M\}}, \quad \text{für ein } \vartheta_0 > 0.$$

Der stochastische Prozess λ erbt nun die rechtsseitige Stetigkeit des Zählprozesses N , ist also - anders als in Satz 2.1.30 gefordert - *nicht* linksseitig stetig. Es gibt jedoch eine linksseitig stetige *Modifikation* der Intensitätsfunktion, das heißt einen linksseitig stetigen stochastischen Prozess, der für jedes $t \geq 0$ fast sicher mit λ übereinstimmt¹⁹:

¹⁹Der Begriff der Modifikation ist nicht mit der Ununterscheidbarkeit zu verwechseln; in letzterem Falle stimmen die stochastischen Prozesse fast sicher *für alle t simultan* überein.

Dazu definieren wir den stochastischen Prozess λ^- durch

$$\lambda^-(t) = \lim_{s \uparrow t} \lambda(s).$$

Dieser Prozess ist offensichtlich linksseitig stetig und es gilt für jedes $t \geq 0$:

$$P(\{\lambda^-(t) \neq \lambda(t)\}) = P\left(\left\{\lim_{s \uparrow t} N_s \neq N_t\right\}\right) = 0,$$

da ein Sprung zum exakten Zeitpunkt t mit Wahrscheinlichkeit 0 auftritt.

Die via λ^- bzw. λ definierten kumulierten Intensitätsfunktionen stimmen offensichtlich überein, so dass der Kompensator in diesem Falle gemäß Satz 2.1.30 gegeben ist durch

$$\Lambda(t) = \int_0^t \vartheta_0 \frac{M}{M - N_s} \mathbb{1}_{\{N_s < M\}} ds,$$

insbesondere ist der Kompensator sogar ein *stetiger* stochastischer Prozess (vergleiche Bemerkung 2.1.29).

2.1.7 Itô-Isometrie für quadratisch integrierbare Martingale

Als nächstes möchten wir uns kurz mit stochastischen Integralen, deren Integrator durch ein *quadratisch integrierbares* Martingal gegeben ist, auseinandersetzen²⁰.

Wir präsentieren einige nützliche Resultate zur Handhabung solcher Itô-Integrale; insbesondere die *Itô-Isometrie für quadratisch integrierbare Martingale* wird in den späteren Abschnitten dieser Arbeit ubiquitär sein.

Ist $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq t_0}, P)$ ein filtrierter Wahrscheinlichkeitsraum, so heißt dabei ein Martingal $M = (M_t, \mathcal{F}_t)_{t \geq t_0}$ *quadratisch integrierbar*, falls für alle $t \geq t_0$ gilt:

$$E(M_t^2) < \infty.$$

Für ein solches quadratisch integrierbares Martingal gilt offensichtlich unter Verwendung der *Jensen'schen Ungleichung für bedingte Erwartungswerte* (siehe etwa Bauer 1991, S. 121):

$$E[M_t^2 | \mathcal{F}_s] \geq (E[M_t | \mathcal{F}_s])^2 = M_s^2, \quad \text{für } t \geq s \geq t_0,$$

es ist folglich $(M_t^2, \mathcal{F}_t)_{t \geq t_0}$ ein nicht-negatives Submartingal. Ist der Prozess $(M_t^2)_{t \geq t_0}$ zudem rechtsseitig stetig, so existiert folglich die Doob-Meyer-Zerlegung aus Satz 2.1.28 für dieses Submartingal.

Diese Überlegung liegt dem nachfolgenden Satz zugrunde.

²⁰Für eine detaillierte Einführung in die Theorie der stochastischen Integration sei an dieser Stelle auf die Monographie Protter 2005 verwiesen.

Satz 2.1.31. (siehe Pang, Talreja und Whitt 2007, S. 211)

Es gelten die Voraussetzungen und Notationen aus Satz 2.1.28, der Doob-Meyer-Zerlegung für nicht-negative Submartingale, wobei der mit dem Zählprozess assoziierte Punktprozess die Bedingung (iii) aus Bemerkung 2.1.4 erfülle.

Ist der Kompensator des Zählprozesses ein *stetiger* Prozess ist, dann gilt:

- (i) Das Martingal $M = (M_t, \mathcal{F}_t)_{t \in t \geq t_0}$ ist quadratisch-integrierbar und rechtsseitig stetig.
- (ii) Der Kompensator aus der Doob-Meyer-Zerlegung des Submartingals $(M_t^2, \mathcal{F}_t)_{t \in t \geq t_0}$ ist gegeben durch den Kompensator Λ des ursprünglichen Zählprozesses.

Beweis:

Der Beweis dieses überraschenden Resultates findet sich in Pang, Talreja und Whitt 2007, S. 265. Dazu wird explizit gezeigt, dass der Prozess $(M_t^2 - \Lambda_t)_{t \geq t_0}$ ein $\{\mathcal{F}_t\}_{t \geq t_0}$ -Martingal ist. Die Behauptung folgt dann aus der Eindeutigkeit der Doob-Meyer-Zerlegung. \square

Das abschließende Resultat des Abschnitts bildet eine Verallgemeinerung der bekannten *Itô-Isometrie*, die etwa in Kuo 2006, S. 48 formuliert wird. Wir möchten nicht näher auf das ursprüngliche Resultat eingehen und formulieren daher die nachfolgende Verallgemeinerung:

Satz 2.1.32. (*Itô-Isometrie für quadratisch integrierbare Martingale*, siehe Kuo 2006, S. 88)

Es sei $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq t_0}, P)$ ein filtrierter Wahrscheinlichkeitsraum und $(M_t, \mathcal{F}_t)_{t \geq t_0}$ ein quadratisch integrierbares, rechtsseitig stetiges Martingal, dessen linksseitige Grenzwerte

$$M_t^- = \lim_{s \uparrow t} M_s, \quad \text{für } t > t_0,$$

fast sicher existieren. Es bezeichne Λ den Kompensator des Submartingals $(M_t^2, \mathcal{F}_t)_{t \geq t_0}$ aus der Doob-Meyer-Zerlegung, siehe Satz 2.1.28.

Schließlich sei $(f_t)_{t \geq t_0}$ ein vorhersagbarer stochastischer Prozess mit

$$\mathbb{E} \left[\int_{t_0}^t |f_s|^2 d\Lambda(s) \right] < \infty, \quad \text{für alle } t \geq t_0. \quad (2.1.38)$$

Dann ist der stochastische Prozess $X = (X_t)_{t \geq t_0}$ mit

$$X_t = \int_{t_0}^t f_s dM(s), \quad \text{für } t \geq t_0,$$

ein $\{\mathcal{F}_t\}_{t \geq t_0}$ -Martingal und es gilt:

$$\mathbb{E} \left[|X_t|^2 \right] = \mathbb{E} \left[\left| \int_{t_0}^t f_s \, dM(s) \right|^2 \right] = \mathbb{E} \left[\int_{t_0}^t |f_s|^2 \, d\Lambda(s) \right]. \quad (2.1.39)$$

Die Gleichung (2.1.39) bezeichnen wir als *verallgemeinerte Itô-Isometrie*.

Beweis:

Der Beweis wird in Kuo 2006, Kapitel 4, für den Spezialfall der *Brownschen Bewegung* geführt. Dabei wird jedoch lediglich ausgenutzt, dass dieser stochastische Prozess ein quadratisch integrierbares, rechtsseitig stetiges Martingal bezüglich einer geeigneten Filtration ist. Der Beweis lässt sich daher auf die vorliegende Situation übertragen. \square

Obgleich das Resultat aus Satz 2.1.32 offensichtlich auch in einem allgemeineren Kontext gilt, werden wir uns meist im Rahmen der Situation aus Satz 2.1.31 bewegen. Wir beachten, dass ein aus der Doob-Meyer-Zerlegung eines adaptierten Zählprozesses gewonnenes Martingal gerade die Voraussetzungen der Itô-Isometrie für quadratisch integrierbare Martingale erfüllt.

In der nachfolgenden Bemerkung halten wir schließlich fest, weshalb die Formel aus Gleichung (2.1.39) als Isometrie bezeichnet wird.

Bemerkung 2.1.33.

In Anlehnung an die Notation aus Kuo 2006 bezeichnen wir die Menge aller vorhersagbaren Prozesse mit Indexmenge $[t_0, t^0]$, die der Bedingung aus (2.1.38) für alle $t \in [t_0, t^0]$ genügen, mit $L^2_{\text{pred}}([t_0, t^0]_{\Lambda} \times \Omega)$. Analog bezeichne $L^2(\Omega, \mathcal{F}, P)$ die Menge der quadratisch integrierbaren Zufallsvariablen über dem Wahrscheinlichkeitsraum (Ω, \mathcal{F}, P) .

Die für $t^0 > t_0$ definierten Abbildungen

$$\mathbb{I}_{t^0} : L^2_{\text{pred}}([t_0, t^0]_{\Lambda} \times \Omega) \longrightarrow L^2(\Omega, \mathcal{F}, P) : f \longmapsto \int_{t_0}^{t^0} f_s \, dM(s)$$

sind dann gemäß Gleichung (2.1.39) isometrisch.

2.1.8 Selbstanregende Punkt- und Zählprozesse

Zum Abschluss von Abschnitt 2.1 möchten wir die Definition eines *selbstanregenden Punktbeziehungsweise Zählprozesses* angeben. Im Beispiel 2.1.7 aus Abschnitt 2.1.3 haben wir den verschobenen Geburtsprozess bereits als einen *selbstanregenden Zählprozess* bezeichnet, ohne diesen Begriff zuvor mathematisch definiert zu haben.

Heuristisch gesprochen möchten wir einen Zählprozess als *selbstanregend* bezeichnen, wenn die bedingte Intensitätsfunktion, wie sie in Gleichung (2.1.37) aus Satz 2.1.30 definiert wird,

von der Vergangenheit des zugehörigen Zählprozesses abhängt:

In diesem Falle können auch Ereignisse wie das Auftreten des ersten Sprunges oder die Wartezeit zwischen dem ersten und dem zweiten Sprung das Verhalten des Zählprozesses selbst langfristig beeinflussen.

Infolgedessen wird in Snyder und Miller 1991, S. 288 die folgende Definition angegeben:

Es sei $t_0 \in \mathbb{R}$ und $(N_t)_{t \geq t_0}$ ein bezüglich einer Filtration $\{\mathcal{F}_t\}_{t \geq t_0}$ adaptierter Zählprozess, der die nachfolgenden Annahmen erfülle:

1. $(N_t)_{t \geq t_0}$ ist bedingt geordnet.
2. Für alle $t \geq t_0$ existiert der Grenzwert

$$\lambda_t = \lim_{\Delta t \rightarrow 0} \frac{P(\{N(t, t + \Delta t) = 1\} | \mathcal{F}_t^-)}{\Delta t}.$$

3. Es gilt $P(\{N_{t_0} = 0\}) = 1$.

Der Zählprozess heißt dann *selbstanregend im Sinne von Snyder & Miller*, wenn der stochastische Prozess $\lambda = (\lambda_t)_{t \geq t_0}$ abhängig von N ist.

Der Begriff des selbstanregenden Punktprozesses (engl. “*self-exciting point process*”) geht zurück auf Alan G. Hawkes, der diese Bezeichnung in Hawkes 1971 für eine Klasse von Zählprozessen einführte, die der obigen Heuristik genügen. Wir möchten diese heute eher als *Hawkes-Prozesse* bekannten Zählprozesse hier nicht näher diskutieren.

Beispiel 2.1.7. (6. Fortsetzung, *Schadensakkumulation*)

In der ersten Fortsetzung dieses Beispiels haben wir bereits das Prinzip der *Lastumverteilung* diskutiert. Da die im Sinne von *Snyder & Miller* selbstanregenden Zählprozesse eine Verallgemeinerung des verschobenen Geburtsprozesses aus Definition 2.1.9 darstellen, eignen sie sich folglich ebenfalls zur Modellierung dieses Sachverhaltes.

Wir möchten nun noch einmal an das zugrunde liegende Belastungsexperiment erinnern, bei dem ein Verbund von Drahtseilen zyklischen Belastungen ausgesetzt wurde:

Bei der Modellierung des mit den Drahtbrüchen assoziierten Zählprozesses durch einen reinen verschobenen Geburtsprozess mit Abbruch bei $M \in \mathbb{N}$ ist die Intensität der Belastung nach $n < M$ Drahtbrüchen *unabhängig* von der jeweiligen Dauer der Belastung in den vorherigen Stadien des Experimentes. Wir möchten diese Problematik anhand der Abbildung 5 verdeutlichen:

- In zwei Wiederholungen des Belastungsexperiments - *Experiment 1* und *Experiment 2* - wurde nach t_1 Lastzyklen der sechste Drahtbruch beobachtet. Die beobachteten Pfade der mit den Drahtbrüchen assoziierten Zählprozesse sind in Abbildung 5 in rot

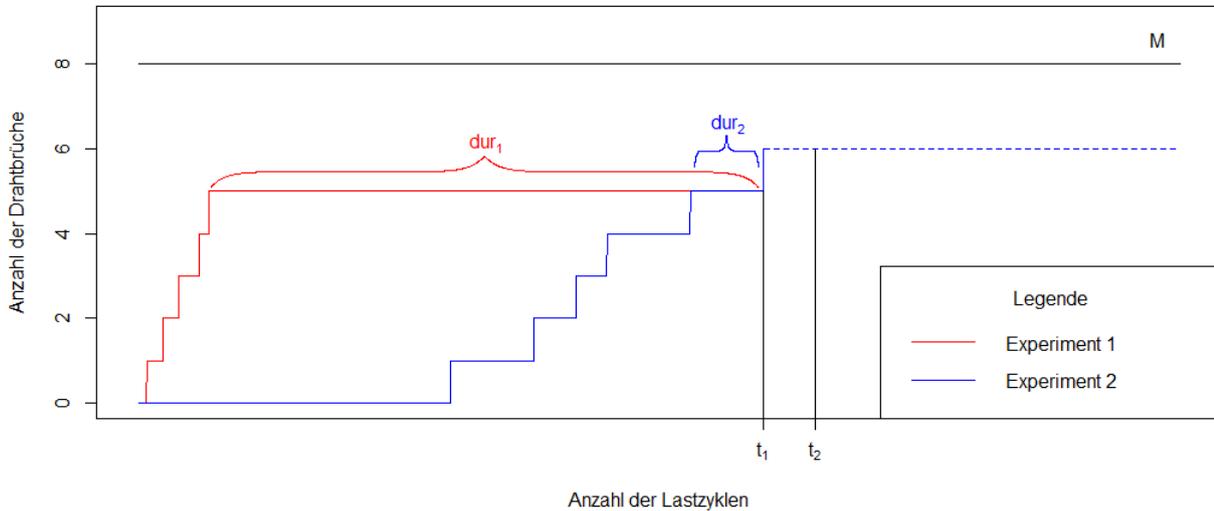


Abbildung 5: Plot zur Veranschaulichung des Prinzips der *Schadensakkumulation*. Diese Grafik wurde mit der statistischen Software *R* erstellt, siehe R Core Team 2015.

(Experiment 1) bzw. blau (Experiment 2) dargestellt.

- Uns interessiert nun die Wahrscheinlichkeit des Auftretens eines weiteren Drahtbruches zwischen dem t_1 -ten und dem t_2 -ten Lastzyklus. Diese Wahrscheinlichkeit lässt sich gemäß Bemerkung 2.1.29 und Satz 2.1.30 durch Integration der bedingten Intensitätsfunktion von t_1 bis t_2 ermitteln.
 - Im Falle eines reinen verschobenen Geburtsprozesses mit Abbruch bei $M = 8$ stimmen die bedingten Intensitätsfunktionen für die beiden Experimente nahe t_1 überein, da in beiden Versuchen der letzte Drahtbruch nach t_1 Lastzyklen erfolgte und die Funktion nur vom *letzten* Drahtbruch abhängt.
- ⇒ Wir erwarten folglich für beide Experimente die gleiche Anzahl Drahtbrüche zwischen dem t_1 -ten und dem t_2 -ten Lastzyklus.

Wir können der Abbildung 5 ferner entnehmen, dass im ersten Experiment die komplette Belastung über eine Dauer von dur_1 Lastzyklen auf die letzten drei verbleibenden Drahtseile ausgeübt wurde, bevor der sechste Drahtbruch nach t_1 Lastzyklen erfolgte. Analog wurde im zweiten Experiment die komplette Belastung über eine Dauer von $dur_2 \ll dur_1$ Lastzyklen von den verbleibenden drei Drahtseilen getragen.

Dieser Gedankengang führt uns zum Prinzip der *Schadensakkumulation*:

In beiden Experimenten bricht das sechste Drahtseil nach t_1 Lastzyklen, die verbleibenden zwei Drahtseile wurden jedoch im ersten Experiment einer insgesamt höheren Belastung ausgesetzt, als im zweiten Experiment. Wir erwarten daher eine Vorschädigung der verbleibenden Drahtseile im ersten Experiment durch die akkumulierte Belastung aus den Vorstadien des Versuches und somit eine im Vergleich zum zweiten Experiment erhöhte Wahrscheinlichkeit für einen weiteren Drahtbruch zwischen dem t_1 -ten und dem t_2 -ten Lastzyklus.

Dies ist jedoch mit obiger Folgerung *nicht* vereinbar, weshalb sich der verschobene Geburtsprozess als nicht geeignet für die Modellierung der Schadensakkumulation erweist.

Da ein selbstanregender Zählprozess im Sinne von *Snyder & Miller* auch von den *Wartezeiten* dur_1 und dur_2 abhängen darf, eignet sich erst die obige Verallgemeinerung zur Beschreibung sowohl der *Lastumverteilung* als auch der *Schadensakkumulation*.

Es ist nun unter Berücksichtigung der *Lastumverteilung* und der *Schadensakkumulation* anschaulich klar, dass unsere Definition der *selbstanregenden Zählprozesse* die obigen mit der Heuristik konform gehenden Prozesse beinhalten sollte.

Wir erachten die Definition im Sinne von *Snyder & Miller* jedoch als zu restriktiv, weshalb wir nachfolgend eine an Stute und Kopperschmidt 2013 angelehnte Definition verwenden möchten.

Definition 2.1.34. (*Selbstanregende Punkt- & Zählprozesse*, vgl. Stute und Kopperschmidt 2013, S. 1275)

Es sei $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq t_0}, P)$ ein filtrierter Wahrscheinlichkeitsraum, der die üblichen Bedingungen erfülle (siehe Anmerkungen im Vorfeld von Lemma 2.1.26).

Ein Zählprozess $N = (N_t)_{t \geq t_0}$ über $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq t_0}, P)$ heißt dann ein *selbstanregender Zählprozess*, falls N adaptiert bezüglich der Filtration $\{\mathcal{F}_t\}_{t \geq t_0}$ ist.

Analog heißt ein Punktprozess $T = (T_n)_{n \in \mathbb{N}}$ über $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq t_0}, P)$ ein *selbstanregender Punktprozess*, falls der assoziierte Zählprozess selbstanregend ist.

Die Motivation für diese Definition entstammt der Doob-Meyer-Zerlegung für nicht-negative Submartingale aus Satz 2.1.28:

Wir nennen demnach all jene Zählprozesse selbstanregend, für die eine Doob-Meyer-Zerlegung existiert. Dies ermöglicht es uns nachfolgend, die Dynamik eines Zählprozesses stets anhand des Kompensators aus der Doob-Meyer-Zerlegung zu modellieren.

Schließlich beinhaltet die obige Definition auch die selbstanregenden Zählprozesse im *Snyder & Miller'schen* Sinne, solange ein filtrierter Wahrscheinlichkeitsraum zugrunde liegt, der die üblichen Bedingungen erfüllt.

Hiermit möchten wir den Abschnitt über die Grundlagen der stochastischen Prozesse abschließen. Im nächsten Abschnitt wenden wir uns den *U-Statistiken* und einer Verallgemeinerung des *starken Gesetzes der großen Zahlen* zu.

2.2 U-Statistiken

2.2.1 Definition und das starke Gesetz der großen Zahlen für U-Statistiken

In diesem Abschnitt werden wir die sogenannten *U-Statistiken* einführen und ein *starkes Gesetz der großen Zahlen* für selbige formulieren. Unter Verwendung des *0-1-Gesetzes von Hewitt-Savage* wird es uns später möglich sein, mittels elementarer Martingaltheorie einen kurzen und eleganten Beweis eben jenes Resultates zu erbringen, welches später eine Vielzahl starker Konvergenzaussagen begründen wird (siehe auch Stute und Kopperschmidt 2013, S. 1294). Der Begriff der *U-Statistik* geht - in Anspielung auf die *Unverzerrtheit* dieser Statistiken - zurück auf *Wassily Hoeffding* (siehe Lee 1990, S. 8), der diese Bezeichnung in dem Artikel *A class of statistics with asymptotically normal distributions* aus dem Jahre 1948 einführte.

Wir gehen in diesem Abschnitt stets von einem zugrunde liegenden Wahrscheinlichkeitsraum (Ω, \mathcal{F}, P) aus, der keine weiteren Voraussetzungen erfüllen muss. Im Hinblick auf die Unterabschnitte 2.1.5 bis 2.1.8 werden wir jedoch in den meisten Fällen auf einen vollständigen Wahrscheinlichkeitsraum zurückgreifen.

Definition 2.2.1. (*U-Statistik*, siehe Hoeffding 1948)

Es sei $X = (X_n)_{n \in \mathbb{N}}$ eine Folge unabhängig identisch verteilter Zufallsvariablen mit Werten in einem Messraum (E, \mathcal{E}) und $\psi : (E^m, \mathcal{E}^m) \rightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ für ein $m \in \mathbb{N}$ eine messbare Funktion, wobei $E^m = E \times \dots \times E$ das m -fache kartesische Produkt von E und \mathcal{E}^m die entsprechende Produkt- σ -Algebra sei (an dieser Stelle sei erneut auf Bauer 1991, S. 58-67 verwiesen).

Ferner bezeichne (n, m) für $n \geq m$ die Menge aller injektiven Abbildungen von $\{1, \dots, m\}$ nach $\{1, \dots, n\}$, das heißt:

$$(n, m) = \{\tau : \{1, \dots, m\} \rightarrow \{1, \dots, n\} \mid \tau(i) \neq \tau(j) \text{ für } i \neq j, 1 \leq i, j \leq m\}.$$

Für $n \in \mathbb{N}$ mit $n \geq m$ ist dann die *U-Statistik* U_n definiert als $U_n = u_n \circ X$, wobei u_n wie folgt definiert sei:

$$u_n((x_k)_{k \in \mathbb{N}}) = \frac{(n-m)!}{n!} \sum_{\tau \in (n, m)} \psi(x_{\tau(1)}, \dots, x_{\tau(m)}).$$

Die Funktion ψ bezeichnen wir als m -dimensionalen *Kern* der U-Statistik U_n .

Bemerkung 2.2.2. Bekanntermaßen finden wir $\binom{n}{m}$ Möglichkeiten, m verschiedene Objekte aus einer n Objekte umfassenden Grundgesamtheit zu ziehen. Beachtet man auch die Reihenfolge der m gezogenen Objekte, so multipliziert sich die Anzahl dieser Möglichkeiten mit $m!$, der Anzahl aller Permutationen von m Objekten. Folglich gilt

$$|(n, m)| = m! \cdot \binom{n}{m} = \frac{n!}{(n-m)!},$$

so dass u_n das Mittel über alle Werte darstellt, die ψ auf den m -elementigen *geordneten* Teilmengen von $\{x_1, \dots, x_n\}$ annimmt (vergleiche auch Hoeffding 1948, S. 296).

Da u_n somit nur von den ersten n Folgengliedern x_1, \dots, x_n abhängt, schreiben wir häufig auch $u_n((x_k)_{k \in \mathbb{N}}) = u_n(x_1, \dots, x_n)$.

Insbesondere ist folglich u_n symmetrisch bezüglich seiner Argumente x_1, \dots, x_n , das heißt für $\tau \in (n, n)$ gilt:

$$u_n(x_{\tau(1)}, \dots, x_{\tau(n)}) = u_n(x_1, \dots, x_n). \quad (2.2.1)$$

Dieser Umstand motiviert schließlich den - von Erfolg gekrönten - Versuch, das *starke Gesetz der großen Zahlen für U-Statistiken* mit Hilfe des *0-1-Gesetzes von Hewitt-Savage* zu beweisen.

An dieser Stelle ist es uns bereits möglich, das *starke Gesetz der großen Zahlen für U-Statistiken* zu formulieren, obgleich wir die Voraussetzungen weiter unten noch etwas verschärfen müssen. Der Beweis wird jedoch weitere Vorarbeit benötigen und somit erst später erfolgen.

Satz 2.2.3. (*Starkes Gesetz der großen Zahlen für U-Statistiken*, siehe Lee 1990, S. 122)

Es gelte $E|\psi(X_1, \dots, X_m)| < \infty$. Dann konvergiert U_n für $n \rightarrow \infty$ fast sicher gegen $E(\psi(X_1, \dots, X_m))$.

Bemerkung 2.2.4.

Wir können das Resultat aus Satz 2.2.3 als Verallgemeinerung des bekannten *starken Gesetzes der großen Zahlen* betrachten (siehe Bauer 1991, S. 86):

Ist $(X_n)_{n \in \mathbb{N}}$ eine Folge unabhängig identisch verteilter Zufallsvariablen mit Werten in $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ und wählen wir den eindimensionalen Kern ψ mit

$$\psi : (\mathbb{R}, \mathcal{B}(\mathbb{R})) \longrightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R})) : x \mapsto x,$$

so erhalten wir gerade die U-Statistik

$$U_n = \frac{1}{n} \sum_{m=1}^n X_m,$$

das heißt U_n ist das arithmetische Mittel der ersten n Zufallsvariablen.

In diesem Falle stimmt das *starke Gesetz der großen Zahlen für U-Statistiken* mit dem gewöhnlichen *starken Gesetz der großen Zahlen* überein.

Wir möchten noch darauf hinweisen, dass die in diesem Abschnitt entwickelte Theorie primär dem Beweis von Satz 2.2.3 dient und keine praktische Relevanz für die restliche Arbeit besitzt.

Es sei dem Leser also freigestellt, sich mit dieser Formulierung des *starken Gesetzes der großen Zahlen für U-Statistiken* zu begnügen und den restlichen Abschnitt 2.2 zu überspringen.

2.2.2 Das 0-1-Gesetz von Hewitt-Savage

Im folgenden Unterabschnitt möchten wir das *0-1-Gesetz von Hewitt-Savage* beweisen, welches wir später für einen Beweis des *starken Gesetzes der großen Zahlen für U-Statistiken* utilisieren werden. Dazu bedarf es einiger topologischer und maßtheoretischer Vorbereitungen. Wir beginnen mit der Definition eines sogenannten *polnischen Raumes*.

Definition 2.2.5. (*Polnische Räume*, siehe Klenke 2013, S. 251)

Ein topologischer Raum (E, \mathcal{T}) heißt *polnischer Raum*, falls er vollständig metrisierbar und separabel ist.

Ein insbesondere im Rahmen dieser Arbeit wichtiges Beispiel eines solchen polnischen Raumes ist der sogenannte *Skorokhod-Raum* $\mathcal{D}(I)$ (s. auch Stute und Kopperschmidt 2013, S. 1294).

Definition 2.2.6. (*Der Skorokhod-Raum*, siehe Billingsley 1968, S. 109)

Es sei $I \subset \mathbb{R}$ ein kompaktes Intervall. Der *Skorokhod-Raum*²¹ $\mathcal{D}(I)$ ist die Menge aller *càdlàg*-Funktionen (von französisch *continue à droite, limite à gauche*, d.h. stetig von rechts mit Grenzwerten von links) auf dem Intervall I , also der Menge aller Funktionen $x : I \rightarrow \mathbb{R}$, für die gilt:

- Für alle $t \in [0, 1)$ existiert der Grenzwert $x(t^+) = \lim_{s \downarrow t} x(s)$ und es gilt $x(t^+) = x(t)$.
- Für alle $t \in (0, 1]$ existiert der Grenzwert $x(t^-) = \lim_{s \uparrow t} x(s)$.

Ist N ein Zählprozess, dessen Indexmenge das Intervall I enthält, so haben wir bereits in Abschnitt 2.1.2 gesehen, dass die Einschränkungen der Pfade von N auf das Intervall I im Raum $\mathcal{D}(I)$ liegen. Wir können also den entsprechend eingeschränkten Zählprozess -

²¹Tatsächlich ist es üblich, den topologischen Raum $(\mathcal{D}(I), \mathcal{T}_d)$, also $\mathcal{D}(I)$ versehen mit der Skorokhod-Topologie \mathcal{T}_d , als *Skorokhod-Raum* zu bezeichnen; für eine Definition siehe Satz 2.2.7.

ähnlich wie in Lemma 2.1.27 - als Skorokhod-Raum-wertige Zufallsvariable auffassen.

Um nun einen polnischen Raum zu erhalten, muss eine (metrisierbare) Topologie auf $\mathcal{D}(I)$ definiert werden. Der nächste Satz liefert das gewünschte Resultat; der Beweis ist jedoch länglich und rein analytisch, so dass an dieser Stelle auf selbigen verzichtet wird.

Satz 2.2.7. (*Skorokhod-Topologie*)

Es bezeichne Γ den Raum der streng monoton wachsenden Homöomorphismen von I auf sich selbst²². Für $\gamma \in \Gamma$ definieren wir:

$$\|\gamma\| = \sup_{s \neq t} \left| \log \frac{\gamma(t) - \gamma(s)}{t - s} \right|.$$

Setzen wir ferner

$$d(x, y) = \inf_{\gamma \in \Gamma} \max \{ \|\gamma\|, \sup_{t \in I} |x(t) - y(\gamma(t))| \},$$

so wird durch d eine Metrik auf $\mathcal{D}(I)$ definiert.

Bezeichnet \mathcal{T}_d die von d erzeugte *Skorokhod-Topologie* auf $\mathcal{D}(I)$, so ist $(\mathcal{D}(I), \mathcal{T}_d)$ ein polnischer Raum.

Beweis:

Für einen Beweis der Vollständigkeit siehe Billingsley 1968, S. 111-116, die Separabilität wird etwa in Ethier und Kurtz 1986, S. 121 f. für den Fall $I = [0, \infty)$ gezeigt. Der Beweis geht jedoch auf Kolmogorov zurück (siehe Kolmogorov 1956) und wird hier in größerer Allgemeinheit behandelt (siehe Ethier und Kurtz 1986, S. 154); motivierend vergleiche man auch Stute und Kopperschmidt 2013, S. 1294. \square

Bemerkung 2.2.8.

Die in 2.2.7 definierte *Skorokhod-Topologie* wird von Skorokhod selbst als die Topologie J_1 auf $\mathcal{D}(I)$ bezeichnet (s. Skorokhod 1956, S. 265)²³. Diese Topologie wird von der Metrik \tilde{d} , die durch

$$\tilde{d}(x, y) = \inf_{\gamma \in \Gamma} \left(\sup_{t \in I} |\gamma(t) - t| + \sup_{t \in I} |x(t) - y(\gamma(t))| \right)$$

definiert wird, erzeugt (siehe Jakubowski 1999, S. 4, ebenso Billingsley 1968, S. 111). Diese Metrik erlaubt eine einfachere Interpretation:

Während sich die Supremumsnorm (mit welcher der Unterraum $C(I)$ von $\mathcal{D}(I)$ zu einem Banachraum wird) im wahrscheinlichkeitstheoretischen Kontext als ungeeignet herausstellt, erscheint die Metrik \tilde{d} als geeignetes Substitut. In dieser Metrik sind zwei Funk-

²²Das heißt alle streng monoton wachsenden, stetigen Abbildungen γ von $I = [I_{\min}, I_{\max}]$ nach I mit $\gamma(I_{\min}) = I_{\min}$ und $\gamma(I_{\max}) = I_{\max}$.

²³Skorokhod betrachtet auch die Topologien J_2 , M_1 und M_2 , die im Rahmen dieser Arbeit nicht behandelt werden (siehe erneut Skorokhod 1956 und Jakubowski 1999).

tion $x, y \in \mathcal{D}(I)$ dann nah beieinander, wenn sie bei einer hinreichend kleinen Störung der Zeit durch $\gamma \in \Gamma$ gleichmäßig (d.h. in der Supremumsnorm) nah beieinander liegen - diese Modifikation des gleichmäßigen Abstandsbegriffes erhält die Eigenschaft der Separabilität für den Raum $\mathcal{D}(I)$, dessen Elemente Unstetigkeitsstellen aufweisen dürfen.

Der Übergang von der Metrik \tilde{d} zur äquivalenten (s. Billingsley 1968, S. 114) Metrik d aus Satz 2.2.7 ist schließlich notwendig, da zwar die erzeugten Topologien übereinstimmen, $\mathcal{D}(I)$ jedoch nur mit der Metrik d zu einem *vollständigen* metrischen Raum wird.

Zukünftig werden wir nun unabhängig identisch verteilte Zufallsvariablen X_n , $n \in \mathbb{N}$, die Werte im Skorokhod-Raum $(\mathcal{D}(I), \mathcal{T}_d)$ oder allgemeiner in einem polnischen Raum annehmen, untersuchen. Im Wertebereich wird dabei stets die von der Topologie (also im konkreten Fall von \mathcal{T}_d) erzeugte, sogenannte *Borel'sche σ -Algebra* betrachtet. Kopperschmidt und Stute bemerken in dem Artikel *The Statistical Analysis of Self-Exciting Point Processes*, dass in dieser Situation bekannte Argumente für reellwertige Zufallsvariablen angewandt werden können (siehe Stute und Kopperschmidt 2013, S. 1294). Die Grundlage dieser Behauptung liefert dabei der - den Begriff eines *Borel'schen Standardraumes* motivierende (siehe Kechris 1995, S. 74) - Satz 2.2.10, dem wir eine Definition voranstellen möchten.

Definition 2.2.9. (*Messraum-Isomorphismen*, siehe Klenke 2013, S. 189)

Zwei Messräume (E, \mathcal{E}) und (E', \mathcal{E}') heißen *isomorph*, falls es eine bijektive Abbildung $\phi : E \rightarrow E'$ gibt, sodass ϕ \mathcal{E} - \mathcal{E}' -messbar und ϕ^{-1} \mathcal{E}' - \mathcal{E} -messbar ist. ϕ heißt dann ein *Messraum-Isomorphismus*.

Definition 2.2.9 erlaubt nun die Formulierung von Satz 2.2.10. Ein Raum, der die Voraussetzungen dieses Satzes erfüllt, wird auch *Borel'scher Standardraum* genannt.

Satz 2.2.10. (siehe Klenke 2013, S. 189)

Es sei (E, \mathcal{T}) ein polnischer Raum und \mathcal{E} die von \mathcal{T} erzeugte Borel'sche σ -Algebra. Wie zuvor bezeichne $\mathcal{B}(\mathbb{R})$ die Borel'sche σ -Algebra über den reellen Zahlen. Dann gibt es eine Borel'sche Menge $B \in \mathcal{B}(\mathbb{R})$, sodass (E, \mathcal{E}) und $(B, \mathcal{B}(B))$ isomorph sind.

Beweis:

Der Satz 2.2.10 ist eine unmittelbare Konsequenz von *Theorem 13.1.1* aus Dudley 2002. Die für unsere Situation angepasste Formulierung dieses Theorems lautet (s. Dudley 2002, S. 487):

Es seien (E, \mathcal{T}_E) und $(E', \mathcal{T}_{E'})$ zwei polnische Räume mit den erzeugten Borel'schen σ -Algebren \mathcal{E} bzw. \mathcal{E}' . Es seien ferner $\Omega \in \mathcal{E}$ und $\Omega' \in \mathcal{E}'$. Dann sind $(\Omega, \mathcal{E}|_{\Omega})$ und $(\Omega', \mathcal{E}'|_{\Omega'})$ isomorph genau dann, wenn Ω und Ω' die gleiche Kardinalität (die entweder endlich, abzählbar oder die Kardinalität des Kontinuums $[0, 1]$ ist) aufweisen.

Die Behauptung folgt dann sofort, indem man $\Omega = E$ und $\Omega' = B \in \mathcal{B}(\mathbb{R})$ mit der entsprechenden Kardinalität wählt und beachtet, dass \mathbb{R} mit der üblichen Topologie bzw. Metrik ein polnischer Raum mit erzeugter σ -Algebra $\mathcal{B}(\mathbb{R})$ ist.

Der Beweis des zitierten Theorems findet sich ebenfalls in Dudley 2002 auf den Seiten 487 bis 492 und erfordert weitere Vorüberlegungen. Wir verzichten deshalb auf eine weitere Ausführung dieses Beweises. \square

Bemerkung 2.2.11.

Auf Grundlage des Satzes 2.2.10 folgen nun zwei wichtige Bemerkungen:

- Da Satz 2.2.10 auf *Theorem 13.1.1* aus Dudley 2002 beruht und selbiges nur Anforderungen an die Kardinalität von $B \in \mathcal{B}(\mathbb{R})$ stellt, können wir im Falle des Skorokhod-Raumes $\mathcal{D}(I)$ einen Isomorphismus ϕ auf $B = I$ oder $B = \mathbb{R}$ finden.
- Es seien nun X_n für $n \in \mathbb{N}$ unabhängig identisch verteilte Zufallsvariablen mit Werten in einem überabzählbaren polnischen Raum E (vergleiche dazu die Ausführungen im Anschluss an die Bemerkung 2.2.8). Gemäß Satz 2.2.10 und dem ersten Teil dieser Bemerkung finden wir dann einen Messraum-Isomorphismus $\phi : E \rightarrow \mathbb{R}$. Durch Übergang zu den reellwertigen Zufallsvariablen $Y_n = \phi \circ X_n$ können wir dann bekannte Argumente für reellwertige Zufallsvariablen anwenden.

Ein erstes Beispiel für die Anwendbarkeit des in Bemerkung 2.2.11 geschilderten Vorgehens und zugleich das zentrale Resultat dieses Unterabschnittes liefert das *0-1-Gesetz von Hewitt-Savage*. Um diesen Satz zu formulieren, führen wir zunächst den Begriff der *permutierbaren Menge* ein.

Definition 2.2.12. (*Permutierbare Mengen*, vgl. Bauer 1991, S. 79f.)

Es sei $(X_n)_{n \in \mathbb{N}}$ eine Folge von Zufallsvariablen auf einem Wahrscheinlichkeitsraum (Ω, \mathcal{F}, P) mit Werten im Messraum (E, \mathcal{E}) . Die Folge definiert die Zufallsvariable

$$X = \bigotimes_{n=1}^{\infty} X_n$$

mit Werten im Messraum $(E^{\mathbb{N}}, \mathcal{E}^{\mathbb{N}})$, wobei X das übliche Produkt der Zufallsvariablen X_n , $E^{\mathbb{N}} = \prod_{n \in \mathbb{N}} E$ und $\mathcal{E}^{\mathbb{N}}$ die entsprechende Produkt- σ -Algebra sei (vgl. erneut [Bauer 1991,

S. 58-67]).

Eine Menge $A \in \mathcal{E}^{\mathbb{N}}$ heißt *permutierbar bezüglich X* , wenn für jede endliche Permutation τ gilt:

$$\{\tau X \in A\} = \{X \in A\} \quad (2.2.2)$$

Eine Bijektion $\tau : \mathbb{N} \rightarrow \mathbb{N}$ werde dabei als *endliche Permutation* bezeichnet, wenn ein $n_0 \in \mathbb{N}$ existiert, so dass $\tau(n) = n$ für alle $n \geq n_0$ gilt.

Mit dieser Definition können wir nun das *0-1-Gesetz von Hewitt-Savage* - zunächst für den Fall reeller Zufallsvariablen - formulieren und beweisen.

Satz 2.2.13. (*0-1-Gesetz von Hewitt-Savage, reelle Version*, siehe Bauer 1991, S. 80)

Es sei $(X_n)_{n \in \mathbb{N}}$ eine Folge unabhängig identisch verteilter reeller Zufallsvariablen, d.h. es gelte $E = \mathbb{R}$, $\mathcal{E} = \mathcal{B}(\mathbb{R})$ (siehe Definition 2.2.12). Man definiere $\mathcal{B} = (\mathcal{B}(\mathbb{R}))^{\mathbb{N}}$. Dann gilt für jede bezüglich X (siehe erneut Definition 2.2.12) permutierbare Menge $A \in \mathcal{B}$

$$P(\{X \in A\}) = 0 \quad \vee \quad P(\{X \in A\}) = 1.$$

Beweis²⁴:

Es bezeichne $\pi_n : \mathbb{R}^{\mathbb{N}} \rightarrow \mathbb{R}$ die n -te kanonische Projektion, wobei $n \in \mathbb{N}$ sei. Gemäß der Definition der Produkt- σ -Algebra $\mathcal{B} = (\mathcal{B}(\mathbb{R}))^{\mathbb{N}}$ gilt dann $\mathcal{B} = \sigma(\pi_n : n \in \mathbb{N})$ (siehe Bauer 1991, S. 60).

Setzen wir $\mathcal{A}_n = \sigma(\pi_1, \dots, \pi_n)$, so erzeugt die Algebra

$$\mathcal{A}_{\infty} = \bigcup_{n=1}^{\infty} \mathcal{A}_n$$

gerade die σ -Algebra \mathcal{B} . Daher können wir eine Approximationseigenschaft aus der Maßtheorie nutzen (siehe Bauer 1992, S. 29, *Satz 5.7*):

Es sei μ ein endliches Maß auf einer σ -Algebra \mathcal{A} in Ω , die von einer Algebra \mathcal{A}_{∞} in Ω erzeugt wird. Dann existiert zu jedem $A \in \mathcal{A}$ eine Folge $(C_k)_{k \in \mathbb{N}}$ in \mathcal{A}_{∞} mit

$$\lim_{k \rightarrow \infty} \mu(A \triangle C_k) = 0. \quad (2.2.3)$$

Dabei bezeichne \triangle die symmetrische Differenz zweier Mengen A und B , d.h. $A \triangle B = (A \setminus B) \cup (B \setminus A)$.

²⁴Der hiesige Beweise entstammt ebenfalls dem Lehrbuch Bauer 1991 und wurde um einige Erläuterungen erweitert.

Unter Verwendung von (2.2.3) erhalten wir in dieser Situation also zu $A \in \mathcal{B}$ eine Folge $(C_k)_{k \in \mathbb{N}}$ mit

$$\lim_{k \rightarrow \infty} P(\{X \in A \Delta C_k\}) = \lim_{k \rightarrow \infty} P_X(A \Delta C_k) = 0, \quad (2.2.4)$$

wobei P_X wie üblich das Bildmaß von P unter X bezeichne, welches als Wahrscheinlichkeitsmaß insbesondere endlich ist. Wegen $C_k \in \mathcal{A}_\infty$ und $\mathcal{A}_1 \subset \mathcal{A}_2 \subset \dots$ existiert dann zu $k \in \mathbb{N}$ ein $n_k \in \mathbb{N}$, sodass $C_k \in \mathcal{A}_{n_k}$ ist²⁵. Nach Definition der σ -Algebra \mathcal{A}_{n_k} existiert somit eine Borel'sche Menge $B_k \in (\mathcal{B}(\mathbb{R}))^{n_k}$ mit

$$C_k = \{(\pi_1, \dots, \pi_{n_k}) \in B_k\}. \quad (2.2.5)$$

Für $k \in \mathbb{N}$ sei nun τ_k die endliche Permutation mit $\tau_k(n) = n$ für $n > 2n_k$ und

$$\tau_k(n) = \begin{cases} n + n_k, & \text{für } n = 1, \dots, n_k, \\ n - n_k, & \text{für } n = n_k + 1, \dots, 2n_k. \end{cases}$$

Ist nun die Menge $A \in \mathcal{B}$ permutierbar bzgl. X , so erhalten wir nach Definition 2.2.12:

$$\{\tau_k X \in A\} = \{X \in A\}. \quad (2.2.6)$$

Motiviert durch Gleichung (2.2.5) definieren wir

$$\tilde{C}_k = \{(\pi_{\tau_k(1)}, \dots, \pi_{\tau_k(n_k)}) \in B_k\} = \{(\pi_{n_k+1}, \dots, \pi_{2n_k}) \in B_k\}. \quad (2.2.7)$$

Aus der Definition der \tilde{C}_k in Gleichung (2.2.7) wird sofort klar, dass $\tilde{C}_k \in \sigma(\pi_n : n \geq n_k + 1)$ ist, während $C_k \in \mathcal{A}_{n_k} = \sigma(\pi_n : n \leq n_k)$ gilt. Da $(X_n)_{n \in \mathbb{N}}$ eine Folge unabhängiger Zufallsvariablen (bzgl. P) ist, ist $(\pi_n)_{n \in \mathbb{N}}$ eine Folge unabhängiger Zufallsvariablen bezüglich des Bildmaßes P_X .

Daher sind C_k und \tilde{C}_k unabhängig²⁶ und somit gilt:

$$P_X(C_k \cap \tilde{C}_k) = P_X(C_k) \cdot P_X(\tilde{C}_k). \quad (2.2.8)$$

Wiederum da $(X_n)_{n \in \mathbb{N}}$ eine Folge unabhängig *identisch* verteilter Zufallsvariablen ist, besitzen X und $\tau_k X$ dieselbe Verteilung P_X - heuristisch gesprochen können wir alle X_n als gleichwertig betrachten, sodass eine Permutation derselbigen keine Veränderung der Verteilung zur Folge hat - und die Gleichungen (2.2.5) und (2.2.7) liefern:

²⁵Die Folge $(n_k)_{k \in \mathbb{N}}$ kann zudem streng monoton steigend gewählt werden.

²⁶Für eine weitere Ausführung dieses Arguments siehe Bauer 1991, S. 45, Satz 6.5.

$$\begin{aligned}
\{X \in C_k\} &= \{(\pi_1 \circ X, \dots, \pi_{n_k} \circ X) \in B_k\} = \{(X_1, \dots, X_{n_k}) \in B_k\}, \\
\{X \in \tilde{C}_k\} &= \{(\pi_{\tau_k(1)} \circ X, \dots, \pi_{\tau_k(n_k)} \circ X) \in B_k\} = \{(X_{\tau_k(1)}, \dots, X_{\tau_k(n_k)}) \in B_k\}.
\end{aligned} \tag{2.2.9}$$

Aus Gleichung (2.2.9) liest man sofort ab, dass damit

$$\{\tau_k X \in C_k\} = \{X \in \tilde{C}_k\}$$

gilt. Mit Gleichung (2.2.6) erhalten wir deshalb

$$\{\tau_k X \in A \Delta C_k\} = \{X \in A \Delta \tilde{C}_k\}$$

und wegen der gleichen Verteilung von X und $\tau_k X$ schließlich für alle $k \in \mathbb{N}$:

$$P_X(A \Delta C_k) = P_X(A \Delta \tilde{C}_k) \tag{2.2.10}$$

Nun beachten wir, dass folgende mengentheoretische Inklusion gültig ist:

$$\begin{aligned}
A \Delta (C_k \cap \tilde{C}_k) &= A \setminus (C_k \cap \tilde{C}_k) \cup (C_k \cap \tilde{C}_k) \setminus A \\
&= A \setminus C_k \cup A \setminus \tilde{C}_k \cup (C_k \cap \tilde{C}_k) \setminus A \\
&\subset A \setminus C_k \cup A \setminus \tilde{C}_k \cup (C_k \cup \tilde{C}_k) \setminus A \\
&= A \setminus C_k \cup A \setminus \tilde{C}_k \cup C_k \setminus A \cup \tilde{C}_k \setminus A \\
&= A \Delta C_k \cup A \Delta \tilde{C}_k.
\end{aligned}$$

Daher folgt mit Gleichung (2.2.10)

$$P_X\left(A \Delta (C_k \cap \tilde{C}_k)\right) \leq P_X(A \Delta C_k) + P_X(A \Delta \tilde{C}_k) = 2P_X(A \Delta C_k),$$

und deshalb nach Wahl der Folge $(C_k)_{k \in \mathbb{N}}$ (siehe Gleichung (2.2.4)):

$$\lim_{k \rightarrow \infty} P_X\left(A \Delta (C_k \cap \tilde{C}_k)\right) = 0. \tag{2.2.11}$$

Aus Gleichung (2.2.11) folgt nun unmittelbar

$$\lim_{k \rightarrow \infty} P_X(C_k \cap \tilde{C}_k) = P_X(A). \tag{2.2.12}$$

Um dies einzusehen, beachte man folgendes Argument (siehe dazu Bauer 1992, S. 30):

Ist μ ein endliches Maß auf einer σ -Algebra \mathcal{A} und sind $C, A \in \mathcal{A}$, dann gilt

$$|\mu(C) - \mu(A)| \leq \mu(A \Delta C) . \quad (2.2.13)$$

Die Gültigkeit dieser Ungleichung wird wie folgt ersichtlich:

$$\begin{aligned} \mu(C) - \mu(A) &\leq \mu(A \cup (C \setminus A)) - \mu(A) \\ &= \mu(A) + \mu(C \setminus A) - \mu(A) \\ &= \mu(C \setminus A) \leq \mu(C \setminus A \cup A \setminus C) \\ &= \mu(A \Delta C) , \end{aligned}$$

und völlig analog

$$\mu(A) - \mu(C) \leq \mu(A \Delta C) .$$

Mit Gleichung (2.2.13) folgt dann einerseits Gleichung (2.2.12) und mit selbigem Argument unter Verwendung von (2.2.10) andererseits auch

$$P_X(A) = \lim_{k \rightarrow \infty} P_X(C_k) = \lim_{k \rightarrow \infty} P_X(\tilde{C}_k) .$$

Schlussendlich erhalten wir dann mit Gleichung (2.2.8):

$$\begin{aligned} P_X(A) &= \lim_{k \rightarrow \infty} P_X(C_k \cap \tilde{C}_k) = \lim_{k \rightarrow \infty} (P_X(C_k) \cdot P_X(\tilde{C}_k)) \\ &= \left(\lim_{k \rightarrow \infty} P_X(C_k) \right) \cdot \left(\lim_{k \rightarrow \infty} P_X(\tilde{C}_k) \right) = P_X(A) \cdot P_X(A) . \end{aligned}$$

Hieraus folgt $P(\{X \in A\}) = P_X(A) \in \{0, 1\}$ und somit die Behauptung. \square

Wie der obige Beweis zeigt, ist es prinzipiell nicht notwendig, sich auf den Spezialfall $E = \mathbb{R}$ und $\mathcal{E} = \mathcal{B}(\mathbb{R})$ zu beschränken:

Um die Approximationseigenschaft aus Gleichung (2.2.3) anzuwenden, genügt es bereits, dass die Folge $(X_n)_{n \in \mathbb{N}}$ unabhängig identisch verteilter Zufallsvariablen Werte in einem überabzählbaren polnischen Raum (E, \mathcal{T}) annimmt. Da ein polnischer Raum dem *zweiten Abzählbarkeitsaxiom* genügt, entspricht dann nämlich die von der Produkttopologie $\mathcal{T}^{\mathbb{N}}$ erzeugte Borel'sche σ -Algebra gerade dem abzählbaren Produkt der von \mathcal{T} erzeugten Borel'schen σ -Algebra (siehe dafür Kerr und Li 2014, S. 211, *Lemma A.15*, oder auch Srivastava 1998, S. 88), so dass die Approximationseigenschaft ihre Gültigkeit behält.

