

TECHNISCHE UNIVERSITÄT DRESDEN
INSTITUT FÜR MATHEMATISCHE STOCHASTIK

DRESDNER SCHRIFTEN ZUR
MATHEMATISCHEN STOCHASTIK

7/2009

Dietmar Ferger

Stochastische Prozesse mit Strukturbrüchen

Herausgeber: Die Professoren des Instituts für Mathematische Stochastik

ISSN 0946–4735

Vorlesungsskript

Stochastische Prozesse mit Strukturbrüchen

Dietmar Ferger
TU Dresden

13. Juli 2009

Inhaltsverzeichnis

1	Einführung	1
1.1	Hazard-Funktion mit Sprung	1
1.2	Regression mit Sprung	6
2	Der Skorokhod-Raum D	10
3	Verteilungskonvergenz in D	13
4	Stetigkeitssätze für das argmax-Funktional	21
5	Erste Anwendung: Regression mit Sprung	34
5.1	Starke Konsistenz des MLE	36
5.2	Fehlerabschätzung	39
5.3	Verteilungskonvergenz	44
5.4	Der Fall a und b unbekannt	50
5.5	Der Schätzer von Dempfle und Stute	59
6	Mehrdimensionale Grenzwertsätze	65
7	Asymptotik des Dempfle-Stute Schätzers in allgemeinen Modellen	82
8	Fehlerschranken für den Dempfle-Stute Schätzer	87
8.1	Verteilungskonvergenz im glatten Fall	108
8.2	Grenzwertsätze im nicht-glatten Fall	118
	Literaturverzeichnis	121

Kapitel 1

Einführung

Zur Motivation zwei Beispiele:

1.1 Hazard-Funktion mit Sprung

Sei $X \geq 0$ eine Zufallsvariable (ZV) mit Verteilungsfunktion (VF) F und Dichte f .

Beispiel.

- X = Zeitpunkt des Todes eines Patienten (gemessen ab OP)
- X = Zeitpunkt des Wiederauftretens von Metastasen (gemessen ab Chemotherapie)
- X = Ausfallzeitpunkt einer technischen Einheit

Sprechweise: X = Lebensdauer eines Teilchens

Definition und Satz 1.1. $T := \{x \geq 0 : F(x) < 1\}$. Dann ist T ein Intervall vom Typ $T = [0, r)$ für ein $r \leq \infty$. Die Funktion $h : T \rightarrow [0, \infty]$ mit:

$$h(t) := \frac{f(t)}{1 - F(t)}, \quad t \in T$$

heißt HAZARD-FUNKTION (*hazard rate, failure rate, mortality rate, force of mortality*).

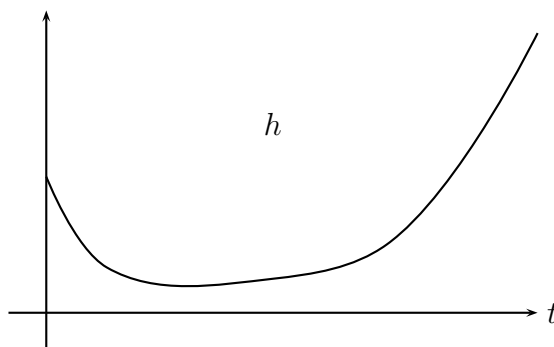
Zur Bedeutung von h betrachte die Wahrscheinlichkeit, dass das Teilchen nach Erreichen der Lebensdauer t innerhalb einer nachfolgenden Zeitspanne Δ stirbt:

$$\begin{aligned} P(X \leq t + \Delta \mid X > t) &= \frac{F(t + \Delta) - F(t)}{1 - F(t)} \\ &= \frac{f(z) \cdot \Delta}{1 - F(t)}, \quad \text{für ein } z \in (t, t + \Delta) \\ &= \frac{f(t)}{1 - F(t)} \cdot \Delta + \frac{f(z) - f(t)}{1 - F(t)} \cdot \Delta. \end{aligned}$$

Falls f stetig in $t \in T$, so folgt:

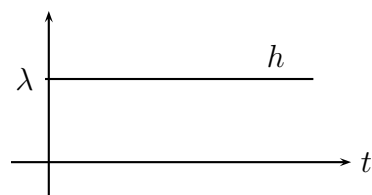
$$P(X \leq t + \Delta \mid X > t) = h(t) \cdot \Delta + o(\Delta), \quad \Delta \rightarrow 0.$$

Beispiel. (Sterblichkeitsrate beim Menschen)



Beispiel.

$$\begin{aligned} f(x) &= \lambda e^{-\lambda x} \Rightarrow 1 - F(x) = e^{-\lambda x} \Rightarrow \\ h(t) &= \lambda \quad \forall t \in T = [0, \infty) \end{aligned}$$



„Nichtalterungseigenschaft“

F bzw. f legen h (bis auf Nullmenge) eindeutig fest. Die Umkehrung gilt auch.

Dazu:

$$\begin{aligned} \frac{d}{dt} \log(1 - F(t)) &= - \frac{f(t)}{1 - F(t)} = -h(t) \Rightarrow \\ \int_0^x h(t) dt &= -\log(1 - F(t)) \Big|_0^x \\ &= -\log(1 - F(x)) \quad \text{wegen } F(0) = 0 \\ &< \infty \quad \forall x \in T \end{aligned}$$

$F(x) = 1 - \exp\left\{-\int_0^x h(t) dt\right\}, \quad x \in T \quad (1.1)$ $f(x) = h(x) \exp\left\{-\int_0^x h(t) dt\right\}, \quad x \in T \quad (1.2)$
--

Ferner wegen $F(x) \rightarrow 1$, $x \rightarrow r$:

$$\int_0^r h(t) dt = \infty.$$

Somit charakterisiert jedes $h : T \rightarrow [0, \infty)$ mit:

$$\int_0^x h(t) dt < \infty \quad \forall x \in T = [0, r)$$

und

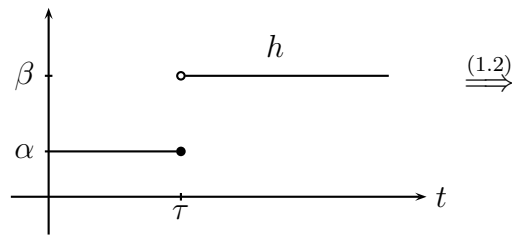
$$\int_0^r h(t) dt = \infty$$

eindeutig eine VF F (einer ZV $X \geq 0$) mittels der Beziehung (1.1) bzw. (1.2).

Beispiel 1.2. $h : [0, \infty) \rightarrow [0, \infty)$

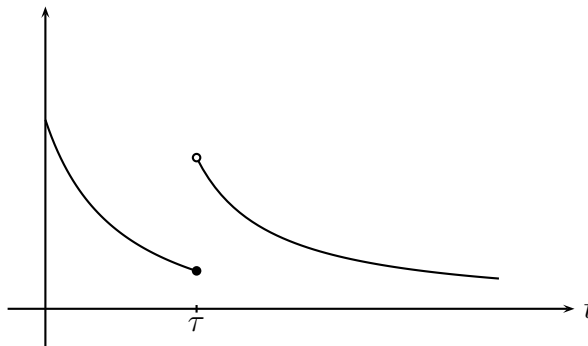
$$h(t) = \begin{cases} \alpha, & 0 \leq t \leq \tau \\ \beta, & t > \tau \end{cases}$$

τ , α und β positiv, $\alpha \neq \beta$.



$$f(t) = f_\tau(t) = \begin{cases} \alpha e^{-\alpha t}, & 0 \leq t \leq \tau \\ \gamma e^{-\beta t}, & t > \tau \end{cases} \quad (1.3)$$

mit $\gamma = \beta e^{(\beta-\alpha)\tau}$



τ = Zeitpunkt einer abrupten Veränderung von Umwelteinflüssen des Teilchens (Schock)

oder auch

τ = kritischer Zeitpunkt nach Transplantation

Schätzen der Sprungstelle

$X_1, \dots, X_n$ i.i.d. mit Dichte $f = f_\tau$ aus (1.3), wobei $\alpha \neq \beta$ bekannt $\Rightarrow$

$$\begin{aligned} \text{Likelihood-Funktion} &= L_n(\tau, \underline{X}_n) = \prod_{i=1}^n f_\tau(X_i) = \\ &= \prod_{\substack{i=1 \\ X_i \leq \tau}}^n f_\tau(X_i) \cdot \prod_{\substack{i=1 \\ X_i > \tau}}^n f_\tau(X_i) \quad \Rightarrow \end{aligned}$$

Log-Likelihood-Funktion $= l_n(\tau, \underline{X}_n) =$

$$\begin{aligned}
 &= \underbrace{\sum_{X_i \leq \tau} (\log \alpha - \alpha X_i)}_{\sum_{i=1}^n \dots - \sum_{X_i > \tau} \dots} + \sum_{X_i > \tau} (\log \beta + (\beta - \alpha)\tau - \beta X_i) = \\
 &= \sum_{X_i > \tau} \left[\log \frac{\beta}{\alpha} + (\beta - \alpha)\tau - (\beta - \alpha)X_i \right] + \sum_{i=1}^n (\log \alpha - \alpha X_i) = \\
 &= \sum_{i=1}^n 1_{\{X_i > \tau\}} \left[\log \frac{\beta}{\alpha} + (\beta - \alpha)(\tau - X_i) \right] + \underbrace{\sum_{i=1}^n (\log \alpha - \alpha X_i)}_{\text{hängt nicht von } \tau \text{ ab}} \quad , \quad \tau > 0 \quad \Rightarrow
 \end{aligned}$$

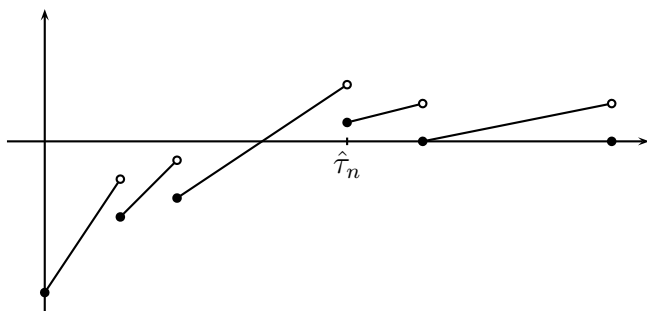
Maximum-Likelihood-Estimator (MLE) für $\hat{\tau}_n$ ist gegeben durch

$$\hat{\tau}_n = \operatorname{argmax}_{t \geq 0} S_n(t),$$

wobei

$$S_n(t) = \frac{1}{n} \sum_{i=1}^n 1_{\{X_i > t\}} \left[\log \frac{\beta}{\alpha} + (\beta - \alpha)(t - X_i) \right].$$

Beachte, die Trajektorien von S_n sind rcll (right-continuous with left-hand limits)



$$\alpha < \beta$$

Man kann zeigen (Details später):

1. Numerische Bestimmung von $\hat{\tau}_n$ unproblematisch
- 2.

$$\sup_{t \geq 0} |S_n(t) - S(t)| \xrightarrow{(n \rightarrow \infty)} 0 \quad \text{fast sicher (f.s.) ,}$$

wobei $S : [0, \infty) \rightarrow \mathbb{R}$ stetig mit

$$\tau = \operatorname{argmax}_{t \geq 0} S(t) \text{ als } \mathbf{eindeutige} \text{ Maximalstelle.}$$

Problem: Gilt Implikation

$$S_n \rightarrow S \text{ gleichmäßig auf } \mathbb{R}_+ \text{ f.s.} \Rightarrow \operatorname{argmax}_{t \geq 0} S_n(t) \rightarrow \operatorname{argmax}_{t \geq 0} S(t) \text{ f.s.}?$$

Lieferte:

$$\hat{\tau}_n \rightarrow \tau \text{ f.s. (starke Konsistenz).}$$

3. Sei Z der lokalisierte Prozess zu S_n , d.h. gemäß Definition

$$Z_n(t) := n \left\{ S_n\left(\tau + \frac{t}{n}\right) - S_n(\tau) \right\}, \quad t \in \mathbb{R}.$$

Es gilt:

$$n(\hat{\tau}_n - \tau) = \operatorname{argmax}_{t \in \mathbb{R}} Z_n(t).$$

(Z_n) kann aufgefasst werden als Folge in einem Funktionenraum $D(\mathbb{R})$, und es gilt:

$$(*) \ Z_n \xrightarrow{\mathcal{L}} Z \text{ in } D(\mathbb{R})$$

Problem: Was genau bedeutet die Verteilungskonvergenz in $(*)$? Und folgt aus $(*)$

$$\operatorname{argmax}_{t \in \mathbb{R}} Z_n(t) \xrightarrow{\mathcal{L}} \operatorname{argmax}_{t \in \mathbb{R}} Z(t) ?$$

Lieferte:

$$n(\hat{\tau}_n - \tau) \xrightarrow{\mathcal{L}} \operatorname{argmax}_{t \in \mathbb{R}} Z(t)$$

.

Ziel der Vorlesung: Möglichst allgemeine Antworten auf Fragen in 2. und 3.

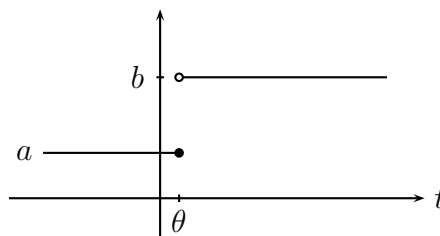
1.2 Regression mit Sprung

(X_i, Y_i) i.i.d. $\sim (X, Y)$ mit

(1) $X \sim F$ unbekannte Verteilungsfunktion

(2) $P(Y \in \cdot | X = x) = \operatorname{Bin}(1, m(x))$

$$(3) \quad m(x) = \begin{cases} a, & x \leq \theta \\ b, & x > \theta \end{cases}$$



θ = Parameter von Interesse, $a \neq b \in (0, 1)$ Störparameter.

Interpretation:

- X = zufälliger Input
- Y = zufälliger Output

$\Rightarrow \theta$ = Schwellenwert (kritischer Wert)

Beispiel. (Herzrhythmusstörung)

X = Variabilität des Herzschlags

$$Y = \begin{cases} 1, & \text{falls plötzlicher Herztod} \\ 0, & \text{sonst} \end{cases}$$

Beispiel. (Atemwegserkrankungen bei Kindern)

X = NO_2 -Emission

$$Y = \begin{cases} 1, & \text{falls Kind mit Bronchitis} \\ 0, & \text{falls Kind ohne Bronchitis} \end{cases}$$

Beachte:

$$\begin{aligned} P(Y \in A) &= E(E(1_{\{Y \in A\}} | X)) \\ &= \int E(1_{\{Y \in A\}} | X = x) F(dx) \\ &= \int P(Y \in A | X = x) F(dx) \\ &\stackrel{(2)}{=} \int \text{Bin}(1, m(x))(A) F(dx) \\ &= \int \{m(x)\delta_1(A) + (1 - m(x))\delta_0(A)\} F(dx) \\ &= \int m dF \cdot \delta_1(A) + \left(1 - \int m dF\right) \cdot \delta_0(A) \\ &= \text{Bin}(1, \int m dF)(A) \Rightarrow \end{aligned}$$

$$(4) \quad Y \sim \text{Bin}(1, \int m \, dF)$$

Ferner:

$$(5) \quad E(Y|X = x) = E(1_{\{Y=1\}}|X = x) = P(Y = 1|X = x) \stackrel{(2)}{=} m(x).$$

D.h. m ist die Faktorisierung der bedingten Erwartung: $m(X) = E(Y|X)$. Setze $\epsilon = Y - m(X) \Rightarrow$

$$(6) \quad Y = m(X) + \epsilon \text{ mit } E(\epsilon|X) = 0.$$

Insbesondere ist ϵ zentriert wegen $E(\epsilon) = E(E(\epsilon|X)) = 0$.

Gleichung (6) heißt REGRESSIONSMODELL. Somit liefern (1)-(3) ein (spezielles) Regressionsmodell mit sogenannter REGRESSIONSFUNKTION $m =$ Treppenfunktion mit Sprung. Zum Schätzen der Sprungstelle θ folgende Idee: Betrachte zu (X_1, Y_1) und (X_2, Y_2) den Erwartungswert

$$\begin{aligned} \rho(t) &:= E(1_{\{X_1 \leq t < X_2\}}(Y_1 - Y_2)) \\ &= E(1_{\{X_1 \leq t\}} 1_{\{X_2 > t\}} Y_1) - E(1_{\{X_1 \leq t\}} 1_{\{X_2 > t\}} Y_2) \\ &= P(X_2 > t) \cdot E(1_{\{X_1 \leq t\}} Y_1) - P(X_1 \leq t) \cdot E(1_{\{X_2 > t\}} Y_2) \quad \text{wg. Unabh.} \\ &\stackrel{\text{Bedingen bzgl } X_1 \text{ bzw. } X_2 \text{ sowie (5)}}{=} (1 - F(t)) \int_{-\infty}^t m(x) F(dx) - F(t) \int_t^{\infty} m(x) F(dx) \\ &= \int_{-\infty}^t m(x) F(dx) - F(t) \overline{m}, \\ &\text{wobei } \overline{m} := \int_{-\infty}^{\infty} m(x) F(dx). \end{aligned}$$

Eine Fallunterscheidung $t \leq \theta$ und $t > \theta$ liefert:

$$(7) \quad \rho(t) = \begin{cases} (a - b)(1 - F(\theta))F(t), & t \leq \theta \\ (a - b)F(\theta)(1 - F(t)), & t > \theta \end{cases}$$

Als Folgerung: Ist F in einer Umgebung von θ streng monoton und $F(\theta) \in (0, 1)$, so ist

$$\theta = \operatorname{argmax}_{t \in \mathbb{R}} \rho(t)$$

eindeutige Maximalstelle von F , falls $a > b$. Ein erwartungstreuer Schätzer für $\rho(t)$ ist gegeben durch die **V-Statistik**

$$\rho_n(t) = \frac{1}{n(n-1)} \sum_{i=1}^n \sum_{\substack{j=1 \\ i \neq j}}^n 1_{\{X_i \leq t < X_j\}} (Y_i - Y_j).$$

Wir erhalten den Schätzer von Dempfle und Stute [DS02]:

$$\hat{\theta}_n := \operatorname{argmax}_{t \in \mathbb{R}} \rho_n(t).$$

Ziele:

1. Zeige starke Konsistenz

$$\hat{\theta}_n \rightarrow \theta \quad \text{f.s. } \forall \theta \in \mathbb{R}$$

mittels f.s. gleichmäßiger Konvergenz von $\rho_n \rightarrow \rho$.

2. Nachweis der Verteilungskonvergenz über lokalisierten Prozess Y_n zu ρ_n :

$$Y_n(t) := n \left\{ \rho_n\left(\theta + \frac{t}{n}\right) - \rho_n(\theta) \right\}, \quad t \in \mathbb{R} \Rightarrow$$

$$n(\hat{\theta}_n - \theta) = \operatorname{argmax}_{t \in \mathbb{R}} Y_n(t).$$

Fazit: Wir haben es zu tun mit ZV $\tau_n = \operatorname{argmax}_{t \in T} M_n(t)$, wobei $M_n = \{M_n(t) : t \in T\}$, $n \in \mathbb{N}$, Folge von stochastischen Prozessen mit Parametermenge $T \subseteq \mathbb{R}$. Problem: Wann zieht eine bestimmte Konvergenz der stochastischen Prozesse die der Maximalstellen nach sich:

$$M_n \rightarrow M \text{ f.s. } \stackrel{?}{\Rightarrow} \operatorname{argmax}(M_n) \rightarrow \operatorname{argmax}(M) \text{ f.s.}$$

bzw.

$$(*) \quad M_n \xrightarrow{\mathcal{L}} M \stackrel{?}{\Rightarrow} \operatorname{argmax}(M_n) \xrightarrow{\mathcal{L}} \operatorname{argmax}(M)$$

In $(*)$ stellt sich zusätzlich die Frage nach dem metrischen Raum S , in dem die Verteilungskonvergenz stattfindet. Beachte, unsere $M_n \in \{S_n, Z_n, \rho_n, Y_n\}$ sind **nicht** stetig, da Sprünge auftauchen. Damit ist der Funktionenraum $C(T)$ zu klein. Als in vielfacher Hinsicht geeignet erweist sich der Skorokhod-Raum.

Kapitel 2

Der Skorokhod-Raum D

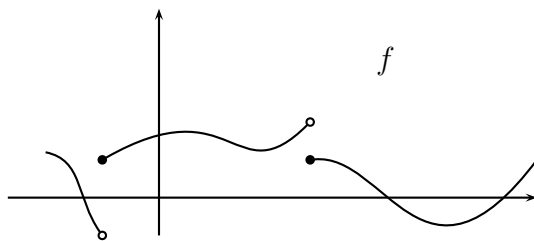
Sei $I := [a, b]$, $a < b$, und

$$D := D[a, b] := \{f : I \rightarrow \mathbb{R} \text{ mit } f(t+) \text{ existiert und} \\ f(t+) = f(t) \ \forall t \in [a, b) \text{ sowie} \\ f(t-) \text{ existiert } \forall t \in (a, b]\}$$

kurz: f rcll (right-continuous with left-hand limits).

Lemma 2.1. $f \in D \Rightarrow$

- (1) f beschränkt.
- (2) f hat höchstens abzählbar viele Sprungstellen.



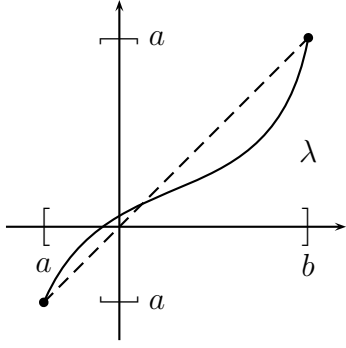
Sprungstellen können sich durchaus häufen!

Wegen Lemma 2.1(1) ist sup-Metrik

$$d(f, g) = \sup_{t \in I} |f(t) - g(t)| =: \|f - g\|, \quad f, g \in D$$

definiert, aber ungeeignet, weil z.B. (D, d) nicht separabel und weil Borel- σ -Algebra $\mathcal{B}_d(D)$ zu groß ($\rightarrow$ Messbarkeitsproblem). Es gibt eine schwächere (= bessere) Metrik. Dazu definiere:

$$H := \{\lambda : I \rightarrow I \mid \lambda(a) = a, \lambda(b) = b, \lambda \text{ stetig + streng } \nearrow\}$$



$$\lambda \in H \Rightarrow \lambda^{-1} \in H$$

$$Id \in H \text{ mit: } Id(x) := x, \forall x \in I$$

$$s(f, g) := \inf_{\lambda \in H} \{ \|f - g \circ \lambda\| + \|\lambda - Id\| \}$$

Man überprüft leicht:

Lemma 2.2. $s : D \times D \rightarrow \mathbb{R}$ ist Metrik auf D (so genannte SKOROKHOD-METRIK).

Eigenschaften von s in

Lemma 2.3. (1) $f_n \rightarrow_d f \Rightarrow f_n \rightarrow_s f$. Umkehrung gilt im Allgemeinen nicht.
Aber:

$$(2) f_n \rightarrow_s f, f \in C(I) \Rightarrow f_n \rightarrow_d f.$$

(3) (eingeschränkte Stetigkeit der Addition)

$$f_n \rightarrow_s f, g_n \rightarrow_s g \in C(I) \Rightarrow f_n + g_n \rightarrow_s f + g.$$

Falls $g \notin C(I)$, muss die Implikation nicht gelten. (D kein topologischer Vektorraum)

$$(4) f_n \rightarrow_s f \Leftrightarrow \exists (\lambda_n) \subseteq H \text{ mit } \|f_n - f \circ \lambda_n\| \rightarrow 0 \text{ und } \|\lambda_n - Id\| \rightarrow 0.$$

(Folgt aus Definition von s und der des Infimums).

Sei

$$\mathcal{B}_s(D) := \sigma(\{G \subseteq D : G \text{ s-offen}\}) = \text{Borel-}\sigma\text{-Algebra bzgl. } s$$

und

$$\pi_t : D \rightarrow \mathbb{R}, f \mapsto \pi_t(f) := f(t), \text{ die Projektion bzgl. } t.$$

Satz 2.4. $\mathcal{B}_s(D) = \sigma(\pi_t : t \in I)$.

Folgerung 2.5. Sei $X : (\Omega, \mathcal{A}) \rightarrow (D, \mathcal{B}_s(D))$ eine Abbildung. Dann:

$$X \text{ } \mathcal{A} - \mathcal{B}_s(D) \text{ messbar} \Leftrightarrow X(t) \text{ } \mathcal{A} - \mathcal{B}(\mathbb{R}) \text{ messbar } \forall t \in I.$$

Beweis: „ $\Rightarrow$ “ $X(t) = \pi_t \circ X =$ Komposition von messbaren Abbildungen.

„ $\Leftarrow$ “ Satz 2.4 und 1.2.11 in Gänsler und Stute [GS77].

□

Sei $T = \{t_1, \dots, t_k\} \subseteq I$. Setze

$$\pi_T(f) := (f(t_1), \dots, f(t_k)) = (f(t))_{t \in T} \Rightarrow \pi_T : D \rightarrow \mathbb{R}^{|T|}.$$

Diese Projektionen erzeugen ebenfalls die Borel- σ -Algebra, d.h.

$$\mathcal{B}_s(D) = \sigma(\pi_T : T \subseteq I, \text{ endlich}) .$$

Mit dem Maßeindeutigkeitssatz folgt:

Satz 2.6. (1) Q und $\tilde{Q}$ W -Maße auf $\mathcal{B}_s(D)$ Dann:

$$Q = \tilde{Q} \Leftrightarrow Q \circ \pi_T^{-1} = \tilde{Q} \circ \pi_T^{-1} \quad \forall T \subseteq I \text{ endlich} .$$

(2) X, Y ZE in (D, s) . Dann:

$$X \stackrel{\mathcal{L}}{=} Y \Leftrightarrow \pi_T(X) \stackrel{\mathcal{L}}{=} \pi_T(Y) \quad \forall T \subseteq I \text{ endlich} .$$

Gemäß Satz 2.6(2) ist also die Verteilung eines ZEs in (D, s) eindeutig durch seine fidis festgelegt (fidis = finite dimensional distribution).

Kapitel 3

Verteilungskonvergenz in D

Ziel: Bereitstellung von Kriterien für Verteilungskonvergenz in D . Erinnere an den Stetigkeitsmodul

$$\omega(f, \delta) = \sup_{\substack{s, t \in I \\ |s-t| \leq \delta}} |f(s) - f(t)|, \quad f \in D.$$

Das Gegenstück für den Raum D ist

$$\omega'(f, \delta) = \inf_{t_i, \delta} \max_{1 \leq i \leq r} \inf_{t_{i-1} \leq s, t < t_i} |f(s) - f(t)|,$$

wobei das Infimum gebildet wird über alle Zerlegungen $a =: t_0 < t_1 < \dots < t_r = b$ mit $t_i - t_{i-1} > \delta \ \forall \ 1 \leq i \leq r$, $r \in \mathbb{N}$.

Satz 3.1. Seien $X, X_n, n \geq 1$, ZE in (D, s) . Falls

$$\pi_T(X_n) \xrightarrow{\mathcal{L}} \pi_T(X) \quad \forall T \subseteq I \text{ endlich} \quad (3.1)$$

und

$$\lim_{\delta \rightarrow 0} \limsup_{n \rightarrow \infty} P(\omega'(X_n, \delta) > \epsilon) = 0 \quad \forall \epsilon > 0, \quad (3.2)$$

so gilt:

$$X_n \xrightarrow{\mathcal{L}} X \text{ in } (D, s).$$

Die obige Aussage lässt sich erweitern, falls ω' durch ω ersetzt wird.

Satz 3.2. Gilt in Satz 3.1 anstelle von (3.2) die Bedingung

$$\lim_{\delta \rightarrow 0} \limsup_{n \rightarrow \infty} P(\omega(X_n, \delta) > \epsilon) = 0 \quad \forall \epsilon > 0, \quad (3.3)$$

so folgt

$$X_n \xrightarrow{\mathcal{L}} X \text{ in } (D, s) \text{ **und** } X \in C(I) \text{ f.s.}$$

Der Nachweis von (3.2) bzw. (3.3) ist im konkreten Fall mitunter schwierig. Es gibt ein handlicheres Kriterium.

Satz 3.3. Seien $X, X_n, n \geq 1$, ZE in (D, s) mit (3.1) und

$$P(X(b-) \neq X(b)) = 0 .$$

Falls eine stetige, monoton wachsende Funktion $F : I \rightarrow \mathbb{R}$ und $\alpha \geq 0, \beta > 1$ existieren mit

$$E|X_n(t_1) - X_n(t)|^\alpha |X_n(t) - X_n(t_2)|^\alpha \leq (F(t_2) - F(t_1))^\beta \quad (3.4)$$

$$\forall a \leq t_1 \leq t \leq t_2 \leq b, \forall n \geq 1,$$

so:

$$X_n \xrightarrow{\mathcal{L}} X \text{ in } (D, s) .$$

Als Beispiel betrachte den so genannten **uniformen empirischen Prozess**

$\alpha_n = \{\alpha_n(t) : t \in [0, 1]\}, n \in \mathbb{N}$, mit

$$\alpha_n(t) := \frac{1}{\sqrt{n}} \sum_{i=1}^n (1_{\{U_i \leq t\}} - t),$$

wobei $U_1, \dots, U_n$ i.i.d. $\sim \mathcal{U}(0, 1)$. Es gilt:

Satz 3.4. (Invarianzprinzip von Donsker [Don52])

$$\alpha_n \xrightarrow{\mathcal{L}} B_0 \text{ in } D[0, 1] ,$$

wobei B_0 eine **Brownsche Brücke** bezeichnet.

(Beachte aus der Definition $B_0(t) := B(t) - tB(1)$, $t \in [0, 1]$, wobei B eine Brownsche Bewegung (BB), folgt, dass B_0 ein zentrierter Gauß-Prozess mit Kovarianzfunktion: $\text{Cov}(B_0(s), B_0(t)) = s \wedge t - st$.)

Beweis: 1.) Konvergenz der fidis: $t_1 < \dots < t_k \Rightarrow$

$$(\alpha_n(t_1), \dots, \alpha_n(t_k))' = \frac{1}{\sqrt{n}} \sum_{i=1}^n \xi_i$$

mit

$$\xi_i = (1_{\{U_i \leq t_1\}} - t_1, \dots, 1_{\{U_i \leq t_k\}} - t_k)' \text{ i.i.d.}$$

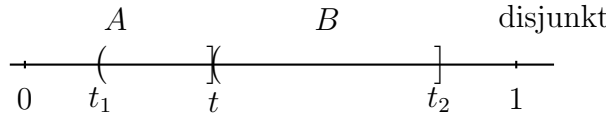
zentrierte ZV in $\mathbb{R}^k$ mit $\Gamma := \text{Cov}(\xi_1) = (t_r \wedge t_s - t_r t_s)_{1 \leq r, s \leq k} \xrightarrow{ZGWS}$

$$(\alpha_n(t_1), \dots, \alpha_n(t_k))' \xrightarrow{\mathcal{L}} N_k(0, \Gamma).$$

Da $(B_0(t_1), \dots, B_0(t_k)) \stackrel{\mathcal{L}}{=} N_k(0, \Gamma)$, folgt (3.1). Wegen B_0 stetig ist somit der 1. Teil von Satz 3.3 erfüllt.

2.) Seien $t_1 < t < t_2 \Rightarrow M_n(t_1, t, t_2) := E|\alpha_n(t_1) - \alpha_n(t)|^2 |\alpha_n(t_2) - \alpha_n(t)|^2 =$

$$\begin{aligned}
 &= E \left(\frac{1}{\sqrt{n}} \sum_{i=1}^n \underbrace{1_{(t_1, t]}(U_i)}_{=:A} - \underbrace{(t - t_1)}_{=\lambda(A)} \right)^2 \left(\frac{1}{\sqrt{n}} \sum_{i=1}^n \underbrace{1_{(t, t_2]}(U_i)}_{=:B} - \underbrace{(t_2 - t)}_{=\lambda(B)} \right)^2 \\
 &= n^{-2} \sum_{1 \leq i, j, k, l \leq n} E \left[\left(1_A(U_i) - \lambda(A) \right) \left(1_A(U_j) - \lambda(A) \right) \right. \\
 &\quad \left. \cdot \left(1_B(U_k) - \lambda(B) \right) \left(1_B(U_l) - \lambda(B) \right) \right]
 \end{aligned}$$



$$= n^{-2} \sum_{1 \leq i, j, k, l \leq n} \mu_{ijkl}$$

Wegen Unabhängigkeit + Zentriertheit gilt: $\mu_{ijkl} = 0 \ \forall \ (i, j, k, l) \in \{1, \dots, n\}^4$ mit mindestens 3 voneinander verschiedenen Komponenten, denn dann existiert mindestens ein Index, der von allen anderen verschieden ist. Aus demselben Grund: $\mu_{jiii} = \mu_{ijii} = \mu_{iiji} = \mu_{iiii} = 0 \ (j \neq i)$

$$\Rightarrow \sum_{i, j, k, l} \mu_{ijkl} = \sum_{i=1}^n \mu_{iiii} + \sum_{i \neq j} \left\{ \mu_{iijj} + \mu_{ijij} + \mu_{ijji} + \mu_{jjii} + \mu_{jiji} + \mu_{jiij} \right\}$$

$$\begin{aligned}
 \mu_{iiii} &= E \left(1_A(U_i) - \lambda(A) \right)^2 \left(1_B(U_i) - \lambda(B) \right)^2 \\
 &= E \left[(1_A(U_i) - 2\lambda(A)1_A(U_i) + \lambda^2(A)) (1_B(U_i) - 2\lambda(B)1_B(U_i) + \lambda^2(B)) \right] \\
 &\stackrel{Disj.}{=} \lambda(A)\lambda^2(B) - 2\lambda^2(A)\lambda^2(B) + \lambda^2(A)\lambda(B) - 2\lambda^2(A)\lambda^2(B) + \lambda^2(A)\lambda^2(B) \\
 &= \lambda(A)\lambda(B) \left(\underbrace{\lambda(B)}_{\leq 1} - \underbrace{3\lambda(A)\lambda(B)}_{\geq 0} + \underbrace{\lambda(A)}_{\leq 1} \right) \\
 &\leq 2(t - t_1)(t_2 - t) .
 \end{aligned}$$

$$\begin{aligned}
 \mu_{iijj} &\stackrel{Unabh.}{=} E\left(1_A(U_i) - \lambda(A)\right)^2 E\left(1_B(U_j) - \lambda(B)\right)^2 = \dots \\
 &= \lambda(A)\lambda(B)\left(1 - \overbrace{\lambda(A)}^{\geq 0}\right)\left(1 - \overbrace{\lambda(B)}^{\geq 0}\right) \\
 &\leq (t - t_1)(t_2 - t)
 \end{aligned}$$

$$\mu_{ijij} = E \left[\begin{array}{|c|c|} \hline (1_A(U_i) - \lambda(A)) & (1_A(U_j) - \lambda(A)) \\ \hline (1_B(U_i) - \lambda(B)) & (1_B(U_j) - \lambda(B)) \\ \hline \end{array} \right]$$

$$\begin{aligned}
 &\boxed{} \text{ und } \boxed{} \text{ sind unabhängig} \\
 &= \{E[\boxed{}]\}^2, \text{ da } U_i \stackrel{\mathcal{L}}{=} U_j \\
 &= \lambda^2(A)\lambda^2(B) \\
 &\leq \lambda(A)\lambda(B) = (t - t_1)(t - t_2)
 \end{aligned}$$

Da $\mu_{jjii} = \mu_{iijj}$ und $\mu_{ijij} =$ restliche μ 's, folgt

$$\begin{aligned}
 M_n(t_1, t, t_2) &\leq (t - t_1)(t_2 - t) \underbrace{\left\{ \frac{2}{n} + n(n-1)\frac{6}{n^2} \right\}}_{= \frac{2}{n} + 6 - \frac{6}{n} \leq 6} \\
 &\leq 6(t_2 - t_1)^2
 \end{aligned}$$

$\Rightarrow$ (3.4) erfüllt mit $a = 2$, $b = 2$, $F(t) = \sqrt{6}t$
 $\stackrel{\text{Satz 3.3}}{\Rightarrow}$ Behauptung.

□

Unsere Prozesse S_n , Z_n , ρ_n und Y_n sind zwar rcll, aber haben nicht $I = [a, b]$ als Definitionsbereich, sondern $I = [0, \infty)$ bzw. $I = \mathbb{R}$. Es gibt auch eine Theorie für $D[0, \infty)$ und $D(\mathbb{R})$. Die zugehörige Skorokhod-Metrik $s = s_\infty$ ist derart, dass sich alles auf den kompakten Fall $I = [a, b]$ zurückführen lässt. (Verzichte deshalb auf die komplizierte Definition von s_∞ .)

Satz 3.5. (1) Seien X , X_n , $n \geq 1$, ZE in $(D(\mathbb{R}), s_\infty)$. Dann gilt:

$$X_n \xrightarrow{\mathcal{L}} X \text{ in } (D(\mathbb{R}), s_\infty)$$

$\Leftrightarrow$

$$\begin{aligned}
 X_n|_{[-a, a]} &\xrightarrow{\mathcal{L}} X|_{[-a, a]} \text{ in } (D[-a, a], s) \\
 &\forall a \in T_X \subseteq \mathbb{R}.
 \end{aligned}$$

Dabei $T_X := \{t \in \mathbb{R} : \pi_t \text{ f.s. stetig in } X\}$.

(2) $X, X_n, n \geq 1$, ZE in $(D[0, \infty), s_\infty)$. Dann

$$X_n \xrightarrow{\mathcal{L}} X \text{ in } (D[0, \infty), s_\infty)$$

$\Leftrightarrow$

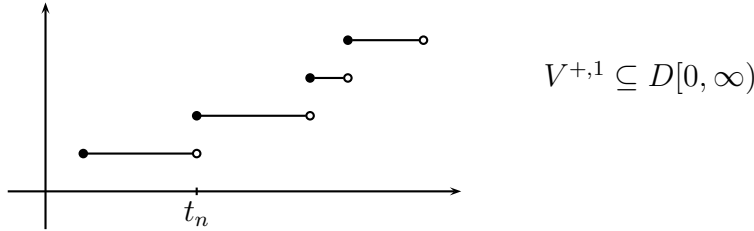
$$\begin{aligned} X_n|_{[0,a]} &\xrightarrow{\mathcal{L}} X|_{[0,a]} \text{ in } (D[0, a], s) \\ \forall a \in T_X &\subseteq [0, \infty) . \end{aligned}$$

Da $\mathcal{B}_{s_\infty}(D(\mathbb{R})) = \sigma(\pi_t : t \in \mathbb{R})$, ist **jeder** stochastische Prozess mit rcll Pfaden automatisch ein ZE in $(D(\mathbb{R}), s_\infty)$. Analoges gilt für $D[0, \infty)$. Damit ist der Messbarkeitsnachweis vollkommen unproblematisch.

In den Anwendungen ist der Nachweis von (3.2), (3.3) oder (3.4) (sogenannte STRAFFHEIT, TIGHTNESS) meist der schwierigere Teil. Dieser entfällt bei einer bestimmten Prozessklasse. Sei $V^{+,1}$ die Menge aller Funktionen $f : [0, \infty) \rightarrow \mathbb{R}$ der Gestalt

$$f(t) = \sum_{n \geq 1} 1_{\{t_n \leq t\}}, \quad t \geq 0 ,$$

wobei $t_n \uparrow \infty$, $t_1 > 0$ und $t_n < t_{n+1}$, falls $t_n < \infty$.



Satz 3.6. Seien $X, X_n, n \geq 1$, ZE in $(D[0, \infty), s_\infty)$ mit Pfaden in $V^{+,1}$. Falls

$$\pi_T(X_n) \xrightarrow{\mathcal{L}} \pi_T(X) \quad \forall T \subseteq \Delta, \quad T \text{ endlich}, \quad (3.5)$$

wobei Δ irgendeine dichte Teilmenge von $[0, \infty)$ ist, dann:

$$X_n \xrightarrow{\mathcal{L}} X \text{ in } (D[0, \infty), s_\infty).$$

Dieses Resultat findet sich in Jacod und Shiryaev [JS03]. Fenger und Vogel [FV07] geben eine Verallgemeinerung an. Dabei dürfen die X_n Sprungprozesse sein mit beliebigen ganzzahligen, positiven Sprunghöhen und nur der Grenzprozess X muss Sprünge sämtlich der Höhe Eins aufweisen. Eine Formulierung dieses Ergebnisses geht auf M. Csörgő und Horváth [CH88] zurück. Deren Beweis ist allerdings

fehlerhaft.

Für eine Anwendung des Satzes betrachte U_i , $i \geq 1$, i.i.d. $\sim \mathcal{U}(0, 1)$ mit empirischer VF

$$F_n(t) = \frac{1}{n} \sum_{i=1}^n 1_{\{U_i \leq t\}}, \quad t \in \mathbb{R}.$$

Dann ist

$$\beta_n(t) := nF_n\left(\frac{t}{n}\right) = \sum_{i=1}^n 1_{\{U_i \leq \frac{t}{n}\}}, \quad t \geq 0$$

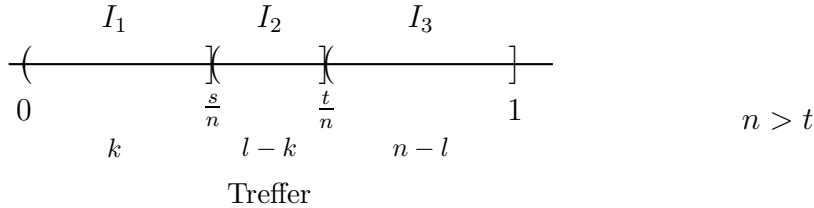
ein stochastischer Prozess (SP), also $\beta_n = \{\beta_n(t) : t \geq 0\}$ ein ZE in $D[0, \infty)$. Es gilt

Satz 3.7. $\beta_n \xrightarrow{\mathcal{L}} N$ in $D[0, \infty)$, wobei N ein Poisson-Prozess ist mit Parameter $\lambda = 1$.

Beweis: $\beta_n \in V^{+,1}$ ($t_i = nU_{i:n}$, $1 \leq i \leq n$, $t_i = +\infty \forall i \geq n+1$) $\forall n \geq 1$ und $N \in V^{+,1}$ ($t_i = i$ -te Eintrittszeit). Also bleibt nach Satz 3.6 nur Nachweis der Konvergenz der fidis.

