

Change-Point-Tests für die Innovationenverteilung in nichtparametrischen Autoregressionsmodellen auf Basis sequentieller empirischer Prozesse

Dissertation

zur Erlangung des Doktorgrades
der Fakultät für Mathematik, Informatik und Naturwissenschaften
der Universität Hamburg

vorgelegt im Fachbereich Mathematik von

LEONIE SELK

aus Hamburg

Hamburg 2011

Als Dissertation angenommen vom Fachbereich Mathematik der Universität Hamburg

Auf Grund der Gutachten von
und

Prof. Dr. Natalie Neumeyer (Universität Hamburg)
Prof. Dr. Jens-Peter Kreiß (Technische Universität Braunschweig)

Hamburg, 19. September 2011

Prof. Dr. Ingenuin Gasser
Leiter des Fachbereichs Mathematik

Inhaltsverzeichnis

1	Einleitung	5
1.1	Change-Points in Zeitreihen	7
1.1.1	Anwendungsbeispiele	8
2	Grundlagen	10
2.1	Das Modell	10
2.1.1	Stationarität	10
2.2	Mischungsbedingung	11
2.3	Konvergenzbegriffe	14
2.3.1	Landau-Symbole	16
3	Die Kernschätzer	18
3.1	Definition der Kernschätzer	18
3.2	Konsistenz und Konvergenzraten der Kernschätzer	19
3.3	Verwandte Ergebnisse	43
4	Der sequentielle empirische Prozess der Residuen	45
4.1	Definition des sequentiellen empirischen Prozesses der Residuen	45
4.2	Stochastische Entwicklung für den sequentiellen empirischen Prozess der Residuen	46
4.3	Schwache Konvergenz des sequentiellen empirischen Prozesses der Residuen	60
5	Test auf einen Change-Point in der Innovationen-Verteilung	70
5.1	Definition der Teststatistik	70
5.2	Verteilung der Teststatistik unter H_0	71
5.3	Konsistenz des Tests	75
5.4	Teststatistik im homoskedastischen Fall	81
5.5	Verwandte Ergebnisse	81
6	Beispiele und Simulationen für den Change-Point-Test	83
6.1	AR(1)-Modelle	83
6.1.1	Normalverteilte Innovationen	83
6.1.2	Skew-Normal verteilte Innovationen	87
6.2	ARCH(1)-Modelle	90
6.2.1	Normalverteilte Innovationen	90
6.2.2	Skew-Normal verteilte Innovationen	92
6.2.3	Student-t-verteilte Innovationen	93
6.3	Homoskedastische Modelle	98
6.4	Simulationstechnik	98
6.5	Fazit der Simulationen	100
6.6	Wahl der Bandbreite	100
6.7	Anwendung auf einen echten Datensatz	102
7	Weitere Tests: Goodness-of-fit-Tests	104
7.1	Test auf Normalverteilung der Innovationen	104

7.2	Tests auf andere Innovationenverteilungen	107
7.3	Tests auf bestimmte Formen der Regressions- oder Skalierungsfunktion	108
7.4	Bootstrap-Verfahren	109
8	Literaturüberblick	114
9	Lemmata	117
10	Technische Lemmata	142
11	Ausblick	172

1 Einleitung

Die vorliegende Arbeit beschäftigt sich mit der Analyse von Zeitreihen. Eine Zeitreihe ist eine Folge von reellwertigen Zufallsvariablen $(X_j)_{j \in \mathbb{Z}}$.

Zeitreihen treten in vielen verschiedenen Bereichen, wie z.B. der (Finanz-) Wirtschaft, der Medizin oder den Naturwissenschaften auf. Hier sollen einige Beispiele für Zeitreihen genannt werden:

- Die täglichen Kurse einer bestimmten Aktie oder eines bestimmten Aktienindex.
- Die monatlichen Arbeitslosenzahlen in Deutschland.
- Die jährlichen minimalen Wasserstände des Nils.

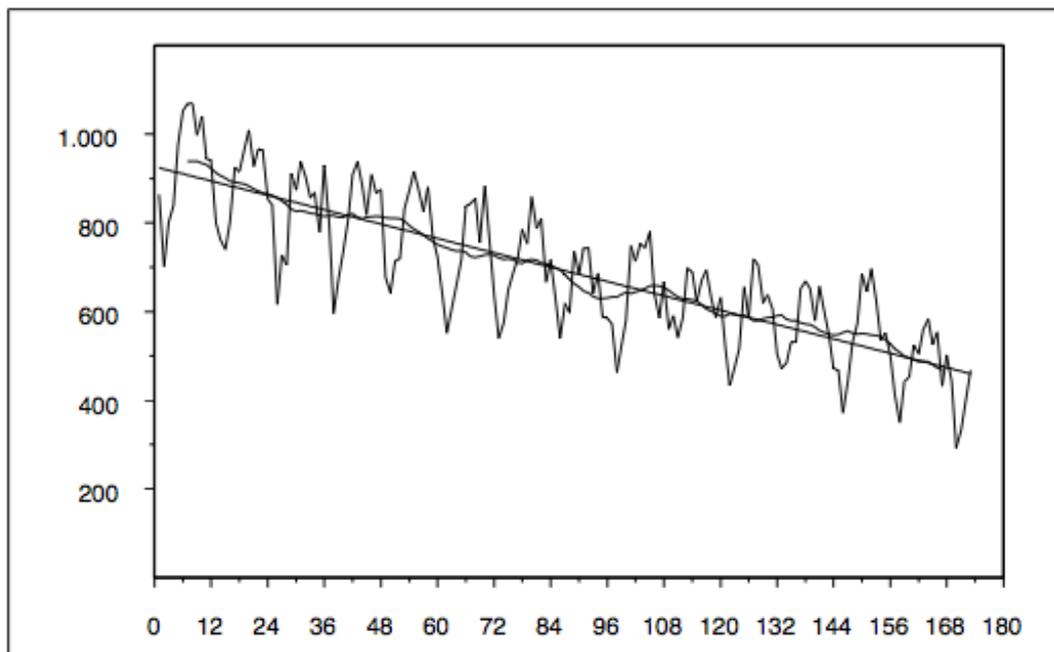


Abbildung 1.1: Monatliche Anzahl der durch Verkehrsunfälle in Deutschland getöteten Personen von Januar 1991 bis Mai 2005 zusammen mit einem geschätzten linearen Trend und einem gemäß des gleitenden Mittels über 12 Monate geschätzten Trend
Quelle: Kreiß and Neuhaus [2006]

In Abbildung 1.1 ist als Beispiel einer Zeitreihe die monatliche Anzahl der durch Verkehrsunfälle in Deutschland getöteten Personen von Januar 1991 bis Mai 2005 dargestellt (Quelle: Kreiß and Neuhaus [2006]). Es ist zu sehen, dass die Zahlen im Laufe der Zeit absinken, die Zeitreihe unterliegt einem sogenannten Trend. Außerdem ist zu sehen, dass die Zahlen von der Saison abhängig sind,

im Sommer treten offensichtlich mehr tödliche Unfälle auf als im Winter.

In der Zeitreihenanalyse, wie sie in dieser Arbeit betrieben wird, wird davon ausgegangen, dass die Zeitreihen keinen solchen Trend- oder Saisoneinflüssen unterliegen. Durch Schätzung des Trends und des saisonalen Anteils ist es (allerdings nur unvollständig und höchstens approximativ) möglich, eine Zeitreihe, wie die in Abbildung 1.1 dargestellte, von Trend- und Saisoneinflüssen zu bereinigen. Für solche Schätzungen gibt es verschiedene Methoden, die hier aber nicht näher ausgeführt werden sollen.

Abbildung 1.2 zeigt die trend- und saisonbereinigte Zeitreihe der tödlichen Unfallopfer in Deutschland von 1991 bis 2005 (Quelle: Kreiß and Neuhaus [2006]).

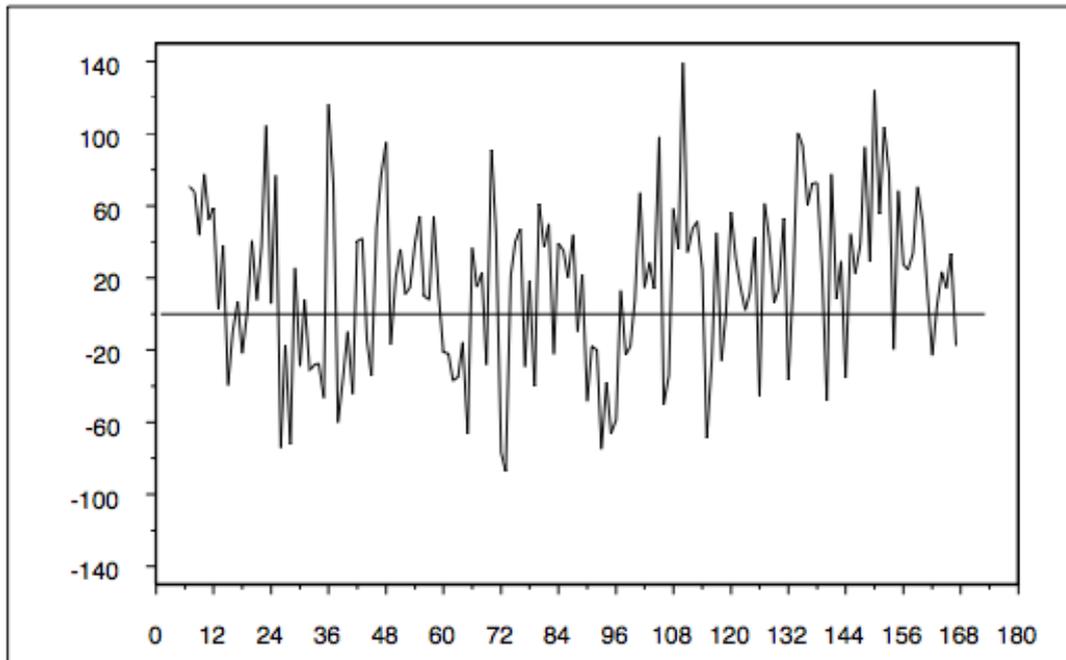


Abbildung 1.2: Trend- und saisonbereinigte Anzahl der tödlich verletzten Unfallopfer in Deutschland von 1991 bis 2005 (vgl. mit Abbildung 1.1)
Quelle: Kreiß and Neuhaus [2006]

Für Zeitreihen, die keinen Trend- und Saisoneinflüssen unterliegen wird in dieser Arbeit mit Methoden der nichtparametrischen Statistik untersucht, ob ein sogenannter Change-Point (Strukturbruch) vorliegt. Im nächsten Abschnitt wird näher beschrieben, was ein Change-Point in einer Zeitreihe ist, und in welchen Situationen Tests darauf, ob ein Change-Point vorliegt, von Interesse sind.

Der Rest der Arbeit ist wie folgt aufgebaut. In Kapitel 2 wird das mathematische Modell vorgestellt und einige Grundlagen zu den Begriffen Mischungsbedingung und Konvergenz von Zufallselementen eingeführt. In Kapitel 3 werden die nichtparametrischen Schätzer definiert und Konsistenz dieser Schätzer bewiesen. In Kapitel 4 wird der sequentielle empirische Prozess der nichtparametrisch geschätzten Residuen definiert, eine stochastische Entwicklung für diesen hergeleitet und schwache Konvergenz des Prozesses bewiesen. In Kapitel 5 wird auf Basis der Ergebnisse aus Kapitel 4 ein Test auf einen Change-Point in der Innovationenverteilung vorgestellt und die Verteilung der Teststatistik sowie Konsistenz des Tests hergeleitet. In Kapitel 6 wird der Test für verschiedene Modelle auf simulierte Daten angewendet und ein echter Datensatz zu einem Aktienindex untersucht. In Kapitel 7 werden weitere mögliche Hypothesentests vorgestellt, für die die Ergebnisse aus Kapitel

4 verwendet werden können. In Kapitel 8 werden zwei Artikel vorgestellt, die sich mit ähnlichen Modellen wie das dieser Arbeit beschäftigen, und es wird herausgearbeitet, warum die Ergebnisse dieser Artikel nicht verwendet werden können. Kapitel 9 und Kapitel 10 beinhalten die für die Beweise in Kapitel 4 und Kapitel 5 notwendigen Lemmata. Zum Schluss gibt Kapitel 11 einen Ausblick über mögliche weiterführende Forschung auf dem Gebiet dieser Arbeit.

1.1 Change-Points in Zeitreihen

Ein sogenannter Change-Point (Strukturbruch) liegt in einer Zeitreihe vor, wenn sich die Verteilung der Zufallsvariablen, aus denen die Zeitreihe besteht, an einem bestimmten Zeitpunkt, dem Change-Point, ändert.

Teilweise ist dies schon optisch in der Graphik einer Realisation der Zeitreihe zu erkennen, wie z.B. in Abbildung 1.3, in der eine Realisation einer Zeitreihe mit einem Change-Point zum Zeitpunkt 100 dargestellt ist.

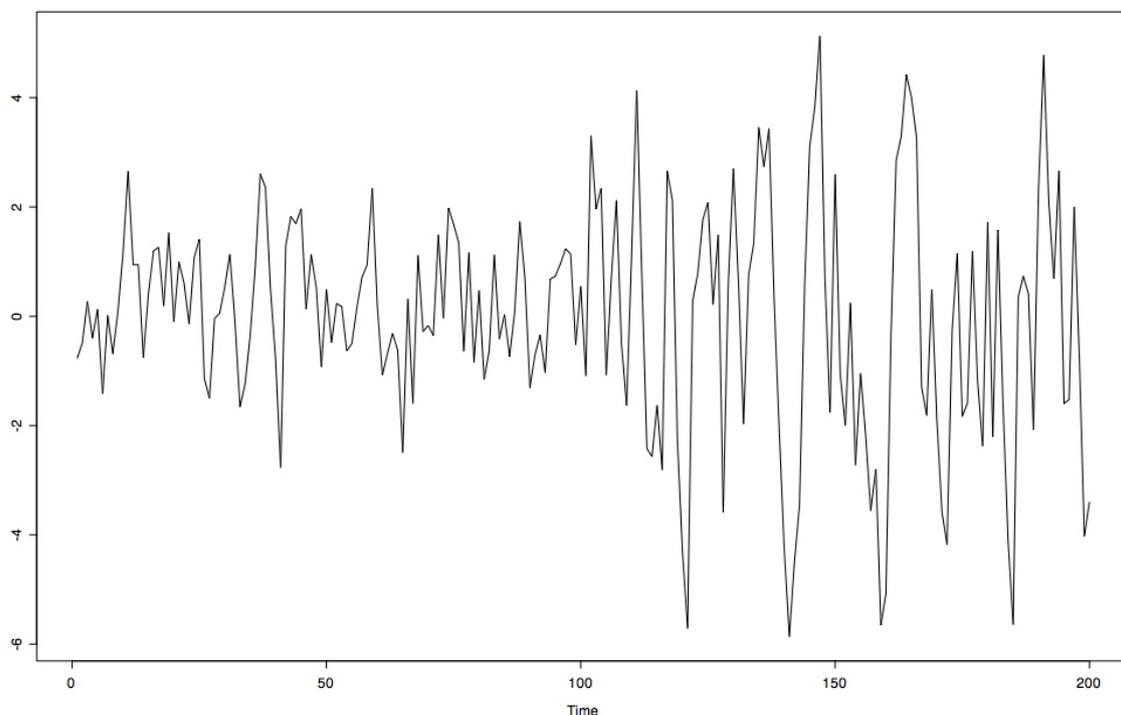


Abbildung 1.3: Realisation einer Zeitreihe mit Change-Point zum Zeitpunkt 100
(Modell (6.0.1) (S. 83) mit $n = 200$, $\zeta = 1$)

Zum Vergleich ist in Abbildung 1.4 eine Realisation einer Zeitreihe ohne Change-Point dargestellt.

In einigen Fällen ist es allerdings nicht eindeutig in der Graphik einer Realisation zu erkennen, ob ein Change-Point vorliegt, oder nicht. Ein Beispiel dafür ist Abbildung 1.5, in der eine Realisation einer Zeitreihe mit einem Change-Point zum Zeitpunkt 100 dargestellt ist, der Change-Point allerdings

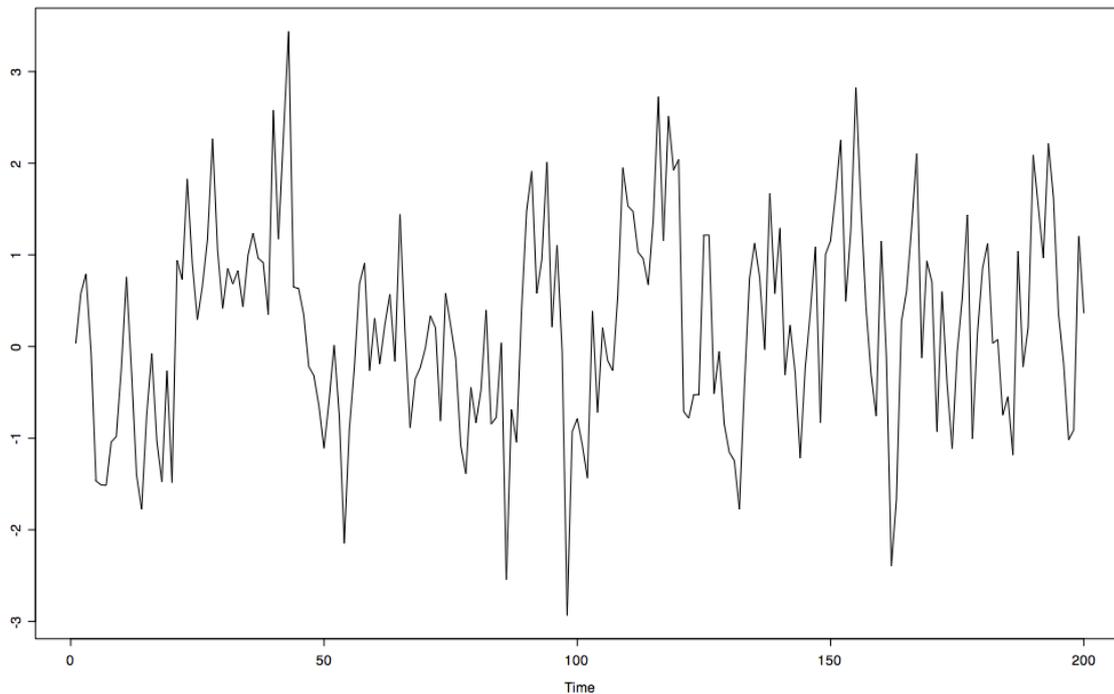


Abbildung 1.4: Realisation einer Zeitreihe ohne Change-Point
(Modell (6.0.1) (S. 83) mit $n = 200$, $\zeta = 0$)

nicht so deutlich zu erkennen ist, wie in Abbildung 1.3.

Für solche Fälle sind Tests auf einen Change-Point, wie der in Kapitel 5 vorgestellte, von Bedeutung.

1.1.1 Anwendungsbeispiele

Die Frage, ob ein Change-Point in einer Zeitreihe vorliegt, ist in vielen Anwendungsbereichen interessant, z.B. in der (Finanz-)Wirtschaft, in der Klimaforschung oder in der Medizin, z.B. bei der Überwachung von Intensiv-Stationen-Patienten. Exemplarisch sollen hier einige konkrete Beispiele vorgestellt werden:

- **Qualitätskontrolle:** Von einer Maschine produzierte Teile sind meist (in einem geringen Maße) zufälligen Abweichungen vom Soll-Wert unterworfen. Eine bestimmte Ausprägung nacheinander produzierter Teile kann also als stochastische Zeitreihe angesehen werden. Ein Change-Point-Test könnte hier überprüfen, ob sich die Abweichung vom Soll-Wert nach einer Zeit statistisch signifikant verändert hat. (Vergleich Page [1954])
- **Finanzzeitreihen:** Mit einem Change-Point-Test kann z.B. getestet werden, ob sich die Volatilität oder die Rendite (z.B. absolute oder quadratische tägliche Rendite) statistisch signifikant verändert hat. (siehe Andreou and Ghysels [2009], Abschnitt 4)
- **Wasserstände:** In der Klimaforschung werden die Wasserstände verschiedener Flüsse aufgezeichnet. Diese Daten können auf einen Change-Point getestet werden. Ein bekanntes Beispiel

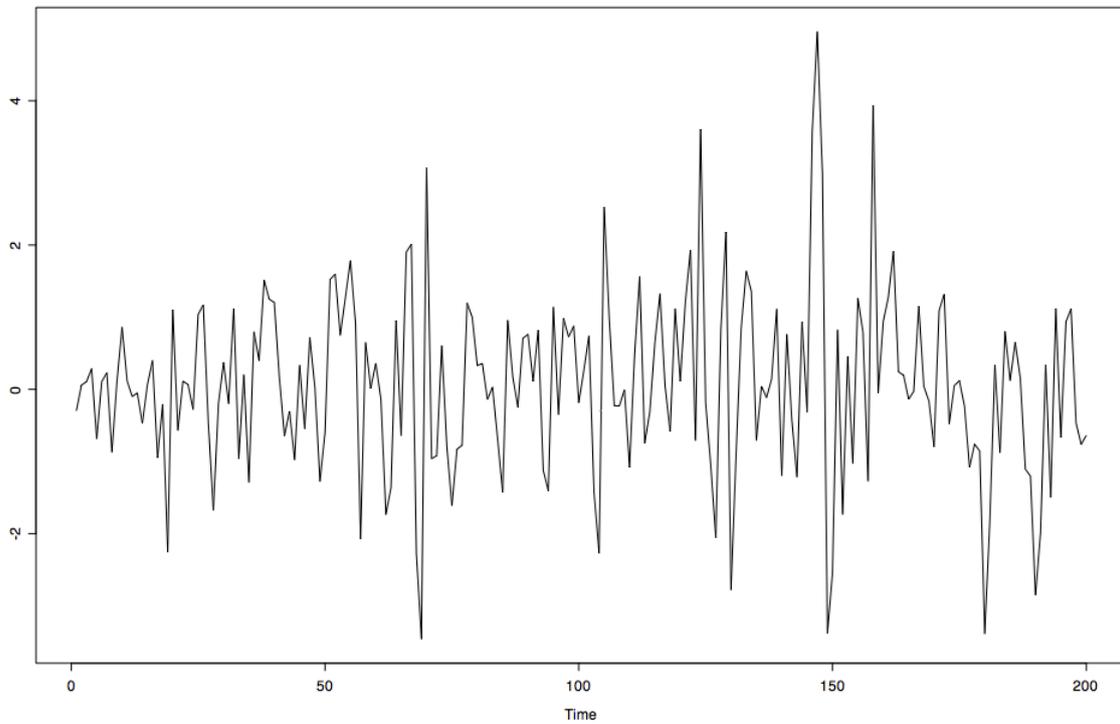


Abbildung 1.5: Realisation einer Zeitreihe mit Change-Point zum Zeitpunkt 100
(Modell (6.0.4) (S. 90) mit $n = 200$, $\zeta = 0.5$)

sind die Wasserstände des Nils, die sich 1899 (Bau des ersten Assuan-Staudamms) statistisch signifikant veränderten. Ein weiteres Beispiel für einen Change-Point-Test auf Wasserstandsdaten wird in Hušková and Antoch [2003] gegeben.

2 Grundlagen

In diesem Kapitel wird zunächst das Modell, das in dieser Arbeit betrachtet wird, definiert (Abschnitt 2.1). Anschließend wird eine Übersicht zu den Grundlagen der Begriffe Mischungsbedingung (Abschnitt 2.2) und Konvergenz von Zufallselementen (Abschnitt 2.3) gegeben.

2.1 Das Modell

Das Modell, das in dieser Arbeit betrachtet wird, ist ein reellwertiger stochastischer Prozess $(X_j)_{j \in \mathbb{Z}}$, der durch die folgende Gleichung bestimmt ist:

$$X_j = m(X_{j-1}) + \sigma(X_{j-1})\epsilon_j, \quad (2.0.1)$$

wobei m und σ geeignet glatte Funktionen sind (siehe Voraussetzung (M), Kapitel 3) und (ϵ_j) unabhängig mit $E[\epsilon_j] = 0$ und $E[\epsilon_j^2] = 1 \forall j$ sind und ϵ_j von allen X_k mit $k \leq j - 1$ stochastisch unabhängig ist. Ein solches Modell wird als heteroskedastisches autoregressives Modell bezeichnet. Die Zufallsvariablen (ϵ_j) heißen Innovationen.

Es wird davon ausgegangen, dass X_0, \dots, X_n eines solchen Prozesses beobachtet wurden. Die Funktionen m und σ sind unbekannt, ebenso die Verteilung der Innovationen (ϵ_j) . Sie werden auf Basis der Beobachtungen geschätzt (siehe Kapitel 3 und Kapitel 4).

Die Funktionen m und σ bezeichnen den bedingten Erwartungswert und die bedingte Streuung der Beobachtungen:

Der Term $E[X_j | X_{j-1} = x]$, $x \in \mathbb{R}$ ist unabhängig von $j = 1, \dots, n$ und wird Regressionsfunktion genannt. Er wird mit $m(x)$ bezeichnet, also

$$\begin{aligned} m(x) &:= E[X_1 | X_0 = x], & x \in \mathbb{R} \\ &= E[X_2 | X_1 = x] = \dots = E[X_n | X_{n-1} = x]. \end{aligned}$$

Ebenso ist $\text{Var}(X_j | X_{j-1} = x)$, $x \in \mathbb{R}$ unabhängig von $j = 1, \dots, n$, und

$$\sigma(x) = \sqrt{\sigma^2(x)} := \sqrt{\text{Var}(X_1 | X_0 = x)}, \quad x \in \mathbb{R}$$

wird Skalierungsfunktion genannt.

2.1.1 Stationarität

Häufig wird in Modellen wie 2.0.1 angenommen, dass die Zeitreihe $(X_j)_{j \in \mathbb{Z}}$ strikt stationär ist, also dass die Verteilung von $(X_{j_1}, \dots, X_{j_k})$ gleich der Verteilung von $(X_{j_1+h}, \dots, X_{j_k+h})$ ist, für alle $h \in \mathbb{Z}$, alle $k \in \mathbb{N}$ und alle $j_1, \dots, j_k \in \mathbb{Z}$.

Hierfür muss insbesondere gelten, dass alle ϵ_j dieselbe Verteilung haben (durch die Unabhängigkeit ist dann die Zeitreihe $(\epsilon_j)_{j \in \mathbb{Z}}$ strikt stationär).

In Kapitel 5 werden Modelle mit einem sogenannte Change-Point in der Verteilung der Innovationen (ϵ_j) untersucht, also Modelle, bei denen sich die Verteilung der ϵ_j zu einem bestimmten Zeitpunkt ändert. Die mit solchen Innovationen entstehenden Zeitreihen $(X_j)_{j \in \mathbb{Z}}$ können nicht strikt stationär sein.

In den Lemmata, die für den Beweis der Konsistenz des Change-Point-Tests notwendig sind (Lemmata 3.3 und 3.4, Lemmata in Kapitel 9), wird daher auf die Annahme strikter Stationarität und auch auf die schwächere Annahme, dass alle X_j dieselbe Verteilung haben, verzichtet.

Für die stochastische Entwicklung des sequentiellen empirischen Prozesses der Residuen in Kapitel 4, aus der die Verteilung des Change-Point-Tests unter der Nullhypothese keines Change-Points folgt, sind die Annahmen, dass alle X_j dieselbe Verteilung haben und dass alle ϵ_j dieselbe Verteilung haben, allerdings notwendig.

Bemerkung 2.1 Aus den Annahmen, dass alle X_j dieselbe Verteilung haben und dass alle ϵ_j dieselbe Verteilung haben, folgt strikte Stationarität des Prozesses $(X_j)_{j \in \mathbb{Z}}$.

Beweis Sei $k \in \mathbb{N}$, $j_1, \dots, j_k \in \mathbb{Z}$ und o.E. $j_1 < j_2 < \dots < j_k$. Die gemeinsame Dichte $f_{X_{j_1}, \dots, X_{j_k}}$ von $(X_{j_1}, \dots, X_{j_k})$ lässt sich darstellen als

$$\begin{aligned} & f_{X_{j_1}, \dots, X_{j_k}}(x_1, \dots, x_k) \\ &= f_{X_{j_1}}(x_1) \cdot f_{X_{j_2}|X_{j_1}}(x_2|x_1) \cdot \dots \cdot f_{X_{j_k}|X_{j_{k-1}}}(x_k|x_{k-1}), \end{aligned}$$

da die durch 2.0.1 definierte Zeitreihe markovsch ist. Weiter gilt für alle $i = 1, \dots, k$

$$\begin{aligned} & f_{X_{j_i}|X_{j_{i-1}}}(x_i|x_{i-1}) \\ &= \begin{cases} \frac{f_{\epsilon_{j_i}}\left(\frac{x_i - m(x_{i-1})}{\sigma(x_{i-1})}\right)}{\sigma(x_{i-1})} & \text{falls } j_{i-1} = j_i - 1 \\ \int \frac{f_{\epsilon_{j_{i-1}+1}}\left(\frac{y_1 - m(x_{i-1})}{\sigma(x_{i-1})}\right)}{\sigma(x_{i-1})} \frac{f_{\epsilon_{j_{i-1}+2}}\left(\frac{y_2 - m(y_1)}{\sigma(y_1)}\right)}{\sigma(y_1)} \dots \frac{f_{\epsilon_{j_i}}\left(\frac{x_i - m(y_{j_i - j_{i-1} - 1})}{\sigma(y_{j_i - j_{i-1} - 1})}\right)}{\sigma(y_{j_i - j_{i-1} - 1})} d(y_1, \dots, y_{j_i - j_{i-1} - 1}) & \text{sonst.} \end{cases} \end{aligned}$$

Damit ist die Verteilung von $(X_{j_1}, \dots, X_{j_k})$ gleich der Verteilung von $(X_{j_1+h}, \dots, X_{j_k+h})$ für alle $h \in \mathbb{Z}$, wenn alle X_j dieselbe Verteilung haben und alle ϵ_j dieselbe Verteilung haben. □

2.2 Mischungsbedingung

Eine Voraussetzung in vielen in dieser Arbeit bewiesenen Sätzen ist, dass die Beobachtungen $(X_j)_{j \in \mathbb{Z}}$ α -mischend sind und die Mischungskoeffizienten bestimmte Konvergenzraten erfüllen.

Definition 2.2 Ein stochastischer Prozess $(X_j)_{j \in \mathbb{Z}}$ ist α -mischend mit Mischungskoeffizienten $\alpha(n)$, wenn

$$\alpha(n) := \sup_{k \in \mathbb{Z}} \sup_{A \in \sigma(X_j: j \leq k), B \in \sigma(X_j: j \geq k+n)} |P(A \cap B) - P(A) \cdot P(B)| \xrightarrow{n \rightarrow \infty} 0, \quad (2.2.1)$$

wobei $\sigma(\cdot)$ hier die von \cdot erzeugte σ -Algebra bezeichnet.
(Siehe z.B. Rio [1995].)

Ist der Prozess $(X_j)_{j \in \mathbb{Z}}$ stationär, so kann die Bedingung (2.2.1) zu

$$\alpha(n) = \sup_{A \in \sigma(X_j: j \leq 0), B \in \sigma(X_j: j \geq n)} |P(A \cap B) - P(A) \cdot P(B)| \xrightarrow{n \rightarrow \infty} 0$$

vereinfacht werden (siehe Fan and Yao [2005] (Abschnitt 2.6.1)), und im Fall eines stationären Markov-Prozesses sogar zu

$$\alpha(n) = \sup_{A \in \sigma(X_0), B \in \sigma(X_n)} |P(A \cap B) - P(A) \cdot P(B)| \xrightarrow{n \rightarrow \infty} 0.$$

(Siehe Fan and Yao [2005] 2.6.1(iv).)

Der hier betrachtete stochastische Prozess

$$X_j = m(X_{j-1}) + \sigma(X_{j-1})\epsilon_j$$

ist markovsch. Allerdings wird er nicht für alle Theoreme als stationär angenommen (siehe Abschnitt 2.1.1).

Die stärkste Forderung an die Konvergenzrate der Mischungskoeffizienten in dieser Arbeit, ist

$$\alpha(n) = O(\bar{\alpha}^{-n}), \text{ für ein } 1 < \bar{\alpha} < \infty,$$

also dass die Mischungskoeffizienten exponentiell-schnell fallen.

Diese Forderung ist erfüllt, wenn die Folge von Zufallsvariablen geometrisch ergodisch ist, wie in Lemma 2.4 gezeigt wird.

Definition 2.3 Ein Markov-Prozess $(X_j)_{j \in \mathbb{Z}}$ ist geometrisch ergodisch, wenn es ein eindeutiges invariantes Maß π gibt, so dass

$$\left\| P^{(n)}(\cdot | x) - \pi \right\|_{Var} \leq G(x) \bar{\beta}^{-n}, \text{ für ein } 1 < \bar{\beta} < \infty$$

für eine Funktion $G > 0$ mit $\int G(x) \pi(dx) < \infty$, wobei $P^{(n)}(\cdot | x)$ die n -Schritt Übergangswahrscheinlichkeit mit Ausgangspunkt x bezeichne und $\|\cdot\|_{Var}$ die Totalvariation.
(Siehe Bhattacharya and Lee [1995] oder Doukhan [1994] (S.89)).

Lemma 2.4 Ein stationärer geometrisch ergodischer Prozess $(X_j)_{j \in \mathbb{Z}}$ ist α -mischend mit exponentiell-schnell fallenden Mischungskoeffizienten.

Das betrachtete Modell

$$X_j = m(X_{j-1}) + \sigma(X_{j-1}) \cdot \epsilon_j, \quad (\epsilon_j) \text{ i.i.d.}$$

ist nach Doukhan [1994] (Abschnitt 2.4.2.3, Proposition 6) geometrisch ergodisch (und damit α -mischend mit exponentiell-schnell fallenden Mischungskoeffizienten), wenn die Verteilung der Innovationen (ϵ_j) absolut stetig bezüglich des Lebesgue-Maßes ist und außerdem die Bedingung

$$\exists \tau \geq 1 : E[|\epsilon_j|^\tau] < \infty, \quad \limsup_{|x| \rightarrow \infty} \frac{|m(x)| + |\sigma(x)| (E[|\epsilon_j|^\tau])^{\frac{1}{\tau}}}{|x|} < 1 \quad (2.4.1)$$

erfüllt ist.

Ähnliche Bedingungen werden auch in Bhattacharya and Lee [1995] angegeben.

Beweis von Lemma 2.4 Ein stationärer geometrisch ergodischer Prozess $(X_j)_{j \in \mathbb{Z}}$ ist nach Fan and Yao [2005] (Abschnitt 2.6.1 Punkt (vi)) oder Doukhan [1994] (S.89) β -mischend mit exponentiell-schnell fallenden Mischungskoeffizienten, d.h.

$$\beta(n) := \sup_{k \in \mathbb{Z}} E \left[\sup_{B \in \sigma(X_j : j \geq k+n)} |P(B) - P(B|\sigma(X_j : j \leq k))| \right] = O(\bar{\beta}^{-n}), \quad \text{für ein } 1 < \bar{\beta} < \infty$$

(Definition siehe z.B. Fan and Yao [2005] (Abschnitt 2.6.1)).

Es gilt

$$\begin{aligned} & \sup_{k \in \mathbb{Z}} E \left[\sup_{B \in \sigma(X_j : j \geq k+n)} |P(B) - P(B|\sigma(X_j : j \leq k))| \right] \\ & \geq \sup_{k \in \mathbb{Z}} \sup_{A \in \sigma(X_j : j \leq k)} E \left[\sup_{B \in \sigma(X_j : j \geq k+n)} |P(B) - P(B|\sigma(X_j : j \leq k))| \cdot 1_A \right], \\ & \text{da } \sup_{B \in \sigma(X_j : j \geq k+n)} |P(B) - P(B|\sigma(X_j : j \leq k))| \geq 0 \\ & \geq \sup_{k \in \mathbb{Z}} \sup_{A \in \sigma(X_j : j \leq k)} \left(\sup_{B \in \sigma(X_j : j \geq k+n)} E[|P(B) - P(B|\sigma(X_j : j \leq k))| \cdot 1_A] \right) \\ & \geq \sup_{k \in \mathbb{Z}} \sup_{A \in \sigma(X_j : j \leq k), B \in \sigma(X_j : j \geq k+n)} |E[(P(B) - P(B|\sigma(X_j : j \leq k))) \cdot 1_A]| \\ & = \sup_{k \in \mathbb{Z}} \sup_{A \in \sigma(X_j : j \leq k), B \in \sigma(X_j : j \geq k+n)} |P(B) \cdot E[1_A] - E[P(B|\sigma(X_j : j \leq k)) \cdot 1_A]| \\ & = \sup_{k \in \mathbb{Z}} \sup_{A \in \sigma(X_j : j \leq k), B \in \sigma(X_j : j \geq k+n)} |P(B) \cdot P(A) - E[E[1_B|\sigma(X_j : j \leq k)] \cdot 1_A]| \\ & = \sup_{k \in \mathbb{Z}} \sup_{A \in \sigma(X_j : j \leq k), B \in \sigma(X_j : j \geq k+n)} |P(B) \cdot P(A) - E[1_B \cdot 1_A]|, \\ & \text{nach Definition der bedingten Erwartung} \\ & = \sup_{k \in \mathbb{Z}} \sup_{A \in \sigma(X_j : j \leq k), B \in \sigma(X_j : j \geq k+n)} |P(B) \cdot P(A) - P(B \cap A)|. \end{aligned}$$

Damit ist $\alpha(n) \leq \beta(n)$, also ist der stochastische Prozess $(X_j)_{j \in \mathbb{Z}}$ α -mischend mit exponentiell-schnell fallenden Mischungskoeffizienten. □

2.3 Konvergenzbegriffe

In dieser Arbeit treten häufig Konvergenzen von Zufallselementen auf. Daher wird hier eine kurze Übersicht über verschiedene Konvergenzarten von Zufallselementen gegeben.

Die folgenden Definitionen und Bemerkungen, sowie ausführlichere Erläuterungen können z.B. in Billingsley [1968], in van der Vaart [2000] oder in van der Vaart and Wellner [1996] gefunden werden.

Die in dieser Arbeit betrachteten Zufallselemente sind zum Einen reellwertige Zufallsvariablen und zum Anderen stochastische Prozesse.

Zunächst werden Konvergenzbegriffe für reellwertige Zufallsvariablen betrachtet, also für Funktionen X auf einem Wahrscheinlichkeitsraum (Ω, \mathcal{A}, P) , mit Werten in \mathbb{R} , $X : \Omega \rightarrow \mathbb{R}$, die $(\mathcal{A}, \mathbb{B})$ -messbar sind, wobei \mathbb{B} die Borel- σ -Algebra auf \mathbb{R} bezeichnet.

Definition 2.5 Sei $(X_n)_{n \in \mathbb{N}}$ eine Folge von reellwertigen Zufallsvariablen und X eine reellwertige Zufallsvariable.

- $(X_n)_{n \in \mathbb{N}}$ konvergiert fast sicher gegen X , wenn

$$P\left(\lim_{n \rightarrow \infty} X_n = X\right) = 1.$$

Notation: $X_n \xrightarrow[n \rightarrow \infty]{f.s.} X$

- $(X_n)_{n \in \mathbb{N}}$ konvergiert in Wahrscheinlichkeit gegen X , wenn

$$\lim_{n \rightarrow \infty} P(|X_n - X| > \eta) = 0, \quad \forall \eta > 0.$$

Notation: $X_n \xrightarrow[n \rightarrow \infty]{P} X$

- $(X_n)_{n \in \mathbb{N}}$ konvergiert in Verteilung gegen X , wenn

$$\lim_{n \rightarrow \infty} P(X_n \leq x) = P(X \leq x), \quad \text{für alle Stetigkeitspunkte } x \text{ von } P(X \leq x).$$

Notation: $X_n \xrightarrow[n \rightarrow \infty]{\mathcal{D}} X$

- $(X_n)_{n \in \mathbb{N}}$ konvergiert schwach gegen X , wenn

$$\lim_{n \rightarrow \infty} E[H(X_n)] = E[H(X)], \quad \text{für alle stetigen und beschränkten Funktionen } H : \mathbb{R} \rightarrow \mathbb{R}.$$

Notation: $X_n \Rightarrow X$

Bemerkung 2.6 Im hier betrachteten Fall der reellwertigen Zufallsvariablen ist schwache Konvergenz äquivalent zu Konvergenz in Verteilung.

Im Folgenden werden Konvergenzbegriffe für stochastische Prozesse betrachtet, also für Familien $(X_t)_{t \in T}$ von reellwertigen Zufallsvariablen, die auf einem gemeinsamen Wahrscheinlichkeitsraum (Ω, \mathcal{A}, P) definiert sind (T : Indexmenge).

Auch stochastische Prozesse können als Abbildungen aufgefasst werden. Sei hierfür zu einer beliebigen Indexmenge T mit $l^\infty(T)$ der Raum der beschränkten Abbildungen $T \rightarrow \mathbb{R}$ bezeichnet. Dann kann $(X_t)_{t \in T}$ als Abbildung

$$\begin{aligned} (X_t)_{t \in T} : \Omega &\rightarrow l^\infty(T) \\ \omega &\mapsto X.(\omega) : T \rightarrow \mathbb{R} \end{aligned}$$

betrachtet werden, falls jeder Pfad von $(X_t)_{t \in T}$ ($t \mapsto X_t(\omega)$) beschränkt ist.

Mit der Supremumsnorm wird $l^\infty(T)$ zum metrischen Raum. Dieser Raum kann zusätzlich mit der σ -Algebra der Borelmengen ausgestattet werden und man erhält ein ähnliches Konstrukt wie im Fall von reellwertigen Zufallsvariablen.

Die Schwierigkeit die auftritt ist allerdings, dass die hier betrachteten Folgen von stochastischen Prozessen (empirische Prozesse) im Allgemeinen nicht messbar sind. Daher wird in zu Definition 2.5 analogen Definitionen für stochastische Prozesse auf Messbarkeit der Folgenglieder verzichtet und nur Messbarkeit des Grenzwerts gefordert.

Dafür müssen zunächst analoge Begriffe zu Erwartungswert und Wahrscheinlichkeitsmaß für nicht messbare Abbildungen definiert werden.

Definition 2.7 (Ω, \mathcal{A}, P) sei ein Wahrscheinlichkeitsraum, $X : \Omega \rightarrow \bar{\mathbb{R}}$ ($= \mathbb{R} \cup \{-\infty, \infty\}$) eine Abbildung. Dann heißt

$$E^*[X] := \inf \{E[Y] : Y \geq X, Y : \Omega \rightarrow \bar{\mathbb{R}} \text{ } (\mathcal{A}, \bar{\mathbb{B}})\text{-messbar, } E[Y] \text{ existiert}\}$$

äußerer Erwartungswert von X .

Für eine beliebige Teilmenge $B \subset \Omega$ ist

$$P^*(B) = \inf \{P(A) : B \subset A, A \in \mathcal{A}\}$$

die äußere Wahrscheinlichkeit von B .

Mit Hilfe dieser Definition können zu Definition 2.5 analoge Definitionen für nicht messbare Abbildungen angegeben werden.

Definition 2.8 S sei ein metrischer Raum mit Metrik d und Borel- σ -Algebra \mathcal{S} , (Ω, \mathcal{A}, P) sei ein Wahrscheinlichkeitsraum und $X_n : \Omega \rightarrow S$, $X : \Omega \rightarrow S$ seien Abbildungen

- $(X_n)_{n \in \mathbb{N}}$ konvergiert fast sicher gegen X , wenn es ein $\Delta_n : \Omega \rightarrow \mathbb{R}$ gibt, das $(\mathcal{A}, \mathbb{B})$ -messbar ist und das fast sicher gegen Null konvergiert, $\Delta_n \xrightarrow[n \rightarrow \infty]{f.s.} 0$, mit

$$d(X_n, X) \leq \Delta_n.$$

Notation: $X_n \xrightarrow[n \rightarrow \infty]{f.s.*} X$

- $(X_n)_{n \in \mathbb{N}}$ konvergiert in Wahrscheinlichkeit gegen X , wenn

$$\lim_{n \rightarrow \infty} P^*(d(X_n, X) > \eta) = 0, \quad \forall \eta > 0.$$

Notation: $X_n \xrightarrow[n \rightarrow \infty]{P} X$

- $(X_n)_{n \in \mathbb{N}}$ konvergiert schwach gegen X , wenn

$$\lim_{n \rightarrow \infty} E[H^*(X_n)] = E[H(X)], \quad \text{für alle stetigen und beschränkten Funktionen } H : S \rightarrow \mathbb{R}.$$

Notation: $X_n \Rightarrow X$

Hier wird vorausgesetzt, dass der Grenzwert X $(\mathcal{A}, \mathcal{S})$ -messbar ist.

2.3.1 Landau-Symbole

Mit Hilfe der Landau-Symbole o und O werden Konvergenzraten angegeben. Es gilt für reelle (deterministische) Folgen $(x_n)_{n \in \mathbb{N}}$ und $(y_n)_{n \in \mathbb{N}}$

$$x_n = o(y_n) \quad :\Leftrightarrow \quad \lim_{n \rightarrow \infty} \left| \frac{x_n}{y_n} \right| = 0$$

und

$$x_n = O(y_n) \quad :\Leftrightarrow \quad \exists \bar{M} < \infty, \bar{N} < \infty : \left| \frac{x_n}{y_n} \right| < \bar{M} \quad \forall n \geq \bar{N}.$$

Häufig wird auch die Schreibweise

$$x_n = \tilde{x}_n + o(y_n) \quad :\Leftrightarrow \quad |x_n - \tilde{x}_n| = o(y_n)$$

(analog mit O anstatt o) verwendet.

Bemerkung 2.9 Aus der Definition der Landau-Symbole lassen sich direkt die folgenden Eigenschaften ableiten

- $x_n = o(y_n) \Rightarrow x_n = O(y_n)$
- $x_n = O(y_n), \tilde{x}_n = O(\tilde{y}_n) \Rightarrow x_n + \tilde{x}_n = O(\max(y_n, \tilde{y}_n)), x_n \tilde{x}_n = O(y_n \tilde{y}_n)$
- $x_n = o(y_n), \tilde{x}_n = o(\tilde{y}_n) \Rightarrow x_n + \tilde{x}_n = o(\max(y_n, \tilde{y}_n))$
- $x_n = O(y_n), \tilde{x}_n = o(\tilde{y}_n) \Rightarrow x_n \tilde{x}_n = o(y_n \tilde{y}_n)$

In dieser Arbeit tauchen häufig Landau-Symbole für stochastische Konvergenzen, O_P und o_P , auf. Analog zur obigen Definition gilt für zufällige reelle Folgen $(X_n)_{n \in \mathbb{N}}$ und $(Y_n)_{n \in \mathbb{N}}$

$$\begin{aligned} X_n = o_P(Y_n) & \quad :\Leftrightarrow \quad \frac{X_n}{Y_n} \xrightarrow[n \rightarrow \infty]{P} 0 \\ & \quad \Leftrightarrow \quad P \left(\left| \frac{X_n}{Y_n} \right| > \eta \right) \xrightarrow[n \rightarrow \infty]{} 0 \quad \forall \eta > 0 \end{aligned}$$

und

$$X_n = O_P(Y_n) \quad :\Leftrightarrow \quad \forall \eta > 0 \quad \exists \bar{M} < \infty, \bar{N} < \infty : P\left(\left|\frac{X_n}{Y_n}\right| > \bar{M}\right) < \eta \quad \forall n \geq \bar{N}.$$

Auch für die stochastischen Landau-Symbole gelten die Implikationen in Bemerkung 2.9 (siehe van der Vaart [2000], Abschnitt 2.2)

3 Die Kernschätzer

In diesem Kapitel werden die Schätzer für die Regressionsfunktion und die Skalierungsfunktion definiert, Konsistenz dieser Schätzer bewiesen (Satz 3.2, S. 21) und die Konvergenzraten der Schätzer hergeleitet (Lemma 3.4, S. 35).

Analoge Ergebnisse für ähnliche Modelle sind in Abschnitt 3.3 (S. 43) zusammengefasst. Allerdings kann keins dieser Ergebnisse für dieses Modell verwendet werden, da hier keine Stationarität gefordert wird.

Das Kapitel beginnt mit einem Abschnitt zur Definition der Schätzer, gefolgt von einem Abschnitt zur Konsistenz und den Konvergenzraten der Schätzer, in dem zum einen in Lemma 3.4 (S. 35) die Konvergenzraten der Schätzer hergeleitet werden (basierend auf Lemma 3.3, S. 21) und zum anderen die Ergebnisse von Lemma 3.4 benutzt werden, um die Konsistenz der Schätzer zu beweisen (Satz 3.2, S. 21).

3.1 Definition der Kernschätzer

Nichtparametrische Schätzer wie z.B. Kernschätzer haben gegenüber parametrischen Schätzern den Vorteil, dass keine spezifischen Annahmen an die zu schätzenden Funktionen gestellt werden müssen. Die einzigen Annahmen, die gemacht werden müssen sind Differenzierbarkeitsannahmen und Annahmen über die Wachstumsrate der Funktionen und ihrer Ableitungen. Forderungen an die genaue Form der Funktion müssen nicht gestellt werden, im Unterschied zur parametrischen Schätzung.

Für die Schätzung wird eine sogenannte Kernfunktion K benötigt, die eine Dichte ist, für die also gilt $K(u) \geq 0$ für alle $u \in \mathbb{R}$ und $\int K(u)du = 1$. Außerdem wird eine positive Nullfolge c_n benötigt, die sogenannte Bandbreite.

Ein Kernschätzer für die Regressionsfunktion $m(x) = E[X_1|X_0 = x]$ ist der Nadaraya-Watson-Schätzer

$$\hat{m}(x) := \frac{\frac{1}{nc_n} \sum_{i=1}^n K\left(\frac{x-X_{i-1}}{c_n}\right) X_i}{\frac{1}{nc_n} \sum_{i=1}^n K\left(\frac{x-X_{i-1}}{c_n}\right)}.$$

Entsprechend ist

$$\hat{\sigma}(x) = \sqrt{\hat{\sigma}^2(x)} := \sqrt{\frac{\frac{1}{nc_n} \sum_{i=1}^n K\left(\frac{x-X_{i-1}}{c_n}\right) X_i^2}{\frac{1}{nc_n} \sum_{i=1}^n K\left(\frac{x-X_{i-1}}{c_n}\right)} - (\hat{m}(x))^2}$$

ein Schätzer für $\sigma(x) = \sqrt{\text{Var}(X_1|X_0 = x)} = \sqrt{E[X_1^2|X_0 = x] - (E[X_1|X_0 = x])^2}$, die Skalierungsfunktion.

Anmerkung:

$$\hat{\sigma} \text{ ist wohldefiniert, da } \frac{\frac{1}{nc_n} \sum_{i=1}^n K\left(\frac{x-X_{i-1}}{c_n}\right) X_i^2}{\frac{1}{nc_n} \sum_{i=1}^n K\left(\frac{x-X_{i-1}}{c_n}\right)} - (\hat{m}(x))^2 = \frac{\frac{1}{nc_n} \sum_{i=1}^n K\left(\frac{x-X_{i-1}}{c_n}\right) (X_i - \hat{m}(x))^2}{\frac{1}{nc_n} \sum_{i=1}^n K\left(\frac{x-X_{i-1}}{c_n}\right)} \geq 0$$

3.2 Konsistenz und Konvergenzraten der Kernschätzer

Kernschätzer, wie die in Abschnitt 3.1 vorgestellten $\hat{m}(x)$ und $\hat{\sigma}(x)$, liefern eher schlechte Schätzwerte für Werte x , die weit von den Beobachtungen X_0, \dots, X_n entfernt liegen. Die Beobachtungen nehmen aber auf Grund der Annahmen an m und σ (die zu langfristiger Stationarität führen, siehe Abschnitt 2.2, S. 11) nur mit sehr kleiner Wahrscheinlichkeit sehr große Werte an. Daher sind generell die Schätzwerte für große x -Werte eher schlecht, im Sinne von entfernt vom wahren Wert. Dieses Problem kann gelöst werden, indem man den Wertebereich der Schätzer auf ein kompaktes Intervall einschränkt, das für $n \rightarrow \infty$ gegen die reelle Zahlengerade konvergiert.

Eine solche Folge von Intervallen ist

$$I_n := [a_n, b_n]$$

mit $a_n \xrightarrow{n \rightarrow \infty} -\infty$ und $b_n \xrightarrow{n \rightarrow \infty} \infty$.

Es wird also nicht $\sup_{x \in \mathbb{R}} |\hat{m}(x) - m(x)|$ und $\sup_{x \in \mathbb{R}} |\hat{\sigma}(x) - \sigma(x)|$, sondern $\sup_{x \in I_n} |\hat{m}(x) - m(x)|$ und $\sup_{x \in I_n} |\hat{\sigma}(x) - \sigma(x)|$ betrachtet.

Konsistenz eines Schätzers bedeutet, dass er in Wahrscheinlichkeit gegen die zu schätzende Funktion konvergiert. Für \hat{m} und $\hat{\sigma}$ wird Konsistenz hier durch $\sup_{x \in I_n} |\hat{m}(x) - m(x)| = o_P(1)$ und $\sup_{x \in I_n} |\hat{\sigma}(x) - \sigma(x)| = o_P(1)$ definiert.

Unter den folgenden Voraussetzungen sind die Schätzer \hat{m} und $\hat{\sigma}$ konsistent.

Annahme (K) Der Kern K habe den kompakten Träger $[-C, C]$. Er sei dreimal differenzierbar und $\sup_{u \in [-C, C]} |K^{(\mu)}(u)| \leq \bar{K}$, $\mu = 0, 1, 2, 3$. Außerdem gelte $K(C) = K(-C) = K'(C) = K'(-C) = 0$ und $\int K(u) u du = 0$.

Annahme (C') Für die Bandbreite c_n gebe es ein r_c mit $|r_c| < \infty$, so dass gelte $c_n \sim n^{-\frac{1}{4}} \log(n)^{r_c}$.

Annahme (I') Für das Intervall $I_n = [a_n, b_n]$, mit $a_n \xrightarrow{n \rightarrow \infty} -\infty$ und $b_n \xrightarrow{n \rightarrow \infty} \infty$, gebe es ein $r_I < \infty$, so dass gelte $(b_n - a_n) = O(\log(n)^{r_I})$.

Annahme (X') Der Prozess $(X_i)_{i \in \mathbb{Z}}$ der Beobachtungen sei α -mischend mit Mischungskoeffizienten, $\alpha(n) = O(n^{-\beta})$ für ein $\beta > 38$ (siehe Lemma 3.3, S. 21).

Die Dichten f_{X_i} der Beobachtungen seien viermal differenzierbar und es gelte

$$\sup_{x \in \mathbb{R}} \frac{1}{n} \sum_{i=1}^n \left| f_{X_{i-1}}^{(\mu)}(x) \right| = O(1), \quad \mu = 0, 1, 2, 3, 4.$$

Außerdem gebe es ein $r_f < \infty$, so dass $\frac{1}{\inf_{x \in I_n} \frac{1}{n} \sum_{i=1}^n f_{X_{i-1}}(x)} =: q_n^f = O(\log(n)^{r_f})$.

Annahme (E') Für die $2b$ -ten Momente der X_1, \dots, X_n gelte $\frac{1}{n} \sum_{i=1}^n E[|X_i|^{2b}] = O(1)$ für ein $b > \frac{8}{3}$, $n \rightarrow \infty$.

Annahme (Z') Für alle $m_n \leq n$, mit $m_n^{-1} = o(1)$ gelte

$$\sup_{x \in J_n} \left((|m(x)| + |\sigma(x)|) \frac{1}{m_n} \max_{0 \leq S \leq n - m_n} \sum_{i=S+1}^{S+m_n} f_{X_{i-1}}(x) \right) = O(1),$$

$$\sup_{x \in J_n} \left((|m(x)| + |\sigma(x)|)^{2k} \frac{1}{m_n} \max_{0 \leq S \leq n - m_n} \sum_{i=S+1}^{S+m_n} (1 + E[\epsilon_i^4])^{k-1} f_{X_{i-1}}(x) \right) = O(1),$$

und es gebe ein $1 \leq j^* < \infty$, so dass

$$\sup_{x, x' \in J_n} \left((|m(x)| + |\sigma(x)|)^k (|m(x')| + |\sigma(x')|)^k \frac{1}{m_n^2} \max_{0 \leq S \leq n - m_n} \sum_{\substack{i, j = S+1 \\ |i-j| > j^*}}^{S+m_n} f_{X_{i-1}, X_{j-1}}(x, x') \right) = O(1) \text{ und}$$

$$\sup_{x \in J_n} \left((|m(x)| + |\sigma(x)|)^{2k} \max_{j^*+1 \leq i \leq n-j^*} \sum_{j=i-j^*}^{i+j^*} \left(1 + E[\epsilon_j^4]\right)^{k-1} f_{X_{j-1}}(x) \right) = O(1)$$

für $k = 1, 2$, $n \rightarrow \infty$ mit $J_n := [a_n - (C + c_n^{-\frac{1}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}})c_n, b_n + (C + c_n^{-\frac{1}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}})c_n]$.

Anmerkung: Nach Annahme (C') gilt $c_n \rightarrow 0$ und $c_n^{-\frac{1}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} \rightarrow 0$ für $n \rightarrow \infty$. Daher gibt es eine positive Konstante $\bar{C} < \infty$, so dass $(C + c_n^{-\frac{1}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}})c_n \leq \bar{C}$ für alle n . Damit gelten die Konvergenzraten in Annahme (Z'), wenn die entsprechenden Konvergenzraten für $\sup_{x \in [a_n - \bar{C}, b_n + \bar{C}]}$ anstatt $\sup_{x \in J_n}$ gelten.

Anmerkung: Annahme (X'), Annahme (E') und Annahme (Z') vereinfachen sich im Fall von stationären Zeitreihen, wie sie in Kapitel 4 betrachtet werden. Lemma 3.3 und Lemma 3.4 werden für den Beweis der Konsistenz des Change-Points-Tests (Abschnitt 5.3, S. 75) benötigt, daher kann für sie keine Stationarität gefordert werden.

Annahme (M) Die Regressionsfunktion m und die Skalierungsfunktion σ seien viermal differenzierbar und es gebe $r_q, r_s < \infty$ und q_n, q_n^σ mit $(q_n)^{-1} = O(1)$, $(q_n^\sigma)^{-1} = O(1)$, so dass gelte $\sup_{x \in [a_n - Cc_n, b_n + Cc_n]} |m^{(\mu)}(x)| = O(q_n)$, $\sup_{x \in [a_n - Cc_n, b_n + Cc_n]} |\sigma^{(\mu)}(x)| = O(q_n)$, $\mu = 0, 1, 2, 3, 4$ sowie $\frac{1}{\inf_{x \in I_n} |\sigma(x)|} = O(q_n^\sigma)$ mit $q_n = O(\log(n)^{r_q})$, $q_n^\sigma = O(\log(n)^{r_s})$.

Anmerkung: Auch hier gelten die Konvergenzraten, wenn sie für $\sup_{x \in [a_n - \bar{C}, b_n + \bar{C}]}$ (siehe Anmerkung zu Annahme (Z')) anstatt $\sup_{x \in [a_n - Cc_n, b_n + Cc_n]}$ gelten ($[a_n - Cc_n, b_n + Cc_n] \subset J_n$).

Anmerkung: Annahme (Z') und Annahme (M) betreffen beide das Wachstumsverhalten der Regressions- und der Skalierungsfunktion. Annahme (Z') beschränkt das Wachstum der Funktionen, durch das Wachstum des Kehrwerts der Dichten $f_{X_{i-1}}^{-1}$ und Annahme (M) beschränkt das Wachstum der Funktionen durch $q_n = O(\log(n)^{r_q})$. Im stationären Fall folgt aus Annahme (Z') und Annahme (X'), dass $\sup_{x \in J_n} (|m(x)| + |\sigma(x)|)^{2k} = O(q_n^f)$, und daraus der erste Teil von Annahme (M) (für $\mu = 0$) mit $q_n = (q_n^f)^{\frac{1}{2k}}$.

Beispiel 3.1 Ein einfaches Beispiel, das diese Annahmen erfüllt, ist das AR(1)-Modell

$$X_j = 0.5 \cdot X_{j-1} + \epsilon_j$$

mit Standardnormalverteilten Innovationen (ϵ_j) .

X_j hat in diesem Fall für alle j die Verteilung $\mathcal{N}\left(0, \frac{1}{1-0.5^2}\right)$ (da $X_j = \sum_{i=0}^{\infty} 0.5^i \epsilon_{j-i}$, siehe z.B. Beispiel 2.7 in Kreiß and Neuhaus [2006]).

Damit erfüllt (X_j) die ersten beiden Punkte in Annahme (X') nach Abschnitt 2.2 (siehe (2.4.1), S. 13). Der dritte Punkt in Annahme (X') bedeutet, dass $\max\left(\exp\left(\frac{a_n^2}{3}\right), \exp\left(\frac{b_n^2}{3}\right)\right) = O(\log(n)^{r_f})$.

Diese Annahme ist erfüllt für $I_n = \left[-\sqrt{\frac{8}{3} \log(\log(n))}, \sqrt{\frac{8}{3} \log(\log(n))}\right]$. Mit diesem I_n sind auch die Annahmen (I') und (M) erfüllt, wobei $q_n = O\left(\sqrt{\log(\log(n))}\right) = O(\log(n))$ und $q_n^\sigma = 1$. Annahme (E') gilt und Annahme (Z') reduziert sich auf $\sup_{x \in J_n} (|m(x)| + |\sigma(x)|)^{2k} f_{X_0}(x) = O(1)$ und

$\sup_{x, x' \in J_n} \left((|m(x)| + |\sigma(x)|)^k (|m(x')| + |\sigma(x')|)^k f_{X_0, X_{j-1}}(x, x') \right) = O(1)$ für alle $j > j^* + 1$, ($n \rightarrow \infty$) und ist durch das Modell ebenfalls erfüllt (siehe Abschnitt 4.2, S. 46 Bemerkung zu Annahme (Z) und Beispiel 4.1).

Mit einem Kern und einer Bandbreite, die die Bedingungen (K) und (C') erfüllen, sind alle Bedingungen für Satz 3.2 erfüllt.

Satz 3.2 Unter den Voraussetzungen (K), (C'), (I'), (X'), (E'), (Z') und (M) gilt

$$\sup_{x \in I_n} |\hat{m}(x) - m(x)| = o_P(1) \quad \text{und} \quad \sup_{x \in I_n} |\hat{\sigma}(x) - \sigma(x)| = o_P(1),$$

die Schätzer \hat{m} und $\hat{\sigma}$ sind also konsistent.

Beweis Nach (3.4.5) (S. 40) im Beweis zu Lemma 3.4 ist

$$\begin{aligned} \sup_{x \in I_n} |\hat{m}(x) - m(x)| &= O_P \left(c_n^{-\frac{1}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} q_n q_n^f + c_n^2 q_n q_n^f \right) \\ &= O_P \left(n^{-\frac{3}{8}} \log(n)^{r_1} + n^{-\frac{1}{2}} \log(n)^{r_2} \right) \quad r_1, r_2 < \infty \\ &= o_P(1) \end{aligned}$$

mit Annahme (X'), (M) und (C').

Genauso ist nach (3.4.6) (S. 42) im Beweis zu Lemma 3.4

$$\begin{aligned} \sup_{x \in I_n} |\hat{\sigma}(x) - \sigma(x)| &= O_P \left(c_n^{-\frac{1}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} q_n^2 (q_n^f)^2 q_n^\sigma + c_n^2 q_n^2 (q_n^f)^2 q_n^\sigma \right) \\ &= O_P \left(n^{-\frac{3}{8}} \log(n)^{r_3} + n^{-\frac{1}{2}} \log(n)^{r_4} \right) \quad r_3, r_4 < \infty \\ &= o_P(1) \end{aligned}$$

mit Annahme (X'), (M) und (C').

□

Das im Beweis benutzte Lemma 3.4 gibt die genauen Konvergenzraten der Schätzer an. Es wird im Folgenden bewiesen. Lemma 3.3 wird für den Beweis von Lemma 3.4 benötigt, und daher vorangestellt.

Lemma 3.3 Die Kernfunktion K sei dreimal differenzierbar und es gebe ein $\bar{K} < \infty$ mit $\sup_{u \in [-C, C]} |K^{(\mu)}(u)| \leq \bar{K}$ für $\mu = 0, 1, 2, 3$. Der Träger von K sei das kompakte Intervall $[-C, C]$.

Für dieses Lemma sei $J_n := [a_n - (C + c_n^{-\frac{1}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}})c_n, b_n + (C + c_n^{-\frac{1}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}})c_n]$.

Für die Dichten der Beobachtungen $f_{X_{i-1}}$, die Innovationen ϵ_i und die Regressions- und die Skalierungsfunktion m und σ gelte damit für alle $m_n \leq n$ mit $m_n^{-1} = o(1)$

$$\sup_{x \in J_n} \left((|m(x)| + |\sigma(x)|) \frac{1}{m_n} \max_{0 \leq S \leq n - m_n} \sum_{i=S+1}^{S+m_n} f_{X_{i-1}}(x) \right) = O(1),$$

$$\sup_{x \in J_n} \left((|m(x)| + |\sigma(x)|)^{2k} \frac{1}{m_n} \max_{0 \leq S \leq n-m_n} \sum_{i=S+1}^{S+m_n} (1 + E[\epsilon_i^4])^{k-1} f_{X_{i-1}}(x) \right) = O(1)$$

für $k = 1, 2$ und es gebe ein $1 \leq j^* < \infty$, so dass

$$\sup_{x, x' \in J_n} \left((|m(x)| + |\sigma(x)|)^k (|m(x')| + |\sigma(x')|)^k \frac{1}{m_n^2} \max_{0 \leq S \leq n-m_n} \sum_{\substack{i, j=S+1 \\ |i-j| > j^*}}^{S+m_n} f_{X_{i-1}, X_{j-1}}(x, x') \right) = O(1) \quad (3.3.1)$$

und

$$\sup_{x \in J_n} \left((|m(x)| + |\sigma(x)|)^{2k} \max_{j^*+1 \leq i \leq n-j^*} \sum_{j=i-j^*}^{i+j^*} (1 + E[\epsilon_j^4])^{k-1} f_{X_{j-1}}(x) \right) = O(1) \quad (3.3.2)$$

für $k = 1, 2$.

Zusätzlich sei

$$\frac{1}{n} \sum_{i=1}^n E \left[\left| X_i^{kb} \right| \right] = O(1), \quad \text{für } k = 1, 2$$

mit einem $b > 2$, für das gilt $D_n(b)^{-1} = o(1)$ mit

$$D_n(b) := n^{\frac{1}{2} - \frac{1}{b}} c_n^{\frac{1}{2}} \log(M_n)^{-1} \log(n)^{-\frac{3}{2}}$$

wobei $M_n := n^{\frac{1}{2}} c_n^{-\frac{1}{2}} \log(n)^{-\frac{1}{2}} (b_n - a_n)$.

Anmerkung: Für $c_n \sim n^{-\frac{1}{4}} \log(n)^{r_1}$ und $(b_n - a_n) = O(\log(n)^{r_2})$ mit $|r_1|, |r_2| < \infty$ bedeutet das $b > \frac{8}{3}$.

Für die Bandbreite c_n gelte außerdem $c_n \log(M_n) = o(1)$.

Desweiteren sei der Prozess der Beobachtungen $(X_i)_{i \in \mathbb{Z}}$ α -mischend mit Mischungskoeffizient $\alpha(n) = O(n^{-\beta})$, wobei β die Bedingungen

$$D_n(b)^{-\beta-1} n^{\frac{3}{2}} c_n^{-\frac{1}{2}} \log(n)^{-\frac{1}{2}} (b_n - a_n) = o(1)$$

und

$$(c_n \log(M_n))^{\beta+1} n^{\frac{3}{2}} c_n^{-\frac{1}{2}} \log(n)^{-\frac{1}{2}} (b_n - a_n) = o(1)$$

erfülle.

Anmerkung: Für $c_n \sim n^{-\frac{1}{4}} \log(n)^{r_1}$ mit $|r_1| < \infty$, $(b_n - a_n) = O(\log(n)^{r_2})$ mit $|r_2| < \infty$ und $b = 3$ bedeutet das $\beta > 38$.

Dann gilt

$$\sup_{x \in I_n} \left| \frac{1}{nc_n} \sum_{i=1}^n K^{(\nu)} \left(\frac{x - X_{i-1}}{c_n} \right) X_i^k - E \left[\frac{1}{nc_n} \sum_{i=1}^n K^{(\nu)} \left(\frac{x - X_{i-1}}{c_n} \right) X_i^k \right] \right| = O_P \left(c_n^{-\frac{1}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} \right)$$

für $\nu = 0, 1, 2$, $k = 0, 1, 2$.

Bemerkung: In Hansen [2008] wird die selbe Konvergenzrate unter der Bedingung, dass $(X_j)_{j \in \mathbb{Z}}$ streng stationär ist gezeigt (siehe Abschnitt 3.3, S. 43).

Beweis Um die Notation zu vereinfachen wird zunächst

$$\bar{\epsilon}_n := c_n^{-\frac{1}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}}$$

definiert, es ist also zu zeigen, dass

$$\sup_{x \in I_n} \left| \frac{1}{nc_n} \sum_{i=1}^n K^{(\nu)} \left(\frac{x - X_{i-1}}{c_n} \right) X_i^k - E \left[\frac{1}{nc_n} \sum_{i=1}^n K^{(\nu)} \left(\frac{x - X_{i-1}}{c_n} \right) X_i^k \right] \right| = O_P(\bar{\epsilon}_n),$$

was aus: $\exists \bar{M} < \infty$, so dass

$$P \left(\sup_{x \in I_n} \left| \frac{1}{nc_n} \sum_{i=1}^n K^{(\nu)} \left(\frac{x - X_{i-1}}{c_n} \right) X_i^k - E \left[\frac{1}{nc_n} \sum_{i=1}^n K^{(\nu)} \left(\frac{x - X_{i-1}}{c_n} \right) X_i^k \right] \right| > \bar{M} \bar{\epsilon}_n \right) = o(1)$$

folgt.

Als erster Schritt des Beweises wird gezeigt, dass sich der Wertebereich der X_i^k auf $|X_i^k| \leq n^{\frac{1}{b}} \log(n)$ einschränken lässt.

$$\begin{aligned} P \left(\max_{1 \leq i \leq n} |X_i^k| > n^{\frac{1}{b}} \log(n) \right) &\leq \sum_{i=1}^n P \left(|X_i^k| > n^{\frac{1}{b}} \log(n) \right) \\ &= \sum_{i=1}^n E \left[I \left\{ |X_i^k| > n^{\frac{1}{b}} \log(n) \right\} \right] \\ &= \frac{1}{\log(n)^b} \cdot \frac{1}{n} \sum_{i=1}^n E \left[n \log(n)^b I \left\{ |X_i^{kb}| > n \log(n)^b \right\} \right] \\ &\leq \frac{1}{\log(n)^b} \cdot \frac{1}{n} \sum_{i=1}^n E \left[|X_i^{kb}| I \left\{ |X_i^{kb}| > n \log(n)^b \right\} \right], \end{aligned}$$

da der Erwartungswert monoton ist und die betrachtete Funktion nur dann von Null verschieden ist, wenn $|X_i^{kb}| > n \log(n)^b$.

Da nach Voraussetzung

$$\frac{1}{n} \sum_{i=1}^n E \left[|X_i^{kb}| \right] = O(1),$$

ist auch

$$\frac{1}{n} \sum_{i=1}^n E \left[|X_i^{kb}| I \left\{ |X_i^{kb}| > n \log(n)^b \right\} \right] = O(1),$$

und daher

$$P \left(\max_{1 \leq i \leq n} |X_i^k| > n^{\frac{1}{b}} \log(n) \right) = O \left(\frac{1}{\log(n)^b} \right) = o(1).$$

Somit gilt

$$P \left(\underbrace{\sup_{x \in I_n} \left| \frac{1}{nc_n} \sum_{i=1}^n K^{(\nu)} \left(\frac{x - X_{i-1}}{c_n} \right) X_i^k - E \left[\frac{1}{nc_n} \sum_{i=1}^n K^{(\nu)} \left(\frac{x - X_{i-1}}{c_n} \right) X_i^k \right] \right|}_{=: B} > \bar{M} \bar{\epsilon}_n \right)$$

$$\begin{aligned}
&= P\left(B \cap \left(\max_{1 \leq i \leq n} |X_i^k| \leq n^{\frac{1}{b}} \log(n)\right)\right) + P\left(B \cap \left(\max_{1 \leq i \leq n} |X_i^k| > n^{\frac{1}{b}} \log(n)\right)\right) \\
&\leq P\left(B \cap \left(\max_{1 \leq i \leq n} |X_i^k| \leq n^{\frac{1}{b}} \log(n)\right)\right) + P\left(\max_{1 \leq i \leq n} |X_i^k| > n^{\frac{1}{b}} \log(n)\right) \\
&= P\left(B \cap \left(\max_{1 \leq i \leq n} |X_i^k| \leq n^{\frac{1}{b}} \log(n)\right)\right) + o(1).
\end{aligned}$$

Im Folgenden kann also davon ausgegangen werden, dass $|X_i^k| \leq n^{\frac{1}{b}} \log(n)$ für alle $i = 1, \dots, n$.

Der nächste Schritt des Beweises besteht darin, das Intervall I_n mit einem Netz, dessen Feinheit gegen Null geht, zu überdecken, und so die Supremum-Bildung durch Maximieren zu ersetzen. Die Stützstellen $x_j \in I_n$ seien so angeordnet, dass I_n durch $M_n^* \leq M_n = (b_n - a_n)(\bar{\epsilon}_n c_n)^{-1}$ Intervalle $\{x \in \mathbb{R} : |x - x_j| \leq \bar{\epsilon}_n c_n\}$ überdeckt werden kann.

Für den Kern K gilt nach dem Mittelwertsatz

$$\left|K^{(\nu)}(\tilde{u}) - K^{(\nu)}(u')\right| \begin{cases} = 0 & \text{für } \tilde{u}, u' \notin [-C, C] \\ \leq \sup_{u \in [-C, C]} |K^{(\nu+1)}(u)| |\tilde{u} - u'| \leq \bar{K} |\tilde{u} - u'| & \text{sonst} \end{cases},$$

da K nach Voraussetzung dreimal differenzierbar ist und $\sup_{u \in [-C, C]} |K^{(\nu+1)}(u)| \leq \bar{K}$ für $\nu = 0, 1, 2$. Damit ist für alle x mit $|x - x_j| \leq \bar{\epsilon}_n c_n$

$$\left|K^{(\nu)}\left(\frac{x - X_{i-1}}{c_n}\right) - K^{(\nu)}\left(\frac{x_j - X_{i-1}}{c_n}\right)\right| \leq \bar{K}_I \left(\frac{x_j - X_{i-1}}{c_n}\right) \frac{|x - x_j|}{c_n} \leq \bar{K}_I \left(\frac{x_j - X_{i-1}}{c_n}\right) \bar{\epsilon}_n,$$

wobei

$$\bar{K}_I \left(\frac{x_j - X_{i-1}}{c_n}\right) = \begin{cases} \bar{K} & \text{für } \left|\frac{x_j - X_{i-1}}{c_n}\right| \leq C + \bar{\epsilon}_n \\ 0 & \text{für } \left|\frac{x_j - X_{i-1}}{c_n}\right| > C + \bar{\epsilon}_n \quad \left(\Rightarrow \frac{x - X_{i-1}}{c_n}, \frac{x_j - X_{i-1}}{c_n} \notin [-C, C]\right), \end{cases}$$

und weiter

$$\begin{aligned}
&\left|\frac{1}{nc_n} \sum_{i=1}^n K^{(\nu)}\left(\frac{x - X_{i-1}}{c_n}\right) X_i^k - \frac{1}{nc_n} \sum_{i=1}^n K^{(\nu)}\left(\frac{x_j - X_{i-1}}{c_n}\right) X_i^k\right| \\
&\leq \frac{1}{nc_n} \sum_{i=1}^n \left|K^{(\nu)}\left(\frac{x - X_{i-1}}{c_n}\right) - K^{(\nu)}\left(\frac{x_j - X_{i-1}}{c_n}\right)\right| |X_i^k| \\
&\leq \bar{\epsilon}_n \frac{1}{nc_n} \sum_{i=1}^n \bar{K}_I \left(\frac{x_j - X_{i-1}}{c_n}\right) |X_i^k|.
\end{aligned}$$

Eine analoge Ungleichung gilt auch für den Erwartungswert, da dieser linear und monoton wachsend ist, und somit folgt

$$\begin{aligned}
&\left|\frac{1}{nc_n} \sum_{i=1}^n K^{(\nu)}\left(\frac{x - X_{i-1}}{c_n}\right) X_i^k - E\left[\frac{1}{nc_n} \sum_{i=1}^n K^{(\nu)}\left(\frac{x - X_{i-1}}{c_n}\right) X_i^k\right]\right| \\
&\leq \left|\frac{1}{nc_n} \sum_{i=1}^n K^{(\nu)}\left(\frac{x_j - X_{i-1}}{c_n}\right) X_i^k - E\left[\frac{1}{nc_n} \sum_{i=1}^n K^{(\nu)}\left(\frac{x_j - X_{i-1}}{c_n}\right) X_i^k\right]\right| \\
&\quad + \bar{\epsilon}_n \frac{1}{nc_n} \sum_{i=1}^n \bar{K}_I \left(\frac{x_j - X_{i-1}}{c_n}\right) |X_i^k| + \bar{\epsilon}_n E\left[\frac{1}{nc_n} \sum_{i=1}^n \bar{K}_I \left(\frac{x_j - X_{i-1}}{c_n}\right) |X_i^k|\right]
\end{aligned}$$

gleichmäßig für alle x mit $|x - x_j| \leq \bar{\epsilon}_n c_n$, und daher

$$\begin{aligned}
& \sup_{x \in I_n} \left| \frac{1}{nc_n} \sum_{i=1}^n K^{(\nu)} \left(\frac{x - X_{i-1}}{c_n} \right) X_i^k - E \left[\frac{1}{nc_n} \sum_{i=1}^n K^{(\nu)} \left(\frac{x - X_{i-1}}{c_n} \right) X_i^k \right] \right| \\
&= \max_{1 \leq j \leq M_n^*} \sup_{x \in I_n: |x - x_j| \leq \bar{\epsilon}_n c_n} \left| \frac{1}{nc_n} \sum_{i=1}^n K^{(\nu)} \left(\frac{x - X_{i-1}}{c_n} \right) X_i^k - E \left[\frac{1}{nc_n} \sum_{i=1}^n K^{(\nu)} \left(\frac{x - X_{i-1}}{c_n} \right) X_i^k \right] \right| \\
&\leq \max_{1 \leq j \leq M_n^*} \left(\left| \frac{1}{nc_n} \sum_{i=1}^n K^{(\nu)} \left(\frac{x_j - X_{i-1}}{c_n} \right) X_i^k - E \left[\frac{1}{nc_n} \sum_{i=1}^n K^{(\nu)} \left(\frac{x_j - X_{i-1}}{c_n} \right) X_i^k \right] \right| \right. \\
&\quad \left. + \bar{\epsilon}_n \frac{1}{nc_n} \sum_{i=1}^n \bar{K}_I \left(\frac{x_j - X_{i-1}}{c_n} \right) |X_i^k| + \bar{\epsilon}_n E \left[\frac{1}{nc_n} \sum_{i=1}^n \bar{K}_I \left(\frac{x_j - X_{i-1}}{c_n} \right) |X_i^k| \right] \right) \\
&\leq \max_{1 \leq j \leq M_n^*} \left| \frac{1}{nc_n} \sum_{i=1}^n K^{(\nu)} \left(\frac{x_j - X_{i-1}}{c_n} \right) X_i^k - E \left[\frac{1}{nc_n} \sum_{i=1}^n K^{(\nu)} \left(\frac{x_j - X_{i-1}}{c_n} \right) X_i^k \right] \right| \\
&\quad + \max_{1 \leq j \leq M_n^*} \bar{\epsilon}_n \left| \frac{1}{nc_n} \sum_{i=1}^n \bar{K}_I \left(\frac{x_j - X_{i-1}}{c_n} \right) |X_i^k| - E \left[\frac{1}{nc_n} \sum_{i=1}^n \bar{K}_I \left(\frac{x_j - X_{i-1}}{c_n} \right) |X_i^k| \right] \right| \\
&\quad + 2\bar{\epsilon}_n \max_{1 \leq j \leq M_n^*} E \left[\frac{1}{nc_n} \sum_{i=1}^n \bar{K}_I \left(\frac{x_j - X_{i-1}}{c_n} \right) |X_i^k| \right].
\end{aligned}$$

Der Summand in der letzten Zeile ist $O(\bar{\epsilon}_n)$, denn mit der Substitutionsregel gilt

$$\begin{aligned}
& \sup_{x \in I_n} E \left[\frac{1}{nc_n} \sum_{i=1}^n \bar{K}_I \left(\frac{x - X_{i-1}}{c_n} \right) |X_i^k| \right] \\
&= \sup_{x \in I_n} \frac{1}{nc_n} \sum_{i=1}^n \int \bar{K}_I \left(\frac{x - y}{c_n} \right) |z^k| f_{X_{i-1}, X_i}(y, z) d(y, z) \\
&= \sup_{x \in I_n} \frac{1}{nc_n} \sum_{i=1}^n \iint \bar{K}_I(u) |z^k| f_{X_{i-1}, X_i}(x - uc_n, z) c_n du dz \\
&= \sup_{x \in I_n} \frac{1}{n} \sum_{i=1}^n \iint \bar{K}_I(u) |z^k| f_{X_i | X_{i-1}}(z | x - uc_n) f_{X_{i-1}}(x - uc_n) dz du \\
&= \sup_{x \in I_n} \frac{1}{n} \sum_{i=1}^n \int \bar{K}_I(u) E \left[|X_i^k| \middle| X_{i-1} = x - uc_n \right] f_{X_{i-1}}(x - uc_n) du \\
&\leq \underbrace{\int \bar{K}_I(u) du}_{=2(C+\bar{\epsilon}_n)\bar{K}} \underbrace{\sup_{x \in I_n} \frac{1}{n} \sum_{i=1}^n E \left[|X_i^k| \middle| X_{i-1} = x \right] f_{X_{i-1}}(x)}_{=O(1) \text{ nach Voraussetzung}} \\
&= O(1).
\end{aligned}$$

Durch den folgenden Hilfssatz kann gezeigt werden, dass

$$\max_{1 \leq j \leq M_n^*} \left| \frac{1}{nc_n} \sum_{i=1}^n K^{(\nu)} \left(\frac{x_j - X_{i-1}}{c_n} \right) X_i^k - E \left[\frac{1}{nc_n} \sum_{i=1}^n K^{(\nu)} \left(\frac{x_j - X_{i-1}}{c_n} \right) X_i^k \right] \right| = O_P(\bar{\epsilon}_n) \quad (3.3.3)$$

und

$$\max_{1 \leq j \leq M_n^*} \left| \frac{1}{nc_n} \sum_{i=1}^n \bar{K}_I \left(\frac{x_j - X_{i-1}}{c_n} \right) |X_i^k| - E \left[\frac{1}{nc_n} \sum_{i=1}^n \bar{K}_I \left(\frac{x_j - X_{i-1}}{c_n} \right) |X_i^k| \right] \right| = O_P(\bar{\epsilon}_n) (= O_P(1)) \quad (3.3.4)$$

und somit der noch fehlende Teil des Beweises erbracht werden.

Hilfssatz 3.3.A Sei $h : \mathbb{R} \rightarrow \mathbb{R}$ eine messbare Funktion für die gilt, dass $|h(X_i)| \leq n^{\frac{1}{b}} \log(n)$ für alle $i = 1, \dots, n$. Zudem sei $K^* : \mathbb{R} \rightarrow \mathbb{R}_0^+$ eine messbare Funktion mit $\sup_{u \in \mathbb{R}} K^*(u) \leq \bar{K}$ und $K^*(u) = 0$ für $u \notin [-C - \bar{\epsilon}_n, C + \bar{\epsilon}_n]$.

