

Martingalmethoden zur Berechnung von
Überschreitungswahrscheinlichkeiten
datenbasierter Sprungprozesse

Dissertation

zur Erlangung des Doktorgrades
an den Naturwissenschaftlichen Fachbereichen
(Mathematik)
der Justus-Liebig-Universität Gießen

vorgelegt von
David Heß

Betreuer
Prof. Dr. Winfried Stute

Gießen, den 18. November 2015

Danksagung

Mein Dank gilt allen voran meinem Betreuer Herrn Prof. Dr. Stute, der mir vor und während der Erstellung meiner Dissertation mit seiner Expertise und langjährigen wissenschaftlichen Erfahrung hilfreich zur Seite stand. Nach jeder unserer zahlreichen und ausgiebigen Diskussionen, die neben den fachlichen Angelegenheiten auch immer wieder die ein oder andere amüsante Anekdote aus Herrn Stutes Lebensweg beinhalteten, war ich stets etwas klüger.

Desweiteren danke ich allen meinen Kollegen am Mathematischen Institut der Justus-Liebig-Universität Gießen für die bemerkenswert gute Arbeitsatmosphäre. Insbesondere danke ich allen aktuellen und ehemaligen Mitgliedern der Arbeitsgruppe Stochastik, die ich immer als freundlich, uneigennützig und liebenswert erlebt habe und die bei sämtlichen Fragen und Problemen ihre Hilfe anboten.

Ein besonderer Dank gilt meinen Eltern Hubert und Renate, meinem Bruder Stefan und seiner Frau Sigrid sowie den übrigen Menschen aus meiner Familie und meinem Freundeskreis. Nur durch diese Menschen bin ich heute der, der ich bin, und nur mit ihnen an meiner Seite war es mir möglich, diese Arbeit zu verfassen.

Danke!

Inhaltsverzeichnis

1	Einführung	1
2	Der Single-Event-Prozess	5
2.1	Doob-Meyer-Zerlegungen für den Single-Event-Prozess	5
2.1.1	Doob-Meyer-Zerlegung des Single-Event-Prozesses	7
2.1.2	Doob-Meyer-Zerlegungen von gewissen Transformationen des Single- Event-Prozesses	9
2.1.3	Martingaltransformationen des Single-Event-Prozesses	16
3	Die Empirische Verteilungsfunktion	21
3.1	Doob-Meyer-Zerlegungen für die Empirische Verteilungsfunktion	22
3.1.1	Doob-Meyer-Zerlegung der Empirischen Verteilungsfunktion	23
3.1.2	Doob-Meyer-Zerlegung von Transformationen der Empirischen Vertei- lungsfunktion	26
3.1.3	Martingaltransformationen der Empirischen Verteilungsfunktion . . .	35
3.1.3.1	Affine Martingaltransformationen der Empirischen Vertei- lungsfunktion	35

Inhaltsverzeichnis

3.1.3.2	Polynomiale Martingaltransformationen der Empirischen Verteilungsfunktion	36
3.1.3.3	Exponentielle Martingaltransformationen der Empirischen Verteilungsfunktion	39
3.1.3.4	Allgemeine Martingaltransformationen der Empirischen Verteilungsfunktion	45
3.2	Reverse Doob-Meyer-Zerlegungen für die Empirische Verteilungsfunktion . . .	54
3.2.1	Reverse Doob-Meyer-Zerlegung der Empirischen Verteilungsfunktion .	54
3.2.2	Reverse Doob-Meyer-Zerlegung von Transformationen der Empirischen Verteilungsfunktion	59
3.2.3	Reverse Martingaltransformationen der Empirischen Verteilungsfunktion	67
3.2.3.1	Affine reverse Martingaltransformationen der Empirischen Verteilungsfunktion	67
3.2.3.2	Polynomiale reverse Martingaltransformationen der Empirischen Verteilungsfunktion	70
3.2.3.3	Exponentielle reverse Martingaltransformationen der Empirischen Verteilungsfunktion	73
3.2.3.4	Allgemeine reverse Martingaltransformationen der Empirischen Verteilungsfunktion	77
4	Überschreitungswahrscheinlichkeiten	87
4.1	Affine Schranken	88
4.2	Polynomiale Schranken	95

Inhaltsverzeichnis

4.3	Überschreitungswahrscheinlichkeiten für die Empirische Verteilungsfunktion .	100
5	Goodness-of-Fit-Tests für Typ-II-zensierte Daten	119
5.1	Goodness-of-Fit-Tests für Typ-II-rechtszensierte Daten	119
5.2	Goodness-of-Fit-Tests für Typ-II-linkszensierte Daten	121

1 Einführung

In vielen Fragestellungen der Statistik ist es hilfreich, die beobachtbaren Zufallsvariablen als die Summe zweier Teile zu betrachten. Dabei repräsentiert ein Teil üblicherweise ein angenommenes Modell und der andere weißes Rauschen. In der Regressionsanalyse schreibt man beispielsweise

$$Y_i = m(X_i) + \epsilon_i, \quad i = 1, \dots, n.$$

Dabei ergibt sich der Output Y_i jeweils aus der Summe des modellbasiert transformierten Inputs $m(X_i)$ und einem zentrierten Fehler ϵ_i (siehe z.B. Draper et al. [7]).

In der Theorie stochastischer Prozesse gilt die Doob-Meyer-Zerlegung (Doob [5, 6]) für viele \mathcal{F} -adaptierte Prozesse, wobei $\mathcal{F} = (\mathcal{F}_t)_{t \in \mathbb{R}}$ eine geeignete Filtration ist. Danach lässt sich solch ein \mathcal{F} -adaptierter Prozess $X = (X_t)_{t \in \mathbb{R}}$ als Summe eines (trendfreien) \mathcal{F} -Martingals $M = (M_t)_{t \in \mathbb{R}}$ und eines (\mathcal{F} -vorhersehbaren) Kompensators $A = (A_t)_{t \in \mathbb{R}}$ schreiben, d.h.

$$X_t = A_t + M_t, \quad t \in \mathbb{R}.$$

Nun lassen sich für das \mathcal{F} -Martingal M mit Hilfe von Stopptechniken Überschreitungswahrscheinlichkeiten berechnen, welche beispielsweise bei der Herleitung kritischer Bereiche sta-

1 Einführung

tistischer Tests Anwendung finden (siehe z.B. Koul und Stute [12]). Für den Prozess X sind derartige Überschreitungswahrscheinlichkeiten in der Regel (ausgenommen den Fall $A \equiv 0$) sehr viel schwieriger oder gar nicht berechenbar. Daher ist es sinnvoll, X geschickt zu transformieren, sodass der transformierte Prozess einerseits ein \mathcal{F} -Martingal ist und andererseits die für das unterstellte Modell relevanten Informationen enthält.

Ein bekanntes Beispiel derartiger sogenannter Martingaltransformationen sind solche von Brownschen Bewegungen. Die Doob-Meyer-Zerlegung bezüglich der natürlichen Filtration $\mathcal{F} = (\mathcal{F}_t)_{t \in [0, \infty)}$ der zugrundeliegenden Brownschen Bewegung $B = (B_t)_{t \in [0, \infty)}$ wird hierbei durch Itô's Formel (Itô [10]) geliefert. Sei dazu $\varphi : [0, \infty) \times \mathbb{R} \rightarrow \mathbb{R}$ stetig differenzierbar in der ersten Komponente, welche die Zeit repräsentiert, sowie zweimal stetig differenzierbar in der zweiten Komponente, welche für den Ort, also den ursprünglichen Prozess B , vorgesehen ist. Der Prozess $X = (X_t)_{t \in [0, \infty)}$ sei gegeben durch $X_t = \varphi(t, B_t)$ für $t \in [0, \infty)$. Dann gilt

$$X_t = X_0 + \int_0^t \left(\frac{\partial \varphi}{\partial s}(s, B_s) + \frac{1}{2} \cdot \frac{\partial^2 \varphi}{\partial x^2}(s, B_s) \right) ds + \int_0^t \frac{\partial \varphi}{\partial x}(s, B_s) dB_s, \quad t \in [0, \infty),$$

wobei $\frac{\partial \varphi}{\partial s}$ die Ableitung von φ nach der Zeitkomponente und $\frac{\partial \varphi}{\partial x}$ bzw. $\frac{\partial^2 \varphi}{\partial x^2}$ die Ableitungen nach der Ortskomponente bezeichnet. Nun ist das zweite Integral ein Martingal bezüglich \mathcal{F} und das erste Integral beschreibt den \mathcal{F} -vorhersehbaren Teil der Doob-Meyer-Zerlegung, welcher oft auch als Trend bezeichnet wird. Falls nun φ die partielle Differentialgleichung

$$\frac{\partial \varphi}{\partial s} + \frac{1}{2} \cdot \frac{\partial^2 \varphi}{\partial x^2} = 0 \tag{1.1}$$

erfüllt, so verschwindet das erste Integral und X ist selbst ein \mathcal{F} -Martingal. Die Gleichung (1.1) wird daher auch als \mathcal{F} -Martingaldifferentialgleichung bezeichnet.

1 Einführung

Die vorliegende Arbeit beschäftigt sich mit Transformationen bzw. Martingaltransformationen der Empirischen Verteilungsfunktion. Im Vergleich zur Brownschen Bewegung, welche stetige Pfade besitzt, werden hierbei also Sprungprozesse betrachtet. Und während die Brownsche Bewegung ein Prozess mit unabhängigen und stationären Zuwächsen ist, trifft dies auf die Empirische Verteilungsfunktion nicht zu. Außerdem ist die Empirische Verteilungsfunktion im Gegensatz zur Brownschen Bewegung kein Martingal bezüglich ihrer natürlichen Filtration. Eine ausführliche Übersicht zu sogenannten Zählprozessen - speziell in Hinsicht auf zensierte Daten - befindet sich in Andersen et al. [1]. In Shorack und Wellner [17] wird umfassend auf die Empirische Verteilungsfunktion und verwandte Prozesse eingegangen.

Sei $\varphi : \mathbb{R}^2 \rightarrow \mathbb{R}$ eine in der ersten Komponenten - der Zeitkomponenten - differenzierbare Abbildung, $F_n = (F_n(t))_{t \in \mathbb{R}}$ die Empirische Verteilungsfunktion von unabhängig und identisch verteilten Zufallsvariablen X_1, \dots, X_n mit $X_1 \sim F$, F stetig, und $X = (X_t)_{t \in \mathbb{R}}$ mit $X_t = \varphi(t, F_n(t))$. Für derartige Prozesse X wird die Doob-Meyer-Zerlegung bezüglich der natürlichen Filtration $\mathcal{F} = (\mathcal{F}_t)_{t \in \mathbb{R}}$ von F_n ebenfalls eine \mathcal{F} -Martingaldifferentialgleichung liefern. Deren Lösungen ergeben Martingaltransformationen bezüglich \mathcal{F} , welche mit Hilfe von Stoptechniken wiederum Überschreitungswahrscheinlichkeiten gewisser Transformationen der Empirischen Verteilungsfunktion ermöglichen. Als Ausgangspunkt für die Berechnung derartiger Überschreitungswahrscheinlichkeiten wird ein alternativer Beweis eines Ergebnisses dienen, welches auf Daniels [4] zurückgeht und von Rényi [15] bewiesen wurde. Die Beweismethoden lehnen an jenen an, welche auch zu gewissen Überschreitungswahrscheinlichkeiten für die Brownsche Bewegung und die Brownsche Brücke führen (vgl. Robbins [16] und Wald [19]).

1 Einführung

Ein weiteres Ziel wird sein, die Klasse der Martingaltransformationen $X = (X_t)_{t \in \mathbb{R}}$ der Form $X_t = \varphi(t, F_n(t))$ bezüglich der natürlichen Filtration \mathcal{F} von F_n vollständig zu charakterisieren. Ähnlich haben dies beispielsweise Nualart und Schoutens [14] schon für eine spezielle Klasse von Lévy Prozessen getan, welche jedoch die sehr einschränkende Forderung beinhaltet, dass der zugrundeliegende Prozess $Z = (Z_t)_{t \in \mathbb{R}}$ ein sogenanntes normales Martingal bezüglich der betrachteten Filtration $\mathcal{F} = (\mathcal{F}_t)_{t \in \mathbb{R}}$ ist. Dies bedeutet, dass für ein $c > 0$ neben Z auch der Prozess $Y = (Y_t)_{t \in \mathbb{R}}$ mit $Y_t = Z_t^2 - ct$ ein \mathcal{F} -Martingal sein muss. In vielen Anwendungen ist es jedoch nicht gegeben, dass der Kompensator der quadratischen Transformation linear oder überhaupt deterministisch ist.

Zu Beginn dieser Arbeit werden in Kapitel 2 anhand des Single-Event-Prozesses $S = (S_t)_{t \in \mathbb{R}}$ - also der Empirischen Verteilungsfunktion für ein einziges Datum - strukturelle Eigenschaften von Transformationen $X = (X_t)_{t \in \mathbb{R}}$ der Form $X_t = \varphi(t, S_t)$ untersucht. Dabei wird die Martingaldifferentialgleichung bezüglich \mathcal{F} für derartige Prozesse hergeleitet und die Menge ihrer Lösungen vollständig charakterisiert. Kapitel 3 enthält die Verallgemeinerung auf den Fall von n unabhängig und identisch verteilten Daten und betrachtet zudem die entsprechenden reversen Martingaltransformationen der Empirischen Verteilungsfunktion bezüglich ihrer natürlichen reversen Filtration $\mathcal{G} = (\mathcal{G}_t)_{t \in \mathbb{R}}$. Auch hierbei wird die Klasse der \mathcal{F} -Martingaltransformationen $X = (X_t)_{t \in \mathbb{R}}$ der Form $X_t = \varphi(t, F_n(t))$ und zudem die entsprechende Klasse der reversen \mathcal{G} -Martingaltransformationen vollständig beschrieben. Kapitel 4 widmet sich der Herleitung diverser Überschreitungswahrscheinlichkeiten der Empirischen Verteilungsfunktion und gewisser Transformationen. Letztere werden in Kapitel 5 genutzt, um Goodness-of-Fit-Tests für Typ-II-zensierte Daten zu entwerfen.

2 Der Single-Event-Prozess

Bevor Eigenschaften der Empirischen Verteilungsfunktion von n unabhängig und identisch verteilten Daten untersucht werden sollen, lohnt es sich zunächst den einfachsten Fall, $n = 1$, gesondert zu betrachten. In diesem Fall wird also genau ein Datum beobachtet. Sei $S = (S_t)_{t \in \mathbb{R}}$ mit $S_t = 1_{\{X \leq t\}}$ der sogenannte Single-Event-Prozess, wobei X eine beliebige Zufallsvariable ist. Der Single-Event-Prozess ist offensichtlich rechtsseitig stetig und monoton wachsend. Genauer gesagt ändert er seinen Wert nur an einer Stelle: Im Datum X springt er von 0 auf 1. Bezüglich seiner natürlichen Filtration $\mathcal{F} = (\mathcal{F}_t)_{t \in \mathbb{R}}$ mit $\mathcal{F}_t = \sigma(S_u : u \leq t)$ ist er aufgrund der Monotonie kein Martingal. Im folgenden wird nun die Doob-Meyer-Zerlegung von (S, \mathcal{F}) selbst sowie von transformierten Prozessen $X = (X_t)_{t \in \mathbb{R}}$ mit $X_t = \varphi(t, S_t)$, welche von Zeit und Ort des Single-Event-Prozesses abhängen, hergeleitet.

2.1 Doob-Meyer-Zerlegungen für den Single-Event-Prozess

Auch wenn der Single-Event-Prozess sehr simpel erscheint, lohnt es sich einen Blick auf seine Eigenschaften zu werfen, um auf einfache Weise einen ersten Eindruck von den Ei-

2 Der Single-Event-Prozess

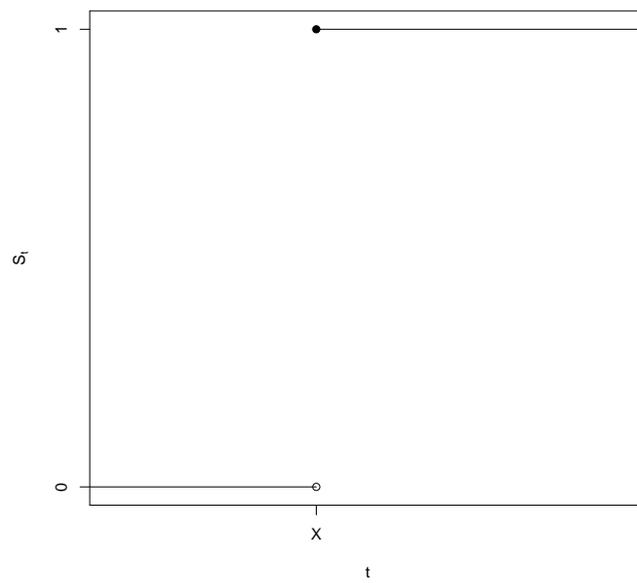


Abbildung 2.1: Der Single-Event-Prozess S der Zufallsvariablen X .

2 Der Single-Event-Prozess

genschaften der Empirischen Verteilungsfunktion zu gewinnen. Hier sollen nun die Doob-Meyer-Zerlegung des Single-Event-Prozesses und Doob-Meyer-Zerlegungen von Transformationen des Single-Event-Prozesses hergeleitet werden. Anhand letzterer lässt sich eine Martingaldifferentialgleichung bezüglich der natürlichen Filtration des Single-Event-Prozesses ableiten. Aufgrund der einfachen Struktur des Single-Event-Prozesses kann die Klasse der \mathcal{F} -Martingaltransformationen $X = (X_t)_{t \in \mathbb{R}}$ der Form $X_t = \varphi(t, S_t)$ auf einfache Weise vollständig charakterisiert werden.

2.1.1 Doob-Meyer-Zerlegung des Single-Event-Prozesses

Zunächst soll nun die Doob-Meyer-Zerlegung des Single-Event-Prozesses selbst hergeleitet werden. Man betrachte dazu eine Zufallsvariable $X \sim F$, wobei F die zugehörige stetige Verteilungsfunktion sei. Für ein festes $t \in \mathbb{R}$ mit $F(t) < 1$ betrachte man die Zerlegung des Intervalls $(-\infty, t]$ mittels $t_0 < \dots < t_K = t$ mit $K \in \mathbb{N}$. Es soll zuerst die Doob-Meyer-Zerlegung des diskretisierten Single-Event-Prozesses gefunden werden.

Satz 2.1. *Sei (S, \mathcal{F}) der Single-Event-Prozess mit natürlicher Filtration. Dann gilt für $k = 1, \dots, K$*

$$\mathbb{E}(S_{t_k} | \mathcal{F}_{t_{k-1}}) = S_{t_{k-1}} + (1 - S_{t_{k-1}}) \frac{F(t_k) - F(t_{k-1})}{1 - F(t_{k-1})}.$$

Beweis. Es gilt

$$\begin{aligned} \mathbb{E}(S_{t_k} | \mathcal{F}_{t_{k-1}}) &= \mathbb{E}(1_{\{X \leq t_k\}} | \mathcal{F}_{t_{k-1}}) \\ &= 1_{\{X \leq t_{k-1}\}} + \mathbb{E}(1_{\{t_{k-1} < X \leq t_k\}} | \mathcal{F}_{t_{k-1}}) \\ &= 1_{\{X \leq t_{k-1}\}} + 1_{\{X > t_{k-1}\}} \mathbb{E}(1_{\{t_{k-1} < X \leq t_k\}} | X > t_{k-1}) \end{aligned}$$

2 Der Single-Event-Prozess

$$\begin{aligned}
&= 1_{\{X \leq t_{k-1}\}} + 1_{\{X > t_{k-1}\}} \mathbb{P}(t_{k-1} < X \leq t_k | X > t_{k-1}) \\
&= 1_{\{X \leq t_{k-1}\}} + 1_{\{X > t_{k-1}\}} \frac{\mathbb{P}(t_{k-1} < X \leq t_k)}{\mathbb{P}(X > t_{k-1})} \\
&= 1_{\{X \leq t_{k-1}\}} + 1_{\{X > t_{k-1}\}} \frac{F(t_k) - F(t_{k-1})}{1 - F(t_{k-1})}.
\end{aligned}$$

□

Man betrachte den beliebigen Startzeitpunkt $t_0 \in \mathbb{R}$. Bezüglich der diskreten Filtration $\mathcal{F}^K = (\mathcal{F}_{t_k}^K)_{k \in \{0, \dots, K\}}$ mit $\mathcal{F}_{t_k}^K = \sigma(S_{t_l} : l \leq k)$ gilt für das Innovationsmartingal $M^{0,K} = (M_{t_k}^{0,K})_{k \in \{0, \dots, K\}}$ der Doob-Meyer-Zerlegung von (S^K, \mathcal{F}^K) , wobei $S^K = (S_{t_k})_{k \in \{0, \dots, K\}}$,

$$\begin{aligned}
M_{t_k}^{0,K} &= M_{t_{k-1}}^{0,K} + S_{t_k} - \mathbb{E}\left(S_{t_k} | \mathcal{F}_{t_{k-1}}^K\right) \\
&= M_{t_{k-1}}^{0,K} + 1_{\{t_{k-1} < X \leq t_k\}} - 1_{\{X > t_{k-1}\}} \frac{F(t_k) - F(t_{k-1})}{1 - F(t_{k-1})}.
\end{aligned}$$

Setze $M_{t_0}^{0,K} := S_{t_0}$. Dann gilt

$$M_{t_k}^{0,K} = S_{t_k} - \sum_{j=1}^k \frac{1 - S_{t_{j-1}}}{1 - F(t_{j-1})} (F(t_j) - F(t_{j-1})) \quad \text{für } k = 0, \dots, K.$$

Für den \mathcal{F}^K -Kompensator $A^{0,K} = (A_{t_k}^{0,K})_{k \in \{0, \dots, K\}}$ gilt demzufolge

$$A_{t_k}^{0,K} = \sum_{j=1}^k \frac{1 - S_{t_{j-1}}}{1 - F(t_{j-1})} (F(t_j) - F(t_{j-1})) \quad \text{für } k = 0, \dots, K.$$

Damit ist die diskrete Version der Doob-Meyer-Zerlegung gefunden und für die zeitstetige Version ergibt sich nach Grenzübergang:

2 Der Single-Event-Prozess

Satz 2.2. *Der Single-Event Prozess mit seiner natürlichen Filtration (S, \mathcal{F}) besitzt das Innovationsmartingal (M^0, \mathcal{F}) und den Kompensator (A^0, \mathcal{F}) , wobei $M^0 = (M_t^0)_{t \in \mathbb{R}}$ mit*

$$M_t^0 = S_t - \int_{(-\infty, t]} \frac{1 - S_{u-}}{1 - F(u)} F(du)$$

und $A^0 = (A_t^0)_{t \in \mathbb{R}}$ mit

$$A_t^0 = \int_{(-\infty, t]} \frac{1 - S_{u-}}{1 - F(u)} F(du).$$

Dabei bezeichnet $S_{t-} := \lim_{u \uparrow t} S_u$ den linksseitigen Grenzwert von S in t .

2.1.2 Doob-Meyer-Zerlegungen von gewissen Transformationen des Single-Event-Prozesses

Für eine beliebige Funktion $\varphi : \mathbb{R}^2 \rightarrow \mathbb{R}$ mit $\varphi(\cdot, x)$ differenzierbar für $x \in \{0, 1\}$ betrachte man den Prozess $X = (X_t)_{t \in \mathbb{R}}$ mit $X_t = \varphi(t, S_t)$, wobei S der Single-Event-Prozess sei. Für ein festes $t \in \mathbb{R}$ mit $F(t) < 1$ betrachte man wiederum die Zerlegung des Intervalls $(-\infty, t]$ mittels $t_0 < \dots < t_K = t$ mit $K \in \mathbb{N}$.

Satz 2.3. *Sei (X, \mathcal{F}) die oben genannte Transformation des Single-Event-Prozesses mit der natürlichen Filtration des Single-Event-Prozesses. Dann gilt für $k = 1, \dots, K$*

$$\mathbb{E}(X_{t_k} | \mathcal{F}_{t_{k-1}}) = \varphi(t_k, S_{t_{k-1}}) + (\varphi(t_k, 1) - \varphi(t_k, 0)) \frac{1 - S_{t_{k-1}}}{1 - F(t_{k-1})} (F(t_k) - F(t_{k-1})).$$

2 Der Single-Event-Prozess

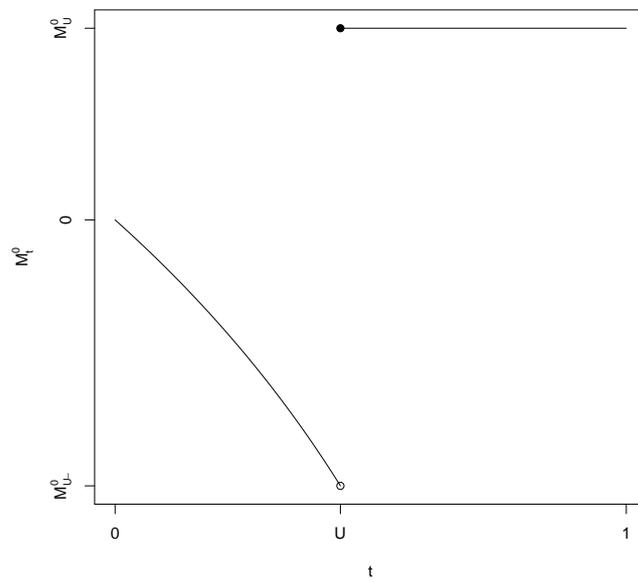


Abbildung 2.2: Das Doob-Meyer-Martingal (M^0, \mathcal{F}) des Single-Event-Prozesses einer auf $[0, 1]$ gleichverteilten Zufallsvariablen U .

2 Der Single-Event-Prozess

Beweis. Es gilt

$$\begin{aligned}
\mathbb{E}(\varphi(t_k, S_{t_k}) | \mathcal{F}_{t_{k-1}}) &= \mathbb{E}(\varphi(t_k, 1_{\{X \leq t_k\}}) | \mathcal{F}_{t_{k-1}}) \\
&= \mathbb{E}(1_{\{X \leq t_{k-1}\}} \varphi(t_k, 1_{\{X \leq t_k\}}) + 1_{\{X > t_{k-1}\}} \varphi(t_k, 1_{\{X \leq t_k\}}) | \mathcal{F}_{t_{k-1}}) \\
&= \mathbb{E}(1_{\{X \leq t_{k-1}\}} \varphi(t_k, 1) + 1_{\{X > t_{k-1}\}} \varphi(t_k, 1_{\{t_{k-1} < X \leq t_k\}}) | \mathcal{F}_{t_{k-1}}) \\
&= 1_{\{X \leq t_{k-1}\}} \varphi(t_k, 1) + 1_{\{X > t_{k-1}\}} \mathbb{E}(\varphi(t_k, 1_{\{t_{k-1} < X \leq t_k\}}) | X > t_{k-1}) \\
&= 1_{\{X \leq t_{k-1}\}} \varphi(t_k, 1) + 1_{\{X > t_{k-1}\}} \frac{\mathbb{E}(1_{\{X > t_{k-1}\}} \varphi(t_k, 1_{\{t_{k-1} < X \leq t_k\}}))}{\mathbb{P}(X > t_{k-1})} \\
&= 1_{\{X \leq t_{k-1}\}} \varphi(t_k, 1) + 1_{\{X > t_{k-1}\}} \frac{\mathbb{E}(1_{\{t_{k-1} < X \leq t_k\}} \varphi(t_k, 1) + 1_{\{X > t_k\}} \varphi(t_k, 0))}{\mathbb{P}(X > t_{k-1})} \\
&= 1_{\{X \leq t_{k-1}\}} \varphi(t_k, 1) + 1_{\{X > t_{k-1}\}} \frac{\varphi(t_k, 1) \mathbb{P}(t_{k-1} < X \leq t_k) + \varphi(t_k, 0) \mathbb{P}(X > t_k)}{\mathbb{P}(X > t_{k-1})} \\
&= 1_{\{X \leq t_{k-1}\}} \varphi(t_k, 1) + 1_{\{X > t_{k-1}\}} \frac{\varphi(t_k, 1) (F(t_k) - F(t_{k-1})) + \varphi(t_k, 0) (1 - F(t_k))}{1 - F(t_{k-1})} \\
&= 1_{\{X \leq t_{k-1}\}} \varphi(t_k, 1) + 1_{\{X > t_{k-1}\}} \varphi(t_k, 1) \frac{F(t_k) - F(t_{k-1})}{1 - F(t_{k-1})} \\
&\quad + 1_{\{X > t_{k-1}\}} \varphi(t_k, 0) \frac{1 - F(t_k)}{1 - F(t_{k-1})} \\
&= 1_{\{X \leq t_{k-1}\}} \varphi(t_k, 1) + 1_{\{X > t_{k-1}\}} \varphi(t_k, 1) \frac{F(t_k) - F(t_{k-1})}{1 - F(t_{k-1})} \\
&\quad - 1_{\{X > t_{k-1}\}} \varphi(t_k, 0) \left(\frac{F(t_k) - F(t_{k-1})}{1 - F(t_{k-1})} - 1 \right) \\
&= 1_{\{X \leq t_{k-1}\}} \varphi(t_k, 1) + 1_{\{X > t_{k-1}\}} \varphi(t_k, 0) \\
&\quad + 1_{\{X > t_{k-1}\}} (\varphi(t_k, 1) - \varphi(t_k, 0)) \frac{F(t_k) - F(t_{k-1})}{1 - F(t_{k-1})} \\
&= \varphi(t_k, 1_{\{X \leq t_{k-1}\}}) + 1_{\{X > t_{k-1}\}} (\varphi(t_k, 1) - \varphi(t_k, 0)) \frac{F(t_k) - F(t_{k-1})}{1 - F(t_{k-1})} \\
&= \varphi(t_k, S_{t_{k-1}}) + (\varphi(t_k, 1) - \varphi(t_k, 0)) \frac{1 - S_{t_{k-1}}}{1 - F(t_{k-1})} (F(t_k) - F(t_{k-1})).
\end{aligned}$$

2 Der Single-Event-Prozess

□

Man betrachte den beliebigen Startzeitpunkt $t_0 \in \mathbb{R}$. Bezüglich der diskreten Filtration $\mathcal{F}^K = (\mathcal{F}_{t_k}^K)_{k \in \{0, \dots, K\}}$ mit $\mathcal{F}_{t_k}^K = \sigma(S_{t_l} : l \leq k)$ gilt für das Innovationsmartingal $M^K = (M_{t_k}^K)_{k \in \{0, \dots, K\}}$ der Doob-Meyer-Zerlegung von (X^K, \mathcal{F}^K) , wobei $X^K = (X_{t_k})_{k \in \{0, \dots, K\}}$,

$$\begin{aligned}
 M_{t_k}^K &= M_{t_{k-1}}^K + X_{t_k} - \mathbb{E}(X_{t_k} | \mathcal{F}_{t_{k-1}}) \\
 &= M_{t_{k-1}}^K + \varphi(t_k, 1_{\{X \leq t_k\}}) - \varphi(t_k, 1_{\{X \leq t_{k-1}\}}) \\
 &\quad - 1_{\{X > t_{k-1}\}} (\varphi(t_k, 1) - \varphi(t_k, 0)) \frac{F(t_k) - F(t_{k-1})}{1 - F(t_{k-1})} \\
 &= M_{t_{k-1}}^K + 1_{\{t_{k-1} < X \leq t_k\}} (\varphi(t_k, 1) - \varphi(t_k, 0)) \\
 &\quad - 1_{\{X > t_{k-1}\}} (\varphi(t_k, 1) - \varphi(t_k, 0)) \frac{F(t_k) - F(t_{k-1})}{1 - F(t_{k-1})} \\
 &= M_{t_{k-1}}^K + (\varphi(t_k, 1) - \varphi(t_k, 0)) \left(1_{\{t_{k-1} < X \leq t_k\}} - 1_{\{X > t_{k-1}\}} \frac{F(t_k) - F(t_{k-1})}{1 - F(t_{k-1})} \right) \\
 &= M_{t_{k-1}}^K + (\varphi(t_k, 1) - \varphi(t_k, 0)) (M_{t_k}^{K,0} - M_{t_{k-1}}^{K,0}).
 \end{aligned}$$

Setze $M_{t_0}^K := X_{t_0}$. Dann gilt für $k = 1, \dots, K$

$$M_{t_k}^K = X_{t_0} + \sum_{j=1}^k (\varphi(t_j, 1) - \varphi(t_j, 0)) (M_{t_j}^{K,0} - M_{t_{j-1}}^{K,0}).$$

Nach Grenzübergang ergibt sich für die zeitstetige Version:

Satz 2.4. *Sei $t_0 \in \mathbb{R}$ ein beliebiger Startwert. Der Prozess $X = (X_t)_{t \in [t_0, \infty)}$ mit $X_t = \varphi(t, S_t)$ besitzt das Innovationsmartingal (M, \mathcal{F}) , wobei $\mathcal{F} = (\mathcal{F}_t)_{t \in [t_0, \infty)}$ mit $\mathcal{F}_t = \sigma(S_u : u \leq t)$*

2 Der Single-Event-Prozess

und $M = (M_t)_{t \in [t_0, \infty)}$ mit

$$M_t = X_{t_0} + \int_{(t_0, t]} \varphi(u, 1) - \varphi(u, 0) dM_u^0.$$

Korollar 2.5. *Es existiere der Grenzwert*

$$x_{-\infty} := \lim_{s \rightarrow -\infty} \varphi(s, 0).$$

Dann besitzt der Prozess $X = (X_t)_{t \in \mathbb{R}}$ mit $X_t = \varphi(t, S_t)$ das Innovationsmartingal (M, \mathcal{F}) , wobei $\mathcal{F} = (\mathcal{F}_t)_{t \in \mathbb{R}}$ mit $\mathcal{F}_t = \sigma(S_u : u \leq t)$ und $M = (M_t)_{t \in \mathbb{R}}$ mit

$$M_t = x_{-\infty} + \int_{(-\infty, t]} \varphi(u, 1) - \varphi(u, 0) dM_u^0.$$

Weiterhin liefert die Doob-Meyer-Zerlegung von (X^K, \mathcal{F}^K) den Kompensator (A^K, \mathcal{F}^K) , wobei $A^K = (A_{t_k}^K)_{k \in \{0, \dots, K\}}$ mit

$$\begin{aligned} A_{t_k}^K &= A_{t_{k-1}}^K - X_{t_{k-1}} + \mathbb{E}(X_{t_k} | \mathcal{F}_{t_{k-1}}) \\ &= A_{t_{k-1}}^K - \varphi(t_{k-1}, S_{t_{k-1}}) + \varphi(t_k, S_{t_{k-1}}) \\ &\quad + (\varphi(t_k, 1) - \varphi(t_k, 0)) \frac{1 - S_{t_{k-1}}}{1 - F(t_{k-1})} (F(t_k) - F(t_{k-1})) \\ &= A_{t_{k-1}}^K + \varphi(t_k, S_{t_{k-1}}) - \varphi(t_{k-1}, S_{t_{k-1}}) + (\varphi(t_k, 1) - \varphi(t_k, 0)) (A_{t_k}^{K,0} - A_{t_{k-1}}^{K,0}). \end{aligned}$$

2 Der Single-Event-Prozess

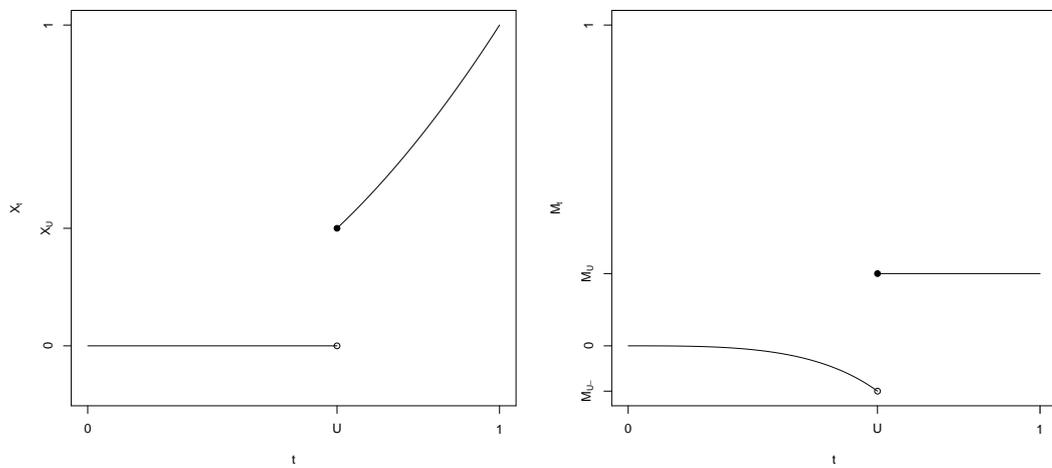


Abbildung 2.3: Der Prozess X mit $X_t = t^2 S_t$ und sein Doob-Meyer-Martingal (M, \mathcal{F}) , wobei die unterliegende Zufallsvariable U gleichverteilt auf $[0, 1]$ ist.