Seien $0 < s < t < \infty \Rightarrow (\beta_n(s), \beta_n(t))$ **diskrete** ZV aus $\mathbb{N}_0^2$. Betrachte deshalb für $k \leq l$ aus $\mathbb{N}_0$:

$$P(\beta_n(s) = k, \beta_n(t) = l) = P(\beta_n(s) = k, \beta_n(t) - \beta_n(s) = l - k) =: q_n$$



$$\begin{aligned} \Rightarrow \beta_n(s) &= \sum 1_{\{U_i \in I_1\}} = \# \text{ Treffer 1. Art } = k \\ \beta_n(t) - \beta_n(s) &= \sum 1_{\{U_i \in I_2\}} = \# \text{ Treffer 2. Art } = l - k \\ &\Rightarrow \sum 1_{\{U_i \in I_3\}} = \# \text{ Treffer 3. Art } = n - l \end{aligned}$$

$$\begin{aligned}
 \Rightarrow q_n &= \text{Mult } (n, \underline{p})(\{k, l-k, n-l\}) \text{ mit } \underline{p} = \left(\frac{s}{n}, \frac{t-s}{n}, 1 - \frac{t}{n}\right) \\
 &= \frac{n!}{k!(l-k)!(n-l)!} \left(\frac{s}{n}\right)^k \left(\frac{t-s}{n}\right)^{l-k} \left(1 - \frac{t}{n}\right)^{n-l} \\
 &= \underbrace{\frac{n \cdot (n-1) \cdot \dots \cdot (n-l+1)}{n^l}}_{\rightarrow 1} \frac{s^k}{k!} \frac{(t-s)^{l-k}}{(l-k)!} \underbrace{\left(1 - \frac{t}{n}\right)^n}_{\rightarrow e^{-t}} \underbrace{\left(1 - \frac{t}{n}\right)^{-l}}_{\rightarrow 1} \\
 &\rightarrow e^{-s} \frac{s^k}{k!} e^{-(t-s)} \frac{(t-s)^{l-k}}{(l-k)!} \\
 &= P(N(s) = k) P(N(t) - N(s) = l-k) \\
 &= P(N(s) = k, N(t) - N(s) = l-k) \quad \text{Wegen Unabhängigkeit der Zuwächse} \\
 &= P(N(s) = k, N(t) = l) \Rightarrow \\
 &\quad (\beta_n(s), \beta_n(t)) \xrightarrow{\mathcal{L}} (N(s), N(t)).
 \end{aligned}$$

Analog behandelt man $k \geq 2$ Punkte $t_1 < \dots < t_k \xrightarrow{\text{Satz 3.6}} \Rightarrow$ Behauptung (mit $\Delta = (0, \infty)$) $\square$

Bemerkung. In der Diplomarbeit von Vogel [Vog05] wird gezeigt: $(\alpha_n, \beta_n) \xrightarrow{\mathcal{L}} (B_0, N)$ in $(D[0, 1] \times D[0, \infty), s \times s_\infty)$, wobei B_0 und N **unabhängig**.

Zum Abschluss eine Variante von Satz 3.6. Ein Beweis findet sich in Jacod und Shiryaev [JS03], p. 345.

Satz 3.8. Seien $X, X_n, n \geq 1$ ZE in $(D[0, \infty), s_\infty)$ mit monoton wachsenden Pfaden. Falls X stetig, dann liefert (3.5) bereits

$$X_n \xrightarrow{\mathcal{L}} X \text{ in } (D[0, \infty), s_\infty).$$

Beispiel. Seien $\xi_i, i \geq 1$, i.i.d. $\sim F$ mit Träger $[0, \infty)$. Setze

$$\alpha_n(s) := \frac{1}{\sqrt{n}} \sum_{i=1}^n (1_{\{\xi_i \leq s\}} - F(s)), \quad s \geq 0.$$

$\Rightarrow (\alpha_n) \subseteq D[0, \infty)$. Mit der Quantil-Transformation und Theorem 16.6 in Billingsley [Bil99] zeigt man, dass

$$\alpha_n \xrightarrow{\mathcal{L}} \bar{\alpha}_n \circ F,$$

wobei $\bar{\alpha}_n$ = uniformer empirischer Prozess.

Die Abbildung $h : (D[0, 1], s) \rightarrow (D[0, \infty), s_\infty)$, $h(f) := f \circ F$ ist stetig auf $C[0, 1]$, denn:

$$(f_n) \subseteq D[0, 1] \text{ mit } f_n \rightarrow_s f \in C[0, 1] \Rightarrow f_n \rightarrow_d f .$$

Es folgt $\forall N > 0$:

$$\begin{aligned} \sup_{0 \leq t \leq N} |h(f_n)(t) - h(f)(t)| &= \sup_{0 \leq t \leq N} |f_n(F(t)) - f(F(t))| \\ &\leq d(f_n, f) \rightarrow 0 \Rightarrow h(f_n) \rightarrow_{\delta_{lu}} h(f) , \end{aligned}$$

wobei δ_{lu} = Metrik der gleichmaigen Konvergenz auf Kompakta (local uniform topology). Da δ_{lu} starker als s_∞ (vgl. Proposition 1.17, p. 328 in Jacod und Shiryaev [JS03]) folgt Stetigkeit von h . Da $P(B_0 \in C[0, 1]) = 1$ folgt mit CMT:

$$\alpha_n \xrightarrow{\mathcal{L}} \alpha_0 \text{ in } D[0, \infty) ,$$

wobei $\alpha_0(s) = B_0(F(s))$, $s \geq 0$. Betrachte

$$\begin{aligned} X_n(t) &:= \int_0^t |\alpha_n(s)| ds, \quad t \geq 0 \\ X(t) &:= \int_0^t |B_0(F(s))| ds, \quad t \geq 0 . \end{aligned}$$

Mit Continuous Mapping Theorem (CMT) folgt (checken!), dass (3.5) erfullt ist. Somit liefert Satz 3.8

$$\left\{ \int_0^t |\alpha_n(s)| ds : t \geq 0 \right\} \xrightarrow{\mathcal{L}} \left\{ \int_0^t |B_0(F(s))| ds : t \geq 0 \right\}$$

in $(D[0, \infty), s_\infty)$.

Kapitel 4

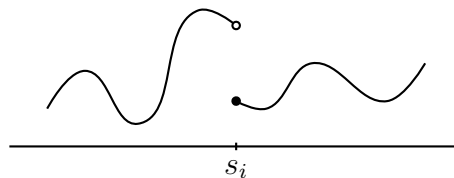
Stetigkeitssätze für das argmax-Funktional

Erinnere an Problem: Wann überträgt sich die Konvergenz von stochastischen Prozessen auf deren Maximalstellen? Dazu zunächst genaue Definition der Maximalstelle $\operatorname{argmax}(f)$ für f rcll. Es folgt aus Lemma 2.1(2), dass jedes $f \in D(\mathbb{R})$ höchstens abzählbar viele Sprungstellen s_i , $i \in I \subseteq \mathbb{N}$ besitzt. Sei

$$\delta := \delta(f) := \inf\{|s_i - s_j| : i \neq j \in I\} .$$

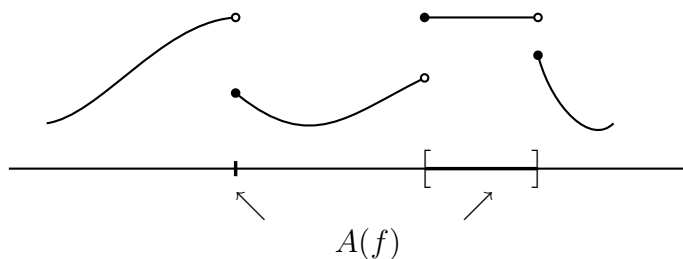
Falls $\delta(f) > 0$, so gilt $\forall s_i$:

$$f \text{ stetig auf } (s_i, s_i + \delta) \text{ und } f \text{ stetig auf } (s_i - \delta, s_i) \quad (4.1)$$



Sei

$$A(f) := \{t \in \mathbb{R} : f(t) \vee f(t-) = \sup_{s \in \mathbb{R}} f(s)\}.$$



Lemma 4.1. $f \in D(\mathbb{R})$ mit $\delta(f) > 0 \Rightarrow A(f)$ abgeschlossen.

Beweis: $A(f) = \emptyset \Rightarrow A(f)$ abgeschlossen. Also o.E. $A(f) \neq \emptyset$. Sei $(t_n) \subseteq A(f)$ mit $t_n \rightarrow t$.

1. Fall: $t \notin \{s_i : i \in I\} =: J \Rightarrow \exists$ aufeinanderfolgende $s < s' \in J$ mit $t \in (s, s') \Rightarrow t_n \in (s, s')$ schließlich $\Rightarrow (*) f(t_n) = f(t_n-) = \sup_{s \in \mathbb{R}} f(s)$ schließlich $\Rightarrow$

$$\begin{aligned} f(t) &= \lim_{n \rightarrow \infty} f(t_n) \quad \text{wegen } f \text{ stetig in } t \\ &= \sup_{s \in \mathbb{R}} f(s) \quad \text{wegen } (*) \text{ und } t_n \in A(f) \end{aligned}$$

$\Rightarrow t \in A(f)$.

2. Fall: $\exists i \in I : t = s_i$. O.E. $t_n \downarrow t$ oder $t_n \uparrow t$

(i) $t_n \downarrow t \Rightarrow t_n \in (t, t + \delta)$ schließlich mit $\delta = \delta(f) \stackrel{(4.1)}{\Rightarrow} f(t_n) = f(t_n-)$ schließlich $\Rightarrow t \in A(f)$ (wie im 1. Fall).

(ii) $t_n \uparrow t \Rightarrow t_n \in (t - \delta, t)$ schließlich $\Rightarrow f(t_n) \stackrel{(4.1)}{=} f(t_n-) \stackrel{t_n \in A(f)}{=} \sup_{s \in \mathbb{R}} f(s)$ schließlich $\Rightarrow f(t-) = \lim_{n \rightarrow \infty} f(t_n) = \sup_{s \in \mathbb{R}} f(s)$
 $\Rightarrow t \in A(f)$.

□

Beachte: $\delta(f) > 0$ ist immer gegeben, solange die Sprungstellen sich nicht häufen. Damit sind nur „pathologische“ f aus der folgenden Definition ausgeschlossen.

Definition 4.2. $f \in D(\mathbb{R})$.

Falls $\emptyset \neq A(f)$ abgeschlossen und nach unten beschränkt (nub), dann

$$\operatorname{argmax}(f) := \min A(f) = \text{kleinste Maximalstelle (Supremalstelle)}.$$

Falls $\emptyset \neq A(f)$ abgeschlossen und nach oben beschränkt, dann

$$\operatorname{Argmax}(f) := \max A(f) = \text{größte Maximalstelle (Supremalstelle)}.$$

Lemma 4.3. $f \in D(\mathbb{R})$, $\tau \in A(f)$. Angenommen für ein $\varepsilon > 0$ gilt:

$$\sup_{t \in \mathbb{R}} f(t) > \sup\{f(t) : |t - \tau| > \varepsilon\} =: a(\varepsilon).$$

Dann folgt:

$$(1) \quad b(\varepsilon) := \frac{1}{3}(\sup_{t \in \mathbb{R}} f(t) - a(\varepsilon)) > 0.$$

(2) $g \in D(\mathbb{R})$, $\sigma \in A(g) \neq \emptyset$. Dann:

$$\|f - g\| \leq b(\varepsilon) \Rightarrow |\tau - \sigma| \leq \varepsilon.$$

Beweis: (1) ist trivial.

(2) Nach Vor. gilt $\sup_{t \in \mathbb{R}} f(t) \in \{f(\tau), f(\tau-)\}$. Sei $t \in \mathbb{R}$ mit $|t - \tau| > \varepsilon$. Wegen

$$(*) \|f - g\| \leq b(\varepsilon) \Leftrightarrow -b(\varepsilon) \leq g(s) - f(s) \leq b(\varepsilon) \quad \forall s \in \mathbb{R}$$

folgt $g(\tau \pm) - f(\tau \pm) \geq -b(\varepsilon)$ und $g(t) \leq f(t) + b(\varepsilon) \leq a(\varepsilon) + b(\varepsilon)$

$$\begin{aligned} \Rightarrow g(\tau \pm) - g(t) &\geq f(\tau \pm) - b(\varepsilon) - a(\varepsilon) - b(\varepsilon) \\ &= \underbrace{\sup_{t \in \mathbb{R}} f(t) - a(\varepsilon) - 2b(\varepsilon)}_{=3b(\varepsilon)} = b(\varepsilon) > 0 \quad \Rightarrow \end{aligned}$$

$$g(\tau \pm) \geq g(t) + b(\varepsilon) \quad \forall t \notin [\tau - \varepsilon, \tau + \varepsilon]$$

$$\begin{array}{ccccccc} \xrightarrow{\quad} & & & & \xrightarrow{\quad} & & \\ s & t & \tau - \varepsilon & \tau & \tau + \varepsilon & s & t \end{array} \quad \text{Grenzübergang } s \uparrow t \text{ liefert:}$$

$$g(\tau \pm) \geq g(t-) + b(\varepsilon) \quad \forall t \notin [\tau - \varepsilon, \tau + \varepsilon]$$

$$\stackrel{b(\varepsilon) > 0}{\Rightarrow} (+) g(t) \vee g(t-) < g(\tau \pm) \quad \forall t \notin [\tau - \varepsilon, \tau + \varepsilon]$$

Angenommen $|\sigma - \tau| > \varepsilon$, also $\sigma \notin [\tau - \varepsilon, \tau + \varepsilon] \Rightarrow$

$$\begin{aligned} \sup_{s \in \mathbb{R}} g(s) &= g(\sigma) \vee g(\sigma-) \\ &\stackrel{(+)}{<} g(\tau \pm) \leq \sup_{s \in \mathbb{R}} g(s). \end{aligned}$$

Widerspruch $\Rightarrow |\sigma - \tau| \leq \varepsilon$. □

Beachte, in Lemma 4.3 keine Eindeutigkeit von τ gefordert.

Folgerung 4.4. f und τ wie in 4.3. Angenommen:

$$(1) \quad \sup_{t \in \mathbb{R}} f(t) > a(\varepsilon) \quad \forall \varepsilon \in \mathbb{Q}.$$

Sei $(f_n) \subseteq D(\mathbb{R})$ und $\tau_n \in A(f_n) \neq \emptyset \quad \forall n$ schließlich. Falls $\|f_n - f\| \rightarrow 0$, dann

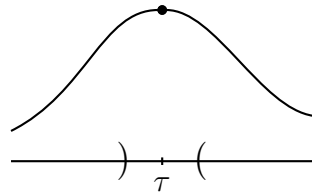
$$\tau_n \rightarrow \tau.$$

Beweis: Sei $\varepsilon > 0$ beliebig. O.E. $\varepsilon \in \mathbb{Q} \cap (0, \varepsilon_0] \Rightarrow b(\varepsilon) > 0 \Rightarrow \exists n_0 = n_0(\varepsilon) : \|f_n - f\| \leq b(\varepsilon) \quad \forall n \geq n_0 \xrightarrow{\text{Lemma 4.3}} |\tau_n - \tau| \leq \varepsilon \quad \forall n \geq n_0(\varepsilon).$ $\square$

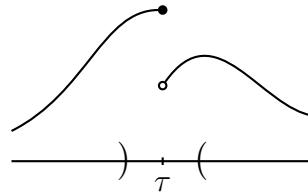
Bemerkung. (1) $\tau \in A(f)$ mit Eigenschaft (1) aus 4.4 ist eindeutige Maximalstelle, denn angenommen $\exists \tilde{\tau} \neq \tau, \tilde{\tau} \in A(f)$. Wegen $\tilde{\tau} \neq \tau : \exists \varepsilon \in \mathbb{Q} : |\tilde{\tau} - \tau| > \varepsilon.$

$$\Rightarrow \underbrace{f(\tilde{\tau} \pm)}_{= \sup(f)} \leq a(\varepsilon) < \sup(f) \quad \text{Wid.}$$

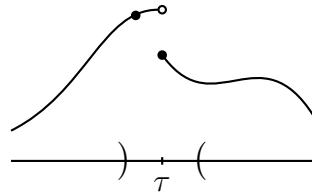
(2) $\tau \in A(f)$ mit 4.4(1) heißt wohlsepariert



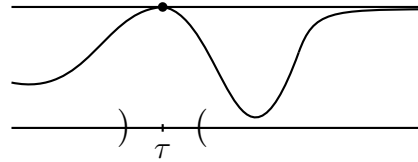
wohlsepariert



wohlsepariert



wohlsepariert

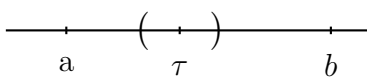


nicht wohlsepariert

Die Aussagen 4.1-4.4 gelten entsprechend für $D(T)$ mit $T = [0, \infty)$, $T = [a, b]$, $T = (a, b)$, etc. In Anwendungen ist f oft stetig und die Wohlsepariertheit durch Kompaktheits- und Monotonieargumente nachweisbar.

Folgerung 4.5. Sei $f \in C[a, b]$, $A(f) = \{\tau\}$, $(f_n) \subseteq D[a, b]$, $\tau_n \in A(f_n) \neq \emptyset \quad \forall n$. Dann gilt:

$$\|f_n - f\| \rightarrow 0 \Rightarrow \tau_n \rightarrow \tau.$$

Beweis: 

$$\varepsilon_0 := \frac{1}{2} \min(\tau - a, b - \tau) > 0$$

$\Rightarrow f$ nimmt auf Kompaktum $[a, \tau - \varepsilon] \cup [\tau + \varepsilon, b] =: K$ das Supremum an.
 $\Rightarrow \exists \sigma \in K : a(\varepsilon) = f(\sigma) \Rightarrow f(\tau) > a(\varepsilon)$, sonst: $f(\tau) = f(\sigma) \Rightarrow \sigma \in A(f)$, $\sigma \neq \tau$
 Widerspruch $\Rightarrow$ Wohlsepariertheit $\xrightarrow{\text{Folg. 4.4}}$ Behauptung. $\square$

Folgerung 4.4 liefert unser 1. CMT:

Satz 4.6. $M, M_n, n \geq 1$, SP über $(\Omega, \mathcal{A}, P)$ mit Pfaden in $D(\mathbb{R})$ und

- (1) $\tau \in A(M)$ f.s.
- (2) $\sup_{t \in \mathbb{R}} M(t) > \sup\{M(t) : |t - \tau| > \varepsilon\}$ f.s. $\forall \varepsilon > 0$.
- (3) $A(M_n) \neq \emptyset$ f.s. $\forall n \geq 1$.
- (4) $\|M_n - M\| \rightarrow 0$ f.s. für $n \rightarrow \infty$.

Für jede messbare Auswahl $\tau_n \in A(M_n)$ folgt:

$$\tau_n \rightarrow \tau \quad \text{f.s.}$$

Beweis: $\Omega_1 := \{\tau \in A(M)\}$,

$$\Omega_2 := \left\{ \sup_{t \in \mathbb{R}} M(t) > a(\varepsilon) \quad \forall \varepsilon \in \mathbb{Q} \right\},$$

$$\Omega_3 := \{A(M_n) \neq \emptyset \quad \forall n \geq 1\},$$

$$\Omega_4 := \{\|M_n - M\| \rightarrow 0\}. \quad \Rightarrow$$

$$\Omega_0 := \Omega_1 \cap \dots \cap \Omega_4 \text{ ist } P\text{-Einsmenge} \stackrel{\text{Folg. 4.4}}{\Rightarrow} \Omega_0 \subseteq \{\tau_n \rightarrow \tau\}$$

□

Analog erhält man aus Folgerung 4.5 einen Stetigkeitssatz.

Beispiel. $X_1, \dots, X_n$ i.i.d., $X_i \sim \beta_\theta$, $\theta \in [a, b]$, $0 \leq a \leq b \leq 1$ bekannt. Dann betrachte

$$\begin{aligned} \hat{\theta}_n &:= \operatorname{argmax}_{a \leq t \leq b} M_n(t) \\ M_n(t) &:= \overline{X}_n \log(t) + (1 - \overline{X}_n) \log(1 - t), \quad t \in [a, b] \\ (\Rightarrow M_n &= \frac{1}{n} l_n, \quad l_n = \text{Log-Likelihood-Funktion}) \end{aligned}$$

$M_n \rightarrow M$ gleichmäßig auf $[a, b]$ f.s., wobei

$$M(t) = \theta \log(t) + (1 - \theta) \log(1 - t)$$

stetig mit $A(M) = \{\theta\} \stackrel{\text{Folg. 4.5}}{\Rightarrow} \hat{\theta}_n \rightarrow \theta$ f.s.

□

Aus Satz 4.6 bzw. Analogon folgt insbesondere

$$\tau_n - \tau \xrightarrow{\mathcal{L}} 0.$$

Problem: Finde (α_n) mit $\alpha_n \rightarrow \infty$, so dass

$$\alpha_n(\tau_n - \tau) \xrightarrow{\mathcal{L}} T,$$

wobei T nicht-degenerierte ZV. Folgendes Lemma liefert die Grundidee

Lemma 4.7. Sei $g \in D(\mathbb{R})$ mit Sprungstellenmenge

$$J(g) := \{s \in \mathbb{R} : |f(s) - f(s-)| > 0\} .$$

Für $\alpha, c > 0$ und $\tau \in \mathbb{R}$ definiere

$$h(t) := c \left\{ g \left(\tau + \frac{t}{\alpha} \right) - g(\tau) \right\}, \quad t \in \mathbb{R} .$$

Dann gilt:

(1) $h \in D(\mathbb{R})$.

(2) Zu jedem $t \in J(h)$ existiert genau ein $s \in J(g)$ und umgekehrt. Für diese gilt:

$$(*) \quad t = \alpha(s - \tau) .$$

(3) Zu jedem $t \in A(h)$ existiert genau ein $s \in A(g)$ und umgekehrt. Für diese gilt (*). Insbesondere: $A(h) = \alpha(A(g) - \tau)$.

(4) $\operatorname{argmax}(h) = \alpha(\operatorname{argmax}(g) - \tau)$ und $\operatorname{Argmax}(h) = \alpha(\operatorname{Argmax}(g) - \tau)$.

Beweis: Die Abbildung $t \mapsto \tau + \frac{t}{\alpha}$ ist homöomorph.

$$\Rightarrow h(t-) = c \left\{ g \left(\left(\tau + \frac{t}{\alpha} \right) - \right) - g(\tau) \right\}$$

1. Fall: $\tau + \frac{t}{\alpha} \notin J(g) \Rightarrow h(t-) = h(t)$

2. Fall: $\tau + \frac{t}{\alpha} \in J(g) \Rightarrow (+) \quad \tau + \frac{t}{\alpha} = s, s \in J(g)$ eindeutig wegen Bijektivität.

$$\begin{aligned} \Rightarrow h(t-) &= c\{g(s-) - g(\tau)\} \\ &\neq c\{g(s) - g(\tau)\} && \text{wegen } s \in J(g) \\ &= h(t) && \text{wegen } (+) \end{aligned}$$

Analog: $h(t+) = h(t)$ stets $\Rightarrow h \in D(\mathbb{R})$. Ferner:

$$\tau + \frac{t}{\alpha} \in J(g) \Leftrightarrow t \in J(h) \quad (++)$$

Dies zeigt (1) und (2).

Zu (3): Sei $t \in A(h)$.

1. Fall: $t \notin J(h)$. Dann:

$$\begin{aligned} h(t) &\geq h(u) \quad \forall u \in \mathbb{R} \Leftrightarrow \\ g\left(\tau + \frac{t}{\alpha}\right) &\geq g\left(\tau + \frac{u}{\alpha}\right) \quad \forall u \in \mathbb{R} \Leftrightarrow \\ g\left(\tau + \frac{t}{\alpha}\right) &\geq g(v) \quad \forall v \in \mathbb{R} \Rightarrow \\ \tau + \frac{t}{\alpha} &\in A(g) \Rightarrow t = \alpha(s - \tau), \quad s \in A(g) \end{aligned}$$

2. Fall: $t \in J(h) \stackrel{++}{\Rightarrow} t = \alpha(s - \tau), \quad s \in J(g) \Rightarrow$

$$\begin{aligned} h(t-) &\geq h(u) \quad \forall u \in \mathbb{R} \Rightarrow \\ g(s-) &\geq g(v) \quad \forall v \in \mathbb{R} \Rightarrow \\ s &\in A(g) \Rightarrow t = \alpha(s - \tau), \quad s \in A(g) \end{aligned}$$

Umkehrung: analog. Dies zeigt (3). Aus (3) folgt:

$$A(h) = \alpha(A(g) - \tau)$$

$\Rightarrow$ (4) (Rechenregeln für min und max) □

Bemerkung 4.8. Es gibt andere Transformationen h von g , die dasselbe wie Lemma 4.7(4) leisten, z.B. für jede positive Konstante c :

$$h(t) = c \cdot \frac{g\left(\tau + \frac{t}{\alpha}\right)}{g(\tau)}.$$

□

Eine Anwendung von Lemma 4.7 auf $g = M_n$, $s = \tau_n \in A(M_n)$, $\tau = \operatorname{argmax}(M)$, $c = c_n$, $\alpha = \alpha_n$ liefert:

$$\alpha_n(\tau_n - \tau) \in A(L_n), \tag{4.2}$$

wobei

$$L_n(t) = c_n \left\{ M_n\left(\tau + \frac{t}{\alpha_n}\right) - M_n(\tau) \right\}, \quad t \in \mathbb{R}.$$

Die Darstellung (4.2) in Kombination mit folgendem 2. CMT liefert die Verteilungskonvergenz.

Satz 4.9. (*Argmax-CMT für Verteilungskonvergenz*) Seien L , L_n , $n \in \mathbb{N}$, SP über $(\Omega, \mathcal{A}, P)$ mit Pfaden in $D(\mathbb{R})$, und sei $\sigma_n \in A(L_n) \neq \emptyset$ messbar. Falls

$$(1) \quad L_n \xrightarrow{\mathcal{L}} L \text{ in } (D[-a, a], s) \quad \forall a > 0,$$

(2) $\emptyset \neq A(L)$ abgeschlossen und beschränkt f.s.

und

(3) $\sigma_n = O_P(1)$, d.h. $\lim_{a \rightarrow \infty} \limsup_{n \rightarrow \infty} P(|\sigma_n| > a) = 0$,

so gilt:

(4) $\liminf_{n \rightarrow \infty} P(\sigma_n \leq x) \geq P(\sigma_{\max} < x) \quad \forall x \in \mathbb{X}$,

(5) $\limsup_{n \rightarrow \infty} P(\sigma_n \leq x) \leq P(\sigma_{\min} \leq x) \quad \forall x \in \mathbb{X}$,

wobei $\sigma_{\min} := \operatorname{argmax}(L)$, $\sigma_{\max} := \operatorname{Argmax}(L)$ und

$$\mathbb{X} := \{x \in \mathbb{R} : L \text{ stetig in } x \text{ f.s.}\}.$$

Ferner gilt:

(6) $\mathbb{R} \setminus \mathbb{X}$ ist höchstens abzählbar, also liegt $\mathbb{X}$ **dicht** in $\mathbb{R}$.

(7) Falls σ_n echte Maximalstelle, so lässt sich (4) verschärfen zu

$$\liminf_{n \rightarrow \infty} P(\sigma_n < x) \geq P(\sigma_{\max} < x) \quad \forall x \in \mathbb{X}.$$

Zum Beweis von Satz 4.9 verwende 2 Lemmata.

Lemma 4.10. $f \in D[a, b]$, $\sigma \in A(f)$ abgeschlossen, $\underline{\sigma} := \operatorname{argmax}(f)$, $\bar{\sigma} := \operatorname{Argmax}(f)$. Dann gilt $\forall x \in [a, b]$:

(1) $\sigma \leq x \Rightarrow \sup_{a \leq t \leq x} f(t) \geq \sup_{x \leq t \leq b} f(t) \Rightarrow \underline{\sigma} \leq x$

(2) $\sigma > x \Rightarrow \sup_{a \leq t \leq x} f(t) \leq \sup_{x \leq t \leq b} f(t) \Rightarrow \bar{\sigma} \geq x$.

Falls σ **echte** Maximalstelle so gilt:

(3) $\sigma \geq x \Rightarrow \sup_{a \leq t \leq x} f(t) \leq \sup_{x \leq t \leq b} f(t) \Rightarrow \bar{\sigma} \geq x$.

Beweis: Setze $M[a, x] := \sup_{a \leq t \leq x} f(t)$, $M[x, b] := \sup_{x \leq t \leq b} f(t)$. Sei $\sigma \leq x$.

1. Fall: $f(\sigma) = M[a, b] \Rightarrow M[a, x] \overset{\sigma \in [a, x]}{\geq} f(\sigma) = M[a, b] \geq M[x, b]$.

Beweis von Satz 4.9: Sei $x \in \mathbb{X}$ beliebig. Finde $a > 0$ mit $|x| < a \Rightarrow$

$$\begin{aligned} \{\sigma_n \leq x\} &= \underbrace{\{-a \leq \sigma_n \leq x\}}_{\substack{\text{Lemma 4.10(1)} \\ \subseteq \{ \sup_{-a \leq t \leq x} L_n(t) \geq \sup_{x \leq t \leq a} L_n(t) \}}} \cup \underbrace{\{\sigma_n < -a\}}_{\subseteq \{|\sigma_n| > a\}} \\ \Rightarrow P(\sigma_n \leq x) &\leq P(\Lambda_x(L_n) \geq 0) + P(|\sigma_n| > a) \text{ mit:} \end{aligned}$$

$$\Lambda_x(f) := \sup_{-a \leq t \leq x} f(t) - \sup_{x \leq t \leq a} f(t)$$

messbar und stetig P_L -f.ü. (denn: $x \in \mathbb{X} \Leftrightarrow P(L \in E_x) = 1$ und $E_x \subseteq C_{\Lambda_x}$ gemäß 4.11)

$$\begin{aligned} &\xRightarrow{\text{CMT}} \Lambda_x(L_n) \xrightarrow{\mathcal{L}} \Lambda_x(L), \quad n \rightarrow \infty \\ &\xRightarrow{\text{Portm.}} \limsup_{n \rightarrow \infty} P(\sigma_n \leq x) \leq P(\Lambda_x(L) \geq 0) + \underbrace{\limsup_{n \rightarrow \infty} P(|\sigma_n| > a)}_{\rightarrow 0 \text{ (} a \rightarrow \infty)} \quad \forall a > |x|. \end{aligned}$$

Sei $\sigma_{\min}(a) := \min A(L|_{[-a,a]})$. Dann folgt aus Lemma 4.10(1), dass $\{\Lambda_x(L) \geq 0\} \subseteq \{\sigma_{\min}(a) \leq x\} \subseteq \{\sigma_{\min} \leq x\}$, wobei die letzte Inklusion wegen $\sigma_{\min} \leq \sigma_{\min}(a) \forall a > 0$ gilt. $\Rightarrow$ (5). Weiter: (4) $\Leftrightarrow$ (Übergang zum Komplement)

$$\begin{aligned} (4)^* \quad \limsup_{n \rightarrow \infty} P(\sigma_n > x) &\leq P(\sigma_{\max} \geq x) \\ \{\sigma_n > x\} &= \underbrace{\{x < \sigma_n \leq a\}}_{\substack{\text{Satz 4.10(2)} \\ \subseteq \{\Lambda_x(L_n) \leq 0\}}} \cup \underbrace{\{\sigma_n > a\}}_{\subseteq \{|\sigma_n| > a\}} \end{aligned}$$

Rest wie oben mit Satz 4.10(2) $\Rightarrow$ (4). Die Aussage (6) beweist man wie Proposition 3.12 in Jacod und Shiryaev [JS03]. (7) folgt analog wie (4) mit Lemma 4.10(3). $\square$

Man erhält in Satz 4.9 sogar Verteilungskonvergenz, falls L f.s. **eindeutige** Maximalstelle besitzt. Zum Nachweis verwende

Satz 4.12. (*Pólya*)

(1) Seien $F, F_n : \mathbb{R} \rightarrow \mathbb{R}$ ($n \in \mathbb{N}$) monoton und $\Delta \subseteq \mathbb{R}$ dicht. Dann:

$$F_n(x) \rightarrow F(x) \quad \forall x \in \Delta \quad \Rightarrow \quad F_n(x) \rightarrow F(x) \quad \forall x \in C_F.$$

(2) Ist F zusätzlich stetig und beschränkt, so:

$$F_n(x) \rightarrow F(x) \quad \forall x \in \Delta \Rightarrow \|F_n - F\| \rightarrow 0.$$

(3) Für $F, F_n : [a, b] \rightarrow \mathbb{R}$ gilt (2) analog.

Beweis: $\rightarrow$ Witting und Müller-Funk [WMF95], p. 70-71. $\square$

Folgerung 4.13. *Gilt in Satz 4.9 anstelle von (2) sogar*

$$(U) \quad A(L) = \{\tau\} \quad \text{f.s.} \quad (\underline{\text{Uniqueness}})$$

mit $\tau : \Omega \rightarrow \mathbb{R}$ ist ZV, so folgt:

$$\sigma_n \xrightarrow{\mathcal{L}} \tau.$$

Beweis: $\sigma_{\min} = \sigma_{\max} = \tau$ f.s. wegen (U). Aus Satz 4.9(4) und (5) folgt:

$$F_n(x) := P(\sigma_n \leq x) \rightarrow F(x) := P(\tau \leq x) \quad \forall x \in \mathbb{X} \cap C_F =: D.$$

Gemäß Satz 4.9(6) ist $\overline{\mathbb{X}} = \mathbb{R} \setminus \mathbb{X}$ höchstens abzählbar und bekanntlich auch $\overline{C_F}$, also auch $\overline{D} := \overline{\mathbb{X}} \cup \overline{C_F} \Rightarrow D$ dicht in $\mathbb{R} \Rightarrow$ Behauptung mit Satz 4.12(1). $\square$

Das Argmax-CMT lässt sich verallgemeinern. Damit kann man später z.B. mehrdimensionale MLE behandeln.

Satz 4.14. (*Extended Argmax-CMT*)

Bezeichnungen wie in Satz 4.9. Ferner seien $Y, Y_n, n \geq 1$ ZV über $(\Omega, \mathcal{A}, P)$ in einen separablen und vollständigen metrischen Raum (S, d) . Annahmen:

(1) $L_n \xrightarrow{\mathcal{L}} L$ in $D[-a, a] \forall a > 0$. (Bemerkung: Es reicht Straffheit von L)

(2) $(\pi_T(L_n), Y_n) \xrightarrow{\mathcal{L}} (\pi_T(L), Y)$ in $\mathbb{R}^{|T|} \times S \forall T \subseteq \mathbb{R}$ endlich.

(3) $\sigma_n = O_P(1)$.

Dann gilt $\forall x \in \mathbb{X} \forall B$ Y -randlos:

(4) $\liminf_{n \rightarrow \infty} P(\sigma_n \leq x, Y_n \in B) \geq P(\sigma_{\max} < x, Y \in B).$
 $<, \text{ falls } \sigma_n \text{ echte Max.-stelle}$

(5) $\limsup_{n \rightarrow \infty} P(\sigma_n \leq x, Y_n \in B) \leq P(\sigma_{\min} \leq x, Y \in B),$

wobei $\mathbb{X}$ dicht in $\mathbb{R}$ gemäß Satz 4.9(6).

Beweis: Schneidet man im Beweis von Satz 4.9 das Ereignis $\{\sigma_n \leq x\}$ mit $\{Y_n \in B\}$, so erhält man analog:

$$(i) \quad P(\sigma_n \leq x, Y_n \in B) \leq P((\Lambda_x(L_n), Y_n) \in [0, \infty) \times \overbrace{B}^{\subseteq \text{cl}(B)}) + P(|\sigma_n| > a)$$

Zeige: (ii) $(L_n, Y_n) \xrightarrow{\mathcal{L}} (L, Y)$ in $D[-a, a] \times S$.

Dazu: (L_n) straff nach Voraussetzung $(L_n \xrightarrow{\mathcal{L}} L \Rightarrow (L_n) \text{ relativ kompakt} \Rightarrow (L_n)$ straff mit Prokhorov (vgl. Theorem 5.2 in [Bil99])). Aus (2) folgt $Y_n \xrightarrow{\mathcal{L}} Y \Rightarrow (Y_n)$ straff wegen Prokhorov $\Rightarrow (L_n, Y_n)$ straff in $D[-a, a] \times S$ (versehen mit Produkttopologie) wegen Satz von Tychonov. Sei (n') beliebige Teilfolge von (n) . Da $D[-a, a] \times S$ separabel, ist gemäß Prokhorov ist (L_n, Y_n) relativ kompakt, also existiert Teilteilfolge (n'') mit

$$(iii) (L_{n''}, Y_{n''}) \xrightarrow{\mathcal{L}} (L', Y')$$

für eine Grenzvariable (L', Y') , die (zunächst noch) von der Teilfolge (n') abhängt. Sei $T_Z := \{t \in [-a, a] : \pi_t \text{ stetig in } Z \text{ f.s.}\}$ für eine ZV Z in $D[-a, a]$. Es gilt: T_Z dicht in $[-a, a]$, da $\overline{T_Z}$ höchstens abzählbar, vgl. Billingsley [Bil99], p.138. Für $T = \{t_1, \dots, t_k\} \subseteq [-a, a]$ gilt: π_T stetig P_Z f.ü. $\Leftrightarrow T \subseteq T_Z$. Somit liefert CMT:

$$(\pi_T(L_{n''}), Y_{n''}) \xrightarrow{\mathcal{L}} (\pi_T(L'), Y') \quad \forall T \subseteq T_{L'}.$$

Andererseits folgt aus (2) (angewendet auf Teilteilfolge (n'')):

$$\begin{aligned} & (\pi_T(L_{n''}), Y_{n''}) \xrightarrow{\mathcal{L}} (\pi_T(L), Y) \quad \forall T \subseteq T_L \\ \Rightarrow & (iv) (\pi_T(L'), Y') \stackrel{\mathcal{L}}{=} (\pi_T(L), Y) \quad \forall T \subseteq T_{L'} \cap T_L \end{aligned}$$

Sei

$$\mathcal{F}_1 := \{\pi_T^{-1}(A) : A \in \mathcal{B}(\mathbb{R}^{|T|}), T \subseteq T_{L'} \cap T_L \text{ endlich}\}.$$

Wegen $\overline{T_{L'} \cap T_L} = \overline{T_{L'}} \cup \overline{T_L}$ höchstens abzählbar, ist $T_0 := T_{L'} \cap T_L$ dicht in $[-a, a]$. Gemäß Theorem 12.5 (iii) in Billingsley [Bil99] gilt:

$$\begin{aligned} \mathcal{B}_s(D) = \sigma(\mathcal{F}_1) & \Rightarrow \\ \mathcal{B}(D \times S) = \mathcal{B}_s(D) \otimes \mathcal{B}_d(S) & \quad \text{gemäß 1.3.12 in Gänsler und Stute [GS77]} \\ = \sigma(\mathcal{F}_1 \times \mathcal{B}_d(S)) & \quad \text{gemäß 22.1 in Bauer [Bau90]} \end{aligned}$$

Die Gleichung (iv) bedeutet, dass die Verteilungen $P_{(L', Y')}$ und $P_{(L, Y)}$ auf $\mathcal{F}_1 \times \mathcal{B}_d(S)$ übereinstimmen. Letzteres ist aber $\cap$ -stabil und erzeugt die Borel- σ -Algebra $\mathcal{B}(D \times S) \Rightarrow (L', Y') \stackrel{\mathcal{L}}{=} (L, Y)$. Dann zeigen (iii) und das Teilfolgenprinzip (Theorem 2.6 in Billingsley, 1999), dass (ii) gilt. Da $[0, \infty) \times cl(B)$ abgeschlossen (in der Produkttopologie) und da $(\Lambda_x, Id) : D \times S \rightarrow \mathbb{R} \times S, (f, y) \mapsto (\Lambda_x(f), y)$ stetig $P_{(L, Y)}$ -f.ü., liefern das CMT und das Portmanteau-Theorem für alle $a > |x|$:

$$\limsup_{n \rightarrow \infty} P\left(\sigma_n \leq x, Y_n \in B\right) \leq P\left(\underbrace{\Lambda_x(L) \geq 0}_{\substack{\text{vgl. Bew. von 4.9(5)} \\ \subseteq \{\sigma_{\min} \leq x\}}}, Y \in cl(B)\right) + \underbrace{o(1)}_{\rightarrow 0, n \rightarrow \infty}$$

$\Rightarrow$ (5), da $B \subseteq cl(B) = Int(B) \cup \partial B \subseteq B \cup \partial B$ und $P(Y \in \partial B) = 0$. Zum Nachweis von (4) beachte:

$$(\Omega \setminus A) \cap E = E \setminus (A \cap E) \quad \forall A, E \subseteq \Omega,$$

somit:

$$\begin{aligned} \liminf_{n \rightarrow \infty} P(\underbrace{\sigma_n \leq x}_{\Omega \setminus \{\sigma_n > x\}}, Y_n \in B) &= \liminf_{n \rightarrow \infty} [P(Y_n \in B) - P(\sigma_n > x, Y_n \in B)] \\ &\geq \liminf_{n \rightarrow \infty} P(Y_n \in B) - \limsup_{n \rightarrow \infty} P(\sigma_n > x, Y_n \in B). \end{aligned}$$

$$(2) \Rightarrow Y_n \xrightarrow{\mathcal{L}} Y \stackrel{\text{Portmanteau}}{\Rightarrow} \liminf_{n \rightarrow \infty} P(Y_n \in B) \geq P(Y \in B),$$

da B Y -randlos.

Ferner wie oben mit Lemma 4.10(2)

$$\limsup_{n \rightarrow \infty} P(\sigma_n > x, Y_n \in B) \leq P(\underbrace{\Lambda_x(L) \leq 0}_{\subseteq \{\sigma_{\max} \geq x\}}, Y \in B).$$

$\Rightarrow$ Behauptung. □

Folgerung 4.15. *Gilt in Satz 4.14 die Bedingung (U), d.h. $\sigma_{\min} = \tau = \sigma_{\max}$ f.s., so:*

$$\lim_{n \rightarrow \infty} P(\sigma_n \leq x, Y_n \in B) = P(\tau \leq x, Y \in B)$$

$\forall x \in C(F_\tau), \forall B$ Y -randlos.