Das nächste Korollar soll nun jedoch Bemerkung 2.2.11 nutzen, um das *0-1-Gesetz von Hewitt-Savage* auf diese allgemeinere Situation anzuwenden.

Korollar 2.2.14. (0-1-Gesetz von Hewitt-Savage, allgemeine Version)

Es sei $(X_n)_{n \in \mathbb{N}}$ eine Folge unabhängig identisch verteilter Zufallsvariablen mit Werten in einem überabzählbaren polnischen Raum (E, \mathcal{T}) mit von der Topologie erzeugter Borel'scher σ -Algebra \mathcal{E} . Dann gilt für jede bezüglich X permutierbare Menge $A \in \mathcal{E}^{\mathbb{N}}$

$$P(\{X \in A\}) = 0 \quad \vee \quad P(\{X \in A\}) = 1.$$

Beweis²⁷:

Es sei $A \in \mathcal{E}^{\mathbb{N}}$ permutierbar bezüglich X . Ferner sei $\phi : (E, \mathcal{E}) \rightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ der Messraum-Isomorphismus aus Bemerkung 2.2.11 und wir definieren die reellen Zufallsvariablen $Y_n = \phi \circ X_n$ für $n \in \mathbb{N}$. Offensichtlich bildet dann $(Y_n)_{n \in \mathbb{N}}$ eine Folge von unabhängig identisch verteilten *reellen* Zufallsvariablen.

Wir stellen fest, dass $\mathcal{E}^{\mathbb{N}}$ die von der Produkttopologie $\mathcal{T}^{\mathbb{N}}$ erzeugte Borel'sche σ -Algebra ist und dass

$$\Phi = \bigotimes_{n=1}^{\infty} \phi : (E^{\mathbb{N}}, \mathcal{E}^{\mathbb{N}}) \rightarrow (\mathbb{R}^{\mathbb{N}}, (\mathcal{B}(\mathbb{R}))^{\mathbb{N}})$$

ein Messraum-Isomorphismus ist. Die Gültigkeit dieser Aussage beruht auf den Ausführungen im Vorfeld dieses Korollars, da folglich die so definierte Abbildung Φ genau dann ein Messraum-Isomorphismus ist, wenn bereits ϕ ein Messraum-Isomorphismus ist (siehe Srivastava 1998, S. 89, *Proposition 3.1.29*). Vergleiche dafür auch *Lemma 7.10* in Biskup 2014.

Setzen wir $\tilde{A} = \Phi(A) \in (\mathcal{B}(\mathbb{R}))^{\mathbb{N}}$, so gilt mit $Y = \bigotimes_{n=1}^{\infty} Y_n$ für eine beliebige endliche Permutation τ :

$$\begin{aligned} \{\tau Y \in \tilde{A}\} &= \{(Y_{\tau(1)}, Y_{\tau(2)}, \dots) \in \Phi(A)\} = \{(\phi \circ X_{\tau(1)}, \phi \circ X_{\tau(2)}, \dots) \in \Phi(A)\} \\ &= \{\Phi((X_{\tau(1)}, X_{\tau(2)}, \dots)) \in \Phi(A)\} = \{(X_{\tau(1)}, X_{\tau(2)}, \dots) \in A\} \\ &= \{\tau X \in A\} = \{X \in A\} \stackrel{\text{analog}}{=} \{Y \in \tilde{A}\}. \end{aligned} \tag{2.2.14}$$

Folglich ist \tilde{A} permutierbar bzgl. Y und das *0-1-Gesetz von Hewitt-Savage* im reellen Fall liefert mit Gleichung (2.2.14)

$$P(\{X \in A\}) = P(\{Y \in \tilde{A}\}) \in \{0, 1\},$$

und damit die Behauptung. □

²⁷Dieser Beweis wurde vom Autor dieser Arbeit selbst geführt.

2.2.3 Rückwärtsmartingale und der Doobsche Konvergenzsatz

Neben der Verwendung des *0-1-Gesetzes von Hewitt-Savage* wird der Beweis des *starken Gesetzes der großen Zahlen für U-Statistiken* Nutzen aus einem weiteren Stützpfiler der Wahrscheinlichkeitstheorie ziehen:

Dem *Doobschen Konvergenzsatz für Rückwärtsmartingale*.

Obgleich wir an dieser Stelle auf den - üblicherweise im Rahmen einer Stochastik-Vorlesung geführten - Beweis dieses Satzes verzichten möchten, da er nur indirekt einen tieferen Einblick in die Thematik ermöglicht, möchten wir dem Leser dennoch eine kurze Wiederholung der elementaren Konzepte anbieten.

Im Anschluss an diese Wiederholung werden wir außerdem im nächsten Unterabschnitt zeigen, dass die Folge der U-Statistiken $(U_n)_{n \geq m}$ (siehe Definition 2.2.1) bezüglich einer geeigneten Folge von σ -Algebren (und unter einer vergleichsweise schwachen zusätzlichen Forderung) zu einem *Rückwärtsmartingal* wird. Zunächst gilt es jedoch, diesen Begriff zu definieren. Wir erinnern dazu an die Definition eines Martingals aus Abschnitt 2.1.5, siehe Definition 2.1.24.

Definition 2.2.15. (*Rückwärtsmartingal*, vgl. Lee 1990, S. 111)

Es sei $(X_n)_{n \in \mathbb{N}}$ eine Folge von Zufallsvariablen über dem Wahrscheinlichkeitsraum (Ω, \mathcal{F}, P) , die bzgl. einer Folge von Unter- σ -Algebren $\{\mathcal{F}_n\}_{n \in \mathbb{N}}$ adaptiert ist, d.h. es gelte $\sigma(X_n) \subset \mathcal{F}_n$ für alle $n \in \mathbb{N}$ (vergleiche auch Definition 2.1.21).

Dann heißt die Folge $(X_n, \mathcal{F}_n)_{n \in \mathbb{N}}$ ein *Rückwärtsmartingal*, falls die nachfolgenden Bedingungen erfüllt sind:

- (i) Es ist $\mathcal{F} \supset \mathcal{F}_{n_0} \supset \mathcal{F}_n$ für alle $n, n_0 \in \mathbb{N}$ mit $n \geq n_0$.
- (ii) Die Zufallsvariablen X_n sind alle integrierbar, d.h. es ist $E|X_n| < \infty$ für alle $n \in \mathbb{N}$.
- (iii) Es ist $E(X_n | \mathcal{F}_{n_0}) = X_{n_0}$ fast sicher für alle $n, n_0 \in \mathbb{N}$ mit $n \leq n_0$.

Bemerkung 2.2.16.

Man kann Rückwärtsmartingale auch als *negativ indizierte Martingale* auffassen:

Ist nämlich die Folge $(X_n, \mathcal{F}_n)_{n \in \mathbb{N}}$ ein Rückwärtsmartingal, so ist die negativ indizierte Folge $(\tilde{X}_n, \tilde{\mathcal{F}}_n)_{n \in -\mathbb{N}}$ mit $\tilde{X}_n = X_{-n}$ und $\tilde{\mathcal{F}}_n = \mathcal{F}_{-n}$ ein Martingal, vergleiche auch Bauer 1991, S. 168. Insbesondere ist die Folge $\{\tilde{\mathcal{F}}_n\}_{n \in \mathbb{N}}$ eine Filtration der σ -Algebra \mathcal{F} .

Obgleich der Begriff des (Rückwärts-)Martingals ein abstraktes Konzept ist, werden wir uns im Folgenden speziell den *reellwertigen* Martingalen zuwenden, also solchen Martingalen, für die $(X_n)_{n \in \mathbb{N}}$ eine Folge *reellwertiger* Zufallsvariablen ist. Dies stellt für das weitere Vorgehen offensichtlich keine Einschränkung dar, da die Folge von U-Statistiken $(U_n)_{n \geq m}$

eine Folge *reeller* Zufallsvariablen ist.

Wir formulieren nun der Vollständigkeit halber die *Doobschen Konvergenzsätze*.

Für uns ist insbesondere der *Zweite Doobsche Konvergenzsatz* für Rückwärtsmartingale von Interesse.

Satz 2.2.17. (*Erster Doobscher Konvergenzsatz*, siehe Bauer 1991, S. 162)

Es sei $(X_n, \mathcal{F}_n)_{n \in \mathbb{N}}$ ein reellwertiges Martingal, das der Bedingung

$$\sup_{n \in \mathbb{N}} \mathbb{E} |X_n| < \infty \quad (2.2.15)$$

genüge. Ferner definieren wir die σ -Algebra $\mathcal{F}_\infty^{(+)}$ durch

$$\mathcal{F}_\infty^{(+)} = \sigma \left(\bigcup_{n \in \mathbb{N}} \mathcal{F}_n \right).$$

Dann existiert eine integrierbare und $\mathcal{F}_\infty^{(+)}$ -messbare reelle Zufallsvariable X_∞ , sodass X_n für $n \rightarrow \infty$ fast sicher gegen X_∞ konvergiert.

Beweis:

Eine ausführliche Einführung in die Theorie der Martingale und einen detaillierten Beweis von Satz 2.2.17 findet man in Bauer 1991, siehe speziell Seite 162f. \square

Satz 2.2.18. (*Zweiter Doobscher Konvergenzsatz*, siehe Bauer 1991, S. 172)

Es sei $(X_n, \mathcal{F}_n)_{n \in \mathbb{N}}$ ein reellwertiges Rückwärtsmartingal²⁸, das der Bedingung (2.2.15) aus Satz 2.2.17 genüge. Ferner definieren wir die σ -Algebra $\mathcal{F}_\infty^{(-)}$ durch

$$\mathcal{F}_\infty^{(-)} = \bigcap_{n \in \mathbb{N}} \mathcal{F}_n.$$

Dann existiert eine integrierbare und $\mathcal{F}_\infty^{(-)}$ -messbare reelle Zufallsvariable X_∞ , sodass X_n für $n \rightarrow \infty$ fast sicher gegen X_∞ konvergiert.

Beweis:

Für einen Beweis sei erneut auf Bauer 1991 verwiesen. Der auf Seite 172 formulierte und bewiesene Satz wurde hierbei in seiner Allgemeinheit etwas vermindert, um die vorliegende Situation präziser zu reflektieren. \square

²⁸Für die in Bauer 1991 verwendete Notation siehe auch Bemerkung 2.2.16.

2.2.4 U-Statistiken als Rückwärtsmartingal

Als letzte Vorbereitung für den Beweis des *starken Gesetzes der großen Zahlen für U-Statistiken* zeigen wir, dass eine Folge $(U_n)_{n \geq m}$ von U-Statistiken unter geeigneten Voraussetzungen zu einem Rückwärtsmartingal wird. In Verbindung mit dem *zweiten Doob'schen Konvergenzsatz* wird dies zunächst die (fast sichere) Konvergenz der Folge sicherstellen, um anschließend mittels des *0-1-Gesetzes von Hewitt-Savage* die (ebenso fast sichere) Konstanz des Grenzwerts zu folgern.

Lemma 2.2.19. (siehe Lee 1990, S. 118)

Es sei $(U_n)_{n \geq m}$ eine Folge von U-Statistiken mit gemeinsamem Kern ψ , der die Bedingung

$$\mathbb{E} |\psi(X_1, \dots, X_m)| < \infty \quad (2.2.16)$$

erfülle. Außerdem sei eine Folge $(\mathcal{F}_n)_{n \geq m}$ von σ -Algebren definiert durch

$$\mathcal{F}_n = \sigma(U_j : j \geq n) . \quad (2.2.17)$$

Dann ist $(U_n, \mathcal{F}_n)_{n \geq m}$ ein Rückwärtsmartingal.

Beweis²⁹:

Nach Konstruktion in (2.2.17) ist die Folge $(U_n)_{n \geq m}$ der U-Statistiken bzgl. $(\mathcal{F}_n)_{n \geq m}$ adaptiert, da sicher $\sigma(U_n) \subset \sigma(U_j : j \geq n)$ gilt.

Ferner gilt für $n \geq n_0 \geq m$ trivialerweise $\mathcal{F}_n = \sigma(U_j : j \geq n) \subset \sigma(U_j : j \geq n_0) = \mathcal{F}_{n_0}$ und damit die erste Bedingung aus Definition 2.2.15.

Für die zweite Bedingung beachte man, dass die Zufallsvariablen X_n für $n \in \mathbb{N}$ unabhängig identisch verteilt sind (vergleiche Definition 2.2.1). Man erhält somit aufgrund der Voraussetzung aus Gleichung (2.2.16) für alle $n \geq m$:

$$\begin{aligned} \mathbb{E} |U_n| &= \mathbb{E} \left| \frac{(n-m)!}{n!} \sum_{\tau \in (n,m)} \psi(X_{\tau(1)}, \dots, X_{\tau(m)}) \right| \\ &\leq \frac{(n-m)!}{n!} \sum_{\tau \in (n,m)} \mathbb{E} |\psi(X_{\tau(1)}, \dots, X_{\tau(m)})| \\ &= \frac{(n-m)!}{n!} \sum_{\tau \in (n,m)} \mathbb{E} |\psi(X_1, \dots, X_m)| \\ &= \mathbb{E} |\psi(X_1, \dots, X_m)| < \infty . \end{aligned}$$

Folglich gilt auch die zweite Bedingung aus Definition 2.2.15.

²⁹Der Beweis findet sich in gekürzter Form in der Monographie Lee 1990.

Durch den Übergang zum Supremum über alle $n \geq m$ erhalten wir sogar ferner auf natürliche Art und Weise die Voraussetzung des *zweiten Doob'schen Konvergenzsatzes*, nämlich

$$\sup_{n \geq m} \mathbb{E}|U_n| \leq \mathbb{E}|\psi(X_1, \dots, X_m)| < \infty. \quad (2.2.18)$$

Da nun $\sigma(U_n) \subset \mathcal{F}_n$ ist, folgert man in Analogie zu obiger Argumentation:

$$\begin{aligned} U_n &= \mathbb{E}(U_n | \mathcal{F}_n) = \frac{(n-m)!}{n!} \sum_{\tau(n,m)} \mathbb{E}\left(\psi(X_{\tau(1)}, \dots, X_{\tau(m)}) | \mathcal{F}_n\right) \\ &= \frac{(n-m)!}{n!} \sum_{\tau(n,m)} \mathbb{E}\left(\psi(X_1, \dots, X_m) | \mathcal{F}_n\right) = \mathbb{E}\left(\psi(X_1, \dots, X_m) | \mathcal{F}_n\right), \end{aligned}$$

und daraus für $m \leq n \leq n_0$ wegen $\mathcal{F}_{n_0} \subset \mathcal{F}_n$ mit der Turmeigenschaft unmittelbar:

$$\begin{aligned} \mathbb{E}(U_n | \mathcal{F}_{n_0}) &= \mathbb{E}\left(\mathbb{E}\left(\psi(X_1, \dots, X_m) | \mathcal{F}_n\right) | \mathcal{F}_{n_0}\right) \\ &= \mathbb{E}\left(\psi(X_1, \dots, X_m) | \mathcal{F}_{n_0}\right) = U_{n_0}. \end{aligned}$$

Damit gilt schließlich auch die dritte Bedingung aus Definition 2.2.15, und ergo ist $(U_n, \mathcal{F}_n)_{n \geq m}$ ein Rückwärtsmartingal. □

2.2.5 Der Beweis des starken Gesetzes der großen Zahlen für U-Statistiken

Mit den Resultaten der vorherigen Abschnitte können wir nun *starke Gesetz der großen Zahlen für U-Statistiken* beweisen, dessen Aussage bereits in Satz 2.2.3 formuliert wurde. Wie im Vorfeld dieses Satzes angekündigt, werden wir auch die Voraussetzungen noch näher spezifizieren:

Satz 2.2.3. (*Starkes Gesetz der großen Zahlen für U-Statistiken*)

Es sei $(U_n)_{n \geq m}$ eine Folge von U-Statistiken mit gemeinsamem m -dimensionalen Kern ψ , die über eine Folge $(X_n)_{n \in \mathbb{N}}$ unabhängig identisch verteilter Zufallsvariablen mit Werten in einem *polnischen Raum* (E, \mathcal{E}) definiert werde, wobei wie üblich \mathcal{E} die Borel'sche σ -Algebra bezeichne.

Ferner gelte $\mathbb{E}|\psi(X_1, \dots, X_m)| < \infty$. Dann konvergiert U_n für $n \rightarrow \infty$ fast sicher gegen $\mathbb{E}(\psi(X_1, \dots, X_m))$.

Beweis³⁰:

Zunächst definieren wir eine Folge $(\mathcal{F}_n)_{n \geq m}$ von σ -Algebren wie in Gleichung (2.2.17), d.h.

$$\mathcal{F}_n = \sigma(U_j : j \geq n) .$$

Gemäß Lemma 2.2.19 ist dann $(U_n, \mathcal{F}_n)_{n \geq m}$ ein Rückwärtsmartingal. Nach dem *Zweiten Doobischen Konvergenzsatz* (siehe Satz 2.2.18) existiert dann eine reelle Zufallsvariable U_∞ , sodass U_n für $n \rightarrow \infty$ fast sicher gegen U_∞ konvergiert.

Folglich ist der Beweis vollendet, wenn wir zeigen können, dass $U_\infty = \mathbb{E}(\psi(X_1, \dots, X_m))$ fast sicher gilt.

Für $\varepsilon > 0$ betrachten wir dafür die Menge³¹

$$U(\varepsilon) = \bigcap_{n=m}^{\infty} \bigcup_{j=n}^{\infty} \left\{ x \in E^{\mathbb{N}} : u_j \circ x > \mathbb{E}(\psi(X_1, \dots, X_m)) + \varepsilon \right\} \in \mathcal{E}^{\mathbb{N}}, \quad (2.2.19)$$

also die Menge aller $x \in E^{\mathbb{N}}$, auf der $u_n \circ x > \mathbb{E}(\psi(X_1, \dots, X_m)) + \varepsilon$ für unendlich viele $n \geq m$ gilt.

Da die Bildung des Durchschnitts in Gleichung (2.2.19) über eine *absteigende* Folge von Mengen erfolgt, stellen wir insbesondere fest, dass für jedes beliebige $N \geq m$ gilt:

$$U(\varepsilon) = \bigcap_{n=N}^{\infty} \bigcup_{j=n}^{\infty} \left\{ x \in E^{\mathbb{N}} : u_j \circ x > \mathbb{E}(\psi(X_1, \dots, X_m)) + \varepsilon \right\}. \quad (2.2.20)$$

Ist nun τ eine endliche Permutation, so setzen wir $N = N(\tau)$ und Gleichung (2.2.1) aus Bemerkung 2.2.2 liefert wegen $j \geq N$:

$$\begin{aligned} \{\tau X \in U(\varepsilon)\} &= \left\{ \tau X \in \bigcap_{n=N}^{\infty} \bigcup_{j=n}^{\infty} \left\{ x \in E^{\mathbb{N}} : u_j \circ x > \mathbb{E}(\psi(X_1, \dots, X_m)) + \varepsilon \right\} \right\} \\ &\stackrel{(2.2.1)}{=} \left\{ \tau X \in \bigcap_{n=N}^{\infty} \bigcup_{j=n}^{\infty} \left\{ x \in E^{\mathbb{N}} : u_j \circ \tau x > \mathbb{E}(\psi(X_1, \dots, X_m)) + \varepsilon \right\} \right\} \\ &= \left\{ X \in \bigcap_{n=N}^{\infty} \bigcup_{j=n}^{\infty} \left\{ x \in E^{\mathbb{N}} : u_j \circ x > \mathbb{E}(\psi(X_1, \dots, X_m)) + \varepsilon \right\} \right\} \\ &= \{X \in U(\varepsilon)\}. \end{aligned} \quad (2.2.21)$$

Da τ eine *beliebige* endliche Permutation war, folgt aus Gleichung (2.2.21) also, dass die Menge $U(\varepsilon)$ permutierbar ist und das *0-1-Gesetz von Hewitt-Savage* (siehe Korollar 2.2.14; hierfür ist die angekündigte Verschärfung der Voraussetzungen notwendig) liefert

³⁰Der Beweis findet sich in ähnlicher Form ebenfalls in der Monographie Lee 1990, siehe Seite 130.

³¹Die Messbarkeit des m -dimensionalen Kerns ψ stellt sicher, dass $U(\varepsilon) \in \mathcal{E}^{\mathbb{N}}$ gilt.

$P(\{X \in U(\varepsilon)\}) = 0$ oder $P(\{X \in U(\varepsilon)\}) = 1$.

Wegen $E(U_n) = E(\psi(X_1, \dots, X_m))$ (s. Beweis von Lemma 2.2.19) kann jedoch nicht $P(\{X \in U(\varepsilon)\}) = 1$ gelten, da andernfalls $E(U_n) > E(\psi(X_1, \dots, X_m))$ mit entsprechendem Widerspruch gälte. Es gilt demnach

$$P(\{X \in U(\varepsilon)\}) = 0. \quad (2.2.22)$$

Betrachten wir weiter die Menge

$$L(\varepsilon) = \bigcap_{n=m}^{\infty} \bigcup_{j=n}^{\infty} \left\{ x \in (E^{\mathbb{N}}, \mathcal{E}^{\mathbb{N}}) : u_j \circ x < E(\psi(X_1, \dots, X_m)) - \varepsilon \right\}, \quad (2.2.23)$$

so erhalten wir völlig analog zu obiger Argumentation auch

$$P(\{X \in L(\varepsilon)\}) = 0. \quad (2.2.24)$$

Schließlich können wir die Menge, auf der U_n für $n \rightarrow \infty$ nicht gegen $E(\psi(X_1, \dots, X_m))$ konvergiert, durch die Mengen $U(\varepsilon)$ und $L(\varepsilon)$ charakterisieren:

$$\{U_{\infty} \neq E(\psi(X_1, \dots, X_m))\} = \bigcup_{j=1}^{\infty} \left\{ X \in L\left(\frac{1}{j}\right) \right\} \cup \left\{ X \in U\left(\frac{1}{j}\right) \right\}. \quad (2.2.25)$$

Aus den Gleichungen (2.2.22) und (2.2.24) folgt dann mit Gleichung (2.2.25) die Behauptung, da gilt:

$$P(\{U_{\infty} \neq E(\psi(X_1, \dots, X_m))\}) \leq \sum_{j=1}^{\infty} P(\{X \in L\left(\frac{1}{j}\right)\}) + P(\{X \in U\left(\frac{1}{j}\right)\}) = 0,$$

und somit $U_{\infty} = E(\psi(X_1, \dots, X_m))$ fast sicher. \square

Hiermit möchten wir auch den Abschnitt 2.2 abschließen. Der nachfolgende und letzte Abschnitt dieses Kapitels wird sich mit der Straffheit von Wahrscheinlichkeitsmaßen über dem Raum der stetigen Funktionen auseinandersetzen.

2.3 Straffheit im Raum der stetigen Funktionen

Zum Abschluss des Kapitels über die mathematischen Grundlagen dieser Arbeit möchten wir den Begriff der *Straffheit* beleuchten, der zunächst als Eigenschaft einer Familie von Wahrscheinlichkeitsmaßen über einem metrischen Raum (S, d) eingeführt wird.

Ist (S, d) ein vollständiger und separabler metrischer Raum, so besagt der hier nicht explizit thematisierte *Satz von Prohorov* (siehe Ethier und Kurtz 1986, S. 104), dass die Begriffe

der *Straffheit* und der *relativen Kompaktheit* äquivalent sind.

Diese Äquivalenz wird deutlich, wenn wir die Straffheit der von einer Folge $C(K)$ -wertiger Zufallsfunktionen induzierten Familie von Wahrscheinlichkeitsmaßen unter Zuhilfenahme des *Satzes von Arzelá-Ascoli* charakterisieren (siehe Korollar 2.3.5).

Diese Charakterisierung liefert die Grundlage eines elementaren Konvergenzresultates, das in Lemma 2.3.6 formuliert und mannigfache Anwendung finden wird.

Schließlich führen wir noch das *Straffheitskriterium von Kolmogorov* ein, das sich als Mittel unserer Wahl zum Nachweis von Straffheit erweisen wird.

Dem voran stellen wir einige Definitionen, um das Konzept der *Straffheit* und den *Raum der stetigen Funktionen* einzuführen.

Definition 2.3.1. (*Straffheit einer Familie von Wahrscheinlichkeitsmaßen*, s. Billingsley 1968, S. 37)

Es sei S ein metrischer Raum mit Metrik d . Ferner bezeichne \mathcal{B}_d die von der Metrik induzierte *Borel'sche σ -Algebra* über S .

Eine Familie Π von Wahrscheinlichkeitsmaßen auf (S, \mathcal{B}_d) heißt dann *straff*, wenn für alle $\varepsilon > 0$ ein Kompaktum $K = K(\varepsilon) \in \mathcal{B}_d$ existiert, so dass gilt:

$$P(K) > 1 - \varepsilon, \quad \text{für alle } P \in \Pi.$$

Wir verallgemeinern den Begriff der Straffheit sodann auf Folgen von Zufallsfunktionen.

Definition 2.3.2. (*Straffheit einer Folge von Zufallsfunktionen*, s. Billingsley 1968, S. 57)

In der Situation aus Definition 2.3.1 sei $(X_n)_{n \in \mathbb{N}}$ eine Folge von Zufallsfunktionen mit Werten im Messraum (S, \mathcal{B}_d) . Die Folge $(X_n)_{n \in \mathbb{N}}$ heißt dann *straff*, wenn die zugehörige Familie der Bildmaße $\Pi = \{P^{X_n}\}_{n \in \mathbb{N}}$ straff ist, das heißt wenn für alle $\varepsilon > 0$ ein Kompaktum K existiert, so dass gilt:

$$P^{X_n}(K) = P(\{X_n \in K\}) > 1 - \varepsilon, \quad \text{für alle } n \in \mathbb{N}.$$

Schließlich führen wir den *Raum der stetigen Funktionen* ein, der in der Einleitung dieses Abschnitts bereits kontextlos verwendet wurde.

Definition 2.3.3. (*Der Raum der stetigen Funktionen $C(X, Y)$* , vgl. Dudley 2002, S. 51)

Es bezeichne X einen kompakten topologischen Raum und Y einen normierten Vektorraum. Der Raum $C(X, Y)$ ist definiert als die Menge aller stetigen Funktionen von X nach

Y . versehen mit der Supremumsnorm $\|\cdot\|_\infty$, die definiert ist via

$$\|f\|_\infty = \sup_{x \in X} \|f(x)\|,$$

wobei $\|\cdot\|$ die Norm auf Y bezeichne, wird $C(X, Y)$ zu einem Banachraum (siehe Dudley 2002, S. 53).

Nachfolgend beschränken wir uns auf den Spezialfall, dass X eine kompakte Teilmenge des \mathbb{R}^d ist und schreiben suggestiv $X = K \subset \mathbb{R}^d$. Außerdem betrachten wir nur den Fall $Y = \mathbb{R}^{\tilde{d}}$, wobei die konkrete Dimension $\tilde{d} \in \mathbb{N}$ sich als unerheblich erweisen wird. Daher werden wir den Raum $C(K, \mathbb{R}^{\tilde{d}})$ zumeist lediglich mit $C(K)$ bezeichnen.

Nach dem *Satz von Weierstraß* (siehe Dudley 2002, S. 54) ist $C(K)$ separabel.

Der *Satz von Arzelá-Ascoli* liefert uns nun Informationen über die kompakten Teilmengen des Raumes $C(K)$. Die nachfolgende, anschaulichere Charakterisierung von Straffheit im Raum der stetigen Funktionen $C(K)$ wird von diesem Satz Gebrauch machen.

Satz 2.3.4. (*Straffheit einer Folge von Wahrscheinlichkeitsmaßen über dem Raum $C(K)$*), siehe Billingsley 1968, S. 55)

Eine Folge $\{P_n\}_{n \in \mathbb{N}}$ von Wahrscheinlichkeitsmaßen auf der Borel'schen σ -Algebra über dem Raum $C(K)$ ist straff genau dann, wenn die nachfolgenden Bedingungen erfüllt sind:

- (i) Es gibt ein $\vartheta_0 \in K$, so dass für jedes $\eta > 0$ ein $a \in \mathbb{R}$ existiert mit

$$P_n(\{f : \|f(\vartheta_0)\| > a\}) \leq \eta, \quad \text{für alle } n \in \mathbb{N}. \quad (2.3.1)$$

- (ii) Für jedes $\eta > 0$, $\varepsilon > 0$ existiert ein $\delta > 0$ und ein $n_0 \in \mathbb{N}$, so dass gilt:

$$P_n\left(\left\{f : \sup_{\|x-y\| < \delta} \|f(x) - f(y)\| \geq \varepsilon\right\}\right) \leq \eta, \quad \text{für alle } n \geq n_0. \quad (2.3.2)$$

Beweis³²:

Zunächst nehmen wir an, die Folge $\{P_n\}_{n \in \mathbb{N}}$ sei straff. Ferner seien $\eta > 0$ und $\varepsilon > 0$ gegeben. Dann wählen wir eine kompakte Teilmenge $\mathcal{K} \subset C(K)$ dergestalt, dass

$$P_n(\mathcal{K}) > 1 - \eta$$

für alle $n \in \mathbb{N}$ erfüllt ist. Nach dem *Satz von Arzelá-Ascoli* (siehe Dudley 2002, S. 52) ist dann \mathcal{K} einerseits gleichmäßig beschränkt, so dass für ein beliebiges $\vartheta_0 \in K$ ein $a \in \mathbb{R}$ mit

$$\|f(\vartheta_0)\| \leq a, \quad \text{für alle } f \in \mathcal{K}, \quad (2.3.3)$$

³²Der Beweis befindet sich für $K = [0, 1]$ in Billingsley 1968 und verläuft völlig analog.

existiert. Andererseits ist \mathcal{K} gleichgradig stetig, das heißt zu gegebenem $\varepsilon > 0$ existiert ein $\delta > 0$ mit

$$\|f(x) - f(y)\| < \varepsilon, \quad \text{für alle } f \in \mathcal{K} \text{ und } x, y \in K \text{ mit } \|x - y\| < \delta. \quad (2.3.4)$$

Die Ungleichung aus (2.3.3) impliziert nun $\mathcal{K} \subset \{f : |f(\vartheta_0)| \leq a\}$. Die Ungleichung aus (2.3.4) wiederum impliziert, dass

$$\mathcal{K} \subset \left\{ f : \sup_{\|x-y\|<\delta} \|f(x) - f(y)\| < \varepsilon \right\}$$

ist, und daher folgt unmittelbar:

$$P_n(\{f : \|f(\vartheta_0)\| \leq a\}) \geq 1 - \eta,$$

$$P_n\left(\left\{f : \sup_{\|x-y\|<\delta} \|f(x) - f(y)\| < \varepsilon\right\}\right) \geq 1 - \eta, \quad \text{für alle } n \in \mathbb{N}.$$

Durch Übergang zum Komplement erhalten wir damit sofort die Bedingungen (i) und (ii).

Nun seien die Bedingungen (i) und (ii) erfüllt.

Da ein einzelnes Wahrscheinlichkeitsmaß P auf $C(K)$ stets straff ist (siehe Billingsley 1968, S. 10, *Theorem 1.4* unter Berücksichtigung von Definition 2.3.3), können wir in Bedingung (ii) stets $n_0 = 1$ fordern:

Sind nämlich $\eta > 0$, $\varepsilon > 0$, $\delta > 0$ und $n_0 > 1$ derart, dass (2.3.2) gilt, so wählen wir für jedes Wahrscheinlichkeitsmaß P_1, \dots, P_{n_0-1} ein δ_i mit

$$P_i\left(\left\{f : \sup_{\|x-y\|<\delta_i} \|f(x) - f(y)\| \geq \varepsilon\right\}\right) \leq 1 - \eta, \quad \text{für alle } i = 1, \dots, n_0 - 1.$$

Übergang zu $\delta^* = \min\{\delta, \delta_1, \dots, \delta_{n_0-1}\}$ liefert dann die Gültigkeit von Bedingung (ii) mit $n_0 = 1$.

Zu gegebenem $\eta > 0$ wählen wir nun aufgrund der Gültigkeit von Bedingung (i) ein $a \in \mathbb{R}$ mit

$$P_n(A) \geq 1 - \frac{\eta}{2}, \quad A = \{f : \|f(\vartheta_0)\| \leq a\}.$$

Weiter wählen wir aufgrund der Gültigkeit von Bedingung (ii) eine Folge $(\delta_k)_{k \in \mathbb{N}}$ mit $\delta_k > 0$ so, dass

$$P_n(A_k) \geq 1 - \frac{\eta}{2^{k+1}} \quad \text{für } A_k = \left\{f : \sup_{\|x-y\|<\delta_k} \|f(x) - f(y)\| < \frac{1}{k}\right\}$$

ist. Wählen wir \mathcal{K} als den Abschluss von

$$A \cap \bigcap_{k \in \mathbb{N}} A_k \subset C(K),$$

so ist \mathcal{K} nach dem *Satz von Arzelá-Ascoli* einerseits kompakt, andererseits gilt

$$\begin{aligned} P_n \left(\left(A \cap \bigcap_{k \in \mathbb{N}} A_k \right)^c \right) &= P_n \left(A^c \cup \bigcup_{k \in \mathbb{N}} A_k^c \right) \\ &\leq P_n(A^c) + \sum_{k \in \mathbb{N}} P_n(A_k^c) \\ &\leq \frac{\eta}{2} + \sum_{k \in \mathbb{N}} \frac{\eta}{2^{k+1}} = \eta, \end{aligned} \tag{2.3.5}$$

und damit via erneutem Übergang zum Komplement

$$P_n(\mathcal{K}) \geq P_n \left(A \cap \bigcap_{k \in \mathbb{N}} A_k \right) \geq 1 - \eta.$$

Ergo ist die Folge $\{P_n\}_{n \in \mathbb{N}}$ straff. □

Als Korollar aus Satz 2.3.4 erhalten wir sofort eine Charakterisierung von Straffheit für eine Folge von $C(K)$ -wertigen Zufallsfunktionen.

Korollar 2.3.5. (*Straffheit einer Folge von $C(K)$ -wertigen Zufallsfunktionen*, siehe Billingsley 1968, S. 58)

Eine Folge $(X_n)_{n \in \mathbb{N}}$ von Zufallsfunktionen mit Werten im Raum $C(K)$ ist straff genau dann, wenn die nachfolgenden Bedingungen erfüllt sind:

- (i) Es gibt ein $\vartheta_0 \in K$, so dass für jedes $\eta > 0$ ein $a \in \mathbb{R}$ existiert mit

$$P(\{\|X_n(\vartheta_0)\| > a\}) \leq \eta, \quad \text{für alle } n \in \mathbb{N}. \tag{2.3.6}$$

- (ii) Für jedes $\eta > 0$, $\varepsilon > 0$ existiert ein $\delta > 0$ und ein $n_0 \in \mathbb{N}$, so dass gilt:

$$P \left(\left\{ \sup_{\|x-y\| < \delta} \|X_n(x) - X_n(y)\| \geq \varepsilon \right\} \right) \leq \eta, \quad \text{für alle } n \geq n_0. \tag{2.3.7}$$

Beweis:

Das Korollar folgt mit Definition 2.3.2 unmittelbar aus Satz 2.3.4. □

Das Korollar 2.3.5 erlaubt erstmalig auch eine heuristische Interpretation des Begriffes

Straffheit im Falle $C(K)$ -wertiger Zufallsfunktionen:

In der Bedingung (2.3.6) wird die punktweise stochastische Beschränktheit der Realisierungen gefordert, während in der Bedingung (2.3.7) verlangt wird, dass zudem die Fluktuationen dieser Realisierungen kontrolliert werden können.

Anschaulich ist also eine Folge $C(K)$ -wertiger Zufallsfunktionen straff genau dann, wenn ihre Realisierungen an einem Punkt beschränkt sind und lokal lediglich geringe Schwankungen aufweisen.

Aufgrund dieser Interpretation mag es uns also nicht überraschen, dass wir die Straffheit vermöge einer *Lipschitz-Bedingung* folgern können (siehe Anhang A.3, Korollar A.3.1).

Im nachfolgenden Unterabschnitt möchten wir jedoch das *Straffheitskriterium von Kolmogorov* diskutieren, das seinen Namen aufgrund der verwendeten *Momenten-Bedingung* aus dem *Satz von Kolmogorov* (siehe etwa Bauer 1991, S. 340) erhielt.

Zuvor präsentieren wir noch ein Konvergenzresultat, das in seiner Allgemeinheit mannigfaltige Anwendungen sieht. Es lässt sich leicht unter Zuhilfenahme von Korollar 2.3.5 beweisen.

Lemma 2.3.6.

Es sei $(X_n)_{n \in \mathbb{N}}$ eine straffe Folge $C(K)$ -wertiger Zufallsfunktionen.

Ferner sei $(\vartheta_n)_{n \in \mathbb{N}}$ eine Folge von Zufallsvariablen mit Werten in K , die fast sicher gegen ein $\vartheta_0 \in K$ konvergiert. Dann gilt:

$$X_n(\vartheta_n) - X_n(\vartheta_0) \xrightarrow{P} 0 \quad (n \rightarrow \infty).$$

Beweis³³:

Wir möchten formal die folgende Aussage zeigen:

$$\forall \varepsilon > 0 : \quad \lim_{n \rightarrow \infty} P(\{|X_n(\vartheta_n) - X_n(\vartheta_0)| \geq \varepsilon\}) = 0. \quad (2.3.8)$$

Es sei also $\eta > 0$ beliebig. Da $(X_n)_{n \in \mathbb{N}}$ nach Voraussetzung straff ist, existiert nach Korollar 2.3.5 ein $n_0 \in \mathbb{N}$ und ein $\delta > 0$ mit

$$P\left(\left\{\sup_{\|x-y\|<\delta} |X_n(x) - X_n(y)| \geq \varepsilon\right\}\right) \leq \frac{\eta}{2}, \quad \text{für alle } n \geq n_0. \quad (2.3.9)$$

Aufgrund der fast sicheren Konvergenz der Folge $(\vartheta_n)_{n \in \mathbb{N}}$ gegen ϑ_0 existiert zu gegebenem $\delta > 0$ ein $n_1 \in \mathbb{N}$ mit (vgl. Bauer 1991, S. 34)

$$P\left(\left\{\|\vartheta_n - \vartheta_0\| < \delta \quad \forall n \geq n_1\right\}\right) \geq 1 - \frac{\eta}{2}. \quad (2.3.10)$$

³³Der Beweis dieses Lemmas wurde vom Autor dieser Arbeit selbst geführt.

Damit erhalten wir für alle $n \geq \max\{n_0, n_1\}$ durch Kombination von (2.3.9) und (2.3.10):

$$\begin{aligned}
& P\left(\left\{|X_n(\vartheta_n) - X_n(\vartheta_0)| \geq \varepsilon\right\}\right) \\
&= P\left(\left\{|X_n(\vartheta_n) - X_n(\vartheta_0)| \geq \varepsilon\right\} \cap \left\{\|\vartheta_m - \vartheta_0\| < \delta \quad \forall m \geq n_1\right\}\right) \\
&\quad + P\left(\left\{|X_n(\vartheta_n) - X_n(\vartheta_0)| \geq \varepsilon\right\} \cap \left\{\|\vartheta_m - \vartheta_0\| < \delta \quad \forall m \geq n_1\right\}^c\right) \\
&\leq P\left(\left\{|X_n(\vartheta_n) - X_n(\vartheta_0)| \geq \varepsilon\right\} \cap \left\{\|\vartheta_n - \vartheta_0\| < \delta\right\}\right) \\
&\quad + P\left(\left\{\|\vartheta_m - \vartheta_0\| < \delta \quad \forall m \geq n_1\right\}^c\right) \\
&\leq P\left(\left\{\sup_{\|x-y\|<\delta} |X_n(x) - X_n(y)| \geq \varepsilon\right\}\right) + \frac{\eta}{2} \leq \eta.
\end{aligned}$$

Die Beliebigkeit von η liefert somit die Behauptung aus Gleichung (2.3.8). \square

2.3.1 Das Straffheitskriterium von Kolmogorov

Wir möchten nun ein handlicheres Kriterium für die Straffheit einer Folge von $C(K)$ -wertigen Zufallsfunktionen formulieren und beweisen. Es wird sich schließlich als geeignetes Mittel zum Nachweis von Straffheit erweisen, obgleich wir es zunächst nur für den Fall zeigen werden, dass K der d -dimensionale abgeschlossene Einheitswürfel im \mathbb{R}^d ist.

Der Beweis beruht auf einem in Billingsley 1968 formulierten Theorem, welches zudem nur den Fall $d = 1$ - also das kompakte Einheitsintervall - umfasst.

Lemma 2.3.7. (vergleiche etwa Billingsley 1968, S. 95)

Eine Folge $(X_n)_{n \in \mathbb{N}}$ von Zufallsfunktionen mit Werten im Raum $C([0, 1]^d)$ ist straff, wenn die nachfolgenden Bedingungen erfüllt sind:

- (i) Es gibt ein $\vartheta_0 \in [0, 1]^d$, so dass für jedes $\eta > 0$ ein $a \in \mathbb{R}$ existiert mit

$$P\left(\left\{\|X_n(\vartheta_0)\| > a\right\}\right) \leq \eta, \quad \text{für alle } n \in \mathbb{N}. \quad (2.3.11)$$

- (ii) Es gilt die *Momenten-Bedingung*³⁴

$$E\|X_n(t) - X_n(s)\|^2 \leq \nu \|t - s\|^{d+\alpha}, \quad \text{für alle } s, t \in [0, 1]^d \text{ und } n \in \mathbb{N}, \quad (2.3.12)$$

mit positiven Konstanten $\nu > 0$ und $\alpha > 0$.

³⁴Wie in Billingsley 1968 im Falle $d = 1$ gezeigt, kann die hier angegebene Momenten-Bedingung weiter verallgemeinert werden. Der Beweis lässt sich leicht übertragen, wir konzentrieren uns an dieser Stelle jedoch aus Gründen der Übersichtlichkeit auf diesen Spezialfall.

Beweis³⁵:

Die Momenten-Bedingung (2.3.12) impliziert aufgrund der *Chebyshev-Markovschen Ungleichung* (siehe etwa Bauer 1992, S. 128), dass für alle $n \in \mathbb{N}$ und $\varepsilon > 0$ gilt:

$$P(\|X_n(t) - X_n(s)\| \geq \varepsilon) \leq \frac{\nu}{\varepsilon^2} \|t - s\|^{d+\alpha}, \quad \text{für alle } s, t \in [0, 1]^d. \quad (2.3.13)$$

Aufgrund der Äquivalenz der Normen im \mathbb{R}^d können wir ohne Beschränkung der Allgemeinheit die Maximumsnorm verwenden, indem gegebenenfalls die Konstante ν modifiziert wird.

Bei gegebenem $0 < \delta \leq 1$ und $n \in \mathbb{N}$ definieren wir für $t \in [0, 1]^d$:

$$Q_t = \left\{ s \in [0, 1]^d : t_i \leq s_i \leq t_i + \delta \quad \forall 1 \leq i \leq d \right\},$$

$$A_t = \left\{ \sup_{s \in Q_t} \|X_n(t) - X_n(s)\| \geq \varepsilon \right\}.$$

Wir wählen eine injektive Abzählung h des Gitters $G_\delta^d = [0, 1]^d \cap \delta\mathbb{Z}^d$ (siehe Lemma A.2.1 im Anhang für eine detaillierte Diskussion). Die Dreiecksungleichung liefert sodann die Gültigkeit der folgenden Inklusion:

$$\left\{ \sup_{\|t-s\|_\infty < \delta} \|X_n(t) - X_n(s)\| \geq 3\varepsilon \right\} \subset \bigcup_{q=1}^{\lfloor 1+\delta^{-1} \rfloor^d} A_{h(q)}. \quad (2.3.14)$$

Dies sieht man wie folgt ein: Für ein Ereignis ω aus dem zugrundeliegenden Wahrscheinlichkeitsraum gelte

$$\sup_{\|t-s\|_\infty < \delta} \|X_n(t) - X_n(s)\| \geq 3\varepsilon,$$

so dass $s, t \in [0, 1]^d$ mit $\|t - s\|_\infty \leq \delta$ und $\|X_n(t) - X_n(s)\| \geq 3\varepsilon$ existieren. Dann existieren natürliche Zahlen q_1, q_2 , die nicht notwendig verschieden seien, mit $s \in Q_{h(q_1)}, t \in Q_{h(q_2)}$. Wegen $\|s - t\|_\infty \leq \delta$ ist $Q_{h(q_1)} \cap Q_{h(q_2)} \neq \emptyset$, da sich die d -dimensionalen Würfel zumindest in einer Ecke überschneiden. Wir wählen $r \in Q_{h(q_1)} \cap Q_{h(q_2)}$ und nehmen an, es gelte $\omega \notin A_{h(q_1)} \cup A_{h(q_2)}$. Dann folgt aber:

$$\|X_n(t) - X_n(s)\| \leq \|X_n(t) - X_n(r)\| + \|X_n(r) - X_n(s)\| < \varepsilon + \varepsilon = 2\varepsilon,$$

³⁵Der Beweis orientiert sich an der angegebenen Literaturquelle, wurde vom Autor jedoch ins Mehrdimensionale transferiert.

und damit der Widerspruch. Es gilt folglich

$$\omega \in A_{h(q_1)} \cup A_{h(q_2)} \subset \bigcup_{q=1}^{(1+\lfloor \delta^{-1} \rfloor)^d} A_{h(q)},$$

und damit die Inklusion aus (2.3.14). Die Monotonie und σ -Subadditivität des Wahrscheinlichkeitsmaßes P liefern unmittelbar die Abschätzung

$$P \left(\left\{ \sup_{\|t-s\|_\infty < \delta} \|X_n(t) - X_n(s)\| \geq 3\varepsilon \right\} \right) \leq P \left(\bigcup_q A_{h(q)} \right) \leq \sum_{q=1}^{(1+\lfloor \delta^{-1} \rfloor)^d} P(A_{h(q)}). \quad (2.3.15)$$

Gemäß Korollar 2.3.5 genügt es also aufgrund der Abschätzung (2.3.15), für ein gegebenes $\eta > 0$ aus der in (2.3.13) formulierten Bedingung für hinreichend kleines $\delta > 0$ folgern zu können, dass

$$\sum_{q=1}^{(1+\lfloor \delta^{-1} \rfloor)^d} P(A_{h(q)}) \leq \eta \quad (2.3.16)$$

gilt. Wir betrachten die Wahrscheinlichkeiten $P(A_{h(q)})$ separat.

Es sei also $1 \leq q \leq (1 + \lfloor \delta^{-1} \rfloor)^d$ und für ein $m \in \mathbb{N}$ wählen wir eine Abzählung ι des Gitters³⁶ $h(q) + \delta G_{\frac{1}{m}}^d$ wie in Lemma A.2.1, so dass gilt:

$$\|\iota(i+1) - \iota(i)\|_\infty = \frac{\delta}{m}, \quad \text{für alle } i = 1, \dots, (m+1)^d - 1. \quad (2.3.17)$$

Ferner definieren wir die Zufallsvariablen

$$\xi_i = X_n(\iota(i+1)) - X_n(\iota(i)), \quad \text{für } i = 1, \dots, (m+1)^d - 1.$$

Mit $u_i = \nu^{\frac{1}{d+\alpha}} \|\iota(i+1) - \iota(i)\|_\infty$ gilt dann für $k > j$ mit Abschätzung (2.3.13):

$$\begin{aligned} P \left(\left\| \sum_{i=1}^k \xi_i - \sum_{i=1}^j \xi_i \right\| \geq \varepsilon \right) &= P \left(\|X_n(\iota(k+1)) - X_n(\iota(j+1))\| \geq \varepsilon \right) \\ &\leq \frac{\nu}{\varepsilon^2} \|\iota(k+1) - \iota(j+1)\|_\infty^{d+\alpha} \leq \frac{1}{\varepsilon^2} \left(\nu^{\frac{1}{d+\alpha}} \|\iota(k+1) - \iota(j+1)\|_\infty \right)^{d+\alpha} \\ &\leq \frac{1}{\varepsilon^2} \left(\sum_{j < i \leq k} u_i \right)^{d+\alpha} \end{aligned} \quad (2.3.18)$$

³⁶Man erhält das Gitter mittels der affinen Transformation $t \mapsto h(q) + \delta t$ aus dem Gitter $G_{\frac{1}{m}}^d$, die Schreibweise wurde entsprechend suggestiv gewählt.

Die Abschätzung aus (2.3.18) erlaubt uns die Anwendung eines technischen Theorems, um die folgende Abschätzung zu erhalten (siehe *Theorem 12.2* und Gleichung (12.43) bzw. (12.44) in Billingsley 1968, S. 94):

$$P \left(\left\{ \max_{1 \leq i \leq (m+1)^{d-1}} \|X_n(\iota(i)) - X_n(h(q))\| \geq \varepsilon \right\} \right) \leq \frac{\tilde{\nu}}{\varepsilon^2} \left(\sum_{i=1}^m u_i \right)^{d+\alpha}. \quad (2.3.19)$$

Nach Gleichung (2.3.17) gilt mit Abschätzung (2.3.19) also:

$$P \left(\left\{ \max_{1 \leq i \leq (m+1)^{d-1}} \|X_n(\iota(i)) - X_n(h(q))\| \geq \varepsilon \right\} \right) \leq \frac{\tilde{\nu}\nu\delta^{d+\alpha}}{\varepsilon^2},$$

und aufgrund der Stetigkeit der Realisierungen von X_n - man beachte, dass die Zufallsfunktionen nach Voraussetzung $C([0, 1]^d)$ -wertig sind - erhalten wir somit für $m \rightarrow \infty$:

$$P(A_{h(q)}) = P \left(\left\{ \sup_{s \in Q_{h(q)}} \|X_n(s) - X_n(h(q))\| \geq \varepsilon \right\} \right) \leq \frac{\tilde{\nu}\nu\delta^{d+\alpha}}{\varepsilon^2}. \quad (2.3.20)$$

Einsetzen von (2.3.20) die linke Seite von (2.3.16) liefert sodann:

$$\sum_{q=1}^{(1+\lfloor \delta^{-1} \rfloor)^d} P(A_{h(q)}) \leq (1 + \lfloor \delta^{-1} \rfloor)^d \cdot \frac{\tilde{\nu}\nu\delta^{d+\alpha}}{\varepsilon^2} = \mathcal{O}(1) \cdot \delta^\alpha \longrightarrow 0 \quad (\delta \rightarrow 0).$$

Da diese Abschätzung unabhängig von n ist, können wir also zu gegebenem $\eta > 0$ ein $\delta > 0$ finden, so dass für alle $n \in \mathbb{N}$ gilt³⁷:

$$\sum_{q=1}^{(1+\lfloor \delta^{-1} \rfloor)^d} P(A_{h(q)}) \leq \eta.$$

Hieraus folgt dann wegen (2.3.15) die Behauptung. □

In der Literatur wird das Lemma 2.3.7 bisweilen bereits als *Straffheitskriterium von Kolmogorov* bezeichnet. Wir verwenden diese Bezeichnung hingegen für das nachfolgende Korollar, welches jedoch leicht aus dem obigen Kriterium gefolgert werden kann. Es zeigt, dass die Einschränkung auf den d -dimensionalen Einheitswürfel modifiziert werden kann, um auch die d -dimensionalen abgeschlossenen Bälle als Kompaktum zuzulassen.