2 Der Single-Event-Prozess

Durch Induktion ergibt sich

$$A_{t_k}^K = \sum_{j=1}^k \varphi(t_j, S_{t_{j-1}}) - \varphi(t_{j-1}, S_{t_{j-1}}) + \sum_{j=1}^k (\varphi(t_j, 1) - \varphi(t_j, 0)) (A_{t_j}^{K,0} - A_{t_{j-1}}^{K,0}).$$

Nach Grenzübergang ergibt sich für den zeitstetigen Kompensator:

Satz 2.6. *Der Prozess $X = (X_t)_{t \in \mathbb{R}}$ mit $X_t = \varphi(t, S_t)$ besitzt den Kompensator (A, \mathcal{F}) , wobei $\mathcal{F} = (\mathcal{F}_t)_{t \in \mathbb{R}}$ mit $\mathcal{F}_t = \sigma(S_u : u \leq t)$ und $A = (A_t)_{t \in \mathbb{R}}$ mit*

$$A_t = \int_{(-\infty, t]} \frac{\partial \varphi}{\partial F}(u, S_{u-}) + (\varphi(u, 1) - \varphi(u, 0)) \frac{1 - S_{u-}}{1 - F(u)} F(du).$$

Die \mathcal{F} -Doob-Meyer-Zerlegung von X liefert also ein Martingal, welches aus einem Startwert und einem Integral bezüglich M^0 - dem \mathcal{F} -Doob-Meyer-Martingal des Single-Event-Prozesses - besteht, und einen Kompensator, welcher als Integral bezüglich F , also einer deterministischen Funktion, geschrieben werden kann. Damit ähnelt sie der Darstellung von Transformationen der Brownschen Bewegung in Itô's Formel (Itô [10]). Dort ergibt sich eine Summe bestehend aus einem Startwert, einem Integral bezüglich der Brownschen Bewegung und einem Lebesgue-Integral. Das Integral bezüglich der Brownschen Bewegung gemeinsam mit dem Startwert ist ein Martingal. Das Lebesgue-Integral repräsentiert den Kompensator. Verschwindet dessen Integrand, so ist der transformierte Prozess selbst ein Martingal. Daraus lässt sich eine Martingaldifferentialgleichung ableiten, deren Lösungen zu eben solchen Martingaltransformationen führen. Dies soll im Folgenden auch für den Single-Event-Prozess geschehen.

2.1.3 Martingaltransformationen des Single-Event-Prozesses

Falls der Integrand im Kompensator aus Satz 2.6 verschwindet, so ist (X, \mathcal{F}) ein Martingal. Daraus leitet sich die folgende \mathcal{F} -Martingaldifferentialgleichung ab, deren Lösungen im Anschluss hergeleitet werden.

Satz 2.7. *Der Prozess $X = (X_t)_{t \in \mathbb{R}}$ mit $X_t = \varphi(t, S_t)$ ist ein Martingal bzgl. $\mathcal{F} = (\mathcal{F}_t)_{t \in \mathbb{R}}$ mit $\mathcal{F}_t = \sigma(S_u : u \leq t)$, falls für alle $t \in \mathbb{R}$ und $x \in \{0, 1\}$ gilt*

$$\frac{\partial \varphi}{\partial F}(t, x) + (\varphi(t, 1) - \varphi(t, 0)) \frac{1 - x}{1 - F(t)} = 0.$$

Im Folgenden werden zunächst affine Lösungen der Martingaldifferentialgleichung aus Satz 2.7 gesucht, wobei affine Lösung bedeutet, dass für $t \in \mathbb{R}$ die Funktion $\varphi(t, \cdot)$ affin ist. Diese Lösungen führen zu folgenden \mathcal{F} -Martingaltransformationen des Single-Event-Prozesses.

Satz 2.8. *Seien $c_0, c_1 \in \mathbb{R}$. Dann ist der Prozess $X = (X_t)_{t \in \mathbb{R}}$ mit*

$$X_t = c_1 \frac{1 - S_t}{1 - F(t)} + c_0$$

ein Martingal bzgl. $\mathcal{F} = (\mathcal{F}_t)_{t \in \mathbb{R}}$ mit $\mathcal{F}_t = \sigma(S_u : u \leq t)$.

Beweis. Sei

$$\varphi(t, x) = a_1(F(t))x + a_0(F(t)).$$

Dann gilt

$$\frac{\partial \varphi}{\partial F}(t, x) = \frac{\partial a_1}{\partial F}(F(t))x + \frac{\partial a_0}{\partial F}(F(t))$$

2 Der Single-Event-Prozess

und $X_t = \varphi(t, S_t)$ ist ein Martingal bzgl. \mathcal{F} , falls

$$\left(\frac{\partial a_1}{\partial F}(F(t)) - \frac{a_1(F(t))}{1-F(t)} \right) x + \frac{\partial a_0}{\partial F}(F(t)) + \frac{a_1(F(t))}{1-F(t)} = 0.$$

Nun gilt

$$\begin{aligned} \frac{\partial a_1}{\partial F}(F(t)) - \frac{a_1(F(t))}{1-F(t)} = 0 &\Leftrightarrow \frac{\partial \ln(|a_1(F(t))|)}{\partial F} = \frac{1}{1-F(t)} \\ \Leftrightarrow \ln(|a_1(F(t))|) &= \ln\left(\frac{1}{1-F(t)}\right) + c \quad \text{für ein } c \in \mathbb{R} \\ \Leftrightarrow a_1(F(t)) &= \frac{c}{1-F(t)} \quad \text{für ein } c \in \mathbb{R} \\ \Leftrightarrow a_1(F(t)) &= -\frac{c_1}{1-F(t)} \quad \text{für ein } c_1 \in \mathbb{R} \end{aligned}$$

sowie

$$\begin{aligned} \frac{\partial a_0}{\partial F}(F(t)) + \frac{a_1(F(t))}{1-F(t)} = 0 &\Leftrightarrow \frac{\partial a_0}{\partial F}(F(t)) = \frac{c_1}{(1-F(t))^2} \quad \text{für ein } c_1 \in \mathbb{R} \\ \Leftrightarrow a_0(F(t)) &= \frac{c_1}{1-F(t)} + c_0 \quad \text{für } c_0, c_1 \in \mathbb{R}. \end{aligned}$$

Also ist für $c_0, c_1 \in \mathbb{R}$

$$X_t = c_1 \frac{1 - S_t}{1 - F(t)} + c_0$$

ein Martingal bzgl. \mathcal{F} .

□

Aufgrund der einfachen Struktur des Single-Event-Prozesses gibt es abgesehen von den genannten affinen \mathcal{F} -Martingaltransformationen keine weiteren \mathcal{F} -Martingale dieser Form.

2 Der Single-Event-Prozess

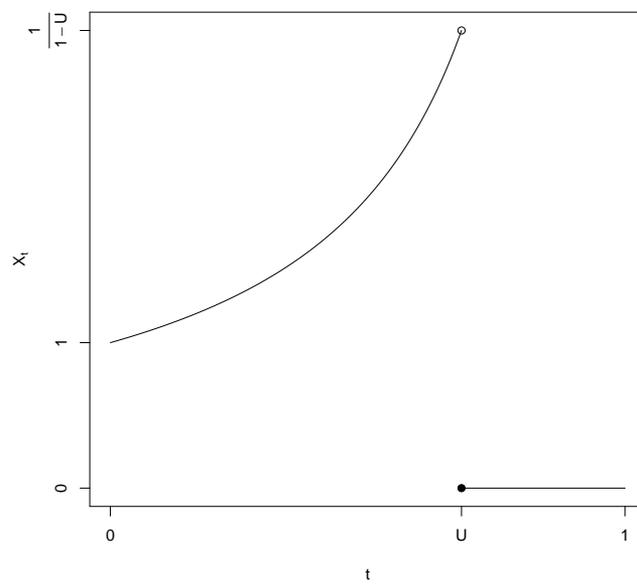


Abbildung 2.4: Das Martingal (X, \mathcal{F}) mit $X_t = \frac{1-S_t}{1-t}$, wobei die unterliegende Zufallsvariable U gleichverteilt ist auf $[0, 1]$.

2 Der Single-Event-Prozess

Satz 2.9. *Alle Martingaltransformationen $X = (X_t)_{t \in \mathbb{R}}$ des Single-Event-Prozesses $S = (S_t)_{t \in \mathbb{R}}$ bzgl. $\mathcal{F} = (\mathcal{F}_t)_{t \in \mathbb{R}}$ haben die Form*

$$X_t = c_1 \frac{1 - S_t}{1 - F(t)} + c_0$$

mit $c_0, c_1 \in \mathbb{R}$.

Beweis. Der Single-Event-Prozess S kennt nur zwei Zustände. Jeder Prozess X mit $X_t = \varphi(t, S_t)$ lässt sich also wie folgt schreiben:

$$X_t = \varphi_0(t)(1 - S_t) + \varphi_1(t)S_t.$$

Nach Eintritt des Events gilt $X_t = \varphi_1(t)$. Von da an ist X also eine deterministische Funktion. Damit X ein Martingal bezüglich \mathcal{F} ist, muss also φ_1 konstant sein. Sei also $\varphi_1 \equiv a \in \mathbb{R}$. Falls X ein \mathcal{F} -Martingal ist, gilt nun $X_t = \varphi_0(t)(1 - S_t) + aS_t$ und damit für $t \geq s$

$$\begin{aligned} \varphi_0(s) &= \mathbb{E}(X_t | X_s = \varphi_0(s)) \\ &= \varphi_0(t) \mathbb{P}(X > t | X > s) + a \mathbb{P}(X \leq t | X > s) \\ &= \varphi_0(t) \frac{\mathbb{P}(X > t)}{\mathbb{P}(X > s)} + a \frac{\mathbb{P}(s < X \leq t)}{\mathbb{P}(X > s)} \\ &= \varphi_0(t) \frac{1 - F(t)}{1 - F(s)} + a \frac{F(t) - F(s)}{1 - F(s)}. \end{aligned}$$

Hierbei muss gelten

$$\varphi_0(t) = \varphi_0(s) \frac{1 - F(s)}{1 - F(t)} - a \frac{F(t) - F(s)}{1 - F(t)}. \quad (2.1)$$

Man betrachte den Startpunkt $s_0 = \sup_{s \in \mathbb{R}} \{F(s) = 0\}$, wobei $s_0 = -\infty$, falls $F(s) > 0$ für alle

2 Der Single-Event-Prozess

$s \in \mathbb{R}$. Nun gilt (2.1) auch für $s = s_0$, also

$$\varphi_0(t) = \varphi_0(s_0) \frac{1}{1 - F(t)} - a \frac{F(t)}{1 - F(t)}$$

bzw. falls $s_0 = -\infty$

$$\varphi_0(t) = \lim_{s \rightarrow -\infty} \varphi_0(s) \frac{1}{1 - F(t)} - a \frac{F(t)}{1 - F(t)}.$$

Setze $\varphi_0(s_0) = b \in \mathbb{R}$, falls $s_0 \in \mathbb{R}$, bzw. $\lim_{s \rightarrow -\infty} \varphi_0(s) = b \in \mathbb{R}$, falls $s_0 = -\infty$. Dann folgt

also

$$\varphi_0(t) = b \frac{1}{1 - F(t)} - a \frac{F(t)}{1 - F(t)}$$

und schließlich

$$X_t = (b - a) \frac{1 - S_t}{1 - F(t)} + a.$$

Dies sind also die einzigen \mathcal{F} -Martingale, welche sich aus S gewinnen lassen.

□

3 Die Empirische Verteilungsfunktion

Der Single-Event-Prozess ist ein Spezialfall der Empirischen Verteilungsfunktion. Für $n \in \mathbb{N}$ unabhängig und identisch verteilte Daten repräsentiert diese die Verteilung, die jedem Datum die Masse $\frac{1}{n}$ zuspricht. Seien also X_1, \dots, X_n unabhängig und identisch verteilte Zufallsvariablen mit $X_1 \sim F$, wobei F eine stetige Verteilungsfunktion sei. Dann ist $F_n = (F_n(t))_{t \in \mathbb{R}}$ mit $F_n(t) = \frac{1}{n} \sum_{i=1}^n 1_{\{X_i \leq t\}}$ die Empirische Verteilungsfunktion. Sie ist ein konsistenter und erwartungstreuer Schätzer der wahren Verteilungsfunktion F (siehe Glivenko [9] und Cantelli [3]).

Die Empirische Verteilungsfunktion ist offensichtlich rechtsseitig stetig und monoton wachsend. Bezüglich ihrer natürlichen Filtration $\mathcal{F} = (\mathcal{F}_t)_{t \in \mathbb{R}}$ mit $\mathcal{F}_t = \sigma(F_n(u) : u \leq t)$ ist sie aufgrund der Monotonie kein Martingal. Im folgenden wird die Doob-Meyer-Zerlegung von (F_n, \mathcal{F}) selbst sowie von Prozessen $X = (X_t)_{t \in \mathbb{R}}$ mit $X_t = \varphi(t, F_n(t))$, welche von Zeit und Ort der Empirischen Verteilungsfunktion abhängen, hergeleitet. Anschließend werden mit Hilfe der Doob-Meyer-Zerlegung bezüglich \mathcal{F} Martingaltransformationen der Empirischen Verteilungsfunktion bestimmt. Das Kapitel schließt mit der Herleitung der reversen Doob-Meyer-Zerlegungen von (F_n, \mathcal{G}) und $(\varphi(\cdot, F_n(\cdot)), \mathcal{G})$, wobei $\mathcal{G} = (\mathcal{G}_t)_{t \in \mathbb{R}}$ mit $\mathcal{G}_t = \sigma(F_n(u) : u \geq t)$ die natürliche reverse Filtration bezeichnet, und den dazugehörigen

3 Die Empirische Verteilungsfunktion

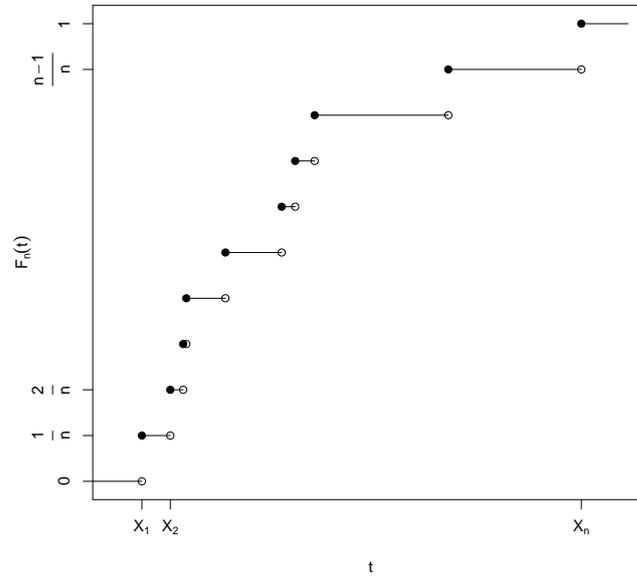


Abbildung 3.1: Die Empirische Verteilungsfunktion F_n der Zufallsvariablen X_1, \dots, X_n .

reversen \mathcal{G} -Martingaltransformationen.

3.1 Doob-Meyer-Zerlegungen für die Empirische Verteilungsfunktion

In diesem Abschnitt werden die Ergebnisse aus Kapitel 2 auf die Empirische Verteilungsfunktion ausgeweitet. Dabei wird erst die \mathcal{F} -Doob-Meyer-Zerlegung der Empirischen Verteilungsfunktion selbst und anschließend die \mathcal{F} -Doob-Meyer-Zerlegung von Transformatio-

3 Die Empirische Verteilungsfunktion

nen $X = (X_t)_{t \in \mathbb{R}}$ der Form $X_t = \varphi(t, F_n(t))$ hergeleitet. Aus letzterer wird sich wiederum eine \mathcal{F} -Martingaldifferentialgleichung ergeben, deren Lösungen zum Abschluss des Abschnitts erarbeitet werden. Dabei wird sich herausstellen, dass sich die Klasse der \mathcal{F} -Martingaltransformationen der genannten Form vollständig charakterisieren lässt.

3.1.1 Doob-Meyer-Zerlegung der Empirischen Verteilungsfunktion

Zunächst soll die Doob-Meyer-Zerlegung der Empirischen Verteilungsfunktion selbst hergeleitet werden. Man betrachte die unabhängig und identisch verteilten Zufallsvariablen X_1, \dots, X_n , $n \in \mathbb{N}$, mit $X_1 \sim F$, wobei F eine stetige Verteilungsfunktion sei. Für ein festes $t \in \mathbb{R}$ mit $F(t) < 1$ betrachte man die Zerlegung des Intervalls $(-\infty, t]$ mittels $t_0 < \dots < t_K = t$ mit $K \in \mathbb{N}$. Zunächst soll die Doob-Meyer-Zerlegung der diskretisierten Empirischen Verteilungsfunktion gefunden werden.

Satz 3.1. *Sei (F_n, \mathcal{F}) die Empirische Verteilungsfunktion mit natürlicher Filtration. Dann gilt für $k = 1, \dots, K$*

$$\mathbb{E}(F_n(t_k) | \mathcal{F}_{t_{k-1}}) = F_n(t_{k-1}) + (1 - F_n(t_{k-1})) \frac{F(t_k) - F(t_{k-1})}{1 - F(t_{k-1})}.$$

Beweis. Die Mengen

$$\{X_i \leq t_0\}, \{t_0 < X_i \leq t_1\}, \dots, \{t_{k-1} < X_i \leq t_k\}, \{X_i > t_k\}$$

bilden eine Partition des Grundraumes für jedes $i = 1, \dots, n$. Da X_1, \dots, X_n unabhängig und identisch verteilt sind, ist die gemeinsame Verteilung der mit n multiplizierten Zuwächse

3 Die Empirische Verteilungsfunktion

der Empirischen Verteilungsfunktion eine Multinomialverteilung. Es gilt also

$$\begin{aligned} & n(F_n(t_0), F_n(t_1) - F_n(t_0), \dots, F_n(t_k) - F_n(t_{k-1}), 1 - F_n(t_k)) \\ & \sim \text{Mult}(n, (F(t_0), F(t_1) - F(t_0), \dots, F(t_k) - F(t_{k-1}), 1 - F(t_k))). \end{aligned}$$

Damit gilt für $0 \leq m_0 \leq \dots \leq m_k \leq n$

$$\begin{aligned} & \mathbb{P}\left(F_n(t_l) = \frac{m_l}{n} \text{ für alle } l = 0, \dots, k\right) \\ & = \mathbb{P}\left(F_n(t_1) = \frac{m_1}{n}, F_n(t_1) - F_n(t_0) = \frac{m_2 - m_1}{n}, \dots, F_n(t_k) - F_n(t_{k-1}) = \frac{m_k - m_{k-1}}{n}\right) \\ & = \binom{n}{m_0 \quad m_1 - m_0 \quad \dots \quad m_k - m_{k-1} \quad n - m_k} \\ & \cdot F(t_0)^{m_1} \prod_{l=0}^{k-1} (F(t_{l+1}) - F(t_l))^{m_{l+1} - m_l} (1 - F(t_k))^{n - m_k} \end{aligned}$$

und für die bedingte Verteilung

$$\begin{aligned} & \mathbb{P}\left(F_n(t_k) = \frac{m_k}{n} \mid F_n(t_l) = \frac{m_l}{n} \text{ für alle } l = 0, \dots, k-1\right) \\ & = \frac{\mathbb{P}\left(F_n(t_l) = \frac{m_l}{n} \text{ für alle } l = 0, \dots, k\right)}{\mathbb{P}\left(F_n(t_l) = \frac{m_l}{n} \text{ für alle } l = 0, \dots, k-1\right)} \\ & = \binom{n - m_{k-1}}{m_k - m_{k-1}} \left(\frac{F(t_k) - F(t_{k-1})}{1 - F(t_{k-1})}\right)^{m_k - m_{k-1}} \left(\frac{1 - F(t_k)}{1 - F(t_{k-1})}\right)^{n - m_k} \\ & = \mathbb{P}\left(F_n(t_k) = \frac{m_k}{n} \mid F_n(t_{k-1}) = \frac{m_{k-1}}{n}\right). \end{aligned}$$

Somit ist F_n ein Markov-Prozess bzgl. seiner natürlichen Filtration. Weiter gilt schließlich

$$\mathbb{E}\left(F_n(t_k) \mid F_n(t_{k-1}) = \frac{m_{k-1}}{n}\right)$$

3 Die Empirische Verteilungsfunktion

$$\begin{aligned}
&= \sum_{m=m_{k-1}}^n \frac{m}{n} \binom{n-m_{k-1}}{m-m_{k-1}} \left(\frac{F(t_k) - F(t_{k-1})}{1 - F(t_{k-1})} \right)^{m-m_{k-1}} \left(\frac{1 - F(t_k)}{1 - F(t_{k-1})} \right)^{n-m} \\
&= \sum_{m=0}^{n-m_{k-1}} \frac{m_{k-1} + m}{n} \binom{n-m_{k-1}}{m} \left(\frac{F(t_k) - F(t_{k-1})}{1 - F(t_{k-1})} \right)^m \left(\frac{1 - F(t_k)}{1 - F(t_{k-1})} \right)^{n-m_{k-1}-m} \\
&= \frac{m_{k-1}}{n} + \left(1 - \frac{m_{k-1}}{n} \right) \frac{F(t_k) - F(t_{k-1})}{1 - F(t_{k-1})}.
\end{aligned}$$

□

Man betrachte einen beliebigen Startzeitpunkt $t_0 \in \mathbb{R}$. Bezüglich der diskreten Filtration $\mathcal{F}^K = (\mathcal{F}_{t_k}^K)_{k \in \{0, \dots, K\}}$ mit $\mathcal{F}_{t_k}^K = \sigma(F_n(t_l) : l \leq k)$ gilt für das Innovationsmartingal $M^{0,K} = (M_{t_k}^{0,K})_{k \in \{0, \dots, K\}}$ der Doob-Meyer-Zerlegung von (F_n^K, \mathcal{F}^K) , wobei $F_n^K = (F_n(t_k))_{k \in \{0, \dots, K\}}$,

$$\begin{aligned}
M_{t_k}^{0,K} &= M_{t_{k-1}}^{0,K} + F_n(t_k) - \mathbb{E}\left(F_n(t_k) \mid \mathcal{F}_{t_{k-1}}^K\right) \\
&= M_{t_{k-1}}^{0,K} + F_n(t_k) - F_n(t_{k-1}) - (1 - F_n(t_{k-1})) \frac{F(t_k) - F(t_{k-1})}{1 - F(t_{k-1})}.
\end{aligned}$$

Setze $M_{t_0}^{0,K} := F_n(t_0)$. Dann gilt

$$M_{t_k}^{0,K} = F_n(t_k) - \sum_{j=1}^k \frac{1 - F_n(t_{k-1})}{1 - F(t_{j-1})} (F(t_j) - F(t_{j-1})) \quad \text{für } k = 0, \dots, K.$$

Für den Kompensator $A^{0,K} = (A_{t_k}^{0,K})_{k \in \{0, \dots, K\}}$ gilt demzufolge

$$A_{t_k}^{0,K} = \sum_{j=1}^k \frac{1 - F_n(t_{k-1})}{1 - F(t_{j-1})} (F(t_j) - F(t_{j-1})) \quad \text{für } k = 0, \dots, K.$$

3 Die Empirische Verteilungsfunktion

Damit ist die diskrete Version der Doob-Meyer-Zerlegung gefunden und für die zeitstetige Version ergibt sich nach Grenzübergang:

Satz 3.2. *Die Empirische Verteilungsfunktion mit ihrer natürlichen Filtration (F_n, \mathcal{F}) besitzt das Innovationsmartingal (M^0, \mathcal{F}) und Kompensator (A^0, \mathcal{F}) , wobei $M^0 = (M_t^0)_{t \in \mathbb{R}}$ mit*

$$M_t^0 = F_n(t) - \int_{(-\infty, t]} \frac{1 - F_n(u-)}{1 - F(u)} F(du)$$

und $A^0 = (A_t^0)_{t \in \mathbb{R}}$ mit

$$A_t^0 = \int_{(-\infty, t]} \frac{1 - F_n(u-)}{1 - F(u)} F(du).$$

3.1.2 Doob-Meyer-Zerlegung von Transformationen der Empirischen Verteilungsfunktion

Für eine beliebige Funktion $\varphi : \mathbb{R}^2 \rightarrow \mathbb{R}$ mit $\varphi(\cdot, x)$ differenzierbar für $x \in \{0, \frac{1}{n}, \dots, 1\}$ betrachte man den Prozess $X = (X_t)_{t \in \mathbb{R}}$ mit $X_t = \varphi(t, F_n(t))$, wobei F_n die Empirische Verteilungsfunktion sei. Für ein festes $t \in \mathbb{R}$ mit $F(t) < 1$ betrachte man wiederum die Zerlegung des Intervalls $(-\infty, t]$ mittels $t_0 < \dots < t_K = t$ mit $K \in \mathbb{N}$.

Satz 3.3. *Sei (X, \mathcal{F}) die oben genannte Transformation der Empirischen Verteilungsfunktion mit der natürlichen Filtration der Empirischen Verteilungsfunktion. Dann gilt für $k = 1, \dots, K$*

$$\mathbb{E}(X_{t_k} | \mathcal{F}_{t_{k-1}}) = \sum_{h=0}^n \sum_{m=0}^{n-h} \sum_{j=0}^h 1_{\{F_n(t_{k-1}) = \frac{m}{n}\}}$$

3 Die Empirische Verteilungsfunktion

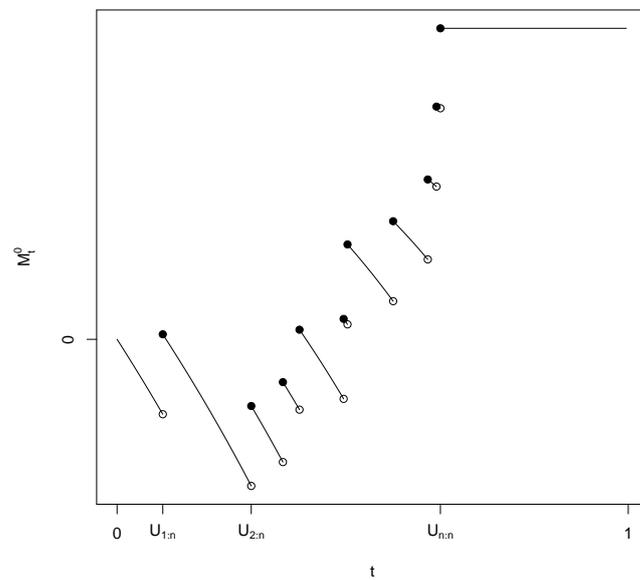


Abbildung 3.2: Das Doob-Meyer-Martingal (M^0, \mathcal{F}) der Empirischen Verteilungsfunktion von unabhängig und identisch auf $[0, 1]$ gleichverteilten Zufallsvariablen U_1, \dots, U_n , $n = 10$.