Dann gibt es für jedes feste $x \in I_n$ ein $\bar{M} < \infty$, so dass für hinreichend großes n

$$\begin{aligned} P \left(\left| \frac{1}{nc_n} \sum_{i=1}^n K^* \left(\frac{x - X_{i-1}}{c_n} \right) h(X_i) - E \left[\frac{1}{nc_n} \sum_{i=1}^n K^* \left(\frac{x - X_{i-1}}{c_n} \right) h(X_i) \right] \right| > \bar{M} \bar{\epsilon}_n \right) \\ \leq 4 \exp \left(- \frac{nc_n \bar{M}^2 \bar{\epsilon}_n^2}{64 N_1 c_n A(N_1) + \frac{16}{3} \bar{M} \bar{K} N_1 n^{\frac{1}{b}} \bar{\epsilon}_n \log(n)} \right) + 4 \frac{n}{N_1} \alpha(N_1), \end{aligned} \quad (3.3.5)$$

wobei $\alpha(\cdot)$ der Mischungskoeffizient von $(X_i)_{i \in \mathbb{Z}}$ ist,

$$N_1 := \lfloor n^{1-\frac{1}{b}} c_n \bar{\epsilon}_n \log(M_n)^{-1} \log(n)^{-2} \rfloor$$

und

$$\begin{aligned} A(m) := & 2(C + \bar{\epsilon}_n) \bar{K}^2 \sup_{x \in J_n} \max_{j^*+1 \leq i \leq n-j^*} \sum_{j=i-j^*}^{i+j^*} \sqrt{E \left[h^2(X_j) \middle| X_{j-1} = x \right] f_{X_{j-1}}(x)} \\ & \cdot \sup_{x \in J_n} \frac{1}{m} \max_{0 \leq S \leq n-m} \sum_{i=S+1}^{S+m} \sqrt{E \left[h^2(X_i) \middle| X_{i-1} = x \right] f_{X_{i-1}}(x)} \\ & + (2(C + \bar{\epsilon}_n) \bar{K})^2 \left(\left(\sup_{x \in J_n} \frac{1}{m} \max_{0 \leq S \leq n-m} \sum_{i=S+1}^{S+m} E \left[|h(X_i)| \middle| X_{i-1} = x \right] f_{X_{i-1}}(x) \right)^2 \right. \\ & \left. + \sup_{x, x' \in J_n} \frac{1}{m^2} \max_{\substack{0 \leq S \leq n-m \\ i, j=S+1 \\ |i-j| > j^*}} \sum_{i, j=S+1}^{S+m} E \left[|h(X_i)| |h(X_j)| \middle| X_{i-1} = x, X_{j-1} = x' \right] f_{X_{i-1}, X_{j-1}}(x, x') \right), \end{aligned}$$

mit $m \leq n$ und j^* wie in (3.3.1) und (3.3.2) (S. 22).

Anmerkung: Damit ist $A(m_n) = O(1)$ für $h(X_i) = X_i^k$ und $h(X_i) = |X_i^k|$ und alle m_n mit $m_n^{-1} = o(1)$, siehe (3.3.8), S. 34.

Falls $n^{\frac{1}{b}-\frac{1}{2}} c_n^{-\frac{3}{2}} \log(n)^{\frac{5}{2}} = O(1)$ bleibt die Ungleichung (3.3.5) gültig, wenn man N_1 durch

$$N_2 := \lfloor n \bar{\epsilon}_n^2 \log(M_n)^{-1} \log(n)^{-1} \rfloor$$

ersetzt.

Anmerkung: Falls b so groß ist, dass $n^{\frac{1}{b}-\frac{1}{2}} c_n^{-\frac{3}{2}} \log(n)^{\frac{5}{2}} = O(1)$, folgt aus (3.3.5) (mit N_1) nicht mehr, dass $P \left(\left| \frac{1}{nc_n} \sum_{i=1}^n K^* \left(\frac{x - X_{i-1}}{c_n} \right) h(X_i) - E \left[\frac{1}{nc_n} \sum_{i=1}^n K^* \left(\frac{x - X_{i-1}}{c_n} \right) h(X_i) \right] \right| > \bar{M} \bar{\epsilon}_n \right) = o(1)$. Daher braucht man in diesem Fall die Ungleichung mit N_2 statt N_1 .

Beweis Messbare Funktionen erhalten nach Fan and Yao [2005] 2.6.1 (ii) Mischungs-Eigenschaften, daher ist

$$Y_i := K^* \left(\frac{x - X_{i-1}}{c_n} \right) h(X_i) - E \left[K^* \left(\frac{x - X_{i-1}}{c_n} \right) h(X_i) \right]$$

für festes $x \in I_n$ und $n \in \mathbb{N}$ α -mischend mit dem selben Mischungskoeffizienten wie $(X_i)_i$. Außerdem ist

$$|Y_i| \leq 2\bar{K}n^{\frac{1}{b}} \log(n)$$

und

$$E[Y_i] = 0,$$

das Theorem 2.1 aus Liebscher [1996] ist also auf $\sum_{i=1}^n Y_i$ anwendbar. Dieses Theorem basiert auf Theorem 5 in Rio [1995].

Mit der Cauchy-Schwarz-Ungleichung kann gezeigt werden, dass für j^* wie in (3.3.1) und (3.3.2) (S. 22), alle $0 \leq T \leq n - 1$ und alle $N \leq n$

$$\begin{aligned} & E \left[\left(\sum_{i=T+1}^{\min(T+N,n)} Y_i \right)^2 \right] \\ &= \sum_{i,j=T+1}^{\min(T+N,n)} E[Y_i Y_j] \\ &= \sum_{i,j=T+1}^{\min(T+N,n)} E \left[K^* \left(\frac{x - X_{i-1}}{c_n} \right) h(X_i) K^* \left(\frac{x - X_{j-1}}{c_n} \right) h(X_j) \right] \\ &\quad - \sum_{i,j=T+1}^{\min(T+N,n)} E \left[K^* \left(\frac{x - X_{i-1}}{c_n} \right) h(X_i) \right] E \left[K^* \left(\frac{x - X_{j-1}}{c_n} \right) h(X_j) \right] \\ &= \sum_{i=T+1}^{\min(T+N,n)} \sum_{j=\max(i-j^*, T+1)}^{\min(i+j^*, T+N,n)} E \left[K^* \left(\frac{x - X_{i-1}}{c_n} \right) h(X_i) K^* \left(\frac{x - X_{j-1}}{c_n} \right) h(X_j) \right] \\ &\quad + \sum_{\substack{i,j=T+1 \\ |i-j| > j^*}}^{\min(T+N,n)} E \left[K^* \left(\frac{x - X_{i-1}}{c_n} \right) h(X_i) K^* \left(\frac{x - X_{j-1}}{c_n} \right) h(X_j) \right] \\ &\quad - \sum_{i,j=T+1}^{\min(T+N,n)} E \left[K^* \left(\frac{x - X_{i-1}}{c_n} \right) h(X_i) \right] E \left[K^* \left(\frac{x - X_{j-1}}{c_n} \right) h(X_j) \right] \\ &\leq \sum_{i=T+1}^{\min(T+N,n)} \sum_{j=\max(i-j^*, T+1)}^{\min(i+j^*, T+N,n)} \sqrt{E \left[K^{*2} \left(\frac{x - X_{i-1}}{c_n} \right) h^2(X_i) \right]} \sqrt{E \left[K^{*2} \left(\frac{x - X_{j-1}}{c_n} \right) h^2(X_j) \right]} \\ &\quad + \sum_{\substack{i,j=T+1 \\ |i-j| > j^*}}^{\min(T+N,n)} E \left[K^* \left(\frac{x - X_{i-1}}{c_n} \right) h(X_i) K^* \left(\frac{x - X_{j-1}}{c_n} \right) h(X_j) \right] \\ &\quad - \sum_{i,j=T+1}^{\min(T+N,n)} E \left[K^* \left(\frac{x - X_{i-1}}{c_n} \right) h(X_i) \right] E \left[K^* \left(\frac{x - X_{j-1}}{c_n} \right) h(X_j) \right] \\ &= \sum_{i=T+1}^{\min(T+N,n)} \sum_{j=\max(i-j^*, T+1)}^{\min(i+j^*, T+N,n)} \sqrt{\int K^{*2} \left(\frac{x - y}{c_n} \right) h^2(z) f_{X_{i-1}, X_i}(y, z) d(y, z)} \end{aligned}$$

$$\begin{aligned}
& \cdot \sqrt{\int K^{*2} \left(\frac{x-y}{c_n} \right) h^2(z) f_{X_{j-1}, X_j}(y, z) d(y, z)} \\
& + \sum_{\substack{i, j=T+1 \\ |i-j| > j^*}}^{\min(T+N, n)} \int K^* \left(\frac{x-y}{c_n} \right) h(z) K^* \left(\frac{x-y'}{c_n} \right) h(z') f_{X_{i-1}, X_i, X_{j-1}, X_j}(y, z, y', z') d(y, z, y', z') \\
& - \left(\sum_{i=T+1}^{\min(T+N, n)} \int K^* \left(\frac{x-y}{c_n} \right) h(z) f_{X_{i-1}, X_i}(y, z) d(y, z) \right)^2
\end{aligned}$$

und weiter durch Substitution

$$\begin{aligned}
& E \left[\left(\sum_{i=T+1}^{\min(T+N, n)} Y_i \right)^2 \right] \\
\leq & \sum_{i=T+1}^{\min(T+N, n)} \sum_{j=\max(i-j^*, T+1)}^{\min(i+j^*, T+N, n)} \sqrt{\int \int K^{*2}(u) h^2(z) f_{X_{i-1}, X_i}(x - uc_n, z) c_n du dz} \\
& \cdot \sqrt{\int \int K^{*2}(u) h^2(z) f_{X_{j-1}, X_j}(x - uc_n, z) c_n du dz} \\
& + \sum_{\substack{i, j=T+1 \\ |i-j| > j^*}}^{\min(T+N, n)} \iiint K^*(u) h(z) K^*(u') h(z') f_{X_{i-1}, X_i, X_{j-1}, X_j}(x - uc_n, z, x - u'c_n, z') c_n^2 du dz du' dz' \\
& - \left(\sum_{i=T+1}^{\min(T+N, n)} \iint K^*(u) h(z) f_{X_{i-1}, X_i}(x - uc_n, z) c_n du dz \right)^2 \\
= & c_n \sum_{i=T+1}^{\min(T+N, n)} \sum_{j=\max(i-j^*, T+1)}^{\min(i+j^*, T+N, n)} \sqrt{\int K^{*2}(u) \int h^2(z) f_{X_{i-1}, X_i}(x - uc_n, z) dz du} \\
& \cdot \sqrt{\int K^{*2}(u) \int h^2(z) f_{X_{j-1}, X_j}(x - uc_n, z) dz du} \\
& + c_n^2 \sum_{\substack{i, j=T+1 \\ |i-j| > j^*}}^{\min(T+N, n)} \iint K^*(u) K^*(u') \iint h(z) h(z') f_{X_{i-1}, X_i, X_{j-1}, X_j}(x - uc_n, z, x - u'c_n, z') dz dz' du du' \\
& - c_n^2 \left(\sum_{i=T+1}^{\min(T+N, n)} \int K^*(u) \int h(z) f_{X_{i-1}, X_i}(x - uc_n, z) dz du \right)^2.
\end{aligned}$$

Da $K^*(u) = 0$ für $u \notin [-C - \bar{\epsilon}_n, C + \bar{\epsilon}_n]$, kann $(x - uc_n)$ durch $\sup_{x \in J_n}(x)$ abgeschätzt werden, und es gilt weiter

$$\begin{aligned}
& E \left[\left(\sum_{i=T+1}^{\min(T+N, n)} Y_i \right)^2 \right] \\
\leq & c_n \int K^{*2}(u) du \sup_{x, x' \in J_n} \sum_{i=T+1}^{\min(T+N, n)} \sum_{j=\max(i-j^*, T+1)}^{\min(i+j^*, T+N, n)} \sqrt{\int h^2(z) f_{X_{i-1}, X_i}(x, z) dz}
\end{aligned}$$

$$\begin{aligned}
& \cdot \sqrt{\int h^2(z) f_{X_{j-1}, X_j}(x', z) dz} \\
& + c_n^2 \int K^*(u) du \int K^*(u') du' \sup_{x, x' \in J_n} \sum_{\substack{i, j=T+1 \\ |i-j| > j^*}}^{\min(T+N, n)} \iint |h(z)| |h(z')| f_{X_{i-1}, X_i, X_{j-1}, X_j}(x, z, x', z') dz dz' \\
& + c_n^2 \left(\int K^*(u) du \sup_{x \in J_n} \sum_{i=T+1}^{\min(T+N, n)} \int |h(z)| f_{X_{i-1}, X_i}(x, z) dz \right)^2 \\
= & c_n \int K^{*2}(u) du \sup_{x, x' \in J_n} \sum_{i=T+1}^{\min(T+N, n)} \sum_{j=\max(i-j^*, T+1)}^{\min(i+j^*, T+N, n)} \sqrt{\int h^2(z) f_{X_i|X_{i-1}}(z|x) f_{X_{i-1}}(x) dz} \\
& \cdot \sqrt{\int h^2(z) f_{X_j|X_{j-1}}(z|x') f_{X_{j-1}}(x') dz} \\
& + c_n^2 \left(\int K^*(u) du \right)^2 \sup_{x, x' \in J_n} \sum_{\substack{i, j=T+1 \\ |i-j| > j^*}}^{\min(T+N, n)} \iint |h(z)| |h(z')| f_{X_i, X_j|X_{i-1}, X_{j-1}}(z, z'|x, x') f_{X_{i-1}, X_{j-1}}(x, x') dz dz' \\
& + c_n^2 \left(\int K^*(u) du \right)^2 \left(\sup_{x \in J_n} \sum_{i=T+1}^{\min(T+N, n)} \int |h(z)| f_{X_i|X_{i-1}}(z|x) f_{X_{i-1}}(x) dz \right)^2 \\
= & c_n \int K^{*2}(u) du \sup_{x, x' \in J_n} \sum_{i=T+1}^{\min(T+N, n)} \left(\sqrt{E \left[h^2(X_i) \middle| X_{i-1} = x \right]} f_{X_{i-1}}(x) \right. \\
& \left. \cdot \sum_{j=\max(i-j^*, T+1)}^{\min(i+j^*, T+N, n)} \sqrt{E \left[h^2(X_j) \middle| X_{j-1} = x' \right]} f_{X_{j-1}}(x') \right) \\
& + c_n^2 \left(\int K^*(u) du \right)^2 \sup_{x, x' \in J_n} \sum_{\substack{i, j=T+1 \\ |i-j| > j^*}}^{\min(T+N, n)} E \left[|h(X_i)| |h(X_j)| \middle| X_{i-1} = x, X_{j-1} = x' \right] f_{X_{i-1}, X_{j-1}}(x, x') \\
& + c_n^2 \left(\int K^*(u) du \right)^2 \left(\sup_{x \in J_n} \sum_{i=T+1}^{\min(T+N, n)} E \left[|h(X_i)| \middle| X_{i-1} = x \right] f_{X_{i-1}}(x) \right)^2 \\
\leq & N c_n 2(C + \bar{\epsilon}_n) \bar{K}^2 \sup_{x \in J_n} \frac{1}{N} \sum_{i=T+1}^{\min(T+N, n)} \sqrt{E \left[h^2(X_i) \middle| X_{i-1} = x \right]} f_{X_{i-1}}(x) \\
& \cdot \sup_{x \in J_n} \max_{j^*+1 \leq i \leq n-j^*} \sum_{j=\max(i-j^*, T+1)}^{\min(i+j^*, T+N)} \sqrt{E \left[h^2(X_j) \middle| X_{j-1} = x \right]} f_{X_{j-1}}(x) \\
& + N^2 c_n^2 (2(C + \bar{\epsilon}_n) \bar{K})^2 \sup_{x, x' \in J_n} \frac{1}{N^2} \sum_{\substack{i, j=T+1 \\ |i-j| > j^*}}^{\min(T+N, n)} E \left[|h(X_i)| |h(X_j)| \middle| X_{i-1} = x, X_{j-1} = x' \right] f_{X_{i-1}, X_{j-1}}(x, x') \\
& + N^2 c_n^2 (2(C + \bar{\epsilon}_n) \bar{K})^2 \left(\sup_{x \in J_n} \frac{1}{N} \sum_{i=1}^{\min(T+N, n)} E \left[|h(X_i)| \middle| X_{i-1} = x \right] f_{X_{i-1}}(x) \right)^2 \\
\leq & (N c_n + N^2 c_n^2) A(N),
\end{aligned}$$

da mit $\sqrt{E \left[h^2(X_i) \middle| X_{i-1} = x \right] f_{X_{i-1}}(x)} \geq 0$,

$E \left[|h(X_i)| |h(X_j)| \middle| X_{i-1} = x, X_{j-1} = x' \right] f_{X_{i-1}, X_{j-1}}(x, x') \geq 0$

und $E \left[|h(X_i)| \middle| X_{i-1} = x \right] f_{X_{i-1}}(x) \geq 0$ folgt, dass

$$\begin{aligned}
& \sum_{i=T+1}^{\min(T+N, n)} \sqrt{E \left[h^2(X_i) \middle| X_{i-1} = x \right] f_{X_{i-1}}(x)} \\
& \leq \max_{0 \leq S \leq n-N} \sum_{i=S+1}^{S+N} \sqrt{E \left[h^2(X_i) \middle| X_{i-1} = x \right] f_{X_{i-1}}(x)}, \\
& \sum_{j=\max(i-j^*, T+1)}^{\min(i+j^*, T+N)} \sqrt{E \left[h^2(X_j) \middle| X_{j-1} = x \right] f_{X_{j-1}}(x)} \\
& \leq \sum_{j=i-j^*}^{i+j^*} \sqrt{E \left[h^2(X_j) \middle| X_{j-1} = x \right] f_{X_{j-1}}(x)}, \\
& \sum_{\substack{i, j=T+1 \\ |i-j| > j^*}}^{\min(T+N, n)} E \left[|h(X_i)| |h(X_j)| \middle| X_{i-1} = x, X_{j-1} = x' \right] f_{X_{i-1}, X_{j-1}}(x, x') \\
& \leq \max_{0 \leq S \leq n-N} \sum_{\substack{i, j=S+1 \\ |i-j| > j^*}}^{S+N} E \left[|h(X_i)| |h(X_j)| \middle| X_{i-1} = x, X_{j-1} = x' \right] f_{X_{i-1}, X_{j-1}}(x, x')
\end{aligned}$$

und

$$\begin{aligned}
& \sum_{i=T+1}^{\min(T+N, n)} E \left[|h(X_i)| \middle| X_{i-1} = x \right] f_{X_{i-1}}(x) \\
& \leq \max_{0 \leq S \leq n-N} \sum_{i=S+1}^{S+N} E \left[|h(X_i)| \middle| X_{i-1} = x \right] f_{X_{i-1}}(x)
\end{aligned}$$

für alle $0 \leq T \leq n-1$ und $N \leq n$.

Damit gilt nach Liebscher [1996]

$$\begin{aligned}
& P \left(\left| \sum_{i=1}^n Y_i \right| > nc_n \bar{M} \bar{\epsilon}_n \right) \\
& \leq 4 \exp \left(- \frac{n^2 c_n^2 \bar{M}^2 \bar{\epsilon}_n^2}{64 \frac{n}{N} (Nc_n + N^2 c_n^2) A(N) + \frac{8}{3} nc_n \bar{M} \bar{\epsilon}_n N 2 \bar{K} n^{\frac{1}{b}} \log(n)} \right) + 4 \frac{n}{N} \alpha(N) \\
& = 4 \exp \left(- \frac{nc_n \bar{M}^2 \bar{\epsilon}_n^2}{64 (1 + Nc_n) A(N) + \frac{16}{3} \bar{M} \bar{K} N n^{\frac{1}{b}} \bar{\epsilon}_n \log(n)} \right) + 4 \frac{n}{N} \alpha(N)
\end{aligned}$$

für alle $N \in \mathbb{N}$ und \bar{M} , die die Bedingungen

$$1 \leq N \leq n \tag{3.3.6}$$

und

$$nc_n \bar{M} \bar{\epsilon}_n > 4N_2 \bar{K} n^{\frac{1}{b}} \log(n) \quad (3.3.7)$$

erfüllen.

Da

$$P\left(\left|\frac{1}{nc_n} \sum_{i=1}^n Y_i\right| > \bar{M} \bar{\epsilon}_n\right) = P\left(\left|\sum_{i=1}^n Y_i\right| > nc_n \bar{M} \bar{\epsilon}_n\right)$$

muss nur noch gezeigt werden, dass N_1 und N_2 die Bedingungen (3.3.6) und (3.3.7) für ein $\bar{M} > \infty$ erfüllen.

Mit der Definition von $\bar{\epsilon}_n$ folgt

$$n > N_1 = \lfloor n^{1-\frac{1}{b}} c_n \bar{\epsilon}_n \log(M_n)^{-1} \log(n)^{-2} \rfloor = \underbrace{\lfloor n^{\frac{1}{2}-\frac{1}{b}} c_n^{\frac{1}{2}} \log(n)^{-\frac{3}{2}} \log(M_n)^{-1} \rfloor}_{\xrightarrow[n \rightarrow \infty]{\infty} \text{nach Voraussetzung}} > 1 \quad \text{für hinr. große } n$$

und

$$n > N_2 = \lfloor n \bar{\epsilon}_n^2 \log(M_n)^{-1} \log(n)^{-1} \rfloor = \underbrace{\lfloor (c_n \log(M_n))^{-1} \rfloor}_{\xrightarrow[n \rightarrow \infty]{\infty} \text{nach Voraussetzung}} > 1 \quad \text{für hinreichend große } n.$$

Außerdem gibt es ein $\bar{M} < \infty$, so dass N_1 die Bedingung (3.3.7) erfüllt, denn

$$\begin{aligned} nc_n \bar{M} \bar{\epsilon}_n &> 4N_1 2\bar{K} n^{\frac{1}{b}} \log(n) = 8\bar{K} n^{\frac{1}{b}} \log(n) \lfloor n^{1-\frac{1}{b}} c_n \bar{\epsilon}_n \log(M_n)^{-1} \log(n)^{-2} \rfloor \\ \Leftrightarrow \bar{M} &> 8\bar{K} n^{-1+\frac{1}{b}} c_n^{-1} \bar{\epsilon}_n^{-1} \log(n) \lfloor n^{1-\frac{1}{b}} c_n \bar{\epsilon}_n \log(M_n)^{-1} \log(n)^{-2} \rfloor \\ &= O(\log(M_n)^{-1} \log(n)^{-1}) \\ &= o(1). \end{aligned}$$

Unter der zusätzliche Bedingung dass $n^{\frac{1}{b}-\frac{1}{2}} c_n^{-\frac{3}{2}} \log(n)^{\frac{5}{2}} = O(1)$ gibt es desweiteren ein $\bar{M} < \infty$, so dass N_2 die Bedingung (3.3.7) erfüllt, denn

$$\begin{aligned} nc_n \bar{M} \bar{\epsilon}_n &> 4N_2 \bar{K} n^{\frac{1}{b}} \log(n) = 8\bar{K} n^{\frac{1}{b}} \log(n) \lfloor n \bar{\epsilon}_n^2 \log(M_n)^{-1} \log(n)^{-1} \rfloor \\ \Leftrightarrow \bar{M} &> 8\bar{K} n^{-1+\frac{1}{b}} c_n^{-1} \bar{\epsilon}_n^{-1} \log(n) \lfloor n \bar{\epsilon}_n^2 \log(M_n)^{-1} \log(n)^{-1} \rfloor \\ &= O\left(n^{\frac{1}{b}} c_n^{-1} \bar{\epsilon}_n \log(M_n)^{-1}\right) \\ &= O\left(n^{\frac{1}{b}-\frac{1}{2}} c_n^{-\frac{3}{2}} \log(n)^{\frac{1}{2}} \log(M_n)^{-1}\right) \\ &= O\left(n^{\frac{1}{b}-\frac{1}{2}} c_n^{-\frac{3}{2}} \log(n)^{\frac{5}{2}}\right) O(\log(n)^{-2} \log(M_n)^{-1}) \\ &= o(1). \end{aligned}$$

□

Fortsetzung Beweis Lemma 3.3:

Mit Hilfssatz 3.3.A können die Abschätzungen (3.3.3) und (3.3.4) bewiesen werden.

Zunächst wird dafür das $A(\cdot)$ aus dem Hilfssatz näher untersucht. Es gilt für alle m_n mit $m_n^{-1} = o(1)$

$$\begin{aligned}
& \sup_{x \in J_n} \frac{1}{m_n} \max_{0 \leq S \leq n-m_n} \sum_{i=S+1}^{S+m_n} E \left[|X_i| \middle| X_{i-1} = x \right] f_{X_{i-1}}(x) \\
&= \sup_{x \in J_n} \frac{1}{m_n} \max_{0 \leq S \leq n-m_n} \sum_{i=S+1}^{S+m_n} E [|m(x) + \sigma(x)\epsilon_i|] f_{X_{i-1}}(x) \\
&\leq \sup_{x \in J_n} \frac{1}{m_n} \max_{0 \leq S \leq n-m_n} \sum_{i=S+1}^{S+m_n} (|m(x)| + |\sigma(x)| E[|\epsilon_i|]) f_{X_{i-1}}(x) \\
&\leq \sup_{x \in J_n} \left((|m(x)| + 2|\sigma(x)|) \frac{1}{m_n} \max_{0 \leq S \leq n-m_n} \sum_{i=S+1}^{S+m_n} f_{X_{i-1}}(x) \right) \\
&\quad \text{da } E[|\epsilon_i|] \leq 1 + E[\epsilon_i^2] \text{ (} |x|^m \leq 1 + |x|^n \text{ für alle } x \in \mathbb{R} \text{ und } m \leq n \text{) und } E[\epsilon_i^2] = 1 \forall i \text{ nach Vss.} \\
&= O(1)
\end{aligned}$$

nach Voraussetzung, sowie analog

$$\begin{aligned}
& \sup_{x \in J_n} \frac{1}{m_n} \max_{0 \leq S \leq n-m_n} \sum_{i=S+1}^{S+m_n} E \left[X_i^2 \middle| X_{i-1} = x \right] f_{X_{i-1}}(x) \\
&= \sup_{x \in J_n} \left((m^2(x) + \sigma^2(x)) \frac{1}{m_n} \max_{0 \leq S \leq n-m_n} \sum_{i=S+1}^{S+m_n} f_{X_{i-1}}(x) \right) \\
&\quad \text{da } E[\epsilon_i] = 0 \text{ und } E[\epsilon_i^2] = 1 \forall i \\
&\leq \sup_{x \in J_n} \left((|m(x)| + |\sigma(x)|)^2 \frac{1}{m_n} \max_{0 \leq S \leq n-m_n} \sum_{i=S+1}^{S+m_n} f_{X_{i-1}}(x) \right) \\
&= O(1)
\end{aligned}$$

nach Voraussetzung.

Damit gilt ebenfalls

$$\begin{aligned}
& \sup_{x \in J_n} \frac{1}{m_n} \max_{0 \leq S \leq n-m_n} \sum_{i=S+1}^{S+m_n} \sqrt{E \left[X_i^2 \middle| X_{i-1} = x \right]} f_{X_{i-1}}(x) \\
&= \sup_{x \in J_n} \frac{1}{m_n} \max_{0 \leq S \leq n-m_n} \left(\sum_{i=S+1}^{S+m_n} \sqrt{E \left[X_i^2 \middle| X_{i-1} = x \right]} f_{X_{i-1}}(x) \right. \\
&\quad \left. \begin{aligned} & \left(E \left[X_i^2 \middle| X_{i-1} = x \right] f_{X_{i-1}}(x) \leq 1 \right) \\ & + \sum_{i=S+1}^{S+m_n} \sqrt{E \left[X_i^2 \middle| X_{i-1} = x \right]} f_{X_{i-1}}(x) \\ & \left. \left(E \left[X_i^2 \middle| X_{i-1} = x \right] f_{X_{i-1}}(x) > 1 \right) \right) \end{aligned} \right)
\end{aligned}$$

$$\begin{aligned}
&\leq \sup_{x \in J_n} \frac{1}{m_n} \max_{0 \leq S \leq n-m_n} \left(m_n + \sum_{i=S+1}^{S+m_n} E \left[X_i^2 \middle| X_{i-1} = x \right] f_{X_{i-1}}(x) \right) \\
&\leq 1 + \sup_{x \in J_n} \frac{1}{m_n} \max_{0 \leq S \leq n-m_n} \sum_{i=S+1}^{S+m_n} E \left[X_i^2 \middle| X_{i-1} = x \right] f_{X_{i-1}}(x) \\
&= O(1),
\end{aligned}$$

und analog dazu auch

$$\sup_{x \in J_n} \frac{1}{m_n} \max_{0 \leq S \leq n-m_n} \sum_{i=S+1}^{S+m_n} \sqrt{E \left[X_i^4 \middle| X_{i-1} = x \right] f_{X_{i-1}}(x)} = O(1),$$

da

$$\begin{aligned}
&\sup_{x \in J_n} \frac{1}{m_n} \max_{0 \leq S \leq n-m_n} \sum_{i=S+1}^{S+m_n} E \left[X_i^4 \middle| X_{i-1} = x \right] f_{X_{i-1}}(x) \\
&\leq \sup_{x \in J_n} \left((|m(x)| + |\sigma(x)|)^4 \frac{1}{m_n} \max_{0 \leq S \leq n-m_n} \sum_{i=S+1}^{S+m_n} (1 + E[\epsilon_i^4]) f_{X_{i-1}}(x) \right) \\
&= O(1).
\end{aligned}$$

Analog folgt für $k = 1, 2$ und j^* wie in (3.3.1) und (3.3.2)

$$\begin{aligned}
&\sup_{x, x' \in J_n} \frac{1}{m_n^2} \max_{0 \leq S \leq n-m_n} \sum_{\substack{i, j=S+1 \\ |i-j| > j^*}}^{S+m_n} E \left[|X_i^k| |X_j^k| \middle| X_{i-1} = x, X_{j-1} = x' \right] f_{X_{i-1}, X_{j-1}}(x, x') \\
&= O \left(\sup_{x, x' \in J_n} \left((|m(x)| + |\sigma(x)|)^k (|m(x')| + |\sigma(x')|)^k \frac{1}{m_n^2} \max_{0 \leq S \leq n-m_n} \sum_{\substack{i, j=S+1 \\ |i-j| > j^*}}^{S+m_n} f_{X_{i-1}, X_{j-1}}(x, x') \right) \right) \\
&= O(1)
\end{aligned}$$

nach Voraussetzung, da ϵ_i und ϵ_j für $i \neq j$ unabhängig sind.

Außerdem gilt ebenfalls für $k = 1, 2$

$$\sup_{x \in J_n} \max_{j^*+1 \leq i \leq n-j^*} \sum_{j=i-j^*}^{i+j^*} \sqrt{E \left[X_j^{2k} \middle| X_{j-1} = x \right] f_{X_{j-1}}(x)} = O(1),$$

da

$$\begin{aligned}
&\sup_{x \in J_n} \max_{j^*+1 \leq i \leq n-j^*} \sum_{j=i-j^*}^{i+j^*} E \left[X_j^{2k} \middle| X_{j-1} = x \right] f_{X_{j-1}}(x) \\
&\leq \sup_{x \in J_n} \left((|m(x)| + |\sigma(x)|)^{2k} \max_{j^*+1 \leq i \leq n-j^*} \sum_{j=i-j^*}^{i+j^*} (1 + E[\epsilon_j^4])^{k-1} f_{X_{j-1}}(x) \right)
\end{aligned}$$

$$= O(1)$$

nach Voraussetzung.

Somit gilt also für $h(X_i) = X_i^k$ und $h(X_i) = |X_i^k|$, $k = 1, 2$ und alle m_n mit $m_n^{-1} = o(1)$

$$A(m_n) = O(1). \quad (3.3.8)$$

Nun wird gezeigt, dass die Abschätzungen (3.3.3) und (3.3.4) gelten.

Mit Hilfssatz 3.3.A und $M_n^* \leq M_n$ folgt

$$\begin{aligned} & P \left(\max_{1 \leq j \leq M_n^*} \left| \frac{1}{nc_n} \sum_{i=1}^n K^{(\nu)} \left(\frac{x_j - X_{i-1}}{c_n} \right) X_i^k - E \left[\frac{1}{nc_n} \sum_{i=1}^n K^{(\nu)} \left(\frac{x_j - X_{i-1}}{c_n} \right) X_i^k \right] \right| > \bar{M} \bar{\epsilon}_n \right) \\ & \leq \sum_{j=1}^{M_n^*} P \left(\left| \frac{1}{nc_n} \sum_{i=1}^n K^{(\nu)} \left(\frac{x_j - X_{i-1}}{c_n} \right) X_i^k - E \left[\frac{1}{nc_n} \sum_{i=1}^n K^{(\nu)} \left(\frac{x_j - X_{i-1}}{c_n} \right) X_i^k \right] \right| > \bar{M} \bar{\epsilon}_n \right) \\ & \leq M_n \left(4 \exp \left(- \frac{nc_n \bar{M}^2 \bar{\epsilon}_n^2}{64 (1 + N_1 c_n) A(N_1) + \frac{16}{3} \bar{M} \bar{K} N_1 n^{\frac{1}{b}} \bar{\epsilon}_n \log(n)} \right) + 4 \frac{n}{N_1} \alpha(N_1) \right) \end{aligned}$$

und genauso

$$\begin{aligned} & P \left(\max_{1 \leq j \leq M_n^*} \left| \frac{1}{nc_n} \sum_{i=1}^n \bar{K}_I \left(\frac{x_j - X_{i-1}}{c_n} \right) |X_i^k| - E \left[\frac{1}{nc_n} \sum_{i=1}^n \bar{K}_I \left(\frac{x_j - X_{i-1}}{c_n} \right) |X_i^k| \right] \right| > \bar{M} \bar{\epsilon}_n \right) \\ & \leq M_n \left(4 \exp \left(- \frac{nc_n \bar{M}^2 \bar{\epsilon}_n^2}{64 (1 + N_1 c_n) A(N_1) + \frac{16}{3} \bar{M} \bar{K} N_1 n^{\frac{1}{b}} \bar{\epsilon}_n \log(n)} \right) + 4 \frac{n}{N_1} \alpha(N_1) \right). \end{aligned}$$

Dies lässt sich weiter abschätzen mit

$$\begin{aligned} & M_n \left(4 \exp \left(- \frac{nc_n \bar{M}^2 \bar{\epsilon}_n^2}{64 (1 + N_1 c_n) A(N_1) + \frac{16}{3} \bar{M} \bar{K} N_1 n^{\frac{1}{b}} \bar{\epsilon}_n \log(n)} \right) + 4 \frac{n}{N_1} \alpha(N_1) \right) \\ & = 4 \exp \left(\log(M_n) - \frac{nc_n \bar{M}^2 \bar{\epsilon}_n^2}{64 (1 + N_1 c_n) A(N_1) + \frac{16}{3} \bar{M} \bar{K} N_1 n^{\frac{1}{b}} \bar{\epsilon}_n \log(n)} \right) \quad (3.3.9) \\ & \quad + O \left(M_n n N_1^{-\beta-1} \right). \end{aligned}$$

Die Zeile (3.3.9) ist $o(1)$, wenn der Exponent gegen minus unendlich geht. Dies ist der Fall, wenn

$$\frac{64 (1 + N_1 c_n) A(N_1) + \frac{16}{3} \bar{M} \bar{K} N_1 n^{\frac{1}{b}} \bar{\epsilon}_n \log(n)}{nc_n \bar{M}^2 \bar{\epsilon}_n^2} = o \left(\frac{1}{\log(M_n)} \right).$$

Mit $N_1 = \lfloor n^{1-\frac{1}{b}} c_n \bar{\epsilon}_n \log(M_n)^{-1} \log(n)^{-2} \rfloor$ gilt

$$\begin{aligned} & \log(M_n) \frac{64 (1 + N_1 c_n) A(N_1) + \frac{16}{3} \bar{M} \bar{K} N_1 n^{\frac{1}{b}} \bar{\epsilon}_n \log(n)}{nc_n \bar{M}^2 \bar{\epsilon}_n^2} \\ & = O \left(n^{-1} c_n^{-1} \bar{\epsilon}_n^{-2} + n^{-\frac{1}{b}} c_n \bar{\epsilon}_n^{-1} \log(n)^{-2} + \log(n)^{-1} \right), \text{ da } A(N_1) = O(1) \text{ nach (3.3.8)} \\ & = O \left(\log(n)^{-1} + n^{-\frac{1}{b} + \frac{1}{2}} c_n^{\frac{3}{2}} \log(n)^{-\frac{5}{2}} + \log(n)^{-1} \right) \\ & = o(1), \end{aligned}$$

falls $n^{-\frac{1}{b}+\frac{1}{2}}c_n^{\frac{3}{2}}\log(n)^{-\frac{5}{2}} = o(1)$. Anderenfalls ist $n^{\frac{1}{b}-\frac{1}{2}}c_n^{-\frac{3}{2}}\log(n)^{\frac{5}{2}} = O(1)$ und N_1 kann durch N_2 ersetzt werden.

Auch dann gilt mit $N_2 := \lfloor n\bar{\epsilon}_n^2 \log(M_n)^{-1} \log(n)^{-1} \rfloor$

$$\begin{aligned} & \log(M_n) \frac{64(1+N_2c_n)A(N_2) + \frac{16}{3}\bar{M}\bar{K}N_2n^{\frac{1}{b}}\bar{\epsilon}_n\log(n)}{nc_n\bar{M}^2\bar{\epsilon}_n^2} \\ &= O\left(n^{-1}c_n^{-1}\bar{\epsilon}_n^{-2} + \log(n)^{-1} + n^{\frac{1}{b}}\bar{\epsilon}_nc_n^{-1}\right) \\ &= O\left(\log(n)^{-1} + n^{\frac{1}{b}-\frac{1}{2}}c_n^{-\frac{3}{2}}\log(n)^{\frac{1}{2}}\right) \\ &= O\left(\log(n)^{-1} + \left(n^{\frac{1}{b}-\frac{1}{2}}c_n^{-\frac{3}{2}}\log(n)^{\frac{5}{2}}\right)\log(n)^{-2}\right) \\ &= o(1). \end{aligned}$$

Es bleibt also zu zeigen, dass

$$M_n n N_1^{-\beta-1} = o(1)$$

und

$$M_n n N_2^{-\beta-1} = o(1).$$

Beides gilt nach Voraussetzung, denn

$$\begin{aligned} M_n n N_1^{-\beta-1} &\leq (b_n - a_n)\bar{\epsilon}_n^{-1}c_n^{-1}nN_1^{-\beta-1} \\ &= O\left((b_n - a_n)n^{\frac{3}{2}}c_n^{-\frac{1}{2}}\log(n)^{-\frac{1}{2}}\left(n^{\frac{1}{2}-\frac{1}{b}}c_n^{\frac{1}{2}}\log(M_n)^{-1}\log(n)^{-\frac{3}{2}}\right)^{-\beta-1}\right) \\ &= O\left((b_n - a_n)n^{\frac{3}{2}}c_n^{-\frac{1}{2}}\log(n)^{-\frac{1}{2}}D_n(b)^{-\beta-1}\right) \\ &= o(1) \end{aligned}$$

und

$$\begin{aligned} M_n n N_2^{-\beta-1} &\leq (b_n - a_n)\bar{\epsilon}_n^{-1}c_n^{-1}nN_2^{-\beta-1} \\ &= O\left((b_n - a_n)n^{\frac{3}{2}}c_n^{-\frac{1}{2}}\log(n)^{-\frac{1}{2}}(c_n^{-1}\log(M_n)^{-1})^{-\beta-1}\right) \\ &= O\left((b_n - a_n)n^{\frac{3}{2}}c_n^{-\frac{1}{2}}\log(n)^{-\frac{1}{2}}(c_n\log(M_n))^{\beta+1}\right) \\ &= o(1). \end{aligned}$$

□

Lemma 3.4 Es seien die Voraussetzungen von Lemma 3.3 gegeben.

Zusätzlich gelte für die Kernfunktion K , dass $\int K(u)udu = 0$ und $K(C) = K(-C) = K'(C) = K'(-C) = 0$.

Für dieses Lemma sei $J_n := [a_n - Cc_n, b_n + Cc_n]$.

Die Regressionsfunktion m und die Skalierungsfunktion σ seien jeweils viermal differenzierbar.

Ebenso seien die Dichten f_{X_i} der Beobachtungen viermal differenzierbar und es gelte $\sup_{x \in J_n} \frac{1}{n} \sum_{i=1}^n |f_{X_{i-1}}^{(\mu)}(x)| = O(1)$ für $\mu = 0, 1, 2, 3, 4$.

Dann gelten die folgenden Konvergenzraten

$$(i) \quad \sup_{x \in I_n} \left| \frac{1}{nc_n} \sum_{i=1}^n K\left(\frac{x-X_{i-1}}{c_n}\right) - \frac{1}{n} \sum_{i=1}^n f_{X_{i-1}}(x) \right| = O_P\left(c_n^{-\frac{1}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} + c_n^2\right)$$

$$(ii) \quad a) \quad \sup_{x \in I_n} \left| \frac{\hat{m}(x) - m(x)}{\sigma(x)} \right| = O_P\left(\left(c_n^{-\frac{1}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} + c_n^2\right) q_n q_n^f q_n^\sigma\right)$$

$$b) \quad \sup_{x \in I_n} \left| \frac{\hat{\sigma}(x) - \sigma(x)}{\sigma(x)} \right| = O_P\left(\left(c_n^{-\frac{1}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} + c_n^2\right) (q_n q_n^f q_n^\sigma)^2\right),$$

falls $\left(c_n^{-\frac{1}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} + c_n^2\right) q_n q_n^f = o(1)$

mit $q_n := \max(\max_{\mu=0,1,2,3,4} \sup_{x \in J_n} |m^{(\mu)}(x)|, \max_{\mu=0,1,2,3,4} \sup_{x \in J_n} |\sigma^{(\mu)}(x)|)$
und $q_n^\sigma := \frac{1}{\inf_{x \in I_n} \sigma(x)}$, sowie $q_n^f := \frac{1}{\inf_{x \in I_n} \frac{1}{n} \sum_{i=1}^n f_{X_{i-1}}(x)}$

$$(iii) \quad a) \quad \sup_{x \in I_n} \left| \frac{\partial}{\partial x} \left(\frac{\hat{m}(x) - m(x)}{\sigma(x)} \right) \right| = o_P(1)$$

$$b) \quad \sup_{x \in I_n} \left| \frac{\partial}{\partial x} \left(\frac{\hat{\sigma}(x) - \sigma(x)}{\sigma(x)} \right) \right| = o_P(1)$$

falls $\left(c_n^{-\frac{3}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} + c_n^2\right) q_n^2 (q_n^f q_n^\sigma)^4 = o(1)$

$$(iv) \quad a) \quad \sup_{x, y \in I_n, x \neq y} \frac{\left| \frac{\partial}{\partial x} \left(\frac{\hat{m}(x) - m(x)}{\sigma(x)} \right) - \frac{\partial}{\partial y} \left(\frac{\hat{m}(y) - m(y)}{\sigma(y)} \right) \right|}{|y-x|^\delta} = o_P(1)$$

$$b) \quad \sup_{x, y \in I_n, x \neq y} \frac{\left| \frac{\partial}{\partial x} \left(\frac{\hat{\sigma}(x) - \sigma(x)}{\sigma(x)} \right) - \frac{\partial}{\partial y} \left(\frac{\hat{\sigma}(y) - \sigma(y)}{\sigma(y)} \right) \right|}{|y-x|^\delta} = o_P(1)$$

für alle δ mit $\left(c_n^{-\frac{3}{2}-\delta} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} + c_n^{2-\delta}\right) q_n^2 (q_n^f q_n^\sigma)^6 = o(1)$.

Anmerkung: Für $\sup_{x \in J_n} |m^{(\mu)}(x)| = O(\log(n)^{r_1})$, $\sup_{x \in J_n} |\sigma^{(\mu)}(x)| = O(\log(n)^{r_2})$, $\mu = 0, 1, 2, 3, 4$,
 $\frac{1}{\inf_{x \in I_n} \sigma(x)} = O(\log(n)^{r_3})$, $\frac{1}{\inf_{x \in I_n} \frac{1}{n} \sum_{j=1}^n f_{X_{j-1}}(x)} = O(\log(n)^{r_4})$ und $c_n \sim n^{-\frac{1}{4}} \log(n)^{r_5}$,
mit $|r_1|, \dots, |r_5| < \infty$, bedeutet die Forderung $\delta < \frac{1}{2}$.

Beweis Sei $g_{k,i}(x) := E[X_i^k | X_{i-1} = x] f_{X_{i-1}}(x)$ und $\hat{g}_k(x) := \frac{1}{nc_n} \sum_{i=1}^n K\left(\frac{x-X_{i-1}}{c_n}\right) X_i^k$ mit $k = 0, 1, 2$ und $i = 1, \dots, n$.

Für $\nu = 0, 1, 2$ gilt für den Schätzer \hat{g}_k mit Lemma 3.3

$$\begin{aligned}
& \sup_{x \in I_n} \left| \frac{\partial^\nu}{\partial x^\nu} \left(\hat{g}_k(x) - \frac{1}{n} \sum_{i=1}^n g_{k,i}(x) \right) \right| \\
& \leq \sup_{x \in I_n} \left| \frac{\partial^\nu}{\partial x^\nu} \left(\hat{g}_k(x) - E[\hat{g}_k(x)] \right) \right| + \sup_{x \in I_n} \left| \frac{\partial^\nu}{\partial x^\nu} \left(E[\hat{g}_k(x)] - \frac{1}{n} \sum_{i=1}^n g_{k,i}(x) \right) \right| \\
& = \frac{1}{c_n^\nu} \sup_{x \in I_n} \left| \frac{1}{nc_n} \sum_{i=1}^n K^{(\nu)} \left(\frac{x - X_{i-1}}{c_n} \right) X_i^k - E \left[\frac{1}{nc_n} \sum_{i=1}^n K^{(\nu)} \left(\frac{x - X_{i-1}}{c_n} \right) X_i^k \right] \right| \\
& \quad + \sup_{x \in I_n} \left| \frac{\partial^\nu}{\partial x^\nu} \left(E[\hat{g}_k(x)] - \frac{1}{n} \sum_{i=1}^n g_{k,i}(x) \right) \right| \\
& = O_P \left(c_n^{-\frac{1}{2}-\nu} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} \right) + \sup_{x \in I_n} \left| \frac{\partial^\nu}{\partial x^\nu} \left(E[\hat{g}_k(x)] - \frac{1}{n} \sum_{i=1}^n g_{k,i}(x) \right) \right|, \tag{3.4.1}
\end{aligned}$$

und weiter

$$\begin{aligned}
& \sup_{x \in I_n} \left| \frac{\partial^\nu}{\partial x^\nu} \left(E[\hat{g}_k(x)] - \frac{1}{n} \sum_{i=1}^n g_{k,i}(x) \right) \right| \\
& = \sup_{x \in I_n} \left| E \left[E \left[\frac{1}{c_n^\nu} \frac{1}{nc_n} \sum_{i=1}^n K^{(\nu)} \left(\frac{x - X_{i-1}}{c_n} \right) X_i^k \middle| X_{i-1} \right] \right] - \frac{1}{n} \sum_{i=1}^n g_{k,i}^{(\nu)}(x) \right| \\
& = \sup_{x \in I_n} \left| \frac{1}{c_n^\nu} \frac{1}{nc_n} \sum_{i=1}^n E \left[K^{(\nu)} \left(\frac{x - X_{i-1}}{c_n} \right) E[X_i^k | X_{i-1}] \right] - \frac{1}{n} \sum_{i=1}^n g_{k,i}^{(\nu)}(x) \right| \\
& = \sup_{x \in I_n} \left| \frac{1}{c_n^\nu} \frac{1}{nc_n} \sum_{i=1}^n \int K^{(\nu)} \left(\frac{x - y}{c_n} \right) E[X_i^k | X_{i-1} = y] f_{X_{i-1}}(y) dy - \frac{1}{n} \sum_{i=1}^n g_{k,i}^{(\nu)}(x) \right|.
\end{aligned}$$

Mit der Substitutionsregel folgt hieraus

$$\begin{aligned}
& \sup_{x \in I_n} \left| \frac{\partial^\nu}{\partial x^\nu} \left(E[\hat{g}_k(x)] - \frac{1}{n} \sum_{i=1}^n g_{k,i}(x) \right) \right| \\
& = \sup_{x \in I_n} \left| \frac{1}{c_n^\nu} \frac{1}{nc_n} \sum_{i=1}^n \int K^{(\nu)}(u) E[X_i^k | X_{i-1} = x - uc_n] f_{X_{i-1}}(x - uc_n) c_n du - \frac{1}{n} \sum_{i=1}^n g_{k,i}^{(\nu)}(x) \right| \\
& = \sup_{x \in I_n} \left| \frac{1}{c_n^\nu} \frac{1}{n} \sum_{i=1}^n \int K^{(\nu)}(u) g_{k,i}(x - uc_n) du - \frac{1}{n} \sum_{i=1}^n g_{k,i}^{(\nu)}(x) \right|. \tag{3.4.2}
\end{aligned}$$

Für den ersten Term von (3.4.2) gilt mit einer Taylorentwicklung von $g_{k,i}(x - uc_n)$ in x und einem $\xi(x, u)$ zwischen x und $x - uc_n$

$$\begin{aligned}
& \frac{1}{c_n^\nu} \frac{1}{n} \sum_{i=1}^n \int K^{(\nu)}(u) g_{k,i}(x - uc_n) du \\
& = \frac{1}{c_n^\nu} \left(\sum_{\mu=0}^{\nu+1} \frac{1}{n} \sum_{i=1}^n g_{k,i}^{(\mu)}(x) \frac{(-c_n)^\mu}{\mu!} \int K^{(\nu)}(u) u^\mu du + \frac{1}{n} \sum_{i=1}^n \int g_{k,i}^{(\nu+2)}(\xi(x, u)) \frac{(-c_n)^{\nu+2}}{(\nu+2)!} K^{(\nu)}(u) u^{\nu+2} du \right) \\
& = \frac{1}{n} \sum_{i=1}^n g_{k,i}^{(\nu)}(x) + \frac{1}{c_n^\nu} \frac{1}{n} \sum_{i=1}^n \int g_{k,i}^{(\nu+2)}(\xi(x, u)) \frac{(-c_n)^{\nu+2}}{(\nu+2)!} K^{(\nu)}(u) u^{\nu+2} du,
\end{aligned}$$

da aus $\int K(u)du = 1$, $\int K(u)udu = 0$ und $K(C) = K(-C) = K'(C) = K'(-C) = 0$ mit partieller Integration folgt

$$\int K^{(\nu)}(u) u^\mu du = 0 \quad \text{für } \mu \leq \nu + 1, \mu \neq \nu$$

und

$$\int K^{(\nu)}(u) u^\nu du = (-1)^\nu \nu!$$

für $\nu = 0, 1, 2$.

Damit folgt aus (3.4.2)

$$\begin{aligned} & \sup_{x \in I_n} \left| \frac{\partial^\nu}{\partial x^\nu} \left(E[\hat{g}_k(x)] - \frac{1}{n} \sum_{i=1}^n g_{k,i}(x) \right) \right| \\ &= \sup_{x \in I_n} \left| \frac{1}{c_n^\nu} \frac{1}{n} \sum_{i=1}^n \int g_{k,i}^{(\nu+2)}(\xi(x,u)) \frac{(-c_n)^{\nu+2}}{(\nu+2)!} K^{(\nu)}(u) u^{\nu+2} du \right| \\ &\leq \frac{c_n^2}{(\nu+2)!} \sup_{x \in J_n} \left| \frac{1}{n} \sum_{i=1}^n g_{k,i}^{(\nu+2)}(x) \right| \left| \int K^{(\nu)}(u) u^{\nu+2} du \right|. \end{aligned} \quad (3.4.3)$$

Es gilt für alle $i = 1, \dots, n$

$$E[X_i | X_{i-1} = x] = m(x)$$

und

$$E[X_i^2 | X_{i-1} = x] = \text{Var}(X_i | X_{i-1} = x) + (E[X_i | X_{i-1} = x])^2 = \sigma^2(x) + m^2(x),$$

daher ist für alle $i = 1, \dots, n$

$$\sup_{x \in J_n} \left| \frac{\partial^\mu}{\partial x^\mu} E[X_i^k | X_{i-1} = x] \right| = O(q_n^k) \quad \text{für } k = 0, 1, 2, \mu = 0, 1, 2, 3, 4,$$

denn $\sup_{x \in J_n} |m^{(\mu)}(x)| = O(q_n)$ und $\sup_{x \in J_n} |\sigma^{(\mu)}(x)| = O(q_n)$ für alle $\mu = 0, 1, 2, 3, 4$ nach Voraussetzung.

Da ebenfalls nach Voraussetzung für $\mu = 0, 1, 2, 3, 4$ gilt $\sup_{x \in J_n} \frac{1}{n} \sum_{i=1}^n |f_{X_{i-1}}^{(\mu)}(x)| = O(1)$, ist

$$\sup_{x \in J_n} \left| \frac{1}{n} \sum_{i=1}^n g_{k,i}^{(\mu)}(x) \right| = O(q_n^k) \quad \text{für } k = 0, 1, 2, \mu = 0, 1, 2, 3, 4,$$

woraus mit (3.4.3) folgt

$$\sup_{x \in I_n} \left| \frac{\partial^\nu}{\partial x^\nu} \left(E[\hat{g}_k(x)] - \frac{1}{n} \sum_{i=1}^n g_{k,i}(x) \right) \right| = O(c_n^2 q_n^k),$$

da sich aus $\int K(u)u^2 du \leq 2C\bar{K}C^2 < \infty$ und $K(C) = K(-C) = K'(C) = K'(-C) = 0$

$$\left| \int K^{(\nu)}(u) u^{\nu+2} du \right| < \infty$$

mit partieller Integration ergibt.

Somit gilt für den Schätzer \hat{g}_k mit (3.4.1) die Konvergenzrate

$$\sup_{x \in I_n} \left| \frac{\partial^\nu}{\partial x^\nu} \left(\hat{g}_k(x) - \frac{1}{n} \sum_{i=1}^n g_{k,i}(x) \right) \right| = O_P \left(n^{-\frac{1}{2}} c_n^{-\frac{1}{2}-\nu} \log(n)^{\frac{1}{2}} + c_n^2 q_n^k \right) \quad (3.4.4)$$

für $\nu = 0, 1, 2$ und $k = 0, 1, 2$.

Damit ist Behauptung (i) bewiesen, denn

$$\begin{aligned} \sup_{x \in I_n} \left| \frac{1}{n c_n} \sum_{i=1}^n K \left(\frac{x - X_{i-1}}{c_n} \right) - \frac{1}{n} \sum_{i=1}^n f_{X_{i-1}}(x) \right| &= \sup_{x \in I_n} \left| \hat{g}_0(x) - \frac{1}{n} \sum_{i=1}^n g_{0,i}(x) \right| \\ &= O_P \left(n^{-\frac{1}{2}} c_n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} + c_n^2 \right). \end{aligned}$$

Der Schätzer \hat{m} ist definiert als

$$\hat{m}(x) = \frac{\frac{1}{n c_n} \sum_{i=1}^n K \left(\frac{x - X_{i-1}}{c_n} \right) X_i}{\frac{1}{n c_n} \sum_{i=1}^n K \left(\frac{x - X_{i-1}}{c_n} \right)} = \frac{\hat{g}_1(x)}{\hat{g}_0(x)}.$$

Damit folgt für ihn aus der eben gezeigten Konvergenz von \hat{g}_1 und \hat{g}_0 ((3.4.4))

$$\begin{aligned} &\sup_{x \in I_n} |\hat{m}(x) - m(x)| \\ &= \sup_{x \in I_n} \left| \frac{\hat{g}_1(x) - m(x) \hat{g}_0(x)}{\hat{g}_0(x)} \right| \\ &= \sup_{x \in I_n} \left| \frac{\hat{g}_1(x) - m(x) \frac{1}{n} \sum_{i=1}^n f_{X_{i-1}}(x) + m(x) \left(\frac{1}{n} \sum_{i=1}^n g_{0,i}(x) - \hat{g}_0(x) \right)}{\hat{g}_0(x)} \right| \\ &= \sup_{x \in I_n} \left| \frac{\hat{g}_1(x) - \frac{1}{n} \sum_{i=1}^n g_{1,i}(x) + m(x) \left(\frac{1}{n} \sum_{i=1}^n f_{X_{i-1}}(x) - \hat{g}_0(x) \right)}{\hat{g}_0(x)} \right| \\ &= \sup_{x \in I_n} \left| \frac{\frac{1}{\frac{1}{n} \sum_{i=1}^n f_{X_{i-1}}(x)} \left(\hat{g}_1(x) - \frac{1}{n} \sum_{i=1}^n g_{1,i}(x) + m(x) \left(\frac{1}{n} \sum_{i=1}^n g_{0,i}(x) - \hat{g}_0(x) \right) \right)}{\frac{\hat{g}_0(x)}{\frac{1}{n} \sum_{i=1}^n g_{0,i}(x)}} \right| \\ &\leq \frac{\frac{1}{\inf_{x \in I_n} \frac{1}{n} \sum_{i=1}^n f_{X_{i-1}}(x)} \sup_{x \in I_n} \left| \hat{g}_1(x) - \frac{1}{n} \sum_{i=1}^n g_{1,i}(x) \right|}{\inf_{x \in I_n} \left| \frac{\hat{g}_0(x)}{\frac{1}{n} \sum_{i=1}^n g_{0,i}(x)} \right|} \\ &\quad + \frac{\frac{1}{\inf_{x \in I_n} \frac{1}{n} \sum_{i=1}^n f_{X_{i-1}}(x)} \sup_{x \in I_n} |m(x)| \sup_{x \in I_n} \left| \frac{1}{n} \sum_{i=1}^n g_{0,i}(x) - \hat{g}_0(x) \right|}{\inf_{x \in I_n} \left| \frac{\hat{g}_0(x)}{\frac{1}{n} \sum_{i=1}^n g_{0,i}(x)} \right|} \\ &= \frac{O_P \left(c_n^{-\frac{1}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} q_n^f + c_n^2 q_n q_n^f \right) + O_P \left(c_n^{-\frac{1}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} q_n q_n^f + c_n^2 q_n q_n^f \right)}{\inf_{x \in I_n} \left| \frac{\hat{g}_0(x)}{\frac{1}{n} \sum_{i=1}^n g_{0,i}(x)} \right|}, \end{aligned}$$

denn nach Voraussetzung gilt $\frac{1}{\inf_{x \in I_n} \frac{1}{n} \sum_{i=1}^n f_{X_{i-1}}(x)} = O(q_n^f)$ und $\sup_{x \in I_n} |m(x)| = O(q_n)$.

Da auch

$$\inf_{x \in I_n} \left| \frac{\hat{g}_0(x)}{\frac{1}{n} \sum_{i=1}^n g_{0,i}(x)} \right| = \inf_{x \in I_n} \left| 1 + \left(\frac{\hat{g}_0(x) - \frac{1}{n} \sum_{i=1}^n g_{0,i}(x)}{\frac{1}{n} \sum_{i=1}^n g_{0,i}(x)} \right) \right|$$

$$\begin{aligned}
&\geq \inf_{x \in I_n} \left| 1 - \left| \frac{\hat{g}_0(x) - \frac{1}{n} \sum_{i=1}^n g_{0,i}(x)}{\frac{1}{n} \sum_{i=1}^n g_{0,i}(x)} \right| \right| \\
&\geq \inf_{x \in I_n} \left(1 - \left| \frac{\hat{g}_0(x) - \frac{1}{n} \sum_{i=1}^n g_{0,i}(x)}{\frac{1}{n} \sum_{i=1}^n g_{0,i}(x)} \right| \right) \\
&= 1 - \sup_{x \in I_n} \left| \frac{\hat{g}_0(x) - \frac{1}{n} \sum_{i=1}^n g_{0,i}(x)}{\frac{1}{n} \sum_{i=1}^n g_{0,i}(x)} \right| \\
&\geq 1 - \frac{\sup_{x \in I_n} \left| \hat{g}_0(x) - \frac{1}{n} \sum_{i=1}^n g_{0,i}(x) \right|}{\inf_{x \in I_n} \left| \frac{1}{n} \sum_{i=1}^n g_{0,i}(x) \right|} \\
&= 1 + O_P \left(c_n^{-\frac{1}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} q_n^f + c_n^2 q_n^f \right)
\end{aligned}$$

gilt

$$\sup_{x \in I_n} |\hat{m}(x) - m(x)| = O_P \left(c_n^{-\frac{1}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} q_n q_n^f + c_n^2 q_n q_n^f \right). \quad (3.4.5)$$

Somit gilt die Behauptung (ii) a)

$$\begin{aligned}
\sup_{x \in I_n} \left| \frac{\hat{m}(x) - m(x)}{\sigma(x)} \right| &\leq \frac{1}{\inf_{x \in I_n} |\sigma(x)|} \sup_{x \in I_n} |\hat{m}(x) - m(x)| \\
&= O_P \left(c_n^{-\frac{1}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} q_n q_n^f q_n^\sigma + c_n^2 q_n q_n^f q_n^\sigma \right),
\end{aligned}$$

da $\frac{1}{\inf_{x \in I_n} |\sigma(x)|} = O(q_n^\sigma)$ nach Voraussetzung.

Der Schätzer $\hat{\sigma}^2$ ist definiert als

$$\hat{\sigma}^2(x) = \frac{\frac{1}{nc_n} \sum_{i=1}^n K\left(\frac{x-X_{i-1}}{c_n}\right) X_i^2}{\frac{1}{nc_n} \sum_{i=1}^n K\left(\frac{x-X_{i-1}}{c_n}\right)} - \left(\frac{\frac{1}{nc_n} \sum_{i=1}^n K\left(\frac{x-X_{i-1}}{c_n}\right) X_i}{\frac{1}{nc_n} \sum_{i=1}^n K\left(\frac{x-X_{i-1}}{c_n}\right)} \right)^2 = \frac{\hat{g}_2(x)}{\hat{g}_0(x)} - \hat{m}^2(x).$$

Damit folgt für ihn analog zur Berechnung für \hat{m} aus (3.4.4)

$$\begin{aligned}
&\sup_{x \in I_n} |\hat{\sigma}^2(x) - \sigma^2(x)| \\
&= \sup_{x \in I_n} \left| \frac{\hat{g}_2(x) - \sigma^2(x) \hat{g}_0(x) - \hat{m}^2(x) \hat{g}_0(x)}{\hat{g}_0(x)} \right| \\
&= \sup_{x \in I_n} \left| \frac{\hat{g}_2(x) - (\sigma^2(x) + m^2(x)) \frac{1}{n} \sum_{i=1}^n f_{X_{i-1}}(x) + (m^2(x) - \hat{m}^2(x)) \hat{g}_0(x)}{\hat{g}_0(x)} \right. \\
&\quad \left. + \frac{(\sigma^2(x) + m^2(x)) \left(\frac{1}{n} \sum_{i=1}^n g_{0,i}(x) - \hat{g}_0(x) \right)}{\hat{g}_0(x)} \right| \\
&\leq \frac{\frac{1}{\inf_{x \in I_n} \frac{1}{n} \sum_{i=1}^n f_{X_{i-1}}(x)} \sup_{x \in I_n} \left| \hat{g}_2(x) - \frac{1}{n} \sum_{i=1}^n g_{2,i}(x) \right|}{\inf_{x \in I_n} \left| \frac{\hat{g}_0(x)}{\frac{1}{n} \sum_{i=1}^n g_{0,i}(x)} \right|} \\
&\quad + \frac{\frac{1}{\inf_{x \in I_n} \frac{1}{n} \sum_{i=1}^n f_{X_{i-1}}(x)} \sup_{x \in I_n} |m^2(x) - \hat{m}^2(x)| \sup_{x \in I_n} |\hat{g}_0(x)|}{\inf_{x \in I_n} \left| \frac{\hat{g}_0(x)}{\frac{1}{n} \sum_{i=1}^n g_{0,i}(x)} \right|}
\end{aligned}$$

$$\begin{aligned}
& + \frac{\frac{1}{\inf_{x \in I_n} \frac{1}{n} \sum_{i=1}^n f_{X_{i-1}}(x)} (\sup_{x \in I_n} |\sigma^2(x)| + \sup_{x \in I_n} |m^2(x)|) \sup_{x \in I_n} \left| \frac{1}{n} \sum_{i=1}^n g_{0,i}(x) - \hat{g}_0(x) \right|}{\inf_{x \in I_n} \left| \frac{\hat{g}_0(x)}{\frac{1}{n} \sum_{i=1}^n g_{0,i}(x)} \right|} \\
= & O_P \left(c_n^{-\frac{1}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} q_n^f + c_n^2 q_n^2 q_n^f \right) \\
& + \frac{\frac{1}{\inf_{x \in I_n} \frac{1}{n} \sum_{i=1}^n f_{X_{i-1}}(x)} \sup_{x \in I_n} |m^2(x) - \hat{m}^2(x)| \sup_{x \in I_n} |\hat{g}_0(x)|}{\inf_{x \in I_n} \left| \frac{\hat{g}_0(x)}{\frac{1}{n} \sum_{i=1}^n g_{0,i}(x)} \right|} \\
& + O_P \left(c_n^{-\frac{1}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} q_n^2 q_n^f + c_n^2 q_n^2 q_n^f \right)
\end{aligned}$$

da nach Voraussetzung ebenfalls gilt $\sup_{x \in I_n} |\sigma(x)| = O(q_n)$.

Die Konvergenzrate für den noch verbleibenden Term ergibt sich mit (3.4.5) aus

$$\begin{aligned}
\sup_{x \in I_n} |\hat{m}^2(x) - m^2(x)| &= \sup_{x \in I_n} \left| (\hat{m}(x) - m(x))^2 + 2m(x) (\hat{m}(x) - m(x)) \right| \\
&\leq \left(\sup_{x \in I_n} |\hat{m}(x) - m(x)| \right)^2 + 2 \sup_{x \in I_n} |m(x)| \sup_{x \in I_n} |\hat{m}(x) - m(x)| \\
&= O \left(\sup_{x \in I_n} |m(x)| \sup_{x \in I_n} |\hat{m}(x) - m(x)| \right) \\
&= O_P \left(c_n^{-\frac{1}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} q_n^2 q_n^f + c_n^2 q_n^2 q_n^f \right),
\end{aligned}$$

da $\sup_{x \in I_n} |m(x)| = O(q_n)$, und mit (3.4.4) aus

$$\begin{aligned}
\sup_{x \in I_n} |\hat{g}_0(x)| &\leq \sup_{x \in I_n} \left| \frac{1}{n} \sum_{i=1}^n f_{X_{i-1}}(x) \right| + \sup_{x \in I_n} \left| \hat{g}_0(x) - \frac{1}{n} \sum_{i=1}^n g_{0,i}(x) \right| \\
&= O(1),
\end{aligned}$$

da $\sup_{x \in I_n} \left| \frac{1}{n} \sum_{i=1}^n f_{X_{i-1}}(x) \right| = O(1)$ nach Voraussetzung.

Es folgt also

$$\begin{aligned}
& \frac{\frac{1}{\inf_{x \in I_n} \frac{1}{n} \sum_{i=1}^n f_{X_{i-1}}(x)} \sup_{x \in I_n} |m^2(x) - \hat{m}^2(x)| \sup_{x \in I_n} |\hat{g}_0(x)|}{\inf_{x \in I_n} \left| \frac{\hat{g}_0(x)}{\frac{1}{n} \sum_{i=1}^n g_{0,i}(x)} \right|} \\
= & O_P \left(c_n^{-\frac{1}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} q_n^2 (q_n^f)^2 + c_n^2 q_n^2 (q_n^f)^2 \right)
\end{aligned}$$

und somit für $\hat{\sigma}^2$

$$\sup_{x \in I_n} |\hat{\sigma}^2(x) - \sigma^2(x)| = O_P \left(c_n^{-\frac{1}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} q_n^2 (q_n^f)^2 + c_n^2 q_n^2 (q_n^f)^2 \right).$$

Für $\hat{\sigma}$ gilt

$$\sup_{x \in I_n} |\hat{\sigma}(x) - \sigma(x)| = \sup_{x \in I_n} \left| \frac{(\hat{\sigma}(x) - \sigma(x)) (\hat{\sigma}(x) + \sigma(x))}{\hat{\sigma}(x) + \sigma(x)} \right|$$

$$\leq \frac{\sup_{x \in I_n} |\hat{\sigma}^2(x) - \sigma^2(x)|}{\inf_{x \in I_n} |\hat{\sigma}(x) + \sigma(x)|}.$$

Da $\hat{\sigma}(x) = \sqrt{\hat{\sigma}^2(x)} \geq 0$ und $\sigma(x) = \sqrt{\sigma^2(x)} \geq 0$, ist

$$\inf_{x \in I_n} |\hat{\sigma}(x) + \sigma(x)| \geq \inf_{x \in I_n} |\sigma(x)|,$$

und somit

$$\begin{aligned} \sup_{x \in I_n} |\hat{\sigma}(x) - \sigma(x)| &\leq \frac{\sup_{x \in I_n} |\hat{\sigma}^2(x) - \sigma^2(x)|}{\inf_{x \in I_n} |\sigma(x)|} \\ &= O_P \left(c_n^{-\frac{1}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} q_n^2 (q_n^f)^2 q_n^\sigma + c_n^2 q_n^2 (q_n^f)^2 q_n^\sigma \right), \end{aligned} \quad (3.4.6)$$

da nach Voraussetzung gilt $\frac{1}{\inf_{x \in I_n} |\sigma(x)|} = O(\log(n))$.

Damit gilt nun auch die Behauptung (ii) b), denn

$$\begin{aligned} \sup_{x \in I_n} \left| \frac{\hat{\sigma}(x) - \sigma(x)}{\sigma(x)} \right| &= \frac{1}{\inf_{x \in I_n} |\sigma(x)|} \sup_{x \in I_n} |\hat{\sigma}(x) - \sigma(x)| \\ &= O_P \left(c_n^{-\frac{1}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} q_n^2 (q_n^f)^2 (q_n^\sigma)^2 + c_n^2 q_n^2 (q_n^f)^2 (q_n^\sigma)^2 \right). \end{aligned}$$

Wie man durch Differenzieren leicht sieht, ist

$$\begin{aligned} &\sup_{x \in I_n} \left| \frac{\partial^\nu}{\partial x^\nu} \left(\frac{\hat{m}(x) - m(x)}{\sigma(x)} \right) \right| \\ &= O \left((q_n^\sigma)^{\nu+1} \sum_{j=0}^{\nu} O \left((q_n^f)^{j+1} \right) \sum_{l=0}^1 O \left(q_n^{1-l} \right) \cdot O \left(\sup_{x \in I_n} \left| \frac{\partial^j}{\partial x^j} \left(\hat{g}_l(x) - \frac{1}{n} \sum_{i=1}^n g_{l,i}(x) \right) \right| \right) \right) \\ &\stackrel{(3.4.4)}{=} O_P \left(\left(c_n^{-\frac{1}{2}-\nu} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} + c_n^2 \right) q_n (q_n^f q_n^\sigma)^{\nu+1} \right), \end{aligned} \quad (3.4.7)$$

da $\sup_{x \in I_n} |m^{(\mu)}(x)| = O(q_n)$, $\sup_{x \in I_n} |\sigma^{(\mu)}(x)| = O(q_n)$ und $\sup_{x \in I_n} \frac{1}{n} \sum_{i=1}^n |f_{X_{i-1}}^{(\mu)}(x)| = O(1)$ für $\mu = 0, 1, 2$, sowie $\frac{1}{\inf_{x \in I_n} |\sigma(x)|} = O(q_n^\sigma)$ und $\frac{1}{\inf_{x \in I_n} \frac{1}{n} \sum_{i=1}^n f_{X_{i-1}}(x)} = O(q_n^f)$.

Genauso gilt für den Schätzer $\hat{\sigma}$

$$\begin{aligned} &\sup_{x \in I_n} \left| \frac{\partial^\nu}{\partial x^\nu} \left(\frac{\hat{\sigma}(x) - \sigma(x)}{\sigma(x)} \right) \right| \\ &= O \left(((q_n^\sigma)^2)^{\nu+1} \sum_{j=0}^{\nu} O \left(((q_n^f)^2)^{j+1} \right) \sum_{l=0}^2 O \left(q_n^{2-l} \right) \cdot O \left(\sup_{x \in I_n} \left| \frac{\partial^j}{\partial x^j} \left(\hat{g}_l(x) - \frac{1}{n} \sum_{i=1}^n g_{l,i}(x) \right) \right| \right) \right) \\ &= O_P \left(\left(c_n^{-\frac{1}{2}-\nu} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} + c_n^2 \right) q_n^2 (q_n^f q_n^\sigma)^{2\nu+2} \right). \end{aligned}$$

Damit gelten die Behauptungen (iii) a) und (iii) b), denn nach Voraussetzung ist

$$\left(c_n^{-\frac{3}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} + c_n^2 \right) q_n^2 (q_n^f q_n^\sigma)^4 = o(1).$$

Es fehlen noch die Beweise für die Behauptungen (iv) a) und (iv) b). Exemplarisch wird der Beweis von (iv) a) notiert, da der Beweis von (iv) b) ganz analog funktioniert. Durch eine Aufteilung in $|x - y| \leq c_n$ und $|x - y| > c_n$ ergibt sich

$$\begin{aligned}
& \sup_{x,y \in I_n, x \neq y} \frac{\left| \frac{\partial}{\partial x} \left(\frac{\hat{m}(x) - m(x)}{\sigma(x)} \right) - \frac{\partial}{\partial y} \left(\frac{\hat{m}(y) - m(y)}{\sigma(y)} \right) \right|}{|y - x|^\delta} \\
& \leq \sup_{x,y \in I_n, |x-y| > c_n} \frac{\left| \frac{\partial}{\partial x} \left(\frac{\hat{m}(x) - m(x)}{\sigma(x)} \right) - \frac{\partial}{\partial y} \left(\frac{\hat{m}(y) - m(y)}{\sigma(y)} \right) \right|}{|y - x|^\delta} \\
& \quad + \sup_{x,y \in I_n, 0 < |x-y| \leq c_n} \frac{\left| \frac{\partial}{\partial x} \left(\frac{\hat{m}(x) - m(x)}{\sigma(x)} \right) - \frac{\partial}{\partial y} \left(\frac{\hat{m}(y) - m(y)}{\sigma(y)} \right) \right|}{|y - x|^\delta} \\
& \leq 2 \cdot \sup_{x \in I_n} \left| \frac{\partial}{\partial x} \left(\frac{\hat{m}(x) - m(x)}{\sigma(x)} \right) \right| c_n^{-\delta} + \sup_{x \in I_n} \left| \frac{\partial^2}{\partial x^2} \left(\frac{\hat{m}(x) - m(x)}{\sigma(x)} \right) \right| \sup_{x,y \in I_n, 0 < |x-y| \leq c_n} |x - y|^{1-\delta} \\
& \leq 2 \cdot \sup_{x \in I_n} \left| \frac{\partial}{\partial x} \left(\frac{\hat{m}(x) - m(x)}{\sigma(x)} \right) \right| c_n^{-\delta} + \sup_{x \in I_n} \left| \frac{\partial^2}{\partial x^2} \left(\frac{\hat{m}(x) - m(x)}{\sigma(x)} \right) \right| c_n^{(1-\delta)},
\end{aligned}$$

und mit der Konvergenzrate (3.4.7) folgt weiter

$$\begin{aligned}
& \sup_{x,y \in I_n, x \neq y} \frac{\left| \frac{\partial}{\partial x} \left(\frac{\hat{m}(x) - m(x)}{\sigma(x)} \right) - \frac{\partial}{\partial y} \left(\frac{\hat{m}(y) - m(y)}{\sigma(y)} \right) \right|}{|y - x|^\delta} \\
& = O_P \left(\left(c_n^{-\frac{3}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} + c_n^2 \right) q_n (q_n^f q_n^\sigma)^2 \right) c_n^{-\delta} \\
& \quad + O_P \left(\left(c_n^{-\frac{5}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} + c_n^2 \right) q_n (q_n^f q_n^\sigma)^3 \right) c_n^{(1-\delta)} \\
& = O_P \left(\left(c_n^{-\frac{3}{2}-\delta} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} + c_n^{2-\delta} \right) q_n (q_n^f q_n^\sigma)^2 \right) \\
& \quad + O_P \left(\left(c_n^{-\frac{3}{2}-\delta} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} + c_n^{3-\delta} \right) q_n (q_n^f q_n^\sigma)^3 \right) \\
& = O_P \left(\left(c_n^{-\frac{3}{2}-\delta} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} + c_n^{2-\delta} \right) q_n (q_n^f q_n^\sigma)^3 \right) \\
& = o_P(1)
\end{aligned}$$

nach Voraussetzung.

□

3.3 Verwandte Ergebnisse

In der Literatur gibt es einige vergleichbare Ergebnisse zu Lemma 3.4 (S. 35) für unterschiedliche Modelle.

In Akritas and van Keilegom [2001] wird ein heteroskedastisches Modell mit unabhängigen Beobachtungen betrachtet. Für dieses Modell werden die gleichmäßigen Konvergenzraten der Nadaraya-Watson-Schätzer für die Regressionsfunktion und die Skalierungsfunktion hergeleitet.

Diese Ergebnisse können hier auf Grund der Annahme unabhängiger Beobachtungen nicht verwendet werden.

In Hansen [2008] werden für abhängige Beobachtungen gleichmäßige Konvergenzraten für Kernschätzer bewiesen. Unter anderem wird die gleichmäßige Konvergenzrate des Nadaraya-Watson-Schätzers für die Regressionsfunktion hergeleitet.

Diese Ergebnisse können hier nicht verwendet werden, da strenge Stationarität der Beobachtungen vorausgesetzt wird. Strenge Stationarität kann hier nicht gefordert werden, da sich nach einem möglichen Change-Point auch die Verteilung der Beobachtungen ändert, und Lemma 3.4 auch für den Beweis der Konsistenz des Change-Point-Tests benötigt wird.

In Dette et al. [2009] wird ein heteroskedastisches Modell mit abhängigen Beobachtungen betrachtet. Für dieses Modell werden die Ergebnisse aus Hansen [2008] verwendet um die gleichmäßigen Konvergenzraten der Nadaraya-Watson-Schätzer für die Regressionsfunktion und die Skalierungsfunktion herzuleiten.

Auch in diesem Artikel wird strenge Stationarität gefordert.

Der Beweis in Hansen [2008] basiert, wie der Beweis von Lemma 3.3 (S. 21), auf Theorem 2.1 in Liebscher [1996]. Die Voraussetzungen in Hansen [2008] und Lemma 3.3 unterscheiden sich hauptsächlich in den Momentebedingungen.

Die Annahmen an die bedingten Momente in Hansen [2008] entsprechen den Bedingungen an die Regressions- und Skalierungsfunktion, die Innovationen und die Dichten der Beobachtungen in Annahme (Z'), die durch die autoregressive Form des Modells entstehen, die in Hansen [2008] nicht vorausgesetzt wird.

Der Unterschied in den Momentebedingungen besteht zum einen darin, dass in Lemma 3.3 die Momentebedingungen nicht nur für X , sondern auch für X^2 gefordert werden. Das liegt daran, dass Lemma 3.3 auch für die Herleitung der Konvergenzrate für den Schätzer der Skalierungsfunktion benötigt wird (diese wird in Hansen [2008] nicht hergeleitet).

Der zweite Unterschied besteht darin, dass in Lemma 3.3 jeweils der Mittelwert über alle X_j betrachtet wird, während die Bedingungen in Hansen [2008] auf X_0 basieren. Das liegt daran, dass in Hansen [2008] strenge Stationarität gefordert wird, in Lemma 3.3 nicht.

4 Der sequentielle empirische Prozess der Residuen

In diesem Kapitel wird der sequentielle empirische Prozess der Residuen definiert (Abschnitt 4.1), eine stochastische Entwicklung für ihn hergeleitet, die aus einem sequentiellen empirischen Prozess unabhängiger Zufallsvariablen besteht (Abschnitt 4.2, S. 46), und mit Hilfe dieser stochastischen Entwicklung schwache Konvergenz des sequentiellen empirischen Prozesses der Residuen bewiesen (Abschnitt 4.3).

Eine solche Darstellung für einen sequentiellen empirischen Residuenprozess in einem nichtparametrischen Autoregressionsmodell gibt es, nach bestem Wissen der Autorin, bisher in der Literatur noch nicht.

4.1 Definition des sequentiellen empirischen Prozesses der Residuen

Die Innovationen $\epsilon_1, \dots, \epsilon_n$ sind nicht beobachtbar. Sollen also Aussagen über die Eigenschaften der Innovationen abgeleitet werden, so müssen zunächst die Innovationen selbst geschätzt werden. Dies geschieht mit Hilfe der in Abschnitt 3.1 definierten Kernschätzer \hat{m} und $\hat{\sigma}$ durch die sogenannten Residuen

$$\hat{\epsilon}_j := \frac{X_j - \hat{m}(X_{j-1})}{\hat{\sigma}(X_{j-1})}.$$

Damit sind die Residuen $\hat{\epsilon}_1, \dots, \hat{\epsilon}_n$ also gerade so definiert, dass $X_j = \hat{m}(X_{j-1}) + \hat{\sigma}(X_{j-1})\hat{\epsilon}_j$ für alle $j = 1, \dots, n$.

Wie in Abschnitt 3.2 beschrieben, liefern die Kernschätzer $\hat{m}(x)$ und $\hat{\sigma}(x)$ für große Werte von x normalerweise eher schlechte Ergebnisse. Daher ist es sinnvoll für weitere Schätzungen nur die Residuen $\hat{\epsilon}_j$ zu verwenden, deren X_{j-1} in einem kompakten Intervall $I_n = [a_n, b_n]$ liegen, für das gilt $a_n \xrightarrow{n \rightarrow \infty} -\infty$ und $b_n \xrightarrow{n \rightarrow \infty} \infty$.

Dies wird mit Hilfe einer stetigen Gewichtsfunktion $w_n : \mathbb{R} \rightarrow [0, 1]$ realisiert, für die gilt

$$w_n(x) := \begin{cases} 1 & x \in [a_n + \kappa, b_n - \kappa] \\ 0 & x \notin [a_n, b_n] \end{cases},$$

für ein festes $\kappa > 0$, das unabhängig von n ist. In $(a_n, a_n + \kappa)$ und in $(b_n - \kappa, b_n)$ wird glatt interpoliert (siehe Annahme (W) in Abschnitt 4.2), wobei auch diese Interpolation nur in Form der Verschiebung auf der x -Achse von n abhängig ist. Um die Notation zu vereinfachen ist

$$w_{nj} := w_n(X_{j-1})$$

und

$$\bar{w}_{jN} := \frac{w_n(X_{j-1})}{\sum_{k=1}^N w_n(X_{k-1})}$$

für $N \leq n$.

Der empirische Prozess der Residuen ist damit

$$\hat{F}_n(t) := \sum_{j=1}^n \bar{w}_{jn} I\{\hat{\epsilon}_j \leq t\}.$$

Er ist ein Schätzer für die Verteilungsfunktion der Residuen und somit für die Verteilungsfunktion der Innovationen.

Der sequentielle empirische Prozess der Residuen ist analog zum empirischen Prozess der Residuen definiert, basiert aber nur auf den ersten $\lfloor ns \rfloor$ Residuen:

$$\hat{F}_{\lfloor ns \rfloor}(t) := \sum_{j=1}^{\lfloor ns \rfloor} \bar{w}_{j\lfloor ns \rfloor} I\{\hat{\epsilon}_j \leq t\}$$

mit $s \leq 1$. Er ist ein Schätzer für die Verteilungsfunktion der ersten $\lfloor ns \rfloor$ Residuen und damit der ersten $\lfloor ns \rfloor$ Innovationen, und wird zum Beispiel für die Aufstellung eines Change-Point-Tests benötigt (siehe Kapitel 5, S. 70).

Bemerkung: Die Verwendung von derartigen Gewichtsfunktionen für empirische Prozesse von nicht-parametrisch geschätzten Residuen in Autoregressionsmodellen ist für nicht sequentielle empirische Prozesse z.B. aus Müller et al. [2009] und Dette et al. [2009] bekannt (siehe Kapitel 8, S. 114).

4.2 Stochastische Entwicklung für den sequentiellen empirischen Prozess der Residuen

Die Differenz $\hat{F}_{\lfloor ns \rfloor} - F$ gibt Auskunft über die Leistung von \hat{F} als Schätzer, wenn die ersten $\lfloor ns \rfloor$ Innovationen die Verteilungsfunktion F haben. Um eine Aussage gleichmäßig über alle $s \in [0, 1]$ zu erhalten, muss diese Differenz noch mit einem Vorfaktor multipliziert werden, um die Division durch Null zu vermeiden

$$\frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \left(\hat{F}_{\lfloor ns \rfloor}(t) - F(t) \right).$$

Für diesen Term gibt Satz 4.2 eine stochastische Entwicklung an.

Für die Gültigkeit von Satz 4.2 müssen die folgenden Bedingungen erfüllt sein.

Annahme (K) Der Kern K habe den kompakten Träger $[-C, C]$. Er sei dreimal differenzierbar und $\sup_{u \in [-C, C]} |K^{(\mu)}(u)| \leq \bar{K}$, $\mu = 0, 1, 2, 3$. Außerdem gelte $K(C) = K(-C) = K'(C) = K'(-C) = 0$ und $\int K(u)u du = 0$.