³⁷Man beachte, dass die A_ε per Definition eigentlich auch von n abhängen.

Korollar 2.3.8. (Straffheitskriterium von Kolmogorov)

Es sei $\varepsilon > 0$ und $\vartheta_0 \in \mathbb{R}^d$. Eine Folge $(X_n)_{n \in \mathbb{N}}$ von Zufallsfunktionen mit Werten im Raum $C(\overline{B_\varepsilon(\vartheta_0)})$ ist straff, wenn die nachfolgenden Bedingungen erfüllt sind:

(i) Es gibt ein $\vartheta \in \overline{B_\varepsilon(\vartheta_0)}$, so dass für jedes $\eta > 0$ ein $a \in \mathbb{R}$ existiert mit

$$P(\{\|X_n(\vartheta)\| > a\}) \leq \eta, \quad \text{für alle } n \in \mathbb{N}. \quad (2.3.21)$$

(ii) Es gilt die *Momenten-Bedingung*

$$E\|X_n(\vartheta_2) - X_n(\vartheta_1)\|^2 \leq \nu \|\vartheta_2 - \vartheta_1\|^{d+\alpha}, \quad \text{für alle } \vartheta_1, \vartheta_2 \in \overline{B_\varepsilon(\vartheta_0)} \text{ und } n \in \mathbb{N}, \quad (2.3.22)$$

mit positiven Konstanten $\nu > 0$ und $\alpha > 0$.

Beweis³⁸:

Aus der Analysis ist die Existenz eines Bilipschitz-Homöomorphismus

$$f_1 : [0, 1]^d \longrightarrow \overline{B_1(0)}$$

bekannt. Als affine Abbildung ist auch

$$f_2 : \overline{B_1(0)} \longrightarrow \overline{B_\varepsilon(\vartheta_0)} : \vartheta \longmapsto \vartheta_0 + \varepsilon\vartheta \quad (2.3.23)$$

ein Bilipschitz-Homöomorphismus, so dass die Komposition $f = f_2 \circ f_1$ einen Bilipschitz-Homöomorphismus von $[0, 1]^d$ nach $\overline{B_\varepsilon(\vartheta_0)}$ liefert.

Sodann definieren wir die $C([0, 1]^d)$ -wertigen Zufallsfunktionen Y_n via

$$Y_n(t) = X_n(f(t)).$$

Damit existiert folglich für jedes $\eta > 0$ ein $a \in \mathbb{R}$ mit

$$P(\{\|Y_n(f^{-1}(\vartheta))\| > a\}) \leq \eta,$$

und es gilt zudem für alle $t_1, t_2 \in [0, 1]^d$:

$$\begin{aligned} P(\|Y_n(t_2) - Y_n(t_1)\| \geq \varepsilon) &= P(\|X_n(f(t_2)) - X_n(f(t_1))\| \geq \varepsilon) \\ &\leq \frac{\nu}{\varepsilon^2} \|f(t_2) - f(t_1)\|^{d+\alpha}. \end{aligned} \quad (2.3.24)$$

Nach Konstruktion ist f Lipschitz-stetig mit einer Lipschitz-Konstante L , und wir erhalten

³⁸Satz und Beweis wurden vom Autor dieser Arbeit formuliert.

somit aus der Abschätzung (2.3.24):

$$P\left(\|Y_n(t_2) - Y_n(t_1)\| \geq \varepsilon\right) \leq \frac{\nu L^{d+\alpha}}{\varepsilon^2} \|t_2 - t_1\|^{d+\alpha},$$

woraus insgesamt die Straffheit der Folge (Y_n) mit Lemma 2.3.7 folgt.

Folglich existiert gemäß Korollar 2.3.5 für jedes $\eta > 0$ und $\varepsilon > 0$ ein $\delta > 0$ und ein $n_0 \in \mathbb{N}$, so dass gilt:

$$P\left(\left\{\sup_{\|t_2 - t_1\| < \delta} \|Y_n(t_2) - Y_n(t_1)\| \geq \varepsilon\right\}\right) \leq \eta, \quad \text{für alle } n \geq n_0.$$

Die gleichmäßige Stetigkeit von f^{-1} ermöglicht es uns, ein $\tilde{\delta} > 0$ dergestalt zu wählen, dass für $\|\vartheta_2 - \vartheta_1\| < \tilde{\delta}$ stets $\|f^{-1}(\vartheta_2) - f^{-1}(\vartheta_1)\| < \delta$ folgt. Somit erhalten wir:

$$\begin{aligned} P\left(\left\{\sup_{\|\vartheta_2 - \vartheta_1\| < \tilde{\delta}} \|X_n(\vartheta_2) - X_n(\vartheta_1)\| \geq \varepsilon\right\}\right) &\leq P\left(\left\{\sup_{\|t_2 - t_1\| < \delta} \|X_n(f(t_2)) - X_n(f(t_1))\| \geq \varepsilon\right\}\right) \\ &= P\left(\left\{\sup_{\|t_2 - t_1\| < \delta} \|Y_n(t_2) - Y_n(t_1)\| \geq \varepsilon\right\}\right) \leq \eta, \quad \text{für alle } n \geq n_0, \end{aligned}$$

und damit erneut nach 2.3.5 die Straffheit der Folge $(X_n)_{n \in \mathbb{N}}$. □

Bemerkung 2.3.9.

- Im Beweis von Lemma 2.3.7 - man betrachte im Speziellen die Abschätzung aus (2.3.15) - haben wir den d -dimensionalen abgeschlossenen Einheitswürfel in Abhängigkeit von einem gegebenen $\delta > 0$ in geeignete kompakte Teilmengen unterteilt. Aus der nebenbei in Abschätzung (2.3.20) gezeigten Straffheit der auf eine derartige Teilmenge eingeschränkten Folge erhielten wir schließlich die Straffheit der auf ganz $[0, 1]^d$ definierten Folge.

Dieses Vorgehen lässt sich nun abstrahieren, um mit einem Überdeckungsargument das *Straffheitskriterium von Kolmogorov* auf beliebige Kompakta zu erweitern. Für das weitere Vorgehen ist dieser Sachverhalt jedoch nicht von Belang und soll hier deshalb lediglich Erwähnung finden.

- In Stute und Kopperschmidt 2013 wird das Straffheitskriterium aus Korollar 2.3.8 stets mit $d + \alpha = 2$ angewandt. Der obige Beweis deckt dann nur den Fall $d = 1$ ab, für den auch das in Billingsley 1968 zu findende Resultat genügt. Um ein von der Dimension d *unabhängiges* Straffheitskriterium zur Hand zu haben, wird deshalb im Anhang dieser Arbeit ein auf einer *Lipschitz-Bedingung* basierendes Alternativkriterium angegeben. Die Anwendung dieses Kriteriums erfordert jedoch mitunter eine Verschärfung der Voraussetzungen.

Aufgrund des zweiten Gesichtspunktes aus Bemerkung 2.3.9 sind die auf dem Straffheitskriterium von Kolmogorov basierenden Beweise bis auf weiteres nur im eindimensionalen Fall gültig.

Es ist möglich, die Beweise unter leicht modifizierten Voraussetzungen mit dem Lipschitz-Straffheitskriterium aus Korollar A.3.1 zu führen; dies würde jedoch den Rahmen dieser Arbeit sprengen.

Wir schließen hiermit das Kapitel der mathematischen Grundlagen ab. Im nächsten Kapitel, das wir dem *Minimum-Distanz-Schätzer* ϑ_n aus dem Artikel Stute und Kopperschmidt 2013 widmen, werden uns zahlreiche Konzepte speziell aus den Unterabschnitten 2.1.6, 2.1.7, 2.2.1 und 2.3.1 begegnen, die retrospektiv den Umfang dieser Einführung rechtfertigen.

3 Der Minimum-Distanz-Schätzer ϑ_n

In der Praxis der Ingenieur- und Finanzwissenschaften kann eine Vielzahl von Ereignissen durch sogenannte selbstanregende Punktprozesse modelliert werden, die auf natürliche Weise mit den entsprechenden Zählprozessen assoziiert sind (siehe Unterabschnitt 2.1.2 und Beispiel 2.1.7 aus Unterabschnitt 2.1.3, sowie Definition 2.1.34).

Unter geeigneten Voraussetzungen kann die Dynamik eines solchen Zählprozesses anschaulich durch die kumulierte Intensitätsfunktion beschrieben werden (siehe Satz 2.1.30 aus Unterabschnitt 2.1.6).

Durch den Einfluss externer Effekte, das heißt sobald eine *feinere* Filtration als die kanonische Filtration des Zählprozesses vorliegt, wird diese kumulierte Intensitätsfunktion jedoch selten vorhersagbar oder auch nur adaptiert bezüglich jener kanonischen Filtration sein (siehe erneut Unterabschnitt 2.1.5 und Stute und Kopperschmidt 2013, S. 1277).

In dieser Situation ist es ohne weitere Annahmen wie etwa der Unabhängigkeit der Zuwächse oder der Stationarität des Prozesses praktisch nicht möglich, eine exakte Verteilung zu bestimmen (siehe Stute und Kopperschmidt 2013, S. 1277).

Verzichten wir folglich auf jegliche Verteilungsannahmen in einer semiparametrischen Modellierung der kumulierten Intensitätsfunktion, so erweisen sich die üblichen Maximum-Likelihood-Methoden als unzureichend (vergleiche auch Stute und Kopperschmidt 2013). Ein alternativer Ansatz ist die Einführung eines *Minimum-Distanz-Schätzers*, dem die Idee der qualitativen Beschreibung des Zählprozesses durch die kumulierte Intensitätsfunktion zugrunde liegt (man rekapituliere die Ausführung aus den Unterabschnitten 2.1.4 sowie 2.1.6).

Um unnötige Restriktionen zu vermeiden, konstruieren wir diesen Schätzer auf Grundlage des Kompensators aus der Doob-Meyer-Zerlegung, der eine Verallgemeinerung der kumulierten Intensitätsfunktion darstellt und zur Vorhersage des Zählprozesses dient (siehe Sätze 2.1.28 und 2.1.30 aus Unterabschnitt 2.1.6, sowie Stute und Kopperschmidt 2013, S. 1278): Beobachten wir unabhängig identisch verteilte Kopien des selbstanregenden Zählprozesses, so wählen wir aus einem gegebenen semiparametrischen Modell für den Kompensator jene Funktion aus, die bezüglich eines geeigneten Abstandsbegriffes den minimalen Abstand zum aggregierten Zählprozess aufweist (siehe Stute und Kopperschmidt 2013, S. 1279).

Die zentralen Resultate dieses Kapitels behandeln die asymptotischen Eigenschaften des so definierten Parameterschätzers:

Der Minimum-Distanz-Schätzer ist einerseits stark Konsistenz (siehe Abschnitt 3.3) und andererseits asymptotisch normalverteilt (siehe Abschnitt 3.4).

Das vorliegende Kapitel orientiert sich in seinem Aufbau an dem Artikel Stute und Kopperschmidt 2013, der als zentrale Referenz in der statistischen Analyse der selbstanregenden

Zählprozesse betrachtet wird.

Wir beginnen mit der Definition des *Minimum-Distanz-Schätzers*, der wir jedoch einige notationsspezifische Konventionen voranstellen möchten.

3.1 Definition des Schätzers ϑ_n

In diesem Unterabschnitt fassen wir zunächst die Notationen, wie sie in dem Artikel *The Statistical Analysis of Self-Exciting Point Processes* von K. Kopperschmidt und W. Stute (siehe Stute und Kopperschmidt 2013) verwendet werden, zusammen, um anschließend den *Minimum-Distanz-Schätzer* ϑ_n formal zu definieren.

Zuvor erweitern wir die Definition eines Zählprozesses (siehe Definition 2.1.5 aus Unterabschnitt 2.1.2), um kompakte Intervalle als Indexmenge zuzulassen.

Definition 2.1.5. (*Zählprozess, Fortsetzung der Definition*)

Es bezeichne $I = [t_0, t^0] \subset \mathbb{R}$ ein kompaktes Intervall und $N = (N_t)_{t \geq t_0}$ einen Zählprozess. Dann heißt der stochastische Prozess $N|_I = (N_t)_{t \in I}$ die *Einschränkung* des Zählprozesses N auf das kompakte Intervall I .

Einen stochastischen Prozess mit Indexmenge I , der die Einschränkung eines Zählprozesses ist, bezeichnen wir suggestiv ebenfalls als *Zählprozess*.

Bemerkung 3.1.1. (*Fortsetzung eines Zählprozesses*)

- Ist erneut $I = [t_0, t^0]$ ein kompaktes Intervall und $N = (N_t)_{t \in I}$ ein Zählprozess im obigen Sinne, so können wir stets einen Zählprozess \tilde{N} mit Indexmenge $[t_0, \infty)$ konstruieren, dessen Einschränkung der stochastische Prozess N ist, indem wir definieren:

$$\tilde{N}_t = N_{\min(t, t^0)}, \quad \text{für } t \geq t_0, \quad (3.1.1)$$

das heißt wir setzen den Prozess N jenseits von t^0 konstant fort.

- Die Resultate aus Abschnitt 2.1 bleiben gültig, indem die Sätze auf die via Gleichung (3.1.1) definierte *Fortsetzung* des Zählprozesses N angewandt werden.
- Im Rahmen der sechsten Fortsetzung von Beispiel 2.1.7 im Vorfeld von Definition 2.1.34 haben wir bereits einen Zählprozess betrachtet, der faktisch auf dem kompakten Intervall $[0, t_1]$ definiert war:

Nach t_1 Lastzyklen wurde das dort thematisierte Experiment *unterbrochen*. So lässt sich die Einschränkung eines Zählprozesses in der Praxis als Unterbrechung oder auch Abbruch interpretieren, da im Anwendungskontext kein Versuch für alle Zeiten fortgeführt wird.

Wir möchten an dieser Stelle eine erste Generalvoraussetzung formulieren, die für den Rest dieser Arbeit gelte. Gleichzeitig möchten wir die Notationen der obigen Definition festhalten.

Fortan bezeichne $I = [t_0, t^0]$ ein kompaktes Intervall und $N = (N_t)_{t \in I}$ einen selbstanregenden Zählprozess über dem filtrierten Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \in I}, P)$, das heißt:

Der filtrierten Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \in I}, P)$ erfüllt die üblichen Bedingungen (siehe Ausführungen im Vorfeld von Lemma 2.1.26) und N ist bezüglich der Filtration $\{\mathcal{F}_t\}_{t \in I}$ adaptiert.

In dieser Situation³⁹ existiert gemäß Satz 2.1.28 die Doob-Meyer-Zerlegung des Zählprozesses N . Es bezeichne $\Lambda = (\Lambda_t)_{t \in I}$ den Kompensator von N und $M = (M_t)_{t \in I}$ das sogenannte *Innovationsmartingal* mit

$$M_t = N_t - \Lambda_t, \quad \text{für } t \in I.$$

Bekanntermaßen ist M ein Martingal bezüglich der Filtration $\{\mathcal{F}_t\}_{t \in I}$. In dieser Situation gilt insbesondere die Identität

$$\mathbb{E}N_t = \mathbb{E}\Lambda_t, \quad \text{für } t \in I, \quad (3.1.2)$$

weshalb wir den Kompensator Λ zur *Vorhersage* des Zählprozesses N nutzen können. Durch diesen Umstand motiviert, setzen wir ein semiparametrisches Modell für die kumulierte Intensitätsfunktion Λ an⁴⁰:

Es sei $d \in \mathbb{N}$ und $\Theta \subset \mathbb{R}^d$ eine beschränkte offene Menge. Das zugehörige Modell bezeichnen wir mit

$$\mathcal{H}_N = \{\Lambda_\vartheta : \vartheta \in \Theta\}, \quad (3.1.3)$$

wobei \mathcal{H}_N von der konkreten Realisierung des Zählprozesses N abhängen kann. Im Folgenden sei stets ϑ_0 der *wahre* Parameter, das heißt es gelte:

$$\Lambda = \Lambda_{\vartheta_0} \quad \text{für ein } \vartheta_0 \in \Theta. \quad (3.1.4)$$

³⁹Wir müssen formal noch die Integrierbarkeit des Zählprozesses N fordern. Dies geschieht retrospektiv am Ende dieses Unterabschnittes.

⁴⁰Die Gültigkeit von Gleichung (3.1.4) wird stets implizit *angenommen*, da sonst das angesetzte Modell schlichtweg nicht zur Beschreibung des Kompensators Λ geeignet ist.

Beispiel 2.1.7. (*7. Fortsetzung*)

Aus vorherigen Fortsetzungen des Beispiels ist bereits das vorläufige Modell

$$\tilde{\mathcal{H}}_N = \left\{ \Lambda_{\vartheta}(t) = \vartheta \int_{t_0}^t \frac{M}{M - N_s} \mathbb{1}_{\{N_s < M\}} ds : \vartheta > 0 \right\}$$

für $M \in \mathbb{N}$ bekannt, wobei wir $\Lambda_{\vartheta}(t)$ anstelle von $(\Lambda_{\vartheta})_t$ schreiben. Üblicherweise gilt $t_0 = 0$. Damit dieses Modell den Forderungen aus (3.1.3) genügt, muss der Parameterraum $(0, \infty)$, dem die Grundintensität ϑ entstammt, durch eine Konstante $\vartheta_{\max} > 0$ beschränkt werden; in der Praxis geschieht dies etwa durch Vorabversuche oder anderweitig erworbene Vorkenntnisse. Wir erhalten dann das Modell

$$\mathcal{H}_N = \left\{ \Lambda_{\vartheta}(t) = \vartheta \int_{t_0}^t \frac{M}{M - N_s} \mathbb{1}_{\{N_s < M\}} ds : \vartheta \in \Theta = (0, \vartheta_{\max}) \right\}. \quad (3.1.5)$$

Um den *wahren*, unbekanntem Parameter ϑ_0 zu schätzen, betrachten wir nachfolgend $n \in \mathbb{N}$ unabhängige Kopien $N^{(1)}, \dots, N^{(n)}$ des Zählprozesses N .

Während der Parameter ϑ_0 nun für alle $1 \leq i \leq n$ identisch sei und somit den über alle Beobachtungen gleich bleibenden Effekt beschreibe, kann das konkrete Modell von der jeweiligen Realisierung des assoziierten Zählprozesses abhängen. Wir fassen die obigen Ausführungen an dieser Stelle formal zusammen:

Es seien $N^{(1)}, \dots, N^{(n)}$ unabhängige, identische Kopien des Zählprozesses N . Ferner sei $\Theta \subset \mathbb{R}^d$ eine beschränkte offene Menge und für jedes $1 \leq i \leq n$ sei

$$\mathcal{H}_{N^{(i)}} = \{ \Lambda_{\vartheta}^{(i)} : \vartheta \in \Theta \}$$

eine Modell für den wahren Kompensator $\Lambda^{(i)}$ des Zählprozesses $N^{(i)}$ dergestalt, dass für *alle* $1 \leq i \leq n$ gilt:

$$\Lambda^{(i)} = \Lambda_{\vartheta_0}^{(i)}. \quad (3.1.6)$$

Wir bezeichnen ϑ_0 als den *wahren* Parameter des Kompensators Λ .

Im Sinne von Gleichung (3.1.6) besteht das vorrangige Interesse nun in der Schätzung des Parameters ϑ_0 . Zur Vermeidung weiterer Verteilungsannahmen in der bewusst vagen Modellierung verzichten wir nachfolgend auf einen in dieser Situation kaum anwendbaren *Maximum-Likelihood-Ansatz* (vgl. auch Stute und Kopperschmidt 2013). Der nachfolgend vorgestellte Schätzer ϑ_n nutzt einen *Minimum-Distanz-Ansatz* und benötigt dabei keine weiteren Verteilungsannahmen.

Wir führen noch einige weitere Definitionen in Anlehnung an den Artikel Stute und Kopperschmidt 2013 ein:

Es sei μ ein endliches Maß auf der Borel'schen σ -Algebra $\mathcal{B}(I)$.

Für $f, g \in L^2(I, \mathcal{B}(I), \mu)$ definieren wir:

$$\langle f, g \rangle_\mu = \int_I fg \, d\mu, \quad (3.1.7)$$

sowie die induzierte Halbnorm

$$\|f\|_\mu = \sqrt{\langle f, f \rangle_\mu} = \left[\int_I f^2 \, d\mu \right]^{\frac{1}{2}}. \quad (3.1.8)$$

Ferner definieren wir den *aggregierten Zählprozess* $\bar{N}^{(n)}$ sowie den *aggregierten Kompensator* $\bar{\Lambda}^{(n)}$ durch

$$\bar{N}^{(n)} = \frac{1}{n} \sum_{i=1}^n N^{(i)}, \quad \bar{\Lambda}_\vartheta^{(n)} = \frac{1}{n} \sum_{i=1}^n \Lambda_\vartheta^{(i)}. \quad (3.1.9)$$

Das zugehörige Innovationsmartingal $\bar{M}^{(n)}$ ist dann gegeben durch

$$\bar{M}^{(n)} = \bar{N}^{(n)} - \bar{\Lambda}_{\vartheta_0}^{(n)}. \quad (3.1.10)$$

Bezeichnet δ_t das *Dirac*-Maß in t und $\mathcal{J}_N = \{t : N_t - \lim_{s \uparrow t} N_s \geq 1\}$ die Menge aller Sprünge eines Zählprozesses N , so definieren wir das durch N induzierte Maß μ_N via

$$\mu_N = \sum_{t \in \mathcal{J}_N} \left(N_t - \lim_{s \uparrow t} N_s \right) \delta_t.$$

Insbesondere können wir für μ in Gleichung (3.1.7) und (3.1.8) das durch den aggregierten Zählprozess $\bar{N}^{(n)}$ induzierte Maß

$$\mu_{\bar{N}^{(n)}} = \frac{1}{n} \sum_{i=1}^n \mu_{N^{(i)}} = \sum_{t \in \mathcal{J}_{\bar{N}^{(n)}}} \left(\bar{N}_t^{(n)} - \lim_{s \uparrow t} \bar{N}_s^{(n)} \right) \delta_t$$

wählen. Hiermit können wir nun ein Maß für die Güte der Anpassung von $\bar{\Lambda}_\vartheta^{(n)}$ an $\bar{N}^{(n)}$ definieren:

$$\|\bar{N}^{(n)} - \bar{\Lambda}_\vartheta^{(n)}\|_{\mu_{\bar{N}^{(n)}}}. \quad (3.1.11)$$

Der Term aus (3.1.11) lässt sich - dem Verständnis dienlich - auch schreiben als

$$\left[\sum_{t \in \mathcal{J}_{\bar{N}^{(n)}}} \left(\bar{N}_t^{(n)} - \bar{\Lambda}_{\vartheta}^{(n)}(t) \right)^2 \left(\bar{N}_t^{(n)} - \lim_{s \uparrow t} \bar{N}_s^{(n)} \right) \right]^{\frac{1}{2}},$$

ergo misst die Größe $\|\bar{N}^{(n)} - \bar{\Lambda}_{\vartheta}^{(n)}\|_{\mu_{\bar{N}^{(n)}}}$ im Wesentlichen die quadratischen Abstände zwischen den aggregierten Prozessen an den Sprungstellen des Prozesses $\bar{N}^{(n)}$, gewichtet mit der Höhe des jeweiligen Sprunges. Anhand dieses Abstandsbegriffes können wir nun den *Minimum-Distanz-Schätzer* ϑ_n definieren.

Definition 3.1.2. (*Der Minimum-Distanz-Schätzer* ϑ_n , siehe Stute und Kopperschmidt 2013, S. 1279)

Der Schätzer ϑ_n ist definiert als das Element des Parameterraumes Θ , welches den Abstand in (3.1.11) minimiert, das heißt⁴¹

$$\vartheta_n = \arg \inf_{\vartheta \in \Theta} \|\bar{N}^{(n)} - \bar{\Lambda}_{\vartheta}^{(n)}\|_{\mu_{\bar{N}^{(n)}}}.$$

Eine technische wenngleich notwendige Forderung ist zudem gegeben durch die Annahme, dass gilt:

$$E_{\vartheta_0} \left[\sup_{\vartheta \in \Theta} \left\{ N(t^0)^i \cdot \Lambda_{\vartheta}(t^0)^j \right\} \right] < \infty, \quad \text{für alle } i, j \in \mathbb{N}_0 \text{ mit } i + j \leq 3, \quad (3.1.12)$$

während in Stute und Kopperschmidt 2013 die bedeutend schwächere Forderung

$$E_{\vartheta_0} [N(t^0)] < \infty \quad \wedge \quad E_{\vartheta_0} [\Lambda_{\vartheta}(t^0)] < \infty \quad \text{für alle } \vartheta \in \Theta$$

gestellt wird, damit die Integrierbarkeit der untersuchten Prozesse und somit die Existenz der Doob-Meyer-Zerlegung gewährleistet ist.

Um die Existenz relevanter Grenzwerte sicherzustellen (man siehe etwa Lemma 3.2.1), verschärfen wir diese Forderung in (3.1.12). Die Annahme wird jedoch obsolet, wenn wir später die bloße Existenz der untersuchten Integrale bereits *voraussetzen*. Dieses Vorgehen mag implizit in Stute und Kopperschmidt 2013 praktiziert worden sein, die stärkere Forderung ruft jedoch keinen praxisrelevanten Nachteil hervor.

⁴¹In der Statistik ist es nicht unüblich, einen Schätzer mit einem Zirkumflex zu versehen. Da wir dem Schätzer häufig in Form eines Index begegnen werden, verzichten wir zugunsten der Übersichtlichkeit darauf, den Schätzer mit $\hat{\vartheta}_n$ zu bezeichnen.

Beispiel 2.1.7. (8. Fortsetzung)

Im Falle eines reinen verschobenen Geburtsprozesses mit Abbruch bei $M \in \mathbb{N}$ (siehe Unterabschnitt 2.1.4) ist bei Gültigkeit des Modells \mathcal{H}_N aus Gleichung (3.1.5) der siebten Fortsetzung dieses Beispiels der Erwartungswert in (3.1.12) beschränkt durch

$$\left[\vartheta_{\max} \cdot M(t^0 - t_0) \right]^3 < \infty,$$

da der Zählprozess selbst durch M und die bedingte Intensitätsfunktion durch $\vartheta_{\max} M$ beschränkt ist. Die obige Forderung ist hier also trivialerweise erfüllt.

Wir bemerken noch, dass wir anstelle der Schreibweise $E_{\vartheta_0}[\dots]$ häufig die kürzere Schreibweise $E[\dots]$ verwenden werden.

In den folgenden Abschnitten werden wir zunächst die *starke Konsistenz* des Schätzers ϑ_n beweisen, bevor wir uns der *asymptotischen Normalverteilung* zuwenden.

3.2 Folgerungen aus dem starken Gesetz der großen Zahlen für U-Statistiken

Als Vorbereitung für den Beweis der starken Konsistenz des Minimum-Distanz Schätzers ϑ_n werden wir in diesem Abschnitt einige Aussagen aus dem Artikel Stute und Kopperschmidt 2013, die unmittelbar aus dem starken Gesetz der großen Zahlen für U-Statistiken (siehe Satz 2.2.3) folgen, formulieren und beweisen.

Dafür führen wir noch einige neue Notationen ein:

Für $\varepsilon > 0$ bezeichne

$$B_\varepsilon(\vartheta_0) = \{\vartheta \in \mathbb{R}^d : \|\vartheta - \vartheta_0\| < \varepsilon\} \tag{3.2.1}$$

den offenen ε -Ball um ϑ_0 . Es sei ferner $r > 0$. Dann definieren wir:

$$B_r^\varepsilon(\vartheta) = \{\vartheta' \in \overline{\Theta} \setminus B_\varepsilon(\vartheta_0) : \|\vartheta' - \vartheta\| < r\}, \tag{3.2.2}$$

wobei $\overline{\Theta}$ den Abschluss von Θ im \mathbb{R}^d bezeichne. $B_r^\varepsilon(\vartheta)$ bezeichnet also den Teil des r -Balls um ϑ , der *nicht* zu $B_\varepsilon(\vartheta_0)$ gehört.

Wie bereits gegen Ende von Abschnitt 3.1 angekündigt, werden wir nachfolgend die abkürzende Notation $E[\dots]$ anstelle der unübersichtlicheren Schreibweise $E_{\vartheta_0}[\dots]$ verwenden.

Schließlich stellen wir noch eine Glattheitsforderung an den Kompensator, die wir im weiteren Verlauf dieser Arbeit noch verschärfen werden⁴². Konkret fordern wir:

⁴²Durch zusätzliche Forderung der Differenzierbarkeit wird rückwirkend der Begriff der *Glattheitsforderung* anstelle einer *Stetigkeitsforderung* motiviert.

Der Prozess

$$(t, \vartheta) \longmapsto \Lambda_{\vartheta}(t) \quad (3.2.3)$$

sei fast sicher stetig und besitze eine stetige Fortsetzung auf $I \times \overline{\Theta}$.

Beispiel 2.1.7. (9. Fortsetzung)

Dass die Forderung der fast sicheren Stetigkeit des Kompensators bezüglich der zeitlichen Komponente gewissermaßen natürlich erscheint, wurde bereits in Bemerkung 2.1.29 erläutert. Darüber hinaus erfüllt der Prozess

$$(t, \vartheta) \longmapsto \Lambda_{\vartheta}(t) = \vartheta \int_{t_0}^t \frac{M}{M - N_s} \mathbb{1}_{\{N_s < M\}} ds$$

durch seine simple Abhängigkeit vom Parameter ϑ offensichtlich die Glattheitsforderung aus (3.2.3), die also mit unserem Praxisbeispiel vereinbar ist.

Lemma 3.2.1. (Siehe Stute und Kopperschmidt 2013, S. 1285)

In der Situation aus Abschnitt 3.1 gelte ferner die Stetigkeitsforderung aus Gleichung (3.2.3). Dann gilt fast sicher:

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \sup_{\vartheta' \in B_r^{\varepsilon}(\vartheta)} \|N^{(i)} - \Lambda_{\vartheta'}^{(i)}\|_{\mu_{N^{(i)}}}^2 = \mathbb{E} \left[\sup_{\vartheta' \in B_r^{\varepsilon}(\vartheta)} \|N - \Lambda_{\vartheta'}\|_{\mu_N}^2 \right], \quad (3.2.4)$$

$$\lim_{n \rightarrow \infty} \frac{1}{n(n-1)} \sum_{\substack{i,j=1 \\ i \neq j}}^n \sup_{\vartheta' \in B_r^{\varepsilon}(\vartheta)} \|N^{(i)} - \Lambda_{\vartheta'}^{(i)}\|_{\mu_{N^{(j)}}}^2 = \mathbb{E} \left[\sup_{\vartheta' \in B_r^{\varepsilon}(\vartheta)} \|N^{(1)} - \Lambda_{\vartheta'}^{(1)}\|_{\mu_{N^{(2)}}}^2 \right], \quad (3.2.5)$$

$$\begin{aligned} & \lim_{n \rightarrow \infty} \frac{1}{n(n-1)} \sum_{\substack{i,j=1 \\ i \neq j}}^n \sup_{\vartheta' \in B_r^{\varepsilon}(\vartheta)} \langle N^{(i)} - \Lambda_{\vartheta'}^{(i)}, N^{(j)} - \Lambda_{\vartheta'}^{(j)} \rangle_{\mu_{N^{(i)}}} \\ &= \mathbb{E} \left[\sup_{\vartheta' \in B_r^{\varepsilon}(\vartheta)} \langle N^{(1)} - \Lambda_{\vartheta'}^{(1)}, N^{(2)} - \Lambda_{\vartheta'}^{(2)} \rangle_{\mu_{N^{(1)}}} \right], \end{aligned} \quad (3.2.6)$$

$$\begin{aligned} & \lim_{n \rightarrow \infty} \frac{1}{n(n-1)(n-2)} \sum_{\substack{i,j,k=1 \\ i \neq j \neq k \neq i}}^n \sup_{\vartheta' \in B_r^{\varepsilon}(\vartheta)} \langle N^{(i)} - \Lambda_{\vartheta'}^{(i)}, N^{(j)} - \Lambda_{\vartheta'}^{(j)} \rangle_{\mu_{N^{(k)}}} \\ &= \mathbb{E} \left[\sup_{\vartheta' \in B_r^{\varepsilon}(\vartheta)} \langle N^{(1)} - \Lambda_{\vartheta'}^{(1)}, N^{(2)} - \Lambda_{\vartheta'}^{(2)} \rangle_{\mu_{N^{(3)}}} \right]. \end{aligned} \quad (3.2.7)$$

Die Aussagen in den Gleichungen (3.2.4) bis (3.2.7) gelten analog, wenn das *Supremum* jeweils durch ein *Infimum* ersetzt wird.

In den Gleichungen (3.2.5), (3.2.6) und (3.2.7) ist weiter die *Unabhängigkeit* der Kopien $N^{(1)}, N^{(2)}, \dots$ des Zählprozesses N wichtig, nicht jedoch die Wahl der konkreten Kopie; die Wahl der *ersten* drei Kopien etwa in Gleichung (3.2.7) ist daher völlig willkürlich.

Beweis⁴³:

Zunächst bemerken wir, dass die wir die stochastischen Prozesse $N^{(i)}$ für $i \in \mathbb{N}$ als Zufallsvariablen mit Werten im Skorokhod-Raum $(\mathcal{D}(I), \mathcal{T}_d)$ der *càdlàg*-Funktionen auffassen können, siehe Definition 2.2.6 in Abschnitt 2.2.

Gemäß Satz 2.2.7 ist der Skorokhod-Raum ein polnischer Raum, und folglich können wir $(N^{(n)})_{n \in \mathbb{N}}$ als Folge unabhängig identisch verteilter Zufallsvariablen mit Werten in einem polnischen Raum auffassen.

Für jede der Gleichungen (3.2.4) bis (3.2.7) lässt sich nun die linke Seite als Grenzwert einer U-Statistik auffassen. Die Kerne dieser U-Statistiken lassen sich dabei leicht ablesen, wobei für den Zusatz das Supremum durch ein Infimum ersetzt wird:

$$\begin{aligned}\psi_1(N) &= \sup_{\vartheta' \in \mathcal{B}_r^\varepsilon(\vartheta)} \|N - \Lambda_{\vartheta'}\|_{\mu_N}^2, \\ \psi_2(N^{(1)}, N^{(2)}) &= \sup_{\vartheta' \in \mathcal{B}_r^\varepsilon(\vartheta)} \|N^{(1)} - \Lambda_{\vartheta'}^{(1)}\|_{\mu_{N^{(2)}}}^2, \\ \psi_3(N^{(1)}, N^{(2)}) &= \sup_{\vartheta' \in \mathcal{B}_r^\varepsilon(\vartheta)} \langle N^{(1)} - \Lambda_{\vartheta'}^{(1)}, N^{(2)} - \Lambda_{\vartheta'}^{(2)} \rangle_{\mu_{N^{(1)}}}, \\ \psi_4(N^{(1)}, N^{(2)}, N^{(3)}) &= \sup_{\vartheta' \in \mathcal{B}_r^\varepsilon(\vartheta)} \langle N^{(1)} - \Lambda_{\vartheta'}^{(1)}, N^{(2)} - \Lambda_{\vartheta'}^{(2)} \rangle_{\mu_{N^{(3)}}}.\end{aligned}$$

Die Behauptung folgt dann sofort aus dem für diese Situation formulierten starken Gesetz der großen Zahlen für U-Statistiken (siehe Satz 2.2.3), insofern die Erwartungswerte in den Gleichungen (3.2.4) bis (3.2.7) existieren. Schlüssel zur Existenz wird dabei die Annahme aus (3.1.12) sein, die die nachfolgenden Abschätzungen unter Beachtung der Unabhängigkeit verschiedener Kopien des Zählprozesses N motiviert, wobei wir bisweilen die Schreibweise $N(t)$ anstelle von N_t präferieren:

$$\begin{aligned}\mathbb{E} |\psi_1(N)| &= \mathbb{E} \left[\sup_{\vartheta' \in \mathcal{B}_r^\varepsilon(\vartheta)} \|N - \Lambda_{\vartheta'}\|_{\mu_N}^2 \right] = \mathbb{E} \left[\sup_{\vartheta' \in \mathcal{B}_r^\varepsilon(\vartheta)} \left\{ \int_I (N - \Lambda_{\vartheta'})^2(t) \, d\mu_N(t) \right\} \right] \\ &\leq \mathbb{E} \left[\sup_{\vartheta' \in \mathcal{B}_r^\varepsilon(\vartheta)} \left\{ \int_I N^2(t) + \Lambda_{\vartheta'}^2(t) \, d\mu_N(t) \right\} \right] \\ &\leq \mathbb{E} \left[\sup_{\vartheta' \in \mathcal{B}_r^\varepsilon(\vartheta)} \left\{ (N^2(t^0) + \Lambda_{\vartheta'}^2(t^0)) \int_I d\mu_N \right\} \right] \\ &= \mathbb{E} \left[\sup_{\vartheta' \in \mathcal{B}_r^\varepsilon(\vartheta)} \left\{ (N^2(t^0) + \Lambda_{\vartheta'}^2(t^0)) \left(\sum_{t \in \mathcal{J}_N} (N_t - \lim_{s \uparrow t} N_s) \right) \right\} \right]\end{aligned}$$

⁴³Der Beweis bedient sich der im Vergleich zu dem Artikel Stute und Kopperschmidt 2013 verschärfte Bedingung (3.1.12) und wurde vom Autor dieser Arbeit selbst geführt. Stute und Kopperschmidt verweisen lediglich auf das starke Gesetz der großen Zahlen für U-Statistiken.

$$\begin{aligned}
&= \mathbb{E} \left[\sup_{\vartheta' \in \mathbb{B}_r^\varepsilon(\vartheta)} \left\{ (N^2(t^0) + \Lambda_{\vartheta'}^2(t^0)) N(t^0) \right\} \right] \\
&\leq \mathbb{E} [N^3(t^0)] + \mathbb{E} \left[\sup_{\vartheta' \in \mathbb{B}_r^\varepsilon(\vartheta)} \left\{ N(t^0) \cdot \Lambda_{\vartheta'}^2(t^0) \right\} \right] \stackrel{(3.1.12)}{<} \infty,
\end{aligned}$$

$$\begin{aligned}
\mathbb{E} \left| \psi_2 \left(N^{(1)}, N^{(2)} \right) \right| &= \mathbb{E} \left[\sup_{\vartheta' \in \mathbb{B}_r^\varepsilon(\vartheta)} \|N^{(1)} - \Lambda_{\vartheta'}^{(1)}\|_{\mu_{N^{(2)}}}^2 \right] = \mathbb{E} \left[\sup_{\vartheta' \in \mathbb{B}_r^\varepsilon(\vartheta)} \left\{ \int_I \left(N^{(1)} - \Lambda_{\vartheta'}^{(1)} \right)^2(t) \, \mathrm{d}\mu_{N^{(2)}}(t) \right\} \right] \\
&\stackrel{\text{analog}}{\leq} \mathbb{E} \left[\sup_{\vartheta' \in \mathbb{B}_r^\varepsilon(\vartheta)} \left\{ \left((N^{(1)})^2(t^0) + (\Lambda_{\vartheta'}^{(1)})^2(t^0) \right) N^{(2)}(t^0) \right\} \right] \\
&\leq \mathbb{E} \left[(N^{(1)})^2(t^0) \right] \cdot \mathbb{E} [N^{(2)}(t^0)] + \mathbb{E} \left[\sup_{\vartheta' \in \mathbb{B}_r^\varepsilon(\vartheta)} \left\{ (\Lambda_{\vartheta'}^{(1)})^2(t^0) \right\} \right] \cdot \mathbb{E} [N^{(2)}(t^0)] \stackrel{(3.1.12)}{<} \infty,
\end{aligned}$$

$$\begin{aligned}
\mathbb{E} \left| \psi_3 \left(N^{(1)}, N^{(2)} \right) \right| &= \mathbb{E} \left[\sup_{\vartheta' \in \mathbb{B}_r^\varepsilon(\vartheta)} \langle N^{(1)} - \Lambda_{\vartheta'}^{(1)}, N^{(2)} - \Lambda_{\vartheta'}^{(2)} \rangle_{\mu_{N^{(1)}}} \right] \\
&= \mathbb{E} \left[\sup_{\vartheta' \in \mathbb{B}_r^\varepsilon(\vartheta)} \left\{ \int_I \left(N^{(1)} - \Lambda_{\vartheta'}^{(1)} \right) \left(N^{(2)} - \Lambda_{\vartheta'}^{(2)} \right) (t) \, \mathrm{d}\mu_{N^{(1)}}(t) \right\} \right] \\
&\leq \mathbb{E} \left[\sup_{\vartheta' \in \mathbb{B}_r^\varepsilon(\vartheta)} \left\{ \int_I \left(N^{(1)} N^{(2)} + \Lambda_{\vartheta'}^{(1)} \Lambda_{\vartheta'}^{(2)} \right) (t) \, \mathrm{d}\mu_{N^{(1)}}(t) \right\} \right] \\
&\leq \mathbb{E} \left[\sup_{\vartheta' \in \mathbb{B}_r^\varepsilon(\vartheta)} \left\{ \left(N^{(1)} N^{(2)} + \Lambda_{\vartheta'}^{(1)} \Lambda_{\vartheta'}^{(2)} \right) (t^0) N^{(1)}(t^0) \right\} \right] \\
&\leq \mathbb{E} \left[(N^{(1)})^2(t^0) \right] \cdot \mathbb{E} [N^{(2)}(t^0)] + \mathbb{E} \left[\sup_{\vartheta' \in \mathbb{B}_r^\varepsilon(\vartheta)} \left\{ N^{(1)}(t^0) \cdot \Lambda_{\vartheta'}^{(1)}(t^0) \right\} \right] \cdot \mathbb{E} \left[\sup_{\vartheta' \in \mathbb{B}_r^\varepsilon(\vartheta)} \left\{ \Lambda_{\vartheta'}^{(2)}(t^0) \right\} \right] \\
&\stackrel{(3.1.12)}{<} \infty,
\end{aligned}$$

$$\begin{aligned}
\mathbb{E} \left| \psi_4 \left(N^{(1)}, N^{(2)}, N^{(3)} \right) \right| &= \mathbb{E} \left[\sup_{\vartheta' \in \mathbb{B}_r^\varepsilon(\vartheta)} \langle N^{(1)} - \Lambda_{\vartheta'}^{(1)}, N^{(2)} - \Lambda_{\vartheta'}^{(2)} \rangle_{\mu_{N^{(3)}}} \right] \\
&\stackrel{\text{analog}}{\leq} \mathbb{E} \left[\sup_{\vartheta' \in \mathbb{B}_r^\varepsilon(\vartheta)} \left\{ \left(N^{(1)} N^{(2)} + \Lambda_{\vartheta'}^{(1)} \Lambda_{\vartheta'}^{(2)} \right) (t^0) N^{(3)}(t^0) \right\} \right] \\
&\leq \mathbb{E} [N^{(1)}(t^0)] \cdot \mathbb{E} [N^{(2)}(t^0)] \cdot \mathbb{E} [N^{(3)}(t^0)] \\
&\quad + \mathbb{E} \left[\sup_{\vartheta' \in \mathbb{B}_r^\varepsilon(\vartheta)} \left\{ \Lambda_{\vartheta'}^{(1)}(t^0) \right\} \right] \cdot \mathbb{E} \left[\sup_{\vartheta' \in \mathbb{B}_r^\varepsilon(\vartheta)} \left\{ \Lambda_{\vartheta'}^{(2)}(t^0) \right\} \right] \cdot \mathbb{E} [N^{(3)}(t^0)] \\
&\stackrel{(3.1.12)}{<} \infty.
\end{aligned}$$

Die Existenz der Erwartungswerte liefert somit unter Verwendung des starken Gesetzes der großen Zahlen für U-Statistiken die Behauptung. \square

Der Beweis von Lemma 3.2.1 liefert unmittelbar auch die Gültigkeit des nachfolgenden Korollars.

Korollar 3.2.2. (Siehe Stute und Kopperschmidt 2013, S. 1285)

In der Situation aus Abschnitt 3.1 gelte ferner die Stetigkeitsforderung aus Gleichung (3.2.3). Dann gilt fast sicher:

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \|N^{(i)} - \Lambda_{\vartheta}^{(i)}\|_{\mu_{N^{(i)}}}^2 = \mathbb{E} [\|N - \Lambda_{\vartheta}\|_{\mu_N}^2], \quad (3.2.8)$$

$$\lim_{n \rightarrow \infty} \frac{1}{n(n-1)} \sum_{\substack{i,j=1 \\ i \neq j}}^n \|N^{(i)} - \Lambda_{\vartheta}^{(i)}\|_{\mu_{N^{(j)}}}^2 = \mathbb{E} [\|N^{(1)} - \Lambda_{\vartheta}^{(1)}\|_{\mu_{N^{(2)}}}^2], \quad (3.2.9)$$

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{n(n-1)} \sum_{\substack{i,j=1 \\ i \neq j}}^n \langle N^{(i)} - \Lambda_{\vartheta}^{(i)}, N^{(j)} - \Lambda_{\vartheta}^{(j)} \rangle_{\mu_{N^{(i)}}} \\ = \mathbb{E} [\langle N^{(1)} - \Lambda_{\vartheta}^{(1)}, N^{(2)} - \Lambda_{\vartheta}^{(2)} \rangle_{\mu_{N^{(1)}}}], \end{aligned} \quad (3.2.10)$$

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{n(n-1)(n-2)} \sum_{\substack{i,j,k=1 \\ i \neq j \neq k \neq i}}^n \langle N^{(i)} - \Lambda_{\vartheta}^{(i)}, N^{(j)} - \Lambda_{\vartheta}^{(j)} \rangle_{\mu_{N^{(k)}}} \\ = \mathbb{E} [\langle N^{(1)} - \Lambda_{\vartheta}^{(1)}, N^{(2)} - \Lambda_{\vartheta}^{(2)} \rangle_{\mu_{N^{(3)}}}]. \end{aligned} \quad (3.2.11)$$

Beweis:

Der Beweis folgt völlig analog zum Beweis von Lemma 3.2.1. \square

Den Grenzwert aus Gleichung (3.2.11) werden wir im folgenden Lemma explizit berechnen.

Lemma 3.2.3. (Siehe Stute und Kopperschmidt 2013, S. 1285)

Unter den Voraussetzungen von Korollar 3.2.2 gilt für den Grenzwert aus Gleichung (3.2.11):

$$\mathbb{E} [\langle N^{(1)} - \Lambda_{\vartheta}^{(1)}, N^{(2)} - \Lambda_{\vartheta}^{(2)} \rangle_{\mu_{N^{(3)}}}] = \|\mathbb{E}\Lambda_{\vartheta_0} - \mathbb{E}\Lambda_{\vartheta}\|_{\mathbb{E}\Lambda_{\vartheta_0}}^2. \quad (3.2.12)$$

Die Integration auf der rechten Seite von Gleichung (3.2.12) sei dabei die klassische *Riemann-Stieltjes*-Integration bezüglich der Funktion $\mathbb{E}\Lambda_{\vartheta_0}$.

Beweis⁴⁴:

Um Gleichung (3.2.12) zu zeigen, berechnen wir zunächst unter Beachtung der Unabhängigkeit von $N^{(1)} - \Lambda_{\vartheta}^{(1)}$ und $N^{(2)} - \Lambda_{\vartheta}^{(2)}$ sowie dem *Satz von Fubini*:

$$\begin{aligned}
& \mathbb{E} \left[\langle N^{(1)} - \Lambda_{\vartheta}^{(1)}, N^{(2)} - \Lambda_{\vartheta}^{(2)} \rangle_{\mu_{N^{(3)}}} \mid N^{(3)} \right] \\
&= \mathbb{E} \left[\int_I \left(N^{(1)} - \Lambda_{\vartheta}^{(1)} \right) \left(N^{(2)} - \Lambda_{\vartheta}^{(2)} \right) d\mu_{N^{(3)}} \mid N^{(3)} \right] \\
&= \int_I \mathbb{E} \left[\left(N^{(1)} - \Lambda_{\vartheta}^{(1)} \right) \left(N^{(2)} - \Lambda_{\vartheta}^{(2)} \right) \right] d\mu_{N^{(3)}} \\
&= \int_I \left(\mathbb{E} \left[N^{(1)} - \Lambda_{\vartheta}^{(1)} \right] \right)^2 d\mu_{N^{(3)}} \\
&= \int_I (\mathbb{E}\Lambda_{\vartheta_0} - \mathbb{E}\Lambda_{\vartheta})^2 d\mu_{N^{(3)}}, \tag{3.2.13}
\end{aligned}$$

wobei $N^{(3)}$ im letzten Glied der Gleichungskette gegeben sei.