3 Die Empirische Verteilungsfunktion

$$\cdot \binom{n-m}{h} \binom{h}{j} (-1)^{h-j} \varphi \left(t_k, \frac{j+m}{n} \right) \left(\frac{F(t_k) - F(t_{k-1})}{1 - F(t_{k-1})} \right)^h.$$

Beweis. Es gilt mit Hilfe des Beweises zu Satz 3.1

$$\begin{aligned} \mathbb{E} \left(X_{t_k} | F_n(t_{k-1}) = \frac{m}{n} \right) &= \mathbb{E} \left(\varphi(t_k, F_n(t_k)) | F_n(t_{k-1}) = \frac{m}{n} \right) \\ &= \sum_{j=m}^n \binom{n-m}{j-m} \varphi \left(t_k, \frac{j}{n} \right) \left(\frac{F(t_k) - F(t_{k-1})}{1 - F(t_{k-1})} \right)^{j-m} \left(1 - \frac{F(t_k) - F(t_{k-1})}{1 - F(t_{k-1})} \right)^{n-j} \\ &= \sum_{j=0}^{n-m} \binom{n-m}{j} \varphi \left(t_k, \frac{j+m}{n} \right) \left(\frac{F(t_k) - F(t_{k-1})}{1 - F(t_{k-1})} \right)^j \left(1 - \frac{F(t_k) - F(t_{k-1})}{1 - F(t_{k-1})} \right)^{n-m-j} \\ &= \sum_{j=0}^{n-m} \binom{n-m}{j} \varphi \left(t_k, \frac{j+m}{n} \right) \left(\frac{F(t_k) - F(t_{k-1})}{1 - F(t_{k-1})} \right)^j \\ &\quad \cdot \sum_{h=0}^{n-m-j} \binom{n-m-j}{h} \left(-\frac{F(t_k) - F(t_{k-1})}{1 - F(t_{k-1})} \right)^h \\ &= \sum_{j=0}^{n-m} \sum_{h=0}^{n-m-j} \binom{n-m}{j} \binom{n-m-j}{h} (-1)^h \varphi \left(t_k, \frac{j+m}{n} \right) \left(\frac{F(t_k) - F(t_{k-1})}{1 - F(t_{k-1})} \right)^{j+h} \\ &= \sum_{j=0}^{n-m} \sum_{h=j}^{n-m} \binom{n-m}{j} \binom{n-m-j}{h-j} (-1)^{h-j} \varphi \left(t_k, \frac{j+m}{n} \right) \left(\frac{F(t_k) - F(t_{k-1})}{1 - F(t_{k-1})} \right)^h \\ &= \sum_{h=0}^{n-m} \sum_{j=0}^h \binom{n-m}{j} \binom{n-m-j}{h-j} (-1)^{h-j} \varphi \left(t_k, \frac{j+m}{n} \right) \left(\frac{F(t_k) - F(t_{k-1})}{1 - F(t_{k-1})} \right)^h \\ &= \sum_{h=0}^{n-m} \sum_{j=0}^h \binom{n-m}{h} \binom{h}{j} (-1)^{h-j} \varphi \left(t_k, \frac{j+m}{n} \right) \left(\frac{F(t_k) - F(t_{k-1})}{1 - F(t_{k-1})} \right)^h. \end{aligned}$$

Damit gilt nun

$$\mathbb{E} (X_{t_k} | \mathcal{F}_{t_{k-1}}) = \sum_{m=0}^n 1_{\{F_n(t_{k-1}) = \frac{m}{n}\}} \mathbb{E} \left(X_{t_k} | F_n(t_{k-1}) = \frac{m}{n} \right)$$

3 Die Empirische Verteilungsfunktion

$$\begin{aligned}
&= \sum_{m=0}^n 1_{\{F_n(t_{k-1}) = \frac{m}{n}\}} \sum_{h=0}^{n-m} \sum_{j=0}^h \binom{n-m}{h} \binom{h}{j} (-1)^{h-j} \varphi\left(t_k, \frac{j+m}{n}\right) \left(\frac{F(t_k) - F(t_{k-1})}{1 - F(t_{k-1})}\right)^h \\
&= \sum_{m=0}^n \sum_{h=0}^{n-m} \sum_{j=0}^h 1_{\{F_n(t_{k-1}) = \frac{m}{n}\}} \binom{n-m}{h} \binom{h}{j} (-1)^{h-j} \varphi\left(t_k, \frac{j+m}{n}\right) \left(\frac{F(t_k) - F(t_{k-1})}{1 - F(t_{k-1})}\right)^h \\
&= \sum_{h=0}^n \sum_{m=0}^{n-h} \sum_{j=0}^h 1_{\{F_n(t_{k-1}) = \frac{m}{n}\}} \binom{n-m}{h} \binom{h}{j} (-1)^{h-j} \varphi\left(t_k, \frac{j+m}{n}\right) \left(\frac{F(t_k) - F(t_{k-1})}{1 - F(t_{k-1})}\right)^h.
\end{aligned}$$

□

Man betrachte einen beliebigen Startzeitpunkt $t_0 \in \mathbb{R}$. Bezüglich der diskreten Filtration $\mathcal{F}^K = (\mathcal{F}_{t_k}^K)_{k \in \{0, \dots, K\}}$ mit $\mathcal{F}_{t_k}^K = \sigma(F_n(t_l) : l \leq k)$ gilt für das Innovationsmartingal $M^K = (M_{t_k}^K)_{k \in \{0, \dots, K\}}$ der Doob-Meyer-Zerlegung von (X^K, \mathcal{F}^K) , wobei $X^K = (X_{t_k})_{k \in \{0, \dots, K\}}$,

$$\begin{aligned}
M_{t_k}^K &= M_{t_{k-1}}^K + X_{t_k} - \mathbb{E}\left(X_{t_k} \mid \mathcal{F}_{t_{k-1}}^K\right) \\
&= M_{t_{k-1}}^K + \varphi(t_k, F_n(t_k)) - \sum_{h=0}^n \sum_{m=0}^{n-h} \sum_{j=0}^h 1_{\{F_n(t_{k-1}) = \frac{m}{n}\}} \\
&\quad \cdot \binom{n-m}{h} \binom{h}{j} (-1)^{h-j} \varphi\left(t_k, \frac{j+m}{n}\right) \left(\frac{F(t_k) - F(t_{k-1})}{1 - F(t_{k-1})}\right)^h \\
&= M_{t_{k-1}}^K + \varphi(t_k, F_n(t_k)) - \varphi(t_k, F_n(t_{k-1})) \\
&\quad - 1_{\{F_n(t_{k-1}) \leq \frac{n-1}{n}\}} n(1 - F_n(t_{k-1})) \\
&\quad \cdot \left(\varphi\left(t_k, F_n(t_{k-1}) + \frac{1}{n}\right) - \varphi(t_k, F_n(t_{k-1}))\right) \left(\frac{F(t_k) - F(t_{k-1})}{1 - F(t_{k-1})}\right) \\
&\quad - \sum_{h=2}^n \sum_{m=0}^{n-h} \sum_{j=0}^h 1_{\{F_n(t_{k-1}) = \frac{m}{n}\}} \\
&\quad \cdot \binom{n-m}{h} \binom{h}{j} (-1)^{h-j} \varphi\left(t_k, \frac{j+m}{n}\right) \left(\frac{F(t_k) - F(t_{k-1})}{1 - F(t_{k-1})}\right)^h
\end{aligned}$$

3 Die Empirische Verteilungsfunktion

$$\begin{aligned}
&= M_{t_{k-1}}^K + \varphi(t_k, F_n(t_k)) - \varphi(t_k, F_n(t_{k-1})) \\
&\quad - n(1 - F_n(t_{k-1})) \left(\varphi\left(t_k, F_n(t_{k-1}) + \frac{1}{n}\right) - \varphi(t_k, F_n(t_{k-1})) \right) \left(\frac{F(t_k) - F(t_{k-1})}{1 - F(t_{k-1})} \right) \\
&\quad - \sum_{h=2}^n \sum_{m=0}^{n-h} \sum_{j=0}^h 1_{\{F_n(t_{k-1}) = \frac{m}{n}\}} \\
&\quad \cdot \binom{n-m}{h} \binom{h}{j} (-1)^{h-j} \varphi\left(t_k, \frac{j+m}{n}\right) \left(\frac{F(t_k) - F(t_{k-1})}{1 - F(t_{k-1})} \right)^h.
\end{aligned}$$

Nun ergibt sich durch Induktion

$$\begin{aligned}
M_{t_k}^K &= X_{t_0} + \sum_{l=1}^k \varphi(t_l, F_n(t_l)) - \varphi(t_l, F_n(t_{l-1})) \\
&\quad - \sum_{l=1}^k n(1 - F_n(t_{l-1})) \left(\varphi\left(t_l, F_n(t_{l-1}) + \frac{1}{n}\right) - \varphi(t_l, F_n(t_{l-1})) \right) \left(\frac{F(t_l) - F(t_{l-1})}{1 - F(t_{l-1})} \right) \\
&\quad - \sum_{l=1}^k \sum_{h=2}^n \sum_{m=0}^{n-h} \sum_{j=0}^h 1_{\{F_n(t_{l-1}) = \frac{m}{n}\}} \\
&\quad \cdot \binom{n-m}{h} \binom{h}{j} (-1)^{h-j} \varphi\left(t_l, \frac{j+m}{n}\right) \left(\frac{F(t_l) - F(t_{l-1})}{1 - F(t_{l-1})} \right)^h.
\end{aligned}$$

Es gilt weiter

$$\begin{aligned}
&\sum_{l=1}^k \varphi(t_l, F_n(t_l)) - \varphi(t_l, F_n(t_{l-1})) \\
&= \sum_{l=1}^k \sum_{m=0}^n 1_{\{F_n(t_{l-1}) = \frac{m}{n}\}} (\varphi(t_l, F_n(t_l)) - \varphi(t_l, F_n(t_{l-1}))) \\
&= \sum_{l=1}^k \sum_{m=0}^{n-1} 1_{\{F_n(t_{l-1}) = \frac{m}{n}\}} (\varphi(t_l, F_n(t_l)) - \varphi(t_l, F_n(t_{l-1})))
\end{aligned}$$

3 Die Empirische Verteilungsfunktion

$$\begin{aligned}
&= \sum_{l=1}^k \sum_{m=0}^{n-1} \sum_{j=m+1}^n \mathbf{1}_{\{F_n(t_{l-1})=\frac{m}{n}\}} \mathbf{1}_{\{F_n(t_l)=\frac{j}{n}\}} \left(\varphi(t_l, F_n(t_l)) - \varphi(t_l, F_n(t_{l-1})) \right) \\
&= \sum_{l=1}^k \sum_{m=0}^{n-1} \sum_{j=m+1}^n \mathbf{1}_{\{F_n(t_{l-1})=\frac{m}{n}\}} \mathbf{1}_{\{F_n(t_l)=\frac{j}{n}\}} \left(\varphi\left(t_l, \frac{j}{n}\right) - \varphi\left(t_l, \frac{m}{n}\right) \right) \\
&= \sum_{l=1}^k \sum_{m=0}^{n-1} \sum_{j=m+1}^n \sum_{h=m}^{j-1} \mathbf{1}_{\{F_n(t_{l-1})=\frac{m}{n}\}} \mathbf{1}_{\{F_n(t_l)=\frac{j}{n}\}} \left(\varphi\left(t_l, \frac{h+1}{n}\right) - \varphi\left(t_l, \frac{h}{n}\right) \right) \\
&= \sum_{l=1}^k \sum_{m=0}^{n-1} \sum_{h=m}^{n-1} \sum_{j=h+1}^n \mathbf{1}_{\{F_n(t_{l-1})=\frac{m}{n}\}} \mathbf{1}_{\{F_n(t_l)=\frac{j}{n}\}} \left(\varphi\left(t_l, \frac{h+1}{n}\right) - \varphi\left(t_l, \frac{h}{n}\right) \right) \\
&= \sum_{l=1}^k \sum_{m=0}^{n-1} \sum_{h=m}^{n-1} \mathbf{1}_{\{F_n(t_{l-1})=\frac{m}{n}\}} \mathbf{1}_{\{F_n(t_l) > \frac{h}{n}\}} \left(\varphi\left(t_l, \frac{h+1}{n}\right) - \varphi\left(t_l, \frac{h}{n}\right) \right) \\
&= \sum_{l=1}^k \sum_{h=0}^{n-1} \sum_{m=0}^h \mathbf{1}_{\{F_n(t_{l-1})=\frac{m}{n}\}} \mathbf{1}_{\{F_n(t_l) > \frac{h}{n}\}} \left(\varphi\left(t_l, \frac{h+1}{n}\right) - \varphi\left(t_l, \frac{h}{n}\right) \right) \\
&= \sum_{l=1}^k \sum_{h=0}^{n-1} \mathbf{1}_{\{F_n(t_{l-1}) \leq \frac{h}{n}\}} \mathbf{1}_{\{F_n(t_l) > \frac{h}{n}\}} \left(\varphi\left(t_l, \frac{h+1}{n}\right) - \varphi\left(t_l, \frac{h}{n}\right) \right) \\
&= \sum_{l=1}^k \sum_{m=1}^n \mathbf{1}_{\{t_{l-1} < X_{m:n} \leq t_l\}} \left(\varphi\left(t_l, \frac{m}{n}\right) - \varphi\left(t_l, \frac{m-1}{n}\right) \right).
\end{aligned}$$

Also gilt schließlich

$$\begin{aligned}
M_{t_k}^K &= X_{t_0} + \sum_{l=1}^k \sum_{m=1}^n \mathbf{1}_{\{t_{l-1} < X_{m:n} \leq t_l\}} \left(\varphi\left(t_l, \frac{m}{n}\right) - \varphi\left(t_l, \frac{m-1}{n}\right) \right) \\
&\quad - \sum_{l=1}^k n(1 - F_n(t_{l-1})) \left(\varphi\left(t_l, F_n(t_{l-1}) + \frac{1}{n}\right) - \varphi\left(t_l, F_n(t_{l-1})\right) \right) \left(\frac{F(t_l) - F(t_{l-1})}{1 - F(t_{l-1})} \right) \\
&\quad - \sum_{l=1}^k \sum_{h=2}^n \sum_{m=0}^{n-h} \sum_{j=0}^h \mathbf{1}_{\{F_n(t_{l-1})=\frac{m}{n}\}} \\
&\quad \cdot \binom{n-m}{h} \binom{h}{j} (-1)^{h-j} \varphi\left(t_l, \frac{j+m}{n}\right) \left(\frac{F(t_l) - F(t_{l-1})}{1 - F(t_{l-1})} \right)^h.
\end{aligned}$$

3 Die Empirische Verteilungsfunktion

Durch Grenzübergang erhält man wieder die zeitstetige Version. Die Zuwächse quadratischer und höherer Ordnung verschwinden hierbei und das resultierende Martingal lässt sich als Integral bezüglich des \mathcal{F} -Doob-Meyer-Martingals M^0 der Empirischen Verteilungsfunktion schreiben.

Satz 3.4. *Sei $t_0 \in \mathbb{R}$ ein beliebiger Startwert. Der Prozess $X = (X_t)_{t \in [t_0, \infty)}$ mit $X_t = \varphi(t, F_n(t))$ besitzt das Innovationsmartingal (M, \mathcal{F}) , wobei $\mathcal{F} = (\mathcal{F}_t)_{t \in [t_0, \infty)}$ mit $\mathcal{F}_t = \sigma(F_n(u) : u \leq t)$, und $M = (M_t)_{t \in [t_0, \infty)}$ mit*

$$\begin{aligned} M_t &= X_{t_0} + \int_{(t_0, t]} n \left(\varphi \left(u, F_n(u-) + \frac{1}{n} \right) - \varphi(u, F_n(u-)) \right) dF_n(u) \\ &\quad - \int_{(t_0, t]} n \left(\varphi \left(u, F_n(u-) + \frac{1}{n} \right) - \varphi(u, F_n(u-)) \right) \frac{1 - F_n(u-)}{1 - F(u)} dF(u) \\ &= X_{t_0} + \int_{(t_0, t]} n \left(\varphi \left(u, F_n(u-) + \frac{1}{n} \right) - \varphi(u, F_n(u-)) \right) dM_u^0. \end{aligned}$$

Satz 3.5. *Es existiere der Grenzwert*

$$x_{-\infty} := \lim_{s \rightarrow -\infty} \varphi(s, 0).$$

Dann besitzt der Prozess $X = (X_t)_{t \in \mathbb{R}}$ mit $X_t = \varphi(t, F_n(t))$ das Innovationsmartingal (M, \mathcal{F}) , wobei $\mathcal{F} = (\mathcal{F}_t)_{t \in \mathbb{R}}$ mit $\mathcal{F}_t = \sigma(F_n(u) : u \leq t)$, und $M = (M_t)_{t \in \mathbb{R}}$ mit

$$\begin{aligned} M_t &= x_{-\infty} + \int_{(-\infty, t]} n \left(\varphi \left(u, F_n(u-) + \frac{1}{n} \right) - \varphi(u, F_n(u-)) \right) dF_n(u) \\ &\quad - \int_{(-\infty, t]} n \left(\varphi \left(u, F_n(u-) + \frac{1}{n} \right) - \varphi(u, F_n(u-)) \right) \frac{1 - F_n(u-)}{1 - F(u)} dF(u) \end{aligned}$$

3 Die Empirische Verteilungsfunktion

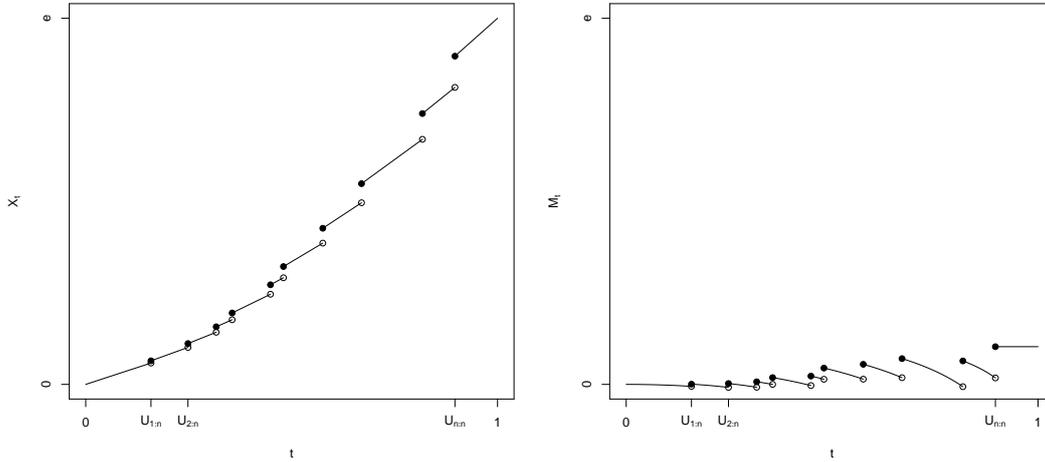


Abbildung 3.3: Der Prozess X mit $X_t = t \exp(F_n(t))$ und sein Doob-Meyer-Martingal (M, \mathcal{F}) , wobei die unterliegenden Zufallsvariablen U_1, \dots, U_n , $n = 10$, unabhängig und identisch gleichverteilt sind auf $[0, 1]$.

$$= x_{-\infty} + \int_{(-\infty, t]} n \left(\varphi \left(u, F_n(u-) + \frac{1}{n} \right) - \varphi(u, F_n(u-)) \right) dM_u^0.$$

Weiterhin liefert die Doob-Meyer-Zerlegung folgenden Kompensator

$$\begin{aligned} A_{t_k}^K &= A_{t_{k-1}}^K - X_{t_{k-1}} + \mathbb{E} \left(X_{t_k} \mid \mathcal{F}_{t_{k-1}}^K \right) \\ &= A_{t_{k-1}}^K - \varphi(t_{k-1}, F_n(t_{k-1})) + \sum_{h=0}^n \sum_{m=0}^{n-h} \sum_{j=0}^h 1_{\{F_n(t_{k-1}) = \frac{m}{n}\}} \\ &\quad \cdot \binom{n-m}{h} \binom{h}{j} (-1)^{h-j} \varphi \left(t_k, \frac{j+m}{n} \right) \left(\frac{F(t_k) - F(t_{k-1})}{1 - F(t_{k-1})} \right)^h \\ &= A_{t_{k-1}} + \varphi(t_k, F_n(t_{k-1})) - \varphi(t_{k-1}, F_n(t_{k-1})) \end{aligned}$$

3 Die Empirische Verteilungsfunktion

$$\begin{aligned}
& + n(1 - F_n(t_{k-1})) \left(\varphi \left(t_k, F_n(t_{k-1}) + \frac{1}{n} \right) - \varphi(t_k, F_n(t_{k-1})) \right) \frac{F(t_k) - F(t_{k-1})}{1 - F(t_{k-1})} \\
& + \sum_{h=2}^n \sum_{m=0}^{n-h} \sum_{j=0}^h 1_{\{F_n(t_{k-1}) = \frac{m}{n}\}} \\
& \cdot \binom{n-m}{h} \binom{h}{j} (-1)^{h-j} \varphi \left(t_k, F_n(t_{k-1}) + \frac{j}{n} \right) \left(\frac{F(t_k) - F(t_{k-1})}{1 - F(t_{k-1})} \right)^h.
\end{aligned}$$

Durch Induktion ergibt sich

$$\begin{aligned}
A_{t_k}^K & = \sum_{l=1}^k \varphi(t_l, F_n(t_{l-1})) - \varphi(t_{l-1}, F_n(t_{l-1})) \\
& + \sum_{l=1}^k n(1 - F_n(t_{l-1})) \left(\varphi \left(t_l, F_n(t_{l-1}) + \frac{1}{n} \right) - \varphi(t_l, F_n(t_{l-1})) \right) \frac{F(t_l) - F(t_{l-1})}{1 - F(t_{l-1})} \\
& + \sum_{l=1}^k \sum_{h=2}^n \sum_{m=0}^{n-h} \sum_{j=0}^h 1_{\{F_n(t_{l-1}) = \frac{m}{n}\}} \\
& \cdot \binom{n-m}{h} \binom{h}{j} (-1)^{h-j} \varphi \left(t_l, \frac{j+m}{n} \right) \left(\frac{F(t_l) - F(t_{l-1})}{1 - F(t_{l-1})} \right)^h.
\end{aligned}$$

Nach Grenzübergang ergibt sich auch der zeitstetige Kompensator:

Satz 3.6. *Der Prozess $(X_t)_{t \in \mathbb{R}}$ mit $X_t = \varphi(t, F_n(t))$ besitzt den Kompensator (A, \mathcal{F}) , wobei $\mathcal{F} = (\mathcal{F}_t)_{t \in \mathbb{R}}$ mit $\mathcal{F}_t = \sigma(F_n(u) : u \leq t)$ und $A = (A_t)_{t \in \mathbb{R}}$ mit*

$$A_t = \int_{(-\infty, t]} \frac{\partial \varphi}{\partial F}(u, F_n(u-)) + n \left(\varphi \left(u, F_n(u-) + \frac{1}{n} \right) - \varphi(u, F_n(u-)) \right) \frac{1 - F_n(u-)}{1 - F(u)} F(du).$$

Die \mathcal{F} -Doob-Meyer-Zerlegung von X liefert also auch hier ein Martingal, welches aus einem Startwert und einem Integral bezüglich M^0 - dem \mathcal{F} -Doob-Meyer-Martingal der Empirischen Verteilungsfunktion - besteht, und einen Kompensator, welcher als Integral bezüglich F ,

3 Die Empirische Verteilungsfunktion

also einer deterministischen Funktion, geschrieben werden kann. Daraus lässt sich wieder eine Martingaldifferentialgleichung ableiten, deren Lösungen zu Martingaltransformationen führen. Dies soll im Folgenden geschehen.

3.1.3 Martingaltransformationen der Empirischen Verteilungsfunktion

Wie auch im Spezialfall des Single-Event-Prozesses ist (X, \mathcal{F}) genau dann ein Martingal, falls der Integrand im Kompensator aus Satz 3.6 verschwindet. Dies führt zu folgender Martingaldifferentialgleichung:

Satz 3.7. *Der Prozess $X = (X_t)_{t \in \mathbb{R}}$ mit $X_t = \varphi(t, F_n(t))$ ist ein Martingal bzgl. $\mathcal{F} = (\mathcal{F}_t)_{t \in \mathbb{R}}$ mit $\mathcal{F}_t = \sigma(F_n(u) : u \leq t)$, falls für alle $t \in \mathbb{R}$ und $x \in \{0, \frac{1}{n}, \dots, \frac{n-1}{n}\}$ gilt*

$$\frac{\partial \varphi}{\partial F}(t, x) + n \left(\varphi \left(t, x + \frac{1}{n} \right) - \varphi(t, x) \right) \frac{1-x}{1-F(t)} = 0.$$

3.1.3.1 Affine Martingaltransformationen der Empirischen Verteilungsfunktion

Wie zuvor sollen nun zunächst die affinen Lösungen der Martingaldifferentialgleichung aus Satz 3.7 gesucht werden, wobei affin wiederum bedeutet, dass für $t \in \mathbb{R}$ die Funktion $\varphi(t, \cdot)$ affin ist. Diese Lösungen führen zu folgenden \mathcal{F} -Martingaltransformationen der Empirischen Verteilungsfunktion.

Satz 3.8. *Seien $c_0, c_1 \in \mathbb{R}$. Dann ist der Prozess $X = (X_t)_{t \in \mathbb{R}}$ mit*

$$X_t = c_1 \frac{1 - F_n(t)}{1 - F(t)} + c_0$$

ein Martingal bzgl. $\mathcal{F} = (\mathcal{F}_t)_{t \in \mathbb{R}}$ mit $\mathcal{F}_t = \sigma(F_n(u) : u \leq t)$.

3 Die Empirische Verteilungsfunktion

Beweis. Sei

$$\varphi(t, x) = a_1(F(t))x + a_0(F(t)).$$

Dann gilt

$$\frac{\partial \varphi}{\partial F}(t, x) = \frac{\partial a_1}{\partial F}(F(t))x + \frac{\partial a_0}{\partial F}(F(t))$$

und $X_t = \varphi(t, F_n(t))$ ist ein Martingal bzgl. \mathcal{F} , falls

$$\left(\frac{\partial a_1}{\partial F}(F(t)) - \frac{a_1(F(t))}{1-F(t)} \right) x + \frac{\partial a_0}{\partial F}(F(t)) + \frac{a_1(F(t))}{1-F(t)} = 0.$$

Analog zum Beweis von Satz 2.8 muss also für $c_0, c_1 \in \mathbb{R}$ gelten, dass

$$a_1(F(t)) = -\frac{c_1}{1-F(t)} \quad \text{und} \quad a_0(F(t)) = \frac{c_1}{1-F(t)} + c_0.$$

Also ist für $c_0, c_1 \in \mathbb{R}$

$$X_t = c_1 \frac{1 - F_n(t)}{1 - F(t)} + c_0$$

ein \mathcal{F} -Martingal.

□

3.1.3.2 Polynomiale Martingaltransformationen der Empirischen Verteilungsfunktion

Die affinen \mathcal{F} -Martingaltransformationen aus dem vorhergehenden Abschnitt sind tatsächlich nur ein Spezialfall aus der Klasse der polynomialen \mathcal{F} -Martingaltransformationen der Empirischen Verteilungsfunktion, wobei polynomial wiederum bedeutet, dass für $t \in \mathbb{R}$ die Funktion $\varphi(t, \cdot)$ ein Polynom ist. Im folgenden Satz werden Vertreter der Klasse der po-

3 Die Empirische Verteilungsfunktion

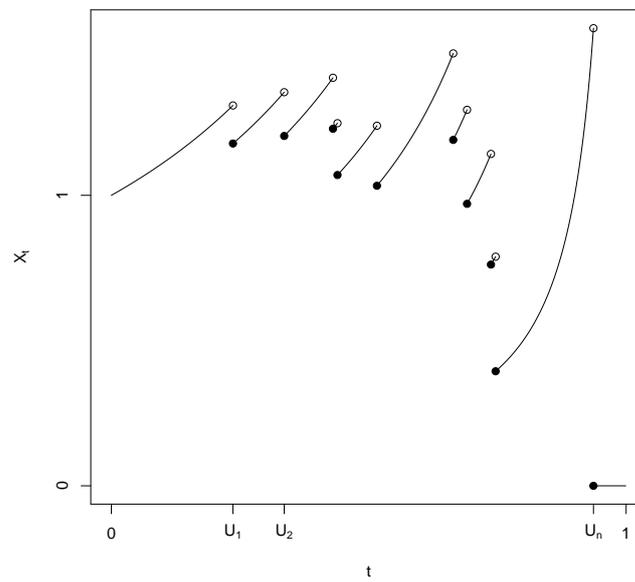


Abbildung 3.4: Das Martingal (X, \mathcal{F}) mit $X_t = \frac{1 - F_n(t)}{1 - t}$, wobei die unterliegenden Zufallsvariablen U_1, \dots, U_n , $n = 10$, unabhängig und identisch gleichverteilt sind auf $[0, 1]$.

3 Die Empirische Verteilungsfunktion

polynomialen \mathcal{F} -Martingaltransformationen der Empirischen Verteilungsfunktion vorgestellt, welche eine Basis aller polynomialen \mathcal{F} -Martingaltransformationen der Empirischen Verteilungsfunktion der Form $X = (X_t)_{t \in \mathbb{R}}$ mit $X_t = \varphi(t, F_n(t))$ sind.

Satz 3.9. Sei $m \in \{0, \dots, n\}$. Dann ist der Prozess $X^m = (X_t^m)_{t \in \mathbb{R}}$ mit

$$X_t^m = \frac{(n-m)!n^m}{n!} \prod_{j=n-(m-1)}^n \frac{\frac{j}{n} - F_n(t)}{1 - F(t)}$$

ein Martingal bzgl. $\mathcal{F} = (\mathcal{F}_t)_{t \in \mathbb{R}}$ mit $\mathcal{F}_t = \sigma(F_n(u) : u \leq t)$.

Beweis. Setze

$$\varphi(t, x) = \frac{\prod_{j=n-(m-1)}^n \left(\frac{j}{n} - x\right)}{(1 - F(t))^m}.$$

Dann gilt

$$\frac{\partial \varphi}{\partial F}(t, x) = \frac{m \prod_{j=n-(m-1)}^n \left(\frac{j}{n} - x\right)}{(1 - F(t))^{m+1}}$$

und

$$\begin{aligned} & n \left(\varphi\left(t, x + \frac{1}{n}\right) - \varphi(t, x) \right) \frac{1-x}{1-F(t)} \\ &= n \left(\frac{\prod_{j=n-(m-1)}^n \left(\frac{j-1}{n} - x\right)}{(1-F(t))^m} - \frac{\prod_{j=n-(m-1)}^n \left(\frac{j}{n} - x\right)}{(1-F(t))^m} \right) \frac{1-x}{1-F(t)} \\ &= \frac{n(1-x) \left(\prod_{j=n-m}^{n-1} \left(\frac{j}{n} - x\right) - \prod_{j=n-(m-1)}^n \left(\frac{j}{n} - x\right) \right)}{(1-F(t))^{m+1}} \end{aligned}$$

3 Die Empirische Verteilungsfunktion

$$\begin{aligned}
 & n(1-x) \left(\binom{n-m}{n} - x \right) - (1-x) \prod_{j=n-(m-1)}^{n-1} \left(\frac{j}{n} - x \right) \\
 = & \frac{\quad}{(1-F(t))^{m+1}} \\
 = & - \frac{m \prod_{j=n-(m-1)}^n \left(\frac{j}{n} - x \right)}{(1-F(t))^{m+1}}.
 \end{aligned}$$

Also erfüllt φ die Martingal-Differentialgleichung aus Satz 3.7 und $X_t = \frac{(n-m)!n^m}{n!} \varphi(t, F_n(t))$ ist ein Martingal bzgl. \mathcal{F} .

□

- Bemerkung 3.10.**
1. Durch den Faktor $\frac{(n-m)!n^m}{n!}$ wird sichergestellt, dass die Martingale aus Satz 3.9 alle im Wert 1 starten.
 2. Die Martingale aus Satz 3.9 sind nichtnegativ und fallen auf den Wert 0, sobald die Empirische Verteilungsfunktion den Wert $1 - \frac{m-1}{n}$ erreicht. Zur vollständigen Kenntnis des Martingals werden also nur die ersten $n - (m - 1)$ Ordnungsstatistiken benötigt. Dadurch werden diese Martingale interessant im Fall von Typ-II-Rechtszensierungen, in denen nur die ersten $n - (m - 1)$ Ordnungsstatistiken beobachtet werden.
 3. Durch Linearkombinationen der Martingale aus Satz 3.9 erhält man sämtliche polynomiale \mathcal{F} -Martingale, die sich aus F_n gewinnen lassen.

3.1.3.3 Exponentielle Martingaltransformationen der Empirischen Verteilungsfunktion

Auf der Suche nach weiteren Martingaltransformationen kann beispielsweise ein exponentieller Ansatz gewählt werden. Auch hier ergeben sich eine ganze Reihe von Martingaltransformationen.

3 Die Empirische Verteilungsfunktion

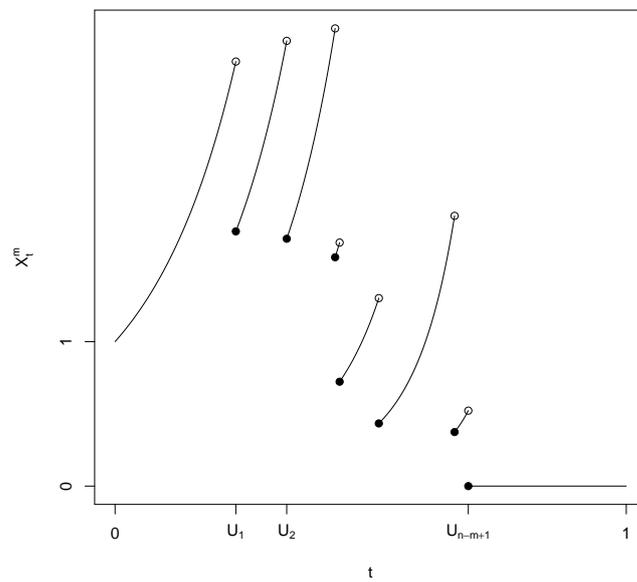


Abbildung 3.5: Das Martingal (X^m, \mathcal{F}) aus Satz 3.9 mit $n = 10$ und $m = 4$, wobei die unterliegenden Zufallsvariablen U_1, \dots, U_n , unabhängig und identisch gleichverteilt sind auf $[0, 1]$.

3 Die Empirische Verteilungsfunktion

Satz 3.11. Seien $c_1, c_2 \in \mathbb{R}$. Dann ist der Prozess $X = (X_t)_{t \in \mathbb{R}}$ mit

$$X_t = c_2 \left(1 + \frac{c_1}{1 - F(t)} \right)^{n(1 - F_n(t))}$$

ein Martingal bzgl. $\mathcal{F} = (\mathcal{F}_t)_{t \in \mathbb{R}}$ mit $\mathcal{F}_t = \sigma(F_n(u) : u \leq t)$.

Beweis. Setze

$$\varphi(t, x) = \exp(a_1(F(t))x + a_0(F(t))).$$

Dann gilt

$$\frac{\partial \varphi}{\partial F}(t, x) = \left(\frac{\partial a_1}{\partial F}(F(t))x + \frac{\partial a_0}{\partial F}(F(t)) \right) \varphi(t, x)$$

sowie

$$\begin{aligned} n \left(\varphi \left(t, x + \frac{1}{n} \right) - \varphi(t, x) \right) &= n \left(\exp \left(a_1(F(t)) \left(x + \frac{1}{n} \right) + a_0(F(t)) \right) - \varphi(t, x) \right) \\ &= n \left(\exp \left(\frac{a_1(F(t))}{n} \right) - 1 \right) \varphi(t, x) \end{aligned}$$

und $X_t = \varphi(t, F_n(t))$ ist ein Martingal bzgl. \mathcal{F} , falls

$$\left(\frac{\partial a_1}{\partial F}(F(t)) - n \frac{\exp \left(\frac{a_1(F(t))}{n} \right) - 1}{1 - F(t)} \right) x + \frac{\partial a_0}{\partial F}(F(t)) + n \frac{\exp \left(\frac{a_1(F(t))}{n} \right) - 1}{1 - F(t)} = 0.$$

Somit gilt

$$\begin{aligned} \frac{\partial a_1}{\partial F}(F(t)) - n \frac{\exp \left(\frac{a_1(F(t))}{n} \right) - 1}{1 - F(t)} = 0 &\Leftrightarrow \frac{\partial a_1}{\partial F}(F(t)) = n \frac{\exp \left(\frac{a_1(F(t))}{n} \right) - 1}{1 - F(t)} \\ \Leftrightarrow a_1(F(t)) = -n \ln \left(1 + \frac{c_1}{1 - F(t)} \right) &\text{ für ein } c_1 \in \mathbb{R} \end{aligned}$$

3 Die Empirische Verteilungsfunktion

sowie

$$\begin{aligned}
 \frac{\partial a_0}{\partial F}(F(t)) + n \frac{\exp\left(\frac{a_1(F(t))}{n}\right) - 1}{1 - F(t)} &= 0 \quad \Leftrightarrow \quad \frac{\partial a_0}{\partial F}(F(t)) = -n \frac{\exp\left(\frac{a_1(F(t))}{n}\right) - 1}{1 - F(t)} \\
 \Leftrightarrow \quad \frac{\partial a_0}{\partial F}(F(t)) &= n \frac{c_1}{(1 - F(t))(1 + c_1 - F(t))} \quad \text{für ein } c_1 \in \mathbb{R} \\
 \Leftrightarrow \quad a_0(F(t)) &= n \ln\left(\frac{1 + c_1 - F(t)}{1 - F(t)}\right) + c_2 \quad \text{für } c_1, c_2 \in \mathbb{R}.
 \end{aligned}$$

Es gilt also

$$\begin{aligned}
 \varphi(t, x) &= \exp(a_1(F(t))x + a_0(F(t))) \\
 &= \exp\left(nx \ln\left(\frac{1 - F(t)}{1 + c_1 - F(t)}\right) + n \ln\left(\frac{1 + c_1 - F(t)}{1 - F(t)}\right) + \tilde{c}_2\right) \quad \text{für } c_1, \tilde{c}_2 \in \mathbb{R} \\
 &= c_2 \left(\frac{1 + c_1 - F(t)}{1 - F(t)}\right)^{n(1-x)} \quad \text{für } c_1, c_2 \in \mathbb{R} \\
 &= c_2 \left(1 + \frac{c_1}{1 - F(t)}\right)^{n(1-x)} \quad \text{für } c_1, c_2 \in \mathbb{R}.
 \end{aligned}$$

□

Bei genauerer Betrachtung sind die in Satz 3.11 beschriebenen exponentiellen Martingale jedoch lediglich spezielle polynomiale \mathcal{F} -Martingale.