Beweis: $0 \leq P(\sigma_{\max} = x, Y \in B) \leq P(\sigma_{\max} = x) \stackrel{(U)}{=} P(\sigma_{\min} = x) = P(\tau = x) = 0$, da $x \in C(F_\tau)$. $\stackrel{\text{Satz 4.14}}{\Rightarrow} F_n(x) := P(\sigma_n \leq x, Y_n \in B) \rightarrow F(x) := P(\tau \leq x, Y \in B), \forall x \in \mathbb{X} \cap C(F_\tau)$. Anwendung von 4.12(1) auf F_n, F und $D = \mathbb{X} \cap C(F_\tau)$ liefert die Behauptung. □

Gilt in Satz 4.9 nur (2) und **nicht** (U), so ist die Schlussfolgerung, z.B.

$$\operatorname{argmax}(L_n) \xrightarrow{\mathcal{L}} \operatorname{argmax}(L) \tag{4.3}$$

zwar verführerisch, aber nicht gesichert. In der Tat lassen sich Beispiele angeben, bei denen Satz 4.9(1)-(3) erfüllt sind, aber (4.3) **nicht** gilt.

Kapitel 5

Erste Anwendung: Regression mit Sprung

Erinnere an Abschnitt 1.2: (X_i, Y_i) , $i \geq 1$, $\sim (X, Y)$ mit

(1) $X \sim F$ und Q Verteilung induziert durch VF F ,

$$(2) P(Y = y|X = x) = \begin{cases} m(x), & y = 1 \\ 1 - m(x), & y = 0 \end{cases}$$

$$(3) m(x) = a1_{\{x \leq \theta\}} + b1_{\{x > \theta\}}, \quad a \neq b.$$

Sei $\mu := \delta_0 + \delta_1$ das Zählmaß auf $\{0, 1\}$. Es folgt $\forall A \subseteq \mathcal{B}(\mathbb{R}), B \subseteq \{0, 1\}$:

$$\begin{aligned} P((X, Y) \in A \times B) &= P(X \in A, Y \in B) \\ &= \int_A P(Y \in B|X = x)Q(dx) \\ &\stackrel{(2)}{=} \int_A \text{Bin}(1, m(x))(B)Q(dx) \\ &= \int_A \int_B m(x)^y (1 - m(x))^{1-y} \mu(dy) Q(dx) \end{aligned}$$

$\stackrel{\text{Fubini}}{\Rightarrow} (X, Y)$ hat $Q \otimes \mu$ -Dichte

$$f_{(X,Y)}(x, y) = m(x)^y (1 - m(x))^{1-y}, \quad x \in \mathbb{R}, y \in \{0, 1\}.$$

$$\Rightarrow \text{Likelihood-Funktion} = \prod_{i=1}^n m(X_i)^{Y_i} (1 - m(X_i))^{1-Y_i}.$$

Daraus erhält man die Log-Likelihood-Funktion:

$$\begin{aligned}
 & \sum_{i=1}^n Y_i \log(m(X_i)) + (1 - Y_i) \log(1 - m(X_i)) \\
 & \stackrel{(3)}{=} \sum_{i=1}^n 1_{\{X_i \leq \theta\}} [Y_i \log(a) + (1 - Y_i) \log(1 - a)] + \\
 & \quad + \sum_{i=1}^n \underbrace{1_{\{X_i > \theta\}}}_{=1-1_{\{X_i \leq \theta\}}} [Y_i \log(b) + (1 - Y_i) \log(1 - b)] \\
 & = \sum_{i=1}^n 1_{\{X_i \leq \theta\}} \left[Y_i \log \frac{a}{b} + (1 - Y_i) \log \frac{1-a}{1-b} \right] + \underbrace{\sum_{i=1}^n Y_i \log(b) + (1 - Y_i) \log(1 - b)}_{\substack{\text{hängt nicht von } \theta \text{ ab} \\ \text{(und auch nicht von } a \text{)}}}
 \end{aligned}$$

Daraus ergibt sich der Maximum-Likelihood-Schätzer:

$$\begin{aligned}
 \hat{\theta}_n &:= \operatorname{argmax}_{t \in \mathbb{R}} S_n(t) \text{ für} \\
 S_n(t) &:= \frac{1}{n} \sum_{i=1}^n 1_{\{X_i \leq t\}} [\alpha Y_i + \beta], \quad t \in \mathbb{R} \text{ mit} \\
 \alpha &:= \log\left(\frac{a}{b} \frac{1-b}{1-a}\right) \quad \text{und} \quad \beta := \log\left(\frac{1-a}{1-b}\right).
 \end{aligned}$$

S_n ist Treppenfunktion mit Sprüngen in $X_1, \dots, X_n$ (und sonst nicht). Wegen $\alpha + \beta = \log \frac{a}{b}$ und $\beta = \log \frac{1-a}{1-b}$ wird entweder nach oben oder nach unten gesprungen, daher ist $A(S_n)$ eine disjunkte Vereinigung von bestimmten Intervallen vom Typ $[X_{i-1:n}, X_{i:n}]$. Betrachte die empirischen Prozesse in t :

$$\begin{aligned}
 H_n(t) &:= \frac{1}{n} \sum_{i=1}^n 1_{\{X_i \leq t\}} Y_i \quad (\text{markierte empirische VF}), \\
 F_n(t) &:= \frac{1}{n} \sum_{i=1}^n 1_{\{X_i \leq t\}} \quad (\text{empirische VF}).
 \end{aligned}$$

Damit kann $S_n(t)$ auch geschrieben werden als

$$S_n(t) = \alpha H_n(t) + \beta F_n(t),$$

und wir erhalten

$$S(t) := E(S_n(t)) = \alpha H(t) + \beta F(t)$$

mit

$$H(t) = E(H_n(t)) = \int_{-\infty}^t m(x) F(dx),$$

wobei sich das letzte Gleichheitszeichen durch Bedingen bzgl X ergibt.

5.1 Starke Konsistenz des MLE

Lemma 5.1. (1) Angenommen F ist strikt monoton wachsend in einer (bel. kleinen) Umgebung $B_r(\theta) = [\theta - r, \theta + r]$, $r > 0$, dann folgt

$$\theta = \underset{t \in \mathbb{R}}{\operatorname{argmax}} S(t)$$

ist eindeutige Maximalstelle von S .

(2) Falls zusätzlich eine Konstante $\underline{L} = \underline{L}(r) > 0$ existiert mit

$$|F(\theta) - F(t)| \geq \underline{L}|\theta - t| \quad \forall t \in B_r(\theta), \quad (5.1)$$

so gilt:

$$S(\theta) - S(t) \geq L|\theta - t| \quad \forall t \in B_r(\theta), \quad (5.2)$$

wobei

$$L := L(r) := \underline{L} \min\{\alpha a + \beta, -(\alpha b + \beta)\} \geq 0.$$

Beweis: Aus $t \leq \theta$ folgt $H(t) = aF(t)$ und aus $t > \theta$ folgt $H(t) = aF(\theta) + b(F(t) - F(\theta))$. Daraus folgt:

$$S(t) = \begin{cases} (\alpha a + \beta)F(t), & t \leq \theta \\ (\alpha b + \beta)F(t) + \alpha(a - b)F(\theta), & t > \theta \end{cases} \quad (5.3)$$

Kann man also zeigen, dass

$$\alpha a + \beta > 0 > \alpha b + \beta, \quad (5.4)$$

so folgen (1) und (2). Dazu:

$$\begin{aligned} \alpha a + \beta &= a \cdot \underbrace{\log \frac{a}{b} \frac{1-b}{1-a}}_{=\log(\frac{a}{b}) - \log(\frac{1-a}{1-b})} + \log \frac{1-a}{1-b} \\ &= a \log \frac{a}{b} + (1-a) \log \frac{1-a}{1-b} \\ &= K(Q, P), \end{aligned}$$

wobei $K(Q, P)$ gerade die Kullback-Leibler-Info zu $Q = \operatorname{Bin}(1, a)$ und $P = \operatorname{Bin}(1, b)$ ist.

Klar: $a \neq b \Rightarrow Q \neq P$. Damit gilt bekanntlich $K(Q, P) > 0$. Analog kann $\alpha b + \beta = -K(P, Q) < 0$ gezeigt werden. Daraus folgt (5.4) und damit folgen auch (1) und (2). $\square$

Lemma 5.2. F stetige VF $\Rightarrow \|S_n - S\| \rightarrow 0$ f.s.

Beweis: $\|S_n - S\| \leq |\alpha| \|H_n - H\| + |\beta| \|F_n - F\|$. $\Omega_t := \{H_n(t) \rightarrow H(t)\}$ P -Einsmenge $\forall t \in \mathbb{R}$ gemäß SGGZ $\Rightarrow \Omega_0 := \bigcap_{t \in \mathbb{Q}} \Omega_t$ P -Einsmenge, da $\mathbb{Q}$ abzählbar.

Schließlich:

$$\Omega_0 \subseteq \{\|H_n - H\| \rightarrow 0\}$$

gemäß Satz 4.12(2), da H_n und H monoton und

$$H(t) = 1_{\{t \leq \theta\}} aF(t) + 1_{\{t > \theta\}} [aF(\theta) + b(F(t) - F(\theta))]$$

stetig und beschränkt.

$$\Rightarrow \|H_n - H\| \rightarrow 0 \quad \text{f.s.}$$

Analog: $\|F_n - F\| \rightarrow 0$ f.s. (so genannter Satz von Glivenko + Cantelli in Glivenko [Gli33]; gilt sogar für beliebige F) $\Rightarrow$ Behauptung. $\square$

Bemerkung 5.3. Lemma 5.2 gilt auch für beliebige VF F .

Beweis: Gemäß Quantiltransformation von Rosenblatt [Ros52] gilt

$$(X, Y) \stackrel{\mathcal{L}}{=} (F^{-1}(U), m^{-1}(V|F^{-1}(U))),$$

wobei (U, V) gleichverteilt auf $(0, 1)^2$ und

$$\begin{aligned} m(y|x) &:= P(Y \leq y | X = x) \\ &= P(\text{Bin}(1, m(x)) \leq y), \quad x, y \in \mathbb{R} \\ &= \begin{cases} 0, & y < 0 \\ 1 - m(x), & 0 \leq y < 1 \\ 1, & y \geq 1 \end{cases} \Rightarrow \\ m^{-1}(v|x) &= \begin{cases} 0, & 0 \leq v \leq 1 - m(x) \\ 1, & 1 - m(x) < v \leq 1 \end{cases}, \quad 0 \leq v \leq 1. \end{aligned}$$

Somit o.E. $(X_i, Y_i)_{i \geq 1} := (F^{-1}(U_i), m^{-1}(V_i|F^{-1}(U_i)))_{i \geq 1}$, (U_i, V_i) i.i.d. $\sim (U, V)$.

$$\Rightarrow H_n(t) = \overline{H}_n(F(t)),$$

$$\begin{aligned} \overline{H}_n(u) &:= \frac{1}{n} \sum_{i=1}^n 1_{\{U_i \leq u\}} m^{-1}(V_i|F^{-1}(U_i)) \\ &\rightarrow \overline{H}(u) \quad \text{f.s. für jedes feste } u \text{ aus } [0, 1]. \end{aligned}$$

$$\overline{H}(u) = E(1_{\{U \leq u\}} m^{-1}(V|F^{-1}(U)))$$

$$\begin{aligned}
 &= \int_0^u E(m^{-1}(V|F^{-1}(U)|U=x))dx \\
 &= \int_0^u E(m^{-1}(V|F^{-1}(x)))dx \quad \text{wegen } U \text{ und } V \text{ unabhängig} \\
 &= \int_0^u E[1_{\{1-m(F^{-1}(x)) < V \leq 1\}}]dx \\
 &= \int_0^u m(F^{-1}(x))dx, \quad 0 \leq u \leq 1 \quad \text{wegen } V \sim U(0,1).
 \end{aligned}$$

$$\begin{aligned}
 \Rightarrow \overline{H}(F(t)) &= \int 1_{\{0 \leq x \leq F(t)\}} m(F^{-1}(x))dx \\
 &= \int 1_{\{F^{-1}(x) \leq t\}} m(F^{-1}(x))dx, \quad \text{da } F(t) \geq x \Leftrightarrow t \geq F^{-1}(x) \\
 &= \int_{-\infty}^t m(y)F(dy) \quad \text{Trafo} \\
 &= H(t) \quad \Rightarrow
 \end{aligned}$$

$$\begin{aligned}
 \sup_{t \in \mathbb{R}} |H_n(t) - H(t)| &= \sup_{t \in \mathbb{R}} |\overline{H}_n(F(t)) - \overline{H}(F(t))| \\
 &\leq \sup_{u \in [0,1]} |H_n(u) - H(u)| \\
 &\rightarrow 0 \quad f.s.,
 \end{aligned}$$

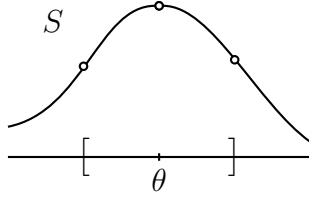
da $\overline{H}$ stetig und beschränkt bzw. $[0, 1]$ kompakt. (analoge Anwendung von Satz 4.12 wie im Beweis von Lemma 5.2) $\Rightarrow$ Behauptung. $\square$

Wir erhalten starke Konsistenz von $(\hat{\theta}_n)$.

Satz 5.4. *F stetig auf $B_r(\theta)$ für ein $r > 0$ und dort streng monoton wachsend*
 $\Rightarrow$

$$\hat{\theta}_n \rightarrow \theta \quad f.s., \quad n \rightarrow \infty \quad \forall \theta \in \mathbb{R}, \quad \forall a \neq b.$$

Beweis: Wegen (5.4), (5.3) und Voraussetzungen an F ist S stetig auf $B_r(\theta)$ und strikt $\nearrow$ bzw. $\searrow$ auf $[\theta - r, \theta]$ bzw. $[\theta, \theta + r]$, insbesondere gilt $A(S) = \{\theta\} \Rightarrow$



$$\begin{aligned} \Rightarrow a(\varepsilon) &:= \max\{S(t) : |t - \theta| > \varepsilon\} \\ &= \max\{S(\theta - \varepsilon), S(\theta + \varepsilon)\} \end{aligned}$$

existiert $\forall \varepsilon \in (0, r)$.

$$\begin{aligned} \Rightarrow b(\varepsilon) &= \frac{1}{3}(S(\theta) - S(\theta - \varepsilon)) \quad \text{oder} \\ b(\varepsilon) &= \frac{1}{3}(S(\theta) - S(\theta + \varepsilon)) \end{aligned}$$

Damit $b(\varepsilon) > 0 \forall 0 < \varepsilon < r \Rightarrow$ Behauptung mit Satz 4.6 und Bemerkung 5.3. $\square$

Numerische Bestimmung des MLE $\hat{\theta}_n$: Seien $X_{1:n} \leq \dots \leq X_{n:n}$ die Ordnungsgrößen zu $X_1, \dots, X_n$ und $Y_{[1:n]}, \dots, Y_{[n:n]}$ die Konkomitanten, d.h. die i-te Konkomitante erfüllt

$$Y_{[i:n]} = Y_j \Leftrightarrow X_{i:n} = X_j.$$

Damit ist

$$\hat{\theta}_n = X_{\hat{l}:n},$$

wobei

$$\hat{l} := \operatorname{argmax}_{1 \leq l \leq n} \sum_{i=1}^l (\alpha Y_{[i:n]} + \beta)$$

der kleinste Index ist, der die Partialsumme maximiert. Beachte, $\hat{\theta}_n$ ist echte Maximalstelle von S_n , d.h. $S_n(\hat{\theta}_n) = \max_{t \in \mathbb{R}} S_n(t)$.

5.2 Fehlerabschätzung

Ziel: Herleitung einer oberen Schranke für die Fehlerwahrscheinlichkeit $P(|\hat{\theta}_n - \theta| \geq d)$, $d > 0$.

Satz 5.5. Angenommen für ein (beliebig kleines) $r > 0$ gilt:

- (1) F strikt monoton wachsend auf $B_r(\theta)$.
- (2) Es existieren Konstanten $\underline{L} = \underline{L}(r)$ und $\bar{L} = \bar{L}(r)$ mit:

$$\underline{L}|s - t| \leq |F(s) - F(t)| \leq \bar{L}|s - t| \quad \forall s, t \in B_r(\theta).$$

Dann existiert Konstante $C = C(r)$ mit

$$\limsup_{n \rightarrow \infty} P\left(n|\hat{\theta}_n - \theta| \geq x\right) \leq Cx^{-1} \quad \forall x > 0. \quad (5.5)$$

Insbesondere liefert $x \rightarrow \infty$:

$$\hat{\theta}_n - \theta = O_P(n^{-1}). \quad (5.6)$$

Beweis: Sei $d > 0$ beliebig $\Rightarrow$

$$(1) \quad P(|\hat{\theta}_n - \theta| \geq d) \leq \underbrace{P(d \leq |\hat{\theta}_n - \theta| \leq r)}_{=0 \text{ für } d > r} + \underbrace{P(|\hat{\theta}_n - \theta| > r)}_{=: a_n(r) \rightarrow 0 \quad \forall r > 0 \text{ gemäß Satz 5.4}}.$$

Daher betrachte nur für $d \leq r$:

$$(2) \quad \{d \leq |\hat{\theta}_n - \theta| \leq r\} \subseteq \bigcup_{d \leq |u - \theta| \leq r} \{S_n(u) \geq S_n(\theta)\},$$

denn sei $\omega \in$ linker Menge und angenommen $\omega \notin$ rechter Menge, dann folgt, dass $S_n(u) < S_n(\theta) \quad \forall d \leq |u - \theta| \leq r$, also insbesondere für $u = \hat{\theta}_n$ folgt $S_n(\hat{\theta}_n) < S_n(\theta)$, was einen Widerspruch darstellt.

Für

$$(3) \quad \bigcup_{d \leq |u - \theta| \leq r} \{\dots\} \subseteq \bigcup_{d \leq u - \theta \leq r} \{\dots\} \cup \bigcup_{d \leq \theta - u \leq r} \{\dots\} \\ =: E_1 \cup E_2$$

mit

$$\begin{aligned} E_1 &= \bigcup_{d \leq v \leq r} \{S_n(v + \theta) \geq S_n(\theta)\} \\ &= \bigcup_{d \leq v \leq r} \{S_n(v + \theta) - S(v + \theta) - (S_n(\theta) - S(\theta)) \geq \underbrace{S(\theta) - S(v + \theta)}_{\substack{(5.2) \\ \geq L \cdot v}}\} \\ &\subseteq \left\{ \sup_{d \leq v \leq r} \frac{|S_n(v + \theta) - S(v + \theta) - (S_n(\theta) - S(\theta))|}{v} \geq L \right\} \end{aligned}$$

folgt mit (2):

$$(4) \quad P(d \leq |\hat{\theta}_n - \theta| \leq r) \leq P(E_1) + P(E_2)$$

mit: (da $S_n = \alpha H_n + \beta F_n$, $S = \alpha H + \beta F$, α und $\beta \neq 0$)

$$\begin{aligned} (5) \quad P(E_1) &\leq P\left(\sup_{d \leq v \leq r} \frac{|H_n(\theta + v) - H(\theta + v) - (H_n(\theta) - H(\theta))|}{v} \geq \frac{L}{2|\alpha|}\right) + \\ &\quad + P\left(\sup_{d \leq v \leq r} \frac{|F_n(\theta + v) - F(\theta + v) - (F_n(\theta) - F(\theta))|}{v} \geq \frac{L}{2|\beta|}\right) \\ &=: A_1 + A_2. \end{aligned}$$

Für die Behandlung von A_1 sei

$$T := T_H := \{t \in \mathbb{R} : H(t) < 1\}.$$

Definiere $\forall n \in \mathbb{N} \forall t \in T$:

$$(6) \quad M_n(t) := \frac{H_n(t) - H(t)}{1 - H(t)} \quad \text{und} \quad \mathcal{F}_n(t) := \sigma(H_n(u) : u \leq t).$$

Da unser $H < 1$, ist hier $T_H = \mathbb{R}$. Man kann zeigen, dass für jedes feste $n \in \mathbb{N}$

$$(M_n(t), \mathcal{F}_n(t) : t \in T_H)$$

ein zentriertes Martingal ist. Wegen (6) ist $H_n - H = (1 - H)M_n$, und es folgt:

$$\begin{aligned} & v^{-1} \{H_n(\theta + v) - H(\theta + v) - (H_n(\theta) - H(\theta))\} \stackrel{(6)}{=} \\ & v^{-1} \{(1 - H(\theta + v))M_n(\theta + v) - (1 - H(\theta))M_n(\theta)\} = \\ & v^{-1} \{H(\theta) - H(\theta + v)\}M_n(\theta + v) + v^{-1}(1 - H(\theta))\{M_n(\theta + v) - M_n(\theta)\} \end{aligned}$$

$$\begin{aligned} \Rightarrow A_1 & \leq P \left(\sup_{d \leq v \leq r} \frac{|H(\theta + v) - H(\theta)|}{v} |M_n(\theta + v)| \geq \frac{L}{4|\alpha|} \right) \\ & \quad + P \left(\sup_{d \leq v \leq r} \frac{|M_n(\theta + v) - M_n(\theta)|}{v} \geq \underbrace{\frac{L}{4|\alpha|(1 - H(\theta))}}_{=: \lambda} \right) \\ & =: A_{11} + A_{12}. \end{aligned}$$

Für die Abschätzung von A_{11} beachte:

$$\begin{aligned} (7) \quad |H(\theta + v) - H(\theta)| &= \int_{\theta}^{\theta+v} m(x) F(dx) \\ &= b \{F(\theta + v) - F(\theta)\} \stackrel{\text{Vor. 5.5(2)}}{\leq} b \bar{L} v \quad \forall d \leq v \leq r. \end{aligned}$$

Außerdem ist $(M_n(\theta + v), \mathcal{F}_n(\theta + v) : d \leq v \leq r)$ ein zentriertes Martingal mit rcl Pfaden. Daher können wir die Doob'sche Martingalungleichung (vgl. Shorack und Wellner [SW86], p. 874) anwenden:

$$\begin{aligned} (8) \quad A_{11} & \stackrel{(7)}{\leq} P \left(\sup_{d \leq v \leq r} |M_n(\theta + v)| \geq \frac{L}{4|\alpha|b\bar{L}} \right) \\ & \stackrel{Doob}{\leq} 16\alpha^2 b^2 \bar{L}^2 L^{-2} \underbrace{E[M_n(\theta + r)^2]}_{\stackrel{(6)}{=} (1 - H(\theta + r))^{-2} \text{Var}(H_n(\theta + r))}. \end{aligned}$$

Zur Berechnung der Varianz, beachte:

$$nH_n(t) = \sum_{i=1}^n \underbrace{1_{\{X_i \leq t\}} Y_i}_{\substack{\text{Bernoulli-Var, da } Y_i \in \{0,1\}, \\ \text{mit Treff.-w. } p=H(t)}} \sim \text{Bin}(n, H(t))$$

$$\begin{aligned} \Rightarrow \text{Var} H_n(\theta + r) &= n^{-2} \text{Var}[\text{Bin}(n, H(\theta + r))] \\ &= n^{-1} H(\theta + r)(1 - H(\theta + r)) \Rightarrow \\ (9) \quad A_{11} &\leq C_1 n^{-1} \end{aligned}$$

mit $C_1 = 16\alpha^2 b^2 \bar{L}^2 L^{-2} H(\theta + r)(1 - H(\theta + r))^{-1}$.

Zur Abschätzung von A_{12} betrachte

$$\overline{M}_n(v) := M_n(\theta + v) - M_n(\theta), \quad d \leq v \leq r \Rightarrow$$

$(\overline{M}_n(v), \mathcal{F}(\theta + v) : d \leq v \leq r)$ ist zentriertes Martingal.

Sei

$$\{v_k := d + (r - d)k2^{-m} : 0 \leq k \leq 2^m\}, \quad m \in \mathbb{N},$$

eine dyadische Zerlegung von $[d, r]$. Setze

$$S_k := \overline{M}_n(v_k), \quad \mathcal{G}_k := \mathcal{F}_n(\theta + v_k), \quad 1 \leq k \leq 2^m, \quad S_0 := 0, \quad \mathcal{G}_0 := \{\emptyset, \Omega\}.$$

$(S_k, \mathcal{G}_k : 0 \leq k \leq 2^m)$ ist (diskretes) zentriertes Martingal $\Rightarrow$

$$\begin{aligned} (10) \quad A_{12} &= P\left(\sup_{d \leq v \leq r} \frac{|\overline{M}_n(v)|}{v} \geq \overbrace{\frac{\lambda}{2}}^{> \lambda/2}\right) \\ &\leq P\left(\sup_{d \leq v \leq r} \frac{|\overline{M}_n(v)|}{v} > \frac{\lambda}{2}\right) \\ &= P\left(\sup_{m \geq 1} \max_{1 \leq k \leq 2^m} \frac{|\overline{M}_n(v_k)|}{v_k} > \frac{\lambda}{2}\right) \\ &\text{da } v \mapsto \frac{|\overline{M}_n(v)|}{v} \text{ rcll auf } [d, r] \\ &= P\left(\bigcup_{m \geq 1} \underbrace{\left\{ \max_{1 \leq k \leq 2^m} \frac{|S_k|}{v_k} > \frac{\lambda}{2} \right\}}_{\text{monoton wachsend}}\right) \\ &= \lim_{m \rightarrow \infty} P\left(\underbrace{\max_{1 \leq k \leq 2^m} \frac{|S_k|}{v_k} > \frac{\lambda}{2}}_{\substack{\leq \frac{4\lambda^{-2}}{2} \sum_{1 \leq k \leq 2^m} v_k^{-2} E[(S_k - S_{k-1})^2]}}\right) \quad \begin{array}{l} \sigma\text{-Stetigkeit} \\ \text{von unten} \end{array} \\ &\text{gemäß Hajek's Martingal-Ungleichung } (\rightarrow \text{z.B. Gänssler und Stute [GS77]}). \blacksquare \end{aligned}$$

$$\begin{aligned}
 (11) \quad E[(S_k - S_{k-1})^2] &= E(S_k^2) - 2 \underbrace{E(S_{k-1}S_k)} + E(S_{k-1}^2) \\
 &= E(S_{k-1} \underbrace{E(S_k | \mathcal{G}_{k-1}))} = E(S_{k-1}^2) \\
 &= S_{k-1}, \text{ da } S_k \text{ Martingal} \\
 &= E(S_k^2) - E(S_{k-1}^2) = E[\overline{M}_n(v_k)^2] - E[\overline{M}_n(v_{k-1})^2],
 \end{aligned}$$

wobei die letzte Gleichung nur für alle $k \geq 2$ gilt wegen der Definition von S_k .

$$\begin{aligned}
 (12) \quad E[\overline{M}_n(v)^2] &\stackrel{\text{Def.}}{=} E[(M_n(\theta + v) - M_n(\theta))^2] \\
 &= E[M_n(\theta + v)^2] - E[M_n(\theta)^2], \text{ da } (M_n(t) : t \in \mathbb{R}) \text{ Martingal.}
 \end{aligned}$$

Weiter mit (6):

$$\begin{aligned}
 E[M_n(t)^2] &\stackrel{(6)}{=} (1 - H(t))^{-2} \text{Var} H_n(t) \stackrel{\text{s.o.}}{=} \\
 &= n^{-1} H(t) (1 - H(t))^{-1} \stackrel{(11)+(12)}{\Rightarrow} \\
 E(S_k^2) - E(S_{k-1}^2) &= E(M_n(\theta + v_k)^2) - E(M_n(\theta + v_{k-1})^2) \quad \forall k \geq 2 \\
 &= n^{-1} \left\{ \frac{H(\theta + v_k)}{1 - H(\theta + v_k)} - \frac{H(\theta + v_{k-1})}{1 - H(\theta + v_{k-1})} \right\} \\
 &= n^{-1} \frac{H(\theta + v_k) - H(\theta + v_{k-1})}{\underbrace{(1 - H(\theta + v_k))}_{\leq H(\theta+r)} \underbrace{(1 - H(\theta + v_{k-1}))}_{\leq H(\theta+r)<1}} \quad \boxed{\frac{\frac{x}{1-x} - \frac{y}{1-y}}{\frac{x-y}{(1-x)(1-y)}} =} \\
 &\leq \underbrace{(1 - H(\theta + r))^{-2}}_{=:D} n^{-1} \underbrace{\{H(\theta + v_k) - H(\theta + v_{k-1})\}}_{\substack{\theta+v_k \\ \theta+v_{k-1}} \int \underbrace{m}_{\leq 1} dF} \\
 &\leq D n^{-1} \underbrace{\{F(\theta + v_k) - F(\theta + v_{k-1})\}}_{\leq \overline{L}(v_k - v_{k-1}) \text{ wegen Vor. 5.5(2)}} \\
 &\leq D \overline{L} n^{-1} (v_k - v_{k-1}).
 \end{aligned}$$

Für $k = 1$ ist $S_{k-1} = S_0 = 0$, also

$$\begin{aligned}
 E(S_k^2) - E(S_{k-1}^2) &= E[M_n(\theta + v_1)^2] - E[M_n(\theta)^2] \\
 &\leq D \overline{L} n^{-1} v_1 \\
 &= D \overline{L} n^{-1} (v_1 - v_0) + D \overline{L} n^{-1} v_0.
 \end{aligned}$$

Da $v_0 = d < v_1$, folgt:

$$A_{12} \leq 4D \overline{L} \lambda^{-1} n^{-1} \lim_{m \rightarrow \infty} \sum_{1 \leq k \leq 2^m} v_k^{-2} (v_k - v_{k-1}) + 4D \overline{L} n^{-1} d^{-1}$$

$$\begin{aligned}
 &= 4D\bar{L}\lambda^{-2}n^{-1} \underbrace{\int_d^r v^{-2}dv}_{=\frac{1}{d}-\frac{1}{r}\leq\frac{1}{d}} + 4D\bar{L}n^{-1}d^{-1} \\
 \Rightarrow A_{12} &\leq C_2n^{-1}d^{-1} \stackrel{(10)}{\Rightarrow} \\
 (13) \quad A_1 &\leq C_1n^{-1} + C_2n^{-1}d^{-1}.
 \end{aligned}$$

Zur Behandlung von A_2 beachte:

$$\left(\frac{F_n(t) - F(t)}{1 - F(t)}, \sigma(F_n(s) : s \leq t) : t \in T_F \right)$$

ist zentriertes Martingal mit rcl Pfaden $\forall n \in \mathbb{N}$ (siehe z.B. Gaenssler und Stute [GS87]).

Analoges Vorgehen liefert:

$$\begin{aligned}
 A_2 &\leq C_3n^{-1} + C_4n^{-1}d^{-1} \stackrel{(5)}{\Rightarrow} \\
 P(E_1) &\leq C_5n^{-1} + C_6n^{-1}d^{-1}.
 \end{aligned}$$

Es ist

$$E_2 = \bigcup_{-r \leq v \leq -d} \{S_n(\theta + v) \leq S_n(\theta)\}.$$

Daher lässt sich $P(E_2)$ analog abschätzen, wenn man berücksichtigt, dass

$$M_n(t) = \frac{H_n(t) - H(t)}{H(t)} \quad \text{bzgl.}$$

$$\mathcal{F}_n(t) = \sigma(H_n(u) : u \geq t), \quad t \in T_{1-H},$$

ein reverses Martingal bildet. Man erhält

$$\begin{aligned}
 P(E_2) &\leq C_7n^{-1} + C_8n^{-1}d^{-1} \stackrel{(1)-(3)}{\Rightarrow} \\
 P(|\hat{\theta}_n - \theta| \geq d) &\leq C_9n^{-1} + C_{10}n^{-1}d^{-1} + a_n,
 \end{aligned}$$

wobei $a_n = a_n(r) = o(1)$, $n \rightarrow \infty$. Sei $d = xn^{-1}$, $x > 0 \Rightarrow (5.5)$. □

5.3 Verteilungskonvergenz

Ziel: Grenzwertaussagen für die VF von $n(\hat{\theta}_n - \theta)$ für $n \rightarrow \infty$. Dazu: Gemäß (4.2) gilt

$$n(\hat{\theta}_n - \theta) = \operatorname{argmax}_{t \in \mathbb{R}} L_n(t), \quad (5.7)$$

wobei

$$L_n(t) := n \left\{ S_n \left(\theta + \frac{t}{n} \right) - S_n(\theta) \right\}.$$

Anwendung von Satz 4.9 erfordert:

- (1) $n(\hat{\theta}_n - \theta) = O_P(1)$
- (2) $L_n \xrightarrow{\mathcal{L}} L$ in $D[-a, a] \forall a > 0$
- (3) $\emptyset \neq A(L)$ abgeschlossen und beschränkt
 $(\Rightarrow \sigma_{\min} = \operatorname{argmax}(L) \text{ und } \sigma_{\max} = \operatorname{Argmax}(L) \text{ existieren})$

Wegen (5.6) ist (1) erfüllt. Bleiben (2) und (3). Zu (2): Es stellt sich heraus, dass L ein zusammengesetzter Poisson Prozess (compound Poisson process) ist, definiert durch:

$$L(t) = \begin{cases} \sum_{i=1}^{N_1(t)} \xi_i, & t \geq 0 \\ - \sum_{i=1}^{N_2(-t)} \eta_i, & t < 0 \end{cases} \quad (5.8)$$

wobei N_1 und N_2 Poisson Prozesse sind mit Parameter $\lambda_1 = F'(\theta+)$ bzw. $\lambda_2 = F'(\theta-)$. Ferner gilt $\forall i$:

$$P(\xi_i = x) = \begin{cases} b, & x = \log \frac{a}{b} = \alpha + \beta \\ 1 - b, & x = \log \frac{1-a}{1-b} = \beta \end{cases}$$

$$P(\eta_i = x) = \begin{cases} a, & x = \log \frac{a}{b} \\ 1 - a, & x = \log \frac{1-a}{1-b} \end{cases}$$

Außerdem sind $\{N_1, N_2, \xi_i, \eta_i : i \geq 1\}$ unabhängig (Wähle den Poisson Prozess N_2 derart, dass die Pfade linksseitig stetig sind mit rechtsseitigen Limiten $\Rightarrow L$ rcll).

Lemma 5.6. $t_1, \dots, t_k \in \mathbb{R} \Rightarrow (L_n(t_1), \dots, L_n(t_k)) \xrightarrow{\mathcal{L}} (L(t_1), \dots, L(t_k))$.

Beweis: o.E. $k = 2$. Seien $s < t \in \mathbb{R}$.

1. Fall: $0 \leq s < t$. Betrachte Zuwächse

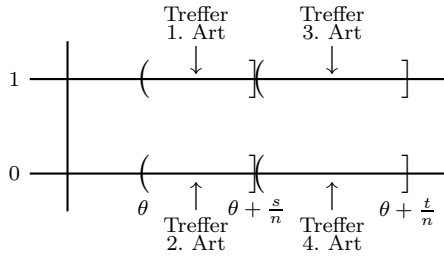
$$\begin{aligned} L_n(s) &= L_n(s) - L_n(0) \\ &= \sum_{i=1}^n 1_{\{\theta < X_i \leq \theta + \frac{s}{n}\}} [\alpha Y_i + \beta] && \text{und} \\ L_n(t) - L_n(s) &= \sum_{i=1}^n 1_{\{\theta + \frac{s}{n} < X_i \leq \theta + \frac{t}{n}\}} [\alpha Y_i + \beta] && \Rightarrow \\ L_n(s) &= (\alpha + \beta) Z_1 + \beta Z_2 \\ L_n(t) - L_n(s) &= (\alpha + \beta) Z_3 + \beta Z_4, && \text{wobei} \end{aligned}$$

$$\begin{aligned}
 Z_1 &:= \sum_{i=1}^n 1_{\{\theta < X_i \leq \theta + \frac{s}{n}, Y_i=1\}} \\
 Z_2 &:= \sum_{i=1}^n 1_{\{\theta < X_i \leq \theta + \frac{s}{n}, Y_i=0\}} \\
 Z_3 &:= \sum_{i=1}^n 1_{\{\theta + \frac{s}{n} < X_i \leq \theta + \frac{t}{n}, Y_i=1\}} \\
 Z_4 &:= \sum_{i=1}^n 1_{\{\theta + \frac{s}{n} < X_i \leq \theta + \frac{t}{n}, Y_i=0\}}
 \end{aligned}$$

Die Zuwächse sind diskret mit Zielbereich $W = \{k(\alpha + \beta) + l\beta : k, l \in \mathbb{N}_0\}$. Seien $u, v \in W$ mit (zunächst eindeutigen) Darstellungen $u = k(\alpha + \beta) + l\beta$, $v = m(\alpha + \beta) + q\beta \Rightarrow$

$$\begin{aligned}
 P(L_n(s) = u, L_n(t) - L_n(s) = v) &\stackrel{\substack{\text{Eind. v.} \\ u, v}}{=} P(Z_1 = k, Z_2 = l, Z_3 = m, Z_4 = q) \\
 &= P(Z_1 = k, Z_2 = l, Z_3 = m, Z_4 = q, Z_5 = \underbrace{n - k - l - m - q}_{=: p}),
 \end{aligned}$$

wobei $Z_5 := n - (Z_1 + \dots + Z_4)$.



$$\begin{aligned}
 \forall n \geq k + l + m + q \Rightarrow \\
 (Z_1, \dots, Z_5) \sim \text{Mult}(n, p_1, \dots, p_5)
 \end{aligned}$$

$$\begin{aligned}
 p_1 &= P(\theta < X < \theta + \frac{s}{n}, Y = 1) = \int_{\theta}^{\theta + \frac{s}{n}} m(x) F(dx) \\
 &= b \left\{ F\left(\theta + \frac{s}{n}\right) - F(\theta) \right\} \\
 p_2 &= (1 - b) \left\{ F\left(\theta + \frac{s}{n}\right) - F(\theta) \right\} \\
 p_3 &= b \left\{ F\left(\theta + \frac{t}{n}\right) - F\left(\theta + \frac{s}{n}\right) \right\}
 \end{aligned}$$

$$p_4 = (1 - b) \left\{ F\left(\theta + \frac{t}{n}\right) - F\left(\theta + \frac{s}{n}\right) \right\}$$

$$p_5 = 1 - (p_1 + \dots + p_5) = 1 - \left\{ F\left(\theta + \frac{t}{n}\right) - F(\theta) \right\} \Rightarrow$$

$$\begin{aligned} np_1 &\rightarrow bsF'(\theta+), & np_2 &\rightarrow (1-b)sF'(\theta+) \\ np_3 &\rightarrow b(t-s)F'(\theta+), & np_4 &\rightarrow (1-b)(t-s)F'(\theta+) \\ n(p_1 + \dots + p_4) &\rightarrow tF'(\theta+) \quad \Rightarrow \end{aligned}$$

$$(1) \quad P(L_n(s) = u, L_n(t) - L_n(s) = v) = \binom{n}{k! \, l! \, m! \, q! \, p!} p_1^k \cdot p_2^l \cdot p_3^m \cdot p_4^q \cdot p_5^p$$

$$\xrightarrow{\text{vgl. Poisson'scher Grenzwertsatz}} \frac{b^{k+m}(1-b)^{l+q}}{k! \cdot l! \cdot m! \cdot q!} (\lambda_1 s)^{k+l} (\lambda_1(t-s))^{m+q} e^{-\lambda_1 t}.$$

Andererseits

$$\begin{aligned} &P(L(s) = u, L(t) - L(s) = v) \stackrel{u=k(\alpha+\beta)+l\beta}{=} \stackrel{v=m(\alpha+\beta)+q\beta}{=} \\ &= P\left(\sum_{i=1}^{k+l} \xi_i = u, N_1(s) = k+l, \sum_{i=k+l+1}^{k+l+m+q} \xi_i = v, N_1(t) - N_1(s) = m+q\right) \\ &= P\left(\sum_{i=1}^{k+l} \xi_i = u\right) \cdot P(N_1(s) = k+l) \\ &\quad \cdot P\left(\sum_{i=k+l+1}^{k+l+m+q} \xi_i = v\right) \cdot P(N_1(t) - N_1(s) = m+q), \end{aligned}$$

wobei die 1. Gleichung gilt, weil

$$\begin{aligned} \left\{ \sum_{i=1}^{N_1(s)} \xi_i = u \right\} &\subseteq \{N_1(s) = k+l\} \\ \left\{ \sum_{i=N_1(s)+1}^{N_1(t)} \xi_i = v \right\} &\subseteq \{N_1(t) - N_1(s) = m+q\}, \end{aligned}$$

und die 2. Gleichung gilt wegen der Unabhängigkeit. Da

$$P\left(\sum_{i=1}^{k+l} \xi_i = u\right) = \binom{k+l}{k} b^k (1-b)^l,$$

$$P(N_1(s) = k+l) = e^{-\lambda_1 s} \frac{(\lambda_1 s)^{k+l}}{(k+l)!},$$

$$P\left(\sum_{i=k+l+1}^{k+l+m+q} \xi_i = v\right) = \binom{m+q}{m} b^m (1-b)^q \quad \text{und}$$

$$P(N_1(t) - N_1(s) = m+q) = e^{-\lambda_1(t-s)} \frac{(\lambda_1(t-s))^{m+q}}{(m+q)!},$$

folgt aus (1)

$$\begin{aligned} (L_n(s), L_n(t) - L_n(s)) &\xrightarrow{\mathcal{L}} (L(s), L(t) - L(s)) \\ &\xRightarrow{\text{CMT}} (L_n(s), L_n(t)) \rightarrow (L(s), L(t)). \end{aligned}$$

Die Eindeutigkeit der Darstellungen $u = k(\alpha + \beta) + l\beta$ und $v = m(\alpha + \beta) + q\beta$ war mehrfach entscheidend. Angenommen z.B. $u = \bar{k}(\alpha + \beta) + \bar{l}\beta$ mit $(\bar{k}, \bar{l}) \neq (k, l)$ ist weitere (von genau 2 Darstellungen) $\Rightarrow \{L_n(s) = u\} = \{Z_1 = k, Z_2 = l\} + \{Z_1 = \bar{k}, Z_2 = \bar{l}\}$. Allgemein erhält man disjunkte Zerlegung in d Mengen bei $d \in \mathbb{N}$ verschiedenen Darstellungen von u . Genauso für $\{L_n(t) - L_n(s) = v\}$, falls v genau $e \in \mathbb{N}$ Darstellungen hat. Es folgt:

$$\begin{aligned} P\{L_n(s) = u, L_n(t) - L_n(s) = v\} &= \\ \sum_{i=1}^d \sum_{j=1}^e &\underbrace{P\{Z_1 = k_i, Z_2 = l_i, Z_3 = m_j, Z_4 = q_j\}}_{\text{lasst sich behandeln wie oben}}. \end{aligned}$$

2. Fall: $s < t \leq 0$: analog

3. Fall: $s < 0 \leq t$: Betrachte $(-L_n(s), L_n(t))$. Mit den gleichen Argumenten wie oben zeige

$$(-L_n(s), L_n(t)) \xrightarrow{\mathcal{L}} (-L(s), L(t)).$$

□

Das nachste Resultat liefert die Straffheit von $(L_n) \subseteq D[-a, a]$.