Mittels Gleichung (3.2.13) können wir nun unter Verwendung der Identität (siehe etwa Bauer 1991, S. 120)

$$\mathbb{E} \left[\langle N^{(1)} - \Lambda_{\vartheta}^{(1)}, N^{(2)} - \Lambda_{\vartheta}^{(2)} \rangle_{\mu_{N^{(3)}}} \right] = \mathbb{E} \left\{ \mathbb{E} \left[\langle N^{(1)} - \Lambda_{\vartheta}^{(1)}, N^{(2)} - \Lambda_{\vartheta}^{(2)} \rangle_{\mu_{N^{(3)}}} \mid N^{(3)} \right] \right\}$$

den Grenzwert berechnen. Zu beachten ist dabei, dass der Integrand aus Gleichung (3.2.13) nach Voraussetzung (3.2.3) stetig ist, so dass wir das Integral unter Verwendung einer *Riemann-Approximation* berechnen können. Es sei dazu $(S_n)_{n \in \mathbb{N}}$ eine Folge von Partitionen des Intervalls I , so dass

$$S_n = \{t_0 = s_{n,0} < s_{n,1} < \dots < s_{n,n} = t^0\}$$

die Bedingung

$$\max_{i=0, \dots, n-1} (s_{n,i+1} - s_{n,i}) \rightarrow 0 \quad (n \rightarrow \infty)$$

erfüllt. Dann gilt unter Verwendung der majorisierten Konvergenz:

$$\begin{aligned}
& \mathbb{E} \left\{ \mathbb{E} \left[\langle N^{(1)} - \Lambda_{\vartheta}^{(1)}, N^{(2)} - \Lambda_{\vartheta}^{(2)} \rangle_{\mu_{N^{(3)}}} \mid N^{(3)} \right] \right\} = \mathbb{E} \left\{ \int_I (\mathbb{E}\Lambda_{\vartheta_0} - \mathbb{E}\Lambda_{\vartheta})^2 d\mu_{N^{(3)}} \right\} \\
&= \mathbb{E} \left\{ \lim_{n \rightarrow \infty} \sum_{i=0}^{n-1} (\mathbb{E}\Lambda_{\vartheta_0} - \mathbb{E}\Lambda_{\vartheta})^2 (s_{n,i}) \left(N^{(3)}(s_{n,i+1}) - N^{(3)}(s_{n,i}) \right) \right\}
\end{aligned}$$

⁴⁴Stute und Kopperschmidt 2013 geben den Hinweis, auf $N^{(3)}$ zu bedingen und den *Satz von Fubini* zu utilisieren. Die detaillierte Ausarbeitung des Beweises geht auf den Autor dieser Arbeit zurück.

$$\begin{aligned}
&= \lim_{n \rightarrow \infty} \mathbb{E} \left\{ \sum_{i=0}^{n-1} (\mathbb{E}\Lambda_{\vartheta_0} - \mathbb{E}\Lambda_{\vartheta})^2 (s_{n,i}) \left(N^{(3)}(s_{n,i+1}) - N^{(3)}(s_{n,i}) \right) \right\} \\
&= \lim_{n \rightarrow \infty} \sum_{i=0}^{n-1} (\mathbb{E}\Lambda_{\vartheta_0} - \mathbb{E}\Lambda_{\vartheta})^2 (s_{n,i}) \left(\mathbb{E}[N^{(3)}](s_{n,i+1}) - \mathbb{E}[N^{(3)}](s_{n,i}) \right) \\
&= \lim_{n \rightarrow \infty} \sum_{i=0}^{n-1} (\mathbb{E}\Lambda_{\vartheta_0} - \mathbb{E}\Lambda_{\vartheta})^2 (s_{n,i}) \left(\mathbb{E}[\Lambda_{\vartheta_0}](s_{n,i+1}) - \mathbb{E}[\Lambda_{\vartheta_0}](s_{n,i}) \right) \\
&= \int_I (\mathbb{E}\Lambda_{\vartheta_0} - \mathbb{E}\Lambda_{\vartheta})^2 d\mathbb{E}\Lambda_{\vartheta_0} = \|\mathbb{E}\Lambda_{\vartheta_0} - \mathbb{E}\Lambda_{\vartheta}\|_{\mathbb{E}\Lambda_{\vartheta_0}}^2.
\end{aligned} \tag{3.2.14}$$

Die konvergente Majorante in Gleichung (3.2.14) erhält man via

$$\begin{aligned}
&\sum_{i=0}^{n-1} (\mathbb{E}\Lambda_{\vartheta_0} - \mathbb{E}\Lambda_{\vartheta})^2 (s_{n,i}) \left(N^{(3)}(s_{n,i+1}) - N^{(3)}(s_{n,i}) \right) \\
&\leq \sum_{i=0}^{n-1} \left([\mathbb{E}\Lambda_{\vartheta_0}(t^0)]^2 + [\mathbb{E}\Lambda_{\vartheta}(t^0)]^2 \right) \left(N^{(3)}(s_{n,i+1}) - N^{(3)}(s_{n,i}) \right) \\
&= \left([\mathbb{E}\Lambda_{\vartheta_0}(t^0)]^2 + [\mathbb{E}\Lambda_{\vartheta}(t^0)]^2 \right) \sum_{i=0}^{n-1} \left(N^{(3)}(s_{n,i+1}) - N^{(3)}(s_{n,i}) \right) \\
&= \left([\mathbb{E}\Lambda_{\vartheta_0}(t^0)]^2 + [\mathbb{E}\Lambda_{\vartheta}(t^0)]^2 \right) N^{(3)}(t^0) \\
&\leq 2 \sup_{\vartheta \in \Theta} \{ \mathbb{E}\Lambda_{\vartheta}(t^0)^2 \} N^{(3)}(t^0),
\end{aligned} \tag{3.2.15}$$

da der Erwartungswert der rechten Seite der Abschätzung aus (3.2.15) wegen - man verwende die Monotonie des Integrals -

$$\sup_{\vartheta \in \Theta} \{ \mathbb{E}\Lambda_{\vartheta}(t^0)^2 \} \leq \mathbb{E} \left[\sup_{\vartheta \in \Theta} \{ \Lambda_{\vartheta}(t^0)^2 \} \right]$$

und somit nach Bedingung (3.1.12) existiert.

Insgesamt gilt folglich Gleichung (3.2.12) und damit Lemma 3.2.3. \square

Bemerkung 3.2.4.

Die fast sichere Stetigkeit eines Prozesses impliziert noch *nicht* die Stetigkeit des assoziierten Erwartungswertes, obgleich die Konstruktion eines Gegenbeispiels nicht trivial ist.

Die im Beweis von Lemma 3.2.3 verwendete Stetigkeit des Prozesses $\mathbb{E}\Lambda_{\vartheta}$ für ein beliebiges $\vartheta \in \overline{\Theta}$ folgt dennoch aus der fast sicheren Stetigkeit des Prozesses Λ_{ϑ} , indem die Bedingung aus (3.1.12) hinzugezogen wird. Wegen $\Lambda_{\vartheta}(t) \leq \Lambda_{\vartheta}(t^0)$ für alle $t \in I$ liefert dann nämlich

bereits die in Stute und Kopperschmidt 2013 gestellte Forderung $E \Lambda_\vartheta(t^0) < \infty$ für alle $\vartheta \in \bar{\Theta}$ die Existenz einer integrierbaren Majorante, so dass via majorisierter Konvergenz die Stetigkeit gezeigt werden kann:

$$\lim_{\substack{t \rightarrow t' \\ t \in I}} E \Lambda_\vartheta(t) = E \left[\lim_{\substack{t \rightarrow t' \\ t \in I}} \Lambda_\vartheta(t) \right] = E \Lambda_\vartheta(t') \quad \text{für alle } t' \in I \text{ und alle } \vartheta \in \bar{\Theta} \quad (3.2.16)$$

3.3 Konsistenz des Schätzers ϑ_n

In diesem Abschnitt werden wir die Konvergenz des Minimum-Distanz-Schätzers ϑ_n zeigen, die letztlich aufgrund von Abschnitt 3.2 als Folgerung aus dem starken Gesetz der großen Zahlen für U-Statistiken erscheint.

Wir wenden uns zunächst dem bereits im Beweis von Lemma 3.2.3 untersuchten Prozess

$$\vartheta \mapsto \int_I \left(N^{(1)} - \Lambda_\vartheta^{(1)} \right) \left(N^{(2)} - \Lambda_\vartheta^{(2)} \right) d\mu_{N^{(3)}} = \langle N^{(1)} - \Lambda_\vartheta^{(1)}, N^{(2)} - \Lambda_\vartheta^{(2)} \rangle_{\mu_{N^{(3)}}} \quad (3.3.1)$$

zu. Im folgenden Lemma untersuchen wir lokale Fluktuationen dieses Prozesses.

Generalvoraussetzung in diesem Abschnitt sei die Situation aus Abschnitt 3.1 sowie die Stetigkeitsforderung aus (3.2.3).

Lemma 3.3.1. (Siehe Stute und Kopperschmidt 2013, S. 1286)

Es seien $\varepsilon > 0$ und $\delta > 0$ gegeben. Dann existiert für jedes $\vartheta \in \bar{\Theta} \setminus B_\varepsilon(\vartheta_0)$ ein $r > 0$, so dass

$$\left| E \left[\sup_{\vartheta' \in B_r^\varepsilon(\vartheta)} \langle N^{(1)} - \Lambda_{\vartheta'}^{(1)}, N^{(2)} - \Lambda_{\vartheta'}^{(2)} \rangle_{\mu_{N^{(3)}}} \right] - \inf_{\vartheta' \in B_r^\varepsilon(\vartheta)} \|E\Lambda_{\vartheta_0} - E\Lambda_{\vartheta'}\|_{E\Lambda_{\vartheta_0}}^2 \right| \leq \delta, \quad (3.3.2)$$

$$\left| E \left[\sup_{\vartheta' \in B_r^\varepsilon(\vartheta)} \langle N^{(1)} - \Lambda_{\vartheta'}^{(1)}, N^{(2)} - \Lambda_{\vartheta'}^{(2)} \rangle_{\mu_{N^{(3)}}} \right] - \|E\Lambda_{\vartheta_0} - E\Lambda_\vartheta\|_{E\Lambda_{\vartheta_0}}^2 \right| \leq \delta, \quad (3.3.3)$$

sowie

$$\left| E \left[\inf_{\vartheta' \in B_r^\varepsilon(\vartheta)} \langle N^{(1)} - \Lambda_{\vartheta'}^{(1)}, N^{(2)} - \Lambda_{\vartheta'}^{(2)} \rangle_{\mu_{N^{(3)}}} \right] - \inf_{\vartheta' \in B_r^\varepsilon(\vartheta)} \|E\Lambda_{\vartheta_0} - E\Lambda_{\vartheta'}\|_{E\Lambda_{\vartheta_0}}^2 \right| \leq \delta, \quad (3.3.4)$$

$$\left| E \left[\inf_{\vartheta' \in B_r^\varepsilon(\vartheta)} \langle N^{(1)} - \Lambda_{\vartheta'}^{(1)}, N^{(2)} - \Lambda_{\vartheta'}^{(2)} \rangle_{\mu_{N^{(3)}}} \right] - \|E\Lambda_{\vartheta_0} - E\Lambda_\vartheta\|_{E\Lambda_{\vartheta_0}}^2 \right| \leq \delta, \quad (3.3.5)$$

gilt.

Beweis⁴⁵:

Die fast sichere Stetigkeit des Prozesses Λ_ϑ für alle $\vartheta \in \bar{\Theta}$ impliziert mittels majorisierter

⁴⁵Der Beweis ist skizzenhaft in Stute und Kopperschmidt 2013 zu finden und wurde vom Autor detailliert ausgearbeitet.

Konvergenz zugleich die Stetigkeit des Prozesses aus Gleichung (3.3.1). Für $r \downarrow 0$ gilt folglich:

$$\sup_{\vartheta' \in B_r^\varepsilon(\vartheta)} \langle N^{(1)} - \Lambda_{\vartheta'}^{(1)}, N^{(2)} - \Lambda_{\vartheta'}^{(2)} \rangle_{\mu_{N^{(3)}}} \downarrow \langle N^{(1)} - \Lambda_{\vartheta}^{(1)}, N^{(2)} - \Lambda_{\vartheta}^{(2)} \rangle_{\mu_{N^{(3)}}}, \quad (3.3.6)$$

$$\inf_{\vartheta' \in B_r^\varepsilon(\vartheta)} \langle N^{(1)} - \Lambda_{\vartheta'}^{(1)}, N^{(2)} - \Lambda_{\vartheta'}^{(2)} \rangle_{\mu_{N^{(3)}}} \uparrow \langle N^{(1)} - \Lambda_{\vartheta}^{(1)}, N^{(2)} - \Lambda_{\vartheta}^{(2)} \rangle_{\mu_{N^{(3)}}}. \quad (3.3.7)$$

Der Erwartungswert der linken Seite aus Gleichung (3.3.6) bzw. (3.3.7) existiert bereits nach dem Beweis von Lemma 3.2.1 und seinem Zusatz. Die monotone Konvergenz liefert sodann mit Lemma 3.2.3:

$$\begin{aligned} & \lim_{r \downarrow 0} \mathbb{E} \left[\sup_{\vartheta' \in B_r^\varepsilon(\vartheta)} \langle N^{(1)} - \Lambda_{\vartheta'}^{(1)}, N^{(2)} - \Lambda_{\vartheta'}^{(2)} \rangle_{\mu_{N^{(3)}}} \right] \\ &= \mathbb{E} \left[\lim_{r \downarrow 0} \sup_{\vartheta' \in B_r^\varepsilon(\vartheta)} \langle N^{(1)} - \Lambda_{\vartheta'}^{(1)}, N^{(2)} - \Lambda_{\vartheta'}^{(2)} \rangle_{\mu_{N^{(3)}}} \right] \\ &= \mathbb{E} \left[\langle N^{(1)} - \Lambda_{\vartheta}^{(1)}, N^{(2)} - \Lambda_{\vartheta}^{(2)} \rangle_{\mu_{N^{(3)}}} \right] = \|\mathbb{E}\Lambda_{\vartheta_0} - \mathbb{E}\Lambda_{\vartheta}\|_{\mathbb{E}\Lambda_{\vartheta_0}}^2, \end{aligned}$$

und völlig analog

$$\lim_{r \downarrow 0} \mathbb{E} \left[\inf_{\vartheta' \in B_r^\varepsilon(\vartheta)} \langle N^{(1)} - \Lambda_{\vartheta'}^{(1)}, N^{(2)} - \Lambda_{\vartheta'}^{(2)} \rangle_{\mu_{N^{(3)}}} \right] = \|\mathbb{E}\Lambda_{\vartheta_0} - \mathbb{E}\Lambda_{\vartheta}\|_{\mathbb{E}\Lambda_{\vartheta_0}}^2.$$

Zu $\delta > 0$ existiert daher ein $r_1 > 0$, so dass für alle $0 < r < r_1$ simultan gilt:

$$\begin{aligned} & \left| \mathbb{E} \left[\sup_{\vartheta' \in B_r^\varepsilon(\vartheta)} \langle N^{(1)} - \Lambda_{\vartheta'}^{(1)}, N^{(2)} - \Lambda_{\vartheta'}^{(2)} \rangle_{\mu_{N^{(3)}}} \right] - \|\mathbb{E}\Lambda_{\vartheta_0} - \mathbb{E}\Lambda_{\vartheta}\|_{\mathbb{E}\Lambda_{\vartheta_0}}^2 \right| \leq \delta, \\ & \left| \mathbb{E} \left[\inf_{\vartheta' \in B_r^\varepsilon(\vartheta)} \langle N^{(1)} - \Lambda_{\vartheta'}^{(1)}, N^{(2)} - \Lambda_{\vartheta'}^{(2)} \rangle_{\mu_{N^{(3)}}} \right] - \|\mathbb{E}\Lambda_{\vartheta_0} - \mathbb{E}\Lambda_{\vartheta}\|_{\mathbb{E}\Lambda_{\vartheta_0}}^2 \right| \leq \delta. \end{aligned}$$

Dies zeigt die Gleichungen (3.3.3) und (3.3.5).

Wiederum aus der fast sicheren Stetigkeit des Prozesses Λ_{ϑ} folgt die Stetigkeit des parametrischen Prozesses

$$\vartheta \mapsto \|\mathbb{E}\Lambda_{\vartheta_0} - \mathbb{E}\Lambda_{\vartheta}\|_{\mathbb{E}\Lambda_{\vartheta_0}}^2.$$

Man betrachte dafür den nachfolgenden, zu Gleichung (3.2.16) weitgehend analogen Stetigkeitsbeweis.

Zunächst gilt - wie in Gleichung (3.2.15) aus dem Beweis von Lemma 3.2.3 - die Abschät-

zung

$$(\mathbb{E}\Lambda_{\vartheta_0} - \mathbb{E}\Lambda_{\vartheta})^2(t) \leq 2 \sup_{\vartheta \in \Theta} \{\mathbb{E}\Lambda_{\vartheta}(t^0)^2\} \leq 2 \mathbb{E} \left[\sup_{\vartheta \in \Theta} \{\Lambda_{\vartheta}(t^0)^2\} \right], \quad \text{für alle } t \in I, \vartheta \in \bar{\Theta}. \quad (3.3.8)$$

Die rechte Seite aus Gleichung (3.3.8) ist somit die gesuchte konvergente Majorante, denn es gilt

$$\begin{aligned} \|\mathbb{E}\Lambda_{\vartheta_0} - \mathbb{E}\Lambda_{\vartheta}\|_{\mathbb{E}\Lambda_{\vartheta_0}}^2 &= \int_I (\mathbb{E}\Lambda_{\vartheta_0} - \mathbb{E}\Lambda_{\vartheta})^2(t) \, d\mathbb{E}\Lambda_{\vartheta_0} \\ &\leq 2 \mathbb{E} \left[\sup_{\vartheta \in \Theta} \{\Lambda_{\vartheta}(t^0)^2\} \right] \int_I d\mathbb{E}\Lambda_{\vartheta_0} \\ &\leq 2 \mathbb{E} \left[\sup_{\vartheta \in \Theta} \{\Lambda_{\vartheta}(t^0)^2\} \right] \mathbb{E}\Lambda_{\vartheta_0}(t^0) \\ &\leq 2 \mathbb{E} \left[\sup_{\vartheta \in \Theta} \{\Lambda_{\vartheta}(t^0)^2\} \right] \mathbb{E} \left[\sup_{\vartheta \in \Theta} \{\Lambda_{\vartheta}(t^0)\} \right], \quad \text{für alle } \vartheta \in \bar{\Theta}. \end{aligned}$$

Wie in Gleichung (3.2.16) folgt damit die Stetigkeit des parametrischen Prozesses.

Für $r \downarrow 0$ liefert dies die Konvergenz

$$\inf_{\vartheta' \in \mathbb{B}_r^c(\vartheta)} \|\mathbb{E}\Lambda_{\vartheta_0} - \mathbb{E}\Lambda_{\vartheta'}\|_{\mathbb{E}\Lambda_{\vartheta_0}}^2 \uparrow \|\mathbb{E}\Lambda_{\vartheta_0} - \mathbb{E}\Lambda_{\vartheta}\|_{\mathbb{E}\Lambda_{\vartheta_0}}^2. \quad (3.3.9)$$

Folglich existiert ein $r_2 > 0$, so dass für alle $0 < r < r_2$ gleichermaßen gilt:

$$\left| \mathbb{E} \left[\sup_{\vartheta' \in \mathbb{B}_r^c(\vartheta)} \langle N^{(1)} - \Lambda_{\vartheta'}^{(1)}, N^{(2)} - \Lambda_{\vartheta'}^{(2)} \rangle_{\mu_{N^{(3)}}} \right] - \|\mathbb{E}\Lambda_{\vartheta_0} - \mathbb{E}\Lambda_{\vartheta}\|_{\mathbb{E}\Lambda_{\vartheta_0}}^2 \right| \leq \frac{\delta}{2}, \quad (3.3.10)$$

$$\left| \mathbb{E} \left[\inf_{\vartheta' \in \mathbb{B}_r^c(\vartheta)} \langle N^{(1)} - \Lambda_{\vartheta'}^{(1)}, N^{(2)} - \Lambda_{\vartheta'}^{(2)} \rangle_{\mu_{N^{(3)}}} \right] - \|\mathbb{E}\Lambda_{\vartheta_0} - \mathbb{E}\Lambda_{\vartheta}\|_{\mathbb{E}\Lambda_{\vartheta_0}}^2 \right| \leq \frac{\delta}{2} \quad (3.3.11)$$

$$\text{und} \quad \left| \|\mathbb{E}\Lambda_{\vartheta_0} - \mathbb{E}\Lambda_{\vartheta}\|_{\mathbb{E}\Lambda_{\vartheta_0}}^2 - \inf_{\vartheta' \in \mathbb{B}_r^c(\vartheta)} \|\mathbb{E}\Lambda_{\vartheta_0} - \mathbb{E}\Lambda_{\vartheta'}\|_{\mathbb{E}\Lambda_{\vartheta_0}}^2 \right| \leq \frac{\delta}{2} \quad (3.3.12)$$

Abschließend liefert die Kombination der Gleichungen (3.3.10) und (3.3.12) via Dreiecksungleichung die Gültigkeit von Gleichung (3.3.2). Analog folgert man aus den Gleichungen (3.3.11) und (3.3.12) die Gültigkeit von Gleichung (3.3.4). Übergang zum Minimum von r_1 und r_2 vollendet dann den Beweis des Lemmas. \square

Die nächsten Lemmata bereiten den Weg für den Beweis des zentralen Theorems dieses Abschnittes, welches die starke Konsistenz des Schätzers ϑ_n zum Inhalt hat.

Lemma 3.3.2. (Siehe Stute und Kopperschmidt 2013, S. 1286)

Für jedes $\vartheta \in \Theta$ gilt für $n \rightarrow \infty$ fast sicher:

$$\|\overline{N}^{(n)} - \overline{\Lambda}_\vartheta^{(n)}\|_{\mu_{\overline{N}^{(n)}}}^2 \rightarrow \|E \Lambda_{\vartheta_0} - E \Lambda_\vartheta\|_{E \Lambda_{\vartheta_0}}^2 \quad (3.3.13)$$

Für $\vartheta = \vartheta_0$ ist der Grenzwert insbesondere gleich Null.

Beweis⁴⁶:

Per Definition der aggregierten Prozesse - siehe Gleichung (3.1.9) - ergibt sich für die linke Seite aus Gleichung (3.3.13):

$$\begin{aligned} \|\overline{N}^{(n)} - \overline{\Lambda}_\vartheta^{(n)}\|_{\mu_{\overline{N}^{(n)}}}^2 &= \int_I \left[\overline{N}^{(n)} - \overline{\Lambda}_\vartheta^{(n)} \right]^2 d\mu_{\overline{N}^{(n)}} \\ &= \frac{1}{n} \sum_{k=1}^n \int_I \left[\frac{1}{n} \sum_{i=1}^n N^{(i)} - \frac{1}{n} \sum_{i=1}^n \Lambda_\vartheta^{(i)} \right]^2 d\mu_{N^{(k)}} \\ &= \frac{1}{n^3} \sum_{k=1}^n \int_I \left[\sum_{i=1}^n (N^{(i)} - \Lambda_\vartheta^{(i)}) \right]^2 d\mu_{N^{(k)}} \\ &= \frac{1}{n^3} \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n \int_I (N^{(i)} - \Lambda_\vartheta^{(i)}) (N^{(j)} - \Lambda_\vartheta^{(j)}) d\mu_{N^{(k)}} \\ &= \frac{1}{n^3} \sum_{i,j,k=1}^n \langle N^{(i)} - \Lambda_\vartheta^{(i)}, N^{(j)} - \Lambda_\vartheta^{(j)} \rangle_{\mu_{N^{(k)}}}. \end{aligned} \quad (3.3.14)$$

Für die Berechnung des Grenzwertes kann nun die Summation in Gleichung (3.3.14) ohne weitere Einschränkung auf die Summanden restringiert werden, die paarweise verschieden sind. Dazu beachte man:

$$\begin{aligned} \frac{1}{n^3} \sum_{i,j,k=1}^n \langle N^{(i)} - \Lambda_\vartheta^{(i)}, N^{(j)} - \Lambda_\vartheta^{(j)} \rangle_{\mu_{N^{(k)}}} &= \frac{1}{n^3} \sum_{\substack{i,j,k=1 \\ i \neq j \neq k \neq i}} \langle N^{(i)} - \Lambda_\vartheta^{(i)}, N^{(j)} - \Lambda_\vartheta^{(j)} \rangle_{\mu_{N^{(k)}}} \\ &+ \frac{1}{n^3} \sum_{\substack{i,j,k=1 \\ i=j \neq k}} \langle N^{(i)} - \Lambda_\vartheta^{(i)}, N^{(j)} - \Lambda_\vartheta^{(j)} \rangle_{\mu_{N^{(k)}}} + \frac{1}{n^3} \sum_{\substack{i,j,k=1 \\ i=k \neq j}} \langle N^{(i)} - \Lambda_\vartheta^{(i)}, N^{(j)} - \Lambda_\vartheta^{(j)} \rangle_{\mu_{N^{(k)}}} \\ &+ \frac{1}{n^3} \sum_{\substack{i,j,k=1 \\ j=k \neq i}} \langle N^{(i)} - \Lambda_\vartheta^{(i)}, N^{(j)} - \Lambda_\vartheta^{(j)} \rangle_{\mu_{N^{(k)}}} + \frac{1}{n^3} \sum_{\substack{i,j,k=1 \\ i=j=k}} \langle N^{(i)} - \Lambda_\vartheta^{(i)}, N^{(j)} - \Lambda_\vartheta^{(j)} \rangle_{\mu_{N^{(k)}}} \end{aligned}$$

⁴⁶Die wesentliche Beweisidee - die Verwendung von Korollar 3.2.2 - geht zurück auf Kopperschmidt und Stute.

$$\begin{aligned}
&= \frac{1}{n^3} \sum_{\substack{i,j,k=1 \\ i \neq j \neq k \neq i}}^n \langle N^{(i)} - \Lambda_{\vartheta}^{(i)}, N^{(j)} - \Lambda_{\vartheta}^{(j)} \rangle_{\mu_{N^{(k)}}} + \frac{1}{n^3} \sum_{\substack{i,k=1 \\ i \neq k}}^n \|N^{(i)} - \Lambda_{\vartheta}^{(i)}\|_{\mu_{N^{(k)}}}^2 \\
&\quad + \frac{2}{n^3} \sum_{\substack{i,j=1 \\ i \neq j}}^n \langle N^{(i)} - \Lambda_{\vartheta}^{(i)}, N^{(j)} - \Lambda_{\vartheta}^{(j)} \rangle_{\mu_{N^{(i)}}} + \frac{1}{n^3} \sum_{i=1}^n \|N^{(i)} - \Lambda_{\vartheta}^{(i)}\|_{\mu_{N^{(i)}}}^2. \tag{3.3.15}
\end{aligned}$$

Gemäß Korollar 3.2.2 können wir die Konvergenzordnung der einzelnen Summanden aus Gleichung (3.3.15) bestimmen und erhalten infolgedessen:

$$\begin{aligned}
&\frac{1}{n^3} \sum_{i,j,k=1}^n \langle N^{(i)} - \Lambda_{\vartheta}^{(i)}, N^{(j)} - \Lambda_{\vartheta}^{(j)} \rangle_{\mu_{N^{(k)}}} \\
&= \frac{1}{n^3} \sum_{\substack{i,j,k=1 \\ i \neq j \neq k \neq i}}^n \langle N^{(i)} - \Lambda_{\vartheta}^{(i)}, N^{(j)} - \Lambda_{\vartheta}^{(j)} \rangle_{\mu_{N^{(k)}}} + \frac{3}{n^3} \mathcal{O}(n^2) + \frac{1}{n^3} \mathcal{O}(n). \tag{3.3.16}
\end{aligned}$$

Asymptotisch sind gemäß Gleichung (3.3.16) also sämtliche Terme, in denen mindestens zwei der drei Indizes übereinstimmen, vernachlässigbar.

Nach Gleichung (3.2.11) konvergiert schließlich der verbleibende Summand gegen den in Lemma 3.2.3 berechneten Grenzwert, das heißt:

$$\frac{1}{n^3} \sum_{\substack{i,j,k=1 \\ i \neq j \neq k \neq i}}^n \langle N^{(i)} - \Lambda_{\vartheta}^{(i)}, N^{(j)} - \Lambda_{\vartheta}^{(j)} \rangle_{\mu_{N^{(k)}}} \rightarrow \|E\Lambda_{\vartheta_0} - E\Lambda_{\vartheta}\|_{E\Lambda_{\vartheta_0}}^2 \quad (n \rightarrow \infty).$$

Aus Gleichung (3.3.14) folgt sodann die Behauptung. □

Lemma 3.3.3. (Siehe Stute und Kopperschmidt 2013, S. 1287)

Für jedes $\varepsilon > 0$ gilt für $n \rightarrow \infty$ fast sicher:

$$\inf_{\vartheta: \|\vartheta - \vartheta_0\| \geq \varepsilon} \|\overline{N}^{(n)} - \overline{\Lambda}_{\vartheta}^{(n)}\|_{\mu_{\overline{N}^{(n)}}}^2 \rightarrow \inf_{\vartheta: \|\vartheta - \vartheta_0\| \geq \varepsilon} \|E\Lambda_{\vartheta_0} - E\Lambda_{\vartheta}\|_{E\Lambda_{\vartheta_0}}^2, \tag{3.3.17}$$

wobei das Infimum nur über jene ϑ aus $\overline{\Theta}$ gebildet werde.

Beweis⁴⁷:

Es sei $\varepsilon > 0$ gegeben. Ferner sei $\delta > 0$ beliebig. Nach Lemma 3.3.1 existiert dann zu jedem $\vartheta \in \overline{\Theta} \setminus B_{\varepsilon}(\vartheta_0)$ ein $r = r(\vartheta) > 0$, so dass die Gleichungen (3.3.2) und (3.3.4) gleichermaßen erfüllt sind. Damit gilt:

$$\bigcup_{\vartheta \in \overline{\Theta} \setminus B_{\varepsilon}(\vartheta_0)} B_{r(\vartheta)}^{\varepsilon}(\vartheta) = \overline{\Theta} \setminus B_{\varepsilon}(\vartheta_0) \tag{3.3.18}$$

⁴⁷Der Beweis findet sich in Stute und Kopperschmidt 2013 und wurde vom Autor dieser Arbeit um hilfreiche Erläuterungen ergänzt.

Bezüglich der Teilraumtopologie des \mathbb{R}^d auf $\overline{\Theta} \setminus B_\varepsilon(\vartheta_0)$ stellt die linke Seite von Gleichung (3.3.18) eine offene Überdeckung dieses Kompaktums dar. Folglich existieren endlich viele $\vartheta_1, \dots, \vartheta_q$, so dass bereits gilt:

$$\bigcup_{p=1}^q B_{r(\vartheta_p)}^\varepsilon(\vartheta_p) = \overline{\Theta} \setminus B_\varepsilon(\vartheta_0) \quad (3.3.19)$$

Aus Gleichung (3.3.19) folgern wir nun mit $r_p = r(\vartheta_p)$:

$$\begin{aligned} & \left| \inf_{\vartheta: \|\vartheta - \vartheta_0\| \geq \varepsilon} \|\overline{N}^{(n)} - \overline{\Lambda}_\vartheta^{(n)}\|_{\mu_{\overline{N}^{(n)}}}^2 - \inf_{\vartheta: \|\vartheta - \vartheta_0\| \geq \varepsilon} \|\mathbb{E} \Lambda_{\vartheta_0} - \mathbb{E} \Lambda_\vartheta\|_{\mathbb{E} \Lambda_{\vartheta_0}}^2 \right| \\ &= \left| \min_{1 \leq p \leq q} \inf_{\vartheta \in B_{r_p}^\varepsilon(\vartheta_p)} \|\overline{N}^{(n)} - \overline{\Lambda}_\vartheta^{(n)}\|_{\mu_{\overline{N}^{(n)}}}^2 - \min_{1 \leq p \leq q} \inf_{\vartheta \in B_{r_p}^\varepsilon(\vartheta_p)} \|\mathbb{E} \Lambda_{\vartheta_0} - \mathbb{E} \Lambda_\vartheta\|_{\mathbb{E} \Lambda_{\vartheta_0}}^2 \right| \\ &\leq \max_{1 \leq p \leq q} \left| \inf_{\vartheta \in B_{r_p}^\varepsilon(\vartheta_p)} \|\overline{N}^{(n)} - \overline{\Lambda}_\vartheta^{(n)}\|_{\mu_{\overline{N}^{(n)}}}^2 - \inf_{\vartheta \in B_{r_p}^\varepsilon(\vartheta_p)} \|\mathbb{E} \Lambda_{\vartheta_0} - \mathbb{E} \Lambda_\vartheta\|_{\mathbb{E} \Lambda_{\vartheta_0}}^2 \right| \quad (3.3.20) \end{aligned}$$

Die letzte Abschätzung in Gleichung (3.3.20) kann wie folgt begründet werden:

Es seien a_1, \dots, a_q sowie b_1, \dots, b_q positive reelle Zahlen. Dann gilt:

$$\left| \min_{1 \leq p \leq q} a_p - \min_{1 \leq p \leq q} b_p \right| \leq \max_{1 \leq p \leq q} |a_p - b_p| \quad (3.3.21)$$

Aus der nachfolgend bewiesenen Gültigkeit von Gleichung (3.3.21) folgt dann unmittelbar auch die Gültigkeit der Abschätzung aus Gleichung (3.3.20).

Es seien also $p_1, p_2 \in \{1, \dots, q\}$ dergestalt, dass

$$a_{p_1} = \min_{1 \leq p \leq q} a_p \quad \text{und} \quad b_{p_2} = \min_{1 \leq p \leq q} b_p \quad (3.3.22)$$

gilt. Wir unterscheiden zwei Fälle.

- Es gelte $a_{p_1} \geq b_{p_2}$. Dann folgt nach der Wahl von p_1 in Gleichung (3.3.22):

$$\begin{aligned} \left| \min_{1 \leq p \leq q} a_p - \min_{1 \leq p \leq q} b_p \right| &= |a_{p_1} - b_{p_2}| = a_{p_1} - b_{p_2} \leq a_{p_2} - b_{p_2} \\ &= |a_{p_2} - b_{p_2}| \leq \max_{1 \leq p \leq q} |a_p - b_p| \quad (3.3.23) \end{aligned}$$

- Es gelte $a_{p_1} < b_{p_2}$. Dann folgt nach der Wahl von p_2 in Gleichung (3.3.22):

$$\begin{aligned} \left| \min_{1 \leq p \leq q} a_p - \min_{1 \leq p \leq q} b_p \right| &= |a_{p_1} - b_{p_2}| = b_{p_2} - a_{p_1} \leq b_{p_1} - a_{p_1} \\ &= |a_{p_1} - b_{p_1}| \leq \max_{1 \leq p \leq q} |a_p - b_p| \end{aligned} \quad (3.3.24)$$

Aus den Gleichungen (3.3.23) und (3.3.24) folgt dann sofort Gleichung (3.3.21).

Um den Beweis zu vollenden, genügt es nun zu zeigen, dass für alle $1 \leq p \leq q$ fast sicher gilt:

$$\limsup_{n \rightarrow \infty} \left| \inf_{\vartheta \in B_{r_p}^\varepsilon(\vartheta_p)} \|\overline{N}^{(n)} - \overline{\Lambda}_\vartheta^{(n)}\|_{\mu_{\overline{N}^{(n)}}}^2 - \inf_{\vartheta \in B_{r_p}^\varepsilon(\vartheta_p)} \|E \Lambda_{\vartheta_0} - E \Lambda_\vartheta\|_{E \Lambda_{\vartheta_0}}^2 \right| \leq \delta \quad (3.3.25)$$

Unter erneuter Verwendung von Gleichung (3.3.14) aus dem Beweis von Lemma 3.3.2 erhalten wir zunächst die Abschätzungen

$$\begin{aligned} \inf_{\vartheta \in B_{r_p}^\varepsilon(\vartheta_p)} \|\overline{N}^{(n)} - \overline{\Lambda}_\vartheta^{(n)}\|_{\mu_{\overline{N}^{(n)}}}^2 &= \inf_{\vartheta \in B_{r_p}^\varepsilon(\vartheta_p)} \frac{1}{n^3} \sum_{i,j,k=1}^n \langle N^{(i)} - \Lambda_\vartheta^{(i)}, N^{(j)} - \Lambda_\vartheta^{(j)} \rangle_{\mu_{N^{(k)}}} \\ &\leq \frac{1}{n^3} \sum_{i,j,k=1}^n \sup_{\vartheta \in B_{r_p}^\varepsilon(\vartheta_p)} \langle N^{(i)} - \Lambda_\vartheta^{(i)}, N^{(j)} - \Lambda_\vartheta^{(j)} \rangle_{\mu_{N^{(k)}}} \cdot \end{aligned} \quad (3.3.26)$$

$$\inf_{\vartheta \in B_{r_p}^\varepsilon(\vartheta_p)} \|\overline{N}^{(n)} - \overline{\Lambda}_\vartheta^{(n)}\|_{\mu_{\overline{N}^{(n)}}}^2 \geq \frac{1}{n^3} \sum_{i,j,k=1}^n \inf_{\vartheta \in B_{r_p}^\varepsilon(\vartheta_p)} \langle N^{(i)} - \Lambda_\vartheta^{(i)}, N^{(j)} - \Lambda_\vartheta^{(j)} \rangle_{\mu_{N^{(k)}}}. \quad (3.3.27)$$

Gemäß Lemma 3.2.1 konvergieren die rechten Seiten der Gleichungen (3.3.26) und (3.3.27) gegen

$$E \left[\sup_{\vartheta \in B_{r_p}^\varepsilon(\vartheta_p)} \langle N^{(1)} - \Lambda_\vartheta^{(1)}, N^{(2)} - \Lambda_\vartheta^{(2)} \rangle_{\mu_{N^{(3)}}} \right] \quad (3.3.28)$$

beziehungsweise

$$E \left[\inf_{\vartheta \in B_{r_p}^\varepsilon(\vartheta_p)} \langle N^{(1)} - \Lambda_\vartheta^{(1)}, N^{(2)} - \Lambda_\vartheta^{(2)} \rangle_{\mu_{N^{(3)}}} \right]. \quad (3.3.29)$$

Folglich erhalten wir durch Kombination der Gleichungen (3.3.26) bis (3.3.29):

$$\begin{aligned}
& \mathbb{E} \left[\inf_{\vartheta \in \mathbb{B}_{r_p}^\varepsilon(\vartheta_p)} \langle N^{(1)} - \Lambda_\vartheta^{(1)}, N^{(2)} - \Lambda_\vartheta^{(2)} \rangle_{\mu_{N^{(3)}}} \right] \\
& \leq \limsup_{n \rightarrow \infty} \inf_{\vartheta \in \mathbb{B}_{r_p}^\varepsilon(\vartheta_p)} \|\overline{N}^{(n)} - \overline{\Lambda}_\vartheta^{(n)}\|_{\mu_{\overline{N}^{(n)}}}^2 \\
& \leq \mathbb{E} \left[\sup_{\vartheta \in \mathbb{B}_{r_p}^\varepsilon(\vartheta_p)} \langle N^{(1)} - \Lambda_\vartheta^{(1)}, N^{(2)} - \Lambda_\vartheta^{(2)} \rangle_{\mu_{N^{(3)}}} \right]. \tag{3.3.30}
\end{aligned}$$

Die Wahl der Radien r_p für $1 \leq p \leq q$ liefert sodann (vergleiche Lemma 3.3.1, insbesondere Gleichungen (3.3.2) und (3.3.4)) aufgrund der Konvexität des Intervalls

$$\mathcal{K}_\delta = \left(\inf_{\vartheta \in \mathbb{B}_{r_p}^\varepsilon(\vartheta_p)} \|\mathbb{E} \Lambda_{\vartheta_0} - \mathbb{E} \Lambda_\vartheta\|_{\mathbb{E} \Lambda_{\vartheta_0}}^2 - \delta, \inf_{\vartheta \in \mathbb{B}_{r_p}^\varepsilon(\vartheta_p)} \|\mathbb{E} \Lambda_{\vartheta_0} - \mathbb{E} \Lambda_\vartheta\|_{\mathbb{E} \Lambda_{\vartheta_0}}^2 + \delta \right)$$

und wegen

$$\begin{aligned}
& \mathbb{E} \left[\sup_{\vartheta \in \mathbb{B}_{r_p}^\varepsilon(\vartheta_p)} \langle N^{(1)} - \Lambda_\vartheta^{(1)}, N^{(2)} - \Lambda_\vartheta^{(2)} \rangle_{\mu_{N^{(3)}}} \right] \in \mathcal{K}_\delta, \\
& \mathbb{E} \left[\inf_{\vartheta \in \mathbb{B}_{r_p}^\varepsilon(\vartheta_p)} \langle N^{(1)} - \Lambda_\vartheta^{(1)}, N^{(2)} - \Lambda_\vartheta^{(2)} \rangle_{\mu_{N^{(3)}}} \right] \in \mathcal{K}_\delta,
\end{aligned}$$

schließlich auch

$$\limsup_{n \rightarrow \infty} \inf_{\vartheta \in \mathbb{B}_{r_p}^\varepsilon(\vartheta_p)} \|\overline{N}^{(n)} - \overline{\Lambda}_\vartheta^{(n)}\|_{\mu_{\overline{N}^{(n)}}}^2 \in \mathcal{K}_\delta$$

und damit die Behauptung, da $\delta > 0$ beliebig gewählt war. \square

Wir utilisieren nun die Lemmata 3.3.2 und 3.3.3, um das zentrale Resultat dieses Abschnittes zu beweisen, welches wir im nächsten Satz formulieren.

Wir erinnern abschließend noch einmal daran, dass die Simplität der hiesigen Beweise aus der Anwendbarkeit des starken Gesetzes der großen Zahlen für U-Statistiken folgt.

Dies rechtfertigt rückwirkend die ausführliche Diskussion dieses Resultates in Abschnitt 2.2.

Satz 3.3.4. (*Starke Konsistenz des Minimum-Distanz-Schätzers* ϑ_n , siehe Stute und Kopperschmidt 2013, S. 1279)

In der Situation aus Abschnitt 3.1 gelte zusätzlich die *Stetigkeitsbedingung* aus Gleichung (3.2.3) sowie die *Identifizierbarkeitsbedingung*

$$\forall \varepsilon > 0 : \quad \inf_{\vartheta: \|\vartheta - \vartheta_0\| \geq \varepsilon} \|\mathbb{E} \Lambda_{\vartheta_0} - \mathbb{E} \Lambda_{\vartheta}\|_{\mathbb{E} \Lambda_{\vartheta_0}} > 0. \quad (3.3.31)$$

Dann gilt fast sicher:

$$\lim_{n \rightarrow \infty} \vartheta_n = \vartheta_0. \quad (3.3.32)$$

Bemerkung 3.3.5.

Wie Kopperschmidt und Stute bemerken, impliziert die Stetigkeitsbedingung aus Gleichung (3.2.3), dass das Infimum in Gleichung (3.3.31) angenommen wird. Folglich ist die *Identifizierbarkeitsbedingung* in diesem Falle äquivalent zu der Bedingung

$$\mathbb{E} \Lambda_{\vartheta_0}(t) \neq \mathbb{E} \Lambda_{\vartheta}(t) \quad \text{für alle } \vartheta \neq \vartheta_0 \quad (3.3.33)$$

und $t \in I_{\vartheta}$, wobei $I_{\vartheta} \subset I$ eine Menge mit $\mu_{\mathbb{E} \Lambda_{\vartheta_0}}(I_{\vartheta}) > 0$ sei. Dabei bezeichne $\mu_{\mathbb{E} \Lambda_{\vartheta_0}}$ das Maß mit kumulativer Funktion $\mathbb{E} \Lambda_{\vartheta_0}$ (siehe Stute und Kopperschmidt 2013, S. 1280).

Heuristisch gesprochen bedeutet die Bedingung aus Gleichung (3.3.31) ergo, dass sich die Funktion $\mathbb{E} \Lambda_{\vartheta}$ für $\vartheta \neq \vartheta_0$ auf einer bezüglich $\mathbb{E} \Lambda_{\vartheta_0}$ *relevanten* Teilmenge des Intervalls I (d.h. die Funktion ist hier *nicht* nur konstant) von $\mathbb{E} \Lambda_{\vartheta_0}$ unterscheidet.

Beweis von Satz 3.3.4⁴⁸:

Es sei $\varepsilon > 0$ gegeben. Dann gilt nach Definition des Schätzers ϑ_n , dass $\|\vartheta_n - \vartheta_0\| \geq \varepsilon$ nur gelten kann, falls das Infimum der Funktion

$$\vartheta \longmapsto \|\overline{N}^{(n)} - \overline{\Lambda}_{\vartheta}^{(n)}\|_{\mu_{\overline{N}^{(n)}}}$$

über alle ϑ aus $\overline{\Theta} \setminus B_{\varepsilon}(\vartheta_0)$ kleiner als das Infimum über alle ϑ aus $B_{\varepsilon}(\vartheta_0)$ ist. Folglich erhalten wir die Mengeneinklusionen

$$\begin{aligned} \{\|\vartheta_n - \vartheta_0\| \geq \varepsilon\} &\subset \left\{ \inf_{\vartheta: \|\vartheta - \vartheta_0\| \geq \varepsilon} \|\overline{N}^{(n)} - \overline{\Lambda}_{\vartheta}^{(n)}\|_{\mu_{\overline{N}^{(n)}}} < \inf_{\vartheta: \|\vartheta - \vartheta_0\| < \varepsilon} \|\overline{N}^{(n)} - \overline{\Lambda}_{\vartheta}^{(n)}\|_{\mu_{\overline{N}^{(n)}}} \right\} \\ &\subset \left\{ \inf_{\vartheta: \|\vartheta - \vartheta_0\| \geq \varepsilon} \|\overline{N}^{(n)} - \overline{\Lambda}_{\vartheta}^{(n)}\|_{\mu_{\overline{N}^{(n)}}} < \|\overline{N}^{(n)} - \overline{\Lambda}_{\vartheta_0}^{(n)}\|_{\mu_{\overline{N}^{(n)}}} \right\}. \end{aligned} \quad (3.3.34)$$

⁴⁸Der Beweis wird in Stute und Kopperschmidt 2013, S. 1287 geführt.

Gemäß Lemma 3.3.3 gilt für $n \rightarrow \infty$ einerseits die Konvergenz

$$\inf_{\vartheta: \|\vartheta - \vartheta_0\| \geq \varepsilon} \|\overline{N}^{(n)} - \overline{\Lambda}_{\vartheta}^{(n)}\|_{\mu_{\overline{N}^{(n)}}} \rightarrow \inf_{\vartheta: \|\vartheta - \vartheta_0\| \geq \varepsilon} \|\mathbb{E} \Lambda_{\vartheta_0} - \mathbb{E} \Lambda_{\vartheta}\|_{\mathbb{E} \Lambda_{\vartheta_0}}, \quad (3.3.35)$$

und der Grenzwert ist nach Voraussetzung (siehe *Identifizierbarkeitsbedingung* aus Gleichung (3.3.31)) positiv.

Andererseits gilt für $n \rightarrow \infty$ nach Lemma 3.3.2 auch

$$\|\overline{N}^{(n)} - \overline{\Lambda}_{\vartheta_0}^{(n)}\|_{\mu_{\overline{N}^{(n)}}} \rightarrow \|\mathbb{E} \Lambda_{\vartheta_0} - \mathbb{E} \Lambda_{\vartheta_0}\|_{\mathbb{E} \Lambda_{\vartheta_0}} = 0, \quad (3.3.36)$$

und daher implizieren die Gleichungen (3.3.35) und (3.3.36), dass die Menge in Gleichung (3.3.34) für $n \rightarrow \infty$ gegen die leere Menge konvergiert. Dementsprechend gilt

$$P\left(\limsup_{n \rightarrow \infty} \{\|\vartheta_n - \vartheta_0\| \geq \varepsilon\}\right) = 0,$$

und somit die starke Konsistenz des Schätzers ϑ_n . □

Beispiel 2.1.7. (10. Fortsetzung)

Analog zur neunten Fortsetzung dieses Beispiels können wir die Identifizierbarkeitsbedingung (3.3.31) vermöge der Struktur des Modells \mathcal{H}_N aus (3.1.5) nachweisen. Es gilt nämlich:

$$\begin{aligned} \|\mathbb{E} \Lambda_{\vartheta_0} - \mathbb{E} \Lambda_{\vartheta}\|_{\mathbb{E} \Lambda_{\vartheta_0}}^2 &= \left\| \mathbb{E} \left[\int_{t_0}^t \vartheta_0 \frac{M}{M - N_s} \mathbb{1}_{\{N_s < M\}} ds \right] - \mathbb{E} \left[\int_{t_0}^t \vartheta \frac{M}{M - N_s} \mathbb{1}_{\{N_s < M\}} ds \right] \right\|_{\mathbb{E} \Lambda_{\vartheta_0}}^2 \\ &= |\vartheta_0 - \vartheta|^2 \int_I \left| \mathbb{E} \left[\int_{t_0}^t \frac{M}{M - N_s} \mathbb{1}_{\{N_s < M\}} ds \right] \right|^2 d\mathbb{E} \Lambda_{\vartheta_0}(t). \end{aligned} \quad (3.3.37)$$

Da nach der Definition eines verschobenen Geburtsprozesses aus Definition 2.1.9 gerade $P(\{N_{t_0} \geq M\}) = 0$ gilt, ist der Integrand aufgrund der Monotonie des Erwartungswertes echt positiv. Aus dem selben Grund ist der Integrator $\mathbb{E} \Lambda_{\vartheta_0}$ nahe t_0 eine streng monoton wachsende Funktion, weshalb das Integral aus Gleichung (3.3.37) eine positive Konstante ist, die wir suggestiv mit $const_{\text{Int}} > 0$ bezeichnen. Damit schließlich erhalten wir für $\varepsilon > 0$:

$$\inf_{\vartheta: \|\vartheta - \vartheta_0\| \geq \varepsilon} \|\mathbb{E} \Lambda_{\vartheta_0} - \mathbb{E} \Lambda_{\vartheta}\|_{\mathbb{E} \Lambda_{\vartheta_0}} = \inf_{\vartheta: \|\vartheta - \vartheta_0\| \geq \varepsilon} \sqrt{const_{\text{Int}}} \cdot |\vartheta_0 - \vartheta| = \sqrt{const_{\text{Int}}} \cdot \varepsilon > 0,$$

und somit die Identifizierbarkeitsbedingung aus (3.3.31).

3.4 Asymptotische Verteilung des Schätzers ϑ_n

In Abschnitt 3.3 haben wir die starke Konsistenz des Schätzers ϑ_n gezeigt. Um jedoch ein - zumindest approximatives - Konfidenzintervall für den wahren Parameter ϑ_0 angeben zu können, benötigen wir Kenntnis über die asymptotische Verteilung von ϑ_n .

Ziel dieses Abschnittes wird es folglich sein, die asymptotische Normalverteiltheit dieses Schätzers zu zeigen.