Satz 3.12. Sei $X = (X_t)_{t \in \mathbb{R}}$ ein exponentielles \mathcal{F} -Martingal mit $c_1, c_2 \in \mathbb{R}$, d.h.

$$X_t = c_2 \left(1 + \frac{c_1}{1 - F(t)}\right)^{n(1 - F_n(t))}$$

3 Die Empirische Verteilungsfunktion

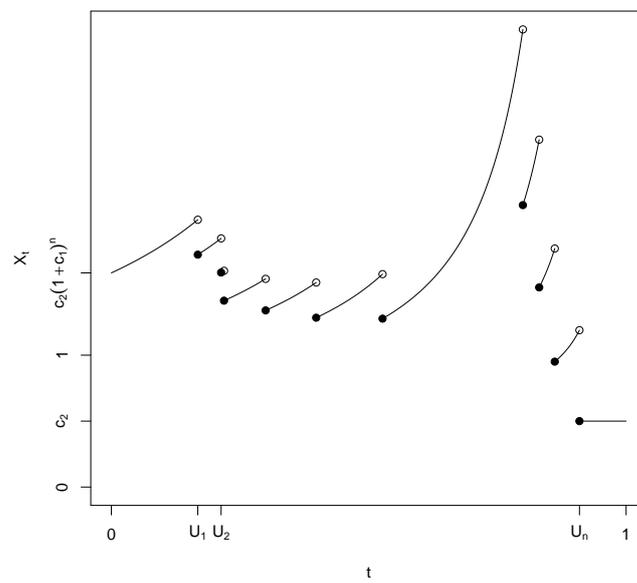


Abbildung 3.6: Das Martingal (X, \mathcal{F}) mit $X_t = c_2 \left(1 + \frac{c_1}{1-t}\right)^{n(1-F_n(t))}$ mit $c_1 = \frac{1}{8}$ und $c_2 = \frac{1}{2}$, wobei die unterliegenden Zufallsvariablen U_1, \dots, U_n , $n = 10$, unabhängig und identisch gleichverteilt sind auf $[0, 1]$.

3 Die Empirische Verteilungsfunktion

und seien $X^m = (X_t^m)_{t \in \mathbb{R}}$, $m \in \{0, \dots, n\}$, polynomiale \mathcal{F} -Martingale mit

$$X_t^m = \frac{(n-m)!n^m}{n!} \prod_{j=n-(m-1)}^n \frac{\frac{j}{n} - F_n(t)}{1 - F(t)}.$$

Dann gilt für alle $t \in \mathbb{R}$

$$X_t = \sum_{m=0}^n a_m X_t^m,$$

wobei

$$a_m = \binom{n}{m} c_1^m c_2.$$

Beweis. Es gilt einerseits

$$\begin{aligned} X_t &= c_2 \left(1 + \frac{c_1}{1 - F(t)} \right)^{n(1-F_n(t))} \\ &= \sum_{m=0}^{n(1-F_n(t))} c_1^m c_2 \binom{n(1-F_n(t))}{m} \frac{1}{(1 - F(t))^m} \end{aligned}$$

und andererseits

$$\begin{aligned} \sum_{m=0}^n a_m X_t^m &= \sum_{m=0}^n a_m \frac{(n-m)!n^m}{n!} \prod_{j=n-(m-1)}^n \frac{\frac{j}{n} - F_n(t)}{1 - F(t)} \\ &= \sum_{m=0}^n a_m \frac{(n-m)!}{n! (1 - F(t))^m} \prod_{j=n-(m-1)}^n (j - nF_n(t)) \\ &= \sum_{m=0}^n a_m \frac{(n-m)! (n(1 - F_n(t)))!}{n! (n(1 - F_n(t)) - m)! (1 - F(t))^m} \\ &= \sum_{m=0}^n \frac{a_m}{\binom{n}{m}} \binom{n(1 - F_n(t))}{m} \frac{1}{(1 - F(t))^m}. \end{aligned}$$

3 Die Empirische Verteilungsfunktion

Für $a_m = \binom{n}{m} c_1^m c_2$ gilt also $X = \sum_{m=0}^n a_m X^m$.

□

3.1.3.4 Allgemeine Martingaltransformationen der Empirischen Verteilungsfunktion

Satz 3.12 hat gezeigt, dass sich die exponentiellen \mathcal{F} -Martingale als Linearkombinationen der polynomialen \mathcal{F} -Martingale darstellen lassen. Im Folgenden wird gezeigt, dass die polynomialen \mathcal{F} -Martingale tatsächlich eine Basis für den Raum aller Martingale (X, \mathcal{F}) mit $X = (\varphi(t, F_n(t)))_{t \in \mathbb{R}}$ und $\mathcal{F} = (\mathcal{F}_t)_{t \in \mathbb{R}}$ mit $\mathcal{F}_t = \sigma(F_n(u) : u \leq t)$ bilden.

Satz 3.13. Sei $X = (X_t)_{t \in \mathbb{R}}$ mit $X_t = \varphi(t, F_n(t))$ ein Martingal bzgl. $\mathcal{F} = (\mathcal{F}_t)_{t \in \mathbb{R}}$. Dann existieren $c_0, \dots, c_n \in \mathbb{R}$, sodass

$$X_t = \sum_{m=0}^n c_m X_t^m,$$

wobei

$$X_t^m = \frac{(n-m)!n^m}{n!} \prod_{j=n-(m-1)}^n \frac{\frac{j}{n} - F_n(t)}{1 - F(t)}, \quad m \in \{0, \dots, n\}.$$

Beweis. Die Empirische Verteilungsfunktion kann nur die Zustände $0, \frac{1}{n}, \dots, 1$ annehmen.

Daher lässt sich jeder Prozess X mit $X_t = \varphi(t, F_n(t))$ wie folgt schreiben:

$$X_t = \sum_{m=0}^n \varphi_m(t) 1_{\{F_n(t) = \frac{m}{n}\}},$$

wobei $\varphi_m(t) = \varphi(t, \frac{m}{n})$ für $t \in \mathbb{R}$ und $m \in \{0, \dots, n\}$. Damit X ein \mathcal{F} -Martingal ist, muss

3 Die Empirische Verteilungsfunktion

für $m \in \{0, \dots, n\}$ und $s \leq t$ gelten, dass

$$\begin{aligned}
 \varphi_m(s) &= \mathbb{E} \left(X_t \mid F_n(s) = \frac{m}{n} \right) \\
 &= \sum_{j=m}^n \varphi_j(t) \mathbb{P} \left(F_n(t) = \frac{j}{n} \mid F_n(s) = \frac{m}{n} \right) \\
 &= \sum_{j=m}^n \varphi_j(t) \binom{n-m}{j-m} \left(\frac{F(t) - F(s)}{1 - F(s)} \right)^{j-m} \left(\frac{1 - F(t)}{1 - F(s)} \right)^{n-j} \\
 &= \varphi_m(t) \left(\frac{1 - F(t)}{1 - F(s)} \right)^{n-m} + \sum_{j=m+1}^n \varphi_j(t) \binom{n-m}{j-m} \left(\frac{F(t) - F(s)}{1 - F(s)} \right)^{j-m} \left(\frac{1 - F(t)}{1 - F(s)} \right)^{n-j}
 \end{aligned}$$

bzw.

$$\begin{aligned}
 \varphi_m(s) &= \varphi_m(t) \left(\frac{1 - F(t)}{1 - F(s)} \right)^{n-m} \\
 &\quad + \sum_{j=m+1}^n \varphi_j(t) \binom{n-m}{j-m} \left(\frac{F(t) - F(s)}{1 - F(s)} \right)^{j-m} \left(\frac{1 - F(t)}{1 - F(s)} \right)^{n-j} \\
 &\Leftrightarrow \varphi_m(t) = \varphi_m(s) \left(\frac{1 - F(s)}{1 - F(t)} \right)^{n-m} - \sum_{j=m+1}^n \varphi_j(t) \binom{n-m}{j-m} \left(\frac{F(t) - F(s)}{1 - F(t)} \right)^{j-m} \\
 &\Leftrightarrow \varphi_m(t) = \varphi_m(s) \left(\frac{1 - F(s)}{1 - F(t)} \right)^{n-m} - \sum_{j=m+1}^n \varphi_j(t) \binom{n-m}{j-m} \left(\frac{1 - F(s)}{1 - F(t)} - 1 \right)^{j-m} \\
 &\Leftrightarrow \varphi_m(t) = \varphi_m(s) \left(\frac{1 - F(s)}{1 - F(t)} \right)^{n-m} \\
 &\quad - \sum_{j=m+1}^n \varphi_j(t) \binom{n-m}{j-m} \sum_{i=0}^{j-m} \binom{j-m}{i} (-1)^i \left(\frac{1 - F(s)}{1 - F(t)} \right)^{j-m-i} \\
 &\Leftrightarrow \varphi_m(t) = \varphi_m(s) \left(\frac{1 - F(s)}{1 - F(t)} \right)^{n-m} \\
 &\quad - \sum_{j=m+1}^n \varphi_j(t) \binom{n-m}{j-m} \sum_{i=m}^j \binom{j-m}{i-m} (-1)^{i-m} \left(\frac{1 - F(s)}{1 - F(t)} \right)^{j-i}.
 \end{aligned}$$

3 Die Empirische Verteilungsfunktion

Man betrachte den Startpunkt $s_0 = \sup_{s \in \mathbb{R}} \{F(s) = 0\}$, wobei $s_0 = -\infty$, falls $F(s) > 0$ für alle $s \in \mathbb{R}$. Dann gilt obige Gleichung auch für $s = s_0$, also

$$\begin{aligned} \varphi_m(t) &= \varphi_m(s_0) \left(\frac{1}{1-F(t)} \right)^{n-m} \\ &\quad - \sum_{j=m+1}^n \varphi_j(t) \binom{n-m}{j-m} \sum_{i=m}^j \binom{j-m}{i-m} (-1)^{i-m} \left(\frac{1}{1-F(t)} \right)^{j-i} \end{aligned}$$

bzw. falls $s_0 = -\infty$

$$\begin{aligned} \varphi_m(t) &= \lim_{s \rightarrow -\infty} \varphi_m(s) \left(\frac{1}{1-F(t)} \right)^{n-m} \\ &\quad - \sum_{j=m+1}^n \varphi_j(t) \binom{n-m}{j-m} \sum_{i=m}^j \binom{j-m}{i-m} (-1)^{i-m} \left(\frac{1}{1-F(t)} \right)^{j-i}. \end{aligned}$$

Setze $\varphi_m(s_0) = a_m \in \mathbb{R}$, falls $s_0 \in \mathbb{R}$, bzw. $\lim_{s \rightarrow -\infty} \varphi_m(s) = a_m \in \mathbb{R}$, falls $s_0 = -\infty$. Dann folgt also

$$\varphi_m(t) = a_m \frac{1}{(1-F(t))^{n-m}} - \sum_{j=m+1}^n \varphi_j(t) \binom{n-m}{j-m} \sum_{i=m}^j \binom{j-m}{i-m} (-1)^{i-m} \frac{1}{(1-F(t))^{j-i}}.$$

Mit vollständiger Induktion zeigt man nun, dass für gewisse b_j^m mit $m \in \{0, \dots, n\}$ und $j \in \{0, \dots, n-m\}$ für die explizite Darstellung gilt

$$\varphi_m(t) = \sum_{j=0}^{n-m} \frac{b_j^m}{(1-F(t))^j}, \quad b_j^m \in \mathbb{R}.$$

Mit $b_0^n = a_n$ ist diese Behauptung für $m = n$ offensichtlich richtig. Die Behauptung gelte

3 Die Empirische Verteilungsfunktion

nun für alle $m > m_0$, wobei $m_0 \in \{0, \dots, n-1\}$. Dann folgt

$$\begin{aligned}
\varphi_{m_0}(t) &= \frac{a_{m_0}}{(1-F(t))^{n-m_0}} - \sum_{m=m_0+1}^n \varphi_m(t) \binom{n-m_0}{m-m_0} \sum_{i=m_0}^m \binom{m-m_0}{i-m_0} \frac{(-1)^{i-m_0}}{(1-F(t))^{m-i}} \\
&= \frac{a_{m_0}}{(1-F(t))^{n-m_0}} - \sum_{m=m_0+1}^n \sum_{j=0}^{n-m} \sum_{i=m_0}^m (-1)^{i-m_0} \binom{n-m_0}{m-m_0} \binom{m-m_0}{i-m_0} \frac{b_j^m}{(1-F(t))^{m+j-i}} \\
&= \frac{a_{m_0}}{(1-F(t))^{n-m_0}} - \sum_{m=m_0+1}^n \sum_{i=m_0}^m \sum_{j=0}^{n-m} (-1)^{i-m_0} \binom{n-m_0}{m-m_0} \binom{m-m_0}{i-m_0} \frac{b_j^m}{(1-F(t))^{m+j-i}} \\
&= \frac{a_{m_0}}{(1-F(t))^{n-m_0}} - \sum_{m=m_0+1}^n \sum_{i=m_0}^m \sum_{j=m-i}^{n-i} (-1)^{i-m_0} \binom{n-m_0}{m-m_0} \binom{m-m_0}{i-m_0} \frac{b_{j-(m-i)}^m}{(1-F(t))^j} \\
&= \frac{a_{m_0}}{(1-F(t))^{n-m_0}} - \sum_{m=m_0+1}^n \sum_{j=m-m_0}^{n-m_0} \binom{n-m_0}{m-m_0} \frac{b_{j-(m-m_0)}^m}{(1-F(t))^j} \\
&\quad - \sum_{m=m_0+1}^n \sum_{i=m_0+1}^m \sum_{j=m-i}^{n-i} (-1)^{i-m_0} \binom{n-m_0}{m-m_0} \binom{m-m_0}{i-m_0} \frac{b_{j-(m-i)}^m}{(1-F(t))^j} \\
&= \frac{a_{m_0}}{(1-F(t))^{n-m_0}} - \sum_{j=1}^{n-m_0} \sum_{m=m_0+1}^{m_0+j} \binom{n-m_0}{m-m_0} \frac{b_{j-(m-m_0)}^m}{(1-F(t))^j} \\
&\quad - \sum_{i=m_0+1}^n \sum_{m=i}^n \sum_{j=m-i}^{n-i} (-1)^{i-m_0} \binom{n-m_0}{m-m_0} \binom{m-m_0}{i-m_0} \frac{b_{j-(m-i)}^m}{(1-F(t))^j} \\
&= \frac{a_{m_0}}{(1-F(t))^{n-m_0}} - \sum_{j=1}^{n-m_0} \sum_{m=m_0+1}^{m_0+j} \binom{n-m_0}{m-m_0} \frac{b_{j-(m-m_0)}^m}{(1-F(t))^j} \\
&\quad - \sum_{i=m_0+1}^n \sum_{j=0}^{n-i} \sum_{m=i}^{i+j} (-1)^{i-m_0} \binom{n-m_0}{m-m_0} \binom{m-m_0}{i-m_0} \frac{b_{j-(m-i)}^m}{(1-F(t))^j} \\
&= \frac{a_{m_0}}{(1-F(t))^{n-m_0}} - \sum_{j=1}^{n-m_0} \sum_{m=m_0+1}^{m_0+j} \binom{n-m_0}{m-m_0} \frac{b_{j-(m-m_0)}^m}{(1-F(t))^j} \\
&\quad - \sum_{j=0}^{n-m_0-1} \sum_{i=m_0+1}^{n-j} \sum_{m=i}^{i+j} (-1)^{i-m_0} \binom{n-m_0}{m-m_0} \binom{m-m_0}{i-m_0} \frac{b_{j-(m-i)}^m}{(1-F(t))^j}.
\end{aligned}$$

3 Die Empirische Verteilungsfunktion

Also gilt

$$\varphi_{m_0}(t) = \sum_{j=0}^{n-m_0} \frac{b_j^{m_0}}{(1-F(t))^j}, \quad b_j^{m_0} \in \mathbb{R}$$

mit

$$b_0^{m_0} = - \sum_{m=m_0+1}^n (-1)^{m-m_0} \binom{n-m_0}{m-m_0} b_0^m,$$

$$b_{n-m_0}^{m_0} = a_{m_0} - \sum_{m=m_0+1}^n \binom{n-m_0}{m-m_0} b_{n-m}^m$$

und

$$b_j^{m_0} = - \sum_{m=m_0+1}^{m_0+j} \binom{n-m_0}{m-m_0} b_{m_0+j-m}^m - \sum_{l=m_0+1}^{n-j} \sum_{m=l}^{l+j} (-1)^{l-m_0} \binom{n-m_0}{m-m_0} \binom{m-m_0}{l-m_0} b_{l+j-m}^m$$

für $j \in \{1, \dots, n-m_0-1\}$.

Mit einer weiteren vollständigen Induktion zeigt man, dass

$$b_j^m = \binom{n-m}{j} \sum_{i=0}^j \binom{j}{i} (-1)^{j-i} a_{n-i}, \quad m = 0, \dots, n, \quad j = 0, \dots, n-m.$$

Offensichtlich gilt die Behauptung für $m = n$. Sei $m_0 \in \{0, \dots, n-1\}$. Die Behauptung gelte für alle $m > m_0$. Nun gilt

$$\begin{aligned} b_0^{m_0} &= - \sum_{m=m_0+1}^n (-1)^{m-m_0} \binom{n-m_0}{m-m_0} b_0^m \\ &= - \sum_{m=m_0+1}^n (-1)^{m-m_0} \binom{n-m_0}{m-m_0} a_n \\ &= a_n \left(1 - \sum_{m=0}^{n-m_0} (-1)^m \binom{n-m_0}{m} \right) \end{aligned}$$

3 Die Empirische Verteilungsfunktion

$$= a_n.$$

Weiter gilt

$$\begin{aligned}
 b_{n-m_0}^{m_0} &= a_{m_0} - \sum_{m=m_0+1}^n \binom{n-m_0}{m-m_0} b_{n-m}^m \\
 &= a_{m_0} - \sum_{m=m_0+1}^n \binom{n-m_0}{m-m_0} \sum_{i=0}^{n-m} \binom{n-m}{i} (-1)^{n-m-i} a_{n-i} \\
 &= a_{m_0} - \sum_{i=0}^{n-m_0-1} a_{n-i} \sum_{m=m_0+1}^{n-i} \binom{n-m_0}{m-m_0} \binom{n-m}{i} (-1)^{n-m-i} \\
 &= a_{m_0} - \sum_{i=0}^{n-m_0-1} \binom{n-m_0}{i} a_{n-i} \sum_{m=m_0+1}^{n-i} \binom{n-m_0-i}{m-m_0} (-1)^{n-m-i} \\
 &= a_{m_0} - \sum_{i=0}^{n-m_0-1} \binom{n-m_0}{i} a_{n-i} \sum_{m=1}^{n-m_0-i} \binom{n-m_0-i}{m} (-1)^{n-m-m_0-i} \\
 &= a_{m_0} - \sum_{i=0}^{n-m_0-1} \binom{n-m_0}{i} (-1)^{n-m_0-i} a_{n-i} \sum_{m=1}^{n-m_0-i} \binom{n-m_0-i}{m} (-1)^m \\
 &= a_{m_0} - \sum_{i=0}^{n-m_0-1} \binom{n-m_0}{i} (-1)^{n-m_0-i} a_{n-i} \left(-1 + \sum_{m=0}^{n-m_0-i} \binom{n-m_0-i}{m} (-1)^m \right) \\
 &= a_{m_0} + \sum_{i=0}^{n-m_0-1} \binom{n-m_0}{i} (-1)^{n-m_0-i} a_{n-i} \\
 &= \sum_{i=0}^{n-m_0} \binom{n-m_0}{i} (-1)^{n-m_0-i} a_{n-i}.
 \end{aligned}$$

Schließlich gilt für $j \in \{1, \dots, n-m_0-1\}$:

$$b_j^{m_0} = - \sum_{m=m_0+1}^{m_0+j} \binom{n-m_0}{m-m_0} b_{m_0+j-m}^m - \sum_{l=m_0+1}^{n-j} \sum_{m=l}^{l+j} (-1)^{l-m_0} \binom{n-m_0}{m-m_0} \binom{m-m_0}{l-m_0} b_{l+j-m}^m$$

3 Die Empirische Verteilungsfunktion

$$\begin{aligned}
&= - \sum_{m=m_0+1}^{m_0+j} \sum_{i=0}^{m_0+j-m} \binom{n-m_0}{m-m_0} \binom{n-m}{m_0+j-m} \binom{m_0+j-m}{i} (-1)^{m_0+j-m-i} a_{n-i} \\
&\quad - \sum_{l=m_0+1}^{n-j} \sum_{m=l}^{l+j} \sum_{i=0}^{l+j-m} \binom{n-m_0}{m-m_0} \binom{m-m_0}{l-m_0} \binom{n-m}{l+j-m} \binom{l+j-m}{i} (-1)^{j-m-i-m_0} a_{n-i} \\
&= - \sum_{i=0}^{j-1} \sum_{m=m_0+1}^{m_0+j-i} \binom{n-m_0}{m-m_0} \binom{n-m}{m_0+j-m} \binom{m_0+j-m}{i} (-1)^{m_0+j-m-i} a_{n-i} \\
&\quad - \sum_{l=m_0+1}^{n-j} \sum_{m=0}^j \sum_{i=0}^{j-m} \binom{n-m_0}{m+l-m_0} \binom{m+l-m_0}{l-m_0} \binom{n-m-l}{j-m} \binom{j-m}{i} \\
&\quad \cdot (-1)^{j-m-l-i-m_0} a_{n-i} \\
&= - \sum_{i=0}^{j-1} \sum_{m=m_0+1}^{m_0+j-i} \binom{n-m_0}{m-m_0} \binom{n-m}{m_0+j-m} \binom{m_0+j-m}{i} (-1)^{m_0+j-m-i} a_{n-i} \\
&\quad - \sum_{i=0}^j \sum_{m=0}^{j-i} \sum_{l=m_0+1}^{n-j} \binom{n-m_0}{m+l-m_0} \binom{m+l-m_0}{l-m_0} \binom{n-m-l}{j-m} \binom{j-m}{i} \\
&\quad \cdot (-1)^{j-m-l-i-m_0} a_{n-i} \\
&= - \sum_{i=0}^{j-1} \sum_{m=1}^{j-i} \binom{n-m_0}{m} \binom{n-m-m_0}{j-m} \binom{j-m}{i} (-1)^{j-m-i} a_{n-i} \\
&\quad - \sum_{i=0}^j \sum_{m=0}^{j-i} \sum_{l=1}^{n-m_0-j} \binom{n-m_0}{m+l} \binom{m+l}{l} \binom{n-m_0-m-l}{j-m} \binom{j-m}{i} (-1)^{j-m-l-i} a_{n-i} \\
&= - \sum_{i=0}^{j-1} \binom{n-m_0}{j} \binom{j}{i} a_{n-i} \sum_{m=1}^{j-i} \binom{j-i}{m} (-1)^{j-i-m} \\
&\quad - \sum_{i=0}^j \sum_{m=0}^{j-i} \binom{n-m_0}{j} \binom{j}{i} \binom{j-i}{m} (-1)^{j-i-m} a_{n-i} \sum_{l=1}^{n-m_0-j} \binom{n-m_0-j}{l} (-1)^{-l} \\
&= \binom{n-m_0}{j} \sum_{i=0}^{j-1} \binom{j}{i} (-1)^{j-i} a_{n-i} \\
&\quad + \sum_{i=0}^j \binom{n-m_0}{j} \binom{j}{i} (-1)^{j-i} a_{n-i} \sum_{m=0}^{j-i} \binom{j-i}{m} (-1)^{-m}
\end{aligned}$$

3 Die Empirische Verteilungsfunktion

$$= \binom{n-m_0}{j} \sum_{i=0}^j \binom{j}{i} (-1)^{j-i} a_{n-i}.$$

Also gilt

$$\varphi_{m_0}(t) = \sum_{j=0}^{n-m_0} \frac{\binom{n-m_0}{j}}{(1-F(t))^j} \sum_{i=0}^j \binom{j}{i} (-1)^{j-i} a_{n-i}, \quad a_i \in \mathbb{R} \text{ für } i \in \{0, \dots, n\}.$$

Sei nun $X = (X_t)_{t \in \mathbb{R}}$ mit $X_t = \varphi(t, F_n(t))$ ein Martingal bzgl. $\mathcal{F} = (\mathcal{F}_t)_{t \in \mathbb{R}}$. Dann gilt

$$X_t = \sum_{m=0}^n \sum_{j=0}^{n-m} \frac{\binom{n-m}{j}}{(1-F(t))^j} \sum_{i=0}^j \binom{j}{i} (-1)^{j-i} a_{n-i} 1_{\{F_n(t) = \frac{m}{n}\}}.$$

Desweiteren gilt

$$\begin{aligned} X_t^{m_0} &= \frac{(n-m_0)! n^{m_0}}{n!} \prod_{j=n-(m_0-1)}^n \frac{j - F_n(t)}{1 - F(t)} \\ &= \sum_{m=0}^{n-m_0} \frac{(n-m_0)!}{n!} \prod_{j=n-(m_0-1)}^n \frac{j-m}{1-F(t)} 1_{\{F_n(t) = \frac{m}{n}\}} \\ &= \sum_{m=0}^{n-m_0} \frac{(n-m_0)! (n-m)!}{n! (n-m-m_0)!} \cdot \frac{1}{(1-F(t))^{m_0}} 1_{\{F_n(t) = \frac{m}{n}\}} \\ &= \frac{1}{(1-F(t))^{m_0}} \sum_{m=0}^{n-m_0} \frac{\binom{n-m_0}{m}}{\binom{n}{m}} 1_{\{F_n(t) = \frac{m}{n}\}}. \end{aligned}$$

Damit folgt

$$\begin{aligned} \sum_{m_0=0}^n c_{m_0} X_t^{m_0} &= \sum_{m_0=0}^n c_{m_0} \frac{1}{(1-F(t))^{m_0}} \sum_{m=0}^{n-m_0} \frac{\binom{n-m_0}{m}}{\binom{n}{m}} 1_{\{F_n(t) = \frac{m}{n}\}} \\ &= \sum_{m=0}^n \sum_{j=0}^{n-m} \frac{\binom{n-j}{m} c_j}{\binom{n}{m} (1-F(t))^j} 1_{\{F_n(t) = \frac{m}{n}\}}. \end{aligned}$$

3 Die Empirische Verteilungsfunktion

Weiter gilt

$$\begin{aligned}
 \sum_{j=0}^{n-m} \frac{\binom{n-m}{j}}{(1-F(t))^j} \sum_{i=0}^j \binom{j}{i} (-1)^{j-i} a_{n-i} &= \sum_{j=0}^{n-m} \frac{\binom{n-j}{m} c_j}{\binom{n}{m} (1-F(t))^j} \quad \forall m \in \{0, \dots, n\} \\
 \Leftrightarrow \binom{n-m}{j} \sum_{i=0}^j \binom{j}{i} (-1)^{j-i} a_{n-i} &= \frac{\binom{n-j}{m} c_j}{\binom{n}{m}} \quad \forall m \in \{0, \dots, n\}, j \in \{0, \dots, n-m\} \\
 \Leftrightarrow c_j = \frac{\binom{n}{m} \binom{n-m}{j}}{\binom{n-j}{m}} \sum_{i=0}^j \binom{j}{i} (-1)^{j-i} a_{n-i} &\quad \forall m \in \{0, \dots, n\}, j \in \{0, \dots, n-m\} \\
 \Leftrightarrow c_j = \binom{n}{j} \sum_{i=0}^j \binom{j}{i} (-1)^{j-i} a_{n-i} &\quad \forall j \in \{0, \dots, n\}.
 \end{aligned}$$

Damit folgt schließlich die Behauptung. □

Dem Beweis des vorhergehenden Satzes ist sogar die konkrete Darstellung eines \mathcal{F} -Martingals $X = (X_t)_{t \in \mathbb{R}}$ mit $X_t = \varphi(t, F_n(t))$ zu entnehmen.

Korollar 3.14. Sei $X = (X_t)_{t \in \mathbb{R}}$ mit $X_t = \varphi(t, F_n(t))$ ein Martingal bzgl. $\mathcal{F} = (\mathcal{F}_t)_{t \in \mathbb{R}}$. Dann existieren $a_0, \dots, a_n \in \mathbb{R}$ mit

$$a_m = \lim_{s \rightarrow -\infty} \varphi\left(s, \frac{m}{n}\right) \quad \text{für } m = 0, \dots, n$$

und es gilt

$$X_t = \sum_{m=0}^n \binom{n}{m} \sum_{i=0}^m \binom{m}{i} (-1)^{m-i} a_{n-i} X_t^m,$$

wobei

$$X_t^m = \frac{(n-m)! n^m}{n!} \prod_{j=n-(m-1)}^n \frac{\frac{j}{n} - F_n(t)}{1 - F(t)}, \quad m \in \{0, \dots, n\}.$$

3.2 Reverse Doob-Meyer-Zerlegungen für die Empirische Verteilungsfunktion

Für spätere Anwendungen ist es hilfreich, die Empirische Verteilungsfunktion und Transformationen $X = (X_t)_{t \in \mathbb{R}}$ der Form $X_t = \varphi(t, F_n(t))$ auch in „umgekehrter Zeit“ zu betrachten, also bezüglich der natürlichen reversen Filtration der Empirischen Verteilungsfunktion $\mathcal{G} = (\mathcal{G}_t)_{t \in \mathbb{R}}$ mit $\mathcal{G}_t = \sigma(F_n(u) : u \geq t)$. Im Folgenden soll zuerst die reverse Doob-Meyer-Zerlegung bezüglich \mathcal{G} der Empirischen Verteilungsfunktion selbst und anschließend die reverse Doob-Meyer-Zerlegung bezüglich \mathcal{G} von Transformationen der genannten Form gefunden werden. Daraus leiten sich wiederum reverse \mathcal{G} -Martingaltransformationen ab, der Lösungen zum Abschluss des Abschnitts hergeleitet werden. Dabei stellt sich heraus, dass die gefundenen reversen \mathcal{G} -Martingaltransformationen eine vollständige Klasse bilden.

3.2.1 Reverse Doob-Meyer-Zerlegung der Empirischen Verteilungsfunktion

Man betrachte die unabhängig und identisch verteilten Zufallsvariablen X_1, \dots, X_n , $n \in \mathbb{N}$, mit $X_1 \sim F$, wobei F eine stetige Verteilungsfunktion ist. Für ein festes $t \in \mathbb{R}$ mit $F(t) > 0$ betrachte die Zerlegung des Intervalls $[t, \infty)$ mittels $t = t_0 < \dots < t_K$ mit $K \in \mathbb{N}$. Zunächst soll die reverse Doob-Meyer-Zerlegung der diskretisierten Empirischen Verteilungsfunktion gefunden werden.