Annahme (I) Für das Intervall $I_n = [a_n, b_n]$, mit $a_n \xrightarrow{n \rightarrow \infty} -\infty$ und $b_n \xrightarrow{n \rightarrow \infty} \infty$, gebe es ein $r_I < \infty$, so dass gelte $(b_n - a_n) = O(\log(n)^{r_I})$. Außerdem sei $\left(\int_{-\infty}^{a_n + \kappa} f_{X_0}(x) dx + \int_{b_n - \kappa}^{\infty} f_{X_0}(x) dx \right) = o(\log(n)^{-1})$, wobei $\kappa > 0$ in Abschnitt 4.1 (S. 45) definiert ist.

Annahme (W) Die Gewichtsfunktion $w_n : \mathbb{R} \rightarrow [0, 1]$ sei dreimal differenzierbar und es gelte $\sup_{x \in \mathbb{R}} |w_n^{(\mu)}(x)| < \infty$ für $\mu = 1, 2, 3$ und alle $n \in \mathbb{N}$ (der Wert $\sup_{x \in \mathbb{R}} |w_n^{(\mu)}(x)|$ ist durch die Form der Funktion w_n unabhängig von n , siehe Abschnitt 4.1).

Annahme (F) Die Innovationen $\epsilon_1, \dots, \epsilon_n$ seien identisch verteilt mit endlichem $2b$ -tem Moment, für ein $b > \frac{8}{3}$. Ihre Dichte werde mit f bezeichnet. Sie sei stetig differenzierbar und es gelte $\sup_{t \in \mathbb{R}} |f(t)t| < \infty$ und $\sup_{t \in \mathbb{R}} |f'(t)t^2| < \infty$.

Annahme: Durch die Stetigkeit der Dichte f und der Ableitung f' gilt damit auch $\sup_{t \in \mathbb{R}} |f(t)| < \infty$, $\sup_{t \in \mathbb{R}} |f'(t)| < \infty$ und $\sup_{t \in \mathbb{R}} |f'(t)t| < \infty$.

Annahme (X) Die Beobachtungen X_0, \dots, X_n seien identisch verteilt und der Prozess $(X_j)_{j \in \mathbb{Z}}$ α -mischend mit einem exponentiell fallendem Mischungskoeffizienten, $\alpha(n) = O(\bar{\alpha}^{-n})$ für ein $\bar{\alpha}$ mit $1 < \bar{\alpha} < \infty$. (*Definition α -mischend siehe Abschnitt 2.2*)

Die Dichte f_{X_0} der Beobachtungen sei beschränkt und viermal differenzierbar und auch die Ableitungen seien beschränkt, $\sup_{x \in \mathbb{R}} |f_{X_0}^{(\mu)}(x)| < \infty$, $\mu = 0, 1, 2, 3, 4$.

Auch gegen Null sei die Dichte auf kompakten Intervallen beschränkt, und es gebe ein $r_f < \infty$, so dass $\frac{1}{\inf_{x \in I_n} f_{X_0}(x)} =: q_n^f = O(\log(n)^{r_f})$.

Annahme (E) Das $2b$ -te Moment der X_0, \dots, X_n sei endlich, $E[|X_0|^{2b}] < \infty$, für ein $b > \frac{8}{3}$ (nicht notwendiger Weise das selbe b wie in Annahme (F)).

Annahme (Z) Es gelte $\sup_{x \in J_n} (|m(x)| + |\sigma(x)|)^{2k} f_{X_0}(x) = O(1)$, und es gebe ein $1 \leq j^* < \infty$, so dass

$\sup_{x, x' \in J_n} (|m(x)| + |\sigma(x)|)^k (|m(x')| + |\sigma(x')|)^k f_{X_0, X_{j-1}}(x, x') = O(1)$ für alle $j > j^* + 1$, für $k = 1, 2$, $n \rightarrow \infty$ mit $J_n := [a_n - (C + c_n^{-\frac{1}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}})c_n, b_n + (C + c_n^{-\frac{1}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}})c_n]$.

Anmerkung: Nach Annahme (C) gilt $c_n \rightarrow 0$ und $c_n^{-\frac{1}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} \rightarrow 0$ für $n \rightarrow \infty$. Daher gibt es eine positive Konstante $\bar{C} < \infty$, so dass $(C + c_n^{-\frac{1}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}})c_n \leq \bar{C}$ für alle n . Damit gelten die Konvergenzraten in Annahme (Z), wenn die entsprechenden Konvergenzraten für $\sup_{x \in [a_n - \bar{C}, b_n + \bar{C}]}$ anstatt $\sup_{x \in J_n}$ gelten.

Anmerkung: Aus Annahme (F) und Annahme (X) folgt strikte Stationarität des Prozesses $(X_j)_{j \in \mathbb{Z}}$ (siehe Bemerkung 2.1, S. 11). Damit impliziert Annahme (Z) die in Annahme (Z') (Abschnitt 3.2, S. 19) geforderten Konvergenzraten.

Annahme (M) Die Regressionsfunktion m und die Skalierungsfunktion σ seien viermal differenzierbar und es gebe $r_q, r_s < \infty$ und q_n, q_n^σ mit $(q_n)^{-1} = O(1)$, $(q_n^\sigma)^{-1} = O(1)$, so dass gelte $\sup_{x \in [a_n - Cc_n, b_n + Cc_n]} |m^{(\mu)}(x)| = O(q_n)$, $\sup_{x \in [a_n - Cc_n, b_n + Cc_n]} |\sigma^{(\mu)}(x)| = O(q_n)$, $\mu = 0, 1, 2, 3, 4$ sowie $\frac{1}{\inf_{x \in I_n} |\sigma(x)|} = O(q_n^\sigma)$ mit $q_n = O(\log(n)^{r_q})$, $q_n^\sigma = O(\log(n)^{r_s})$.

Anmerkungen: Da $I_n \subset [a_n - Cc_n, b_n + Cc_n]$ gilt damit auch $\sup_{x \in I_n} |m^{(\mu)}(x)| = O(q_n)$ und $\sup_{x \in I_n} |\sigma^{(\mu)}(x)| = O(q_n)$, $\mu = 0, 1, 2, 3, 4$.

Auch hier gelten die Konvergenzraten, wenn sie für $\sup_{x \in [a_n - \bar{C}, b_n + \bar{C}]}$ (siehe Anmerkung zu Annahme (Z)) anstatt $\sup_{x \in [a_n - Cc_n, b_n + Cc_n]}$ gelten ($[a_n - \bar{C}, b_n + \bar{C}] \subset J_n$).

Anmerkung: Annahme (Z) und Annahme (M) betreffen beide das Wachstumsverhalten der Regressions- und der Skalierungsfunktion. Annahme (Z) beschränkt das Wachstum der Funktionen, durch das Wachstum des Kehrwerts der Dichte $f_{X_0}^{-1}$ und Annahme (M) beschränkt das Wachstum der Funktionen durch $q_n = O(\log(n)^{r_q})$. Aus Annahme (Z) und Annahme (X) folgt, dass $\sup_{x \in J_n} (|m(x)| + |\sigma(x)|)^{2k} = O(q_n^f)$, und daraus der erste Teil von Annahme (M) (für $\mu = 0$) mit $q_n = (q_n^f)^{\frac{1}{2k}}$.

Annahme (C) Für die Bandbreite c_n gebe es ein $\eta_c > 0$, so dass gelte $c_n \sim n^{-\frac{1}{4}} \log(n)^{-2r_q - 2r_s - \frac{1}{2}r_f - \eta_c}$.

Anmerkung: Die Annahme (C) ist damit strenger als die Annahme (C') in Abschnitt 3.2 (S. 19). Es wird zusätzlich gefordert, dass die Bandbreite schneller als $n^{-\frac{1}{4}} \log(n)^{-2r_q - 2r_s - \frac{1}{2}r_f}$ gegen Null geht, damit $c_n^2 (q_n q_n^\sigma)^4 q_n^f = O\left(n^{-\frac{1}{2}} \log(n)^{-2\eta_c}\right) = o\left(n^{-\frac{1}{2}}\right)$ gilt, was im Beweis von Satz 4.2 benötigt wird.

Beispiel 4.1 Ein einfaches Beispiel, das diese Annahmen erfüllt, ist, analog zu Beispiel 3.1 (S. 20) das AR(1)-Modell

$$X_j = 0.5 \cdot X_{j-1} + \epsilon_j$$

mit Standardnormalverteilten Innovationen (ϵ_j). Damit ist Annahme (F) erfüllt.

X_j hat in diesem Fall für alle j die Verteilung $\mathcal{N}\left(0, \frac{1}{1-0.5^2}\right)$ (da $X_j = \sum_{i=0}^{\infty} 0.5^i \epsilon_{j-i}$, siehe z.B. Beispiel 2.7 in Kreiß and Neuhaus [2006]).

Damit erfüllt (X_j) die ersten beiden Punkte in Annahme (X) nach Abschnitt 2.2 (siehe (2.4.1), S. 13). Der dritte Punkt in Annahme (X) bedeutet, dass

$$\max\left(\exp\left(\frac{a_n^2}{\frac{8}{3}}\right), \exp\left(\frac{b_n^2}{\frac{8}{3}}\right)\right) = O(\log(n)^{r_f}). \quad (4.1.1)$$

Eine weitere Forderung an a_n und b_n ergibt sich aus dem zweiten Punkt in Annahme (I). Er bedeutet, dass

$$\left(\int_{-\infty}^{a_n + \kappa} \exp\left(-\frac{x^2}{\frac{8}{3}}\right) dx + \int_{b_n - \kappa}^{\infty} \exp\left(-\frac{x^2}{\frac{8}{3}}\right) dx\right) = o(\log(n)^{-1}) \quad (4.1.2)$$

Die Forderungen (4.1.1) und (4.1.2) sind erfüllt für $I_n = \left[-\sqrt{\frac{8}{3}} \log(\log(n)^2) - \kappa, \sqrt{\frac{8}{3}} \log(\log(n)^2) + \kappa\right]$, wie im Folgenden gezeigt wird:

Für Forderung (4.1.1) gilt

$$\begin{aligned} \exp\left(\frac{a_n^2}{\frac{8}{3}}\right) &= \exp\left(\frac{b_n^2}{\frac{8}{3}}\right) \\ &= \exp\left(\frac{\left(\sqrt{\frac{8}{3}} \log(\log(n)^2) + \kappa\right)^2}{\frac{8}{3}}\right) \\ &= \exp\left(\frac{\left(\sqrt{\frac{8}{3}} \log(\log(n)^2)\right)^2}{\frac{8}{3}}\right) \underbrace{\exp\left(\frac{2\sqrt{\frac{8}{3}} \log(\log(n)^2) \kappa}{\frac{8}{3}}\right)}_{=O\left(\exp\left(\frac{\left(\sqrt{\frac{8}{3}} \log(\log(n)^2)\right)^2}{\frac{8}{3}}\right)\right)} \underbrace{\exp\left(\frac{\kappa^2}{\frac{8}{3}}\right)}_{< \infty} \\ &= O\left(\left(\exp(\log(\log(n)^2))\right)^2\right) = O(\log(n)^4). \end{aligned}$$

Für Forderung (4.1.2) gilt

$$\begin{aligned} \int_{b_n - \kappa}^{\infty} \exp\left(-\frac{x^2}{\frac{8}{3}}\right) dx &= \int_{-\infty}^{a_n + \kappa} \exp\left(-\frac{x^2}{\frac{8}{3}}\right) dx \\ &\leq \int_{-\infty}^{a_n + \kappa} \left(-\frac{6}{8}x\right) \cdot \exp\left(-\frac{x^2}{\frac{8}{3}}\right) dx \quad \text{für hinreichend großes } n \end{aligned}$$

$$\begin{aligned}
&= \exp\left(-\frac{(a_n + \kappa)^2}{\frac{8}{3}}\right) \\
&= O\left(\exp(-\log(\log(n)^2))\right) \\
&= O(\log(n)^{-2}) = o(\log(n)^{-1}).
\end{aligned}$$

Mit diesem I_n sind auch der erste Punkt in Annahme (I) und Annahme (M) erfüllt, wobei $q_n = O\left(\sqrt{\log(\log(n)^2)} + c_n\right) = O(\log(n))$ und $q_n^\sigma = 1$.

Die Annahme (E) gilt und auch die Annahme (Z), da

$$\begin{aligned}
&\sup_{x \in J_n} \left((|m(x)| + |\sigma(x)|)^{2k} f_{X_0}(x) \right) \\
&\leq \sup_{x \in \mathbb{R}} \left((|m(x)| + |\sigma(x)|)^{2k} f_{X_0}(x) \right) \\
&= \sup_{x \in \mathbb{R}} \left((0.5 \cdot |x| + 1)^{2k} \frac{1}{\sqrt{2\pi \frac{4}{3}}} \exp\left(-\frac{x^2}{\frac{8}{3}}\right) \right) \\
&< \infty \quad \text{da } \sup_{x \in \mathbb{R}} |x|^{2k} \exp(-x^2) < \infty
\end{aligned}$$

und analog auch die zweite Bedingung folgt, indem man die gemeinsame Dichte $f_{X_0, X_{j-1}}$ wie im Beweis zu Bemerkung 2.1 (S. 11) schreibt:

$$\begin{aligned}
&f_{X_0, X_{j-1}}(x, x') \\
&= \frac{1}{\sqrt{2\pi \frac{4}{3}}} \exp\left(-\frac{x^2}{\frac{8}{3}}\right) \\
&\int \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(y_1 - 0.5x)^2}{2}\right) \\
&\int \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(y_2 - 0.5y_1)^2}{2}\right) \\
&\vdots \\
&\int \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(y_{j-2} - 0.5y_{j-3})^2}{2}\right) \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(x' - 0.5y_{j-2})^2}{2}\right) dy_1 \dots dy_{j-2}.
\end{aligned}$$

Mit einem Kern und einer Gewichtsfunktion, die die Bedingungen (K) und (W) erfüllen, und einer Bandbreite c_n mit $c_n \sim n^{-\frac{1}{4}} \log(n)^{-2-\frac{1}{2} \cdot 4 - n^c}$ sind alle Bedingungen für Satz 4.2 erfüllt.

Satz 4.2 Unter den Voraussetzungen (K), (I), (W), (F), (X), (E), (Z), (M) und (C) gilt die Entwicklung

$$\begin{aligned}
&\frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \left(\sum_{j=1}^{\lfloor ns \rfloor} \bar{w}_{j \lfloor ns \rfloor} I\{\hat{\epsilon}_j \leq t\} - F(t) \right) \\
&= \frac{1}{n} \sum_{j=1}^{\lfloor ns \rfloor} (I\{\epsilon_j \leq t\} - F(t)) + \frac{\lfloor ns \rfloor}{n} f(t) \frac{1}{n} \sum_{j=1}^n \epsilon_j + \frac{\lfloor ns \rfloor}{n} f(t) t \frac{1}{2n} \sum_{j=1}^n (\epsilon_j^2 - 1) + o_p\left(\frac{1}{\sqrt{n}}\right)
\end{aligned}$$

gleichmäßig in $t \in \mathbb{R}$ und $s \in [0, 1]$.

Bemerkung 4.3 Mit Satz 4.2 folgt, dass $\sqrt{n} \left(\frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \left(\hat{F}_{\lfloor ns \rfloor} - F \right) \right)$, und insbesondere, für $s = 1$, $\sqrt{n} \left(\hat{F}_n - F \right)$ schwach gegen einen zentrierten Gaußprozess konvergiert, siehe Abschnitt 4.3 (S. 60).

Beweis von Satz 4.2 Mit Lemma 9.1 (S.117) und Lemma 9.3 (S. 133) gilt gleichmäßig in $t \in \mathbb{R}$ und $s \in [0, 1]$

$$\begin{aligned}
& \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \left(\hat{F}_{\lfloor ns \rfloor}(t) - F(t) \right) \\
= & \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \left(\sum_{j=1}^{\lfloor ns \rfloor} \bar{w}_{j\lfloor ns \rfloor} I\{\hat{\epsilon}_j \leq t\} - F(t) \right) \\
= & \frac{1}{n} \sum_{j=1}^{\lfloor ns \rfloor} (I\{\epsilon_j \leq t\} - F(t)) \\
& + \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \left(\sum_{j=1}^{\lfloor ns \rfloor} \bar{w}_{j\lfloor ns \rfloor} I\{\epsilon_j \leq t\} - F(t) \right) - \frac{1}{n} \sum_{j=1}^{\lfloor ns \rfloor} (I\{\epsilon_j \leq t\} - F(t)) \\
& + \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \sum_{j=1}^{\lfloor ns \rfloor} \bar{w}_{j\lfloor ns \rfloor} \left(F \left(t \frac{\hat{\sigma}}{\sigma}(X_{j-1}) + \frac{\hat{m} - m}{\sigma}(X_{j-1}) \right) - F(t) \right) \\
& + \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \sum_{j=1}^{\lfloor ns \rfloor} \bar{w}_{j\lfloor ns \rfloor} \left(I\{\hat{\epsilon}_j \leq t\} - F \left(t \frac{\hat{\sigma}}{\sigma}(X_{j-1}) + \frac{\hat{m} - m}{\sigma}(X_{j-1}) \right) - I\{\epsilon_j \leq t\} + F(t) \right) \\
= & \frac{1}{n} \sum_{j=1}^{\lfloor ns \rfloor} (I\{\epsilon_j \leq t\} - F(t)) \\
& + o_p \left(\frac{1}{\sqrt{n}} \right) \text{ nach Lemma 9.3 auf S. 133} \\
& + \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \sum_{j=1}^{\lfloor ns \rfloor} \bar{w}_{j\lfloor ns \rfloor} \left(F \left(t \frac{\hat{\sigma}}{\sigma}(X_{j-1}) + \frac{\hat{m} - m}{\sigma}(X_{j-1}) \right) - F(t) \right) \\
& + o_p \left(\frac{1}{\sqrt{n}} \right) \text{ nach Lemma 9.1 auf S. 117.}
\end{aligned}$$

Es bleibt also noch zu zeigen, dass

$$\begin{aligned}
& \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \sum_{j=1}^{\lfloor ns \rfloor} \bar{w}_{j\lfloor ns \rfloor} \left(F \left(t \frac{\hat{\sigma}}{\sigma}(X_{j-1}) + \frac{\hat{m} - m}{\sigma}(X_{j-1}) \right) - F(t) \right) \\
& = \frac{\lfloor ns \rfloor}{n} f(t) \frac{1}{n} \sum_{j=1}^n \epsilon_j + \frac{\lfloor ns \rfloor}{n} f(t) t \cdot \frac{1}{2n} \sum_{j=1}^n (\epsilon_j^2 - 1) + o_p \left(\frac{1}{\sqrt{n}} \right)
\end{aligned}$$

gleichmäßig in $t \in \mathbb{R}$ und $s \in [0, 1]$.

Hierfür wird zunächst für $F \left(t \frac{\hat{\sigma}}{\sigma}(X_{j-1}) + \frac{\hat{m} - m}{\sigma}(X_{j-1}) \right)$ in $F(t)$ eine Taylor-Entwicklung durchgeführt. Mit einem $\xi(t, j)$ zwischen t und $t \frac{\hat{\sigma}}{\sigma}(X_{j-1}) + \frac{\hat{m} - m}{\sigma}(X_{j-1})$ gilt

$$F \left(t \frac{\hat{\sigma}}{\sigma}(X_{j-1}) + \frac{\hat{m} - m}{\sigma}(X_{j-1}) \right) - F(t) = f(t) \left(t \frac{\hat{\sigma}}{\sigma}(X_{j-1}) + \frac{\hat{m} - m}{\sigma}(X_{j-1}) - t \right)$$

$$\begin{aligned}
& + \frac{1}{2} f'(\xi(t, j)) \left(t \frac{\hat{\sigma}}{\sigma}(X_{j-1}) + \frac{\hat{m} - m}{\sigma}(X_{j-1}) - t \right)^2 \\
& = f(t) \left(t \frac{\hat{\sigma}}{\sigma}(X_{j-1}) + \frac{\hat{m} - m}{\sigma}(X_{j-1}) - t \right) \\
& \quad + \frac{1}{2} f'(\xi(t, j)) t^2 \left(\frac{\hat{\sigma}}{\sigma}(X_{j-1}) - 1 \right)^2 \\
& \quad + f'(\xi(t, j)) t \left(\frac{\hat{\sigma}}{\sigma}(X_{j-1}) - 1 \right) \left(\frac{\hat{m} - m}{\sigma}(X_{j-1}) \right) \\
& \quad + \frac{1}{2} f'(\xi(t, j)) \left(\frac{\hat{m} - m}{\sigma}(X_{j-1}) \right)^2.
\end{aligned} \tag{4.3.1}$$

Da nach Lemma 3.4 gilt

$$\sup_{x \in I_n} \left| \frac{\hat{m} - m}{\sigma}(x) \right| = O_P \left(\left(c_n^{-\frac{1}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} + c_n^2 \right) q_n q_n^f q_n^\sigma \right)$$

und

$$\sup_{x \in I_n} \left| \frac{\hat{\sigma}}{\sigma}(x) - 1 \right| = O_P \left(\left(c_n^{-\frac{1}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} + c_n^2 \right) \left(q_n q_n^f q_n^\sigma \right)^2 \right),$$

ist

$$\begin{aligned}
& \sup_{t \in \mathbb{R}, s \in [0,1]} \left| \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \sum_{j=1}^{\lfloor ns \rfloor} \bar{w}_{j \lfloor ns \rfloor} \left(\frac{1}{2} f'(\xi(t, j)) t^2 \left(\frac{\hat{\sigma}}{\sigma}(X_{j-1}) - 1 \right)^2 \right. \right. \\
& \quad \left. \left. + f'(\xi(t, j)) t \left(\frac{\hat{\sigma}}{\sigma}(X_{j-1}) - 1 \right) \left(\frac{\hat{m} - m}{\sigma}(X_{j-1}) \right) \right. \right. \\
& \quad \left. \left. + \frac{1}{2} f'(\xi(t, j)) \left(\frac{\hat{m} - m}{\sigma}(X_{j-1}) \right)^2 \right) \right| \\
& = O \left(\sup_{t \in \mathbb{R}, s \in [0,1]} \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \sum_{j=1}^{\lfloor ns \rfloor} \bar{w}_{j \lfloor ns \rfloor} (|f'(\xi(t, j)) t^2| + |f'(\xi(t, j)) t| + |f'(\xi(t, j))|) \right) \\
& \quad \cdot O_P \left(\left(c_n^{-\frac{1}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} + c_n^2 \right)^2 \left(\left(q_n q_n^f q_n^\sigma \right)^2 + \left(q_n q_n^f q_n^\sigma \right)^3 + \left(q_n q_n^f q_n^\sigma \right)^4 \right) \right) \\
& = O \left(\sup_{t \in \mathbb{R}, s \in [0,1]} \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \sum_{j=1}^{\lfloor ns \rfloor} \bar{w}_{j \lfloor ns \rfloor} (|f'(\xi(t, j)) t^2| + |f'(\xi(t, j)) t| + |f'(\xi(t, j))|) \right) \cdot o_p \left(\frac{1}{\sqrt{n}} \right)
\end{aligned} \tag{4.3.2}$$

für alle c_n mit $c_n \sim n^{-\frac{1}{4}} \log(n)^r$, für ein $|r| < \infty$, also auch für ein c_n wie in Annahme (C).

Da nach Voraussetzung gilt, dass $\sup_{t \in \mathbb{R}} |f'(t)| < \infty$, $\sup_{t \in \mathbb{R}} |f'(t)t| < \infty$ und $\sup_{t \in \mathbb{R}} |f'(t)t^2| < \infty$, ist

$$\begin{aligned}
& \sup_{t \in \mathbb{R}, s \in [0,1]} \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \sum_{j=1}^{\lfloor ns \rfloor} \bar{w}_{j \lfloor ns \rfloor} (|f'(\xi(t, j)) t^2| + |f'(\xi(t, j)) t| + |f'(\xi(t, j))|) \\
& = \sup_{t \in \mathbb{R}, s \in [0,1]} \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \sum_{j=1}^{\lfloor ns \rfloor} \bar{w}_{j \lfloor ns \rfloor} \left(\left| f'(\xi(t, j)) \xi(t, j)^2 \left(\frac{t}{\xi(t, j)} \right)^2 \right| \right)
\end{aligned}$$

$$\begin{aligned}
& + \left| f'(\xi(t, j)) \xi(t, j) \frac{t}{\xi(t, j)} \right| + |f'(\xi(t, j))| \\
& = O \left(\sup_{t \in \mathbb{R}, s \in [0, 1]} \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \sum_{j=1}^{\lfloor ns \rfloor} \bar{w}_{j \lfloor ns \rfloor} \left(\left| \left(\frac{t}{\xi(t, j)} \right)^2 \right| + \left| \frac{t}{\xi(t, j)} \right| + 1 \right) \right) \\
& = O_P(1),
\end{aligned}$$

da $\xi(t, j)$ zwischen t und $t \frac{\hat{\sigma}}{\sigma}(X_{j-1}) + \frac{\hat{m}-m}{\sigma}(X_{j-1})$ liegt, und $t \frac{\hat{\sigma}}{\sigma}(x) + \frac{\hat{m}-m}{\sigma}(x) \xrightarrow[n \rightarrow \infty]{P} t$ für alle $x \in I_n$, sowie $\sup_{s \in [0, 1]} \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \sum_{j=1}^{\lfloor ns \rfloor} \bar{w}_{j \lfloor ns \rfloor} \leq 1$ ist.

Damit ist (4.3.2) = $o_p\left(\frac{1}{\sqrt{n}}\right)$ und daher mit (4.3.1)

$$\begin{aligned}
& \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \sum_{j=1}^{\lfloor ns \rfloor} \bar{w}_{j \lfloor ns \rfloor} \left(F \left(t \frac{\hat{\sigma}}{\sigma}(X_{j-1}) + \frac{\hat{m}-m}{\sigma}(X_{j-1}) \right) - F(t) \right) \\
& = \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} f(t) \sum_{j=1}^{\lfloor ns \rfloor} \bar{w}_{j \lfloor ns \rfloor} \frac{\hat{m}-m}{\sigma}(X_{j-1}) + \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} f(t) t \sum_{j=1}^{\lfloor ns \rfloor} \bar{w}_{j \lfloor ns \rfloor} \frac{\hat{\sigma}-\sigma}{\sigma}(X_{j-1}) + o_p\left(\frac{1}{\sqrt{n}}\right)
\end{aligned}$$

gleichmäßig in $t \in \mathbb{R}$ und $s \in [0, 1]$.

Nun kann man zeigen, dass

$$\frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \sum_{j=1}^{\lfloor ns \rfloor} \bar{w}_{j \lfloor ns \rfloor} \frac{\hat{m}-m}{\sigma}(X_{j-1}) = \frac{\lfloor ns \rfloor}{n} \frac{1}{n} \sum_{j=1}^n \epsilon_j + o_p\left(\frac{1}{\sqrt{n}}\right), \quad (4.3.3)$$

indem man den Nadaraya-Watson-Schätzer für \hat{m} (siehe 3.1, S. 18) einsetzt.

Es gilt

$$\begin{aligned}
& \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \sum_{j=1}^{\lfloor ns \rfloor} \bar{w}_{j \lfloor ns \rfloor} \frac{\hat{m}-m}{\sigma}(X_{j-1}) \\
& = \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \sum_{j=1}^{\lfloor ns \rfloor} \frac{\bar{w}_{j \lfloor ns \rfloor}}{\sigma(X_{j-1})} \left(\frac{\frac{1}{nc_n} \sum_{i=1}^n K\left(\frac{X_{j-1}-X_{i-1}}{c_n}\right) X_i}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{X_{j-1}-X_{k-1}}{c_n}\right)} - m(X_{j-1}) \right) \\
& = \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \sum_{j=1}^{\lfloor ns \rfloor} \frac{\bar{w}_{j \lfloor ns \rfloor}}{\sigma(X_{j-1})} \left(\frac{\frac{1}{nc_n} \sum_{i=1}^n K\left(\frac{X_{j-1}-X_{i-1}}{c_n}\right) X_i - \frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{X_{j-1}-X_{k-1}}{c_n}\right) m(X_{j-1})}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{X_{j-1}-X_{k-1}}{c_n}\right)} \right) \\
& = \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \sum_{j=1}^{\lfloor ns \rfloor} \frac{\bar{w}_{j \lfloor ns \rfloor}}{\sigma(X_{j-1})} \left(\frac{\frac{1}{nc_n} \sum_{i=1}^n K\left(\frac{X_{j-1}-X_{i-1}}{c_n}\right) (X_i - m(X_{i-1}))}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{X_{j-1}-X_{k-1}}{c_n}\right)} \right) \\
& \quad + \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \sum_{j=1}^{\lfloor ns \rfloor} \frac{\bar{w}_{j \lfloor ns \rfloor}}{\sigma(X_{j-1})} \left(\frac{\frac{1}{nc_n} \sum_{i=1}^n K\left(\frac{X_{j-1}-X_{i-1}}{c_n}\right) (m(X_{i-1}) - m(X_{j-1}))}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{X_{j-1}-X_{k-1}}{c_n}\right)} \right) \\
& = \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \sum_{j=1}^{\lfloor ns \rfloor} \frac{\bar{w}_{j \lfloor ns \rfloor}}{\sigma(X_{j-1})} \left(\frac{\frac{1}{nc_n} \sum_{i=1}^n K\left(\frac{X_{j-1}-X_{i-1}}{c_n}\right) (X_i - m(X_{i-1}))}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{X_{j-1}-X_{k-1}}{c_n}\right)} \right)
\end{aligned}$$

$$+o_p\left(\frac{1}{\sqrt{n}}\right),$$

da die Funktionen $\frac{1}{q_n}m(\cdot)$ und $\frac{1}{(q_n^\sigma)^3(q_n)^2}\sigma(\cdot)$ die Bedingungen an die Funktionen v_n und t_n in Lemma 10.4 (S. 166) erfüllen, und damit

$$\begin{aligned} & \sup_{s \in [0,1]} \left| \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \sum_{j=1}^{\lfloor ns \rfloor} \frac{\bar{w}_{j\lfloor ns \rfloor}}{\sigma(X_{j-1})} \left(\frac{\frac{1}{nc_n} \sum_{i=1}^n K\left(\frac{X_{j-1}-X_{i-1}}{c_n}\right) (m(X_{i-1}) - m(X_{j-1}))}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{X_{j-1}-X_{k-1}}{c_n}\right)} \right) \right| \\ &= O_P\left((q_n)^3(q_n^\sigma)^3 \left(c_n^{\frac{1}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} + c_n^2\right) q_n^f\right) \\ &= o_p\left(\frac{1}{\sqrt{n}}\right) \quad (\text{siehe Annahme (C)}) \end{aligned}$$

gilt.

Weiter ist mit Lemma 9.2 (S. 130), Lemma 10.2 (S. 145) und Lemma 10.3 (S. 158)

$$\begin{aligned} & \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \sum_{j=1}^{\lfloor ns \rfloor} \frac{\bar{w}_{j\lfloor ns \rfloor}}{\sigma(X_{j-1})} \left(\frac{\frac{1}{nc_n} \sum_{i=1}^n K\left(\frac{X_{j-1}-X_{i-1}}{c_n}\right) (X_i - m(X_{i-1}))}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{X_{j-1}-X_{k-1}}{c_n}\right)} \right) \\ &= \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \sum_{j=1}^{\lfloor ns \rfloor} \frac{\bar{w}_{j\lfloor ns \rfloor}}{\sigma(X_{j-1})} \left(\frac{\frac{1}{nc_n} \sum_{i=1}^n K\left(\frac{X_{j-1}-X_{i-1}}{c_n}\right) \epsilon_i \sigma(X_{i-1})}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{X_{j-1}-X_{k-1}}{c_n}\right)} \right) \\ &= \frac{\lfloor ns \rfloor}{n} \frac{1}{n} \sum_{i=1}^n \epsilon_i \\ & \quad + \frac{1}{n} \sum_{i=1}^n \epsilon_i \left(\frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \sum_{j=1}^{\lfloor ns \rfloor} \frac{\bar{w}_{j\lfloor ns \rfloor}}{c_n} K\left(\frac{X_{j-1}-X_{i-1}}{c_n}\right) \frac{\sigma(X_{i-1})}{\sigma(X_{j-1})} - \frac{\lfloor ns \rfloor}{n} \right) \\ &= \frac{\lfloor ns \rfloor}{n} \frac{1}{n} \sum_{i=1}^n \epsilon_i \\ & \quad + \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \frac{1}{n} \sum_{i=1}^n \epsilon_i \sum_{j=1}^{\lfloor ns \rfloor} \frac{\bar{w}_{j\lfloor ns \rfloor}}{\sigma(X_{j-1})} \left(\frac{\frac{1}{c_n} K\left(\frac{X_{j-1}-X_{i-1}}{c_n}\right) (\sigma(X_{i-1}) - \sigma(X_{j-1}))}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{X_{j-1}-X_{k-1}}{c_n}\right)} \right) \\ & \quad + \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \frac{1}{n} \sum_{i=1}^n \epsilon_i \sum_{j=1}^{\lfloor ns \rfloor} \bar{w}_{j\lfloor ns \rfloor} \left(\frac{\frac{1}{c_n} K\left(\frac{X_{j-1}-X_{i-1}}{c_n}\right) - \frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{X_{j-1}-X_{k-1}}{c_n}\right)}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{X_{j-1}-X_{k-1}}{c_n}\right)} \right) \\ & \quad + \frac{1}{n} \sum_{i=1}^n \epsilon_i \left(\frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \sum_{j=1}^{\lfloor ns \rfloor} \bar{w}_{j\lfloor ns \rfloor} - \frac{\lfloor ns \rfloor}{n} \right) \\ &= \frac{\lfloor ns \rfloor}{n} \frac{1}{n} \sum_{i=1}^n \epsilon_i + o_p\left(\frac{1}{\sqrt{n}}\right), \end{aligned}$$

da

$$\sup_{s \in [0,1]} \left| \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \frac{1}{n} \sum_{i=1}^n \epsilon_i \sum_{j=1}^{\lfloor ns \rfloor} \frac{\bar{w}_{j\lfloor ns \rfloor}}{\sigma(X_{j-1})} \left(\frac{\frac{1}{c_n} K\left(\frac{X_{j-1}-X_{i-1}}{c_n}\right) (\sigma(X_{i-1}) - \sigma(X_{j-1}))}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{X_{j-1}-X_{k-1}}{c_n}\right)} \right) \right| = o_p\left(\frac{1}{\sqrt{n}}\right)$$

nach Lemma 10.3 (S. 158), und

$$\sup_{s \in [0,1]} \left| \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \frac{1}{n} \sum_{i=1}^n \epsilon_i \sum_{j=1}^{\lfloor ns \rfloor} \bar{w}_{j \lfloor ns \rfloor} \left(\frac{\frac{1}{c_n} K \left(\frac{X_{j-1} - X_{i-1}}{c_n} \right) - \frac{1}{nc_n} \sum_{k=1}^n K \left(\frac{X_{j-1} - X_{k-1}}{c_n} \right)}{\frac{1}{nc_n} \sum_{k=1}^n K \left(\frac{X_{j-1} - X_{k-1}}{c_n} \right)} \right) \right| = o_p \left(\frac{1}{\sqrt{n}} \right)$$

nach Lemma 10.2 (S. 145), sowie

$$\sup_{s \in [0,1]} \left| \frac{1}{n} \sum_{i=1}^n \epsilon_i \left(\frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \sum_{j=1}^{\lfloor ns \rfloor} \bar{w}_{j \lfloor ns \rfloor} - \frac{\lfloor ns \rfloor}{n} \right) \right| = o_p \left(\frac{1}{\sqrt{n}} \right),$$

da $\frac{1}{n} \sum_{i=1}^n \epsilon_i = O_P \left(\frac{1}{\sqrt{n}} \right)$ (Summe von unabhängigen zentrierten Zufallsvariablen) und

$$\sup_{s \in [0,1]} \left| \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \underbrace{\sum_{j=1}^{\lfloor ns \rfloor} \bar{w}_{j \lfloor ns \rfloor}}_{=1} - \frac{\lfloor ns \rfloor}{n} \right| = o_P(1) \text{ nach Lemma 9.2 (S. 130).}$$

Damit ist gezeigt, dass (4.3.3) gilt.

Um zu zeigen, dass

$$\frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \sum_{j=1}^{\lfloor ns \rfloor} \bar{w}_{j \lfloor ns \rfloor} \frac{\hat{\sigma} - \sigma}{\sigma} (X_{j-1}) = \frac{\lfloor ns \rfloor}{n} \frac{1}{2n} \sum_{j=1}^n (\epsilon_j^2 - 1) + o_p \left(\frac{1}{\sqrt{n}} \right) \quad (4.3.4)$$

muss man zunächst noch etwas umformen. Es gilt

$$\begin{aligned} \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \sum_{j=1}^{\lfloor ns \rfloor} \bar{w}_{j \lfloor ns \rfloor} \frac{\hat{\sigma} - \sigma}{\sigma} (X_{j-1}) &= \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \sum_{j=1}^{\lfloor ns \rfloor} \bar{w}_{j \lfloor ns \rfloor} \left(\frac{\hat{\sigma}^2 - \sigma^2}{2\sigma^2} (X_{j-1}) - \frac{(\hat{\sigma} - \sigma)^2}{2\sigma^2} (X_{j-1}) \right) \\ &= \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \sum_{j=1}^{\lfloor ns \rfloor} \bar{w}_{j \lfloor ns \rfloor} \frac{\hat{\sigma}^2 - \sigma^2}{2\sigma^2} (X_{j-1}) + o_p \left(\frac{1}{\sqrt{n}} \right), \end{aligned}$$

da nach Lemma 3.4 auf S. 35

$$\begin{aligned} \sup_{s \in [0,1]} \left| \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \sum_{j=1}^{\lfloor ns \rfloor} \bar{w}_{j \lfloor ns \rfloor} \frac{(\hat{\sigma} - \sigma)^2}{2\sigma^2} (X_{j-1}) \right| &\leq \left(\sup_{x \in I_n} \left| \frac{\hat{\sigma} - \sigma}{2\sigma} (x) \right| \right)^2 \\ &= O_P \left(\left(\left(c_n^{-\frac{1}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} + c_n^2 \right) (q_n q_n^f q_n^\sigma)^2 \right)^2 \right) \\ &= o_p \left(\frac{1}{\sqrt{n}} \right), \end{aligned}$$

für $c_n \sim n^{-\frac{1}{4}} \log(n)^r$, für alle $|r| < \infty$.

Weiter ergibt sich mit $\hat{\sigma}^2$ wie in 3.1 (S. 18) definiert

$$\frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \frac{1}{2} \sum_{j=1}^{\lfloor ns \rfloor} \bar{w}_{j \lfloor ns \rfloor} \frac{\hat{\sigma}^2 - \sigma^2}{\sigma^2} (X_{j-1})$$

$$\begin{aligned}
&= \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \frac{1}{2} \sum_{j=1}^{\lfloor ns \rfloor} \frac{\bar{w}_j^{\lfloor ns \rfloor}}{\sigma^2(X_{j-1})} \left(\frac{\frac{1}{nc_n} \sum_{i=1}^n K\left(\frac{X_{j-1}-X_{i-1}}{c_n}\right) X_i^2}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{X_{j-1}-X_{k-1}}{c_n}\right)} - \sigma^2(X_{j-1}) - \hat{m}^2(X_{j-1}) \right) \\
&= \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \frac{1}{2} \sum_{j=1}^{\lfloor ns \rfloor} \frac{\bar{w}_j^{\lfloor ns \rfloor}}{\sigma^2(X_{j-1})} \left(\frac{\frac{1}{nc_n} \sum_{i=1}^n K\left(\frac{X_{j-1}-X_{i-1}}{c_n}\right) (m(X_{i-1}) + \sigma(X_{i-1})\epsilon_i)^2}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{X_{j-1}-X_{k-1}}{c_n}\right)} \right) \\
&\quad + \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \frac{1}{2} \sum_{j=1}^{\lfloor ns \rfloor} \frac{\bar{w}_j^{\lfloor ns \rfloor}}{\sigma^2(X_{j-1})} \left(\frac{\frac{1}{nc_n} \sum_{i=1}^n K\left(\frac{X_{j-1}-X_{i-1}}{c_n}\right) (-\sigma^2(X_{j-1}) - \hat{m}^2(X_{j-1}))}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{X_{j-1}-X_{k-1}}{c_n}\right)} \right) \\
&= \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \frac{1}{2} \sum_{j=1}^{\lfloor ns \rfloor} \frac{\bar{w}_j^{\lfloor ns \rfloor}}{\sigma^2(X_{j-1})} \left(\frac{\frac{1}{nc_n} \sum_{i=1}^n K\left(\frac{X_{j-1}-X_{i-1}}{c_n}\right) (m^2(X_{i-1}) + 2m\sigma(X_{i-1})\epsilon_i + \sigma^2(X_{i-1})\epsilon_i^2)}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{X_{j-1}-X_{k-1}}{c_n}\right)} \right) \\
&\quad + \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \frac{1}{2} \sum_{j=1}^{\lfloor ns \rfloor} \frac{\bar{w}_j^{\lfloor ns \rfloor}}{\sigma^2(X_{j-1})} \left(\frac{\frac{1}{nc_n} \sum_{i=1}^n K\left(\frac{X_{j-1}-X_{i-1}}{c_n}\right) (-\sigma^2(X_{j-1}) - \hat{m}^2(X_{j-1}))}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{X_{j-1}-X_{k-1}}{c_n}\right)} \right) \\
&= \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \frac{1}{2} \sum_{j=1}^{\lfloor ns \rfloor} \frac{\bar{w}_j^{\lfloor ns \rfloor}}{\sigma^2(X_{j-1})} \left(\frac{\frac{1}{nc_n} \sum_{i=1}^n K\left(\frac{X_{j-1}-X_{i-1}}{c_n}\right) (\sigma^2(X_{i-1})\epsilon_i^2 - \sigma^2(X_{j-1}))}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{X_{j-1}-X_{k-1}}{c_n}\right)} \right) \\
&\quad + \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \frac{1}{2} \sum_{j=1}^{\lfloor ns \rfloor} \frac{\bar{w}_j^{\lfloor ns \rfloor}}{\sigma^2(X_{j-1})} \left(\frac{\frac{1}{nc_n} \sum_{i=1}^n K\left(\frac{X_{j-1}-X_{i-1}}{c_n}\right) (2m\sigma(X_{i-1})\epsilon_i + m^2(X_{i-1}) - \hat{m}^2(X_{j-1}))}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{X_{j-1}-X_{k-1}}{c_n}\right)} \right).
\end{aligned}$$

Als nächstes wird gezeigt, dass

$$\begin{aligned}
&\frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \frac{1}{2} \sum_{j=1}^{\lfloor ns \rfloor} \frac{\bar{w}_j^{\lfloor ns \rfloor}}{\sigma^2(X_{j-1})} \left(\frac{\frac{1}{nc_n} \sum_{i=1}^n K\left(\frac{X_{j-1}-X_{i-1}}{c_n}\right) (\sigma^2(X_{i-1})\epsilon_i^2 - \sigma^2(X_{j-1}))}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{X_{j-1}-X_{k-1}}{c_n}\right)} \right) \\
&= \frac{\lfloor ns \rfloor}{n} \frac{1}{2n} \sum_{i=1}^n (\epsilon_i^2 - 1) + o_p\left(\frac{1}{\sqrt{n}}\right) \tag{4.3.5}
\end{aligned}$$

gilt, und anschließend, dass

$$\begin{aligned}
&\frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \frac{1}{2} \sum_{j=1}^{\lfloor ns \rfloor} \frac{\bar{w}_j^{\lfloor ns \rfloor}}{\sigma^2(X_{j-1})} \left(\frac{\frac{1}{nc_n} \sum_{i=1}^n K\left(\frac{X_{j-1}-X_{i-1}}{c_n}\right) (2m\sigma(X_{i-1})\epsilon_i + m^2(X_{i-1}) - \hat{m}^2(X_{j-1}))}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{X_{j-1}-X_{k-1}}{c_n}\right)} \right) \\
&= o_p\left(\frac{1}{\sqrt{n}}\right) \tag{4.3.6}
\end{aligned}$$

gilt.

Die Gleichung (4.3.5) gilt, denn

$$\begin{aligned}
&\frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \frac{1}{2} \sum_{j=1}^{\lfloor ns \rfloor} \frac{\bar{w}_j^{\lfloor ns \rfloor}}{\sigma^2(X_{j-1})} \left(\frac{\frac{1}{nc_n} \sum_{i=1}^n K\left(\frac{X_{j-1}-X_{i-1}}{c_n}\right) (\sigma^2(X_{i-1})\epsilon_i^2 - \sigma^2(X_{j-1}))}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{X_{j-1}-X_{k-1}}{c_n}\right)} \right) \\
&= \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \frac{1}{2} \sum_{j=1}^{\lfloor ns \rfloor} \frac{\bar{w}_j^{\lfloor ns \rfloor}}{\sigma^2(X_{j-1})} \left(\frac{\frac{1}{nc_n} \sum_{i=1}^n K\left(\frac{X_{j-1}-X_{i-1}}{c_n}\right) \sigma^2(X_{i-1})(\epsilon_i^2 - 1)}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{X_{j-1}-X_{k-1}}{c_n}\right)} \right)
\end{aligned}$$

$$\begin{aligned}
& + \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \frac{1}{2} \sum_{j=1}^{\lfloor ns \rfloor} \frac{\bar{w}_j^{\lfloor ns \rfloor}}{\sigma^2(X_{j-1})} \left(\frac{\frac{1}{nc_n} \sum_{i=1}^n K\left(\frac{X_{j-1}-X_{i-1}}{c_n}\right) (\sigma^2(X_{i-1}) - \sigma^2(X_{j-1}))}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{X_{j-1}-X_{k-1}}{c_n}\right)} \right) \\
& = \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \frac{1}{2} \sum_{j=1}^{\lfloor ns \rfloor} \frac{\bar{w}_j^{\lfloor ns \rfloor}}{\sigma^2(X_{j-1})} \left(\frac{\frac{1}{nc_n} \sum_{i=1}^n K\left(\frac{X_{j-1}-X_{i-1}}{c_n}\right) \sigma^2(X_{i-1}) (\epsilon_i^2 - 1)}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{X_{j-1}-X_{k-1}}{c_n}\right)} \right) \\
& + o_p\left(\frac{1}{\sqrt{n}}\right),
\end{aligned}$$

da die Funktionen $\frac{1}{(q_n)^2} \sigma^2(\cdot)$ und $\frac{1}{(q_n^{\sigma})^4 (q_n)^2 \sigma^2(\cdot)}$ die Bedingungen an die Funktionen v_n und t_n in Lemma 10.4 (S. 166) erfüllen, und damit

$$\begin{aligned}
& \sup_{s \in [0,1]} \left| \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \frac{1}{2} \sum_{j=1}^{\lfloor ns \rfloor} \frac{\bar{w}_j^{\lfloor ns \rfloor}}{\sigma^2(X_{j-1})} \left(\frac{\frac{1}{nc_n} \sum_{i=1}^n K\left(\frac{X_{j-1}-X_{i-1}}{c_n}\right) (\sigma^2(X_{i-1}) - \sigma^2(X_{j-1}))}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{X_{j-1}-X_{k-1}}{c_n}\right)} \right) \right| \\
& = O_P\left((q_n)^4 (q_n^{\sigma})^4 \left(c_n^{\frac{1}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} + c_n^2\right) q_n^f\right) \\
& = o_p\left(\frac{1}{\sqrt{n}}\right)
\end{aligned}$$

gilt.

Weiter ist mit Lemma 9.2 (S. 130), Lemma 10.2 (S. 145) und Lemma 10.3 (S. 158)

$$\begin{aligned}
& \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \frac{1}{2} \sum_{j=1}^{\lfloor ns \rfloor} \frac{\bar{w}_j^{\lfloor ns \rfloor}}{\sigma^2(X_{j-1})} \left(\frac{\frac{1}{nc_n} \sum_{i=1}^n K\left(\frac{X_{j-1}-X_{i-1}}{c_n}\right) \sigma^2(X_{i-1}) (\epsilon_i^2 - 1)}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{X_{j-1}-X_{k-1}}{c_n}\right)} \right) \\
& = \frac{1}{2n} \sum_{i=1}^n (\epsilon_i^2 - 1) \left(\frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \sum_{j=1}^{\lfloor ns \rfloor} \frac{\bar{w}_j^{\lfloor ns \rfloor}}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{X_{j-1}-X_{k-1}}{c_n}\right)} K\left(\frac{X_{j-1}-X_{i-1}}{c_n}\right) \frac{\sigma^2(X_{i-1})}{\sigma^2(X_{j-1})} \right) \\
& = \frac{\lfloor ns \rfloor}{n} \frac{1}{2n} \sum_{i=1}^n (\epsilon_i^2 - 1) \\
& + \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \frac{1}{2n} \sum_{i=1}^n (\epsilon_i^2 - 1) \sum_{j=1}^{\lfloor ns \rfloor} \bar{w}_j^{\lfloor ns \rfloor} \left(\frac{\frac{1}{c_n} K\left(\frac{X_{j-1}-X_{i-1}}{c_n}\right)}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{X_{j-1}-X_{k-1}}{c_n}\right)} \right) \left(\frac{\sigma^2(X_{i-1})}{\sigma^2(X_{j-1})} - \frac{\sigma^2(X_{j-1})}{\sigma^2(X_{j-1})} \right) \\
& + \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \frac{1}{2n} \sum_{i=1}^n (\epsilon_i^2 - 1) \sum_{j=1}^{\lfloor ns \rfloor} \bar{w}_j^{\lfloor ns \rfloor} \left(\frac{\frac{1}{c_n} K\left(\frac{X_{j-1}-X_{i-1}}{c_n}\right) - \frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{X_{j-1}-X_{k-1}}{c_n}\right)}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{X_{j-1}-X_{k-1}}{c_n}\right)} \right) \\
& + \frac{1}{n} \sum_{i=1}^n (\epsilon_i^2 - 1) \left(\frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \sum_{j=1}^{\lfloor ns \rfloor} \bar{w}_j^{\lfloor ns \rfloor} - \frac{\lfloor ns \rfloor}{n} \right) \\
& = \frac{\lfloor ns \rfloor}{n} \frac{1}{2n} \sum_{i=1}^n (\epsilon_i^2 - 1) + o_p\left(\frac{1}{\sqrt{n}}\right),
\end{aligned}$$

da

$$\sup_{s \in [0,1]} \left| \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \frac{1}{2n} \sum_{i=1}^n (\epsilon_i^2 - 1) \sum_{j=1}^{\lfloor ns \rfloor} \bar{w}_j^{\lfloor ns \rfloor} \left(\frac{\frac{1}{c_n} K\left(\frac{X_{j-1}-X_{i-1}}{c_n}\right)}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{X_{j-1}-X_{k-1}}{c_n}\right)} \right) \left(\frac{\sigma^2(X_{i-1})}{\sigma^2(X_{j-1})} - \frac{\sigma^2(X_{j-1})}{\sigma^2(X_{j-1})} \right) \right|$$

$$= o_p\left(\frac{1}{\sqrt{n}}\right)$$

nach Lemma 10.3 (S. 158), und

$$\sup_{s \in [0,1]} \left| \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \frac{1}{n} \sum_{i=1}^n (\epsilon_i^2 - 1) \sum_{j=1}^{\lfloor ns \rfloor} \bar{w}_{j \lfloor ns \rfloor} \left(\frac{\frac{1}{c_n} K\left(\frac{X_{j-1} - X_{i-1}}{c_n}\right) - \frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{X_{j-1} - X_{k-1}}{c_n}\right)}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{X_{j-1} - X_{k-1}}{c_n}\right)} \right) \right| = o_p\left(\frac{1}{\sqrt{n}}\right)$$

nach Lemma 10.2 (S. 145), sowie

$$\sup_{s \in [0,1]} \left| \frac{1}{n} \sum_{i=1}^n (\epsilon_i^2 - 1) \left(\frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \sum_{j=1}^{\lfloor ns \rfloor} \bar{w}_{j \lfloor ns \rfloor} - \frac{\lfloor ns \rfloor}{n} \right) \right| = o_p\left(\frac{1}{\sqrt{n}}\right),$$

da $\frac{1}{n} \sum_{i=1}^n (\epsilon_i^2 - 1) = O_P\left(\frac{1}{\sqrt{n}}\right)$ (Summe von zentrierten i.i.d. Zufallsvariablen) und

$$\sup_{s \in [0,1]} \left| \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \underbrace{\sum_{j=1}^{\lfloor ns \rfloor} \bar{w}_{j \lfloor ns \rfloor}}_{=1} - \frac{\lfloor ns \rfloor}{n} \right| = o_P(1) \text{ nach Lemma 9.2 (S. 130).}$$

Damit gilt (4.3.5) und es bleibt zu zeigen, dass (4.3.6) ebenfalls gilt.

Dies ist der Fall, denn

$$\begin{aligned} & \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \frac{1}{2} \sum_{j=1}^{\lfloor ns \rfloor} \frac{\bar{w}_{j \lfloor ns \rfloor}}{\sigma^2(X_{j-1})} \left(\frac{\frac{1}{nc_n} \sum_{i=1}^n K\left(\frac{X_{j-1} - X_{i-1}}{c_n}\right) (2m\sigma(X_{i-1})\epsilon_i + m^2(X_{i-1}) - \hat{m}^2(X_{j-1}))}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{X_{j-1} - X_{k-1}}{c_n}\right)} \right) \\ &= \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \frac{1}{2} \sum_{j=1}^{\lfloor ns \rfloor} \frac{\bar{w}_{j \lfloor ns \rfloor}}{\sigma^2(X_{j-1})} \left(\frac{\frac{1}{nc_n} \sum_{i=1}^n K\left(\frac{X_{j-1} - X_{i-1}}{c_n}\right) 2m\sigma(X_{i-1})\epsilon_i}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{X_{j-1} - X_{k-1}}{c_n}\right)} \right) \\ &+ \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \frac{1}{2} \sum_{j=1}^{\lfloor ns \rfloor} \frac{\bar{w}_{j \lfloor ns \rfloor}}{\sigma^2(X_{j-1})} \left(\frac{\frac{1}{nc_n} \sum_{i=1}^n K\left(\frac{X_{j-1} - X_{i-1}}{c_n}\right) (m^2(X_{j-1}) - \hat{m}^2(X_{j-1}))}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{X_{j-1} - X_{k-1}}{c_n}\right)} \right) \\ &+ \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \frac{1}{2} \sum_{j=1}^{\lfloor ns \rfloor} \frac{\bar{w}_{j \lfloor ns \rfloor}}{\sigma^2(X_{j-1})} \left(\frac{\frac{1}{nc_n} \sum_{i=1}^n K\left(\frac{X_{j-1} - X_{i-1}}{c_n}\right) (m^2(X_{i-1}) - m^2(X_{j-1}))}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{X_{j-1} - X_{k-1}}{c_n}\right)} \right) \end{aligned}$$

Auch die Funktionen $\frac{1}{(q_n)^2} m^2(\cdot)$ und $\frac{1}{(q_n^\sigma)^4 (q_n)^2 \sigma^2(\cdot)}$ erfüllen die Bedingungen an die Funktionen v_n und t_n in Lemma 10.4 (S. 166), daher gilt

$$\begin{aligned} & \sup_{s \in [0,1]} \left| \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \frac{1}{2} \sum_{j=1}^{\lfloor ns \rfloor} \frac{\bar{w}_{j \lfloor ns \rfloor}}{\sigma^2(X_{j-1})} \left(\frac{\frac{1}{nc_n} \sum_{i=1}^n K\left(\frac{X_{j-1} - X_{i-1}}{c_n}\right) (m^2(X_{i-1}) - m^2(X_{j-1}))}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{X_{j-1} - X_{k-1}}{c_n}\right)} \right) \right| \\ &= O_P\left((q_n)^4 (q_n^\sigma)^4 \left(c_n^{\frac{1}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} + c_n^2\right) q_n^f\right) \\ &= o_p\left(\frac{1}{\sqrt{n}}\right). \end{aligned}$$

Außerdem ist

$$\begin{aligned}
& \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \frac{1}{2} \sum_{j=1}^{\lfloor ns \rfloor} \frac{\bar{w}_j \lfloor ns \rfloor}{\sigma^2(X_{j-1})} \left(\frac{\frac{1}{nc_n} \sum_{i=1}^n K\left(\frac{X_{j-1}-X_{i-1}}{c_n}\right) (m^2(X_{j-1}) - \hat{m}^2(X_{j-1}))}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{X_{j-1}-X_{k-1}}{c_n}\right)} \right) \\
&= \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \frac{1}{2} \sum_{j=1}^{\lfloor ns \rfloor} \frac{\bar{w}_j \lfloor ns \rfloor}{\sigma^2(X_{j-1})} \left(\frac{\frac{1}{nc_n} \sum_{i=1}^n K\left(\frac{X_{j-1}-X_{i-1}}{c_n}\right) (2m(m - \hat{m})(X_{j-1}) - (m - \hat{m})^2(X_{j-1}))}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{X_{j-1}-X_{k-1}}{c_n}\right)} \right) \\
&= \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \frac{1}{2} \sum_{j=1}^{\lfloor ns \rfloor} \frac{\bar{w}_j \lfloor ns \rfloor}{\sigma^2(X_{j-1})} \left(\frac{\frac{1}{nc_n} \sum_{i=1}^n K\left(\frac{X_{j-1}-X_{i-1}}{c_n}\right) 2m(m - \hat{m})(X_{j-1})}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{X_{j-1}-X_{k-1}}{c_n}\right)} \right) \\
&\quad - \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \frac{1}{2} \sum_{j=1}^{\lfloor ns \rfloor} \frac{\bar{w}_j \lfloor ns \rfloor}{\sigma^2(X_{j-1})} \left(\frac{\frac{1}{nc_n} \sum_{i=1}^n K\left(\frac{X_{j-1}-X_{i-1}}{c_n}\right) (m - \hat{m})^2(X_{j-1})}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{X_{j-1}-X_{k-1}}{c_n}\right)} \right)
\end{aligned}$$

und

$$\begin{aligned}
& \sup_{s \in [0,1]} \left| \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \frac{1}{2} \sum_{j=1}^{\lfloor ns \rfloor} \frac{\bar{w}_j \lfloor ns \rfloor}{\sigma^2(X_{j-1})} \left(\frac{\frac{1}{nc_n} \sum_{i=1}^n K\left(\frac{X_{j-1}-X_{i-1}}{c_n}\right) (m - \hat{m})^2(X_{j-1})}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{X_{j-1}-X_{k-1}}{c_n}\right)} \right) \right| \\
&= \sup_{s \in [0,1]} \left| \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \frac{1}{2} \sum_{j=1}^{\lfloor ns \rfloor} \frac{\bar{w}_j \lfloor ns \rfloor}{\sigma^2(X_{j-1})} (m - \hat{m})^2(X_{j-1}) \right| \\
&\leq \frac{1}{2} \left(\sup_{x \in I_n} \left| \frac{m - \hat{m}}{\sigma}(x) \right| \right)^2 \\
&= O_P \left(\left(\left(c_n^{-\frac{1}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} + c_n^2 \right) q_n q_n^\sigma q_n^f \right)^2 \right) \\
&= o_p \left(\frac{1}{\sqrt{n}} \right)
\end{aligned}$$

nach Lemma 3.4 (S. 35) mit $c_n \sim n^{-\frac{1}{4}} \log(n)^r$, für alle $|r| < \infty$.

Damit bleibt die Konvergenzrate zu zeigen für

$$\begin{aligned}
& \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \frac{1}{2} \sum_{j=1}^{\lfloor ns \rfloor} \frac{\bar{w}_j \lfloor ns \rfloor}{\sigma^2(X_{j-1})} \left(\frac{\frac{1}{nc_n} \sum_{i=1}^n K\left(\frac{X_{j-1}-X_{i-1}}{c_n}\right) 2m\sigma(X_{i-1})\epsilon_i}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{X_{j-1}-X_{k-1}}{c_n}\right)} \right) \\
&+ \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \frac{1}{2} \sum_{j=1}^{\lfloor ns \rfloor} \frac{\bar{w}_j \lfloor ns \rfloor}{\sigma^2(X_{j-1})} \left(\frac{\frac{1}{nc_n} \sum_{i=1}^n K\left(\frac{X_{j-1}-X_{i-1}}{c_n}\right) 2m(m - \hat{m})(X_{j-1})}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{X_{j-1}-X_{k-1}}{c_n}\right)} \right) \\
&= \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \frac{1}{2} \sum_{j=1}^{\lfloor ns \rfloor} \frac{\bar{w}_j \lfloor ns \rfloor}{\sigma^2(X_{j-1})} \left(\frac{\frac{1}{nc_n} \sum_{i=1}^n K\left(\frac{X_{j-1}-X_{i-1}}{c_n}\right) 2m\sigma(X_{i-1})\epsilon_i}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{X_{j-1}-X_{k-1}}{c_n}\right)} \right) \\
&+ \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \frac{1}{2} \sum_{j=1}^{\lfloor ns \rfloor} \frac{\bar{w}_j \lfloor ns \rfloor}{\sigma^2(X_{j-1})} \left(\frac{\frac{1}{nc_n} \sum_{i=1}^n K\left(\frac{X_{j-1}-X_{i-1}}{c_n}\right) 2m^2(X_{j-1})}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{X_{j-1}-X_{k-1}}{c_n}\right)} \right. \\
&\quad \left. - 2m(X_{j-1}) \frac{\frac{1}{nc_n} \sum_{i=1}^n K\left(\frac{X_{j-1}-X_{i-1}}{c_n}\right) X_i}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{X_{j-1}-X_{k-1}}{c_n}\right)} \right)
\end{aligned}$$

$$\begin{aligned}
&= \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \frac{1}{2} \sum_{j=1}^{\lfloor ns \rfloor} \frac{\bar{w}_j^{\lfloor ns \rfloor}}{\sigma^2(X_{j-1})} \left(\frac{\frac{1}{nc_n} \sum_{i=1}^n K\left(\frac{X_{j-1}-X_{i-1}}{c_n}\right) 2m\sigma(X_{i-1})\epsilon_i}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{X_{j-1}-X_{k-1}}{c_n}\right)} \right) \\
&\quad + \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \frac{1}{2} \sum_{j=1}^{\lfloor ns \rfloor} \frac{\bar{w}_j^{\lfloor ns \rfloor}}{\sigma^2(X_{j-1})} \left(\frac{\frac{1}{nc_n} \sum_{i=1}^n K\left(\frac{X_{j-1}-X_{i-1}}{c_n}\right) 2m^2(X_{j-1})}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{X_{j-1}-X_{k-1}}{c_n}\right)} \right. \\
&\quad \quad \left. - 2m(X_{j-1}) \frac{\frac{1}{nc_n} \sum_{i=1}^n K\left(\frac{X_{j-1}-X_{i-1}}{c_n}\right) (m(X_{i-1}) + \sigma(X_{i-1})\epsilon_i)}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{X_{j-1}-X_{k-1}}{c_n}\right)} \right) \\
&= \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \sum_{j=1}^{\lfloor ns \rfloor} \frac{\bar{w}_j^{\lfloor ns \rfloor}}{\sigma^2(X_{j-1})} \left(\frac{\frac{1}{nc_n} \sum_{i=1}^n K\left(\frac{X_{j-1}-X_{i-1}}{c_n}\right) (m\sigma(X_{i-1})\epsilon_i - m(X_{j-1})\sigma(X_{i-1})\epsilon_i)}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{X_{j-1}-X_{k-1}}{c_n}\right)} \right) \\
&\quad + \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \sum_{j=1}^{\lfloor ns \rfloor} \frac{\bar{w}_j^{\lfloor ns \rfloor}}{\sigma^2(X_{j-1})} \left(\frac{\frac{1}{nc_n} \sum_{i=1}^n K\left(\frac{X_{j-1}-X_{i-1}}{c_n}\right) (m^2(X_{j-1}) - m(X_{j-1})m(X_{i-1}))}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{X_{j-1}-X_{k-1}}{c_n}\right)} \right) \\
&= \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \sum_{j=1}^{\lfloor ns \rfloor} \frac{\bar{w}_j^{\lfloor ns \rfloor}}{\sigma^2(X_{j-1})} \left(\frac{\frac{1}{nc_n} \sum_{i=1}^n K\left(\frac{X_{j-1}-X_{i-1}}{c_n}\right) \sigma(X_{i-1})\epsilon_i (m(X_{i-1}) - m(X_{j-1}))}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{X_{j-1}-X_{k-1}}{c_n}\right)} \right) \\
&\quad + \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \sum_{j=1}^{\lfloor ns \rfloor} \frac{\bar{w}_j^{\lfloor ns \rfloor}}{\sigma^2(X_{j-1})} \frac{m(X_{j-1})}{\sigma^2(X_{j-1})} \left(\frac{\frac{1}{nc_n} \sum_{i=1}^n K\left(\frac{X_{j-1}-X_{i-1}}{c_n}\right) (m(X_{j-1}) - m(X_{i-1}))}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{X_{j-1}-X_{k-1}}{c_n}\right)} \right).
\end{aligned}$$

Weiter gilt wieder nach Lemma 10.4 (S. 166)

$$\begin{aligned}
&\sup_{s \in [0,1]} \left| \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \sum_{j=1}^{\lfloor ns \rfloor} \frac{\bar{w}_j^{\lfloor ns \rfloor}}{\sigma^2(X_{j-1})} \frac{m(X_{j-1})}{\sigma^2(X_{j-1})} \left(\frac{\frac{1}{nc_n} \sum_{i=1}^n K\left(\frac{X_{j-1}-X_{i-1}}{c_n}\right) (m(X_{i-1}) - m(X_{j-1}))}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{X_{j-1}-X_{k-1}}{c_n}\right)} \right) \right| \\
&= O_P \left((q_n)^4 (q_n^\sigma)^4 \left(c_n^{\frac{1}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} + c_n^2 \right) q_n^f \right), \text{ da die Funktionen } \frac{1}{q_n} m(\cdot) \text{ und } \frac{1}{(q_n^\sigma)^4 (q_n)^3} \frac{m(\cdot)}{\sigma^2(\cdot)} \\
&\quad \text{die Bedingungen an die Funktionen } v_n \text{ und } t_n \text{ erfüllen,} \\
&= o_P \left(\frac{1}{\sqrt{n}} \right),
\end{aligned}$$

und nach Lemma 10.3 (S. 158)

$$\begin{aligned}
&\sup_{s \in [0,1]} \left| \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \sum_{j=1}^{\lfloor ns \rfloor} \frac{\bar{w}_j^{\lfloor ns \rfloor}}{\sigma^2(X_{j-1})} \left(\frac{\frac{1}{nc_n} \sum_{i=1}^n K\left(\frac{X_{j-1}-X_{i-1}}{c_n}\right) \sigma(X_{i-1})\epsilon_i (m(X_{i-1}) - m(X_{j-1}))}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{X_{j-1}-X_{k-1}}{c_n}\right)} \right) \right| \\
&= o_P \left(\frac{1}{\sqrt{n}} \right).
\end{aligned}$$

Somit ist gezeigt, dass (4.3.6) gilt, und damit auch, dass (4.3.4) gilt.

Insgesamt gilt daher gleichmaig in $t \in \mathbb{R}$ und $s \in [0, 1]$ die Entwicklung

$$\frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \left(\sum_{j=1}^{\lfloor ns \rfloor} \bar{w}_j^{\lfloor ns \rfloor} I\{\hat{\epsilon}_j \leq t\} - F(t) \right) = \frac{1}{n} \sum_{j=1}^{\lfloor ns \rfloor} (I\{\epsilon_j \leq t\} - F(t))$$

$$\begin{aligned}
& + \frac{\lfloor ns \rfloor}{n} f(t) \frac{1}{n} \sum_{i=1}^n \epsilon_i + f(t) \cdot o_p \left(\frac{1}{\sqrt{n}} \right) \\
& + \frac{\lfloor ns \rfloor}{n} f(t) t \frac{1}{2n} \sum_{i=1}^n (\epsilon_i^2 - 1) + f(t) t \cdot o_p \left(\frac{1}{\sqrt{n}} \right) \\
& + o_p \left(\frac{1}{\sqrt{n}} \right) \\
& = \frac{1}{n} \sum_{j=1}^{\lfloor ns \rfloor} (I\{\epsilon_j \leq t\} - F(t)) \\
& + \frac{\lfloor ns \rfloor}{n} f(t) \frac{1}{n} \sum_{i=1}^n \epsilon_i \\
& + \frac{\lfloor ns \rfloor}{n} f(t) t \frac{1}{2n} \sum_{i=1}^n (\epsilon_i^2 - 1) \\
& + o_p \left(\frac{1}{\sqrt{n}} \right),
\end{aligned}$$

da nach Voraussetzung $\sup_{t \in \mathbb{R}} |f(t)| < \infty$ und $\sup_{t \in \mathbb{R}} |f(t)t| < \infty$.

□

4.3 Schwache Konvergenz des sequentiellen empirischen Prozesses der Residuen

In diesem Abschnitt wird aus der stochastischen Entwicklung in Abschnitt 4.2 die schwache Konvergenz der Folge von stochastischen Prozessen $\sqrt{n} \left(\frac{\sum_{k=1}^{\lfloor n \cdot \rfloor} w_{nk}}{n} \left(\hat{F}_{\lfloor n \cdot \rfloor} - F \right) \right)$ abgeleitet.

Für die weitere Anwendung, z.B. Goodness-of-fit-Tests wie in Kapitel 7, ist besonders der klassische Fall des empirischen Prozesses von Interesse, in dem die Innovationen identisch verteilt sind mit Verteilungsfunktion F und dieses F durch den empirischen Prozess der Residuen \hat{F}_n , wie in Abschnitt 4.1 definiert, geschätzt wird.

Schwache Konvergenz von $\sqrt{n} \left(\hat{F}_n - F \right)$ wurde in der Literatur bereits für verschiedene Modelle gezeigt, z.B. für heteroskedastische nichtparametrische Regressionsmodelle mit unabhängigen Daten in Akritas and van Keilegom [2001]. Für homoskedastische nichtparametrische Autoregressionsmodelle wird in Müller et al. [2009] eine stochastische Entwicklung für den empirischen Prozess der Residuen gezeigt, aus der sich die schwache Konvergenz von $\sqrt{n} \left(\hat{F}_n - F \right)$ folgern lässt.

In der vorliegenden Arbeit werden diese Resultate darauf ausgeweitet, dass schwache Konvergenz von $\sqrt{n} \left(\hat{F}_n - F \right)$ für heteroskedastische nichtparametrische Autoregressionsmodelle gezeigt wird. Ein verwandtes Resultat wird in Dette et al. [2009] hergeleitet, wo für ein heteroskedastisches nichtparametrisches Regressionsmodell mit abhängigen Daten eine Darstellung für den empirischen Prozess der Residuen gezeigt wird, aus der sich eine stochastische Entwicklung für den empirischen Prozess der Residuen, die aus einem empirischen Prozess unabhängiger Zufallvariablen besteht, herleiten lässt, und aus der damit auch schwache Konvergenz von $\sqrt{n} \left(\hat{F}_n - F \right)$ ableitbar ist. Die schwache

Konvergenz von $\sqrt{n}(\hat{F}_n - F)$ wird dort allerdings nicht direkt gezeigt.
(Vergleich Kapitel 8, S. 114)

Korollar 4.5 zeigt die schwache Konvergenz von $\sqrt{n}(\hat{F}_n - F)$ für heteroskedastische nichtparametrische Autoregressionsmodelle und ist eine direkte Folgerung aus Satz 4.4.

Satz 4.4 Unter den Voraussetzungen von Satz 4.2 (S. 49) gilt

$$\sqrt{n} \left(\frac{\sum_{k=1}^{\lfloor n \cdot \rfloor} w_{nk}}{n} (\hat{F}_{\lfloor n \cdot \rfloor} - F) \right) \Rightarrow \mathbb{K}_F,$$

wobei $(\mathbb{K}_F(s, t))_{s \in [0, 1], t \in \mathbb{R}}$ ein Gauß-Prozess, also ein stochastischer Prozess mit normalverteilten Randverteilungen ist, für den gilt

$$E[\mathbb{K}_F(s, t)] = 0 \quad \forall s \in [0, 1], t \in \mathbb{R}$$

und

$$\begin{aligned} \text{Cov}(\mathbb{K}_F(s_1, t_1), \mathbb{K}_F(s_2, t_2)) &= \min(s_1, s_2) (F(\min(t_1, t_2)) - F(t_1)F(t_2)) \\ &\quad + s_1 s_2 \left(f(t_1) (E[\epsilon_1 I\{\epsilon_1 \leq t_2\}] + t_1 E[(\epsilon_1^2 - 1) I\{\epsilon_1 \leq t_2\}]) \right. \\ &\quad \quad \quad \left. + f(t_2) (E[\epsilon_1 I\{\epsilon_1 \leq t_1\}] + t_2 E[(\epsilon_1^2 - 1) I\{\epsilon_1 \leq t_1\}]) \right. \\ &\quad \quad \quad \left. + f(t_1) f(t_2) (1 + (t_1 + t_2) E[\epsilon_1^3] + t_1 t_2 (E[\epsilon_1^4] - 1)) \right). \end{aligned}$$

Korollar 4.5 Unter den Voraussetzungen von Satz 4.2 (S. 49) gilt

$$\sqrt{n}(\hat{F}_n - F) \Rightarrow \mathbb{G}_F,$$

wobei $(\mathbb{G}_F(t))_{t \in \mathbb{R}}$ ein Gauß-Prozess, also ein stochastischer Prozess mit normalverteilten Randverteilungen ist, für den gilt

$$E[\mathbb{G}_F(t)] = 0 \quad \forall t \in \mathbb{R}$$

und

$$\begin{aligned} \text{Cov}(\mathbb{G}_F(t_1), \mathbb{G}_F(t_2)) &= F(\min(t_1, t_2)) - F(t_1)F(t_2) \\ &\quad + f(t_1) (E[\epsilon_1 I\{\epsilon_1 \leq t_2\}] + t_1 E[(\epsilon_1^2 - 1) I\{\epsilon_1 \leq t_2\}]) \\ &\quad + f(t_2) (E[\epsilon_1 I\{\epsilon_1 \leq t_1\}] + t_2 E[(\epsilon_1^2 - 1) I\{\epsilon_1 \leq t_1\}]) \\ &\quad + f(t_1) f(t_2) (1 + (t_1 + t_2) E[\epsilon_1^3] + t_1 t_2 (E[\epsilon_1^4] - 1)). \end{aligned}$$

Beweis von Korollar 4.5 Es gilt

$$\sqrt{n}(\hat{F}_n - F) = \underbrace{\frac{n}{\sum_{k=1}^n w_{nk}}}_{=1+o_P(1)} \sqrt{n} \left(\frac{\sum_{k=1}^{\lfloor n \cdot \rfloor} w_{nk}}{n} (\hat{F}_{\lfloor n \cdot \rfloor} - F) \right).$$

nach Lemma 9.2 (S. 130)

Daher gilt mit dem Lemma von Slutsky

$$\sqrt{n}(\hat{F}_n - F) \Rightarrow \mathbb{G}_F$$

wenn

$$\sqrt{n} \left(\frac{\sum_{k=1}^{\lfloor n \cdot 1 \rfloor} w_{nk}}{n} \left(\hat{F}_{\lfloor n \cdot 1 \rfloor} - F \right) \right) \Rightarrow \mathbb{G}_F,$$

was eine direkte Folgerung aus Satz 4.4, da die Projektion auf $s = 1$ eine stetige Abbildung ist und somit das Continuous-Mapping-Theorem angewendet werden kann.

□

Beweis von Satz 4.4 Mit Satz 4.2 gilt nach dem Lemma von Slutsky

$$\sqrt{n} \left(\frac{\sum_{k=1}^{\lfloor n \cdot \cdot \rfloor} w_{nk}}{n} \left(\hat{F}_{\lfloor n \cdot \cdot \rfloor} - F \right) \right) \Rightarrow \mathbb{K}_F$$

wenn

$$\left(\sqrt{n} \left(\frac{1}{n} \sum_{j=1}^{\lfloor ns \rfloor} (I \{ \epsilon_j \leq t \} - F(t)) + \frac{\lfloor ns \rfloor}{n} f(t) \frac{1}{n} \sum_{j=1}^n \epsilon_j + \frac{\lfloor ns \rfloor}{n} f(t) t \frac{1}{2n} \sum_{j=1}^n (\epsilon_j^2 - 1) \right) \right)_{s \in [0,1], t \in \mathbb{R}} \Rightarrow \mathbb{K}_F. \quad (4.5.1)$$

Um zu zeigen, dass die zweite Konvergenz gilt, wird das Theorem 2.11.9 in van der Vaart and Wellner [1996] verwendet. Hierfür wird zunächst die Notation

$$\begin{aligned} & \sqrt{n} \left(\frac{1}{n} \sum_{j=1}^{\lfloor ns \rfloor} (I \{ \epsilon_j \leq t \} - F(t)) + \frac{\lfloor ns \rfloor}{n} f(t) \frac{1}{n} \sum_{j=1}^n \epsilon_j + \frac{\lfloor ns \rfloor}{n} f(t) t \frac{1}{2n} \sum_{j=1}^n (\epsilon_j^2 - 1) \right) \\ &= \sum_{j=1}^n (Z_{nj}(s, t, f(t), f(t)t) - E[Z_{nj}(s, t, f(t), f(t)t)]) \end{aligned} \quad (4.5.2)$$

eingeführt, mit

$$Z_{nj}(s, t, u, v) := n^{-\frac{1}{2}} \left(I \{ \epsilon_j \leq t \} I \left\{ \frac{j}{n} \leq s \right\} + \frac{\lfloor ns \rfloor}{n} u \epsilon_j + \frac{\lfloor ns \rfloor}{n} \frac{1}{2} v (\epsilon_j^2 - 1) \right), \quad (s, t, u, v) \in \mathcal{F}$$

und

$$\mathcal{F} := \left\{ (s, t, u, v) : s \in [0, 1], t \in \mathbb{R}, u \in \left[0, \sup_{t \in \mathbb{R}} f(t) \right], v \in \left[-\sup_{t \in \mathbb{R}} |f(t)t|, \sup_{t \in \mathbb{R}} |f(t)t| \right] \right\}$$

wobei nach Voraussetzung $\sup_{t \in \mathbb{R}} f(t) < \infty$ und $\sup_{t \in \mathbb{R}} |f(t)t| < \infty$.

Damit ist \mathcal{F} ein total beschränkter Raum bezüglich der Semimetrik

$$\rho((s, t, u, v), (s', t', u', v')) := |s - s'| + |F(t) - F(t')| + |u - u'| + |v - v'|.$$

Mit Hilfe von Theorem 2.11.9 in van der Vaart and Wellner [1996] wird im Folgenden gezeigt, dass

$$\left(\sum_{j=1}^n (Z_{nj}(s, t, u, v) - E[Z_{nj}(s, t, u, v)]) \right)_{(s,t,u,v) \in \mathcal{F}} \quad (4.5.3)$$

schwach gegen einen Gauß-Prozess konvergiert. Mit dem Continuous-Mapping-Theorem kann daraus durch die Projektionen $u = f(t)$ und $v = f(t)t$ gefolgert werden, dass auch

$$\left(\sum_{j=1}^n (Z_{nj}(s, t, f(t), f(t)t) - E[Z_{nj}(s, t, f(t), f(t)t)]) \right)_{(s,t) \in [0,1] \times \mathbb{R}}$$

schwach gegen einen Gauß-Prozess konvergiert. Die genaue Form des Grenzprozesses \mathbb{K}_F ergibt sich aus der Berechnung der Kovarianz, die am Ende dieses Beweises durchgeführt wird.

Nach (4.5.1) und (4.5.2) gilt damit dann auch die schwache Konvergenz in der Behauptung des Satzes.

Für die Anwendung des Theorems 2.11.9 in van der Vaart and Wellner [1996] müssen drei Bedingungen überprüft werden.

Die erste Bedingung gilt, da

$$\begin{aligned} & \sum_{j=1}^n E \left[\sup_{(s,t,u,v) \in \mathcal{F}} |Z_{nj}(s, t, u, v)| I \left\{ \sup_{(s,t,u,v) \in \mathcal{F}} |Z_{nj}(s, t, u, v)| > \eta \right\} \right] \\ & \leq \sum_{j=1}^n E \left[n^{-\frac{1}{2}} \left(1 + \sup_{t \in \mathbb{R}} f(t) |\epsilon_j| + \sup_{t \in \mathbb{R}} |f(t)t| |\epsilon_j^2 - 1| \right) \right. \\ & \quad \left. \cdot I \left\{ n^{-\frac{1}{2}} \left(1 + \sup_{t \in \mathbb{R}} f(t) |\epsilon_j| + \sup_{t \in \mathbb{R}} |f(t)t| |\epsilon_j^2 - 1| \right) > \eta \right\} \right] \\ & = E \left[n^{\frac{1}{2}} \left(1 + \sup_{t \in \mathbb{R}} f(t) |\epsilon_1| + \sup_{t \in \mathbb{R}} |f(t)t| |\epsilon_1^2 - 1| \right) \right. \\ & \quad \left. \cdot I \left\{ n^{-\frac{1}{2}} \left(1 + \sup_{t \in \mathbb{R}} f(t) |\epsilon_1| + \sup_{t \in \mathbb{R}} |f(t)t| |\epsilon_1^2 - 1| \right) > \eta \right\} \right] \\ & = \frac{1}{\eta} E \left[\eta n^{\frac{1}{2}} \left(1 + \sup_{t \in \mathbb{R}} f(t) |\epsilon_1| + \sup_{t \in \mathbb{R}} |f(t)t| |\epsilon_1^2 - 1| \right) \right. \\ & \quad \left. \cdot I \left\{ \left(1 + \sup_{t \in \mathbb{R}} f(t) |\epsilon_1| + \sup_{t \in \mathbb{R}} |f(t)t| |\epsilon_1^2 - 1| \right) > \eta n^{\frac{1}{2}} \right\} \right] \\ & \leq \frac{1}{\eta} E \left[\left(1 + \sup_{t \in \mathbb{R}} f(t) |\epsilon_1| + \sup_{t \in \mathbb{R}} |f(t)t| |\epsilon_1^2 - 1| \right)^2 \right. \\ & \quad \left. \cdot I \left\{ \left(1 + \sup_{t \in \mathbb{R}} f(t) |\epsilon_1| + \sup_{t \in \mathbb{R}} |f(t)t| |\epsilon_1^2 - 1| \right) > \eta n^{\frac{1}{2}} \right\} \right] \\ & \xrightarrow{n \rightarrow \infty} 0 \quad \forall \eta > 0, \end{aligned}$$

da nach Voraussetzung

$$E \left[\left(\underbrace{1 + \sup_{t \in \mathbb{R}} f(t) |\epsilon_1|}_{< \infty} + \underbrace{\sup_{t \in \mathbb{R}} |f(t)t| |\epsilon_1^2 - 1|}_{< \infty} \right)^2 \right] < \infty, \quad \text{da } E[\epsilon_1^4] < \infty. \quad (4.5.4)$$

Damit gilt

$$\sum_{j=1}^n E \left[\sup_{(s,t,u,v) \in \mathcal{F}} |Z_{nj}(s, t, u, v)| I \left\{ \sup_{(s,t,u,v) \in \mathcal{F}} |Z_{nj}(s, t, u, v)| > \eta \right\} \right] \xrightarrow{n \rightarrow \infty} 0 \quad \forall \eta > 0$$

was gerade die erste Bedingung im verwendeten Theorem ist.