Generalvoraussetzung in diesem Abschnitt wird erneut die Situation aus Abschnitt 3.1 sein. Weitere Voraussetzungen werden sowohl die Identifizierbarkeitsbedingung aus (3.3.31) als auch die Stetigkeitsforderung aus (3.2.3) sein, so dass wir insbesondere die starke Konsistenz des Schätzers ϑ_n annehmen können. Weiterhin verschärfen wir, wie im Vorfeld von (3.2.3) angekündigt, die Glattheitsforderung an den Kompensator, indem wir zusätzlich fordern:

Der Prozess

$$(t, \vartheta) \longmapsto \Lambda_\vartheta(t) \quad (3.4.1)$$

sei für alle $t \in I$ bezüglich ϑ in einer Umgebung von ϑ_0 fast sicher zweimal stetig differenzierbar.

Diese Forderung liefert sodann ein erstes Resultat, das wir im nachfolgenden Lemma formulieren möchten.

Lemma 3.4.1. (Siehe Stute und Kopperschmidt 2013, S. 1288)

Für hinreichend großes $n \in \mathbb{N}$ gilt die Identität:

$$\begin{aligned} & \int_I [\bar{\Lambda}_{\vartheta_n}^{(n)} - \bar{\Lambda}_{\vartheta_0}^{(n)}] \frac{\partial}{\partial \vartheta} \bar{\Lambda}_\vartheta^{(n)} \Big|_{\vartheta=\vartheta_n} d\mu_{\bar{N}^{(n)}} \\ &= \int_I \bar{M}^{(n)} \frac{\partial}{\partial \vartheta} \bar{\Lambda}_\vartheta^{(n)} \Big|_{\vartheta=\vartheta_n} d\bar{M}^{(n)} + \int_I \bar{M}^{(n)} \frac{\partial}{\partial \vartheta} \bar{\Lambda}_\vartheta^{(n)} \Big|_{\vartheta=\vartheta_n} d\bar{\Lambda}_{\vartheta_0}^{(n)}, \end{aligned} \quad (3.4.2)$$

wobei die Integrale auf der rechten Seite von Gleichung (3.4.2) als *Itô-Integrale* bezüglich $\bar{M}^{(n)}$ beziehungsweise $\bar{\Lambda}_{\vartheta_0}^{(n)}$ zu interpretieren sind.

Beweis⁴⁹:

Zunächst liefert die Gleichung (3.1.10) aus Abschnitt 3.1 die Gültigkeit der Identität

$$d\bar{M}^{(n)} = d\mu_{\bar{N}^{(n)}} - d\bar{\Lambda}_{\vartheta_0}^{(n)}, \quad (3.4.3)$$

⁴⁹In Stute und Kopperschmidt 2013 wird lediglich darauf verwiesen, dass die Ableitung aus (3.4.8) in ϑ_n verschwindet. Die weitere Ausarbeitung geht zurück auf den Autor dieser Arbeit.

so dass sich die rechte Seite aus Gleichung (3.4.2) schreiben lässt als

$$\begin{aligned} & \int_I \overline{M}^{(n)} \frac{\partial}{\partial \vartheta} \overline{\Lambda}_\vartheta^{(n)} \Big|_{\vartheta=\vartheta_n} d\overline{M}^{(n)} + \int_I \overline{M}^{(n)} \frac{\partial}{\partial \vartheta} \overline{\Lambda}_\vartheta^{(n)} \Big|_{\vartheta=\vartheta_n} d\overline{\Lambda}_{\vartheta_0}^{(n)} \\ &= \int_I \overline{M}^{(n)} \frac{\partial}{\partial \vartheta} \overline{\Lambda}_\vartheta^{(n)} \Big|_{\vartheta=\vartheta_n} d\mu_{\overline{N}^{(n)}}. \end{aligned} \quad (3.4.4)$$

Aufgrund der Glattheitsforderung (3.4.1) existiert ein $\varepsilon > 0$, so dass $\overline{\Lambda}_\vartheta^{(n)}(t)$ in $B_\varepsilon(\vartheta_0)$ fast sicher zweimal stetig differenzierbar ist.

Die starke Konsistenz des Schätzers ϑ_n liefert nun für $n \in \mathbb{N}$ hinreichend groß, dass fast sicher $\vartheta_n \in B_{\frac{\varepsilon}{2}}(\vartheta_0)$ gilt. Wir betrachten das Parameterintegral

$$\vartheta \longmapsto \|\overline{N}^{(n)} - \overline{\Lambda}_\vartheta^{(n)}\|_{\mu_{\overline{N}^{(n)}}}^2 = \int_I [\overline{N}^{(n)} - \overline{\Lambda}_\vartheta^{(n)}]^2 d\mu_{\overline{N}^{(n)}}. \quad (3.4.5)$$

Der Integrand aus (3.4.5) ist nach Voraussetzung fast sicher zweimal stetig differenzierbar in $B_{\frac{\varepsilon}{2}}(\vartheta_0)$ und es gilt

$$\begin{aligned} & \left\| \frac{\partial}{\partial \vartheta} [\overline{N}^{(n)}(t) - \overline{\Lambda}_\vartheta^{(n)}(t)]^2 \right\| = 2 \left\| [\overline{N}^{(n)}(t) - \overline{\Lambda}_\vartheta^{(n)}(t)] \frac{\partial}{\partial \vartheta} \overline{\Lambda}_\vartheta^{(n)}(t) \right\| \\ & \leq 2 \sup_{(t,\vartheta) \in I \times B_{\frac{\varepsilon}{2}}(\vartheta_0)} \left\{ \left(\overline{N}^{(n)}(t^0) + \overline{\Lambda}_\vartheta^{(n)}(t^0) \right) \left\| \frac{\partial}{\partial \vartheta} \overline{\Lambda}_\vartheta^{(n)}(t) \right\| \right\}. \end{aligned} \quad (3.4.6)$$

Da in der Abschätzung (3.4.6) das Supremum einer auf $I \times B_\varepsilon(\vartheta_0)$ stetigen Funktion betrachtet wird, existieren ein $t_1 \in I$ und ein $\vartheta_1 \in \overline{B_{\frac{\varepsilon}{2}}(\vartheta_0)}$ mit

$$\left\| \frac{\partial}{\partial \vartheta} [\overline{N}^{(n)}(t) - \overline{\Lambda}_\vartheta^{(n)}(t)]^2 \right\| \leq 2 \left(\overline{N}^{(n)}(t^0) + \overline{\Lambda}_{\vartheta_1}^{(n)}(t^0) \right) \left\| \frac{\partial}{\partial \vartheta} \overline{\Lambda}_\vartheta^{(n)}(t_1) \Big|_{\vartheta=\vartheta_1} \right\|. \quad (3.4.7)$$

Die rechte Seite aus Abschätzung (3.4.7) ist konstant und somit eine gleichmäßige Majorante für den Integranden des Parameterintegrals aus (3.4.5). Nach dem *Satz über die Differenzierbarkeit von Parameterintegralen* erhalten wir also

$$\frac{\partial}{\partial \vartheta} \|\overline{N}^{(n)} - \overline{\Lambda}_\vartheta^{(n)}\|_{\mu_{\overline{N}^{(n)}}}^2 = 2 \int_I [\overline{N}^{(n)} - \overline{\Lambda}_\vartheta^{(n)}] \frac{\partial}{\partial \vartheta} \overline{\Lambda}_\vartheta^{(n)} d\mu_{\overline{N}^{(n)}}. \quad (3.4.8)$$

Nun nimmt das Parameterintegral in ϑ_n definitionsgemäß ein Minimum an, weshalb die Ableitung nach ϑ an dieser Stelle verschwinden muss. Aus Gleichung (3.4.8) folgt daher

schließlich (der Faktor 2 kann offensichtlich vernachlässigt werden):

$$\begin{aligned}
0 &= \frac{\partial}{\partial \vartheta} \|\bar{N}^{(n)} - \bar{\Lambda}_{\vartheta}^{(n)}\|_{\mu_{\bar{N}^{(n)}}}^2 \Big|_{\vartheta=\vartheta_n} = \int_I [\bar{N}^{(n)} - \bar{\Lambda}_{\vartheta_n}^{(n)}] \frac{\partial}{\partial \vartheta} \bar{\Lambda}_{\vartheta}^{(n)} \Big|_{\vartheta=\vartheta_n} d\mu_{\bar{N}^{(n)}} \\
&= \int_I [\bar{N}^{(n)} - \bar{\Lambda}_{\vartheta_0}^{(n)}] \frac{\partial}{\partial \vartheta} \bar{\Lambda}_{\vartheta}^{(n)} \Big|_{\vartheta=\vartheta_n} d\mu_{\bar{N}^{(n)}} + \int_I [\bar{\Lambda}_{\vartheta_0}^{(n)} - \bar{\Lambda}_{\vartheta_n}^{(n)}] \frac{\partial}{\partial \vartheta} \bar{\Lambda}_{\vartheta}^{(n)} \Big|_{\vartheta=\vartheta_n} d\mu_{\bar{N}^{(n)}} \\
&= \int_I \bar{M}^{(n)} \frac{\partial}{\partial \vartheta} \bar{\Lambda}_{\vartheta}^{(n)} \Big|_{\vartheta=\vartheta_n} d\mu_{\bar{N}^{(n)}} + \int_I [\bar{\Lambda}_{\vartheta_0}^{(n)} - \bar{\Lambda}_{\vartheta_n}^{(n)}] \frac{\partial}{\partial \vartheta} \bar{\Lambda}_{\vartheta}^{(n)} \Big|_{\vartheta=\vartheta_n} d\mu_{\bar{N}^{(n)}}
\end{aligned}$$

Diese Gleichung ist äquivalent zu Gleichung (3.4.2); die Behauptung folgt unmittelbar. \square

3.4.1 Die Hilfsprozesse $\{\alpha_n\}$, $\{\beta_n\}$ und $\{\gamma_n\}$

Das obige Lemma 3.4.1 motiviert nun die Definition der parametrischen Prozesse, die in Bemerkung 3.4.2 definiert werden. Sie werden eine wichtige Rolle im Beweis der asymptotischen Normalverteiltheit des Schätzers ϑ_n spielen.

Bemerkung 3.4.2. (siehe Stute und Kopperschmidt 2013, S. 1288)

Durch Gleichung (3.4.2) motiviert definieren wir die parametrischen Prozesse

$$\begin{aligned}
\alpha_n(\vartheta) &= \sqrt{n} \int_I [\bar{\Lambda}_{\vartheta}^{(n)} - \bar{\Lambda}_{\vartheta_0}^{(n)}] \frac{\partial}{\partial \vartheta} \bar{\Lambda}_{\vartheta}^{(n)} d\bar{M}^{(n)}, \\
\beta_n(\vartheta) &= \sqrt{n} \int_I \bar{M}^{(n)} \frac{\partial}{\partial \vartheta} \bar{\Lambda}_{\vartheta}^{(n)} d\bar{M}^{(n)}, \\
\gamma_n(\vartheta) &= \sqrt{n} \int_I \bar{M}^{(n)} \frac{\partial}{\partial \vartheta} \bar{\Lambda}_{\vartheta}^{(n)} d\bar{\Lambda}_{\vartheta_0}^{(n)}.
\end{aligned} \tag{3.4.9}$$

Die Aussage von Lemma 3.4.1 liefert damit:

$$\begin{aligned}
\beta_n(\vartheta_n) + \gamma_n(\vartheta_n) &= \sqrt{n} \int_I [\bar{\Lambda}_{\vartheta_n}^{(n)} - \bar{\Lambda}_{\vartheta_0}^{(n)}] \frac{\partial}{\partial \vartheta} \bar{\Lambda}_{\vartheta}^{(n)} \Big|_{\vartheta=\vartheta_n} d\mu_{\bar{N}^{(n)}} \\
&= \alpha_n(\vartheta_n) + \sqrt{n} \int_I [\bar{\Lambda}_{\vartheta_n}^{(n)} - \bar{\Lambda}_{\vartheta_0}^{(n)}] \frac{\partial}{\partial \vartheta} \bar{\Lambda}_{\vartheta}^{(n)} \Big|_{\vartheta=\vartheta_n} d\bar{\Lambda}_{\vartheta_0}^{(n)}.
\end{aligned} \tag{3.4.10}$$

Wir bemerken an dieser Stelle, dass bedingt durch die Glattheitsforderung (3.4.1) für jede kompakte Teilmenge $K \subset \Theta$ durch

$$(\alpha_n(\cdot)|_K)_{n \in \mathbb{N}}$$

und analog für β_n sowie γ_n eine Folge $C(K)$ -wertiger Zufallsfunktionen definiert wird. Können wir die Straffheit dieser Folge für eine kompakte Menge K mit $\vartheta_0 \in K \subset \Theta$ zeigen (siehe Definition 2.3.2), so liefert die starke Konsistenz des Schätzers ϑ_n unter Verwendung

von Lemma 2.3.6 sofort, dass gilt:

$$\alpha_n(\vartheta_n) \xrightarrow{P} 0 \quad (n \rightarrow \infty).$$

Diese Begebenheit motiviert die Darstellungssätze der obigen parametrischen Prozesse, die in den nachfolgenden Paragraphen thematisiert werden.

Darstellungssatz für die parametrischen Prozesse $\{\alpha_n\}$

In diesem Paragraphen zeigen wir, wie bereits im Rahmen der Bemerkung 3.4.2 vorweggenommen, die Straffheit der Folge $(\alpha_n(\cdot)|_K)_{n \in \mathbb{N}}$ als Familie von $C(K)$ -wertigen Zufallsfunktionen, die unter gewissen Regularitätsbedingungen gilt. Bevor wir also das Ziel dieses Paragraphen in Form des nachfolgenden Lemmas formulieren, erweitern wir erneut - und zum vorletzten mal - die Generalvoraussetzungen:

Es existiere eine Umgebung $U \subset \Theta$ von ϑ_0 , so dass die *Regularitätsbedingung*

$$\mathbb{E} \left[\sup_{(t, \vartheta) \in I \times U} \left\{ \left\| \frac{\partial^j}{\partial \vartheta^j} \Lambda_\vartheta(t) \right\| \right\} \right] < \infty \quad \text{für } j \in \{1, 2\} \quad (3.4.11)$$

erfüllt sei.

In Bemerkung 3.4.3 diskutieren wir zunächst diese Regularitätsbedingung im Kontext des Artikels von Stute und Kopperschmidt 2013 und werden erkennen, dass die hiesige Forderung eine Verschärfung der dort formulierten Voraussetzungen darstellt. Aus diesem Grunde werden wir später den Beweis des Lemmas 3.4.5 wesentlich erweitern:

Wir verwenden zunächst die Regularitätsforderung aus (3.4.11) und zeigen darauffolgend, inwiefern diese situativ abgeschwächt werden kann, um mit Stute und Kopperschmidt 2013 d'accord zu gehen.

Bemerkung 3.4.3.

- In dem Artikel Stute und Kopperschmidt 2013 wird etwa auf Seite 1280 implizit die Forderung

$$\mathbb{E} \left[\sup_{\vartheta \in U} \left\{ \left\| \frac{\partial^j}{\partial \vartheta^j} \Lambda_\vartheta(t) \right\| \right\} \right] < \infty \quad \text{für alle } t \in I, j \in \{1, 2\}$$

gestellt. Hieraus folgert man leicht eine Abschwächung der Regularitätsbedingung aus (3.4.11):

$$\sup_{t \in I} \mathbb{E} \left[\sup_{\vartheta \in U} \left\{ \left\| \frac{\partial^j}{\partial \vartheta^j} \Lambda_\vartheta(t) \right\| \right\} \right] < \infty \quad \text{für } j \in \{1, 2\}.$$

In der Praxis sind zumeist beide Forderungen erfüllt.

- Betrachtet man die Forderung (3.1.12) aus Abschnitt 3.1, so lässt sich diese für $i = 0$ formulieren als

$$\mathbb{E} \left[\sup_{(t, \vartheta) \in I \times \Theta} \left\{ \Lambda_{\vartheta}(t)^j \right\} \right] < \infty \quad \text{für } j \in \{1, 2, 3\}.$$

In Bemerkung 3.2.4 haben wir die Stetigkeitsforderung aus (3.2.3) verwendet, um hieraus die Stetigkeit des Prozesses

$$t \mapsto \mathbb{E} \Lambda_{\vartheta}(t)$$

zu folgern. Analog können wir die Glattheitsforderung aus (3.4.1) utilisieren, um aus der Regularitätsforderung (3.4.11) mittels des *Satzes über die Differenzierbarkeit von Parameterintegralen* zu folgern, dass der Prozess

$$(t, \vartheta) \mapsto \mathbb{E} \Lambda_{\vartheta}(t)$$

ebenfalls für alle $t \in I$ bezüglich ϑ in einer Umgebung U von ϑ_0 zweimal stetig differenzierbar ist. Wir können die verschärfte Generalvoraussetzung folglich auch als Erweiterung der Momentenbedingung aus (3.1.12) auffassen.

- Wie eingangs erwähnt, weichen wir an dieser Stelle geringfügig von der in Stute und Kopperschmidt 2013 gegebenen Situation ab. Bevor wir diesen Schritt begründen, seien hier zunächst die auf Seite 1281 des Artikels formulierten Voraussetzungen angegeben:

Es gelte die *Regularitätsbedingung*

$$\left\| \frac{\partial}{\partial \vartheta} (\mathbb{E} \Lambda_{\vartheta}(t) - \mathbb{E} \Lambda_{\vartheta_0}(t)) \mathbb{E} \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta}(t)^T \right\| \leq g(t) \quad (3.4.12)$$

für alle ϑ in einer Umgebung von ϑ_0 , wobei die Majorante g integrierbar bezüglich $\mathbb{E} \Lambda_{\vartheta_0}$ sei. Ferner sei zudem die Funktion

$$\varphi : I \longrightarrow \mathbb{R}^d : t \mapsto \varphi(t) = \int_{[t, t^0]} \mathbb{E} \left[\frac{\partial}{\partial \vartheta} \Lambda_{\vartheta} \Big|_{\vartheta = \vartheta_0} \right] d\mathbb{E} \Lambda_{\vartheta_0} \quad (3.4.13)$$

bezüglich $\mathbb{E} \Lambda_{\vartheta_0}$ quadratisch integrierbar.

Die erste Forderung wird gestellt, um in der Definition der erst später relevanten

Funktion (siehe Stute und Kopperschmidt 2013, S. 1280)

$$\Phi_0(\vartheta) = \frac{\partial}{\partial \vartheta} \int_I (\mathbb{E}\Lambda_\vartheta - \mathbb{E}\Lambda_{\vartheta_0}) \mathbb{E} \frac{\partial}{\partial \vartheta} \Lambda_\vartheta^T d\mathbb{E}\Lambda_{\vartheta_0} \quad (3.4.14)$$

die Integration und Differentiation zu vertauschen. Um dies zu gewährleisten, muss der Differentialoperator in der Regularitätsbedingung aus (3.4.12) auf den *gesamten* Integranden angewandt werden; die Notation im Artikel Stute und Kopperschmidt 2013 ist hier jedoch ambig.

Die Notwendigkeit der zweiten Forderung wird ersichtlich, sobald wir im Beweis von 3.4.10 nach Anwendung des *Satzes von Fubini* das Integral

$$\sqrt{n} \int_I \varphi d\overline{M}^{(n)} \quad (3.4.15)$$

erhalten. Sie stellt sicher, dass das Integral

$$\mathbb{E} \left\| \int_I \varphi dM \right\|^2$$

existiert und gewährleistet letztlich die Anwendung des *zentralen Grenzwertsatzes*.

- Aus dem letzten Gesichtspunkt wird klar, dass die angegebenen Voraussetzungen für die nachfolgenden Beweise *notwendig* erscheinen.

Im Rahmen der nachfolgenden Lemmata stellen sie sich jedoch als nicht *hinreichend* für die Existenz wichtiger Integrale heraus.

- Die zu beschränkende Funktion aus (3.4.12) ist unter der Regularitätsforderung (3.4.11) stetig bezüglich t . Folglich ist die Existenz einer integrierbaren Majorante g trivialerweise erfüllt. Analog lässt sich auch die Gültigkeit der Forderung (3.4.13) herleiten.
- Die modifizierte Regularitätsbedingung dient lediglich dazu, die *Existenz* der in Lemma 3.4.5 betrachteten Integrale nachzuweisen. Sie wird folglich obsolet, wenn wir den Integranden stets im Sinne der Bedingung (3.4.12) majorisieren können. Dies erlaubt es uns insbesondere, die Regularitätsbedingung aus (3.4.11) situativ abzuschwächen (man erinnere die Bemerkung im Anschluss an die Forderung (3.1.12)).

Wir formulieren nun das zentrale Resultat dieses Paragraphen.

Lemma 3.4.4. (*Darstellungssatz für die Folge* $(\alpha_n(\cdot)|_K)_{n \in \mathbb{N}}$, siehe Stute und Kopperschmidt 2013, S. 1289)

Es gelten die üblichen Identifizierbarkeits- und Glattheitsbedingungen. Unter hinreichenden Regularitätsbedingungen existiert dann ein $\varepsilon > 0$, so dass der Prozess α_n auf $B_\varepsilon(\vartheta_0)$

die Darstellung

$$\alpha_n(\vartheta) = \sqrt{n} \int_I \left[\mathbb{E}\Lambda_\vartheta - \mathbb{E}\Lambda_{\vartheta_0} \right] \mathbb{E} \frac{\partial}{\partial \vartheta} \Lambda_\vartheta \, d\overline{M}^{(n)} + R_1(\vartheta) \quad (3.4.16)$$

besitzt, wobei der führende Term für alle kompakten Teilmengen $K \subset B_\varepsilon(\vartheta_0)$ mit $\vartheta_0 \in K$ straff in $C(K)$ ist. Ferner gilt gleichmäßig auf K die Konvergenz

$$R_1(\vartheta) \xrightarrow{P} 0 \quad (n \rightarrow \infty).$$

Die Regularitätsbedingung aus (3.4.26) stellt eine hinreichende Erweiterung der obigen Regularitätsbedingung aus (3.4.11) dar.

Der Beweis des Lemmas 3.4.4 wird in drei Schritten erfolgen. Zunächst führen wir jedoch einige Notationen ein, die durch den Artikel Stute und Kopperschmidt 2013 motiviert sind. Für eine kompaktere Schreibweise definieren wir dazu:

$$\begin{aligned} \alpha_n^{(1)}(\vartheta) &= \sqrt{n} \int_I \left[\overline{\Lambda}_{\vartheta_0}^{(n)} - \mathbb{E}\Lambda_{\vartheta_0} \right] \mathbb{E} \frac{\partial}{\partial \vartheta} \Lambda_\vartheta \, d\overline{M}^{(n)}, \\ \alpha_n^{(2)}(\vartheta) &= \sqrt{n} \int_I \left[\overline{\Lambda}_\vartheta^{(n)} - \mathbb{E}\Lambda_\vartheta \right] \mathbb{E} \frac{\partial}{\partial \vartheta} \Lambda_\vartheta \, d\overline{M}^{(n)}, \\ \alpha_n^{(3)}(\vartheta) &= \sqrt{n} \int_I \left[\overline{\Lambda}_\vartheta^{(n)} - \overline{\Lambda}_{\vartheta_0}^{(n)} \right] \left[\frac{\partial}{\partial \vartheta} \overline{\Lambda}_\vartheta^{(n)} - \mathbb{E} \frac{\partial}{\partial \vartheta} \Lambda_\vartheta \right] \, d\overline{M}^{(n)}, \\ \tilde{\alpha}_n(\vartheta) &= \sqrt{n} \int_I \left[\mathbb{E}\Lambda_\vartheta - \mathbb{E}\Lambda_{\vartheta_0} \right] \mathbb{E} \frac{\partial}{\partial \vartheta} \Lambda_\vartheta \, d\overline{M}^{(n)}. \end{aligned} \quad (3.4.17)$$

Damit gilt offensichtlich:

$$\begin{aligned} \tilde{\alpha}_n(\vartheta) - \alpha_n^{(1)}(\vartheta) + \alpha_n^{(2)}(\vartheta) + \alpha_n^{(3)}(\vartheta) &= \tilde{\alpha}_n(\vartheta) + \sqrt{n} \int_I \left[\overline{\Lambda}_\vartheta^{(n)} - \overline{\Lambda}_{\vartheta_0}^{(n)} \right] \mathbb{E} \frac{\partial}{\partial \vartheta} \Lambda_\vartheta \, d\overline{M}^{(n)} \\ &\quad + \sqrt{n} \int_I \left[\mathbb{E}\Lambda_{\vartheta_0} - \mathbb{E}\Lambda_\vartheta \right] \mathbb{E} \frac{\partial}{\partial \vartheta} \Lambda_\vartheta \, d\overline{M}^{(n)} + \alpha_n^{(3)}(\vartheta) \\ &= \sqrt{n} \int_I \left[\overline{\Lambda}_\vartheta^{(n)} - \overline{\Lambda}_{\vartheta_0}^{(n)} \right] \mathbb{E} \frac{\partial}{\partial \vartheta} \Lambda_\vartheta \, d\overline{M}^{(n)} + \alpha_n^{(3)}(\vartheta) \\ &= \sqrt{n} \int_I \left[\overline{\Lambda}_\vartheta^{(n)} - \overline{\Lambda}_{\vartheta_0}^{(n)} \right] \frac{\partial}{\partial \vartheta} \overline{\Lambda}_\vartheta^{(n)} \, d\overline{M}^{(n)} = \alpha_n(\vartheta). \end{aligned} \quad (3.4.18)$$

Nach Voraussetzung existiert eine Umgebung $U \subset \Theta$ von ϑ_0 , so dass auf U simultan die Glattheitsforderung aus (3.4.1) und die Regularitätsbedingung aus (3.4.11) erfüllt sind. Wir wählen $\varepsilon > 0$ dergestalt, dass $B_\varepsilon(\vartheta_0) \subset U$ ist. Ist $K \subset B_\varepsilon(\vartheta_0)$ kompakt, so existiert ein $0 < r < \varepsilon$ mit $K \subset \overline{B_r(\vartheta_0)} \subset B_\varepsilon(\vartheta_0)$. Folglich müssen wir für den Beweis nur die

konzentrischen abgeschlossenen r -Bälle mit $0 < r < \varepsilon$ um ϑ_0 betrachten.

Aufgrund von Gleichung (3.4.18) genügt es nun, die nachfolgenden Aussagen zu zeigen, um Lemma 3.4.4 zu beweisen:

- Für jedes $0 < r < \varepsilon$ sind die Folgen $\left(\tilde{\alpha}_n|_{\overline{B_r(\vartheta_0)}}\right)_{n \in \mathbb{N}}$ und $\left(\alpha_n^{(m)}|_{\overline{B_r(\vartheta_0)}}\right)_{n \in \mathbb{N}}$ für $m = 1, 2, 3$ straff in $C(\overline{B_r(\vartheta_0)})$.
- Für jedes feste $\vartheta \in B_\varepsilon(\vartheta_0)$ gilt:

$$\alpha_n^{(m)}(\vartheta) \xrightarrow{P} 0, \quad m = 1, 2, 3.$$

Wir bemerken, dass diese Situation die Anwendung des Straffheitskriteriums von Kolmogorov, wie es ist Korollar 2.3.8 formuliert ist, erlaubt.

Im Lichte der Bemerkung 2.3.9 gelten die nachfolgenden Beweise daher nur im Falle eines eindimensionalen Parameterraumes, wie er etwa in Beispiel 2.1.7 betrachtet wird.

Für eine Erweiterung der Beweise bietet sich das im Anhang formulierte Lipschitz- Straffheitskriterium an, das jedoch eine Verschärfung der Regularitätsforderungen aus (3.4.11) beziehungsweise (3.4.26) erfordert, wovon wir hier absehen möchten.

Der Beweis der obigen Aussagen wird sich über die nächsten drei Lemmata erstrecken. Speziell der Beweis des nächsten Lemmas wird sich dabei hinsichtlich der Bemerkung 3.4.3 als ergiebig herausstellen.

Lemma 3.4.5. (*Straffheit der Folge* $\left(\tilde{\alpha}_n(\cdot)|_{\overline{B_r(\vartheta_0)}}\right)_{n \in \mathbb{N}}$, siehe Stute und Kopperschmidt 2013, S. 1290)

Für jedes $0 < r < \varepsilon$ ist die Folge $\left(\tilde{\alpha}_n(\cdot)|_{\overline{B_r(\vartheta_0)}}\right)_{n \in \mathbb{N}}$ straff in $C(\overline{B_r(\vartheta_0)})$.

Beweis ⁵⁰:

Für den Beweis verwenden wir das Straffheitskriterium von Kolmogorov aus Korollar 2.3.8. Wegen $\vartheta_0 \in \overline{B_r(\vartheta_0)}$ für alle $0 < r < \varepsilon$ und wegen $\tilde{\alpha}_n(\vartheta_0) = 0$ für alle $n \in \mathbb{N}$ (siehe Gleichung (3.4.17)) ist die erste Bedingung

$$P(\{\|\tilde{\alpha}_n(\vartheta_0)\| > a\}) \leq \eta \quad \text{für alle } n \in \mathbb{N}$$

mit einem beliebigen $a > 0$ trivialerweise für alle $\eta > 0$ erfüllt. Es bleibt also zu zeigen, dass eine Konstante $\nu > 0$ existiert mit

$$E\|\tilde{\alpha}_n(\vartheta) - \tilde{\alpha}_n(\vartheta')\|^2 \leq \nu\|\vartheta - \vartheta'\|^2, \quad \text{für alle } n \in \mathbb{N} \text{ und } \vartheta, \vartheta' \in \overline{B_r(\vartheta_0)}. \quad (3.4.19)$$

⁵⁰Ein skizzenhafter Beweis ist in Stute und Kopperschmidt 2013, S. 1290 zu finden. Der Autor dieser Arbeit bietet hier einen ausführlichen Beweis an, der von der Regularitätsbedingung aus (3.4.11) Gebrauch macht.

Einsetzen der Definition aus Gleichung (3.4.17) liefert sofort die Identität⁵¹:

$$\begin{aligned}
\tilde{\alpha}_n(\vartheta) - \tilde{\alpha}_n(\vartheta') &= \sqrt{n} \int_I [\mathbb{E}\Lambda_\vartheta - \mathbb{E}\Lambda_{\vartheta_0}] \mathbb{E} \frac{\partial}{\partial \vartheta} \Lambda_\vartheta \, d\bar{M}^{(n)} \\
&\quad - \sqrt{n} \int_I [\mathbb{E}\Lambda_{\vartheta'} - \mathbb{E}\Lambda_{\vartheta_0}] \mathbb{E} \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta'} \, d\bar{M}^{(n)} \\
&= \sqrt{n} \int_I [\mathbb{E}\Lambda_\vartheta - \mathbb{E}\Lambda_{\vartheta'}] \mathbb{E} \frac{\partial}{\partial \vartheta} \Lambda_\vartheta \, d\bar{M}^{(n)} \\
&\quad + \sqrt{n} \int_I [\mathbb{E}\Lambda_{\vartheta'} - \mathbb{E}\Lambda_{\vartheta_0}] \mathbb{E} \frac{\partial}{\partial \vartheta} \Lambda_\vartheta \, d\bar{M}^{(n)} \\
&\quad - \sqrt{n} \int_I [\mathbb{E}\Lambda_{\vartheta'} - \mathbb{E}\Lambda_{\vartheta_0}] \mathbb{E} \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta'} \, d\bar{M}^{(n)} \\
&= \sqrt{n} \int_I [\mathbb{E}\Lambda_\vartheta - \mathbb{E}\Lambda_{\vartheta'}] \mathbb{E} \frac{\partial}{\partial \vartheta} \Lambda_\vartheta \, d\bar{M}^{(n)} \\
&\quad + \sqrt{n} \int_I [\mathbb{E}\Lambda_{\vartheta'} - \mathbb{E}\Lambda_{\vartheta_0}] \left[\mathbb{E} \frac{\partial}{\partial \vartheta} \Lambda_\vartheta - \mathbb{E} \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta'} \right] \, d\bar{M}^{(n)} \quad (3.4.20)
\end{aligned}$$

Hier und nachfolgend verwenden wir für eine differenzierbare Funktion f die abkürzende Notation

$$\frac{\partial}{\partial \vartheta} f(\vartheta) \Big|_{\vartheta=\vartheta'} = \frac{\partial}{\partial \vartheta} f(\vartheta').$$

Die im Anhang A.1 in ähnlicher Form mannigfaltig angewandte Abschätzung $\|a + b\|^2 \leq 2(\|a\|^2 + \|b\|^2)$ ermöglicht es uns, die Terme auf der rechten Seite von Gleichung (3.4.20) separat zu betrachten.

In beiden Fällen ist der Integrand eine deterministische Funktion, und für eine solche Funktion f gilt:

- Die Integrale $\int_I f \, dM^{(i)}$ sind stochastisch unabhängig.
- Es gilt

$$\mathbb{E} \left[\int_I f \, dM^{(i)} \right] = 0. \quad (3.4.21)$$

Die stochastische Unabhängigkeit der Integrale folgt sofort aus der stochastischen Unabhängigkeit der Martingale $M^{(i)}$.

Völlig analog zum Vorgehen aus Gleichung (3.2.14) des Beweises von Lemma 3.2.3 erhalten wir mit einer entsprechend gewählten Folge $\{t_0 = s_{n,0} < s_{n,1} < \dots < s_{n,n} = t^0\}_{n \in \mathbb{N}}$ von Partitionen des Intervalls I :

⁵¹Unter Beachtung der Regularitätsbedingung aus (3.4.11) ist dieser Zwischenschritt *nicht* notwendig, wie der aufmerksame Leser beobachten wird. Der hiesige Beweis orientiert sich jedoch an der Form des Originalbeweises, um später die Natürlichkeit der Regularitätsbedingung aus (3.4.12) einzusehen.

$$\begin{aligned}
\mathbb{E} \left[\int_I f \, dM^{(i)} \right] &= \mathbb{E} \left[\lim_{n \rightarrow \infty} \sum_{i=0}^{n-1} f(s_{n,i}) \left(M^{(i)}(s_{n,i+1}) - M^{(i)}(s_{n,i}) \right) \right] \\
&= \lim_{n \rightarrow \infty} \mathbb{E} \left[\sum_{i=0}^{n-1} f(s_{n,i}) \left(M^{(i)}(s_{n,i+1}) - M^{(i)}(s_{n,i}) \right) \right] \\
&= \lim_{n \rightarrow \infty} \sum_{i=0}^{n-1} f(s_{n,i}) \mathbb{E} \left[M^{(i)}(s_{n,i+1}) - M^{(i)}(s_{n,i}) \right] \\
&= \lim_{n \rightarrow \infty} \sum_{i=0}^{n-1} f(s_{n,i}) \mathbb{E} \left[\mathbb{E} \left[M^{(i)}(s_{n,i+1}) - M^{(i)}(s_{n,i}) \mid \mathcal{F}_{s_{n,i}} \right] \right] \\
&= \lim_{n \rightarrow \infty} \sum_{i=0}^{n-1} f(s_{n,i}) \mathbb{E} \left[M^{(i)}(s_{n,i}) - M^{(i)}(s_{n,i}) \right] = 0, \tag{3.4.22}
\end{aligned}$$

und dies zeigt Gleichung (3.4.21), wobei wir für die vorletzte Gleichheit ausgenutzt haben, dass $M^{(i)}$ nach Konstruktion aus der *Doob-Meyer-Zerlegung* ein Martingal ist.

Als konvergente Majorante für die Anwendung der majorisierten Konvergenz in Gleichung (3.4.22) eignet sich etwa die Funktion

$$\sup_{t \in I} f(t) [N(t^0) + \Lambda_{\partial_0}(t^0)],$$

wovon man sich beim Vergleich mit (3.2.15) leicht überzeugt.

Aus den obigen Aussagen leiten wir nun die folgende Gleichungskette her:

$$\begin{aligned}
\mathbb{E} \left\| \sqrt{n} \int_I f \, d\bar{M}^{(n)} \right\|^2 &= \mathbb{E} \left\{ \left[\sqrt{n} \int_I f \, d\bar{M}^{(n)} \right]^T \left[\sqrt{n} \int_I f \, d\bar{M}^{(n)} \right] \right\} \\
&= \mathbb{E} \left\{ \left[\frac{1}{\sqrt{n}} \sum_{i=1}^n \int_I f \, dM^{(i)} \right]^T \left[\frac{1}{\sqrt{n}} \sum_{i=1}^n \int_I f \, dM^{(i)} \right] \right\} \\
&= \frac{1}{n} \sum_{i=1}^n \mathbb{E} \left\| \int_I f \, dM^{(i)} \right\|^2 + \frac{1}{n} \sum_{i \neq j} \mathbb{E} \left\{ \left[\int_I f \, dM^{(i)} \right]^T \left[\int_I f \, dM^{(j)} \right] \right\} \\
&= \mathbb{E} \left\| \int_I f \, dM \right\|^2 + \frac{1}{n} \sum_{i \neq j} \mathbb{E} \left[\int_I f \, dM^{(i)} \right]^T \mathbb{E} \left[\int_I f \, dM^{(j)} \right] \\
&= \mathbb{E} \left\| \int_I f \, dM \right\|^2.
\end{aligned}$$

Da M per Konstruktion aus der *Doob-Meyer-Zerlegung* ein quadratisch integrierbares Martingal und der Integrand f als deterministische Funktion insbesondere vorhersagbar ist,

liefert die verallgemeinerte Itô-Isometrie aus Satz 2.1.32:

$$\mathbb{E} \left\| \int_I f \, dM \right\|^2 = \mathbb{E} \left[\int_I \|f\|^2 \, d\Lambda_{\vartheta_0} \right]. \quad (3.4.23)$$

Da in der rechten Seite von (3.4.23) die Funktion f deterministisch ist, liefert erneut das Vorgehen aus (3.2.14) mittels einer *Riemann-Approximation* die Abschätzung:

$$\mathbb{E} \left\| \sqrt{n} \int_I f \, d\bar{M}^{(n)} \right\|^2 = \mathbb{E} \left[\int_I \|f\|^2 \, d\Lambda_{\vartheta_0} \right] = \int_I \|f\|^2 \, d\mathbb{E}\Lambda_{\vartheta_0}.$$

Wir wenden nun die soeben gezeigte Abschätzung nacheinander auf die beiden Summanden der rechten Seite aus Gleichung (3.4.20) an.

Unabhängig von der konkreten Regularitätsbedingung - also der in (3.4.11) formulierten oder einer geeigneten Abschwächung - möchten wir den *Satz von Taylor* für eine Entwicklung um ϑ' anwenden. Wir diskutieren zunächst den Fall der modifizierten Regularitätsbedingung aus (3.4.11), welche sodann die nachfolgende Abschätzung ermöglicht:

$$\begin{aligned} & \mathbb{E} \left\| \sqrt{n} \int_I [\mathbb{E}\Lambda_{\vartheta} - \mathbb{E}\Lambda_{\vartheta'}] \mathbb{E} \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta} \, d\bar{M}^{(n)} \right\|^2 \leq \int_I \left\| [\mathbb{E}\Lambda_{\vartheta} - \mathbb{E}\Lambda_{\vartheta'}] \mathbb{E} \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta} \right\|^2 \, d\mathbb{E}\Lambda_{\vartheta_0} \\ & \leq \int_I \sup_{\vartheta \in B_r(\vartheta_0)} \left\| \frac{\partial}{\partial \vartheta} \left[(\mathbb{E}\Lambda_{\vartheta} - \mathbb{E}\Lambda_{\vartheta'}) \mathbb{E} \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta}^T \right] \right\|^2 \|\vartheta - \vartheta'\|^2 \, d\mathbb{E}\Lambda_{\vartheta_0} \\ & \leq \|\vartheta - \vartheta'\|^2 \left[\sup_{(t, \vartheta) \in I \times B_r(\vartheta_0)} \left\| \frac{\partial}{\partial \vartheta} \left[(\mathbb{E}\Lambda_{\vartheta} - \mathbb{E}\Lambda_{\vartheta'}) \mathbb{E} \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta}^T \right] \right\|^2 \int_I \, d\mathbb{E}\Lambda_{\vartheta_0} \right] \\ & \leq \|\vartheta - \vartheta'\|^2 \left[K_0 \cdot \mathbb{E}\Lambda_{\vartheta_0}(t^0) \right] = \nu \|\vartheta - \vartheta'\|^2, \end{aligned} \quad (3.4.24)$$

mit einer Konstanten K_0 , wobei wir ausgenutzt haben, dass hier das Supremum einer *stetigen* Funktion über dem Kompaktum $I \times \overline{B_r(\vartheta_0)}$ ermittelt wird. Die Konstante K_0 kann und darf dabei von r abhängen.

Wie in Bemerkung 3.4.3 angekündigt, haben wir die Regularitätsbedingung (3.4.11) nur verwendet, um die Existenz des Integrals aus der obigen Abschätzung sicherzustellen. Dies erlaubt es uns, den Beweis des Lemmas mit einer abgeschwächten Regularitätsbedingung zu führen:

Um eine Abschätzung im Sinne von (3.4.24) zu gewährleisten, ist es schließlich völlig ausreichend, die Existenz des Integrals

$$\int_I \sup_{\vartheta \in B_r(\vartheta_0)} \left\| \frac{\partial}{\partial \vartheta} \left[(\mathbb{E}\Lambda_{\vartheta} - \mathbb{E}\Lambda_{\vartheta'}) \mathbb{E} \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta}^T \right] \right\|^2 \, d\mathbb{E}\Lambda_{\vartheta_0}$$

vorauszusetzen. Offensichtlich existiert das Integral, wenn eine bezüglich $\mathbb{E}\Lambda_{\vartheta_0}$ *quadratisch*

integrierbare Funktion g_1 existiert mit

$$\left\| \frac{\partial}{\partial \vartheta} \left[(\mathbb{E}\Lambda_{\vartheta}(t) - \mathbb{E}\Lambda_{\vartheta'}(t)) \mathbb{E} \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta}(t)^T \right] \right\| \leq g_1(t)$$

für alle $\vartheta \in B_{\varepsilon}(\vartheta_0)$. Die Regularitätsforderung aus (3.4.12) alleine genügt offenbar nicht, um diese Abschätzung für arbiträre $\vartheta' \in B_r(\vartheta_0)$ zu folgern.

Erweitert man jedoch die in Stute und Kopperschmidt 2013 formulierte Regularitätsbedingung auf natürliche Art und Weise, um die Majorisierung sämtlicher auftretender Integranden mit einer *quadratisch* integrierbaren Funktion sicherzustellen, so kann die in (3.4.11) formulierte Verschärfung der Regularitätsbedingung abgeschwächt werden.

Wir möchten den Beweis des Lemmas nun abschließen. Unter der Regularitätsforderung (3.4.11) erhalten wir mit einem völlig analogen Vorgehen wiederum über die Stetigkeit des Integranden schließlich ebenfalls:

$$\begin{aligned} & \mathbb{E} \left\| \sqrt{n} \int_I [\mathbb{E}\Lambda_{\vartheta'} - \mathbb{E}\Lambda_{\vartheta_0}] \left[\mathbb{E} \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta} - \mathbb{E} \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta'} \right] d\bar{M}^{(n)} \right\|^2 \\ & \leq \nu_2 \|\vartheta - \vartheta'\|^2. \end{aligned} \tag{3.4.25}$$

Hier genügt es nun alternativ, die Existenz des Integrals

$$\int_I \sup_{\vartheta \in B_r(\vartheta_0)} \left\| \frac{\partial}{\partial \vartheta} \left[(\mathbb{E}\Lambda_{\vartheta'} - \mathbb{E}\Lambda_{\vartheta_0}) \left(\mathbb{E} \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta} - \mathbb{E} \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta'} \right)^T \right] \right\|^2 d\mathbb{E}\Lambda_{\vartheta_0}$$

vorauszusetzen. Das Integral existiert bereits, wenn eine bezüglich $\mathbb{E}\Lambda_{\vartheta_0}$ quadratisch integrierbare Funktion g_2 existiert mit

$$\left\| \mathbb{E} \left[\frac{\partial^j}{\partial \vartheta^j} \Lambda_{\vartheta}(t) \right] \right\| \leq g_2(t), \quad \text{für } j \in \{1, 2\}$$

und alle $\vartheta \in B_r(\vartheta_0)$. Eine einfache wenngleich umfangreiche Rechnung offenbart, dass diese Forderung gemeinsam mit der Regularitätsforderung aus (3.4.12) sogar die Existenz der obigen Majorante g_1 impliziert, falls die Majorante g aus dieser Forderung ebenfalls *quadratisch* integrierbar ist.

In jedem Falle liefert die Kombination der Abschätzungen (3.4.24) und (3.4.25) schlussendlich die Gültigkeit von (3.4.19) und damit die Behauptung. \square

Im Beweis von Lemma 3.4.5 waren die Integranden stets deterministische Funktionen. Wir möchten die Regularitätsbedingung aus (3.4.11) geeignet anpassen, um später auch Integrale kontrollieren zu können, deren Integranden weder vorhersagbar noch unabhängig

vom Integrator sind.

Um also den Beweis von Lemma 3.4.4 fertigzustellen, muss die Generalvoraussetzung letztmalig verschärft werden, indem die Regularitätsbedingung aus (3.4.11) und die Momentenbedingung aus (3.1.12) kombiniert werden.

Es existiere eine Umgebung $U \subset \Theta$ von ϑ_0 , so dass die nachfolgende *kombinierte Regularitätsbedingung* erfüllt sei:

$$\mathbb{E} \left[\sup_{(t, \vartheta) \in I \times U} \left\{ N(t^0)^i \cdot \Lambda_{\vartheta}(t^0)^j \cdot \left\| \frac{\partial^k}{\partial \vartheta^k} \Lambda_{\vartheta}(t) \right\|^l \right\} \right] < \infty \quad \text{für } i + j \leq 4, k \in \{1, 2\}, l \leq 2. \quad (3.4.26)$$

Wir werden ohne Beschränkung der Allgemeinheit annehmen, dass $B_{\varepsilon}(\vartheta_0) \subset U$ gilt. Andernfalls wählen wir retrospektiv ein hinreichend kleines $\varepsilon > 0$, so dass diese Bedingung erfüllt sei.

Im Rahmen des Beweises von Lemma 3.4.5 haben wir exemplarisch gezeigt, inwieweit die Regularitätsbedingung aus (3.4.11) abgeschwächt werden kann. Analog kann auch in den nachfolgenden Lemmata die kombinierte Regularitätsbedingung aus (3.4.26) abgeschwächt werden, wir verzichten jedoch darauf, dieses Argument stets wieder aufzugreifen.

In diesem Sinne möchten wir nun mit dem nächsten Schritt des Beweises von Lemma 3.4.4 fortfahren:

Lemma 3.4.6. (Asymptotik der Folgen $(\alpha_n^{(m)}(\vartheta))_{n \in \mathbb{N}}$ für $m = 1, 2, 3$, siehe Stute und Kopperschmidt 2013, S. 1290)

Für jedes feste $\vartheta \in B_{\varepsilon}(\vartheta_0)$ gilt:

$$\alpha_n^{(m)}(\vartheta) \xrightarrow{P} 0, \quad m = 1, 2, 3. \quad (3.4.27)$$

Beweis⁵²:

Wir zeigen die Behauptung für $m = 1, 2$ simultan. Dazu erinnern wir an dieser Stelle zunächst an die Definition von $\alpha_n^{(1)}$ und $\alpha_n^{(2)}$ aus Gleichung (3.4.17):

$$\begin{aligned} \alpha_n^{(1)} &= \sqrt{n} \int_I \left[\bar{\Lambda}_{\vartheta_0}^{(n)} - \mathbb{E} \Lambda_{\vartheta_0} \right] \mathbb{E} \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta} \, d \bar{M}^{(n)}, \\ \alpha_n^{(2)} &= \sqrt{n} \int_I \left[\bar{\Lambda}_{\vartheta}^{(n)} - \mathbb{E} \Lambda_{\vartheta} \right] \mathbb{E} \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta} \, d \bar{M}^{(n)}. \end{aligned}$$

⁵²Ein skizzenhafter Beweis ist wiederum in Stute und Kopperschmidt 2013 zu finden. Die detaillierte Ausarbeitung, die insbesondere die Anwendbarkeit der Hilfssätze aus dem Anhang diskutiert, geht zurück auf den Autor dieser Arbeit.

Die Aussage folgt also sofort, wenn wir zeigen, dass für alle $\vartheta, \vartheta' \in B_\varepsilon(\vartheta_0)$ die Konvergenz

$$\sqrt{n} \int_I \left[\bar{\Lambda}_{\vartheta'}^{(n)} - E\Lambda_{\vartheta'} \right] E \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta} d\bar{M}^{(n)} \xrightarrow{P} 0 \quad (n \rightarrow \infty) \quad (3.4.28)$$

gilt. Definieren wir die vektorwertigen Zufallsvariablen U_{ki} durch

$$U_{ki} = \int_I \left[\Lambda_{\vartheta'}^{(k)} - E\Lambda_{\vartheta'} \right] E \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta} dM^{(i)}, \quad k, i \in \{1, \dots, n\}, \quad (3.4.29)$$

so gilt für das Integral aus Gleichung (3.4.28):

$$\begin{aligned} & \sqrt{n} \int_I \left[\bar{\Lambda}_{\vartheta'}^{(n)} - E\Lambda_{\vartheta'} \right] E \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta} d\bar{M}^{(n)} \\ &= \frac{\sqrt{n}}{n^2} \sum_{k=1}^n \sum_{i=1}^n \int_I \left[\Lambda_{\vartheta'}^{(k)} - E\Lambda_{\vartheta'} \right] E \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta} dM^{(i)} = n^{-\frac{3}{2}} \sum_{i,k=1}^n U_{ki}. \end{aligned} \quad (3.4.30)$$

Da die $\Lambda_{\vartheta}^{(k)}$ gemäß der Stetigkeitsforderung aus (3.2.3) stetig in t sind, folgt sofort, dass diese Prozesse vorhersagbar sind (vergleiche Definition 2.1.22).

Als Komposition deterministischer und vorhersagbarer Prozesse sind folglich die Integranden aus der Definition der Zufallsvariablen U_{ki} vorhersagbar.

Ist allgemein f ein solcher vorhersagbarer Prozess, so ist aufgrund von Satz 2.1.32 der durch

$$t \mapsto \int_{t_0}^t f dM^{(i)}$$

definierte stochastische Prozess für $i \in \mathbb{N}$ ein Martingal bezüglich der gleichen Filtration wie $M^{(i)}$. Hieraus folgt dann trivialerweise

$$E \left[\int_{t_0}^{t^0} f dM^{(i)} \right] = E \left[\int_I f dM^{(i)} \right] = 0, \quad \text{für alle } i \in \mathbb{N}, \quad (3.4.31)$$

und insbesondere $E(U_{ki}) = 0$ für alle $k, i \in \{1, \dots, n\}$ gilt.