Satz 3.15. *Sei (F_n, \mathcal{G}) die Empirische Verteilungsfunktion mit natürlicher reverser Filtra-*

3 Die Empirische Verteilungsfunktion

tion. Dann gilt für $k = 0, \dots, K - 1$

$$\mathbb{E}(F_n(t_k) | \mathcal{G}_{t_{k+1}}) = F_n(t_{k+1}) - \frac{F_n(t_{k+1})}{F(t_{k+1})} (F(t_{k+1}) - F(t_k)).$$

Beweis. Die Mengen

$$\{X_i \leq t_k\}, \{t_k < X_i \leq t_{k+1}\}, \dots, \{t_{K-1} < X_i \leq t_K\}, \{X_i > t_K\}$$

bilden eine Partition des Grundraumes für jedes $i = 1, \dots, n$. Da X_1, \dots, X_n unabhängig und identisch verteilt sind, ist die gemeinsame Verteilung der Zuwächse der Empirischen Verteilungsfunktion eine Multinomialverteilung. Es gilt also

$$\begin{aligned} & n(F_n(t_k), F_n(t_{k+1}) - F_n(t_k), \dots, F_n(t_K) - F_n(t_{K-1}), 1 - F_n(t_K)) \\ & \sim \text{Mult}(n, (F_n(t_k), F_n(t_{k+1}) - F_n(t_k), \dots, F_n(t_K) - F_n(t_{K-1}), 1 - F_n(t_K))). \end{aligned}$$

Damit gilt für $0 \leq m_k \leq \dots \leq m_K \leq n$

$$\begin{aligned} & \mathbb{P}\left(F_n(t_l) = \frac{m_l}{n} \quad \text{für alle } l = k, \dots, K\right) \\ & = \mathbb{P}\left(F_n(t_k) = \frac{m_k}{n}, F_n(t_{l+1}) - F_n(t_l) = \frac{m_{l+1} - m_l}{n} \quad \text{für alle } l = k, \dots, K - 1\right) \\ & = \binom{n}{m_k \quad m_{k+1} - m_k \quad \dots \quad m_K - m_{K-1} \quad n - m_K} \\ & \quad \cdot F(t_k)^{m_k} \prod_{l=k}^{K-1} (F(t_{l+1}) - F(t_l))^{m_{l+1} - m_l} (1 - F(t_K))^{n - m_K} \end{aligned}$$

3 Die Empirische Verteilungsfunktion

und für die bedingte Verteilung

$$\begin{aligned}
& \mathbb{P}\left(F_n(t_k) = \frac{m_k}{n} \mid F_n(t_l) = \frac{m_l}{n} \text{ für alle } l = k+1, \dots, K\right) \\
&= \frac{\mathbb{P}\left(F_n(t_l) = \frac{m_l}{n} \text{ für alle } l = k, \dots, K\right)}{\mathbb{P}\left(F_n(t_l) = \frac{m_l}{n} \text{ für alle } l = k+1, \dots, K\right)} \\
&= \binom{m_{k+1}}{m_k} \left(\frac{F(t_k)}{F(t_{k+1})}\right)^{m_k} \left(1 - \frac{F(t_k)}{F(t_{k+1})}\right)^{m_{k+1}-m_k} \\
&= \mathbb{P}\left(F_n(t_k) = \frac{m_k}{n} \mid F_n(t_{k+1}) = \frac{m_{k+1}}{n}\right).
\end{aligned}$$

Somit ist F_n ein reverser Markov-Prozess bzgl. seiner natürlichen reversen Filtration. Weiter gilt schließlich

$$\begin{aligned}
& \mathbb{E}\left(F_n(t_k) \mid F_n(t_{k+1}) = \frac{m_{k+1}}{n}\right) \\
&= \sum_{m=0}^{m_{k+1}} \frac{m}{n} \binom{m_{k+1}}{m} \left(\frac{F(t_k)}{F(t_{k+1})}\right)^m \left(1 - \frac{F(t_k)}{F(t_{k+1})}\right)^{m_{k+1}-m} \\
&= \sum_{m=1}^{m_{k+1}} \frac{m}{n} \binom{m_{k+1}}{m} \left(\frac{F(t_k)}{F(t_{k+1})}\right)^m \left(1 - \frac{F(t_k)}{F(t_{k+1})}\right)^{m_{k+1}-m} \\
&= \frac{m_{k+1}}{n} \cdot \frac{F(t_k)}{F(t_{k+1})} \sum_{m=0}^{m_{k+1}-1} \binom{m_{k+1}-1}{m} \left(\frac{F(t_k)}{F(t_{k+1})}\right)^m \left(1 - \frac{F(t_k)}{F(t_{k+1})}\right)^{m_{k+1}-1-m} \\
&= \frac{m_{k+1}}{n} \cdot \frac{F(t_k)}{F(t_{k+1})} \\
&= \frac{m_{k+1}}{n} - \frac{m_{k+1}}{n} \cdot \frac{F(t_{k+1}) - F(t_k)}{F(t_{k+1})}.
\end{aligned}$$

□

Man betrachte einen beliebigen Startzeitpunkt $t_K \in \mathbb{R}$. Bezüglich der diskreten reversen Filtration $\mathcal{G}^K = (\mathcal{G}_{t_k}^K)_{k \in \{0, \dots, K\}}$ mit $\mathcal{G}_{t_k}^K = \sigma(F_n(t_l) : l \geq k)$ gilt für das reverse Innovati-

3 Die Empirische Verteilungsfunktion

onsmartingal $M^{K,0} = (M_{t_k}^{K,0})_{k \in \{0, \dots, K\}}$ der reversen Doob-Meyer-Zerlegung von (F_n^K, \mathcal{G}^K) , wobei $F_n^K = (F_n(t_k))_{k \in \{0, \dots, K\}}$,

$$\begin{aligned} M_{t_k}^{0,K} &= M_{t_{k+1}}^{0,K} + F_n(t_k) - \mathbb{E}\left(F_n(t_k) \mid \mathcal{G}_{t_{k+1}}^K\right) \\ &= M_{t_{k+1}}^{0,K} + F_n(t_k) - \left(F_n(t_{k+1}) - \frac{F_n(t_{k+1})}{F(t_{k+1})}(F(t_{k+1}) - F(t_k))\right). \end{aligned}$$

Setze $M_{t_K}^{K,0} := F_n(t_K)$. Dann gilt

$$M_{t_k}^{K,0} = F_n(t_k) + \sum_{j=k}^{K-1} \frac{F_n(t_{j+1})}{F(t_{j+1})}(F(t_{j+1}) - F(t_j)) \quad \text{für } k = 0, \dots, K.$$

Für den reversen Kompensator $A^{K,0} = (A_{t_k}^{K,0})_{k \in \{0, \dots, K\}}$ gilt demzufolge

$$A_{t_k}^{K,0} = - \sum_{j=k}^{K-1} \frac{F_n(t_{j+1})}{F(t_{j+1})}(F(t_{j+1}) - F(t_j)) \quad \text{für } k = 0, \dots, K.$$

Damit ist die diskrete Version der reversen Doob-Meyer-Zerlegung gefunden und für die zeitstetige Version ergibt sich nach Grenzübergang:

Satz 3.16. *Die Empirische Verteilungsfunktion mit ihrer natürlichen reversen Filtration (F_n, \mathcal{G}) besitzt das reverse Innovationsmartingal (M^0, \mathcal{G}) und den reversen Kompensator (A^0, \mathcal{G}) , wobei $M^0 = (M_t^0)_{t \in \mathbb{R}}$ mit*

$$M_t^0 = F_n(t) + \int_{[t, \infty)} \frac{F_n(u)}{F(u)} F(du)$$

3 Die Empirische Verteilungsfunktion

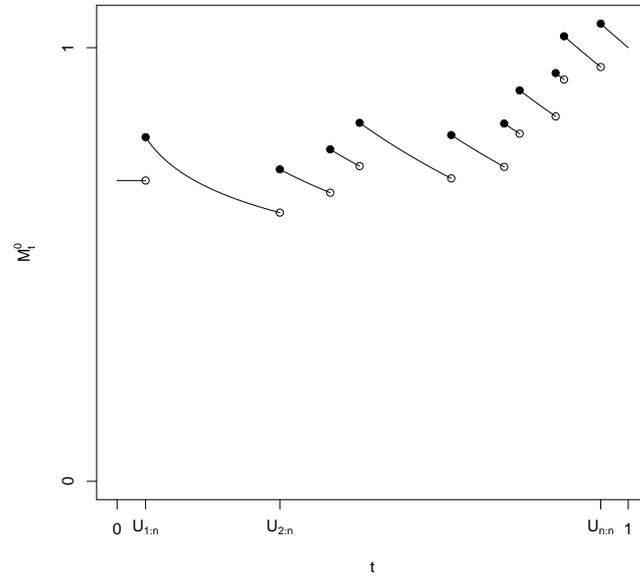


Abbildung 3.7: Das reverse Doob-Meyer-Martingal (M^0, \mathcal{G}) der Empirischen Verteilungsfunktion von unabhängig und identisch auf $[0, 1]$ gleichverteilten Zufallsvariablen U_1, \dots, U_n , $n = 10$.

und $A^0 = (A_t^0)_{t \in \mathbb{R}}$ mit

$$A_t^0 = - \int_{[t, \infty)} \frac{F_n(u)}{F(u)} F(du).$$

3.2.2 Reverse Doob-Meyer-Zerlegung von Transformationen der Empirischen Verteilungsfunktion

Für eine beliebige Funktion $\varphi : \mathbb{R}^2 \rightarrow \mathbb{R}$ mit $\varphi(\cdot, x)$ differenzierbar für $x \in \{0, \frac{1}{n}, \dots, 1\}$ betrachte man den Prozess $X = (X_t)_{t \in \mathbb{R}}$ mit $X_t = \varphi(t, F_n(t))$, wobei F_n die Empirische Verteilungsfunktion ist. Für ein festes $t \in \mathbb{R}$ mit $F(t) > 0$ betrachte wiederum die Zerlegung des Intervalls $[t, \infty)$ mittels $t = t_0 < \dots < t_K$ mit $K \in \mathbb{N}$.

Satz 3.17. *Sei (X, \mathcal{G}) die oben genannte Transformation der Empirischen Verteilungsfunktion mit der natürlichen reversen Filtration der Empirischen Verteilungsfunktion. Dann gilt für $k = 0, \dots, K - 1$*

$$\begin{aligned} \mathbb{E}(X_{t_k} | \mathcal{G}_{t_{k+1}}) &= \sum_{h=0}^n \sum_{m=h}^n \sum_{j=0}^h 1_{\{F_n(t_{k+1}) = \frac{m}{n}\}} \\ &\cdot \binom{m}{h} \binom{h}{j} (-1)^{h-j} \varphi\left(t_k, \frac{m-j}{n}\right) \left(\frac{F(t_{k+1}) - F(t_k)}{F(t_{k+1})}\right)^h. \end{aligned}$$

Beweis. Es gilt

$$\begin{aligned} \mathbb{E}\left(X_{t_k} | F_n(t_{k+1}) = \frac{m}{n}\right) &= \mathbb{E}\left(\varphi(t_k, F_n(t_k)) | F_n(t_{k+1}) = \frac{m}{n}\right) \\ &= \sum_{j=0}^m \binom{m}{j} \varphi\left(t_k, \frac{j}{n}\right) \left(\frac{F(t_k)}{F(t_{k+1})}\right)^j \left(1 - \frac{F(t_k)}{F(t_{k+1})}\right)^{m-j} \\ &= \sum_{j=0}^m \binom{m}{j} \varphi\left(t_k, \frac{j}{n}\right) \left(1 - \frac{F(t_{k+1}) - F(t_k)}{F(t_{k+1})}\right)^j \left(\frac{F(t_{k+1}) - F(t_k)}{F(t_{k+1})}\right)^{m-j} \\ &= \sum_{j=0}^m \binom{m}{j} \varphi\left(t_k, \frac{j}{n}\right) \left(\frac{F(t_{k+1}) - F(t_k)}{F(t_{k+1})}\right)^{m-j} \sum_{h=0}^j \binom{j}{h} \left(-\frac{F(t_{k+1}) - F(t_k)}{F(t_{k+1})}\right)^h \end{aligned}$$

3 Die Empirische Verteilungsfunktion

$$\begin{aligned}
&= \sum_{j=0}^m \sum_{h=0}^j \binom{m}{j} \binom{j}{h} (-1)^h \varphi \left(t_k, \frac{j}{n} \right) \left(\frac{F(t_{k+1}) - F(t_k)}{F(t_{k+1})} \right)^{m-j+h} \\
&= \sum_{j=0}^m \sum_{h=m-j}^m \binom{m}{j} \binom{j}{h-m+j} (-1)^{h-m+j} \varphi \left(t_k, \frac{j}{n} \right) \left(\frac{F(t_{k+1}) - F(t_k)}{F(t_{k+1})} \right)^h \\
&= \sum_{h=0}^m \sum_{j=m-h}^m \binom{m}{j} \binom{j}{m-h} (-1)^{h-m+j} \varphi \left(t_k, \frac{j}{n} \right) \left(\frac{F(t_{k+1}) - F(t_k)}{F(t_{k+1})} \right)^h.
\end{aligned}$$

Damit gilt nun

$$\begin{aligned}
\mathbb{E} \left(X_{t_k} \mid \mathcal{G}_{t_{k+1}} \right) &= \sum_{m=0}^n 1_{\{F_n(t_{k+1}) = \frac{m}{n}\}} \mathbb{E} \left(X_{t_k} \mid F_n(t_{k+1}) = \frac{m}{n} \right) \\
&= \sum_{m=0}^n 1_{\{F_n(t_{k+1}) = \frac{m}{n}\}} \\
&\quad \cdot \sum_{h=0}^m \sum_{j=m-h}^m \binom{m}{j} \binom{j}{m-h} (-1)^{h-m+j} \varphi \left(t_k, \frac{j}{n} \right) \left(\frac{F(t_{k+1}) - F(t_k)}{F(t_{k+1})} \right)^h \\
&= \sum_{m=0}^n \sum_{h=0}^m \sum_{j=0}^h 1_{\{F_n(t_{k+1}) = \frac{m}{n}\}} \\
&\quad \cdot \binom{m}{j+m-h} \binom{j+m-h}{m-h} (-1)^j \varphi \left(t_k, \frac{j+m-h}{n} \right) \left(\frac{F(t_{k+1}) - F(t_k)}{F(t_{k+1})} \right)^h \\
&= \sum_{h=0}^n \sum_{m=h}^n \sum_{j=0}^h 1_{\{F_n(t_{k+1}) = \frac{m}{n}\}} \\
&\quad \cdot \binom{m}{j+m-h} \binom{j+m-h}{m-h} (-1)^j \varphi \left(t_k, \frac{j+m-h}{n} \right) \left(\frac{F(t_{k+1}) - F(t_k)}{F(t_{k+1})} \right)^h \\
&= \sum_{h=0}^n \sum_{m=h}^n \sum_{j=0}^h 1_{\{F_n(t_{k+1}) = \frac{m}{n}\}} \binom{m}{h} \binom{h}{j} (-1)^j \varphi \left(t_k, \frac{j+m-h}{n} \right) \left(\frac{F(t_{k+1}) - F(t_k)}{F(t_{k+1})} \right)^h \\
&= \sum_{h=0}^n \sum_{m=h}^n \sum_{j=0}^h 1_{\{F_n(t_{k+1}) = \frac{m}{n}\}} \binom{m}{h} \binom{h}{j} (-1)^{h-j} \varphi \left(t_k, \frac{m-j}{n} \right) \left(\frac{F(t_{k+1}) - F(t_k)}{F(t_{k+1})} \right)^h.
\end{aligned}$$

3 Die Empirische Verteilungsfunktion

□

Man betrachte einen beliebigen Startzeitpunkt $t_K \in \mathbb{R}$. Bezüglich der diskreten reversen Filtration $\mathcal{G}^K = (\mathcal{G}_{t_k}^K)_{k \in \{0, \dots, K\}}$ mit $\mathcal{G}_{t_k}^K = \sigma(F_n(t_l) : l \geq k)$ gilt für das reverse Innovationsmartingal $M^K = (M_{t_k}^K)_{k \in \{0, \dots, K\}}$ der reversen Doob-Meyer-Zerlegung von (X^K, \mathcal{G}^K) , wobei $X^K = (X_{t_k})_{k \in \{0, \dots, K\}}$,

$$\begin{aligned}
 M_{t_k}^K &= M_{t_{k+1}}^K + X_{t_k} - \mathbb{E}\left(X_{t_k} \mid \mathcal{G}_{t_{k+1}}^K\right) \\
 &= M_{t_{k+1}}^K + \varphi(t_k, F_n(t_k)) \\
 &\quad - \sum_{h=0}^n \sum_{m=h}^n \sum_{j=0}^h \mathbf{1}_{\{F_n(t_{k+1}) = \frac{m}{n}\}} \binom{m}{h} \binom{h}{j} (-1)^{h-j} \varphi\left(t_k, \frac{m-j}{n}\right) \left(\frac{F(t_{k+1}) - F(t_k)}{F(t_{k+1})}\right)^h \\
 &= M_{t_{k+1}}^K + \varphi(t_k, F_n(t_k)) - \varphi(t_k, F_n(t_{k+1})) \\
 &\quad + \mathbf{1}_{\{F_n(t_{k+1}) \geq \frac{1}{n}\}} \\
 &\quad \cdot n F_n(t_{k+1}) \left(\varphi(t_k, F_n(t_{k+1})) - \varphi\left(t_k, F_n(t_{k+1}) - \frac{1}{n}\right)\right) \frac{F(t_{k+1}) - F(t_k)}{F(t_{k+1})} \\
 &\quad - \sum_{h=2}^n \sum_{m=h}^n \sum_{j=0}^h \mathbf{1}_{\{F_n(t_{k+1}) = \frac{m}{n}\}} \binom{m}{h} \binom{h}{j} (-1)^{h-j} \varphi\left(t_k, \frac{m-j}{n}\right) \left(\frac{F(t_{k+1}) - F(t_k)}{F(t_{k+1})}\right)^h \\
 &= M_{t_{k+1}}^K - (\varphi(t_k, F_n(t_{k+1})) - \varphi(t_k, F_n(t_k))) \\
 &\quad + n F_n(t_{k+1}) \left(\varphi(t_k, F_n(t_{k+1})) - \varphi\left(t_k, F_n(t_{k+1}) - \frac{1}{n}\right)\right) \frac{F(t_{k+1}) - F(t_k)}{F(t_{k+1})} \\
 &\quad - \sum_{h=2}^n \sum_{m=h}^n \sum_{j=0}^h \mathbf{1}_{\{F_n(t_{k+1}) = \frac{m}{n}\}} \binom{m}{h} \binom{h}{j} (-1)^{h-j} \varphi\left(t_k, \frac{m-j}{n}\right) \left(\frac{F(t_{k+1}) - F(t_k)}{F(t_{k+1})}\right)^h.
 \end{aligned}$$

3 Die Empirische Verteilungsfunktion

Nun ergibt sich durch Induktion

$$\begin{aligned}
 M_{t_k}^K &= M_{t_k}^K - \sum_{l=k}^{K-1} \varphi(t_l, F_n(t_{l+1})) - \varphi(t_l, F_n(t_l)) \\
 &+ \sum_{l=k}^{K-1} n F_n(t_{l+1}) \left(\varphi(t_l, F_n(t_{l+1})) - \varphi\left(t_l, F_n(t_{l+1}) - \frac{1}{n}\right) \right) \frac{F(t_{l+1}) - F(t_l)}{F(t_{l+1})} \\
 &- \sum_{l=k}^{K-1} \sum_{h=2}^n \sum_{m=h}^n \sum_{j=0}^h 1_{\{F_n(t_{l+1}) = \frac{m}{n}\}} \\
 &\cdot \binom{m}{h} \binom{h}{j} (-1)^{h-j} \varphi\left(t_l, \frac{m-j}{n}\right) \left(\frac{F(t_{l+1}) - F(t_l)}{F(t_{l+1})} \right)^h.
 \end{aligned}$$

Es gilt

$$\begin{aligned}
 &\sum_{l=k}^{K-1} \varphi(t_l, F_n(t_{l+1})) - \varphi(t_l, F_n(t_l)) \\
 &= \sum_{l=k}^{K-1} \sum_{m=0}^n 1_{\{F_n(t_{l+1}) = \frac{m}{n}\}} (\varphi(t_l, F_n(t_{l+1})) - \varphi(t_l, F_n(t_l))) \\
 &= \sum_{l=k}^{K-1} \sum_{m=1}^n 1_{\{F_n(t_{l+1}) = \frac{m}{n}\}} (\varphi(t_l, F_n(t_{l+1})) - \varphi(t_l, F_n(t_l))) \\
 &= \sum_{l=k}^{K-1} \sum_{m=1}^n \sum_{j=0}^{m-1} 1_{\{F_n(t_l) = \frac{j}{n}\}} 1_{\{F_n(t_{l+1}) = \frac{m}{n}\}} (\varphi(t_l, F_n(t_{l+1})) - \varphi(t_l, F_n(t_l))) \\
 &= \sum_{l=k}^{K-1} \sum_{m=1}^n \sum_{j=0}^{m-1} 1_{\{F_n(t_l) = \frac{j}{n}\}} 1_{\{F_n(t_{l+1}) = \frac{m}{n}\}} \left(\varphi\left(t_l, \frac{m}{n}\right) - \varphi\left(t_l, \frac{j}{n}\right) \right) \\
 &= \sum_{l=k}^{K-1} \sum_{m=1}^n \sum_{j=0}^{m-1} \sum_{h=j}^{m-1} 1_{\{F_n(t_l) = \frac{j}{n}\}} 1_{\{F_n(t_{l+1}) = \frac{m}{n}\}} \left(\varphi\left(t_l, \frac{h+1}{n}\right) - \varphi\left(t_l, \frac{h}{n}\right) \right) \\
 &= \sum_{l=k}^{K-1} \sum_{m=1}^n \sum_{h=0}^{m-1} \sum_{j=0}^h 1_{\{F_n(t_l) = \frac{j}{n}\}} 1_{\{F_n(t_{l+1}) = \frac{m}{n}\}} \left(\varphi\left(t_l, \frac{h+1}{n}\right) - \varphi\left(t_l, \frac{h}{n}\right) \right)
 \end{aligned}$$

3 Die Empirische Verteilungsfunktion

$$\begin{aligned}
&= \sum_{l=k}^{K-1} \sum_{m=1}^n \sum_{h=0}^{m-1} 1_{\{F_n(t_l) \leq \frac{h}{n}\}} 1_{\{F_n(t_{l+1}) = \frac{m}{n}\}} \left(\varphi \left(t_l, \frac{h+1}{n} \right) - \varphi \left(t_l, \frac{h}{n} \right) \right) \\
&= \sum_{l=k}^{K-1} \sum_{h=0}^{n-1} \sum_{m=h+1}^n 1_{\{F_n(t_l) \leq \frac{h}{n}\}} 1_{\{F_n(t_{l+1}) = \frac{m}{n}\}} \left(\varphi \left(t_l, \frac{h+1}{n} \right) - \varphi \left(t_l, \frac{h}{n} \right) \right) \\
&= \sum_{l=k}^{K-1} \sum_{h=0}^{n-1} 1_{\{F_n(t_l) \leq \frac{h}{n}\}} 1_{\{F_n(t_{l+1}) > \frac{h}{n}\}} \left(\varphi \left(t_l, \frac{h+1}{n} \right) - \varphi \left(t_l, \frac{h}{n} \right) \right) \\
&= \sum_{l=k}^{K-1} \sum_{m=1}^n 1_{\{t_l < X_{m:n} \leq t_{l+1}\}} \left(\varphi \left(t_l, \frac{m}{n} \right) - \varphi \left(t_l, \frac{m-1}{n} \right) \right).
\end{aligned}$$

Setze $M_{t_K}^K := X_{t_K}$. Dann gilt

$$\begin{aligned}
M_{t_k}^K &= X_{t_k} - \sum_{l=k}^{K-1} \sum_{m=1}^n 1_{\{t_l < X_{m:n} \leq t_{l+1}\}} \left(\varphi \left(t_l, \frac{m}{n} \right) - \varphi \left(t_l, \frac{m-1}{n} \right) \right) \\
&\quad + \sum_{l=k}^{K-1} n F_n(t_{l+1}) \left(\varphi \left(t_l, F_n(t_{l+1}) \right) - \varphi \left(t_l, F_n(t_{l+1}) - \frac{1}{n} \right) \right) \frac{F(t_{l+1}) - F(t_l)}{F(t_{l+1})} \\
&\quad - \sum_{l=k}^{K-1} \sum_{h=2}^n \sum_{m=h}^n \sum_{j=0}^h 1_{\{F_n(t_{l+1}) = \frac{m}{n}\}} \\
&\quad \cdot \binom{m}{h} \binom{h}{j} (-1)^{h-j} \varphi \left(t_l, \frac{m-j}{n} \right) \left(\frac{F(t_{l+1}) - F(t_l)}{F(t_{l+1})} \right)^h.
\end{aligned}$$

Nun erhält man durch Grenzübergang wieder die zeitstetige Version. Die Zuwächse quadratischer und höherer Ordnung verschwinden hierbei und das resultierende reverse Martingal lässt sich als Integral bezüglich des reversen \mathcal{G} -Doob-Meyer-Martingals M^0 der Empirischen Verteilungsfunktion schreiben.

Satz 3.18. *Sei $t_K \in \mathbb{R}$ ein beliebiger Startwert. Der Prozess $X = (X_t)_{t \in (-\infty, t_K]}$ mit $X_t = \varphi(t, F_n(t))$ besitzt das reverse Innovationsmartingal (M, \mathcal{G}) , wobei $\mathcal{G} = (\mathcal{G}_t)_{t \in \mathbb{R}}$ mit $\mathcal{G}_t =$*

3 Die Empirische Verteilungsfunktion

$\sigma(F_n(u) : u \geq t)$, und $M = (M_t)_{t \in (-\infty, t_K]}$ mit

$$\begin{aligned} M_t &= X_{t_K} + \int_{[t, t_K)} n \left(\varphi(u, F_n(u)) - \varphi\left(u, F_n(u) - \frac{1}{n}\right) \right) dF_n(u) \\ &\quad - \int_{[t, t_K)} n \left(\varphi(u, F_n(u)) - \varphi\left(u, F_n(u) - \frac{1}{n}\right) \right) \frac{F_n(u)}{F(u)} dF(u) \\ &= X_{t_K} + \int_{[t, t_K)} n \left(\varphi(u, F_n(u)) - \varphi\left(u, F_n(u) - \frac{1}{n}\right) \right) dM_u^0. \end{aligned}$$

Satz 3.19. *Es existiere der Grenzwert*

$$x_\infty := \lim_{s \rightarrow \infty} \varphi(s, 1).$$

Dann besitzt der Prozess $X = (X_t)_{t \in \mathbb{R}}$ mit $X_t = \varphi(t, F_n(t))$ das reverse Innovationsmartingal (M, \mathcal{G}) , wobei $\mathcal{G} = (\mathcal{G}_t)_{t \in \mathbb{R}}$ mit $\mathcal{G}_t = \sigma(F_n(u) : u \geq t)$, und $M = (M_t)_{t \in \mathbb{R}}$ mit

$$\begin{aligned} M_t &= x_\infty + \int_{[t, \infty)} n \left(\varphi(u, F_n(u)) - \varphi\left(u, F_n(u) - \frac{1}{n}\right) \right) dF_n(u) \\ &\quad - \int_{[t, \infty)} n \left(\varphi(u, F_n(u)) - \varphi\left(u, F_n(u) - \frac{1}{n}\right) \right) \frac{F_n(u)}{F(u)} dF(u) \\ &= x_\infty + \int_{[t, \infty)} n \left(\varphi(u, F_n(u)) - \varphi\left(u, F_n(u) - \frac{1}{n}\right) \right) dM_u^0. \end{aligned}$$

Weiterhin liefert die Doob-Meyer-Zerlegung folgenden Kompensator

$$\begin{aligned} A_{t_k}^K &= A_{t_{k+1}}^K - X_{t_{k+1}} + \mathbb{E}(X_{t_k} | \mathcal{G}_{t_{k-1}}) \\ &= A_{t_{k+1}}^K - \varphi(t_{k+1}, F_n(t_{k+1})) \end{aligned}$$

3 Die Empirische Verteilungsfunktion

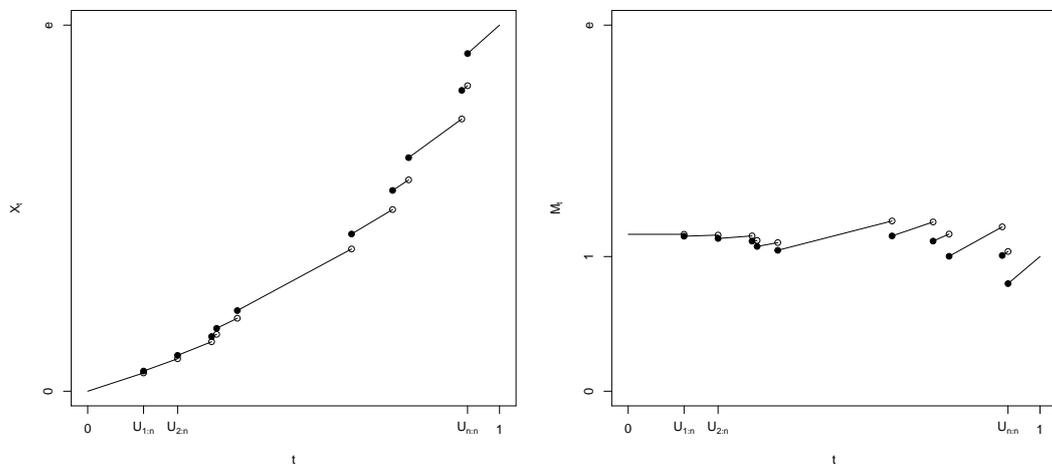


Abbildung 3.8: Der Prozess X mit $X_t = t \exp(F_n(t))$ und sein reverses Doob-Meyer-Martingal (M, \mathcal{G}) , wobei die unterliegenden Zufallsvariablen U_1, \dots, U_n , $n = 10$, unabhängig und identisch gleichverteilt sind auf $[0, 1]$.

3 Die Empirische Verteilungsfunktion

$$\begin{aligned}
& + \sum_{h=0}^n \sum_{m=h}^n \sum_{j=0}^h 1_{\{F_n(t_{k+1}) = \frac{m}{n}\}} \binom{m}{h} \binom{h}{j} (-1)^{h-j} \varphi\left(t_k, \frac{m-j}{n}\right) \left(\frac{F(t_{k+1}) - F(t_k)}{F(t_{k+1})}\right)^h \\
& = A_{t_{k+1}}^K - \varphi(t_{k+1}, F_n(t_{k+1})) + \varphi(t_k, F_n(t_{k+1})) \\
& \quad - nF_n(t_{k+1}) \left(\varphi(t_k, F_n(t_{k+1})) - \varphi\left(t_k, F_n(t_{k+1}) - \frac{1}{n}\right)\right) \frac{F(t_{k+1}) - F(t_k)}{F(t_{k+1})} \\
& \quad + \sum_{h=2}^n \sum_{m=h}^n \sum_{j=0}^h 1_{\{F_n(t_{k+1}) = \frac{m}{n}\}} \\
& \quad \cdot \binom{m}{h} \binom{h}{j} (-1)^{h-j} \varphi\left(t_k, F_n(t_{k+1}) - \frac{j}{n}\right) \left(\frac{F(t_{k+1}) - F(t_k)}{F(t_{k+1})}\right)^h.
\end{aligned}$$

Durch Induktion ergibt sich

$$\begin{aligned}
A_{t_k}^K & = - \sum_{l=k}^{K-1} (\varphi(t_{l+1}, F_n(t_{l+1})) - \varphi(t_l, F_n(t_{l+1}))) \\
& \quad - \sum_{l=k}^{K-1} nF_n(t_{l+1}) \left(\varphi(t_l, F_n(t_{l+1})) - \varphi\left(t_l, F_n(t_{l+1}) - \frac{1}{n}\right)\right) \frac{F(t_{l+1}) - F(t_l)}{F(t_{l+1})} \\
& \quad + \sum_{l=k}^{K-1} \sum_{h=2}^n \sum_{m=h}^n \sum_{j=0}^h 1_{\{F_n(t_{l+1}) = \frac{m}{n}\}} \\
& \quad \cdot \binom{m}{h} \binom{h}{j} (-1)^{h-j} \varphi\left(t_l, F_n(t_{l+1}) - \frac{j}{n}\right) \left(\frac{F(t_{l+1}) - F(t_l)}{F(t_{l+1})}\right)^h.
\end{aligned}$$

Nach Grenzübergang ergibt sich wiederum der zeitstetige reverse Kompensator:

Satz 3.20. *Der Prozess $(X_t)_{t \in \mathbb{R}}$ mit $X_t = \varphi(t, F_n(t))$ besitzt den reversen Kompensator (A, \mathcal{G}) , wobei $\mathcal{G} = (\mathcal{G}_t)_{t \in \mathbb{R}}$ mit $\mathcal{G}_t = \sigma(F_n(u) : u \geq t)$ und $A = (A_t)_{t \in \mathbb{R}}$ mit*

$$A_t = - \int_{[t, \infty)} \frac{\partial \varphi}{\partial F}(u, F_n(u)) + n \left(\varphi(u, F_n(u)) - \varphi\left(u, F_n(u) - \frac{1}{n}\right)\right) \frac{F_n(u)}{F(u)} F(du).$$

3 Die Empirische Verteilungsfunktion

Die reverse \mathcal{G} -Doob-Meyer-Zerlegung von X liefert wiederum ein reverses Martingal, welches aus einem Startwert und einem Integral bezüglich M^0 - dem reversen \mathcal{G} -Doob-Meyer-Martingal der Empirischen Verteilungsfunktion - besteht, und einen reversen Kompensator, welcher als Integral bezüglich F , also einer deterministischen Funktion, geschrieben werden kann. Daraus lässt sich eine reverse Martingaldifferentialgleichung ableiten, deren Lösungen zu reversen Martingaltransformationen führen. Dies soll im Folgenden geschehen.

3.2.3 Reverse Martingaltransformationen der Empirischen Verteilungsfunktion

Wie zuvor ist (X, \mathcal{G}) genau dann ein reverses Martingal, falls der Integrand im reversen Kompensator aus Satz 3.20 verschwindet. Dies führt zu folgender reverser Martingaldifferentialgleichung, deren Lösungen im Anschluss hergeleitet werden.

Satz 3.21. *Der Prozess $X = (X_t)_{t \in \mathbb{R}}$ mit $X_t = \varphi(t, F_n(t))$ ist ein reverses Martingal bzgl. $\mathcal{G} = (\mathcal{G}_t)_{t \in \mathbb{R}}$ mit $\mathcal{G}_t = \sigma(F_n(u) : u \geq t)$, falls für alle $t \in \mathbb{R}$ und $x \in \{\frac{1}{n}, \dots, \frac{n-1}{n}, 1\}$ gilt*

$$\frac{\partial \varphi}{\partial F}(t, x) + n \left(\varphi(t, x) - \varphi\left(t, x - \frac{1}{n}\right) \right) \frac{x}{F(t)} = 0.$$

3.2.3.1 Affine reverse Martingaltransformationen der Empirischen Verteilungsfunktion

Wie auch im vorhergehenden Abschnitt sollen zunächst die affinen Lösungen der reversen Martingaldifferentialgleichung aus Satz 3.21 gefunden werden, wobei affin bedeutet, dass für $t \in \mathbb{R}$ die Funktion $\varphi(t, \cdot)$ affin ist. Diese Lösungen führen zu folgenden affinen reversen \mathcal{G} -Martingaltransformationen.

3 Die Empirische Verteilungsfunktion

Satz 3.22. Seien $c_0, c_1 \in \mathbb{R}$. Dann ist der Prozess $X = (X_t)_{t \in \mathbb{R}}$ mit

$$X_t = c_1 \frac{F_n(t)}{F(t)} + c_0$$

ein reverses Martingal bzgl. $\mathcal{G} = (\mathcal{G}_t)_{t \in \mathbb{R}}$ mit $\mathcal{G}_t = \sigma(F_n(u) : u \geq t)$.

Beweis. Sei

$$\varphi(t, x) = a_1(F(t))x + a_0(F(t)).$$

Dann gilt

$$\frac{\partial \varphi}{\partial F}(t, x) = \frac{\partial a_1}{\partial F}(F(t))x + \frac{\partial a_0}{\partial F}(F(t))$$

und $X_t = \varphi(t, F_n(t))$ ist ein Martingal bzgl. \mathcal{G} , falls

$$\left(\frac{\partial a_1}{\partial F}(F(t)) + \frac{a_1(F(t))}{F(t)} \right) x + \frac{\partial a_0}{\partial F}(F(t)) = 0.$$

Nun muss also gelten

$$a_1(F(t)) = \frac{c_1}{F(t)} \quad \text{und} \quad a_0(F(t)) = c_0.$$

Also ist für $c_0, c_1 \in \mathbb{R}$

$$X_t = c_1 \frac{F_n(t)}{F(t)} + c_0$$

ein reverses \mathcal{G} -Martingal.