Für die zweite und dritte Bedingung wird der folgende Term benötigt.

$$\begin{aligned}
n (Z_{nj}(s, t, u, v) - Z_{nj}(s', t', u', v'))^2 &= \left((I \{\epsilon_j \leq t\} - I \{\epsilon_j \leq t'\}) I \left\{ \frac{j}{n} \leq s \right\} \right. \\
&\quad + \frac{\lfloor ns \rfloor}{n} (u - u') \epsilon_j \\
&\quad + \frac{\lfloor ns \rfloor}{n} \frac{1}{2} (v - v') (\epsilon_j^2 - 1) \\
&\quad + I \{\epsilon_j \leq t'\} \left(I \left\{ \frac{j}{n} \leq s \right\} - I \left\{ \frac{j}{n} \leq s' \right\} \right) \\
&\quad \left. + \left(\frac{\lfloor ns \rfloor}{n} - \frac{\lfloor ns' \rfloor}{n} \right) \left(u' \epsilon_j + \frac{1}{2} v' (\epsilon_j^2 - 1) \right) \right)^2.
\end{aligned} \tag{4.5.5}$$

Durch Ausmultiplizieren von (4.5.5) ergibt sich für die zweite Bedingung

$$\begin{aligned}
&\sum_{j=1}^n E \left[(Z_{nj}(s, t, u, v) - Z_{nj}(s', t', u', v'))^2 \right] \\
&= \frac{1}{n} \sum_{j=1}^n E \left[n (Z_{nj}(s, t, u, v) - Z_{nj}(s', t', u', v'))^2 \right] \\
&\leq \frac{1}{n} \sum_{j=1}^n \left(|F(t) - F(t')| + |u - u'| E[|\epsilon_j|] + |v - v'| E[|\epsilon_j^2 - 1|] \right. \\
&\quad \left. + \left| I \left\{ \frac{j}{n} \leq s \right\} - I \left\{ \frac{j}{n} \leq s' \right\} \right| + \left| \frac{\lfloor ns \rfloor}{n} - \frac{\lfloor ns' \rfloor}{n} \right| (|u'| E[|\epsilon_j|] + |v'| E[|\epsilon_j^2 - 1|]) \right) \\
&\quad + |u - u'| E[|\epsilon_j|] + (u - u')^2 E[\epsilon_j^2] + |u - u'| |v - v'| E[|\epsilon_j(\epsilon_j^2 - 1)|] \\
&\quad + |u - u'| E[|\epsilon_j|] + |u - u'| (|u'| E[\epsilon_j^2] + |v'| E[|\epsilon_j(\epsilon_j^2 - 1)|]) \\
&\quad + |v - v'| E[|\epsilon_j^2 - 1|] + |u - u'| |v - v'| E[|\epsilon_j(\epsilon_j^2 - 1)|] + (v - v')^2 E[(\epsilon_j^2 - 1)^2] \\
&\quad + |v - v'| E[|\epsilon_j^2 - 1|] + |v - v'| (|u'| E[|\epsilon_j(\epsilon_j^2 - 1)|] + |v'| E[(\epsilon_j^2 - 1)^2]) \\
&\quad + |F(t) - F(t')| + |u - u'| E[|\epsilon_j|] + |v - v'| E[|\epsilon_j^2 - 1|] \\
&\quad + \left| I \left\{ \frac{j}{n} \leq s \right\} - I \left\{ \frac{j}{n} \leq s' \right\} \right| + \left| \frac{\lfloor ns \rfloor}{n} - \frac{\lfloor ns' \rfloor}{n} \right| (|u'| E[|\epsilon_j|] + |v'| E[|\epsilon_j^2 - 1|]) \\
&\quad + |F(t) - F(t')| (|u'| E[|\epsilon_j|] + |v'| E[|\epsilon_j^2 - 1|]) + |u - u'| (|u'| E[\epsilon_j^2] + |v'| E[|\epsilon_j(\epsilon_j^2 - 1)|]) \\
&\quad + |v - v'| (|u'| E[|\epsilon_j(\epsilon_j^2 - 1)|] + |v'| E[(\epsilon_j^2 - 1)^2]) \\
&\quad + \left| I \left\{ \frac{j}{n} \leq s \right\} - I \left\{ \frac{j}{n} \leq s' \right\} \right| (|u'| E[|\epsilon_j|] + |v'| E[|\epsilon_j^2 - 1|]) \\
&\quad \left. + \left(\frac{\lfloor ns \rfloor}{n} - \frac{\lfloor ns' \rfloor}{n} \right)^2 E \left[(u' \epsilon_j + v' (\epsilon_j^2 - 1))^2 \right] \right) \\
&= O \left(|F(t) - F(t')| + |u - u'| + |v - v'| + \frac{1}{n} \sum_{j=1}^n \left| I \left\{ \frac{j}{n} \leq s \right\} - I \left\{ \frac{j}{n} \leq s' \right\} \right| + \left| \frac{\lfloor ns \rfloor}{n} - \frac{\lfloor ns' \rfloor}{n} \right| \right),
\end{aligned}$$

da $E[\epsilon_j^4] < \infty$ für alle j , $|u| < \infty$, $|u'| < \infty$ und $|v| < \infty$, $|v'| < \infty$ nach Voraussetzung.

Weiter gilt

$$\begin{aligned}
& \frac{1}{n} \sum_{j=1}^n \left| I \left\{ \frac{j}{n} \leq s \right\} - I \left\{ \frac{j}{n} \leq s' \right\} \right| \\
&= \frac{1}{n} \left(0 + \sum_{j=\min(\lfloor ns \rfloor, \lfloor ns' \rfloor)+1}^{\max(\lfloor ns \rfloor, \lfloor ns' \rfloor)} 1 + 0 \right) \\
&\quad \text{wobei } \sum_{j=\min(\lfloor ns \rfloor, \lfloor ns' \rfloor)+1}^{\max(\lfloor ns \rfloor, \lfloor ns' \rfloor)} := 0 \text{ falls } \lfloor ns \rfloor = \lfloor ns' \rfloor \\
&= \frac{1}{n} (\max(\lfloor ns \rfloor, \lfloor ns' \rfloor) - \min(\lfloor ns \rfloor, \lfloor ns' \rfloor)) \\
&= \left| \frac{\lfloor ns \rfloor}{n} - \frac{\lfloor ns' \rfloor}{n} \right| \\
&\leq \frac{1}{n} (\max(ns, ns') + 1 - \min(ns, ns')) \\
&= |s - s'| + \frac{1}{n} \tag{4.5.6}
\end{aligned}$$

und damit ist

$$\begin{aligned}
& \sum_{j=1}^n E \left[(Z_{nj}(s, t, u, v) - Z_{nj}(s', t', u', v'))^2 \right] \\
&= O \left(|F(t) - F(t')| + |u - u'| + |v - v'| + |s - s'| + \frac{1}{n} \right)
\end{aligned}$$

also

$$\begin{aligned}
& \sup_{\substack{s, t, u, v, s', t', u', v': \\ \rho((s, t, u, v), (s', t', u', v')) < \bar{\epsilon}_n}} \sum_{j=1}^n E \left[(Z_{nj}(s, t, u, v) - Z_{nj}(s', t', u', v'))^2 \right] \\
&= O(\bar{\epsilon}_n + \frac{1}{n}) \\
&= o(1)
\end{aligned}$$

für alle $\bar{\epsilon}_n \searrow 0$. Dies ist gerade die zweite Bedingung in Theorem 2.11.9 in van der Vaart and Wellner [1996].

Für die dritte Bedingung muss die Klammerungszahl $N_{[]}(\eta, \mathcal{F}, L_n^2)$ bestimmt werden, wobei die Klammern in Abhängigkeit von n gewählt werden können und so bestimmt werden, dass jede Klammer $\mathcal{F}_{h,i,k,l} := [s_h, s_{h+1}) \times [t_i, t_{i+1}) \times [u_k, u_{k+1}) \times [v_l, v_{l+1})$

$$\sum_{j=1}^n E \left[\sup_{(s, t, u, v), (s', t', u', v') \in \mathcal{F}_{h,i,k,l}} |Z_{nj}(s, t, u, v) - Z_{nj}(s', t', u', v')|^2 \right] \leq \eta^2$$

erfüllt.

Es folgt mit (4.5.5) analog zum Ausmultiplizieren für die zweite Bedingung

$$\sum_{j=1}^n E \left[\sup_{(s, t, u, v), (s', t', u', v') \in \mathcal{F}_{h,i,k,l}} |Z_{nj}(s, t, u, v) - Z_{nj}(s', t', u', v')|^2 \right]$$

$$\begin{aligned}
&= \frac{1}{n} \sum_{j=1}^n E \left[\sup_{(s,t,u,v),(s',t',u',v') \in \mathcal{F}_{h,i,k,l}} n (Z_{nj}(s,t,u,v) - Z_{nj}(s',t',u',v'))^2 \right] \\
&\leq \frac{1}{n} \sum_{j=1}^n E \left[\sup_{t,t' \in [t_i, t_{i+1}]} |I\{\epsilon_j \leq t\} - I\{\epsilon_j \leq t'\}| \right] \\
&\quad + O \left(\sup_{u,u' \in [u_k, u_{k+1}]} |u - u'| + \sup_{v,v' \in [v_l, v_{l+1}]} |v - v'| + \sup_{s,s' \in [s_h, s_{h+1}]} |s - s'| + \frac{1}{n} \right) \\
&\leq \frac{1}{n} \sum_{j=1}^n E [I\{\epsilon_j \leq t_{i+1}\} - I\{\epsilon_j \leq t_i\}] \\
&\quad + O \left(\sup_{u,u' \in [u_k, u_{k+1}]} |u - u'| + \sup_{v,v' \in [v_l, v_{l+1}]} |v - v'| + \sup_{s,s' \in [s_h, s_{h+1}]} |s - s'| + \frac{1}{n} \right) \\
&= (F(t_{i+1}) - F(t_i)) + O \left(\sup_{u,u' \in [u_k, u_{k+1}]} |u - u'| + \sup_{v,v' \in [v_l, v_{l+1}]} |v - v'| + \sup_{s,s' \in [s_h, s_{h+1}]} |s - s'| + \frac{1}{n} \right).
\end{aligned}$$

Ist $\frac{1}{n} \leq \eta^2$ und werden die Klammern so gewählt, dass $s_{h+1} - s_h \leq \eta^2$, $u_{k+1} - u_k \leq \eta^2$, $v_{l+1} - v_l \leq \eta^2$ und $F(t_{i+1}) - F(t_i) \leq \eta^2$, dann gilt damit

$$\sum_{j=1}^n E \left[\sup_{(s,t,u,v),(s',t',u',v') \in \mathcal{F}_{h,i,k,l}} |Z_{nj}(s,t,u,v) - Z_{nj}(s',t',u',v')|^2 \right] = O(\eta^2),$$

die Anzahl der Klammern, die nötig sind, um \mathcal{F} zu überdecken, ist also $O((\eta^{-2})^4) = O(\eta^{-8})$.

Ist $\frac{1}{n} > \eta^2$, so müssen die Klammern für $s \in [0, 1]$ anders gewählt werden. Wählt man die Intervallgrenzen $\{\frac{0}{n}, \frac{1}{n}, \dots, \frac{n}{n}\}$, so werden $n < \eta^{-2}$ Klammern zur Überdeckung von $[0, 1]$ benötigt, und für alle Klammern $[\tilde{s}_h, \tilde{s}_{h+1})$ gilt

$$\sup_{s,s' \in [\tilde{s}_h, \tilde{s}_{h+1})} |[ns] - [ns']| = 0,$$

also wird die Bedingung

$$\sum_{j=1}^n E \left[\sup_{(s,t,u,v),(s',t',u',v') \in \tilde{\mathcal{F}}_{h,i,k,l}} |Z_{nj}(s,t,u,v) - Z_{nj}(s',t',u',v')|^2 \right] = O(\eta^2)$$

mit $\tilde{\mathcal{F}}_{h,i,k,l} := [\tilde{s}_h, \tilde{s}_{h+1}) \times [t_i, t_{i+1}) \times [u_k, u_{k+1}) \times [v_l, v_{l+1})$ für $\frac{1}{n} > \eta^2$ ebenfalls erfüllt.

Damit sind in beiden Fällen $O(\eta^{-8})$ Klammern zur Überdeckung nötig.

Weiter gilt

$$\int_0^{\bar{\epsilon}_n} \sqrt{\log(N_{[]}(\eta, \mathcal{F}, L_n^2))} d\eta = O \left(\int_0^{\bar{\epsilon}_n} \sqrt{\log\left(\frac{1}{\eta}\right)} d\eta \right) \xrightarrow{n \rightarrow \infty} 0 \quad \text{für } \bar{\epsilon}_n \searrow 0,$$

was gerade die dritte Bedingung in Theorem 2.11.9 von der Vaart and Wellner [1996] ist.

Damit sind alle drei Bedingungen erfüllt.

Das Theorem besagt nun, dass die schwache Konvergenz von Term (4.5.3) $(\sum_{j=1}^n (Z_{nj} - E[Z_{nj}]))$ aus der Konvergenz der endlich dimensional Randverteilungen folgt.

Um zu zeigen, dass die endlich dimensionalen Randverteilungen konvergieren, werde zunächst eine neue Notation eingeführt. Es sei $(\tilde{\mathbb{K}}_F(s, t, u, v))_{(s,t,u,v) \in \mathcal{F}}$ der Grenzprozess von Term (4.5.3), also

$$\left(\sum_{j=1}^n (Z_{nj}(s, t, u, v) - E[Z_{nj}(s, t, u, v)]) \right)_{(s,t,u,v) \in \mathcal{F}} \Rightarrow (\tilde{\mathbb{K}}_F(s, t, u, v))_{(s,t,u,v) \in \mathcal{F}}.$$

Damit bleibt also zu zeigen, dass

$$\left(\sum_{j=1}^n (Z_{nj}(s_1, t_1, u_1, v_1) - E[Z_{nj}(s_1, t_1, u_1, v_1)]), \dots, \sum_{j=1}^n (Z_{nj}(s_k, t_k, u_k, v_k) - E[Z_{nj}(s_k, t_k, u_k, v_k)]) \right) \xrightarrow{\mathcal{D}} (\tilde{\mathbb{K}}_F(s_1, t_1, u_1, v_1), \dots, \tilde{\mathbb{K}}_F(s_k, t_k, u_k, v_k)) \quad (4.5.7)$$

für alle $k \in \mathbb{N}$, $(s_1, t_1, u_1, v_1), \dots, (s_k, t_k, u_k, v_k) \in \mathcal{F}$.

Dies ist nach dem Satz von Cramér-Wold gleichbedeutend mit

$$\sum_{i=1}^k \lambda_i \sum_{j=1}^n (Z_{nj}(s_i, t_i, u_i, v_i) - E[Z_{nj}(s_i, t_i, u_i, v_i)]) \xrightarrow{\mathcal{D}} \sum_{i=1}^k \lambda_i \tilde{\mathbb{K}}_F(s_i, t_i, u_i, v_i) \quad (4.5.8)$$

für alle $(\lambda_1, \dots, \lambda_k) \in \mathbb{R}^k$, $k \in \mathbb{N}$, $(s_1, t_1, u_1, v_1), \dots, (s_k, t_k, u_k, v_k) \in \mathcal{F}$.

Dass (4.5.8) gilt wird bewiesen, indem gezeigt wird, dass die Folge

$$\left(\sum_{i=1}^k \lambda_i (Z_{nj}(s_i, t_i, u_i, v_i) - E[Z_{nj}(s_i, t_i, u_i, v_i)]) \right)_{j=1, \dots, n}$$

die Lindeberg-Bedingung erfüllt für alle $(\lambda_1, \dots, \lambda_k) \in \mathbb{R}^k$, $k \in \mathbb{N}$, $(s_1, t_1, u_1, v_1), \dots, (s_k, t_k, u_k, v_k) \in \mathcal{F}$.

Es ist für alle $\eta > 0$ und alle $(\lambda_1, \dots, \lambda_k) \in \mathbb{R}^k$, $k \in \mathbb{N}$, $(s_1, t_1, u_1, v_1), \dots, (s_k, t_k, u_k, v_k) \in \mathcal{F}$

$$\begin{aligned} & \sum_{j=1}^n E \left[\left(\sum_{i=1}^k \lambda_i (Z_{nj}(s_i, t_i, u_i, v_i) - E[Z_{nj}(s_i, t_i, u_i, v_i)]) \right)^2 \right. \\ & \quad \cdot I \left\{ \left| \sum_{i=1}^k \lambda_i (Z_{nj}(s_i, t_i, u_i, v_i) - E[Z_{nj}(s_i, t_i, u_i, v_i)]) \right| > \eta \right\} \\ & \leq \sum_{j=1}^n E \left[\sup_{(s,t,u,v) \in \mathcal{F}} |Z_{nj}(s, t, u, v) - E[Z_{nj}(s, t, u, v)]|^2 \left(\sum_{i=1}^k |\lambda_i| \right)^2 \right. \\ & \quad \cdot I \left\{ \sup_{(s,t,u,v) \in \mathcal{F}} |Z_{nj}(s, t, u, v) - E[Z_{nj}(s, t, u, v)]| \sum_{i=1}^k |\lambda_i| > \eta \right\} \\ & \leq \sum_{j=1}^n E \left[n^{-1} \left(1 + \sup_{t \in \mathbb{R}} f(t) |\epsilon_j| + \sup_{t \in \mathbb{R}} |f(t)t| |\epsilon_j^2 - 1| \right)^2 \left(\sum_{i=1}^k |\lambda_i| \right)^2 \right. \\ & \quad \cdot I \left\{ n^{-\frac{1}{2}} \left(1 + \sup_{t \in \mathbb{R}} f(t) |\epsilon_j| + \sup_{t \in \mathbb{R}} |f(t)t| |\epsilon_j^2 - 1| \right) \sum_{i=1}^k |\lambda_i| > \eta \right\} \end{aligned}$$

$$\begin{aligned}
&= \left(\sum_{i=1}^k |\lambda_i| \right)^2 E \left[\left(1 + \sup_{t \in \mathbb{R}} f(t) |\epsilon_1| + \sup_{t \in \mathbb{R}} |f(t)t| |\epsilon_1^2 - 1| \right)^2 \right. \\
&\quad \left. \cdot I \left\{ \left(1 + \sup_{t \in \mathbb{R}} f(t) |\epsilon_1| + \sup_{t \in \mathbb{R}} |f(t)t| |\epsilon_1^2 - 1| \right) \sum_{i=1}^k |\lambda_i| > \eta n^{\frac{1}{2}} \right\} \right] \\
&= o(1), \quad \text{für } n \rightarrow \infty \text{ (siehe (4.5.4))}
\end{aligned}$$

womit die Lindeberg-Bedingung für alle $(\lambda_1, \dots, \lambda_k) \in \mathbb{R}^k$, $k \in \mathbb{N}$, $(s_1, t_1, u_1, v_1), \dots, (s_k, t_k, u_k, v_k) \in \mathcal{F}$ erfüllt ist, und somit nach dem Zentralen Grenzwertsatz (4.5.8) und damit auch (4.5.7), also die Konvergenz der endlich dimensionalen Randverteilungen gilt.

Aus Theorem 2.11.9 in van der Vaart and Wellner [1996] folgt daher, dass die schwache Konvergenz von Term (4.5.3) gilt und somit die schwache Konvergenz in der Behauptung des Satzes.

Die Kovarianzfunktion ergibt sich aus

$$\begin{aligned}
&\text{Cov} \left(\sqrt{n} \left(\frac{1}{n} \sum_{j=1}^{\lfloor ns_1 \rfloor} (I \{\epsilon_j \leq t_1\} - F(t_1)) + \frac{\lfloor ns_1 \rfloor}{n} f(t_1) \frac{1}{n} \sum_{j=1}^n \epsilon_j + \frac{\lfloor ns_1 \rfloor}{n} f(t_1) t_1 \frac{1}{2n} \sum_{j=1}^n (\epsilon_j^2 - 1) \right), \right. \\
&\quad \left. \sqrt{n} \left(\frac{1}{n} \sum_{j=1}^{\lfloor ns_2 \rfloor} (I \{\epsilon_j \leq t_2\} - F(t_2)) + \frac{\lfloor ns_2 \rfloor}{n} f(t_2) \frac{1}{n} \sum_{j=1}^n \epsilon_j + \frac{\lfloor ns_2 \rfloor}{n} f(t_2) t_2 \frac{1}{2n} \sum_{j=1}^n (\epsilon_j^2 - 1) \right) \right) \\
&= E \left[\sqrt{n} \left(\frac{1}{n} \sum_{j=1}^{\lfloor ns_1 \rfloor} (I \{\epsilon_j \leq t_1\} - F(t_1)) + \frac{\lfloor ns_1 \rfloor}{n} f(t_1) \frac{1}{n} \sum_{j=1}^n \epsilon_j + \frac{\lfloor ns_1 \rfloor}{n} f(t_1) t_1 \frac{1}{2n} \sum_{j=1}^n (\epsilon_j^2 - 1) \right) \right. \\
&\quad \left. \cdot \sqrt{n} \left(\frac{1}{n} \sum_{j=1}^{\lfloor ns_2 \rfloor} (I \{\epsilon_j \leq t_2\} - F(t_2)) + \frac{\lfloor ns_2 \rfloor}{n} f(t_2) \frac{1}{n} \sum_{j=1}^n \epsilon_j + \frac{\lfloor ns_2 \rfloor}{n} f(t_2) t_2 \frac{1}{2n} \sum_{j=1}^n (\epsilon_j^2 - 1) \right) \right] \\
&= (\star)
\end{aligned}$$

und weiter durch Ausmultiplizieren, da ϵ_i, ϵ_j unabhängig sind für $i \neq j$, $E[\epsilon_j] = 0 \forall j$, $E[\epsilon_j^2 - 1] = 0 \forall j$ und $E[I \{\epsilon_j \leq t\}] = F(t) \forall j, t$

$$\begin{aligned}
(\star) &= \frac{1}{n} \sum_{j=1}^{\min(\lfloor ns_1 \rfloor, \lfloor ns_2 \rfloor)} (E[I \{\epsilon_j \leq t_1\} I \{\epsilon_j \leq t_2\}] - F(t_1)F(t_2)) \\
&\quad + \frac{\lfloor ns_2 \rfloor}{n} f(t_2) \frac{1}{n} \sum_{j=1}^{\lfloor ns_1 \rfloor} E[I \{\epsilon_j \leq t_1\} \epsilon_j] \\
&\quad + \frac{\lfloor ns_2 \rfloor}{n} f(t_2) t_2 \frac{1}{n} \sum_{j=1}^{\lfloor ns_1 \rfloor} E[I \{\epsilon_j \leq t_1\} (\epsilon_j^2 - 1)] \\
&\quad + \frac{\lfloor ns_1 \rfloor}{n} f(t_1) \frac{1}{n} \sum_{j=1}^{\lfloor ns_2 \rfloor} E[\epsilon_j I \{\epsilon_j \leq t_2\}] \\
&\quad + \frac{\lfloor ns_1 \rfloor}{n} \frac{\lfloor ns_2 \rfloor}{n} f(t_1) f(t_2) \frac{1}{n} \sum_{j=1}^n E[\epsilon_j^2]
\end{aligned}$$

$$\begin{aligned}
& + \frac{\lfloor ns_1 \rfloor}{n} \frac{\lfloor ns_2 \rfloor}{n} f(t_1) f(t_2) t_2 \frac{1}{n} \sum_{j=1}^n E[\epsilon_j^3] \\
& + \frac{\lfloor ns_1 \rfloor}{n} f(t_1) t_1 \frac{1}{n} \sum_{j=1}^{\lfloor ns_2 \rfloor} E[(\epsilon_j^2 - 1) I\{\epsilon_j \leq t_2\}] \\
& + \frac{\lfloor ns_1 \rfloor}{n} \frac{\lfloor ns_2 \rfloor}{n} f(t_1) t_1 f(t_2) \frac{1}{n} \sum_{j=1}^n E[\epsilon_j^3] \\
& + \frac{\lfloor ns_1 \rfloor}{n} \frac{\lfloor ns_2 \rfloor}{n} f(t_1) t_1 f(t_2) \frac{1}{n} \sum_{j=1}^n (E[\epsilon_j^4] - 2 E[\epsilon_j^2] + 1) \\
= & \frac{\min(\lfloor ns_1 \rfloor, \lfloor ns_2 \rfloor)}{n} (F(\min(t_1, t_2)) - F(t_1)F(t_2)) \\
& + \frac{\lfloor ns_1 \rfloor}{n} \frac{\lfloor ns_2 \rfloor}{n} \left(f(t_1) (E[\epsilon_1 I\{\epsilon_1 \leq t_2\}] + t_1 E[(\epsilon_1^2 - 1) I\{\epsilon_1 \leq t_2\}]) \right. \\
& \quad + f(t_2) (E[\epsilon_1 I\{\epsilon_1 \leq t_1\}] + t_2 E[(\epsilon_1^2 - 1) I\{\epsilon_1 \leq t_1\}]) \\
& \quad \left. + f(t_1) f(t_2) (1 + (t_1 + t_2) E[\epsilon_1^3] + t_1 t_2 (E[\epsilon_1^4] - 1)) \right),
\end{aligned}$$

da alle ϵ_j dieselbe Verteilung haben und $E[\epsilon_1^2] = 1$.

Für $n \rightarrow \infty$ gilt

$$\frac{\lfloor ns \rfloor}{n} = \frac{ns}{n} - O\left(\frac{1}{n}\right) \xrightarrow{n \rightarrow \infty} s,$$

damit gilt die behauptete asymptotische Kovarianz.

□

5 Test auf einen Change-Point in der Innovationen-Verteilung

In diesem Kapitel wird ein Change-Point-Test vorgestellt, der darauf testet, ob sich die Verteilung der Innovationen zu einem bestimmten Zeitpunkt ändert.

Solche Tests können z.B. in der Finanzwirtschaft von Interesse sein, um auf einen Change-Point in der Verteilung der Rendite zu testen, siehe Andreou and Ghysels [2009] Abschnitt 4.5.

5.1 Definition der Teststatistik

Ein sogenannter Change-Point in der Verteilung der Innovationen liegt vor, wenn es ein $\theta_0 \in (0, 1)$ gibt, so dass

$$\epsilon_1, \dots, \epsilon_{\lfloor n\theta_0 \rfloor} \sim F, \quad \epsilon_{\lfloor n\theta_0 \rfloor + 1}, \dots, \epsilon_n \sim \tilde{F}, \quad F \neq \tilde{F}. \quad (5.0.1)$$

Um zu testen, ob ein solcher Change-Point vorliegt, wird ein Hypothesentest durchgeführt für die Nullhypothese

$$H_0 : \quad \epsilon_1, \dots, \epsilon_n \sim F$$

gegen die Alternative (5.0.1)

$$H_1 : \quad \exists \theta_0 \in (0, 1) : \quad \epsilon_1, \dots, \epsilon_{\lfloor n\theta_0 \rfloor} \sim F, \quad \epsilon_{\lfloor n\theta_0 \rfloor + 1}, \dots, \epsilon_n \sim \tilde{F}, \quad F \neq \tilde{F}.$$

Der sequentielle empirische Prozess

$$\hat{F}_{\lfloor n\theta_0 \rfloor}(t) = \sum_{j=1}^{\lfloor n\theta_0 \rfloor} \bar{w}_{j\lfloor n\theta_0 \rfloor} I\{\hat{\epsilon}_j \leq t\},$$

wie in Abschnitt 4.1 (S. 45) definiert, ist ein Schätzer für $F(t)$, sowohl unter der Nullhypothese, als auch unter der Alternative.

Der sequentielle empirische Prozess

$$\hat{F}_{n-\lfloor n\theta_0 \rfloor}^*(t) := \sum_{j=\lfloor n\theta_0 \rfloor + 1}^n \frac{w_{nj}}{\sum_{k=\lfloor n\theta_0 \rfloor + 1}^n w_{nk}} I\{\hat{\epsilon}_j \leq t\}$$

ist unter der Nullhypothese ebenfalls ein Schätzer für $F(t)$, unter der Alternative allerdings ein Schätzer für $\tilde{F}(t)$.

Daher betrachtet man die Differenz von beiden, um zu überprüfen, ob ein Change-Point vorliegt. Der der Teststatistik zugrundeliegende Prozess hat also die Form

$$\hat{T}_n(s, t) := \sqrt{n} \cdot \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \frac{\sum_{k=\lfloor ns \rfloor + 1}^n w_{nk}}{n} \left(\hat{F}_{\lfloor ns \rfloor}(t) - \hat{F}_{n-\lfloor ns \rfloor}^*(t) \right)$$

wobei der Vorfaktor wieder, wie in Abschnitt 4.1, die Division durch Null bei Aussagen gleichmäßig über alle $s \in [0, 1]$ verhindern soll.

Eine auf diesem Prozess aufbauende Teststatistik ist die Kolmogorov-Smirnov Teststatistik

$$T_{KS}(X_0, \dots, X_n) := \sup_{s \in [0, 1], t \in \mathbb{R}} \left| \hat{T}_n(s, t) \right|.$$

Um einen Test zu einem bestimmten Niveau aufstellen zu können, wird die Verteilung der Teststatistik unter der Nullhypothese benötigt. Im nächsten Abschnitt wird gezeigt, dass die hier vorgestellte Teststatistik unter H_0 asymptotisch verteilungsfrei ist, was bedeutet, dass die Verteilung der Teststatistik nicht von der Verteilung der dem Test zugrundeliegenden Zufallsvariablen X_0, \dots, X_n abhängt.

5.2 Verteilung der Teststatistik unter H_0

Ein Hypothesentest mit einer Teststatistik T und einem kritischen Wert L hat ein Niveau γ , wenn gilt

$$P\left(T > L \mid H_0\right) \leq \gamma.$$

Eine Folge von Hypothesentests mit Teststatistiken $(T_n)_{n \in \mathbb{N}}$ und kritischem Wert L hat asymptotisches Niveau γ , wenn gilt

$$\limsup_{n \rightarrow \infty} P\left(T_n > L \mid H_0\right) \leq \gamma.$$

Das Niveau ist eine wichtige Kenngröße eines Tests, da es die Wahrscheinlichkeit für einen Fehler erster Art, also für eine Entscheidung gegen H_0 obwohl H_0 gilt, beschränkt. Um einen Test zu einem gegebenen (asymptotischen) Niveau γ aufstellen zu können, benötigt man also die (asymptotische) Verteilung der Teststatistik unter der Nullhypothese, um die Größe des kritischen Werts L bestimmen zu können.

Satz 5.1 Unter den Voraussetzungen von Satz 4.2 (S. 49), und damit unter der Nullhypothese H_0 dass kein Change-Point vorliegt, gilt für die in Abschnitt 5.1 vorgestellte Teststatistik T_{KS}

$$T_{KS}(X_0, \dots, X_n) \xrightarrow[n \rightarrow \infty]{\mathcal{D}} \sup_{s \in [0, 1], z \in [0, 1]} |\mathbb{G}(s, z)|,$$

wobei $(\mathbb{G}(s, z))_{s \in [0, 1], z \in [0, 1]}$ ein sogenannter Gaußprozess, also ein stochastischer Prozess mit normalverteilten Randverteilungen, ist, für den gilt

$$E[\mathbb{G}(s, z)] = 0 \quad \forall s \in [0, 1], z \in [0, 1]$$

und

$$\text{Cov}(\mathbb{G}(s_1, z_1), \mathbb{G}(s_2, z_2)) = (\min(s_1, s_2) - s_1 s_2)(\min(z_1, z_2) - z_1 z_2).$$

Bemerkung 5.2 Damit ist die Verteilung von \mathbb{G} unabhängig von der Verteilung der Zufallsvariablen X_0, \dots, X_n und $\epsilon_1, \dots, \epsilon_n$. Dies ist eine für die Anwendung des Tests vorteilhafte und ungewöhnliche Eigenschaft, da die Verteilung von Tests, die auf empirischen Prozessen von nichtparametrisch geschätzten Residuen basieren, normalerweise stark von der unbekanntenen Verteilung der Innovationen abhängt. (Siehe z.B. Neumeyer and van Keilegom [2009] Theoreme 2, 3 & 4; van Keilegom et al. [2008] Korollar 3.2, Dette et al. [2007] Korollar 3.3 oder Neumeyer and Dette [2007] Korollar 3.2)

Bemerkung 5.3 Satz 5.1 sagt aus, dass die asymptotische Verteilung der Teststatistik T_{KS} unter H_0 der Verteilung von $\sup |\mathbb{G}|$ entspricht. Für die Verteilung von $\sup |\mathbb{G}|$ wurde in Picard [1985] eine Tabelle erstellt, der man die Größen des kritischen Werts L für das jeweils gewünschte asymptotische Niveau γ entnehmen kann.

Für den Beweis von Satz 5.1 wird die folgende Konvergenzrate für den sequentiellen empirischen Prozess der Innovationen benötigt.

Lemma 5.4 Für den sequentiellen empirischen Prozess der Innovationen gilt, wenn die Innovationen i.i.d. sind,

$$\sup_{s \in [0,1], t \in \mathbb{R}} \left| \frac{1}{n} \sum_{j=1}^{\lfloor ns \rfloor} (I\{\epsilon_j \leq t\} - F(t)) \right| = O_P \left(\frac{1}{\sqrt{n}} \right).$$

Beweis Nach Theorem 2.12.1 in van der Vaart and Wellner [1996] konvergiert $\frac{1}{\sqrt{n}} \sum_{j=1}^{\lfloor ns \rfloor} (I\{\epsilon_j \leq t\} - F(t))$ schwach gegen eine straffe Zufallsfunktion $\mathbb{K}(s, t)$, einen sogenannten Kiefer-Prozess, da nach Example 2.1.3 in van der Vaart and Wellner [1996] die Funktionenklasse $\{I\{\cdot \leq t\}, t \in \mathbb{R}\}$ Donsker ist.

Dass die Zufallsfunktion $\mathbb{K}(s, t)$ straff ist bedeutet, dass $\sup_{s \in [0,1], t \in \mathbb{R}} |\mathbb{K}(s, t)| = O_P(1)$.

Aus der schwachen Konvergenz von $\frac{1}{\sqrt{n}} \sum_{j=1}^{\lfloor ns \rfloor} (I\{\epsilon_j \leq t\} - F(t))$ folgt

$$\sup_{s \in [0,1], t \in \mathbb{R}} \left| \frac{1}{\sqrt{n}} \sum_{j=1}^{\lfloor ns \rfloor} (I\{\epsilon_j \leq t\} - F(t)) \right| \xrightarrow[n \rightarrow \infty]{\mathcal{D}} \sup_{s \in [0,1], t \in \mathbb{R}} |\mathbb{K}(s, t)|$$

und damit $\sup_{s \in [0,1], t \in \mathbb{R}} \left| \frac{1}{\sqrt{n}} \sum_{j=1}^{\lfloor ns \rfloor} (I\{\epsilon_j \leq t\} - F(t)) \right| = O_P(1)$, also

$$\sup_{s \in [0,1], t \in \mathbb{R}} \left| \frac{1}{n} \sum_{j=1}^{\lfloor ns \rfloor} (I\{\epsilon_j \leq t\} - F(t)) \right| = O_P \left(\frac{1}{\sqrt{n}} \right).$$

Anmerkung: Die schwache Konvergenz des sequentiellen empirischen Prozesses der Innovationen gegen einen Kiefer-Prozess wurde ebenfalls in Theorem 1, Kapitel 3.5 in Shorack and Wellner [1986] gezeigt, das auf der Arbeit Bickel and Wichura [1971] basiert.

□

Beweis von Satz 5.1 Für den Prozess \hat{T}_n gilt

$$\hat{T}_n(s, t)$$

$$\begin{aligned}
&= \sqrt{n} \cdot \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \frac{\sum_{k=\lfloor ns \rfloor+1}^n w_{nk}}{n} \left(\hat{F}_{\lfloor ns \rfloor}(t) - \hat{F}_{n-\lfloor ns \rfloor}^*(t) \right) \\
&= \sqrt{n} \left(\frac{\sum_{k=\lfloor ns \rfloor+1}^n w_{nk}}{n} \frac{1}{n} \sum_{j=1}^{\lfloor ns \rfloor} w_{nj} I \{ \hat{\epsilon}_j \leq t \} - \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \frac{1}{n} \sum_{j=\lfloor ns \rfloor+1}^n w_{nj} I \{ \hat{\epsilon}_j \leq t \} \right) \\
&= \sqrt{n} \left(\left(\frac{\sum_{k=\lfloor ns \rfloor+1}^n w_{nk}}{n} + \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \right) \frac{1}{n} \sum_{j=1}^{\lfloor ns \rfloor} w_{nj} I \{ \hat{\epsilon}_j \leq t \} \right. \\
&\quad \left. - \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \left(\frac{1}{n} \sum_{j=\lfloor ns \rfloor+1}^n w_{nj} I \{ \hat{\epsilon}_j \leq t \} + \frac{1}{n} \sum_{j=1}^{\lfloor ns \rfloor} w_{nj} I \{ \hat{\epsilon}_j \leq t \} \right) \right) \\
&= \sqrt{n} \left(\frac{\sum_{k=1}^n w_{nk}}{n} \frac{1}{n} \sum_{j=1}^{\lfloor ns \rfloor} w_{nj} I \{ \hat{\epsilon}_j \leq t \} - \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \frac{1}{n} \sum_{j=1}^n w_{nj} I \{ \hat{\epsilon}_j \leq t \} \right) \\
&= \sqrt{n} \left(\frac{\sum_{k=1}^n w_{nk}}{n} \frac{1}{n} \sum_{j=1}^{\lfloor ns \rfloor} w_{nj} (I \{ \hat{\epsilon}_j \leq t \} - F(t)) - \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \frac{1}{n} \sum_{j=1}^n w_{nj} (I \{ \hat{\epsilon}_j \leq t \} - F(t)) \right) \\
&= \sqrt{n} \left(\frac{\sum_{k=1}^n w_{nk}}{n} \left(\frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} (\hat{F}_{\lfloor ns \rfloor}(t) - F(t)) \right) - \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \left(\frac{\sum_{k=1}^n w_{nk}}{n} (\hat{F}_n(t) - F(t)) \right) \right). \tag{5.4.1}
\end{aligned}$$

Mit Satz 4.2 (S. 49) gilt

$$\begin{aligned}
&\frac{\sum_{k=1}^n w_{nk}}{n} \left(\frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} (\hat{F}_{\lfloor ns \rfloor}(t) - F(t)) \right) \\
&= \frac{\sum_{k=1}^n w_{nk}}{n} \left(\frac{1}{n} \sum_{j=1}^{\lfloor ns \rfloor} (I \{ \epsilon_j \leq t \} - F(t)) + \frac{\lfloor ns \rfloor}{n} f(t) \frac{1}{n} \sum_{j=1}^n \epsilon_j + \frac{\lfloor ns \rfloor}{n} f(t) t \frac{1}{n} \sum_{j=1}^n (\epsilon_j^2 - 1) + o_p \left(\frac{1}{\sqrt{n}} \right) \right) \\
&= \frac{1}{n} \sum_{j=1}^{\lfloor ns \rfloor} (I \{ \epsilon_j \leq t \} - F(t)) + \frac{\lfloor ns \rfloor}{n} f(t) \frac{1}{n} \sum_{j=1}^n \epsilon_j + \frac{\lfloor ns \rfloor}{n} f(t) t \frac{1}{n} \sum_{j=1}^n (\epsilon_j^2 - 1) + o_p \left(\frac{1}{\sqrt{n}} \right) \\
&\quad + \left(\frac{\sum_{k=1}^n w_{nk}}{n} - 1 \right) \\
&\quad \cdot \left(\frac{1}{n} \sum_{j=1}^{\lfloor ns \rfloor} (I \{ \epsilon_j \leq t \} - F(t)) + \frac{\lfloor ns \rfloor}{n} f(t) \frac{1}{n} \sum_{j=1}^n \epsilon_j + \frac{\lfloor ns \rfloor}{n} f(t) t \frac{1}{n} \sum_{j=1}^n (\epsilon_j^2 - 1) + o_p \left(\frac{1}{\sqrt{n}} \right) \right). \tag{5.4.2}
\end{aligned}$$

Nach Lemma 9.2 (S. 130) ist $\left(\frac{\sum_{k=1}^n w_{nk}}{n} - 1 \right) = o_P(1)$.

Da $\epsilon_1, \dots, \epsilon_n$ i.i.d. sind und $E[\epsilon_i] = 0$ und $E[\epsilon_i^4] < \infty$ für alle i gilt, ist $\frac{1}{n} \sum_{j=1}^n \epsilon_j = O_P \left(\frac{1}{\sqrt{n}} \right)$, sowie $\frac{1}{n} \sum_{j=1}^n (\epsilon_j^2 - 1) = O_P \left(\frac{1}{\sqrt{n}} \right)$.

Außerdem gilt nach Lemma 5.4 $\sup_{s \in [0,1], t \in \mathbb{R}} \left| \frac{1}{n} \sum_{j=1}^{\lfloor ns \rfloor} (I \{ \epsilon_j \leq t \} - F(t)) \right| = O_P \left(\frac{1}{\sqrt{n}} \right)$.

Damit folgt aus (5.4.2) gleichmäßig über alle $s \in [0, 1]$ und $t \in \mathbb{R}$ (mit Lemma 9.2 und Lemma 5.4)

$$\begin{aligned} & \frac{\sum_{k=1}^n w_{nk}}{n} \left(\frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \left(\hat{F}_{\lfloor ns \rfloor}(t) - F(t) \right) \right) \\ &= \frac{1}{n} \sum_{j=1}^{\lfloor ns \rfloor} (I\{\epsilon_j \leq t\} - F(t)) + \frac{\lfloor ns \rfloor}{n} f(t) \frac{1}{n} \sum_{j=1}^n \epsilon_j + \frac{\lfloor ns \rfloor}{n} f(t) t \frac{1}{n} \sum_{j=1}^n (\epsilon_j^2 - 1) + o_p\left(\frac{1}{\sqrt{n}}\right), \end{aligned}$$

und genauso

$$\begin{aligned} & \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \left(\frac{\sum_{k=1}^n w_{nk}}{n} \left(\hat{F}_n(t) - F(t) \right) \right) \\ &= \frac{\lfloor ns \rfloor}{n} \left(\frac{1}{n} \sum_{j=1}^n (I\{\epsilon_j \leq t\} - F(t)) + f(t) \frac{1}{n} \sum_{j=1}^n \epsilon_j + f(t) t \frac{1}{n} \sum_{j=1}^n (\epsilon_j^2 - 1) + o_p\left(\frac{1}{\sqrt{n}}\right) \right). \end{aligned}$$

Damit folgt aus (5.4.1)

$$\begin{aligned} \hat{T}_n(s, t) &= \sqrt{n} \left(\frac{1}{n} \sum_{j=1}^{\lfloor ns \rfloor} (I\{\epsilon_j \leq t\} - F(t)) - \frac{\lfloor ns \rfloor}{n} \frac{1}{n} \sum_{j=1}^n (I\{\epsilon_j \leq t\} - F(t)) + o_p\left(\frac{1}{\sqrt{n}}\right) \right) \\ &= \sqrt{n} \left(\frac{1}{n} \sum_{j=1}^{\lfloor ns \rfloor} I\{\epsilon_j \leq t\} - \frac{\lfloor ns \rfloor}{n} \frac{1}{n} \sum_{j=1}^n I\{\epsilon_j \leq t\} + o_p\left(\frac{1}{\sqrt{n}}\right) \right) \\ &= \sqrt{n} \left(\left(1 - \frac{\lfloor ns \rfloor}{n}\right) \frac{1}{n} \sum_{j=1}^{\lfloor ns \rfloor} I\{\epsilon_j \leq t\} - \frac{\lfloor ns \rfloor}{n} \frac{1}{n} \sum_{j=\lfloor ns \rfloor+1}^n I\{\epsilon_j \leq t\} + o_p\left(\frac{1}{\sqrt{n}}\right) \right) \\ &= \sqrt{n} \frac{\lfloor ns \rfloor}{n} \left(1 - \frac{\lfloor ns \rfloor}{n}\right) \left(\frac{1}{\lfloor ns \rfloor} \sum_{j=1}^{\lfloor ns \rfloor} I\{\epsilon_j \leq t\} - \frac{1}{n - \lfloor ns \rfloor} \sum_{j=\lfloor ns \rfloor+1}^n I\{\epsilon_j \leq t\} \right) + o_P(1) \end{aligned} \tag{5.4.3}$$

gleichmäßig über alle $s \in [0, 1]$ und alle $t \in \mathbb{R}$.

Nach Neumeyer and van Keilegom [2009] (Remark 2) gilt

$$\begin{aligned} & \sup_{s \in [0, 1], t \in \mathbb{R}} \left| \sqrt{n} \frac{\lfloor ns \rfloor}{n} \left(1 - \frac{\lfloor ns \rfloor}{n}\right) \left(\frac{1}{\lfloor ns \rfloor} \sum_{j=1}^{\lfloor ns \rfloor} I\{\epsilon_j \leq t\} - \frac{1}{n - \lfloor ns \rfloor} \sum_{j=\lfloor ns \rfloor+1}^n I\{\epsilon_j \leq t\} \right) - \mathbb{G}_n(s, F(t)) \right| \\ &= o_P(1), \end{aligned}$$

wobei $((\mathbb{G}_n(s, F(t)))_{s \in [0, 1], t \in \mathbb{R}})_{n \in \mathbb{N}}$ eine Folge von Gaußprozessen ist, die für alle $n \geq 1$ in Verteilung gleich ist mit $(\mathbb{G}(s, F(t)))_{s \in [0, 1], t \in \mathbb{R}}$, wobei \mathbb{G} wie in Satz 5.1 definiert ist.

Diese Gleichung lässt sich ableiten aus Theorem 3.1 in Csörgö et al. [1997].

Daraus folgt mit (5.4.3)

$$\sup_{s \in [0, 1], t \in \mathbb{R}} \left| \hat{T}_n(s, t) \right| \xrightarrow[n \rightarrow \infty]{\mathcal{D}} \sup_{s \in [0, 1], t \in \mathbb{R}} |\mathbb{G}(s, F(t))|,$$

und damit die Behauptung, denn

$$T_{KS}(X_0, \dots, X_n) = \sup_{s \in [0,1], t \in \mathbb{R}} \left| \hat{T}_n(s, t) \right| = \sup_{s \in [0,1], z \in [0,1]} \left| \hat{T}_n(s, F^{-1}(z)) \right| \xrightarrow[n \rightarrow \infty]{\mathcal{D}} \sup_{s \in [0,1], z \in [0,1]} |\mathbb{G}(s, z)|.$$

□

5.3 Konsistenz des Tests

Ein Test heißt konsistent, wenn

$$P \left(\text{„Der Test entscheidet für } H_1 \text{“} \middle| H_1 \right) \xrightarrow[n \rightarrow \infty]{} 1,$$

d.h. wenn die Wahrscheinlichkeit für einen Fehler zweiter Art, also dafür, dass die Nullhypothese nicht verworfen wird, obwohl die Alternative gilt, mit $n \rightarrow \infty$ gegen Null geht.

Für den in Abschnitt 5.1 vorgestellten Change-Point-Test bedeutet Konsistenz also, dass

$$P \left(\sup_{s \in [0,1], t \in \mathbb{R}} \left| \hat{T}_n(s, t) \right| > L \middle| H_1 \right) \xrightarrow[n \rightarrow \infty]{} 1,$$

wobei L durch das Niveau des Tests bestimmt ist, wie in Abschnitt 5.2 beschrieben.

Unter den folgenden Voraussetzungen gilt diese Konsistenz (vgl. Annahmen in Abschnitt 3.2, S. 19).

Annahme (K) Der Kern K habe den kompakten Träger $[-C, C]$. Er sei dreimal differenzierbar und $\sup_{u \in [-C, C]} |K^{(\mu)}(u)| \leq \bar{K}$, $\mu = 0, 1, 2, 3$. Außerdem gelte $K(C) = K(-C) = K'(C) = K'(-C) = 0$ und $\int K(u)udu = 0$.

Annahme (C') Für die Bandbreite c_n gebe es ein r_c mit $|r_c| < \infty$, so dass gelte $c_n \sim n^{-\frac{1}{4}} \log(n)^{r_c}$.

Annahme (P') Für das Intervall $I_n = [a_n, b_n]$, mit $a_n \xrightarrow[n \rightarrow \infty]{} -\infty$ und $b_n \xrightarrow[n \rightarrow \infty]{} \infty$, gebe es ein $r_I < \infty$, so dass gelte $(b_n - a_n) = O(\log(n)^{r_I})$. Außerdem gelte für alle $m_n \leq n$ mit $m_n^{-1} = o(1)$ $\frac{1}{m_n} \max_{0 \leq S \leq n - m_n} \sum_{i=S+1}^{S+m_n} \left(\int_{-\infty}^{a_n + \kappa} f_{X_{j-1}}(x)dx + \int_{b_n - \kappa}^{\infty} f_{X_{j-1}}(x)dx \right) = o\left(\frac{1}{\log(n)}\right)$, wobei $\kappa > 0$ in Abschnitt 4.1 (S. 45) definiert ist, und $f_{X_{j-1}}$ die Dichten der Beobachtungen bezeichnen.

Annahme (F') Für die Dichten f_{ϵ_j} der Innovationen gelte $\frac{1}{n} \sum_{j=1}^n \sup_{t \in \mathbb{R}} |f_{\epsilon_j}(t)| = O(1)$ und $\frac{1}{n} \sum_{j=1}^n \sup_{t \in \mathbb{R}} |f_{\epsilon_j}(t)t| = O(1)$ für $n \rightarrow \infty$.

Annahme (X') Der Prozess $(X_i)_{i \in \mathbb{Z}}$ der Beobachtungen sei α -mischend mit Mischungskoeffizienten, $\alpha(n) = O(n^{-\beta})$ für ein $\beta > 38$ (siehe Lemma 3.3, S. 21).

Die Dichten f_{X_i} der Beobachtungen seien viermal differenzierbar und es gelte

$$\sup_{x \in \mathbb{R}} \frac{1}{n} \sum_{i=1}^n \left| f_{X_{i-1}}^{(\mu)}(x) \right| = O(1), \mu = 0, 1, 2, 3, 4.$$

Außerdem gebe es ein $r_f < \infty$, so dass $\frac{1}{\inf_{x \in I_n} \frac{1}{n} \sum_{i=1}^n f_{X_{i-1}}(x)} =: q_n^f = O(\log(n)^{r_f})$.

Annahme (E') Für die $2b$ -ten Momente der X_1, \dots, X_n gelte $\frac{1}{n} \sum_{i=1}^n E[|X_i|^{2b}] = O(1)$ für ein $b > \frac{8}{3}, n \rightarrow \infty$.

Annahme (Z') Für alle $m_n \leq n$, mit $m_n^{-1} = o(1)$ gelte

$$\sup_{x \in J_n} \left((|m(x)| + |\sigma(x)|) \frac{1}{m_n} \max_{0 \leq S \leq n-m_n} \sum_{i=S+1}^{S+m_n} f_{X_{i-1}}(x) \right) = O(1),$$

$$\sup_{x \in J_n} \left((|m(x)| + |\sigma(x)|)^{2k} \frac{1}{m_n} \max_{0 \leq S \leq n-m_n} \sum_{i=S+1}^{S+m_n} (1 + E[\epsilon_i^4])^{k-1} f_{X_{i-1}}(x) \right) = O(1),$$

und es gebe ein $1 \leq j^* < \infty$, so dass

$$\sup_{x, x' \in J_n} \left((|m(x)| + |\sigma(x)|)^k (|m(x')| + |\sigma(x')|)^k \frac{1}{m_n^2} \max_{0 \leq S \leq n-m_n} \sum_{\substack{i, j=S+1 \\ |i-j| > j^*}}^{S+m_n} f_{X_{i-1}, X_{j-1}}(x, x') \right) = O(1) \text{ und}$$

$$\sup_{x \in J_n} \left((|m(x)| + |\sigma(x)|)^{2k} \max_{j^*+1 \leq i \leq n-j^*} \sum_{j=i-j^*}^{i+j^*} (1 + E[\epsilon_j^4])^{k-1} f_{X_{j-1}}(x) \right) = O(1)$$

für $k = 1, 2, n \rightarrow \infty$ mit $J_n := [a_n - (C + c_n^{-\frac{1}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}})c_n, b_n + (C + c_n^{-\frac{1}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}})c_n]$.

Annahme (M) Die Regressionsfunktion m und die Skalierungsfunktion σ seien viermal differenzierbar und es gebe $r_q, r_s < \infty$ und q_n, q_n^σ mit $(q_n)^{-1} = O(1), (q_n^\sigma)^{-1} = O(1)$, so dass gelte $\sup_{x \in [a_n - Cc_n, b_n + Cc_n]} |m^{(\mu)}(x)| = O(q_n), \sup_{x \in [a_n - Cc_n, b_n + Cc_n]} |\sigma^{(\mu)}(x)| = O(q_n), \mu = 0, 1, 2, 3, 4$ sowie $\frac{1}{\inf_{x \in I_n} |\sigma(x)|} = O(q_n^\sigma)$ mit $q_n = O(\log(n)^{r_q}), q_n^\sigma = O(\log(n)^{r_s})$.

Beispiel 5.5 Als Beispiel, welches diese Annahmen erfüllt, wird hier ein homoskedastisches Modell angegeben, obwohl in dieser Arbeit eigentlich heteroskedastische Modelle betrachtet werden. Die Begründung hierfür ist, dass ein Change-Point in einem heteroskedastischen Modell nicht durch eine Veränderung in der Varianz erfolgen kann, sondern nur durch Veränderungen in höheren Momenten, und daher zu komplizierteren Darstellungen führt. (Siehe Kapitel 6, insbesondere Abschnitt 6.3, S. 98)

Ein einfaches homoskedastisches Beispiel, das diese Annahmen erfüllt, ist also, analog zu Beispiel 3.1 (S. 20) und Beispiel 4.1 (S. 48) das AR(1)-Modell

$$X_j = 0.5 \cdot X_{j-1} + \epsilon_j$$

mit einem Change-Point zum Zeitpunkt $\lfloor \frac{n}{2} \rfloor$ in der Verteilung der Innovationen (ϵ_j) , so dass die Innovationen bis zum Change-Point Standardnormalverteilt sind, und nach dem Change-Point $\mathcal{N}(0, 2)$ -verteilt. Damit ist Annahme (F') erfüllt.

X_j hat in diesem Fall für alle $j = 1, \dots, \lfloor \frac{n}{2} \rfloor$ die Verteilung $\mathcal{N}\left(0, \frac{1}{1-0.5^2}\right)$ (da $X_j = \sum_{i=0}^{\infty} 0.5^i \epsilon_{j-i}$, siehe z.B. Beispiel 2.7 in Kreiß and Neuhaus [2006]). Auch nach dem Change-Point ist (X_j) langfristig stationär und hat (langfristig) die Verteilung $\mathcal{N}\left(0, \frac{2}{1-0.5^2}\right)$. Das arithmetische Mittel $\frac{1}{n} \sum_{i=1}^n f_{X_{i-1}}(x)$ konvergiert also gegen das arithmetische Mittel der Dichten von $\mathcal{N}\left(0, \frac{1}{1-0.5^2}\right)$ und $\mathcal{N}\left(0, \frac{2}{1-0.5^2}\right)$,

$$\text{also gegen } \frac{1}{\sqrt{2\pi}} \frac{\sqrt{\frac{3}{4}} \exp\left(-\frac{x^2}{\frac{3}{8}}\right) + \sqrt{\frac{3}{8}} \exp\left(-\frac{x^2}{\frac{16}{3}}\right)}{2}.$$

Damit erfüllt (X_j) die ersten beiden Punkte in Annahme (X') nach Abschnitt 2.2 (siehe (2.4.1), S. 13). Der dritte Punkt in Annahme (X') ist erfüllt, wenn

$$\max \left(\exp\left(\frac{a_n^2}{\frac{8}{3}}\right), \exp\left(\frac{b_n^2}{\frac{8}{3}}\right) \right) = O(\log(n)^{r_f}). \quad (5.5.1)$$

Eine weitere Forderung an a_n und b_n ergibt sich aus dem zweiten Punkt in Annahme (I'). Auch hier konvergiert das arithmetische Mittel über alle $i = S + 1, \dots, S + m_n$, und zwar je nach der

Gestalt von m_n gegen die Integrale über die Dichte von $\mathcal{N}\left(0, \frac{1}{1-0.5^2}\right)$, gegen die Integrale über die Dichte von $\mathcal{N}\left(0, \frac{2}{1-0.5^2}\right)$, oder gegen ein Mittel aus beiden. Der zweite Punkt in Annahme (I') ist also erfüllt, wenn

$$\left(\int_{-\infty}^{a_n+\kappa} \exp\left(-\frac{x^2}{\frac{16}{3}}\right) dx + \int_{b_n-\kappa}^{\infty} \exp\left(-\frac{x^2}{\frac{16}{3}}\right) dx \right) = o(\log(n)^{-1}) \quad (5.5.2)$$

Die Forderungen (5.5.1) und (5.5.2) sind erfüllt für $I_n = \left[-\sqrt{\frac{16}{3} \log(\log(n)^2)} - \kappa, \sqrt{\frac{16}{3} \log(\log(n)^2)} + \kappa\right]$, wie im Folgenden gezeigt wird:

Für Forderung (5.5.1) gilt

$$\begin{aligned} \exp\left(\frac{a_n^2}{\frac{8}{3}}\right) &= \exp\left(\frac{b_n^2}{\frac{8}{3}}\right) \\ &= \exp\left(\frac{\left(\sqrt{\frac{16}{3} \log(\log(n)^2)} + \kappa\right)^2}{\frac{8}{3}}\right) \\ &= \exp\left(\frac{\left(\sqrt{\frac{16}{3} \log(\log(n)^2)}\right)^2}{\frac{8}{3}}\right) \underbrace{\exp\left(\frac{2\sqrt{\frac{16}{3} \log(\log(n)^2)}\kappa}{\frac{8}{3}}\right)}_{=O\left(\exp\left(\frac{\left(\sqrt{\frac{16}{3} \log(\log(n)^2)}\right)^2}{\frac{16}{3}}\right)\right)} \underbrace{\exp\left(\frac{\kappa^2}{\frac{8}{3}}\right)}_{<\infty} \\ &= O\left(\left(\exp(2 \log(\log(n)^2))\right)^2\right) = O(\log(n)^8). \end{aligned}$$

Für Forderung (5.5.2) gilt

$$\begin{aligned} \int_{b_n-\kappa}^{\infty} \exp\left(-\frac{x^2}{\frac{16}{3}}\right) dx &= \int_{-\infty}^{a_n+\kappa} \exp\left(-\frac{x^2}{\frac{16}{3}}\right) dx \\ &= o(\log(n)^{-1}) \end{aligned}$$

analog zum Vorgehen in Beispiel 4.1 (S. 48).

Mit diesem I_n sind auch der erste Punkt in Annahme (I') und Annahme (M) erfüllt, wobei $q_n = O\left(\sqrt{\log(\log(n)^2)} + c_n\right) = O(\log(n))$ und $q_n^\sigma = 1$.

Analog zu den Überlegungen für das Mittel über die Dichten und den zweiten Punkt in (I') kann gefolgert werden, dass sich die Annahmen (E') und (Z') ebenfalls auf Annahmen "ohne Summe" reduzieren, die analog zum Vorgehen in Beispiel 4.1 behandelt werden können.

Mit einem Kern und einer Bandbreite, die die Bedingungen (K) und (C') erfüllen, sind alle Bedingungen für Satz 5.6 erfüllt.

Satz 5.6 Unter den Voraussetzungen (K), (C'), (I'), (F'), (X'), (E'), (Z'), (M) und H_1 , also einem Change-Point in $[n\theta_0]$, gilt

$$\sup_{t \in \mathbb{R}} \left| \frac{\sum_{k=1}^{\lfloor n\theta_0 \rfloor} w_{nk}}{n} \left(\hat{F}_{\lfloor n\theta_0 \rfloor}(t) - F(t) \right) \right| = o_P(1)$$

und

$$\sup_{t \in \mathbb{R}} \left| \frac{\sum_{k=\lfloor n\theta_0 \rfloor+1}^n w_{nk}}{n} \left(\hat{F}_{n-\lfloor n\theta_0 \rfloor}^*(t) - \tilde{F}(t) \right) \right| = o_P(1)$$

die Schätzer $\hat{F}_{\lfloor n\theta_0 \rfloor}$ und $\hat{F}_{n-\lfloor n\theta_0 \rfloor}^*$ sind also konsistent.

Korollar 5.7 Mit Satz 5.6 gilt unter H_1

$$\begin{aligned}
& \sup_{t \in \mathbb{R}} \left| \frac{\sum_{k=1}^{\lfloor n\theta_0 \rfloor} w_{nk}}{n} \frac{\sum_{k=\lfloor n\theta_0 \rfloor+1}^n w_{nk}}{n} \left(\hat{F}_{\lfloor n\theta_0 \rfloor}(t) - \hat{F}_{n-\lfloor n\theta_0 \rfloor}^*(t) \right) \right| \\
& \geq \sup_{t \in \mathbb{R}} \left| \frac{\sum_{k=1}^{\lfloor n\theta_0 \rfloor} w_{nk}}{n} \frac{\sum_{k=\lfloor n\theta_0 \rfloor+1}^n w_{nk}}{n} \left(F(t) - \tilde{F}(t) \right) \right| \\
& \quad - \sup_{t \in \mathbb{R}} \left| \frac{\sum_{k=1}^{\lfloor n\theta_0 \rfloor} w_{nk}}{n} \underbrace{\frac{\sum_{k=\lfloor n\theta_0 \rfloor+1}^n w_{nk}}{n}}_{\leq 1} \left(\hat{F}_{\lfloor n\theta_0 \rfloor}(t) - F(t) \right) \right| \\
& \quad - \sup_{t \in \mathbb{R}} \left| \underbrace{\frac{\sum_{k=1}^{\lfloor n\theta_0 \rfloor} w_{nk}}{n}}_{\leq 1} \frac{\sum_{k=\lfloor n\theta_0 \rfloor+1}^n w_{nk}}{n} \left(\tilde{F}(t) - \hat{F}_{n-\lfloor n\theta_0 \rfloor}^*(t) \right) \right| \\
& = \sup_{t \in \mathbb{R}} \left| \frac{\sum_{k=1}^{\lfloor n\theta_0 \rfloor} w_{nk}}{n} \frac{\sum_{k=\lfloor n\theta_0 \rfloor+1}^n w_{nk}}{n} \underbrace{\left(F(t) - \tilde{F}(t) \right)}_{>0} \right| + o_P(1).
\end{aligned}$$

Damit gilt

$$\sup_{t \in \mathbb{R}} \left| \hat{T}_n(\theta_0, t) \right| = \sqrt{n} \sup_{t \in \mathbb{R}} \left| \frac{\sum_{k=1}^{\lfloor n\theta_0 \rfloor} w_{nk}}{n} \frac{\sum_{k=\lfloor n\theta_0 \rfloor+1}^n w_{nk}}{n} \left(\hat{F}_{\lfloor n\theta_0 \rfloor}(t) - \hat{F}_{n-\lfloor n\theta_0 \rfloor}^*(t) \right) \right| \xrightarrow[n \rightarrow \infty]{P} \infty,$$

also auch

$$\sup_{s \in [0,1], t \in \mathbb{R}} \left| \hat{T}_n(s, t) \right| \xrightarrow[n \rightarrow \infty]{P} \infty$$

unter H_1 . Damit ist der Test konsistent.

Beweis von Satz 5.6 Mit Lemma 9.1 (S.117), Lemma 9.3 (S. 133) und Lemma 9.4 (S. 139) gilt wie im Beweis zu Satz 4.2 (S. 49) gleichmaig in $t \in \mathbb{R}$

$$\begin{aligned}
& \frac{\sum_{k=1}^{\lfloor n\theta_0 \rfloor} w_{nk}}{n} \left(\hat{F}_{\lfloor n\theta_0 \rfloor}(t) - F(t) \right) \\
& = \frac{\sum_{k=1}^{\lfloor n\theta_0 \rfloor} w_{nk}}{n} \left(\sum_{j=1}^{\lfloor n\theta_0 \rfloor} \bar{w}_{j\lfloor n\theta_0 \rfloor} I\{\hat{\epsilon}_j \leq t\} - F(t) \right) \\
& = \frac{1}{n} \sum_{j=1}^{\lfloor n\theta_0 \rfloor} (I\{\epsilon_j \leq t\} - F(t)) \\
& \quad + \frac{\sum_{k=1}^{\lfloor n\theta_0 \rfloor} w_{nk}}{n} \left(\sum_{j=1}^{\lfloor n\theta_0 \rfloor} \bar{w}_{j\lfloor n\theta_0 \rfloor} I\{\epsilon_j \leq t\} - F(t) \right) - \frac{1}{n} \sum_{j=1}^{\lfloor n\theta_0 \rfloor} (I\{\epsilon_j \leq t\} - F(t)) \\
& \quad + \frac{\sum_{k=1}^{\lfloor n\theta_0 \rfloor} w_{nk}}{n} \sum_{j=1}^{\lfloor n\theta_0 \rfloor} \bar{w}_{j\lfloor n\theta_0 \rfloor} \left(F \left(t \frac{\hat{\sigma}}{\sigma} (X_{j-1}) + \frac{\hat{m} - m}{\sigma} (X_{j-1}) \right) - F(t) \right) \\
& \quad + \frac{\sum_{k=1}^{\lfloor n\theta_0 \rfloor} w_{nk}}{n} \sum_{j=1}^{\lfloor n\theta_0 \rfloor} \bar{w}_{j\lfloor n\theta_0 \rfloor} \left(I\{\hat{\epsilon}_j \leq t\} - F \left(t \frac{\hat{\sigma}}{\sigma} (X_{j-1}) + \frac{\hat{m} - m}{\sigma} (X_{j-1}) \right) - I\{\epsilon_j \leq t\} + F(t) \right)
\end{aligned}$$

$$\begin{aligned}
&= O_P\left(n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}}\right) \text{ nach Lemma 9.4 auf S. 139} \\
&+ o_P\left(\frac{1}{\sqrt{n}}\right) \text{ nach Lemma 9.3 auf S. 133} \\
&+ \frac{\sum_{k=1}^{\lfloor n\theta_0 \rfloor} w_{nk}}{n} \sum_{j=1}^{\lfloor n\theta_0 \rfloor} \bar{w}_{j\lfloor n\theta_0 \rfloor} \left(F\left(t \frac{\hat{\sigma}}{\sigma}(X_{j-1}) + \frac{\hat{m}-m}{\sigma}(X_{j-1})\right) - F(t) \right) \\
&+ o_P\left(\frac{1}{\sqrt{n}}\right) \text{ nach Lemma 9.1 auf S. 117.}
\end{aligned}$$

Damit gilt gleichmaig in $t \in \mathbb{R}$

$$\begin{aligned}
&\frac{\sum_{k=1}^{\lfloor n\theta_0 \rfloor} w_{nk}}{n} \left(\hat{F}_{\lfloor n\theta_0 \rfloor}(t) - F(t) \right) \\
&= o_P(1) \\
&+ \frac{1}{n} \sum_{j=1}^{\lfloor n\theta_0 \rfloor} w_{nj} \left(F\left(t \frac{\hat{\sigma}}{\sigma}(X_{j-1}) + \frac{\hat{m}-m}{\sigma}(X_{j-1})\right) - F(t) \right). \tag{5.7.1}
\end{aligned}$$

Nach Lemma 3.4 (S. 35) gilt

$$\sup_{x \in I_n} \left| \frac{\hat{m}-m}{\sigma}(x) \right| = O_P\left(\left(c_n^{-\frac{1}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} + c_n^2\right) q_n q_n^f q_n^\sigma\right) = o_P(1)$$

und

$$\sup_{x \in I_n} \left| \frac{\hat{\sigma}-\sigma}{\sigma}(x) \right| = O_P\left(\left(c_n^{-\frac{1}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} + c_n^2\right) (q_n q_n^f q_n^\sigma)^2\right) = o_P(1).$$

Damit gilt fur (5.7.1) mit dem Mittelwertsatz fur ein $\xi(t, X_{j-1})$ zwischen $t \frac{\hat{\sigma}}{\sigma}(X_{j-1}) + \frac{\hat{m}-m}{\sigma}(X_{j-1})$ und t

$$\begin{aligned}
&\sup_{t \in \mathbb{R}} \left| \frac{1}{n} \sum_{j=1}^{\lfloor n\theta_0 \rfloor} w_{nj} \left(F\left(t \frac{\hat{\sigma}}{\sigma}(X_{j-1}) + \frac{\hat{m}-m}{\sigma}(X_{j-1})\right) - F(t) \right) \right| \\
&= \sup_{t \in \mathbb{R}} \left| \frac{1}{n} \sum_{j=1}^{\lfloor n\theta_0 \rfloor} w_{nj} f(\xi(t, X_{j-1})) \left(t \frac{\hat{\sigma}}{\sigma}(X_{j-1}) + \frac{\hat{m}-m}{\sigma}(X_{j-1}) - t \right) \right| \\
&\leq \sup_{x \in I_n} \left| \frac{\hat{m}-m}{\sigma}(x) \right| \sup_{t \in \mathbb{R}} |f(t)| \\
&\quad + \sup_{t \in \mathbb{R}} \left| \frac{1}{n} \sum_{j=1}^{\lfloor n\theta_0 \rfloor} w_{nj} f(\xi(t, X_{j-1})) t \left(\frac{\hat{\sigma}}{\sigma}(X_{j-1}) - 1 \right) \right| \\
&= o_P(1) \sup_{t \in \mathbb{R}} |f(t)| \\
&\quad + \sup_{t \in \mathbb{R}} \left| \frac{1}{n} \sum_{j=1}^{\lfloor n\theta_0 \rfloor} w_{nj} f(\xi(t, X_{j-1})) \xi(t, X_{j-1}) \frac{t}{\xi(t, X_{j-1})} \left(\frac{\hat{\sigma}}{\sigma}(X_{j-1}) - 1 \right) \right| \\
&\leq o_P(1) \sup_{t \in \mathbb{R}} |f(t)|
\end{aligned}$$

$$\begin{aligned}
& + \sup_{x \in I_n} \left| \frac{\hat{\sigma} - \sigma}{\sigma}(x) \right| \sup_{t \in \mathbb{R}, x \in I_n} \left| \frac{t}{\xi(t, x)} \right| \sup_{t \in \mathbb{R}} |f(t)t| \\
& = o_P(1) \cdot O_P(1),
\end{aligned}$$

da $\sup_{t \in \mathbb{R}} |f(t)| < \infty$ und $\sup_{t \in \mathbb{R}} |f(t)t| < \infty$ nach Voraussetzung und $\xi(t, x)$ zwischen t und $t \frac{\hat{\sigma}}{\sigma}(x) + \frac{\hat{m} - m}{\sigma}(x) \xrightarrow[n \rightarrow \infty]{P} t$ gleichmäßig über alle $x \in I_n$.

Somit ist die erste Behauptung des Satzes bewiesen.

Für die zweite Behauptung folgt analog

$$\begin{aligned}
& \frac{\sum_{k=\lfloor n\theta_0 \rfloor + 1}^n w_{nk}}{n} \left(\hat{F}_{n-\lfloor n\theta_0 \rfloor}^*(t) - \tilde{F}(t) \right) \\
& = \frac{1}{n} \sum_{j=\lfloor n\theta_0 \rfloor + 1}^n w_{nj} \left(I\{\hat{\epsilon}_j \leq t\} - \tilde{F}(t) \right) \\
& = \frac{1}{n} \sum_{j=\lfloor n\theta_0 \rfloor + 1}^n \left(I\{\epsilon_j \leq t\} - \tilde{F}(t) \right) \\
& \quad + \frac{1}{n} \sum_{j=\lfloor n\theta_0 \rfloor + 1}^n w_{nj} \left(I\{\epsilon_j \leq t\} - \tilde{F}(t) \right) - \frac{1}{n} \sum_{j=\lfloor n\theta_0 \rfloor + 1}^n \left(I\{\epsilon_j \leq t\} - \tilde{F}(t) \right) \\
& \quad + \frac{1}{n} \sum_{j=\lfloor n\theta_0 \rfloor + 1}^n w_{nj} \left(\tilde{F} \left(t \frac{\hat{\sigma}}{\sigma}(X_{j-1}) + \frac{\hat{m} - m}{\sigma}(X_{j-1}) \right) - \tilde{F}(t) \right) \\
& \quad + \frac{1}{n} \sum_{j=\lfloor n\theta_0 \rfloor + 1}^n w_{nj} \left(I\{\hat{\epsilon}_j \leq t\} - \tilde{F} \left(t \frac{\hat{\sigma}}{\sigma}(X_{j-1}) + \frac{\hat{m} - m}{\sigma}(X_{j-1}) \right) - I\{\epsilon_j \leq t\} + \tilde{F}(t) \right) \\
& = \frac{1}{n} \sum_{j=1}^n \left(I\{\epsilon_j \leq t\} - F_{\epsilon_j}(t) \right) \\
& \quad - \frac{1}{n} \sum_{j=1}^{\lfloor n\theta_0 \rfloor} \left(I\{\epsilon_j \leq t\} - F(t) \right) \\
& \quad + \frac{1}{n} \sum_{j=1}^n w_{nj} \left(I\{\epsilon_j \leq t\} - F_{\epsilon_j}(t) \right) - \frac{1}{n} \sum_{j=1}^n \left(I\{\epsilon_j \leq t\} - F_{\epsilon_j}(t) \right) \\
& \quad - \left(\frac{1}{n} \sum_{j=1}^{\lfloor n\theta_0 \rfloor} w_{nj} \left(I\{\epsilon_j \leq t\} - F(t) \right) - \frac{1}{n} \sum_{j=1}^{\lfloor n\theta_0 \rfloor} \left(I\{\epsilon_j \leq t\} - F(t) \right) \right) \\
& \quad + \frac{1}{n} \sum_{j=\lfloor n\theta_0 \rfloor + 1}^n w_{nj} \left(\tilde{F} \left(t \frac{\hat{\sigma}}{\sigma}(X_{j-1}) + \frac{\hat{m} - m}{\sigma}(X_{j-1}) \right) - \tilde{F}(t) \right) \\
& \quad + \frac{1}{n} \sum_{j=1}^n w_{nj} \left(I\{\hat{\epsilon}_j \leq t\} - F_{\epsilon_j} \left(t \frac{\hat{\sigma}}{\sigma}(X_{j-1}) + \frac{\hat{m} - m}{\sigma}(X_{j-1}) \right) - I\{\epsilon_j \leq t\} + F_{\epsilon_j}(t) \right) \\
& \quad - \frac{1}{n} \sum_{j=1}^{\lfloor n\theta_0 \rfloor} w_{nj} \left(I\{\hat{\epsilon}_j \leq t\} - F \left(t \frac{\hat{\sigma}}{\sigma}(X_{j-1}) + \frac{\hat{m} - m}{\sigma}(X_{j-1}) \right) - I\{\epsilon_j \leq t\} + F(t) \right) \\
& = O_P \left(n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} \right) + O_P \left(n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} \right) \text{ nach Lemma 9.4 auf S. 139}
\end{aligned}$$

$$\begin{aligned}
& +o_p\left(\frac{1}{\sqrt{n}}\right) + o_p\left(\frac{1}{\sqrt{n}}\right) \text{ nach Lemma 9.3 auf S. 133} \\
& + \frac{1}{n} \sum_{j=\lfloor n\theta_0 \rfloor + 1}^n w_{nj} \left(\tilde{F} \left(t \frac{\hat{\sigma}}{\sigma}(X_{j-1}) + \frac{\hat{m} - m}{\sigma}(X_{j-1}) \right) - \tilde{F}(t) \right) \\
& + o_p\left(\frac{1}{\sqrt{n}}\right) + o_p\left(\frac{1}{\sqrt{n}}\right) \text{ nach Lemma 9.1 auf S. 117.}
\end{aligned}$$

und wie oben mit dem Mittelwertsatz für ein $\xi(t, X_{j-1})$ zwischen $t \frac{\hat{\sigma}}{\sigma}(X_{j-1}) + \frac{\hat{m} - m}{\sigma}(X_{j-1})$ und t

$$\begin{aligned}
& \frac{1}{n} \sum_{j=\lfloor n\theta_0 \rfloor + 1}^n w_{nj} \left(\tilde{F} \left(t \frac{\hat{\sigma}}{\sigma}(X_{j-1}) + \frac{\hat{m} - m}{\sigma}(X_{j-1}) \right) - \tilde{F}(t) \right) \\
& = \sup_{t \in \mathbb{R}} \left| \frac{1}{n} \sum_{j=\lfloor n\theta_0 \rfloor + 1}^n w_{nj} \tilde{f}(\xi(t, X_{j-1})) \left(t \frac{\hat{\sigma}}{\sigma}(X_{j-1}) + \frac{\hat{m} - m}{\sigma}(X_{j-1}) - t \right) \right| \\
& = o_P(1).
\end{aligned}$$

Damit ist auch die zweite Behauptung des Satzes bewiesen.

□

5.4 Teststatistik im homoskedastischen Fall

Auch für homoskedastische Modelle der Form

$$X_j = m(X_{j-1}) + \epsilon_j$$

kann auf die gleiche Weise wie in Abschnitt 5.1 (S. 70) ein Test auf einen Change-Point in der Innovationenverteilung durchgeführt werden. In solchen Modellen muss für die Innovationen gelten $E[\epsilon_j] = 0$ und $\text{Var}(\epsilon_j) < \infty$ für alle j , im Gegensatz zu $\text{Var}(\epsilon_j) = 1$ im heteroskedastischen Modell. Daher können im homoskedastischen Fall auch Modelle mit einem Change-Point in der Größe der Varianz betrachtet werden (siehe Abschnitt 6.3, S. 98).

Um den in Abschnitt 5.1 (S. 70) vorgestellten Test auf ein homoskedastisches Modell anwenden zu können, muss lediglich die Schätzung der Innovationen, also die Residuenbildung verändert werden. Da die Skalierungsfunktion σ nicht geschätzt werden muss (sie ist im homoskedastischen Modell $\equiv 1$), reduziert sich die Schätzung der Innovationen auf

$$\hat{\epsilon}_j = X_j - \hat{m}(X_{j-1})$$

(Vergleich Abschnitt 4.1, S. 45). Mit diesen Residuen kann die Teststatistik genauso wie in Abschnitt 5.1 (S. 70) beschrieben gebildet werden.

5.5 Verwandte Ergebnisse

Sequentielle empirische Prozesse von parametrisch geschätzten Residuen und darauf aufbauende Change-Point-Test für die Verteilung der Innovationen wurden in der Literatur bereits für verschiedene Modelle betrachtet. So wird z.B. das asymptotische Verhalten des sequentiellen empirischen

Prozesses der Residuen und korrespondierende Change-Point-Tests in Bai [1994] für klassische parametrische ARMA Modelle, in Koul [1996] für nichtlineare parametrische Zeitreihenmodelle, und in Ling [1998] für nichtstationäre parametrische Autoregressionsmodelle betrachtet.

In Hlávka et al. [2010] wird ebenfalls ein Change-Point-Test für ein parametrisches Modell, ein AR(p)-Modell, betrachtet, allerdings nicht auf Basis eines sequentiellen empirischen Prozesses von Residuen, sondern durch die Verwendung der empirischen charakteristischen Funktion der Residuen. Außerdem wird hier nicht auf Basis von n Beobachtungen getestet, ob in diesen ein Change-Point vorliegt, sondern ab einem bestimmten Zeitpunkt wird für jede neu hinzukommende Beobachtung getestet, ob die Nullhypothese keines Change-Points noch gilt.

Sequentielle empirische Prozesse von nichtparametrisch geschätzten Residuen und darauf aufbauende Change-Point-Tests wurden in der Literatur bereits für Regressionsmodelle mit unabhängigen Daten betrachtet. So werden in Neumeyer and van Keilegom [2009] sowohl Testmethoden für eine Veränderung der Innovationenverteilung zu einem bestimmten Zeitpunkt, als auch abhängig vom Wert der Kovariablen, für nichtparametrische Regressionsmodelle mit unabhängigen Daten vorgestellt.

Die Resultate dieser Arbeit weiten die Betrachtung von sequentiellen empirischen Residuenprozessen und darauf aufbauenden Change-Point-Tests für die Innovationenverteilung auf nichtparametrische Autoregressionsmodelle aus.

6 Beispiele und Simulationen für den Change-Point-Test

In diesem Kapitel werden verschiedene (oft verwendete) Zeitreihenmodelle betrachtet.

Der in Kapitel 5 vorgestellte Change-Point-Test wird auf Daten, die nach diesen Modellen erzeugt wurden, angewendet. Die in Kapitel 5 bewiesenen Eigenschaften des Tests (Verteilung unter H_0 und Konsistenz) werden durch diese Simulationen bestätigt.

In Abschnitt 6.7 (S. 102) wird der Change-Point-Test außerdem auf einen echten Datensatz angewendet.

6.1 AR(1)-Modelle

Ein AR(1)-Modell ist ein autoregressives Modell der Form

$$X_j = a \cdot X_{j-1} + \epsilon_j, \quad a \in \mathbb{R}.$$

Für $|a| < 1$ ist dieser stochastische Prozess geometrisch ergodisch, also langfristig stationär und α -mischend mit einem exponentiell abfallenden Mischungskoeffizienten (siehe Abschnitt 2.2, S. 11).

6.1.1 Normalverteilte Innovationen

Es sei das spezielle AR(1)-Modell

$$X_j = 0.5 \cdot X_{j-1} + \epsilon_j, \quad \epsilon_1, \dots, \epsilon_{\lfloor \frac{n}{2} \rfloor} \sim \mathcal{N}(0, 1), \quad \epsilon_{\lfloor \frac{n}{2} \rfloor + 1}, \dots, \epsilon_n \sim \tilde{F}_1 \quad (6.0.1)$$

gegeben, wobei \tilde{F}_1 die Verteilungsfunktion einer Zufallsvariable sei, die mit Wahrscheinlichkeit 0.5 die Verteilung $\mathcal{N}(-2\zeta, 1)$ und mit Wahrscheinlichkeit 0.5 die Verteilung $\mathcal{N}(2\zeta, 1)$ hat.

Die Dichte einer solchen Zufallsvariable ergibt sich aus $\frac{1}{2} \left(\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(x+2\zeta)^2}{2}\right) + \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(x-2\zeta)^2}{2}\right) \right)$ und ist in Abbildung 6.1 für $\zeta = 0.5$ und in Abbildung 6.2 für $\zeta = 0.8$ graphisch dargestellt.

Eine Simulation des Change-Point-Tests zum Niveau $\gamma = 5\%$ bei 500 Wiederholungen hat für dieses Modell die Ablehnungswahrscheinlichkeiten

%	$\zeta = 0$	$\zeta = 0.1$	$\zeta = 0.2$	$\zeta = 0.3$	$\zeta = 0.4$	$\zeta = 0.5$
$n = 100$	5	6.2	7	9.2	10.2	13
$n = 200$	5.8	9.8	10.6	12.4	16.2	28.2

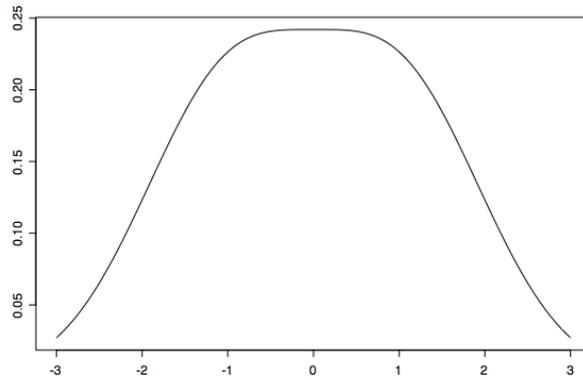


Abbildung 6.1: Dichte für Modell (6.0.1) mit $\zeta = 0.5$

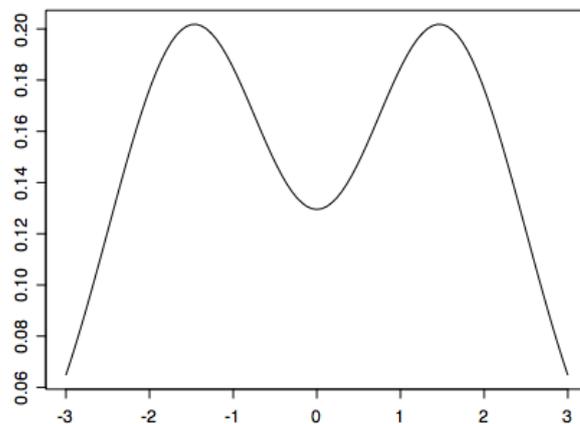


Abbildung 6.2: Dichte für Modell (6.0.1) mit $\zeta = 0.8$

%	$\zeta = 0.6$	$\zeta = 0.7$	$\zeta = 0.8$	$\zeta = 0.9$	$\zeta = 1$
$n = 100$	16.4	21	29	35.4	41
$n = 200$	40.2	59	78.2	88	92.4

ergeben. Die Ergebnisse für $\zeta = 0$ (Nullhypothese) zeigen, dass das Niveau $\gamma = 5\%$ ungefähr eingehalten wird. Außerdem zeigen die Ergebnisse, dass sich die Ablehnungswahrscheinlichkeit mit wachsendem ζ erhöht, und ebenso mit wachsendem n .

In Abbildung 6.3 sind diese Werte graphisch dargestellt.

Eine weitere Möglichkeit in welcher Art ein Change-Point in einem solchen AR(1)-Modell auftreten könnte, ist durch das folgende Modell gegeben

$$X_j = 0.5 \cdot X_{j-1} + \epsilon_j, \quad \epsilon_1, \dots, \epsilon_{\lfloor \frac{n}{2} \rfloor} \sim \mathcal{N}(0, 1), \quad \epsilon_{\lfloor \frac{n}{2} \rfloor + 1}, \dots, \epsilon_n \sim \tilde{F}_2, \quad (6.0.2)$$

wobei \tilde{F}_2 die Verteilungsfunktion einer Zufallsvariable sei, die mit Wahrscheinlichkeit 0.5 die Verteilung $\mathcal{N}(0, (1 - \zeta)^2)$ und mit Wahrscheinlichkeit 0.5 die Verteilung $\mathcal{N}(0, 2 - (1 - \zeta)^2)$ hat.