Betrachten wir nun also in Gleichung (3.4.30) die Summe über alle $i = k$, so erhalten wir eine Summe *unabhängig identisch verteilter* verteilter Zufallsvariablen mit Erwartungswert 0. Das *starke Gesetz der großen Zahlen*, das wir als Spezialfall des in Satz 2.2.3 formulierten Resultates kennengelernt haben, liefert also unmittelbar die Konvergenz:

$$\frac{1}{n} \sum_{i=1}^n U_{ii} \xrightarrow{P} 0 \quad (n \rightarrow \infty),$$

und somit gilt auch

$$n^{-\frac{3}{2}} \sum_{\substack{i,k=1 \\ i=k}}^n U_{ik} \xrightarrow{P} 0 \quad (n \rightarrow \infty).$$

Um die Summe über alle $i \neq k$ zu beschränken, möchten wir Lemma A.1.2 (siehe Anhang) verwenden. Im weiteren Verlauf dieser Arbeit werden auch die Lemmata A.1.1 und A.1.3 aus dem Anhang utlisiert, wobei jeweils die Voraussetzungen zu überprüfen sind. Da die verwendeten Argumente stets gleich sind, möchten wir die Gültigkeit der Voraussetzungen an dieser Stelle exemplarisch demonstrieren, um die benötigten Methoden zu dokumentieren und nachfolgend auf das Vorgehen zu verweisen.

- Für $i, k \in \{1, \dots, n\}$ gilt $\mathbb{E}\|U_{ki}\|^2 < \infty$: Die verallgemeinerte Itô-Isometrie, die schon in der Abschätzung (3.4.23) zur Anwendung kam, bleibt gültig, wenn der Integrand ein vorhersagbarer Prozess ist. Wir erhalten also sofort:

$$\mathbb{E}\|U_{ki}\|^2 \leq \mathbb{E} \left[\int_I \left\| \left[\Lambda_{\vartheta'}^{(k)} - \mathbb{E}\Lambda_{\vartheta'} \right] \mathbb{E} \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta} \right\|^2 d\Lambda_{\vartheta_0}^{(i)} \right].$$

Der Integrand ist aufgrund der Glattheitsforderung aus (3.4.1) stetig in t , und folglich gilt

$$\begin{aligned} \mathbb{E} \left[\int_I \left\| \left[\Lambda_{\vartheta'}^{(k)} - \mathbb{E}\Lambda_{\vartheta'} \right] \mathbb{E} \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta} \right\|^2 d\Lambda_{\vartheta_0}^{(i)} \right] &\leq \mathbb{E} \left[\sup_I \left\| \left[\Lambda_{\vartheta'}^{(k)} - \mathbb{E}\Lambda_{\vartheta'} \right] \mathbb{E} \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta} \right\|^2 \Lambda_{\vartheta_0}^{(i)}(t^0) \right] \\ &\leq \sup_I \left\| \mathbb{E} \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta} \right\|^2 \mathbb{E} \left[2 \left(\Lambda_{\vartheta'}^{(k)}(t^0)^2 + \mathbb{E}\Lambda_{\vartheta'}(t^0)^2 \right) \Lambda_{\vartheta_0}^{(i)}(t^0) \right] \\ &\leq \sup_I \left\| \mathbb{E} \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta} \right\|^2 \left[2\mathbb{E} \left[\Lambda_{\vartheta'}^{(k)}(t^0)^2 \Lambda_{\vartheta_0}^{(i)}(t^0) \right] + 2\mathbb{E}\Lambda_{\vartheta'}(t^0)^2 \mathbb{E}\Lambda_{\vartheta_0}(t^0) \right]. \end{aligned} \quad (3.4.32)$$

Die Behauptung folgt nun aus der Annahme aus (3.1.12).

- Es gilt die Bedingung aus (A.1.6): Die Überprüfung dieser Voraussetzung wird sich als länglich erweisen und ist der Grund, warum wir das Vorgehen nur an dieser Stelle in vollem Umfang aufzeigen möchten. Wir möchten zeigen, dass $\mathbb{E} [U_{ki}^T U_{lj}] = 0$ gilt, sobald $k \notin \{i, l, j\}$, $i \notin \{k, l, j\}$, $l \notin \{k, i, j\}$ oder $j \notin \{k, i, l\}$ ist. Aus Symmetriegründen genügt es, die ersten beiden Fälle zu untersuchen:

– $k \notin \{i, l, j\}$: Offensichtlich gilt die Identität:

$$\mathbb{E} \left[\left(\Lambda_{\vartheta'}^{(k)} - \mathbb{E}\Lambda_{\vartheta'} \right) \mathbb{E} \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta} \right] = \left[(\mathbb{E}\Lambda_{\vartheta'} - \mathbb{E}\Lambda_{\vartheta'}) \mathbb{E} \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta} \right] = 0. \quad (3.4.33)$$

Betrachten wir also eine *Riemann-Approximation* von

$$\begin{aligned} \mathbb{E} [U_{ki}^T U_{lj}] &= \mathbb{E} \left[\int_I \int_I \left[\left(\Lambda_{\vartheta'}^{(k)}(s) - \mathbb{E} \Lambda_{\vartheta'}(s) \right) \mathbb{E} \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta}(s) \right]^T \right. \\ &\quad \left. \left[\left(\Lambda_{\vartheta'}^{(l)}(t) - \mathbb{E} \Lambda_{\vartheta'}(t) \right) \mathbb{E} \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta}(t) \right] dM^{(i)}(s) dM^{(j)}(t) \right], \end{aligned}$$

so hat diese mit entsprechend gewählten Folgen $\{t_0 = s_{n,0} < s_{n,1} < \dots < s_{n,n} = t^0\}_{n \in \mathbb{N}}$, $\{t_0 = t_{n,0} < t_{n,1} < \dots < t_{n,n} = t^0\}_{n \in \mathbb{N}}$ von Partitionen des Intervalls I die Gestalt

$$\begin{aligned} \mathbb{E} \left[\lim_{n \rightarrow \infty} \lim_{m \rightarrow \infty} \sum_{p=0}^{n-1} \sum_{q=0}^{m-1} \left[\left(\Lambda_{\vartheta'}^{(k)}(s_{n,p}) - \mathbb{E} \Lambda_{\vartheta'}(s_{n,p}) \right) \mathbb{E} \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta}(s_{n,p}) \right]^T \right. \\ \left. \left[\left(\Lambda_{\vartheta'}^{(l)}(t_{m,q}) - \mathbb{E} \Lambda_{\vartheta'}(t_{m,q}) \right) \mathbb{E} \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta}(t_{m,q}) \right] \left(M^{(i)}(s_{n,p+1}) - M^{(i)}(s_{n,p}) \right) \right. \\ \left. \left(M^{(j)}(t_{m,q+1}) - M^{(j)}(t_{m,q}) \right) \right]. \end{aligned} \quad (3.4.34)$$

Nutzen wir wiederum die majorisierte Konvergenz und beachten die Unabhängigkeit des Prozesses $\Lambda_{\vartheta'}^{(k)}$ von $\Lambda_{\vartheta'}^{(l)}$, $M^{(i)}$ und $M^{(j)}$, so ist der Ausdruck aus (3.4.34) gleich

$$\begin{aligned} \lim_{n \rightarrow \infty} \lim_{m \rightarrow \infty} \sum_{p=0}^{n-1} \sum_{q=0}^{m-1} \mathbb{E} \left[\left(\Lambda_{\vartheta'}^{(k)}(s_{n,p}) - \mathbb{E} \Lambda_{\vartheta'}(s_{n,p}) \right) \mathbb{E} \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta}(s_{n,p}) \right]^T \\ \mathbb{E} \left[\left[\left(\Lambda_{\vartheta'}^{(l)}(t_{m,q}) - \mathbb{E} \Lambda_{\vartheta'}(t_{m,q}) \right) \mathbb{E} \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta}(t_{m,q}) \right] \left(M^{(i)}(s_{n,p+1}) - M^{(i)}(s_{n,p}) \right) \right. \\ \left. \left(M^{(j)}(t_{m,q+1}) - M^{(j)}(t_{m,q}) \right) \right] \end{aligned} \quad (3.4.35)$$

und damit nach Gleichung (3.4.33) gerade 0.

- $i \notin \{k, l, j\}$: Wir verfahren zunächst wie im vorherigen Fall. Anstelle des Terms (3.4.35) erreichen wir schließlich den Ausdruck

$$\begin{aligned} \lim_{n \rightarrow \infty} \lim_{m \rightarrow \infty} \sum_{p=0}^{n-1} \sum_{q=0}^{m-1} \mathbb{E} \left[\left[\left(\Lambda_{\vartheta'}^{(k)}(s_{n,p}) - \mathbb{E} \Lambda_{\vartheta'}(s_{n,p}) \right) \mathbb{E} \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta}(s_{n,p}) \right]^T \right. \\ \left. \left[\left(\Lambda_{\vartheta'}^{(l)}(t_{m,q}) - \mathbb{E} \Lambda_{\vartheta'}(t_{m,q}) \right) \mathbb{E} \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta}(t_{m,q}) \right] \left(M^{(j)}(t_{m,q+1}) - M^{(j)}(t_{m,q}) \right) \right] \\ \cdot \mathbb{E} \left[M^{(i)}(s_{n,p+1}) - M^{(i)}(s_{n,p}) \right] \end{aligned} \quad (3.4.36)$$

Da $M^{(i)}$ ein Martingal ist, folgt

$$\mathbb{E} \left[M^{(i)}(s_{n,p+1}) - M^{(i)}(s_{n,p}) \right] = 0.$$

Somit ist auch der Ausdruck aus (3.4.36) gerade 0.

Insgesamt folgt aus den beiden Fällen die Gültigkeit der Bedingung aus (A.1.6).

Die Verwendung von Lemma A.1.2 liefert nun:

$$\frac{1}{n^3} \mathbb{E} \left\| \sum_{\substack{k,i=1 \\ k \neq i}}^n U_{ki} \right\|^2 \leq \frac{2}{n^3} \sum_{\substack{k,i=1 \\ k \neq i}}^n \mathbb{E} \|U_{ki}\|^2. \quad (3.4.37)$$

Da die U_{ki} für $k \neq i$ identisch verteilt sind, besteht die Summe auf der rechten Seite der Abschätzung aus (3.4.37) aus $n^2 - n$ identischen Summanden, und wir erhalten:

$$\frac{1}{n^3} \mathbb{E} \left\| \sum_{\substack{k,i=1 \\ k \neq i}}^n U_{ki} \right\|^2 \leq \frac{2(n^2 - n)}{n^3} \text{const} \rightarrow 0 \quad (n \rightarrow \infty). \quad (3.4.38)$$

Hierin wird uns sodann die Sinnhaftigkeit der Hilfssätze aus Abschnitt A.1 gewahrt:

Während wir durch obige Anwendung des Lemmas $\mathcal{O}(n^2)$ endliche und identische Summanden erhielten, resultierte das naive Ausmultiplizieren der linken Seite aus Abschätzung (3.4.38) in $\mathcal{O}(n^4)$ gleichermaßen endlichen - wenngleich nicht notwendig identischen - Summanden, die durch den Faktor n^{-3} nicht unmittelbar egalisiert werden können.

Die *Chebyshev-Markovschen Ungleichung* (siehe etwa Bauer 1992, S. 128) liefert damit für $\eta > 0$:

$$P \left(\left\| n^{-\frac{3}{2}} \sum_{k \neq i} U_{ki} \right\| \geq \eta \right) \leq \frac{1}{\eta^2} \mathbb{E} \left\| n^{-\frac{3}{2}} \sum_{k \neq i} U_{ki} \right\|^2 = \frac{1}{\eta^2 n^3} \mathbb{E} \left\| \sum_{k \neq i} U_{ki} \right\|^2.$$

Die rechte Seite dieser Abschätzung geht nun nach (3.4.38) für $n \rightarrow \infty$ gegen 0, und somit gilt schließlich die Behauptung aus (3.4.28).

Es verbleibt der Fall $m = 3$. Wir verfahren analog zu den Fällen $m = 1, 2$ und definieren

$$U_{pki} = \int_I \left[\Lambda_{\vartheta}^{(p)} - \Lambda_{\vartheta_0}^{(p)} \right] \left[\frac{\partial}{\partial \vartheta} \Lambda_{\vartheta}^{(k)} - \mathbb{E} \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta} \right] dM^{(i)}.$$

Durch Einsetzen der Definition sehen wir unmittelbar:

$$\begin{aligned}
\alpha_n^{(3)} &= \sqrt{n} \int_I \left[\bar{\Lambda}_\vartheta^{(n)} - \bar{\Lambda}_{\vartheta_0}^{(n)} \right] \left[\frac{\partial}{\partial \vartheta} \bar{\Lambda}_\vartheta^{(n)} - \mathbb{E} \frac{\partial}{\partial \vartheta} \Lambda_\vartheta \right] d\bar{M}^{(n)} \\
&= \frac{\sqrt{n}}{n^3} \sum_{p=1}^n \sum_{k=1}^n \sum_{i=1}^n \int_I \left[\Lambda_\vartheta^{(p)} - \Lambda_{\vartheta_0}^{(p)} \right] \left[\frac{\partial}{\partial \vartheta} \Lambda_\vartheta^{(k)} - \mathbb{E} \frac{\partial}{\partial \vartheta} \Lambda_\vartheta \right] dM^{(i)} \\
&= n^{-\frac{5}{2}} \sum_{p,k,i=1}^n U_{pki}
\end{aligned} \tag{3.4.39}$$

Wir teilen die Summe aus (3.4.39) in geeignete Untersummen auf, die jeweils mit den zuvor illustrierten Verfahren gehandhabt werden können:

$$\begin{aligned}
n^{-\frac{5}{2}} \sum_{p,k,i=1}^n U_{pki} &= n^{-\frac{5}{2}} \sum_{\substack{p,k,i=1 \\ p=k=i}}^n U_{pki} + n^{-\frac{5}{2}} \sum_{\substack{p,k,i=1 \\ p=k \neq i}}^n U_{pki} + n^{-\frac{5}{2}} \sum_{\substack{p,k,i=1 \\ k=i \neq p}}^n U_{pki} \\
&\quad + n^{-\frac{5}{2}} \sum_{\substack{p,k,i=1 \\ i=p \neq k}}^n U_{pki} + n^{-\frac{5}{2}} \sum_{\substack{p,k,i=1 \\ p \neq k \neq i \neq p}}^n U_{pki}
\end{aligned} \tag{3.4.40}$$

Wir zeigen für jede der fünf Summen aus (3.4.40) individuell die stochastische Konvergenz gegen 0.

$p = k = i$: Die Unabhängigkeit der Summanden liefert sofort:

$$\mathbb{E} \left\| n^{-\frac{5}{2}} \sum_{\substack{p,k,i=1 \\ p=k=i}}^n U_{pki} \right\|^2 = \mathbb{E} \left\| n^{-\frac{5}{2}} \sum_{i=1}^n U_{iii} \right\|^2 = \frac{1}{n^5} \sum_{i=1}^n \mathbb{E} \|U_{iii}\|^2. \tag{3.4.41}$$

Da die U_{iii} für $i \in \{1, \dots, n\}$ identisch verteilt sind, erfolgt die Summation also über $\mathcal{O}(n)$ identische Summanden.

Die Existenz des Integrals $\mathbb{E} \|U_{iii}\|^2$ folgert man dabei wie zuvor unter Verwendung der Itô-Isometrie aus der Regularitätsbedingung (3.4.11) und der Momentenbedingung (3.1.12), vergleiche (3.4.31).

Der Faktor n^{-5} liefert daher in (3.4.41) die Konvergenz gegen 0, und wie gehabt folgert man die stochastische Konvergenz gegen 0 aus der *Chebyshev-Markovschen Ungleichung*.

$p = k \neq i$: Anders als in (3.4.37) führt hier das naive Ausmultiplizieren zum Ziel, da der Faktor n^{-5} die Konvergenz sicherstellt.

Wir erhalten zunächst:

$$\begin{aligned}
\mathbb{E} \left\| n^{-\frac{5}{2}} \sum_{\substack{p,k,i=1 \\ p=k \neq i}}^n U_{pki} \right\|^2 &= \mathbb{E} \left\| n^{-\frac{5}{2}} \sum_{\substack{k,i=1 \\ k \neq i}}^n U_{kki} \right\|^2 = \frac{1}{n^5} \sum_{\substack{k,i=1 \\ k \neq i}}^n \mathbb{E} \|U_{kki}\|^2 \\
&+ \frac{1}{n^5} \sum_{\substack{k \neq i, l \neq j, \\ k \neq l, k \neq j, \\ i \neq l, i \neq j}} \mathbb{E}(U_{kki}^T U_{llj}) + \frac{1}{n^5} \sum_{\substack{k \neq i, l \neq j, \\ k=l, i \neq j}} \mathbb{E}(U_{kki}^T U_{llj}) + \frac{1}{n^5} \sum_{\substack{k \neq i, l \neq j, \\ k \neq l, i=j}} \mathbb{E}(U_{kki}^T U_{llj}) \\
&+ \frac{1}{n^5} \sum_{\substack{k \neq i, l \neq j, \\ k=j, i \neq l}} \mathbb{E}(U_{kki}^T U_{llj}) + \frac{1}{n^5} \sum_{\substack{k \neq i, l \neq j, \\ k \neq j, i=l}} \mathbb{E}(U_{kki}^T U_{llj}) + \frac{1}{n^5} \sum_{\substack{k \neq i, l \neq j, \\ k=j, i=l}} \mathbb{E}(U_{kki}^T U_{llj}).
\end{aligned} \tag{3.4.42}$$

Die auf den ersten Blick unnötig kompliziert anmutende Schreibweise aus (3.4.42) unterteilt die Summe in weitere Untersummen, die durch die Relation der Indizes zueinander klassifiziert werden. Diese Aufspaltung hat den inhärenten Vorteil, dass jede Summe auf der rechten Seite von (3.4.42) aus *identischen* Summanden besteht. Die Gleichung vereinfacht sich also wie folgt:

$$\begin{aligned}
\mathbb{E} \left\| n^{-\frac{5}{2}} \sum_{\substack{p,k,i=1 \\ p=k \neq i}}^n U_{pki} \right\|^2 &= \frac{1}{n^5} \left[\mathcal{O}(n^2) \mathbb{E} \|U_{112}\|^2 + \mathcal{O}(n^4) \mathbb{E}(U_{112}^T U_{334}) \right. \\
&+ \mathcal{O}(n^3) \mathbb{E}(U_{112}^T U_{113}) + \mathcal{O}(n^3) \mathbb{E}(U_{112}^T U_{332}) + \mathcal{O}(n^3) \mathbb{E}(U_{112}^T U_{331}) \\
&\left. + \mathcal{O}(n^3) \mathbb{E}(U_{112}^T U_{223}) + \mathcal{O}(n^2) \mathbb{E}(U_{112}^T U_{221}) \right],
\end{aligned} \tag{3.4.43}$$

und aus der Existenz der Erwartungswerte folgt wiederum die Konvergenz gegen 0. Mit der *Chebyshev-Markovschen Ungleichung* erhalten wir erneut die stochastische Konvergenz gegen 0.

$k = i \neq p$: Der Beweis erfolgt analog zum Fall $p = k \neq i$.

$i = p \neq k$: Auch hier kommt das obige Beweisverfahren zur Anwendung.

$p \neq k \neq i \neq p$: Wir verwenden hier das Lemma A.1.3 und erhalten:

$$\mathbb{E} \left\| n^{-\frac{5}{2}} \sum_{\substack{p,k,i=1 \\ p \neq k \neq i \neq p}}^n U_{pki} \right\|^2 \leq \frac{32}{n^5} \sum \mathbb{E}(U_{pki}^T U_{qki}), \tag{3.4.44}$$

wobei die Summation über alle Indexkombinationen erfolgt, in denen p und q die selbe Position relativ zu i und k einnehmen (siehe A.1.3 für eine ausführlichere Erläuterung). Wir erhalten offensichtlich $\mathcal{O}(n^4)$ Summanden, deren Endlichkeit man

wie zuvor beweist⁵³. Die Behauptung folgt dann leicht mit den obigen Argumenten.

Insgesamt folgt also in (3.4.39) die stochastische Konvergenz gegen 0 und damit auch im Falle $m = 3$ die Behauptung. \square

Das Vorgehen aus dem Beweis von Lemma 3.4.6 wird sich auch für das letzte Lemma, das zum Beweis des Lemmas 3.4.4 vonnöten ist, als zielführend erweisen.

Lemma 3.4.7. (Straffheit der Folgen $\left(\alpha_n^{(m)}(\cdot)|_{\overline{B_r(\vartheta_0)}}\right)_{n \in \mathbb{N}}$ **für** $m = 1, 2, 3$, siehe Stute und Kopperschmidt 2013, S. 1291)

Für jedes $0 < r < \varepsilon$ und $m = 1, 2, 3$ ist die Folge $\left(\alpha_n^{(m)}(\cdot)|_{\overline{B_r(\vartheta_0)}}\right)_{n \in \mathbb{N}}$ straff in $C(\overline{B_r(\vartheta_0)})$.

Beweis⁵⁴:

Für den Straffheitsbeweis werden wir erneut das *Kriterium von Kolmogorov* verwenden. Die Gültigkeit der ersten Bedingung für ein beliebiges $\vartheta \in \overline{B_r(\vartheta_0)}$ folgt trivialerweise aus Lemma 3.4.6 und der Definition der stochastischen Konvergenz.

Wie in Lemma 3.4.5 bleibt also zu zeigen, dass für $m = 1, 2, 3$ eine Konstante $\nu_m > 0$ existiert mit

$$\mathbb{E} \left\| \alpha_n^{(m)}(\vartheta) - \alpha_n^{(m)}(\vartheta') \right\|^2 \leq \nu_m \|\vartheta - \vartheta'\|^2, \quad \text{für alle } \vartheta, \vartheta' \in \overline{B_r(\vartheta_0)} \text{ und } m = 1, 2, 3, n \in \mathbb{N}. \quad (3.4.45)$$

Wie in Lemma 3.4.6 werden wir zunächst die Fälle $m = 1, 2$ behandeln. Dazu definieren wir wie in Gleichung (3.4.29) für $\vartheta, \vartheta' \in \overline{B_r(\vartheta_0)}$:

$$U_{ki}^{(\vartheta', \vartheta)} = \int_I \left[\Lambda_{\vartheta'}^{(k)} - \mathbb{E} \Lambda_{\vartheta'} \right] \mathbb{E} \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta} \, dM^{(i)}, \quad k, i \in \{1, \dots, n\}. \quad (3.4.46)$$

Definieren wir ferner

$$V_{ki}^{(1)} = U_{ki}^{(\vartheta_0, \vartheta)} - U_{ki}^{(\vartheta_0, \vartheta')}, \quad V_{ki}^{(2)} = U_{ki}^{(\vartheta, \vartheta)} - U_{ki}^{(\vartheta', \vartheta')},$$

dann gilt wie zuvor:

$$\begin{aligned} \alpha_n^{(1)}(\vartheta) - \alpha_n^{(1)}(\vartheta') &= n^{-\frac{3}{2}} \sum_{i,k=1}^n \left(U_{ki}^{(\vartheta_0, \vartheta)} - U_{ki}^{(\vartheta_0, \vartheta')} \right) = n^{-\frac{3}{2}} \sum_{i,k=1}^n V_{ki}^{(1)}, \\ \alpha_n^{(2)}(\vartheta) - \alpha_n^{(2)}(\vartheta') &= n^{-\frac{3}{2}} \sum_{i,k=1}^n \left(U_{ki}^{(\vartheta, \vartheta)} - U_{ki}^{(\vartheta', \vartheta')} \right) = n^{-\frac{3}{2}} \sum_{i,k=1}^n V_{ki}^{(2)}. \end{aligned}$$

⁵³Man erhält *identische* Summanden, wenn die Summe weiterhin in zwei Untersummen mit $p = q$ beziehungsweise $p \neq q$ unterteilt werden.

⁵⁴Der Beweis wird in Stute und Kopperschmidt 2013 übersprungen, folgt er doch nahezu unmittelbar aus dem vorherigen Beweis. Die wichtigsten Argumente sollen hier dennoch aufgezeigt werden.

Wir können die Summen über $k = i$ und $k \neq i$ separat betrachten, denn es gilt:

$$\left\| n^{-\frac{3}{2}} \sum_{i,k=1}^n V_{ki}^{(m)} \right\|^2 \leq 2 \left[\left\| n^{-\frac{3}{2}} \sum_{i \neq k} V_{ki}^{(m)} \right\|^2 + \left\| n^{-\frac{3}{2}} \sum_{i=k} V_{ki}^{(m)} \right\|^2 \right], \quad m = 1, 2. \quad (3.4.47)$$

Wir beginnen mit der Summe über $i = k$. Da die Summanden unabhängig identisch verteilt und zentriert sind, folgt unmittelbar:

$$\mathbb{E} \left\| n^{-\frac{3}{2}} \sum_{i=1}^n V_{ii}^{(m)} \right\|^2 = \frac{1}{n^3} \sum_{i=1}^n \mathbb{E} \|V_{ii}^{(m)}\|^2 = \frac{1}{n^2} \mathbb{E} \|V_{11}^{(m)}\|^2 \leq \mathbb{E} \|V_{11}^{(m)}\|^2.$$

Nun gilt mit dem üblichen Vorgehen, da der Integrand ein vorhersagbarer Prozess ist:

$$\begin{aligned} \mathbb{E} \|V_{11}^{(1)}\|^2 &= \mathbb{E} \left\| U_{11}^{(\vartheta_0, \vartheta)} - U_{11}^{(\vartheta_0, \vartheta')} \right\|^2 \\ &= \mathbb{E} \left\| \int_I \left[\left(\Lambda_{\vartheta_0}^{(1)} - \mathbb{E} \Lambda_{\vartheta_0} \right) \mathbb{E} \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta} - \left(\Lambda_{\vartheta_0}^{(1)} - \mathbb{E} \Lambda_{\vartheta_0} \right) \mathbb{E} \frac{\partial}{\partial \vartheta'} \Lambda_{\vartheta'} \right] dM^{(1)} \right\|^2 \\ &= \mathbb{E} \left[\int_I \left\| \left(\Lambda_{\vartheta_0}^{(1)} - \mathbb{E} \Lambda_{\vartheta_0} \right) \mathbb{E} \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta} - \left(\Lambda_{\vartheta_0}^{(1)} - \mathbb{E} \Lambda_{\vartheta_0} \right) \mathbb{E} \frac{\partial}{\partial \vartheta'} \Lambda_{\vartheta'} \right\|^2 d\Lambda_{\vartheta_0}^{(1)} \right], \end{aligned}$$

und eine Taylor-Approximation um ϑ' sowie die Verwendung der kombinierten Regularitätsbedingung (3.4.26) liefern alsdann:

$$\begin{aligned} &\mathbb{E} \left[\int_I \left\| \left(\Lambda_{\vartheta_0}^{(1)} - \mathbb{E} \Lambda_{\vartheta_0} \right) \mathbb{E} \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta} - \left(\Lambda_{\vartheta_0}^{(1)} - \mathbb{E} \Lambda_{\vartheta_0} \right) \mathbb{E} \frac{\partial}{\partial \vartheta'} \Lambda_{\vartheta'} \right\|^2 d\Lambda_{\vartheta_0}^{(1)} \right] \\ &\leq \mathbb{E} \left[\|\vartheta - \vartheta'\|^2 \int_I \sup_{\vartheta \in \mathbb{B}_r(\vartheta_0)} \left\| \frac{\partial}{\partial \vartheta} \left[\left(\Lambda_{\vartheta_0}^{(1)} - \mathbb{E} \Lambda_{\vartheta_0} \right) \mathbb{E} \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta}^T - \left(\Lambda_{\vartheta_0}^{(1)} - \mathbb{E} \Lambda_{\vartheta_0} \right) \mathbb{E} \frac{\partial}{\partial \vartheta'} \Lambda_{\vartheta'}^T \right] \right\|^2 d\Lambda_{\vartheta_0}^{(1)} \right] \\ &\leq \|\vartheta - \vartheta'\|^2 \mathbb{E} \left[\sup_{I \times \mathbb{B}_r(\vartheta_0)} \left\{ \left\| \frac{\partial}{\partial \vartheta} \left[\left(\Lambda_{\vartheta_0}^{(1)} - \mathbb{E} \Lambda_{\vartheta_0} \right) \mathbb{E} \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta}^T - \left(\Lambda_{\vartheta_0}^{(1)} - \mathbb{E} \Lambda_{\vartheta_0} \right) \mathbb{E} \frac{\partial}{\partial \vartheta'} \Lambda_{\vartheta'}^T \right] \right\|^2 \Lambda_{\vartheta_0}^{(1)}(t^0) \right\} \right] \\ &= \nu_1^{(=)} \|\vartheta - \vartheta'\|^2, \end{aligned}$$

indem wiederholt die Dreiecksungleichung angewandt wird.

Die gleichen Argumente finden auch für $m = 2$ Anwendung, so dass wir gleichermaßen folgern können:

$$\mathbb{E} \|V_{11}^{(2)}\|^2 \leq \nu_2^{(=)} \|\vartheta - \vartheta'\|^2.$$

Für die Summe über $i \neq k$ liefert die Verwendung von Lemma A.1.2:

$$\mathbb{E} \left\| n^{-\frac{3}{2}} \sum_{i \neq k} V_{ki}^{(m)} \right\|^2 \leq \frac{2}{n^3} \sum_{i \neq k} \mathbb{E} \|V_{ki}^{(m)}\|^2 = 2 \frac{n^2 - n}{n^3} \mathbb{E} \|V_{12}^{(m)}\|^2 \leq \mathbb{E} \|V_{12}^{(m)}\|^2, \quad (3.4.48)$$

wobei die letzte Ungleichung induktiv gezeigt werden kann. Die Voraussetzungen aus (A.1.6) sind dabei erfüllt, da die $V_{ki}^{(m)}$ als Differenz der Zufallsvariablen U_{ki} geschrieben werden können, welche die Bedingung gemäß Lemma 3.4.6 erfüllen. Wir illustrieren das Argument kurz im Falle $m = 1$; hier gilt:

$$\begin{aligned} \mathbb{E} \left[V_{pk}^{(1)T} V_{ql}^{(1)} \right] &= \mathbb{E} \left[\left(U_{pk}^{(\vartheta_0, \vartheta)} - U_{pk}^{(\vartheta_0, \vartheta')} \right)^T \left(U_{ql}^{(\vartheta_0, \vartheta)} - U_{ql}^{(\vartheta_0, \vartheta')} \right) \right] \\ &= \mathbb{E} \left[U_{pk}^{(\vartheta_0, \vartheta)T} U_{ql}^{(\vartheta_0, \vartheta)} \right] + \mathbb{E} \left[U_{pk}^{(\vartheta_0, \vartheta)T} U_{ql}^{(\vartheta_0, \vartheta')} \right] \\ &\quad + \mathbb{E} \left[U_{pk}^{(\vartheta_0, \vartheta')T} U_{ql}^{(\vartheta_0, \vartheta)} \right] + \mathbb{E} \left[U_{pk}^{(\vartheta_0, \vartheta')T} U_{ql}^{(\vartheta_0, \vartheta')} \right], \end{aligned} \quad (3.4.49)$$

wobei jeder dieser Summanden verschwindet, sobald sich ein Index von den anderen unterscheidet (dies wurde ausführlich in Lemma 3.4.6 diskutiert). Folglich ist die Bedingung auch für die $V_{ki}^{(m)}$ erfüllt.

Nun kann das obige Vorgehen wiederholt werden, um weiterhin

$$\mathbb{E} \|V_{12}^{(m)}\|^2 \leq \nu_m^{(\neq)} \|\vartheta - \vartheta'\|^2, \quad m = 1, 2,$$

mit geeigneten Konstanten $\nu_m^{(\neq)}$ folgern zu können. Aus der Abschätzung (3.4.47) folgt dann schließlich

$$\mathbb{E} \|\alpha_n^{(m)}(\vartheta) - \alpha_n^{(m)}(\vartheta')\|^2 \leq 2 \left(\nu_m^{(=)} + \nu_m^{(\neq)} \right) \|\vartheta - \vartheta'\|^2 = \nu_m \|\vartheta - \vartheta'\|^2, \quad \text{für } m = 1, 2,$$

und damit die Bedingung aus (3.4.45).

Für den Fall $m = 3$ verfahren wir ähnlich, indem wir wiederum definieren:

$$\begin{aligned} U_{pki}^{(\vartheta)} &= \int_I \left[\Lambda_{\vartheta}^{(p)} - \Lambda_{\vartheta_0}^{(p)} \right] \left[\frac{\partial}{\partial \vartheta} \Lambda_{\vartheta}^{(k)} - \mathbb{E} \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta} \right] dM^{(i)}, \\ V_{pki} &= U_{pki}^{(\vartheta)} - U_{pki}^{(\vartheta')}. \end{aligned}$$

Analog zu den obigen Fällen folgt dann unter Verwendung von Lemma A.1.3 auch hier

$$\mathbb{E} \|\alpha_n^{(3)}(\vartheta) - \alpha_n^{(3)}(\vartheta')\|^2 \leq 2 \left(\nu_3^{(=)} + \nu_3^{(\neq)} \right) \|\vartheta - \vartheta'\|^2 = \nu_3 \|\vartheta - \vartheta'\|^2,$$

wobei ebenfalls die kombinierte Regularitätsbedingung aus (3.4.26) angewandt wird.

Ergo ist auch die Folge $\left(\alpha_n^{(3)}(\cdot) \Big|_{\overline{B_r(\vartheta_0)}} \right)_{n \in \mathbb{N}}$ straff und somit gilt die Behauptung. \square

Wir haben nun die Aussagen bewiesen, die gemäß der Bemerkungen im Vorfeld von Lemma 3.4.5 genügen, um das zentrale Resultat dieses Abschnittes, welches in Lemma 3.4.4 formuliert wurde, zu zeigen. Der Beweis wird zwar in Stute und Kopperschmidt 2013 ausgelassen,

soll hier jedoch ausführlich demonstriert werden. Dies ermöglicht dem Leser schließlich die Erkenntnis, wie sich die Puzzleteile der vorangegangenen Lemmata zu einem konsistenten Gesamtbild anordnen lassen.

Beweis von Lemma 3.4.4⁵⁵:

In Gleichung (3.4.18) haben wir bereits gezeigt, dass wir den parametrischen Prozess α_n aufspalten können in die Summe

$$\begin{aligned}\alpha_n(\vartheta) &= \tilde{\alpha}_n(\vartheta) - \alpha_n^{(1)}(\vartheta) + \alpha_n^{(2)}(\vartheta) + \alpha_n^{(3)}(\vartheta) \\ &= \sqrt{n} \int_I [\mathbb{E}\Lambda_\vartheta - \mathbb{E}\Lambda_{\vartheta_0}] \mathbb{E} \frac{\partial}{\partial \vartheta} \Lambda_\vartheta d\overline{M}^{(n)} - \alpha_n^{(1)}(\vartheta) + \alpha_n^{(2)}(\vartheta) + \alpha_n^{(3)}(\vartheta),\end{aligned}$$

wobei der führende (Integral-)Term gemäß Lemma 3.4.5 für alle kompakten Teilmengen $K \subset B_\varepsilon(\vartheta_0)$ mit $\vartheta_0 \in K$ straff in $C(K)$ ist.

Es verbleibt also lediglich zu zeigen, dass gleichmäßig auf K die Konvergenz

$$-\alpha_n^{(1)} + \alpha_n^{(2)} + \alpha_n^{(3)} \xrightarrow{P} 0 \quad (n \rightarrow \infty)$$

gilt. Dafür genügt es offensichtlich für $m \in \{1, 2, 3\}$ individuell zu zeigen, dass gleichmäßig auf K gilt:

$$\alpha_n^{(m)} \xrightarrow{P} 0 \quad (n \rightarrow \infty).$$

Es sei also $m \in \{1, 2, 3\}$. Formal möchten wir beweisen, dass für beliebige $\varepsilon, \eta > 0$ ein $n_0 \in \mathbb{N}$ existiert mit

$$P \left(\left\{ \sup_{\vartheta \in K} |\alpha_n^{(m)}(\vartheta)| \geq \varepsilon \right\} \right) \leq \eta, \quad \text{für alle } n \geq n_0. \quad (3.4.50)$$

Für gegebene $\varepsilon, \eta > 0$ liefert zunächst die Straffheit aus Lemma 3.4.7 mit Korollar 2.3.5, dass ein $\delta > 0$ und ein $\tilde{n}_0 \in \mathbb{N}$ existieren mit

$$P \left(\left\{ \sup_{\|\vartheta - \vartheta'\| < \delta} |\alpha_n^{(m)}(\vartheta) - \alpha_n^{(m)}(\vartheta')| \geq \frac{\varepsilon}{2} \right\} \right) \leq \frac{\eta}{2}, \quad \text{für alle } n \geq \tilde{n}_0. \quad (3.4.51)$$

Nun gilt offensichtlich

$$K \subset \bigcup_{\vartheta \in K} (B_\delta(\vartheta) \cap K), \quad (3.4.52)$$

und da K kompakt ist, gibt es eine endliche Teilüberdeckung dieser offenen Überdeckung.

⁵⁵Dieser Beweis wurde in Stute und Kopperschmidt 2013 ausgelassen und vom Autor dieser Arbeit selbst geführt.

Es existieren also $\vartheta^{(1)}, \dots, \vartheta^{(l)} \in K$ mit

$$K \subset B_\delta(\vartheta^{(1)}) \cup \dots \cup B_\delta(\vartheta^{(l)}).$$

Gemäß Lemma 3.4.6 gibt es $n_1, \dots, n_l \in \mathbb{N}$, so dass

$$P\left(\left\{\left|\alpha_n^{(m)}(\vartheta_i)\right| \geq \frac{\varepsilon}{2}\right\}\right) \leq \frac{\eta}{2}, \quad \text{für alle } n \geq n_i, i \in \{1, \dots, l\} \quad (3.4.53)$$

ist. Wegen $\alpha_n^{(m)}(\cdot) \in C(K)$ schließlich existiert nun zu $n \in \mathbb{N}$ ein (zufällig bestimmtes) $\vartheta_{(n)} \in K$ dergestalt⁵⁶, dass

$$\left|\alpha_n^{(m)}(\vartheta_{(n)})\right| = \sup_{\vartheta \in K} \left|\alpha_n^{(m)}(\vartheta)\right|$$

gilt und wegen (3.4.52) ein (ebenfalls zufällig bestimmtes) $j(n) \in \{1, \dots, l\}$ mit

$$\vartheta_{(n)} \in B_\delta(\vartheta^{(j(n))}).$$

Damit schließlich erhalten wir aufgrund der Abschätzungen (3.4.51) und (3.4.53) für $n \geq n_0 = \max\{\tilde{n}_0, n_1, \dots, n_l\}$:

$$\begin{aligned} P\left(\left\{\sup_{\vartheta \in K} \left|\alpha_n^{(m)}(\vartheta)\right| \geq \varepsilon\right\}\right) &= P\left(\left\{\omega \in \Omega : \left|\alpha_n^{(m)}(\omega)(\vartheta_{(n)}(\omega))\right| \geq \varepsilon\right\}\right) \\ &\leq P\left(\left\{\omega \in \Omega : \left|\alpha_n^{(m)}(\omega)(\vartheta_{(n)}(\omega)) - \alpha_n^{(m)}(\omega)(\vartheta^{(j(n)(\omega))})\right| \right. \right. \\ &\quad \left. \left. + \left|\alpha_n^{(m)}(\omega)(\vartheta^{(j(n)(\omega))})\right| \geq \varepsilon\right\}\right) \\ &\leq P\left(\left\{\omega \in \Omega : \left|\alpha_n^{(m)}(\omega)(\vartheta_{(n)}(\omega)) - \alpha_n^{(m)}(\omega)(\vartheta^{(j(n)(\omega))})\right| \geq \frac{\varepsilon}{2}\right\} \right. \\ &\quad \left. \cup \left\{\omega \in \Omega : \left|\alpha_n^{(m)}(\omega)(\vartheta^{(j(n)(\omega))})\right| \geq \frac{\varepsilon}{2}\right\}\right) \\ &\leq P\left(\left\{\sup_{\|\vartheta - \vartheta'\| < \delta} \left|\alpha_n^{(m)}(\vartheta) - \alpha_n^{(m)}(\vartheta')\right| \geq \frac{\varepsilon}{2}\right\}\right) \\ &\quad + P\left(\left\{\omega \in \Omega : \left|\alpha_n^{(m)}(\omega)(\vartheta^{(j(n)(\omega))})\right| \geq \frac{\varepsilon}{2}\right\}\right) \\ &\leq \frac{\eta}{2} + \frac{\eta}{2} = \eta, \end{aligned} \quad (3.4.54)$$

und somit ist die Bedingung aus (3.4.50) erfüllt. Hieraus folgt letztlich die Behauptung aus Lemma 3.4.4. \square

⁵⁶Dieses $\vartheta_{(n)}$ ist nicht mit dem Minimum-Distanz-Schätzer ϑ_n zu verwechseln!

Darstellungssatz für die parametrischen Prozesse $\{\beta_n\}$

In Bemerkung 3.4.2 haben wir neben den parametrischen Prozessen α_n auch die Prozesse β_n und γ_n eingeführt. Aus Gleichung (3.4.10) wird der Zusammenhang zwischen diesen Prozessen ersichtlich, der uns letztlich den Beweis der asymptotischen Normalverteiltheit des Schätzers ϑ_n ermöglicht. Dazu ist es notwendig, auch die verbleibenden parametrischen Prozesse zu studieren. Wir beginnen mit dem Prozess

$$\beta_n(\vartheta) = \sqrt{n} \int_I \overline{M}^{(n)} \frac{\partial}{\partial \vartheta} \overline{\Lambda}_\vartheta^{(n)} d\overline{M}^{(n)}.$$

Lemma 3.4.8. (*Darstellungssatz für die Folge $(\beta_n(\cdot)|_K)_{n \in \mathbb{N}}$* , siehe Stute und Kopperschmidt 2013, S. 1291)

Es gelten die üblichen Identifizierbarkeits- und Glattheitsbedingungen. Unter der Regularitätsbedingung aus (3.4.26) existiert dann ein $\varepsilon > 0$, so dass die Folge $(\beta_n(\cdot)|_K)_{n \in \mathbb{N}}$ für alle kompakten Teilmengen $K \subset B_\varepsilon(\vartheta_0)$ mit $\vartheta_0 \in K$ straff in $C(K)$ ist.

Ferner gilt für alle $\vartheta \in K$:

$$\beta_n(\vartheta) \xrightarrow{P} 0 \quad (n \rightarrow \infty),$$

und somit

$$\beta_n \xrightarrow{P} 0 \quad \text{gleichmäßig auf } K.$$

Beweis ⁵⁷:

Wir verwenden die selben Methoden, die sich schon für die parametrischen Prozesse α_n als zielführend erwiesen haben. Das Vorgehen spiegelt infolgedessen die Beweise der Lemmata 3.4.6 und 3.4.7 wider.

Erneut betrachten wir nur die Kompakta $\overline{B}_r(\vartheta_0)$ für $0 < r < \varepsilon$.

Definieren wir hier wiederum vektorwertige Zufallsvariablen U_{pki} via

$$U_{pki} = \int_I M^{(k)} \frac{\partial}{\partial \vartheta} \Lambda_\vartheta^{(p)} dM^{(i)}, \quad (3.4.55)$$

so erhalten wir definitionsgemäß:

$$\begin{aligned} \beta_n(\vartheta) &= \sqrt{n} \int_I \overline{M}^{(n)} \frac{\partial}{\partial \vartheta} \overline{\Lambda}_\vartheta^{(n)} d\overline{M}^{(n)} \\ &= n^{-\frac{5}{2}} \sum_{p=1}^n \sum_{k=1}^n \sum_{i=1}^n \int_I M^{(k)} \frac{\partial}{\partial \vartheta} \Lambda_\vartheta^{(p)} dM^{(i)} = n^{-\frac{5}{2}} \sum_{p,k,i=1}^n U_{pki}. \end{aligned}$$

⁵⁷Der Beweis orientiert sich an Stute und Kopperschmidt 2013, S. 1292 und wurde vom Autor dieser Arbeit geeignet ergänzt.

Die Existenz der zweiten Momente der U_{pki} folgt aus der kombinierten Regularitätsbedingung aus (3.4.26), wie wir kurz anhand des Falles $p = k = i$ demonstrieren möchten.

Es gilt dabei nach Voraussetzung, indem man $M^{(1)} = N^{(1)} - \Lambda_{\vartheta_0}^{(1)}$ schreibt und sodann die Abschätzung $\|M^{(1)}\|^2 \leq 2 \left[\|N^{(1)}\|^2 + \|\Lambda_{\vartheta_0}^{(1)}\|^2 \right]$ verwendet:

$$\begin{aligned} \mathbb{E} \|U_{111}\|^2 &= \mathbb{E} \left\| \int_I M^{(1)} \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta}^{(1)} dM^{(1)} \right\|^2 \leq 2 \mathbb{E} \left[\sup_{I \times \overline{B_r(\vartheta_0)}} \left\{ \left\| M^{(1)} \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta}^{(1)} \right\|^2 \left(N^{(1)}(t^0)^2 + \Lambda_{\vartheta_0}^{(1)}(t^0)^2 \right) \right\} \right] \\ &= 2 \mathbb{E} \left[\sup_{I \times \overline{B_r(\vartheta_0)}} \left\{ \left(N^{(1)}(t^0)^2 + \Lambda_{\vartheta_0}^{(1)}(t^0)^2 \right) \left\| \left(N^{(1)} - \Lambda_{\vartheta_0}^{(1)} \right) \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta}^{(1)} \right\|^2 \right\} \right] \\ &\leq 4 \mathbb{E} \left[\sup_{I \times \overline{B_r(\vartheta_0)}} \left\{ \left(N^{(1)}(t^0)^2 + \Lambda_{\vartheta_0}^{(1)}(t^0)^2 \right)^2 \left\| \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta}^{(1)} \right\|^2 \right\} \right] \\ &\leq 8 \mathbb{E} \left[\sup_{I \times \overline{B_r(\vartheta_0)}} \left\{ N^{(1)}(t^0)^4 \left\| \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta}^{(1)} \right\|^2 \right\} \right] + 8 \mathbb{E} \left[\sup_{I \times \overline{B_r(\vartheta_0)}} \left\{ \Lambda_{\vartheta_0}^{(1)}(t^0)^4 \left\| \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta}^{(1)} \right\|^2 \right\} \right] < \infty, \end{aligned}$$

wobei die Regularitätsbedingung (3.4.11) hier mit $i = 4, j = 0, k = 1, l = 2$ bzw. $i = 0, j = 4, k = 1, l = 2$ angewandt wurde.

Der Faktor $n^{-\frac{5}{2}}$ stellt nun wie zuvor sicher, dass für die punktweise stochastische Konvergenz gegen 0 nur die Untersumme über $p \neq k \neq i \neq p$ betrachtet werden muss, vergleiche etwa die Diskussion im Anschluss an Gleichung (3.4.40).

Um diese Untersumme zu beschränken, soll erneut das Lemma A.1.3 herangezogen werden. Die Bedingung aus (A.1.8) ist erfüllt, wobei zu beachten ist, dass die Indizes p, k und i in (3.4.55) nicht willkürlich zugewiesen wurden und daher auch *nicht* vertauscht werden dürfen.

Wir erhalten sodann

$$\mathbb{E} \left\| n^{-\frac{5}{2}} \sum_{p,k,i}^n U_{pki} \right\|^2 \leq \frac{32}{n^5} \sum \mathbb{E}(U_{pki}^T U_{qki}), \quad (3.4.56)$$

wobei die Summation über $\mathcal{O}(n^4)$ endliche und von n unabhängige Summanden erfolgt. Die *Chebyshev-Markovsche Ungleichung* liefert daher schließlich für alle $\vartheta \in \overline{B_r(\vartheta_0)}$

$$\beta_n(\vartheta) \xrightarrow{P} 0 \quad (n \rightarrow \infty)$$

und somit die punktweise stochastische Konvergenz gegen 0.

Wenden wir im Folgenden also erneut das *Straffheitskriterium von Kolmogorov* an, so müssen wir lediglich die Existenz einer Konstanten $\nu > 0$ mit

$$\mathbb{E} \|\beta_n(\vartheta) - \beta_n(\vartheta')\|^2 \leq \nu \|\vartheta - \vartheta'\|^2, \quad \text{für alle } n \in \mathbb{N} \text{ und } \vartheta, \vartheta' \in \overline{B_r(\vartheta_0)}$$

nachweisen. Der Beweis läuft hier völlig analog zum Beweis von Lemma 3.4.7; es genügt also vollkommen, die gemischten Erwartungswerte für gewisse $p, q, k, l, i, j \in \{1, \dots, n\}$ zu beschränken. Genauer möchten wir zeigen, dass

$$\begin{aligned} & \mathbb{E} \left[\int_I \int_I M^{(k)}(t) M^{(l)}(s) \left[\frac{\partial}{\partial \vartheta} \Lambda_{\vartheta}^{(p)}(t)^T - \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta'}^{(p)}(t)^T \right] \left[\frac{\partial}{\partial \vartheta} \Lambda_{\vartheta}^{(q)}(s) - \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta'}^{(q)}(s) \right] dM^{(i)}(t) dM^{(j)}(s) \right] \\ & \leq \nu \|\vartheta - \vartheta'\|^2 \end{aligned} \quad (3.4.57)$$

gilt. Dazu beachten wir, dass mit der *Cauchy-Schwarzschen Ungleichung* gilt:

$$\begin{aligned} & \left[\frac{\partial}{\partial \vartheta} \Lambda_{\vartheta}^{(p)}(t)^T - \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta'}^{(p)}(t)^T \right] \left[\frac{\partial}{\partial \vartheta} \Lambda_{\vartheta}^{(q)}(s) - \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta'}^{(q)}(s) \right] \\ & \leq \left\| \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta}^{(p)}(t)^T - \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta'}^{(p)}(t)^T \right\| \cdot \left\| \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta}^{(q)}(s)^T - \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta'}^{(q)}(s)^T \right\|. \end{aligned}$$

Eine Taylor-Entwicklung um ϑ' in *beiden* Faktoren liefert damit den benötigten Faktor $\|\vartheta - \vartheta'\|^2$:

$$\begin{aligned} & \left\| \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta}^{(p)}(t)^T - \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta'}^{(p)}(t)^T \right\| \cdot \left\| \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta}^{(q)}(s)^T - \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta'}^{(q)}(s)^T \right\| \\ & \leq \sup_{\vartheta \in \overline{B}_r(\vartheta_0)} \left\{ \left\| \frac{\partial^2}{\partial \vartheta^2} \Lambda_{\vartheta}^{(p)}(t)^T \right\| \cdot \left\| \frac{\partial^2}{\partial \vartheta^2} \Lambda_{\vartheta}^{(q)}(s)^T \right\| \right\} \|\vartheta - \vartheta'\|^2. \end{aligned} \quad (3.4.58)$$

Einsetzen von (3.4.58) in (3.4.57) liefert dann die Behauptung, insofern das Integral existiert. Für die Anwendung der Dreiecksungleichung für Integrale spaltet man dafür die Martingale $M^{(i)}$ erneut in eine Summe aus $N^{(i)}$ und $\Lambda_{\vartheta_0}^{(1)}$ auf.