Lemma 5.7. *Angenommen es existiert ein $r > 0$ mit $|F(s) - F(t)| \leq \bar{L}|s - t| \forall s, t \in B_r(\theta)$. Seien $a > 0$ und $-a \leq t_1 < t < t_2 \leq a$ beliebig. Dann existiert eine Konstante C , so dass $\forall n \geq \frac{a}{r}$ gilt:*

$$E|L_n(t) - L_n(t_1)||L_n(t_2) - L_n(t)| \leq C^2(t - t_1)(t_2 - t) \leq (C(t_2 - t_1))^2.$$

Beweis: $|Y_i| \leq 1 \ \forall i \Rightarrow |\alpha Y_i + \beta| \leq c = |\alpha + \beta| + |\beta| \Rightarrow$

$$\begin{aligned}
 E|\dots||\dots| &= E \left| \sum_{i=1}^n 1_{\{\theta + \frac{t_1}{n} < X_i \leq \theta + \frac{t}{n}\}} (\alpha Y_i + \beta) \right| \left| \sum_{j=1}^n 1_{\{\theta + \frac{t}{n} < X_j \leq \theta + \frac{t_2}{n}\}} (\alpha Y_j + \beta) \right| \\
 &\leq c^2 \sum_{\substack{i,j \\ i \neq j}} P \left(\theta + \frac{t_1}{n} < X_i \leq \theta + \frac{t}{n}, \theta + \frac{t}{n} < X_j \leq \theta + \frac{t_2}{n} \right) \\
 &\quad \text{da } \underbrace{\left(\theta + \frac{t_1}{n}, \theta + \frac{t}{n} \right) \cap \left(\theta + \frac{t}{n}, \theta + \frac{t_2}{n} \right) = \emptyset}_{\text{disj.}} \\
 &\stackrel{\text{Unabh.}}{=} c^2 \sum_{i \neq j} \underbrace{\left\{ F \left(\theta + \frac{t}{n} \right) - F \left(\theta + \frac{t_1}{n} \right) \right\}}_{\leq \overline{L} \frac{t-t_1}{n}} \underbrace{\left\{ F \left(\theta + \frac{t_2}{n} \right) - F \left(\theta + \frac{t}{n} \right) \right\}}_{\leq \overline{L} \frac{t-t_2}{n}} \quad (5.2) \\
 &\Rightarrow \text{Behauptung.}
 \end{aligned}$$

□

Da $P(N_1(a-) = N_1(a)) = 1$ (vgl. 7.5.6 in Gaenssler und Stute [GS77]) und

$$\left\{ \sum_{i=1}^{N_1(a-)} \xi_i = \sum_{i=1}^{N_1(a)} \xi_i \right\} \supseteq \{N_1(a-) = N_1(a)\},$$

folgt: $P(L_n(a-) \neq L_n(a)) = 0$.

Somit folgt mit Satz 3.3, Lemma 5.6 und Lemma 5.7:

Satz 5.8. Angenommen $\exists r > 0$ mit $|F(s) - F(t)| \leq \overline{L}|s - t| \ \forall s, t \in B_r(\theta)$ und angenommen $F'(\theta+), F'(\theta-) \in (0, \infty)$ existieren $\Rightarrow$

$$L_n \xrightarrow{L} L \text{ in } D[-a, a] \ \forall a > 0.$$

Damit ist auch (2) erfüllt. Bleibt (3). Dazu:

$$E(\xi_1) = \alpha b + \beta < 0 < \alpha a + \beta = E(\eta_1)$$

gemäß (5.4). Bekanntlich $N_1(t) \rightarrow \infty, t \rightarrow \infty$ und $N_2(-t) \rightarrow \infty, t \rightarrow -\infty$ f.s. $\Rightarrow L(t) \rightarrow -\infty, |t| \rightarrow \infty$ f.s. mit SGGZ $\Rightarrow A(L) \neq \emptyset$. Da $\delta(L) > 0$ f.s. gilt wegen Lemma 4.1: $A(L)$ abgeschlossen $\Rightarrow$ (3). Somit liefert Satz 4.9:

Theorem 5.9. Unter den Voraussetzungen von Satz 5.5 gelte zusätzlich, dass $F'(\theta+)$ und $F'(\theta-)$ existieren. Dann folgt:

$$(1) \liminf_{n \rightarrow \infty} P \left(n(\hat{\theta}_n - \theta) < x \right) \geq P(\sigma < x) \text{ und}$$

$$(2) \limsup_{n \rightarrow \infty} P\left(n(\hat{\theta}_n - \theta) \leq x\right) \leq P(\tau \leq x), \forall x \in \mathbb{X} = \mathbb{R},$$

wobei σ und τ die größte bzw. kleinste Maximalstelle von L aus (5.8).

Bemerkung 5.10. (1) Da L Treppenfunktion mit $\delta(L) > 0$ f.s., ist $\sigma > \tau$ f.s. Insbesondere gilt **nicht**:

$$\sigma \stackrel{\mathcal{L}}{=} \tau \quad (\text{denn sonst: } \sigma \stackrel{\mathcal{L}}{=} \tau \Rightarrow 0 = E(\underbrace{\sigma - \tau}_{\geq 0}) \Rightarrow \sigma = \tau \text{ f.s.})$$

(2) Sei $(a, b) \in \mathbb{R}^2$ mit $a < b$ eine Lösung der Ungleichung

$$P(\sigma < b) - P(\tau \leq a) \geq 1 - \alpha,$$

und sei

$$\hat{I}_n(\alpha) := \left(\hat{\theta}_n - \frac{b}{n}, \hat{\theta}_n - \frac{a}{n} \right) \Rightarrow$$

$$\begin{aligned} \liminf_{n \rightarrow \infty} P(\theta \in \hat{I}_n(\alpha)) &= \liminf_{n \rightarrow \infty} P(a < n(\hat{\theta}_n - \theta) < b) \\ &\geq \underbrace{\liminf_{n \rightarrow \infty} P(n(\hat{\theta}_n - \theta) < b)}_{\stackrel{(2)}{\geq} P(\sigma < b)} - \underbrace{\limsup_{n \rightarrow \infty} P(n(\hat{\theta}_n - \theta) \leq a)}_{\stackrel{(2)}{\leq} P(\tau \leq a)} \\ &\geq 1 - \alpha \end{aligned}$$

$\Rightarrow \hat{I}_n(\alpha)$ ist asymptotisches KI für θ zum Niveau $1 - \alpha$.

□

5.4 Der Fall a und b unbekannt

Dann klar: a und b sind zu schätzen. Dazu: Angenommen θ wäre bekannt. Erinnere an Log-Likelihood-Fkt.

$$\begin{aligned} l_n(\theta, a, b) &= \sum_{i=1}^n 1_{\{X_i \leq \theta\}} [Y_i \log(a) + (1 - Y_i) \log(1 - a)] + \\ &\quad + \sum_{i=1}^n 1_{\{X_i > \theta\}} [Y_i \log(b) + (1 - Y_i) \log(1 - b)] . \end{aligned}$$

Diese ist für festes θ partiell differenzierbar nach a und $b \Rightarrow$

$$0 = \frac{\partial l_n}{\partial a} = \sum_{i=1}^n 1_{\{X_i \leq \theta\}} \underbrace{\left[\frac{Y_i}{a} - \frac{1 - Y_i}{1 - a} \right]}_{\frac{Y_i - a}{a(1-a)}} \Leftrightarrow$$

$$a = \hat{a}_n(\theta) = \frac{\sum_{i=1}^n Y_i 1_{\{X_i \leq \theta\}}}{\sum_{i=1}^n 1_{\{X_i \leq \theta\}}} \quad (\theta \geq X_{1:n})$$

Analog: $0 = \frac{\partial l_n}{\partial b} \Leftrightarrow b = \hat{b}_n(\theta) = \frac{\sum_{i=1}^n Y_i 1_{\{X_i > \theta\}}}{\sum_{i=1}^n 1_{\{X_i > \theta\}}} \quad (\theta < X_{n:n}).$

Mit dem Lemma von Borel-Cantelli folgt leicht, dass der MLE $(\hat{a}_n(\theta), \hat{b}_n(\theta))$ für (a, b) (bei bekanntem θ) wohldefiniert ist für schließlich alle $n \in \mathbb{N}$ f.s. Da θ aber eben unbekannt, betrachte beliebiges Paar $(\hat{a}_n(s), \hat{b}_n(t))$ mit $s, t \in \mathbb{R}$.

Lemma 5.11. Falls $F(s) > 0$ und $F(t) < 1$, so gilt:

$$\hat{a}_n(s) \rightarrow \frac{1}{F(s)} \int_{-\infty}^s m(x) F(dx) =: \hat{a}(s) \quad f.s.$$

$$\hat{b}_n(t) \rightarrow \frac{1}{1-F(t)} \int_t^{\infty} m(x) F(dx) =: \hat{b}(t) \quad f.s.$$

Beweis: SGGZ

□

Folgerung 5.12. Falls $s \leq \theta \leq t$, $F(s) > 0$ und $F(t) < 1$, so

$$(\hat{a}_n(s), \hat{b}_n(t)) \rightarrow (a, b) \quad f.s.$$

D.h. man muss θ nicht exakt kennen, um (a, b) konsistent zu schätzen. Es reicht Kenntnis eines Intervalls $[s, t]$, das θ enthält. Klar: Je kleiner s bzw. je größer t , desto schlechter sollte die „performance“ des MLE sein. Mathematische Rechtfertigung in

Satz 5.13. (Asymptotische Normalität) Voraussetzungen von Folgerung 5.12 $\Rightarrow$

$$\sqrt{n} \left\{ \begin{pmatrix} \hat{a}_n(s) \\ \hat{b}_n(t) \end{pmatrix} - \begin{pmatrix} a \\ b \end{pmatrix} \right\} \xrightarrow{\mathcal{L}} \begin{pmatrix} U \\ V \end{pmatrix},$$

wobei $U \sim N(0, \sigma^2)$ und $V \sim N(0, \tau^2)$ unabhängig mit $\sigma^2 = \sigma^2(s) = \frac{a(1-a)}{F(s)}$ und $\tau^2 = \tau^2(t) = \frac{b(1-b)}{1-F(t)}$.

Beweis: Erinnere an Abschnitt 1.2 (6) $\Rightarrow$

$$\begin{aligned} Y_i &= m(X_i) + \varepsilon_i, \quad E(\varepsilon_i|X) = 0, \\ m(x) &= a1_{\{x \leq \theta\}} + b1_{\{x > \theta\}}, \\ H_n(s) &= \frac{1}{n} \sum_{i=1}^n 1_{\{X_i \leq s\}}(a + \varepsilon_i) \\ &= aF_n(s) + \frac{1}{n} \sum_{i=1}^n 1_{\{X_i \leq s\}} \varepsilon_i \quad \Rightarrow \end{aligned}$$

$$\sqrt{n}\{\hat{a}_n(s) - a\} = \frac{n^{-\frac{1}{2}} \sum_{i=1}^n 1_{\{X_i \leq s\}} \varepsilon_i}{F_n(s)}.$$

Analog

$$\sqrt{n}\{\hat{b}_n(t) - a\} = \frac{n^{-\frac{1}{2}} \sum_{i=1}^n 1_{\{X_i > t\}} \varepsilon_i}{1 - F_n(t)}.$$

Sei

$$S_n := n^{-\frac{1}{2}} \sum_{i=1}^n \underbrace{(1_{\{X_i \leq s\}}, 1_{\{X_i > t\}})}_{=: Z_i} \varepsilon_i.$$

Wegen $E(\varepsilon_i|X_i) = 0$ sind Z_i i.i.d. und zentriert. Mit ZGWS folgt sofort

$$S_n \xrightarrow{\mathcal{L}} N_2(0, \Gamma) =: S =: (\xi, \eta)^t, \text{ wobei}$$

$$\begin{aligned} \Gamma_{11} &= a(1-a)F(s) \\ \Gamma_{22} &= b(1-b)(1-F(t)) \end{aligned}$$

und $\Gamma_{12} = \Gamma_{21} = 0$ wegen $1_{\{X_i \leq s\}}1_{\{X_i > t\}} = 0$. Berechne z.B.

$$\begin{aligned} \Gamma_{11} &= E[1_{\{X \leq s\}} \varepsilon^2] \\ &= E[1_{\{X \leq s\}} (Y - a)^2] \\ &= \int_{-\infty}^s E[(Y - a)^2 | X = x] F(dx) \\ &= \int_{-\infty}^s \underbrace{\left[\int (y - a)^2 \text{Bin}(1, a)(dy) \right]}_{= \text{Var}(\text{Bin}(1, a)) = a(1-a)} F(dx). \end{aligned}$$

Mit SGGZ: $(F_n(s), 1 - F_n(t)) \rightarrow (F(s), 1 - F(t))$ f.s. Daher mit Cramér-Slutsky:

$$(S_n, F_n(s), 1 - F_n(t)) \xrightarrow{\mathcal{L}} (\xi, \eta, F(s), 1 - F(t)) ,$$

wobei $\xi \sim N(0, \Gamma_{11})$ und $\eta \sim N(0, \Gamma_{22})$ unabhängig wegen Γ diagonal. Da

$$\sqrt{n} \begin{pmatrix} \hat{a}_n - a \\ \hat{b}_n - b \end{pmatrix} = h(S_n, F_n(s), 1 - F_n(t))$$

und h stetig f.ü., folgt die Behauptung mit CMT. $\square$

Beachte, $\sigma^2 \rightarrow \infty$ für $s \rightarrow -\infty$ und $\tau^2 \rightarrow \infty$ für $t \rightarrow \infty$. D.h. für zu große Intervalle $[s, t]$ haben die Schätzer $\hat{a}_n(s)$ und $\hat{b}_n(t)$ u.U. inakzeptabel große Varianzen. Andererseits sind σ^2 und τ^2 minimal für $s = \theta = t$, d.h. $\hat{a}_n(\theta)$ und $\hat{b}_n(\theta)$ sind in diesem Sinne optimal. Problem (immer noch): θ unbekannt. Idee (Plug-in): Ersetze θ durch Schätzer θ_n^* , der nicht von a und b abhängt (Vorsicht: $\hat{\theta}_n = \hat{\theta}_n(a, b)$ ist daher nicht geeignet, denn hier „würde sich die Katze in den Schwanz beißen“.)

Erinnere an Dempfle-Stute Schätzer:

$$\bar{\theta}_n^+ := \operatorname{argmax}_{t \in \mathbb{R}} \rho_n(t)$$

$$\rho_n(t) = \frac{1}{n(n-1)} \sum_{1 \leq i \neq j \leq n} 1_{\{X_i \leq t < X_j\}} (Y_i - Y_j) .$$

Zweiseitige Variante davon ist

$$\bar{\theta}_n := \operatorname{argmax}_{t \in \mathbb{R}} |\rho_n(t)| .$$

Dempfle und Stute [DS02] zeigen

$$\bar{\theta}_n^+ - \theta = O_P(n^{-1}), \quad \text{falls } a > b .$$

Es lässt sich ähnlich wie in Satz 5.4 zeigen (Beweis später):

$$\hat{\theta}_n \rightarrow \theta \quad \text{f.s.,} \quad \text{falls } a > b$$

Vermutung (Beweis später):

$$\begin{aligned} \bar{\theta}_n - \theta &= O_P(n^{-1}), & \text{falls } a \neq b \\ \bar{\theta}_n &\rightarrow \theta \quad \text{f.s.,} & \text{falls } a = b. \end{aligned}$$

Sei nun θ_n^* irgendein Schätzer mit

$$\theta_n^* \rightarrow \theta \quad \text{f.s.} \quad \forall a \neq b \tag{5.9}$$

bzw.

$$\theta_n^* - \theta = O_P(n^{-1}) \quad \forall a \neq b. \tag{5.10}$$

Betrachte: $a_n^* := \hat{a}_n(\theta_n^*)$ und $b_n^* = \hat{b}_n(\theta_n^*)$

Satz 5.14. F stetig auf $B_r(\theta)$, $0 < F(\theta) < 1 \Rightarrow$

$$(a_n^*, b_n^*) \rightarrow (a, b) \quad \text{f.s.}$$

Beweis: $\hat{a}_n(s) = \frac{H_n(s)}{F_n(s)}$, $s \in [X_{1:n}, \infty)$. Beweis von Lemma 5.2 zeigt für stetige F :

$$(1) \quad \|H_n - H\| \rightarrow 0 \quad \text{f.s.}$$

Dies gilt auch für beliebige F (vgl. Bemerkung 5.3). Bekanntlich gemäß Glivenko-Cantelli [Gli33]:

$$(2) \quad \|F_n - F\| \rightarrow 0 \quad \text{f.s.}$$

Da

$$\hat{a}_n - \hat{a} = \frac{H_n}{F_n} - \frac{H}{F} = \frac{H_n - H}{F_n} - \frac{H(F_n - F)}{F F_n},$$

folgt mit (1) und (2):

$$\sup_{s \in B_r(\theta)} |\hat{a}_n(s) - \hat{a}(s)| \leq \frac{\|H_n - H\|}{F_n(\theta - r)} + \frac{\overbrace{H(\infty)}^{\leq 1} \|F_n - F\|}{F(\theta - r) F_n(\theta - r)} \rightarrow 0 \quad \text{f.s.},$$

denn $F(\theta - r) > 0$, falls r klein genug. Wegen (5.9) gilt:

$$\theta_n^* \in B_r(\theta) \text{ für schließlich alle } n \geq 1 \text{ f.s.}$$

$$\begin{aligned} \Rightarrow |a_n^* - a| &= |\hat{a}_n(\theta_n^*) - \hat{a}(\theta)| \\ &\leq |\hat{a}_n(\theta_n^*) - \hat{a}(\theta_n^*)| + |\hat{a}(\theta_n^*) - \hat{a}(\theta)| \\ &\leq \underbrace{\sup_{s \in B_r(\theta)} |\hat{a}_n(s) - \hat{a}(s)|}_{\rightarrow 0 \text{ f.s.}} + \underbrace{|a(\theta_n^*) - \hat{a}(\theta)|}_{\rightarrow 0 \text{ denn } \hat{a} \text{ stetig auf } B_r(\theta)} \end{aligned}$$

Analog: $b_n^* \rightarrow b$ f.s. □

Satz 5.15. Es gelten Voraussetzungen (1) und (2) aus Satz 5.5. Falls zusätzlich $F'(\theta+)$ und $F'(\theta-)$ existieren, so:

$$\sqrt{n} \left\{ \begin{pmatrix} a_n^* \\ b_n^* \end{pmatrix} - \begin{pmatrix} a \\ b \end{pmatrix} \right\} \xrightarrow{\mathcal{L}} \begin{pmatrix} U \\ V \end{pmatrix},$$

wobei $U \sim N(0, \sigma^2(\theta))$ und $V \sim N(0, \tau^2(\theta))$ unabhängig (mit den minimalen Varianzen).

Gemäß Poisson'schem Grenzwertsatz gilt:

$$\begin{aligned} \sum_{i=1}^n 1_{\{\theta < X_i \leq \theta + \frac{M}{n}\}} &\stackrel{\mathcal{L}}{=} \text{Bin} \left(n, F \left(\theta + \frac{M}{n} \right) - F(\theta) \right) \stackrel{\mathcal{L}}{\rightarrow} \text{Poisson}(MF'(\theta+)) \\ &\stackrel{\text{Slutsky}}{\Rightarrow} A_n(M) \rightarrow 0, \quad n \rightarrow \infty \quad \forall M > 0. \end{aligned}$$

Analog: $B_n(M) \rightarrow 0 \quad \forall M > 0$. Schließlich wegen (5.10):

$$\lim_{M \rightarrow \infty} \limsup_{n \rightarrow \infty} C_n(M) = 0.$$

Es folgt (1). Analog erhält man (2). $\square$

Erinnere an $\text{MLS} = \hat{\theta}_n = \hat{\theta}_n(a, b)$. Hängt von a und b ab. Ersetze (a, b) durch Schätzer (a_n^*, b_n^*) Erhalte:

$$\begin{aligned} \hat{\theta}_n^* &:= \operatorname{argmax}_{t \in \mathbb{R}} S_n^*(t) \\ S_n^*(t) &:= \frac{1}{n} \sum_{i=1}^n 1_{\{X_i \leq t\}} [\alpha_n^* Y_i + \beta_n^*] \\ \alpha_n^* &:= \log \frac{a_n^*(1 - b_n^*)}{b_n^*(1 - a_n^*)}, \quad \beta_n^* := \log \frac{1 - a_n^*}{1 - b_n^*} \end{aligned}$$

Numerische Bestimmung wieder unproblematisch

$$\begin{aligned} \hat{\theta}_n^* &= X_{l_n^*, n} \\ l_n^* &= \operatorname{argmax}_{1 \leq l \leq n} \left\{ \alpha_n^* \sum_{i=1}^l Y_{[i:n]} + l \beta_n^* \right\} \end{aligned}$$

Satz 5.16. F stetig auf $B_r(\theta)$ und streng monoton dort $\Rightarrow$

$$\hat{\theta}_n^* \rightarrow \theta \quad \text{f.s.}$$

Beweis: $S_n^* = \alpha_n^* H_n + \beta_n^* F_n$ gemäß Definition. $(\alpha_n^*, \beta_n^*) \xrightarrow{\text{f.s.}} (\alpha, \beta)$ wegen Stetigkeit und Satz 5.14.

Ferner (vgl. Lemma 5.2 bzw. Bemerkung 5.3):

$$\begin{aligned} \|H_n - H\| &\rightarrow 0, \quad \|F_n - F\| \rightarrow 0 \quad \text{f.s.} \Rightarrow \\ \|S_n^* - S\| &\rightarrow 0 \quad \text{f.s.} \Rightarrow \text{Behauptung mit 4.6 (vgl. 5.4).} \end{aligned}$$

$\square$

Abschließend der Nachweis, dass sich auch die (beinahe) Verteilungskonvergenz in Theorem 5.9 auf $\hat{\theta}_n^*$ überträgt. Erfolgt wieder mit unserem Argmax-CMT 4.9.

Satz 5.17. *Voraussetzungen aus Satz 5.5 $\Rightarrow \exists C > 0$:*

$$\limsup_{n \rightarrow \infty} P(n|\hat{\theta}_n^* - \theta| \geq x) \leq Cx^{-1} \quad \forall x > 0.$$

Insbesondere: $n(\hat{\theta}_n^* - \theta) = O_P(1)$.

Beweis: Wir orientieren uns am Beweis von Satz 5.5. Sei $d > 0$ beliebig $\Rightarrow$

$$(1) \quad P(|\hat{\theta}_n^* - \theta| \geq d) \leq P(d \leq |\hat{\theta}_n^* - \theta| \leq r) + \underbrace{P(|\hat{\theta}_n^* - \theta| > r)}_{\rightarrow 0 \quad \forall r > 0 \text{ wegen Satz 5.16}}.$$

Für o.E. $d \leq r$ gilt:

$$P(|\hat{\theta}_n^* - \theta| \geq d) \leq P(E_1^*) + P(E_2^*)$$

mit z.B.

$$(2) \quad P(E_1^*) \leq P\left(\sup_{d \leq v \leq r} \frac{|S_n^*(\theta + v) - S(\theta + v) - (S_n^*(\theta) - S(\theta))|}{v} \geq L\right).$$

Da $S_n^* = \alpha_n^* H_n + \beta_n^* F_n$ und $S = \alpha H + \beta F$, folgt
 $S_n^* - S = \alpha_n^*(H_n - H) + (\alpha_n^* - \alpha)H + \beta_n^*(F_n - F) + (\beta_n^* - \beta)F$,
 und daraus folgt

$$\begin{aligned} P(E_1^*) &\leq P\left(\sup_{d \leq v \leq r} \frac{|\alpha_n^*||H_n(\theta + v) - H(\theta + v) - (H_n(\theta) - H(\theta))|}{v} \geq \frac{L}{4}\right) \\ &\quad + P\left(\sup_{d \leq v \leq r} \frac{|\alpha_n^* - \alpha||H(\theta + v) - H(\theta)|}{v} \geq \frac{L}{4}\right) \\ &\quad + P\left(\sup_{d \leq v \leq r} \frac{|\beta_n^*||F_n(\theta + v) - F(\theta + v) - (F_n(\theta) - F(\theta))|}{v} \geq \frac{L}{4}\right) \\ &\quad + P\left(\sup_{d \leq v \leq r} \frac{|\beta_n^* - \beta||F(\theta + v) - F(\theta)|}{v} \geq \frac{L}{4}\right) \\ &=: A_1^* + A_2^* + A_3^* + A_4^*. \end{aligned}$$

Zerlegung des zugehörigen Ereignisses liefert:

$$\begin{aligned} A_1^* &\leq P\left(\sup_{d \leq v \leq r} \frac{|H_n(\theta + v) - H(\theta + v) - (H_n(\theta) - H(\theta))|}{v} \geq \frac{L}{8|\alpha|}\right) \\ &\quad + P(|\alpha_n^*| > 2|\alpha|) \\ &=: A_{11}^* + A_{12}^*. \end{aligned}$$

Dieselbe Herleitung wie von (13) im Beweis von Satz 5.5 zeigt:

$$(3) \quad A_{11}^* \leq C_1^* n^{-1} + C_2^* n^{-1} d^{-1}.$$

Weiter

$$(4) \quad A_{12}^* \leq P\left(\left||\alpha_n^*| - |\alpha|\right| \geq \overbrace{|\alpha|}^{>0}\right) \rightarrow 0 \quad \text{wegen Satz 5.14.}$$

Da $|H(\theta + v) - H(\theta)| \leq b\bar{L}v$ gemäß Satz 5.5 (2), folgt

$$(5) \quad A_2^* \leq P\left(|\alpha_n^* - \alpha| \geq \frac{L}{4b\bar{L}}\right) \rightarrow 0 \quad \text{wegen Satz 5.14.}$$

Analog behandelt man A_3^* (vgl. A_2 in Satz 5.5) und $A_4^* \Rightarrow$ Behauptung. $\square$

Bleibt funktionaler Grenzwertsatz für

$$L_n^*(t) := n \left\{ S_n^* \left(\theta + \frac{t}{n} \right) - S_n^*(\theta) \right\} \quad , \quad t \in \mathbb{R} ,$$

denn gemäß (4.2) haben wir wieder

$$n(\hat{\theta}_n^* - \theta) = \operatorname{argmax}_{t \in \mathbb{R}} L_n^*(t).$$

Satz 5.18. *Voraussetzungen aus Satz 5.8 $\Rightarrow$*

$$L_n^* \xrightarrow{\mathcal{L}} L \text{ in } D[-a, a] \quad \forall a > 0.$$

Beweis: Gemäß Definition gilt z.B. für $t \geq 0$

$$L_n^*(t) - L_n(t) = (\alpha_n^* - \alpha) \sum_{i=1}^n 1_{\{\theta < X_i \leq \theta + t\}} \overbrace{Y_i}^{|\cdot| \leq 1} + (\beta_n^* - \beta) \sum_{i=1}^n 1_{\{\theta < X_i \leq \theta + \frac{t}{n}\}}$$

$$\begin{aligned} \Rightarrow \sup_{-a \leq t \leq a} |L_n^*(t) - L_n(t)| &\leq \\ &\{|\alpha_n^* - \alpha| + |\beta_n^* - \beta|\} \sum_{i=1}^n 1_{\{\theta - \frac{a}{n} < X_i \leq \theta + \frac{a}{n}\}}. \end{aligned}$$

Die Summen konvergieren in Verteilung gemäß dem Grenzwertsatz von Poisson, während die Faktoren stochastisch gegen Null konvergieren gemäß 5.16, also konvergiert die obere Schranke ebenfalls stochastisch gegen Null.

$\Rightarrow$ Behauptung mit Slutsky + Satz 5.8. $\square$

Aus Argmax-CMT 4.9 folgt

Theorem 5.19. *Sind die Voraussetzungen aus Theorem 5.9 erfüllt, so gilt für alle $x \in \mathbb{R}$.*

$$(1) \liminf_{n \rightarrow \infty} P \left(n(\hat{\theta}_n^* - \theta) < x \right) \geq P(\sigma < x)$$

$$(2) \limsup_{n \rightarrow \infty} P \left(n(\hat{\theta}_n^* - \theta) \leq x \right) \leq P(\tau \leq x).$$

Bemerkung. (1) gilt $\forall G = (-\infty, x)$, $x \in \mathbb{R}$, aber nicht notwendig für alle offenen $G \subseteq \mathbb{R}$. In diesem Fall hätte man gemäß Portmanteau-Theorem, dass $n(\hat{\theta}_n^* - \theta) \xrightarrow{\mathcal{L}} \sigma$. Analog gilt (2) nur für alle abgeschlossene $(-\infty, x]$, $x \in \mathbb{R}$. Insofern gilt Portmanteau-Theorem nur beinahe (und auch für 2 verschiedene ZV σ und τ). Daher Sprechweise: $n(\hat{\theta}_n^* - \theta)$ konvergiert beinahe in Verteilung (oder fast schwach, almost weakly, a.w.) gegen (σ, τ) :

$$n(\hat{\theta}_n^* - \theta) \rightarrow_{\text{a.w.}} (\sigma, \tau).$$

5.5 Der Schätzer von Dempfle und Stute

$$\begin{aligned} \rho_n(t) &:= \frac{1}{n(n-1)} \sum_{1 \leq i \neq j \leq n} 1_{\{X_i \leq t < X_j\}} (Y_i - Y_j) \\ \bar{\theta}_n^\pm &:= \operatorname{argmax}_{t \in \mathbb{R}} \pm \rho_n(t) \\ \bar{\theta}_n &:= \operatorname{argmax}_{t \in \mathbb{R}} |\rho_n(t)| \end{aligned}$$

Idee am Rande: Ersetze die Funktion $h(Y_i, Y_j) = Y_i - Y_j$ durch beliebige messbare, antisymmetrische Funktion h .

Bestätige zunächst Vermutung, dass $\bar{\theta}_n$ Schätzer mit (5.9) und (5.10). Dazu

Lemma 5.20. $a \neq b$ und F beliebige Verteilungsfunktion $\Rightarrow$

$$\begin{aligned} \|\rho_n - \rho\| &\rightarrow 0 \quad \text{fast sicher, wobei} \\ \rho(t) &= (a - b) \begin{cases} (1 - F(\theta))F(t), & t \leq \theta \\ F(\theta)(1 - F(t)), & t > \theta \end{cases} \end{aligned}$$

Beweis: Eine einfache Umformung liefert

$$\begin{aligned} \rho_n(t) &= \frac{1}{n-1} \sum_{i=1}^n 1_{\{X_i \leq t\}} (Y_i - \bar{Y}_n) \\ &= \frac{n}{n-1} \{H_n(t) - \bar{Y}_n F_n(t)\}. \end{aligned}$$

$\|F_n - F\| \rightarrow 0$ f.s. gemäß Glivenko-Cantelli. Im Beweis von Bemerkung 5.3 wurde gezeigt:

$$\|H_n - H\| \rightarrow 0 \quad \text{f.s.}$$

Klar mit SGGZ: $\bar{Y}_n \rightarrow \int m dF$ f.s. $\Rightarrow$ Behauptung. $\square$

Folgerung aus Satz 4.6, falls F auf $B_r(\theta)$ sowohl strikt monoton wachsend als auch stetig:

$$\begin{aligned} \bar{\theta}_n^+ &\rightarrow \theta \quad \text{f.s., falls } a > b \\ \bar{\theta}_n^- &\rightarrow \theta \quad \text{f.s., falls } a < b \end{aligned}$$

Beachte: $\| |\rho_n| - |\rho| \| \leq \|\rho_n - \rho\| \rightarrow 0$ f.s.

$$\xrightarrow{\text{Satz 4.6}} \bar{\theta}_n \rightarrow \theta \quad \text{f.s., falls } a \neq b.$$

Dies zeigt, dass $\bar{\theta}_n$ (5.9) erfüllt. Bleibt z.z. dass $\bar{\theta}_n$ auch (5.10) erfüllt, also

$$\bar{\theta}_n - \theta = O_P(n^{-1}).$$

Folgt aus

$$P(\bar{\theta}_n \neq \bar{\theta}_n^+) \rightarrow 0, \quad a > b \quad (5.11)$$

und

$$P(\bar{\theta}_n \neq \bar{\theta}_n^-) \rightarrow 0, \quad a < b, \quad (5.12)$$

denn:

$$\begin{aligned} 0 \leq P(n|\bar{\theta}_n - \theta| > d) &\stackrel{\text{Zerlegung}}{\leq} P(n|\bar{\theta}_n^\pm - \theta| > d) + P(\bar{\theta}_n^\pm \neq \bar{\theta}_n) \stackrel{[\text{DS02}]}{\Rightarrow} \\ &\lim_{d \rightarrow \infty} \limsup_{n \rightarrow \infty} P(n|\bar{\theta}_n - \theta| > d) = 0 \Leftrightarrow (5.10). \end{aligned}$$

Nachweis von (5.11) und (5.12) mit folgendem

Lemma 5.21. $f \in D(\mathbb{R})$, $A(f) \neq \emptyset$. Falls

$$-\inf_{s \in \mathbb{R}} f(s) < \sup_{s \in \mathbb{R}} f(s), \quad (5.13)$$

so folgt

$$A(f) = A(|f|). \quad (5.14)$$

Insbesondere, falls zusätzlich $A(f)$ abgeschlossen+nub:

$$\operatorname{argmax}(f) = \operatorname{argmax}(|f|). \quad (5.15)$$

Beweis: (1) Zunächst 1. Vorbetrachtung:

$$\begin{aligned} t \in A(f) &\stackrel{\text{Def.}}{\Leftrightarrow} \sup f(s) = \max\{f(t), f(t-)\} \\ &\Leftrightarrow f(s) \leq \max\{f(t), f(t-)\} \quad \forall s \in \mathbb{R}, \end{aligned}$$

denn: „ $\Rightarrow$ “ trivial;

„ $\Leftarrow$ “

1. Fall: $f(t) \geq f(t-) \stackrel{\max=f(t)}{\Rightarrow} f(s) \leq f(t) \quad \forall s \in \mathbb{R}$

$$\begin{aligned} \sup f(s) &\leq f(t) \stackrel{\text{trivial}}{\leq} \sup f(s) \Rightarrow \\ \sup f(s) &= f(t) = \max\{f(t), f(t-)\} \end{aligned}$$

2. Fall: $f(t-) \geq f(t) \Rightarrow f(s) \leq f(t-) \quad \forall s \in \mathbb{R}$

$$\begin{aligned} \Rightarrow \sup f(s) &\leq f(t-) = \lim_{u \uparrow t} \underbrace{f(u)}_{\leq \sup f(s)} \leq \sup f(s) \\ \Rightarrow \sup f(s) &= f(t-) = \max\{f(t), f(t-)\} \end{aligned}$$

(2) Vorbetrachtung: $t \in A(f) \Rightarrow$

$$(*) \quad f(t) \vee f(t-) = |f(t)| \vee |f(t-)|, \quad \text{denn:}$$

1. Fall: $f(t)$ und $f(t-) \geq 0 \Rightarrow (*)$

2. Fall: $f(t) < 0 \leq f(t-) \Rightarrow$

$$(+)\quad |f(t)| \leq |f(t-)|$$

(sonst:

$$\begin{aligned} \underbrace{|f(t)|}_{=-f(t)} &> |f(t-)| = f(t-) = \max\{f(t), f(t-)\} \\ &= \sup f(s) \stackrel{(5.13)}{\geq} -\inf f(s) \end{aligned}$$

$\Rightarrow f(t) < \inf f(s)$ Widerspruch)

$$\Rightarrow |f(t)| \vee |f(t-)| \stackrel{(+)}{=} |f(t-)| \stackrel{2. \text{ Fall}}{=} f(t-) \stackrel{2. \text{ Fall}}{=} f(t) \vee f(t-)$$

$\Rightarrow (*)$

3. Fall: $f(t-) < 0 \leq f(t) \Rightarrow$

$$(++)\quad |f(t-)| \leq |f(t)|$$

(sonst:

$$\begin{aligned} -f(t-) &= |f(t-)| > |f(t)| = f(t) \\ &= f(t) \vee f(t-) = \sup f(s) \stackrel{(5.13)}{\geq} -\inf f(s) \end{aligned}$$

$$\Rightarrow f(t-) < \inf f(s) \Rightarrow \exists \varepsilon > 0:$$

$$f(t-) + \varepsilon < \inf f(s)$$

$$\text{Zu } \varepsilon > 0 \exists u < t : |f(u) - f(t-)| \leq \varepsilon \Rightarrow$$

$$f(u) \leq f(t-) + \varepsilon < \inf f(s) \text{ Widerspruch}$$

$$\begin{aligned} \Rightarrow |f(t)| \vee |f(t-)| &\stackrel{(++)}{=} |f(t)| \stackrel{3. \text{ Fall}}{=} f(t) \stackrel{3. \text{ Fall}}{=} f(t) \vee f(t-) \\ \Rightarrow (*) \end{aligned}$$

4. Fall: $f(t)$ und $f(t-) < 0$ nicht möglich, denn dann:

$$\sup f(s) < 0 \Rightarrow -\inf f(s) \stackrel{(5.13)}{<} \sup f(s) < 0 \Rightarrow \inf f(s) > 0 > \sup f(s)$$

$\Rightarrow$ Widerspruch.

Jetzt Beweis von (5.14):

„ $\subseteq$ “: Sei $t \in A(f)$ und $s \in \mathbb{R}$ beliebig.

1. Fall: $f(s) \geq 0 \Rightarrow$

$$|f(s)| = f(s) \stackrel{(1)}{\leq} f(t) \vee f(t-) \stackrel{(2)}{=} |f(t)| \vee |f(t-)|$$

2. Fall: $f(s) < 0 \Rightarrow$

$$\begin{aligned} |f(s)| &= -f(s) \leq -\inf f(s) \stackrel{(5.13)}{\leq} \sup f(s) \\ &= f(t) \vee f(t-) \stackrel{(2)}{=} |f(t)| \vee |f(t-)|. \end{aligned}$$

Somit: $|f(s)| \leq |f(t)| \vee |f(t-)| \quad \forall s \in \mathbb{R} \stackrel{(1)}{\Rightarrow} t \in A(|f|).$

(Wenden in der letzten Implikation (1) auf $|f|$ an mit $|f|(s) = |f(s)|$. Beachte, dass wegen der Stetigkeit der Betragsfunktion gilt:

$$\begin{aligned} |f|(t-) &= \lim_{u \uparrow t} |f|(u) = \lim_{u \uparrow t} |f(u)| \\ &\stackrel{\text{Stet.}}{=} \left| \lim_{u \uparrow t} f(u) \right| = |f(t-)| \end{aligned}$$

„ $\supset$ “: $t \in A(|f|) \stackrel{(1)}{\Rightarrow}$

$$\sup |f(s)| = |f(t)| \vee |f(t-)|$$

1. Fall: $\sup |f| = |f(t)| \Rightarrow f(t) \geq 0$

(sonst: $\sup(-f(s)) \geq -f(t) = |f(t)| = \sup |f(s)| \geq \sup(-f(s)) \Rightarrow \sup |f(s)| = \sup(-f(s)) = -\inf f(s) \stackrel{(5.13)}{<} \sup f(s) \leq \sup |f(s)|$ Widerspruch)

Es folgt für $f(s) \geq 0$:

$$f(s) = |f(s)| \leq |f(t)| = f(t) \leq f(t) \vee f(t-)$$

und für $f(s) < 0$:

$$f(s) < 0 \leq f(t) \leq f(t) \vee f(t-) \Rightarrow$$

$$f(s) \leq f(t) \vee f(t-) \quad \forall s \in \mathbb{R} \Rightarrow t \in A(f)$$

2. Fall: $\sup |f(s)| = |f(t-)| \Rightarrow f(t-) \geq 0$

(sonst:

$$\begin{aligned} \sup(-f(s)) &\geq \underbrace{-f(t-)}_{=\lim_{u \uparrow t} \underbrace{-f(u)}_{\leq \sup(-f(s))}} = |f(t-)| = \sup |f(s)| \geq \sup(-f(s)) \Rightarrow \\ &\sup |f(s)| = \sup(-f(s)) = -\inf f(s) \stackrel{(5.13)}{<} \sup f(s) \leq \sup |f(s)| \quad \text{Widerspruch} \end{aligned}$$

Es folgt für $f(s) \geq 0$:

$$f(s) = |f(s)| \leq |f(t-)| = f(t-) \leq f(t) \vee f(t-)$$

und für $f(s) < 0$:

$$\begin{aligned} f(s) < 0 \leq f(t-) \leq f(t) \vee f(t-) \Rightarrow \\ f(s) \leq f(t) \vee f(t-) \quad \forall s \in \mathbb{R} \stackrel{(1)}{\Rightarrow} t \in A(f) \end{aligned}$$

Dies zeigt (5.14), woraus mit Definition 4.2 die Gleichung (5.15) folgt. $\square$

Lemma 5.22.

$$P(\bar{\theta}_n \neq \bar{\theta}_n^\pm) \leq 2P(\|\rho_n - \rho\| > \delta^\pm),$$

wobei $\delta^\pm = \pm \frac{1}{2}\rho(\theta) = \pm \frac{1}{2}(a-b)F(\theta)(1-F(\theta))$ positiv.