Die Dichte einer solchen Zufallsvariable ergibt sich aus

$$\frac{1}{2} \left(\frac{1}{\sqrt{2\pi(1-\zeta)^2}} \exp\left(-\frac{x^2}{2(1-\zeta)^2}\right) + \frac{1}{\sqrt{2\pi(2-(1-\zeta)^2)}} \exp\left(-\frac{x^2}{2(2-(1-\zeta)^2)}\right) \right) \text{ und ist in Abbildung 6.4 für } \zeta =$$

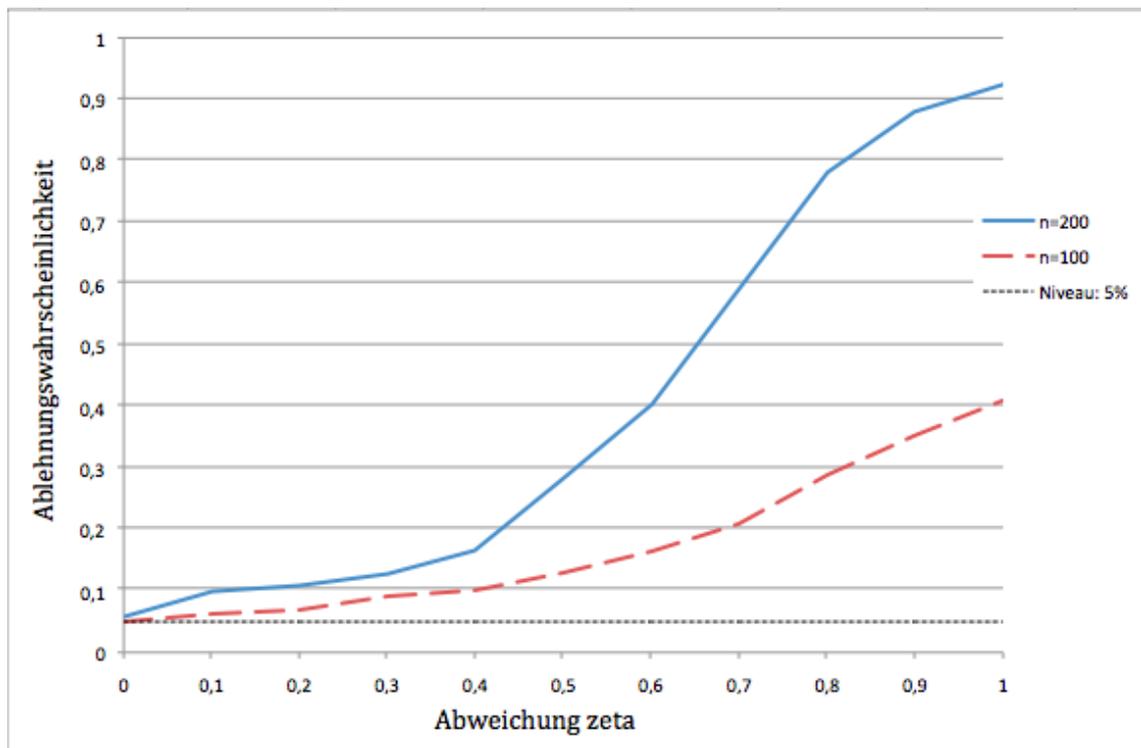


Abbildung 6.3: Ablehnungswahrscheinlichkeit für Modell (6.0.1)

0.5 und in Abbildung 6.5 für $\zeta = 0.8$ graphisch dargestellt. Mit dieser Dichte berechnet sich die Varianz zu $\frac{1}{2} \left((1 - \zeta)^2 + 2 - (1 - \zeta)^2 \right) = 1$, die Bedingung des heteroskedastischen Modells (Varianz= 1) ist also erfüllt.

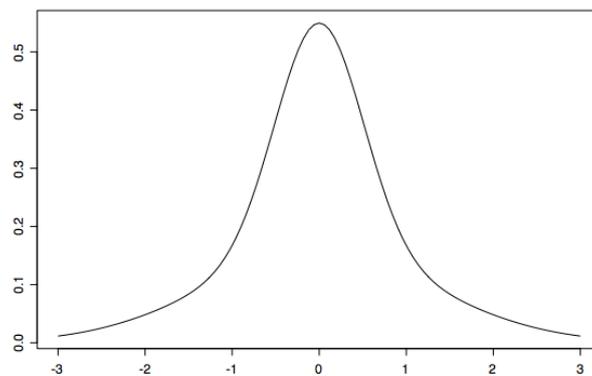


Abbildung 6.4: Dichte für Modell (6.0.2) mit $\zeta = 0.5$

Eine Simulation des Change-Point-Tests zum Niveau $\gamma = 5\%$ bei 500 Wiederholungen hat für dieses Modell die Ablehnungswahrscheinlichkeiten

%	$\zeta = 0$	$\zeta = 0.1$	$\zeta = 0.2$	$\zeta = 0.3$	$\zeta = 0.4$	$\zeta = 0.5$
$n = 100$	5	6.8	6.4	7	7.2	8.6
$n = 200$	5.8	8.8	9.8	9.6	11.2	13.2

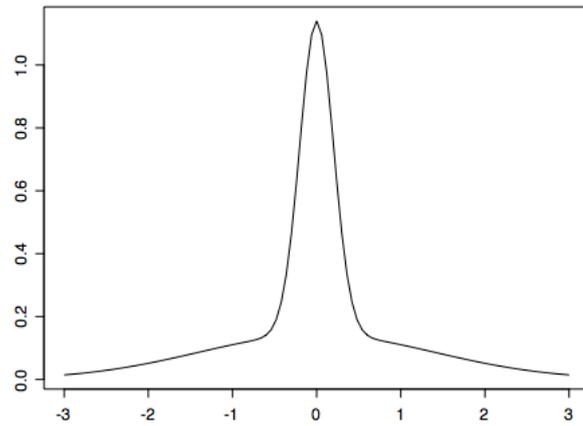


Abbildung 6.5: Dichte für Modell (6.0.2) mit $\zeta = 0.8$

%	$\zeta = 0.6$	$\zeta = 0.7$	$\zeta = 0.8$	$\zeta = 0.9$	$\zeta = 0.99$
$n = 100$	10.8	13.2	18	23	28.2
$n = 200$	18.4	30.4	41.6	54.4	63.4

ergeben. Auch hier zeigen die Ergebnisse, dass die Ablehnungswahrscheinlichkeiten für wachsendes ζ , genauso wie für wachsendes n , steigen.

In Abbildung 6.6 sind diese Werte graphisch dargestellt.

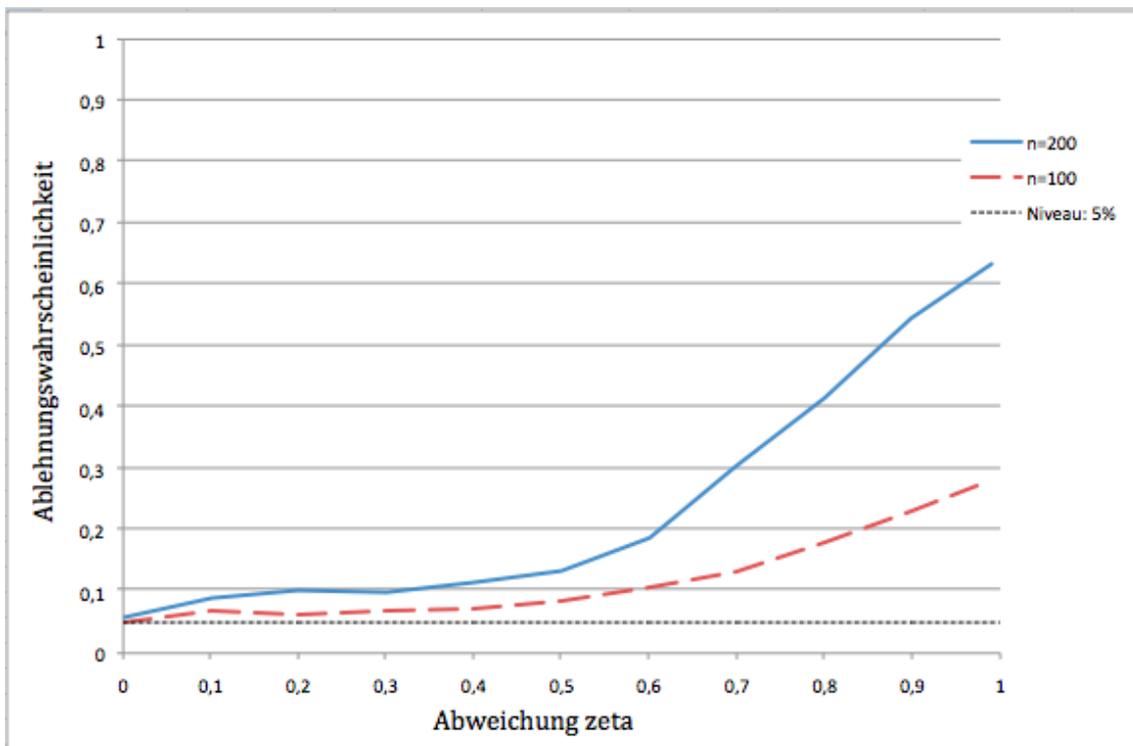


Abbildung 6.6: Ablehnungswahrscheinlichkeit für Modell (6.0.2)

In beiden Modellen ist für $n = 200$ ein deutlicher Anstieg der Ablehnungswahrscheinlichkeit für

positive ζ zu erkennen, man kann also darauf schließen, dass der Test konsistent ist, also dass die Ablehnungswahrscheinlichkeit bei einem Change-Point für $n \rightarrow \infty$ gegen 1 konvergiert. Außerdem ist zu sehen, dass die Ablehnungswahrscheinlichkeit beim ersten Modell (Modell (6.0.1)) stärker ansteigt als beim zweiten Modell (Modell (6.0.2)).

6.1.2 Skew-Normal verteilte Innovationen

Eine weitere Möglichkeit für ein Modell mit einem Change-Point, ist ein AR(1)-Modell mit ‘‘Skew-Normal’’-verteilten Innovationen

$$X_j = 0.5 \cdot X_{j-1} + \epsilon_j, \quad \epsilon_1, \dots, \epsilon_{\lfloor \frac{n}{2} \rfloor} \sim \mathcal{N}(0, 1), \quad \epsilon_{\lfloor \frac{n}{2} \rfloor + 1}, \dots, \epsilon_n \sim \tilde{F}_3, \quad (6.0.3)$$

wobei \tilde{F}_3 die ‘‘Skew-Normal’’-Verteilung $SN\left(-\sqrt{\frac{2\pi((10\zeta)^2+(10\zeta)^4)}{\pi^2+(2\pi^2-2\pi)\cdot(10\zeta)^2+(\pi^2-2\pi)\cdot(10\zeta)^4}}, \sqrt{\frac{\pi(1+(10\zeta)^2)}{\pi+(\pi-2)\cdot(10\zeta)^2}}, 10\zeta\right)$ sei.

Eine Zufallsvariable Z ist $SN(\lambda_1, \lambda_2, \lambda_3)$ verteilt, wenn $Z = \lambda_1 + \lambda_2 \cdot Z_0$ und Z_0 die Dichte $2\phi(x)\Phi(\lambda_3 x)$ hat, wobei ϕ die Dichte und Φ die Verteilungsfunktion der Standardnormalverteilung ist (siehe Azzalini and Capitanio [1999]).

Der Erwartungswert von einer solchen Zufallsvariablen Z errechnet sich aus $\lambda_1 + \lambda_2 \sqrt{\frac{2}{\pi}} \cdot \frac{\lambda_3}{\sqrt{1+\lambda_3^2}}$ und die Varianz aus $\lambda_2^2 \left(1 - \frac{2}{\pi} \cdot \frac{\lambda_3^2}{1+\lambda_3^2}\right)$. Daraus ergeben sich die Parameter für die Verteilung nach dem Change-Point, da der Erwartungswert der Innovationen immer 0 und die Varianz immer 1 bleiben muss.

Eine Simulation des Change-Point-Tests zum Niveau $\gamma = 5\%$ bei 500 Wiederholungen hat für dieses Modell die Ablehnungswahrscheinlichkeiten

%	$\zeta = 0$	$\zeta = 0.1$	$\zeta = 0.2$	$\zeta = 0.3$	$\zeta = 0.4$	$\zeta = 0.5$
$n = 100$	5	7	7.2	8.6	11.4	11.6
$n = 200$	5.8	8.4	10.2	14.8	15.2	17.4

%	$\zeta = 0.6$	$\zeta = 0.7$	$\zeta = 0.8$	$\zeta = 0.9$	$\zeta = 1$
$n = 100$	9.8	12.2	12.2	11.6	10.8
$n = 200$	17.2	18.8	20.8	19.2	21.6

ergeben.

In Abbildung 6.7 sind diese Werte graphisch dargestellt.

Es ist zu sehen, dass die Ablehnungswahrscheinlichkeit ab etwa $\zeta = 0.5$ nur noch langsam weiter ansteigt, während sie in den beiden vorher betrachteten Modellen gerade in diesem Bereich stark ansteigt.

Diese Beobachtung ist damit zu erklären, dass die betrachtete Skew-Normal-Verteilung für wachsendes ζ konvergiert.

Der erste Parameter $-\sqrt{\frac{2\pi((10\zeta)^2+(10\zeta)^4)}{\pi^2+(2\pi^2-2\pi)\cdot(10\zeta)^2+(\pi^2-2\pi)\cdot(10\zeta)^4}}$ (siehe (6.0.3)) konvergiert für wachsendes ζ gegen $-\sqrt{\frac{2\pi}{\pi^2-2\pi}}$ und der zweite Parameter $\sqrt{\frac{\pi(1+(10\zeta)^2)}{\pi+(\pi-2)\cdot(10\zeta)^2}}$ gegen $\sqrt{\frac{\pi}{\pi-2}}$. Die Dichte $2\phi(x)\Phi(10\zeta x)$ von Z_0 konvergiert für wachsendes ζ gegen 0 für $x < 0$ und gegen $2\phi(x)$ für $x > 0$.

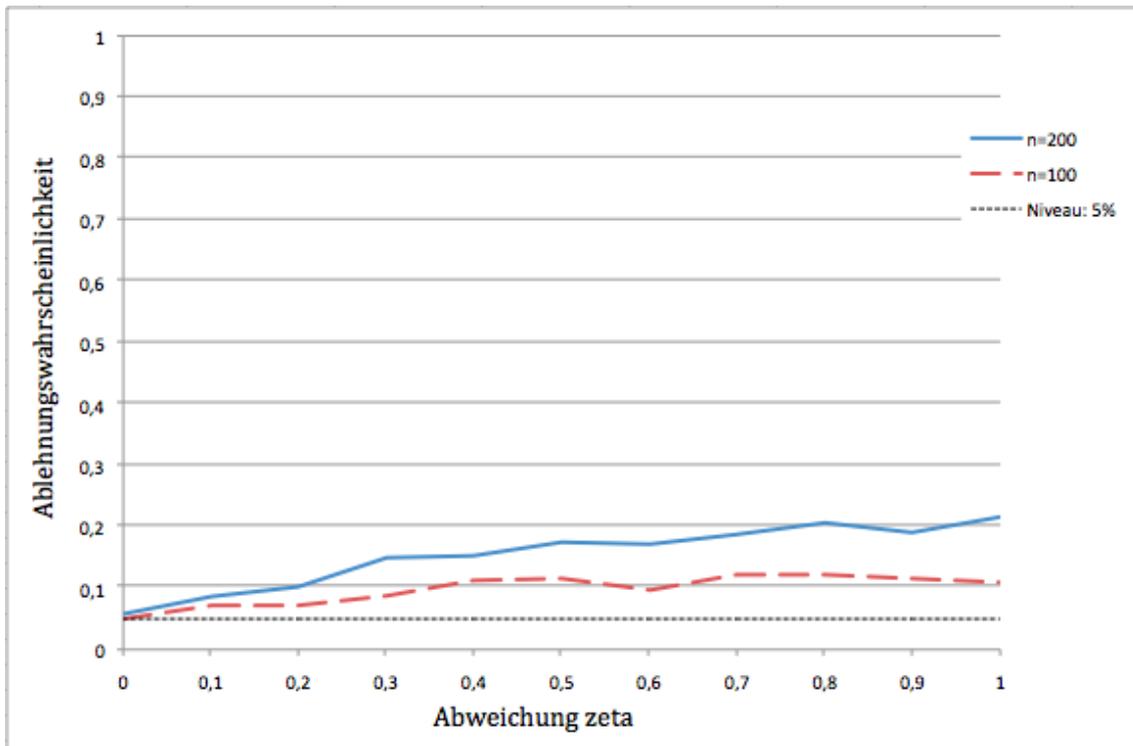


Abbildung 6.7: Ablehnungswahrscheinlichkeit für Modell (6.0.3)

In den Abbildungen 6.8, 6.9 und 6.10 sind die Dichten für $\zeta = 0.2$, $\zeta = 0.6$ und $\zeta = 1$ zu sehen. Diese Graphiken wurden unter <http://azzalini.stat.unipd.it/SN/plot-SN1.html> erstellt.

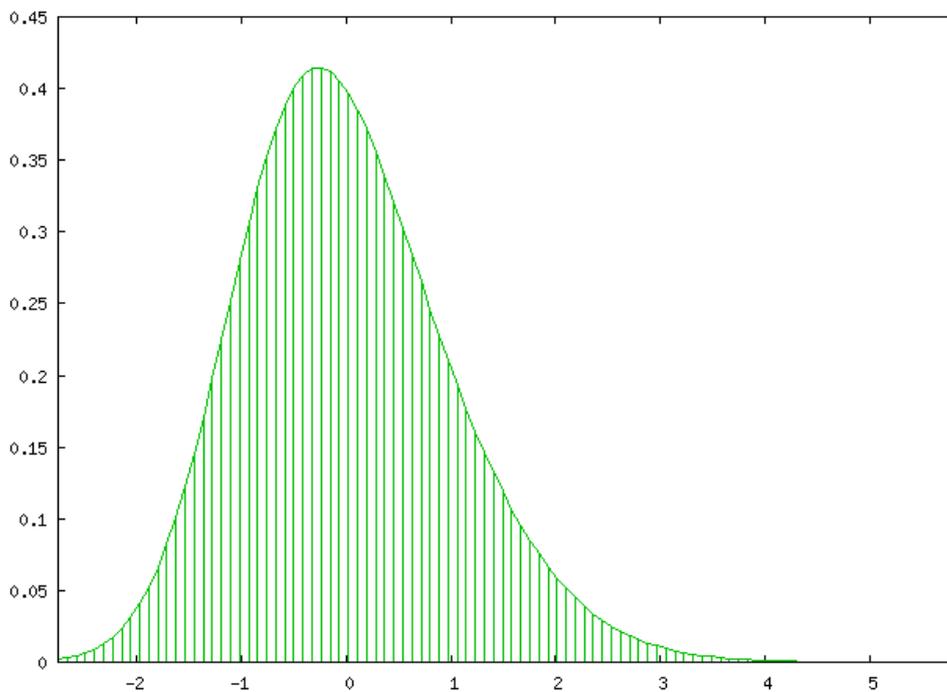


Abbildung 6.8: Dichte der Skew-Normal-Verteilung für Modell (6.0.3) mit $\zeta = 0.2$

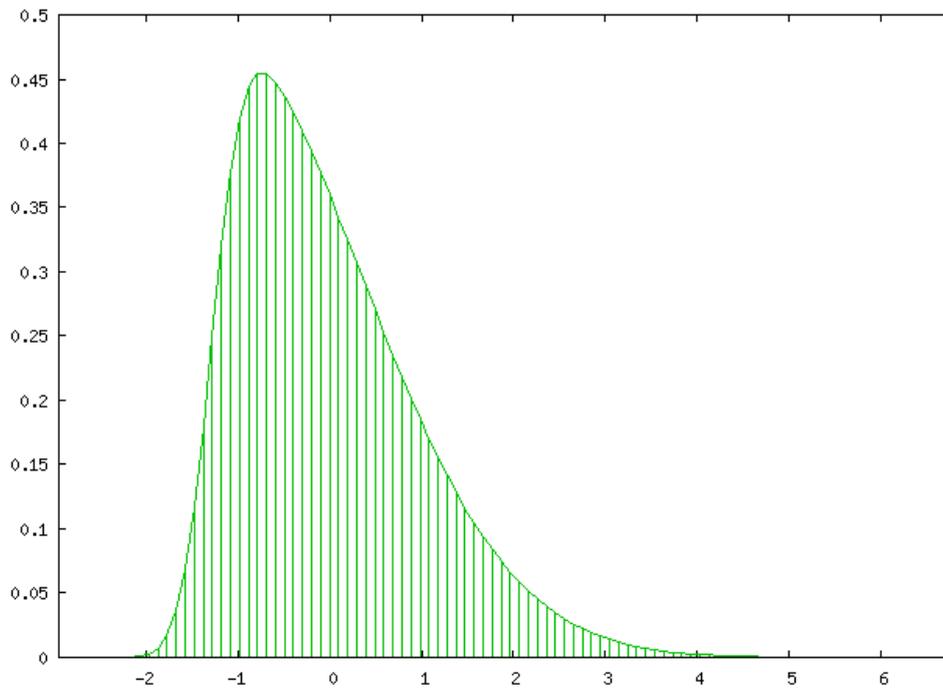


Abbildung 6.9: Dichte der Skew-Normal-Verteilung für Modell (6.0.3) mit $\zeta = 0.6$

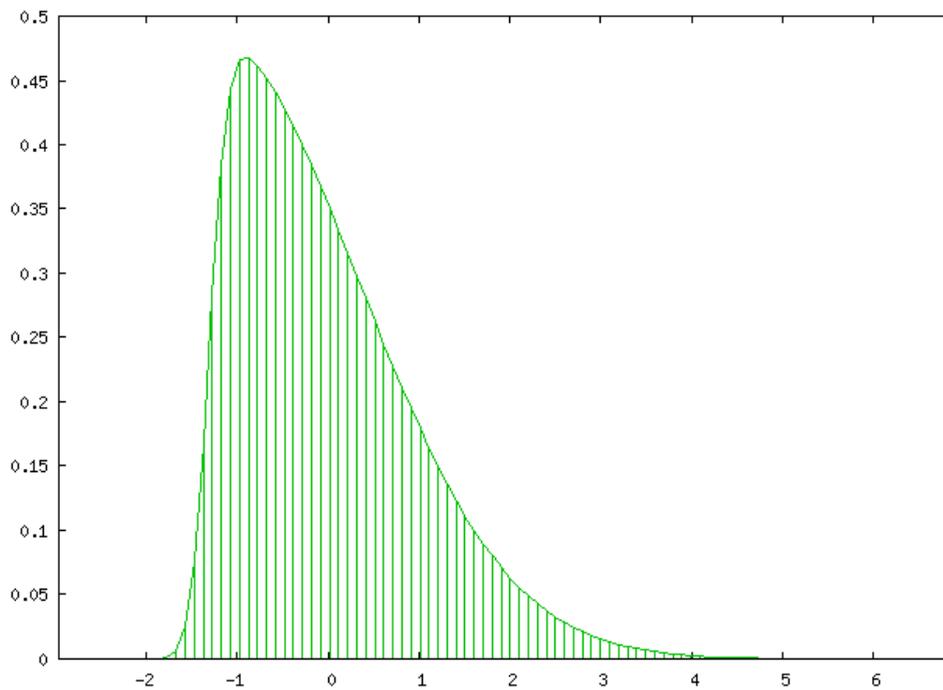


Abbildung 6.10: Dichte der Skew-Normal-Verteilung für Modell (6.0.3) mit $\zeta = 1$

Ein optischer Vergleich dieser Dichten zeigt, dass der Unterschied zwischen der Dichte mit $\zeta = 0.2$ und der Dichte mit $\zeta = 0.6$ größer ist, als der Unterschied zwischen der Dichte mit $\zeta = 0.6$ und der Dichte mit $\zeta = 1$.

Auch optisch ist also zu sehen, dass die Unterschiede zwischen den Dichten mit wachsendem ζ geringer werden.

Dies ist der Grund dafür, dass die Ablehnungswahrscheinlichkeit nicht in der gleichen Weise wie in den Modellen (6.0.1) und (6.0.2) ansteigt.

6.2 ARCH(1)-Modelle

Ein ARCH(1)-Modell ist ein autoregressives Modell der Form

$$X_j = \sqrt{a_0 + a_1 X_{j-1}^2} \cdot \epsilon_j, \quad a_0, a_1 \in \mathbb{R}.$$

Für $a_0 > 0$ und $1 > a_1 > 0$ ist dieser stochastische Prozess geometrisch ergodisch, also langfristig stationär und α -mischend mit einem exponentiell abfallenden Mischungskoeffizienten (siehe Doukhan [1994], 2.4.2.3 Example 5a).

6.2.1 Normalverteilte Innovationen

Es sei das spezielle ARCH(1)-Modell

$$X_j = \sqrt{0.75 + 0.25 X_{j-1}^2} \cdot \epsilon_j, \quad \epsilon_1, \dots, \epsilon_{\lfloor \frac{n}{2} \rfloor} \sim \mathcal{N}(0, 1), \quad \epsilon_{\lfloor \frac{n}{2} \rfloor + 1}, \dots, \epsilon_n \sim \tilde{F}_1 \quad (6.0.4)$$

gegeben, wobei, wie in Modell (6.0.1), \tilde{F}_1 die Verteilungsfunktion einer Zufallsvariable sei, die mit Wahrscheinlichkeit 0.5 die Verteilung $\mathcal{N}(-2\zeta, 1)$ und mit Wahrscheinlichkeit 0.5 die Verteilung $\mathcal{N}(2\zeta, 1)$ hat.

Eine Simulation des Change-Point-Tests zum Niveau $\gamma = 5\%$ bei 500 Wiederholungen hat für dieses Modell die Ablehnungswahrscheinlichkeiten

%	$\zeta = 0$	$\zeta = 0.1$	$\zeta = 0.2$	$\zeta = 0.3$	$\zeta = 0.4$	$\zeta = 0.5$
$n = 100$	4.8	6.4	6.6	7.8	8	11.2
$n = 200$	5	8.4	9.2	11.2	14.6	27

%	$\zeta = 0.6$	$\zeta = 0.7$	$\zeta = 0.8$	$\zeta = 0.9$	$\zeta = 1$
$n = 100$	15.6	20.8	29.6	35.2	56.8
$n = 200$	37.2	47.4	55.4	61.4	80

ergeben. Die Ergebnisse zeigen, dass unter der Nullhypothese das Niveau $\gamma = 5\%$ sehr gut eingehalten wird, und dass die Ablehnungswahrscheinlichkeit für wachsendes ζ , genauso wie für wachsendes n , ansteigt.

In Abbildung 6.11 sind diese Werte graphisch dargestellt.

Auch für das ARCH(1)-Modell kann ein Change-Point analog zu Modell (6.0.2) betrachtet werden mit

$$X_j = \sqrt{0.75 + 0.25 X_{j-1}^2} \cdot \epsilon_j, \quad \epsilon_1, \dots, \epsilon_{\lfloor \frac{n}{2} \rfloor} \sim \mathcal{N}(0, 1), \quad \epsilon_{\lfloor \frac{n}{2} \rfloor + 1}, \dots, \epsilon_n \sim \tilde{F}_2, \quad (6.0.5)$$

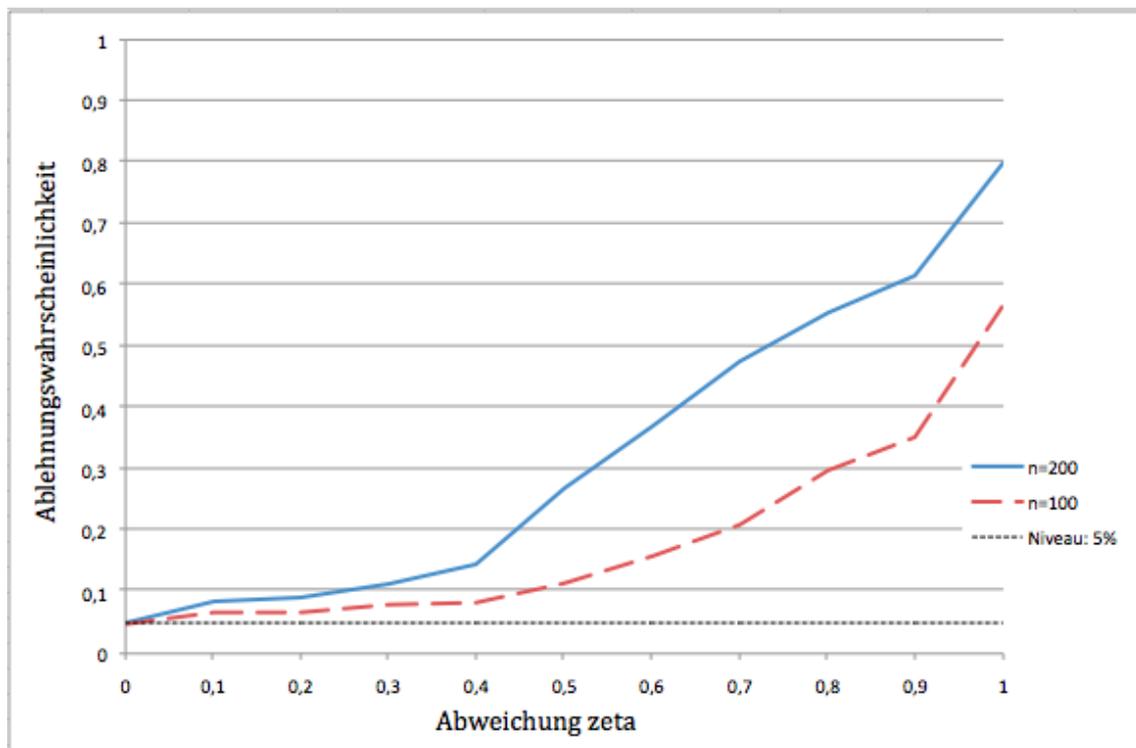


Abbildung 6.11: Ablehnungswahrscheinlichkeit für Modell (6.0.4)

wobei \tilde{F}_2 die Verteilungsfunktion einer Zufallsvariable sei, die mit Wahrscheinlichkeit 0.5 die Verteilung $\mathcal{N}(0, (1 - \zeta)^2)$ und mit Wahrscheinlichkeit 0.5 die Verteilung $\mathcal{N}(0, 2 - (1 - \zeta)^2)$ hat.

Eine Simulation des Change-Point-Tests zum Niveau $\gamma = 5\%$ bei 500 Wiederholungen hat für dieses Modell die Ablehnungswahrscheinlichkeiten

%	$\zeta = 0$	$\zeta = 0.1$	$\zeta = 0.2$	$\zeta = 0.3$	$\zeta = 0.4$	$\zeta = 0.5$
$n = 100$	4.8	6.8	7.6	8	7.4	8.4
$n = 200$	5	8.2	10	8.6	11.4	12

%	$\zeta = 0.6$	$\zeta = 0.7$	$\zeta = 0.8$	$\zeta = 0.9$	$\zeta = 0.99$
$n = 100$	10.2	11.6	17.6	27.2	42.2
$n = 200$	19.4	28.8	43.2	61.6	78.8

ergeben. Auch hier zeigen die Ergebnisse, dass die Ablehnungswahrscheinlichkeiten für wachsendes ζ , genauso wie für wachsendes n , steigen.

In Abbildung 6.12 sind diese Werte graphisch dargestellt.

Im Vergleich zu den Simulationen im AR(1)-Modell ist zu sehen, dass die Ablehnungswahrscheinlichkeiten im zweiten Modell im ARCH(1)-Modell deutlich stärker ansteigen als im AR(1)-Modell (Modell (6.0.2) und (6.0.5)).

Für $n = 200$ sind die Kurvenverläufe im Modell (6.0.4) (erstes Modell) und im Modell (6.0.5) recht ähnlich. Für $n = 100$ steigt die Ablehnungswahrscheinlichkeit im Modell (6.0.4) mit wachsendem ζ stärker an als im in Modell (6.0.5).

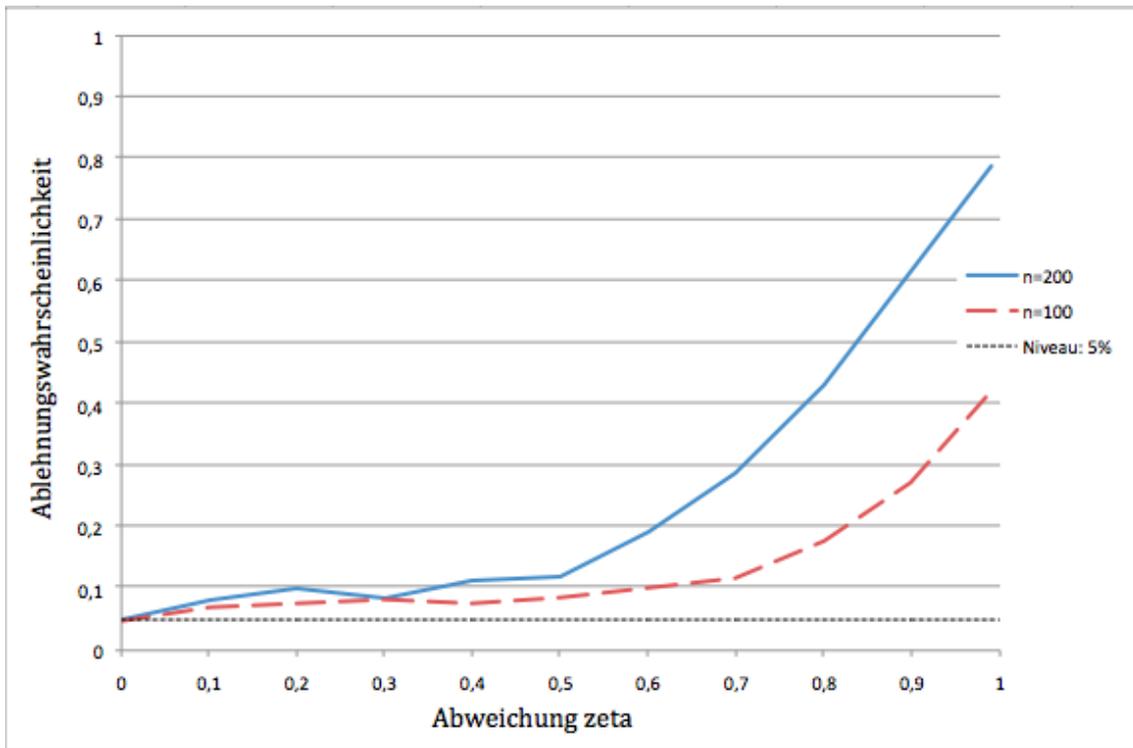


Abbildung 6.12: Ablehnungswahrscheinlichkeit für Modell (6.0.5)

6.2.2 Skew-Normal verteilte Innovationen

Eine weitere Möglichkeit für ein Modell mit einem Change-Point, ist analog zu Modell (6.0.3) ein ARCH(1)-Modell mit “Skew-Normal”-verteilten Innovationen

$$X_j = \sqrt{0.75 + 0.25X_{j-1}^2} \cdot \epsilon_j, \quad \epsilon_1, \dots, \epsilon_{\lfloor \frac{n}{2} \rfloor} \sim \mathcal{N}(0, 1), \quad \epsilon_{\lfloor \frac{n}{2} \rfloor + 1}, \dots, \epsilon_n \sim \tilde{F}_3, \quad (6.0.6)$$

wobei \tilde{F}_3 die “Skew-Normal”-Verteilung $SN\left(-\sqrt{\frac{2\pi((10\zeta)^2 + (10\zeta)^4)}{\pi^2 + (2\pi^2 - 2\pi) \cdot (10\zeta)^2 + (\pi^2 - 2\pi) \cdot (10\zeta)^4}}, \sqrt{\frac{\pi(1 + (10\zeta)^2)}{\pi + (\pi - 2) \cdot (10\zeta)^2}}, 10\zeta\right)$ sei.

Eine Simulation des Change-Point-Tests zum Niveau $\gamma = 5\%$ bei 500 Wiederholungen hat für dieses Modell die Ablehnungswahrscheinlichkeiten

%	$\zeta = 0$	$\zeta = 0.1$	$\zeta = 0.2$	$\zeta = 0.3$	$\zeta = 0.4$	$\zeta = 0.5$
$n = 100$	4.8	7.6	8	10	11.8	9.8
$n = 200$	5	7.8	11.8	12	15.2	17.6

%	$\zeta = 0.6$	$\zeta = 0.7$	$\zeta = 0.8$	$\zeta = 0.9$	$\zeta = 1$
$n = 100$	11	10.8	12	11	12.2
$n = 200$	18.8	18.4	20.6	18.4	18.6

ergeben.

In Abbildung 6.13 sind diese Werte graphisch dargestellt.

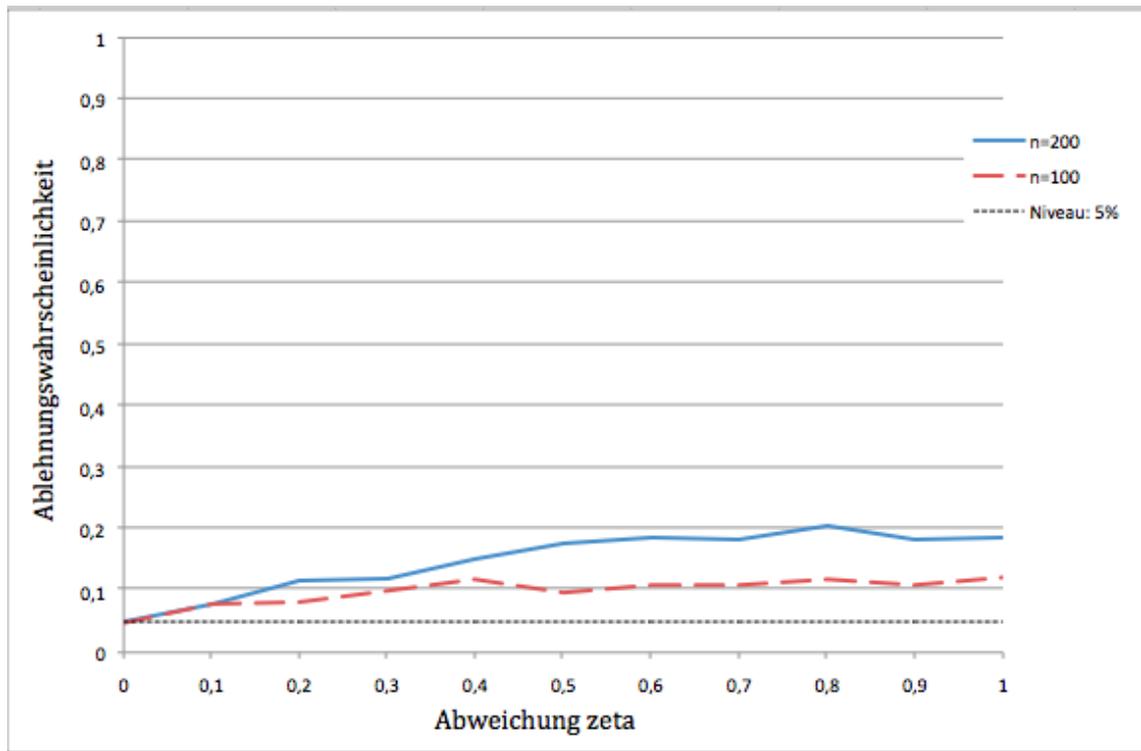


Abbildung 6.13: Ablehnungswahrscheinlichkeit für Modell (6.0.6)

Der Kurvenverlauf ist sehr ähnlich zu dem von Modell (6.0.3) und kann mit den selben Argumenten begründet werden.

6.2.3 Student-t-verteilte Innovationen

Ein weiteres Modell von Interesse ist ein ARCH(1)-Modell mit Innovationen, die einer standardisierten Student-t-Verteilung unterliegen. Nach Bollerslev [1987] ist diese Art der Modellwahl besser für die Modellierung von spekulativen Preisänderungen und Ertragsarten geeignet als ARCH(1)-Modelle mit normalverteilten Innovationen, da die Student-t-Verteilung schwerere Ränder hat, als die Normalverteilung.

Auf Grund der Annahme $\text{Var}(\epsilon_j) = 1$ für alle j muss die Student-t-Verteilung standardisiert werden, da die Varianz der eigentlichen Student-t-Verteilung durch $\frac{\nu}{\nu-2}$ berechnet wird, wobei $\nu \in \mathbb{N}$, $\nu > 2$ den Freiheitsgrad bezeichnet.

Die Dichte der standardisierten Student-t-Verteilung mit $\nu > 2$ Freiheitsgraden ($\mathcal{St}(\nu)$) ist gegeben durch

$$t \mapsto \frac{1}{\sqrt{(\nu-2)\pi}} \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})} \left(1 + \frac{t^2}{\nu-2}\right)^{-\frac{\nu+1}{2}}.$$

In Abbildung 6.14 ist das Randverhalten der $\mathcal{St}(\nu)$ -Verteilung für verschiedene Werte von ν im Vergleich zur $\mathcal{N}(0,1)$ -Verteilung dargestellt, und es ist zu sehen, dass die Ränder, besonders für kleine ν , schwerer sind, als bei der $\mathcal{N}(0,1)$ -Verteilung.

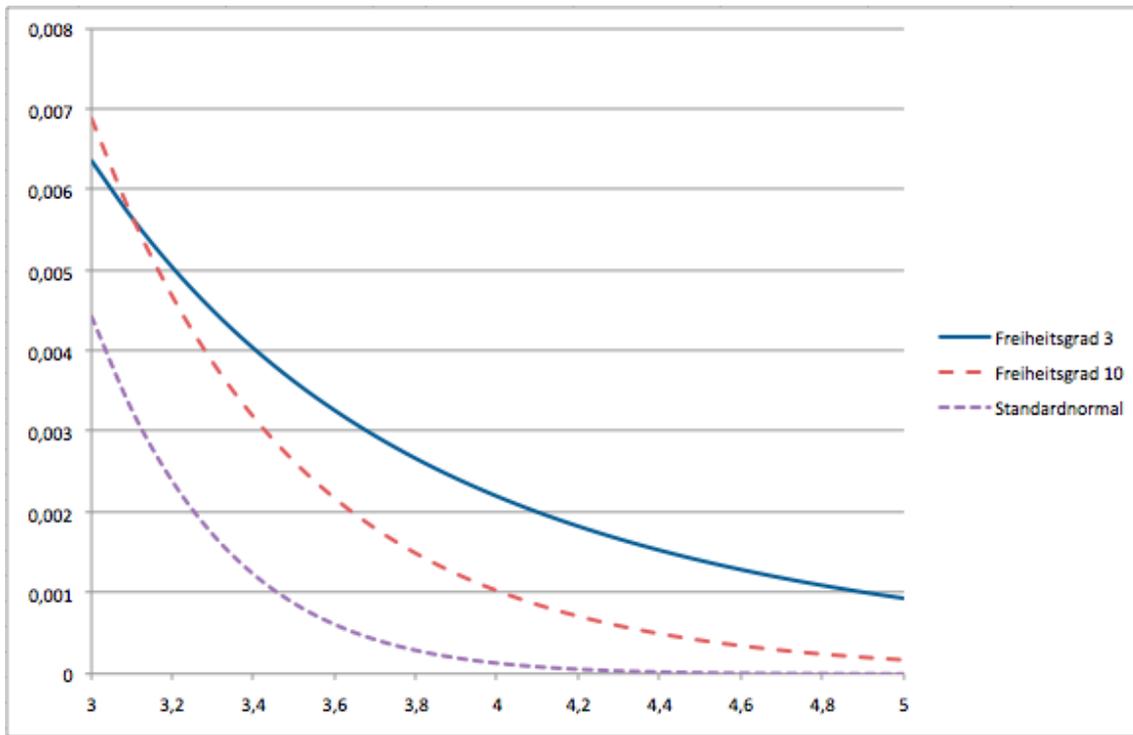


Abbildung 6.14: Randverhalten der $\mathcal{St}(\nu)$ -Verteilung für $\nu = 3$, $\nu = 10$ und der $\mathcal{N}(0, 1)$ -Verteilung

Ein Modell mit dieser Innovationen-Verteilung und einem Change-Point ist z.B.

$$X_j = \sqrt{0.75 + 0.25X_{j-1}^2} \cdot \epsilon_j, \quad \epsilon_1, \dots, \epsilon_{\lfloor \frac{n}{2} \rfloor} \sim \mathcal{St}(3), \quad \epsilon_{\lfloor \frac{n}{2} \rfloor + 1}, \dots, \epsilon_n \sim \mathcal{St}(3 + 10 \cdot \zeta). \quad (6.0.7)$$

Eine Simulation des Change-Point-Tests zum Niveau $\gamma = 5\%$ bei 500 Wiederholungen hat für dieses Modell die Ablehnungswahrscheinlichkeiten

%	$\zeta = 0$	$\zeta = 0.1$	$\zeta = 0.2$	$\zeta = 0.3$	$\zeta = 0.4$	$\zeta = 0.5$
$n = 100$	5.4	6.8	8.2	8.4	10.6	10
$n = 200$	6	9.6	12.8	14.6	17.2	14.4

%	$\zeta = 0.6$	$\zeta = 0.7$	$\zeta = 0.8$	$\zeta = 0.9$	$\zeta = 1$
$n = 100$	10	10	9.6	11.6	9.4
$n = 200$	17	17.4	18.8	19.6	19.8

ergeben. Es ist zu sehen, dass auch für diese Innovationenverteilung unter der Nullhypothese das Niveau von 5% ungefähr eingehalten wird.

In Abbildung 6.15 sind diese Werte graphisch dargestellt.

Hier ist ein sehr ähnlicher Verlauf wie im Fall der Skew-Normal verteilten Innovationen in Abbildung 6.7 und 6.13 zu beobachten. Ab etwa $\zeta = 0.4$ ist kein starker Anstieg mehr zu beobachten, was auch hier mit der Konvergenz für wachsendes ζ erklärt werden kann.

Die Student-t-Verteilung und damit auch die standardisierte Student-t-Verteilung $\mathcal{St}(\nu)$ konvergiert für wachsendes ν gegen die Standardnormalverteilung. Wie in Abbildung 6.16 zu sehen ist, wird der

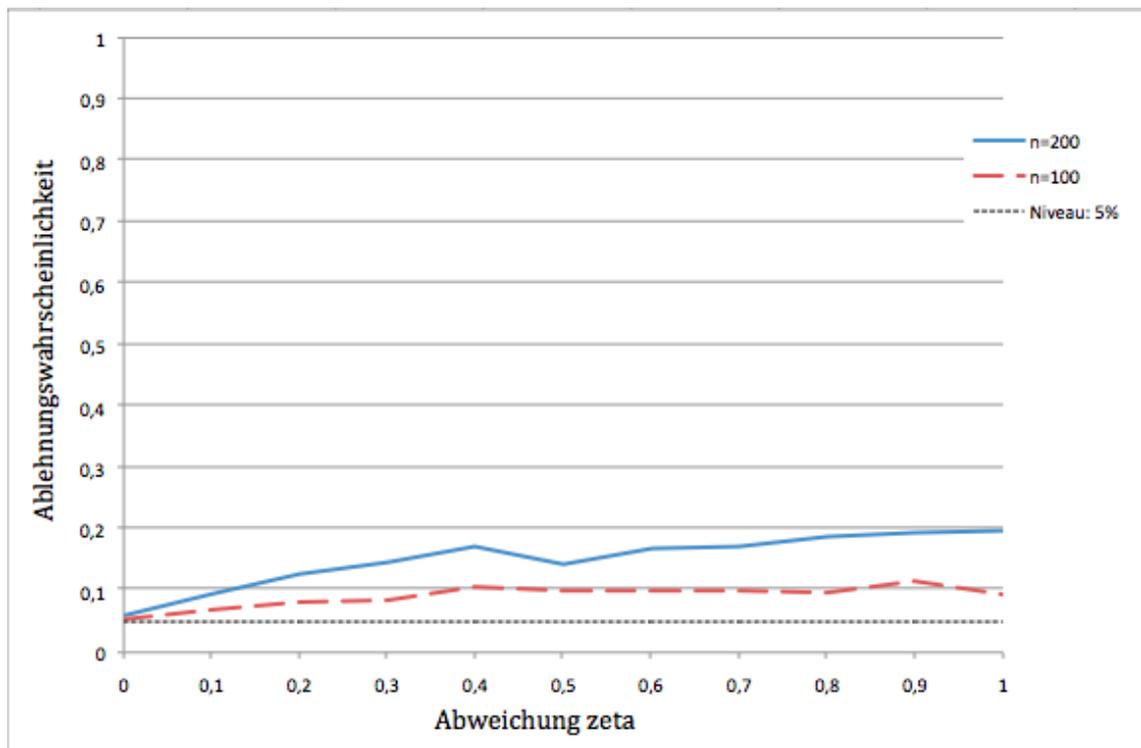


Abbildung 6.15: Ablehnungswahrscheinlichkeit für Modell (6.0.7)

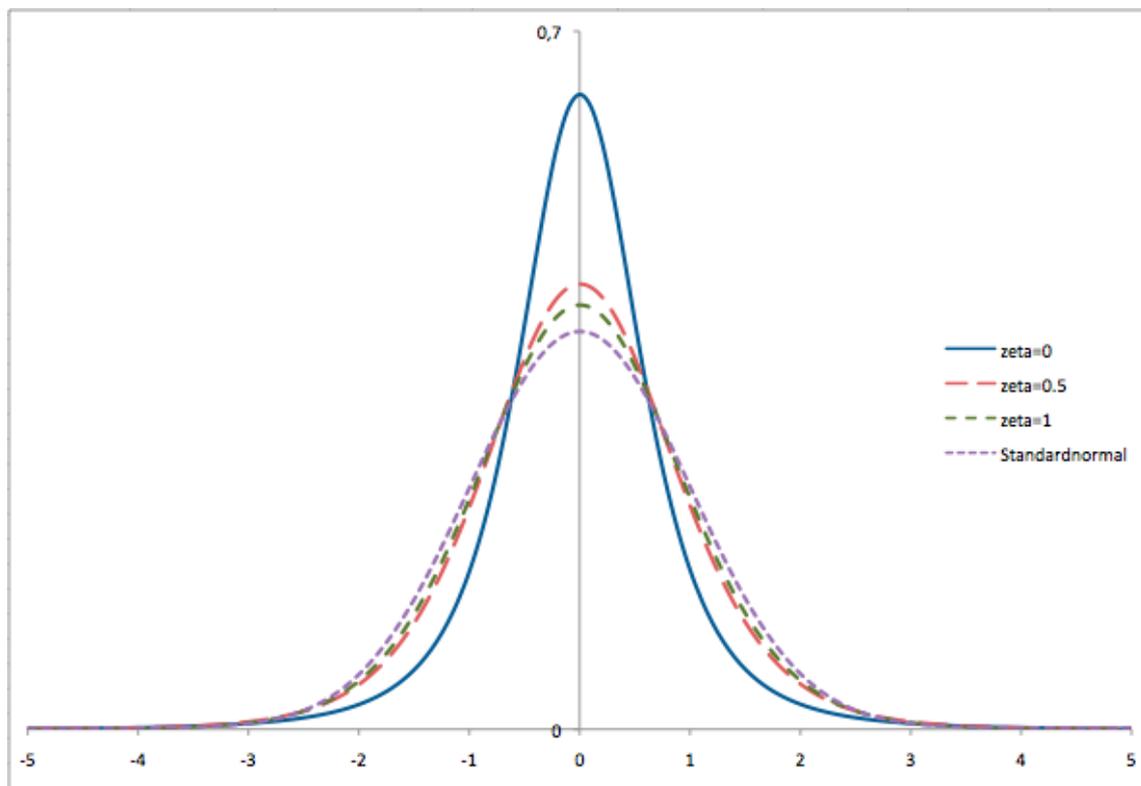


Abbildung 6.16: Dichten der $\mathcal{St}(3+10\cdot\zeta)$ -Verteilung für $\zeta = 0$, $\zeta = 0.5$, $\zeta = 1$ und $\mathcal{N}(0, 1)$ -Verteilung

Unterschied zwischen den einzelnen Dichten mit wachsendem ζ kleiner.

Daher ist der geringe Anstieg der Ablehnungswahrscheinlichkeit analog zum Skew-Normal-Fall zu erklären.

Auch im Fall der Student-t-verteilten Innovationen können natürlich Modelle mit einem Change-Point wie im Fall der Normalverteilten Innovationen ((6.0.4) und (6.0.5)) betrachtet werden.

So ein Modell ist z.B.

$$X_j = \sqrt{0.75 + 0.25X_{j-1}^2} \cdot \epsilon_j, \quad \epsilon_1, \dots, \epsilon_{\lfloor \frac{n}{2} \rfloor} \sim \mathcal{St}(3), \quad \epsilon_{\lfloor \frac{n}{2} \rfloor + 1}, \dots, \epsilon_n \sim \tilde{F}_4, \quad (6.0.8)$$

wobei \tilde{F}_4 die Verteilungsfunktion einer Zufallsvariable sei, die sich mit Wahrscheinlichkeit 0.5 aus der Addition einer $\mathcal{St}(3)$ -verteilten Zufallsvariable und 2ζ , und mit Wahrscheinlichkeit 0.5 aus der Addition einer $\mathcal{St}(3)$ -verteilten Zufallsvariable und -2ζ ergibt.

Eine Simulation des Change-Point-Tests zum Niveau $\gamma = 5\%$ bei 500 Wiederholungen hat für dieses Modell die Ablehnungswahrscheinlichkeiten

%	$\zeta = 0$	$\zeta = 0.1$	$\zeta = 0.2$	$\zeta = 0.3$	$\zeta = 0.4$	$\zeta = 0.5$
$n = 100$	5.4	7	7.4	12.2	20.2	29.6
$n = 200$	6	9.2	12.4	25	51.2	72.4

%	$\zeta = 0.6$	$\zeta = 0.7$	$\zeta = 0.8$	$\zeta = 0.9$	$\zeta = 1$
$n = 100$	32.8	32.2	34	34.4	62.4
$n = 200$	82	85	78	73.6	84.6

ergeben.

In Abbildung 6.17 sind diese Werte graphisch dargestellt.

Im Vergleich zu Modell (6.0.4), also der selben Art von Change-Point mit Normalverteilung, ist zu sehen, dass bereits für kleinere Abweichungen ζ ein starker Anstieg der Ablehnungswahrscheinlichkeit, besonders für $n = 200$, zu beobachten ist.

Eine weitere Möglichkeit für einen Change-Point in diesem Modell ist analog zu (6.0.5)

$$X_j = \sqrt{0.75 + 0.25X_{j-1}^2} \cdot \epsilon_j, \quad \epsilon_1, \dots, \epsilon_{\lfloor \frac{n}{2} \rfloor} \sim \mathcal{St}(3), \quad \epsilon_{\lfloor \frac{n}{2} \rfloor + 1}, \dots, \epsilon_n \sim \tilde{F}_5, \quad (6.0.9)$$

wobei \tilde{F}_5 die Verteilungsfunktion einer Zufallsvariable sei, die sich mit Wahrscheinlichkeit 0.5 aus der Multiplikation einer $\mathcal{St}(3)$ -verteilten Zufallsvariable und $(1 - \zeta)$, und mit Wahrscheinlichkeit 0.5 aus der Multiplikation einer $\mathcal{St}(3)$ -verteilten Zufallsvariable und $\sqrt{2 - (1 - \zeta)^2}$ ergibt.

Eine Simulation des Change-Point-Tests zum Niveau $\gamma = 5\%$ bei 500 Wiederholungen hat für dieses Modell die folgenden Ablehnungswahrscheinlichkeiten ergeben, die in Abbildung 6.18 graphisch dargestellt sind.

%	$\zeta = 0$	$\zeta = 0.1$	$\zeta = 0.2$	$\zeta = 0.3$	$\zeta = 0.4$	$\zeta = 0.5$
$n = 100$	5.4	8	8.8	8.2	8.4	8.2
$n = 200$	6	8.6	10.2	9.2	11.6	12.8

%	$\zeta = 0.6$	$\zeta = 0.7$	$\zeta = 0.8$	$\zeta = 0.9$	$\zeta = 1$
$n = 100$	11.2	12	18.6	26.8	37.8
$n = 200$	15	26.8	39.8	58.6	77.4

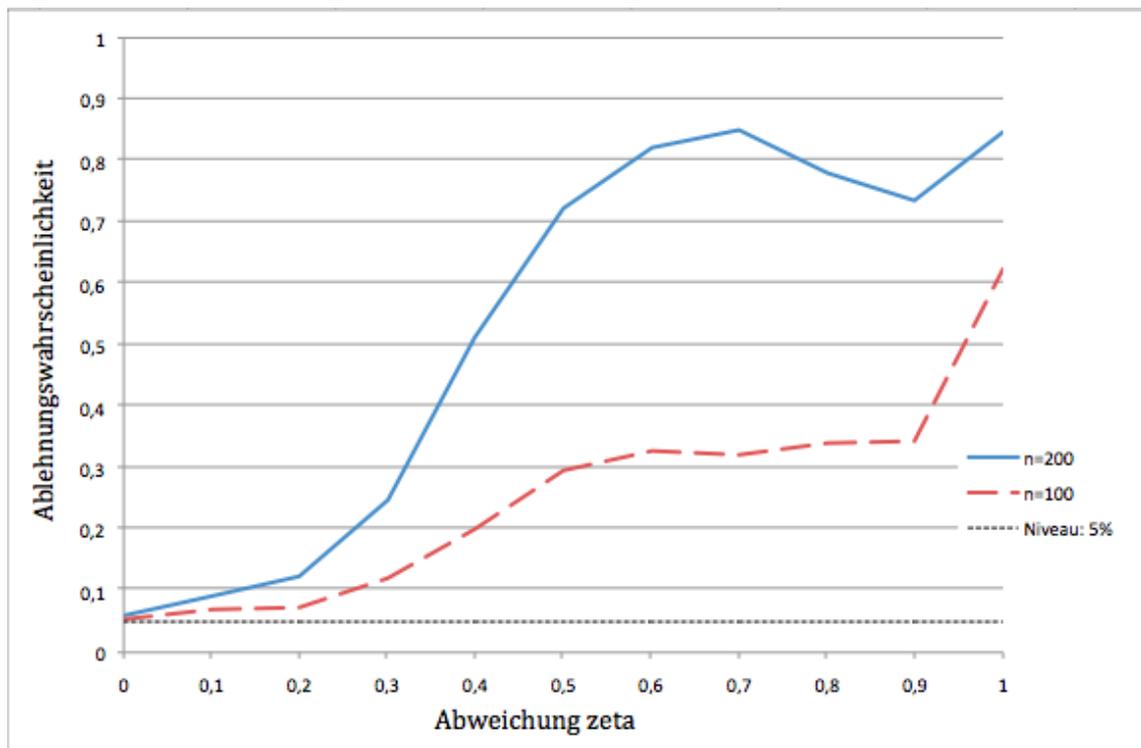


Abbildung 6.17: Ablehnungswahrscheinlichkeit für Modell 6.0.8

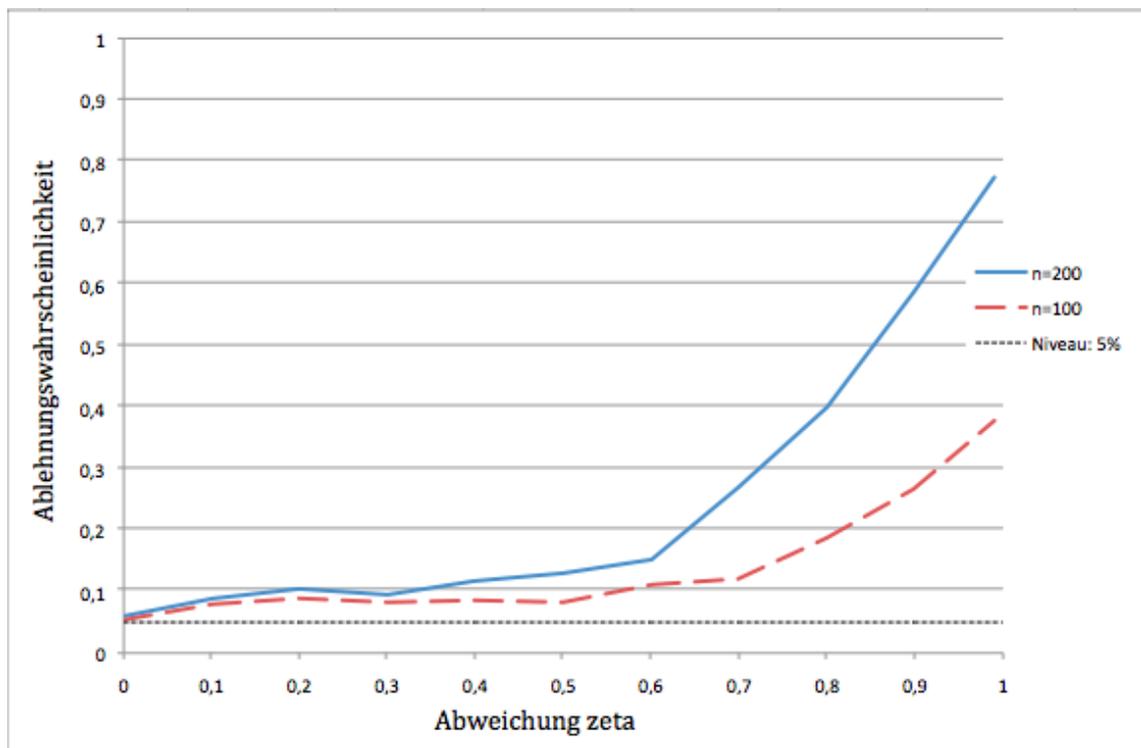


Abbildung 6.18: Ablehnungswahrscheinlichkeit für Modell (6.0.9)

Es ist zu sehen, dass die Kurvenverläufe sehr ähnlich zu den Ergebnissen von Modell (6.0.5), also

der selben Art von Change-Point mit Normalverteilung, sind.

6.3 Homoskedastische Modelle

Der betrachtete Change-Point-Test kann natürlich auch auf homoskedastische Modelle der Form

$$X_j = m(X_{j-1}) + \epsilon_j$$

angewendet werden (siehe Abschnitt 5.4, S. 81). In solchen Modellen muss für die Innovationen gelten $E[\epsilon_j] = 0$ und $\text{Var}(\epsilon_j) < \infty$ für alle j , im Gegensatz zu $\text{Var}(\epsilon_j) = 1$ im heteroskedastischen Modell. Daher können im homoskedastischen Fall auch Modelle mit einem Change-Point in der Größe der Varianz betrachtet werden.

Ein solches Modell ist z.B. das AR(1)-Modell

$$X_j = 0.5 \cdot X_{j-1} + \epsilon_j, \quad \epsilon_1, \dots, \epsilon_{\lfloor \frac{n}{2} \rfloor} \sim \mathcal{N}(0, 0.5^2), \quad \epsilon_{\lfloor \frac{n}{2} \rfloor + 1}, \dots, \epsilon_n \sim \mathcal{N}(0, (0.5 + \zeta)^2). \quad (6.0.10)$$

Eine Simulation des Change-Point-Tests zum Niveau $\gamma = 5\%$ bei 500 Wiederholungen hat für dieses Modell die Ablehnungswahrscheinlichkeiten

%	$\zeta = 0$	$\zeta = 0.1$	$\zeta = 0.2$	$\zeta = 0.3$	$\zeta = 0.4$	$\zeta = 0.5$
$n = 100$	4.6	6.8	8.6	9.8	14.4	16.2
$n = 200$	5.6	9	17.2	25.2	40.2	51.4

%	$\zeta = 0.6$	$\zeta = 0.7$	$\zeta = 0.8$	$\zeta = 0.9$	$\zeta = 1$
$n = 100$	20	24.4	29.4	36.4	39
$n = 200$	68.2	77.2	85.2	89.6	95.2

ergeben. Diese Ergebnisse zeigen, dass auch hier das Niveau $\gamma = 5\%$ ungefähr eingehalten wird. In Abbildung 6.19 sind diese Werte graphisch dargestellt.

Es ist zu sehen, dass die Ablehnungswahrscheinlichkeit für kleine Abweichungen ζ stärker ansteigt als in den heteroskedastischen Modellen. Dies ist damit zu erklären, dass der Unterschied zwischen den Verteilungen vor und nach dem Change-Point in dem hier betrachteten Modell stärker ist als in allen heteroskedastischen. Durch die Vorgabe $\text{Var}(\epsilon_j) = 1$ ist im heteroskedastischen Fall bereits festgelegt, dass sich die Verteilungen vor und nach dem Change-Point in den ersten beiden Momenten gleichen, während sie sich im hier betrachteten homoskedastischen Modell im zweiten Moment unterscheiden.

6.4 Simulationstechnik

Für die Simulationen wurden jeweils $10 \cdot n$ Beobachtungen X_j erzeugt, von denen die ersten $9,5 \cdot n$ durch Innovationen mit Verteilung vor dem Change Point und die letzten $0,5 \cdot n$ durch Innovationen mit Verteilung nach dem Change-Point entstanden. Für die Teststatistik wurden die letzten n Beobachtungen benutzt. Der Vorlauf wurde erzeugt um zu gewährleisten, dass sich der Prozess im Betrachtungszeitraum im Gleichgewicht befindet.

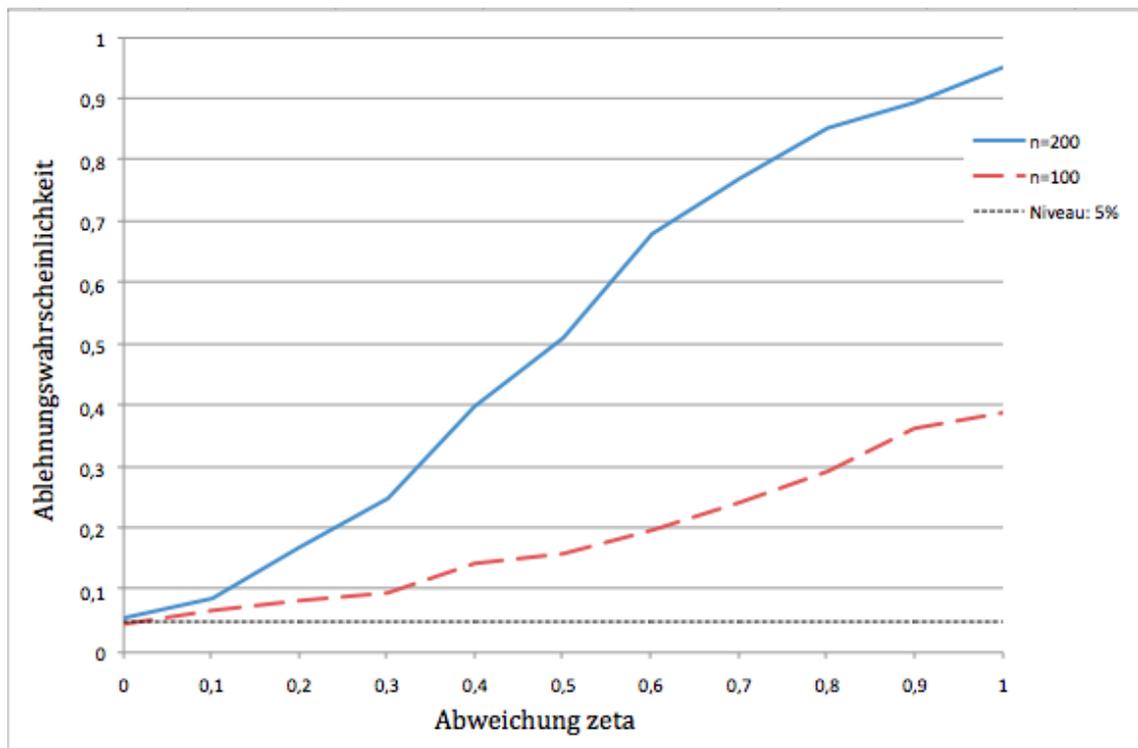


Abbildung 6.19: Ablehnungswahrscheinlichkeit für Modell (6.0.10)

Die Nadaraya-Watson Schätzer \hat{m} und $\hat{\sigma}$ wurden mit der Dichte der Standardnormal-Verteilung als Kern und Bandbreite $c_n = n^{-\frac{1}{4}}$ berechnet. Diese Wahl entspricht nicht allen Voraussetzungen, z.B. ist der Träger des Kerns nicht kompakt. Da jedoch die Dichte der Standardnormal-Verteilung an den Rändern exponentiell schnell klein wird, hat diese Abweichung nur sehr geringe, und damit vernachlässigbare, Auswirkungen auf die Simulationen.

Auch in der Wahl der Bandbreite $c_n = n^{-\frac{1}{4}}$ wurde die Voraussetzung $c_n \sim n^{-\frac{1}{4}} \log(n)^{-r}$ ($0 < r < \infty$) nicht eingehalten. Dies geschah, da für $n = 100$ oder $n = 200$ der Einfluss von $\log(n)^{-r}$ im Vergleich zum Einfluss von $n^{-\frac{1}{4}}$ zu groß ist. Der wichtige Term in der Bandbreitengröße ist $n^{-\frac{1}{4}}$, der für $n \rightarrow \infty$ auch der bestimmende ist. Für die relativ kleinen Stichprobenumfänge in den Simulationen wurde der Logarithmus-Teil daher weggelassen.

Die empirischen Prozesse der Residuen wurden in den Simulationen ohne die Gewichtsfunktion w_n gebildet. Das bedeutet, dass das Intervall I_n als die reelle Zahlengerade gewählt wurde. Die Voraussetzung $I_n = O(\log(n)^{r_I})$ ($r_I < \infty$) wurde also nicht eingehalten.

Die trotzdem sehr guten Ergebnisse der Simulationen lassen vermuten, dass die Gewichtsfunktion für die Theorie zwar notwendig ist, die Anwendung des Tests aber auch ohne sie funktioniert. Ein Grund hierfür könnte sein, dass sehr große Beobachtungswerte, die durch die Gewichtsfunktion aus der Berechnung des empirischen Prozesses entfernt werden sollen, in den Simulationen praktisch nicht vorkommen.

Die Simulationen wurden alle mit der Statistiksoftware R erstellt.

6.5 Fazit der Simulationen

In allen simulierten Modellen ist zu sehen, dass der Test das Niveau (ungefähr) einhält.

Für mehrere verschiedene Arten von Change-Points ist in den Modellen jeweils zu sehen, dass die Ablehnungswahrscheinlichkeit mit wachsendem n wächst. Für $n \rightarrow \infty$ geht die Ablehnungswahrscheinlichkeit gegen 1. Für die Modelle mit einem Change-Point im Erwartungswert oder in der Varianz sieht man dies auch schon für $n = 200$, für die Modelle mit einem Change-Point im dritten Parameter der Skew-Normal-Verteilung oder im Freiheitsgrad der Student-t-Verteilung müsste n noch größer gewählt werden, um die Konvergenz gegen 1 sehen zu können.

Die zuvor bewiesenen Eigenschaften des Tests (die Verteilung unter H_0 und die Konsistenz) bestätigen sich also in den Simulationen.

6.6 Wahl der Bandbreite

In den bisher gezeigten Simulationsergebnissen wurde als Bandbreite jeweils $c_n = n^{-\frac{1}{4}}$ gewählt. Wie in Abschnitt 6.4 (S. 98) beschrieben wurde, ist dies eine geeignete Bandbreitenwahl um den Modellannahmen zu genügen.

Durch die Annahmen ist vorgegeben, dass $n^{-\frac{1}{4}}$ der bestimmende Term in der Bandbreite sein muss. Allerdings kann die Bandbreite durch Multiplikation mit Konstanten verändert werden. Um die Auswirkungen der Größe der Bandbreite auf den Test zu untersuchen, wurden daher beispielhaft für das Modell (6.0.1), mit $n = 200$, Simulationen mit $c_n = c \cdot n^{-\frac{1}{4}}$ für verschiedene $c \in \mathbb{R}_{>0}$ durchgeführt.

Die Ergebnisse sind in Abbildung 6.20 dargestellt.

Es ist zu sehen, dass besonders ab einer Abweichung von $\zeta = 0.5$ die Ablehnungswahrscheinlichkeit mit wachsender Bandbreite ansteigt.

Der Bereich der Abweichung ζ zwischen 0 und 0.5 ist in Abbildung 6.21 noch einmal gesondert dargestellt.

Es ist zu sehen, dass die Ablehnungswahrscheinlichkeit auch unter der Nullhypothese keines Change-Points mit wachsendem c ansteigt, dass das Niveau also nicht mehr (ungefähr) eingehalten wird. Unter diesem Gesichtspunkt werden die Ergebnisse also schlechter, wenn die Bandbreite vergrößert wird. Betrachtet man den Bereich der Abweichungen von 0.5 bis 1, so führt unter dem Gesichtspunkt der Power des Tests, also dem Ansteigen der Ablehnungswahrscheinlichkeit für wachsendes ζ , eine größere Bandbreite zu besseren Ergebnissen. Zwischen Abweichungen von ca. 0.05 und 0.5 ist kein eindeutiger Trend erkennbar.

Eindeutig ist allerdings erkennbar, dass die Multiplikation mit $c = 0.5$ gegenüber $c = 1$ nach den Simulationsergebnissen nicht zu empfehlen ist, da sie zu einer geringeren Power führt, und auch für $c = 1$ das Niveau (ungefähr) eingehalten wird.

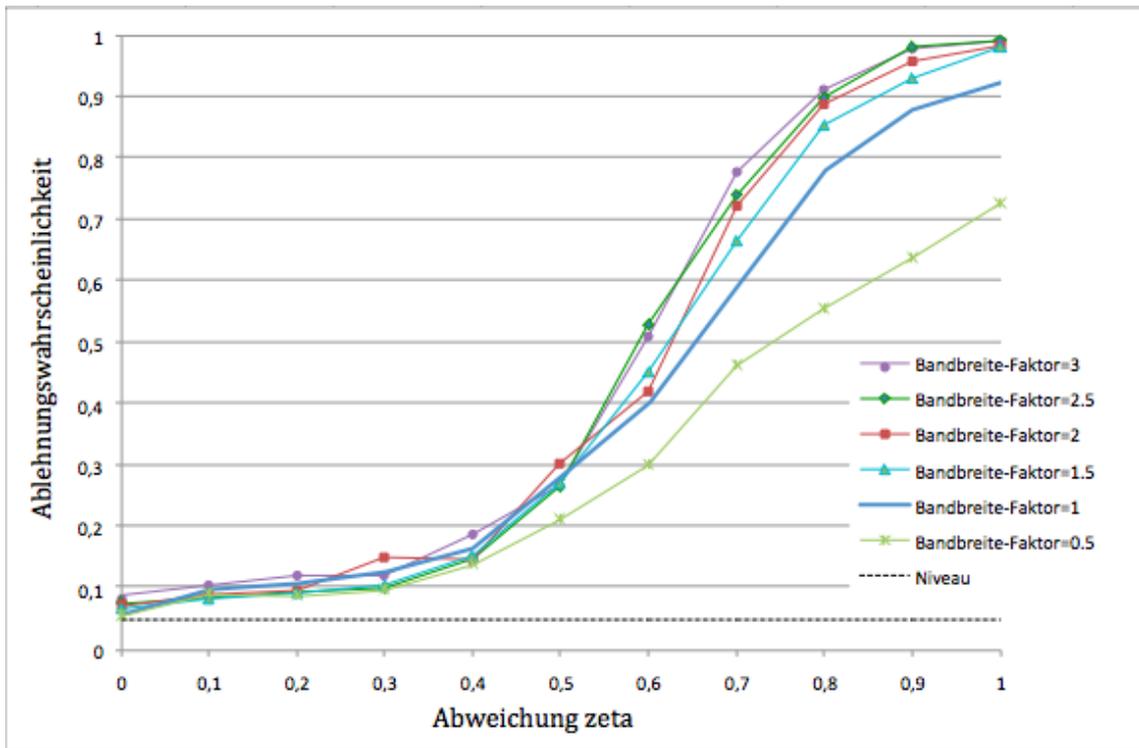


Abbildung 6.20: Ablehnungsw. für Modell (6.0.1) mit $c = 0.5, c = 1, c = 1.5, c = 2, c = 2.5, c = 3$

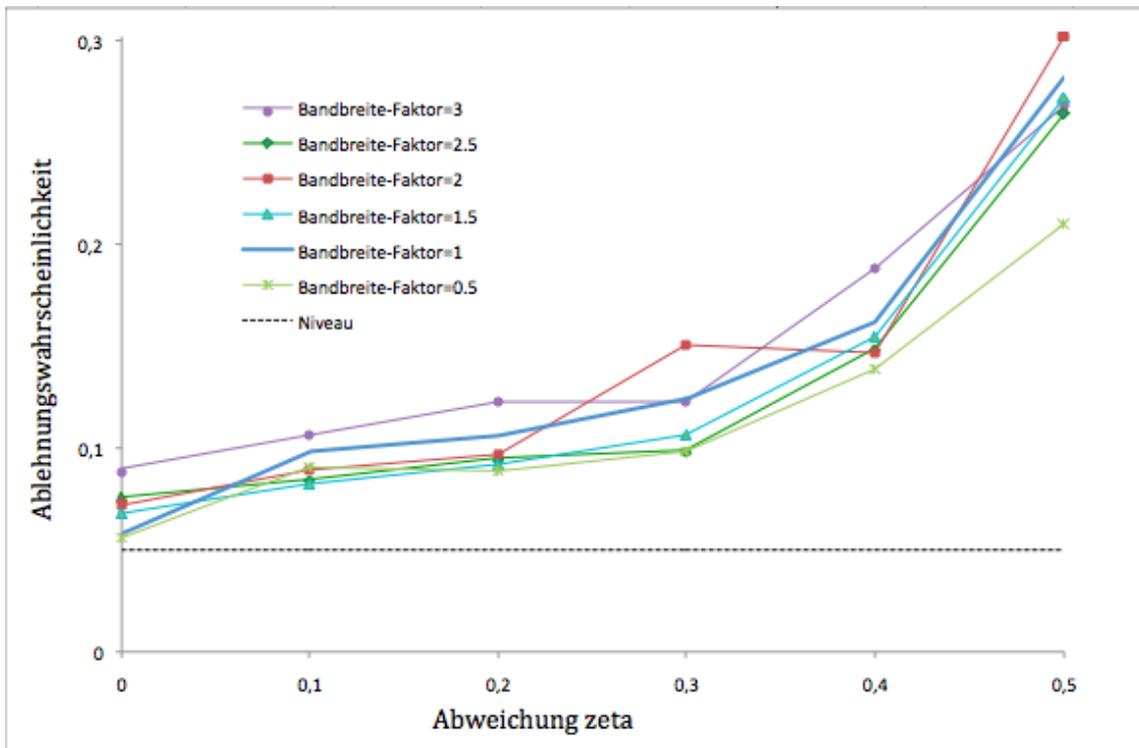


Abbildung 6.21: Ablehnungsw. für Modell (6.0.1) mit $c = 0.5, c = 1, c = 1.5, c = 2, c = 2.5, c = 3$

6.7 Anwendung auf einen echten Datensatz

Der in Kapitel 5 vorgestellte Test hat vielfältige Anwendungsmöglichkeiten in der (Finanz-)Wirtschaft, der Medizin oder der Klimaforschung, siehe Abschnitt 1.1.1 (S. 8).

Als Beispiel wird hier die Anwendung auf die Rendite des S&P 500 Index vorgestellt.

Der Aktienindex S&P 500 (Standard & Poor's 500) umfasst die Aktien von 500 der größten, börsennotierten US-amerikanischen Unternehmen und gehört zu den meistbeachteten Aktienindizes der Welt.

Die hier untersuchten Daten sind die täglichen Kursveränderungen des S&P 500 Index vom 01. Juli 1998 bis zum 30. Juni 2006 (2013 Daten). Sie werden in Form von sogenannten log-returns angegeben

$$X_j = \log(P_j) - \log(P_{j-1}), \quad j = 1, 2, \dots, \quad (6.0.11)$$

wobei P_j den Kurs zum Zeitpunkt j bezeichnet.

(Übliche Weise Renditen anzugeben, Vergleich z.B. Kapitel 14 Kreiß and Neuhaus [2006])

Die Daten wurden heruntergeladen von <http://finance.yahoo.com/>.

Für die täglichen log-returns des S&P 500 Index von Anfang Juli 1998 bis Ende Juni 2006 lehnt der Change-Point-Test die Nullhypothese keines Change-Points zum Niveau 5% ab.

In Abbildung 6.22 ist die durch (6.0.11) definierte Zeitreihe zu sehen. Der senkrechte Strich markiert den Zeitpunkt $[ns] = 1211$ für den der Testprozess $\hat{T}_n(s, t)$ (Vergleich Abschnitt 5.1, S. 70) maximal ist.

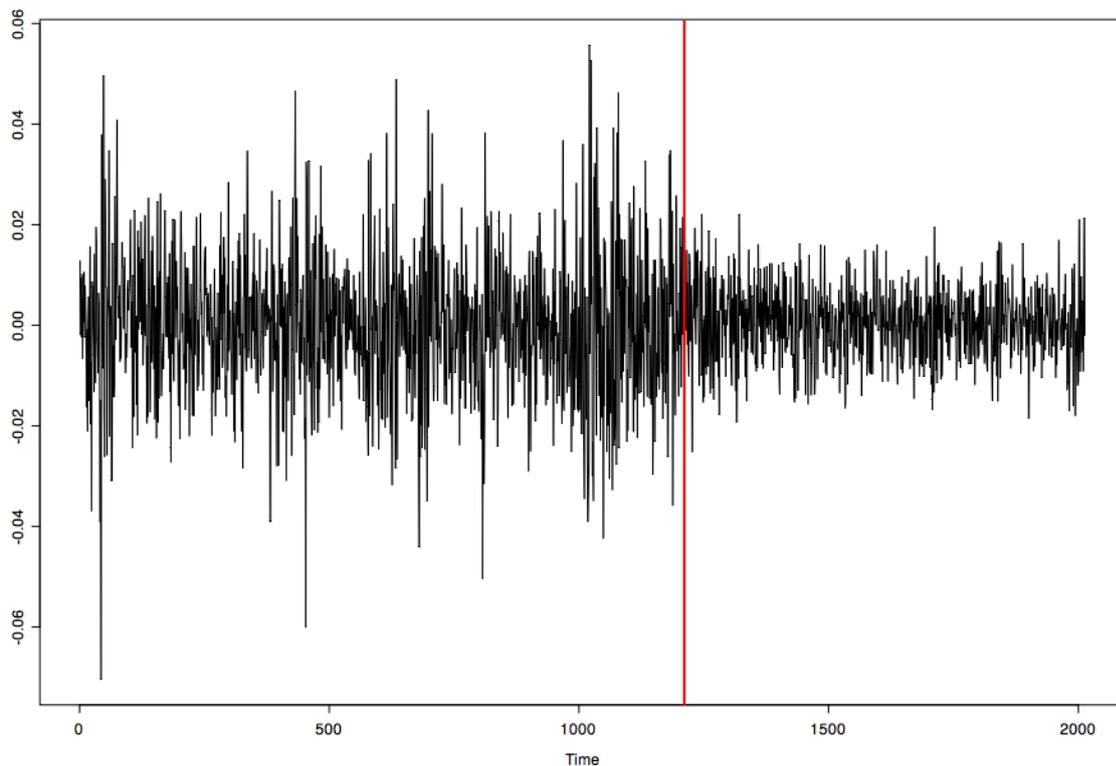


Abbildung 6.22: Tagesliche log-returns des S&P 500 Index von Anfang Juli 1998 bis Ende Juni 2006

Es kann also angenommen werden, dass ein Change-Point am 25. April 2003 aufgetreten ist.

In Kirch and Tadjuidje Kamgaing [2011] werden ebenfalls die Kursveränderung des S&P 500 Index von Juli 1998 bis Juni 2006 untersucht, allerdings als ein parametrisches Modell. Außerdem werden dort sogenannte transformed squared-log-returns betrachtet, eine Abwandlung der log-returns, die kleine Veränderungen stärker gewichtet.

Der dort verwandte Test ist ein sogenannter CUSUM Test, ein Test für einen Change-Point in der Verteilung der Beobachtungen, der auf Partialsummen von parametrisch geschätzten Residuen basiert.

Das Ergebnis in Kirch and Tadjuidje Kamgaing [2011] ist ebenfalls, dass die Nullhypothese keines Change-Points abgelehnt wird. Der Zeitpunkt des Change-Points wird dort auf den 23. Juli 2003 geschätzt.

7 Weitere Tests: Goodness-of-fit-Tests

In diesem Kapitel werden Goodness-of-fit-Tests, also Tests darauf, ob bestimmte Modellannahmen zutreffen, für Autoregressionsmodelle vorgestellt. Die asymptotischen Verteilungen solcher Teststatistiken können aus Korollar 4.5 (S. 61) (schwache Konvergenz) abgeleitet werden.

In Abschnitt 7.1 wird ein Test auf Normalverteilung der Innovationen in Autoregressionsmodellen vorgestellt.