□

3 Die Empirische Verteilungsfunktion

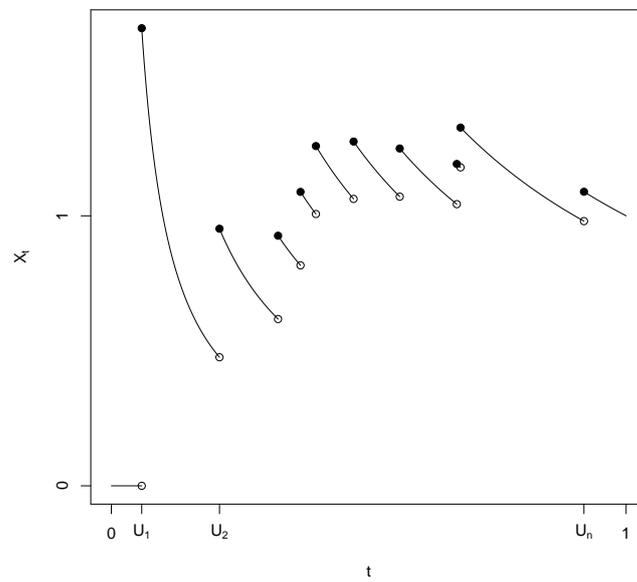


Abbildung 3.9: Das reverse Martingal (X, \mathcal{G}) mit $X_t = \frac{F_n(t)}{t}$, wobei die unterliegenden Zufallsvariablen U_1, \dots, U_n , $n = 10$, unabhängig und identisch gleichverteilt sind auf $[0, 1]$.

3.2.3.2 Polynomiale reverse Martingaltransformationen der Empirischen Verteilungsfunktion

Wie auch im nicht-reversen Fall sind die affinen reversen Martingaltransformationen aus Satz 3.22 ein Spezialfall aus der Klasse der polynomialen reversen \mathcal{G} -Martingaltransformationen der Empirischen Verteilungsfunktion. Auch hier ist mit polynomial gemeint, dass für $t \in \mathbb{R}$ die Funktion $\varphi(t, \cdot)$ ein Polynom ist. Im folgenden Satz werden Vertreter der Klasse der polynomialen reversen \mathcal{G} -Martingaltransformationen der Empirischen Verteilungsfunktion vorgestellt, welche eine Basis aller polynomialen reversen \mathcal{G} -Martingaltransformationen der Empirischen Verteilungsfunktion der Form $X = (X_t)_{t \in \mathbb{R}}$ mit $X_t = \varphi(t, F_n(t))$ sind.

Satz 3.23. *Sei $m \in \{0, \dots, n\}$. Dann ist der Prozess $X^m = (X_t^m)_{t \in \mathbb{R}}$ mit*

$$X_t^m = \frac{(n-m)!n^m}{n!} \prod_{j=0}^{m-1} \frac{F_n(t) - \frac{j}{n}}{F(t)}$$

ein reverses Martingal bzgl. $\mathcal{G} = (\mathcal{G}_t)_{t \in \mathbb{R}}$ mit $\mathcal{G}_t = \sigma(F_n(u) : u \geq t)$.

Beweis. Setze

$$\varphi(t, x) = \frac{\prod_{j=0}^{m-1} (x - \frac{j}{n})}{F(t)^m}.$$

Dann gilt

$$\frac{\partial \varphi}{\partial F}(t, x) = -\frac{m \prod_{j=0}^{m-1} (x - \frac{j}{n})}{F(t)^{m+1}}$$

3 Die Empirische Verteilungsfunktion

und

$$\begin{aligned}
 & n \left(\varphi(t, x) - \varphi\left(t, x - \frac{1}{n}\right) \right) \frac{x}{F(t)} \\
 &= n \left(\frac{\prod_{j=0}^{m-1} \left(x - \frac{j}{n}\right)}{F(t)^m} - \frac{\prod_{j=0}^{m-1} \left(x - \frac{j+1}{n}\right)}{F(t)^m} \right) \frac{x}{F(t)} \\
 &= \frac{nx \left(\prod_{j=0}^{m-1} \left(x - \frac{j}{n}\right) - \prod_{j=1}^m \left(x - \frac{j}{n}\right) \right)}{F(t)^{m+1}} \\
 &= \frac{mx \prod_{j=1}^{m-1} \left(x - \frac{j}{n}\right)}{F(t)^{m+1}} \\
 &= \frac{m \prod_{j=0}^{m-1} \left(x - \frac{j}{n}\right)}{F(t)^{m+1}}.
 \end{aligned}$$

Damit erfüllt φ die reverse Martingaldifferentialgleichung aus Satz 3.21 und

$$X_t = \frac{(n-m)!n^m}{n!} \varphi(t, F_n(t))$$

ist ein reverses Martingal bzgl. \mathcal{G} .

□

Bemerkung 3.24. 1. Durch den Faktor $\frac{(n-m)!n^m}{n!}$ wird sichergestellt, dass die reversen Martingale aus Satz 3.23 alle im Wert 1 starten.

2. Die Martingale aus Satz 3.23 sind nichtnegativ und fallen auf den Wert 0, sobald die Empirische Verteilungsfunktion den Wert $\frac{m-1}{n}$ erreicht. Zur vollständigen Kennt-

3 Die Empirische Verteilungsfunktion

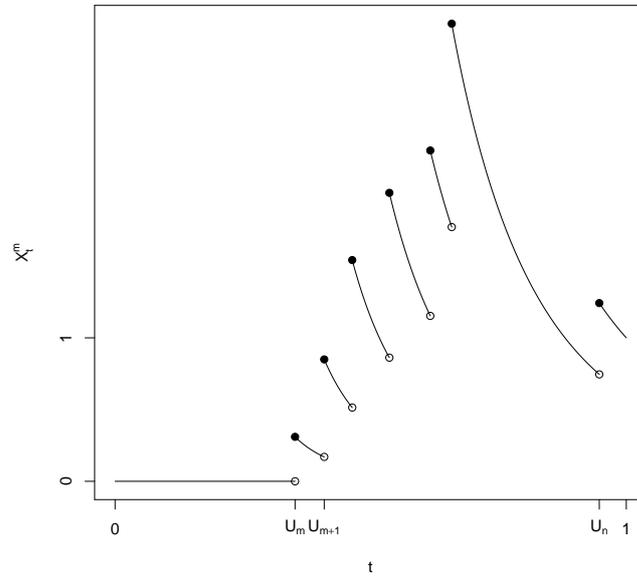


Abbildung 3.10: Das reverse Martingal (X^m, \mathcal{G}) aus Satz 3.23 mit $n = 10$ und $m = 4$, wobei die unterliegenden Zufallsvariablen U_1, \dots, U_n unabhängig und identisch gleichverteilt sind auf $[0, 1]$.

nis des Martingals werden also nur die letzten $n - (m - 1)$ Ordnungsstatistiken benötigt. Dadurch werden diese reversen Martingale interessant im Fall von Typ-II-Linkszensierungen, in denen nur die letzten $n - (m - 1)$ Ordnungsstatistiken beobachtet werden.

3. Durch Linearkombinationen der reversen Martingale aus Satz 3.23 erhält man sämtliche polynomiale reverse \mathcal{G} -Martingale, die sich aus F_n gewinnen lassen.

3.2.3.3 Exponentielle reverse Martingaltransformationen der Empirischen Verteilungsfunktion

Nun sollen mit einem exponentiellen Ansatz weitere reverse Martingaltransformationen gefunden werden. Die entsprechenden Lösungen führen zu folgenden reversen Martingalen.

Satz 3.25. Seien $c_1, c_2 \in \mathbb{R}$. Dann ist der Prozess $X = (X_t)_{t \in \mathbb{R}}$ mit

$$X_t = c_2 \left(1 + \frac{c_1}{F(t)} \right)^{nF_n(t)}$$

ein reverses Martingal bzgl. $\mathcal{G} = (\mathcal{G}_t)_{t \in \mathbb{R}}$ mit $\mathcal{G}_t = \sigma(F_n(u) : u \geq t)$.

Beweis. Setze

$$\varphi(t, x) = \exp(a_1(F(t))x + a_0(F(t))).$$

Dann gilt

$$\frac{\partial \varphi}{\partial F}(t, x) = \left(\frac{\partial a_1}{\partial F}(F(t))x + \frac{\partial a_0}{\partial F}(F(t)) \right) \varphi(t, x)$$

sowie

$$\begin{aligned} n \left(\varphi(t, x) - \varphi\left(t, x - \frac{1}{n}\right) \right) &= n \left(\varphi(t, x) - \exp\left(a_1(F(t))\left(x - \frac{1}{n}\right) + a_0(F(t))\right) \right) \\ &= n \left(1 - \exp\left(-\frac{a_1(F(t))}{n}\right) \right) \varphi(t, x) \end{aligned}$$

und $X_t = \varphi(t, F_n(t))$ ist ein reverses Martingal bzgl. \mathcal{G} , falls

$$\left(\frac{\partial a_1}{\partial F}(F(t)) + n \frac{1 - \exp\left(-\frac{a_1(F(t))}{n}\right)}{F(t)} \right) x + \frac{\partial a_0}{\partial F}(F(t)) = 0.$$

3 Die Empirische Verteilungsfunktion

Nun gilt

$$\begin{aligned} \frac{\partial a_1}{\partial F}(F(t)) + n \frac{1 - \exp\left(-\frac{a_1(F(t))}{n}\right)}{F(t)} = 0 &\Leftrightarrow \frac{\partial a_1}{\partial F}(F(t)) = -n \frac{1 - \exp\left(-\frac{a_1(F(t))}{n}\right)}{F(t)} \\ \Leftrightarrow a_1(F(t)) = n \ln\left(1 + \frac{c_1}{F(t)}\right) \end{aligned}$$

sowie

$$\frac{\partial a_0}{\partial F}(F(t)) = 0 \Leftrightarrow a_0(F(t)) = c_2.$$

Damit gilt also

$$\begin{aligned} \varphi(t, x) &= \exp(a_1(F(t))x + a_0(F(t))) \\ &= \exp\left(nx \ln\left(1 + \frac{c_1}{F(t)}\right) + \tilde{c}_2\right) \\ &= c_2 \left(1 + \frac{c_1}{F(t)}\right)^{nx}. \end{aligned}$$

□

Bei genauerer Betrachtung sind die in Satz 3.25 beschriebenen exponentiellen reversen Martingale wieder lediglich spezielle polynomiale reverse \mathcal{G} -Martingale.

Satz 3.26. Sei $X = (X_t)_{t \in \mathbb{R}}$ ein exponentielles reverses \mathcal{G} -Martingal mit $c_1, c_2 \in \mathbb{R}$, d.h.

$$X_t = c_2 \left(1 + \frac{c_1}{F(t)}\right)^{nF_n(t)}$$

3 Die Empirische Verteilungsfunktion

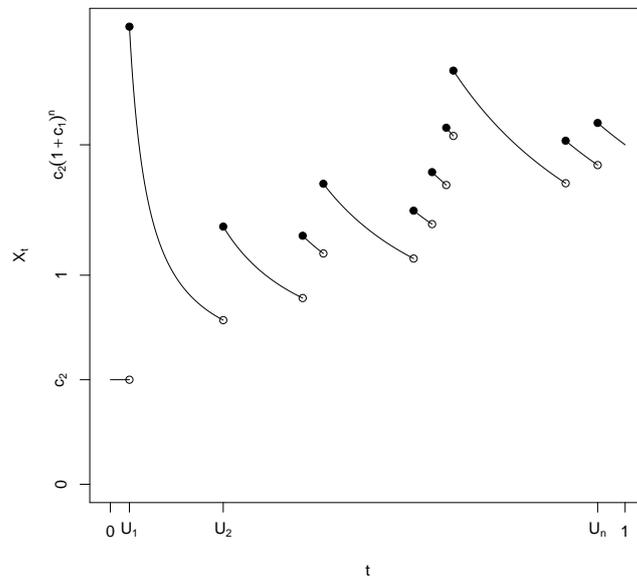


Abbildung 3.11: Das reverse Martingal (X, \mathcal{G}) mit $X_t = c_2 \left(1 + \frac{c_1}{t}\right)^{nF_n(t)}$ mit $c_1 = \frac{1}{8}$ und $c_2 = \frac{1}{2}$, wobei die unterliegenden Zufallsvariablen U_1, \dots, U_n , $n = 10$, unabhängig und identisch gleichverteilt sind auf $[0, 1]$.

3 Die Empirische Verteilungsfunktion

und seien $X^m = (X_t^m)_{t \in \mathbb{R}}$, $m \in \{0, \dots, n\}$, polynomiale reverse \mathcal{G} -Martingale mit

$$X_t^m = \frac{(n-m)!n^m}{n!} \prod_{j=0}^{m-1} \frac{F_n(t) - \frac{j}{n}}{F(t)}.$$

Dann gilt für alle $t \in \mathbb{R}$

$$X_t = \sum_{m=0}^n a_m X_t^m,$$

wobei

$$a_m = \binom{n}{m} c_1^m c_2.$$

Beweis. Es gilt einerseits

$$\begin{aligned} X_t &= c_2 \left(1 + \frac{c_1}{F(t)} \right)^{nF_n(t)} \\ &= \sum_{m=0}^{nF_n(t)} c_1^m c_2 \binom{nF_n(t)}{m} \frac{1}{F(t)^m} \end{aligned}$$

und andererseits

$$\begin{aligned} \sum_{m=0}^n a_m X_t^m &= \sum_{m=0}^n a_m \frac{(n-m)!n^m}{n!} \prod_{j=0}^{m-1} \frac{F_n(t) - \frac{j}{n}}{F(t)} \\ &= \sum_{m=0}^n a_m \frac{(n-m)!}{n! F(t)^m} \prod_{j=0}^{m-1} (nF_n(t) - j) \\ &= \sum_{m=0}^n a_m \frac{(n-m)! (nF_n(t))!}{n! (nF_n(t) - m)! F(t)^m} \\ &= \sum_{m=0}^n \frac{a_m}{\binom{n}{m}} \binom{nF_n(t)}{m} \frac{1}{F(t)^m}. \end{aligned}$$

3 Die Empirische Verteilungsfunktion

Für $a_m = \binom{n}{m} c_1^m c_2$ gilt also $X = \sum_{m=0}^n a_m X^m$.

□

3.2.3.4 Allgemeine reverse Martingaltransformationen der Empirischen Verteilungsfunktion

Satz 3.26 hat gezeigt, dass sich die exponentiellen reversen \mathcal{G} -Martingale als Linearkombinationen der polynomialen reversen \mathcal{G} -Martingale darstellen lassen. Im Folgenden wird gezeigt, dass die polynomialen reversen \mathcal{G} -Martingale eine Basis für den Raum aller reversen Martingale (X, \mathcal{G}) mit $X = (\varphi(t, F_n(t)))_{t \in \mathbb{R}}$ und $\mathcal{G} = (\mathcal{G}_t)_{t \in \mathbb{R}}$ mit $\mathcal{G}_t = \sigma(F_n(u) : u \geq t)$ bilden.

Satz 3.27. Sei $X = (X_t)_{t \in \mathbb{R}}$ mit $X_t = \varphi(t, F_n(t))$ ein reverses Martingal bzgl. $\mathcal{G} = (\mathcal{G}_t)_{t \in \mathbb{R}}$. Dann existieren $c_0, \dots, c_n \in \mathbb{R}$, sodass

$$X_t = \sum_{m=0}^n c_m X_t^m,$$

wobei

$$X_t^m = \frac{(n-m)! n^m}{n!} \prod_{j=0}^{m-1} \frac{F_n(t) - \frac{j}{n}}{F(t)}, \quad m \in \{0, \dots, n\}.$$

Beweis. Die Empirische Verteilungsfunktion kann nur die Zustände $0, \frac{1}{n}, \dots, 1$ annehmen. Daher lässt sich jeder Prozess X mit $X_t = \varphi(t, F_n(t))$ wie folgt schreiben:

$$X_t = \sum_{m=0}^n \varphi_m(t) 1_{\{F_n(t) = \frac{m}{n}\}}.$$

3 Die Empirische Verteilungsfunktion

Damit X ein reverses \mathcal{G} -Martingal ist, muss für $m \in \{0, \dots, n\}$ gelten, dass

$$\begin{aligned}
 \varphi_m(u) &= \mathbb{E} \left(X_t \mid F_n(u) = \frac{m}{n} \right) \\
 &= \sum_{j=0}^m \varphi_j(t) \mathbb{P} \left(F_n(t) = \frac{j}{n} \mid F_n(u) = \frac{m}{n} \right) \\
 &= \sum_{j=0}^m \varphi_j(t) \binom{m}{j} \left(\frac{F(t)}{F(u)} \right)^j \left(1 - \frac{F(t)}{F(u)} \right)^{m-j} \\
 &= \varphi_m(t) \left(\frac{F(t)}{F(u)} \right)^m + \sum_{j=0}^{m-1} \varphi_j(t) \binom{m}{j} \left(\frac{F(t)}{F(u)} \right)^j \left(1 - \frac{F(t)}{F(u)} \right)^{m-j}.
 \end{aligned}$$

bzw.

$$\begin{aligned}
 \varphi_m(u) &= \varphi_m(t) \left(\frac{F(t)}{F(u)} \right)^m + \sum_{j=0}^{m-1} \varphi_j(t) \binom{m}{j} \left(\frac{F(t)}{F(u)} \right)^j \left(\frac{F(u) - F(t)}{F(u)} \right)^{m-j} \\
 \Leftrightarrow \varphi_m(t) &= \varphi_m(u) \left(\frac{F(u)}{F(t)} \right)^m - \sum_{j=0}^{m-1} \varphi_j(t) \binom{m}{j} \left(\frac{F(u) - F(t)}{F(t)} \right)^{m-j} \\
 \Leftrightarrow \varphi_m(t) &= \varphi_m(u) \left(\frac{F(u)}{F(t)} \right)^m - \sum_{j=0}^{m-1} \varphi_j(t) \binom{m}{j} \left(\frac{F(u)}{F(t)} - 1 \right)^{m-j} \\
 \Leftrightarrow \varphi_m(t) &= \varphi_m(u) \left(\frac{F(u)}{F(t)} \right)^m \\
 &\quad - \sum_{j=0}^{m-1} \varphi_j(t) \binom{m}{j} \sum_{i=0}^{m-j} \binom{m-j}{i} (-1)^{m-j-i} \left(\frac{F(u)}{F(t)} \right)^i \\
 \Leftrightarrow \varphi_m(t) &= \varphi_m(u) \left(\frac{F(u)}{F(t)} \right)^m \\
 &\quad - \sum_{j=0}^{m-1} \varphi_j(t) \binom{m}{j} \sum_{i=j}^m \binom{m-j}{m-i} (-1)^{m-i} \left(\frac{F(u)}{F(t)} \right)^{i-j}.
 \end{aligned}$$

Man betrachte den Startpunkt $u_1 = \inf_{u \in \mathbb{R}} \{F(u) = 1\}$, wobei $u_1 = \infty$, falls $F(u) < 1$ für alle

3 Die Empirische Verteilungsfunktion

$u \in \mathbb{R}$. Nun gilt obige Gleichung auch für $u = u_1$, also

$$\varphi_m(t) = \varphi_m(u_1) \left(\frac{1}{F(t)} \right)^m - \sum_{j=0}^{m-1} \varphi_j(t) \binom{m}{j} \sum_{i=j}^m \binom{m-j}{m-i} (-1)^{m-i} \left(\frac{1}{F(t)} \right)^{i-j}$$

bzw. falls $u_1 = \infty$

$$\varphi_m(t) = \lim_{u \rightarrow \infty} \varphi_m(u) \left(\frac{1}{F(t)} \right)^m - \sum_{j=0}^{m-1} \varphi_j(t) \binom{m}{j} \sum_{i=j}^m \binom{m-j}{m-i} (-1)^{m-i} \left(\frac{1}{F(t)} \right)^{i-j}.$$

Setze $\varphi_m(u_1) = a_m \in \mathbb{R}$, falls $u_1 \in \mathbb{R}$, bzw. $\lim_{u \rightarrow \infty} \varphi_m(u) = a_m \in \mathbb{R}$, falls $u_1 = \infty$. Dann folgt also

$$\varphi_m(t) = \frac{a_m}{F(t)^m} - \sum_{j=0}^{m-1} \varphi_j(t) \binom{m}{j} \sum_{i=j}^m \binom{m-j}{m-i} (-1)^{m-i} \frac{1}{F(t)^{i-j}}.$$

Mit vollständiger Induktion zeigt man nun, dass für gewisse b_j^m mit $m \in \{0, \dots, n\}$ und $j \in \{0, \dots, n-m\}$ für die explizite Darstellung gilt

$$\varphi_m(t) = \sum_{j=0}^m \frac{b_j^m}{F(t)^j}, \quad b_j^m \in \mathbb{R}.$$

Mit $b_0^0 = a_0$ ist diese Behauptung für $m = 0$ offensichtlich richtig. Die Behauptung gelte nun für alle $m < m_0$, wobei $m_0 \in \{1, \dots, n\}$. Dann folgt

$$\begin{aligned} \varphi_{m_0}(t) &= \frac{a_{m_0}}{F(t)^{m_0}} - \sum_{m=0}^{m_0-1} \varphi_m(t) \binom{m_0}{m} \sum_{i=m}^{m_0} \binom{m_0-m}{m_0-i} (-1)^{m_0-i} \frac{1}{F(t)^{i-m}} \\ &= \frac{a_{m_0}}{F(t)^{m_0}} - \sum_{m=0}^{m_0-1} \sum_{j=0}^m \sum_{i=m}^{m_0} (-1)^{m_0-i} \binom{m_0}{m} \binom{m_0-m}{m_0-i} \frac{b_j^m}{F(t)^{i-m+j}} \\ &= \frac{a_{m_0}}{F(t)^{m_0}} - \sum_{m=0}^{m_0-1} \sum_{i=m}^{m_0} \sum_{j=0}^m (-1)^{m_0-i} \binom{m_0}{m} \binom{m_0-m}{m_0-i} \frac{b_j^m}{F(t)^{i-m+j}} \end{aligned}$$

3 Die Empirische Verteilungsfunktion

$$\begin{aligned}
&= \frac{a_{m_0}}{F(t)^{m_0}} - \sum_{m=0}^{m_0-1} \sum_{i=m}^{m_0} \sum_{j=i-m}^i (-1)^{m_0-i} \binom{m_0}{m} \binom{m_0-m}{m_0-i} \frac{b_{j-(i-m)}^m}{F(t)^j} \\
&= \frac{a_{m_0}}{F(t)^{m_0}} - \sum_{m=0}^{m_0-1} \sum_{j=m_0-m}^{m_0} \binom{m_0}{m} \frac{b_{j-(m_0-m)}^m}{F(t)^j} \\
&\quad - \sum_{m=0}^{m_0-1} \sum_{i=m}^{m_0-1} \sum_{j=i-m}^i (-1)^{m_0-i} \binom{m_0}{m} \binom{m_0-m}{m_0-i} \frac{b_{j-(i-m)}^m}{F(t)^j} \\
&= \frac{a_{m_0}}{F(t)^{m_0}} - \sum_{j=1}^{m_0} \sum_{m=m_0-j}^{m_0-1} \binom{m_0}{m} \frac{b_{j-(m_0-m)}^m}{F(t)^j} \\
&\quad - \sum_{i=0}^{m_0-1} \sum_{m=0}^i \sum_{j=i-m}^i (-1)^{m_0-i} \binom{m_0}{m} \binom{m_0-m}{m_0-i} \frac{b_{j-(i-m)}^m}{F(t)^j} \\
&= \frac{a_{m_0}}{F(t)^{m_0}} - \sum_{j=1}^{m_0} \sum_{m=m_0-j}^{m_0-1} \binom{m_0}{m} \frac{b_{j-(m_0-m)}^m}{F(t)^j} \\
&\quad - \sum_{i=0}^{m_0-1} \sum_{j=0}^i \sum_{m=i-j}^i (-1)^{m_0-i} \binom{m_0}{m} \binom{m_0-m}{m_0-i} \frac{b_{j-(i-m)}^m}{F(t)^j} \\
&= \frac{a_{m_0}}{F(t)^{m_0}} - \sum_{j=1}^{m_0} \sum_{m=m_0-j}^{m_0-1} \binom{m_0}{m} \frac{b_{j-(m_0-m)}^m}{F(t)^j} \\
&\quad - \sum_{j=0}^{m_0-1} \sum_{i=j}^{m_0-1} \sum_{m=i-j}^i (-1)^{m_0-i} \binom{m_0}{m} \binom{m_0-m}{m_0-i} \frac{b_{j-(i-m)}^m}{F(t)^j}.
\end{aligned}$$

Also gilt

$$\varphi_{m_0}(t) = \sum_{j=0}^{m_0} \frac{b_j^{m_0}}{F(t)^j}, \quad b_j^{m_0} \in \mathbb{R}$$

mit

$$b_0^{m_0} = - \sum_{m=0}^{m_0-1} (-1)^{m_0-m} \binom{m_0}{m} b_0^m,$$

$$b_{m_0}^{m_0} = a_{m_0} - \sum_{m=0}^{m_0-1} \binom{m_0}{m} b_m^m$$

3 Die Empirische Verteilungsfunktion

und

$$b_j^{m_0} = - \sum_{m=m_0-j}^{m_0-1} \binom{m_0}{m} b_{m-(m_0-j)}^m - \sum_{l=j}^{m_0-1} \sum_{m=l-j}^l (-1)^{m_0-l} \binom{m_0}{m} \binom{m_0-m}{m_0-l} b_{m-(l-j)}^m$$

für $j \in \{1, \dots, m_0 - 1\}$.

Mit einer weiteren vollständigen Induktion zeigt man, dass

$$b_j^m = \binom{m}{j} \sum_{i=0}^j \binom{j}{i} (-1)^{j-i} a_i, \quad m = 0, \dots, n, \quad j = 0, \dots, m.$$

Offensichtlich gilt die Behauptung für $m = 0$. Sei $m_0 \in \{1, \dots, n\}$. Die Behauptung gelte für alle $m < m_0$. Nun gilt

$$\begin{aligned} b_0^{m_0} &= - \sum_{m=0}^{m_0-1} (-1)^{m_0-m} \binom{m_0}{m} b_0^m \\ &= - \sum_{m=0}^{m_0-1} (-1)^{m_0-m} \binom{m_0}{m} a_0 \\ &= a_0 \left(1 - \sum_{m=0}^{m_0} (-1)^{m_0-m} \binom{m_0}{m} \right) \\ &= a_0. \end{aligned}$$

Weiter gilt

$$\begin{aligned} b_{m_0}^{m_0} &= a_{m_0} - \sum_{m=0}^{m_0-1} \binom{m_0}{m} b_m^m \\ &= a_{m_0} - \sum_{m=0}^{m_0-1} \binom{m_0}{m} \sum_{i=0}^m \binom{m}{i} (-1)^{m-i} a_i \end{aligned}$$

3 Die Empirische Verteilungsfunktion

$$\begin{aligned}
&= a_{m_0} - \sum_{i=0}^{m_0-1} a_i \sum_{m=i}^{m_0-1} \binom{m_0}{m} \binom{m}{i} (-1)^{m-i} \\
&= a_{m_0} - \sum_{i=0}^{m_0-1} \binom{m_0}{i} a_i \sum_{m=i}^{m_0-1} \binom{m_0-i}{m-i} (-1)^{m-i} \\
&= a_{m_0} - \sum_{i=0}^{m_0-1} \binom{m_0}{i} a_i \sum_{m=0}^{m_0-i-1} \binom{m_0-i}{m} (-1)^m \\
&= a_{m_0} + \sum_{i=0}^{m_0-1} \binom{m_0}{i} (-1)^{m_0-i} a_i \\
&= \sum_{i=0}^{m_0} \binom{m_0}{i} (-1)^{m_0-i} a_i.
\end{aligned}$$

Schließlich gilt für $j \in \{1, \dots, m_0 - 1\}$:

$$\begin{aligned}
b_j^{m_0} &= - \sum_{m=m_0-j}^{m_0-1} \binom{m_0}{m} b_{m-(m_0-j)}^m - \sum_{l=j}^{m_0-1} \sum_{m=l-j}^l (-1)^{m_0-l} \binom{m_0}{m} \binom{m_0-m}{m_0-l} b_{m-(l-j)}^m \\
&= - \sum_{m=m_0-j}^{m_0-1} \sum_{i=0}^{m-(m_0-j)} \binom{m_0}{m} \binom{m}{m-(m_0-j)} \binom{m-(m_0-j)}{i} (-1)^{m-(m_0-j)-i} a_i \\
&\quad - \sum_{l=j}^{m_0-1} \sum_{m=l-j}^l \sum_{i=0}^{m-(l-j)} \binom{m_0}{m} \binom{m_0-m}{m_0-l} \binom{m}{m-(l-j)} \binom{m-(l-j)}{i} (-1)^{j+m-i+m_0} a_i \\
&= - \sum_{i=0}^{j-1} \sum_{m=m_0-(j-i)}^{m_0-1} \binom{m_0}{m} \binom{m}{m-(m_0-j)} \binom{m-(m_0-j)}{i} (-1)^{m-(m_0-j)-i} a_i \\
&\quad - \sum_{l=j}^{m_0-1} \sum_{m=l-j}^l \sum_{i=0}^{m-(l-j)} \binom{m_0}{l} \binom{l}{m} \binom{m}{i} \binom{m-i}{l-j} (-1)^{j+m-i+m_0} a_i \\
&= - \sum_{i=0}^{j-1} \sum_{m=m_0-(j-i)}^{m_0-1} \binom{m_0}{j} \binom{j}{i} \binom{j-i}{m_0-m} (-1)^{m-(m_0-j)-i} a_i \\
&\quad - \sum_{i=0}^j \sum_{m=0}^{j-i} \sum_{l=j}^{m_0-1} \binom{m_0}{i} \binom{m_0-i}{m_0-j+m} \binom{m_0-j+m}{m} \binom{m_0-j}{l-j} (-1)^{j+l-m-i+m_0} a_i
\end{aligned}$$

3 Die Empirische Verteilungsfunktion

$$\begin{aligned}
&= - \sum_{i=0}^{j-1} \binom{m_0}{j} \binom{j}{i} a_i \sum_{m=0}^{j-i-1} \binom{j-i}{m} (-1)^m \\
&\quad - \sum_{i=0}^j \sum_{m=0}^{j-i} (-1)^{m_0-m-i} \binom{m_0}{i} \binom{m_0-i}{m_0-j+m} \binom{m_0-j+m}{m} a_i \sum_{l=0}^{m_0-j-1} \binom{m_0-j}{l} (-1)^l \\
&= \sum_{i=0}^{j-1} \binom{m_0}{j} \binom{j}{i} (-1)^{j-i} a_i \\
&\quad + \sum_{i=0}^j \sum_{m=0}^{j-i} (-1)^{j-m-i} \binom{m_0}{i} \binom{m_0-i}{m_0-j+m} \binom{m_0-j+m}{m} a_i \\
&= \binom{m_0}{j} \sum_{i=0}^{j-1} \binom{j}{i} (-1)^{j-i} a_i + \sum_{i=0}^j \binom{m_0}{j} \binom{j}{i} a_i \sum_{m=0}^{j-i} \binom{j-i}{m} (-1)^{j-i-m} \\
&= \binom{m_0}{j} \sum_{i=0}^j \binom{j}{i} (-1)^{j-i} a_i.
\end{aligned}$$

Also gilt

$$\varphi_{m_0}(t) = \sum_{j=0}^{m_0} \frac{\binom{m_0}{j}}{F(t)^j} \sum_{i=0}^j \binom{j}{i} (-1)^{j-i} a_i, \quad a_i \in \mathbb{R} \text{ für } i \in \{0, \dots, n\}.$$

Sei nun $X = (X_t)_{t \in \mathbb{R}}$ mit $X_t = \varphi(t, F_n(t))$ ein reverses Martingal bzgl. $\mathcal{G} = (\mathcal{G}_t)_{t \in \mathbb{R}}$. Dann gilt

$$X_t = \sum_{m=0}^n \sum_{j=0}^m \frac{\binom{m}{j}}{F(t)^j} \sum_{i=0}^j \binom{j}{i} (-1)^{j-i} a_i 1_{\{F_n(t) = \frac{m}{n}\}}.$$

Desweiteren gilt

$$\begin{aligned}
X_t^{m_0} &= \frac{(n-m_0)! n^{m_0}}{n!} \prod_{j=0}^{m_0-1} \frac{F_n(t) - \frac{j}{n}}{F(t)} \\
&= \sum_{m=m_0}^n \frac{(n-m_0)!}{n!} \prod_{j=0}^{m_0-1} \frac{m-j}{F(t)} 1_{\{F_n(t) = \frac{m}{n}\}}
\end{aligned}$$

3 Die Empirische Verteilungsfunktion

$$\begin{aligned}
 &= \sum_{m=m_0}^n \frac{(n-m_0)!m!}{n!(m-m_0)!} \cdot \frac{1}{F(t)^{m_0}} 1_{\{F_n(t)=\frac{m}{n}\}} \\
 &= \frac{1}{F(t)^{m_0}} \sum_{m=m_0}^n \frac{\binom{n-m_0}{m-m_0}}{\binom{n}{m}} 1_{\{F_n(t)=\frac{m}{n}\}}.
 \end{aligned}$$

Damit folgt

$$\begin{aligned}
 \sum_{m_0=0}^n c_{m_0} X_t^{m_0} &= \sum_{m_0=0}^n c_{m_0} \frac{1}{F(t)^{m_0}} \sum_{m=m_0}^n \frac{\binom{n-m_0}{m-m_0}}{\binom{n}{m}} 1_{\{F_n(t)=\frac{m}{n}\}} \\
 &= \sum_{m=0}^n \sum_{j=0}^m \frac{\binom{n-j}{n-m} c_j}{\binom{n}{m} F(t)^j} 1_{\{F_n(t)=\frac{m}{n}\}}.
 \end{aligned}$$

Weiter gilt

$$\begin{aligned}
 \sum_{j=0}^m \frac{\binom{m}{j}}{F(t)^j} \sum_{i=0}^j \binom{j}{i} (-1)^{j-i} a_i &= \sum_{j=0}^m \frac{\binom{n-j}{n-m} c_j}{\binom{n}{m} F(t)^j} \quad \forall m \in \{0, \dots, n\} \\
 \Leftrightarrow \binom{m}{j} \sum_{i=0}^j \binom{j}{i} (-1)^{j-i} a_i &= \frac{\binom{n-j}{n-m} c_j}{\binom{n}{m}} \quad \forall m \in \{0, \dots, n\}, j \in \{0, \dots, m\} \\
 \Leftrightarrow c_j &= \frac{\binom{n}{m} \binom{m}{j}}{\binom{n-m}{n-j}} \sum_{i=0}^j \binom{j}{i} (-1)^{j-i} a_i \quad \forall m \in \{0, \dots, n\}, j \in \{0, \dots, m\} \\
 \Leftrightarrow c_j &= \frac{\binom{n}{m} \binom{m}{j}}{\binom{n-m}{n-j}} \sum_{i=0}^j \binom{j}{i} (-1)^{j-i} a_i \quad \forall m \in \{0, \dots, n\}, j \in \{0, \dots, m\} \\
 \Leftrightarrow c_j &= \binom{n}{j} \sum_{i=0}^j \binom{j}{i} (-1)^{j-i} a_{n-i} \quad \forall j \in \{0, \dots, n\}.
 \end{aligned}$$

Damit folgt schließlich die Behauptung.