Beweis: Mit Lemma 5.21 folgt:

$$\begin{aligned}
 \{\bar{\theta}_n \neq \bar{\theta}_n^+\} &\subseteq \left\{ -\inf_{s \in \mathbb{R}} \rho_n(s) \geq \underbrace{\sup_{s \in \mathbb{R}} \rho_n(s)}_{\geq \rho_n(\theta)} \right\} \\
 \Rightarrow P(\bar{\theta}_n \neq \bar{\theta}_n^+) &\leq P(-\inf \rho_n(t) \geq \rho_n(\theta), \overbrace{|\rho_n(\theta) - \rho(\theta)| \leq \delta^+}^{\subseteq \{\rho_n(\theta) \geq \rho(\theta) - \delta^+\}}) \\
 &\quad + P(|\rho_n(\theta) - \rho(\theta)| > \delta^+) =: P_n + Q_n.
 \end{aligned}$$

Da

$$\begin{aligned}
 -\inf \rho_n(t) &= -\underbrace{\inf \{\rho_n(t) - \rho(t) + \rho(t)\}}_{\geq \inf \{\rho_n - \rho\} + \inf \rho} \\
 &\leq -\inf \{\rho_n - \rho\} - \underbrace{\inf \rho}_{= 0 \text{ (in unserem Spezialfall Bin-modell)}} \\
 &= \sup \{-(\rho_n - \rho)\} \\
 &\leq \|\rho_n - \rho\|,
 \end{aligned}$$

folgt

$$P_n \leq P(\|\rho_n - \rho\| \geq \underbrace{\rho(\theta) - \delta^+}_{=\frac{1}{2}\rho(\theta) = \delta^+}).$$

Klar: $Q_n \leq P(\|\rho_n - \rho\| \geq \delta^+) \Rightarrow$ Behauptung für $\bar{\theta}_n^+$. Analog für $\bar{\theta}_n^-$, wenn ρ_n, ρ durch $-\rho_n, -\rho$ ersetzt werden. $\square$

Aus den Lemmata 5.20 und 5.22 folgen sofort (5.11) und (5.12). Damit ist unsere Vermutung bestätigt und $\theta_n^* := \bar{\theta}_n$ geeignet für Plug-In. Beachte, dass $\bar{\theta}_n$ nichtparametrischer Schätzer ist, d.h. benötigt nicht Werte von a und b .

Kapitel 6

Mehrdimensionale Grenzwertsätze

Was bisher geschah:

$$\begin{aligned}\hat{\theta}_n &= \hat{\theta}_n(a, b) = \text{MLS für } \theta \text{ bei bekannten } a, b \\ (\hat{a}_n(\theta), \hat{b}_n(\theta)) &= \text{MLS für } (a, b) \text{ bei bekanntem } \theta \\ \theta_n^* &= \text{Schätzer für } \theta \text{ (unabhängig von } (a, b), \text{ z.B. } \bar{\theta}_n) \\ (a_n^*, b_n^*) &= (\hat{a}_n(\theta_n^*), \hat{b}_n(\theta_n^*)) \\ \hat{\theta}_n^* &= \hat{\theta}_n(a_n^*, b_n^*)\end{aligned}$$

Bislang gezeigt:

$$\begin{aligned}(i) \quad n(\hat{\theta}_n^* - \theta) &\rightarrow_{a.w.} (\sigma, \tau) && \text{(Theorem 5.19)} \\ (ii) \quad \sqrt{n} \left\{ \begin{pmatrix} a_n^* \\ b_n^* \end{pmatrix} - \begin{pmatrix} a \\ b \end{pmatrix} \right\} &\xrightarrow{\mathcal{L}} \begin{pmatrix} U \\ V \end{pmatrix} && \text{(Satz 5.15)}\end{aligned}$$

Ziel: Gemeinsame (beinahe) Verteilungskonvergenz der ZVen in (i) und (ii).

Dazu:

$$n(\hat{\theta}_n^* - \theta) = \operatorname{argmax}(L_n^*) \quad \text{gemäß (4.2).}$$

Ferner gemäß (1), (2) im Beweis von Satz 5.15

$$\begin{aligned}\sqrt{n} \left\{ \begin{pmatrix} a_n^* \\ b_n^* \end{pmatrix} - \begin{pmatrix} a \\ b \end{pmatrix} \right\} - \sqrt{n} \left\{ \begin{pmatrix} \hat{a}_n(\theta) \\ \hat{b}_n(\theta) \end{pmatrix} - \begin{pmatrix} a \\ b \end{pmatrix} \right\} &= o_P(1) \Rightarrow \\ \xi_n^* &:= (n(\hat{\theta}_n^* - \theta), \sqrt{n}(a_n^* - a), \sqrt{n}(b_n^* - b)) \\ &= (\operatorname{argmax}(L_n^*), \underbrace{\sqrt{n}(\hat{a}_n(\theta) - a), \sqrt{n}(\hat{b}_n(\theta) - b)}_{=: Y_n}) + o_P(1).\end{aligned}$$

Wir werden sehen, dass ein einfaches Argument á la Cramér-Slutsky zeigt, dass es ausreicht,

$$(\operatorname{argmax}(L_n^*), Y_n)$$

für $n \rightarrow \infty$ zu untersuchen.

Ziel: Anwendung des Extended CMT Satz 4.14. Überprüfe Voraussetzungen (1)-(3) dort. Zu (1): $L_n^* \xrightarrow{\mathcal{L}} L$ in $D[-a, a] \forall a > 0$. Ist erfüllt gemäß Satz 5.18. Zu (3): $\sigma_n = \operatorname{argmax}(L_n^*) = O_P(1)$. Ist erfüllt gemäß Satz 5.17. Bleibt (2) in Satz 4.14, d.h. zu zeigen

$$(\pi_T(L_n^*), Y_n) \xrightarrow{\mathcal{L}} (\pi_T(L), Y) \text{ in } \mathbb{R}^{|T|} \times \mathbb{R}^2 \quad (6.1)$$

$\forall T \subseteq \mathbb{R}$ endlich, wobei $Y = (U, V)$ mit $U \sim N(0, \sigma^2(\theta))$, $V \sim N(0, \tau^2(\theta))$ unabhängig. Dazu: Der Beweis von Satz 5.18 zeigt sogar:

$$\sup_{-a \leq t \leq a} |L_n^*(t) - L_n(t)| = o_P(1) \quad \forall a > 0.$$

Daher betrachte (Notation aus 5.13):

$$(\pi_T(L_n), Y_n) = H(\pi_T(L_n), S_n, F_n(\theta), 1 - F_n(\theta)) ,$$

wobei $H := (Id, h)$, denn $Y_n = h(S_n, F_n(\theta), 1 - F_n(\theta))$. Wegen CMT und Satz von Slutsky reicht es für Nachweis von (6.1) zu zeigen

$$(\pi_T(L_n), S_n) \xrightarrow{\mathcal{L}} (\pi_T(L), S) \text{ in } \mathbb{R}^{|T|} \times \mathbb{R}^2 \quad (6.2)$$

$\forall T \subseteq \mathbb{R}$ endlich, wobei $S = (\xi, \eta) \sim N_2(0, \Gamma)$ mit Γ wie im Beweis von Satz 5.13 mit $s = t = \theta$.

Zeige (6.2) mittels Konvergenz der charakteristischen Funktionen (CF). Starte mit Spezialfall, der allerdings später verwendet werden kann. $|T| = 2$.

1. Fall: $0 \leq s < t$. Sei

$$\begin{aligned} \varphi_n(u_1, u_2) &:= \text{CF von } (L_n(s), L_n(t) - L_n(s)) \\ &= E \left[\exp \left\{ iu_1 \sum_{j=1}^n 1_{\{\theta < X_j \leq \theta + \frac{s}{n}\}} (\alpha Y_j + \beta) \right. \right. \\ &\quad \left. \left. + iu_2 \sum_{j=1}^n 1_{\{\theta + \frac{s}{n} < X_j \leq \theta + \frac{t}{n}\}} (\alpha Y_j + \beta) \right\} \right] \\ &= \{E[\exp\{iu_1 U_1 + iu_2 U_2\}]\}^n \end{aligned}$$

mit

$$\begin{aligned} U_1 &= 1_{\{\theta < X \leq \theta + \frac{s}{n}\}} r(Y) , \\ U_2 &= 1_{\{\theta + \frac{s}{n} < X \leq \theta + \frac{t}{n}\}} r(Y) , \\ r(y) &:= \alpha y + \beta, \quad y \in \{0, 1\}. \end{aligned}$$

Es gilt

$$\begin{aligned} e^{it} &= \sum_{k \geq 0} \frac{(it)^k}{k!} \\ &= 1 + it - \frac{1}{2}t^2 - \frac{1}{6}it^3 \dots, \quad t \in \mathbb{R}. \end{aligned}$$

Setze

$$\begin{aligned} U &:= u_1 U_1 + u_2 U_2 \quad \Rightarrow \\ \{\dots\} &= E(e^{iU}) = E\left(\sum_{k \geq 0} \frac{i^k U^k}{k!}\right) \\ &= 1 + E\left(\sum_{k \geq 1} \frac{i^k}{k!} U^k\right). \end{aligned}$$

Mit dem Binomischen Lehrsatz folgt:

$$U^k = \sum_{l=0}^k \binom{k}{l} 1_{\{\theta < X \leq \theta + \frac{s}{n}\}}^l r(Y)^l \cdot u_1^l \cdot 1_{\{\theta + \frac{s}{n} < X \leq \theta + \frac{t}{n}\}}^{k-l} r(Y)^{k-l} \cdot u_2^{k-l}$$

Es gilt: $1_A^l = 1_A \ \forall l \in \mathbb{N}$, $1_A^0 = 1$, $1_A 1_B = 0 \ \forall A, B$ disjunkt $\Rightarrow$

Es verschwinden die Summanden für $l \in \{1, \dots, k-1\} \Rightarrow$

$$\begin{aligned} U^k &= 1_{\{\theta < X \leq \theta + \frac{s}{n}\}} (u_1 r(Y))^k + 1_{\{\theta + \frac{s}{n} < X \leq \theta + \frac{t}{n}\}} (u_2 r(Y))^k \\ &= 1_{I_1} (u_1 r(Y))^k + 1_{I_2} (u_2 r(Y))^k \quad \Rightarrow \end{aligned}$$

$$E\left(\sum_{k \geq 1} \frac{i^k}{k!} U^k\right) = E\left(\sum_{k \geq 1} 1_{I_1} \frac{(iu_1 r(Y))^k}{k!}\right) + E\left(\sum_{k \geq 1} 1_{I_2} \frac{(iu_2 r(Y))^k}{k!}\right) =: R_1 + R_2.$$

Betrachte zunächst ersten Erwartungswert. Summation ab $k = 0$ liefert wieder die Reihendarstellung der e-Funktion. Dies führt zu:

$$\begin{aligned} R_1 &= E\left[1_{I_1} \sum_{k \geq 0} \frac{(iu_1 r(Y))^k}{k!} - 1_{I_1}\right] \\ &= E[1_{I_1} \exp\{iu_1 r(Y)\}] - P(I_1) \\ &=: R_{11} - R_{12}. \end{aligned}$$

Klar: $R_{12} = P(\theta < X \leq \theta + \frac{s}{n}) = F(\theta + \frac{s}{n}) - F(\theta)$. Bedingen bezüglich X liefert (mit $I_1 = I_1(X)$):

$$\begin{aligned} R_{11} &= E \left[\underbrace{1_{I_1(X)}}_{\sigma(X) - \text{mb}} E(\exp\{iu_1 r(Y)\} | X) \right] \\ &= \int_{I_1(x)} E(\exp\{iu_1 r(Y)\} | X = x) F(dx). \end{aligned}$$

Erinnere an

$$E[T(X, Y) | X = x] = \int T(x, y) P(Y \in dy | X = x) \Rightarrow$$

$$\begin{aligned} R_{11} &= \int_{(\theta, \theta + \frac{s}{n}]} \underbrace{\left[\int e^{iu_1 r(y)} \underbrace{P(Y \in dy | X = x)}_{= \text{Bin}(1, b)(dy), \text{ da } x > \theta} \right]}_{= E[\exp\{iu_1 r(B)\}], \ B \sim \text{Bin}(1, b)} F(dx) \\ &= \varphi_{r(B)}(u_1) \left\{ F\left(\theta + \frac{s}{n}\right) - F(\theta) \right\} \Rightarrow \\ R_1 &= \left\{ F\left(\theta + \frac{s}{n}\right) - F(\theta) \right\} \{\varphi_{r(B)}(u_1) - 1\}. \end{aligned}$$

Analog:

$$R_2 = \left\{ F\left(\theta + \frac{t}{n}\right) - F\left(\theta + \frac{s}{n}\right) \right\} \{\varphi_{r(B)}(u_2) - 1\} \Rightarrow$$

$$\varphi_n(u_1, u_2) = \{1 + R_1 + R_2\}^n = \left(1 + \frac{Z_n}{n}\right)^n \quad \text{mit}$$

$$Z_n = n(R_1 + R_2)$$

$$nR_1 = s \cdot \frac{F\left(\theta + \frac{s}{n}\right) - F(\theta)}{\frac{s}{n}} \{\varphi_{r(B)}(u_1) - 1\}$$

$$\rightarrow sF'(\theta+) \{\varphi_{r(B)}(u_1) - 1\}$$

$$nR_2 = \left(t \cdot \frac{F\left(\theta + \frac{t}{n}\right) - F(\theta)}{\frac{t}{n}} - s \cdot \frac{F\left(\theta + \frac{s}{n}\right) - F(\theta)}{\frac{s}{n}} \right) \{\varphi_{r(B)}(u_2) - 1\}$$

$$\rightarrow (t - s)F'(\theta+) \{\varphi_{r(B)}(u_2) - 1\} \Rightarrow$$

$$\begin{aligned} \varphi_n(u_1, u_2) &\rightarrow \exp\{sF'(\theta+)(\varphi_{r(B)}(u_1) - 1)\} \cdot \\ &\quad \exp\{(t - s)F'(\theta+)(\varphi_{r(B)}(u_2) - 1)\} \quad \forall u_1, u_2 \in \mathbb{R} \\ &= \varphi(u_1, u_2) \end{aligned} \tag{6.3}$$

Zeige:

$$\varphi = \text{CF von } (L(s), L(t) - L(s)). \quad (6.4)$$

Für Beweis von (6.4) verwende

Lemma 6.1. *Seien das Paar $(N, M) \in \mathbb{N}_0^2$ und die Folge $(X_i)_{i \in \mathbb{N}}$ unabhängig. Ferner gelte:*

- (i) $M \leq N$ f.s.
- (ii) $M, N - M$ unabhängig
- (iii) $X_i, i \geq 1$, i.i.d., $X_i \sim X$.

Dann gilt für $S_n := \sum_{i=1}^n X_i, n \in \mathbb{N}_0$:

- (a) $\varphi_{S_N}(\mu) = \Phi_N(\varphi_X(\mu)) \quad \forall \mu \in \mathbb{R}$,
wobei $\Phi_N(z) := \sum_{l \geq 0} z^l P(N = l), \quad z \in \mathbb{C}, |z| \leq 1$,
die probability generating function von N ist.

- (b) $\varphi_{S_N - S_M}(\mu) = \Phi_{N-M}(\varphi_X(\mu)) \quad \forall \mu \in \mathbb{R}$.

- (c) $\varphi_{S_M, S_N - S_M}(\lambda, \mu) = \Phi_M(\varphi_X(\lambda)) \Phi_{N-M}(\varphi_X(\mu)) = \varphi_{S_M}(\lambda) \cdot \varphi_{S_N - S_M}(\mu)$ für alle $\lambda, \mu \in \mathbb{R}$.

- (d) $S_M = \sum_{i=1}^M X_i$ und $S_N - S_M = \sum_{i=M+1}^N X_i$ sind unabhängig.

Beweis:

$$\begin{aligned}
 \varphi_{S_M, S_N - S_M}(\lambda, \mu) &\stackrel{\text{Def.}}{=} E \left[\exp \left\{ i\lambda \sum_{j=1}^M X_j + i\mu \sum_{j=M+1}^N X_j \right\} \right] \\
 &\stackrel{\text{Zerlegung von } \Omega}{=} E \left[\sum_{m \geq 0} \sum_{n \geq 0} \overbrace{1_{\{M=m, N=n\}}}^{\text{Zerlegung von } \Omega} \exp \left\{ i\lambda \sum_{j=1}^m X_j + i\mu \sum_{j=m+1}^n X_j \right\} \right] \\
 &\stackrel{\text{Lebesgue}}{=} \sum_{m \geq 0} \sum_{n \geq 0} E \left[1_{\{M=m, N=n\}} \cdot \exp \{ \dots \} \right]
 \end{aligned}$$

$$\begin{aligned}
 &= \sum_{m \geq 0} \sum_{n \geq 0} P(M = m, N = n) \cdot \underbrace{E \left[e^{i\lambda S_m} e^{i\mu(S_n - S_m)} \right]}_{\stackrel{(iii)}{=} \varphi_{S_m}(\lambda) \varphi_{S_n - S_m}(\mu)} \\
 &\stackrel{(iii)}{=} \{\varphi_X(\lambda)\}^m \cdot \{\varphi_X(\mu)\}^{n-m}
 \end{aligned}$$

da $1_{\{M=m, N=n\}}$, $\exp\{\dots\}$ unabhängig nach Voraussetzung.

Ferner: $\{M = m, N = n\} = \{M = m, N - M = n - m\} \stackrel{(i)}{\Rightarrow} P(M = m, N = n) = 0$ für alle $n < m$ wegen (i), und für alle $n \geq m$ gilt: $P(M = m, N = n) = P(M = m) \cdot P(N - M = n - m)$ wegen (ii) $\Rightarrow$

$$\begin{aligned}
 (e) \quad &\varphi_{S_M, S_N - S_M}(\lambda, \mu) \\
 &= \sum_{m \geq 0} \sum_{n \geq m} \underbrace{\{\varphi_X(\lambda)\}^m \cdot P(M = m)}_{=: a_m} \cdot \underbrace{\{\varphi_X(\mu)\}^{n-m} \cdot P(N - M = n - m)}_{=: b_{n-m}} \\
 &= \left(\sum_{m \geq 0} a_m \right) \left(\sum_{l \geq 0} b_l \right) \quad \text{Cauchy'sches Faltungsprodukt} \\
 &= \sum_{m \geq 0} \{\varphi_X(\lambda)\}^m P(M = m) \cdot \sum_{l \geq 0} \{\varphi_X(\mu)\}^l P(N - M = l).
 \end{aligned}$$

$$M = 0 \Rightarrow S_M = 0 \text{ und } S_N - S_M = S_N \Rightarrow \varphi_{S_N}(\mu) = \varphi_{S_M, S_N - S_M}(\lambda, \mu) \stackrel{(e)}{=} \sum_{l \geq 0} \{\varphi_X(\mu)\}^l \cdot P(N = l) = \Phi_N(\varphi_X(\mu)) \Rightarrow (a).$$

$\lambda = 0$ in (e) liefert:

$$\varphi_{S_N - S_M}(\mu) = \varphi_{S_M, S_N - S_M}(0, \mu) \stackrel{(e)}{=} \Phi_{N-M}(\varphi_X(\mu)) \Rightarrow (b).$$

Nochmal (e) liefert sofort 1. Gleichung in (c), woraus wiederum aus (a) und (b) die 2. Gleichung in (c) folgt.

(d) folgt aus (c) (vgl. z.B. Aufgabe 5, p. 195, in Bauer [Bau90]). $\square$

Zur Identifikation von φ in (6.3) erinnere an:

$$\begin{aligned}
 N &\sim \text{Poisson}(\lambda) \Rightarrow \\
 \Phi_N(z) &= e^{\lambda(z-1)} \quad \forall |z| \leq 1
 \end{aligned} \tag{6.5}$$

(Beachte: $N(s) - N(t) \sim \text{Poisson}(\lambda(s - t))$ für Poissonprozess N mit Parameter λ .)

Damit folgt:

$$\begin{aligned}
 \varphi(u_1, u_2) &\stackrel{(6.3)}{=} \Phi_{N_1(s)}(\varphi_{r(B)}(u_1)) \cdot \Phi_{N_1(t) - N_1(s)}(\varphi_{r(B)}(u_2)) \\
 &\stackrel{\text{Lemma 6.1(c)}}{=} \text{CF von} \left(\sum_{j=1}^{N_1(s)} r(B_j), \sum_{j=N_1(s)+1}^{N_1(t)} r(B_j) \right),
 \end{aligned}$$

wobei $B_j, j \geq 1$, i.i.d. $\sim \text{Bin}(1, b)$. Da

$$r(B) = \alpha B + \beta \stackrel{\mathcal{L}}{=} \xi_1,$$

folgt (vgl. (5.8)):

$$\varphi = \text{CF von } (L(s), L(t) - L(s)).$$

Somit liefert (6.3) mit Stetigkeitssatz für CF'en:

$$\begin{aligned} (L_n(s), L_n(t) - L_n(s)) &\xrightarrow{\mathcal{L}} (L(s), L(t) - L(s)) \\ &\stackrel{\text{CMT}}{\Rightarrow} (L_n(s), L_n(t)) \xrightarrow{\mathcal{L}} (L(s), L(t)). \end{aligned}$$

2. Fall: $s < 0 \leq t$. Betrachte hier

$$(-L_n(s), L_n(t)) = \left(\sum_{i=1}^n 1_{\{\theta + \frac{s}{n} < X_i \leq \theta\}} r(Y_i), \sum_{i=1}^n 1_{\{\theta < X_i \leq \theta + \frac{t}{n}\}} r(Y_i) \right).$$

Jetzt analog wie im 1. Fall mit

$$\begin{aligned} U_1 &= 1_{\{\theta + \frac{s}{n} < X_i \leq \theta\}} r(Y) = 1_{I_1} r(Y) \\ U_2 &= 1_{\{\theta < X_i \leq \theta + \frac{t}{n}\}} r(Y) = 1_{I_2} r(Y). \end{aligned}$$

Wegen $I_1 \cap I_2 = \emptyset$ folgt wieder für $U = u_1 U_1 + u_2 U_2$:

$$\begin{aligned} U^k &= 1_{I_1} (u_1 r(Y))^k + 1_{I_2} (u_2 r(Y))^k \quad \forall k \geq 1 \\ R_{11} &= \int_{(\theta + \frac{s}{n}, \theta]} \left[\int e^{iu_1 r(y)} \underbrace{P(Y \in dy | X = x)}_{=\text{Bin}(1, a), \text{ da } x \leq \theta} F(dx) \right] \\ &= \varphi_{r(A)}(u_1) \left\{ F(\theta) - F\left(\theta + \frac{s}{n}\right) \right\}, \quad A \sim \text{Bin}(1, a) \\ R_{12} &= P(I_1) = F(\theta) - F\left(\theta + \frac{s}{n}\right) \Rightarrow \\ R_1 &= - \left\{ F\left(\theta + \frac{s}{n}\right) - F(\theta) \right\} \{ \varphi_{r(A)}(u_1) - 1 \}. \end{aligned}$$

Analog:

$$\begin{aligned} R_2 &= \left\{ F\left(\theta + \frac{t}{n}\right) - F(\theta) \right\} \{ \varphi_{r(B)}(u_2) - 1 \} \\ \varphi_n &:= \text{CF von } (-L_n(s), L_n(t)) \Rightarrow \\ \varphi_n(u_1, u_2) &= \left(1 + \frac{Z_n}{n} \right)^n, \quad Z_n = n(R_1 + R_2) \\ nR_1 &= -s \frac{F\left(\theta + \frac{s}{n}\right) - F(\theta)}{\frac{s}{n}} \{ \varphi_{r(A)}(u_1) - 1 \} \\ &\rightarrow (-s) F'(\theta-) \{ \varphi_{r(A)}(u_1) - 1 \} \\ nR_2 &\rightarrow t F'(\theta+) \{ \varphi_{r(B)}(u_2) - 1 \} \end{aligned}$$

Daraus folgt

$$\begin{aligned}
 \varphi_n(u_1, u_2) &\rightarrow \exp \{(-s)F'(\theta-)(\varphi_{r(A)}(u_1) - 1)\} \cdot \exp \{tF'(\theta+)(\varphi_{r(B)}(u_2) - 1)\} \\
 &\stackrel{(6.3), (6.5)}{=} \underbrace{\Phi_{N_2(-s)}(\varphi_{r(A)}(u_1))}_{\substack{\text{Lemma 6.1(a)} \\ \text{CF von } \sum_{i=1}^{N_2(-s)} r(A_i)}} \cdot \underbrace{\Phi_{N_1(t)}(\varphi_{r(B)}(u_2))}_{\substack{\text{Lemma 6.1(a)} \\ \text{CF von } \sum_{i=1}^{N_1(t)} r(B_i)}} \\
 &\stackrel{A_i, i.i.d. \sim A}{=} \text{CF von } \left(\sum_{i=1}^{N_2(-s)} r(A_i), \sum_{i=1}^{N_1(t)} r(B_i) \right) \\
 &\stackrel{(5.8)}{=} \text{CF von } (-L(s), L(t))
 \end{aligned}$$

für (A_i) , (B_i) , N_1 , N_2 unabhängig. Mit dem CMT folgt

$$(L_n(s), L_n(t)) \xrightarrow{\mathcal{L}} (L(s), L(t)).$$

3. Fall: $s < t < 0$. Betrachte

$$(-L_n(t), L_n(t) - L_n(s))$$

und gehe vor wie im 1. Fall.

Dies zeigt

$$\pi_T(L_n) \xrightarrow{\mathcal{L}} \pi_T(L) \quad \forall T \subseteq \mathbb{R}, |T| = 2.$$

Beweisskizze im Fall $|T| \geq 2$: Seien endl. viele Punkte $s_l < s_{l-1} < \dots < s_1 < s_0 := 0 =: t_0 < t_1 < t_2 < \dots < t_m$:

$$\begin{array}{cccccccccccccccc}
 | & & | & & | & & | & & | & & | & & | & & | & & | \\
 s_l & \cdots & s_i & & s_{i-1} & \cdots & s_2 & & s_1 & & 0 & & t_1 & & t_2 & \cdots & t_{i-1} & & t_i & \cdots & t_m
 \end{array}$$

Betrachte Vektor der Zuwächse

$$\begin{aligned}
 &\left(L_n(t_i) - L_n(t_{i-1}), 1 \leq i \leq m; L_n(s_{i-1}) - L_n(s_i), 1 \leq i \leq l \right) = \\
 &\left(\sum_{j=1}^n 1_{\{\theta + \frac{t_{i-1}}{n} < X_j \leq \theta + \frac{t_i}{n}\}} r(Y_j), 1 \leq i \leq m; \sum_{j=1}^n 1_{\{\theta + \frac{s_i}{n} < X_j \leq \theta + \frac{s_{i-1}}{n}\}} r(Y_j), 1 \leq i \leq l \right)
 \end{aligned}$$

$$\begin{aligned}
 I_i &= I_i(X) := \left\{ \theta + \frac{t_{i-1}}{n} < X \leq \theta + \frac{t_i}{n} \right\}, \quad 1 \leq i \leq m, \\
 I_{m+i} &= I_{m+i}(X) := \left\{ \theta + \frac{s_i}{n} < X \leq \theta + \frac{s_{i-1}}{n} \right\}, \quad 1 \leq i \leq l, \\
 U_k &:= 1_{I_k(X)} r(Y), \quad 1 \leq k \leq m+l \quad \Rightarrow
 \end{aligned}$$

Vektor der Zuwächse = $\left(\sum_{j=1}^n 1_{I_k} r(Y_j) \right)_{1 \leq k \leq m+l}$ hat CF

$$\begin{aligned}
 \varphi_n(\underline{u}) &= E \left[\exp \left\{ i \sum_{k=1}^{m+l} u_k \left(\sum_{j=1}^n 1_{I_k} r(Y_j) \right) \right\} \right] \\
 &= E \left[\exp \left\{ i \sum_{j=1}^n \sum_{k=1}^{m+l} u_k 1_{I_k} r(Y_j) \right\} \right] \\
 &= \left\{ E \left[\exp \left\{ i \sum_{k=1}^{m+l} u_k 1_{I_k} r(Y) \right\} \right] \right\}^n \\
 &= \left\{ E [\exp \{ iU \}] \right\}^n, \quad \underline{u} = (u_1, \dots, u_{m+l}) \\
 U &:= \sum_{k=1}^{m+l} u_k U_k, \quad U_k := 1_{I_k} r(Y) \\
 &= \left\{ 1 + E \left[\sum_{k \geq 1} \frac{i^k}{k!} U^k \right] \right\}^n.
 \end{aligned}$$

Wegen der paarweisen Disjunktheit der I_k 's folgt mit dem Polynomischen Lehrsatz

$$\begin{aligned}
 U^k &= \sum_{j=1}^{m+l} 1_{I_j} (u_j r(Y))^k \Rightarrow \\
 \sum_{k \geq 1} \frac{i^k}{k!} U^k &= \sum_{j=1}^{m+l} 1_{I_j} \underbrace{\sum_{k \geq 1} \frac{(i u_j r(Y))^k}{k!}}_{=\exp\{i u_j r(Y)\} - 1} \\
 &= \sum_{j=1}^{m+l} 1_{I_j} \exp\{i u_j r(Y)\} - 1_{I_j} \Rightarrow \\
 1 + E \left[\sum_{k \geq 1} \frac{i^k}{k!} U^k \right] &= 1 + \sum_{j=1}^{m+l} \left\{ E \left(1_{I_j} \exp\{i u_j r(Y)\} \right) - P(I_j) \right\} \Rightarrow
 \end{aligned}$$

$$\begin{aligned}
 \varphi_n(\underline{u}) &= \left\{ 1 + \sum_{j=1}^m \left\{ F \left(\theta + \frac{t_j}{n} \right) - F \left(\theta + \frac{t_{j-1}}{n} \right) \right\} \{ \varphi_{r(B)}(u_j) - 1 \} \right. \\
 &\quad \left. + \sum_{j=m+1}^{m+l} \left\{ F \left(\theta + \frac{s_{j-1}}{n} \right) - F \left(\theta + \frac{s_j}{n} \right) \right\} \{ \varphi_{r(A)}(u_j) - 1 \} \right\}^n
 \end{aligned}$$

$$\begin{aligned}
 & \rightarrow \exp \left\{ \sum_{j=1}^m F'(\theta+)(t_j - t_{j-1})(\varphi_{r(B)}(u_j) - 1) \right. \\
 & \quad \left. + \sum_{j=m+1}^{m+l} F'(\theta-) \underbrace{(s_{j-1} - s_j)}_{=-s_j - (-s_{j-1})} (\varphi_{r(A)}(u_j) - 1) \right\} \\
 & \stackrel{(5.8), (6.5)}{=} \prod_{j=1}^m \Phi_{N_1(t_j) - N_1(t_{j-1})}(\varphi_{r(B)}(u_j)) \cdot \prod_{j=m+1}^{m+l} \Phi_{N_2(-s_j) - N_2(-s_{j-1})}(\varphi_{r(A)}(u_j)) \\
 & = \text{CF von } \left(L(t_j) - L(t_{j-1}), 1 \leq j \leq m; L(s_{j-1}) - L(s_j), 1 \leq j \leq l \right).
 \end{aligned}$$

Die letzte Gleichung folgt aus Erweiterung von Lemma 6.1(c) zu

Lemma 6.2. $(N_1, \dots, N_k) \in \mathbb{N}_0^k$ und $(X_i)_{i \geq 1}$ seien unabhängig. Ferner gelte

- (i) $0 =: N_0 \leq N_1 \leq \dots \leq N_k$ f.s.
- (ii) $N_j - N_{j-1}$, $1 \leq j \leq k$ unabhängig.
- (iii) X_i , $i \geq 1$, i.i.d. $\sim X$.

Dann gilt für alle $\underline{\lambda} = (\lambda_1, \dots, \lambda_k) \in \mathbb{R}^k$

$$\begin{aligned}
 (a) \quad \varphi(\lambda_1, \dots, \lambda_k) &= \prod_{j=1}^k \Phi_{N_j - N_{j-1}}(\varphi_X(\lambda_j)) \\
 &= \prod_{j=1}^k \varphi_{S_{N_j} - S_{N_{j-1}}}(\lambda_j),
 \end{aligned}$$

wobei $\varphi = \text{CF}$ des Vektors der Zuwächse $(S_{N_j} - S_{N_{j-1}} : 1 \leq j \leq k)$.

Aus (a) folgt:

$$(b) \quad S_{N_j} - S_{N_{j-1}}, \quad 1 \leq j \leq k, \quad \text{unabhängig.}$$

Beweis: analog wie Lemma 6.1. □

Bemerkung. Aus Lemma 6.2(b) folgt, dass der Grenzprozess L unabhängige Zuwächse besitzt. Vgl. auch Schmidt [Sch96].

Haben nunmehr gezeigt

$$\begin{aligned}
 & \left(L_n(t_j) - L_n(t_{j-1}), 1 \leq j \leq m; L_n(s_{j-1}) - L_n(s_j), 1 \leq j \leq l \right) \\
 & \xrightarrow{\mathcal{L}} \left(L(t_j) - L(t_{j-1}), 1 \leq j \leq m; L(s_{j-1}) - L(s_j), 1 \leq j \leq l \right) \xrightarrow{\text{CMT}} \\
 & \pi_T(L_n) \xrightarrow{\mathcal{L}} \pi_T(L) \quad \forall T \subseteq \mathbb{R}, |T| < \infty.
 \end{aligned} \tag{6.6}$$

Damit Lemma 5.6 nochmals bewiesen. Zurück zum Beweis von (6.2). Wegen CMT betrachte

$$\begin{aligned}\Delta_n &:= \left(\sum_{j=1}^n 1_{I_k} r(Y_j), 1 \leq k \leq m+l, S_n \right) \\ &= \sum_{j=1}^n \left(1_{I_k(X_j)} r(Y_j), 1 \leq k \leq m+l, \underbrace{n^{-\frac{1}{2}} 1_{\{X_j \leq \theta\}} \varepsilon_j}_{=V_1(X_j, Y_j)}, \underbrace{n^{-\frac{1}{2}} 1_{\{X_j > \theta\}} \varepsilon_j}_{=V_2(X_j, Y_j)} \right).\end{aligned}$$

Wegen Unabhängigkeit der Summanden hat Δ_n CF

$$\varphi_n(\underline{u}, \underline{v}) = \left\{ \varphi_{(\underline{U}, \underline{V})}(\underline{u}, \underline{v}) \right\}^n \quad \text{mit}$$

$$\begin{aligned}\underline{u} &= (u_1, \dots, u_{m+l}) \in \mathbb{R}^{m+l}, & \underline{v} &= (v_1, v_2) \in \mathbb{R}^2 \\ \underline{U} &= (U_1, \dots, U_{m+l}), & U_j &= 1_{I_j(X)} r(Y), \quad 1 \leq j \leq m+l \\ \underline{V} &= (V_1, V_2), \\ V_1 &= V_1(X, Y), & V_2 &= V_2(X, Y).\end{aligned}$$

$$\begin{aligned}\varphi_{(\underline{U}, \underline{V})}(\underline{u}, \underline{v}) &= E\left(\exp\{i(\underline{u}'\underline{U} + \underline{v}'\underline{V})\}\right) \\ U &:= \underline{u}'\underline{U}, \quad V := \underline{v}'\underline{V} \quad \Rightarrow\end{aligned}$$

$$\varphi_{(\underline{U}, \underline{V})}(\underline{u}, \underline{v}) = E\left(\exp\{i(U + V)\}\right) = 1 + \sum_{k \geq 1} \frac{i^k}{k!} E((U + V))^k.$$

$$\underline{k=1}: E(U + V) = E(U) + \overbrace{E(V)}^{=0} = E(U)$$

$$\underline{k=2}: E(U + V)^2 = E(U^2) + E(V^2) + 2E(UV)$$

$$\begin{aligned}E(V^2) &= n^{-1} \underline{v}' \Gamma \underline{v} \quad \text{wegen Disjunktheit, wobei} \\ \Gamma &= \text{Diag}(a(1-a)F(\theta), b(1-b)(1-F(\theta))),\end{aligned}$$

$$|E(UV)| \leq \sum_{j=1}^{m+l} \sum_{i=1}^2 u_j v_i E(1_{I_j(X)} |r(Y)| |V_i|)$$

Da $|V_i| \leq n^{-\frac{1}{2}} \forall i$ und $|r(Y)| \leq |\alpha + \beta| \vee |\beta| =: c$ folgt

$$|E(UV)| \leq cn^{-\frac{1}{2}} \sum_{i=1}^2 \sum_{j=1}^{m+l} u_j v_i P(I_j(X)).$$

$$\lim_{n \rightarrow \infty} nP(I_j(X)) = \begin{cases} (t_j - t_{j-1})F'(\theta+), & 1 \leq j \leq m \\ (s_{j-1} - s_j)F'(\theta-), & m+1 \leq j \leq m+l \end{cases}$$

$$\Rightarrow 2E(UV) = O(n^{-\frac{3}{2}}) = o(n^{-1}).$$

$k \geq 3$:

$$\begin{aligned} E(U+V)^k &= \sum_{q=0}^k \binom{k}{q} E(U^{k-q}V^q) \\ &= E(U^k) + \sum_{q=1}^k \binom{k}{q} E(U^{k-q}V^q) \Rightarrow \\ \varphi(\underline{U}, \underline{V}) &= 1 + iE(U) - \frac{1}{2}E(U^2) + \sum_{k \geq 3} \frac{i^k}{k!} E(U^k) \\ &\quad - \frac{1}{2} \underline{v}' \Gamma \underline{v} \cdot n^{-1} + o(n^{-1}) + \underbrace{\sum_{k \geq 3} \frac{i^k}{k!} \sum_{q=1}^k \binom{k}{q} E(U^{k-q}V^q)}_{=: R_n} \\ &= 1 + \sum_{k \geq 1} \frac{i^k}{k!} E(U^k) - \frac{1}{2} \underline{v}' \Gamma \underline{v} \cdot n^{-1} + o(n^{-1}) + R_n \end{aligned}$$

Zeige:

$$R_n = o(n^{-1}). \quad (6.7)$$

Dazu: $R_n = R_{n1} + R_{n2}$ (Abspalten des Summanden $q = k$) mit

$$\begin{aligned} R_{n1} &= \sum_{k \geq 3} \frac{i^k}{k!} E(V^k), \\ R_{n2} &= \sum_{k \geq 3} \frac{i^k}{k!} \sum_{q=1}^{k-1} \binom{k}{q} E(U^{k-q}V^q). \end{aligned}$$

$V = n^{-\frac{1}{2}} \varepsilon (v_1 1_{\{X \leq \theta\}} + v_2 1_{\{X > \theta\}}) = n^{-\frac{1}{2}} \tilde{V}$, wobei $\tilde{V}$ beschränkte ZV, denn

$$\begin{aligned} |\tilde{V}| &\leq \max(|v_1|, |v_2|) =: \tilde{m} \Rightarrow \\ |R_{n1}| &\leq \sum_{k \geq 3} \frac{1}{k!} \underbrace{n^{-\frac{k}{2}}}_{\leq n^{-\frac{3}{2}}} E(|\tilde{V}|^k) \\ &\leq n^{-\frac{3}{2}} \sum_{k \geq 0} \frac{1}{k!} E(|\tilde{V}|^k) \\ &\leq n^{-\frac{3}{2}} \underbrace{E(e^{|\tilde{V}|})}_{< \infty} = o(n^{-1}). \end{aligned}$$

und Δ (Vektor der Zuwächse von L) ist unabhängig von $N_2(0, \Gamma)$. Daraus folgt mit dem CMT

$$(\pi_T(L_n), S_n) \xrightarrow{\mathcal{L}} (\pi_T(L), N_2(0, \Gamma)) = (\pi_T(L), S). \quad (6.8)$$

Damit (6.2) gezeigt, wobei überdies gilt, dass $\pi_T(L)$ und $S = (\xi, \eta)$ unabhängig $\forall T \subseteq \mathbb{R}$ endlich

$$\begin{aligned} \xRightarrow{\text{CMT}} (\pi_T(L_n), Y_n) &\xrightarrow{\mathcal{L}} H(\pi_T(L), S, F(\theta), 1 - F(\theta)) \\ &= (\pi_T(L), Y), \end{aligned}$$

wobei $Y = h(S, F(\theta), 1 - F(\theta)) = (U, V)$ unabhängig von $\pi_T(L) \forall T \subseteq \mathbb{R}$ endlich. Da $\mathcal{B}(D(\mathbb{R})) = \sigma(\pi_T : T \subseteq \mathbb{R} \text{ endlich})$ gemäß Theorem 12.5(iii) in Billingsley [Bil99], folgt $L, (U, V)$ unabhängig, woraus wegen U, V unabhängig folgt, dass L, U und V unabhängig sind. Damit auch (2) in Satz 4.14 gezeigt. Zur Erinnerung

$$U \sim N(0, \sigma^2) \text{ und } V \sim N(0, \tau^2)$$

mit

$$\begin{aligned} \sigma^2 &= \sigma^2(\theta) = \frac{a(1-a)}{F(\theta)} \quad \text{und} \\ \tau^2 &= \tau^2(\theta) = \frac{b(1-b)}{1-F(\theta)}. \end{aligned}$$

Es folgt nunmehr mit Satz 4.14:

Theorem 6.3. *Für alle $x, y, z \in \mathbb{R}$ gilt:*

$$\begin{aligned} (1) \quad \liminf_{n \rightarrow \infty} P \left(\begin{array}{l} n(\hat{\theta}_n^* - \theta) < x \\ \sqrt{n}(a_n^* - a) < y \\ \sqrt{n}(b_n^* - b) < z \end{array} \right) &\geq P(\sigma_{\max} < x) \Phi\left(\frac{y}{\sigma}\right) \Phi\left(\frac{z}{\tau}\right), \\ (2) \quad \limsup_{n \rightarrow \infty} P \left(\begin{array}{l} n(\hat{\theta}_n^* - \theta) \leq x \\ \sqrt{n}(a_n^* - a) \leq y \\ \sqrt{n}(b_n^* - b) \leq z \end{array} \right) &\leq P(\sigma_{\min} \leq x) \Phi\left(\frac{y}{\sigma}\right) \Phi\left(\frac{z}{\tau}\right). \end{aligned}$$

Dabei $\sigma_{\max} = \text{Argmax}(L)$, $\sigma_{\min} = \text{argmax}(L)$ und L aus (5.8). Beachte $\mathbb{X} = \mathbb{R}$.