In Abschnitt 7.2 und Abschnitt 7.3 wird eine Idee zu weiteren Goodness-of-fit-Tests in Autoregressionsmodellen und eine Übersicht über Literatur zu Goodness-of-fit-Tests (hauptsächlich in Regressionsmodellen) gegeben.

Abschnitt 7.4 gibt eine Literaturübersicht zum Thema Bootstrap-Verfahren. Bootstrap-Verfahren werden benötigt, um die asymptotischen Verteilungen der Teststatistiken zu implementieren (siehe Abschnitt 7.2 und 7.3).

7.1 Test auf Normalverteilung der Innovationen

Eine häufig getroffene Annahme für autoregressive Zeitreihen ist, dass die Innovationen normalverteilt sind. Im heteroskedastischen Autoregressionsmodell bedeutet dies die Standardnormalverteilung der Innovationen. Die Verteilungsfunktion der Standardnormalverteilung wird mit Φ und ihre Dichte mit ϕ bezeichnet.

Mit Hilfe von Korollar 4.5 (S. 61) kann ein Goodness-of-fit-Test für diese Hypothese durchgeführt werden. Für diesen Test lautet die Nullhypothese

$$H_0 : \quad \epsilon_1, \dots, \epsilon_n \text{ i.i.d. } \sim \mathcal{N}(0, 1),$$

die z.B. mit der Teststatistik

$$n \int \left(\hat{F}_n - \Phi \right)^2 d\Phi \quad (\text{Cramér-von-Mises-Teststatistik}) \quad (7.0.1)$$

überprüft werden kann.

Satz 7.1 Unter der Nullhypothese gilt

$$\sqrt{n} \left(\hat{F}_n - \Phi \right) \Rightarrow \mathbb{G}_\Phi$$

wobei \mathbb{G}_Φ ein zentrierter Gaußprozess ist, für den gilt

$$\text{Cov}(\mathbb{G}_\Phi(t_1), \mathbb{G}_\Phi(t_2)) = \Phi(\min(t_1, t_2)) - \Phi(t_1)\Phi(t_2) - \phi(t_1)\phi(t_2).$$

Korollar 7.2 Aus Satz 7.1 folgen für die asymptotische Verteilung von (7.0.1) unter der Nullhypothese die folgenden Quantile - und damit kritischen Werte für den Goodness-of-fit-Test:

Niveau des Tests γ	0.15	0.1	0.05	0.025	0.01
Kritischer Wert L	0.118	0.135	0.165	0.196	0.237

(siehe Stephens [1976], Abschnitt 10).

Beweis von Satz 7.1 Nach Korollar 4.5 (S. 61) gilt

$$\sqrt{n} \left(\hat{F}_n - \Phi \right) \stackrel{H_0}{\Rightarrow} \sqrt{n} \left(\hat{F}_n - F \right) \Rightarrow \mathbb{G}_\Phi$$

für einen zentrierten Gaußprozess \mathbb{G}_Φ mit

$$\begin{aligned} \text{Cov}(\mathbb{G}_\Phi(t_1), \mathbb{G}_\Phi(t_2)) &= \Phi(\min(t_1, t_2)) - \Phi(t_1)\Phi(t_2) \\ &+ \phi(t_1) \left(E[\epsilon_1 I\{\epsilon_1 \leq t_2\}] + t_1 E[(\epsilon_1^2 - 1)I\{\epsilon_1 \leq t_2\}] \right) \\ &+ \phi(t_2) \left(E[\epsilon_1 I\{\epsilon_1 \leq t_1\}] + t_2 E[(\epsilon_1^2 - 1)I\{\epsilon_1 \leq t_1\}] \right) \\ &+ \phi(t_1)\phi(t_2) \left(1 + (t_1 + t_2)E[\epsilon_1^3] + t_1 t_2 (E[\epsilon_1^4] - 1) \right). \end{aligned}$$

Für eine standardnormalverteilte Zufallsvariable ϵ gilt

$$\begin{aligned} E[\epsilon I\{\epsilon \leq t\}] &= \int_{-\infty}^{\infty} y I\{y \leq t\} \phi(y) dy \\ &= \int_{-\infty}^t y \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y^2}{2}\right) dy \\ &= \left[-\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y^2}{2}\right) \right]_{y=-\infty}^t \\ &= -\phi(t) \end{aligned}$$

für alle $t \in \mathbb{R}$ und mit partieller Integration

$$\begin{aligned} E[(\epsilon^2 - 1)I\{\epsilon \leq t\}] &= \int_{-\infty}^t (y^2 - 1) \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y^2}{2}\right) dy \\ &= \left[y \left(-\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y^2}{2}\right) \right) \right]_{y=-\infty}^t \\ &= -t\phi(t). \end{aligned}$$

Da die Dichte ϕ symmetrisch ist, ist $E[\epsilon^3] = 0$ für standardnormalverteilte Zufallsvariablen ϵ .

Ebenfalls mit partieller Integration ergibt sich

$$\begin{aligned} E[\epsilon^4] &= \int_{-\infty}^{\infty} y^4 \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y^2}{2}\right) dy \\ &= \left[y^3 \left(-\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y^2}{2}\right) \right) \right]_{y=-\infty}^{\infty} - \int_{-\infty}^{\infty} 3y^2 \left(-\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y^2}{2}\right) \right) dy \end{aligned}$$

$$\begin{aligned}
&= 3E[\epsilon^2] \\
&= 3.
\end{aligned}$$

Damit folgt für die Kovarianzstruktur

$$\begin{aligned}
&\Phi(\min(t_1, t_2)) - \Phi(t_1)\Phi(t_2) \\
&+ \phi(t_1) (E[\epsilon_1 I\{\epsilon_1 \leq t_2\}] + t_1 E[(\epsilon_1^2 - 1)I\{\epsilon_1 \leq t_2\}]) \\
&+ \phi(t_2) (E[\epsilon_1 I\{\epsilon_1 \leq t_1\}] + t_2 E[(\epsilon_1^2 - 1)I\{\epsilon_1 \leq t_1\}]) \\
&+ \phi(t_1)\phi(t_2) (1 + (t_1 + t_2)E[\epsilon_1^3] + t_1 t_2 (E[\epsilon_1^4] - 1)) \\
&= \Phi(\min(t_1, t_2)) - \Phi(t_1)\Phi(t_2) \\
&- 2\phi(t_1)\phi(t_2) - 2t_1 t_2 \phi(t_1)\phi(t_2) \\
&+ \phi(t_1)\phi(t_2) + 2t_1 t_2 \phi(t_1)\phi(t_2) \\
&= \Phi(\min(t_1, t_2)) - \Phi(t_1)\Phi(t_2) - \phi(t_1)\phi(t_2).
\end{aligned}$$

□

Beweis von Korollar 7.2 Nach Satz 7.1 gilt $\sqrt{n}(\hat{F}_n - \Phi) \Rightarrow \mathbb{G}_\Phi$, woraus mit dem Continuous-Mapping-Theorem

$$n \int (\hat{F}_n - \Phi)^2 d\Phi \xrightarrow[n \rightarrow \infty]{\mathcal{D}} \int \mathbb{G}_\Phi^2 d\Phi$$

folgt.

Mit der Substitutionsregel folgt weiter, dass

$$\begin{aligned}
&\int \mathbb{G}_\Phi^2 d\Phi \\
&= \int_{-\infty}^{\infty} (\mathbb{G}_\Phi(t))^2 \phi(t) dt \\
&= \int_0^1 (\mathbb{G}_\Phi(\Phi^{-1}(u)))^2 du.
\end{aligned}$$

Die asymptotische Verteilung der Teststatistik $n \int (\hat{F}_n - \Phi)^2 d\Phi$ ((7.0.1)) entspricht also der Verteilung von $\int_0^1 (\mathbb{G}_\Phi(\Phi^{-1}(u)))^2 du$.

Für den Prozess $\mathbb{G}_\Phi(\Phi^{-1}(u))$ gilt nach Satz 7.1 $E[\mathbb{G}_\Phi(\Phi^{-1}(u))] = 0$ für alle $u \in [0, 1]$ und

$$\begin{aligned}
&\text{Cov}(\mathbb{G}_\Phi(\Phi^{-1}(u_1)), \mathbb{G}_\Phi(\Phi^{-1}(u_2))) \\
&= \Phi(\min(\Phi^{-1}(u_1), \Phi^{-1}(u_2))) - \Phi(\Phi^{-1}(u_1))\Phi(\Phi^{-1}(u_2)) - \phi(\Phi^{-1}(u_1))\phi(\Phi^{-1}(u_2)) \\
&= \min(u_1, u_2) - u_1 u_2 - \phi(\Phi^{-1}(u_1))\phi(\Phi^{-1}(u_2)),
\end{aligned}$$

da Φ^{-1} streng monoton steigend ist.

Diese Verteilung entspricht nach Stephens [1976] Abschnitt 7 genau der Verteilung des Prozesses $(Y(u))_{u \in [0,1]}$, für den gilt, dass $\int_0^1 Y^2(u) du$ der Grenzwert von

$n \int \left(\frac{1}{n} \sum_{j=1}^n I\{Z_j \leq \cdot\} - F_{\hat{\mu}} \right)^2 dF_{\hat{\mu}}$ ist, wobei Z_1, \dots, Z_n i.i.d. Normal verteilt sind mit bekannter

Varianz und unbekanntem Erwartungswert, und $F_{\hat{\mu}}$ die Verteilungsfunktion der Normalverteilung $\mathcal{N}(\hat{\mu}, \text{Var}(Z_1))$, $\hat{\mu} = \frac{1}{n} \sum_{j=1}^n Z_j$ bezeichnet.

Die Quantile von $\int_0^1 Y^2(u) du$, die damit auch die Quantile von $\int_0^1 (\mathbb{G}_{\Phi}(\Phi^{-1}(u)))^2 du$, also der asymptotischen Verteilung von $n \int (\hat{F}_n - \Phi)^2 d\Phi$, sind, sind in Stephens [1976] Abschnitt 10 tabelliert.

Anmerkung: Die asymptotische Verteilung von $n \int (\hat{F}_n - \Phi)^2 d\Phi$ entspricht damit also der asymptotischen Verteilung eines Goodness-of-fit-Tests auf Normalverteilung von Beobachtungen mit geschätztem Erwartungswert und bekannter Varianz. Dies erscheint eventuell merkwürdig, da in \hat{F}_n nicht nur die Schätzung von m (also dem bedingten Erwartungswert), sondern auch die Schätzung von σ (also der Wurzel aus der bedingten Varianz) enthalten ist. Allerdings ist im Beweis zu Satz 7.1 zu sehen, dass die Terme, die in der Kovarianz durch die Schätzung von σ entstehen, im Fall der Normalverteilung gerade wegfallen.

□

7.2 Tests auf andere Innovationenverteilungen

Neben dem Test auf Normalverteilung der Innovationen wie in Abschnitt 7.1 (S. 104) kann es ebenfalls von Interesse sein, auf andere spezielle Verteilungen zu testen. Im Allgemeinen ist die Nullhypothese für so eine Art von Test

$$H_0 : \quad \exists \vartheta_0 \in \Theta : \epsilon_1, \dots, \epsilon_n \sim F_{\vartheta_0},$$

wobei Θ ein Parameterraum ist und $\{F_{\vartheta} : \vartheta \in \Theta\}$ eine Menge von Funktionen, die sich nur durch den Parameter ϑ unterscheiden.

Konkrete Testszenarien wären z.B. ein Test auf standardisiert Student-t-verteilte Innovationen (siehe Abschnitt 6.2.3, S. 93)

$$\exists \nu_0 \in \{3, 4, \dots\} : \epsilon_1, \dots, \epsilon_n \sim St(\nu_0),$$

oder auf Skew-Normal verteilte Innovationen (siehe Abschnitt 6.1.2, S. 87)

$$\exists \lambda_0 \in \mathbb{R} : \epsilon_1, \dots, \epsilon_n \sim SN \left(-\sqrt{\frac{2\pi(\lambda_0^2 + \lambda_0^4)}{\pi^2 + (2\pi^2 - 2\pi)\lambda_0^2 + (\pi^2 - 2\pi)\lambda_0^4}}, \sqrt{\frac{\pi(1 + \lambda_0^2)}{\pi + (\pi - 2)\lambda_0^2}}, \lambda_0 \right).$$

Ein solcher Test könnte mit einer Cramér-von-Mises-Teststatistik wie in Abschnitt 7.1

$$n \int (\hat{F}_n - F_{\hat{\vartheta}_n})^2 dF_{\hat{\vartheta}_n}$$

oder mit einer Kolmogorov-Smirnov-Teststatistik

$$\sqrt{n} \sup_{t \in \mathbb{R}} |\hat{F}_n(t) - F_{\hat{\vartheta}_n}(t)|$$

durchgeführt werden, wobei $\hat{\vartheta}_n$ jeweils ein Schätzer für den unbekannt Parameter ϑ_0 ist, der auf den nichtparametrisch geschätzten Residuen basiert.

Um die asymptotische Verteilung der Teststatistik unter der Nullhypothese zu bestimmen, kann die Aufteilung

$$\hat{F}_n - F_{\hat{\vartheta}_n} = \left(\hat{F}_n - F_{\vartheta_0} \right) - \left(F_{\hat{\vartheta}_n} - F_{\vartheta_0} \right)$$

verwendet werden. Für den ersten Teil gilt die stochastische Entwicklung aus Satz 4.2 (49), da $\hat{F}_n(t) - F_{\vartheta_0}(t) \stackrel{H_0}{=} \hat{F}_n(t) - F(t)$. Es kann angenommen werden, dass für den zweiten Teil ebenfalls eine stochastische Entwicklung bestimmt werden kann, die zusammen mit der Entwicklung des ersten Teils zu schwacher Konvergenz von $\sqrt{n} \left(\hat{F}_n - F_{\hat{\vartheta}_n} \right)$ führt (vgl. z.B. van der Vaart [2000] Abschnitt 19.3). Im Fall von nichtparametrischen homoskedastischen Regressionsmodellen mit unabhängigen Beobachtungen wurde dies in Neumeyer et al. [2006] gezeigt.

Die so erlangte asymptotische Verteilung der Teststatistiken ist im Allgemeinen nicht verteilungsfrei, d.h. sie hängt von der unbekanntem Verteilung F_{ϑ_0} der Innovationen ab (siehe z.B. Neumeyer et al. [2006] Theorem 3.1). Daher können die kritischen Werte für die Tests im Allgemeinen nicht direkt bestimmt werden. Dieses Problem kann mit Hilfe von Bootstrap-Verfahren, wie in Abschnitt 7.4 (S. 109) beschrieben, gelöst werden.

7.3 Tests auf bestimmte Formen der Regressions- oder Skalierungsfunktion

Weitere Goodness-of-fit-Tests entstehen durch die Frage nach der Form der Regressionsfunktion m oder der Skalierungsfunktion σ . Auch hier ist es von Interesse, auf bestimmte parametrische Formen zu testen, d.h. Nullhypothesen der Form

$$H_0 : \quad \exists \vartheta_0 \in \Theta : m = m_{\vartheta_0}$$

oder

$$H_0 : \quad \exists \vartheta_0 \in \Theta : \sigma = \sigma_{\vartheta_0}$$

zu überprüfen, wobei Θ ein Parameterraum ist und $\{m_{\vartheta} : \vartheta \in \Theta\}$ bzw. $\{\sigma_{\vartheta} : \vartheta \in \Theta\}$ eine Menge von Funktionen, die sich nur durch den Parameter ϑ unterscheiden.

Beispiele für solche Tests wären z.B. ein Test auf ein AR(1)-Modell (siehe Abschnitt 6.1, S. 83)

$$\begin{aligned} & \exists a_0 \in (-1, 1) : m(x) = a_0 \cdot x \quad \forall x \in \mathbb{R} \\ \text{und} & \quad \sigma \equiv 1, \end{aligned}$$

oder auf ein ARCH(1)-Modell (siehe Abschnitt 6.2, S. 90)

$$\begin{aligned} & \exists a_0 \in \mathbb{R}_{>0}, a_1 \in (0, 1) : \sigma(x) = \sqrt{a_0 + a_1 \cdot x^2} \quad \forall x \in \mathbb{R} \\ \text{und} & \quad m \equiv 0. \end{aligned}$$

Ein solcher Test kann unter anderem durch den Vergleich der empirischen Verteilungsfunktion der nichtparametrisch geschätzten Residuen und der empirischen Verteilungsfunktion der unter der Nullhypothese geschätzten Residuen durchgeführt werden. Dieses Vorgehen wurde z.B. für Regressionsmodelle mit unabhängigen Beobachtungen in van Keilegom et al. [2008] für Tests bzgl. m und in Dette et al. [2007] für Tests bzgl. σ angewendet.

Der Testprozess hat in diesem Fall die Form

$$\sqrt{n} \left(\hat{F}_n - \hat{F}_{n0} \right),$$

wobei \hat{F}_{n0} der empirische Prozess der unter der Nullhypothese geschätzten Residuen

$$\hat{\epsilon}_{0j} := \frac{X_j - m_{\hat{\vartheta}_n}(X_{j-1})}{\hat{\sigma}(X_{j-1})} \quad \text{bzw.} \quad \hat{\epsilon}_{0j} := \frac{X_j - \hat{m}(X_{j-1})}{\sigma_{\hat{\vartheta}_n}(X_{j-1})}$$

ist und $\hat{\vartheta}_n$ jeweils ein Schätzer für den unbekannt Parameter ϑ_0 , der auf den Beobachtungen (X_j) basiert.

Auch in Dette et al. [2009] wurde ein Testprozess dieser Art benutzt, um für ein Regressionsmodell mit abhängigen Beobachtungen zu testen, ob m proportional zu σ ist.

Mögliche Teststatistiken, die auf diesem Prozess basieren, sind, analog zu Abschnitt 7.2, Cramér-von-Mises oder Kolmogorov-Smirnov Teststatistiken.

Die asymptotische Verteilung dieser Teststatistiken kann analog zu Abschnitt 7.2 durch die Aufteilung

$$\hat{F}_n - \hat{F}_{n0} = \left(\hat{F}_n - F \right) - \left(\hat{F}_{n0} - F \right)$$

mit Satz 4.2 (S. 49) bestimmt werden, indem zusätzlich eine stochastische Entwicklung für $\left(\hat{F}_{n0} - F \right)$ hergeleitet wird. Dieses Vorgehen wurde auch in van Keilegom et al. [2008], in Dette et al. [2007] und in Dette et al. [2009] angewendet.

Auch für diese Art von Goodness-of-fit-Tests ist die asymptotische Verteilung der Teststatistiken im Allgemeinen nicht verteilungsfrei (Vergleich Abschnitt 7.2; siehe z.B. van Keilegom et al. [2008] Korrolar 3.2, Dette et al. [2007] Korrolar 3.3, Dette et al. [2009] Korrolar 1) und Bootstrap-Verfahren sollten angewendet werden.

7.4 Bootstrap-Verfahren

Tests mit Teststatistiken, deren (asymptotische) Verteilungen von unbekannt Faktoren abhängen, sind in der Praxis meist sehr schwierig zu implementieren. Daher werden in diesen Fällen häufig sogenannte Bootstrap-Verfahren angewendet, um die kritischen Werte für die Tests zu approximieren.

Die grundlegende Idee eines Bootstrap-Verfahrens ist es, aus den beobachteten Daten neue Stichproben zu generieren, und mit Hilfe dieser neuen Stichproben z.B. Quantile der Verteilung einer Teststatistik zu schätzen.

Zunächst werden Bootstrap-Verfahren für die in Abschnitt 7.2 und Abschnitt 7.3 zitierten Beispiele für Regressionsmodelle vorgestellt. Anschließend werden Bootstrap-Verfahren in Autoregressionsmodellen betrachtet, wobei dort Dette et al. [2009] das einzige Beispiel für Bootstrap in Goodness-of-fit-Tests basierend auf der Differenz von nichtparametrisch geschätzten Residuenprozessen liefert.

Bootstrap-Verfahren in Regressionsmodellen

Zunächst wird also vorgestellt, wie Bootstrap-Verfahren in einigen Beispielen für Regressionsmodelle mit unabhängigen Daten, also für Modelle der Form

$$Y_j = m(X_j) + \sigma(X_j)\epsilon_j, \quad (X_1, Y_1), \dots, (X_n, Y_n) \text{ i.i.d.}$$

benutzt werden.

Für Goodness-of-fit-Tests bezüglich der Regressions- oder Skalierungsfunktion, wie in Abschnitt 7.3, können sogenannte Bootstrap-Stichproben z.B. wie in Dette et al. [2007] oder in van Keilegom et al. [2008] generiert werden, indem die nichtparametrisch geschätzten Residuen standardisiert werden, und auf Basis dieser standardisierten Residuen Bootstrap-Fehler $\epsilon_1^*, \dots, \epsilon_n^*$ mit einer geglätteten Verteilung gebildet werden. Konkret wird dies z.B. in Dette et al. [2007] gemacht, indem aus den standardisierten Residuen mit Zurücklegen eine Stichprobe $\tilde{\epsilon}_1^*, \dots, \tilde{\epsilon}_n^*$ gezogen wird und daraus Bootstrap-Fehler mit

$$\epsilon_j^* := \tilde{\epsilon}_j^* + v_n Z_j, \quad \text{wobei } v_n \text{ ein Glättungsparameter} \\ \text{und } (Z_j) \text{ i.i.d. Zufallsvariablen (z.B. } \mathcal{N}(0,1)\text{-verteilt)}$$

gebildet werden.

Aus diesen Bootstrap-Fehlern werden Bootstrap-Daten durch

$$Y_j^* = m_{\hat{\vartheta}_n}(X_j) + \hat{\sigma}(X_j)\epsilon_j^* \quad (\text{falls bezüglich der Regressionsfunktion getestet wird})$$

bzw.

$$Y_j^* = \hat{m}(X_j) + \sigma_{\hat{\vartheta}_n}(X_j)\epsilon_j^* \quad (\text{falls bezüglich der Skalierungsfunktion getestet wird})$$

gebildet.

Für Goodness-of-fit-Tests bezüglich der Verteilung der Innovationen, wie in Abschnitt 7.2, kann eine Bootstrap-Stichprobe z.B. wie in Neumeyer et al. [2006] generiert werden, indem Bootstrap-Fehler $\epsilon_1^*, \dots, \epsilon_n^*$ mit Verteilung $F_{\hat{\vartheta}_n}$ erzeugt werden durch

$$\epsilon_j^* = F_{\hat{\vartheta}_n}^{-1}(U_j), \quad \text{wobei } U_1, \dots, U_n \text{ i.i.d. } \sim \mathcal{U}[0, 1],$$

und damit

$$Y_j^* = \hat{m}(X_j) + \hat{\sigma}(X_j)\epsilon_j^*$$

gebildet werden.

Dieses Verfahren wird jeweils B -mal durchgeführt, so dass B Bootstrap-Stichproben

$$(X_1, Y_1^{*(1)}), \dots, (X_n, Y_n^{*(1)}), \dots, (X_1, Y_1^{*(B)}), \dots, (X_n, Y_n^{*(B)})$$

entstehen. Für jede dieser Stichproben $(X_1, Y_1^*), \dots, (X_n, Y_n^*)$ wird die Teststatistik T , wie vorher für die originalen Beobachtungen, berechnet. Die Ergebnisse werden der Größe nach geordnet zu $T_{(1)}^*, \dots, T_{(B)}^*$, und daraus ein Test zum Niveau γ bestimmt, indem die Nullhypothese abgelehnt wird, wenn

$$T((X_1, Y_1), \dots, (X_n, Y_n)) > T_{(\lfloor B(1-\gamma) \rfloor)}^*.$$

Bootstrap-Verfahren in Autoregressionsmodellen

In den in dieser Arbeit betrachteten Autoregressionsmodellen

$$X_j = m(X_{j-1}) + \sigma(X_{j-1})\epsilon_j \tag{7.2.1}$$

können ebenfalls Bootstrap-Verfahren angewendet werden.

In Dette et al. [2009] wird z.B. die selbe Prozedur wie in Dette et al. [2007] angewendet, wobei in Dette et al. [2009] angenommen wird, dass die Beobachtungen abhängig sein dürfen, das Modell (7.2.1) also enthalten ist.

Weitere Beispiele für Bootstrap-Verfahren in Autoregressionsmodellen können für Goodness-of-fit-Tests, die nicht auf der Differenz von empirischen Prozessen von nichtparametrisch geschätzten Residuen basieren, in der Literatur gefunden werden.

In Kreiß et al. [2008] werden z.B. Bootstrap-Tests für verschiedene Hypothesen über die Regressionsfunktion vorgestellt, unter anderem Tests auf eine bestimmte parametrische Form (wie in Abschnitt 7.3). Anstelle des Vergleichs der empirischen Residuenprozesse wird hier eine Teststatistik verwendet, die auf $\frac{1}{nc_n} \sum_{j=1}^n K\left(\frac{\cdot - X_{j-1}}{c_n}\right) (X_j - \tilde{m}(X_{j-1}))$ basiert, wobei \tilde{m} die unter der Nullhypothese geschätzte Regressionsfunktion ist. Es wird gezeigt, dass unter der Nullhypothese (asymptotisch) $(X_j - \tilde{m}(X_{j-1}))$ durch ϵ_j ersetzt werden kann. Basierend auf diesem Ergebnis kann die Bootstrap-Approximation so gewählt werden, dass sie immer gegen die Verteilung der Teststatistik konvergiert, auch wenn die Nullhypothese nicht gilt.

Der Bootstrap wird konkret in Form eines Wild-Bootstrap generiert, indem Bootstrap-Innovationen $\epsilon_1^*, \dots, \epsilon_n^*$ erzeugt werden durch

$$\epsilon_j^* := \hat{\epsilon}_j \cdot Z_j,$$

wobei $(\hat{\epsilon}_j)$ nichtparametrisch geschätzt und (Z_j) i.i.d. Zufallsvariablen mit Erwartungswert 0 und Varianz 1 sind. Mit diesen Bootstrap-Innovationen wird dann die entsprechende Teststatistik auf Basis von $\frac{1}{nc_n} \sum_{j=1}^n K\left(\frac{\cdot - X_{j-1}}{c_n}\right) \epsilon_j^*$ berechnet, und wie im Fall der Regressionsmodelle die Nullhypothese abgelehnt, wenn die originale Teststatistik größer als das empirische $(1 - \gamma)$ -Quantil der Bootstrap-Teststatistik ist.

Anmerkung: Die Nutzung des Wild-Bootstrap erfolgt hier auf Grund der Heteroskedastizität, die nicht in Form von σ ausgedrückt wird (wie in dieser Arbeit), sondern allgemeiner indem eine Abhängigkeit der (ϵ_j) erlaubt wird.

Die Bestimmung der Bootstrap-Teststatistiken erfolgt in Kreiß et al. [2008] also ebenfalls (wie in Dette et al. [2009]) analog zum Vorgehen im regressiven Fall bei einem solchen Modell.

Dieses Vorgehen der Nutzung von Bootstrapmethoden aus der Regression in der Autoregression wird in Neumann and Kreiß [1998] und in Franke et al. [2002] untersucht.

In Neumann and Kreiß [1998] wird gezeigt, dass ein autoregressives Modell

$$X_j = m(X_{j-1}) + \epsilon_j, \quad (\text{wobei } (\epsilon_j) \text{ wie in Kreiß et al. [2008] abhängig sein dürfen})$$

durch ein regressives Modell mit festem Design

$$Y_j = m(x_{j-1}) + \tilde{\epsilon}_j, \quad (\text{wobei } (\tilde{\epsilon}_j) \text{ unabhängig sind } \tilde{\epsilon}_j \sim \epsilon_j | X_{j-1} = x_{j-1} \\ \text{und } x_0, \dots, x_n \text{ Realisation von } X_0, \dots, X_n)$$

approximiert werden kann, in dem Sinne, dass ein (nichtparametrischer) lokal polynomieller Schätzer für die Regressionsfunktion im zweiten Modell mit einer schnellen Rate gegen den selben Schätzer im ersten Modell konvergiert.

Dieses Ergebnis wird benutzt, um zu zeigen, dass der stochastische Teil des lokal polynomiellen Schätzers $(\sum_{j=1}^n \text{“Gewichtsfunktion”}(x, X_{j-1}, (X_0, \dots, X_n)) \cdot \epsilon_j)$ im autoregressiven Modell durch die Wild-Bootstrap-Version $(\sum_{j=1}^n \text{“Gewichtsfunktion”}(x, x_{j-1}, (x_0, \dots, x_n)) \cdot \epsilon_j^*)$ (wie in Kreiß et al.

[2008]) approximiert werden kann.

Mit Hilfe dieses Wild-Bootstrap werden Konfidenzintervalle für die Regressionsfunktion und Goodness-of-fit-Tests für die Regressionsfunktion hergeleitet.

In Franke et al. [2002] werden drei verschiedene Arten von Bootstrap-Verfahren betrachtet, die für ein Modell wie (7.2.1) die Verteilungen von Nadaraya-Watson-Schätzern \hat{m} und $\hat{\sigma}$ konsistent approximieren. Eine solche Approximation kann z.B. (wie in Neumann and Kreiß [1998]) benutzt werden, um Konfidenzintervalle für die Regressionsfunktion oder die Skalierungsfunktion herzuleiten.

Als erstes wird ein Autoregressions-Bootstrap betrachtet, der neue Daten unter Berücksichtigung der autoregressiven Struktur erzeugt:

$$X_j^* = \tilde{m}(X_{j-1}^*) + \tilde{\sigma}(X_{j-1}^*)\epsilon_j^*,$$

wobei \tilde{m} und $\tilde{\sigma}$ Schätzer für m und σ sind, die gewisse Bedingungen erfüllen, und ϵ_j^* Bootstrap-Innovationen, die (gegeben die ursprüngliche Stichprobe) i.i.d. sind, und asymptotisch die gleiche Verteilung haben, wie die originalen Innovationen.

Als zweites wird ein Regressions-Bootstrap betrachtet, der die autoregressive Struktur nicht berücksichtigt, sondern die Beobachtungen X_0, \dots, X_n als Design-Punkte verwendet:

$$X_j^* = \tilde{m}(X_{j-1}) + \tilde{\sigma}(X_{j-1})\epsilon_j^*,$$

wobei die (ϵ_j^*) wieder wie im Fall des Autoregressions-Bootstrap erzeugt werden.

Als drittes wird ein Wild-Bootstrap wie in Kreiß et al. [2008] und Neumann and Kreiß [1998] betrachtet:

$$X_j^* = \tilde{m}(X_{j-1}) + \epsilon_j^*,$$

wobei $\epsilon_j^* = (X_j - \hat{m}(X_{j-1})) \cdot Z_j$, (Z_j) i.i.d. mit Erwartungswert 0 und Varianz 1.

Für alle drei Modelle wird gezeigt, dass die Bootstrap-Approximation der Verteilung von $\sqrt{nc_n}(\hat{m} - m)$ und $\sqrt{nc_n}(\hat{\sigma}^2 - \sigma^2)$ unter geeigneten Voraussetzungen konsistent ist.

Ein weiteres Beispiel, in dem die autoregressive Struktur beim Erzeugen der Bootstrap-Stichproben berücksichtigt wird, findet sich in Neumann and Paparoditis [2008]. Hier werden für Modelle wie (7.2.1) Goodness-of-fit-Test für parametrische Formen der Regressions- und der Skalierungsfunktion betrachtet, die auf der Differenz zwischen der vollkommen nichtparametrisch geschätzten Übergangswahrscheinlichkeit und der unter der Nullhypothese geschätzten Übergangswahrscheinlichkeit basieren:

$$\sup_{x,y \in \mathbb{R}} \left| \frac{1}{\sqrt{n}} \sum_{j=1}^n I\{X_{j-1} \leq x\} \left(I\{X_j \leq y\} - \hat{F}_{X,H_0}(y) \right) \right|,$$

wobei \hat{F}_{X,H_0} mit Hilfe des empirischen Prozesses der unter der Nullhypothese geschätzten Residuen gebildet wird.

Für diese Teststatistik wird die Verteilung durch ein Bootstrap-Verfahren approximiert, das, wie das erste Verfahren in Franke et al. [2002], die autoregressive Struktur berücksichtigt:

$$X_j^* = m_{\hat{\vartheta}_n}(X_{j-1}^*) + \sigma_{\hat{\vartheta}_n}(X_{j-1}^*)\epsilon_j^*,$$

wobei die ϵ_j^* durch den empirischen Prozess der unter der Nullhypothese geschätzten, standardisierten Residuen erzeugt werden.

Es ist also für autoregressive Modelle wie (7.2.1) sowohl möglich, direkt die Bootstrap-Verfahren aus der Regression zu verwenden, indem man die Beobachtungen X_0, \dots, X_n als Design-Punkte für

die Bootstrap-Stichproben verwendet, als auch Bootstrap-Verfahren zu verwenden, die die autoregressive Struktur beibehalten.

Für neue Teststatistiken, wie die in Abschnitt 7.2 und in Abschnitt 7.3 vorgestellten, muss allerdings jeweils erst bewiesen werden, dass bestimmte Bootstrap-Verfahren für das betrachtete Modell funktionieren.

8 Literaturüberblick

In diesem Kapitel werden zwei Artikel vorgestellt, die eng verwandt mit den Ergebnissen dieser Arbeit sind. Allerdings befassen sich beide Artikel mit nicht-sequentiellen Prozessen, so dass z.B. eine Anwendung auf Change-Point-Tests mit den Ergebnissen der Artikel nicht möglich ist.

Müller, Schick und Wefelmeyer (2009)

In Müller et al. [2009] werden Modelle der Form

$$X_j = m(X_{j-1}) + \epsilon_j$$

betrachtet, wobei ϵ_j unabhängig von X_{j-1}, X_{j-2}, \dots ist. Das heteroskedastische Modell in dieser Arbeit ist in Müller et al. [2009] also nicht enthalten. Die Regressionsfunktion m wird, wie in dieser Arbeit, ebenfalls nichtparametrisch geschätzt, wobei allerdings kein lokal konstanter Schätzer, wie der Nadaraya-Watson-Schätzer in dieser Arbeit, sondern ein lokal linearer Schätzer verwendet wird. Das zentrale Ergebnis des Artikels ist eine stochastische Entwicklung von $(\hat{F}_n - F)$. Es wird gezeigt, dass

$$\hat{F}_n(t) - F(t) = \frac{1}{n} \sum_{j=1}^n (I\{\epsilon_j \leq t\} - F(t)) + f(t) \frac{1}{n} \sum_{j=1}^n \epsilon_j + o_p\left(\frac{1}{\sqrt{n}}\right).$$

Diese Entwicklung entspricht der Entwicklung in dieser Arbeit bis auf den Term $f(t)t \frac{1}{2n} \sum_{j=1}^n (\epsilon_j^2 - 1)$, der aus der Heteroskedastizität entsteht.

Die Voraussetzungen in Müller et al. [2009] sind sehr ähnlich zu denen in dieser Arbeit. Eine auf den ersten Blick unterschiedliche Bedingung ist die geometrische Ergodizität (in Müller et al. [2009]) und die Mischungsbedingung (in dieser Arbeit). Wie in Abschnitt 2.2 (S. 11) gezeigt, folgt aus der geometrischen Ergodizität allerdings α -mischend mit exponentiell schnell fallenden Mischungskoeffizienten, also die Voraussetzung in dieser Arbeit.

Dette, Pardo-Fernández und van Keilegom (2009)

In diesem Artikel (Dette et al. [2009]) geht es um Goodness-of-fit-Tests für multiplikative Modelle mit abhängigen Daten. Es werden Modelle der Form

$$Y_j = m(X_j) + \sigma(X_j)\epsilon_j$$

betrachtet, wobei der Fall $X_j = Y_{j-1}$ (wie in dieser Arbeit) enthalten ist. Die Funktionen m und σ werden, wie in dieser Arbeit, nichtparametrisch geschätzt.

Das zentrale Ergebnis des Artikels ist eine stochastische Entwicklung für $(\hat{F}_n - \hat{F}_{n0})$ und daraus

folgend schwache Konvergenz von $\sqrt{n}(\hat{F}_n - \hat{F}_{n0})$, wobei \hat{F}_{n0} der empirische Prozess der Residuen unter der Nullhypothese

$$H_0 : m(\cdot) = c\sigma(\cdot), \quad c \in \mathbb{R}_+$$

ist.

Lemma 1 des Artikels gibt die folgende stochastische Entwicklung für $(\hat{F}_n - F)$:

$$\begin{aligned} \hat{F}_n(t) - F(t) &= \frac{1}{E[w(X_1)]} \frac{1}{n} \sum_{j=1}^n w(X_j) (I\{\epsilon_j \leq t\} - F(t)) \\ &\quad + \frac{f(t)}{E[w(X_1)]} \int \frac{w(x)}{\sigma(x)} (t(\hat{\sigma}(x) - \sigma(x)) + \hat{m}(x) - m(x)) f_{X_1}(x) dx \\ &\quad + o_p\left(\frac{1}{\sqrt{n}}\right). \end{aligned}$$

Diese Darstellung ist vergleichbar mit

$$\begin{aligned} \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \left(\hat{F}_{\lfloor ns \rfloor}(t) - F(t) \right) &= \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \left(\sum_{j=1}^{\lfloor ns \rfloor} \bar{w}_{j\lfloor ns \rfloor} I\{\epsilon_j \leq t\} - F(t) \right) \\ &\quad + \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} f(t) \sum_{j=1}^{\lfloor ns \rfloor} \bar{w}_{j\lfloor ns \rfloor} \left(\frac{\hat{m} - m}{\sigma}(X_{j-1}) + t \frac{\hat{\sigma} - \sigma}{\sigma}(X_{j-1}) \right) \\ &\quad + o_p\left(\frac{1}{\sqrt{n}}\right) \end{aligned}$$

aus dem Beweis von Satz 4.2 (S. 50), was für $s = 1$

$$\begin{aligned} \hat{F}_n(t) - F(t) &= \sum_{j=1}^n \bar{w}_{jn} I\{\epsilon_j \leq t\} - F(t) \\ &\quad + f(t) \sum_{j=1}^n \bar{w}_{jn} \left(\frac{\hat{m} - m}{\sigma}(X_{j-1}) + t \frac{\hat{\sigma} - \sigma}{\sigma}(X_{j-1}) \right) \\ &\quad + o_p\left(\frac{1}{\sqrt{n}}\right) \\ &= \frac{1}{\sum_{k=1}^n w_n(X_{j-1})} \sum_{j=1}^n w_n(X_{j-1}) (I\{\epsilon_j \leq t\} - F(t)) \\ &\quad + \frac{f(t)}{\sum_{k=1}^n w_n(X_{j-1})} \sum_{j=1}^n \frac{w_n(X_{j-1})}{\sigma(X_{j-1})} (t(\hat{\sigma}(X_{j-1}) - \sigma(X_{j-1})) + \hat{m}(X_{j-1}) - m(X_{j-1})) \\ &\quad + o_p\left(\frac{1}{\sqrt{n}}\right) \end{aligned}$$

entspricht.

Diese beiden stochastischen Entwicklungen für $(\hat{F}_n - F)$ unterscheiden sich zum Einen in dem Faktor

$$\frac{1}{n \cdot E[w(X_1)]} \text{ (Dette et al. [2009]),} \quad \frac{1}{\sum_{k=1}^n w_n(X_{j-1})} \text{ (diese Arbeit)}$$

und zum Anderen im zweiten Faktor der zweiten Zeile

$$n \cdot \int \frac{w(x)}{\sigma(x)} (t(\hat{\sigma}(x) - \sigma(x)) + \hat{m}(x) - m(x)) f_{X_1}(x) dx \text{ (Dette et al. [2009])}$$

$$\sum_{j=1}^n \frac{w_n(X_{j-1})}{\sigma(X_{j-1})} (t(\hat{\sigma}(X_{j-1}) - \sigma(X_{j-1})) + \hat{m}(X_{j-1}) - m(X_{j-1})) \quad (\text{diese Arbeit}).$$

Die Mittelwert-Bildung ($\sum_{j=1}^n$) über alle X_{j-1} in dieser Arbeit wird also in Dette et al. [2009] durch den Erwartungswert bzgl. X_{j-1} ersetzt.

Das in dieser Arbeit insbesondere durch Lemma 9.1 (S. 117) Gezeigte unterscheidet sich also in der genauen Darstellung von dem in Dette et al. [2009] Gezeigtem.

Außerdem wurde die Darstellung in Dette et al. [2009] dafür verwendet, eine stochastische Entwicklung für $(\hat{F}_n - \hat{F}_{n0})$ zu bekommen, nicht für $(\hat{F}_n - F)$, wie in dieser Arbeit. Eine stochastische Entwicklung von $(\hat{F}_n - F)$, die aus einem empirischen Prozess unabhängiger Zufallsvariablen besteht, wie in in dieser Arbeit (siehe Satz 4.2, S. 49), ist in Dette et al. [2009] nicht enthalten.

Allerdings wäre es auch mit der Darstellung aus Dette et al. [2009] möglich (wenn w durch w_n ersetzt würde), die Darstellung

$$\hat{F}_n(t) - F(t) = \frac{1}{n} \sum_{j=1}^n (I\{\epsilon_j \leq t\} - F(t)) + f(t) \frac{1}{n} \sum_{j=1}^n \epsilon_j + f(t)t \frac{1}{2n} \sum_{j=1}^n (\epsilon_j^2 - 1) + o_p\left(\frac{1}{\sqrt{n}}\right)$$

zu bekommen, indem analog zum Beweis von Satz 4.2 vorgegangen wird, und die Integraldarstellung durch eine Taylorentwicklung aufgelöst wird.

Die Gewichtsfunktionen w (in Dette et al. [2009]) und w_n (in dieser Arbeit) unterscheiden sich darin, dass die Gewichtsfunktion w einen kompakten Träger R_w hat, der nicht abhängig von n ist. Das Intervall, das vorgibt, welche X_j für die Schätzung benutzt werden, wird also unabhängig von n festgesetzt und somit mit wachsendem n nicht größer, im Gegensatz zum Ansatz in dieser Arbeit.

Das hat zu Folge, dass in der stochastischen Entwicklung von $(\hat{F}_n - \hat{F}_{n0})$ in Dette et al. [2009] die Gewichtsfunktion w enthalten ist, genau wie in der asymptotischen Verteilung von $\sqrt{n}(\hat{F}_n - \hat{F}_{n0})$ und in den asymptotischen Verteilungen der darauf basierenden Teststatistiken.

In dieser Arbeit taucht die Gewichtsfunktion w_n in der stochastischen Entwicklung und den asymptotischen Verteilungen nicht auf, was den Vorteil hat, dass z.B. die asymptotische Verteilung der Change-Point-Teststatistik in 5.2 (S. 71) eine bekannte, tabellierte Verteilung ist.

Ein auffälliger Unterschied in den Voraussetzungen von Dette et al. [2009] und dieser Arbeit ist die Mischungsbedingung.

In Dette et al. [2009] wird gefordert, dass die $(Y_j, X_j)_{j \in \mathbb{Z}}$ β -mischend sind mit $\beta(n) = O(n^{-\beta})$ für ein $\beta > 2$.

In dieser Arbeit wird gefordert, dass die $(X_j)_{j \in \mathbb{Z}}$ α -mischend sind mit $\alpha(n) = O(\alpha^n)$ für ein $\alpha > 1$.

In dieser Arbeit wird also für die Mischungskoeffizienten eine wesentlich schnellere Konvergenzrate gefordert. Diese exponentielle Konvergenzrate wird für die Beweise von Lemma 10.2 (S. 145), Lemma 10.3 (S. 158) und Lemma 10.4 (S. 166) benötigt. Für den Fall $s = 1$, der in Dette et al. [2009] betrachtet wird, können diese Lemmata mit Hilfe von Resultaten über Martingalzuwächse bzw. mit der Tchebychev-Ungleichung ohne jegliche Mischungsbedingung bewiesen werden. Für diesen Fall könnte also auch in dieser Arbeit die Forderung an die Konvergenzrate der Mischungskoeffizienten auf eine polynomiale Rate verringert werden.

9 Lemmata

Lemma 9.1 Für die Dichten f_{ϵ_j} der Innovationen ϵ_j gelte $\frac{1}{n} \sum_{j=1}^n \sup_{t \in \mathbb{R}} |f_{\epsilon_j}(t)| = O(1)$, sowie $\frac{1}{n} \sum_{j=1}^n \sup_{t \in \mathbb{R}} |f_{\epsilon_j}(t)t| = O(1)$.

Für die Schätzer \hat{m} und $\hat{\sigma}$ und die Funktionen m und σ gelte

- $\sup_{x \in I_n} \left| \frac{\hat{m}(x) - m(x)}{\sigma(x)} \right| = O_P(z_{n1}) = o_P(1)$ und $\sup_{x \in I_n} \left| \frac{\hat{\sigma}(x) - \sigma(x)}{\sigma(x)} \right| = O_P(z_{n2}) = o_P(1)$
- $\sup_{x \in I_n} \left| \frac{\partial}{\partial x} \left(\frac{\hat{m}(x) - m(x)}{\sigma(x)} \right) \right| = o_P(1)$ und $\sup_{x \in I_n} \left| \frac{\partial}{\partial x} \left(\frac{\hat{\sigma}(x) - \sigma(x)}{\sigma(x)} \right) \right| = o_P(1)$
- $\sup_{x, y \in I_n, x \neq y} \frac{\left| \frac{\partial}{\partial x} \left(\frac{\hat{m}(x) - m(x)}{\sigma(x)} \right) - \frac{\partial}{\partial y} \left(\frac{\hat{m}(y) - m(y)}{\sigma(y)} \right) \right|}{|y-x|^\delta} = o_P(1)$
und $\sup_{x, y \in I_n, x \neq y} \frac{\left| \frac{\partial}{\partial x} \left(\frac{\hat{\sigma}(x) - \sigma(x)}{\sigma(x)} \right) - \frac{\partial}{\partial y} \left(\frac{\hat{\sigma}(y) - \sigma(y)}{\sigma(y)} \right) \right|}{|y-x|^\delta} = o_P(1)$

mit δ , z_{n1} und z_{n2} für die gilt

$$n^{\frac{1}{2}(\frac{1}{1+\delta}-1)} \log(n)^{\frac{1}{1+\delta}} (b_n - a_n) = o(1) \quad \text{und} \quad n^{\frac{1}{2} \frac{1}{1+\delta}} (b_n - a_n) \log(n)^{\frac{1}{1+\delta}+1} (z_{n1} + z_{n2}) = o(1),$$

wobei $I_n = [a_n, b_n]$.

Anmerkung: Unter den Voraussetzungen von Lemma 3.4 (S. 35) und für $\sup_{x \in I_n} |m^\mu(x)| = O(\log(n)^{r_1})$, $\sup_{x \in I_n} |\sigma^\mu(x)| = O(\log(n)^{r_2})$, $\mu = 0, 1, 2, 3, 4$, $\frac{1}{\inf_{x \in I_n} \sigma(x)} = O(\log(n)^{r_3})$,

$\frac{1}{\inf_{x \in I_n} \frac{1}{n} \sum_{j=1}^n f_{X_{j-1}}(x)} = O(\log(n)^{r_4})$, $(b_n - a_n) = O(\log(n)^{r_5})$ und $c_n \sim n^{-\frac{1}{4}} \log(n)^{r_6}$, mit

$|r_1|, \dots, |r_6| < \infty$, ist $z_{n1} + z_{n2} = O(n^{-\frac{3}{8}} \log(n)^{r_7})$, für ein $|r_7| < \infty$ (Lemma 3.4), und damit bedeutet die Forderung $\delta > \frac{1}{3}$.

Sind also zusätzlich zu den Bedingungen an die Dichten f_{ϵ_j} die Voraussetzungen von Lemma 3.4 erfüllt und erfüllen die Funktionen μ und σ , die Dichten $f_{X_{j-1}}$ und das Intervall $I_n = [a_n, b_n]$ die obigen Konvergenzraten, so sind die Voraussetzungen dieses Lemmas (9.1) erfüllt.

Dann gilt

$$\sup_{s \in [0,1]} \sup_{t \in \mathbb{R}} \left| \frac{1}{n} \sum_{j=1}^{[ns]} w_{nj} \left(I\{\hat{\epsilon}_j \leq t\} - F_{\epsilon_j} \left(t \frac{\hat{\sigma}}{\sigma}(X_{j-1}) + \frac{\hat{m} - m}{\sigma}(X_{j-1}) \right) - I\{\epsilon_j \leq t\} + F_{\epsilon_j}(t) \right) \right| = o_p \left(\frac{1}{\sqrt{n}} \right).$$

Beweis Der Ausdruck lässt sich in folgender Weise darstellen,

$$\sup_{s \in [0,1]} \sup_{t \in \mathbb{R}} \left| \frac{1}{n} \sum_{j=1}^n w_{nj} \left(I\{\epsilon_j \leq t \cdot \hat{d}_{n2}(X_{j-1}) + \hat{d}_{n1}(X_{j-1})\} - I\{\epsilon_j \leq t\} \right) \right|$$

$$-F_{\epsilon_j}(t \cdot \hat{d}_{n2}(X_{j-1}) + \hat{d}_{n1}(X_{j-1})) + F_{\epsilon_j}(t) \Big| I \left\{ \frac{j}{n} \leq s \right\} \Big|,$$

mit $\hat{d}_{n1}(X_{j-1}) := \frac{\hat{m}(X_{j-1}) - m(X_{j-1})}{\hat{\sigma}(X_{j-1})}$ und $\hat{d}_{n2}(X_{j-1}) := \frac{\hat{\sigma}(X_{j-1})}{\hat{\sigma}(X_{j-1})}$, und die Notation sich mit

$$H(s, t, d_1, d_2) := \frac{1}{n} \sum_{j=1}^n w_{nj} (I\{\epsilon_j \leq t \cdot d_2(X_{j-1}) + d_1(X_{j-1})\} - F_{\epsilon_j}(td_2(X_{j-1}) + d_1(X_{j-1}))) I \left\{ \frac{j}{n} \leq s \right\}$$

zu

$$\sup_{s \in [0,1]} \sup_{t \in \mathbb{R}} |H(s, t, \hat{d}_{n1}, \hat{d}_{n2}) - H(s, t, 0, 1)|$$

vereinfachen.

Damit ist zu zeigen, dass für alle $\eta > 0$

$$P \left(\sqrt{n} \sup_{s \in [0,1]} \sup_{t \in \mathbb{R}} |H(s, t, \hat{d}_{n1}, \hat{d}_{n2}) - H(s, t, 0, 1)| > \eta \right) = o(1)$$

gilt.

Für die Innovationen ϵ_j gilt

$$\begin{aligned} P \left(\max_{1 \leq j \leq n} |\epsilon_j| > \sqrt{n} \log(n) \right) &\leq \sum_{j=1}^n P(|\epsilon_j| > \sqrt{n} \log(n)) \\ &= \frac{1}{\log(n)} \frac{1}{n} \sum_{j=1}^n n \log(n) P(|\epsilon_j| > \sqrt{n} \log(n)) \\ &= \frac{1}{\log(n)} \frac{1}{n} \sum_{j=1}^n n \log(n) P(\epsilon_j^2 > n \log(n)^2) \\ &= \frac{1}{\log(n)} \frac{1}{n} \sum_{j=1}^n E[n \log(n) I\{\epsilon_j^2 > n \log(n)^2\}] \\ &\leq \frac{1}{\log(n)} \frac{1}{n} \sum_{j=1}^n E[\epsilon_j^2 I\{\epsilon_j^2 > n \log(n)^2\}], \end{aligned}$$

da der Erwartungswert monoton wachsend ist, und die Funktion nur ungleich Null, wenn $\epsilon_j^2 > n \log(n)^2$. Da außerdem mit $\frac{1}{n} \sum_{j=1}^n E[\epsilon_j^2] = 1 = O(1)$ auch $\frac{1}{n} \sum_{j=1}^n E[\epsilon_j^2 I\{\epsilon_j^2 > n \log(n)^2\}] = O(1)$ gilt, ist

$$P \left(\max_{1 \leq j \leq n} |\epsilon_j| > \sqrt{n} \log(n) \right) = O \left(\frac{1}{\log(n)} \right) = o(1).$$

Es gilt ebenfalls

$$P \left(\sup_{x \in I_n} |\hat{d}_{n1}(x)| \geq 1 \right) = o(1) \quad \text{und} \quad P \left(\inf_{x \in I_n} \hat{d}_{n2}(x) \leq \frac{1}{2} \right) = o(1),$$

da $\sup_{x \in I_n} |\hat{d}_{n1}(x)| = o_P(1)$ und $\sup_{x \in I_n} |\hat{d}_{n2}(x) - 1| = o_P(1)$ nach Voraussetzung.

Da also für $B := \left\{ \sqrt{n} \sup_{s \in [0,1]} \sup_{t \in \mathbb{R}} |H(s, t, \hat{d}_{n1}, \hat{d}_{n2}) - H(s, t, 0, 1)| > \eta \right\}$ gilt

$$P(B) = P \left(B \cap \left(\max_{1 \leq j \leq n} |\epsilon_j| \leq \sqrt{n} \log(n), \sup_{x \in I_n} |\hat{d}_{n1}(x)| < 1, \inf_{x \in I_n} \hat{d}_{n2}(x) > \frac{1}{2} \right) \right)$$

$$\begin{aligned}
& +P\left(B \cap \left(\left(\max_{1 \leq j \leq n} |\epsilon_j| > \sqrt{n} \log(n)\right) \cup \left(\sup_{x \in I_n} |\hat{d}_{n1}(x)| \geq 1\right) \cup \left(\inf_{x \in I_n} \hat{d}_{n2}(x) \leq \frac{1}{2}\right)\right)\right) \\
& \leq P\left(B \cap \left(\max_{1 \leq j \leq n} |\epsilon_j| \leq \sqrt{n} \log(n), \sup_{x \in I_n} |\hat{d}_{n1}(x)| < 1, \inf_{x \in I_n} \hat{d}_{n2}(x) > \frac{1}{2}\right)\right) \\
& \quad +P\left(\left(\max_{1 \leq j \leq n} |\epsilon_j| > \sqrt{n} \log(n)\right) \cup \left(\sup_{x \in I_n} |\hat{d}_{n1}(x)| \geq 1\right) \cup \left(\inf_{x \in I_n} \hat{d}_{n2}(x) \leq \frac{1}{2}\right)\right) \\
& \leq P\left(B \cap \left(\max_{1 \leq j \leq n} |\epsilon_j| \leq \sqrt{n} \log(n), \sup_{x \in I_n} |\hat{d}_{n1}(x)| < 1, \inf_{x \in I_n} \hat{d}_{n2}(x) > \frac{1}{2}\right)\right) \\
& \quad +P\left(\max_{1 \leq j \leq n} |\epsilon_j| > \sqrt{n} \log(n)\right) + P\left(\sup_{x \in I_n} |\hat{d}_{n1}(x)| \geq 1\right) + P\left(\inf_{x \in I_n} \hat{d}_{n2}(x) \leq \frac{1}{2}\right) \\
& = P\left(B \cap \left(\max_{1 \leq j \leq n} |\epsilon_j| \leq \sqrt{n} \log(n), \sup_{x \in I_n} |\hat{d}_{n1}(x)| < 1, \inf_{x \in I_n} \hat{d}_{n2}(x) > \frac{1}{2}\right)\right) \\
& \quad +o(1),
\end{aligned}$$

kann man für die weitere Berechnung davon ausgehen, dass $|\epsilon_j| \leq \sqrt{n} \log(n)$ für alle $j = 1, \dots, n$, $\sup_{x \in I_n} |\hat{d}_{n1}(x)| < 1$ und $\inf_{x \in I_n} \hat{d}_{n2}(x) > \frac{1}{2}$.

Unter diesen Voraussetzungen ist für alle $j = 1, \dots, n$ ($n > 4$)

$$-\sqrt{n} \log(n) \cdot \hat{d}_{n2}(X_{j-1}) + \hat{d}_{n1}(X_{j-1}) < -\sqrt{n} \log(n) \cdot \frac{1}{2} + 1 < \epsilon_j,$$

sowie

$$\epsilon_j < \sqrt{n} \log(n) \cdot \frac{1}{2} - 1 < \sqrt{n} \log(n) \cdot \hat{d}_{n2}(X_{j-1}) + \hat{d}_{n1}(X_{j-1}),$$

und damit für alle $t > \sqrt{n} \log(n)$

$$\begin{aligned}
1 & \geq I\{\epsilon_j \leq t \cdot \hat{d}_{n2}(X_{j-1}) + \hat{d}_{n1}(X_{j-1})\} \geq I\{\epsilon_j \leq \sqrt{n} \log(n) \cdot \hat{d}_{n2}(X_{j-1}) + \hat{d}_{n1}(X_{j-1})\} \\
& \geq I\{\epsilon_j \leq \sqrt{n} \log(n) \cdot \frac{1}{2} - 1\} = 1 = I\{\epsilon_j \leq t\},
\end{aligned}$$

und für alle $t < -\sqrt{n} \log(n)$

$$\begin{aligned}
0 & \leq I\{\epsilon_j \leq t \cdot \hat{d}_{n2}(X_{j-1}) + \hat{d}_{n1}(X_{j-1})\} \leq I\{\epsilon_j \leq -\sqrt{n} \log(n) \cdot \hat{d}_{n2}(X_{j-1}) + \hat{d}_{n1}(X_{j-1})\} \\
& \leq I\{\epsilon_j \leq -\sqrt{n} \log(n) \cdot \frac{1}{2} + 1\} = 0 = I\{\epsilon_j \leq t\},
\end{aligned}$$

also

$$I\{\epsilon_j \leq t \cdot \hat{d}_{n2}(X_{j-1}) + \hat{d}_{n1}(X_{j-1})\} - I\{\epsilon_j \leq t\} = 0 \quad \forall |t| > \sqrt{n} \log(n), \quad j = 1, \dots, n.$$

Hieraus folgt

$$\begin{aligned}
& \sup_{s \in [0,1]} \sup_{|t| > \sqrt{n} \log(n)} |H(s, t, \hat{d}_{n1}, \hat{d}_{n2}) - H(s, t, 0, 1)| \\
& = \sup_{s \in [0,1]} \sup_{|t| > \sqrt{n} \log(n)} \left| \frac{1}{n} \sum_{j=1}^n w_{nj} \left(I\{\epsilon_j \leq t \cdot \hat{d}_{n2}(X_{j-1}) + \hat{d}_{n1}(X_{j-1})\} - I\{\epsilon_j \leq t\} \right. \right. \\
& \quad \left. \left. - F_{\epsilon_j}(t \cdot \hat{d}_{n2}(X_{j-1}) + \hat{d}_{n1}(X_{j-1})) + F_{\epsilon_j}(t) \right) I\left\{\frac{j}{n} \leq s\right\} \right|
\end{aligned}$$

$$\begin{aligned}
&= \sup_{s \in [0,1]} \sup_{|t| > \sqrt{n} \log(n)} \left| \frac{1}{n} \sum_{j=1}^n w_{nj} \left(-F_{\epsilon_j}(t \cdot \hat{d}_{n2}(X_{j-1}) + \hat{d}_{n1}(X_{j-1})) + F_{\epsilon_j}(t) \right) I \left\{ \frac{j}{n} \leq s \right\} \right| \\
&\leq \sup_{s \in [0,1]} \sup_{|t| > \sqrt{n} \log(n)} \frac{1}{n} \sum_{j=1}^n w_{nj} \left| -F_{\epsilon_j}(t \cdot \hat{d}_{n2}(X_{j-1}) + \hat{d}_{n1}(X_{j-1})) + F_{\epsilon_j}(t) \right| I \left\{ \frac{j}{n} \leq s \right\}.
\end{aligned}$$

Da

$$\begin{aligned}
|F_{\epsilon_j}(x) - F_{\epsilon_j}(y)| &= \max\{F_{\epsilon_j}(x), F_{\epsilon_j}(y)\} - \min\{F_{\epsilon_j}(x), F_{\epsilon_j}(y)\} \leq 1 - \min\{F_{\epsilon_j}(x), F_{\epsilon_j}(y)\} \\
&\text{und } \leq \max\{F_{\epsilon_j}(x), F_{\epsilon_j}(y)\} - 0 \quad \forall x, y \in \mathbb{R}, j = 1, \dots, n
\end{aligned}$$

und außerdem

$$\sup_{|t| > \sqrt{n} \log(n)} |\dots| \leq \sup_{t > \sqrt{n} \log(n)} |\dots| + \sup_{t < -\sqrt{n} \log(n)} |\dots|,$$

folgt weiter

$$\begin{aligned}
&\sup_{s \in [0,1]} \sup_{|t| > \sqrt{n} \log(n)} |H(s, t, \hat{d}_{n1}, \hat{d}_{n2}) - H(s, t, 0, 1)| \\
&\leq \sup_{s \in [0,1]} \sup_{|t| > \sqrt{n} \log(n)} \frac{1}{n} \sum_{j=1}^n w_{nj} \left| F_{\epsilon_j}(t) - F_{\epsilon_j}(t \cdot \hat{d}_{n2}(X_{j-1}) + \hat{d}_{n1}(X_{j-1})) \right| I \left\{ \frac{j}{n} \leq s \right\} \\
&\leq \sup_{s \in [0,1]} \sup_{t > \sqrt{n} \log(n)} \frac{1}{n} \sum_{j=1}^n w_{nj} (1 - \min\{F_{\epsilon_j}(t \cdot \hat{d}_{n2}(X_{j-1}) + \hat{d}_{n1}(X_{j-1})), F_{\epsilon_j}(t)\}) I \left\{ \frac{j}{n} \leq s \right\} \\
&\quad + \sup_{s \in [0,1]} \sup_{t < -\sqrt{n} \log(n)} \frac{1}{n} \sum_{j=1}^n w_{nj} \max\{F_{\epsilon_j}(t \cdot \hat{d}_{n2}(X_{j-1}) + \hat{d}_{n1}(X_{j-1})), F_{\epsilon_j}(t)\} I \left\{ \frac{j}{n} \leq s \right\}.
\end{aligned}$$

Da die Verteilungsfunktionen F_{ϵ_j} monoton wachsend sind und

$$\frac{\sqrt{n} \log(n)}{2} - 1 < \min\{t \cdot \hat{d}_{n2}(X_{j-1}) + \hat{d}_{n1}(X_{j-1}), t\} \quad \forall t > \sqrt{n} \log(n)$$

sowie

$$-\frac{\sqrt{n} \log(n)}{2} + 1 > \max\{t \cdot \hat{d}_{n2}(X_{j-1}) + \hat{d}_{n1}(X_{j-1}), t\} \quad \forall t < -\sqrt{n} \log(n),$$

ist weiter

$$\begin{aligned}
&\sup_{s \in [0,1]} \sup_{|t| > \sqrt{n} \log(n)} |H(s, t, \hat{d}_{n1}, \hat{d}_{n2}) - H(s, t, 0, 1)| \\
&\leq \sup_{s \in [0,1]} \sup_{t > \sqrt{n} \log(n)} \frac{1}{n} \sum_{j=1}^n w_{nj} (1 - \min\{F_{\epsilon_j}(t \cdot \hat{d}_{n2}(X_{j-1}) + \hat{d}_{n1}(X_{j-1})), F_{\epsilon_j}(t)\}) I \left\{ \frac{j}{n} \leq s \right\} \\
&\quad + \sup_{s \in [0,1]} \sup_{t < -\sqrt{n} \log(n)} \frac{1}{n} \sum_{j=1}^n w_{nj} \max\{F_{\epsilon_j}(t \cdot \hat{d}_{n2}(X_{j-1}) + \hat{d}_{n1}(X_{j-1})), F_{\epsilon_j}(t)\} I \left\{ \frac{j}{n} \leq s \right\} \\
&\leq \sup_{s \in [0,1]} \frac{1}{n} \sum_{j=1}^n w_{nj} I \left\{ \frac{j}{n} \leq s \right\} \left(1 - F_{\epsilon_j} \left(\frac{\sqrt{n} \log(n)}{2} - 1 \right) \right) \\
&\quad + \sup_{s \in [0,1]} \frac{1}{n} \sum_{j=1}^n w_{nj} I \left\{ \frac{j}{n} \leq s \right\} F_{\epsilon_j} \left(-\frac{\sqrt{n} \log(n)}{2} + 1 \right)
\end{aligned}$$

$$\leq \frac{1}{n} \sum_{j=1}^n \left(1 - F_{\epsilon_j} \left(\frac{\sqrt{n} \log(n)}{2} - 1 \right) \right) + \frac{1}{n} \sum_{j=1}^n F_{\epsilon_j} \left(-\frac{\sqrt{n} \log(n)}{2} + 1 \right). \quad (9.1.1)$$

Für die Verteilungsfunktionen F_{ϵ_j} der Innovationen gilt

$$\begin{aligned} t > 1 : \quad \frac{1}{n} \sum_{j=1}^n (1 - F_{\epsilon_j}(t)) &= \frac{1}{n} \sum_{j=1}^n P(\epsilon_j > t) \\ &= \frac{1}{t^2} \frac{1}{n} \sum_{j=1}^n E[t^2 I\{\epsilon_j > t\}] \\ &\leq \frac{1}{t^2} \frac{1}{n} \sum_{j=1}^n E[\epsilon_j^2 I\{\epsilon_j^2 > t^2\}] \\ &= O\left(\frac{1}{t^2}\right), \end{aligned}$$

da mit $\frac{1}{n} \sum_{j=1}^n E[\epsilon_j^2] = O(1)$ auch $\frac{1}{n} \sum_{j=1}^n E[\epsilon_j^2 I\{\epsilon_j^2 > t^2\}] = O(1)$ gilt. Ebenso ist

$$\begin{aligned} t < -1 : \quad \frac{1}{n} \sum_{j=1}^n F_{\epsilon_j}(t) &= \frac{1}{n} \sum_{j=1}^n P(\epsilon_j \leq t) \\ &= \frac{1}{t^2} \frac{1}{n} \sum_{j=1}^n E[t^2 I\{\epsilon_j \leq t\}] \\ &= \frac{1}{t^2} \frac{1}{n} \sum_{j=1}^n E[t^2 I\{\epsilon_j^2 \geq t^2\}] \\ &\leq \frac{1}{t^2} \frac{1}{n} \sum_{j=1}^n E[\epsilon_j^2 I\{\epsilon_j^2 \geq t^2\}] \\ &= O\left(\frac{1}{t^2}\right). \end{aligned}$$

Damit folgt aus (9.1.1)

$$\begin{aligned} &\sup_{s \in [0,1]} \sup_{|t| > \sqrt{n} \log(n)} |H(s, t, \hat{d}_{n1}, \hat{d}_{n2}) - H(s, t, 0, 1)| \\ &\leq \frac{1}{n} \sum_{j=1}^n \left(1 - F_{\epsilon_j} \left(\frac{\sqrt{n} \log(n)}{2} - 1 \right) \right) + \frac{1}{n} \sum_{j=1}^n F_{\epsilon_j} \left(-\frac{\sqrt{n} \log(n)}{2} + 1 \right) \\ &= O\left(\frac{1}{n \log(n)^2}\right) \\ &= o\left(\frac{1}{\sqrt{n}}\right). \end{aligned}$$

Es bleibt also zu zeigen, dass

$$\sup_{s \in [0,1]} \sup_{|t| \leq \sqrt{n} \log(n)} |H(s, t, \hat{d}_{n1}, \hat{d}_{n2}) - H(s, t, 0, 1)| = o_p\left(\frac{1}{\sqrt{n}}\right).$$

Für ein beschränktes Intervall $I \subset \mathbb{R}$ werden die Klassen von differenzierbaren Funktionen

$$C_i^{1+\delta}(I) := \left\{ d : I \rightarrow \mathbb{R} \left| \max\left\{\sup_{x \in I} |d(x)|, \sup_{x \in I} |d'(x)|\right\} + \sup_{x, y \in I, x \neq y} \frac{|d'(x) - d'(y)|}{|y - x|^\delta} \leq i \right. \right\}, \text{ mit } i = 1, 2$$

und

$$\tilde{C}_2^{1+\delta}(I) := \left\{ d : I \rightarrow \mathbb{R} \left| \max\left\{\sup_{x \in I} |d(x)|, \sup_{x \in I} |d'(x)|\right\} + \sup_{x, y \in I, x \neq y} \frac{|d'(x) - d'(y)|}{|y - x|^\delta} \leq 2, \inf_{x \in I} d(x) \geq \frac{1}{2} \right. \right\}$$

betrachtet. Um die Klassen $C_1^{1+\delta}(I_n)$ und $\tilde{C}_2^{1+\delta}(I_n)$ zu überdecken, braucht man nach Theorem 2.7.1 in van der Vaart and Wellner [1996]

$$M_{n1} := N(\bar{\epsilon}_n, C_1^{1+\delta}(I_n), \|\cdot\|_{I_n}) \leq \exp\left(K_1(2 + b_n - a_n)\bar{\epsilon}_n^{-\frac{1}{1+\delta}}\right)$$

bzw.

$$\begin{aligned} M_{n2} &:= N(\bar{\epsilon}_n, \tilde{C}_2^{1+\delta}(I_n), \|\cdot\|_{I_n}) \\ &\leq N(\bar{\epsilon}_n, C_2^{1+\delta}(I_n), \|\cdot\|_{I_n}) \\ &= N\left(\frac{\bar{\epsilon}_n}{2}, \left\{\frac{d}{2} \mid d \in C_2^{1+\delta}(I_n)\right\}, \|\cdot\|_{I_n}\right) \\ &\leq N\left(\frac{\bar{\epsilon}_n}{2}, C_1^{1+\delta}(I_n), \|\cdot\|_{I_n}\right) \\ &\leq \exp\left(K_2(2 + b_n - a_n)\bar{\epsilon}_n^{-\frac{1}{1+\delta}}\right) \end{aligned}$$

Kugeln vom Radius $\bar{\epsilon}_n > 0$ bzgl. der Pseudonorm $\|\cdot\|_{I_n}$, wobei $\|d\|_{I_n} := \sup_{x \in I_n} |d(x)|$. Die Konstanten K_1 und K_2 hängen nur von δ ab ($K_2 = K_1 \cdot 2^{\frac{1}{1+\delta}}$).

Für die Klassen

$$\begin{aligned} C_1^{1+\delta}(I_n)^n &:= \{d \in C_1^{1+\delta}(I_n), \|d\|_{I_n} \leq z_{n1} \log(n)\}, \\ \tilde{C}_2^{1+\delta}(I_n)^n &:= \{d \in \tilde{C}_2^{1+\delta}(I_n), \|d - 1\|_{I_n} \leq z_{n2} \log(n)\} \end{aligned}$$

gelten die selben Abschätzungen für die Überdeckungszahlen.

Mit $d_{11}, \dots, d_{1M_{n1}}$ bzw. $d_{21}, \dots, d_{2M_{n2}}$ seien die Funktionen bezeichnet, die die Mittelpunkte der Kugeln zur Überdeckung von $C_1^{1+\delta}(I_n)^n$ bzw. $\tilde{C}_2^{1+\delta}(I_n)^n$ bilden.

Damit gilt für alle $k = 1, \dots, M_{n1}$ und $l = 1, \dots, M_{n2}$, dass die Funktionen d_{1k} bzw. d_{2l} jeweils einen Abstand $\leq \bar{\epsilon}_n$ (bzgl. $\sup_{x \in I_n}$) zu einer Funktion aus $C_1^{1+\delta}(I_n)^n$ bzw. $\tilde{C}_2^{1+\delta}(I_n)^n$ haben, und daher z.B. $\sup_{x \in I_n} |d_{1k}(x)| \leq 1 + \bar{\epsilon}_n$ und $d_{2l}(x) \in [\frac{1}{2} - \bar{\epsilon}_n, 2 + \bar{\epsilon}_n]$ für alle $x \in I_n$.

Außerdem seien die Intervalle $[0, 1]$ und $[-\sqrt{n} \log(n), \sqrt{n} \log(n)]$ durch die Punkte $0 = s_1 < \dots < s_{M_{n3}} = 1$ bzw. $-\sqrt{n} \log(n) = t_1 < \dots < t_{M_{n4}} = \sqrt{n} \log(n)$ in Abschnitte der Größe $\bar{\epsilon}_n$ aufgeteilt, so dass $M_{n3} \leq \frac{1}{\bar{\epsilon}_n}$ und $M_{n4} \leq \frac{2\sqrt{n} \log(n)}{\bar{\epsilon}_n}$.

Sei $\bar{\epsilon}_n = \min\{\frac{1}{8}, n^{-\frac{1}{2}} \log(n)^{-1}\}$.

Aus $\|\hat{d}_{n1}\|_{I_n} = O_P(z_{n1})$, was nach Voraussetzung gilt, folgt $\|\hat{d}_{n1}\|_{I_n} = o_P(z_{n1} \log(n))$, und genauso $\|\hat{d}_{n2} - 1\|_{I_n} = o_P(z_{n2} \log(n))$ aus $\|\hat{d}_{n2} - 1\|_{I_n} = O_P(z_{n2})$.

Da also $P(\hat{d}_{n1} \notin C_1^{1+\delta}(I_n)^n) = o(1)$ und $P(\hat{d}_{n2} \notin \tilde{C}_2^{1+\delta}(I_n)^n) = o(1)$ gilt die Behauptung, falls

$$\sup_{s \in [0,1], |t| \leq \sqrt{n} \log(n), d_1 \in C_1^{1+\delta}(I_n)^n, d_2 \in \tilde{C}_2^{1+\delta}(I_n)^n} |H(s, t, d_1, d_2) - H(s, t, 0, 1)| = o_p\left(\frac{1}{\sqrt{n}}\right).$$

Mit Hilfe der Überdeckungspunkte kann man dies abschätzen:

$$\begin{aligned} & \sup_{\substack{s \in [0,1], |t| \leq \sqrt{n} \log(n), \\ d_1 \in C_1^{1+\delta}(I_n)^n, d_2 \in \tilde{C}_2^{1+\delta}(I_n)^n}} |H(s, t, d_1, d_2) - H(s, t, 0, 1)| \\ & \leq \max_{h,i,k,l} |H(s_h, t_i, d_{1k}, d_{2l}) - H(s_h, t_i, 0, 1)| \\ & \quad + \max_h \sup_{\substack{|s-s_h| \leq \bar{\epsilon}_n, |t| \leq \sqrt{n} \log(n), \\ d_1 \in C_1^{1+\delta}(I_n)^n, d_2 \in \tilde{C}_2^{1+\delta}(I_n)^n}} |H(s, t, d_1, d_2) - H(s_h, t, d_1, d_2)| \end{aligned} \quad (9.1.2)$$

$$+ \max_{h,i,k,l} \sup_{\substack{|t-t_i| \leq \bar{\epsilon}_n, \\ \|d_1 - d_{1k}\|_{I_n} \leq \bar{\epsilon}_n, \|d_2 - d_{2l}\|_{I_n} \leq \bar{\epsilon}_n}} |H(s_h, t, d_1, d_2) - H(s_h, t_i, d_{1k}, d_{2l})| \quad (9.1.3)$$

$$+ \max_h \sup_{|s-s_h| \leq \bar{\epsilon}_n, |t| \leq \sqrt{n} \log(n)} |H(s_h, t, 0, 1) - H(s, t, 0, 1)| \quad (9.1.4)$$

$$+ \max_{h,i} \sup_{|t-t_i| \leq \bar{\epsilon}_n} |H(s_h, t_i, 0, 1) - H(s_h, t, 0, 1)|, \quad (9.1.5)$$

mit $h = 1, \dots, M_{n3}$, $i = 1, \dots, M_{n4}$, $k = 1, \dots, M_{n1}$ und $l = 1, \dots, M_{n2}$.

Als nächstes werden die Terme (9.1.2) - (9.1.5) weiter abgeschätzt. Dazu zunächst eine Vorüberlegung: Für $|t - t_i| \leq \bar{\epsilon}_n$, $\|d_1 - d_{1k}\|_{I_n} \leq \bar{\epsilon}_n$, $d_2 \in \tilde{C}_2^{1+\delta}(I_n)^n$ ist

$$(t_i - \bar{\epsilon}_n)d_2(x) + d_{1k}(x) - \bar{\epsilon}_n \leq td_2(x) + d_1(x) \leq (t_i + \bar{\epsilon}_n)d_2(x) + d_{1k}(x) + \bar{\epsilon}_n \quad \forall x \in I_n,$$

da $d_2(x) > 0$ für alle $x \in I_n$. Für eine weitere Abschätzung mit $\|d_2 - d_{2l}\|_{I_n} \leq \bar{\epsilon}_n$ muss man zwischen den Fällen

$$(t_i - \bar{\epsilon}_n) \geq 0 \quad (t_i \geq \bar{\epsilon}_n)$$

$$\Rightarrow (t_i - \bar{\epsilon}_n)(d_{2l}(x) - \bar{\epsilon}_n) + d_{1k}(x) - \bar{\epsilon}_n \leq td_2(x) + d_1(x) \leq (t_i + \bar{\epsilon}_n)(d_{2l}(x) + \bar{\epsilon}_n) + d_{1k}(x) + \bar{\epsilon}_n \quad \forall x \in I_n,$$

$$(t_i - \bar{\epsilon}_n) < 0 < (t_i + \bar{\epsilon}_n) \quad (t_i \in (-\bar{\epsilon}_n, \bar{\epsilon}_n))$$

$$\Rightarrow (t_i - \bar{\epsilon}_n)(d_{2l}(x) + \bar{\epsilon}_n) + d_{1k}(x) - \bar{\epsilon}_n \leq td_2(x) + d_1(x) \leq (t_i + \bar{\epsilon}_n)(d_{2l}(x) - \bar{\epsilon}_n) + d_{1k}(x) + \bar{\epsilon}_n \quad \forall x \in I_n$$

$$\text{und } (t_i + \bar{\epsilon}_n) \leq 0 \quad (t_i \leq -\bar{\epsilon}_n)$$

$$\Rightarrow (t_i - \bar{\epsilon}_n)(d_{2l}(x) + \bar{\epsilon}_n) + d_{1k}(x) - \bar{\epsilon}_n \leq td_2(x) + d_1(x) \leq (t_i + \bar{\epsilon}_n)(d_{2l}(x) - \bar{\epsilon}_n) + d_{1k}(x) + \bar{\epsilon}_n \quad \forall x \in I_n$$

unterscheiden.

Damit kann (9.1.3) weiter abgeschätzt werden. Die Berechnung hierfür wird exemplarisch für den Fall $t_i \geq \bar{\epsilon}_n$ vorgeführt, da die beiden anderen Fälle analog funktionieren. Aus den obigen Ungleichungen folgt für alle $h = 1, \dots, M_{n3}$, $i = 1, \dots, M_{n4}$, $k = 1, \dots, M_{n1}$ und $l = 1, \dots, M_{n2}$ mit $d^j := d(X_{j-1})$

$$\sup_{\substack{|t-t_i| \leq \bar{\epsilon}_n, \\ \|d_1 - d_{1k}\|_{I_n} \leq \bar{\epsilon}_n, \|d_2 - d_{2l}\|_{I_n} \leq \bar{\epsilon}_n}} |H(s_h, t, d_1, d_2) - H(s_h, t_i, d_{1k}, d_{2l})|$$

$$\begin{aligned}
&\leq \sup_{\|d_1 - d_{1k}\|_{I_n} \leq \bar{\epsilon}_n, \|d_2 - d_{2l}\|_{I_n} \leq \bar{\epsilon}_n} \frac{1}{n} \sum_{j=1}^{\lfloor ns_h \rfloor} w_{nj} |I\{\epsilon_j \leq td_2^j + d_1^j\} - I\{\epsilon_j \leq td_{2l}^j + d_{1k}^j\}| \\
&\quad + \sup_{\|d_1 - d_{1k}\|_{I_n} \leq \bar{\epsilon}_n, \|d_2 - d_{2l}\|_{I_n} \leq \bar{\epsilon}_n} \frac{1}{n} \sum_{j=1}^{\lfloor ns_h \rfloor} w_{nj} |F_{\epsilon_j}(td_{2l}^j + d_{1k}^j) - F_{\epsilon_j}(td_2^j + d_1^j)| \\
&\leq \sup_{\|d_1 - d_{1k}\|_{I_n} \leq \bar{\epsilon}_n, \|d_2 - d_{2l}\|_{I_n} \leq \bar{\epsilon}_n} \frac{1}{n} \sum_{j=1}^n w_{nj} |I\{\epsilon_j \leq td_2^j + d_1^j\} - I\{\epsilon_j \leq td_{2l}^j + d_{1k}^j\}| \\
&\quad + \sup_{\|d_1 - d_{1k}\|_{I_n} \leq \bar{\epsilon}_n, \|d_2 - d_{2l}\|_{I_n} \leq \bar{\epsilon}_n} \frac{1}{n} \sum_{j=1}^n w_{nj} |F_{\epsilon_j}(td_{2l}^j + d_{1k}^j) - F_{\epsilon_j}(td_2^j + d_1^j)| \\
&\leq \frac{1}{n} \sum_{j=1}^n w_{nj} \left(I\{\epsilon_j \leq (t_i + \bar{\epsilon}_n)(d_{2l}^j + \bar{\epsilon}_n) + d_{1k}^j + \bar{\epsilon}_n\} - I\{\epsilon_j \leq (t_i - \bar{\epsilon}_n)(d_{2l}^j - \bar{\epsilon}_n) + d_{1k}^j - \bar{\epsilon}_n\} \right) \\
&\quad + \frac{1}{n} \sum_{j=1}^n w_{nj} \left(F_{\epsilon_j}((t_i + \bar{\epsilon}_n)(d_{2l}^j + \bar{\epsilon}_n) + d_{1k}^j + \bar{\epsilon}_n) - F_{\epsilon_j}((t_i - \bar{\epsilon}_n)(d_{2l}^j - \bar{\epsilon}_n) + d_{1k}^j - \bar{\epsilon}_n) \right) \\
&= \frac{1}{n} \sum_{j=1}^n w_{nj} \left(I\{\epsilon_j \leq (t_i + \bar{\epsilon}_n)(d_{2l}^j + \bar{\epsilon}_n) + d_{1k}^j + \bar{\epsilon}_n\} - F_{\epsilon_j}((t_i + \bar{\epsilon}_n)(d_{2l}^j + \bar{\epsilon}_n) + d_{1k}^j + \bar{\epsilon}_n) \right) \\
&\quad - \frac{1}{n} \sum_{j=1}^n w_{nj} \left(I\{\epsilon_j \leq (t_i - \bar{\epsilon}_n)(d_{2l}^j - \bar{\epsilon}_n) + d_{1k}^j - \bar{\epsilon}_n\} - F_{\epsilon_j}((t_i - \bar{\epsilon}_n)(d_{2l}^j - \bar{\epsilon}_n) + d_{1k}^j - \bar{\epsilon}_n) \right) \\
&\quad + \frac{2}{n} \sum_{j=1}^n w_{nj} \left(F_{\epsilon_j}((t_i + \bar{\epsilon}_n)(d_{2l}^j + \bar{\epsilon}_n) + d_{1k}^j + \bar{\epsilon}_n) - F_{\epsilon_j}((t_i - \bar{\epsilon}_n)(d_{2l}^j - \bar{\epsilon}_n) + d_{1k}^j - \bar{\epsilon}_n) \right) \\
&= H(1, t_i + \bar{\epsilon}_n, d_{1k} + \bar{\epsilon}_n, d_{2l} + \bar{\epsilon}_n) - H(1, t_i - \bar{\epsilon}_n, d_{1k} - \bar{\epsilon}_n, d_{2l} - \bar{\epsilon}_n) \\
&\quad + \frac{2}{n} \sum_{j=1}^n w_{nj} \left(F_{\epsilon_j}((t_i + \bar{\epsilon}_n)(d_{2l}^j + \bar{\epsilon}_n) + d_{1k}^j + \bar{\epsilon}_n) - F_{\epsilon_j}((t_i - \bar{\epsilon}_n)(d_{2l}^j - \bar{\epsilon}_n) + d_{1k}^j - \bar{\epsilon}_n) \right).
\end{aligned}$$

Weiter gilt nach dem Mittelwertsatz für ein $\xi(t_i, X_{j-1}, k, l)$ zwischen $(t_i + \bar{\epsilon}_n)(d_{2l}(X_{j-1}) + \bar{\epsilon}_n) + d_{1k}(X_{j-1}) + \bar{\epsilon}_n$ und $(t_i - \bar{\epsilon}_n)(d_{2l}(X_{j-1}) - \bar{\epsilon}_n) + d_{1k}(X_{j-1}) - \bar{\epsilon}_n$

$$\begin{aligned}
&\frac{1}{n} \sum_{j=1}^n w_{nj} \left(F_{\epsilon_j}((t_i + \bar{\epsilon}_n)(d_{2l}^j + \bar{\epsilon}_n) + d_{1k}^j + \bar{\epsilon}_n) - F_{\epsilon_j}((t_i - \bar{\epsilon}_n)(d_{2l}^j - \bar{\epsilon}_n) + d_{1k}^j - \bar{\epsilon}_n) \right) \\
&= \frac{1}{n} \sum_{j=1}^n w_{nj} f_{\epsilon_j}(\xi(t_i, X_{j-1}, k, l)) \left((t_i + \bar{\epsilon}_n)(d_{2l}^j + \bar{\epsilon}_n) + d_{1k}^j + \bar{\epsilon}_n - (t_i - \bar{\epsilon}_n)(d_{2l}^j - \bar{\epsilon}_n) - d_{1k}^j - \bar{\epsilon}_n \right) \\
&= \frac{1}{n} \sum_{j=1}^n w_{nj} f_{\epsilon_j}(\xi(t_i, X_{j-1}, k, l)) \left(2t_i \bar{\epsilon}_n + 2d_{2l}^j \bar{\epsilon}_n + 2\bar{\epsilon}_n \right) \\
&\leq \left(\sup_{x \in I_n} |d_{2l}(x)| \bar{\epsilon}_n + 2\bar{\epsilon}_n \right) \frac{1}{n} \sum_{j=1}^n \sup_{t \in \mathbb{R}} |f_{\epsilon_j}(t)| + 2\bar{\epsilon}_n \frac{1}{n} \sum_{j=1}^n w_{nj} f_{\epsilon_j}(\xi(t_i, X_{j-1}, k, l)) t_i \\
&= O(\bar{\epsilon}_n) + 2\bar{\epsilon}_n \frac{1}{n} \sum_{j=1}^n w_{nj} f_{\epsilon_j}(\xi(t_i, X_{j-1}, k, l)) t_i, \tag{9.1.6}
\end{aligned}$$

da $\sup_{x \in I_n} |d_{2l}(x)| \leq 2 + \bar{\epsilon}_n \leq \frac{17}{8} < \infty$ und $\frac{1}{n} \sum_{j=1}^n \sup_{t \in \mathbb{R}} |f_{\epsilon_j}(t)| = O(1)$ nach Voraussetzung.