Dies demonstrieren wir anhand des Falles $i = j$ und $k = l$ aus (3.4.56), in welchem man etwa für $p = q$ erhält (andernfalls nutze man die Unabhängigkeit der Faktoren aus):

$$\begin{aligned} & \mathbb{E} \left[\int_I \int_I \left| M^{(k)}(t) M^{(k)}(s) \right| \sup_{\vartheta \in \overline{B}_r(\vartheta_0)} \left\{ \left\| \frac{\partial^2}{\partial \vartheta^2} \Lambda_{\vartheta}^{(p)}(t)^T \right\| \cdot \left\| \frac{\partial^2}{\partial \vartheta^2} \Lambda_{\vartheta}^{(q)}(s)^T \right\| \right\} \|\vartheta - \vartheta'\|^2 dN^{(i)}(t) dN^{(i)}(s) \right] \\ & \leq \|\vartheta - \vartheta'\|^2 \mathbb{E} \left[\sup_{t \in I} \left\{ M^{(k)}(t)^2 \right\} \cdot \sup_{(t, \vartheta) \in I \times \overline{B}_r(\vartheta_0)} \left\{ \left\| \frac{\partial^2}{\partial \vartheta^2} \Lambda_{\vartheta}^{(p)}(t)^T \right\|^2 \right\} \cdot N^{(i)}(t^0)^2 \right] \\ & \leq \|\vartheta - \vartheta'\|^2 \mathbb{E} \left[4 \left(N^{(k)}(t^0)^2 + \Lambda_{\vartheta_0}^{(k)}(t^0)^2 \right) \cdot N^{(i)}(t^0)^2 \cdot \sup_{I \times \overline{B}_r(\vartheta_0)} \left\{ \left\| \frac{\partial^2}{\partial \vartheta^2} \Lambda_{\vartheta}^{(p)}(t)^T \right\|^2 \right\} \right], \end{aligned}$$

woraus mit der kombinierten Regularitätsbedingung (3.4.26) die Behauptung folgt. Obgleich in diesem Falle $i \neq k \neq p$ gilt, haben wir für die obige Abschätzung *nicht* die Unabhängigkeit der Faktoren benötigt. Ergo liefert obiges Vorgehen exemplarisch auch in den verbleibenden Fällen ein gültiges Vorgehen, um die Existenz des Erwartungswertes

nachzuweisen.

Insgesamt sind folglich die Bedingungen aus dem *Straffheitskriterium von Kolmogorov* erfüllt. Die Straffheit der Folge $\left(\beta_n(\cdot)|_{\overline{B_r(\vartheta_0)}}\right)_{n \in \mathbb{N}}$ liefert gemeinsam mit der punktweisen stochastischen Konvergenz die gleichmäßige stochastische Konvergenz gegen 0 auf $\overline{B_r(\vartheta_0)}$ und damit auf jeder kompakten Teilmenge $K \subset B_\varepsilon(\vartheta_0)$ mit $\vartheta_0 \in K$, siehe Beweis von Lemma 3.4.4. Dies vollendet den Beweis. \square

Darstellungssatz für die parametrischen Prozesse $\{\gamma_n\}$

In diesem Paragraphen untersuchen wir schließlich die parametrischen Prozesse

$$\gamma_n(\vartheta) = \sqrt{n} \int_I \overline{M}^{(n)} \frac{\partial}{\partial \vartheta} \overline{\Lambda}_{\vartheta}^{(n)} d\overline{\Lambda}_{\vartheta_0}^{(n)}.$$

Lemma 3.4.9. (*Darstellungssatz für die Folge $\left(\gamma_n(\cdot)|_K\right)_{n \in \mathbb{N}}$* , siehe Stute und Kopperschmidt 2013, S. 1292)

Es gelten die üblichen Identifizierbarkeits- und Glattheitsbedingungen. Unter der Regularitätsbedingung aus (3.4.26) existiert dann ein $\varepsilon > 0$, so dass der Prozess γ_n auf $B_\varepsilon(\vartheta_0)$ die Darstellung

$$\gamma_n(\vartheta) = \sqrt{n} \int_I \overline{M}^{(n)} E \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta} dE \Lambda_{\vartheta_0} + R_2(\vartheta) \quad (3.4.59)$$

besitzt, wobei der führende Term für alle kompakten Teilmengen $K \subset B_\varepsilon(\vartheta_0)$ mit $\vartheta_0 \in K$ straff in $C(K)$ ist. Ferner gilt gleichmäßig auf K die Konvergenz

$$R_2(\vartheta) \xrightarrow{P} 0 \quad (n \rightarrow \infty).$$

Beweis⁵⁸:

Der Beweis wird sich der Methoden der vorherigen Lemmata bedienen. Analog zum Beweis von Lemma 3.4.4 definieren wir auch hier geeignete Hilfsprozesse:

$$\begin{aligned} \tilde{\gamma}_n(\vartheta) &= \sqrt{n} \int_I \overline{M}^{(n)} E \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta} dE \Lambda_{\vartheta_0}, \\ \gamma_n^{(1)}(\vartheta) &= \sqrt{n} \int_I \overline{M}^{(n)} \left[\frac{\partial}{\partial \vartheta} \overline{\Lambda}_{\vartheta}^{(n)} - E \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta} \right] dE \Lambda_{\vartheta_0}, \\ \gamma_n^{(2)}(\vartheta) &= \sqrt{n} \int_I \overline{M}^{(n)} \frac{\partial}{\partial \vartheta} \overline{\Lambda}_{\vartheta}^{(n)} d \left[\overline{\Lambda}_{\vartheta_0}^{(n)} - E \Lambda_{\vartheta_0} \right]. \end{aligned}$$

⁵⁸Der Beweis orientiert sich an Stute und Kopperschmidt 2013, S. 1293 und wurde vom Autor dieser Arbeit geeignet ergänzt.

Dann gilt offensichtlich

$$\gamma_n(\vartheta) = \tilde{\gamma}_n(\vartheta) + \gamma_n^{(1)}(\vartheta) + \gamma_n^{(2)}(\vartheta),$$

und daher genügt es für den Beweis dieses Lemmas, die nachfolgenden Aussagen zu zeigen:

- Für jedes $0 < r < \varepsilon$ sind die Folgen $\left(\tilde{\gamma}_n(\cdot)|_{\overline{B_r(\vartheta_0)}}\right)_{n \in \mathbb{N}}$ und $\left(\gamma_n^{(m)}(\cdot)|_{\overline{B_r(\vartheta_0)}}\right)_{n \in \mathbb{N}}$ für $m = 1, 2$ straff in $C(\overline{B_r(\vartheta)})$.
- Für jedes feste $\vartheta \in B_\varepsilon(\vartheta_0)$ gilt:

$$\gamma_n^{(m)}(\vartheta) \xrightarrow{P} 0, \quad m = 1, 2.$$

Wir skizzieren kurz die wichtigsten Beweisschritte.

Zunächst gilt

$$\tilde{\gamma}^{(n)}(\vartheta) = n^{-\frac{1}{2}} \sum_{k=1}^n \int_I M^{(k)} \mathbb{E} \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta} d\mathbb{E} \Lambda_{\vartheta_0}.$$

Die Summation erfolgt über unabhängig identisch verteilte und zentrierte Zufallsvektoren, sodass aus dem *Zentralen Grenzwertsatz* für ein festes $\vartheta \in B_\varepsilon(\vartheta_0)$ und $n \rightarrow \infty$ die Konvergenz gegen einen Zufallsvektor $\tilde{\gamma}^\infty(\vartheta)$ folgt. Dieser Zufallsvektor induziert nun seinerseits ein Bildmaß auf \mathbb{R}^d , das als einzelnes Wahrscheinlichkeitsmaß über einem separablen und vollständigen metrischen Raum straff ist. Dies liefert bei Anwendung des *Straffheitskriteriums von Kolmogorov* die Gültigkeit der ersten Bedingung.

Die zweite Bedingung zeigt man wie zuvor, indem man ähnlich wie in Lemma 3.4.8 eine Taylor-Entwicklung um ϑ' für die Funktion

$$\mathbb{E} \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta} - \mathbb{E} \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta'}$$

ansetzt und die kombinierte Regularitätsbedingung aus (3.4.26) für die Existenz der Integrale ausnutzt. Hieraus folgt schließlich die Straffheit der Folge $\left(\tilde{\gamma}_n(\cdot)|_{\overline{B_r(\vartheta_0)}}\right)_{n \in \mathbb{N}}$ für $0 < r < \varepsilon$.

Wir wenden uns den Prozessen $\gamma_n^{(1)}$ und $\gamma_n^{(2)}$ zu und definieren dafür

$$U_{ki}^{(1)} = \int_I M^{(k)} \left[\frac{\partial}{\partial \vartheta} \Lambda_{\vartheta}^{(i)} - \mathbb{E} \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta} \right] d\mathbb{E} \Lambda_{\vartheta_0},$$

$$U_{pki}^{(2)} = \int_I M^{(k)} \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta}^{(p)} d \left[\Lambda_{\vartheta_0}^{(i)} - \mathbb{E} \Lambda_{\vartheta_0} \right],$$

so dass wir wie zuvor schreiben können:

$$\gamma_n^{(1)}(\vartheta) = n^{-\frac{3}{2}} \sum_{k,i=1}^n U_{ki}^{(1)}, \quad \gamma_n^{(2)}(\vartheta) = n^{-\frac{5}{2}} \sum_{p,k,i=1}^n U_{pki}^{(2)}.$$

Die Zufallsvektoren $U_{ki}^{(1)}$ und $U_{pki}^{(2)}$ sind dabei dergestalt, dass die Voraussetzungen des Lemmas A.1.2 beziehungsweise A.1.3 erfüllt sind; dies zeigt man wie in Lemma 3.4.6. Unter Anwendung der besagten Lemmata zeigt man nun zunächst für $m = 1, 2$ und alle $\vartheta \in B_\varepsilon(\vartheta_0)$ die stochastische Konvergenz von $\gamma^{(m)}(\vartheta)$ gegen 0 und folgert anschließend erneut mit dem *Straffheitskriterium von Kolmogorov* die Straffheit der Folgen $\left(\gamma_n^{(m)}(\cdot) \Big|_{\overline{B_r(\vartheta_0)}}\right)_{n \in \mathbb{N}}$. Abschließend kombiniert man die Ergebnisse wie im ausführlichen Beweis von Lemma 3.4.4, um das gewünschte Resultat zu erhalten. \square

3.4.2 Die asymptotische Normalverteiltheit von ϑ_n

Wir möchten schließlich das zentrale Resultat des Abschnittes 3.4 formulieren, welches zugleich das Hauptresultat des Artikels Stute und Kopperschmidt 2013 darstellt. Dazu möchten wir hier noch einmal sämtliche Voraussetzungen komprimiert darstellen.

In der Situation aus Abschnitt 3.1 gelten die nachfolgenden Forderungen:

Stetigkeit: Der Prozess

$$(t, \vartheta) \longmapsto \Lambda_\vartheta(t) \tag{3.2.3}$$

sei fast sicher stetig und besitze eine stetige Fortsetzung auf $I \times \overline{\Theta}$.

Identifizierbarkeit: Es gelte die Bedingung

$$\forall \varepsilon > 0 : \quad \inf_{\vartheta: \|\vartheta - \vartheta_0\| \geq \varepsilon} \|\mathbb{E} \Lambda_{\vartheta_0} - \mathbb{E} \Lambda_\vartheta\|_{\mathbb{E} \Lambda_{\vartheta_0}} > 0. \tag{3.3.31}$$

Glattheit: In einer Umgebung von ϑ_0 und für alle $t \in I$ sei der Prozess

$$(t, \vartheta) \longmapsto \Lambda_\vartheta(t) \tag{3.4.1}$$

bezüglich ϑ fast sicher zweimal stetig differenzierbar.

Regularität: In einer Umgebung U von ϑ_0 gelte die Bedingung

$$\mathbb{E} \left[\sup_{(t, \vartheta) \in I \times U} \left\{ N(t^0)^i \cdot \Lambda_\vartheta(t^0)^j \cdot \left\| \frac{\partial^k}{\partial \vartheta^k} \Lambda_\vartheta(t) \right\|^l \right\} \right] < \infty \tag{3.4.26}$$

für $i + j \leq 4$, $k \in \{1, 2\}$ und $l \leq 2$. Für $l = 0$ sei $U = \Theta$ zulässig.

Beispiel 2.1.7. (11. Fortsetzung)

Wir haben in den vorigen Fortsetzungen dieses Beispiels bereits gezeigt, dass ein reiner verschobener Geburtsprozess mit Abbruch, dessen Kompensator aus der Doob-Meyer-Zerlegung dem Modell \mathcal{H}_N aus (3.1.5) genügt, die obigen Stetigkeits- und Identifizierbarkeitsforderung erfüllt. Die Glattheit der Funktion

$$(t, \vartheta) \mapsto \Lambda_{\vartheta}(t) = \vartheta \int_{t_0}^t \frac{M}{M - N_s} \mathbb{1}_{\{N_s < M\}} ds$$

bei Differentiation bezüglich ϑ ist evident.

Ähnlich wie zuvor lässt sich der Erwartungswert aus (3.4.26) in diesem Falle beschränken durch

$$\vartheta_{\max}^4 [M(t^0 - t_0)]^6 < \infty, \quad (3.4.60)$$

weshalb der nachfolgende Satz 3.4.10 im Kontext unseres Beispiels anwendbar ist.

Wir erinnern an dieser Stelle die in Gleichung (3.4.14) aus Bemerkung 3.4.3 eingeführte Funktion

$$\Phi_0(\vartheta) = \frac{\partial}{\partial \vartheta} \int_I (\mathbb{E}\Lambda_{\vartheta} - \mathbb{E}\Lambda_{\vartheta_0}) \mathbb{E} \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta}^T d\mathbb{E}\Lambda_{\vartheta_0},$$

und definieren sodann:

$$\Phi_n(\vartheta) = \frac{\partial}{\partial \vartheta} \int_I [\bar{\Lambda}_{\vartheta}^{(n)} - \bar{\Lambda}_{\vartheta_0}^{(n)}] \frac{\partial}{\partial \vartheta} \bar{\Lambda}_{\vartheta}^{(n)T} d\bar{\Lambda}_{\vartheta_0}^{(n)}. \quad (3.4.61)$$

Außerdem sei an die in (3.4.13) definierte Funktion ϕ mit

$$\varphi(t) = \int_{[t, t^0]} \mathbb{E} \left[\frac{\partial}{\partial \vartheta} \Lambda_{\vartheta} \Big|_{\vartheta=\vartheta_0} \right] d\mathbb{E}\Lambda_{\vartheta_0}$$

erinnert, wobei wir $\varphi = (\varphi_1, \dots, \varphi_d)^T$ schreiben.

Unter den obigen Voraussetzungen gilt nun das nachfolgende Resultat:

Satz 3.4.10. (*Asymptotische Normalverteiltheit des Schätzers ϑ_n* , siehe Stute und Kopperschmidt 2013, S. 1281)

Für $n \rightarrow \infty$ gilt die Verteilungskonvergenz

$$\sqrt{n} \Phi_0(\vartheta_0)(\vartheta_n - \vartheta_0) \longrightarrow \mathcal{N}_d(0, \mathcal{G}(\vartheta_0)). \quad (3.4.62)$$

Dabei bezeichne $\mathcal{G}(\vartheta_0)$ die $d \times d$ -Matrix mit den Einträgen

$$\mathcal{G}_{ij}(\vartheta_0) = \int_I \varphi_i(t) \varphi_j(t) d\mathbb{E}\Lambda_{\vartheta_0}(t). \quad (3.4.63)$$

Bevor wir den Beweis von Satz 3.4.10 führen, möchten wir uns noch den im Vorfeld dieses Satzes definierten Prozessen Φ_n zuwenden, die wir als Approximation der Standardisierungsmatrix Φ_0 verstehen können. Dies wird Gegenstand des nächsten Unterabschnittes sein.

3.4.3 Asymptotik der approximativen Standardisierungsmatrix Φ_n

Wir beweisen zunächst den nachfolgenden Darstellungssatz für die parametrischen Prozesse Φ_n , der die Grundlage für das Korollar 3.4.12 bilden wird.

Lemma 3.4.11. (*Darstellungssatz für die Folge $(\Phi_n(\cdot)|_K)_{n \in \mathbb{N}}$* , siehe Stute und Kopperschmidt 2013, S. 1293)

Es gelten die üblichen Identifizierbarkeits- und Glattheitsbedingungen. Unter der Regularitätsbedingung aus (3.4.26) existiert dann ein $\varepsilon > 0$, so dass der Prozess Φ_n auf $B_\varepsilon(\vartheta_0)$ die Darstellung

$$\Phi_n(\vartheta) = \Phi_0(\vartheta) + R_3(\vartheta) \quad (3.4.64)$$

besitzt. Dabei gilt auf allen kompakten Teilmengen $K \subset B_\varepsilon(\vartheta_0)$ gleichmäßig die Konvergenz

$$R_3(\vartheta) \xrightarrow{P} 0 \quad (n \rightarrow \infty).$$

Beweis⁵⁹:

Nach Voraussetzung sind die $\bar{\Lambda}_\vartheta^{(n)}$ als Stichprobenmittelwerte der unabhängig identisch verteilten Prozesse $\Lambda_\vartheta^{(i)}$ definiert. Da diese Prozesse den Kompensator eines Zählprozesses modellieren sollen, sind sie gemäß Konstruktion monoton wachsend. Ähnlich wie im *Satz von Glivenko-Cantelli* möchten wir daraus folgern, dass fast sicher gleichmäßig auf I gilt:

$$\bar{\Lambda}_\vartheta^{(n)} \longrightarrow E\Lambda_\vartheta \quad (n \rightarrow \infty). \quad (3.4.65)$$

Das Schlüsselargument ist dabei der aus der Analysis bekannte *Satz von Dini*; wir werden den Beweis dieses Satzes dazu auf unsere Situation übertragen.

Zunächst beachten wir die Stetigkeit der deterministischen Funktion $E\Lambda_\vartheta$ und den Umstand, dass gemäß der Forderung (3.1.12) zudem

$$E[\Lambda_\vartheta(t^0)] < \infty$$

gilt. Wir können daher zu gegebenem $\eta > 0$ Zeitpunkte $t_0 < t_1 < \dots < t_m = t^0$ dergestalt

⁵⁹Der Beweis wurde vom Autor selbst geführt. In Stute und Kopperschmidt 2013 wird lediglich auf den *Satz von Glivenko-Cantelli* verwiesen.

wählen, dass

$$\left| \mathbb{E} [\Lambda_{\vartheta}(t_{i+1})] - \mathbb{E} [\Lambda_{\vartheta}(t_i)] \right| < \frac{\eta}{2}, \quad \text{für alle } i \in \{0, 1, \dots, m-1\} \quad (3.4.66)$$

ist. Aus dem starken Gesetz der großen Zahlen folgt sofort, dass für $i \in \{0, 1, \dots, m\}$ fast sicher die Konvergenz gilt:

$$\bar{\Lambda}_{\vartheta}^{(n)}(t_i) \longrightarrow \mathbb{E} [\Lambda_{\vartheta}(t_i)] \quad (n \rightarrow \infty).$$

Da endliche Vereinigungen von Nullmengen wiederum Nullmengen sind, gilt diese Konvergenz somit auch fast sicher gleichmäßig auf der Menge $\{t_0, t_1, \dots, t_m\}$. Wir finden also ein $n_0 \in \mathbb{N}$, so dass für alle $i \in \{0, 1, \dots, m\}$ fast sicher gilt:

$$\left| \bar{\Lambda}_{\vartheta}^{(n)}(t_i) - \mathbb{E} [\Lambda_{\vartheta}(t_i)] \right| < \frac{\eta}{2}, \quad \text{für alle } n \geq n_0. \quad (3.4.67)$$

Aus der Monotonie von $\bar{\Lambda}_{\vartheta}^{(n)}$ können wir folgern:

$$\bar{\Lambda}_{\vartheta}^{(n)}(t) \in \left[\bar{\Lambda}_{\vartheta}^{(n)}(t_i), \bar{\Lambda}_{\vartheta}^{(n)}(t_{i+1}) \right], \quad \text{für } t \in [t_i, t_{i+1}],$$

und aus (3.4.66) und (3.4.67) folgt daher mit der Monotonie von $\mathbb{E} \Lambda_{\vartheta}$ für alle $n \geq n_0$ fast sicher:

$$\bar{\Lambda}_{\vartheta}^{(n)}(t) \in \left(\mathbb{E} [\Lambda_{\vartheta}(t_i)] - \frac{\eta}{2}, \mathbb{E} [\Lambda_{\vartheta}(t_{i+1})] + \frac{\eta}{2} \right) \subset \left(\mathbb{E} [\Lambda_{\vartheta}(t)] - \eta, \mathbb{E} [\Lambda_{\vartheta}(t)] + \eta \right), \quad \text{für } t \in [t_i, t_{i+1}].$$

Für jene $n \geq n_0$ gilt deshalb ebenfalls fast sicher

$$\sup_{t \in I} \left| \bar{\Lambda}_{\vartheta}^{(n)}(t) - \mathbb{E} [\Lambda_{\vartheta}(t)] \right| \leq \max_{i \in \{0, 1, \dots, m-1\}} \sup_{t \in [t_i, t_{i+1}]} \left| \bar{\Lambda}_{\vartheta}^{(n)}(t) - \mathbb{E} [\Lambda_{\vartheta}(t)] \right| < \eta,$$

und damit die Behauptung aus (3.4.65).

Für ein $\nu > 0$ und $\delta > 0$ erhalten wir nun mit der *Markovschen Ungleichung*, da die $\Lambda_{\vartheta}^{(i)}$ unabhängig identisch verteilt sind:

$$\begin{aligned} P \left(\left\{ \sup_{\|\vartheta - \vartheta'\| < \delta} \sup_{t \in I} \left| \bar{\Lambda}_{\vartheta}^{(n)}(t) - \bar{\Lambda}_{\vartheta'}^{(n)}(t) \right| > \nu \right\} \right) &\leq P \left(\left\{ \frac{1}{n} \sum_{i=1}^n \sup_{\|\vartheta - \vartheta'\| < \delta} \sup_{t \in I} \left| \Lambda_{\vartheta}^{(i)}(t) - \Lambda_{\vartheta'}^{(i)}(t) \right| > \nu \right\} \right) \\ &\leq \frac{1}{\nu} \mathbb{E} \left[\frac{1}{n} \sum_{i=1}^n \sup_{\|\vartheta - \vartheta'\| < \delta} \sup_{t \in I} \left| \Lambda_{\vartheta}^{(i)}(t) - \Lambda_{\vartheta'}^{(i)}(t) \right| \right] \\ &\leq \frac{1}{\nu} \mathbb{E} \left[\sup_{\|\vartheta - \vartheta'\| < \delta} \sup_{t \in I} \left| \Lambda_{\vartheta}(t) - \Lambda_{\vartheta'}(t) \right| \right], \quad \text{für alle } n \in \mathbb{N}. \end{aligned} \quad (3.4.68)$$

Eine gleichmäßige Majorante für den Integranden aus (3.4.68) ist gemäß (3.1.12) gegeben durch

$$2 \sup_{\vartheta \in \Theta} \Lambda_{\vartheta}(t^0),$$

und somit folgt für $\delta \rightarrow 0$ aus der Stetigkeitsforderung (3.2.3) die Konvergenz gegen 0. Zu $\varepsilon > 0$ existiert also ein $\delta > 0$ mit

$$P \left(\left\{ \sup_{\|\vartheta - \vartheta'\| < \delta} \sup_{t \in I} |\bar{\Lambda}_{\vartheta}^{(n)}(t) - \bar{\Lambda}_{\vartheta'}^{(n)}(t)| > \frac{\nu}{3} \right\} \right) < \varepsilon, \quad \text{für alle } n \in \mathbb{N}. \quad (3.4.69)$$

Weiter folgern wir für ein gegebenes $\vartheta' \in B_{\varepsilon}(\vartheta_0)$:

$$\begin{aligned} P \left(\left\{ \sup_{\|\vartheta - \vartheta'\| < \delta} \sup_{t \in I} |\bar{\Lambda}_{\vartheta}^{(n)}(t) - \mathbb{E}[\Lambda_{\vartheta}(t)]| > \nu \right\} \right) &\leq P \left(\left\{ \sup_{\|\vartheta - \vartheta'\| < \delta} \sup_{t \in I} |\bar{\Lambda}_{\vartheta}^{(n)}(t) - \bar{\Lambda}_{\vartheta'}^{(n)}(t)| \right. \right. \\ &\quad \left. \left. + \sup_{t \in I} |\bar{\Lambda}_{\vartheta'}^{(n)}(t) - \mathbb{E}[\Lambda_{\vartheta'}(t)]| + \sup_{\|\vartheta - \vartheta'\| < \delta} \sup_{t \in I} |\mathbb{E}[\Lambda_{\vartheta}(t)] - \mathbb{E}[\Lambda_{\vartheta'}(t)]| > \nu \right\} \right) \\ &\leq P \left(\left\{ \sup_{\|\vartheta - \vartheta'\| < \delta} \sup_{t \in I} |\bar{\Lambda}_{\vartheta}^{(n)}(t) - \bar{\Lambda}_{\vartheta'}^{(n)}(t)| > \frac{\nu}{3} \right\} \right) + P \left(\left\{ \sup_{t \in I} |\bar{\Lambda}_{\vartheta'}^{(n)}(t) - \mathbb{E}[\Lambda_{\vartheta'}(t)]| > \frac{\nu}{3} \right\} \right) \\ &\quad + P \left(\left\{ \sup_{\|\vartheta - \vartheta'\| < \delta} \sup_{t \in I} |\mathbb{E}[\Lambda_{\vartheta}(t)] - \mathbb{E}[\Lambda_{\vartheta'}(t)]| > \frac{\nu}{3} \right\} \right). \end{aligned}$$

Für hinreichend kleines $\delta > 0$ können wir gemäß (3.4.69) den ersten Summanden beschränken; außerdem ist der letzte Summand für δ hinreichend klein gerade identisch 0. Der zweite Summand verschwindet gleichermaßen gemäß (3.4.65) für n hinreichend groß, so dass wir ein $\delta > 0$ und ein $n_1 \in \mathbb{N}$ finden können mit

$$P \left(\left\{ \sup_{\|\vartheta - \vartheta'\| < \delta} \sup_{t \in I} |\bar{\Lambda}_{\vartheta}^{(n)}(t) - \mathbb{E}[\Lambda_{\vartheta}(t)]| > \nu \right\} \right) < \varepsilon, \quad \text{für alle } n \geq n_1.$$

Mittels des *Überdeckungssatzes von Heine-Borel* können wir hieraus leicht folgern, dass für beliebige Kompakta $K \subset B_{\varepsilon}(\vartheta_0)$ gleichmäßig auf $I \times K$ die (stochastische) Konvergenz aus (3.4.65) gilt.

Ferner gilt dann gleichmäßig auf $I \times K$ sogar die stochastische Konvergenz

$$\frac{\partial^j}{\partial \vartheta^j} \bar{\Lambda}_{\vartheta}^{(n)} \xrightarrow{P} \mathbb{E} \frac{\partial^j}{\partial \vartheta^j} \Lambda_{\vartheta} \quad (n \rightarrow \infty), \quad \text{für } j \in \{1, 2\}, \quad (3.4.70)$$

da die Folgen $\left(\frac{\partial^j}{\partial \vartheta^j} \bar{\Lambda}_{\vartheta}^{(n)}(\cdot) \Big|_{I \times K} \right)_{n \in \mathbb{N}}$ straff in $C(I \times K)$ sind, wie man mit einem zu (3.4.68) analogen Argument unter Verwendung der kombinierten Regularitätsbedingung aus (3.4.26) einsieht. Anders als zuvor erhalten wir dabei zwar lediglich die gleichmäßige *stochastische*

Konvergenz des differenzierten Prozesses auf I , der obige Beweis lässt sich für diesen Fall jedoch völlig analog führen.

Wir beachten nun, dass die Glattheitsforderung aus (3.4.1) die Vertauschung von Differentiation und Integration in der Definition der Funktionen Φ_n aus (3.4.61) gestattet. Man erhält also auf jeder kompakten Teilmengen $K \subset B_\varepsilon(\vartheta_0)$:

$$\Phi_n(\vartheta) = \int_I \left[\bar{\Lambda}_\vartheta^{(n)} - \bar{\Lambda}_{\vartheta_0}^{(n)} \right] \frac{\partial^2}{\partial \vartheta^2} \bar{\Lambda}_\vartheta^{(n)T} d\bar{\Lambda}_{\vartheta_0}^{(n)} + \int_I \frac{\partial}{\partial \vartheta} \bar{\Lambda}_\vartheta^{(n)} \frac{\partial}{\partial \vartheta} \bar{\Lambda}_\vartheta^{(n)T} d\bar{\Lambda}_{\vartheta_0}^{(n)}, \quad \vartheta \in K.$$

Analog folgern wir aus der stetigen Differenzierbarkeit des Integranden der Funktion Φ_0 (vergleiche auch Stute und Kopperschmidt 2013, S. 1280 beziehungsweise Bemerkung 3.4.3):

$$\Phi_0(\vartheta) = \int_I [\mathbb{E}\Lambda_\vartheta - \mathbb{E}\Lambda_{\vartheta_0}] \mathbb{E} \frac{\partial^2}{\partial \vartheta^2} \Lambda_\vartheta^T d\mathbb{E}\Lambda_{\vartheta_0} + \int_I \mathbb{E} \frac{\partial}{\partial \vartheta} \Lambda_\vartheta \mathbb{E} \frac{\partial}{\partial \vartheta} \Lambda_\vartheta^T d\mathbb{E}\Lambda_{\vartheta_0}, \quad \vartheta \in K.$$

Um den Beweis dieses Lemmas zu vollenden, genügt es nun also zu zeigen, dass auf jeder kompakten Teilmenge $K \subset B_\varepsilon(\vartheta_0)$ gleichmäßig die Konvergenzen

$$\int_I \left[\bar{\Lambda}_\vartheta^{(n)} - \bar{\Lambda}_{\vartheta_0}^{(n)} \right] \frac{\partial^2}{\partial \vartheta^2} \bar{\Lambda}_\vartheta^{(n)T} d\bar{\Lambda}_{\vartheta_0}^{(n)} - \int_I [\mathbb{E}\Lambda_\vartheta - \mathbb{E}\Lambda_{\vartheta_0}] \mathbb{E} \frac{\partial^2}{\partial \vartheta^2} \Lambda_\vartheta^T d\mathbb{E}\Lambda_{\vartheta_0} \xrightarrow{P} 0 \quad (n \rightarrow \infty) \quad (3.4.71)$$

sowie

$$\int_I \frac{\partial}{\partial \vartheta} \bar{\Lambda}_\vartheta^{(n)} \frac{\partial}{\partial \vartheta} \bar{\Lambda}_\vartheta^{(n)T} d\bar{\Lambda}_{\vartheta_0}^{(n)} - \int_I \mathbb{E} \frac{\partial}{\partial \vartheta} \Lambda_\vartheta \mathbb{E} \frac{\partial}{\partial \vartheta} \Lambda_\vartheta^T d\mathbb{E}\Lambda_{\vartheta_0} \xrightarrow{P} 0 \quad (n \rightarrow \infty) \quad (3.4.72)$$

gelten. Wir bemerken zunächst, dass aus den Konvergenzen (3.4.65) und (3.4.70) bereits die gleichmäßige stochastische Konvergenz der Integranden folgt. Es genügt folglich, Integrale der Gestalt

$$\int_I f_\vartheta^{(n)} d\bar{\Lambda}_{\vartheta_0}^{(n)}$$

zu betrachten, wobei die Zufallsvektoren $f_\vartheta^{(n)}$ für beliebige Kompakta $K \subset B_\varepsilon(\vartheta_0)$ gleichmäßig auf $I \times K$ stochastisch gegen einen Zufallsvektor f_ϑ konvergieren.

Wie schon in der Abschätzung aus (3.4.68) folgern wir mit der *Markovschen Ungleichung* für $a > 0$:

$$P \left(\left\{ \bar{\Lambda}_{\vartheta_0}^{(n)}(t^0) > a \right\} \right) \leq \frac{1}{a} \mathbb{E} [\Lambda_{\vartheta_0}(t^0)] = \frac{const}{a}, \quad \text{für alle } n \in \mathbb{N}, \quad (3.4.73)$$

wobei die Konstante *nicht* von n abhängt.

Zu gegebenem $\varepsilon > 0$ können wir daher ein $a_0 > 0$ finden, so dass gilt:

$$P\left(\left\{\bar{\Lambda}_{\vartheta_0}^{(n)}(t^0) > a_0\right\}\right) < \frac{\varepsilon}{4}, \quad \text{für alle } n \in \mathbb{N}. \quad (3.4.74)$$

Aus der Definition der stochastischen Konvergenz erhalten wir für $\nu > 0$ zudem ein $n_2 \in \mathbb{N}$, so dass für ein Kompaktum K gilt:

$$P\left(\left\{\sup_{(t,\vartheta) \in I \times K} \|f_{\vartheta}^{(n)}(t) - f_{\vartheta}(t)\| > \frac{\nu}{2a_0}\right\}\right) < \frac{\varepsilon}{4}, \quad \text{für alle } n \geq n_2. \quad (3.4.75)$$

Sodann betrachten wir die folgende Abschätzung:

$$\begin{aligned} & P\left(\left\{\sup_{\vartheta \in K} \left\| \int_I f_{\vartheta}^{(n)} d\bar{\Lambda}_{\vartheta_0}^{(n)} - \int_I f_{\vartheta} dE\Lambda_{\vartheta_0} \right\| > \nu\right\}\right) \\ & \leq P\left(\left\{\sup_{\vartheta \in K} \left\| \int_I f_{\vartheta}^{(n)} d\bar{\Lambda}_{\vartheta_0}^{(n)} - \int_I f_{\vartheta} d\bar{\Lambda}_{\vartheta_0}^{(n)} \right\| + \sup_{\vartheta \in K} \left\| \int_I f_{\vartheta} d\bar{\Lambda}_{\vartheta_0}^{(n)} - \int_I f_{\vartheta} dE\Lambda_{\vartheta_0} \right\| > \nu\right\}\right) \\ & \leq P\left(\left\{\sup_{\vartheta \in K} \left\| \int_I f_{\vartheta}^{(n)} d\bar{\Lambda}_{\vartheta_0}^{(n)} - \int_I f_{\vartheta} d\bar{\Lambda}_{\vartheta_0}^{(n)} \right\| > \frac{\nu}{2}\right\}\right) \\ & \quad + P\left(\left\{\sup_{\vartheta \in K} \left\| \int_I f_{\vartheta} d\bar{\Lambda}_{\vartheta_0}^{(n)} - \int_I f_{\vartheta} dE\Lambda_{\vartheta_0} \right\| > \frac{\nu}{2}\right\}\right). \end{aligned} \quad (3.4.76)$$

Wir betrachten die Summanden aus (3.4.76) separat. Für den ersten Summanden erhalten wir mit (3.4.74) und (3.4.75):

$$\begin{aligned} & P\left(\left\{\sup_{\vartheta \in K} \left\| \int_I f_{\vartheta}^{(n)} d\bar{\Lambda}_{\vartheta_0}^{(n)} - \int_I f_{\vartheta} d\bar{\Lambda}_{\vartheta_0}^{(n)} \right\| > \frac{\nu}{2}\right\}\right) = P\left(\left\{\sup_{\vartheta \in K} \left\| \int_I (f_{\vartheta}^{(n)} - f_{\vartheta}) d\bar{\Lambda}_{\vartheta_0}^{(n)} \right\| > \frac{\nu}{2}\right\}\right) \\ & \leq P\left(\left\{\sup_{(t,\vartheta) \in I \times K} \|f_{\vartheta}^{(n)}(t) - f_{\vartheta}(t)\| \cdot \bar{\Lambda}_{\vartheta_0}^{(n)}(t^0) > \frac{\nu}{2}\right\}\right) \\ & = P\left(\left\{\sup_{(t,\vartheta) \in I \times K} \|f_{\vartheta}^{(n)}(t) - f_{\vartheta}(t)\| \cdot \bar{\Lambda}_{\vartheta_0}^{(n)}(t^0) > \frac{\nu}{2}\right\} \cap \left\{\bar{\Lambda}_{\vartheta_0}^{(n)}(t^0) \leq a_0\right\}\right) \\ & \quad + P\left(\left\{\sup_{(t,\vartheta) \in I \times K} \|f_{\vartheta}^{(n)}(t) - f_{\vartheta}(t)\| \cdot \bar{\Lambda}_{\vartheta_0}^{(n)}(t^0) > \frac{\nu}{2}\right\} \cap \left\{\bar{\Lambda}_{\vartheta_0}^{(n)}(t^0) > a_0\right\}\right) \\ & \leq P\left(\left\{\sup_{(t,\vartheta) \in I \times K} \|f_{\vartheta}^{(n)}(t) - f_{\vartheta}(t)\| > \frac{\nu}{2a_0}\right\}\right) + \frac{\varepsilon}{4} < \frac{\varepsilon}{2}, \quad \text{für alle } n \geq n_2. \end{aligned}$$

Um den zweiten Summanden zu beschränken, möchten wir den *Satz von Helly-Bray für Stieltjes-Integrale*, wie er etwa in Graves 1956, S. 282 als *Theorem 22* zu finden ist, anwenden.

Da der Integrand f_{ϑ} für alle $\vartheta \in B_{\varepsilon}(\vartheta_0)$ stetig bezüglich der Integrationsvariablen ist und gemäß (3.4.65) bereits die fast sichere gleichmäßige Konvergenz der Integratoren $\bar{\Lambda}_{\vartheta_0}^{(n)}$ gegen

$E\Lambda_{\vartheta_0}$ für $n \rightarrow \infty$ gewährleistet ist, müssen wir dazu lediglich eine *gleichmäßige* Schranke für die totale Variation der Integratoren voraussetzen. Hierfür verwenden wir erneut die Abschätzung aus (3.4.74) und erhalten mit $a_1 = \max\{a_0, E\Lambda_{\vartheta_0}(t^0)\}$:

$$\begin{aligned}
& P\left(\left\{\sup_{\vartheta \in K} \left\| \int_I f_{\vartheta} d\bar{\Lambda}_{\vartheta_0}^{(n)} - \int_I f_{\vartheta} dE\Lambda_{\vartheta_0} \right\| > \frac{\nu}{2}\right\}\right) \\
&= P\left(\left\{\sup_{\vartheta \in K} \left\| \int_I f_{\vartheta} d\bar{\Lambda}_{\vartheta_0}^{(n)} - \int_I f_{\vartheta} dE\Lambda_{\vartheta_0} \right\| > \frac{\nu}{2}\right\} \cap \left\{\bar{\Lambda}_{\vartheta_0}^{(n)} \leq a_1\right\}\right) \\
&\quad + P\left(\left\{\sup_{\vartheta \in K} \left\| \int_I f_{\vartheta} d\bar{\Lambda}_{\vartheta_0}^{(n)} - \int_I f_{\vartheta} dE\Lambda_{\vartheta_0} \right\| > \frac{\nu}{2}\right\} \cap \left\{\bar{\Lambda}_{\vartheta_0}^{(n)} > a_1\right\}\right) \\
&\leq P\left(\left\{\sup_{\vartheta \in K} \left\| \int_I f_{\vartheta} d\bar{\Lambda}_{\vartheta_0}^{(n)} - \int_I f_{\vartheta} dE\Lambda_{\vartheta_0} \right\| > \frac{\nu}{2}\right\} \cap \left\{\bar{\Lambda}_{\vartheta_0}^{(n)} \leq a_1\right\}\right) + \frac{\varepsilon}{4}. \quad (3.4.77)
\end{aligned}$$

Für fast alle $\omega \in \{\bar{\Lambda}_{\vartheta_0}^{(n)} \leq a_1\}$ gilt nun aufgrund des *Satzes von Helly-Bray für Stieltjes-Integrale* die Konvergenz

$$\sup_{\vartheta \in K} \left\| \int_I f_{\vartheta} d\bar{\Lambda}_{\vartheta_0}^{(n)} - \int_I f_{\vartheta} dE\Lambda_{\vartheta_0} \right\| \rightarrow 0 \quad (n \rightarrow \infty),$$

und daher existiert ein $n_3 \in \mathbb{N}$ mit

$$P\left(\left\{\sup_{\vartheta \in K} \left\| \int_I f_{\vartheta} d\bar{\Lambda}_{\vartheta_0}^{(n)} - \int_I f_{\vartheta} dE\Lambda_{\vartheta_0} \right\| > \frac{\nu}{2}\right\} \cap \left\{\bar{\Lambda}_{\vartheta_0}^{(n)} \leq a_1\right\}\right) < \frac{\varepsilon}{4}, \quad \text{für alle } n \geq n_3.$$

Gemäß der Abschätzung aus (3.4.77) können wir also auch den zweiten Summanden aus (3.4.76) durch $\frac{\varepsilon}{2}$ beschränken.

Insgesamt kann somit die Summe aus (3.4.76) für alle $n \geq \max\{n_2, n_3\}$ durch ε beschränkt werden, woraus die geforderte stochastische Konvergenz folgt. Somit erhalten wir schließlich die Behauptung aus (3.4.71) und (3.4.72), was den Beweis vollendet. \square

Die Relevanz des Lemmas 3.4.11 wird erst in Bemerkung 3.4.13 im Anschluss an das folgende Korollar ersichtlich.

Korollar 3.4.12.

Es gilt die Konvergenz:

$$\Phi_n(\vartheta_n) \xrightarrow{P} \Phi_0(\vartheta_0) \quad (n \rightarrow \infty). \quad (3.4.78)$$

Beweis⁶⁰:

Der Beweis erfolgt mit den üblichen Methoden und soll daher lediglich skizziert werden.

⁶⁰Die Beweisskizze wurde vom Autor dieser Arbeit angefertigt.

Aufgrund der Dreiecksungleichung genügt es, die Konvergenzen

$$|\Phi_n(\vartheta_n) - \Phi_0(\vartheta_n)| \xrightarrow{P} 0, \quad |\Phi_0(\vartheta_n) - \Phi_0(\vartheta_0)| \xrightarrow{P} 0 \quad (n \rightarrow \infty) \quad (3.4.79)$$

nachzuweisen. Die erste Konvergenz folgt sofort aus Lemma 3.4.11, da die fast sichere Konvergenz von ϑ_n gegen ϑ_0 sicherstellt, dass abseits einer P -Nullmenge für ein $0 < r < \varepsilon$ und hinreichend große $n \in \mathbb{N}$ stets $\vartheta_n \in \overline{B_r(\vartheta_0)}$ gilt.

Die zweite Konvergenz folgt leicht, da Φ_0 eine stetige Funktion des Parameters ϑ ist. Der Beweis läuft dann analog zu dem Vorgehen aus Lemma 2.3.6. \square

Bemerkung 3.4.13. (Vergleiche Stute und Kopperschmidt 2013, S. 1293)

Betrachten wir die Matrizen $\Phi_0(\vartheta_0)$ und $\mathcal{G}(\vartheta_0)$ aus Satz 3.4.10, so stellen wir fest, dass für die Ermittlung der Verteilung aus (3.4.62) zugleich die Kenntnis des unbekanntes und zu schätzenden Parameters ϑ_0 erforderlich ist.

Das *Lemma von Slutsky* impliziert nun mittels des obigen Korollars 3.4.12, dass wir in der Situation einer endlichen Stichprobenmenge, das heißt für endliches $n \in \mathbb{N}$, die Standardisierungsmatrix $\Phi_0(\vartheta_0)$ durch die Approximation $\Phi_n(\vartheta_n)$ ersetzen können, ohne die asymptotische Verteilung zu verändern.

Ersetzen wir in (3.4.61) zudem ϑ_0 durch den Schätzer ϑ_n , so lässt sich der Beweis von Lemma 3.4.11 ganz analog führen, weshalb wir die Standardisierungsmatrix $\Phi_0(\vartheta_0)$ in der Praxis durch den Ausdruck

$$\Phi^{(n)}(\vartheta_n) = \frac{\partial}{\partial \vartheta} \int_I \left[\overline{\Lambda}_{\vartheta}^{(n)} - \overline{\Lambda}_{\vartheta_n}^{(n)} \right] \frac{\partial}{\partial \vartheta} \overline{\Lambda}_{\vartheta}^{(n)T} d\overline{\Lambda}_{\vartheta_n}^{(n)} \Big|_{\vartheta=\vartheta_n}$$

ersetzen können. Ebenso kann auch $\mathcal{G}(\vartheta_0)$ durch das Stichproben-Analogon $\mathcal{G}^{(n)}(\vartheta_n)$ mit

$$\mathcal{G}_{ij}^{(n)}(\vartheta_n) = \int_I \varphi_i^{(n)}(t) \varphi_j^{(n)}(t) d\overline{\Lambda}_{\vartheta_n}^{(n)}(t), \quad \varphi^{(n)}(t) = \int_{[t, t^0]} \frac{\partial}{\partial \vartheta} \overline{\Lambda}_{\vartheta_n}^{(n)}(x) d\overline{\Lambda}_{\vartheta_n}^{(n)}(x)$$

ersetzt werden, wie man ähnlich zeigen kann.

3.4.4 Beweis der asymptotischen Normalverteiltheit

Wir haben nun sämtliche Hilfsmittel beisammen, um den Beweis von Satz 3.4.10 zu führen. Obleich er verhältnismäßig kurz anmuten mag, möchten wir daran erinnern, dass die Hauptarbeit in der Analyse der parametrischen Prozesse aus Unterabschnitt 3.4.1 liegt.

Beweis⁶¹ von Satz 3.4.10:

Wir erinnern zunächst an die Gleichung (3.4.10):

$$\begin{aligned}\beta_n(\vartheta_n) + \gamma_n(\vartheta_n) &= \sqrt{n} \int_I \left[\bar{\Lambda}_{\vartheta_n}^{(n)} - \bar{\Lambda}_{\vartheta_0}^{(n)} \right] \frac{\partial}{\partial \vartheta} \bar{\Lambda}_{\vartheta}^{(n)} \Big|_{\vartheta=\vartheta_n} d\mu_{\bar{N}^{(n)}} \\ &= \alpha_n(\vartheta_n) + \sqrt{n} \int_I \left[\bar{\Lambda}_{\vartheta_n}^{(n)} - \bar{\Lambda}_{\vartheta_0}^{(n)} \right] \frac{\partial}{\partial \vartheta} \bar{\Lambda}_{\vartheta}^{(n)} \Big|_{\vartheta=\vartheta_n} d\bar{\Lambda}_{\vartheta_0}^{(n)}.\end{aligned}\quad (3.4.10)$$

Unter Verwendung der Lemmata 3.4.4, 3.4.8 und 3.4.9 folgt aus dieser Gleichung:

$$\begin{aligned}\beta_n(\vartheta_n) + \sqrt{n} \int_I \bar{M}^{(n)} E \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta} \Big|_{\vartheta=\vartheta_n} dE\Lambda_{\vartheta_0} + R_2(\vartheta_n) \\ &= \alpha_n(\vartheta_n) + \sqrt{n} \int_I \left[\bar{\Lambda}_{\vartheta_n}^{(n)} - \bar{\Lambda}_{\vartheta_0}^{(n)} \right] \frac{\partial}{\partial \vartheta} \bar{\Lambda}_{\vartheta}^{(n)} \Big|_{\vartheta=\vartheta_n} d\bar{\Lambda}_{\vartheta_0}^{(n)} \\ &= \sqrt{n} \int_I \left[E\Lambda_{\vartheta_n} - E\Lambda_{\vartheta_0} \right] E \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta} \Big|_{\vartheta=\vartheta_n} d\bar{M}^{(n)} + R_1(\vartheta_n) \\ &\quad + \sqrt{n} \int_I \left[\bar{\Lambda}_{\vartheta_n}^{(n)} - \bar{\Lambda}_{\vartheta_0}^{(n)} \right] \frac{\partial}{\partial \vartheta} \bar{\Lambda}_{\vartheta}^{(n)} \Big|_{\vartheta=\vartheta_n} d\bar{\Lambda}_{\vartheta_0}^{(n)}\end{aligned}\quad (3.4.80)$$

Nun konvergiert ϑ_n nach Satz 3.3.4 fast sicher gegen ϑ_0 . Die in Lemma 3.4.4 gezeigte Straffheit der Folge

$$\left(\sqrt{n} \int_I \left[E\Lambda_{\vartheta_n} - E\Lambda_{\vartheta_0} \right] E \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta} \Big|_{\vartheta=\vartheta_n} d\bar{M}^{(n)} \right)_{n \in \mathbb{N}}$$

liefert also mit Lemma 2.3.6 die Konvergenz

$$\sqrt{n} \int_I \left[E\Lambda_{\vartheta_n} - E\Lambda_{\vartheta_0} \right] E \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta} \Big|_{\vartheta=\vartheta_n} d\bar{M}^{(n)} \xrightarrow{P} 0 \quad (n \rightarrow \infty).$$

Analog erhalten wir aufgrund von Lemma 3.4.9

$$\sqrt{n} \int_I \bar{M}^{(n)} E \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta} \Big|_{\vartheta=\vartheta_n} dE\Lambda_{\vartheta_0} - \sqrt{n} \int_I \bar{M}^{(n)} E \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta} \Big|_{\vartheta=\vartheta_0} dE\Lambda_{\vartheta_0} \xrightarrow{P} 0 \quad (n \rightarrow \infty).$$

Definieren wir ferner

$$\tilde{R}(\vartheta) = \beta_n(\vartheta) + R_2(\vartheta) - R_1(\vartheta),$$

so gilt wiederum gleichmäßig auf kompakten Teilmengen $K \subset B_\varepsilon(\vartheta_0)$ mit $\vartheta_0 \in K$ die Konvergenz

$$\tilde{R}(\vartheta) \xrightarrow{P} 0 \quad (n \rightarrow \infty).$$

⁶¹Der Beweis orientiert sich an der Vorlage aus Stute und Kopperschmidt 2013, S. 1293. Die wesentlichen Ergänzungen seitens des Autors betreffen die Ausführungen nach Gleichung (3.4.82), insbesondere die Berechnung der Kovarianzmatrix.