Beweis: Seien

$$\begin{aligned} \zeta_n &:= n(\hat{\theta}_n^* - \theta) \\ Y_{n1} &:= \sqrt{n}(\hat{a}_n(\theta) - a) \\ Y_{n2} &:= \sqrt{n}(\hat{b}_n(\theta) - b) \\ Y_{n1}^* &:= \sqrt{n}(a_n^* - a) \\ Y_{n2}^* &:= \sqrt{n}(b_n^* - b). \end{aligned}$$

Gemäß 4.14 gilt:

$$(i) \quad \liminf_{n \rightarrow \infty} P(\zeta_n < x, Y_{n1} < y, Y_{n2} < z) \geq P(\sigma_{max} < x) \Phi\left(\frac{y}{\sigma}\right) \Phi\left(\frac{z}{\tau}\right).$$

Gemäß (1) und (2) im Beweis von 5.15 gilt

$$\varepsilon_n := (Y_{n1}^* - Y_{n1}, Y_{n2}^* - Y_{n2}) = o_P(1).$$

$$\begin{aligned} & P(\zeta_n < x, Y_{n1}^* < y, Y_{n2}^* < z) \\ \geq & P(\zeta_n < x, Y_{n1} + \varepsilon_{n1} < y, Y_{n2} + \varepsilon_{n2} < z, |\varepsilon_{n1}| \leq \varepsilon, |\varepsilon_{n2}| \leq \varepsilon) \\ \geq & P(\underbrace{\zeta_n < x, Y_{n1} < y - \varepsilon, Y_{n2} < z - \varepsilon}_{=:A}, \underbrace{|\varepsilon_{n1}| \leq \varepsilon, |\varepsilon_{n2}| \leq \varepsilon}_{=:B}) \\ = & P(A \cap B) \\ = & P(A) + P(B) - \underbrace{P(A \cup B)}_{\leq 1} \\ \geq & P(A) - (1 - P(B)) \\ = & P(A) - P(CB) \\ \geq & P(\zeta_n < x, Y_{n1} < y - \varepsilon, Y_{n2} < z - \varepsilon) - P(|\varepsilon_{n1}| > \varepsilon) - P(|\varepsilon_{n2}| > \varepsilon) \end{aligned}$$

für alle $\varepsilon > 0$.

Grenzübergang $n \rightarrow \infty$ und anschließend $\varepsilon \rightarrow 0$ liefert mit (i) die Behauptung (1), da Φ stetig. Für Nachweis von (2) verwende man

$$\begin{aligned} & P(\zeta_n \leq x, Y_{n1}^* \leq y, Y_{n2}^* \leq z) \\ \leq & P(\zeta_n \leq x, Y_{n1} \leq y + \varepsilon, Y_{n2} \leq z + \varepsilon) \\ & + P(|\varepsilon_{n1}| > \varepsilon) + P(|\varepsilon_{n2}| > \varepsilon). \end{aligned}$$

□

Bemerkung. Eine unwesentlich modifizierte Argumentation liefert für beliebige Intervalle $I_1, I_2 \subseteq \mathbb{R}$ und $z \in \mathbb{R}$:

$$\begin{aligned} \liminf P(\zeta_n < z, \sqrt{n}(a_n^* - a) \in I_1, \sqrt{n}(b_n^* - b) \in I_2) \\ > P(\sigma_{max} < z) P(N(0, \sigma^2) \in I_1) P(N(0, \tau^2) \in I_2) \end{aligned}$$

und entsprechend für $\limsup$. Dabei ist es unerheblich, ob Intervalle offen, halb-offen, abgeschlossen usw.

Mit Theorem 6.3 lässt sich Konfidenzgebiet für (θ, a, b) angeben. Dazu

$$\hat{\sigma}_n^2 := \frac{a_n^*(1 - a_n^*)}{F_n(\theta_n^*)}, \quad \hat{\tau}_n^2 := \frac{b_n^*(1 - b_n^*)}{1 - F_n(\theta_n^*)}$$

Aus (5.9), Satz 5.14 und Glivenko-Cantelli [Gli33] folgt:

$$\hat{\sigma}_n^2 \rightarrow \sigma^2, \quad \hat{\tau}_n^2 \rightarrow \tau^2 \quad \text{f.s.} \quad (6.9)$$

Theorem 6.4. Für $\zeta_n := n(\hat{\theta}_n^* - \theta)$ und

$$\hat{Y}_n^* := \left(\sqrt{n} \frac{(a_n^* - a)}{\hat{\sigma}_n}, \sqrt{n} \frac{(b_n^* - b)}{\hat{\tau}_n} \right)' = (\hat{Y}_{n1}^*, \hat{Y}_{n2}^*)'$$

gilt:

$$(1) \liminf_{n \rightarrow \infty} P(\zeta_n < z, \hat{Y}_n^* \in I_1 \times I_2) \geq P(\sigma_{\max} < z) \cdot P(N(0, 1) \in I_1) \cdot P(N(0, 1) \in I_2)$$

und

$$(2) \limsup_{n \rightarrow \infty} P(\zeta_n \leq z, \hat{Y}_n^* \in I_1 \times I_2) \leq P(\sigma_{\min} \leq z) \cdot P(N(0, 1) \in I_1) \cdot P(N(0, 1) \in I_2)$$

$\forall z \in \mathbb{R}$ und $\forall$ Intervalle $I_1, I_2 \subseteq \mathbb{R}$.

Beweis: Ähnlich wie Beweis von Theorem 6.3, daher nur Skizze für $I_1 = (x_1, y_1]$, $I_2 = (x_2, y_2]$:

$$\begin{aligned} P \left(\zeta_n < z, \frac{\sqrt{n}(a_n^* - a)}{\hat{\sigma}_n} \in I_1, \frac{\sqrt{n}(b_n^* - b)}{\hat{\tau}_n} \in I_2 \right) \\ &\geq P(\zeta_n < z, \hat{\sigma}_n x_1 < \sqrt{n}(a_n^* - a) \leq \hat{\sigma}_n y_1, \hat{\tau}_n x_2 < \sqrt{n}(b_n^* - b) \leq \hat{\tau}_n y_2, \\ &\quad |\hat{\sigma}_n - \sigma| \leq \varepsilon, |\hat{\tau}_n - \tau| \leq \varepsilon) \\ &\geq P(\zeta_n < z, (\sigma + \varepsilon)x_1 < \sqrt{n}(a_n^* - a) \leq (\sigma - \varepsilon)y_1, (\tau + \varepsilon)x_2 < \sqrt{n}(b_n^* - b) \\ &\quad \leq (\tau - \varepsilon)y_2, |\hat{\sigma}_n - \sigma| \leq \varepsilon, |\hat{\tau}_n - \tau| \leq \varepsilon) \\ &\geq P(\zeta_n < z, Y_{n1}^* \in ((\sigma + \varepsilon)x_1, (\sigma - \varepsilon)y_1], Y_{n2}^* \in ((\tau + \varepsilon)x_2, (\tau - \varepsilon)y_2]) \\ &\quad - P(|\hat{\sigma}_n - \sigma| > \varepsilon) - P(|\hat{\tau}_n - \tau| > \varepsilon)). \end{aligned}$$

Rest des Beweises ist klar. □

Konstruiere mit Theorem 6.4 ein Konfidenzgebiet:

$$\begin{aligned} \hat{G}_n &:= \left(\hat{\theta}_n^* - \frac{c_1}{n}, \hat{\theta}_n^* + \frac{d_1}{n} \right) \times \left(a_n^* - \frac{c_2 \hat{\sigma}_n}{\sqrt{n}}, a_n^* + \frac{d_2 \hat{\sigma}_n}{\sqrt{n}} \right) \times \left(b_n^* - \frac{c_3 \hat{\tau}_n}{\sqrt{n}}, b_n^* + \frac{d_3 \hat{\tau}_n}{\sqrt{n}} \right) \\ &\Rightarrow P((\theta, a, b) \in \hat{G}_n) = P(\underbrace{-d_1 < \zeta_n < c_1}_{=: A_n \setminus B_n}, \underbrace{-d_2 < \hat{Y}_{n1}^* < c_2, -d_3 < \hat{Y}_{n2}^* < c_3}_{=: C_n}) \\ &\quad (A_n \setminus B_n) \cap C_n = (A_n \cap C_n) \setminus (B_n \cap C_n) \Rightarrow \end{aligned}$$

$$\begin{aligned}
 \liminf_{n \rightarrow \infty} P((\theta, a, b) \in \hat{G}_n) &= \liminf_{n \rightarrow \infty} \left\{ P(\zeta_n < c_1, C_n) - P(\zeta \leq -d_1, C_n) \right\} \\
 &\geq \underbrace{\liminf_{n \rightarrow \infty} P(\zeta_n < c_1, C_n)}_{\substack{\text{Theorem 6.4} \\ \geq}} - \underbrace{\limsup_{n \rightarrow \infty} P(\zeta_n \leq -d_1, C_n)}_{\substack{\text{Theorem 6.4} \\ \leq}} \\
 &\geq \left\{ P(\sigma_{\max} < c_1) - P(\sigma_{\min} \leq -d_1) \right\} \cdot \left\{ \Phi(c_2) - \Phi(-d_2) \right\} \cdot \left\{ \Phi(c_3) - \Phi(-d_3) \right\} \\
 &= 1 - \alpha
 \end{aligned}$$

bei geeigneter Wahl von $c_1, c_2, c_3, d_1, d_2, d_3$.

Beispiel. $d_i = c_i$, $i = 2, 3$; $c_2 = c_3 := \Phi^{-1}\left(\frac{1}{2}(\beta + 1)\right)$

$$\begin{aligned}
 c_1 &= \frac{1}{2}(1 + \beta) - \text{Quantil von } \sigma_{\max} \\
 -d_1 &= \frac{1}{2}(1 - \beta) - \text{Quantil von } \sigma_{\min} \\
 \beta &:= (1 - \alpha)^{\frac{1}{3}}
 \end{aligned}$$

c_1, d_1 unbekannt. Können über Monte-Carlo Methode geschätzt werden.

Kapitel 7

Asymptotik des Dempfle-Stute Schätzers in allgemeinen Modellen

$(X_i, Y_i) \sim (X, Y)$, $1 \leq i \leq n$, i.i.d., $X \sim F$, $E|Y| < \infty$, $m(x) := E(Y|X = x)$, $x \in \mathbb{R}$.

$$Y = m(X) + \epsilon \text{ mit } E(\epsilon|X) = 0 \quad \text{f.s.} \quad (7.1)$$

$$\bar{\theta}_n^\pm = \operatorname{argmax}_{t \in \mathbb{R}} \pm \rho_n(t)$$

$$\bar{\theta}_n = \operatorname{argmax}_{t \in \mathbb{R}} |\rho_n(t)|$$

$$\rho_n(t) = \frac{n}{n-1} \{H_n(t) - \bar{Y}_n F_n(t)\}$$

$$H_n(t) = \frac{1}{n} \sum_{i=1}^n 1_{\{X_i \leq t\}} Y_i \rightarrow H(t) = \int_{-\infty}^t m(x) F(dx) \quad \text{f.s.}$$

Der Beweis von Bemerkung 5.3 lässt sich problemlos auf obige allgemeine Situation übertragen, so dass gilt (vgl. auch Stute (1997):

$$\|H_n - H\| \rightarrow 0 \quad \text{f.s.}$$

Bekanntlich: $\|F_n - F\| \rightarrow 0$ f.s. Damit folgt:

Satz 7.1. *Sei F beliebig. $\Rightarrow \|\rho_n - \rho\| \rightarrow 0$ f.s., wobei*

$$\rho(t) = \int_{-\infty}^t (m(x) - \bar{m}) F(dx)$$

und

$$\bar{m} = \int_{\mathbb{R}} m(x) F(dx).$$

Eigenschaften der Grenzfunktion ρ in

Lemma 7.2. (1) $\rho \in D(\mathbb{R})$, d.h. ρ ist rcll.

(2) ρ ist stetig in $t \Leftrightarrow m(t) = \bar{m}$ oder F stetig in t .

(3) ρ monoton wachsend auf $(-\infty, \theta)$, falls

$$m(t) \geq \bar{m} \quad \forall t < \theta. \quad (7.2)$$

ρ monoton fallend auf (θ, ∞) , falls

$$m(t) \leq \bar{m} \quad \forall t > \theta. \quad (7.3)$$

Insbesondere gilt $\theta \in A(\rho)$.

(4) Falls $m(t) \neq \bar{m} \quad \forall t \in \mathbb{R}$, dann $C_\rho = C_F$.

(5) Sei F streng monoton in $B_\epsilon(\theta)$, und es gelte (7.2) und (7.3) mit strikter Ungleichung für $t \in B_\epsilon(\theta) \setminus \{\theta\}$ für ein $\epsilon > 0$, d.h.

$$m(t) > \bar{m} \quad \forall \theta - \epsilon < t < \theta, \quad (7.4)$$

$$m(t) < \bar{m} \quad \forall \theta < t < \theta + \epsilon. \quad (7.5)$$

Dann ist ρ strikt monoton in $B_\epsilon(\theta)$. Insbesondere gilt: $A(\rho) = \{\theta\}$.

Beweis: (1) $h(x) := m(x) - \bar{m}$. Sei $t \leq \theta$ und $t_m \uparrow t$ beliebig

$$\rho(t_m) = \int_{\mathbb{R}} \underbrace{1_{(-\infty, t_m]}(x)}_{\uparrow 1_{(-\infty, t]}} \cdot h(x) F(dx).$$

Mit Satz von Lebesgue (mit integrierbarer Majorante $|m(x)| + |\bar{m}|$) folgt:

$$\lim_{m \rightarrow \infty} \rho(t_m) = \int_{(-\infty, t)} h(x) F(dx) = \rho(t-).$$

Da $1_{(-\infty, t_m]}(x) \rightarrow 1_{(-\infty, t]}(x)$ für jede Folge $t_m \downarrow t$, folgt wiederum mit dem Satz von Lebesgue:

$$\rho(t+) = \int_{(-\infty, t]} h(x) F(dx) = \rho(t)$$

$\Rightarrow \rho$ rcll $\Rightarrow$ (1)

(2)

$$\begin{aligned}\rho \text{ stetig in } t &\stackrel{(1)}{\Leftrightarrow} 0 = \rho(t) - \rho(t-) \\ &\Leftrightarrow 0 = \int 1_{\{t\}}(x)h(x)F(dx) \\ &\Leftrightarrow 0 = (m(t) - \overline{m}) \cdot \{F(t) - F(t-)\}\end{aligned}$$

$\Rightarrow$ (2)

(3) $t < \theta \Rightarrow \rho(t) = \int \underbrace{1_{(-\infty, t]}(x)h(x)}_{\nearrow \text{ in } t} F(dx) \Rightarrow \rho(t)$ monoton wachsend wegen

Monotonie des Integrals.

$$t > \theta \text{ analog mit } \rho(t) = \rho(\theta) + \int \underbrace{1_{(\theta, t]}(x)h(x)}_{\searrow \text{ in } t \text{ wegen (7.3)}} F(dx).$$

(4) folgt aus (2).

(5) Angenommen es existieren $u < v < \theta$, $u, v \in B_\epsilon(\theta)$ mit

$$\begin{aligned}0 &= \rho(v) - \rho(u) \\ &= \int \underbrace{1_{(u, v]}(x)h(x)}_{\geq 0 \text{ wegen (7.2)}} F(dx).\end{aligned}$$

Sei Q = Wahrscheinlichkeitsmaß zu F = Vtlg. von X

$$\Rightarrow (+) \quad Q(\{x \in \mathbb{R} : 1_{(u, v]}(x)h(x) > 0\}) = 0$$

$$\Rightarrow (*) \quad \{x \in (u, v] : h(x) = 0\} \neq \emptyset,$$

denn angenommen: $(**) \{ \dots \} = \emptyset \Rightarrow$

$$\begin{aligned}(u, v] &= \underbrace{\{x \in (u, v] : h(x) < 0\}}_{=\emptyset \text{ wegen } h \geq 0} + \underbrace{\{x \in (u, v] : h(x) = 0\}}_{=\emptyset \text{ wegen } (**)} \\ &\quad + \underbrace{\{x \in (u, v] : h(x) > 0\}}_{=\{x \in \mathbb{R} : 1_{(u, v]}(x)h(x) > 0\}}.\end{aligned}$$

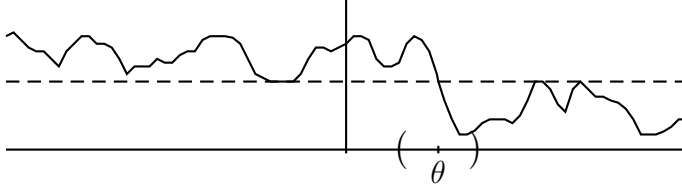
Und damit

$$\begin{aligned}Q((u, v]) &= Q(\{ \dots \}) + Q(\{ \dots \}) + Q(\{ \dots \}) \\ &= Q(\{x \in \mathbb{R} : 1_{(u, v]}(x)h(x) > 0\}) \\ &= 0 \\ &= F(v) - F(u).\end{aligned}$$

Widerspruch zu F streng monoton $\Rightarrow (*) \Rightarrow \exists x \in (u, v]; h(x) = 0$, d.h.
 $m(x) = \bar{m} \Rightarrow$ Widerspruch zu (7.4). $\Rightarrow \rho$ streng monoton wachsend auf
 $(\theta - \epsilon, \theta)$.
 Analog für Bereich $(\theta, \theta + \epsilon)$.

□

Satz 7.3. Sei m mit (7.2)-(7.5).



(m ansonsten sehr
allgemein, z.B. keine
Stetigkeit verlangt)

Falls F auf $B_\epsilon(\theta)$ stetig und strikt monoton, dann:

$$\bar{\theta}_n^+ \rightarrow \theta \quad f.s.$$

Beweis: Mit Satz 7.1: $\|\rho_n - \rho\| \rightarrow 0$ f.s. Mit Lemma 7.2: $A(\rho) = \{\theta\}$, ρ stetig und
 strikt monoton in $B_\epsilon(\theta)$. Die Behauptung folgt jetzt mit Satz 4.6. □

Lasse im Folgenden auch zu:

$$m(t) \leq \bar{m} \quad \forall t < \theta, \quad (7.6)$$

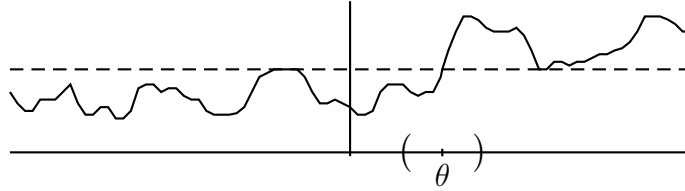
$$m(t) \geq \bar{m} \quad \forall t > \theta \quad (7.7)$$

mit strikter Ungleichung in einer Umgebung $B_\epsilon(\theta)$, d.h.:

$$m(t) < \bar{m} \quad \forall t \in (\theta - \epsilon, \theta), \quad (7.8)$$

$$m(t) > \bar{m} \quad \forall t \in (\theta, \theta + \epsilon) \quad (7.9)$$

für ein $\epsilon > 0$.



Lemma 7.4. Falls entweder $(+) := (7.2)-(7.5)$ oder $(-) := (7.6)-(7.9)$ gilt, so:

$$P(\bar{\theta}_n \neq \bar{\theta}^\pm) \leq 2P(\|\rho_n - \rho\| \geq \delta^\pm),$$

wobei $\delta^\pm = \pm \frac{1}{2}\rho(\theta) > 0$.

Beweis: Beweis von Lemma 5.22 kann wortwörtlich übernommen werden. $\square$

Folgerung 7.5.

$$\begin{aligned} P(\bar{\theta}_n \neq \bar{\theta}_n^+) &\rightarrow 0, & \text{falls } (+) \text{ gilt.} \\ P(\bar{\theta}_n \neq \bar{\theta}_n^-) &\rightarrow 0, & \text{falls } (-) \text{ gilt.} \end{aligned}$$

Klar: Satz 7.3 gilt analog für $\bar{\theta}_n^-$, d.h.

$$\bar{\theta}_n^- \rightarrow \theta \quad \text{f.s.,}$$

falls (7.6)-(7.9) gelten. Bleibt der „zweiseitige“ Schätzer: $\bar{\theta}_n$

Satz 7.6. *Angenommen, es gelten (7.2)-(7.5) =: (A1) oder (7.6)-(7.9) =: (A2). Falls F auf $B_\epsilon(\theta)$ stetig und streng monoton, so gilt:*

$$\bar{\theta}_n \rightarrow \theta \quad \text{f.s.}$$

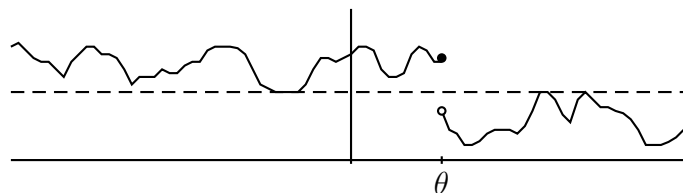
Beweis: $\| |\rho_n| - |\rho| \| \leq \|\rho_n - \rho\| \rightarrow 0$ f.s. gemäß Satz 7.1. Aus Lemma 7.2 folgt $A(|\rho|) = \{\theta\}$, $|\rho|$ stetig und monoton auf $B_\epsilon(\theta)$. Satz 4.6 liefert Behauptung. $\square$

Kapitel 8

Fehlerschranken für den Dempfle-Stute Schätzer

Es wird sich zeigen, dass das asymptotische Verhalten von $\bar{\theta}_n$ hinsichtlich einer Konvergenz der Verteilungen entscheidend von der „Glattheit“ bzw. „Nicht-Glattheit“ von m in einer Umgebung von θ abhängt.

Beispiel 8.1. (nicht-glatte Fall) m mit Sprung in θ , ansonsten mit (A1) oder (A2).



Ferner sei m stetig auf $\mathbb{R} \setminus \{0\}$.

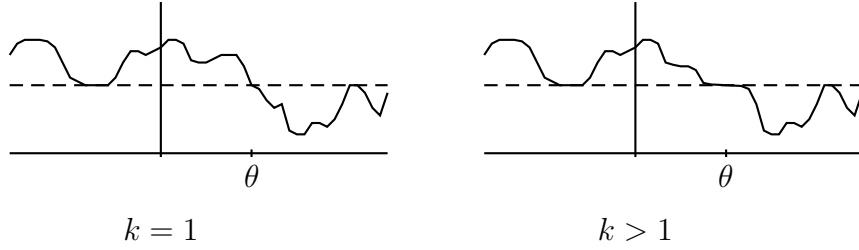
Beispiel 8.2. (glatter Fall) m mit (A1) oder (A2) und $m(\theta) = \overline{m}$, sowie für ein $k \in \mathbb{N}$

$$(*) \quad 0 = m'(\theta) = \dots = m^{(k-1)}(\theta)$$

mit

$$\begin{aligned} & -m^{(k)}(\theta+) > 0, \quad (-1)^k m^{(k)}(\theta-) > 0, \quad \text{falls (A1),} \\ \text{bzw.} \quad & -m^{(k)}(\theta+) < 0, \quad (-1)^k m^{(k)}(\theta-) < 0, \quad \text{falls (A2).} \end{aligned}$$

Für $k = 1$ entfällt die Forderung (*).



θ heißt split-point.

Lemma 8.3. Sei m wie in Beispiel 8.1 und $f = F'$ Dichte von F mit $f(\theta \pm) > 0$. Dann existiert $\varepsilon > 0$ und eine Konstante $L = L(\varepsilon) > 0$ mit

$$\rho(\theta) - \rho(t) \geq L \cdot |t - \theta| \quad \forall t \in B_\varepsilon(\theta).$$

Beweis: (A1): $\rho(t) = \int_{-\infty}^t h(x)f(x)dx$, $\rho'(t) = h(t)f(t)$.

Sei z.B. $t > \theta$. Mittelwertsatz liefert: $\exists z \in (\theta, t)$ mit

$$\begin{aligned} \rho(\theta) - \rho(t) &= -\rho'(z)(t - \theta) \geq \underbrace{\inf_{\theta \leq u \leq \theta + \varepsilon} (-\rho'(u))}_{=: L_1(\varepsilon)} |t - \theta| \\ &=: L_1(\varepsilon) \uparrow \underbrace{(\overline{m} - m(\theta +))}_{> 0 \text{ wegen Sprung}} f(\theta +) \end{aligned}$$

$t < \theta$ analog $\Rightarrow$ Behauptung.

(A2): analog. □

Lemma 8.4. Sei m wie in Beispiel 8.2 und f k -mal stetig diff-bar mit $f(\theta \pm) > 0$. Dann existieren $\varepsilon > 0$ und $L = L(\varepsilon) > 0$ mit

$$\rho(\theta) - \rho(t) \geq L|t - \theta|^{k+1} \quad \forall t \in B_\varepsilon(\theta).$$

Beweis: $\rho'(t) = h(t)f(t)$. Für $2 \leq l \leq k$ folgt mit der Leibniz'schen Formel:

$$\begin{aligned} \rho^{(l)}(\theta) &= (\rho'(\theta))^{(l-1)} = (h(\theta)f(\theta))^{(l-1)} \\ &= \sum_{j=0}^{l-1} \binom{l-1}{j} \underbrace{h^{(j)}(\theta)}_{=0 \quad \forall 0 \leq j \leq l-1 \text{ nach Vor.}} f^{(l-1-j)}(\theta) \\ &= 0. \end{aligned}$$

Ferner nochmal mit Leibniz:

$$\begin{aligned} \rho^{(k+1)}(z) &= \sum_{j=0}^k \binom{k}{j} h^{(j)}(z) f^{(k-j)}(z) \\ &\rightarrow h^{(k)}(\theta \pm) f(\theta \pm) = m^{(k)}(\theta \pm) f(\theta \pm) \quad (z \rightarrow \theta \pm). \end{aligned}$$

Mit Taylor-Formel folgt:

$$\rho(t) = \rho(\theta) + \frac{1}{(k+1)!} \rho^{(k+1)}(z)(t - \theta)^{(k+1)}.$$

1. Fall: $t \in [\theta - \varepsilon, \theta)$

$$(t - \theta)^{k+1} = (-1)^{k+1} |\theta - t|^{k+1} \Rightarrow$$

$$\begin{aligned} \rho(\theta) - \rho(t) &= -\frac{1}{(k+1)!} \rho^{(k+1)}(z)(-1)^{(k+1)} |\theta - t|^{k+1} \\ &\geq \frac{1}{(k+1)!} \underbrace{\inf_{\theta - \varepsilon \leq u < \theta} ((-1)^k \rho^{(k+1)}(u))}_{\rightarrow (-1)^k m^{(k)}(\theta -) f(\theta -) \quad (\varepsilon \downarrow 0)} |\theta - t|^{k+1} \end{aligned}$$

2. Fall: $t \in (\theta, \theta + \varepsilon]$ analog.

$\Rightarrow$ Behauptung. □

Im Folgenden wieder Herleitung einer oberen Schranke für die Fehlerwahrscheinlichkeit $P(|\bar{\theta}_n^+ - \theta| > d)$.

Dazu:

$$\bar{\theta}_n^+ = \operatorname{argmax}_{t \in \mathbb{R}} r_n(t),$$

wobei

$$r_n(t) := H_n(t) - \bar{Y}_n F_n(t).$$

Betrachte (vgl. Beweis von Satz 5.5):

$$E := \{d \leq |\bar{\theta}_n^+ - \theta| \leq \varepsilon\} \subseteq E_1 \cup E_2$$

$$E_1 = \bigcup_{d \leq u \leq \varepsilon} \{r_n(\theta + u) - r_n(\theta) \geq 0\}$$

$$E_2 = \bigcup_{-\varepsilon \leq u \leq -d} \{r_n(\theta + u) - r_n(\theta) \geq 0\}.$$

Es gilt:

$$\begin{aligned} r_n(t) &= \underbrace{H_n(t) - H(t)}_{=: U_n(t)} - \underbrace{\bar{Y}_n(F_n(t) - F(t))}_{=: V_n(t)} + \underbrace{H(t) - \bar{Y}_n F(t)}_{= \underbrace{H(t) - \bar{m} F(t)}_{=: \rho(t)} + \underbrace{(\bar{m} - \bar{Y}_n) F(t)}_{=: Z_n(t)}} \\ &= U_n(t) - V_n(t) + Z_n(t) + \rho(t). \end{aligned}$$

Daraus folgt

$$\begin{aligned} r_n(\theta + u) - r_n(\theta) &= \{U_n(\theta + u) - U_n(\theta)\} - \{V_n(\theta + u) - V_n(\theta)\} \\ &\quad + \underbrace{\{Z_n(\theta + u) - Z_n(\theta)\}}_{=: (\bar{m} - \bar{Y}_n)(F(\theta + u) - F(\theta))} - \underbrace{\{\rho(\theta) - \rho(\theta + u)\}}_{\geq Lu^{k+1}}. \end{aligned}$$

($k = 0 \Leftrightarrow m$ hat Sprungstelle in θ)

$$\begin{aligned} \Rightarrow P(E_1) &\leq P\left(\sup_{d \leq u \leq \varepsilon} \frac{|U_n(\theta + u) - U_n(\theta)|}{u^{k+1}} \geq \frac{1}{3}L\right) \\ &\quad + \left(\sup_{d \leq u \leq \varepsilon} \frac{|V_n(\theta + u) - V_n(\theta)|}{u^{k+1}} \geq \frac{1}{3}L\right) \\ &\quad + \left(|\bar{m} - \bar{Y}_n| \sup_{d \leq u \leq \varepsilon} \frac{|F(\theta + u) - F(\theta)|}{u^{k+1}} \geq \frac{1}{3}L\right) \\ &=: P_1 + P_2 + P_3. \end{aligned}$$

Abschätzung von P_1 :

$$\begin{aligned} U_n(t) &= \frac{1}{n} \sum_{i=1}^n 1_{\{X_i \leq t\}} (Y_i - m(X_i)) + \frac{1}{n} \sum_{i=1}^n (1_{\{X_i \leq t\}} m(X_i) - H(t)) \\ &=: S_n(t) + L_n(t). \\ P_1 &\leq P\left(\sup_{d \leq u \leq \varepsilon} \frac{|S_n(\theta + u) - S_n(\theta)|}{u^{k+1}} \geq \frac{1}{6}L\right) \\ &\quad + P\left(\sup_{d \leq u \leq \varepsilon} \frac{|L_n(\theta + u) - L_n(\theta)|}{u^{k+1}} \geq \frac{1}{6}L\right) \\ &=: q_1 + q_2. \end{aligned}$$

Setze

$$\begin{aligned} \Delta_n(u) &:= S_n(\theta + u) - S_n(\theta) \\ &= \frac{1}{n} \sum_{i=1}^n 1_{\{\theta < X_i \leq \theta + u\}} (Y_i - m(X_i)) \\ &= \frac{1}{n} \sum_{i=1}^n 1_{\{\theta < X_{i:n} \leq \theta + u\}} (Y_{[i:n]} - m(X_{i:n})) \end{aligned}$$

wegen des Kommutativgesetzes. Erinnere an

$$(*) \quad X_{i:n} = F_n^{-1}\left(\frac{i}{n}\right), \quad 1 \leq i \leq n,$$

wobei F_n^{-1} = verallgemeinerte Inverse zu F_n (Quantilfunktion). Für jede VF G gilt (vgl. z.B. Shorack und Wellner [SW86], p. 5):

$$(**) \quad G(x_1) < t \leq G(x_2) \Leftrightarrow x_1 < G^{-1}(t) \leq x_2 \quad \forall x_1 < x_2.$$

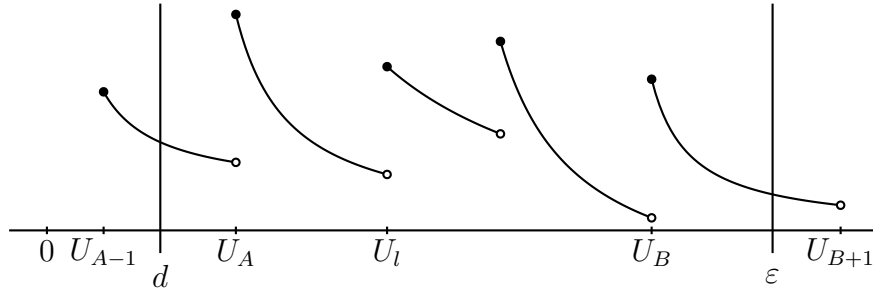
Damit folgt:

$$\theta < X_{[i:n]} \stackrel{(*)}{=} F_n^{-1}\left(\frac{i}{n}\right) \leq \theta + u \stackrel{(**)}{\Leftrightarrow} nF_n(\theta) < i \leq nF_n(\theta + u)$$

$$\Rightarrow \Delta_n(u) = \frac{1}{n} \sum_{nF_n(\theta) < i \leq nF_n(\theta+u)} Y_{[i:n]} - m(X_{i:n}).$$

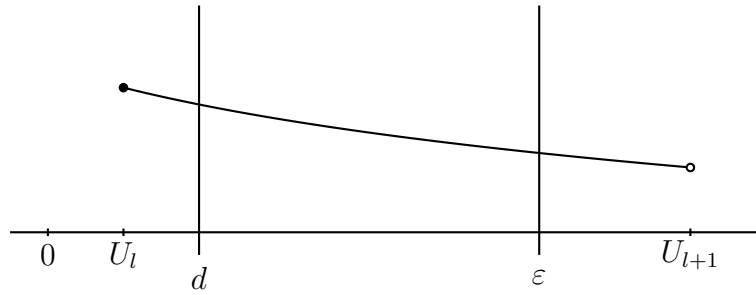
Der Prozess $\Delta_n = \{\Delta_n(u) : u \in [d, \varepsilon]\}$ ist stückweise konstant und springt genau in den Punkten $U_l := X_{l:n} - \theta$. Dasselbe gilt auch für $|\Delta_n| \geq 0$. Da $u \mapsto u^{-(k+1)}$ monoton fallend auf $[d, \varepsilon] \subseteq (0, \infty)$, folgt, dass der Prozess $\left\{ \frac{|\Delta_n(u)|}{u^{k+1}} : u \in [d, \varepsilon] \right\}$ in den Punkten U_l springt und dazwischen monoton fallend ist:

Bild 1:



mindestens ein U_l in $(d, \varepsilon]$.

Bild 2:



kein U_l in $(d, \varepsilon]$.

Setze

$$\begin{aligned} \Omega_n &:= \bigcup_{l=1}^n \{U_l \in (d, \varepsilon]\} \\ &= \bigcup_{l=1}^n \{X_{l:n} \in (\theta + d, \theta + \varepsilon]\}. \end{aligned}$$

Für $\omega \in \Omega_n$ definiere

$$\begin{aligned} A &:= A(\underline{X}_n) = \min\{1 \leq l \leq n : X_{l:n} \in (d + \theta, \theta + \varepsilon]\} \\ B &:= B(\underline{X}_n) = \max\{1 \leq l \leq n : X_{l:n} \in (d + \theta, \theta + \varepsilon]\}, \end{aligned}$$

wobei $\underline{X}_n := (X_{1:n}, \dots, X_{n:n})$ (Vektor der Ordnungsgrößen),

$$C := C(\underline{X}_n) := nF_n(\theta)$$

$$\forall \omega \in \Omega_n \quad \sup_{d \leq u \leq \varepsilon} \frac{|\Delta_n(u)|}{u^{k+1}} \leq \max \left\{ \frac{|\Delta_n(d)|}{d^{k+1}}, \max_{A \leq l \leq B} \frac{|\Delta_n(U_l)|}{U_l^{k+1}} \right\}.$$

Für $\omega \notin \Omega_n$ gilt:

$$\sup_{d \leq u \leq \varepsilon} \frac{|\Delta_n(u)|}{u^{k+1}} = \frac{|\Delta_n(d)|}{d^{k+1}}.$$

Damit

$$\begin{aligned} q_1 &= P \left(\left\{ \sup_{d \leq u \leq \varepsilon} \frac{|\Delta_n(u)|}{u^{k+1}} \geq \frac{1}{6}L \right\} \cap \Omega_n \right) + P \left(\left\{ \sup_{d \leq u \leq \varepsilon} \frac{|\Delta_n(u)|}{u^{k+1}} \geq \frac{1}{6}L \right\} \cap C\Omega_n \right) \\ &=: a_1 + a_2. \end{aligned}$$

$$\begin{aligned} a_1 &\leq P \left(\frac{|\Delta_n(d)|}{d^{k+1}} \geq \frac{1}{6}L \right) + P \left(\left\{ \max_{A \leq l \leq B} \frac{|\Delta_n(U_l)|}{U_l^{k+1}} \geq \frac{1}{6}L \right\} \cap \Omega_n \right) \\ &=: b_1 + b_2. \end{aligned}$$

$$b_1 = P \left(\left| \sum_{i=1}^n 1_{\{\theta < X_i \leq \theta+d\}} (Y_i - m(X_i)) \right| \geq \frac{1}{6}nLd^{k+1} \right),$$

wobei $1_{\{\theta < X_i \leq \theta+d\}} (Y_i - m(X_i))$ i.i.d. (klar) und zentriert, denn:

$$\begin{aligned} E \left[1_{\{\theta < X \leq \theta+d\}} (Y - m(X)) \right] &= E \left(E(1_{\{\theta < X \leq \theta+d\}} \varepsilon | X) \right) \\ &= E \left[1_{\{\theta < X \leq \theta+d\}} \underbrace{E(\varepsilon | X)}_{=0 \text{ f.s.}} \right] = 0. \end{aligned}$$

Die Ungleichung von Tschebyscheff liefert:

$$\begin{aligned} b_1 &\leq 36n^{-2}L^{-2}d^{-2(k+1)}n \text{Var} \left(1_{\{\theta < X \leq \theta+d\}} \varepsilon \right) \\ \text{Var} \left(1_{\{\theta < X \leq \theta+d\}} \varepsilon \right) &\stackrel{\text{Zentr.}}{=} E \left(1_{\{\theta < X \leq \theta+d\}} (Y - m(X))^2 \right) \\ &\stackrel{\text{Bed.}}{=} E \left[1_{\{\theta < X \leq \theta+d\}} \underbrace{E((Y - m(X))^2 | X)}_{=: V(X)} \right] \\ &= \int_{\theta}^{\theta+d} V(x) f(x) dx. \end{aligned}$$

$$b_1 \leq C n^{-1} d^{-(2k+1)} \underbrace{\frac{\int_{\theta}^{\theta+d} V(x) f(x) dx}{d}}_{\rightarrow V(\theta) f(\theta) \text{ für } d \rightarrow 0}. \quad (8.1)$$

Sei $d := d_n := x n^{-\alpha} \Rightarrow n^{-1} d^{-(2k+1)} = x^{-(2k+1)} n^{-1+\alpha(2k+1)} = x^{-(2k+1)}$, falls $\alpha = \frac{1}{2k+1}$. Betrachte ab jetzt $d = x n^{-\frac{1}{2k+1}}$. Schreibe deshalb genauer $b_1 \equiv b_1^{(n)}(x)$. Analog für die Wahrscheinlichkeiten $q_1 \equiv q_1^{(n)}(x)$ etc. Es folgt

$$\limsup_{n \rightarrow \infty} b_1^{(n)}(x) \leq C x^{-(2k+1)} \quad (8.2)$$

für alle $x > 0$.

Für die Abschätzung von b_2 ist folgendes Lemma sehr wichtig.

Lemma 8.5. (Stute und Wang [SW93]) (X_i, Y_i) , $1 \leq i \leq n$, i.i.d., $\sim (X, Y)$ mit beliebiger bivariater Verteilung, $\underline{X}_n = (X_{1:n}, \dots, X_{n:n})$, $\underline{Y}_n = (Y_{1:n}, \dots, Y_{n:n})$. Dann gilt:

- (a) $P \left(\bigcap_{i=1}^n \{Y_{[i:n]} \leq y_i\} | \underline{X}_n = \underline{x}_n \right) = \prod_{i=1}^n P(Y_{[i:n]} \leq y_i | \underline{X}_n = \underline{x}_n)$
für alle $\underline{y}_n \in \mathbb{R}^n$ und für $P \circ \underline{X}_n^{-1}$ fast alle $\underline{x}_n \in \mathbb{R}^n$. D.h. $Y_{[1:n]}, \dots, Y_{[n:n]}$ sind bedingt $\underline{X}_n$ unabhängig.
- (b) $P(Y_{[i:n]} \leq y | \underline{X}_n = \underline{x}_n) = P(Y_{[i:n]} \leq y | X_{i:n} = x_i)$
für alle $y \in \mathbb{R}$, für $P \circ \underline{X}_n^{-1}$ fast alle $\underline{x}_n \in \mathbb{R}^n$ und für jedes feste $1 \leq i \leq n$.
- (c) $P(Y_{[i:n]} \leq y | X_{i:n} = x) = P(Y \leq y | X = x)$
für alle $y \in \mathbb{R}$, für $P \circ X_{i:n}^{-1}$ fast alle $x \in \mathbb{R}$ und für jedes feste $1 \leq i \leq n$.

Bemerkung. (a) Ausarbeitung des Beweises vorhanden (siehe Diplomarbeit Klotsche [Klo02]).