□

Dem Beweis des vorhergehenden Satzes ist wiederum die konkrete Darstellung eines reversen

3 Die Empirische Verteilungsfunktion

\mathcal{G} -Martingals $X = (X_t)_{t \in \mathbb{R}}$ mit $X_t = \varphi(t, F_n(t))$ zu entnehmen.

Korollar 3.28. Sei $X = (X_t)_{t \in \mathbb{R}}$ mit $X_t = \varphi(t, F_n(t))$ ein reverses Martingal bzgl. $\mathcal{G} = (\mathcal{G}_t)_{t \in \mathbb{R}}$. Dann existieren $a_0, \dots, a_n \in \mathbb{R}$ mit

$$a_m = \lim_{u \rightarrow \infty} \varphi\left(u, \frac{m}{n}\right) \quad \text{für } m = 0, \dots, n$$

und es gilt

$$X_t = \sum_{m=0}^n \binom{n}{m} \sum_{i=0}^m \binom{m}{i} (-1)^{m-i} a_{n-i} X_t^m,$$

wobei

$$X_t^m = \frac{(n-m)! n^m}{n!} \prod_{j=0}^{m-1} \frac{F_n(t) - \frac{j}{n}}{F(t)}, \quad m \in \{0, \dots, n\}.$$

3 Die Empirische Verteilungsfunktion

4 Überschreitungswahrscheinlichkeiten

Für die Anwendung in statistischen Tests ist es von Interesse, mit welcher Wahrscheinlichkeit die Empirische Verteilungsfunktion und gewisse Transformationen der Empirischen Verteilungsfunktion bestimmte Schranken überschreiten. Derartige Überschreitungswahrscheinlichkeiten lassen sich für Martingale mittels Stopptechniken berechnen. Die Martingale bzw. reversen Martingale, die in Kapitel 3 hergeleitet wurden, sollen als Grundlage für die Berechnung solcher Überschreitungswahrscheinlichkeiten dienen. Als Ausgangspunkt dieses Kapitels fungiert ein klassisches Resultat einer Überschreitungswahrscheinlichkeit, welches auf Daniels [4] zurückgeht. Hier wird es nun unter Zuhilfenahme von Stopptechniken bewiesen. Die eingesetzten Methoden werden anschließend für die Berechnung weiterer Überschreitungswahrscheinlichkeiten eingesetzt. Dabei werden zunächst Überschreitungswahrscheinlichkeiten gewisser Transformationen der Empirischen Verteilungsfunktion berechnet und davon schließlich Überschreitungswahrscheinlichkeiten der Empirischen Verteilungsfunktion selbst bezüglich verschiedener Schranken abgeleitet.

4.1 Affine Schranken

Das folgende klassische Resultat geht auf Daniels [4] zurück. Hier wird es mittels des reversen Martingals aus Satz 3.22 unter Zuhilfenahme von Stopptechniken bewiesen.

Satz 4.1. *Für die Empirische Verteilungsfunktion $(F_n(t))_{t \in \mathbb{R}}$ der unabhängig und identisch verteilten Zufallsvariablen X_1, \dots, X_n mit stetiger Verteilungsfunktion F gilt für $c > 1$*

$$\mathbb{P}(F_n(t) > cF(t) \quad \text{für ein } t \in \mathbb{R}) = c^{-1}.$$

Beweis. Sei $(\bar{F}_n(t))_{t \in \mathbb{R}}$ die (uniforme) Empirische Verteilungsfunktion der Zufallsvariablen $F(X_1), \dots, F(X_n)$. Dann gilt

$$\{F_n(t) > cF(t) \quad \text{für ein } t \in \mathbb{R}\} = \{\bar{F}_n(t) > ct \quad \text{für ein } t \in [0, 1]\}.$$

Die Ungleichung ist f.s. niemals erfüllt für $t = 0$ und $t = 1$, daher gilt

$$\{\bar{F}_n(t) > ct \quad \text{für ein } t \in [0, 1]\} = \{\bar{F}_n(t) > ct \quad \text{für ein } t \in (0, 1)\}.$$

Da $(F(X_1), \dots, F(X_n))$ absolutstetig verteilt ist auf $(0, 1)$, gilt weiter

$$\{\bar{F}_n(t) > ct \quad \text{für ein } t \in (0, 1)\} \stackrel{\mathcal{L}}{=} \{\bar{F}_n(t) \geq ct \quad \text{für ein } t \in (0, 1)\}$$

Schließlich gilt

$$\{\bar{F}_n(t) \geq ct \quad \text{für ein } t \in (0, 1)\} = \left\{ \sup_{t \in (0, 1)} \frac{\bar{F}_n(t)}{t} \geq c \right\}.$$

4 Überschreitungswahrscheinlichkeiten

Nun ist der Prozess $X = (X_t)_{t \in (0,1)}$ mit $X_t = \frac{\bar{F}_n(t)}{t}$ ein reverses Martingal bzgl. $\mathcal{G} = (\mathcal{G}_t)_{t \in (0,1)}$ mit $\mathcal{G}_t = \sigma(\bar{F}_n(u) : u \geq t)$. Man betrachte die reverse Stoppzeit

$$\tau = \sup \{t \in (0,1) : X_t \geq c\},$$

wobei $\tau = 0$, falls $X_t < c$ für alle $t \in (0,1)$. Hier gilt

$$\left\{ \sup_{t \in (0,1)} \frac{\bar{F}_n(t)}{t} \geq c \right\} = \{\tau > 0\}.$$

Auch der gestoppte Prozess $(X_{t \vee \tau})_{t \in (0,1)}$ ist ein reverses Martingal bzgl. \mathcal{G} und es gilt

$$\mathbb{E}(X_{t \vee \tau}) = \mathbb{E}(X_{1 \vee \tau}) = X_1 = 1.$$

Da $X_{\cdot \vee \tau}$ stückweise stetig ist und lediglich Sprünge nach oben hat, gilt

$$0 \leq X_{t \vee \tau} \leq c$$

und auf $\{\tau > 0\}$

$$X_{t \vee \tau} \xrightarrow{t \rightarrow 0} X_\tau \quad \text{und} \quad X_\tau = c.$$

Unter Verwendung des Konvergenzsatzes von Lebesgue gilt nun

$$\begin{aligned} c\mathbb{P}(\tau > 0) &= \int_{\{\tau > 0\}} c \, d\mathbb{P} \\ &= \int_{\{\tau > 0\}} X_\tau \, d\mathbb{P} \end{aligned}$$

4 Überschreitungswahrscheinlichkeiten

$$\begin{aligned}
 &= \int_{\{\tau > 0\}} \lim_{t \rightarrow 0} X_{t \vee \tau} d\mathbb{P} \\
 &= \lim_{t \rightarrow 0} \int_{\{\tau > 0\}} X_{t \vee \tau} d\mathbb{P} \\
 &= \lim_{t \rightarrow 0} \left(1 - \int_{\{\tau = 0\}} X_{t \vee \tau} d\mathbb{P} \right) \\
 &= 1 - \lim_{t \rightarrow 0} \int_{\{\tau = 0\}} X_{t \vee \tau} d\mathbb{P} \\
 &= 1 - \int_{\{\tau = 0\}} \lim_{t \rightarrow 0} X_{t \vee \tau} d\mathbb{P} \\
 &= 1 - \int_{\{\tau = 0\}} \lim_{t \rightarrow 0} X_t d\mathbb{P}.
 \end{aligned}$$

Nun gilt $F(X_{1:n}) > 0$ f.s. und damit $\lim_{t \rightarrow 0} X_t = 0$ f.s. Daraus folgt also

$$\mathbb{P}(\tau > 0) = c^{-1}.$$

□

Eine ähnliche Aussage lässt sich mittels des Martingals aus Satz 3.8 beweisen.

Satz 4.2. Für die Empirische Verteilungsfunktion $(F_n(t))_{t \in \mathbb{R}}$ der unabhängig und identisch verteilten Zufallsvariablen X_1, \dots, X_n mit stetiger Verteilungsfunktion F gilt für $c > 1$

$$\mathbb{P}(F_n(t) < 1 - c(1 - F(t)) \quad \text{für ein } t \in \mathbb{R}) = c^{-1}.$$

Beweis. Sei $(\bar{F}_n(t))_{t \in \mathbb{R}}$ die (uniforme) Empirische Verteilungsfunktion der Zufallsvariablen

4 Überschreitungswahrscheinlichkeiten

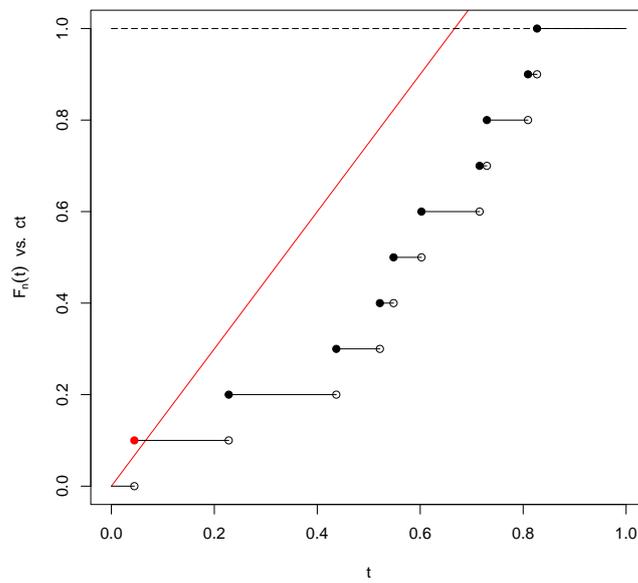


Abbildung 4.1: Die Empirische Verteilungsfunktion F_n der unabhängig und identisch auf $[0, 1]$ gleichverteilten Zufallsvariablen U_1, \dots, U_n , $n = 10$, und die affine Schranke ct mit $c = \frac{3}{2}$, welche mit Wahrscheinlichkeit c^{-1} an mindestens einer Stelle überschritten wird.

4 Überschreitungswahrscheinlichkeiten

$F(X_1), \dots, F(X_n)$. Dann gilt analog zum Beweis von Satz 4.1

$$\{1 - \bar{F}_n(t) \geq c(1-t) \text{ für ein } t \in (0,1)\} = \left\{ \sup_{t \in (0,1)} \frac{1 - \bar{F}_n(t)}{1-t} \geq c \right\}.$$

Der Prozess $X = (X_t)_{t \in (0,1)}$ mit $X_t = \frac{1 - \bar{F}_n(t)}{1-t}$ ist ein Martingal bzgl. $\mathcal{F} = (\mathcal{F}_t)_{t \in (0,1)}$ mit $\mathcal{F}_t = \sigma(\bar{F}_n(s) : s \leq t)$. Man betrachte die Stoppzeit

$$\tau = \inf \{t \in (0,1) : X_t \geq c\},$$

wobei $\tau = 1$, falls $X_t < c$ für alle $t \in (0,1)$. Hier gilt

$$\left\{ \sup_{t \in (0,1)} \frac{1 - \bar{F}_n(t)}{1-t} \geq c \right\} = \{\tau < 1\}.$$

Auch der gestoppte Prozess $(X_{t \wedge \tau})_{t \in (0,1)}$ ist ein Martingal bzgl. \mathcal{F} und es gilt

$$\mathbb{E}(X_{t \wedge \tau}) = \mathbb{E}(X_{0 \wedge \tau}) = X_0 = 1.$$

Da $X_{\cdot \wedge \tau}$ stückweise stetig ist und lediglich Sprünge nach unten hat, gilt

$$0 \leq X_{t \wedge \tau} \leq c$$

und auf $\{\tau < 1\}$

$$X_{t \wedge \tau} \xrightarrow{t \rightarrow 1} X_\tau \quad \text{und} \quad X_\tau = c.$$

4 Überschreitungswahrscheinlichkeiten

Unter Verwendung des Konvergenzsatzes von Lebesgue gilt nun

$$\begin{aligned} c\mathbb{P}(\tau < 1) &= \int_{\{\tau < 1\}} c \, d\mathbb{P} \\ &= \int_{\{\tau < 1\}} X_\tau \, d\mathbb{P} \\ &= \int_{\{\tau < 1\}} \lim_{t \rightarrow 1} X_{t \wedge \tau} \, d\mathbb{P} \\ &= \lim_{t \rightarrow 1} \int_{\{\tau < 1\}} X_{t \wedge \tau} \, d\mathbb{P} \\ &= \lim_{t \rightarrow 1} \left(1 - \int_{\{\tau = 1\}} X_{t \wedge \tau} \, d\mathbb{P} \right) \\ &= 1 - \lim_{t \rightarrow 1} \int_{\{\tau = 1\}} X_{t \wedge \tau} \, d\mathbb{P} \\ &= 1 - \int_{\{\tau = 1\}} \lim_{t \rightarrow 1} X_{t \wedge \tau} \, d\mathbb{P} \\ &= 1 - \int_{\{\tau = 1\}} \lim_{t \rightarrow 1} X_t \, d\mathbb{P}. \end{aligned}$$

Es gilt $F(X_{n:n}) < 1$ f.s. und damit $\lim_{t \rightarrow 1} X_t = 0$ f.s. Daraus folgt also

$$\mathbb{P}(\tau < 1) = c^{-1}.$$

□

4 Überschreitungswahrscheinlichkeiten

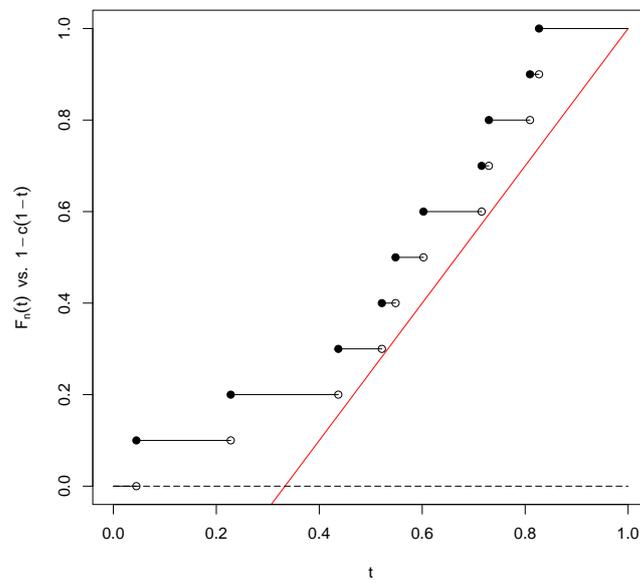


Abbildung 4.2: Die Empirische Verteilungsfunktion F_n der unabhängig und identisch auf $[0, 1]$ gleichverteilten Zufallsvariablen U_1, \dots, U_n , $n = 10$, und die affine Schranke $1 - c(1 - t)$ mit $c = \frac{3}{2}$, welche mit Wahrscheinlichkeit c^{-1} an mindestens einer Stelle unterschritten wird.

4.2 Polynomiale Schranken

Mittels der reversen Martingale aus Satz 3.23 und der Martingale aus Satz 3.9 lassen sich die Ergebnisse aus Satz 4.1 und Satz 4.2 verallgemeinern.

Satz 4.3. Für die Empirische Verteilungsfunktion $(F_n(t))_{t \in \mathbb{R}}$ der unabhängig und identisch verteilten Zufallsvariablen X_1, \dots, X_n mit stetiger Verteilungsfunktion F gilt für $c > 1$ und $m \in \{1, \dots, n\}$

$$\mathbb{P} \left(\frac{(n-m)!n^m}{n!} \prod_{j=0}^{m-1} \left(F_n(t) - \frac{j}{n} \right) > cF(t)^m \quad \text{für ein } t \in \mathbb{R} \right) = c^{-1}.$$

Beweis. Sei $(\bar{F}_n(t))_{t \in \mathbb{R}}$ die (uniforme) Empirische Verteilungsfunktion der Zufallsvariablen $F(X_1), \dots, F(X_n)$. Dann gilt analog zum Beweis von Satz 4.1

$$\left\{ \frac{(n-m)!n^m}{n!} \prod_{j=0}^{m-1} \left(\bar{F}_n(t) - \frac{j}{n} \right) \geq ct^m \quad \text{für ein } t \in (0, 1) \right\} = \left\{ \sup_{t \in (0, 1)} \bar{X}_t^m \geq c \right\},$$

wobei $\bar{X}^m = (\bar{X}_t^m)_{t \in (0, 1)}$ mit $\bar{X}_t^m = \frac{(n-m)!n^m}{n!} \prod_{j=0}^{m-1} \frac{\bar{F}_n(t) - \frac{j}{n}}{t}$ ein reverses Martingal bzgl. $\mathcal{G} = (\mathcal{G}_t)_{t \in (0, 1)}$ mit $\mathcal{G}_t = \sigma(\bar{F}_n(u) : u \geq t)$ ist. Man betrachte die reverse Stoppzeit

$$\tau = \sup \{ t \in (0, 1) : \bar{X}_t^m \geq c \},$$

wobei $\tau = 0$, falls $\bar{X}_t^m < c$ für alle $t \in (0, 1)$. Hier gilt

$$\left\{ \sup_{t \in (0, 1)} \bar{X}_t^m \geq c \right\} = \{ \tau > 0 \}.$$

4 Überschreitungswahrscheinlichkeiten

Auch der gestoppte Prozess $(\bar{X}_{t \vee \tau}^m)_{t \in (0,1)}$ ist ein reverses Martingal bzgl. \mathcal{G} und es gilt

$$\mathbb{E}(\bar{X}_{t \vee \tau}^m) = \mathbb{E}(\bar{X}_{1 \vee \tau}^m) = \bar{X}_1^m = 1.$$

Da $\bar{X}_{\cdot \vee \tau}^m$ stückweise stetig ist und lediglich Sprünge nach oben hat, gilt

$$0 \leq \bar{X}_{t \vee \tau}^m \leq c$$

und auf $\{\tau > 0\}$

$$\bar{X}_{t \vee \tau}^m \xrightarrow{t \rightarrow 0} \bar{X}_\tau^m \quad \text{und} \quad \bar{X}_\tau^m = c.$$

Unter Verwendung des Konvergenzsatzes von Lebesgue gilt nun

$$\begin{aligned} c\mathbb{P}(\tau > 0) &= \int_{\{\tau > 0\}} c \, d\mathbb{P} \\ &= \int_{\{\tau > 0\}} \bar{X}_\tau^m \, d\mathbb{P} \\ &= \int_{\{\tau > 0\}} \lim_{t \rightarrow 0} \bar{X}_{t \vee \tau}^m \, d\mathbb{P} \\ &= \lim_{t \rightarrow 0} \int_{\{\tau > 0\}} \bar{X}_{t \vee \tau}^m \, d\mathbb{P} \\ &= \lim_{t \rightarrow 0} \left(1 - \int_{\{\tau = 0\}} \bar{X}_{t \vee \tau}^m \, d\mathbb{P} \right) \\ &= 1 - \lim_{t \rightarrow 0} \int_{\{\tau = 0\}} \bar{X}_{t \vee \tau}^m \, d\mathbb{P} \end{aligned}$$

4 Überschreitungswahrscheinlichkeiten

$$\begin{aligned}
 &= 1 - \int_{\{\tau=0\}} \lim_{t \rightarrow 0} \bar{X}_{t \vee \tau}^m d\mathbb{P} \\
 &= 1 - \int_{\{\tau=0\}} \lim_{t \rightarrow 0} \bar{X}_t^m d\mathbb{P}.
 \end{aligned}$$

Nun gilt $F(X_{m:n}) > 0$ f.s. und damit $\lim_{t \rightarrow 0} \bar{X}_t^m = 0$ f.s. Daraus folgt also

$$\mathbb{P}(\tau > 0) = c^{-1}.$$

□

Satz 4.4. Für die Empirische Verteilungsfunktion $(F_n(t))_{t \in \mathbb{R}}$ der unabhängig und identisch verteilten Zufallsvariablen X_1, \dots, X_n mit stetiger Verteilungsfunktion F gilt für $c > 1$ und $m \in \{1, \dots, n\}$

$$\mathbb{P} \left(\frac{(n-m)!n^m}{n!} \prod_{j=n-(m-1)}^n \left(\frac{j}{n} - F_n(t) \right) > c(1-F(t))^m \quad \text{für ein } t \in \mathbb{R} \right) = c^{-1}.$$

Beweis. Sei $(\bar{F}_n(t))_{t \in \mathbb{R}}$ die (uniforme) Empirische Verteilungsfunktion der Zufallsvariablen $F(X_1), \dots, F(X_n)$. Dann gilt analog zum Beweis von Satz 4.1

$$\begin{aligned}
 &\left\{ \frac{(n-m)!n^m}{n!} \prod_{j=n-(m-1)}^n \left(\frac{j}{n} - F_n(t) \right) > c(1-F(t))^m \quad \text{für ein } t \in (0,1) \right\} \\
 &= \left\{ \sup_{t \in (0,1)} \bar{X}_t^m \geq c \right\}.
 \end{aligned}$$

wobei $\bar{X}^m = (\bar{X}_t^m)_{t \in (0,1)}$ mit $\bar{X}_t^m = \frac{(n-m)!n^m}{n!} \prod_{j=n-(m-1)}^n \frac{j - \bar{F}_n(t)}{1-t}$ ein Martingal bzgl. $\mathcal{F} =$

4 Überschreitungswahrscheinlichkeiten

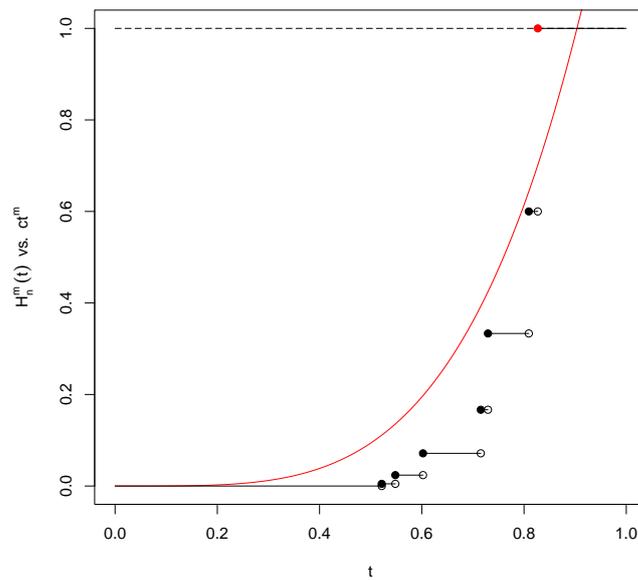


Abbildung 4.3: Der Prozess H_n^m mit $H_n^m(t) = \frac{(n-m)!n^m}{n!} \prod_{j=0}^{m-1} (F_n(t) - \frac{j}{n})$, $n = 10$ und $m = 4$, wobei die unterliegenden Zufallsvariablen U_1, \dots, U_n unabhängig und identisch gleichverteilt sind auf $[0, 1]$ und die Schranke ct^m mit $c = \frac{3}{2}$, welche mit Wahrscheinlichkeit c^{-1} an mindestens einer Stelle überschritten wird.

4 Überschreitungswahrscheinlichkeiten

$(\mathcal{F}_t)_{t \in (0,1)}$ mit $\mathcal{F}_t = \sigma(\bar{F}_n(s) : s \leq t)$ ist. Man betrachte die Stoppzeit

$$\tau = \inf \{t \in (0, 1) : \bar{X}_t^m \geq c\},$$

wobei $\tau = 1$, falls $\bar{X}_t^m < c$ für alle $t \in (0, 1)$. Hier gilt

$$\left\{ \sup_{t \in (0,1)} \bar{X}_t^m \geq c \right\} = \{\tau < 1\}.$$

Auch der gestoppte Prozess $(\bar{X}_{t \wedge \tau}^m)_{t \in (0,1)}$ ist ein Martingal bzgl. \mathcal{F} und es gilt

$$\mathbb{E}(\bar{X}_{t \wedge \tau}^m) = \mathbb{E}(\bar{X}_{0 \wedge \tau}^m) = \bar{X}_0^m = 1.$$

Da $\bar{X}_{\cdot \wedge \tau}^m$ stückweise stetig ist und lediglich Sprünge nach unten hat, gilt

$$0 \leq \bar{X}_{t \wedge \tau}^m \leq c$$

und auf $\{\tau < 1\}$

$$\bar{X}_{t \wedge \tau}^m \xrightarrow{t \rightarrow 1} \bar{X}_\tau^m \quad \text{und} \quad \bar{X}_\tau^m = c.$$

Unter Verwendung des Konvergenzsatzes von Lebesgue gilt nun

$$\begin{aligned} c\mathbb{P}(\tau < 1) &= \int_{\{\tau < 1\}} c \, d\mathbb{P} \\ &= \int_{\{\tau < 1\}} \bar{X}_\tau^m \, d\mathbb{P} \end{aligned}$$

4 Überschreitungswahrscheinlichkeiten

$$\begin{aligned}
&= \int_{\{\tau < 1\}} \lim_{t \rightarrow 1} \bar{X}_{t \wedge \tau}^m d\mathbb{P} \\
&= \lim_{t \rightarrow 1} \int_{\{\tau < 1\}} \bar{X}_{t \wedge \tau}^m d\mathbb{P} \\
&= \lim_{t \rightarrow 1} \left(1 - \int_{\{\tau = 1\}} \bar{X}_{t \wedge \tau}^m d\mathbb{P} \right) \\
&= 1 - \lim_{t \rightarrow 1} \int_{\{\tau = 1\}} \bar{X}_{t \wedge \tau}^m d\mathbb{P} \\
&= 1 - \int_{\{\tau = 1\}} \lim_{t \rightarrow 1} \bar{X}_{t \wedge \tau}^m d\mathbb{P} \\
&= 1 - \int_{\{\tau = 1\}} \lim_{t \rightarrow 1} \bar{X}_t^m d\mathbb{P}.
\end{aligned}$$

Es gilt $F(X_{(n-m+1):n}) < 1$ f.s. und damit $\lim_{t \rightarrow 1} \bar{X}_t^m = 0$ f.s. Daraus folgt also

$$\mathbb{P}(\tau < 1) = c^{-1}.$$

□

4.3 Überschreitungswahrscheinlichkeiten für die Empirische Verteilungsfunktion

Satz 4.1 zeigt die Überschreitungswahrscheinlichkeit der Empirischen Verteilungsfunktion bezüglich einer affinen Schranke auf. Durch Abwandlung der Überschreitungswahrscheinlichkeiten

4 Überschreitungswahrscheinlichkeiten

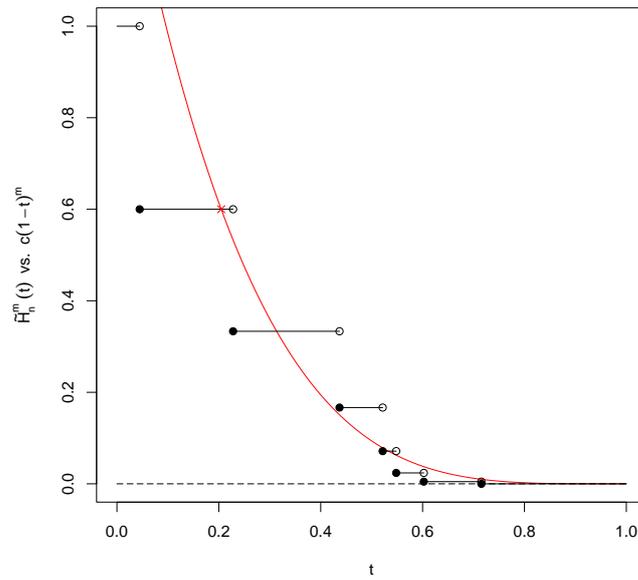


Abbildung 4.4: Der Prozess \tilde{H}_n^m mit $\tilde{H}_n^m(t) = \frac{(n-m)!n^m}{n!} \prod_{j=n-(m-1)}^n \left(\frac{j}{n} - F_n(t)\right)$, $n = 10$ und $m = 4$, wobei die unterliegenden Zufallsvariablen U_1, \dots, U_n unabhängig und identisch gleichverteilt sind auf $[0, 1]$ und die Schranke $c(1-t)^m$ mit $c = \frac{3}{2}$, welche mit Wahrscheinlichkeit c^{-1} an mindestens einer Stelle überschritten wird.

4 Überschreitungswahrscheinlichkeiten

lichkeiten aus Satz 4.3 lassen sich alternative Schranken mit der gleichen Überschreitungswahrscheinlichkeit finden.

Korollar 4.5. *Für die Empirische Verteilungsfunktion $(F_n(t))_{t \in \mathbb{R}}$ der unabhängig und identisch verteilten Zufallsvariablen X_1, \dots, X_n mit stetiger Verteilungsfunktion F gilt für $c > 1$*

$$\mathbb{P} \left(F_n(t) > \frac{1 + \sqrt{4(n-1)ncF(t)^2 + 1}}{2n} \quad \text{für ein } t \in \mathbb{R} \right) = c^{-1}.$$

Beweis. Nach Satz 4.3 gilt

$$\mathbb{P} \left(\frac{n}{n-1} F_n(t) \left(F_n(t) - \frac{1}{n} \right) > cF(t)^2 \quad \text{für ein } t \in \mathbb{R} \right) = c^{-1}.$$

Für $x \in \{0, \frac{1}{n}, \dots, 1\}$ gilt jedoch

$$\frac{n}{n-1} x \left(x - \frac{1}{n} \right) > cF(t)^2 \quad \Leftrightarrow \quad x > \frac{1 + \sqrt{4(n-1)ncF(t)^2 + 1}}{2n}.$$

□

Korollar 4.6. *Für die Empirische Verteilungsfunktion $(F_n(t))_{t \in \mathbb{R}}$ der unabhängig und identisch verteilten Zufallsvariablen X_1, \dots, X_n mit stetiger Verteilungsfunktion F gilt für $c > 1$*

$$\mathbb{P}(F_n(t) > f_0(F(t)) \quad \text{für ein } t \in \mathbb{R}) = c^{-1},$$

wobei

$$f_0(t) = f_1(t) \cdot 1_{\left[0, \sqrt[3]{\frac{2(n-3)!}{3\sqrt{3n!c}}}\right]}(t) + f_2(t) \cdot 1_{\left(\sqrt[3]{\frac{2(n-3)!}{3\sqrt{3n!c}}}, 1\right]}(t)$$

4 Überschreitungswahrscheinlichkeiten

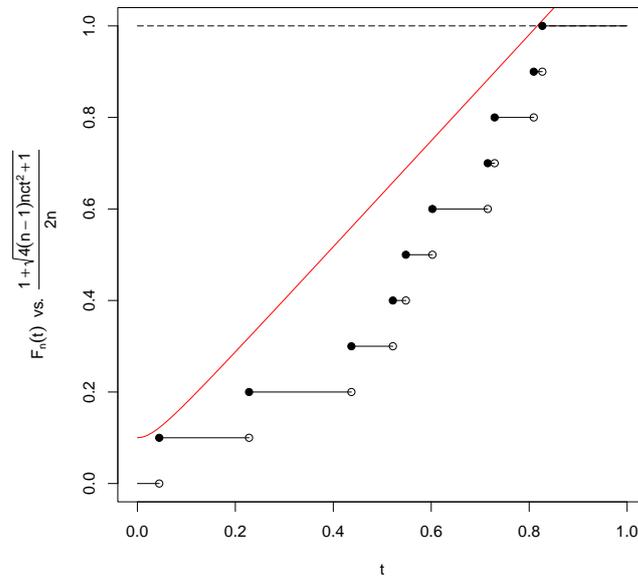


Abbildung 4.5: Die Empirische Verteilungsfunktion F_n der unabhängig und identisch auf $[0, 1]$ gleichverteilten Zufallsvariablen U_1, \dots, U_n , $n = 10$, und die Schranke $\frac{1 + \sqrt{4(n-1)nc^2 + 1}}{2n}$ mit $c = \frac{3}{2}$, welche mit Wahrscheinlichkeit c^{-1} an mindestens einer Stelle überschritten wird.