Mit der Voraussetzung $\frac{1}{n} \sum_{j=1}^n \sup_{t \in \mathbb{R}} |f_{\epsilon_j}(t)t| = O(1)$ folgt außerdem

$\frac{1}{n} \sum_{j=1}^n w_{nj} f_{\epsilon_j}(\xi(t_i, X_{j-1}, k, l)) t_i = O(1)$, was im Folgenden gezeigt wird.

Für $|t_i| > \frac{\sup_{x \in I_n} |d_{1k}(x)| + \bar{\epsilon}_n}{\inf_{x \in I_n} |d_{2l}(x)| - \bar{\epsilon}_n} + \bar{\epsilon}_n$ gilt entweder $(t_i + \bar{\epsilon}_n)(d_{2l}(x) + \bar{\epsilon}_n) + d_{1k}(x) + \bar{\epsilon}_n > 0$ und $(t_i - \bar{\epsilon}_n)(d_{2l}(x) - \bar{\epsilon}_n) + d_{1k}(x) - \bar{\epsilon}_n > 0$ oder $(t_i + \bar{\epsilon}_n)(d_{2l}(x) + \bar{\epsilon}_n) + d_{1k}(x) + \bar{\epsilon}_n < 0$ und $(t_i - \bar{\epsilon}_n)(d_{2l}(x) - \bar{\epsilon}_n) + d_{1k}(x) - \bar{\epsilon}_n < 0$ für alle $x \in I_n$. Damit ist $\xi(t_i, x, k, l) \neq 0$ für alle $x \in I_n$ und daher weiter

$$\begin{aligned} & \frac{1}{n} \sum_{j=1}^n w_{nj} f_{\epsilon_j}(\xi(t_i, X_{j-1}, k, l)) t_i \\ &= \frac{1}{n} \sum_{j=1}^n w_{nj} f_{\epsilon_j}(\xi(t_i, j, k, l)) \xi(t_i, X_{j-1}, k, l) \frac{t_i}{\xi(t_i, X_{j-1}, k, l)} \\ &\leq \sup_{x \in I_n} \left| \frac{t_i}{\xi(t_i, x, k, l)} \right| \frac{1}{n} \sum_{j=1}^n \sup_{t \in \mathbb{R}} |f_{\epsilon_j}(t)t| \\ &= O(1), \end{aligned}$$

da $\xi(t_i, x, k, l)$ zwischen $(t_i + \bar{\epsilon}_n)(d_{2l}(x) + \bar{\epsilon}_n) + d_{1k}(x) + \bar{\epsilon}_n \xrightarrow{n \rightarrow \infty} t_i d_{2l}(x) + d_{1k}(x)$ und $(t_i - \bar{\epsilon}_n)(d_{2l}(x) - \bar{\epsilon}_n) + d_{1k}(x) - \bar{\epsilon}_n \xrightarrow{n \rightarrow \infty} t_i d_{2l}(x) + d_{1k}(x)$ liegt, und $|d_{1k}(x)| \leq 1 + \bar{\epsilon}_n \leq \frac{9}{8}$ sowie $d_{2l}(x) \in [\frac{1}{2} - \bar{\epsilon}_n, 2 + \bar{\epsilon}_n] \subset [\frac{3}{8}, \frac{17}{8}]$ für alle $x \in I_n$.

Für $|t_i| \leq \frac{\sup_{x \in I_n} |d_{1k}(x)| + \bar{\epsilon}_n}{\inf_{x \in I_n} |d_{2l}(x)| - \bar{\epsilon}_n} + \bar{\epsilon}_n \leq \frac{1 + \frac{1}{8} + \frac{1}{8}}{\frac{1}{2} - \frac{1}{8} - \frac{1}{8}} + \frac{1}{8} = \frac{41}{8} < \infty$ folgt $\frac{1}{n} \sum_{j=1}^n w_{nj} f_{\epsilon_j}(\xi(t_i, X_{j-1}, k, l)) t_i = O(1)$ direkt aus der Voraussetzung $\frac{1}{n} \sum_{j=1}^n \sup_{t \in \mathbb{R}} |f_{\epsilon_j}(t)| = O(1)$.

Aus (9.1.6) folgt also

$$\begin{aligned} \frac{1}{n} \sum_{j=1}^n w_{nj} \left(F_{\epsilon_j}((t_i + \bar{\epsilon}_n)(d_{2l}^j + \bar{\epsilon}_n) + d_{1k}^j + \bar{\epsilon}_n) - F_{\epsilon_j}((t_i - \bar{\epsilon}_n)(d_{2l}^j - \bar{\epsilon}_n) + d_{1k}^j - \bar{\epsilon}_n) \right) &= O(\bar{\epsilon}_n) \\ &= o\left(\frac{1}{\sqrt{n}}\right) \end{aligned}$$

und damit für (9.1.3)

$$\begin{aligned} & \max_{h, i, k, l} \sup_{\substack{|t-t_i| \leq \bar{\epsilon}_n, \\ \|d_1 - d_{1k}\|_{I_n} \leq \bar{\epsilon}_n, \|d_2 - d_{2l}\|_{I_n} \leq \bar{\epsilon}_n}} |H(s_h, t, d_1, d_2) - H(s_h, t_i, d_{1k}, d_{2l})| \\ &\leq \max_{i, k, l} |H(1, t_i + \bar{\epsilon}_n, d_{1k} + \bar{\epsilon}_n, d_{2l} + \bar{\epsilon}_n) - H(1, t_i - \bar{\epsilon}_n, d_{1k} - \bar{\epsilon}_n, d_{2l} - \bar{\epsilon}_n)| + O(\bar{\epsilon}_n) \\ &= \max_{i, k, l} |H(1, t_i + \bar{\epsilon}_n, d_{1k} + \bar{\epsilon}_n, d_{2l} + \bar{\epsilon}_n) - H(1, t_i - \bar{\epsilon}_n, d_{1k} - \bar{\epsilon}_n, d_{2l} - \bar{\epsilon}_n)| + o\left(\frac{1}{\sqrt{n}}\right). \quad (9.1.7) \end{aligned}$$

Auf die gleiche Weise folgt für (9.1.5)

$$\begin{aligned} & \max_{h, i} \sup_{|t-t_i| \leq \bar{\epsilon}_n} |H(s_h, t_i, 0, 1) - H(s_h, t, 0, 1)| \\ &\leq \max_i |H(1, t_i + \bar{\epsilon}_n, 0, 1) - H(1, t_i - \bar{\epsilon}_n, 0, 1)| + o\left(\frac{1}{\sqrt{n}}\right). \quad (9.1.8) \end{aligned}$$

Für (9.1.2) gilt

$$\max_h \sup_{\substack{|s-s_h| \leq \bar{\epsilon}_n, |t| \leq \sqrt{n} \log(n), \\ d_1 \in C_1^{1+\delta}(I_n)^n, d_2 \in C_2^{1+\delta}(I_n)^n}} |H(s, t, d_1, d_2) - H(s_h, t, d_1, d_2)|$$

$$\begin{aligned}
&\leq \max_h \sup_{\substack{|s-s_h| \leq \bar{\epsilon}_n, |t| \leq \sqrt{n} \log(n), \\ d_1 \in C_1^{1+\delta}(I_n)^n, d_2 \in \tilde{C}_2^{1+\delta}(I_n)^n}} \frac{1}{n} \sum_{j=1}^n w_{nj} \left| I\{\epsilon_j \leq t \cdot d_2^j + d_1^j\} - F(td_2^j + d_1^j) \right| \left| I\left\{\frac{j}{n} \leq s\right\} - I\left\{\frac{j}{n} \leq s_h\right\} \right| \\
&\leq \max_h \sup_{|s-s_h| \leq \bar{\epsilon}_n} \frac{1}{n} \sum_{j=1}^n \left| I\left\{\frac{j}{n} \leq s\right\} - I\left\{\frac{j}{n} \leq s_h\right\} \right| \\
&\leq \max_h \sup_{|s-s_h| \leq \bar{\epsilon}_n} \left(|s - s_h| + \frac{1}{n} \right) \quad \text{mit (4.5.6) (S. 65) im Beweis zu Satz 4.4} \\
&\leq \bar{\epsilon}_n + \frac{1}{n} \\
&= o\left(\frac{1}{\sqrt{n}}\right)
\end{aligned} \tag{9.1.9}$$

und genauso für (9.1.4)

$$\max_h \sup_{|s-s_h| \leq \bar{\epsilon}_n, |t| \leq \sqrt{n} \log(n)} |H(s_h, t, 0, 1) - H(s, t, 0, 1)| = o\left(\frac{1}{\sqrt{n}}\right).$$

Mit (9.1.7) und (9.1.8) gilt also insgesamt

$$\begin{aligned}
&\sup_{\substack{s \in [0,1], |t| \leq \sqrt{n} \log(n), \\ d_1 \in C_1^{1+\delta}(I_n)^n, d_2 \in \tilde{C}_2^{1+\delta}(I_n)^n}} |H(s, t, d_1, d_2) - H(s, t, 0, 1)| \\
&\leq \max_{h,i,k,l} |H(s_h, t_i, d_{1k}, d_{2l}) - H(s_h, t_i, 0, 1)| \\
&\quad + \max_{i,k,l} |H(1, t_i + \bar{\epsilon}_n, d_{1k} + \bar{\epsilon}_n, d_{2l} + \bar{\epsilon}_n) - H(1, t_i - \bar{\epsilon}_n, d_{1k} - \bar{\epsilon}_n, d_{2l} - \bar{\epsilon}_n)| \\
&\quad + \max_i |H(1, t_i + \bar{\epsilon}_n, 0, 1) - H(1, t_i - \bar{\epsilon}_n, 0, 1)| \\
&\quad + o\left(\frac{1}{\sqrt{n}}\right),
\end{aligned}$$

und es bleibt zu zeigen, dass

$$\max_{h,i,k,l} |H(s_h, t_i, d_{1k}, d_{2l}) - H(s_h, t_i, 0, 1)| = o_p\left(\frac{1}{\sqrt{n}}\right), \tag{9.1.10}$$

sowie

$$\max_{i,k,l} |H(1, t_i + \bar{\epsilon}_n, d_{1k} + \bar{\epsilon}_n, d_{2l} + \bar{\epsilon}_n) - H(1, t_i - \bar{\epsilon}_n, d_{1k} - \bar{\epsilon}_n, d_{2l} - \bar{\epsilon}_n)| = o_p\left(\frac{1}{\sqrt{n}}\right) \tag{9.1.11}$$

und

$$\max_i |H(1, t_i + \bar{\epsilon}_n, 0, 1) - H(1, t_i - \bar{\epsilon}_n, 0, 1)| = o_p\left(\frac{1}{\sqrt{n}}\right) \tag{9.1.12}$$

für $h = 1, \dots, M_{n3}$, $i = 1, \dots, M_{n4}$, $k = 1, \dots, M_{n1}$ und $l = 1, \dots, M_{n2}$.

Dies kann man mit Hilfe des folgenden Hilfssatzes, der aus einer Ungleichung für Martingalzuwächse resultiert.

Hilfssatz 9.1.A Seien $s \in [0, 1]$, t_1, t_2 reelle Zahlen und $d_{11}, d_{21}, d_{12}, d_{22} : \mathbb{R} \rightarrow \mathbb{R}$ Funktionen. Dann gilt für alle $\eta > 0$

$$P(\sqrt{n} \cdot |H(s, t_1, d_{11}, d_{21}) - H(s, t_2, d_{12}, d_{22})| > \eta) \leq 2 \exp\left(-\frac{\eta^2 n}{4\eta\sqrt{n} + 2nA_{t_1, d_{11}, d_{21}, t_2, d_{12}, d_{22}}(n)}\right),$$

wobei

$$A_{t_1, d_{11}, d_{21}, t_2, d_{12}, d_{22}}(n) := \frac{1}{n} \sum_{j=1}^n \sup_{x \in I_n} |F_{\epsilon_j}(t_1 d_{21}(x) + d_{11}(x)) - F_{\epsilon_j}(t_2 d_{22}(x) + d_{12}(x))|$$

Beweis Sei

$$Y_j := w_{nj} I\left\{\frac{j}{n} \leq s\right\} (I\{\epsilon_j \leq t_1 d_{21}(X_{j-1}) + d_{11}(X_{j-1})\} - I\{\epsilon_j \leq t_2 d_{22}(X_{j-1}) + d_{12}(X_{j-1})\} - F_{\epsilon_j}(t_1 d_{21}(X_{j-1}) + d_{11}(X_{j-1})) + F_{\epsilon_j}(t_2 d_{22}(X_{j-1}) + d_{12}(X_{j-1}))),$$

dann ist $|Y_j| \leq 2$ und es gilt

$$E[Y_j | X_0, \dots, X_{j-1}] = 0 = E[-Y_j | X_0, \dots, X_{j-1}],$$

also ist Y_1, \dots, Y_n ($-Y_1, \dots, -Y_n$) eine beschränkte Folge von Martingalzuwächsen. Für eine solche gilt nach Freedman [1975] (Theorem 1.6)

$$P\left(\sum_{j=1}^n Y_j \geq \eta\sqrt{n}, \sum_{j=1}^n E[Y_j^2 | X_0, \dots, X_{j-1}] \leq nA\right) \leq \exp\left(-\frac{\eta^2 n}{4\eta\sqrt{n} + 2nA}\right) \quad \forall \eta, A > 0.$$

Für die zweite Summe ergibt sich durch Ausmultiplizieren

$$\begin{aligned} & \sum_{j=1}^n E[Y_j^2 | X_0, \dots, X_{j-1}] \\ &= \sum_{j=1}^n E[Y_j^2 | X_{j-1}] \\ &\leq \sum_{j=1}^n E\left[\left(I\{\epsilon_j \leq t_1 d_{21}(X_{j-1}) + d_{11}(X_{j-1})\} - I\{\epsilon_j \leq t_2 d_{22}(X_{j-1}) + d_{12}(X_{j-1})\} - F_{\epsilon_j}(t_1 d_{21}(X_{j-1}) + d_{11}(X_{j-1})) + F_{\epsilon_j}(t_2 d_{22}(X_{j-1}) + d_{12}(X_{j-1}))\right)^2 w_{nj}^2 \middle| X_{j-1}\right] \\ &= \sum_{j=1}^n E\left[\left(I\{\epsilon_j \leq t_1 d_{21}(X_{j-1}) + d_{11}(X_{j-1})\} - I\{\epsilon_j \leq t_2 d_{22}(X_{j-1}) + d_{12}(X_{j-1})\}\right)^2 w_{nj}^2 \middle| X_{j-1}\right] \\ &\quad + \sum_{j=1}^n E\left[\left(F_{\epsilon_j}(t_1 d_{21}(X_{j-1}) + d_{11}(X_{j-1})) - F_{\epsilon_j}(t_2 d_{22}(X_{j-1}) + d_{12}(X_{j-1}))\right)^2 w_{nj}^2 \middle| X_{j-1}\right] \\ &\quad - 2 \sum_{j=1}^n E\left[\left(I\{\epsilon_i \leq t e(X_{i-1}) + d(X_{i-1})\} - I\{\epsilon_j \leq u h(X_{i-1}) + g(X_{i-1})\}\right) \cdot \left(F_{\epsilon_i}(t e(X_{i-1}) + d(X_{i-1})) - F_{\epsilon_i}(u h(X_{i-1}) + g(X_{i-1}))\right) w_{nj}^2 \middle| X_{j-1}\right] \end{aligned}$$

$$\begin{aligned}
&= \sum_{j=1}^n nE \left[\left(I\{\epsilon_j \leq t_1 d_{21}(X_{j-1}) + d_{11}(X_{j-1})\} - I\{\epsilon_j \leq t_2 d_{22}(X_{j-1}) + d_{12}(X_{j-1})\} \right)^2 w_{nj}^2 \middle| X_{j-1} \right] \\
&\quad - \sum_{j=1}^n E \left[\left(F_{\epsilon_j}(t_1 d_{21}(X_{j-1}) + d_{11}(X_{j-1})) - F_{\epsilon_j}(t_2 d_{22}(X_{j-1}) + d_{12}(X_{j-1})) \right)^2 w_{nj}^2 \middle| X_{j-1} \right] \\
&\leq \sum_{j=1}^n E \left[\left(I\{\epsilon_j \leq t_1 d_{21}(X_{j-1}) + d_{11}(X_{j-1})\} - I\{\epsilon_j \leq t_2 d_{22}(X_{j-1}) + d_{12}(X_{j-1})\} \right)^2 w_{nj}^2 \middle| X_{j-1} \right] \\
&= \sum_{j=1}^n E \left[\left| I\{\epsilon_j \leq t_1 d_{21}(X_{j-1}) + d_{11}(X_{j-1})\} - I\{\epsilon_j \leq t_2 d_{22}(X_{j-1}) + d_{12}(X_{j-1})\} \right| w_{nj}^2 \middle| X_{j-1} \right] \\
&= \sum_{j=1}^n |F_{\epsilon_j}(t_1 d_{21}(X_{j-1}) + d_{11}(X_{j-1})) - F_{\epsilon_j}(t_2 d_{22}(X_{j-1}) + d_{12}(X_{j-1}))| w_{nj}^2 \\
&\leq \sum_{j=1}^n \sup_{x \in I_n} |F_{\epsilon_j}(t_1 d_{21}(x) + d_{11}(x)) - F_{\epsilon_j}(t_2 d_{22}(x) + d_{12}(x))| \\
&= nA_{t_1, d_{11}, d_{21}, t_2, d_{12}, d_{22}}(n).
\end{aligned}$$

Daher folgt nach Freedman [1975]

$$\begin{aligned}
&P(\sqrt{n} \cdot |H(s, t_1, d_{11}, d_{21}) - H(s, t_2, d_{12}, d_{22})| > \eta) \\
&= P\left(\sqrt{n} \left| \frac{1}{n} \sum_{j=1}^n Y_j \right| > \eta, \sum_{j=1}^n E[Y_j^2 | X_0, \dots, X_{j-1}] \leq nA_{t_1, d_{11}, d_{21}, t_2, d_{12}, d_{22}}(n)\right) \\
&\leq P\left(\sqrt{n} \left| \frac{1}{n} \sum_{j=1}^n Y_j \right| \geq \eta, \sum_{j=1}^n E[Y_j^2 | X_0, \dots, X_{j-1}] \leq nA_{t_1, d_{11}, d_{21}, t_2, d_{12}, d_{22}}(n)\right) \\
&= P\left(\sum_{j=1}^n Y_j \geq \eta\sqrt{n}, \sum_{j=1}^n E[Y_j^2 | X_0, \dots, X_{j-1}] \leq nA_{t_1, d_{11}, d_{21}, t_2, d_{12}, d_{22}}(n)\right) \\
&\quad + P\left(\sum_{j=1}^n -Y_j \geq \eta\sqrt{n}, \sum_{j=1}^n E[Y_j^2 | X_0, \dots, X_{j-1}] \leq nA_{t_1, d_{11}, d_{21}, t_2, d_{12}, d_{22}}(n)\right) \\
&\leq 2 \exp\left(-\frac{\eta^2 n}{4\eta\sqrt{n} + 2nA_{t_1, d_{11}, d_{21}, t_2, d_{12}, d_{22}}(n)}\right).
\end{aligned}$$

□

Fortsetzung Beweis Lemma 9.1

Mit diesem Hilfssatz gilt für (9.1.10) für alle $\eta > 0$

$$\begin{aligned}
&P\left(\sqrt{n} \max_{h,i,k,l} |H(s_h, t_i, d_{1k}, d_{2l}) - H(s_h, t_i, 0, 1)| > \eta\right) \\
&\leq \sum_{h=1}^{M_{n3}} \sum_{i=1}^{M_{n4}} \sum_{k=1}^{M_{n1}} \sum_{l=1}^{M_{n2}} P(\sqrt{n} |H(s_h, t_i, d_{1k}, d_{2l}) - H(s_h, t_i, 0, 1)| > \eta)
\end{aligned}$$

$$\leq M_{n3} \sum_{i=1}^{M_{n4}} \sum_{k=1}^{M_{n1}} \sum_{l=1}^{M_{n2}} \left(2 \exp \left(- \frac{\eta^2 n}{4\eta\sqrt{n} + 2nA_{t_i, d_{1k}, d_{2l}, t_i, 0, 1}(n)} \right) \right). \quad (9.1.13)$$

Für $A_{t_i, d_{1k}, d_{2l}, t_i, 0, 1}(n)$ gilt mit dem Mittelwertsatz für ein $\xi(t_i, x, k, l)$ zwischen $t_i d_{2l}(x) + d_{1k}(x)$ und t_i

$$\begin{aligned} A_{t_i, d_{1k}, d_{2l}, t_i, 0, 1}(n) &= \frac{1}{n} \sum_{j=1}^n \sup_{x \in I_n} |F_{\epsilon_j}(t_i d_{2l}(x) + d_{1k}(x)) - F_{\epsilon_j}(t_i(x))| \\ &= \frac{1}{n} \sum_{j=1}^n \sup_{x \in I_n} |f_{\epsilon_j}(\xi(t_i, x, k, l))(t_i(d_{2l}(x) - 1) + d_{1k}(x))| \end{aligned}$$

Mit den Voraussetzungen $\frac{1}{n} \sum_{j=1}^n \sup_{t \in \mathbb{R}} |f_{\epsilon_j}(t)| = O(1)$ und $\frac{1}{n} \sum_{j=1}^n \sup_{t \in \mathbb{R}} |f_{\epsilon_j}(t)t| = O(1)$ wird dies im Folgenden weiter abgeschätzt.

Für $|t_i| > \frac{\sup_{x \in I_n} |d_{1k}(x)|}{\inf_{x \in I_n} |d_{2l}(x)|}$ gilt entweder $t_i d_{2l}(x) + d_{1k}(x) > 0$ und $t_i d_{2l}(x) + d_{1k}(x) > 0$ oder $t_i d_{2l}(x) + d_{1k}(x) < 0$ und $t_i d_{2l}(x) + d_{1k}(x) < 0$ für alle $x \in I_n$. Damit ist $\xi(t_i, x, k, l) \neq 0$ für alle $x \in I_n$ und daher weiter

$$\begin{aligned} &\frac{1}{n} \sum_{j=1}^n \sup_{x \in I_n} |f_{\epsilon_j}(\xi(t_i, x, k, l))(t_i(d_{2l}(x) - 1) + d_{1k}(x))| \\ &\leq \sup_{x \in I_n} |d_{1k}(x)| \frac{1}{n} \sum_{j=1}^n \sup_{t \in \mathbb{R}} |f_{\epsilon_j}(t)| \\ &\quad + \frac{1}{n} \sum_{j=1}^n \sup_{x \in I_n} \left| f_{\epsilon_j}(\xi(t_i, x, k, l)) \xi(t_i, x, k, l) \frac{t_i}{\xi(t_i, x, k, l)} (d_{2l}(x) - 1) \right| \\ &\leq \sup_{x \in I_n} |d_{1k}(x)| \frac{1}{n} \sum_{j=1}^n \sup_{t \in \mathbb{R}} |f_{\epsilon_j}(t)| \\ &\quad + \sup_{x \in I_n} |d_{2l}(x) - 1| \sup_{x \in I_n} \left| \frac{t_i}{\xi(t_i, x, k, l)} \right| \frac{1}{n} \sum_{j=1}^n \sup_{t \in \mathbb{R}} |f_{\epsilon_j}(t)t|. \end{aligned}$$

Da $\xi(t_i, x, k, l)$ zwischen $t_i d_{2l}(x) + d_{1k}(x)$ und t_i liegt und $|d_{1k}(x)| \leq 1 + \bar{\epsilon}_n \leq \frac{9}{8}$ sowie $d_{2l}(x) \in [\frac{1}{2} - \bar{\epsilon}_n, 2 + \bar{\epsilon}_n] \subset [\frac{3}{8}, \frac{17}{8}]$ für alle $x \in I_n$, kann $\sup_{x \in I_n} \left| \frac{t_i}{\xi(t_i, x, k, l)} \right| \frac{1}{n} \sum_{j=1}^n \sup_{t \in \mathbb{R}} |f_{\epsilon_j}(t)t|$ für hinreichend großes n durch eine von i, k, l unabhängige Konstante noch oben abgeschätzt werden. Das selbe gilt für $\frac{1}{n} \sum_{j=1}^n \sup_{t \in \mathbb{R}} |f_{\epsilon_j}(t)|$ nach Voraussetzung

Für $|t_i| \leq \frac{\sup_{x \in I_n} |d_{1k}(x)|}{\inf_{x \in I_n} |d_{2l}(x)|} \leq \frac{1 + \frac{1}{8}}{\frac{1}{2} - \frac{1}{8}} = 3 < \infty$ gilt $\frac{1}{n} \sum_{j=1}^n \sup_{x \in I_n} |f_{\epsilon_j}(\xi(t_i, x, k, l))(t_i(d_{2l}(x) - 1) + d_{1k}(x))| \leq (\sup_{x \in I_n} |d_{1k}(x)| + 3 \sup_{x \in I_n} |d_{2l}(x) - 1|) \frac{1}{n} \sum_{j=1}^n \sup_{t \in \mathbb{R}} |f_{\epsilon_j}(t)|$, und der zweite Faktor kann ebenfalls für hinreichend großes n durch eine Konstante nach oben abgeschätzt werden.

Da d_{1k} und d_{2l} einen Abstand $\leq \bar{\epsilon}_n$ zu einer Funktion aus $C_1^{1+\delta}(I_n)^n$ bzw. $\tilde{C}_2^{1+\delta}(I_n)^n$ haben (bezüglich $\sup_{x \in I_n}$), ist $\sup_{x \in I_n} |d_{1k}(x)| \leq z_{n1} \log(n) + \bar{\epsilon}_n$ und $\sup_{x \in I_n} |d_{2l}(x) - 1| \leq z_{n2} \log(n) + \bar{\epsilon}_n$ für alle k, l , also gibt es für hinreichend große n ein von i, k, l unabhängiges $K_3 < \infty$, so dass

$$A_{t_i, d_{1k}, d_{2l}, t_i, 0, 1}(n) \leq K_3(z_{n1} \log(n) + z_{n2} \log(n) + \bar{\epsilon}_n) \quad \forall i, k, l.$$

Hieraus folgt mit (9.1.13) für hinreichend große n

$$P \left(\sqrt{n} \max_{h, i, k, l} |H(s_h, t_i, d_{1k}, d_{2l}) - H(s_h, t_i, 0, 1)| > \eta \right)$$

$$\begin{aligned}
&= M_{n3} \sum_{i=1}^{M_{n4}} \sum_{k=1}^{M_{n1}} \sum_{l=1}^{M_{n2}} \left(2 \exp \left(- \frac{\eta^2 n}{4\eta\sqrt{n} + 2nK_3(z_{n1} \log(n) + z_{n2} \log(n) + \bar{\epsilon}_n)} \right) \right) \\
&= M_{n3} M_{n4} M_{n1} M_{n2} \left(2 \exp \left(- \frac{\eta^2 n}{4\eta\sqrt{n} + 2nK_3(z_{n1} \log(n) + z_{n2} \log(n) + \bar{\epsilon}_n)} \right) \right) \\
&= 2 \exp \left(\log \left(\frac{\sqrt{n} \log(n)}{\bar{\epsilon}_n^2} \right) + (K_1 + K_2)(2 + b_n - a_n) \bar{\epsilon}_n^{-\frac{1}{1+\delta}} \right. \\
&\quad \left. - \frac{n\eta^2}{4\eta\sqrt{n} + 2nK_3(z_{n1} \log(n) + z_{n2} \log(n) + \bar{\epsilon}_n)} \right).
\end{aligned}$$

Dieser Ausdruck ist $o(1)$, wenn der Exponent gegen minus unendlich geht, also wenn

$$\frac{4\eta\sqrt{n} + 2nK_3(z_{n1} \log(n) + z_{n2} \log(n) + \bar{\epsilon}_n)}{n\eta^2} = o \left(\frac{1}{\log \left(\frac{\sqrt{n} \log(n)}{\bar{\epsilon}_n^2} \right) + (K_1 + K_2)(2 + b_n - a_n) \bar{\epsilon}_n^{-\frac{1}{1+\delta}}} \right). \quad (9.1.14)$$

Es gilt

$$\begin{aligned}
&\left(\log \left(\frac{\sqrt{n} \log(n)}{\bar{\epsilon}_n^2} \right) + (K_1 + K_2)(2 + b_n - a_n) \bar{\epsilon}_n^{-\frac{1}{1+\delta}} \right) \frac{4\eta\sqrt{n} + 2nK_3(z_{n1} \log(n) + z_{n2} \log(n) + \bar{\epsilon}_n)}{n\eta^2} \\
&= O \left(\left(\log(n) + (b_n - a_n) \bar{\epsilon}_n^{-\frac{1}{1+\delta}} \right) \left(n^{-\frac{1}{2}} + \bar{\epsilon}_n + (z_{n1} \log(n) + z_{n2} \log(n)) \right) \right) \\
&= O \left((b_n - a_n) n^{\frac{1}{2} \frac{1}{1+\delta}} \log(n)^{\frac{1}{1+\delta}} \left(n^{-\frac{1}{2}} + n^{-\frac{1}{2}} \log(n)^{-1} + (z_{n1} \log(n) + z_{n2} \log(n)) \right) \right) \\
&= O \left(n^{\frac{1}{2}(\frac{1}{1+\delta}-1)} \log(n)^{\frac{1}{1+\delta}} (b_n - a_n) \right) + O \left(n^{\frac{1}{2} \frac{1}{1+\delta}} (b_n - a_n) \log(n)^{\frac{1}{1+\delta}+1} (z_{n1} + z_{n2}) \right) \\
&= o(1),
\end{aligned}$$

nach Voraussetzung.

Damit gilt die Gleichung (9.1.14) und somit

$$P \left(\sqrt{n} \max_{h,i,k,l} |H(s_h, t_i, d_{1k}, d_{2l}) - H(s_h, t_i, 0, 1)| > \eta \right) = o(1).$$

Also gilt die Gleichung (9.1.10) und analog auch die Gleichungen (9.1.11) und (9.1.12). □

Lemma 9.2 Für die Gewichtsfunktion $w_n : \mathbb{R} \rightarrow [0, 1]$ gilt

$$\sup_{s \in [0,1]} \left| \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} - \frac{\lfloor ns \rfloor}{n} \right| = o_P(1).$$

Beweis Es gilt

$$\sup_{s \in [0,1]} \left| \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} - \frac{\lfloor ns \rfloor}{n} \right| = \sup_{s \in [0,1]} \left| \frac{1}{n} \sum_{j=1}^{\lfloor ns \rfloor} (w_n(X_{j-1}) - 1) \right|$$

$$= \frac{1}{n} \sum_{j=1}^n (1 - w_n(X_{j-1})) \quad (w_n : \mathbb{R} \rightarrow [0, 1])$$

und

$$\begin{aligned} & \frac{1}{n} \sum_{j=1}^n (1 - w_n(X_{j-1})) \\ & \leq \left| 1 - E \left[\frac{1}{n} \sum_{j=1}^n w_n(X_{j-1}) \right] \right| + \left| \frac{1}{n} \sum_{j=1}^n w_n(X_{j-1}) - E \left[\frac{1}{n} \sum_{j=1}^n w_n(X_{j-1}) \right] \right|. \end{aligned}$$

Weiter ist

$$\begin{aligned} & \left| 1 - \frac{1}{n} \sum_{j=1}^n E[w_n(X_{j-1})] \right| \\ & = \left| \frac{1}{n} \sum_{j=1}^n E[1 - w_n(X_{j-1})] \right| \\ & = \left| \frac{1}{n} \sum_{j=1}^n \left(\int_{-\infty}^{a_n + \kappa} (1 - w_n(x)) f_{X_{j-1}}(x) dx + \int_{b_n - \kappa}^{\infty} (1 - w_n(x)) f_{X_{j-1}}(x) dx \right) \right| \end{aligned}$$

da $w_n(x) = 1$ für alle $x \in [a_n + \kappa, b_n - \kappa]$ nach Definition der Gewichtsfunktion (siehe 4.1, S. 45).

Da $0 \leq (1 - w_n(x)) \leq 1$ für alle x gilt weiter

$$\begin{aligned} & \left| 1 - \frac{1}{n} \sum_{j=1}^n E[w_n(X_{j-1})] \right| \\ & \leq \frac{1}{n} \sum_{j=1}^n \left(\int_{-\infty}^{a_n + \kappa} f_{X_{j-1}}(x) dx + \int_{b_n - \kappa}^{\infty} f_{X_{j-1}}(x) dx \right) \\ & = o(1), \end{aligned}$$

da $a_n \xrightarrow{n \rightarrow \infty} -\infty$, $b_n \xrightarrow{n \rightarrow \infty} \infty$ und $f_{X_{j-1}}(x) \xrightarrow{|x| \rightarrow \infty} 0$ für alle $j = 1, \dots, n$.

Damit gilt für die Varianz der Summe

$$\begin{aligned} \text{Var} \left(\frac{1}{n} \sum_{j=1}^n w_n(X_{j-1}) \right) & = \text{Var} \left(\frac{1}{n} \sum_{j=1}^n (1 - w_n(X_{j-1})) \right) \\ & = E \left[\left(\frac{1}{n} \sum_{j=1}^n (1 - w_n(X_{j-1})) \right)^2 \right] - \left(\frac{1}{n} \sum_{j=1}^n E[1 - w_n(X_{j-1})] \right)^2 \\ & = E \left[\left(\frac{1}{n} \sum_{j=1}^n (1 - w_n(X_{j-1})) \right)^2 \right] + o(1). \end{aligned}$$

Weiter ist

$$E \left[\left(\frac{1}{n} \sum_{j=1}^n (1 - w_n(X_{j-1})) \right)^2 \right]$$

$$\begin{aligned}
&= \left(\frac{1}{n^2} \sum_{k=1}^n \sum_{i=1}^n E[(1 - w_n(X_{k-1}))(1 - w_n(X_{i-1}))] \right) \\
&= \frac{1}{n^2} \sum_{i=1}^n \underbrace{\int (1 - w_n(x_i))^2 f_{X_{i-1}}(x_i) dx_i}_{\leq 1} \\
&\quad + \frac{1}{n^2} \sum_{k=1}^n \sum_{\substack{i=1 \\ i \neq k}}^n \int \int (1 - w_n(x_k))(1 - w_n(x_i)) f_{X_{k-1}, X_{i-1}}(x_k, x_i) dx_k dx_i \\
&\leq \frac{1}{n} \\
&\quad + \frac{1}{n^2} \sum_{k=1}^n \sum_{\substack{i=1 \\ i \neq k}}^n \int_{-\infty}^{a_n + \kappa} \int_{-\infty}^{a_n + \kappa} f_{X_{k-1}, X_{i-1}}(x_k, x_i) dx_k dx_i \\
&\quad + \frac{1}{n^2} \sum_{k=1}^n \sum_{\substack{i=1 \\ i \neq k}}^n \int_{b_n - \kappa}^{\infty} \int_{b_n - \kappa}^{\infty} f_{X_{k-1}, X_{i-1}}(x_k, x_i) dx_k dx_i \\
&= \frac{1}{n} \\
&\quad + \frac{1}{n^2} \sum_{k=1}^n \sum_{\substack{i=1 \\ i \neq k}}^n \int_{-\infty}^{a_n + \kappa} \underbrace{\int_{-\infty}^{a_n + \kappa} f_{X_{k-1}|X_{i-1}}(x_k|x_i) dx_k}_{\leq 1} f_{X_{i-1}}(x_i) dx_i \\
&\quad + \frac{1}{n^2} \sum_{k=1}^n \sum_{\substack{i=1 \\ i \neq k}}^n \int_{b_n - \kappa}^{\infty} \underbrace{\int_{b_n - \kappa}^{\infty} f_{X_{k-1}|X_{i-1}}(x_k|x_i) dx_k}_{\leq 1} f_{X_{i-1}}(x_i) dx_i \\
&\leq \frac{1}{n} \\
&\quad + \frac{1}{n^2} \sum_{k=1}^n \sum_{\substack{i=1 \\ i \neq k}}^n \int_{-\infty}^{a_n + \kappa} f_{X_{i-1}}(x_i) dx_i \\
&\quad + \frac{1}{n^2} \sum_{k=1}^n \sum_{\substack{i=1 \\ i \neq k}}^n \int_{b_n - \kappa}^{\infty} f_{X_{i-1}}(x_i) dx_i \\
&\leq \frac{1}{n} \\
&\quad + \frac{1}{n} \sum_{i=1}^n \left(\int_{-\infty}^{a_n + \kappa} f_{X_{i-1}}(x) dx + \int_{b_n - \kappa}^{\infty} f_{X_{i-1}}(x) dx \right) \\
&= o(1).
\end{aligned}$$

Damit gilt mit der Tchebychev-Ungleichung für alle $\eta > 0$

$$\begin{aligned}
P \left(\left| \frac{1}{n} \sum_{j=1}^n w_n(X_{j-1}) - E \left[\frac{1}{n} \sum_{j=1}^n w_n(X_{j-1}) \right] \right| > \eta \right) &\leq \frac{\text{Var} \left(\frac{1}{n} \sum_{j=1}^n w_n(X_{j-1}) \right)}{\eta^2} \\
&= o(1).
\end{aligned}$$

Somit ist die Behauptung bewiesen. □

Lemma 9.3 Der Prozess $(X_j)_{j \in \mathbb{Z}}$ sei α -mischend mit $\alpha(n) = O(n^{-\beta})$ für ein $\beta > 4$.

Für die Dichten f_{ϵ_j} der Innovationen ϵ_j gelte $\frac{1}{n} \sum_{j=1}^n \sup_{t \in \mathbb{R}} |f_{\epsilon_j}(t)| = O(1)$.

Für das Intervall $I_n = [a_n, b_n]$ gelte für alle $m_n \leq n$ mit $m_n^{-1} = o(1)$

$$\frac{1}{m_n} \max_{0 \leq S \leq n - m_n} \sum_{i=S+1}^{S+m_n} \left(\int_{-\infty}^{a_n + \kappa} f_{X_{j-1}}(x) dx + \int_{b_n - \kappa}^{\infty} f_{X_{j-1}}(x) dx \right) = o\left(\frac{1}{\log(n)}\right),$$

wobei $\kappa > 0$ in 4.1 (S. 45) definiert ist.

Dann ist

$$\sup_{s \in [0,1], t \in \mathbb{R}} \left| \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \left(\sum_{j=1}^{\lfloor ns \rfloor} \bar{w}_j \{I\{\epsilon_j \leq t\} - F_{\epsilon_j}(t)\} \right) - \frac{1}{n} \sum_{j=1}^{\lfloor ns \rfloor} (I\{\epsilon_j \leq t\} - F_{\epsilon_j}(t)) \right| = o_p\left(\frac{1}{\sqrt{n}}\right).$$

Beweis Es ist zu zeigen, dass für alle $\eta > 0$ gilt

$$P \left(\sqrt{n} \sup_{s \in [0,1], t \in \mathbb{R}} \left| \frac{1}{n} \sum_{j=1}^{\lfloor ns \rfloor} (w_n(X_{j-1}) - 1) (I\{\epsilon_j \leq t\} - F_{\epsilon_j}(t)) \right| > \eta \right) = o(1).$$

Als erstes wird gezeigt, dass es ausreicht $\sup_{t \in [-\sqrt{n}, \sqrt{n}]}$ anstatt $\sup_{t \in \mathbb{R}}$ zu betrachten.

Analog zum Vorgehen im Beweis von Lemma 9.1 ergibt sich für die Innovationen ϵ_j

$$P(\max_{1 \leq j \leq n} |\epsilon_j| > \sqrt{n} \log(n)) = O\left(\frac{1}{\log(n)}\right) = o(1).$$

Daher kann man für die weitere Berechnung davon ausgehen, dass $|\epsilon_j| \leq \sqrt{n} \log(n)$ für alle $j = 1, \dots, n$.

Damit gilt

$$\begin{aligned} & \sup_{\substack{s \in [0,1], \\ t \notin [-\sqrt{n} \log(n), \sqrt{n} \log(n)]}} \left| \frac{1}{n} \sum_{j=1}^{\lfloor ns \rfloor} (w_n(X_{j-1}) - 1) (I\{\epsilon_j \leq t\} - F_{\epsilon_j}(t)) \right| \\ & \leq \sup_{\substack{s \in [0,1], \\ t > \sqrt{n} \log(n)}} \left| \frac{1}{n} \sum_{j=1}^n I\left\{\frac{j}{n} \leq s\right\} (w_n(X_{j-1}) - 1) (1 - F_{\epsilon_j}(t)) \right| \\ & \quad + \sup_{\substack{s \in [0,1], \\ t < -\sqrt{n} \log(n)}} \left| \frac{1}{n} \sum_{j=1}^n I\left\{\frac{j}{n} \leq s\right\} (w_n(X_{j-1}) - 1) (0 - F_{\epsilon_j}(t)) \right| \\ & \leq \frac{1}{n} \sum_{j=1}^n (1 - w_n(X_{j-1})) (1 - F_{\epsilon_j}(\sqrt{n} \log(n))) \\ & \quad + \frac{1}{n} \sum_{j=1}^n (1 - w_n(X_{j-1})) F_{\epsilon_j}(-\sqrt{n} \log(n)), \end{aligned}$$

$$\begin{aligned}
& \text{da } F_{\epsilon_j} \text{ monoton wachsend sind,} \\
& \leq \frac{1}{n} \sum_{j=1}^n (1 - F_{\epsilon_j}(\sqrt{n} \log(n))) \\
& \quad + \frac{1}{n} \sum_{j=1}^n F_{\epsilon_j}(-\sqrt{n} \log(n)). \tag{9.3.1}
\end{aligned}$$

Weiter folgt analog zum Vorgehen im Beweis von Lemma 9.1

$$\frac{1}{n} \sum_{j=1}^n (1 - F_{\epsilon_j}(t)) = O\left(\frac{1}{t^2}\right) \quad \text{für } t > 1$$

und

$$\frac{1}{n} \sum_{j=1}^n F_{\epsilon_j}(t) = O\left(\frac{1}{t^2}\right) \quad \text{für } t < -1,$$

und damit aus (9.3.1)

$$\begin{aligned}
\sup_{\substack{s \in [0,1], \\ t \notin [-\sqrt{n} \log(n), \sqrt{n} \log(n)]}} \left| \frac{1}{n} \sum_{j=1}^{\lfloor ns \rfloor} (w_n(X_{j-1}) - 1) (I\{\epsilon_j \leq t\} - F_{\epsilon_j}(t)) \right| &= O\left(\frac{1}{n \log(n)^2}\right) \\
&= o\left(\frac{1}{\sqrt{n}}\right).
\end{aligned}$$

Es bleibt also zu zeigen, dass

$$\sup_{\substack{s \in [0,1], \\ t \in [-\sqrt{n} \log(n), \sqrt{n} \log(n)]}} \left| \frac{1}{n} \sum_{j=1}^{\lfloor ns \rfloor} (w_n(X_{j-1}) - 1) (I\{\epsilon_j \leq t\} - F_{\epsilon_j}(t)) \right| = o_p\left(\frac{1}{\sqrt{n}}\right).$$

Sei $\bar{\epsilon}_n = n^{-\frac{1}{2}} \log(n)^{-1}$ und das Intervall $[0, 1]$ sei durch die Punkte $0 = s_1 < \dots < s_{M_{n,1}} = 1$ in Abschnitte unterteilt, so dass $s_h - s_{h-1} \leq \bar{\epsilon}_n$ für alle $i = 2, \dots, M_{n,1}$ und $M_{n,1} \leq \frac{1}{\bar{\epsilon}_n}$. Das Intervall $[-\sqrt{n} \log(n), \sqrt{n} \log(n)]$ werde durch die Punkte $-\sqrt{n} \log(n) = t_1 < \dots < t_{M_{n,2}} = \sqrt{n} \log(n)$ ebenfalls in Abschnitte unterteilt, so dass $t_i - t_{i-1} \leq \bar{\epsilon}_n$ für alle $i = 2, \dots, M_{n,2}$ und $M_{n,2} \leq \frac{2\sqrt{n} \log(n)}{\bar{\epsilon}_n}$. Damit ist

$$\begin{aligned}
& \sup_{\substack{s \in [0,1], \\ t \in [-\sqrt{n} \log(n), \sqrt{n} \log(n)]}} \left| \frac{1}{n} \sum_{j=1}^{\lfloor ns \rfloor} (w_n(X_{j-1}) - 1) (I\{\epsilon_j \leq t\} - F_{\epsilon_j}(t)) \right| \\
& \leq \max_{\substack{1 \leq h \leq M_{n,1}, \\ 1 \leq i \leq M_{n,2}}} \left| \frac{1}{n} \sum_{j=1}^{\lfloor ns_h \rfloor} (w_n(X_{j-1}) - 1) (I\{\epsilon_j \leq t_i\} - F_{\epsilon_j}(t_i)) \right| \\
& \quad + \max_{1 \leq h \leq M_{n,1}} \sup_{\substack{|s - s_h| \leq \bar{\epsilon}_n, \\ t \in [-\sqrt{n} \log(n), \sqrt{n} \log(n)]}} \left| \frac{1}{n} \sum_{j=1}^{\lfloor ns \rfloor} (w_n(X_{j-1}) - 1) (I\{\epsilon_j \leq t\} - F_{\epsilon_j}(t)) \right. \\
& \qquad \qquad \qquad \left. - \frac{1}{n} \sum_{j=1}^{\lfloor ns_h \rfloor} (w_n(X_{j-1}) - 1) (I\{\epsilon_j \leq t\} - F_{\epsilon_j}(t)) \right| \tag{9.3.2}
\end{aligned}$$

$$\begin{aligned}
& + \max_{\substack{1 \leq h \leq M_{n,1}, \\ 1 \leq i \leq M_{n,2}}} \sup_{|t-t_i| \leq \bar{\epsilon}_n} \left| \frac{1}{n} \sum_{j=1}^{\lfloor ns_h \rfloor} (w_n(X_{j-1}) - 1) (I\{\epsilon_j \leq t\} - F_{\epsilon_j}(t)) \right. \\
& \qquad \qquad \qquad \left. - \frac{1}{n} \sum_{j=1}^{\lfloor ns_h \rfloor} (w_n(X_{j-1}) - 1) (I\{\epsilon_j \leq t_i\} - F_{\epsilon_j}(t_i)) \right|. \quad (9.3.3)
\end{aligned}$$

Als nächstes wird gezeigt, dass der Term (9.3.2) von der Ordnung $O(\bar{\epsilon}_n)$ ist, und der Term (9.3.3) abgeschätzt werden kann durch

$$\begin{aligned}
& \max_{1 \leq i \leq M_{n,2}} \left| \frac{1}{n} \sum_{j=1}^n (w_n(X_{j-1}) - 1) (I\{\epsilon_j \leq t_i + \bar{\epsilon}_n\} - F_{\epsilon_j}(t_i + \bar{\epsilon}_n)) \right| \\
& + \max_{1 \leq i \leq M_{n,2}} \left| \frac{1}{n} \sum_{j=1}^n (w_n(X_{j-1}) - 1) (I\{\epsilon_j \leq t_i - \bar{\epsilon}_n\} - F_{\epsilon_j}(t_i - \bar{\epsilon}_n)) \right| \\
& + O(\bar{\epsilon}_n).
\end{aligned}$$

Der Term (9.3.2) kann abgeschätzt werden durch

$$\begin{aligned}
& \max_{1 \leq h \leq M_{n,1}} \sup_{\substack{|s-s_h| \leq \bar{\epsilon}_n, \\ t \in [-\sqrt{n} \log(n), \sqrt{n} \log(n)]}} \frac{1}{n} \sum_{j=1}^n |w_n(X_{j-1}) - 1| |I\{\epsilon_j \leq t\} - F_{\epsilon_j}(t)| \cdot \left| I\left\{\frac{j}{n} \leq s\right\} - I\left\{\frac{j}{n} \leq s_h\right\} \right| \\
& = O(\bar{\epsilon}_n) \quad \text{siehe (9.1.9) (S. 126) im Beweis zu Lemma 9.1.}
\end{aligned}$$

Für (9.3.3) gilt

$$\begin{aligned}
& \max_{\substack{1 \leq h \leq M_{n,1}, \\ 1 \leq i \leq M_{n,2}}} \sup_{|t-t_i| \leq \bar{\epsilon}_n} \left| \frac{1}{n} \sum_{j=1}^{\lfloor ns_h \rfloor} (w_n(X_{j-1}) - 1) (I\{\epsilon_j \leq t\} - F_{\epsilon_j}(t)) \right. \\
& \qquad \qquad \qquad \left. - \frac{1}{n} \sum_{j=1}^{\lfloor ns_h \rfloor} (w_n(X_{j-1}) - 1) (I\{\epsilon_j \leq t_i\} - F_{\epsilon_j}(t_i)) \right| \\
& = \max_{\substack{1 \leq h \leq M_{n,1}, \\ 1 \leq i \leq M_{n,2}}} \sup_{|t-t_i| \leq \bar{\epsilon}_n} \left| \frac{1}{n} \sum_{j=1}^{\lfloor ns_h \rfloor} (w_n(X_{j-1}) - 1) ((I\{\epsilon_j \leq t\} - I\{\epsilon_j \leq t_i\}) + (F_{\epsilon_j}(t_i) - F_{\epsilon_j}(t))) \right| \\
& \leq \max_{\substack{1 \leq h \leq M_{n,1}, \\ 1 \leq i \leq M_{n,2}}} \frac{1}{n} \sum_{j=1}^{\lfloor ns_h \rfloor} (1 - w_n(X_{j-1})) ((I\{\epsilon_j \leq t_i + \bar{\epsilon}_n\} - I\{\epsilon_j \leq t_i - \bar{\epsilon}_n\}) + (F_{\epsilon_j}(t_i + \bar{\epsilon}_n) - F_{\epsilon_j}(t_i - \bar{\epsilon}_n))) \\
& = \max_{1 \leq i \leq M_{n,2}} \frac{1}{n} \sum_{j=1}^n (1 - w_n(X_{j-1})) ((I\{\epsilon_j \leq t_i + \bar{\epsilon}_n\} - I\{\epsilon_j \leq t_i - \bar{\epsilon}_n\}) + (F_{\epsilon_j}(t_i + \bar{\epsilon}_n) - F_{\epsilon_j}(t_i - \bar{\epsilon}_n))) \\
& \leq \max_{1 \leq i \leq M_{n,2}} \left| \frac{1}{n} \sum_{j=1}^n (w_n(X_{j-1}) - 1) (I\{\epsilon_j \leq t_i + \bar{\epsilon}_n\} - F_{\epsilon_j}(t_i + \bar{\epsilon}_n)) \right| \\
& \quad + \max_{1 \leq i \leq M_{n,2}} \left| \frac{1}{n} \sum_{j=1}^n (w_n(X_{j-1}) - 1) (I\{\epsilon_j \leq t_i - \bar{\epsilon}_n\} - F_{\epsilon_j}(t_i - \bar{\epsilon}_n)) \right|
\end{aligned}$$

$$+2 \max_{1 \leq i \leq M_{n,2}} \frac{1}{n} \sum_{j=1}^n (F_{\epsilon_j}(t_i + \bar{\epsilon}_n) - F_{\epsilon_j}(t_i - \bar{\epsilon}_n)).$$

Da mit dem Mittelwertsatz weiter folgt

$$\max_{1 \leq i \leq M_{n,2}} \frac{1}{n} \sum_{j=1}^n (F_{\epsilon_j}(t_i + \bar{\epsilon}_n) - F_{\epsilon_j}(t_i - \bar{\epsilon}_n)) \leq 2\bar{\epsilon}_n \underbrace{\frac{1}{n} \sum_{j=1}^n \sup_{t \in \mathbb{R}} |f_{\epsilon_j}(t)|}_{=O(1) \text{ nach Voraussetzung}}$$

kann (9.3.3) durch

$$\begin{aligned} & \max_{1 \leq i \leq M_{n,2}} \left| \frac{1}{n} \sum_{j=1}^n (w_n(X_{j-1}) - 1) (I\{\epsilon_j \leq t_i + \bar{\epsilon}_n\} - F_{\epsilon_j}(t_i + \bar{\epsilon}_n)) \right| \\ & + \max_{1 \leq i \leq M_{n,2}} \left| \frac{1}{n} \sum_{j=1}^n (w_n(X_{j-1}) - 1) (I\{\epsilon_j \leq t_i - \bar{\epsilon}_n\} - F_{\epsilon_j}(t_i - \bar{\epsilon}_n)) \right| \\ & + O(\bar{\epsilon}_n) \end{aligned}$$

abgeschätzt werden, und es bleibt zu zeigen, dass

$$\max_{\substack{1 \leq h \leq M_{n,1}, \\ 1 \leq i \leq M_{n,2}}} \left| \frac{1}{n} \sum_{j=1}^{\lfloor ns_h \rfloor} (w_n(X_{j-1}) - 1) (I\{\epsilon_j \leq t_i\} - F_{\epsilon_j}(t_i)) \right| = o_p\left(\frac{1}{\sqrt{n}}\right), \quad (9.3.4)$$

sowie

$$\max_{1 \leq i \leq M_{n,2}} \left| \frac{1}{n} \sum_{j=1}^n (w_n(X_{j-1}) - 1) (I\{\epsilon_j \leq t_i + \bar{\epsilon}_n\} - F_{\epsilon_j}(t_i + \bar{\epsilon}_n)) \right| = o_p\left(\frac{1}{\sqrt{n}}\right) \quad (9.3.5)$$

und

$$\max_{1 \leq i \leq M_{n,2}} \left| \frac{1}{n} \sum_{j=1}^n (w_n(X_{j-1}) - 1) (I\{\epsilon_j \leq t_i - \bar{\epsilon}_n\} - F_{\epsilon_j}(t_i - \bar{\epsilon}_n)) \right| = o_p\left(\frac{1}{\sqrt{n}}\right). \quad (9.3.6)$$

Dies geschieht mit Hilfe des folgenden Hilfssatzes, der auf einer Ungleichung vom Bernstein-Typ für mischende Zufallsvariablen basiert.

Hilfssatz 9.3.A Der Prozess $(X_j)_{j \in \mathbb{Z}}$ sei α -mischend mit Mischungskoeffizient $\alpha(\cdot)$.

Dann gilt für alle festen $s \in [0, 1]$ und $t \in \mathbb{R}$, sowie alle $\eta > 0$

$$\begin{aligned} & P \left(\sqrt{n} \cdot \left| \frac{1}{n} \sum_{j=1}^{\lfloor ns \rfloor} (w_n(X_{j-1}) - 1) (I\{\epsilon_j \leq t\} - F_{\epsilon_j}(t)) \right| > \eta \right) \\ & \leq 4 \exp \left(- \frac{n\eta^2}{64n\omega(\lfloor n^{\frac{1}{2}} \log(n)^{-2} \rfloor, n) + \frac{8}{3}\sqrt{n}\eta \lfloor n^{\frac{1}{2}} \log(n)^{-2} \rfloor} \right) + 4 \frac{n}{\lfloor n^{\frac{1}{2}} \log(n)^{-2} \rfloor} \alpha(\lfloor n^{\frac{1}{2}} \log(n)^{-2} \rfloor), \end{aligned}$$

wobei $\omega(m, n) := \frac{1}{m} \max_{0 \leq S \leq n-m} \sum_{j=S+1}^{S+m} E[(w_n(X_{j-1}) - 1)^2]$, $m, n \in \mathbb{N}$.

Beweis Messbare Funktionen erhalten nach Fan and Yao [2005] 2.6.1 (ii) Mischungs-Eigenschaften, daher ist $(w_n(X_{j-1}) - 1) (I\{\epsilon_j \leq t\} - F_{\epsilon_j}(t))$ α -mischend mit dem selben Mischungskoeffizienten wie (X_j) . Der Mischungskoeffizient von

$$Y_j = (w_n(X_{j-1}) - 1) (I\{\epsilon_j \leq t\} - F_{\epsilon_j}(t)) I \left\{ \frac{j}{n} \leq s \right\}$$

kann für festes $s \in [0, 1]$ und $n \in \mathbb{N}$ durch den Mischungskoeffizienten $\alpha(\cdot)$ abgeschätzt werden.

Außerdem ist

$$E[Y_j] = 0$$

und

$$|Y_j| \leq 1.$$

Damit ist das Theorem 2.1 aus Liebscher [1996] auf $\sum_{j=1}^n Y_j$ anwendbar. Dieses Theorem basiert auf Theorem 5 in Rio [1995].

Sei

$$N = \lfloor n^{\frac{1}{2}} \log(n)^{-2} \rfloor,$$

dann ist $1 \leq N \leq n$ und $\sqrt{n}\eta > 4N$ für hinreichend große n .

Da ϵ_j unabhängig von allen X_i und ϵ_i mit $i < j$ ist, gilt für $0 \leq T \leq n - 1$

$$\begin{aligned} & E \left[\left(\sum_{j=T+1}^{\min(T+N,n)} Y_j \right)^2 \right] \\ &= \sum_{j=T+1}^{\min(T+N,n)} \sum_{i=T+1}^{\min(T+N,n)} E [Y_j Y_i] \\ &= \sum_{j=T+1}^{\min(T+N,n)} \sum_{i=T+1}^{j-1} E [E [Y_j Y_i | X_k, \epsilon_k \ k \leq j-1]] \\ &\quad + \sum_{j=T+1}^{\min(T+N,n)} E [Y_j^2] \\ &\quad + \sum_{j=T+1}^{\min(T+N,n)} \sum_{i=j+1}^{\min(T+N,n)} E [E [Y_j Y_i | X_k, \epsilon_k \ k \leq i-1]] \\ &= \sum_{j=T+1}^{\min(T+N,n)} E [Y_j^2], \end{aligned}$$

da für alle $i \leq j - 1$ gilt

$$\begin{aligned} & E [E [Y_j Y_i | X_k, \epsilon_k \ k \leq j-1]] \\ &= E [E [I\{\epsilon_j \leq t\} - F_{\epsilon_j}(t)] (w_n(X_{j-1}) - 1) (w_n(X_{i-1}) - 1) (I\{\epsilon_i \leq t\} - F_{\epsilon_i}(t))] I \left\{ \frac{j}{n} \leq s \right\} \\ &= 0. \end{aligned}$$

Damit ist

$$E \left[\left(\sum_{j=T+1}^{\min(T+N,n)} Y_j \right)^2 \right]$$

$$\begin{aligned}
&= \sum_{j=T+1}^{\min(T+N,n)} E \left[(w_n(X_{j-1}) - 1)^2 (I\{\epsilon_j \leq t\} - F_{\epsilon_j}(t))^2 \right] I \left\{ \frac{j}{n} \leq s \right\} \\
&= \sum_{j=T+1}^{\min(T+N,n)} E \left[(w_n(X_{j-1}) - 1)^2 \right] E \left[(I\{\epsilon_j \leq t\} - F_{\epsilon_j}(t))^2 \right] I \left\{ \frac{j}{n} \leq s \right\} \\
&\leq \sum_{j=T+1}^{\min(T+N,n)} E \left[(w_n(X_{j-1}) - 1)^2 \right] \\
&\leq N \frac{1}{N} \max_{0 \leq S \leq n-N} \sum_{j=S+1}^{S+N} E \left[(w_n(X_{j-1}) - 1)^2 \right] \\
&= N\omega(N, n),
\end{aligned}$$

und nach Liebscher [1996] gilt

$$\begin{aligned}
&P \left(\sqrt{n} \cdot \left| \frac{1}{n} \sum_{j=1}^n Y_j \right| > \eta \right) \\
&\leq 4 \exp \left(- \frac{n\eta^2}{64 \frac{n}{N} N\omega(N, n) + \frac{8}{3} \sqrt{n}\eta N} \right) + 4 \frac{n}{N} \alpha(N) \\
&= 4 \exp \left(- \frac{n\eta^2}{64n\omega(\lfloor n^{\frac{1}{2}} \log(n)^{-2} \rfloor, n) + \frac{8}{3} \sqrt{n}\eta \lfloor n^{\frac{1}{2}} \log(n)^{-2} \rfloor} \right) + 4 \frac{n}{\lfloor n^{\frac{1}{2}} \log(n)^{-2} \rfloor} \alpha(\lfloor n^{\frac{1}{2}} \log(n)^{-2} \rfloor)
\end{aligned}$$

□

Fortsetzung Beweis Lemma 9.3

Mit diesem Hilfssatz gilt nun für (9.3.4) für alle $\eta > 0$

$$\begin{aligned}
&P \left(\sqrt{n} \max_{\substack{1 \leq h \leq M_{n,1}, \\ 1 \leq i \leq M_{n,2}}} \left| \frac{1}{n} \sum_{j=1}^{\lfloor ns_h \rfloor} (w_n(X_{j-1}) - 1) (I\{\epsilon_j \leq t_i\} - F_{\epsilon_j}(t_i)) \right| > \eta \right) \\
&\leq \sum_{h=1}^{M_{n,1}} \sum_{i=1}^{M_{n,2}} P \left(\sqrt{n} \cdot \left| \frac{1}{n} \sum_{j=1}^{\lfloor ns_h \rfloor} (w_n(X_{j-1}) - 1) (I\{\epsilon_j \leq t_i\} - F_{\epsilon_j}(t_i)) \right| > \eta \right) \\
&\leq M_{n,1} M_{n,2} \left(4 \exp \left(- \frac{n\eta^2}{64n\omega(\lfloor n^{\frac{1}{2}} \log(n)^{-2} \rfloor, n) + \frac{8}{3} \sqrt{n}\eta \lfloor n^{\frac{1}{2}} \log(n)^{-2} \rfloor} \right) \right. \\
&\quad \left. + 4 \frac{n}{\lfloor n^{\frac{1}{2}} \log(n)^{-2} \rfloor} \alpha(\lfloor n^{\frac{1}{2}} \log(n)^{-2} \rfloor) \right) \\
&= 4 \exp \left(\log(M_{n,1} M_{n,2}) - \frac{n\eta^2}{64n\omega(\lfloor n^{\frac{1}{2}} \log(n)^{-2} \rfloor, n) + \frac{8}{3} \sqrt{n}\eta \lfloor n^{\frac{1}{2}} \log(n)^{-2} \rfloor} \right) \tag{9.3.7} \\
&\quad + O \left(M_{n,1} M_{n,2} \frac{n}{\lfloor n^{\frac{1}{2}} \log(n)^{-2} \rfloor} \alpha(\lfloor n^{\frac{1}{2}} \log(n)^{-2} \rfloor) \right).
\end{aligned}$$

Die Zeile (9.3.7) ist $o(1)$, wenn der Exponent gegen minus unendlich geht, also wenn

$$\frac{64n\omega(\lfloor n^{\frac{1}{2}} \log(n)^{-2} \rfloor, n) + \frac{8}{3}\sqrt{n}\eta \lfloor n^{\frac{1}{2}} \log(n)^{-2} \rfloor}{n\eta^2} = o\left(\frac{1}{\log(M_{n,1}M_{n,2})}\right).$$

Dies ist er Fall, da

$$\begin{aligned} & \log(M_{n,1}M_{n,2}) \frac{64n\omega(\lfloor n^{\frac{1}{2}} \log(n)^{-2} \rfloor, n) + \frac{8}{3}\sqrt{n}\eta \lfloor n^{\frac{1}{2}} \log(n)^{-2} \rfloor}{n\eta^2} \\ &= O\left(\log\left(\frac{\sqrt{n}\log(n)}{\bar{\epsilon}_n^2}\right) \left(\omega(\lfloor n^{\frac{1}{2}} \log(n)^{-2} \rfloor, n) + \log(n)^{-2}\right)\right) \\ &= O\left(\log(n) \left(\omega(\lfloor n^{\frac{1}{2}} \log(n)^{-2} \rfloor, n) + \log(n)^{-2}\right)\right) \\ &= o(1), \end{aligned}$$

da

$$\begin{aligned} & \omega(\lfloor n^{\frac{1}{2}} \log(n)^{-2} \rfloor, n) \\ &= \frac{1}{\lfloor n^{\frac{1}{2}} \log(n)^{-2} \rfloor} \max_{0 \leq S \leq n - \lfloor n^{\frac{1}{2}} \log(n)^{-2} \rfloor} \sum_{j=S+1}^{S+\lfloor n^{\frac{1}{2}} \log(n)^{-2} \rfloor} E[(w_n(X_{j-1}) - 1)^2] \\ &\leq \frac{1}{\lfloor n^{\frac{1}{2}} \log(n)^{-2} \rfloor} \max_{0 \leq S \leq n - \lfloor n^{\frac{1}{2}} \log(n)^{-2} \rfloor} \sum_{j=S+1}^{S+\lfloor n^{\frac{1}{2}} \log(n)^{-2} \rfloor} \left(\int_{-\infty}^{a_n+\kappa} f_{X_{j-1}}(x) dx + \int_{b_n-\kappa}^{\infty} f_{X_{j-1}}(x) dx \right), \\ & \text{da } w_n(x) = 1 \quad \forall x \in [a_n + \kappa, b_n - \kappa] \\ &= o\left(\frac{1}{\log(n)}\right) \end{aligned}$$

nach Voraussetzung.

Außerdem gilt

$$\begin{aligned} & O\left(M_{n,1}M_{n,2} \frac{n}{\lfloor n^{\frac{1}{2}} \log(n)^{-2} \rfloor} \alpha(\lfloor n^{\frac{1}{2}} \log(n)^{-2} \rfloor)\right) \\ &= O\left(\frac{\sqrt{n}\log(n)}{\bar{\epsilon}_n^2} n^{\frac{1}{2}} n^{-\frac{\beta}{2}} \log(n)^{2\beta}\right) \\ &= O\left(n^{\frac{1}{2}+1+\frac{1}{2}-\frac{\beta}{2}} \log(n)^{3+2\beta}\right) \\ &= O\left(n^{2-\frac{\beta}{2}} \log(n)^{3+2\beta}\right) \\ &= o(1), \end{aligned}$$

da $\beta > 4$ nach Voraussetzung.

Damit gilt also die Gleichung (9.3.4) und analog auch die Gleichungen (9.3.5) und (9.3.6).

□

Lemma 9.4 Für die Dichten f_{ϵ_j} der Innovationen ϵ_j gelte $\frac{1}{n} \sum_{j=1}^n \sup_{t \in \mathbb{R}} |f_{\epsilon_j}(t)| = O(1)$.

Dann ist

$$\sup_{s \in [0,1], t \in \mathbb{R}} \left| \frac{1}{n} \sum_{j=1}^{\lfloor ns \rfloor} (I\{\epsilon_j \leq t\} - F_{\epsilon_j}(t)) \right| = O_P \left(n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} \right).$$

Beweis Analog zum Beweis von Lemma 9.3 gilt

$$\begin{aligned} & \sup_{s \in [0,1], t \in \mathbb{R}} \left| \frac{1}{n} \sum_{j=1}^{\lfloor ns \rfloor} (I\{\epsilon_j \leq t\} - F_{\epsilon_j}(t)) \right| \\ & \leq \max_{\substack{1 \leq h \leq M_{n,1}, \\ 1 \leq i \leq M_{n,2}}} \left| \frac{1}{n} \sum_{j=1}^{\lfloor ns_h \rfloor} (I\{\epsilon_j \leq t_i\} - F_{\epsilon_j}(t_i)) \right| \\ & \quad + \max_{1 \leq i \leq M_{n,2}} \left| \frac{1}{n} \sum_{j=1}^n (I\{\epsilon_j \leq t_i + \bar{\epsilon}_n\} - F_{\epsilon_j}(t_i + \bar{\epsilon}_n)) \right| \\ & \quad + \max_{1 \leq i \leq M_{n,2}} \left| \frac{1}{n} \sum_{j=1}^n (I\{\epsilon_j \leq t_i - \bar{\epsilon}_n\} - F_{\epsilon_j}(t_i - \bar{\epsilon}_n)) \right| \\ & \quad + o\left(\frac{1}{\sqrt{n}}\right), \end{aligned}$$

wobei $\bar{\epsilon}_n = n^{-\frac{1}{2}} \log(n)^{-1}$ und das Intervall $[0, 1]$ durch die Punkte $0 = s_1 < \dots < s_{M_{n,1}} = 1$ in Abschnitte unterteilt sei, so dass $s_h - s_{h-1} \leq \bar{\epsilon}_n$ für alle $i = 2, \dots, M_{n,1}$ und $M_{n,1} \leq \frac{1}{\bar{\epsilon}_n}$, sowie das Intervall $[-\sqrt{n} \log(n), \sqrt{n} \log(n)]$ durch die Punkte $-\sqrt{n} \log(n) = t_1 < \dots < t_{M_{n,2}} = \sqrt{n} \log(n)$ ebenfalls in Abschnitte unterteilt werde, so dass $t_i - t_{i-1} \leq \bar{\epsilon}_n$ für alle $i = 2, \dots, M_{n,2}$ und $M_{n,2} \leq \frac{2\sqrt{n} \log(n)}{\bar{\epsilon}_n}$.

Nach der Bernstein-Ungleichung für unabhängige Zufallsvariablen gilt für alle festen $s \in [0, 1]$ und $t \in \mathbb{R}$ und alle $0 < \bar{M} < \infty$

$$\begin{aligned} & P \left(n^{\frac{1}{2}} \log(n)^{-\frac{1}{2}} \left| \frac{1}{n} \sum_{j=1}^{\lfloor ns \rfloor} (I\{\epsilon_j \leq t\} - F_{\epsilon_j}(t)) \right| > \bar{M} \right) \\ & = P \left(\left| \sum_{j=1}^n (I\{\epsilon_j \leq t\} - F_{\epsilon_j}(t)) I \left\{ \frac{j}{n} \leq s \right\} \right| > \bar{M} n^{\frac{1}{2}} \log(n)^{\frac{1}{2}} \right) \\ & = P \left(\sum_{j=1}^n (I\{\epsilon_j \leq t\} - F_{\epsilon_j}(t)) I \left\{ \frac{j}{n} \leq s \right\} > \bar{M} n^{\frac{1}{2}} \log(n)^{\frac{1}{2}} \right) \\ & \quad + P \left(\sum_{j=1}^n - (I\{\epsilon_j \leq t\} - F_{\epsilon_j}(t)) I \left\{ \frac{j}{n} \leq s \right\} > \bar{M} n^{\frac{1}{2}} \log(n)^{\frac{1}{2}} \right) \\ & \leq 2 \exp \left(- \frac{\bar{M}^2 n \log(n)}{n + \frac{1}{3} \bar{M} n^{\frac{1}{2}} \log(n)^{\frac{1}{2}}} \right), \end{aligned}$$

da $\left| (I\{\epsilon_j \leq t\} - F_{\epsilon_j}(t)) I \left\{ \frac{j}{n} \leq s \right\} \right| \leq 1$.

Damit folgt für alle $0 < \bar{M} < \infty$

$$\begin{aligned}
& P \left(n^{\frac{1}{2}} \log(n)^{-\frac{1}{2}} \max_{\substack{1 \leq h \leq M_{n,1}, \\ 1 \leq i \leq M_{n,2}}} \left| \frac{1}{n} \sum_{j=1}^{\lfloor ns_h \rfloor} (I\{\epsilon_j \leq t_i\} - F_{\epsilon_j}(t_i)) \right| > \bar{M} \right) \\
& \leq \sum_{h=1}^{M_{n,1}} \sum_{i=1}^{M_{n,2}} P \left(n^{\frac{1}{2}} \log(n)^{-\frac{1}{2}} \left| \frac{1}{n} \sum_{j=1}^{\lfloor ns_h \rfloor} (I\{\epsilon_j \leq t_i\} - F_{\epsilon_j}(t_i)) \right| > \bar{M} \right) \\
& \leq M_{n,1} M_{n,2} 2 \exp \left(- \frac{\bar{M}^2 n \log(n)}{n + \frac{1}{3} \bar{M} n^{\frac{1}{2}} \log(n)^{\frac{1}{2}}} \right) \\
& = 2 \exp \left(\log(M_{n,1} M_{n,2}) - \frac{\bar{M}^2 n \log(n)}{n + \frac{1}{3} \bar{M} n^{\frac{1}{2}} \log(n)^{\frac{1}{2}}} \right) \\
& = 2 \exp \left(\log(M_{n,1} M_{n,2}) - \underbrace{\frac{\bar{M}^2 \log(n)}{1 + \frac{1}{3} \bar{M} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}}}}_{\xrightarrow{n \rightarrow \infty} 0} \right).
\end{aligned}$$

Da $\log(M_{n,1} M_{n,2}) = O(\log(n))$ gibt es für alle $\eta > 0$ ein $\bar{M} < \infty$, so dass

$$P \left(n^{\frac{1}{2}} \log(n)^{-\frac{1}{2}} \max_{\substack{1 \leq h \leq M_{n,1}, \\ 1 \leq i \leq M_{n,2}}} \left| \frac{1}{n} \sum_{j=1}^{\lfloor ns_h \rfloor} (I\{\epsilon_j \leq t_i\} - F_{\epsilon_j}(t_i)) \right| > \bar{M} \right) < \eta.$$

Damit gilt also

$$\max_{\substack{1 \leq h \leq M_{n,1}, \\ 1 \leq i \leq M_{n,2}}} \left| \frac{1}{n} \sum_{j=1}^{\lfloor ns_h \rfloor} (I\{\epsilon_j \leq t_i\} - F_{\epsilon_j}(t_i)) \right| = O_P \left(n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} \right),$$

und genauso

$$\max_{1 \leq i \leq M_{n,2}} \left| \frac{1}{n} \sum_{j=1}^n (I\{\epsilon_j \leq t_i + \bar{\epsilon}_n\} - F_{\epsilon_j}(t_i + \bar{\epsilon}_n)) \right| = O_P \left(n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} \right)$$

und

$$\max_{1 \leq i \leq M_{n,2}} \left| \frac{1}{n} \sum_{j=1}^n (I\{\epsilon_j \leq t_i - \bar{\epsilon}_n\} - F_{\epsilon_j}(t_i - \bar{\epsilon}_n)) \right| = O_P \left(n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} \right).$$

□

10 Technische Lemmata

Lemma 10.1 Es seien die Voraussetzungen von Lemma 3.3 (S. 21) gegeben.

Für die Dichten der Beobachtungen gelte $\sup_{x \in [a_n - Cc_n, b_n + Cc_n]} \left| \frac{1}{n} \sum_{j=1}^n f_{X_{j-1}}(x) \right| = O(1)$.

Seien außerdem $g_n : \mathbb{R} \rightarrow \mathbb{R}$ Funktionen, die dreimal differenzierbar sind und $\sup_{x \in [a_n - Cc_n, b_n + Cc_n]} |g_n^{(\mu)}(x)| = O(1)$ für $\mu = 0, 1, 2, 3$ und (Z_i) unabhängige und identisch verteilte Zufallsvariablen mit $E[Z_i] = 0$ für alle i , die jeweils auch von X_{i-1} unabhängig sind, und für die $X_i = Z_i$ die Bedingungen an X_i für $k = 1$ in Lemma 3.3 erfüllt.

Dann ist für $\nu = 0, 1, 2$, $k_1 = 0, 1$ und $k_2 = 0, 1$

$$\begin{aligned} & \sup_{x \in I_n} \left| \frac{\partial^\nu}{\partial x^\nu} \left(\frac{1}{nc_n} \sum_{i=1}^n g_n(X_{i-1}) Z_i^{k_2} \left(\frac{x - X_{i-1}}{c_n} \right)^{k_1} K \left(\frac{x - X_{i-1}}{c_n} \right) \right. \right. \\ & \quad \left. \left. - E \left[\frac{1}{nc_n} \sum_{i=1}^n g_n(X_{i-1}) Z_i^{k_2} \left(\frac{x - X_{i-1}}{c_n} \right)^{k_1} K \left(\frac{x - X_{i-1}}{c_n} \right) \right] \right) \right| \\ &= O_P \left(c_n^{-\frac{1}{2}-\nu} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} + c_n^{2-\nu} \right). \end{aligned}$$

Beweis Aus Lemma 3.3 folgt für $\nu = 0, 1, 2$, $k_2 = 0, 1$ und $0 < m < \infty$

$$\begin{aligned} & \sup_{x \in I_n} \left| \frac{1}{nc_n} \sum_{i=1}^n Z_i^{k_2} \left(\frac{x - X_{i-1}}{c_n} \right)^m K^{(\nu)} \left(\frac{x - X_{i-1}}{c_n} \right) \right. \\ & \quad \left. - E \left[\frac{1}{nc_n} \sum_{i=1}^n Z_i^{k_2} \left(\frac{x - X_{i-1}}{c_n} \right)^m K^{(\nu)} \left(\frac{x - X_{i-1}}{c_n} \right) \right] \right| \\ &= O_P \left(c_n^{-\frac{1}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} \right), \end{aligned} \tag{10.1.1}$$

denn die Funktion $K^{(\nu)}$ in 3.3 kann durch $u^m K^{(\nu)}(u)$ ersetzt werden, da $u^m K(u)$ ebenfalls den Träger $[-C, C]$ hat und aus $\sup_{u \in [-C, C]} |K^{(\nu)}(u)| < \infty$ auch $\sup_{u \in [-C, C]} |u^m K^{(\nu)}(u)| < \infty$ folgt, sowie $\sup_{u \in [-C, C]} \left| \frac{\partial}{\partial u} (u^m K^{(\nu)}(u)) \right| < \infty$ aus $\sup_{u \in [-C, C]} |K^{(\nu+1)}(u)| < \infty$.

Um dies anwenden zu können, muss zunächst umgeformt werden.

Es gilt

$$\begin{aligned} & \frac{\partial^\nu}{\partial x^\nu} \left(E \left[\frac{1}{nc_n} \sum_{i=1}^n g_n(X_{i-1}) Z_i^{k_2} \left(\frac{x - X_{i-1}}{c_n} \right)^{k_1} K \left(\frac{x - X_{i-1}}{c_n} \right) \right] \right) \\ &= \begin{cases} \frac{1}{c_n^2} \frac{\partial^\nu}{\partial x^\nu} \left(\int c_n \frac{1}{n} \sum_{i=1}^n g_n(y) \left(\frac{x-y}{c_n} \right)^{k_1} K \left(\frac{x-y}{c_n} \right) f_{X_{i-1}}(y) dy \right) & \text{für } k_2 = 0 \\ 0 & \text{für } k_2 = 1, \end{cases} \end{aligned}$$

da Z_i und X_{i-1} unabhängig und $E[Z_i] = 0$ nach Voraussetzung.

Da für die Funktionen $h_n(x) := c_n \frac{1}{n} \sum_{i=1}^n g_n(y) f_{X_{i-1}}(y) \frac{\partial^\nu}{\partial x^\nu} \left(\left(\frac{x-y}{c_n} \right)^{k_1} K \left(\frac{x-y}{c_n} \right) \right)$ mit $y \in [a_n - Cc_n, b_n + Cc_n]$ nach dem Mittelwertsatz gilt

$$\begin{aligned} & |h_n(x_1) - h_n(x_2)| \\ & \leq \sup_{z \in I_n} \left| c_n \frac{1}{n} \sum_{i=1}^n g_n(y) f_{X_{i-1}}(y) \frac{\partial^{\nu+1}}{\partial x^{\nu+1}} \left(\left(\frac{x-y}{c_n} \right)^{k_1} K \left(\frac{x-y}{c_n} \right) \right) \Big|_{x=z} \right| \cdot |x_1 - x_2| \\ & \leq \underbrace{\sup_{u \in [-C, C]} \left| \frac{\partial^{\nu+1}}{\partial u^{\nu+1}} \left(u^{k_1} K(u) \right) \right| \sup_{y \in [a_n - Cc_n, b_n + Cc_n]} |g_n(y)| \sup_{y \in [a_n - Cc_n, b_n + Cc_n]} \left| \frac{1}{n} \sum_{i=1}^n f_{X_{i-1}}(y) \right|}_{=O(1), \text{ unabhängig von } x_1, x_2} |x_1 - x_2| \end{aligned}$$

für alle $x_1, x_2 \in I_n$, sind sie gleichstetig, und damit gilt nach Heuser [1988] S. 575

$$\begin{aligned} & \frac{\partial^\nu}{\partial x^\nu} \left(\int c_n \frac{1}{n} \sum_{i=1}^n g_n(y) \left(\frac{x-y}{c_n} \right)^{k_1} K \left(\frac{x-y}{c_n} \right) f_{X_{i-1}}(y) dy \right) \\ & = \int \frac{\partial^\nu}{\partial x^\nu} \left(c_n \frac{1}{n} \sum_{i=1}^n g_n(y) \left(\frac{x-y}{c_n} \right)^{k_1} K \left(\frac{x-y}{c_n} \right) f_{X_{i-1}}(y) \right) dy \end{aligned}$$

für $\nu = 0, 1, 2$.

Damit gilt also für $k_2 = 0$

$$\begin{aligned} & \frac{\partial^\nu}{\partial x^\nu} \left(E \left[\frac{1}{nc_n} \sum_{i=1}^n g_n(X_{i-1}) Z_i^{k_2} \left(\frac{x - X_{i-1}}{c_n} \right)^{k_1} K \left(\frac{x - X_{i-1}}{c_n} \right) \right] \right) \\ & = \frac{1}{c_n^2} \int \frac{\partial^\nu}{\partial x^\nu} \left(c_n \frac{1}{n} \sum_{i=1}^n g_n(y) \left(\frac{x-y}{c_n} \right)^{k_1} K \left(\frac{x-y}{c_n} \right) f_{X_{i-1}}(y) \right) dy \\ & = E \left[\frac{\partial^\nu}{\partial x^\nu} \left(\frac{1}{nc_n} \sum_{i=1}^n g_n(X_{i-1}) Z_i^{k_2} \left(\frac{x - X_{i-1}}{c_n} \right)^{k_1} K \left(\frac{x - X_{i-1}}{c_n} \right) \right) \right] \end{aligned}$$

und auch für $k_2 = 1$

$$\begin{aligned} & \frac{\partial^\nu}{\partial x^\nu} \left(E \left[\frac{1}{nc_n} \sum_{i=1}^n g_n(X_{i-1}) Z_i^{k_2} \left(\frac{x - X_{i-1}}{c_n} \right)^{k_1} K \left(\frac{x - X_{i-1}}{c_n} \right) \right] \right) \\ & = 0 \\ & = E \left[\frac{\partial^\nu}{\partial x^\nu} \left(\frac{1}{nc_n} \sum_{i=1}^n g_n(X_{i-1}) Z_i^{k_2} \left(\frac{x - X_{i-1}}{c_n} \right)^{k_1} K \left(\frac{x - X_{i-1}}{c_n} \right) \right) \right]. \end{aligned}$$

Da für alle $\nu = 0, 1, 2$ und $k_1 = 0, 1$

$$\begin{aligned} & \frac{\partial^\nu}{\partial x^\nu} \left(\frac{1}{nc_n} \sum_{i=1}^n g_n(X_{i-1}) Z_i^{k_2} \left(\frac{x - X_{i-1}}{c_n} \right)^{k_1} K \left(\frac{x - X_{i-1}}{c_n} \right) \right) \\ & = \frac{1}{nc_n} \sum_{i=1}^n g_n(X_{i-1}) Z_i^{k_2} \left(\frac{1}{c_n} \right)^\nu \left(k_1 \nu K^{(\nu-1)} \left(\frac{x - X_{i-1}}{c_n} \right) + \left(\frac{x - X_{i-1}}{c_n} \right)^{k_1} K^{(\nu)} \left(\frac{x - X_{i-1}}{c_n} \right) \right) \end{aligned}$$

ist

$$\begin{aligned}
& \sup_{x \in I_n} \left| \frac{\partial^\nu}{\partial x^\nu} \left(\frac{1}{nc_n} \sum_{i=1}^n g_n(X_{i-1}) Z_i^{k_2} \left(\frac{x - X_{i-1}}{c_n} \right)^{k_1} K \left(\frac{x - X_{i-1}}{c_n} \right) \right. \right. \\
& \quad \left. \left. - E \left[\frac{1}{nc_n} \sum_{i=1}^n g_n(X_{i-1}) Z_i^{k_2} \left(\frac{x - X_{i-1}}{c_n} \right)^{k_1} K \left(\frac{x - X_{i-1}}{c_n} \right) \right] \right) \right| \\
&= \frac{1}{c_n^\nu} k_1^\nu \sup_{x \in I_n} \left| \frac{1}{nc_n} \sum_{i=1}^n g_n(X_{i-1}) Z_i^{k_2} K^{(\nu-1)} \left(\frac{x - X_{i-1}}{c_n} \right) \right. \\
& \quad \left. - E \left[\frac{1}{nc_n} \sum_{i=1}^n g_n(X_{i-1}) Z_i^{k_2} K^{(\nu-1)} \left(\frac{x - X_{i-1}}{c_n} \right) \right] \right| \\
& \quad + \frac{1}{c_n^\nu} \sup_{x \in I_n} \left| \frac{1}{nc_n} \sum_{i=1}^n g_n(X_{i-1}) Z_i^{k_2} \left(\frac{x - X_{i-1}}{c_n} \right)^{k_1} K^{(\nu)} \left(\frac{x - X_{i-1}}{c_n} \right) \right. \\
& \quad \left. - E \left[\frac{1}{nc_n} \sum_{i=1}^n g_n(X_{i-1}) Z_i^{k_2} \left(\frac{x - X_{i-1}}{c_n} \right)^{k_1} K^{(\nu)} \left(\frac{x - X_{i-1}}{c_n} \right) \right] \right| \\
\end{aligned} \tag{10.1.2}$$

für $\nu = 0, 1, 2$, $k_1 = 0, 1$ und $k_2 = 0, 1$.