(b) Für (X, Y) mit bivariater Dichte wurden (a) und (b) von Yang [Yan77] bewiesen.

$$b_2 = P \left(\left\{ \max_{A \leq l \leq B} \frac{|\Delta_n(U_l)|}{U_l^{k+1}} \geq \frac{L}{6} \right\} \cap \Omega_n \right) = E \left(1_{\Omega_n} 1_{\underbrace{\left\{ \max_{A \leq l \leq B} \frac{|\Delta_n(U_l)|}{U_l^{k+1}} \geq \frac{L}{6} \right\}}_{=: H(\underline{X}_n, \underline{Y}_n)}} \right)$$

mit

$$H(\underline{x}_n, \underline{y}_n) = 1_{\left\{ \max_{A(\underline{x}_n) \leq l \leq B(\underline{x}_n)} (x_l - \theta)^{-(k+1)} \left| \sum_{j=C(\underline{x}_n)+1}^l (y_j - m(x_j)) \right| \geq \frac{L_n}{6} \right\}},$$

$$C(\underline{x}_n) := \sum_{i=1}^n 1_{\{x_i \leq \theta\}},$$

denn $C(\underline{X}_n) = nF_n(\theta)$ und $nF_n(\theta + U_l) = nF_n(X_{l:n}) = l$. Also:

$$b_2 = E(1_{\Omega_n} \cdot H(\underline{X}_n, \underline{Y}_n)).$$

Sei $T_n := \{\underline{x}_n \in \mathbb{R}^n : x_1 \leq x_2 \leq \dots \leq x_n\}$ = Träger von $Q_n := P \circ \underline{X}_n^{-1}$ und

$$R_n := \{\underline{x}_n \in T_n : \exists 1 \leq i \leq n \text{ mit } x_i \in (\theta + d, \theta + \varepsilon + d]\}.$$

Dann folgt $\Omega_n = \underline{X}_n^{-1}(R_n) \in \sigma(\underline{X}_n) \Rightarrow 1_{\Omega_n}$ ist $\sigma(\underline{X}_n)$ - messbar.

$$\begin{aligned} b_2 &= E(1_{\Omega_n} E(H(\underline{X}_n, \underline{Y}_n) | \underline{X}_n)) \\ &\stackrel{\text{Trafo}}{=} \int_{R_n} \underbrace{E(H(\underline{X}_n, \underline{Y}_n) | \underline{X}_n = \underline{x}_n)}_{=: I(\underline{x}_n)} Q_n(d\underline{x}_n). \end{aligned} \quad (8.3)$$

Beachte für $\underline{x}_n \in R_n$ sind $A(\underline{x}_n) = \min\{1 \leq l \leq n : x_l \in (\theta + d, \theta + \varepsilon]\} =: a$ und $b := B(\underline{x}_n) = \max\{1 \leq l \leq n : x_l \in (\theta + d, \theta + \varepsilon]\}$ wohldefiniert.

$$\Rightarrow I(\underline{x}_n) = \int H(\underline{x}_n, \underline{y}_n) P(\underline{Y}_n \in d\underline{y}_n | \underline{X}_n = \underline{x}_n)$$

(vgl. Alsmeyer [Als98], Satz 53.6). Sei $Q(x, \cdot) := P(Y \in \cdot | X = x)$. Dann folgt mit Lemma 8.5

$$P(\underline{Y}_n \in d\underline{y}_n | \underline{X}_n = \underline{x}_n) = \bigotimes_{i=1}^n Q(x_i, dy_i).$$

Somit

$$I(\underline{x}_n) = \int H(\underline{x}_n, \underline{y}_n) \bigotimes_{i=1}^n Q(x_i, dy_i).$$

Seien $Z_1, \dots, Z_n$ unabhängig (o.E. über $(\Omega, \mathcal{A}, P)$) mit $Z_i \sim Q(x_i, \cdot)$, $1 \leq i \leq n \Rightarrow$

$$P \circ (Z_1, \dots, Z_n)^{-1} = \bigotimes_{i=1}^n Q(x_i, \cdot) \Rightarrow$$

$$\begin{aligned} I(\underline{x}_n) &= E(H(\underline{x}_n, Z_1, \dots, Z_n)) \\ &= P\left(\max_{a \leq l \leq b} \frac{1}{(x - \theta)^{k+1}} \left| \sum_{i=c+1}^l (Z_i - m(x_i)) \right| \geq \frac{Ln}{6}\right). \end{aligned}$$

Jetzt nächstes Hilfsmittel:

Satz 8.6. (*Hajek-Rényi-Ungleichung*) Seien ξ_i , $1 \leq i \leq n$, unabhängig und zentriert. Für $1 \leq m \leq n$ gelte $b_m \geq b_{m+1} \geq \dots \geq b_n$. Dann gilt

$$P\left(\max_{m \leq k \leq n} b_k \left| \sum_{i=1}^k \xi_i \right| \geq \lambda\right) \leq \lambda^{-2} \left\{ b_m^2 \sum_{i=1}^m \text{Var}(\xi_i) + \sum_{i=m+1}^n b_i^2 \text{Var}(\xi_i) \right\} \quad \forall \lambda > 0.$$

Beweis: Siehe Gänssler und Stute [GS77] für $m = 1$. Übertragung des Beweises auf $m \geq 1$ unproblematisch (siehe Diplomarbeit Höhne [Höh01], Lemma A.1) $\square$

Bemerkung. Verallgemeinerung auf Martingaldifferenzenfolge $(\xi_i, \mathcal{F}_i)$ in Shorack und Wellner [SW86], p. 873.

Mittels Indextransformation erhalten wir

$$I(\underline{x}_n) = P \left(\max_{a-c \leq r \leq b-c} (x_{r+c} - \theta)^{-(k+1)} \left| \sum_{i=1}^r \underbrace{(Z_{i+c} - m(x_{i+c}))}_{=: \xi_i \text{ unabh.}} \right| \geq \frac{Ln}{6} \right).$$

Ferner:

$$\begin{aligned} E(\xi_i) &= E(Z_{i+c} - m(x_{i+c})) \\ &= \underbrace{\int y \underbrace{Q(x_{i+c}, dy)}_{=P(Y \in dy | X=x_{i+c})} - m(x_{i+c})}_{\substack{[\text{Als98}] \\ \downarrow \\ E(Y | X=x_{i+c})=m(x_{i+c})}} \\ &\Rightarrow E(\xi_i) = 0. \end{aligned}$$

Außerdem: $b_r = (x_{r+c} - \theta)^{-(k+1)}$ monoton fallend in r , da $\underline{x}_n \in R_n \subseteq T_n$. Es folgt mit Satz 8.6:

$$\begin{aligned} I(\underline{x}_n) &\leq \frac{36}{n^2 L^2} \left\{ (x_a - \theta)^{-2(k+1)} \sum_{i=c+1}^a \text{Var}(Z_i - m(x_i)) \right. \\ &\quad \left. + \sum_{i=a+1}^b (x_i - \theta)^{-2(k+1)} \text{Var}(Z_i - m(x_i)) \right\}. \end{aligned}$$

Beachte: $\text{Var}(Z_i - m(x_i)) = E[(Z_i - m(x_i))^2] = \int (y - m(x_i))^2 Q(x_i, dy) = V(x_i)$. Im Folgenden spalte in der ersten Summe den Summanden für $i = a$ ab und ergänze damit die zweite Summe. Wir erhalten

$$I(\underline{x}_n) \leq \frac{36}{n^2 L^2} \{I_1(\underline{x}_n) + I_2(\underline{x}_n)\},$$

wobei

$$\begin{aligned} I_1(\underline{x}_n) &= (x_a - \theta)^{-2(k+1)} \sum_{i=c+1}^{a-1} V(x_i), \\ I_2(\underline{x}_n) &= \sum_{i=a}^b (x_i - \theta)^{-2(k+1)} V(x_i). \end{aligned}$$

Es folgt aus (8.3):

$$b_2 \stackrel{(8.3)}{=} \int_{R_n} I(\underline{x}_n) Q_n(d\underline{x}_n) \leq \frac{36}{n^2 L^2} \sum_{j=1}^2 \int_{R_n} I_j(\underline{x}_n) Q_n(d\underline{x}_n). \quad (8.4)$$

Verwende im Folgenden die Bezeichnung

$$\|g\|_{\varepsilon, \theta} := \sup\{|g(t)| : t \in B_\varepsilon(\theta)\}$$

für die Abbildung $g : B_\varepsilon(\theta) \rightarrow \mathbb{R}$. Mit Transformationssatz folgt

$$\begin{aligned} \int_{R_n} I_1(\underline{x}_n) Q_n(d\underline{x}_n) &= E \left(1_{\Omega_n} \underbrace{(X_{A:n} - \theta)}_{>d}^{-2(k+1)} \sum_{i=C+1}^{A-1} V(X_{i:n}) \right) \\ &\leq d^{-2(k+1)} E \left(1_{\Omega_n} \sum_{i=1}^n 1_{\{C+1 \leq i \leq A-1\}} V(X_{i:n}) \right). \end{aligned}$$

Es gilt auf Ω_n (d.h. $\forall \omega \in \Omega_n$):

$$(*) \quad C+1 \leq i \leq A-1 \Leftrightarrow \theta < X_{i:n} \leq \theta + d.$$

Beweis von (*): „ $\Rightarrow$ “: $i < A \Rightarrow X_{i:n} \leq \theta + d$, sonst Widerspruch zur Minimalität von A . Weiter: $C =$ „Anzahl der $X_{i:n} \leq \theta$ “. Angenommen $X_{i:n} \leq \theta$. Da $i \geq C+1$, folgt $X_{C+1:n} \leq X_{i:n} \leq \theta \Rightarrow$ mindestens $C+1$ Ordnungsgrößen $X_{i:n}$ sind $\leq \theta$. Widerspruch zur Definition von C .

„ $\Leftarrow$ “: Angenommen $i \geq A \Rightarrow X_{i:n} \geq X_{A:n} > \theta + d$ Widerspruch.

Angenommen $i \leq C \Rightarrow X_{i:n} \leq X_{C:n} \leq \theta$ Widerspruch. □

Es folgt

$$\begin{aligned} \int_{R_n} I_1(\underline{x}_n) Q_n(d\underline{x}_n) &\leq d^{-2(k+1)} E \left(\sum_{i=1}^n 1_{\{\theta < X_{i:n} \leq \theta + d\}} V(X_{i:n}) \right) \\ &\stackrel{\text{Komm.}}{=} d^{-2(k+1)} n \int_0^d \underbrace{V(\theta + u) f(\theta + u)}_{\leq \|V\|_{\varepsilon, \theta} \|f\|_{\varepsilon, \theta}} du, \end{aligned}$$

also

$$\int_{R_n} I_1(\underline{x}_n) Q_n(d\underline{x}_n) \leq C n d^{-2(k+1)}. \quad (8.5)$$

C ist hier und im Folgenden eine generische Konstante. Schließlich:

$$\int_{R_n} I_2(\underline{x}_n) Q_n(d\underline{x}_n) = E \left(1_{\Omega_n} \sum_{i=1}^n 1_{\{A \leq i \leq B\}} (X_{i:n} - \theta)^{-2(k+1)} V(X_{i:n}) \right).$$

Da $\Omega_n \cap \{A \leq i \leq B\} \subseteq \{d < X_{i:n} - \theta \leq \varepsilon\}$, ergibt sich folgende obere Schranke:

$$\begin{aligned} \int_{R_n} I_2(\underline{x}_n) Q_n(d\underline{x}_n) &\leq E \left(\sum_{i=1}^n 1_{\{d < X_{i:n} - \theta \leq \varepsilon\}} (X_{i:n} - \theta)^{-2(k+1)} V(X_{i:n}) \right) \\ &\stackrel{\text{Komm.}}{=} n \int_d^\varepsilon u^{-2(k+1)} V(\theta + u) f(\theta + u) du, \end{aligned}$$

also

$$\int_{R_n} I_2(\underline{x}_n) Q_n(d\underline{x}_n) \leq n C d^{-(2k+1)}. \quad (8.6)$$

Mit (8.4)-(8.6) folgt: $b_2 \leq C n^{-1} d^{-(2k+1)}$. Für $d = x n^{-\frac{1}{2k+1}}$ folgt insbesondere ($b_2 \equiv b_2^{(n)}(x)$ etc.):

$$\begin{aligned} \limsup_{n \rightarrow \infty} b_2^{(n)}(x) &\leq C x^{-(2k+1)} \quad \forall x > 0 \\ \Rightarrow \limsup_{n \rightarrow \infty} a_1^{(n)}(x) &\leq C x^{-(2k+1)} \quad \forall x > 0 \\ a_2 \equiv a_2^{(n)}(x) &\leq P \left(\frac{|\Delta_n(d)|}{d^{k+1}} \leq \frac{1}{6} \right) = b_1^{(n)}(x) \\ \limsup_{n \rightarrow \infty} q_1(x) &\leq C x^{-(2k+1)} \quad \forall x > 0. \end{aligned} \quad (8.7)$$

Bleibt $q_2 = q_2^{(n)}(x)$. Dazu verwende die so genannte Doob-Meyer-Zerlegung des Prozesses $\hat{L}_n(t)$ in

$$(+)\quad L_n(t) = \underbrace{\frac{1}{n} \sum_{i=1}^n 1_{\{X_i \leq t\}} m(X_i) - H(t)}_{=: \hat{L}_n(t)}, \quad t \in \mathbb{R}.$$

Es gilt

Satz 8.7. $X_1, \dots, X_n$ i.i.d. $\sim F$ beliebige VF, m beliebige messbare Abbildung. Dann gilt:

$$\hat{L}_n(t) = M_n(t) + A_n(t),$$

wobei

$$(M_n(t), \mathcal{F}_n(t) : t \in \mathbb{R})$$

ein zentriertes Martingal ist bzgl. der Filtration

$$\mathcal{F}_n(t) = \sigma \left(\bigcup_{i=1}^n \sigma(\{X_i \leq r\} : r \leq t) \right).$$

Ferner gilt

$$A_n(t) = \int_{(-\infty, t]} \frac{1 - F_n(x-)}{1 - F(x-)} m(x) F(dx), \quad t \in \mathbb{R}.$$

(M_n heißt Innovations-Martingal und A_n Kompensator).

Beweis:

$$\begin{aligned} M_n(t) &= \hat{L}_n(t) - A_n(t) \\ &= \frac{1}{n} \sum_{i=1}^n \left[1_{\{X_i \leq t\}} m(X_i) - \int_{(-\infty, t]} \frac{1_{\{X_i \geq u\}}}{1 - F(u-)} m(u) F(du) \right] \end{aligned} \quad (8.8)$$

Es ist zu zeigen

$$E(M_n(t) | \mathcal{F}_n(s)) = M_n(s) \quad \text{f.s.} \quad \forall s < t \quad (8.9)$$

Folgendes Argument reduziert die Betrachtung auf den Fall $n = 1$: Es bezeichne

$$\mathcal{G}_i(t) := \sigma\left(\{X_i \leq r\} : r \leq t\right), \quad 1 \leq i \leq n.$$

Daraus folgt

$$\begin{aligned} \mathcal{F}_n(s) &= \sigma\left(\bigcup_{j=1}^n \mathcal{G}_j(s)\right) \\ &= \sigma\left(\bigcup_{\substack{j=1 \\ j \neq i}}^n \mathcal{G}_j(s) \cup \mathcal{G}_i(s)\right) \\ &= \sigma\left(\sigma\left(\bigcup_{\substack{j=1 \\ j \neq i}}^n \mathcal{G}_j(s)\right) \cup \mathcal{G}_i(s)\right) \end{aligned} \quad (8.10)$$

für jedes $i \in \{1, \dots, n\}$, wobei die letzte Gleichung aus der Monotonie des σ -Operators folgt. Mit (8.8) und der Linearität der bedingten Erwartung ergibt sich:

$$\begin{aligned} E(M_n(t) | \mathcal{F}_n(s)) &= \frac{1}{n} \sum_{i=1}^n E \left[\underbrace{1_{\{X_i \leq t\}} m(X_i) - \int_{(-\infty, t]} \frac{1_{\{X_i \geq u\}}}{1 - F(u-)} m(u) F(du)}_{=: \bar{M}_t(X_i)} \middle| \mathcal{F}_n(s) \right] \\ &= \frac{1}{n} \sum_{i=1}^n E [\bar{M}_t(X_i) | \mathcal{G}_i(s)] \end{aligned} \quad (8.11)$$

wobei die letzte Gleichung eine Folgerung ist von (8.10) und der Eigenschaft der bedingten Erwartung, dass „unabhängige Anteile gestrichen werden dürfen“, vgl. z.B. Abschnitt 9.7 in Williams [Wil91]. Fixiere $1 \leq i \leq n$ und unterdrücke i.f. den Index i in der Notation. Zeige für $\overline{M}_t \equiv \overline{M}_t(X)$:

$$E(\overline{M}_t | \mathcal{G}(s)) = \overline{M}_s \quad \text{f.s.} \quad (8.12)$$

Dazu beachte, dass wegen $(-\infty, t] = (-\infty, s] + (s, t]$ folgt:

$$\overline{M}_t = \overline{M}_s + \left[1_{\{s < X \leq t\}} m(X) - \int_{(s,t]} 1_{\{X \geq u\}} \frac{m(u)}{1 - F(u-)} F(du) \right].$$

Daher reicht es zu zeigen:

$$E[1_{\{s < X \leq t\}} m(X) | \mathcal{G}(s)] = E \left[\int_{(s,t]} 1_{\{X \geq u\}} \frac{m(u)}{1 - F(u-)} F(du) | \mathcal{G}(s) \right] \quad \text{f.s.} \quad (8.13)$$

Gemäß Satz 1.7.10 (i) in Gänsler/Stute [GS77] ist dies gleichbedeutend mit

$$\int_G 1_{\{s < X \leq t\}} m(X) dP = \int_G \left[\int_{(s,t]} 1_{\{X \geq u\}} \frac{m(u)}{1 - F(u-)} F(du) \right] dP \quad (8.14)$$

für alle $G \in \mathcal{G}(s) = \sigma(\mathcal{E}(s))$, wobei

$$\mathcal{E}(s) := \{\Omega\} \cup \{\{X \leq r\} : r \leq s\}$$

ein durchschnittsstabiler Erzeuger. Da o.E. $m \geq 0$ angenommen werden kann (sonst $m = m^+ - m^-$ und Linearität der involvierten Integrale ausnutzen), kann der Maßeindeutigkeitssatz angewendet werden, d.h. es reicht (8.14) für Mengen $G \in \mathcal{E}(s)$ nachzuweisen.

1.Fall: $G = \{X \leq r\}$, $r \leq s \Rightarrow$ Auf beiden Seiten in (8.14) steht Null.

2.Fall: $G = \Omega \Rightarrow$

$$\begin{aligned}
 \text{linke Seite in (8.14)} & \stackrel{\text{Transfo}}{=} \int_{(s,t]} m(x) F(dx). \\
 \text{rechte Seite in (8.14)} & = \int_{\Omega} \int_{(s,t]} 1_{\{X \geq u\}} \frac{m(u)}{1 - F(u-)} F(du) dP \\
 & \stackrel{\text{Fubini}}{=} \int_{(s,t]} \int_{\Omega} 1_{\{X \geq u\}} \underbrace{\frac{m(u)}{1 - F(u-)}}_{\text{konstant in } \omega \in \Omega} dP F(du) \\
 & = \int_{(s,t]} \frac{m(u)}{1 - F(u-)} \underbrace{P(X \geq u)}_{1 - F(u-)} F(du) \\
 & = \int_{(s,t]} m(u) F(du).
 \end{aligned}$$

Dies zeigt (8.12) und überdies die Zentriertheit von $M_n(t)$.

Anwendung von (8.12) auf i -ten Summanden in (8.11) liefert

$$E[M_n(t) | \mathcal{F}_n(s)] = \frac{1}{n} \sum_{i=1}^n \overline{M}_s(X_i) = M_n(s) \quad \text{f.s.}$$

$\Rightarrow$ (8.9) wie gewünscht.

Schließlich gilt Adaptiertheit, d.h.

$$M_n(t) \text{ ist } \mathcal{F}_n(t) - \text{messbar } \forall t \in \mathbb{R}. \quad (8.15)$$

Mit Standardargumenten reduziert sich der Nachweis von (8.15) auf den von

$$1_{\{X \leq t\}} m(X) \text{ ist } \mathcal{G}(t) - \text{messbar } \forall t \in \mathbb{R}. \quad (8.16)$$

Es ist $X : (\Omega, \mathcal{A}) \rightarrow \mathbb{R}$ eine Stoppzeit bzgl. $\mathcal{G}(t)$, denn offensichtlich gilt $\{X \leq t\} \in \mathcal{G}(t)$. Sei

$$\mathcal{G}_X := \{A \in \mathcal{A} : A \cap \{X \leq t\} \in \mathcal{G}(t) \forall t \in \mathbb{R}\}$$

die so genannte σ -Algebra der X -Vergangenheit. Dann ist $\mathcal{G}_X$ eine σ -Algebra (klar) und X ist $\mathcal{G}_X$ -messbar, denn für alle $a \in \mathbb{R}$ gilt

$$\{X \leq a\} \cap \{X \leq t\} = \{X \leq a \wedge t\} \in \mathcal{G}(a \wedge t) \subseteq \mathcal{G}(t) \Rightarrow$$

$$\{X \leq a\} \in \mathcal{G}_X \quad \forall a \in \mathbb{R} \Rightarrow$$

Messbarkeit gemäß Satz 1.2.7 in Gänssler/Stute [GS77]. Jetzt Beweis von (8.16) per algebraischer Induktion:

1.Schritt: $m = 1_B$, $B \subseteq \mathbb{R}$ borelsch $\Rightarrow 1_{\{X \leq t\}} m(X) = 1_{\{X \in B\} \cap \{X \leq t\}}$
 $\{X \in B\} \in \mathcal{G}_X$ wegen X $\mathcal{G}_X$ -messbar $\Rightarrow \{X \in B\} \cap \{X \leq t\} \in \mathcal{G}(t)$
 Gemäß Definition von $\mathcal{G}_X \Rightarrow 1_{\{X \leq t\} \cap \{X \in B\}}$ ist $\mathcal{G}(t)$ -messbar
 $\Rightarrow$ (8.16) gilt für $m = 1_B$.

2.Schritt: Sei $m = \sum_{k=1}^l b_k 1_{B_k}$ eine einfache Funktion $\Rightarrow$

$$1_{\{X \leq t\}} m(X) = \sum_{k=1}^n b_k \underbrace{1_{\{X \leq t\}} 1_{B_k}(X)}_{\mathcal{G}(t)\text{-mb } \forall k \text{ gem 1.Schritt}}$$

Rest ist Standard.

$\Rightarrow$ (8.16) $\Rightarrow$ Adaptiertheit (8.15). Damit ist der Beweis von Satz 8.7 vollständig. $\square$

Da

$$\begin{aligned} A_n(t) - H(t) &= A_n(t) - \int_{(-\infty, t]} m(x) F(dx) \\ &= \int_{(-\infty, t]} \frac{1 - F_n(x-) - (1 - F(x-))}{1 - F(x-)} m(x) F(dx), \end{aligned}$$

folgt

$$\begin{aligned} \hat{L}_n(t) &= M_n(t) + H(t) - \int_{(-\infty, t]} \frac{F_n(x-) - F(x-)}{1 - F(x-)} m(x) F(dx) \\ \stackrel{(+)}{\Rightarrow} L_n(t) &= M_n(t) - \underbrace{\int_{(-\infty, t]} \frac{F_n(x-) - F(x-)}{1 - F(x-)} m(x) F(dx)}_{=: D_n(t)} \end{aligned}$$

Daraus folgt:

$$\begin{aligned} q_2 &\stackrel{\text{Def.}}{=} P \left(\sup_{d \leq u \leq \varepsilon} \frac{|L_n(\theta + u) - L_n(\theta)|}{u^{k+1}} \geq \frac{1}{6} L \right) \\ &\leq P \left(\sup_{d \leq u \leq \varepsilon} \frac{|M_n(\theta + u) - M_n(\theta)|}{u^{k+1}} \geq \frac{1}{12} L \right) \\ &\quad + P \left(\sup_{d \leq u \leq \varepsilon} \frac{|D_n(\theta + u) - D_n(\theta)|}{u^{k+1}} \geq \frac{1}{12} L \right) \\ &=: c_1 + c_2. \end{aligned} \tag{8.17}$$

Zur Abschätzung von c_1 betrachte

$$\overline{M}_n(u) := M_n(\theta + u) - M_n(\theta) , \quad d \leq u \leq \varepsilon.$$

$\Rightarrow (\overline{M}_n(u), \mathcal{F}_n(\theta + u) : d \leq u \leq \varepsilon)$ ist wieder ein zentriertes Martingal. Sei $\{u_l := d + (\varepsilon - d)l2^{-m} : 0 \leq l \leq 2^m\}$ die dyadische Zerlegung des Intervalls $[d, \varepsilon]$.

Setze $S_l := \overline{M}_n(u_l)$, $\mathcal{G}_l := \mathcal{F}_n(\theta + u_l)$, $1 \leq l \leq 2^m$, $S_0 := 0$, $\mathcal{G}_0 := \{\emptyset, \Omega\} \Rightarrow (S_l, \mathcal{G}_l : 0 \leq l \leq 2^m)$ ist diskretes zentriertes Martingal.

Eine analoge Argumentation wie beim Herleiten von (10) im Beweis von Satz 5.5 liefert:

$$c_1 \leq \lim_{m \rightarrow \infty} 24^2 L^{-2} \sum_{1 \leq l \leq 2^m} u_l^{-2(k+1)} E[(S_l - S_{l-1})^2] .$$

Da (S_l) Martingal, folgt für $l \geq 1$:

$$\begin{aligned} E(S_l - S_{l-1})^2 &= E(S_l^2) - E(S_{l-1}^2) \\ &= E(\overline{M}_n(u_l)^2) - E(\overline{M}_n(u_{l-1})^2) , \text{ falls } l \geq 2, \\ E(\overline{M}_n(u)^2) &= E((M_n(\theta + u) - M_n(\theta))^2) \\ &= E(M_n(\theta + u)^2) - E(M_n(\theta)^2) . \end{aligned}$$

Man muss also $E(M_n(t)^2)$ berechnen. Dazu:

$$\begin{aligned} M_n(t) &\stackrel{\text{Satz 8.7}}{=} \hat{L}_n(t) - \int_{(-\infty, t]} \frac{1 - F_n(x-)}{1 - F(x-)} m(x) F(dx) \\ &= \frac{1}{n} \sum_{i=1}^n \underbrace{\left[1_{\{X_i \leq t\}} m(X_i) - \int_{(-\infty, t]} \frac{1_{\{X_i \geq x\}}}{1 - F(x-)} m(x) F(dx) \right]}_{=: \xi_i(t) \text{ i.i.d. und zentriert, da } M_1(t) \text{ zentriert}} \\ \Rightarrow EM_n(t)^2 &= \frac{1}{n} \underbrace{E \xi_1(t)^2}_{=: \mu(t)} \end{aligned}$$

$$\begin{aligned} \mu(t) &= E(1_{\{X \leq t\}} m(X)^2) + E \left[\left(\int_{(-\infty, t]} \frac{1_{\{X \geq x\}}}{1 - F(x-)} m(x) F(dx) \right)^2 \right] \\ &\quad - 2E \left[1_{\{X \leq t\}} m(X) \cdot \int_{(-\infty, t]} \frac{1_{\{X \geq x\}}}{1 - F(x-)} m(x) F(dx) \right] \end{aligned}$$

$$\begin{aligned}
 1. \text{ E-wert} & \stackrel{\text{Trafo}}{=} \int_{(-\infty, t]} m(x)^2 F(dx) =: K(t). \\
 2. \text{ E-wert} & = E \left[\int_{(-\infty, t]} \int_{(-\infty, t]} \frac{1_{\{X \geq x\}} m(x) 1_{\{X \geq y\}} m(y)}{(1 - F(x-))(1 - F(y-))} F(dx) F(dy) \right] \\
 & \stackrel{\text{Fubini}}{=} \int_{(-\infty, t]} \int_{(-\infty, t]} \frac{[1 - F((x \vee y)-)] m(x) m(y)}{(1 - F(x-))(1 - F(y-))} F(dx) F(dy) \\
 & \stackrel{\text{Zerleg.}}{=} \int_{(-\infty, t]} \int_{(-\infty, t]} 1_{\{y \leq x\}} \frac{\ddot{\cdot}}{\ddot{\cdot}} F(dx) F(dy) + \int_{(-\infty, t]} \int_{(-\infty, t]} 1_{\{y > x\}} \frac{\ddot{\cdot}}{\ddot{\cdot}} F(dx) F(dy) \\
 & =: I_1 + I_2. \\
 I_1 & = \int_{(-\infty, t]} \int_{(-\infty, t]} 1_{\{y \leq x\}} \frac{m(x) m(y)}{1 - F(y-)} F(dx) F(dy) \\
 & = \int_{(-\infty, t]} \frac{m(y)}{1 - F(y-)} \left[\underbrace{\int_{\mathbb{R}} 1_{\{y \leq x \leq t\}} m(x) F(dx)}_{=H(t)-H(y-) \text{ vgl. Beweis von Lemma 7.2}} \right] F(dy) \\
 & = \int_{(-\infty, t]} \frac{m(y)}{1 - F(y-)} (H(t) - H(y-)) F(dy). \\
 I_2 & \stackrel{\text{analog}}{=} \int_{(-\infty, t]} \frac{m(x)}{1 - F(x-)} (H(t) - H(x)) F(dx). \\
 3. \text{ E-wert} & = \int_{(-\infty, t]} \frac{m(x)}{1 - F(x-)} (H(t) - H(x-)) F(dx) = I_1. \\
 H(x) & = \int_{(-\infty, x]} m(u) F(du) = \int_{(-\infty, x)} m(x) F(du) + \int_{\{x\}} m(u) F(du) \\
 & = H(x-) + m(x) \cdot F(\{x\}). \\
 I_2 & = \int_{(-\infty, t]} \frac{m(x)}{1 - F(x-)} (H(t) - H(x-) - m(x) F(\{x\})) F(dx).
 \end{aligned}$$

Damit folgt

$$\mu(t) = K(t) + I_1 + I_1 - \int_{(-\infty, t]} \frac{F(\{x\})}{1 - F(x-)} m(x)^2 F(dx) - 2I_1$$

$$= K(t) - \int_{-(\infty, t]} \frac{F(\{x\})}{1 - F(x-)} m(x)^2 F(dx).$$

Für alle $l \geq 2$ gilt:

$$\begin{aligned} & E(S_l - S_{l-1})^2 \\ &= E(M_n(\theta + u_l)^2) - E(M_n(\theta)^2) - E(M_n(\theta + u_{l-1})^2) + E(M_n(\theta)^2) \\ &= \frac{1}{n} \{\mu(\theta + u_l) - \mu(\theta + u_{l-1})\} \\ &= \frac{1}{n} \{K(\theta + u_l) - K(\theta + u_{l-1})\} - \underbrace{\frac{1}{n} \int_{(\theta + u_{l-1}, \theta + u_l]} \frac{F(\{x\})}{1 - F(x-)} m(x)^2 F(dx)}_{\geq 0} \\ &\leq \frac{1}{n} \int_{(\theta + u_{l-1}, \theta + u_l]} m(x)^2 F(dx) \end{aligned}$$

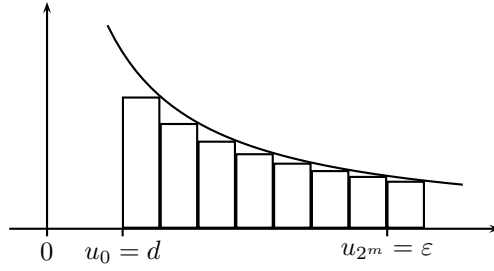
(denn m ist lokal beschränkt: $\|m\|_{\varepsilon, \theta} < \infty$)

$$\begin{aligned} &\leq \frac{1}{n} \|m\|_{\varepsilon, \theta}^2 \underbrace{\{F(\theta + u_l) - F(\theta + u_{l-1})\}}_{\leq \|f\|_{\varepsilon, \theta}(u_l - u_{l-1})} \\ &\leq C n^{-1} (u_l - u_{l-1}) \quad \forall 2 \leq l \leq 2^m. \end{aligned}$$

Nun $l = 1$:

$$\begin{aligned} E(S_1 - S_0)^2 &= E S_1^2 = E(M_n(\theta + u_1)^2) - E(M_n(\theta)^2) \\ &= \frac{1}{n} \{\mu(\theta + u_1) - \mu(\theta)\} \\ &\leq C \frac{1}{n} u_1 \end{aligned}$$

$$\begin{aligned} \Rightarrow \sum_{1 \leq l \leq 2^m} u_l^{-2(k+1)} E(S_l - S_{l-1})^2 &\leq C n^{-1} \sum_{2 \leq l \leq 2^m} u_l^{-2(k+1)} (u_l - u_{l-1}) + C n^{-1} u_1^{-2(k+1)} u_1 \\ &\leq C n^{-1} \underbrace{\sum_{1 \leq l \leq 2^m} u_l^{-2(k+1)} (u_l - u_{l-1})}_{\rightarrow \int_d^{\varepsilon} x^{-2(k+1)} dx \quad (m \rightarrow \infty)} + C n^{-1} d^{-(2k+1)}, \text{ da } u_1 > u_0 = d. \end{aligned}$$



$$\Rightarrow c_1 \leq \underset{\substack{\uparrow \\ \text{hängt nicht von } d \text{ ab}}}{C} n^{-1} d^{-(2k+1)} \quad (\text{mit } d = xn^{-\frac{1}{2k+1}})$$

und daraus folgt

$$c_1 \equiv c_1^{(n)}(x) \leq Cx^{-(2k+1)} \quad \forall x > 0 \quad \forall n \geq 1. \quad (8.18)$$

Zur Abschätzung von $c_2 \equiv c_2^{(n)}(x)$ beachte:

$$\begin{aligned} & \sup_{d \leq u \leq \varepsilon} \frac{|D_n(\theta + u) - D_n(\theta)|}{u^{k+1}} \\ &= \sup_{d \leq u \leq \varepsilon} \frac{\left| \int_{(\theta, \theta+u]} \frac{F_n(x-) - F(x-)}{1 - F(x-)} m(x) F(dx) \right|}{u^{k+1}} \\ &\leq \sup_{d \leq u \leq \varepsilon} \frac{\int_{(\theta, \theta+u]} \frac{|F_n(x-) - F(x-)|}{1 - F(x-)} |m(x)| F(dx)}{u^{k+1}} \\ &\leq \|m\|_{\varepsilon, \theta} \frac{1}{1 - F(\theta + \varepsilon)} \|F_n - F\| \sup_{d \leq u \leq \varepsilon} \frac{|F(\theta + u) - F(\theta)|}{u^{k+1}} \\ &\leq Cd^{-k} \|F_n - F\|, \quad \text{da } |F(\theta + u) - F(\theta)| \leq \|f\|_{\varepsilon, \theta} u. \end{aligned}$$

Jetzt Anwendung von

Satz 8.8. (*Dvoretzky-Kiefer-Wolfowitz-Ungleichung*)

$$P\left(\|F_n - F\| \geq \lambda\right) \leq 2 \exp\{-2n\lambda^2\} \quad \forall \lambda > 0.$$

Beweis: vgl. Shorack und Wellner [SW86].

□

$$\Rightarrow c_2 \leq P\left(\|F_n - F\| \geq \underbrace{\frac{1}{12}LC^{-1}d^k}_{=: \lambda > 0}\right) \stackrel{\text{Satz 8.8}}{\leq} 2 \exp\left\{-2n \frac{1}{144} L^2 C^{-2} d^{2k}\right\}$$

$$\stackrel{d=xn^{-\frac{1}{2k+1}}}{\stackrel{!}{=}} 2 \exp\left\{-Cn^{1-\frac{2k}{2k+1}} x^{2k}\right\}$$

$$\Rightarrow \limsup_{n \rightarrow \infty} c_2^{(n)}(x) = 0 \quad \forall x > 0$$

$$\stackrel{(8.17)+(8.18)}{\Rightarrow} \limsup_{n \rightarrow \infty} q_2^{(n)}(x) \leq Cx^{-(2k+1)} \quad \forall x > 0$$

Mit (8.7) folgt

$$\limsup_{n \rightarrow \infty} P_1^{(n)}(x) \leq Cx^{-(2k+1)} \quad \forall x > 0 \quad (8.19)$$

$$\begin{aligned} P_2(x) &= P_2^{(n)}(x) \\ &= P\left(|\bar{Y}_n| \sup_{d \leq u \leq \varepsilon} \frac{|F_n(\theta + u) - F(\theta + u) - (F_n(\theta) - F(\theta))|}{u^{k+1}} \geq \frac{1}{3}L\right). \end{aligned}$$

Da $|\bar{Y}_n| \leq |\bar{m}| + 1$ auf $\underbrace{\{|\bar{Y}_n - \bar{m}| \leq 1\}}_{\geq |\bar{Y}_n| - |\bar{m}|}$ liefert Zerlegung

$$\begin{aligned} P_2(x) &\leq P\left(\sup_{d \leq u \leq \varepsilon} \frac{|F_n(\theta + u) - F(\theta + u) - (F_n(\theta) - F(\theta))|}{u^{k+1}} \geq \frac{1}{3(|\bar{m}| + 1)}L\right) \\ &\quad + P(|\bar{Y}_n - \bar{m}| > 1) \\ &=: d_1 + d_2. \end{aligned}$$

Klar: $d_2 \rightarrow 0$ (SGGZ). Da Satz 8.7 auch für die konstante Funktion $m = 1$ gilt, lässt sich d_1 vollkommen analog zu q_2 nach oben abschätzen. $\Rightarrow$

$$\limsup_{n \rightarrow \infty} P_2^{(n)}(x) \leq Cx^{-(2k+1)} \quad (8.20)$$

$$\begin{aligned} P_3 &= P_3^{(n)}(x) \\ &= P\left(|\bar{Y}_n - \bar{m}| \sup_{d \leq u \leq \varepsilon} \frac{\overbrace{F(\theta + u) - F(\theta)}^{\leq \|f\|_{\varepsilon, \theta} u}}{u^{k+1}} \geq \frac{1}{3}L\right) \\ &\leq P\left(Cd^{-k}|\bar{Y}_n - \bar{m}| \geq \frac{1}{3}L\right) \end{aligned}$$

$$\begin{aligned}
&= P \left(\underbrace{\left[\frac{1}{\sqrt{n}} \sum_{i=1}^n (Y_i - E(Y_i)) \right]}_{\xrightarrow{L} |N(0,*)| \text{ mit VF } G \text{ stetig}} \geq \underbrace{\frac{L}{3C} d^k n^{1/2}}_{=: x_n = \tilde{C} n^{\frac{1}{2(2k+1)}} \xrightarrow{n \rightarrow \infty}} \right) \\
&= 1 - G_n(x_n -) - (1 - G(x_n)) + 1 - G(x_n) \\
&\leq \underbrace{\|G_n - G\|}_{\rightarrow 0 \text{ (Pólya)}} + 1 - \underbrace{G(x_n)}_{\rightarrow 1} \rightarrow 0 \\
&\Rightarrow \lim_{n \rightarrow \infty} P_3^{(n)}(x) = 0 \quad \forall x > 0.
\end{aligned}$$

Aus (8.19) und (8.20) folgt für $E_1 = E_1^{(n)}(x)$:

$$\limsup_{n \rightarrow \infty} P \left(E_1^{(n)}(x) \right) \leq C x^{-(2k+1)} \quad \forall x > 0.$$

Analog behandelt man E_2 (vgl. auch D. Tillich (2007) [Til07]). Dies zeigt:

$$(12) \quad \limsup_{n \rightarrow \infty} P \left(n^{\frac{1}{2k+1}} |\bar{\theta}_n^+ - \theta| \geq x \right) \leq C x^{-(2k+1)} \quad \forall x > 0.$$

Theorem 8.9. *Sei F diff-bar in $B_\varepsilon(\theta) \setminus \{\theta\}$ mit $F'(\theta \pm) > 0$ und F' beschränkt auf $B_\varepsilon(\theta) \setminus \{\theta\}$.*

(1) *Falls die Voraussetzungen von Beispiel 8.1 (nicht glatter Fall) erfüllt sind, so*

$$\limsup_{n \rightarrow \infty} P \left(n |\bar{\theta}_n - \theta| \geq x \right) \leq C x^{-1} \quad \forall x > 0.$$

(2) *Falls die Voraussetzungen von Beispiel 8.2 erfüllt sind, so*

$$\limsup_{n \rightarrow \infty} P \left(n^{\frac{1}{2k+1}} |\bar{\theta}_n - \theta| \geq x \right) \leq C x^{-(2k+1)} \quad \forall x > 0.$$

Beweis: Falls z.B. (A1) gilt, so:

$$P \left(n^{\frac{1}{2k+1}} |\bar{\theta}_n - \theta| \geq x \right) \leq P \left(n^{\frac{1}{2k+1}} |\bar{\theta}_n^+ - \theta| \geq x \right) + \underbrace{P \left(\bar{\theta}_n \neq \bar{\theta}_n^+ \right)}_{\rightarrow 0}$$

$\stackrel{(12)}{\Rightarrow}$ Behauptung. □

8.1 Verteilungskonvergenz im glatten Fall

Ziel: Herleitung der Grenzverteilung von $n^{\frac{1}{2k+1}}(\bar{\theta}_n - \theta)$. Wegen Folgerung 7.5 betrachte

$$n^{\frac{1}{2k+1}}(\bar{\theta}_n^+ - \theta) = \operatorname{argmax}_{t \in \mathbb{R}} Z_n(t),$$

$$Z_n(t) := n\beta_n \left\{ r_n \left(\theta + \frac{t}{\alpha_n} \right) - r_n(\theta) \right\}, \quad t \in \mathbb{R}.$$

Die Folgen $\alpha_n := n^{\frac{1}{2k+1}}$ und $\beta_n := \alpha_n^{-k}$ erweisen sich als "Verteilungskonvergenz erzwingend". Beachte $\beta_n \rightarrow 0$, da $k \geq 1$. Dies ist später wichtig und zeigt Unterschied zu $k = 0$ (nicht glatter Fall).

$$r_n(u) = \frac{1}{n} \sum_{i=1}^n 1_{\{X_i \leq u\}} (Y_i - \bar{Y}_n).$$

Sei $t \geq 0$:

$$\begin{aligned} Z_n(t) &= \beta_n \sum_{i=1}^n 1_{\{\theta < X_i \leq \theta + \frac{t}{\alpha_n}\}} (Y_i - m(X_i)) + \beta_n \sum_{i=1}^n 1_{\{\theta < X_i \leq \theta + \frac{t}{\alpha_n}\}} (m(X_i) - \bar{m}) \\ &\quad + \beta_n \sum_{i=1}^n 1_{\{\theta < X_i \leq \theta + \frac{t}{\alpha_n}\}} (\bar{m} - \bar{Y}_n) \\ &=: \Gamma_n(t) + \bar{\Gamma}_n(t) + R_n(t). \end{aligned}$$

Für $t < 0$:

$$\begin{aligned} \Gamma_n(t) &= -\beta_n \sum_{i=1}^n 1_{\{\theta + \frac{t}{\alpha_n} < X_i \leq \theta\}} (Y_i - m(X_i)) \\ \bar{\Gamma}_n(t) &= -\beta_n \sum_{i=1}^n 1_{\{\theta + \frac{t}{\alpha_n} < X_i \leq \theta\}} (m(X_i) - \bar{m}) \\ R_n(t) &= -\beta_n \sum_{i=1}^n 1_{\{\theta + \frac{t}{\alpha_n} < X_i \leq \theta\}} (\bar{m} - \bar{Y}_n) \\ &\Rightarrow Z_n = \Gamma_n + \bar{\Gamma}_n + R_n. \end{aligned}$$

Zeige:

$$\Gamma_n \xrightarrow{\mathcal{L}} \Gamma \text{ in } D[-a, a] \quad \forall a > 0 \quad (8.21)$$

und identifiziere den Grenzprozess Γ .

$$\sup_{-a \leq t \leq a} |\bar{\Gamma}_n(t) - \bar{\Gamma}(t)| \rightarrow 0 \text{ f.s. } \forall a > 0 \quad (8.22)$$

und identifiziere die Grenzfunktion $\bar{\Gamma}$.

$$\sup_{-a \leq t \leq a} |R_n(t)| \xrightarrow{P} 0 \quad \forall a > 0. \quad (8.23)$$

Beweis von (8.21): Anwendung von Satz 3.3

(i) Konvergenz der fidis (mit ZGWS).