Insbesondere folgt demnach

$$\begin{aligned} R_0(\vartheta_n) &= \tilde{R}(\vartheta_n) - \sqrt{n} \int_I \left[E\Lambda_{\vartheta_n} - E\Lambda_{\vartheta_0} \right] E \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta} \Big|_{\vartheta=\vartheta_n} d\bar{M}^{(n)} \\ &\quad + \sqrt{n} \int_I \bar{M}^{(n)} E \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta} \Big|_{\vartheta=\vartheta_n} dE\Lambda_{\vartheta_0} - \sqrt{n} \int_I \bar{M}^{(n)} E \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta} \Big|_{\vartheta=\vartheta_0} dE\Lambda_{\vartheta_0} \xrightarrow{P} 0 \quad (n \rightarrow \infty). \end{aligned} \quad (3.4.81)$$

Durch Einsetzen in Gleichung (3.4.80) erhalten wir sodann:

$$\sqrt{n} \int_I \left[\bar{\Lambda}_{\vartheta_n}^{(n)} - \bar{\Lambda}_{\vartheta_0}^{(n)} \right] \frac{\partial}{\partial \vartheta} \bar{\Lambda}_{\vartheta}^{(n)} \Big|_{\vartheta=\vartheta_n} d\bar{\Lambda}_{\vartheta_0}^{(n)} = \sqrt{n} \int_I \bar{M}^{(n)} E \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta} \Big|_{\vartheta=\vartheta_0} dE\Lambda_{\vartheta_0} + R_0(\vartheta_n) \quad (3.4.82)$$

Eine Taylor-Entwicklung der linken Seite von (3.4.82) liefert weiter

$$\begin{aligned} &\sqrt{n} \int_I \left[\bar{\Lambda}_{\vartheta_n}^{(n)} - \bar{\Lambda}_{\vartheta_0}^{(n)} \right] \frac{\partial}{\partial \vartheta} \bar{\Lambda}_{\vartheta}^{(n)} \Big|_{\vartheta=\vartheta_n} d\bar{\Lambda}_{\vartheta_0}^{(n)} \\ &= \frac{\partial}{\partial \vartheta} \left(\sqrt{n} \int_I \left[\bar{\Lambda}_{\vartheta}^{(n)} - \bar{\Lambda}_{\vartheta_0}^{(n)} \right] \frac{\partial}{\partial \vartheta} \bar{\Lambda}_{\vartheta}^{(n)} d\bar{\Lambda}_{\vartheta_0}^{(n)} \right) \Big|_{\vartheta=\tilde{\vartheta}_n} (\vartheta_n - \vartheta_0) \\ &= \sqrt{n} \Phi_n(\tilde{\vartheta}_n) (\vartheta_n - \vartheta_0) \end{aligned}$$

für eine geeignete Zwischenstelle $\tilde{\vartheta}_n$ auf der Verbindungsstrecke zwischen ϑ_n und ϑ_0 .

Kennen wir die asymptotische Verteilung der rechten Seite dieser Gleichung, so können wir später aufgrund von Korollar 3.4.12 unter Verwendung des *Lemmas von Slutsky* $\Phi_n(\tilde{\vartheta}_n)$ durch $\Phi_0(\vartheta_0)$ ersetzen, ohne die asymptotische Verteilung zu verändern. Für den Beweis des Satzes genügt es also, die Verteilungskonvergenz der Folge

$$\sqrt{n} \Phi_n(\tilde{\vartheta}_n) (\vartheta_n - \vartheta_0)$$

zu untersuchen. Dazu betrachten wir die rechte Seite von (3.4.82) und erhalten unter Verwendung des *Satzes von Fubini* wegen $\bar{M}^{(n)}(t_0) = 0$:

$$\begin{aligned} \sqrt{n} \int_I \bar{M}^{(n)} E \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta} \Big|_{\vartheta=\vartheta_0} dE\Lambda_{\vartheta_0} &= \sqrt{n} \int_I \int_{[t_0, t]} d\bar{M}^{(n)}(x) E \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta} \Big|_{\vartheta=\vartheta_0}(t) dE\Lambda_{\vartheta_0}(t) \\ &= \sqrt{n} \int_I \int_{[t_0, t]} E \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta} \Big|_{\vartheta=\vartheta_0}(t) d\bar{M}^{(n)}(x) dE\Lambda_{\vartheta_0}(t) \\ &= \sqrt{n} \int_I \int_{[x, t^0]} E \frac{\partial}{\partial \vartheta} \Lambda_{\vartheta}(t) \Big|_{\vartheta=\vartheta_0} dE\Lambda_{\vartheta_0}(t) d\bar{M}^{(n)}(x) \end{aligned}$$

$$\begin{aligned}
&= \sqrt{n} \int_I \varphi(x) \, d\bar{M}^{(n)}(x) \\
&= \frac{1}{\sqrt{n}} \sum_{i=1}^n \int_I \varphi(x) \, dM^{(i)}(x) \tag{3.4.83}
\end{aligned}$$

Die zweiten Momente der rechten Seite von (3.4.83) existieren für alle $n \in \mathbb{N}$; da die $M^{(i)}$ unabhängig identisch verteilt sind, gilt nämlich wie zuvor unter Verwendung der verallgemeinerten Itô-Isometrie, siehe Satz 2.1.32:

$$\mathbb{E} \left\| \frac{1}{\sqrt{n}} \sum_{i=1}^n \int_I \varphi(x) \, dM^{(i)}(x) \right\|^2 = \mathbb{E} \left\| \int_I \varphi(x) \, dM(x) \right\|^2 = \int_I \|\varphi(x)\|^2 \, d\mathbb{E}\Lambda_{\vartheta_0}(x),$$

und die Behauptung folgt aus der Stetigkeit des Integranden. An dieser Stelle sei noch einmal an Bemerkung 3.4.3 erinnert:

In Stute und Kopperschmidt 2013 wird dabei lediglich gefordert, dass φ bezüglich $\mathbb{E}\Lambda_{\vartheta_0}$ quadratisch integrierbar sei, und diese Voraussetzung genügt hier offensichtlich.

Da die Integrale aus (3.4.83) gemäß Satz 2.1.32 zudem zentriert sind, können wir schließlich den *Zentralen Grenzwertsatz* anwenden und folgern somit die asymptotische Normalverteiltheit.

Für die Kovarianzmatrix beachten wir erneut die Unabhängigkeit der Martingale $M^{(i)}$ und erhalten wie oben:

$$\begin{aligned}
\mathbb{E} \left[\left(\frac{1}{\sqrt{n}} \sum_{i=1}^n \int_I \varphi \, dM^{(i)} \right) \left(\frac{1}{\sqrt{n}} \sum_{i=1}^n \int_I \varphi \, dM^{(i)} \right)^T \right] &= \mathbb{E} \left[\left(\int_I \varphi \, dM \right) \left(\int_I \varphi \, dM \right)^T \right] \\
&= \mathbb{E} \left[\left(\int_I \varphi_i \, dM \right) \left(\int_I \varphi_j \, dM \right) \right]_{1 \leq i, j \leq d} \tag{3.4.84}
\end{aligned}$$

Für $1 \leq i, j \leq d$ erhalten wir unter Verwendung der Identität $4ab = (a+b)^2 - (a-b)^2$ sowie der verallgemeinerten Itô-Isometrie, da die Integranden deterministisch sind:

$$\begin{aligned}
\mathbb{E} \left[\left(\int_I \varphi_i \, dM \right) \left(\int_I \varphi_j \, dM \right) \right] &= \frac{1}{4} \mathbb{E} \left[\left(\int_I (\varphi_i + \varphi_j) \, dM \right)^2 - \left(\int_I (\varphi_i - \varphi_j) \, dM \right)^2 \right] \\
&= \frac{1}{4} \left(\int_I (\varphi_i + \varphi_j)^2 \, d\mathbb{E}\Lambda_{\vartheta_0} - \int_I (\varphi_i - \varphi_j)^2 \, d\mathbb{E}\Lambda_{\vartheta_0} \right) = \int_I \varphi_i \varphi_j \, d\mathbb{E}\Lambda_{\vartheta_0}. \tag{3.4.85}
\end{aligned}$$

Setzen wir Gleichung (3.4.85) in die Gleichung (3.4.84) ein, so erhalten wir schließlich das gewünschte Ergebnis:

$$\mathbb{E} \left[\left(\int_I \varphi_i \, dM \right) \left(\int_I \varphi_j \, dM \right) \right]_{1 \leq i, j \leq d} = \left[\int_I \varphi_i \varphi_j \, d\mathbb{E}\Lambda_{\vartheta_0} \right]_{1 \leq i, j \leq d} = \mathcal{G}(\vartheta_0). \tag{3.4.86}$$

Letztlich folgt somit die Behauptung aus (3.4.82), da mit erneuter Anwendung des *Lemmas von Slutsky* in Verteilung folgt:

$$\sqrt{n}\Phi_n(\tilde{\vartheta}_n)(\vartheta_n - \vartheta_0) \longrightarrow \mathcal{N}_d(0, \mathcal{G}(\vartheta_0)) \quad (n \rightarrow \infty).$$

Dies vollendet den Beweis von Satz 3.4.10. □

Mit diesem Resultat schließen wir die Analyse der asymptotischen Eigenschaften des Minimum-Distanz-Schätzers ϑ_n und damit auch das dritte Kapitel dieser Arbeit ab.

Im nachfolgenden und abschließenden Kapitel möchten wir die wesentlichen Aussagen und Probleme noch einmal resümierend diskutieren.

Außerdem greifen wir die Argumente aus Bemerkung 3.4.13 auf und liefern einen Ausblick auf die weiterführende Verwendung des zentralen Beispiels 2.1.7 dieser Arbeit.

4 Diskussion und Ausblick

Um einen genehmen Abschluss der vorliegenden Arbeit zu gewährleisten, möchten wir die zurückliegenden Kapitel an dieser Stelle noch einmal Revue passieren lassen und den Leser auf die Probleme aufmerksam machen, die sich im Kontext der jeweiligen Abschnitte ergaben:

Kapitel 1: In der Einleitung haben wir die praktische Grundlage für das zentrale Beispiel 2.1.7 dieser Arbeit kennengelernt. Rückblickend mag uns auffallen, dass wir im Kontext dieses Beispiels im dritten Kapitel dieser Arbeit stets angenommen haben, dass der Versuch unter *gleichen* Bedingungen wiederholt wurde, während in der Praxis jede Iteration des Versuches mit einer unterschiedlichen Grundintensität der Belastung vorgenommen wurde. Ein Lösungsansatz dieses Problems beinhaltet die Interpretation der Grundintensität als *Zufallsgröße*, was jedoch in der Notwendigkeit einer Anpassung des Modells \mathcal{H}_N aus Abschnitt 3.1 resultiert.

Kapitel 2: Das zweite und damit das erste inhaltlich relevante Kapitel dieser Arbeit befasste sich mit den mathematischen Grundlagen, die für das Verständnis der nachfolgenden Theorie unabdingbar waren.

Die eminente Aufgabe dieses Kapitels war es, ein praktisch motiviertes Beispiel hervorzubringen, das den technischen Forderungen des dritten Kapitels genügt:

Mit Beispiel 2.1.7 ist es dem Autor gelungen, eine Brücke zwischen der in Kapitel 1 diskutierten Praxis und der Theorie aus Kapitel 3 zu schlagen und gleichermaßen eine neuartige Klasse von Zählprozessen einzuführen, die das Prinzip der Lastumverteilung zu modellieren vermag; im Rahmen von Unterabschnitt 2.1.8 haben wir jedoch gesehen, dass wir eine weitere Verallgemeinerung dieser Klasse benötigen, möchten wir auch die Schadensakkumulation berücksichtigen.

Des Weiteren konnte in diesem Kapitel die Definition des selbstanregenden Punktprozesses über den üblichen Rahmen hinaus motiviert und die Mehrdeutigkeit dieses Begriffes thematisiert werden.

Schließlich ist es dem Autor gelungen, das starke Gesetz der großen Zahlen in einer allgemeinen Version zu beweisen und das bekannte Straffheitskriterium von Kolmogorov ins Mehrdimensionale zu transferieren.

Kapitel 3: Das dritte und zentrale Kapitel behandelte die asymptotischen Eigenschaft des in Stute und Kopperschmidt 2013 eingeführten Minimum-Distanz-Schätzers ϑ_n .

Während des Studiums der mitunter skizzenhaften Beweise dieses Artikels wurden wir mit zwei zentralen Problemen konfrontiert, von denen eines nach wie vor seiner Lösung harret:

Einerseits stellten sich die von Stute und Kopperschmidt formulierten Voraussetzungen als nicht hinreichend heraus, während andererseits kein Beweis für das verwendete Staffheitskriterium von Kolmogorov im Falle mehrdimensionaler Parameterräume vorlag. Der Autor dieser Arbeit verfolgte deshalb die nachfolgenden Lösungsansätze zur Beseitigung dieser Komplikationen:

- Die Voraussetzungen wurden verschärft, um mit den dargelegten Beweisen konform zu gehen; ferner wurde anhand des Beispiels 2.1.7 gezeigt, dass diese Verschärfung keinen praxisrelevanten Nachteil hat.
- Das ins Mehrdimensionale transferierte Staffheitskriterium von Kolmogorov genügt den im dritten Kapitel präsentierten Beweisen nur im Falle eines eindimensionalen Parameterraumes.
Daher wird im Anhang ein alternatives Lipschitz-Kriterium vorgestellt, welches das Problem perspektivisch lösen soll.

Abschließend soll ein kurzer Ausblick die weiterführende Arbeit des Autors im Rahmen der selbstanregenden Zählprozesse skizzieren:

Wie wir in der letzten Fortsetzung des Beispiels 2.1.7 gesehen haben, eignet sich das Modell \mathcal{H}_N für den Kompensator eines reinen verschobenen Geburtsprozesses mit Abbruch zur Anwendung der Theorie aus dem dritten Kapitel dieser Arbeit. Insbesondere ist es uns möglich, anhand der asymptotischen Verteilung des Parameterschätzers ϑ_n ein approximatives Konfidenzintervall für den wahren Parameter ϑ_0 zu ermitteln. Auf Grundlage der Identität

$$E [N(t)] = E [\Lambda_{\vartheta_0}(t)] , \quad \text{für alle } t \in I ,$$

möchten wir sodann ein approximatives Prognoseintervall für die Anzahl der Sprünge des Zählprozesses N - im Kontext von Beispiel 2.1.7 also die Anzahl der Drahtbrüche - aufstellen.

Hierzu muss die obige Problematik des ersten Kapitels umgangen werden, da andernfalls noch keine konkreten Daten, auf deren Grundlage die Berechnung erfolgen könnte, existieren.

Möchten wir schließlich das Modell \mathcal{H}_N erweitern, um etwa weitere Parameter zuzulassen oder die Schadensakkumulation zu implementieren, so bedarf es weitergehender Untersuchung der *selbstanregenden Zählprozesse im Snyder & Miller'schen Sinne*, zu denen auch der verschobene Geburtsprozess gehört, sowie einer Adaption der im dritten Kapitel präsentierten Beweise für die Benutzung des im nachfolgenden Anhang vorgestellten Lipschitz-Staffheitskriteriums.

A Anhang

A.1 Elementare Hilfssätze

In diesem Abschnitt werden die Hilfssätze, die bereits im Anhang des Artikels Stute und Kopperschmidt 2013 zu finden sind, formuliert und bewiesen. Ihre Bewandtnis für diese Arbeit wird offenbar, wenn wir die Straffheit diverser Folgen $C(K)$ -wertiger Zufallsfunktion mittels des Straffheitskriteriums von Kolmogorov nachweisen, siehe Abschnitt 3.4.

Lemma A.1.1. (siehe Stute und Kopperschmidt 2013, S. 1295)

Für $p, k = 1, \dots, n$ seien U_{pk} d -variate Zufallsvektoren mit $E\|U_{pk}\|^2 < \infty$. Ferner gelte

$$E[U_{pk}^T U_{ql}] = 0 \quad \text{für } k \notin \{p, q, l\} \text{ oder } l \notin \{p, q, k\}. \quad (\text{A.1.1})$$

Dann gilt:

$$E\left\|\sum_{\substack{p,k=1 \\ p \neq k}}^n U_{pk}\right\|^2 \leq 2 \sum_{k \gtrless p,q} E[U_{pk}^T U_{qk}]. \quad (\text{A.1.2})$$

Beweis⁶²:

Aus der Abschätzung $(a+b)^2 \leq 2(a^2 + b^2)$ für $a, b \in \mathbb{R}$ und der Monotonie sowie Linearität des Erwartungswertes erhalten wir:

$$\begin{aligned} E\left\|\sum_{\substack{p,k=1 \\ p \neq k}}^n U_{pk}\right\|^2 &= E\left\|\sum_{\substack{p,k=1 \\ k < p}}^n U_{pk} + \sum_{\substack{p,k=1 \\ k > p}}^n U_{pk}\right\|^2 \leq E\left[\left\|\sum_{k < p} U_{pk}\right\| + \left\|\sum_{k > p} U_{pk}\right\|\right]^2 \\ &\leq 2E\left\|\sum_{k < p} U_{pk}\right\|^2 + 2E\left\|\sum_{k > p} U_{pk}\right\|^2. \end{aligned} \quad (\text{A.1.3})$$

Ferner gilt unter Verwendung der Bedingung (A.1.1):

$$\begin{aligned} E\left\|\sum_{k < p} U_{pk}\right\|^2 &= E\left[\left(\sum_{k < p} U_{pk}\right)^T \left(\sum_{l < q} U_{ql}\right)\right] = \sum_{k < p} \sum_{l < q} E[U_{pk}^T U_{ql}] \\ &= \sum_{k < p, q} E[U_{pk}^T U_{qk}], \end{aligned} \quad (\text{A.1.4})$$

da für $k \neq l$ unmittelbar $k \notin \{p, q, l\}$ oder $l \notin \{p, q, k\}$ folgt:

- Ist $k < l$, so gilt wegen $l < q$ auch $k < q$ und daher $k \neq p, q, l$.
- Ist $k > l$, so gilt wegen $k < p$ auch $l < p$ und daher $l \neq p, q, k$.

⁶²Der Beweis findet sich in der angegebenen Quelle.

Völlig analog erhalten wir zudem:

$$\mathbb{E} \left\| \sum_{k>p} U_{pk} \right\|^2 = \sum_{k>p,q} \mathbb{E} [U_{pk}^T U_{qk}] . \quad (\text{A.1.5})$$

Insgesamt folgt somit durch Einsetzen der Gleichungen (A.1.4) und (A.1.5) in die Abschätzung (A.1.3) die Behauptung. \square

Lemma A.1.2. (Siehe Stute und Kopperschmidt 2013, S. 1295)

Für $p, k = 1, \dots, n$ seien U_{pk} d -variate Zufallsvektoren mit $\mathbb{E} \|U_{pk}\|^2 < \infty$. Ferner gelte

$$\mathbb{E} [U_{pk}^T U_{ql}] = 0, \quad \text{wann immer sich ein Index von den restlichen unterscheidet.} \quad (\text{A.1.6})$$

Dann gilt:

$$\mathbb{E} \left\| \sum_{\substack{p,k=1 \\ p \neq k}}^n U_{pk} \right\|^2 \leq 2 \sum_{\substack{p,k=1 \\ p \neq k}}^n \mathbb{E} \|U_{pk}\|^2 . \quad (\text{A.1.7})$$

Beweis⁶³:

Wie im Beweis des vorherigen Lemmas A.1.1 zeigt man:

$$\mathbb{E} \left\| \sum_{\substack{p,k=1 \\ p \neq k}}^n U_{pk} \right\|^2 \leq 2\mathbb{E} \left\| \sum_{k<p} U_{pk} \right\|^2 + 2\mathbb{E} \left\| \sum_{k>p} U_{pk} \right\|^2 .$$

Aus der Bedingung (A.1.6) folgt gleichermaßen analog (man beachte, dass die Bedingung auch die Gültigkeit von Bedingung (A.1.1) impliziert):

$$\begin{aligned} \mathbb{E} \left\| \sum_{k<p} U_{pk} \right\|^2 &= \sum_{k<p} \sum_{l<q} \mathbb{E} [U_{pk}^T U_{ql}] = \sum_{k<p,q} \mathbb{E} [U_{pk}^T U_{qk}] \\ &= \sum_{k<p} \mathbb{E} [U_{pk}^T U_{pk}] = \sum_{k<p} \mathbb{E} \|U_{pk}\|^2 , \end{aligned}$$

da jeder Summand mit $p \neq q$ nach Voraussetzung verschwindet, sowie

$$\mathbb{E} \left\| \sum_{k>p} U_{pk} \right\|^2 = \sum_{k>p} \mathbb{E} \|U_{pk}\|^2 .$$

Hieraus folgt unmittelbar die Behauptung. \square

⁶³Der Beweis findet sich ebenfalls in der angegebenen Quelle.

Lemma A.1.3. (Siehe Stute und Kopperschmidt 2013, S. 1296)

Für $p, k, i = 1, \dots, n$ seien U_{pki} d -variate Zufallsvektoren mit $E\|U_{pki}\|^2 < \infty$. Ferner gelte

$$E[U_{pki}^T U_{qlj}] = 0, \quad \text{wann immer sich } k, i, l \text{ oder } j \text{ von den restlichen Indizes unterscheidet.} \quad (\text{A.1.8})$$

Dann gilt⁶⁴

$$E \left\| \sum_{\substack{p,k,i=1 \\ p \neq k \neq i \neq p}}^n U_{pki} \right\|^2 \leq 32 \sum E[U_{pki}^T U_{qki}], \quad (\text{A.1.9})$$

wobei die Summation rechts über alle Indexkombinationen erfolgt, in denen p und q die selbe Position relativ zu i und k einnehmen, das heißt:

$$\begin{aligned} p, q < k < i, & \quad p, q < i < k, & \quad k < p, q < i \\ i < p, q < k, & \quad k < i < p, q, & \quad i < k < p, q. \end{aligned} \quad (\text{A.1.10})$$

Beweis⁶⁵:

Wir benötigen für den Beweis dieses Lemmas die Verallgemeinerung der Ungleichung $(a + b)^2 \leq 2(a^2 + b^2)$ mit $a, b \in \mathbb{R}$, die wir bereits für den Beweis von Lemma A.1.1 verwendet haben:

Für $a_1, \dots, a_l \in \mathbb{R}^d$ gilt die Abschätzung:

$$\left\| \sum_{j=1}^l a_j \right\|^2 \leq 2^{l-1} \sum_{j=1}^l \|a_j\|^2. \quad (\text{A.1.11})$$

Der Beweis ist elementar und beruht auf dem obigen Spezialfall $l = 2, d = 1$.

Durch sukzessives Anwenden dieser Ungleichung erhält man die Behauptung:

$$\begin{aligned} \left\| \sum_{j=1}^l a_j \right\|^2 &\leq \left(\sum_{j=1}^l \|a_j\| \right)^2 = \left(\sum_{j=1}^{l-1} \|a_j\| + \|a_l\| \right)^2 \\ &\leq 2 \left[\left(\sum_{j=1}^{l-1} \|a_j\| \right)^2 + \|a_l\|^2 \right] \leq 2 \left[2 \left[\left(\sum_{j=1}^{l-2} \|a_j\| \right)^2 + \|a_{l-1}\|^2 \right] + \|a_l\|^2 \right] \end{aligned}$$

⁶⁴In Stute und Kopperschmidt 2013 wird als konstanter Faktor die Zahl 64 verwendet, wir werden jedoch sehen, dass der Faktor 32 genügt.

⁶⁵Der Beweis findet sich in der angegebenen Quelle, wurde jedoch vom Autor dieser Arbeit um wesentliche Argumente erweitert.

$$\begin{aligned}
&\leq 2\|a_l\|^2 + 4\|a_{l-1}\|^2 + \dots + 2^{l-2}\|a_3\|^2 + 2^{l-1}(\|a_2\|^2 + \|a_1\|^2) \\
&\leq 2^{l-1} \sum_{j=1}^l \|a_j\|^2.
\end{aligned}$$

Wie in der Abschätzung A.1.3 aus Lemma A.1.1 werden wir die verallgemeinerte Ungleichung (A.1.11) bereits im ersten Schritt des Beweises utilisieren:

$$\begin{aligned}
\mathbb{E} \left\| \sum_{\substack{p,k,i=1 \\ p \neq k \neq i \neq p}}^n U_{pki} \right\|^2 &= \mathbb{E} \left\| \sum_{p < k < i} U_{pki} + \sum_{p < i < k} U_{pki} + \sum_{k < p < i} U_{pki} \right. \\
&\quad \left. + \sum_{i < p < k} U_{pki} + \sum_{k < i < p} U_{pki} + \sum_{i < k < p} U_{pki} \right\|^2 \\
&\leq 32 \left[\mathbb{E} \left\| \sum_{p < k < i} U_{pki} \right\|^2 + \mathbb{E} \left\| \sum_{p < i < k} U_{pki} \right\|^2 + \mathbb{E} \left\| \sum_{k < p < i} U_{pki} \right\|^2 \right. \\
&\quad \left. + \mathbb{E} \left\| \sum_{i < p < k} U_{pki} \right\|^2 + \mathbb{E} \left\| \sum_{k < i < p} U_{pki} \right\|^2 + \mathbb{E} \left\| \sum_{i < k < p} U_{pki} \right\|^2 \right]. \quad (\text{A.1.12})
\end{aligned}$$

Für den ersten Summanden aus Gleichung (A.1.12) erhalten wir:

$$\begin{aligned}
\mathbb{E} \left\| \sum_{p < k < i} U_{pki} \right\|^2 &= \mathbb{E} \left[\left(\sum_{p < k < i} U_{pki} \right)^T \left(\sum_{p < k < i} U_{pki} \right) \right] \\
&= \mathbb{E} \left[\sum_{p < k < i} \sum_{q < l < j} U_{pki}^T U_{qlj} \right] = \sum_{p < k < i} \sum_{q < l < j} \mathbb{E} [U_{pki}^T U_{qlj}]. \quad (\text{A.1.13})
\end{aligned}$$

Jeder Summand mit $i \neq j$ verschwindet nach Voraussetzung, da in diesen Fällen wie zuvor bereits $i \notin \{p, k, q, l, j\}$ oder $j \notin \{p, k, i, q, l\}$ folgt. Aus Gleichung (A.1.13) folgt also:

$$\mathbb{E} \left\| \sum_{p < k < i} U_{pki} \right\|^2 = \sum_{\substack{p,k,i=1 \\ p < k < i}}^n \sum_{\substack{q,l=1 \\ q < l < i}}^n \mathbb{E} [U_{pki}^T U_{qli}]. \quad (\text{A.1.14})$$

Da nun mit der gleichen Argumentation jeder Summand mit $l \neq k$ verschwindet, folgt aus Gleichung schließlich:

$$\mathbb{E} \left\| \sum_{p < k < i} U_{pki} \right\|^2 = \sum_{\substack{p,q,k,i=1 \\ p,q < k < i}}^n \mathbb{E} [U_{pki}^T U_{qki}]. \quad (\text{A.1.15})$$

Völlig analog dazu liefert die Betrachtung der verbleibenden Summanden:

$$\begin{aligned}
\mathbb{E} \left\| \sum_{p < i < k} U_{pki} \right\|^2 &= \sum_{\substack{p,q,k,i=1 \\ p,q < i < k}}^n \mathbb{E} [U_{pki}^T U_{qki}] , & \mathbb{E} \left\| \sum_{k < p < i} U_{pki} \right\|^2 &= \sum_{\substack{p,q,k,i=1 \\ k < p,q < i}}^n \mathbb{E} [U_{pki}^T U_{qki}] , \\
\mathbb{E} \left\| \sum_{i < p < k} U_{pki} \right\|^2 &= \sum_{\substack{p,q,k,i=1 \\ i < p,q < k}}^n \mathbb{E} [U_{pki}^T U_{qki}] , & \mathbb{E} \left\| \sum_{k < i < p} U_{pki} \right\|^2 &= \sum_{\substack{p,q,k,i=1 \\ k < i < p,q}}^n \mathbb{E} [U_{pki}^T U_{qki}] , \\
\mathbb{E} \left\| \sum_{i < p < k} U_{pki} \right\|^2 &= \sum_{\substack{p,q,k,i=1 \\ i < k < p,q}}^n \mathbb{E} [U_{pki}^T U_{qki}] . & & \tag{A.1.16}
\end{aligned}$$

Einsetzen der Gleichungen (A.1.15) und (A.1.16) in die Abschätzung (A.1.12) liefert letztlich die Behauptung. \square

A.2 Eine Abzählung des Gitters $[0, 1]^d \cap \delta\mathbb{Z}^d$

Im Beweis des Straffheitskriteriums von Kolmogorov für den d -dimensionalen Einheitswürfel aus Lemma 2.3.7 verwenden wir eine spezielle *Abzählung* des Gitters $[0, 1]^d \cap \delta\mathbb{Z}^d$, in der aufeinanderfolgende Punkte stets benachbart sind. Dass eine solche Abzählung tatsächlich existiert, besagt das nachfolgende technische Lemma.

Lemma A.2.1.

Es sei $0 < \delta \leq 1$. Das d -dimensionale Gitter G_δ^d sei definiert durch

$$G_\delta^d = [0, 1]^d \cap \delta\mathbb{Z}^d = \left\{ (x_1, \dots, x_d) \in \mathbb{R}^d : \forall i \in \{1, \dots, d\} \exists j_i \in \{0, 1, \dots, \lfloor \delta^{-1} \rfloor\} \text{ mit } x_i = \delta j_i \right\}.$$

Dann existiert eine injektive Abzählung h aller $(1 + \lfloor \delta^{-1} \rfloor)^d$ Elemente von G_δ^d mit

$$\|h(k+1) - h(k)\|_\infty = \delta, \quad \text{für alle } k = 1, \dots, (1 + \lfloor \delta^{-1} \rfloor)^d - 1, \tag{A.2.1}$$

wobei $\|\cdot\|_\infty$ die Maximumsnorm auf \mathbb{R}^d bezeichne⁶⁶.

Beweis⁶⁷:

Wir zeigen die Behauptung durch vollständige Induktion nach der Dimension d . Um die Notation übersichtlicher zu gestalten, definieren wir $m = \lfloor \delta^{-1} \rfloor$, und multiplizieren alle Elemente des Gitters G_δ^d mit $\lfloor \delta^{-1} \rfloor$, um ganzzahlige Knotenpunkte zu erhalten. Aus der so konstruierten injektiven Abzählung erhält man sodann mittels Reskalierung die gefor-

⁶⁶Die Wahl der Norm ist weitestgehend willkürlich, der Beweis funktioniert beispielsweise auch für die Euklidische Norm.

⁶⁷Dieses Lemma wurde vom Autor dieser Arbeit formuliert und bewiesen.

derte Abbildung h . Um Verwirrung zu vermeiden, werden die im Beweis konstruierten Abzählungen daher stets mit \tilde{h} bezeichnet werden.

$d = 1$: Dieser Fall ist trivial mit $\tilde{h}(k) = k$. Reskalierung liefert die injektive Abzählung $h(k) = \delta\tilde{h}(k) = \delta k$.

$d = 2$: Da für den Induktionsbeweis nicht die bloße Existenz einer solchen Abzählung im Falle $d = 2$ genügt, unterscheiden wir zunächst, ob m gerade oder ungerade ist, und zeigen dementsprechend *unterschiedliche* Aussagen:

1. Falls m gerade ist, existiert eine injektive Abzählung \tilde{h} mit $\tilde{h}(1) = (0, 0)$ und $\tilde{h}((m+1)^2) = (m, m)$.
2. Falls m ungerade ist, existiert eine injektive Abzählung \tilde{h} mit $\tilde{h}(1) = (0, 0)$ und $\tilde{h}((m+1)^2) = (m, 0)$.

Es sei zunächst m gerade. Wir definieren \tilde{h} sukzessive für $k = 1, \dots, m^2 + 2m$:

$$\tilde{h}(1) = (0, 0)$$

$$\tilde{h}(k+1) = \begin{cases} \tilde{h}(k) + (0, 1), & \text{falls } k \equiv 1, \dots, m \pmod{2(m+1)}, \\ \tilde{h}(k) + (1, 0), & \text{falls } k \equiv 0, m+1 \pmod{2(m+1)}, \\ \tilde{h}(k) - (0, 1), & \text{falls } k \equiv m+2, \dots, 2m+1 \pmod{2(m+1)}. \end{cases}$$

Da die obige Definition auf den ersten Blick umständlich erscheinen mag, illustrieren wir das Vorgehen anhand einer Skizze (siehe Abbildung 6) im Falle $m = 4$.

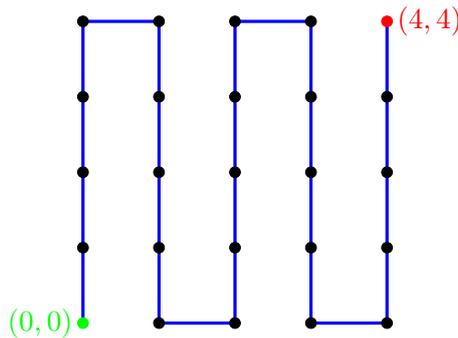


Abbildung 6: Definition der injektiven Abzählung im Falle $d = 2$, $m = 4$.

Die Abbildung ist nach Konstruktion einerseits injektiv und wegen

$$\tilde{h}(k+1) - \tilde{h}(k) \in \{(1, 0), \pm(0, 1)\}$$

für alle $k = 1, \dots, m^2 + 2m$ folgt andererseits auch

$$\|\tilde{h}(k+1) - \tilde{h}(k)\|_\infty = 1$$

und damit nach Reskalierung die Behauptung.

Nun sei m ungerade. Wir können erneut die obige Abzählung \tilde{h} verwenden, um das gewünschte Ergebnis zu erzielen.

Das Vorgehen wird wiederum anhand einer Skizze (siehe Abbildung 7) verdeutlicht, hier im Falle $m = 5$.

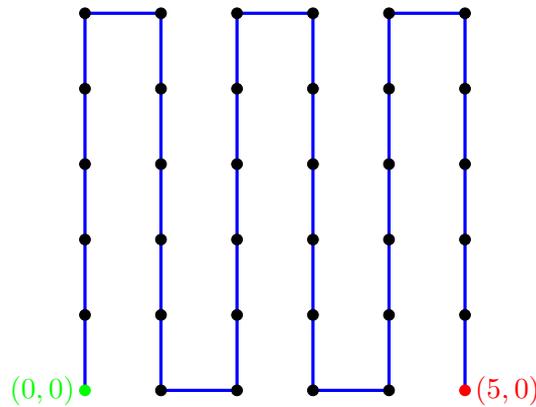


Abbildung 7: Definition der injektiven Abzählung im Falle $d = 2$, $m = 5$.

Wie zuvor liefert Reskalierung also die Behauptung und vollendet den Beweis in dem Fall $d = 2$, der unseren Induktionsanfang darstellen wird.

$d > 2$: Es sei nun $d \geq 3$ beliebig gewählt, und die nachfolgende Behauptung sei bereits bewiesen:

1. Falls m gerade ist, existiert eine injektive Abzählung \tilde{h} des Gitters G_δ^{d-1} mit $\tilde{h}(1) = (0, \dots, 0)$ und $\tilde{h}((m+1)^{d-1}) = (m, \dots, m)$.
2. Falls m ungerade ist, existiert eine injektive Abzählung \tilde{h} des Gitters G_δ^{d-1} mit $\tilde{h}(1) = (0, \dots, 0)$ und $\tilde{h}((m+1)^{d-1}) = (m, 0, \dots, 0)$.

Wir bemerken, dass dies für $d = 3$ (also $d - 1 = 2$) gerade die soeben gezeigten Aussagen sind.

Erneut unterscheiden wir zwei Fälle:

Zunächst sei m gerade. Wir beginnen im Punkt $(0, \dots, 0)$ und fixieren die erste Komponente. Gemäß der Induktionsvoraussetzung können wir alle Elemente des Gitters, deren erste Komponente 0 ist, durchlaufen und erreichen schließlich den Punkt

$(0, m, \dots, m)$. Von hier aus erreichen wir den Punkt $(1, m, \dots, m)$ und durchlaufen die Abzählung *rückwärts*, um zum Punkt $(1, 0, \dots, 0)$ zu gelangen. Definieren wir dann den Punkt $(2, 0, \dots, 0)$ als nachfolgendes Gitterelement, so können wir sukzessive das selbe Verfahren anwenden, um - da m gerade ist - schließlich die Ecke $(m, 0, \dots, 0)$ zu erreichen. Durchlaufen wir letztmalig die Abzählung, so erreichen wir schlussendlich den Punkt (m, \dots, m) , und haben somit eine injektive Abzählung des Gitters $G_\delta^{(d)}$ konstruiert, welche die Gültigkeit der *ersten* Aussage auch in diesem Falle zeigt.

Nun sei m ungerade. Wir verfahren analog zum geraden Fall und gelangen so vom Punkt $(0, \dots, 0)$ bis an den Punkt $(m-1, 0, \dots, 0)$, da m ungerade ist. Wir durchlaufen die Abzählung erneut und erreichen den Punkt $(m-1, m, 0, \dots, 0)$, von wo aus der Punkt $(m, m, 0, \dots, 0)$ erreicht wird. Durchlaufen wir letztmalig die Abzählung *rückwärts*, so erreichen wir hier schlussendlich den Punkt $(m, 0, \dots, 0)$, und haben somit wiederum eine injektive Abzählung des Gitters $G_\delta^{(d)}$ konstruiert, welche die Gültigkeit der *zweiten* Aussage auch in diesem Falle zeigt.

Insgesamt folgt somit durch Reskalierung die Behauptung für ein beliebiges d und damit die Aussage des Lemmas. \square

A.3 Das Lipschitz-Straffheitskriterium

In Bemerkung 2.3.9 haben wir bereits auf eine Alternative zum Straffheitskriterium von Kolmogorov aus Abschnitt 2.3.1 hingewiesen. Dieses Kriterium möchten wir im nachfolgenden Korollar formulieren, das leicht aus Korollar 2.3.5 folgt.

Korollar A.3.1.

Eine Folge $(X_n)_{n \in \mathbb{N}}$ von Zufallsfunktionen mit Werten im Raum $C(K)$ ist straff, wenn die nachfolgenden Bedingungen erfüllt sind:

- (i) Es gibt ein $\vartheta_0 \in K$, so dass für jedes $\eta > 0$ ein $a \in \mathbb{R}$ existiert mit

$$P(\{\|X_n(\vartheta_0)\| > a\}) \leq \eta, \quad \text{für alle } n \in \mathbb{N}.$$

- (ii) Es existiert eine Folge $(L_n)_{n \in \mathbb{N}}$ reeller Zufallsvariablen und eine Konstante $\nu > 0$ mit

$$EL_n \leq \nu, \quad \text{für alle } n \in \mathbb{N}, \tag{A.3.1}$$

so dass die *Lipschitz-Bedingung*

$$\|X_n(x) - X_n(y)\| \leq L_n \|x - y\|, \quad \text{für alle } x, y \in K, \tag{A.3.2}$$

fast sicher erfüllt ist.

Beweis⁶⁸:

Der Beweis folgt unter Verwendung der *Markov'schen Ungleichung* aus Korollar 2.3.5. Zu gegebenem $\eta > 0, \varepsilon > 0$ müssen wir dabei lediglich ein $\delta > 0$ und ein $n_0 \in \mathbb{N}$ finden, so dass gilt:

$$P \left(\left\{ \sup_{\|x-y\|<\delta} \|X_n(x) - X_n(y)\| \geq \varepsilon \right\} \right) \leq \eta, \quad \text{für alle } n \geq n_0.$$

Aufgrund der Bedingungen aus (A.3.1) und (A.3.2) gilt jedoch

$$\begin{aligned} P \left(\left\{ \sup_{\|x-y\|<\delta} \|X_n(x) - X_n(y)\| \geq \varepsilon \right\} \right) &\leq P \left(\left\{ \sup_{\|x-y\|<\delta} L_n \|x - y\| \geq \varepsilon \right\} \right) \\ &\leq P \left(\left\{ L_n \geq \frac{\varepsilon}{\delta} \right\} \right) \leq \frac{\delta}{\varepsilon} \mathbb{E} L_n \leq \frac{\nu}{\varepsilon} \delta \longrightarrow 0 \quad (\delta \rightarrow 0), \quad \text{für alle } n \in \mathbb{N}, \end{aligned}$$

woraus sofort die Behauptung folgt. □

Bemerkung A.3.2.

- Fordern wir anstelle von (A.3.1), dass eine Konstante $\nu > 0$ existiert mit

$$\mathbb{E} L_n^2 \leq \nu, \quad \text{für alle } n \in \mathbb{N},$$

so folgert man aus der Bedingung (A.3.2) leicht die Abschätzung

$$\mathbb{E} \|X_n(x) - X_n(y)\|^2 \leq \nu \|x - y\|^2, \quad \text{für alle } x, y \in K. \quad (\text{A.3.3})$$

Diese Abschätzung erinnert an die Momenten-Bedingung aus dem Straffheitskriterium von Kolmogorov, siehe Korollar 2.3.8. Für den Fall $d = 1$ impliziert die Lipschitz-Bedingung dann sogar die Momenten-Bedingung des Kolmogorov'schen Straffheitskriteriums, weshalb die Voraussetzungen des Lipschitz-Kriteriums als eine Verschärfung interpretiert werden können.

- Mittels der Forderung aus dem ersten Teil dieser Bemerkung ist es möglich, eine Modifikation des Kriteriums aus Korollar A.3.1 zu formulieren, die eher an das *Straffheitskriterium von Kolmogorov* erinnert. Dazu verändern wir die zweite Bedingung des vorherigen Korollars wie folgt:

⁶⁸Das Korollar wurde vom Autor dieser Arbeit formuliert und bewiesen.

(ii*) Es existiert eine Folge $(L_n)_{n \in \mathbb{N}}$ reeller Zufallsvariablen und eine Konstante $\nu > 0$ mit

$$EL_n^2 \leq \nu, \quad \text{für alle } n \in \mathbb{N},$$

so dass die *Lipschitz-Bedingung*

$$\|X_n(x) - X_n(y)\|^2 \leq L_n \|x - y\|^2, \quad \text{für alle } x, y \in K,$$

fast sicher erfüllt ist.

Dann kann in Korollar A.3.1 die Bedingung (ii) durch die Bedingung (ii*) ersetzt werden.

Literatur

- [Bau91] Heinz Bauer. *Wahrscheinlichkeitstheorie*. 4., völlig überarb. und neugestaltete Aufl. Berlin und New York: Walter de Gruyter, 1991. ISBN: 3-11-012191-3.
- [Bau92] Heinz Bauer. *Maß- und Integrationstheorie*. 2., überarb. Aufl. Berlin und New York: Walter de Gruyter, 1992. ISBN: 3-11-013625-2.
- [Bil68] Patrick Billingsley. *Convergence of Probability Measures*. New York u. a.: John Wiley & Sons, Inc., 1968.
- [Bis14] Marek Biskup. *Standard Borel spaces. Infinite product spaces*. 2014. URL: <http://www.math.ucla.edu/~biskup/275b.1.14w/PDFs/Standard-Borel-Spaces.pdf> (besucht am 08.02.2017).
- [DVJ03] D.J. Daley und D. Vere-Jones. *An Introduction to the Theory of Point Processes. Volume I: Elementary Theory and Methods*. Second Edition. Probability and Its Applications. New York, Berlin und Heidelberg: Springer-Verlag, 2003. ISBN: 0-387-95541-0.
- [DVJ08] D.J. Daley und D. Vere-Jones. *An Introduction to the Theory of Point Processes. Volume II: General Theory and Structure*. Second Edition. Probability and Its Applications. New York, Berlin und Heidelberg: Springer-Verlag, 2008. ISBN: 978-0-387-21337-8.
- [Dud02] Richard M. Dudley. *Real Analysis and Probability*. Band 74 aus *Cambridge Studies in Advanced Mathematics*. Revidierter Nachdruck der Originalausgabe von 1989. Cambridge: Cambridge University Press, 2002.
- [EK86] Stewart N. Ethier und Thomas G. Kurtz. *Markov Processes. Characterization and convergence*. Wiley series in probability and mathematical statistics. New York u. a.: John Wiley & Sons, Inc., 1986. ISBN: 0-471-08186-8.
- [Gra56] Lawrence M. Graves. *The Theory of Functions of Real Variables*. Second Edition. International Series in Pure and Applied Mathematics. New York, Toronto und London: McGraw-Hill Book Company, Inc., 1956.
- [Haw71] Alan G. Hawkes. “Spectra of Some Self-Exciting and Mutually Exciting Point Processes”. In: *Biometrika* 58 (1971), S. 83–90.
- [Hoe48] Wassily Hoeffding. “A class of statistics with asymptotically normal distributions”. In: *Annals of Statistics* 19 (1948), S. 293–325.
- [Jac06] Martin Jacobsen. *Point Process Theory and Applications. Marked Point and Piecewise Deterministic Processes*. Probability and its Applications. Boston: Birkhäuser, 2006. ISBN: 0-8176-4215-3.

- [Jak99] Adam Jakubowski. *The Skorokhod space in functional convergence: a short introduction*. 1999. URL: <http://kpbc.umk.pl/Content/39953/kievtopologies.pdf> (besucht am 08.02.2017).
- [KL14] David Kerr und Hanfeng Li. *Ergodic Theory. Independence and Dichotomies*. Appendix A. Polish spaces and standard Borel spaces. 2014. URL: <http://www.math.tamu.edu/~kerr/book/appendixA.pdf> (besucht am 08.02.2017).
- [Kec95] Alexander S. Kechris. *Classical Descriptive Set Theory*. Bd. 156. Graduate Texts in Mathematics. New York, Berlin und Heidelberg: Springer-Verlag, 1995. ISBN: 0-387-94374-9.
- [Kle13] Achim Klenke. *Wahrscheinlichkeitstheorie*. 3. Auflage. Berlin und Heidelberg: Springer-Verlag, 2013. ISBN: 978-3-642-36017-6.
- [Kol56] A. N. Kolmogorov. “On Skorokhod convergence”. In: *Theor. Probability Appl.* 1 (1956), S. 213–222.
- [Kuo06] Hui-Hsiung Kuo. *Introduction to Stochastic Integration*. Universitext. Springer-Verlag, 2006. ISBN: 0-387-28720-5.
- [Lee90] A. J. Lee. *U-Statistics. Theory and Practice*. Bd. 110. Statistics: textbooks and monographs. New York: Marcel Dekker, Inc., 1990. ISBN: 0-8247-8253-4.
- [MSM16] Christine H. Müller, Sebastian Szugat und Reinhard Maurer. *Simulation free prediction intervals for a state dependent failure process using accelerated lifetime experiments*. 2016. URL: https://www.statistik.tu-dortmund.de/fileadmin/user_upload/DP_4516_SFB823_Mueller_Szugat_Maurer.pdf (besucht am 08.02.2017).
- [PTW07] Guodong Pang, Rishi Talreja und Ward Whitt. “Martingale proofs of many-server heavy-traffic limits for Markovian queues”. In: *Probability Surveys* 4 (2007), S. 193–267.
- [Pro05] Philip E. Protter. *Stochastic Integration and Differential Equations*. Second Edition, Version 2.1. Bd. 21. Stochastic Modelling and Applied Probability. Berlin, Heidelberg und New York: Springer-Verlag, 2005. ISBN: 3-540-00313-4.
- [SK13] Winfried Stute und Kai Kopperschmidt. “The Statistical Analysis of Self-Exciting Point Processes”. In: *Statistica Sinica* 23 (2013), S. 1273–1298.
- [SM91] Donald L. Snyder und Michael I. Miller. *Random Point Processes in Time and Space*. Second Edition. New York: Springer-Verlag, 1991. ISBN: 978-1-4612-7821-4.
- [Sko56] Anatolij Skorokhod. “Limit Theorems for Stochastic Processes”. In: *Theor. Probability Appl.* 3 (1956), S. 261–290.

- [Sri98] S.M. Srivastava. *A Course on Borel Sets*. Bd. 180. Graduate Texts in Mathematics. New York, Berlin und Heidelberg: Springer-Verlag, 1998. ISBN: 0-387-98412-7.
- [Szu+16] Sebastian Szugat u. a. “Prediction Intervals for the Failure Time of Prestressed Concrete Beams”. In: *Advances in Materials Science and Engineering* 2016 (2016). URL: <https://www.hindawi.com/journals/amse/2016/9605450/>.
- [R C15] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria, 2015. URL: <http://www.R-project.org/>.

Eidesstattliche Versicherung

Name, Vorname

Matr.-Nr.

Ich versichere hiermit an Eides statt, dass ich die vorliegende Bachelorarbeit/Masterarbeit* mit dem Titel

selbstständig und ohne unzulässige fremde Hilfe erbracht habe. Ich habe keine anderen als die angegebenen Quellen und Hilfsmittel benutzt sowie wörtliche und sinngemäße Zitate kenntlich gemacht. Die Arbeit hat in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde vorgelegen.

Ort, Datum

Unterschrift

*Nichtzutreffendes bitte streichen

Belehrung:

Wer vorsätzlich gegen eine die Täuschung über Prüfungsleistungen betreffende Regelung einer Hochschulprüfungsordnung verstößt, handelt ordnungswidrig. Die Ordnungswidrigkeit kann mit einer Geldbuße von bis zu 50.000,00 € geahndet werden. Zuständige Verwaltungsbehörde für die Verfolgung und Ahndung von Ordnungswidrigkeiten ist der Kanzler/die Kanzlerin der Technischen Universität Dortmund. Im Falle eines mehrfachen oder sonstigen schwerwiegenden Täuschungsversuches kann der Prüfling zudem exmatrikuliert werden. (§ 63 Abs. 5 Hochschulgesetz - HG -)

Die Abgabe einer falschen Versicherung an Eides statt wird mit Freiheitsstrafe bis zu 3 Jahren oder mit Geldstrafe bestraft.

Die Technische Universität Dortmund wird gfls. elektronische Vergleichswerkzeuge (wie z.B. die Software „turnitin“) zur Überprüfung von Ordnungswidrigkeiten in Prüfungsverfahren nutzen.

Die oben stehende Belehrung habe ich zur Kenntnis genommen:

Ort, Datum

Unterschrift