4 Überschreitungswahrscheinlichkeiten

mit

$$f_1(t) = \frac{1}{n} \left(1 + \frac{2}{\sqrt{3}} \cos \left(\frac{\arccos \left(\frac{3\sqrt{3}n!ct^3}{2(n-3)!} \right)}{3} \right) \right)$$

und

$$f_2(t) = \frac{1}{n} + \frac{1}{n} \sqrt[3]{\frac{3\sqrt{3} \frac{n!}{(n-3)!} ct^3 + \sqrt{27 \left(\frac{n!}{(n-3)!} \right)^2 c^2 t^6 - 4}}{6\sqrt{3}}} \\ + \frac{1}{n} \sqrt[3]{\frac{3\sqrt{3} \frac{n!}{(n-3)!} ct^3 - \sqrt{27 \left(\frac{n!}{(n-3)!} \right)^2 c^2 t^6 - 4}}{6\sqrt{3}}}.$$

Beweis. Nach Satz 4.3 gilt

$$\mathbb{P} \left(\frac{n^2}{(n-2)(n-1)} F_n(t) \left(F_n(t) - \frac{1}{n} \right) \left(F_n(t) - \frac{2}{n} \right) > cF(t)^3 \quad \text{für ein } t \in \mathbb{R} \right) = c^{-1}.$$

Für festes $t \in \mathbb{R}$ betrachte die Funktion $g : \mathbb{R} \rightarrow \mathbb{R}$ mit

$$g(x) = x \left(x - \frac{1}{n} \right) \left(x - \frac{2}{n} \right) - \frac{(n-2)(n-1)cF(t)^3}{n^2}.$$

Hierbei handelt es sich um ein Polynom dritten Grades, welches mindestens eine und maximal drei reelle Nullstellen besitzt. Für die ersten beiden Ableitungen gilt

$$g'(x) = 3 \left(x - \frac{1}{n} \right)^2 - \frac{1}{n^2}$$

bzw.

$$g''(x) = 6x - \frac{6}{n}.$$

4 Überschreitungswahrscheinlichkeiten

Nun gilt also

$$g'(x) = 0 \Leftrightarrow x = \frac{1}{n} \left(1 - \frac{1}{\sqrt{3}}\right) \vee x = \frac{1}{n} \left(1 + \frac{1}{\sqrt{3}}\right)$$

und

$$g''\left(\frac{1}{n} \left(1 - \frac{1}{\sqrt{3}}\right)\right) < 0 \quad \text{bzw.} \quad g''\left(\frac{1}{n} \left(1 + \frac{1}{\sqrt{3}}\right)\right) > 0.$$

Also hat g ein lokales Maximum in $\frac{1}{n} \left(1 - \frac{1}{\sqrt{3}}\right)$, wobei

$$g\left(\frac{1}{n} \left(1 - \frac{1}{\sqrt{3}}\right)\right) = \frac{2\sqrt{3}}{9n^3} - \frac{(n-2)(n-1)cF(t)^3}{n^2}.$$

Damit hat g drei Nullstellen, falls

$$F(t) < \sqrt[3]{\frac{2(n-3)!}{3\sqrt{3}n!c}},$$

genau zwei Nullstellen, falls

$$F(t) = \sqrt[3]{\frac{2(n-3)!}{3\sqrt{3}n!c}}$$

und genau eine Nullstelle, falls

$$F(t) > \sqrt[3]{\frac{2(n-3)!}{3\sqrt{3}n!c}}.$$

Man betrachte zunächst den Fall $F(t) = 0$. Hierbei hat g offensichtlich die drei Nullstellen

4 Überschreitungswahrscheinlichkeiten

$0, \frac{1}{n}$ und $\frac{2}{n}$. Damit gilt

$$g(x) > 0 \Leftrightarrow x \in \left(0, \frac{1}{n}\right) \cup \left(\frac{2}{n}, \infty\right).$$

Für $x \in \{0, \frac{1}{n}, \dots, 1\}$ gilt hier also

$$g(x) > 0 \Leftrightarrow x > \frac{2}{n}.$$

Falls $F(t) \in \left(0, \sqrt[3]{\frac{2(n-3)!}{3\sqrt{3n!c}}}\right)$ liegen die beiden kleineren Nullstellen im Intervall $(0, \frac{1}{n})$. Bezeichne x_0 die größte Nullstelle, dann gilt also für $x \in \{0, \frac{1}{n}, \dots, 1\}$

$$g(x) > 0 \Leftrightarrow x > x_0.$$

Um für den Fall $F(t) \in \left(0, \sqrt[3]{\frac{2(n-3)!}{3\sqrt{3n!c}}}\right)$ die Nullstellen zu berechnen, betrachte g in der Form

$$g(x) = x^3 - \frac{3}{n}x^2 + \frac{2}{n^2}x - \frac{(n-2)(n-1)cF(t)^3}{n^2}.$$

Um den quadratischen Term zu eliminieren, betrachte die Funktion $g_1 : \mathbb{R} \rightarrow \mathbb{R}$ mit $g_1(x) = g\left(x + \frac{1}{n}\right)$. Damit gilt nun

$$g_1(x) = x^3 - \frac{1}{n^2}x - \frac{(n-2)(n-1)cF(t)^3}{n^2}.$$

Nun betrachte $g_2 : [0, \infty) \times [0, 2\pi) \rightarrow \mathbb{R}$ mit $g_2(r, \alpha) = g_1(r \cos(\alpha))$. Mit Hilfe des Additi-

4 Überschreitungswahrscheinlichkeiten

onstheorems lässt sich zeigen, dass

$$\cos(3\alpha) = 4 \cos^3(\alpha) - 3 \cos(\alpha).$$

Damit folgt

$$g_2(r, \alpha) = \frac{r^3 \cos(3\alpha)}{4} + \left(\frac{3r^2}{4} - \frac{1}{n^2} \right) r \cos(\alpha) - \frac{(n-2)(n-1) cF(t)^3}{n^2}$$

und für $r = \frac{2}{\sqrt{3n}}$

$$g_2\left(\frac{2}{\sqrt{3n}}, \alpha\right) = \frac{2 \cos(3\alpha)}{3\sqrt{3}n^3} - \frac{(n-2)(n-1) cF(t)^3}{n^2}.$$

Es gilt

$$g_2\left(\frac{2}{\sqrt{3n}}, \alpha\right) = 0 \Leftrightarrow \cos(\alpha) \in \left\{ \cos\left(\frac{\arccos\left(\frac{3\sqrt{3}n!cF(t)^3}{2(n-3)!}\right)}{3} + \frac{2k}{3}\pi\right) : k \in \{0, 1, 2\} \right\},$$

wobei

$$\begin{aligned} \cos\left(\frac{\arccos\left(\frac{3\sqrt{3}n!cF(t)^3}{2(n-3)!}\right)}{3}\right) &\in \left(\frac{\sqrt{3}}{2}, 1\right), \\ \cos\left(\frac{\arccos\left(\frac{3\sqrt{3}n!cF(t)^3}{2(n-3)!}\right)}{3} + \frac{2}{3}\pi\right) &\in \left(-\frac{\sqrt{3}}{2}, -\frac{1}{2}\right) \end{aligned}$$

und

$$\cos\left(\frac{\arccos\left(\frac{3\sqrt{3}n!cF(t)^3}{2(n-3)!}\right)}{3} + \frac{4}{3}\pi\right) \in \left(-\frac{1}{2}, 0\right).$$

4 Überschreitungswahrscheinlichkeiten

Daher gilt

$$x_0 = \frac{1}{n} \left(1 + \frac{2}{\sqrt{3}} \cos \left(\frac{\arccos \left(\frac{3\sqrt{3n!cF(t)^3}}{2(n-3)!} \right)}{3} \right) \right).$$

Im Fall $F(t) = \sqrt[3]{\frac{2(n-3)!}{3\sqrt{3n!c}}}$ fallen die beiden kleineren Nullstellen zusammen und für die größere Nullstelle gilt

$$x_0 = \frac{1}{n} \left(1 + \frac{2}{\sqrt{3}} \right).$$

Falls $F(t) \in \left(\sqrt[3]{\frac{2(n-3)!}{3\sqrt{3n!c}}}, 1 \right]$, hat g genau eine Nullstelle. Man betrachte wiederum g_1 und setze $g_3 : \mathbb{R}^2 \rightarrow \mathbb{R}$ mit $g_3(u, v) = g_1(u + v)$. Dann gilt

$$g_3(u, v) = \left(3uv - \frac{1}{n^2} \right) (u + v) + u^3 + v^3 - \frac{(n-2)(n-1)cF(t)^3}{n^2}.$$

Falls

$$3uv - \frac{1}{n^2} = 0 \quad \text{und} \quad u^3 + v^3 - \frac{(n-2)(n-1)cF(t)^3}{n^2} = 0,$$

gilt also $g_3(u, v) = 0$. Nun gilt

$$3uv - \frac{1}{n^2} = 0 \quad \Leftrightarrow \quad v^3 = \frac{1}{27n^6u^3}$$

und damit

$$\begin{aligned} u^3 + \frac{1}{27n^6u^3} - \frac{(n-2)(n-1)cF(t)^3}{n^2} = 0 &\Leftrightarrow u^6 - \frac{(n-2)(n-1)cF(t)^3}{n^2}u^3 + \frac{1}{27n^6} = 0 \\ \Leftrightarrow \left(u^3 - \frac{(n-2)(n-1)cF(t)^3}{2n^2} \right)^2 &= \frac{27(n-2)^2(n-1)^2n^2c^2F(t)^6 - 4}{108n^6} \end{aligned}$$

4 Überschreitungswahrscheinlichkeiten

$$\Leftrightarrow u^3 \in \left\{ \frac{(n-2)(n-1)cF(t)^3}{2n^2} + s \frac{\sqrt{27 \left(\frac{n!}{(n-3)!}\right)^2 c^2 F(t)^6 - 4}}{6\sqrt{3}n^3} : s \in \{-1, 1\} \right\}.$$

Wähle

$$u = \sqrt[3]{\frac{3\sqrt{3}n!cF(t)^3 + \sqrt{27(n!)^2 c^2 F(t)^6 - 4((n-3)!)^2}}{6\sqrt{3}n^3(n-3)!}},$$

dann gilt

$$v = \sqrt[3]{\frac{3\sqrt{3}n!cF(t)^3 - \sqrt{27(n!)^2 c^2 F(t)^6 - 4((n-3)!)^2}}{6\sqrt{3}n^3(n-3)!}}$$

und somit ist

$$x_0 = \frac{1}{n} + \frac{1}{n} \sqrt[3]{\frac{3\sqrt{3}n!cF(t)^3 + \sqrt{27(n!)^2 c^2 F(t)^6 - 4((n-3)!)^2}}{6\sqrt{3}(n-3)!}} \\ + \frac{1}{n} \sqrt[3]{\frac{3\sqrt{3}n!cF(t)^3 - \sqrt{27(n!)^2 c^2 F(t)^6 - 4((n-3)!)^2}}{6\sqrt{3}(n-3)!}}$$

die Nullstelle von g .

□

Korollar 4.7. Für die Empirische Verteilungsfunktion $(F_n(t))_{t \in \mathbb{R}}$ der unabhängig und identisch verteilten Zufallsvariablen X_1, \dots, X_n mit stetiger Verteilungsfunktion F gilt für $c > 1$

$$\mathbb{P} \left(F_n(t) > \frac{3 + \sqrt{5 + 4\sqrt{\frac{n!}{(n-4)!} c F(t)^4 + 1}}}{2n} \text{ für ein } t \in \mathbb{R} \right) = c^{-1}.$$

4 Überschreitungswahrscheinlichkeiten

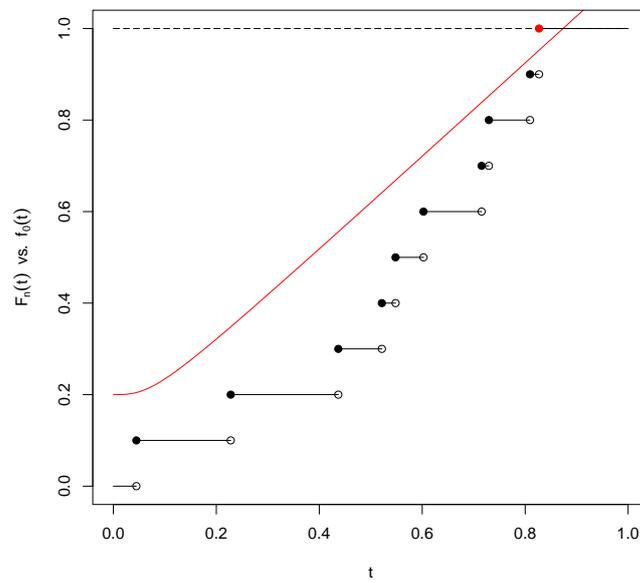


Abbildung 4.6: Die Empirische Verteilungsfunktion F_n der unabhängig und identisch auf $[0, 1]$ gleichverteilten Zufallsvariablen U_1, \dots, U_n , $n = 10$, und die Schranke $f_0(t)$ aus Korollar 4.6 mit $c = \frac{3}{2}$, welche mit Wahrscheinlichkeit c^{-1} an mindestens einer Stelle überschritten wird.

4 Überschreitungswahrscheinlichkeiten

Beweis. Nach Satz 4.3 gilt

$$\begin{aligned} \mathbb{P} \left(\frac{(n-4)!n^4}{n!} F_n(t) \left(F_n(t) - \frac{1}{n} \right) \left(F_n(t) - \frac{2}{n} \right) \left(F_n(t) - \frac{3}{n} \right) > cF(t)^4 \quad \text{für ein } t \in \mathbb{R} \right) \\ = c^{-1}. \end{aligned}$$

Man betrachte die Funktion

$$g(x) = x \left(x - \frac{1}{n} \right) \left(x - \frac{2}{n} \right) \left(x - \frac{3}{n} \right) - \frac{n!cF(t)^4}{(n-4)!n^4}.$$

Hierbei handelt es sich um ein Polynom vierten Grades, welches maximal vier reelle Nullstellen besitzt. Für die ersten beiden Ableitungen gilt

$$g'(x) = 4x^3 - \frac{18}{n}x^2 + \frac{22}{n^2}x - \frac{6}{n^3}$$

bzw.

$$g''(x) = 12x^2 - \frac{36}{n}x + \frac{22}{n^2}.$$

Um die Nullstellen der ersten Ableitung zu erhalten betrachte $h_1 : \mathbb{R} \rightarrow \mathbb{R}$ mit $h_1(x) = \frac{g'(x + \frac{3}{2n})}{4}$. Hierbei gilt

$$h_1(x) = x^3 + \frac{49}{4n^2}x$$

und es ergibt sich

$$g'(x) = 4 \left(x - \frac{3 - \sqrt{5}}{2n} \right) \left(x - \frac{3}{2n} \right) \left(x - \frac{3 + \sqrt{5}}{2n} \right).$$

4 Überschreitungswahrscheinlichkeiten

Damit gilt also

$$g'(x) = 0 \Leftrightarrow x = \frac{3 - \sqrt{5}}{2n} \quad \vee \quad x = \frac{3}{2n} \quad \vee \quad x = \frac{3 + \sqrt{5}}{2n}$$

und

$$g''\left(\frac{3 - \sqrt{5}}{2n}\right) = \frac{10}{n^2} > 0, \quad g''\left(\frac{3}{2n}\right) = -\frac{5}{n^2} < 0 \quad \text{bzw.} \quad g''\left(\frac{3 + \sqrt{5}}{2n}\right) = \frac{10}{n^2} > 0.$$

Also hat g ein lokales Minimum in $\frac{3 - \sqrt{5}}{2n}$, wobei

$$g\left(\frac{3 - \sqrt{5}}{2n}\right) = -\frac{n!cF(t)^4 + (n-4)!}{(n-4)!n^4} < 0,$$

ein lokales Maximum in $\frac{3}{2n}$, wobei

$$g\left(\frac{3}{2n}\right) = \frac{9}{16n^4} - \frac{n!cF(t)^4}{(n-4)!n^4}$$

und ein lokales Minimum in $\frac{3 + \sqrt{5}}{2n}$, wobei

$$g\left(\frac{3 + \sqrt{5}}{2n}\right) = -\frac{n!cF(t)^4 + (n-4)!}{(n-4)!n^4} < 0.$$

Damit hat g vier Nullstellen, falls

$$F(t) < \sqrt[4]{\frac{9(n-4)!}{16n!c}},$$

4 Überschreitungswahrscheinlichkeiten

genau drei Nullstellen, falls

$$F(t) = \sqrt[4]{\frac{9(n-4)!}{16n!c}}$$

und genau zwei Nullstellen, falls

$$F(t) > \sqrt[4]{\frac{9(n-4)!}{16n!c}}.$$

Man betrachte zunächst den Fall $F(t) = 0$. Hierbei hat g offensichtlich die vier Nullstellen $0, \frac{1}{n}, \frac{2}{n}$ und $\frac{3}{n}$. Damit gilt

$$g(x) > 0 \quad \Leftrightarrow \quad x \in \left(\frac{1}{n}, \frac{2}{n}\right) \cup \left(\frac{3}{n}, \infty\right).$$

Für $x \in \{0, \frac{1}{n}, \dots, 1\}$ gilt hier also

$$g(x) > 0 \quad \Leftrightarrow \quad x > \frac{3}{n}.$$

Falls $F(t) \in \left(0, \sqrt[4]{\frac{9(n-4)!}{16n!c}}\right)$, ist die kleinste Nullstelle negativ und die beiden mittleren Nullstellen liegen im Intervall $\left(\frac{1}{n}, \frac{2}{n}\right)$. Bezeichne x_0 die größte Nullstelle, dann gilt also für $x \in \{0, \frac{1}{n}, \dots, 1\}$

$$g(x) > 0 \quad \Leftrightarrow \quad x > x_0.$$

Im Fall $F(t) = \sqrt[4]{\frac{9(n-4)!}{16n!c}}$ fallen die beiden mittleren Nullstellen zusammen und im Fall $F(t) \in \left(\sqrt[4]{\frac{9(n-4)!}{16n!c}}, 1\right]$ gibt es nur die negative kleinere Nullstelle und die größere Nullstelle

4 Überschreitungswahrscheinlichkeiten

x_0 . Es gilt

$$\begin{aligned}
 g(x) &= x \left(x - \frac{1}{n}\right) \left(x - \frac{2}{n}\right) \left(x - \frac{3}{n}\right) - \frac{n!cF(t)^4}{(n-4)!n^4} \\
 &= x^4 - \frac{6}{n}x^3 + \frac{11}{n^2}x^2 - \frac{6}{n^3}x - \frac{n!cF(t)^4}{(n-4)!n^4} \\
 &= \left(x^2 - \frac{3}{n}x + \frac{1}{n^2}\right)^2 - \frac{1}{n^4} - \frac{n!cF(t)^4}{(n-4)!n^4} \\
 &= \left(\left(x - \frac{3}{2n}\right)^2 - \frac{5}{4n^2}\right)^2 - \frac{n!cF(t)^4 + (n-4)!}{(n-4)!n^4}.
 \end{aligned}$$

Damit ergibt sich

$$x_0 = \frac{3 + \sqrt{5 + 4\sqrt{\frac{n!}{(n-4)!}cF(t)^4 + 1}}}{2n}.$$

□

Bemerkung 4.8. Da es im Allgemeinen nicht möglich ist, die Wurzeln von Polynomen fünften Grades oder höher durch Radikale darzustellen, können die Ereignisse aus Satz 4.3 im Fall $m \geq 5$ nicht nach F_n umgestellt werden. Wie in den vorherigen Korollaren ist aber auch in diesen Fällen die größte Nullstelle ausschlaggebend. Mit numerischen Methoden können damit die entsprechenden Kurven angenähert werden.

Bemerkung 4.9. Für die Schranke cF aus dem klassischen Resultat von Daniels [4] gilt

$$cF(t) \geq 1 \quad \forall t \in [F^{-1}(c^{-1}), \infty).$$

Da $F_n(t) \leq 1$ für alle $t \in \mathbb{R}$ gilt, kann die Empirische Verteilungsfunktion die Schranke für „große“ t nicht überschreiten. Der Fokus liegt also mehr auf den kleineren Zeitpunkten. Für

4 Überschreitungswahrscheinlichkeiten

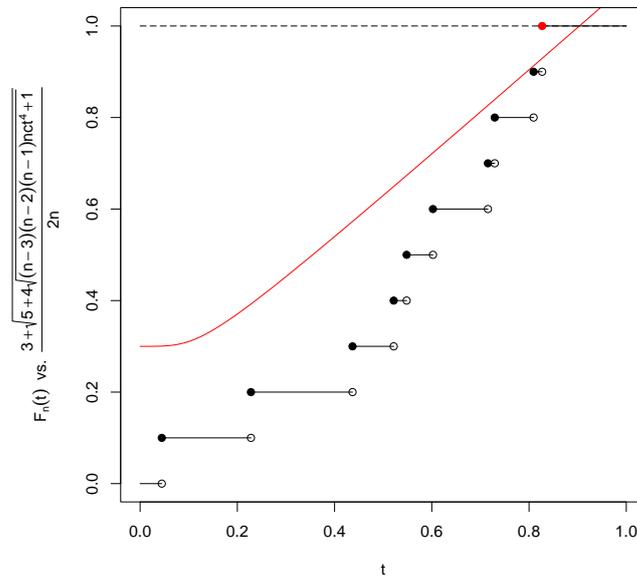


Abbildung 4.7: Die Empirische Verteilungsfunktion F_n der unabhängig und identisch auf $[0, 1]$ gleichverteilten Zufallsvariablen U_1, \dots, U_n , $n = 10$, und die Schran-

ke $\frac{3 + \sqrt{5 + 4\sqrt{\frac{n!}{(n-4)!} ct^4 + 1}}}{2n}$ mit $c = \frac{3}{2}$, welche mit Wahrscheinlichkeit c^{-1} an mindestens einer Stelle überschritten wird.

4 Überschreitungswahrscheinlichkeiten

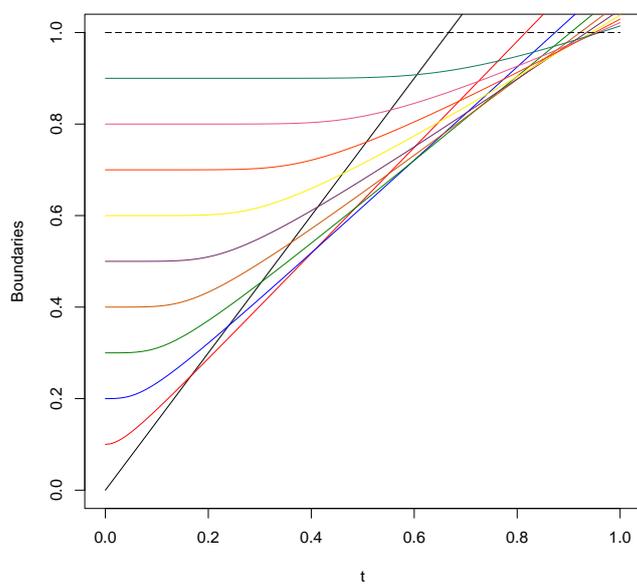


Abbildung 4.8: Zehn verschiedene Schranken, die von der Empirischen Verteilungsfunktion, welcher unabhängig und identisch auf $[0, 1]$ gleichverteilte Zufallsvariablen U_1, \dots, U_{10} zugrundeliegen, jeweils mit der Wahrscheinlichkeit c^{-1} , $c = \frac{3}{2}$, an mindestens einer Stelle überschritten werden.

4 Überschreitungswahrscheinlichkeiten

die Schranke aus Korollar 4.5 gilt

$$\frac{1 + \sqrt{4(n-1)ncF(t)^2 + 1}}{2n} \geq 1 \quad \forall t \in \left[F^{-1}\left(c^{-\frac{1}{2}}\right), \infty \right).$$

Hier ist es also möglich, dass F_n die Schranke nach dem Zeitpunkt $F^{-1}(c^{-1})$ überschreitet, jedoch nicht nach dem Zeitpunkt $F^{-1}\left(c^{-\frac{1}{2}}\right)$. Da die Schranke allerdings im Wert $\frac{1}{n}$ startet, kann sie erst mit dem zweiten Sprung der Empirischen Verteilungsfunktion überschritten werden. Abbildung 4.8 macht deutlich, dass sich diese „Verschiebung des Überschreitungszeitpunktes“ für wachsendes m fortsetzt, wobei $m \leq n$ den Index des der Schranke zugrundeliegenden reversen Martingals aus Satz 4.3 bezeichnet. Die jeweilige Schranke startet in $\frac{m-1}{n}$ und erreicht den Wert 1 zum Zeitpunkt $F^{-1}\left(c^{-\frac{1}{m}}\right)$.

Dieser ungleiche Einfluss der Zeitpunkte auf die Überschreitungswahrscheinlichkeit findet sich beispielsweise auch im Zusammenhang mit der Dvoretzky–Kiefer–Wolfowitz-Ungleichung (siehe Dvoretzky, Kiefer und Wolfowitz [8] sowie Massart [13]) wieder. Dort wird für $\epsilon > 0$ das Ereignis

$$\left\{ \sup_{t \in \mathbb{R}} |F_n(t) - F(t)| > \epsilon \right\}$$

betrachtet. Eine Überschreitung ist hier in der Umgebung des Medians der Verteilung wahrscheinlicher als in den Tails.

4 Überschreitungswahrscheinlichkeiten

5 Goodness-of-Fit-Tests für Typ-II-zensierte Daten

Die Martingale aus Satz 3.9 und die reversen Martingale aus Satz 3.23 haben die spezielle Eigenschaft, dass ihre Realisierung nur von der Realisierung der kleinsten bzw. größten Werte der unterliegenden Stichprobe abhängt. Daher eignen sich diese Prozesse für Goodness-of-Fit-Tests im Fall Typ-II-zensierter Daten, wo eben gerade nur die ersten (Typ-II-Rechtszensierung) bzw. letzten (Typ-II-Linkszensierung) Ordnungsstatistiken beobachtet werden können. Die genannten Tests gehören zu der Klasse der nichtparametrischen Tests für zensierte Daten, welche beispielsweise in Bagdonavicius, Kruopis und Nikulin [2] behandelt werden.

5.1 Goodness-of-Fit-Tests für Typ-II-rechtszensierte Daten

Man betrachte ein Szenario, in dem die Verteilung unabhängiger und identisch verteilter Typ-II-rechtszensierter Daten X_1, \dots, X_n untersucht werden soll, wobei $X_1 \sim F$ mit unbe-

5 Goodness-of-Fit-Tests für Typ-II-zensierte Daten

kannter stetiger Verteilungsfunktion F . Dabei seien nur die k kleinsten Daten beobachtbar. Für eine feste Verteilungsfunktion F_0 betrachte man die Hypothese

$$H_0 : F \geq F_0.$$

Aufgrund der Typ-II-Rechtszensierung wird nun eine Teststatistik benötigt, die nur von den ersten k Ordnungsstatistiken abhängt. Das Martingal X^{n-k+1} aus Satz 3.9 erfüllt diese Voraussetzung und es gilt:

Korollar 5.1. *Für die Empirische Verteilungsfunktion $(F_n(t))_{t \in \mathbb{R}}$ der unabhängig und identisch verteilten Zufallsvariablen X_1, \dots, X_n mit stetiger Verteilungsfunktion F und für $\alpha \in (0, 1)$ und $k \in \{1, \dots, n\}$ gilt unter der Hypothese H_0*

$$\mathbb{P} \left(\frac{(k-1)!n^{n-k}}{(n-1)!} \prod_{j=k}^n \left(\frac{j}{n} - F_n(t) \right) > \frac{(1-F_0(t))^{n-k+1}}{\alpha} \quad \text{für ein } t \in \mathbb{R} \right) \leq \alpha.$$

Bemerkung 5.2. Kann die Menge der in Frage kommenden Verteilungsfunktionen derart eingeschränkt werden, dass die Alternative

$$H_1 : F < F_0$$

lautet, so gilt unter H_1

$$\mathbb{P} \left(\frac{(k-1)!n^{n-k}}{(n-1)!} \prod_{j=k}^n \left(\frac{j}{n} - F_n(t) \right) > \frac{(1-F_0(t))^{n-k+1}}{\alpha} \quad \text{für ein } t \in \mathbb{R} \right) \geq \alpha.$$

5.2 Goodness-of-Fit-Tests für Typ-II-linksensierte Daten

Nun betrachte man ein Szenario, in dem die Verteilung unabhängiger und identisch verteilter Typ-II-linksensierter Daten X_1, \dots, X_n untersucht werden soll, wobei $X_1 \sim F$ mit unbekannter stetiger Verteilungsfunktion F . Hierbei seien nur die k größten Daten beobachtbar. Für eine feste Verteilungsfunktion F_0 betrachte man die Hypothese

$$H_0 : F \leq F_0.$$

Aufgrund der Typ-II-Linksensierung wird nun eine Teststatistik benötigt, die nur von den letzten k Ordnungsstatistiken abhängt. Das reverse Martingal X^{n-k+1} aus Satz 3.23 erfüllt diese Voraussetzung und es gilt:

Korollar 5.3. *Für die Empirische Verteilungsfunktion $(F_n(t))_{t \in \mathbb{R}}$ der unabhängig und identisch verteilten Zufallsvariablen X_1, \dots, X_n mit stetiger Verteilungsfunktion F und für $\alpha \in (0, 1)$ und $k \in \{1, \dots, n\}$ gilt unter der Hypothese H_0*

$$\mathbb{P} \left(\frac{(k-1)!n^{n-k}}{(n-1)!} \prod_{j=0}^{n-k} \left(F_n(t) - \frac{j}{n} \right) > \frac{F_0(t)^{n-k+1}}{\alpha} \quad \text{für ein } t \in \mathbb{R} \right) \leq \alpha.$$

Bemerkung 5.4. Kann die Menge der in Frage kommenden Verteilungsfunktionen derart eingeschränkt werden, dass die Alternative

$$H_1 : F > F_0$$

5 Goodness-of-Fit-Tests für Typ-II-zensierte Daten

lautet, so gilt unter H_1

$$\mathbb{P} \left(\frac{(k-1)!n^{n-k}}{(n-1)!} \prod_{j=0}^{n-k} \left(F_n(t) - \frac{j}{n} \right) > \frac{F_0(t)^{n-k+1}}{\alpha} \quad \text{für ein } t \in \mathbb{R} \right) \geq \alpha.$$

Auch der Kolmogorov-Smirnov-Test (siehe Kolmogorov [11] und Smirnov [18]) basiert auf dem Vergleich einer festen Verteilungsfunktion F_0 mit der wahren Verteilungsfunktion F und liefert einen nichtparametrischen Test (für unzensierte Daten). Während in der Kolmogorov-Smirnov-Statistik allerdings die Differenz zwischen F_n und F als Vergleich dient, spielen bei den Tests in diesem und dem vorangehenden Abschnitt Quotienten die entscheidende Rolle.

Literaturverzeichnis

- [1] P.K. Andersen *et al.*: *Statistical Models Based on Counting Processes*. Springer New York, 1993.
- [2] V. Bagdonavicus, J. Kruopis und M. Nikulin: *Nonparametric Tests for Censored Data*. ISTE. Wiley, 2011.
- [3] F.P. Cantelli: *Sulla Determinazione Empirica di una Legge di Distribuzione*. Giorn. Ist. Ital. Attuari, 4:421–424, 1933.
- [4] H.E. Daniels: *The Statistical Theory of the Strength of Bundles of Threads. I*. Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, 183(995):405–435, 1945.
- [5] J. Doob: *Stochastic Processes*. Wiley Publications in Statistics. John Wiley & Sons, 1953.
- [6] J.L. Doob: *Notes on Martingale Theory*. In: *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 2: Contributions to Probability Theory*, S. 95–102, Berkeley, Calif., 1961. University of California Press.

Literaturverzeichnis

- [7] N.R. Draper und H. Smith: *Applied Regression Analysis*. Wiley Series in Probability and Mathematical Statistics. Wiley, New York, 2. Aufl., 1981.
- [8] A. Dvoretzky, J. Kiefer und J. Wolfowitz: *Asymptotic Minimax Character of the Sample Distribution Function and of the Classical Multinomial Estimator*. Ann. Math. Statist., 27(3):642–669, Sep. 1956.
- [9] V.I. Glivenko: *Sulla Determinazione Empirica di una Legge di Distribuzione*. Giorn. Ist. Ital. Attuari, 4:92–99, 1933.
- [10] K. Itô: *Stochastic integral*. Proc. Imp. Acad., 20(8):519–524, 1944.
- [11] A.N. Kolmogorov: *Sulla Determinazione Empirica di una Legge di Distribuzione*. Giorn. Ist. Ital. Attuari, 4:83–91, 1933.
- [12] H.L. Koul und W. Stute: *Nonparametric Model Checks for Time Series*. Ann. Statist., 27(1):204–236, März 1999.
- [13] P. Massart: *The Tight Constant in the Dvoretzky-Kiefer-Wolfowitz Inequality*. Ann. Probab., 18(3):1269–1283, Juli 1990.
- [14] D. Nualart und W. Schoutens: *Chaotic and Predictable Representations for Lévy Processes*. Stochastic Processes and their Applications, 90(1):109–122, 2000.
- [15] A. Rényi: *On a Group of Problems in the Theory of Ordered Samples*. Selected Transl. Math. Statist. Prob, 13:289–297, 1973.
- [16] H. Robbins: *Statistical Methods Related to the Law of the Iterated Logarithm*. Ann. Math. Statist., 41(5):1397–1409, Okt. 1970.

Literaturverzeichnis

- [17] G. Shorack und J. Wellner: *Empirical Processes with Applications to Statistics*. Classics in Applied Mathematics. Society for Industrial and Applied Mathematics (SIAM, 3600 Market Street, Floor 6, Philadelphia, PA 19104), 2009.
- [18] N. Smirnov: *Approximate Distribution Laws for Random Variables, Constructed from Empirical Data*. Uspekhi Fiziologicheskikh Nauk, 10:179–206, 1944. (In Russian).
- [19] A. Wald: *Sequential Analysis*. John Wiley and Sons, 1. Aufl., 1947.

Eidesstattliche Erklärung

Ich erkläre: Ich habe die vorgelegte Dissertation selbständig und ohne unerlaubte fremde Hilfe und nur mit den Hilfen angefertigt, die ich in der Dissertation angegeben habe.

Alle Textstellen, die wörtlich oder sinngemäß aus veröffentlichten Schriften entnommen sind, und alle Angaben, die auf mündlichen Auskünften beruhen, sind als solche kenntlich gemacht.

Bei den von mir durchgeführten und in der Dissertation erwähnten Untersuchungen habe ich die Grundsätze guter wissenschaftlicher Praxis, wie sie in der „Satzung der Justus-Liebig-Universität Gießen zur Sicherung guter wissenschaftlicher Praxis“ niedergelegt sind, eingehalten.

Gießen, den 18. November 2015

David Heß