Sei $t \geq 0$:

$$\begin{aligned} \Gamma_n(t) &= \sum_{i=1}^n \underbrace{\beta_n 1_{\{\theta < X_i \leq \theta + \frac{t}{\alpha_n}\}} (Y_i - m(X_i))}_{=: \xi_{in} \text{ i.i.d. und zentriert}} \\ \sum_{i=1}^n E(\xi_{in}^2) &= n\beta_n^2 E\left(1_{\{\theta < X \leq \theta + \frac{t}{\alpha_n}\}} \varepsilon^2\right) \\ &= n\beta_n^2 \int_{\theta}^{\theta + \frac{t}{\alpha_n}} E((Y - m(X))^2 | X = x) F(dx) \\ &= n\beta_n^2 \int_{\theta}^{\theta + \frac{t}{\alpha_n}} V(x) f(x) dx \\ &= \underbrace{\frac{n\beta_n^2}{\alpha_n}}_{=1} \cdot t \cdot \underbrace{\frac{\int_{\theta}^{\theta + \frac{t}{\alpha_n}} V(x) f(x) dx}{\frac{t}{\alpha_n}}}_{\rightarrow V(\theta+) f(\theta+) =: \sigma_+^2}. \end{aligned}$$

Ljapunov-Bedingung:

$$\begin{aligned} L_n &:= \sum_{i=1}^n E|\xi_{in}|^p = n\beta_n^p \int_{\theta}^{\theta + \frac{t}{\alpha_n}} E[|Y - m(X)|^p | X = x] F(dx) \\ |a + b|^p &\leq 2^{p-1} \{|a|^p + |b|^p\} \quad \forall a, b > 0 \\ \Rightarrow |Y - m(X)|^p &\leq C \{ |Y|^p + \underbrace{|E(Y|X)|^p}_{\substack{\text{Jensen} \\ \downarrow \\ \leq E[|Y|^p | X]}} \} \\ \Rightarrow \text{Integrand} &\leq 2C \underbrace{E[|Y|^p | X = x]}_{=: M_p(x)} \end{aligned}$$

$$\begin{aligned}
 \Rightarrow L_n &\leq 2Cn\beta_n^p \int_{\theta}^{\theta+\frac{t}{\alpha_n}} M_p(x)f(x)dx \\
 &= 2C \underbrace{\frac{\beta_n^2 n}{\alpha_n}}_{=1} \underbrace{\beta_n^{p-2} t}_{\rightarrow 0} \underbrace{\frac{\int_{\theta}^{\theta+\frac{t}{\alpha_n}} M_p(x)f(x)dx}{\frac{t}{\alpha_n}}}_{\rightarrow M_p(\theta+)f(\theta+)<\infty}
 \end{aligned}$$

$$\stackrel{\text{ZGWS}}{\Rightarrow} \Gamma_n(t) \xrightarrow{\mathcal{L}} N(0, \sigma_+^2 t) = \sigma_+ N(0, t)$$

Im mehrdimensionalen Fall betrachte Zuwächse, z.B. für $0 \leq s < t$

$$(\Gamma_n(s), \Gamma_n(t) - \Gamma_n(s)) = \sum_{i=1}^n \xi_{in}$$

und verwende mult. ZGWS:

$$\begin{aligned}
 \xi_{in} &= \beta_n \begin{pmatrix} 1_{\{\theta < X_i \leq \theta + \frac{s}{\alpha_n}\}} \varepsilon_i \\ 1_{\{\theta + \frac{s}{\alpha_n} < X_i \leq \theta + \frac{t}{\alpha_n}\}} \varepsilon_i \end{pmatrix} \\
 C_n &:= \sum_{i=1}^n \text{Cov}(\xi_{in}) \in M(2 \times 2) \\
 (C_n)_{11} &= n\beta_n^2 \int_{\theta}^{\theta+\frac{s}{\alpha_n}} \underbrace{E[(Y - m(X))^2 | X = x]}_{=V(x)} F(dx) \\
 &= \frac{n\beta_n^2}{\alpha_n} s \frac{\int_{\theta}^{\theta+\frac{s}{\alpha_n}} V(x)f(x)dx}{\frac{s}{\alpha_n}} \rightarrow sV(\theta+)f(\theta+) \\
 (C_n)_{22} &= n\beta_n^2 \int_{\theta+\frac{s}{\alpha_n}}^{\theta+\frac{t}{\alpha_n}} V(x)f(x)dx \\
 &= \frac{n\beta_n^2}{\alpha_n} \left(t \frac{\int_{\theta}^{\theta+\frac{t}{\alpha_n}} V(x)f(x)dx}{\frac{t}{\alpha_n}} - s \frac{\int_{\theta}^{\theta+\frac{s}{\alpha_n}} V(x)f(x)dx}{\frac{s}{\alpha_n}} \right) \\
 &\rightarrow V(\theta+)f(\theta+)t - V(\theta+)f(\theta+)s \\
 &= \sigma_+^2(t - s) \\
 (C_n)_{12} &= 0,
 \end{aligned}$$

wegen Disjunktheit der beiden Indikatormengen.

$$\Rightarrow C_n \rightarrow C = \begin{pmatrix} \sigma_+^2 s & 0 \\ 0 & \sigma_+^2 (t-s) \end{pmatrix},$$

Sei $\Gamma(t) := \sigma_+ B_1(t)$, $t \geq 0$, wobei B_1 eine BB auf $\mathbb{R} \xrightarrow{\text{ZGWS}}$

$$(\Gamma_n(s), \Gamma_n(t) - \Gamma_n(s)) \xrightarrow{\mathcal{L}} (\Gamma(s), \Gamma(t) - \Gamma(s)).$$

Sei nun $s < 0 \leq t$:

$$(-\Gamma_n(s), \Gamma_n(t)) = \sum_{i=1}^n \xi_{in}$$

$$\begin{aligned} \xi_{in} &= \beta_n (1_{\{\theta + \frac{s}{\alpha_n} < X_i \leq \theta\}} \varepsilon_i, 1_{\{\theta < X_i \leq \theta + \frac{t}{\alpha_n}\}} \varepsilon_i) \\ C_n &:= \sum_{i=1}^n \text{Cov}(\xi_{in}) \rightarrow C = \begin{pmatrix} -\sigma_-^2 s & 0 \\ 0 & \sigma_+^2 t \end{pmatrix}, \end{aligned}$$

wobei $\sigma_-^2 = V(\theta-)f(\theta-)$. Sei $\Gamma(t) := -\sigma_- B_2(-t)$, $t \leq 0$, und B_2 eine BB, unabhängig von $B_1 \Rightarrow$

$$C = \text{Cov} \begin{pmatrix} -\Gamma(s) & 0 \\ 0 & \Gamma(t) \end{pmatrix}.$$

Nachweis der Ljapunov Bedingung geht analog zum 1-dimensionalen Fall (Man verwende die L_1 -Norm $\|(x, y)\| = |x| + |y|$). Analog höhere fidis. Insbesondere starke Vermutung:

$$\Gamma(t) := \begin{cases} \sigma_+ B_1(t), & t \geq 0 \\ -\sigma_- B_2(-t), & t < 0 \end{cases} \quad (8.24)$$

ist unser gesuchter Grenzprozess.

(ii) Bleibt Straffheit von (Γ_n) : Nachweis mit Satz 3.3. Seien $s < t < u$ aus $[-a, a]$, z.B. $s < 0 < t < u$:

$$\begin{aligned} & E \left[\underbrace{|\Gamma_n(t) - \Gamma_n(s)|^2}_{= \beta_n \sum_{i=1}^n 1_{\{\theta + \frac{s}{\alpha_n} < X_i \leq \theta + \frac{t}{\alpha_n}\}}} \cdot \underbrace{|\Gamma_n(u) - \Gamma_n(t)|^2}_{= \beta_n \sum_{i=1}^n 1_{\{\theta + \frac{t}{\alpha_n} < X_i \leq \theta + \frac{u}{\alpha_n}\}}} \right] \\ &= E \left[\beta_n^4 \left(\sum_{i=1}^n 1_I(X_i) \varepsilon_i \right)^2 \left(\sum_{j=1}^n 1_J(X_j) \varepsilon_j \right)^2 \right], \end{aligned}$$

wobei $I := \left(\theta + \frac{s}{\alpha_n}, \theta + \frac{t}{\alpha_n}\right]$, $J := \left(\theta + \frac{t}{\alpha_n}, \theta + \frac{u}{\alpha_n}\right]$, $I \cap J = \emptyset$

$$= \beta_n^4 \sum_{1 \leq i, j, k, l \leq n} \underbrace{E \left[1_I(X_i) \varepsilon_i \cdot 1_I(X_j) \varepsilon_j \cdot 1_J(X_k) \varepsilon_k \cdot 1_J(X_l) \varepsilon_l \right]}_{=: \mu_{ijkl}}.$$

Sei $(i, j, k, l) \in \{1, \dots, n\}^4$ mit mindestens 3 verschiedenen Komponenten. Dann existiert mind. eine Komponente, die von allen anderen verschieden ist. Der Erwartungswert dieser Komponente kann wegen der Unabhängigkeit als Faktor abgespalten werden, z.B. der i -te:

$$E(1_I(X_i) \varepsilon_i) = 0$$

(bedingte Zentriertheit der ε_i). Damit $\mu_{ijkl} = 0$, d.h. für diese Tupel verschwinden die μ_{ijkl} . Aus demselben Grund gilt:

$$\mu_{iiij} = \mu_{iiji} = \mu_{ijji} = \mu_{jjii} = 0.$$

Weiter gilt wegen der Disjunktheit von I und J :

$$\mu_{iiii} = \mu_{ijij} = \mu_{ijji} = 0.$$

$$\begin{aligned} \Rightarrow \sum_{1 \leq i, j, k, l \leq n} &= \sum_{1 \leq i \neq j \leq n} \{\mu_{iiij} + \mu_{jjii}\} \\ \mu_{iiij} &\stackrel{\text{Unabh.}}{=} E \left[1_I(X_i) \varepsilon_i^2 1_J(X_j) \varepsilon_j^2 \right] \\ &= \int_{\theta + \frac{s}{\alpha_n}}^{\theta + \frac{t}{\alpha_n}} V(x) f(x) dx \cdot \int_{\theta + \frac{t}{\alpha_n}}^{\theta + \frac{u}{\alpha_n}} V(x) f(x) dx. \end{aligned}$$

Da $s, t, u \in [-a, a]$, liegen die Integrationsbereiche I und J in $B_\varepsilon(\theta) \forall n \geq n_0(a, \varepsilon) \in \mathbb{N} \Rightarrow$

$$\begin{aligned} &E \left[|\Gamma_n(t) - \Gamma_n(s)|^2 |\Gamma_n(u) - \Gamma_n(t)|^2 \right] \\ &\leq 2 \underbrace{\frac{\beta_n^4 n^2}{\alpha_n^2}}_{=1} \|V\|_{\varepsilon, \theta}^2 \|f\|_{\varepsilon, \theta}^2 (t-s)(u-t) \\ &\leq C^2 (u-s)^2 = (Cu - Cs)^2. \end{aligned}$$

Andere Konstellationen von s, t, u analog.

$$\stackrel{\text{Satz 3.3}}{\Rightarrow} \Gamma_n \xrightarrow{\mathcal{L}} \Gamma \quad \text{in } D[-a, a] \quad \forall a > 0,$$

wobei Γ aus (8.24). Dies zeigt (8.21).

Zu (8.22): Fixiere $t \in [0, a]$.

$$\bar{\Gamma}_n(t) = \underbrace{\bar{\Gamma}_n(t) - E(\bar{\Gamma}_n(t))}_{=: V_n(t)} + E(\bar{\Gamma}_n(t)).$$

Zeige

$$V_n(t) \rightarrow 0 \quad \text{f.s. } \forall t \in [0, a] \quad (8.25)$$

$$V_n(t) = \beta_n \underbrace{\sum_{i=1}^n \left(\underbrace{1_{\{\theta < X_i \leq \theta + \frac{t}{\alpha_n}\}}}_{=: Z_i} [m(X_i) - \bar{m}] - E(Z_i) \right)}_{=: \rho_i \text{ i.i.d und zentriert}}$$

$$\begin{aligned} P(|V_n(t)| > \varepsilon) &\stackrel{\text{Markov}}{\leq} \varepsilon^{-p} E|V_n(t)|^p \quad \forall p > 2 \\ &= \varepsilon^{-p} \beta_n^p E \left| \sum_{i=1}^n \rho_i \right|^p. \end{aligned}$$

Satz 8.10. (Ungleichung von Marcinkiewicz-Zygmund) Seien $\xi_1, \dots, \xi_n$ unabhängig und zentriert und p -fach integrierbar ($p > 2$). Dann:

$$E \left| \sum_{i=1}^n \xi_i \right|^p \leq C_p n^{\frac{p-2}{2}} \sum_{i=1}^n E|\xi_i|^p.$$

Falls die (ξ_i) zusätzlich identisch verteilt, so:

$$E \left| \sum_{i=1}^n \xi_i \right|^p \leq C_p n^{\frac{p}{2}} E|\xi_1|^p.$$

Beweis: vgl. Chow und Teicher [CT97], p.386. □

$$\Rightarrow E|V_n(t)|^p \leq C_p \beta_n^p n^{\frac{p}{2}} E|\rho_1|^p$$

$$\begin{aligned} E|\rho_1|^p &= E \left(\underbrace{|Z_1 - E(Z_1)|^p}_{\leq 2^{p-1} \{ |Z_1|^p + |E(Z_1)|^p \}} \right) \leq 2^{p-1} \{ E|Z_1|^p + \underbrace{|E(Z_1)|^p}_{\leq E|Z_1|^p} \} \\ &\leq 2^p E|Z_1|^p. \end{aligned}$$

$$\begin{aligned}
E|Z_1|^p &= E \left[1_{\{\theta < X \leq \theta + \frac{t}{\alpha_n}\}} |m(X) - \underbrace{m(\theta)}_{=\overline{m}}|^p \right] \\
&= \int_{\theta}^{\theta + \frac{t}{\alpha_n}} \underbrace{|m(x) - \overline{m}|^p}_{\stackrel{8.2}{=} \frac{1}{k!} m^{(k)}(z)(x-\theta)^k \quad (z \in B_\varepsilon(\theta))} f(x) dx \\
&\leq C \int_{\theta}^{\theta + \frac{t}{\alpha_n}} (x - \theta)^{pk} dx = C \int_0^{\frac{t}{\alpha_n}} u^{pk} du \\
&\leq C \frac{t^{pk+1}}{\alpha_n^{pk+1}} \\
&\stackrel{t \leq a}{\leq} C \alpha_n^{-pk-1}
\end{aligned}$$

$$\begin{aligned}
\Rightarrow E|V_n(t)|^p &\leq C n^{\frac{p}{2}} \beta_n^p \alpha_n^{-pk-1} = C n^{\frac{p}{2}} \alpha_n^{-(2pk+1)} \\
&= C n^{-\left(\frac{2pk+1}{2k+1} - \frac{p}{2}\right)}
\end{aligned}$$

$$\Rightarrow \sum_{n \geq 1} P(|V_n(t)| > \varepsilon) < \infty \quad \forall \varepsilon > 0 ,$$

falls

$$\frac{2pk+1}{2k+1} - \frac{p}{2} > 1 \Leftrightarrow p > \frac{4k}{2k-1} \quad (\text{Forderung an } p).$$

Mit 1. Borel-Cantelli-Lemma folgt (8.25), also:

$$V_n(t) \rightarrow 0 \quad \text{f.s. } \forall t \in [0, a]$$

$$\begin{aligned}
E\overline{\Gamma}_n(t) &= n\beta_n E \left[1_{\{\theta < X \leq \theta + \frac{t}{\alpha_n}\}} (m(X) - \overline{m}) \right] \\
&= n\beta_n \int_{\theta}^{\theta + \frac{t}{\alpha_n}} m(x) - \overline{m} F(dx) \\
&= n\beta_n \left\{ \rho \left(\theta + \frac{t}{\alpha_n} \right) - \rho(\theta) \right\} \\
&\stackrel{\text{vgl. Bew. von 8.4}}{=} \frac{n\beta_n}{\alpha_n^{k+1}} \underbrace{\frac{\rho^{(k+1)}(z_n)}{(k+1)!}}_{\rightarrow \frac{m^{(k)}(\theta+)}{(k+1)!} f(\theta+), \text{ da } z_n \downarrow \theta \text{ für } n \rightarrow \infty} |t|^{k+1} \quad \text{für } z_n \in (\theta, \theta + \frac{t}{\alpha_n}) \quad \forall t \in [0, a]
\end{aligned}$$

$$\stackrel{(8.25)}{\Rightarrow} \bar{\Gamma}_n(t) \rightarrow \bar{\Gamma}(t) := \frac{m^{(k)}(\theta+)}{(k+1)!} f(\theta+) t^{k+1} \quad \text{f.s. für } n \rightarrow \infty \quad \forall t \in [0, a] \quad \forall a > 0.$$

Wegen (7.3) ist $\bar{\Gamma}_n$ monoton fallend auf $[0, a]$. Da $m^{(k)}(\theta+) < 0$, ist $\bar{\Gamma}$ ebenfalls monoton fallend und als Monom stetig. Somit folgt mit Satz (4.12):

$$\sup_{0 \leq t \leq a} |\bar{\Gamma}_n(t) - \bar{\Gamma}(t)| \rightarrow 0 \quad \text{f.s. } \forall a > 0.$$

Analog für $[-a, 0] \Rightarrow$

$$\sup_{-a \leq t \leq a} |\bar{\Gamma}_n(t) - \bar{\Gamma}(t)| \rightarrow 0 \quad \text{f.s. } \forall a > 0,$$

wobei

$$\bar{\Gamma}(t) = \begin{cases} \frac{1}{(k+1)!} f(\theta+) m^{(k)}(\theta+) |t|^{k+1}, & t \geq 0 \\ -\frac{1}{(k+1)!} f(\theta-) (-1)^k m^{(k)}(\theta-) |t|^{k+1}, & t < 0 \end{cases}$$

$\Rightarrow$ (8.22).

Beweis von (8.23):

$$\begin{aligned} |R_n(t)| &\leq |\bar{Y}_n - \bar{m}| \beta_n \underbrace{\sum_{i=1}^n 1_{\{\theta < X_i \leq \theta + \frac{a}{\alpha_n}\}}}_{=: S_n \geq 0} \quad \forall t \in [0, a] \\ &\Rightarrow \sup_{0 \leq t \leq a} |R_n(t)| \leq \underbrace{n^\gamma |\bar{Y}_n - \bar{m}|}_{=: U_n} \underbrace{n^{-\gamma} \beta_n S_n}_{=: V_n}. \end{aligned}$$

Wegen $\{U_n V_n \geq \varepsilon\} \subseteq \{U_n \geq \sqrt{\varepsilon}\} \cup \{V_n \geq \sqrt{\varepsilon}\} \quad \forall \varepsilon > 0$ folgt:

$$P\left(\sup_{0 \leq t \leq a} |R_n(t)| \geq \varepsilon\right) \leq \underbrace{P\left(|U_n| \geq \sqrt{\varepsilon}\right)}_{=: A_n} + \underbrace{P\left(|V_n| \geq \sqrt{\varepsilon}\right)}_{=: B_n}$$

mit

$$\begin{aligned} A_n &= P\left(\left|\sum_{i=1}^n (Y_i - \bar{m})\right| > n^{1-\gamma} \sqrt{\varepsilon}\right) \\ &\stackrel{\text{Tschebyscheff}}{\leq} \varepsilon^{-1} n^{2\gamma-2} n \text{Var}(Y) \\ &\rightarrow 0, \text{ falls } 2\gamma - 1 < 0 \Leftrightarrow \gamma < \frac{1}{2}. \end{aligned}$$

$$\begin{aligned} S_n &= S_n - E(S_n) + \underbrace{E(S_n)}_{=n\{F(\theta + \frac{a}{\alpha_n}) - F(\theta)\}} \\ &\leq S_n - E(S_n) + Cn\alpha_n^{-1} \end{aligned}$$

$$\Rightarrow B_n \leq P\left(S_n - E(S_n) > \frac{1}{2}n^\gamma \beta_n^{-1} \sqrt{\varepsilon}\right) + P\left(\underbrace{\left\{Cn^{1-\gamma} \alpha_n^{-1} \beta_n > \frac{1}{2} \sqrt{\varepsilon}\right\}}_{\text{unabhängig von } \omega}\right).$$

$\{Cn^{1-\gamma} \alpha_n^{-1} \beta_n > \frac{1}{2} \sqrt{\varepsilon}\} = \emptyset \ \forall \ n \geq n_0(\varepsilon) \in \mathbb{N}$, denn:

$$n^{1-\gamma} \alpha_n^{-1} \beta_n = n^{1-\gamma-\frac{1}{2k+1}-\frac{k}{2k+1}} = n^{1-\gamma-\frac{k+1}{2k+1}} \rightarrow 0,$$

falls

$$1 - \gamma - \frac{k+1}{2k+1} < 0 \Leftrightarrow \gamma > 1 - \frac{k+1}{2k+1} = \frac{k}{2k+1}.$$

Klar: (i) $\gamma \in (\frac{k}{2k+1}, \frac{1}{2}) \neq \emptyset$ existiert für alle $k \geq 1 \Rightarrow 2$. W-keit in B_n ist sogar gleich 0.

$$\begin{aligned} 1. \text{ W-keit in } B_n &\stackrel{\text{Tsehebyscheff}}{\leq} 4n^{-2\gamma} \beta_n^2 \varepsilon^{-1} \text{Var}(S_n) \\ &\leq Cn^{1-2\gamma-\frac{2k}{2k+1}-\frac{1}{2k+1}} \\ &= Cn^{-2\gamma} \rightarrow 0. \end{aligned}$$

(ii) Wähle γ aus (i) $\Rightarrow A_n + B_n \rightarrow 0 \ \forall \ \varepsilon > 0$

$$\Rightarrow \sup_{0 \leq t \leq a} |R_n(t)| \xrightarrow{P} 0 \quad \forall a > 0.$$

Analog zeigt man

$$\sup_{-a \leq t \leq 0} |R_n(t)| \xrightarrow{P} 0 \quad \forall a > 0.$$

$\Rightarrow$ (8.23) wie gewünscht.

$$Z_n = \underset{\downarrow \mathcal{L}}{\Gamma_n} + \underset{\downarrow \overline{\Gamma}}{\overline{\Gamma}_n} + \underset{\downarrow 0}{R_n} \in D[-a, a]$$

Aus (8.21)-(8.23) ergibt sich mit Slutskys Lemma:

Satz 8.11. $Z_n \xrightarrow{\mathcal{L}} Z$ in $D[-a, a] \ \forall a > 0$, wobei

$$Z(t) = \Gamma(t) + \overline{\Gamma}(t) = \begin{cases} \sigma_+ B_1(t) + \frac{m^{(k)}(\theta_+) f(\theta_+)}{(k+1)!} |t|^{k+1}, & t \geq 0 \\ -\sigma_- B_2(-t) - \frac{(-1)^k m^{(k)}(\theta_-) f(\theta_-)}{(k+1)!} |t|^{k+1}, & t < 0 \end{cases}$$

$$\sigma_{\pm} = f(\theta_{\pm}) V(\theta_{\pm}),$$

mit $V(x) = E[(Y - m(X))^2 | X = x]$.

Bemerkung: Wegen $B_2 \stackrel{\mathcal{L}}{=} -B_2$ kann man das Minuszeichen bei σ_- auch weglassen.

Frage: Besitzt Z eine eindeutige Maximalstelle (f.s.)?

Satz 8.12. (*Lifshits [Lif82]*) Sei $M = \{M(t) : t \in \mathbb{R}^d\}$ ein Gauß-Prozess mit fast sicher stetigen Pfaden. Falls

- (1) $E[M(t)] \rightarrow -\infty$ für $\|t\| \rightarrow \infty$ und
- (2) $E[(M(s) - M(t))^2] \neq 0 \forall s \neq t$,

dann besitzt M f.s. eine eindeutige Maximalstelle.

In unserem Fall: $E(Z(t)) = \bar{\Gamma}(t) \rightarrow -\infty$ für $|t| \rightarrow \infty$ wegen $m^{(k)}(\theta+) < 0$ und $(-1)^k m^{(k)}(\theta-) > 0$ nach Annahme $\Rightarrow$ (1).

Zu (2): Zeige

$$K(s, t) := E[(Z(s) - Z(t))^2] = 0 \Rightarrow s = t.$$

Dazu (i) $s \geq t \geq 0$:

$$\begin{aligned} 0 &= K(s, t) = E[\{\sigma_+(B_1(s) - B_1(t)) + \bar{\Gamma}(s) - \bar{\Gamma}(t)\}^2] \\ &= \sigma_+^2 \underbrace{E(B_1(s) - B_1(t))^2}_{=(s-t) \geq 0} + 0 + (\bar{\Gamma}(s) - \bar{\Gamma}(t))^2 \\ &\geq \sigma_+(s - t) \geq 0 \quad \Rightarrow s = t. \end{aligned}$$

(ii) $s \geq 0 \geq t$:

$$\begin{aligned} 0 &= K(s, t) = E[\{\sigma_+ B_1(s) - \sigma_- B_2(-t) + \bar{\Gamma}(s) - \bar{\Gamma}(t)\}^2] \\ &= E[(\sigma_+ B_1(s) - \sigma_- B_2(-t))^2] + (\bar{\Gamma}(s) - \bar{\Gamma}(t))^2 \\ &= \sigma_+^2 s - 2 \underbrace{E(B_1(s) B_2(t-))}_{=0 \text{ wegen } B_1, B_2 \text{ unabh.}} + \sigma_+ \sigma_- + \sigma_-^2 (-t) + \underbrace{(\bar{\Gamma}(s) - \bar{\Gamma}(t))^2}_{\geq 0} \\ &\geq (\sigma_+^2 \wedge \sigma_-^2)(s - t) \geq 0 \quad \Rightarrow s = t. \end{aligned}$$

(iii) $0 \geq s \geq t$: wie (i)

Somit liefert Satz 8.12:

Satz 8.13. Z besitzt f.s. eine eindeutige Maximalstelle $\tau = \operatorname{argmax}_{t \in \mathbb{R}} Z(t)$.

Theorem 8.9, Satz 8.11 und Satz 8.13 ermöglichen die Anwendung des Argmax-CMT in der Form von Folgerung 4.13. Es folgt

$$n^{\frac{1}{2k+1}}(\bar{\theta}_n^+ - \theta) \xrightarrow{\mathcal{L}} \tau, \quad (n \rightarrow \infty). \quad (8.26)$$

Mit Folgerung 7.5 und Portmanteau-Theorem überträgt sich das Ergebnis auf $\bar{\theta}_n$:
Sei $F \subseteq \mathbb{R}$ abgeschlossen $\Rightarrow$

$$\begin{aligned} \limsup_{n \rightarrow \infty} \underbrace{P\left(n^{\frac{1}{2k+1}}(\bar{\theta}_n - \theta) \in F\right)}_{\leq P\left(n^{\frac{1}{2k+1}}(\bar{\theta}_n^+ - \theta) \in F\right) + P(\bar{\theta}_n^+ \neq \bar{\theta})} &\leq \limsup_{n \rightarrow \infty} P\left(n^{\frac{1}{2k+1}}(\bar{\theta}_n^+ - \theta) \in F\right) \\ &\stackrel{(8.26)}{\leq} P(\tau \in F), \end{aligned}$$

und es folgt

Theorem 8.14. *Sei m wie in Beispiel 8.2. In der beliebig kleinen gelochten Umgebung $B_\varepsilon(\theta) \setminus \{\theta\}$ gelte:*

- (1) f k -mal stetig diff-bar und $f(\theta \pm) > 0$.
- (2) $V(x) := E((Y - m(X))^2 | X = x)$ beschränkt.
- (3) $\exists p > 2 : M_p(x) := E[|Y|^p | X = x]$ beschränkt.

Dann folgt:

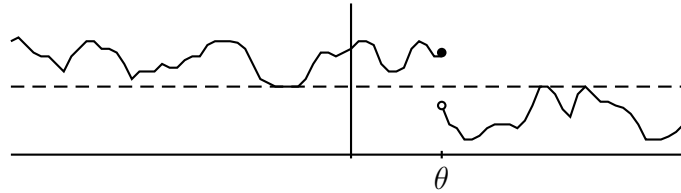
$$n^{\frac{1}{2k+1}}(\bar{\theta}_n - \theta) \xrightarrow[t \in \mathbb{R}]{\mathcal{L}} \operatorname{argmax} Z(t).$$

Bemerkung. • $k = 1 \rightarrow$ „cube root asymptotics“ (in Kim und Pollard [KP90]), Z = Brownsche Bewegung mit parabolischem Drift, Grenzverteilung von τ analytisch bekannt (siehe Groenebloom [Gro89]).

- $k \geq 1 \rightarrow$ „ $(2k + 1)$ -th root asymptotics“.

8.2 Grenzwertsätze im nicht-glaten Fall

Es gelten die Voraussetzungen von Beispiel 8.1



$$n(\bar{\theta}_n^+ - \theta) = \operatorname{argmax}_{t \in \mathbb{R}} Z_n(t)$$

$$\begin{aligned} Z_n(t) &= n \left\{ r_n \left(\theta + \frac{t}{n} \right) - r_n(\theta) \right\} \\ &= \begin{cases} \sum_{i=1}^n 1_{\{\theta < X_i \leq \theta + \frac{t}{n}\}} (Y_i - \bar{Y}_n), & t \leq 0 \\ \sum_{i=1}^n 1_{\{\theta + \frac{t}{n} < X_i \leq \theta\}} (Y_i - \bar{Y}_n), & t < 0 \end{cases} \end{aligned}$$

Betrachte Z_n mit $\bar{Y}_n$ ersetzt durch $\bar{m}$, also

$$\tilde{Z}_n(t) := \begin{cases} \sum_{i=1}^n 1_{\{\theta < X_i \leq \theta + \frac{t}{n}\}} (Y_i - \bar{m}), & t \leq 0 \\ \sum_{i=1}^n 1_{\{\theta + \frac{t}{n} < X_i \leq \theta\}} (Y_i - \bar{m}), & t < 0 \end{cases}$$

Dann gilt für $R_n := Z_n - \tilde{Z}_n$:

$$\sup_{-a \leq t \leq a} |R_n(t)| \leq \max \left\{ |\bar{Y}_n - \bar{m}| \sum_{i=1}^n 1_{\{\theta < X_i \leq \theta + \frac{a}{n}\}}, |\bar{Y}_n - \bar{m}| \sum_{i=1}^n 1_{\{\theta - \frac{a}{n} < X_i \leq \theta\}} \right\}.$$

Die beiden Summen konvergieren in Verteilung gemäß Poissonschen Grenzwertsatz. Klar $\bar{Y}_n - \bar{m} \xrightarrow{P} 0 \Rightarrow$

$$\sup_{-a \leq t \leq a} |R_n(t)| \xrightarrow{P} 0. \quad (8.27)$$

Wegen Cramér betrachte $(\tilde{Z}_n)$. Man kann zeigen (Anwendung von Satz 3.3, dabei Konvergenz der fidis mittels charakteristischer Funktionen, z.B. wie im Nachweis von (6.2), alternativ auch mittels Konkomitanten):

$$\tilde{Z}_n \xrightarrow{\mathcal{L}} Z \quad \text{in } D[-a, a] \quad \forall a > 0, \quad (8.28)$$

dabei

$$Z(t) = \begin{cases} \sum_{i=1}^{N_1(t)} (\xi_i - \bar{m}), & t \geq 0 \\ - \sum_{i=1}^{N_2(-t)} (\eta_i - \bar{m}), & t < 0 \end{cases}$$

wobei $\{N_1, N_2, (\xi_i), (\eta_i)\}$ unabhängig sind.

$N_1 =$ Poissonprozess zu $f(\theta+)$, η_i i.i.d. $\sim Q_1(\theta, \cdot)$,

$N_2 =$ Poissonprozess zu $f(\theta-)$, ξ_i i.i.d. $\sim Q_2(\theta, \cdot)$.

$Q_j = \{Q_j(x, \cdot) : x \in \mathbb{R}\}$, $j = 1, 2$ sind zwei Familien von W-maßen mit

$$Q(x, \cdot) := P(Y \in \cdot | X = x) = \begin{cases} Q_1(x, \cdot), & x < \theta \\ Q_2(x, \cdot), & x > \theta \end{cases}$$

Damit

$$m(x) = E[Y|X = x] = \int yQ(x, dy). \quad (8.29)$$

Beachte, dass $Q = \{Q(x, \cdot) : x \in \mathbb{R}\}$ die Regressionsfunktion m über (8.29) eindeutig festlegt. Die Umkehrung gilt nicht, weil im Allgemeinen die ersten Momente nicht die Verteilung festlegen.

Beispiel.

$$Q(x, \cdot) = N(m(x), \sigma^2(x))$$

$$\tilde{Q}(x, \cdot) = \text{Bin}(1, m(x))$$

$$\Rightarrow \int yQ(x, dy) = m(x) = \int y\tilde{Q}(x, dy) \quad \forall x \in \mathbb{R},$$

aber $Q(x, \cdot) \neq \tilde{Q}(x, \cdot)$ sogar für alle $x \in \mathbb{R}$.

Wegen (8.27) und (8.28) gilt

$$Z_n \xrightarrow{\mathcal{L}} Z \quad \text{in } D[-a, a] \quad \forall a > 0.$$

Gemäß Theorem 8.9 gilt

$$n(\bar{\theta}_n^+ - \theta) = O_P(1). \quad (8.30)$$

Da

$$\begin{aligned} E(\xi_1 - \bar{m}) &= \int yQ_2(\theta, dy) - \bar{m} \\ &= m(\theta+) - \bar{m} < 0 \text{ nach Annahme in Beispiel 8.1} \end{aligned}$$

und

$$\begin{aligned} E(\eta_1 - \bar{m}) &= \int yQ_1(\theta, dy) - \bar{m} \\ &= m(\theta-) - \bar{m} > 0 \text{ nach Annahme in Beispiel 8.1,} \end{aligned}$$

folgt dass

$$Z(t) \rightarrow -\infty \quad \text{f.s. für } |t| \rightarrow \infty.$$

Mit Satz 4.9 folgt $\emptyset \neq A(Z)$ abgeschlossen (wegen $\delta(Z) > 0$) und beschränkt
Folg. 7.5+(8.30)
 $\Rightarrow$

Theorem 8.15. *Es gelten die Voraussetzungen von Beispiel 8.1. Dann folgt:*

$$n(\bar{\theta}_n - \theta) \rightarrow_{a.w.} (\sigma, \tau),$$

wobei:

$$\sigma = \text{kleinste Max.-stelle von } \pm Z,$$

$$\tau = \text{größte Max.-stelle von } \pm Z,$$

je nachdem ob $(+) = (A1)$ oder $(-) = (A2)$ gilt.

Literaturverzeichnis

- [Als98] Gerold Alsmeyer. *Wahrscheinlichkeitstheorie*. Münster: Inst. f. Math. Statistik, 1998.
- [Bau90] Heinz Bauer. *Maß- und Integrationstheorie*. de Gruyter Lehrbuch. Walter de Gruyter & Co., Berlin, 1990.
- [Bil99] Patrick Billingsley. *Convergence of probability measures*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons Inc., New York, second edition, 1999. A Wiley-Interscience Publication.
- [CH88] M. Csörgő and L. Horváth. Convergence of the empirical and quantile distributions to Poisson measures. *Statist. Decisions*, 6(1-2):129–136, 1988.
- [CT97] Yuan Shih Chow and Henry Teicher. *Probability theory*. Springer Texts in Statistics. Springer-Verlag, New York, third edition, 1997.
- [Don52] Monroe D. Donsker. Justification and extension of Doob’s heuristic approach to the Kolmogorov-Smirnov theorems. *Ann. Math. Statistics*, 23:277–281, 1952.
- [DS02] Astrid Dempfle and Winfried Stute. Nonparametric estimation of a discontinuity in regression. *Statist. Neerlandica*, 56(2):233–242, 2002.
- [FV07] Dietmar Ferger and Daniel Vogel. Weak convergence of the empirical process and the rescaled empirical process in the skorokhod product space, 2007. Preprint.
- [Gli33] V. Glivenko. Sulla determinazione empirica della leggi di probabilita. *Giorn. Ist. Ital. Attuari*, IV:92–99, 1933.
- [Gro89] Piet Groeneboom. Brownian motion with a parabolic drift and Airy functions. *Probab. Theory Related Fields*, 81(1):79–109, 1989.

- [GS77] Peter Gänsler and Winfried Stute. *Wahrscheinlichkeitstheorie*. Springer-Verlag, Berlin, 1977.
- [GS87] Peter Gaenssler and Winfried Stute. *Seminar on empirical processes*, volume 9 of *DMV Seminar*. Birkhäuser Verlag, Basel, 1987.
- [Höh01] Cornelia Höhne. Konsistenzraten für Change Point Schätzer, 2001. Diplomarbeit, Technische Universität Dresden.
- [JS03] Jean Jacod and Albert N. Shiryaev. *Limit theorems for stochastic processes*, volume 288 of *Grundlehren der Mathematischen Wissenschaften*. Springer-Verlag, Berlin, second edition, 2003.
- [Klo02] Jens Klotzsche. Eigenschaften von nichtparametrischen Schätzern von Changepoints in der Regression, 2002. Diplomarbeit, Technische Universität Dresden.
- [KP90] Jean Kyung Kim and David Pollard. Cube root asymptotics. *Ann. Statist.*, 18(1):191–219, 1990.
- [Lif82] Mikhail A. Lifshits. Absolute continuity of the distributions of functionals of random processes. *Teor. Veroyatnost. i Primenen.*, 27(3):559–566, 1982.
- [Ros52] Murray Rosenblatt. Remarks on a multivariate transformation. *Ann. Math. Statistics*, 23:470–472, 1952.
- [Sch96] Klaus D. Schmidt. *Lectures on risk theory*. Teubner Skripten zur Mathematischen Stochastik. B. G. Teubner, Stuttgart, 1996.
- [SW86] Galen R. Shorack and Jon A. Wellner. *Empirical processes with applications to statistics*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons Inc., New York, 1986.
- [SW93] Winfried Stute and Jane-Ling Wang. The strong law under random censorship. *Ann. Statist.*, 21(3):1591–1607, 1993.
- [Til07] Daniel Tillich. Schätzung der Sprungstelle bei einstufigen Regressionsfunktionen, 2007. Diplomarbeit, Technische Universität Dresden.
- [Vog05] Daniel Vogel. Weak convergence of the empirical process and the rescaled empirical distribution function in the skorokhod product space, 2005. Diplomarbeit, Technische Universität Dresden.
- [Wil91] David Williams. *Probability with martingales*. Cambridge Mathematical Textbooks. Cambridge University Press, Cambridge, 1991.

- [WMF95] Hermann Witting and Ulrich Müller-Funk. *Mathematische Statistik. II*. B. G. Teubner, Stuttgart, 1995. Asymptotische Statistik: parametrische Modelle und nichtparametrische Funktionale.
- [Yan77] Shie-Shien Yang. General distribution theory of the concomitants of order statistics. *Ann. Statist.*, 5(5):996–1002, 1977.

DRESDNER SCHRIFTEN ZUR MATHEMATISCHEN STOCHASTIK

- 6/2009 **Anja Voß-Böhme**
A Core for Generators of Interacting Particle Systems
- 5/2009 **Dietmar Ferger**
Minimum distance estimation in normed linear spaces with Donsker-classes
- 4/2009 **Anja Voß-Böhme, Andreas Deutsch**
The cellular basis of cell sorting kinetics
- 3/2009 **Tobias Klauß, Anja Voß-Böhme**
Two types of phase transitions in an interacting particle system for collective migration
- 2/2009 **Anja Voß-Böhme**
Gibbsian characterization for the reversible measures of interacting particle systems
- 1/2009 **Dietmar Ferger, Jens Klotsche**
Estimation of split-points in binary regression trees
- 3/2008 **Dietmar Ferger**
Argmax-stable marked empirical processes
- 2/2008 **Lorenz Goenner**
Ungleichungen für charakteristische Funktionen
- 1/2008 **Dietmar Ferger, Michael Scholz**
Limit Distributions of V- and U-Statistics in Terms of Multiple Stochastic Wiener-Type Integrals
- 2/2007 **Robert Knobloch, Lothar Partzsch**
Uniform Conditional Ergodicity and Intrinsic Ultracontractivity
- 1/2007 **Dietmar Ferger, Daniel Vogel**
Weak Convergence of the empirical process and the rescaled empirical distribution function in the Skorokhod product space
- 1/2006 **Jens Klotsche, Lars Pieper, Dietmar Ferger**
Estimation of an optimal cut-off for a continuous risk factor for estimating clinically valid thresholds
- 3/2005 **Mark Freidlin, Matthias Weber**
Small Diffusion Asymptotics for Exit Problems on Graphs
- 2/2005 **Jürgen Franz**
On Posterior Distribution and Loss Functions for Parameter Estimation in Weibull Processes
- 1/2005 **Dietmar Hudak, Matthias Weber**
Generators of Diffusion Processes on Trees, their Resolvent and their Extension to Self-Adjoint Operators
- 4/2004 **Mark Freidlin, Matthias Weber**
On stochasticity of solutions of differential equations with a small delay
- 3/2004 **Dietmar Ferger, Wiltrud Kuhlisch**
A test for a jump point in the Weibull distribution with application to a problem in the materials research