Mit einer Taylorentwicklung von $g_n(X_{i-1})$ in x gilt für ein $\xi_n(i)$ zwischen X_{i-1} und x

$$\begin{aligned}
& \frac{1}{nc_n} \sum_{i=1}^n g_n(X_{i-1}) Z_i^{k_2} \left(\frac{x - X_{i-1}}{c_n} \right)^{k_1} K^{(\nu)} \left(\frac{x - X_{i-1}}{c_n} \right) \\
&= g_n(x) \frac{1}{nc_n} \sum_{i=1}^n Z_i^{k_2} \left(\frac{x - X_{i-1}}{c_n} \right)^{k_1} K^{(\nu)} \left(\frac{x - X_{i-1}}{c_n} \right) \\
& \quad + g'_n(x) \frac{1}{nc_n} \sum_{i=1}^n Z_i^{k_2} c_n \left(\frac{x - X_{i-1}}{c_n} \right)^{k_1+1} K^{(\nu)} \left(\frac{x - X_{i-1}}{c_n} \right) \\
& \quad + \frac{1}{2} g''_n(x) \frac{1}{nc_n} \sum_{i=1}^n Z_i^{k_2} c_n^2 \left(\frac{x - X_{i-1}}{c_n} \right)^{k_1+2} K^{(\nu)} \left(\frac{x - X_{i-1}}{c_n} \right) \\
& \quad + \frac{1}{6} \frac{1}{nc_n} \sum_{i=1}^n Z_i^{k_2} g'''_n(\xi_n(i)) c_n^3 \left(\frac{x - X_{i-1}}{c_n} \right)^{k_1+3} K^{(\nu)} \left(\frac{x - X_{i-1}}{c_n} \right).
\end{aligned}$$

Damit ist für alle $\nu = 0, 1, 2$ und $k_1 = 0, 1$

$$\begin{aligned}
& \sup_{x \in I_n} \left| \frac{1}{nc_n} \sum_{i=1}^n g_n(X_{i-1}) Z_i^{k_2} \left(\frac{x - X_{i-1}}{c_n} \right)^{k_1} K^{(\nu)} \left(\frac{x - X_{i-1}}{c_n} \right) \right. \\
& \quad \left. - E \left[\frac{1}{nc_n} \sum_{i=1}^n g_n(X_{i-1}) Z_i^{k_2} \left(\frac{x - X_{i-1}}{c_n} \right)^{k_1} K^{(\nu)} \left(\frac{x - X_{i-1}}{c_n} \right) \right] \right| \\
& \leq \sup_{x \in I_n} |g_n(x)| \sup_{x \in I_n} \left| \frac{1}{nc_n} \sum_{i=1}^n Z_i^{k_2} \left(\frac{x - X_{i-1}}{c_n} \right)^{k_1} K^{(\nu)} \left(\frac{x - X_{i-1}}{c_n} \right) \right|
\end{aligned}$$

$$\begin{aligned}
& -E \left[\frac{1}{nc_n} \sum_{i=1}^n Z_i^{k_2} \left(\frac{x - X_{i-1}}{c_n} \right)^{k_1} K^{(\nu)} \left(\frac{x - X_{i-1}}{c_n} \right) \right] \Big| \\
& + c_n \sup_{x \in I_n} |g'_n(x)| \sup_{x \in I_n} \left| \frac{1}{nc_n} \sum_{i=1}^n Z_i^{k_2} \left(\frac{x - X_{i-1}}{c_n} \right)^{k_1+1} K^{(\nu)} \left(\frac{x - X_{i-1}}{c_n} \right) \right. \\
& \quad \left. - E \left[\frac{1}{nc_n} \sum_{i=1}^n Z_i^{k_2} \left(\frac{x - X_{i-1}}{c_n} \right)^{k_1+1} K^{(\nu)} \left(\frac{x - X_{i-1}}{c_n} \right) \right] \right| \\
& + \frac{1}{2} c_n^2 \sup_{x \in I_n} |g''_n(x)| \sup_{x \in I_n} \left| \frac{1}{nc_n} \sum_{i=1}^n Z_i^{k_2} \left(\frac{x - X_{i-1}}{c_n} \right)^{k_1+2} K^{(\nu)} \left(\frac{x - X_{i-1}}{c_n} \right) \right. \\
& \quad \left. - E \left[\frac{1}{nc_n} \sum_{i=1}^n Z_i^{k_2} \left(\frac{x - X_{i-1}}{c_n} \right)^{k_1+2} K^{(\nu)} \left(\frac{x - X_{i-1}}{c_n} \right) \right] \right| \\
& + \frac{1}{6} c_n^3 \sup_{x \in I_n} \left| \frac{1}{nc_n} \sum_{i=1}^n g_n'''(\xi_n(i)) Z_i^{k_2} \left(\frac{x - X_{i-1}}{c_n} \right)^{k_1+3} K^{(\nu)} \left(\frac{x - X_{i-1}}{c_n} \right) \right| \\
& + \frac{1}{6} c_n^3 \sup_{x \in I_n} \left| E \left[\frac{1}{nc_n} \sum_{i=1}^n g_n'''(\xi_n(i)) Z_i^{k_2} \left(\frac{x - X_{i-1}}{c_n} \right)^{k_1+3} K^{(\nu)} \left(\frac{x - X_{i-1}}{c_n} \right) \right] \right| \\
& = O_P \left(c_n^{-\frac{1}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} \right) + O(c_n^2), \tag{10.1.3}
\end{aligned}$$

nach Gleichung (10.1.1), da $\sup_{x \in I_n} |g_n^{(\mu)}(x)| = O(1)$ für $\mu = 0, 1, 2$ nach Voraussetzung, und außerdem

$$\begin{aligned}
& c_n \sup_{x \in I_n} \left| \frac{1}{nc_n} \sum_{i=1}^n g_n'''(\xi_n(i)) Z_i^{k_2} \left(\frac{x - X_{i-1}}{c_n} \right)^{k_1+3} K^{(\nu)} \left(\frac{x - X_{i-1}}{c_n} \right) \right| \\
& \leq \sup_{x \in [a_n - 2Cc_n, b_n + 2Cc_n]} |g_n'''(x)| \sup_{u \in [-C, C]} |u^{k_1+3} K^{(\nu)}(u)| \frac{1}{n} \sum_{i=1}^n |Z_i^{K_2}| \\
& = O(1),
\end{aligned}$$

genau wie $c_n \sup_{x \in I_n} \left| E \left[\frac{1}{nc_n} \sum_{i=1}^n g_n'''(\xi_n(i)) Z_i^{k_2} \left(\frac{x - X_{i-1}}{c_n} \right)^{k_1+3} K^{(\nu)} \left(\frac{x - X_{i-1}}{c_n} \right) \right] \right| = O(1)$, da auch $\sup_{x \in [a_n - Cc_n, b_n + Cc_n]} |g_n'''(x)| = O(1)$ nach Voraussetzung.

Aus Gleichung (10.1.2) und (10.1.3) folgt das gewünschte

$$\begin{aligned}
& \sup_{x \in I_n} \left| \frac{\partial^\nu}{\partial x^\nu} \left(\frac{1}{nc_n} \sum_{i=1}^n g_n(X_{i-1}) Z_i^{k_2} \left(\frac{x - X_{i-1}}{c_n} \right)^{k_1} K \left(\frac{x - X_{i-1}}{c_n} \right) \right. \right. \\
& \quad \left. \left. - E \left[\frac{1}{nc_n} \sum_{i=1}^n g_n(X_{i-1}) Z_i^{k_2} \left(\frac{x - X_{i-1}}{c_n} \right)^{k_1} K \left(\frac{x - X_{i-1}}{c_n} \right) \right] \right) \right| \\
& = O_P \left(c_n^{-\frac{1}{2}-\nu} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} + c_n^{2-\nu} \right).
\end{aligned}$$

□

Lemma 10.2 Es seien die Voraussetzungen von Lemma 3.4 (S. 35) gegeben.

Anmerkung: Damit müssen auch die Voraussetzungen von Lemma 3.3 (S. 21) erfüllt sein.

Außerdem sei der Prozess $(X_i)_{i \in \mathbb{Z}}$ α -mischend mit $\alpha(n) = O(\bar{\alpha}^{-n})$ für ein $\bar{\alpha}$ mit $1 < \bar{\alpha} < \infty$.

Zusätzlich seien die Beobachtungen X_0, \dots, X_n identisch verteilt und es sei $q_n^f := \frac{1}{\inf_{x \in I_n} f_{X_{i-1}}(x)}$ für alle i .

Anmerkung: Durch die Voraussetzungen von Lemma 3.4 gilt für die Dichte der Beobachtungen auch $\sup_{x \in \mathbb{R}} f_{X_0}(x) < \infty$, und damit auch $(q_n^f)^{-1} = O(1)$.

Die Gewichtsfunktion w_n sei zweimal differenzierbar und es gelte $\sup_{x \in \mathbb{R}} |w_n^{(\mu)}(x)| < \infty$ für $\mu = 1, 2$.

Für die Bandbreite c_n gelte $nc_n^5 = O(1)$ und $(nc_n^3)^{-1}(q_n^f)^4 \log(n)^2 = o(1)$ und es gebe ein δ , so dass $c_n^{-\frac{3}{2}-\delta} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} (q_n^f)^3 = o(1)$ und $(b_n - a_n)(q_n^f)^2 \log(n)^{3+\frac{1}{1+\delta}} c_n^{-\frac{1}{2}} n^{-\frac{1}{2}+\frac{1}{2} \cdot \frac{1}{1+\delta}} = o(1)$.

Anmerkung: Wenn $c_n \sim n^{-\frac{1}{4}} \log(n)^{r_1}$ für ein $|r_1| < \infty$, $(b_n - a_n) = O(\log(n)^{r_2})$ für ein $|r_2| < \infty$ und $q_n^f = O(\log(n)^{r_3})$ für ein $|r_3| < \infty$, gilt dies für alle $\delta \in (\frac{1}{3}, \frac{1}{2})$.

Dann gilt für unabhängige und identisch verteilte Zufallsvariablen (Z_i) mit $E[Z_i] = 0$ für alle i , die jeweils auch von X_{i-1}, X_{i-2}, \dots unabhängig sind, und für die $X_i = Z_i$ die Bedingungen an X_i für $k = 1$ in Lemma 3.3 erfüllt,

$$\sup_{s \in [0,1]} \left| \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \frac{1}{n} \sum_{i=1}^n Z_i \sum_{j=1}^{\lfloor ns \rfloor} \bar{w}_{j \lfloor ns \rfloor} \left(\frac{\frac{1}{c_n} K\left(\frac{X_{j-1}-X_{i-1}}{c_n}\right) - \frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{X_{j-1}-X_{k-1}}{c_n}\right)}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{X_{j-1}-X_{k-1}}{c_n}\right)} \right) \right| = o_p\left(\frac{1}{\sqrt{n}}\right).$$

Anmerkung: Es reicht, dass $X_i = Z_i$ die Bedingungen für $k = 1$ erfüllt, da auch nur die Aussage für $k = 1$ benötigt wird. Da angenommen wird, dass die Zufallsvariablen (Z_i) i.i.d. sind, reduzieren sich die Voraussetzungen an $X_i = Z_i$ in Lemma 3.3 auf $E[|Z_i|^b] < \infty$ für alle i . (Damit gilt auch $E[Z_i^2] < \infty$.)

Beweis Mit

$$\hat{d}_n(x) := w_n(x) \frac{1}{n} \sum_{i=1}^n Z_i \left(\frac{\frac{1}{c_n} K\left(\frac{x-X_{i-1}}{c_n}\right) - \frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{x-X_{k-1}}{c_n}\right)}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{x-X_{k-1}}{c_n}\right)} \right)$$

ist zu zeigen, dass

$$\sup_{s \in [0,1]} \left| \frac{1}{n} \sum_{j=1}^{\lfloor ns \rfloor} \hat{d}_n(X_{j-1}) \right| = o_p\left(\frac{1}{\sqrt{n}}\right). \quad (10.2.1)$$

Um diese Konvergenz zu beweisen muss zunächst der Ausdruck $\hat{d}_n(X_{j-1})$ bezüglich X_{j-1} zentriert werden. Dafür wird gezeigt, dass

$$\left| \int \hat{d}_n(x) f_{X_0}(x) dx \right| = o_p\left(\frac{1}{\sqrt{n}}\right). \quad (10.2.2)$$

$$\begin{aligned} \left(\text{Es gilt } \sup_{s \in [0,1]} \left| \frac{1}{n} \sum_{j=1}^{\lfloor ns \rfloor} \int \hat{d}_n(x) f_{X_{j-1}}(x) dx \right| = \sup_{s \in [0,1]} \left| \frac{\lfloor ns \rfloor}{n} \int \hat{d}_n(x) f_{X_0}(x) dx \right| \right. \\ \left. = \left| \int \hat{d}_n(x) f_{X_0}(x) dx \right| \right) \end{aligned}$$

Als erstes wird gezeigt, dass der zufällige Nenner ersetzt werden kann. Es gilt

$$\begin{aligned} & \left| \int \hat{d}_n(x) f_{X_0}(x) dx \right| \\ & \leq \left| \frac{1}{n} \sum_{i=1}^n Z_i \int w_n(x) \left(\frac{\frac{1}{c_n} K\left(\frac{x-X_{i-1}}{c_n}\right) - \frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{x-X_{k-1}}{c_n}\right)}{f_{X_0}(x)} \right) f_{X_0}(x) dx \right| \\ & \quad + \left| \frac{1}{n} \sum_{i=1}^n Z_i \int w_n(x) \left(\frac{f_{X_0}(x) - \frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{x-X_{k-1}}{c_n}\right)}{f_{X_0}(x)} \right) \right. \\ & \quad \quad \left. \cdot \left(\frac{\frac{1}{c_n} K\left(\frac{x-X_{i-1}}{c_n}\right) - \frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{x-X_{k-1}}{c_n}\right)}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{x-X_{k-1}}{c_n}\right)} \right) f_{X_0}(x) dx \right|. \\ & = \left| \frac{1}{n} \sum_{i=1}^n Z_i \int w_n(x) \left(\frac{1}{c_n} K\left(\frac{x-X_{i-1}}{c_n}\right) - \frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{x-X_{k-1}}{c_n}\right) \right) dx \right| \\ & \quad + \left| \frac{1}{n} \sum_{i=1}^n Z_i \int w_n(x) \left(\frac{f_{X_0}(x) - \frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{x-X_{k-1}}{c_n}\right)}{f_{X_0}(x)} \right) \right. \\ & \quad \quad \left. \cdot \left(\frac{\frac{1}{c_n} K\left(\frac{x-X_{i-1}}{c_n}\right) - \frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{x-X_{k-1}}{c_n}\right)}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{x-X_{k-1}}{c_n}\right)} \right) f_{X_0}(x) dx \right|. \end{aligned} \tag{10.2.3}$$

Für den zweiten Summanden gilt

$$\begin{aligned} & \left| \frac{1}{n} \sum_{i=1}^n Z_i \int w_n(x) \left(\frac{f_{X_0}(x) - \frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{x-X_{k-1}}{c_n}\right)}{f_{X_0}(x)} \right) \right. \\ & \quad \quad \left. \cdot \left(\frac{\frac{1}{c_n} K\left(\frac{x-X_{i-1}}{c_n}\right) - \frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{x-X_{k-1}}{c_n}\right)}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{x-X_{k-1}}{c_n}\right)} \right) f_{X_0}(x) dx \right| \\ & = \left| \int w_n(x) \left(\frac{f_{X_0}(x) - \frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{x-X_{k-1}}{c_n}\right)}{f_{X_0}(x)} \right) \right. \\ & \quad \quad \left. \cdot \frac{1}{n} \sum_{i=1}^n Z_i \left(\frac{\frac{1}{c_n} K\left(\frac{x-X_{i-1}}{c_n}\right)}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{x-X_{k-1}}{c_n}\right)} - 1 \right) f_{X_0}(x) dx \right| \end{aligned}$$

$$\leq \frac{1}{\inf_{x \in I_n} f_{X_0}(x)} \sup_{x \in I_n} \left| f_{X_0}(x) - \frac{1}{nc_n} \sum_{k=1}^n K \left(\frac{x - X_{k-1}}{c_n} \right) \right| \left(\sup_{x \in I_n} \left| \frac{1}{n} \sum_{i=1}^n \frac{\frac{1}{c_n} K \left(\frac{x - X_{i-1}}{c_n} \right) Z_i}{\frac{1}{nc_n} \sum_{k=1}^n K \left(\frac{x - X_{k-1}}{c_n} \right)} \right| + \left| \frac{1}{n} \sum_{i=1}^n Z_i \right| \right). \quad (10.2.4)$$

Da Z_1, \dots, Z_n unabhängige und identisch verteilte Zufallsvariablen sind und $E[Z_1] = 0$, sowie $E[Z_1^2] < \infty$, gilt $|\frac{1}{n} \sum_{i=1}^n Z_i| = O_P(n^{-\frac{1}{2}})$.

Der Term

$$\sup_{x \in I_n} \left| \frac{1}{n} \sum_{i=1}^n \frac{\frac{1}{c_n} K \left(\frac{x - X_{i-1}}{c_n} \right) Z_i}{\frac{1}{nc_n} \sum_{k=1}^n K \left(\frac{x - X_{k-1}}{c_n} \right)} \right|$$

entspricht $\sup_{x \in I_n} |\hat{m}(x) - \tilde{m}(x)|$, wobei \hat{m} den Nadaraya-Watson-Schätzer für die spezielle Autoegression $X_i = \tilde{m}(X_{i-1}) + Z_i$ mit $\tilde{m} \equiv 0$ bezeichnet. Damit ist

$$\sup_{x \in I_n} \left| \frac{1}{n} \sum_{i=1}^n \frac{\frac{1}{c_n} K \left(\frac{x - X_{i-1}}{c_n} \right) Z_i}{\frac{1}{nc_n} \sum_{k=1}^n K \left(\frac{x - X_{k-1}}{c_n} \right)} \right| = O_P \left(\left(c_n^{-\frac{1}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} + c_n^2 \right) q_n^f \right)$$

nach Lemma 3.4 (mit $q_n = 1$ und $q_n^\sigma = 1$), da $X_i = \tilde{m}(X_{i-1}) + Z_i$ die Bedingungen aus Lemma 3.3 nach Voraussetzung erfüllt.

Ebenfalls nach Lemma 3.4 gilt

$$\sup_{x \in I_n} \left| f_{X_0}(x) - \frac{1}{nc_n} \sum_{k=1}^n K \left(\frac{x - X_{k-1}}{c_n} \right) \right| = O_P \left(c_n^{-\frac{1}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} + c_n^2 \right).$$

Daher folgt mit (10.2.3) und (10.2.4)

$$\begin{aligned} & \left| \int \hat{d}_n(x) f_{X_0}(x) dx \right| \\ & \leq \left| \frac{1}{n} \sum_{i=1}^n Z_i \int w_n(x) \left(\frac{1}{c_n} K \left(\frac{x - X_{i-1}}{c_n} \right) - \frac{1}{nc_n} \sum_{k=1}^n K \left(\frac{x - X_{k-1}}{c_n} \right) \right) dx \right| \\ & \quad + O_P \left(q_n^f \left(c_n^{-\frac{1}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} + c_n^2 \right) \left(\left(c_n^{-\frac{1}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} + c_n^2 \right) q_n^f + n^{-\frac{1}{2}} \right) \right) \\ & = \left| \frac{1}{n} \sum_{i=1}^n Z_i \int w_n(x) \left(\frac{1}{c_n} K \left(\frac{x - X_{i-1}}{c_n} \right) - \frac{1}{nc_n} \sum_{k=1}^n K \left(\frac{x - X_{k-1}}{c_n} \right) \right) dx \right| \\ & \quad + o_p \left(\frac{1}{\sqrt{n}} \right) \end{aligned}$$

da mit $nc_n^5 = O(1)$ gilt $\left(q_n^f \left(c_n^{-\frac{1}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} + c_n^2 \right) \right)^2 = O \left(\left(q_n^f c_n^{-\frac{1}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} \right)^2 \right) = O \left(n^{-\frac{1}{2}} (q_n^f)^2 \log(n) n^{-\frac{1}{2}} c_n^{-1} \right) = o(n^{-\frac{1}{2}})$, da $c_n \cdot (nc_n^3)^{-1} (q_n^f)^4 \log(n)^2 = o(1) \cdot o(1)$ nach Voraussetzung.

Mit den selben Argumenten folgt weiter

$$\begin{aligned}
& \left| \frac{1}{n} \sum_{i=1}^n Z_i \int w_n(x) \left(\frac{1}{c_n} K \left(\frac{x - X_{i-1}}{c_n} \right) - \frac{1}{nc_n} \sum_{k=1}^n K \left(\frac{x - X_{k-1}}{c_n} \right) \right) dx \right| \\
& \leq \left| \frac{1}{n} \sum_{i=1}^n Z_i \int w_n(x) \left(\frac{1}{c_n} K \left(\frac{x - X_{i-1}}{c_n} \right) - f_{X_0}(x) \right) dx \right| \\
& \quad + \left| \frac{1}{n} \sum_{i=1}^n Z_i \int w_n(x) \left(\frac{f_{X_0}(x) - \frac{1}{nc_n} \sum_{k=1}^n K \left(\frac{x - X_{k-1}}{c_n} \right)}{f_{X_0}(x)} \right) f_{X_0}(x) dx \right| \\
& = \left| \frac{1}{n} \sum_{i=1}^n Z_i \int w_n(x) \left(\frac{1}{c_n} K \left(\frac{x - X_{i-1}}{c_n} \right) - f_{X_0}(x) \right) dx \right| \\
& \quad + o_p \left(\frac{1}{\sqrt{n}} \right).
\end{aligned}$$

Sei nun

$$\begin{aligned}
\bar{h}(x) & := \int w_n(y) \left(\frac{1}{c_n} K \left(\frac{y - x}{c_n} \right) - f_{X_0}(y) \right) dy, \\
& \left(= \int w_n(uc_n + x) K(u) du - \int w_n(y) f_{X_0}(y) dy \right)
\end{aligned}$$

dann ist (10.2.2) gleichbedeutend mit

$$\left| \frac{1}{n} \sum_{i=1}^n Z_i \bar{h}(X_{i-1}) \right| = o_p \left(\frac{1}{\sqrt{n}} \right). \quad (10.2.5)$$

Um zu zeigen, dass (10.2.5) gilt, werden der Erwartungswert und die Varianz berechnet, um dann die Tchebychev-Ungleichung anwenden zu können.

Da die Funktion \bar{h} beschränkt ist, ist der Erwartungswert von $\frac{1}{n} \sum_{i=1}^n Z_i \bar{h}(X_{i-1})$ Null, denn $E[Z_i] = 0$ und X_{i-1} und Z_i sind stochastisch unabhängig für alle i .

Da Z_i nach Voraussetzung unabhängig von Z_j und X_j für alle $j \leq i - 1$, gilt für die Varianz

$$\begin{aligned}
\text{Var} \left(\frac{1}{n} \sum_{i=1}^n Z_i \bar{h}(X_{i-1}) \right) & = E \left[\left(\frac{1}{n} \sum_{i=1}^n Z_i \bar{h}(X_{i-1}) \right)^2 \right] \\
& = \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n E [Z_i Z_j \bar{h}(X_{i-1}) \bar{h}(X_{j-1})] \\
& = \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^{i-1} E [E [Z_i Z_j \bar{h}(X_{i-1}) \bar{h}(X_{j-1}) | X_k, Z_k, k \leq i-1]] \\
& \quad + \frac{1}{n^2} \sum_{i=1}^n E [Z_i^2 (\bar{h}(X_{i-1}))^2] \\
& \quad + \frac{1}{n^2} \sum_{i=1}^n \sum_{j=i+1}^n E [E [Z_i Z_j \bar{h}(X_{i-1}) \bar{h}(X_{j-1}) | X_k, Z_k, k \leq j-1]] \\
& = \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^{i-1} E \left[Z_j \bar{h}(X_{i-1}) \bar{h}(X_{j-1}) \underbrace{E [Z_i]}_{=0} \right]
\end{aligned}$$

$$\begin{aligned}
& + \frac{1}{n^2} \sum_{i=1}^n E \left[Z_i^2 (\bar{h}(X_{i-1}))^2 \right] \\
& + \frac{1}{n^2} \sum_{i=1}^n \sum_{j=i+1}^n E \left[Z_i \bar{h}(X_{i-1}) \bar{h}(X_{j-1}) \underbrace{E[Z_j]}_{=0} \right] \\
& = \frac{1}{n^2} \sum_{i=1}^n E \left[Z_i^2 (\bar{h}(X_{i-1}))^2 \right] \\
& = \frac{1}{n} \left(\frac{1}{n} \sum_{i=1}^n E [Z_i^2] E [(\bar{h}(X_{i-1}))^2] \right) \\
& = \frac{1}{n} E [Z_1^2] E [(\bar{h}(X_0))^2]. \tag{10.2.6}
\end{aligned}$$

Mit einer Taylorentwicklung gilt für ein $\xi_n(u, x)$ zwischen $uc_n + x$ und x

$$\begin{aligned}
E[\bar{h}(X_0)] & = \int \int w_n(uc_n + x) K(u) du f_{X_0}(x) dx - \int w_n(y) f_{X_0}(y) dy \\
& = \int w_n(x) \underbrace{\int K(u) du}_{=1} f_{X_0}(x) dx \\
& \quad + \int \int w'_n(\xi_n(u, x)) uc_n K(u) du f_{X_0}(x) dx \\
& \quad - \int w_n(y) f_{X_0}(y) dy \\
& = c_n \int \int w'_n(\xi_n(u, x)) u K(u) du f_{X_0}(x) dx \\
& = o(1),
\end{aligned}$$

da

$$\left| c_n \int \int w'_n(\xi_n(u, x)) u K(u) du f_{X_0}(x) dx \right| \leq c_n \underbrace{\sup_{x \in \mathbb{R}} |w'_n(x)|}_{=O(1) \text{ für } n \rightarrow \infty} \underbrace{\int |u| K(u) du}_{\leq 2C \cdot C\bar{K} < \infty} \underbrace{\int f_{X_0}(x) dx}_{=1}.$$

Weiter gilt

$$\begin{aligned}
& E [(\bar{h}(X_0))^2] \\
& = \left(\int \left(\int w_n(uc_n + x) K(u) du \right)^2 f_{X_0}(x) dx - \int \int w_n(uc_n + x) K(u) du f_{X_0}(x) dx \int w_n(y) f_{X_0}(y) dy \right) \\
& \quad + \underbrace{\left(\left(\int w_n(y) f_{X_0}(y) dy \right)^2 - \int \int w_n(uc_n + x) K(u) du f_{X_0}(x) dx \int w_n(y) f_{X_0}(y) dy \right)}_{= \int w_n(y) f_{X_0}(y) dy \cdot (-E[\bar{h}(X_0)])} \\
& = \left(\int \left(\int w_n(uc_n + x) K(u) du \right)^2 f_{X_0}(x) dx - \int \int w_n(uc_n + x) K(u) du f_{X_0}(x) dx \int w_n(y) f_{X_0}(y) dy \right) \\
& \quad + o(1), \quad \text{da } \int w_n(y) f_{X_0}(y) dy \leq 1
\end{aligned}$$

und weiter

$$\begin{aligned}
& \left| \left(\int \left(\int w_n(uc_n + x)K(u) du \right)^2 f_{X_0}(x) dx - \int \int w_n(uc_n + x)K(u) du f_{X_0}(x) dx \int w_n(y)f_{X_0}(y) dy \right) \right| \\
&= \left| \int \int w_n(uc_n + x)K(u) du \left(\int w_n(uc_n + x)K(u) du - \int w_n(y)f_X(y) dy \right) f_{X_0}(x) dx \right| \\
&\leq \sup_{x \in \mathbb{R}} \underbrace{\left| \int w_n(uc_n + x)K(u) du \right|}_{=1} \left| \int w_n(uc_n + x)K(u) du - \int w_n(y)f_X(y) dy \right| f_{X_0}(x) dx \\
&\leq \int \underbrace{\left| \int w_n(uc_n + x)K(u) du - \int w_n(y)f_X(y) dy \right|}_{\leq 1} \underbrace{f_{X_0}(x)}_{< \infty} dx \\
&= o(1),
\end{aligned}$$

da $\int w_n(uc_n + x)K(u) du = 1$ für $x \in [a_n + Cc_n, b_n - Cc_n]$ und $\int w_n(y)f_{X_0}(y) dy \xrightarrow{n \rightarrow \infty} 1$, also $\left| \int w_n(uc_n + x)K(u) du - \int w_n(y)f_{X_0}(y) dy \right| \xrightarrow{n \rightarrow \infty} 0$ für $x \in [a_n + Cc_n, b_n - Cc_n]$ und $f_{X_0}(x) \rightarrow 0$ für $x \rightarrow \pm\infty$.

Also gilt

$$E [(\bar{h}(X_0))^2] = o(1),$$

und daher mit (10.2.6)

$$\text{Var} \left(\frac{1}{n} \sum_{i=1}^n Z_i \bar{h}(X_{i-1}) \right) = o \left(\frac{1}{n} \right),$$

denn $E [Z_1^2] < \infty$.

Daraus folgt mit der Tchebychev-Ungleichung

$$\left| \frac{1}{n} \sum_{i=1}^n Z_i \bar{h}(X_{i-1}) \right| = o_p \left(\frac{1}{\sqrt{n}} \right)$$

und damit gilt auch die Gleichung (10.2.5), also auch (10.2.2):

$$\left| \int \hat{d}_n(x) f_{X_0}(x) dx \right| = o_p \left(\frac{1}{\sqrt{n}} \right).$$

Der Term $\hat{d}_n(X_{j-1})$ kann also bezüglich X_{j-1} zentriert werden.

Sei

$$\hat{d}_{En}(x) = \hat{d}_n(x) - \int \hat{d}_n(y) f_{X_0}(y) dy.$$

Dann gilt (10.2.1), wenn

$$\sup_{s \in [0,1]} \left| \frac{1}{n} \sum_{j=1}^{\lfloor ns \rfloor} \hat{d}_{En}(X_{j-1}) \right| = o_p \left(\frac{1}{\sqrt{n}} \right). \quad (10.2.7)$$

und es bleibt zu zeigen, dass (10.2.7) gilt.

Die Idee hierfür ist, zu zeigen, dass \hat{d}_{En} zu einer Funktionenklasse gehört, und dann das Supremum über diese Funktionenklasse zu betrachten.

Für die Funktion \hat{d}_{En} gilt

$$\begin{aligned} \sup_{x \in \mathbb{R}} |\hat{d}_{En}(x)| &\leq \sup_{x \in I_n} |\hat{d}_{En}(x)| \quad \text{da } \hat{d}_n(x) = 0 \text{ für } x \notin I_n \\ &\leq 2 \sup_{x \in I_n} |\hat{d}_n(x)| \end{aligned}$$

und für die Ableitungen

$$\sup_{x \in \mathbb{R}} |\hat{d}'_{En}(x)| = \sup_{x \in I_n} |\hat{d}'_n(x)| \quad \text{und} \quad \sup_{x \in \mathbb{R}} |\hat{d}''_{En}(x)| = \sup_{x \in I_n} |\hat{d}''_n(x)|.$$

Für die Funktion \hat{d}_n gilt

$$\hat{d}_n(x) = w_n(x) \left(\hat{m}(x) - \tilde{m}(x) \right) - w_n(x) \frac{1}{n} \sum_{i=1}^n Z_i$$

wobei \hat{m} , wie oben, den Nadaraya-Watson-Schätzer für die spezielle Autoregression $X_i = \tilde{m}(X_{i-1}) + Z_i$ mit $\tilde{m} \equiv 0$ bezeichnet.

Da $X_i = \tilde{m}(X_{i-1}) + Z_i$ die Bedingungen aus Lemma 3.3 (S. 21) nach Voraussetzung erfüllt, gilt mit Lemma 3.4 (S. 35)

$$\begin{aligned} \sup_{x \in I_n} |\hat{d}_n(x)| &\leq \sup_{x \in I_n} |\hat{m}(x) - \tilde{m}(x)| + \left| \frac{1}{n} \sum_{i=1}^n Z_i \right| \\ &\stackrel{\text{Lemma 3.4}}{=} O_P \left(\left(c_n^{-\frac{1}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} + c_n^2 \right) q_n^f \right) + \left| \frac{1}{n} \sum_{i=1}^n Z_i \right| \\ &= O_P \left(\left(c_n^{-\frac{1}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} + c_n^2 \right) q_n^f + n^{-\frac{1}{2}} \right) \\ &= O_P \left(c_n^{-\frac{1}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} q_n^f \right). \end{aligned}$$

Im Beweis von Lemma 3.4 sind in (3.4.7) (S. 42) die genauen Konvergenzraten für die Ableitungen angegeben:

$$\sup_{x \in I_n} \left| \frac{\partial^\nu}{\partial x^\nu} \left(\hat{m}(x) - \tilde{m}(x) \right) \right| = O_P \left(c_n^{-\frac{1}{2}-\nu} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} (q_n^f)^{\nu+1} \right).$$

Damit gilt für die Ableitung von \hat{d}

$$\begin{aligned} \sup_{x \in I_n} |\hat{d}'_n(x)| &\leq \sup_{x \in I_n} \left| \frac{\partial}{\partial x} \left(\hat{m}(x) - \tilde{m}(x) \right) \right| + \sup_{x \in I_n} |w'_n(x)| \left(\sup_{x \in I_n} |\hat{m}(x) - \tilde{m}(x)| + \left| \frac{1}{n} \sum_{i=1}^n Z_i \right| \right) \\ &= o_P(1) \end{aligned}$$

da $\sup_{x \in I_n} |w'_n(x)| = O(1)$ und $c_n^{-\frac{3}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} (q_n^f)^2 = \left((nc_n^3)^{-1} \log(n) (q_n^f)^4 \right)^{\frac{1}{2}} = o(1)$ nach Voraussetzung.

Außerdem ist für alle δ

$$\sup_{x, y \in J_n, x \neq y} \frac{|\hat{d}'_n(x) - \hat{d}'_n(y)|}{|x - y|^\delta} \leq \sup_{x, y \in J_n, |x - y| > c_n} \frac{|\hat{d}'_n(x) - \hat{d}'_n(y)|}{|x - y|^\delta} + \sup_{x, y \in J_n, 0 < |x - y| \leq c_n} \frac{|\hat{d}'_n(x) - \hat{d}'_n(y)|}{|x - y|^\delta}$$

$$\begin{aligned}
&\leq 2 \sup_{x \in J_n} |\hat{d}'_n(x)| c_n^{-\delta} + \sup_{x \in J_n} |\hat{d}''_n(x)| c_n^{1-\delta} \\
&= O_P \left(c_n^{-\frac{3}{2}-\delta} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} (q_n^f)^3 \right)
\end{aligned}$$

Damit gilt für alle δ mit $c_n^{-\frac{3}{2}-\delta} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} (q_n^f)^3 = o(1)$

$$P \left(\hat{d}_{En} \notin C_1^{1+\delta}(I_n)^n \right) = o(1)$$

für die Funktionenklasse

$$C_1^{1+\delta}(I_n)^n := \left\{ d \in C_1^{1+\delta}(I_n), \sup_{x \in I_n} |d(x)| \leq z_n, \int d(y) f_{X_0}(y) dy = 0 \right\},$$

wobei $z_n := c_n^{-\frac{1}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} q_n^f \log(n)^{\frac{1}{2}}$ und

$$C_1^{1+\delta}(I_n) := \left\{ d : I_n \rightarrow \mathbb{R} \left| \max \left(\sup_{x \in I_n} |d(x)|, \sup_{x \in I_n} |d'_s(x)| \right) + \sup_{x, y \in I_n, x \neq y} \frac{|d'(x) - d'(y)|}{|x - y|^\delta} \leq 1 \right. \right\}$$

sei.

Somit gilt (10.2.7), wenn

$$\sup_{s \in [0,1]} \sup_{d \in C_1^{1+\delta}(I_n)^n} \left| \frac{1}{n} \sum_{j=1}^{\lfloor ns \rfloor} d(X_{j-1}) \right| = o_p \left(\frac{1}{\sqrt{n}} \right). \quad (10.2.8)$$

Um die Funktionenklasse $C_1^{1+\delta}(I_n)$ zu überdecken, braucht man nach Theorem 2.7.1 in van der Vaart and Wellner [1996]

$$M_{n2} := N \left(\frac{\bar{\epsilon}_n}{2}, C_1^{1+\delta}(I_n), \|\cdot\|_{J_n} \right) \leq \exp \left(2^{\frac{1}{1+\delta}} \bar{K} (2 + b_n - a_n) \bar{\epsilon}_n^{-\frac{1}{1+\delta}} \right)$$

Kugeln vom Radius $\frac{\bar{\epsilon}_n}{2} > 0$ bzgl. der Pseudonorm $\|\cdot\|_{I_n}$, wobei $\|d\|_{I_n} := \sup_{x \in I_n} |d(x)|$. Die Konstante \bar{K} hängt nur von δ ab. Für die Klassen $C_1^{1+\delta}(I_n)^n$ gelten die selben Abschätzungen für die Überdeckungszahlen.

Nach Pollard [1990] (S. 10) kann die Anzahl von Funktionen aus $C_1^{1+\delta}(I_n)^n$, die zur Überdeckung von $C_1^{1+\delta}(I_n)^n$ mit Kugeln vom Radius $\bar{\epsilon}_n > 0$ nötig sind, ebenfalls durch M_{n2} abgeschätzt werden.

Mit $d_1, \dots, d_{M_{n2}} \in C_1^{1+\delta}(I_n)^n$ seien die Funktionen bezeichnet, die die Mittelpunkte der Kugeln zur Überdeckung von $C_1^{1+\delta}(J_n)^n$ bilden.

Seien $0 = s_1 < \dots < s_{M_{n1}} = 1$ die Punkte zur Überdeckung von $[0, 1]$ mit Abschnitten der Größe $\bar{\epsilon}_n^*$, so dass $M_{n1} \leq \frac{1}{\bar{\epsilon}_n^*}$.

Es sei $\bar{\epsilon}_n = n^{-\frac{1}{2}} \log(n)^{-1}$ und $\bar{\epsilon}_n^* = \bar{\epsilon}_n \cdot z_n^{-1}$. Dann ist

$$\sup_{s \in [0,1]} \sup_{d \in C_1^{1+\delta}(I_n)^n} \left| \frac{1}{n} \sum_{j=1}^{\lfloor ns \rfloor} d(X_{j-1}) \right|$$

$$\begin{aligned}
&= \max_{1 \leq h \leq M_{n1}} \max_{1 \leq k \leq M_{n2}} \sup_{s \in [0,1] \text{ mit } |s-s_h| \leq \bar{\epsilon}_n^*} \sup_{d \in C_1^{1+\delta}(I_n)^n \text{ mit } \|d-d_k\|_{I_n} \leq \bar{\epsilon}_n} \left| \frac{1}{n} \sum_{j=1}^{\lfloor ns \rfloor} d(X_{j-1}) \right| \\
&\leq \max_h \max_k \left| \frac{1}{n} \sum_{j=1}^{\lfloor ns_h \rfloor} d_k(X_{j-1}) \right| \\
&\quad + \max_h \sup_{s \in [0,1] \text{ mit } |s-s_h| \leq \bar{\epsilon}_n^*} \sup_{d \in C_1^{1+\delta}(I_n)^n} \left| \frac{1}{n} \sum_{j=1}^n d(X_{j-1}) \left(I \left\{ \frac{j}{n} \leq s \right\} - I \left\{ \frac{j}{n} \leq s_h \right\} \right) \right| \\
&\quad + \max_h \max_k \sup_{d \in C_1^{1+\delta}(J_n)^n \text{ mit } \|d-d_k\|_{I_n} \leq \bar{\epsilon}_n} \left| \frac{1}{n} \sum_{j=1}^{\lfloor ns_h \rfloor} (d(X_{j-1}) - d_k(X_{j-1})) \right| \\
&\leq \max_h \max_k \left| \frac{1}{n} \sum_{j=1}^{\lfloor ns_h \rfloor} d_k(X_{j-1}) \right| \\
&\quad + \max_h \sup_{s \in [0,1] \text{ mit } |s-s_h| \leq \bar{\epsilon}_n^*} \sup_{d \in C_1^{1+\delta}(I_n)^n} \sup_{x \in I_n} |d(x)| \frac{1}{n} \sum_{j=1}^n \left| I \left\{ \frac{j}{n} \leq s \right\} - I \left\{ \frac{j}{n} \leq s_h \right\} \right| \\
&\quad + \max_h \max_k \sup_{d \in C_1^{1+\delta}(J_n)^n \text{ mit } \|d-d_k\|_{I_n} \leq \bar{\epsilon}_n} \sup_{x \in I_n} |d(x) - d_k(x)| \frac{\lfloor ns_h \rfloor}{n} \\
&\leq \max_h \max_k \left| \frac{1}{n} \sum_{j=1}^{\lfloor ns_h \rfloor} d_k(X_{j-1}) \right| \\
&\quad + z_n \cdot \max_h \sup_{s \in [0,1] \text{ mit } |s-s_h| \leq \bar{\epsilon}_n^*} \left(\underbrace{|s-s_h|}_{\leq \bar{\epsilon}_n z_n^{-1}} + \frac{1}{n} \right) \quad \text{mit (4.5.6) (S. 65) im Beweis zu Satz 4.4} \\
&\quad + \max_k \sup_{d \in C_1^{1+\delta}(J_n)^n \text{ mit } \|d-d_k\|_{I_n} \leq \bar{\epsilon}_n} \underbrace{\sup_{x \in I_n} |d(x) - d_k(x)|}_{\leq \bar{\epsilon}_n} \\
&\leq \max_{1 \leq h \leq M_{n1}} \max_{1 \leq k \leq M_{n2}} \left| \frac{1}{n} \sum_{j=1}^{\lfloor ns_h \rfloor} d_k(X_{j-1}) \right| \\
&\quad + o\left(\frac{1}{\sqrt{n}}\right).
\end{aligned}$$

Damit gilt (10.2.8), wenn

$$\max_{1 \leq h \leq M_{n1}} \max_{1 \leq k \leq M_{n2}} \left| \frac{1}{n} \sum_{j=1}^{\lfloor ns_h \rfloor} d_k(X_{j-1}) \right| = o_p\left(\frac{1}{\sqrt{n}}\right), \quad (10.2.9)$$

und es bleibt zu zeigen, dass (10.2.9) gilt.

Dies kann mit Hilfe des folgenden Hilfssatz, der auf einer Ungleichung vom Bernstein-Typ für mischende Zufallsvariablen basiert, gezeigt werden.

Hilfssatz 10.2.A Sei d eine messbare Funktionen mit $\sup_{x \in \mathbb{R}} |d(x)| \leq z_n$ und $\int d(y) f_{X_0}(y) dy = 0$ und s ein fester Wert in $[0, 1]$. Außerdem sei der Prozess $(X_i)_{i \in \mathbb{Z}}$ α -mischend. Dann gibt es für

hinreichend großes n ein $\bar{M} < \infty$, so dass

$$P \left(\left| \frac{1}{n} \sum_{j=1}^{\lfloor ns \rfloor} d(X_{j-1}) \right| > \bar{M} \bar{\epsilon}_n \right) \\ \leq 4 \exp \left(- \frac{n^2 \bar{M}^2 \bar{\epsilon}_n^2}{64n \lfloor n \bar{\epsilon}_n c_n^{\frac{1}{2}} \rfloor z_n^2 + \frac{8}{3} n \bar{M} \bar{\epsilon}_n \lfloor n \bar{\epsilon}_n c_n^{\frac{1}{2}} \rfloor z_n} \right) + 4 \frac{n}{\lfloor n \bar{\epsilon}_n c_n^{\frac{1}{2}} \rfloor} \alpha(\lfloor n \bar{\epsilon}_n c_n^{\frac{1}{2}} \rfloor),$$

wobei $\alpha(\cdot)$ der Mischungskoeffizient von (X_i) ist.

Beweis Messbare Funktionen erhalten nach Fan and Yao [2005] 2.6.1 (ii) Mischungseigenschaften, daher ist $d(X_{i-1})$ α -mischend mit dem selben Mischungskoeffizienten wie (X_i) . Der Mischungskoeffizient von

$$Y_i := d(X_{i-1}) I \left\{ \frac{i}{n} \leq s \right\}$$

kann für festes $s \in [0, 1]$ und $n \in \mathbb{N}$ durch den Mischungskoeffizienten $\alpha(\cdot)$ abgeschätzt werden.

Außerdem ist

$$|Y_i| \leq z_n$$

und da X_0, \dots, X_n identisch verteilt sind

$$E[Y_i] = \int d(y) I \left\{ \frac{i}{n} \leq s \right\} f_{X_{i-1}}(y) dy = I \left\{ \frac{i}{n} \leq s \right\} \int d(y) f_{X_0}(y) dy = 0.$$

Das Theorem 2.1 aus Liebscher [1996] ist also auf $\sum_{i=1}^n Y_i$ anwendbar. Dieses Theorem basiert auf Theorem 5 in Rio [1995].

Sei

$$N := \lfloor n \bar{\epsilon}_n c_n^{\frac{1}{2}} \rfloor = \lfloor n^{\frac{1}{2}} c_n^{\frac{1}{2}} \log(n)^{-1} \rfloor = \underbrace{\lfloor (nc_n^3)^{-\frac{1}{2}} \log(n) \rfloor^{-1} c_n^{-1}}_{\xrightarrow{n \rightarrow \infty} \infty}$$

dann sind die Bedingungen

$$1 \leq N \leq n$$

und

$$n \bar{M} \bar{\epsilon}_n > 4N z_n = 4 \lfloor n \bar{\epsilon}_n c_n^{\frac{1}{2}} \rfloor c_n^{-\frac{1}{2}} \underbrace{n^{-\frac{1}{2}} \log(n) q_n^f}_{= c_n^{\frac{3}{2}} (nc_n^3)^{-\frac{1}{2}} \log(n) q_n^f = o(1)}$$

nach Voraussetzung für hinreichend große n erfüllt.

Weiter gilt für alle $0 \leq T \leq n - 1$

$$E \left[\left(\sum_{i=T+1}^{\min(T+N, n)} Y_i \right)^2 \right] \\ \leq N^2 (\sup_{x \in \mathbb{R}} |d(x)|)^2$$

$$\leq N^2 z_n^2$$

und damit nach Liebscher [1996]

$$\begin{aligned} & P\left(\left|\sum_{i=1}^n Y_i\right| > n\bar{M}\bar{\epsilon}_n\right) \\ & \leq 4 \exp\left(-\frac{n^2 \bar{M}^2 \bar{\epsilon}_n^2}{64 \frac{n}{N} N^2 z_n^2 + \frac{8}{3} n \bar{M} \bar{\epsilon}_n N z_n}\right) + 4 \frac{n}{N} \alpha(N) \\ & = 4 \exp\left(-\frac{n^2 \bar{M}^2 \bar{\epsilon}_n^2}{64 n \lfloor n \bar{\epsilon}_n c_n^{\frac{1}{2}} \rfloor z_n^2 + \frac{8}{3} n \bar{M} \bar{\epsilon}_n \lfloor n \bar{\epsilon}_n c_n^{\frac{1}{2}} \rfloor z_n}\right) + 4 \frac{n}{\lfloor n \bar{\epsilon}_n c_n^{\frac{1}{2}} \rfloor} \alpha(\lfloor n \bar{\epsilon}_n c_n^{\frac{1}{2}} \rfloor). \end{aligned}$$

Somit gilt die Behauptung, denn

$$\begin{aligned} P\left(\left|\sum_{i=1}^n Y_i\right| > n\bar{M}\bar{\epsilon}_n\right) &= P\left(\left|\frac{1}{n} \sum_{i=1}^n Y_i\right| > \bar{M}\bar{\epsilon}_n\right) \\ &= P\left(\left|\frac{1}{n} \sum_{j=1}^{\lfloor ns \rfloor} d(X_{j-1})\right| > \bar{M}\bar{\epsilon}_n\right). \end{aligned}$$

□

Fortsetzung Beweis Lemma 10.2

Mit diesem Ergebnis gilt für (10.2.9)

$$\begin{aligned} & P\left(\max_{1 \leq h \leq M_{n1}} \max_{1 \leq k \leq M_{n2}} \left|\frac{1}{n} \sum_{j=1}^{\lfloor ns_h \rfloor} d_k(X_{j-1})\right| > \bar{M}\bar{\epsilon}_n\right) \\ & \leq \sum_{h=1}^{M_{n1}} \sum_{k=1}^{M_{n2}} P\left(\left|\frac{1}{n} \sum_{j=1}^{\lfloor ns_h \rfloor} d_k(X_{j-1})\right| > \bar{M}\bar{\epsilon}_n\right) \\ & \leq M_{n1} M_{n2} \left(4 \exp\left(-\frac{n^2 \bar{M}^2 \bar{\epsilon}_n^2}{64 n \lfloor n \bar{\epsilon}_n c_n^{\frac{1}{2}} \rfloor z_n^2 + \frac{8}{3} n \bar{M} \bar{\epsilon}_n \lfloor n \bar{\epsilon}_n c_n^{\frac{1}{2}} \rfloor z_n}\right) + 4 \frac{n}{\lfloor n \bar{\epsilon}_n c_n^{\frac{1}{2}} \rfloor} \alpha(\lfloor n \bar{\epsilon}_n c_n^{\frac{1}{2}} \rfloor)\right) \\ & \leq 4 \exp\left(\log((\bar{\epsilon}_n^*)^{-1}) + 2^{\frac{1}{1+\delta}} \bar{K}(2 + b_n - a_n) \bar{\epsilon}_n^{-\frac{1}{1+\delta}} - \frac{n^2 \bar{M}^2 \bar{\epsilon}_n^2}{64 n \lfloor n \bar{\epsilon}_n c_n^{\frac{1}{2}} \rfloor z_n^2 + \frac{8}{3} n \bar{M} \bar{\epsilon}_n \lfloor n \bar{\epsilon}_n c_n^{\frac{1}{2}} \rfloor z_n}\right) \\ & \quad + 4 \exp\left(\log((\bar{\epsilon}_n^*)^{-1}) + 2^{\frac{1}{1+\delta}} \bar{K}(2 + b_n - a_n) \bar{\epsilon}_n^{-\frac{1}{1+\delta}} + \log\left(\frac{n}{\lfloor n \bar{\epsilon}_n c_n^{\frac{1}{2}} \rfloor}\right) + \log\left(\alpha(\lfloor n \bar{\epsilon}_n c_n^{\frac{1}{2}} \rfloor)\right)\right). \end{aligned} \tag{10.2.10}$$

Die Zeile (10.2.10) ist $o(1)$, wenn der Exponent gegen minus unendlich geht, also wenn

$$\frac{64n \lfloor n\bar{\epsilon}_n c_n^{\frac{1}{2}} \rfloor z_n^2 + \frac{8}{3} n \bar{M} \bar{\epsilon}_n \lfloor n\bar{\epsilon}_n c_n^{\frac{1}{2}} \rfloor z_n}{n^2 \bar{M}^2 \bar{\epsilon}_n^2} = o\left(\frac{1}{2^{\frac{1}{1+\delta}} \bar{K} (2 + b_n - a_n) \bar{\epsilon}_n^{-\frac{1}{1+\delta}} + \log((\bar{\epsilon}_n^*)^{-1})}\right).$$

Dies gilt, denn $\log((\bar{\epsilon}_n^*)^{-1}) = \log(\bar{\epsilon}_n^{-1} z_n) = O(\log(n))$ und außerdem

$$\begin{aligned} & \left(2^{\frac{1}{1+\delta}} \bar{K} (2 + b_n - a_n) \bar{\epsilon}_n^{-\frac{1}{1+\delta}} + \log(n)\right) \frac{64n \lfloor n\bar{\epsilon}_n c_n^{\frac{1}{2}} \rfloor z_n^2 + \frac{8}{3} n \bar{M} \bar{\epsilon}_n \lfloor n\bar{\epsilon}_n c_n^{\frac{1}{2}} \rfloor z_n}{n^2 \bar{M}^2 \bar{\epsilon}_n^2} \\ &= O\left(\left((b_n - a_n) \bar{\epsilon}_n^{-\frac{1}{1+\delta}} + \log(n)\right) \left(\bar{\epsilon}_n^{-1} c_n^{\frac{1}{2}} z_n^2 + c_n^{\frac{1}{2}} z_n\right)\right) \\ &= O\left(\left((b_n - a_n) n^{\frac{1}{2} \cdot \frac{1}{1+\delta}} \log(n)^{\frac{1}{1+\delta}} + \log(n)\right)\right. \\ &\quad \left.\cdot \left(n^{\frac{1}{2}} \log(n) c_n^{\frac{1}{2}} c_n^{-1} n^{-1} \log(n)^2 (q_n^f)^2 + c_n^{\frac{1}{2}} c_n^{-\frac{1}{2}} n^{-\frac{1}{2}} \log(n) q_n^f\right)\right) \\ &= O\left((b_n - a_n) (q_n^f)^2 \log(n)^{3+\frac{1}{1+\delta}} c_n^{-\frac{1}{2}} n^{-\frac{1}{2}+\frac{1}{2} \cdot \frac{1}{1+\delta}}\right) \\ &\quad + O\left((b_n - a_n) q_n^f \log(n)^{1+\frac{1}{1+\delta}} n^{-\frac{1}{2}+\frac{1}{2} \cdot \frac{1}{1+\delta}}\right) \\ &\quad + O\left((q_n^f)^2 \log(n)^4 c_n^{-\frac{1}{2}} n^{-\frac{1}{2}}\right) \\ &\quad + O\left(q_n^f \log(n)^2 n^{-\frac{1}{2}}\right) \\ &= O\left((b_n - a_n) (q_n^f)^2 \log(n)^{3+\frac{1}{1+\delta}} c_n^{-\frac{1}{2}} n^{-\frac{1}{2}+\frac{1}{2} \cdot \frac{1}{1+\delta}}\right) \\ &\quad + O\left((q_n^f)^2 \log(n)^4 c_n^{-\frac{1}{2}} n^{-\frac{1}{2}}\right) \\ &= o(1), \end{aligned}$$

da nach Voraussetzung

$$\begin{aligned} & (b_n - a_n) (q_n^f)^2 \log(n)^{3+\frac{1}{1+\delta}} c_n^{-\frac{1}{2}} n^{-\frac{1}{2}+\frac{1}{2} \cdot \frac{1}{1+\delta}} \\ &= o(1), \end{aligned}$$

sowie

$$\begin{aligned} (q_n^f)^2 \log(n)^4 c_n^{-\frac{1}{2}} n^{-\frac{1}{2}} &= c_n \log(n)^3 (nc_n^3)^{-\frac{1}{2}} (q_n^f)^2 \log(n) \\ &= (nc_n^5)^{\frac{1}{5}} n^{-\frac{1}{5}} \log(n)^3 (nc_n^3)^{-\frac{1}{2}} (q_n^f)^2 \log(n) \\ &= O(1) \cdot o(1) \cdot o(1) = o(1). \end{aligned}$$

Die Zeile (10.2.11) ist ebenfalls $o(1)$, wenn der Exponent gegen minus unendlich geht, also wenn

$$\frac{1}{-\log\left(\alpha(\lfloor n\bar{\epsilon}_n c_n^{\frac{1}{2}} \rfloor)\right)} = o\left(\frac{1}{2^{\frac{1}{1+\delta}} \bar{K} (2 + b_n - a_n) \bar{\epsilon}_n^{-\frac{1}{1+\delta}} + \log\left(\frac{n}{\lfloor n\bar{\epsilon}_n c_n^{\frac{1}{2}} \rfloor}\right) + \log((\bar{\epsilon}_n^*)^{-1})}\right).$$

Auch diese Gleichung gilt, da $\left(\log\left(\alpha(\lfloor n\bar{\epsilon}_n c_n^{\frac{1}{2}} \rfloor)\right)\right)^{-1} = O\left(\left(\lfloor n\bar{\epsilon}_n c_n^{\frac{1}{2}} \rfloor \log(\bar{\alpha})\right)^{-1}\right)$, und weiter

$$\begin{aligned} & \frac{2^{\frac{1}{1+\delta}} \bar{K} (2 + b_n - a_n) \bar{\epsilon}_n^{-\frac{1}{1+\delta}} + \log(n)}{\lfloor n\bar{\epsilon}_n c_n^{\frac{1}{2}} \rfloor \log(\bar{\alpha})} \\ &= O\left((b_n - a_n) \bar{\epsilon}_n^{-\frac{1}{1+\delta}-1} n^{-1} c_n^{-\frac{1}{2}} + \log(n) \bar{\epsilon}_n^{-1} n^{-1} c_n^{-\frac{1}{2}}\right) \\ &= O\left((b_n - a_n) \log(n) \bar{\epsilon}_n^{-\frac{1}{1+\delta}-1} n^{-1} c_n^{-\frac{1}{2}}\right) \\ &= O\left((b_n - a_n) \log(n)^{2+\frac{1}{1+\delta}} c_n^{-\frac{1}{2}} n^{-\frac{1}{2}+\frac{1}{1+\delta}}\right) \\ &= o(1), \end{aligned}$$

nach Voraussetzung.

Damit gilt (10.2.9), denn

$$P\left(\max_{1 \leq h \leq M_{n1}} \max_{1 \leq k \leq M_{n2}} \left| \frac{1}{n} \sum_{j=1}^{\lfloor ns_h \rfloor} d_k(X_{j-1}) \right| > \bar{M} \bar{\epsilon}_n\right) = o(1),$$

also

$$\max_{1 \leq h \leq M_{n1}} \max_{1 \leq k \leq M_{n2}} \left| \frac{1}{n} \sum_{j=1}^{\lfloor ns_h \rfloor} d_k(X_{j-1}) \right| = O_P(\bar{\epsilon}_n) = o_p\left(\frac{1}{\sqrt{n}}\right).$$

□

Lemma 10.3 Es seien die Voraussetzungen von Lemma 10.2 (S. 145) und Lemma 10.1 (S. 142) gegeben.

Anmerkung: Damit müssen auch die Voraussetzungen von Lemma 3.3 (S. 21) und Lemma 3.4 (S. 35) erfüllt sein.

Der Prozess $(X_i)_{i \in \mathbb{Z}}$ sei also α -mischend mit $\alpha(n) = O(\bar{\alpha}^{-n})$ für ein $\bar{\alpha}$ mit $1 < \bar{\alpha} < \infty$ und alle X_i haben dieselbe Verteilung.

Für die Bandbreite c_n gelte $nc_n^5 = O(1)$ und es seien q_n^t, q_n^u und q_n^v mit $(q_n^t)^{-1} = O(1), (q_n^u)^{-1} = O(1), (q_n^v)^{-1} = O(1)$ gegeben, so dass $(nc_n^3)^{-1} (q_n^f)^8 (q_n^t)^4 (q_n^u)^4 (q_n^v)^4 \log(n)^2 = o(1)$ und es gebe ein $\delta > 0$, so dass $c_n^{-\frac{3}{2}-\delta} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} (q_n^f)^3 q_n^t q_n^u q_n^v = o(1)$ und $(b_n - a_n) (q_n^f)^2 (q_n^t)^2 (q_n^u)^2 (q_n^v)^2 \log(n)^{3+\frac{1}{1+\delta}} c_n^{-\frac{1}{2}} n^{-\frac{1}{2}+\frac{1}{1+\delta}} = o(1)$.

Anmerkung: Wenn $c_n \sim n^{-\frac{1}{4}} \log(n)^{r_1}$ für ein $|r_1| < \infty$, $(b_n - a_n) = O(\log(n)^{r_2})$ für ein $|r_2| < \infty$, $q_n^f = O(\log(n)^{r_3})$ für ein $|r_3| < \infty$, $q_n^t = O(\log(n)^{r_4})$ für ein $|r_4| < \infty$, $q_n^u = O(\log(n)^{r_5})$ für ein $|r_5| < \infty$ und $q_n^v = O(\log(n)^{r_6})$ für ein $|r_6| < \infty$, gilt dies für alle $\delta \in (\frac{1}{3}, \frac{1}{2})$.

Dann gilt für i.i.d. Zufallsvariablen (Z_i) wie in 10.2 und zweimal differenzierbare Funktionen $t_n : \mathbb{R} \rightarrow \mathbb{R}$ mit $\sup_{x \in I_n} |t_n^{(\mu)}(x)| = O(q_n^t)$, $\mu = 0, 1, 2$ und dreimal differenzierbare Funktionen $v_n :$

$\mathbb{R} \rightarrow \mathbb{R}$ und $u_n : \mathbb{R} \rightarrow \mathbb{R}$ mit $\sup_{x \in [a_n - Cc_n, b_n + Cc_n]} |v_n^{(\mu)}(x)| = O(q_n^v)$ und $\sup_{x \in [a_n - Cc_n, b_n + Cc_n]} |u_n^{(\mu)}(x)| = O(q_n^u)$, $\mu = 0, 1, 2, 3$

$$\sup_{s \in [0,1]} \left| \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk} \frac{1}{n} \sum_{i=1}^n Z_i \sum_{j=1}^{\lfloor ns \rfloor} \bar{w}_{j \lfloor ns \rfloor} t_n(X_{j-1}) \left(\frac{\frac{1}{c_n} K \left(\frac{X_{j-1} - X_{i-1}}{c_n} \right) (v_n(X_{i-1}) - v_n(X_{j-1})) u_n(X_{i-1})}{\frac{1}{nc_n} \sum_{k=1}^n K \left(\frac{X_{j-1} - X_{k-1}}{c_n} \right)} \right)} \right|$$

$$= o_p \left(\frac{1}{\sqrt{n}} \right).$$

Beweis Als erstes wird gezeigt, dass der zufällige Nenner ersetzt werden kann. Es gilt

$$\begin{aligned} & \sup_{s \in [0,1]} \left| \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk} \frac{1}{n} \sum_{i=1}^n Z_i \sum_{j=1}^{\lfloor ns \rfloor} \bar{w}_{j \lfloor ns \rfloor} t_n(X_{j-1}) \cdot \left(\frac{\frac{1}{c_n} K \left(\frac{X_{j-1} - X_{i-1}}{c_n} \right) (v_n(X_{i-1}) - v_n(X_{j-1})) u_n(X_{i-1})}{\frac{1}{nc_n} \sum_{k=1}^n K \left(\frac{X_{j-1} - X_{k-1}}{c_n} \right)} \right)} \right| \\ & \leq \sup_{s \in [0,1]} \left| \frac{1}{n} \sum_{i=1}^n Z_i \frac{1}{n} \sum_{j=1}^{\lfloor ns \rfloor} w_{nj} t_n(X_{j-1}) \left(\frac{\frac{1}{c_n} K \left(\frac{X_{j-1} - X_{i-1}}{c_n} \right) (v_n(X_{i-1}) - v_n(X_{j-1})) u_n(X_{i-1})}{f_{X_0}(X_{j-1})} \right) \right| \\ & \quad + \sup_{s \in [0,1]} \left| \frac{1}{n} \sum_{i=1}^n Z_i \frac{1}{n} \sum_{j=1}^{\lfloor ns \rfloor} w_{nj} t_n(X_{j-1}) \left(\frac{\frac{1}{c_n} K \left(\frac{X_{j-1} - X_{i-1}}{c_n} \right) (v_n(X_{i-1}) - v_n(X_{j-1})) u_n(X_{i-1})}{\frac{1}{nc_n} \sum_{k=1}^n K \left(\frac{X_{j-1} - X_{k-1}}{c_n} \right)} \right) \right. \\ & \quad \left. \cdot \left(\frac{f_{X_0}(X_{j-1}) - \frac{1}{nc_n} \sum_{k=1}^n K \left(\frac{X_{j-1} - X_{k-1}}{c_n} \right)}{f_{X_0}(X_{j-1})} \right) \right|. \end{aligned} \tag{10.3.1}$$

Als nächstes wird gezeigt, dass der zweite Summand $o_p \left(\frac{1}{\sqrt{n}} \right)$ ist. Es gilt

$$\begin{aligned} & \sup_{s \in [0,1]} \left| \frac{1}{n} \sum_{i=1}^n Z_i \frac{1}{n} \sum_{j=1}^{\lfloor ns \rfloor} w_{nj} t_n(X_{j-1}) \left(\frac{\frac{1}{c_n} K \left(\frac{X_{j-1} - X_{i-1}}{c_n} \right) (v_n(X_{i-1}) - v_n(X_{j-1})) u_n(X_{i-1})}{\frac{1}{nc_n} \sum_{k=1}^n K \left(\frac{X_{j-1} - X_{k-1}}{c_n} \right)} \right) \right. \\ & \quad \left. \cdot \left(\frac{f_{X_0}(X_{j-1}) - \frac{1}{nc_n} \sum_{k=1}^n K \left(\frac{X_{j-1} - X_{k-1}}{c_n} \right)}{f_{X_0}(X_{j-1})} \right) \right| \\ & = \sup_{s \in [0,1]} \left| \frac{1}{n} \sum_{j=1}^{\lfloor ns \rfloor} w_{nj} t_n(X_{j-1}) \left(\frac{f_{X_0}(X_{j-1}) - \frac{1}{nc_n} \sum_{k=1}^n K \left(\frac{X_{j-1} - X_{k-1}}{c_n} \right)}{f_{X_0}(X_{j-1})} \right) \right. \\ & \quad \left. \cdot \frac{1}{n} \sum_{i=1}^n Z_i \left(\frac{\frac{1}{c_n} K \left(\frac{X_{j-1} - X_{i-1}}{c_n} \right) (v_n(X_{i-1}) - v_n(X_{j-1})) u_n(X_{i-1})}{\frac{1}{nc_n} \sum_{k=1}^n K \left(\frac{X_{j-1} - X_{k-1}}{c_n} \right)} \right) \right| \\ & \leq \frac{\sup_{x \in I_n} |t_n(x)|}{\inf_{x \in I_n} f_{X_0}(x)} \underbrace{\sup_{x \in I_n} \left| f_{X_0}(x) - \frac{1}{nc_n} \sum_{k=1}^n K \left(\frac{x - X_{k-1}}{c_n} \right) \right|}_{=O_P \left(c_n^{-\frac{1}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} + c_n^2 \right)} \text{ nach Lemma 3.4} \end{aligned}$$

$$\begin{aligned}
& \cdot \sup_{x \in I_n} \left| \frac{\frac{1}{n} \sum_{i=1}^n \frac{1}{c_n} K\left(\frac{x-X_{i-1}}{c_n}\right) Z_i(v_n(X_{i-1}) - v_n(x)) u_n(X_{i-1})}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{x-X_{k-1}}{c_n}\right)} \right| \\
& = O(q_n^t q_n^f) \cdot O_P\left(c_n^{-\frac{1}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} + c_n^2\right) \\
& \quad \cdot O\left(\sup_{x \in I_n} \left| \frac{\frac{1}{n} \sum_{i=1}^n \frac{1}{c_n} K\left(\frac{x-X_{i-1}}{c_n}\right) Z_i(v_n(X_{i-1}) - v_n(x)) u_n(X_{i-1})}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{x-X_{k-1}}{c_n}\right)} \right|\right), \tag{10.3.2}
\end{aligned}$$

da $\sup_{x \in I_n} |t_n(x)| = O(q_n^t)$ und $\frac{1}{\inf_{x \in I_n} \frac{1}{n} \sum_{k=1}^n f_{X_{k-1}}(x)} = O(q_n^f)$ nach Voraussetzung.

Mit einer Taylorentwicklung von $u_n(X_{i-1})$ in x ist für ein $\xi_{n,1}(x, i)$ zwischen X_{i-1} und x

$$\begin{aligned}
& \sup_{x \in I_n} \left| \frac{\frac{1}{n} \sum_{i=1}^n \frac{1}{c_n} K\left(\frac{x-X_{i-1}}{c_n}\right) Z_i(v_n(X_{i-1}) - v_n(x)) u_n(X_{i-1})}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{x-X_{k-1}}{c_n}\right)} \right| \\
& \leq \sup_{x \in I_n} |u_n(x)| \sup_{x \in I_n} \left| \frac{\frac{1}{n} \sum_{i=1}^n \frac{1}{c_n} K\left(\frac{x-X_{i-1}}{c_n}\right) Z_i(v_n(X_{i-1}) - v_n(x))}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{x-X_{k-1}}{c_n}\right)} \right| \\
& \quad + \sup_{x \in I_n} \left| \frac{\frac{1}{n} \sum_{i=1}^n \frac{1}{c_n} K\left(\frac{x-X_{i-1}}{c_n}\right) Z_i(v_n(X_{i-1}) - v_n(x)) u'_n(\xi_{n,1}(x, i)) (X_{i-1} - x)}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{x-X_{k-1}}{c_n}\right)} \right|
\end{aligned}$$

und weiter mit einer Taylorentwicklung von $v_n(X_{i-1})$ in x und einem $\xi_{n,2}(x, i)$ zwischen X_{i-1} und x

$$\begin{aligned}
& \sup_{x \in I_n} \left| \frac{\frac{1}{n} \sum_{i=1}^n \frac{1}{c_n} K\left(\frac{x-X_{i-1}}{c_n}\right) Z_i(v_n(X_{i-1}) - v_n(x)) u_n(X_{i-1})}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{x-X_{k-1}}{c_n}\right)} \right| \\
& \leq \sup_{x \in I_n} |u_n(x)| \sup_{x \in I_n} |v'_n(x)| \sup_{x \in I_n} \left| \frac{\frac{1}{n} \sum_{i=1}^n \frac{1}{c_n} K\left(\frac{x-X_{i-1}}{c_n}\right) Z_i(X_{i-1} - x)}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{x-X_{k-1}}{c_n}\right)} \right| \\
& \quad + \sup_{x \in I_n} |u_n(x)| \sup_{x \in I_n} \left| \frac{\frac{1}{n} \sum_{i=1}^n \frac{1}{c_n} K\left(\frac{x-X_{i-1}}{c_n}\right) Z_i v''_n(\xi_{n,2}(x, i)) (X_{i-1} - x)^2}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{x-X_{k-1}}{c_n}\right)} \right| \\
& \quad + \sup_{x \in I_n} |v'_n(x)| \sup_{x \in I_n} \left| \frac{\frac{1}{n} \sum_{i=1}^n \frac{1}{c_n} K\left(\frac{x-X_{i-1}}{c_n}\right) Z_i u'_n(\xi_{n,1}(x, i)) (X_{i-1} - x)^2}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{x-X_{k-1}}{c_n}\right)} \right| \\
& \quad + \sup_{x \in I_n} \left| \frac{\frac{1}{n} \sum_{i=1}^n \frac{1}{c_n} K\left(\frac{x-X_{i-1}}{c_n}\right) Z_i v''_n(\xi_{n,2}(x, i)) u'_n(\xi_{n,1}(x, i)) (X_{i-1} - x)^3}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{x-X_{k-1}}{c_n}\right)} \right|
\end{aligned}$$

und da die Summanden nur dann ungleich Null sind, wenn $|X_{i-1} - x| \leq Cc_n$ gilt weiter mit $J_n := [a_n - Cc_n, b_n + Cc_n]$

$$\sup_{x \in I_n} \left| \frac{\frac{1}{n} \sum_{i=1}^n \frac{1}{c_n} K\left(\frac{x-X_{i-1}}{c_n}\right) Z_i(v_n(X_{i-1}) - v_n(x)) u_n(X_{i-1})}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{x-X_{k-1}}{c_n}\right)} \right|$$

$$\begin{aligned}
&\leq \sup_{x \in I_n} |u_n(x)| \sup_{x \in I_n} |v'_n(x)| \sup_{x \in I_n} \left| \frac{1}{n} \sum_{i=1}^n \frac{\frac{1}{c_n} K\left(\frac{x-X_{i-1}}{c_n}\right) Z_i(X_{i-1}-x)}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{x-X_{k-1}}{c_n}\right)} \right| \\
&\quad + \sup_{x \in I_n} |u_n(x)| \sup_{x \in J_n} |v''_n(x)| \sup_{z \in [-C, C]} K(z) C^2 c_n \sup_{x \in I_n} \left| \frac{1}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{x-X_{k-1}}{c_n}\right)} \right| \frac{1}{n} \sum_{i=1}^n |Z_i| \\
&\quad + \sup_{x \in J_n} |u'_n(x)| \sup_{x \in I_n} |v'_n(x)| \sup_{z \in [-C, C]} K(z) C^2 c_n \sup_{x \in I_n} \left| \frac{1}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{x-X_{k-1}}{c_n}\right)} \right| \frac{1}{n} \sum_{i=1}^n |Z_i| \\
&\quad + \sup_{x \in J_n} |u'_n(x)| \sup_{x \in J_n} |v''_n(x)| \underbrace{\sup_{z \in [-C, C]} K(z) C^2 c_n^2}_{\leq \bar{K} < \infty} \sup_{x \in I_n} \left| \frac{1}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{x-X_{k-1}}{c_n}\right)} \right| \underbrace{\frac{1}{n} \sum_{i=1}^n |Z_i|}_{=O_P(1)},
\end{aligned}$$

was mit $\sup_{x \in J_n} |u_n^{(\mu)}(x)| = O(q_n^u)$ und $\sup_{x \in J_n} |v_n^{(\mu)}(x)| = O(q_n^v)$, $\mu = 0, 1, 2$, bedeutet, dass

$$\begin{aligned}
&\sup_{x \in I_n} \left| \frac{1}{n} \sum_{i=1}^n \frac{\frac{1}{c_n} K\left(\frac{x-X_{i-1}}{c_n}\right) Z_i(v_n(X_{i-1}) - v_n(x)) u_n(X_{i-1})}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{x-X_{k-1}}{c_n}\right)} \right| \\
&= O(q_n^u q_n^v) \cdot O\left(\sup_{x \in I_n} \left| \frac{1}{n} \sum_{i=1}^n \frac{\frac{1}{c_n} K\left(\frac{x-X_{i-1}}{c_n}\right) Z_i(X_{i-1}-x)}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{x-X_{k-1}}{c_n}\right)} \right| \right) \\
&\quad + O_P\left(c_n q_n^f q_n^u q_n^v\right), \tag{10.3.3}
\end{aligned}$$

da analog zum Vorgehen im Beweis von Lemma 3.4 gilt

$$\begin{aligned}
\sup_{x \in I_n} \left| \frac{1}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{x-X_{k-1}}{c_n}\right)} \right| &= \sup_{x \in I_n} \left| \frac{\frac{1}{\frac{1}{n} \sum_{k=1}^n f_{X_{k-1}}(x)}}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{x-X_{k-1}}{c_n}\right)} \right| \\
&= \frac{1}{\inf_{x \in I_n} \frac{1}{n} \sum_{k=1}^n f_{X_{k-1}}(x)} \\
&= 1 + O_P\left(c_n^{-\frac{1}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} q_n^f + c_n^2 q_n^f\right) \\
&= O_P\left(q_n^f\right).
\end{aligned}$$

Weiter ist

$$\begin{aligned}
&\sup_{x \in I_n} \left| \frac{1}{n} \sum_{i=1}^n \frac{\frac{1}{c_n} K\left(\frac{x-X_{i-1}}{c_n}\right) Z_i(X_{i-1}-x)}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{x-X_{k-1}}{c_n}\right)} \right| \\
&= c_n \sup_{x \in I_n} \left| \frac{1}{n} \sum_{i=1}^n \frac{1}{c_n} K\left(\frac{x-X_{i-1}}{c_n}\right) \left(\frac{x-X_{i-1}}{c_n}\right) Z_i \right| \sup_{x \in I_n} \left| \frac{1}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{x-X_{k-1}}{c_n}\right)} \right|
\end{aligned}$$

$$\begin{aligned}
&= c_n \sup_{x \in I_n} \left| \frac{1}{n} \sum_{i=1}^n \frac{1}{c_n} K\left(\frac{x - X_{i-1}}{c_n}\right) \left(\frac{x - X_{i-1}}{c_n}\right) Z_i - E \left[\frac{1}{n} \sum_{i=1}^n \frac{1}{c_n} K\left(\frac{x - X_{i-1}}{c_n}\right) \left(\frac{x - X_{i-1}}{c_n}\right) Z_i \right] \right| \\
&\qquad \qquad \qquad \cdot O_P\left(q_n^f\right), \\
&\qquad \qquad \qquad (10.3.4)
\end{aligned}$$

da Z_i von X_{i-1} unabhängig ist, und $E[Z_i] = 0$ für alle i .

Auf den Term

$$\sup_{x \in I_n} \left| \frac{1}{n} \sum_{i=1}^n \frac{1}{c_n} K\left(\frac{x - X_{i-1}}{c_n}\right) \left(\frac{x - X_{i-1}}{c_n}\right) Z_i - E \left[\frac{1}{n} \sum_{i=1}^n \frac{1}{c_n} K\left(\frac{x - X_{i-1}}{c_n}\right) \left(\frac{x - X_{i-1}}{c_n}\right) Z_i \right] \right|$$

kann Lemma 10.1 angewendet werden, da $X_i = Z_i$ die Voraussetzungen an X_i in Lemma 3.3 nach Voraussetzung erfüllt.

Damit ist

$$\begin{aligned}
&\sup_{x \in I_n} \left| \frac{1}{n} \sum_{i=1}^n \frac{1}{c_n} K\left(\frac{x - X_{i-1}}{c_n}\right) \left(\frac{x - X_{i-1}}{c_n}\right) Z_i - E \left[\frac{1}{n} \sum_{i=1}^n \frac{1}{c_n} K\left(\frac{x - X_{i-1}}{c_n}\right) \left(\frac{x - X_{i-1}}{c_n}\right) Z_i \right] \right| \\
&= O_P\left(c_n^{-\frac{1}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}}\right) = O_P(1),
\end{aligned}$$

und daher folgt aus (10.3.2), (10.3.3) und (10.3.4)

$$\begin{aligned}
&\sup_{s \in [0,1]} \left| \frac{1}{n} \sum_{i=1}^n Z_i \frac{1}{n} \sum_{j=1}^{\lfloor ns \rfloor} w_{nj} t_n(X_{j-1}) \left(\frac{\frac{1}{c_n} K\left(\frac{X_{j-1} - X_{i-1}}{c_n}\right) (v_n(X_{i-1}) - v_n(X_{j-1})) u_n(X_{i-1})}{\frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{X_{j-1} - X_{k-1}}{c_n}\right)} \right) \right. \\
&\qquad \qquad \qquad \left. \cdot \left(\frac{f_{X_0}(X_{j-1}) - \frac{1}{nc_n} \sum_{k=1}^n K\left(\frac{X_{j-1} - X_{k-1}}{c_n}\right)}{f_{X_0}(X_{j-1})} \right) \right| \\
&= O(q_n^t q_n^f) O_P\left(c_n^{-\frac{1}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} + c_n^2\right) O_P\left(c_n q_n^f q_n^u q_n^v\right) \\
&= O_P\left(c_n^{\frac{1}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} (q_n^f)^2 q_n^t q_n^u q_n^v\right) \\
&= O_P\left(\frac{1}{\sqrt{n}}\right) \cdot O_P\left(c_n^{\frac{1}{2}} \log(n)^{\frac{1}{2}} (q_n^f)^2 q_n^t q_n^u q_n^v\right) \\
&= o_p\left(\frac{1}{\sqrt{n}}\right),
\end{aligned}$$

da $c_n^{\frac{1}{2}} \log(n)^{\frac{1}{2}} (q_n^f)^2 q_n^t q_n^u q_n^v = \left((nc_n^5) ((nc_n^3)^{-1} \log(n)^2 (q_n^f)^8 (q_n^t)^4 (q_n^u)^4 (q_n^v)^4) \right)^{\frac{1}{4}} = O(1) \cdot o(1)$ nach Voraussetzung.

Mit (10.3.1) bleibt also zu zeigen, dass

$$\begin{aligned}
&\sup_{s \in [0,1]} \left| \frac{1}{n} \sum_{i=1}^n Z_i \frac{1}{n} \sum_{j=1}^{\lfloor ns \rfloor} w_{nj} t_n(X_{j-1}) \left(\frac{\frac{1}{c_n} K\left(\frac{X_{j-1} - X_{i-1}}{c_n}\right) (v_n(X_{i-1}) - v_n(X_{j-1})) u_n(X_{i-1})}{f_{X_0}(X_{j-1})} \right) \right| \\
&= o_p\left(\frac{1}{\sqrt{n}}\right).
\end{aligned}$$

Sei

$$\hat{d}_n(x) := w_n(x)t_n(x)\frac{1}{nc_n}\sum_{i=1}^n K\left(\frac{x-X_{i-1}}{c_n}\right)Z_i(v_n(X_{i-1})-v_n(x))u_n(X_{i-1}).$$

Damit bleibt zu zeigen, dass

$$\sup_{s\in[0,1]}\left|\frac{1}{n}\sum_{j=1}^{\lfloor ns\rfloor}\hat{d}_n(X_{j-1})\right|=o_p\left(\frac{1}{\sqrt{n}}\right). \quad (10.3.5)$$

Dies geschieht analog zum Beweis von Lemma 10.2, indem zunächst gezeigt wird, dass $\hat{d}_n(X_{j-1})$ bezüglich X_{j-1} zentriert werden kann, also dass

$$\left|\int\hat{d}_n(x)f_{X_0}(x)dx\right|=o_p\left(\frac{1}{\sqrt{n}}\right). \quad (10.3.6)$$

Sei hierfür

$$\bar{h}(x):=\int w_n(y)t_n(y)\frac{1}{c_n}K\left(\frac{y-x}{c_n}\right)(v_n(x)-v_n(y))u_n(x)dy$$

dann ist (10.3.6) gleichbedeutend mit

$$\left|\frac{1}{n}\sum_{i=1}^n Z_i\bar{h}(X_{i-1})\right|=o_p\left(\frac{1}{\sqrt{n}}\right).$$

Da

$$\bar{h}(x)=\int w_n(zc_n+x)t_n(zc_n+x)K(z)(v_n(x)-v_n(zc_n+x))u_n(x)dz.$$

gilt

$$\begin{aligned} \sup_{x\in\mathbb{R}}|\bar{h}(x)| &\leq \underbrace{\sup_{x\in[a_n-2Cc_n,b_n+2Cc_n]}|t_n(x)|c_n}_{=O(q_n^t)} \underbrace{\sup_{x\in[a_n-2Cc_n,b_n+2Cc_n]}|v_n(x)|\sup_{x\in J_n}|u_n(x)|}_{=O(q_n^u q_n^v)} \underbrace{\int K(z)|z|dz}_{\leq 2C\cdot C\bar{K}<\infty} \\ &= O(c_n q_n^t q_n^u q_n^v) \\ &= O\left(\left((nc_n^5)(nc_n^3)^{-1}(q_n^t)^2(q_n^u)^2(q_n^v)^2\right)^{\frac{1}{2}}\right) \\ &= O\left(\left((nc_n^5)(nc_n^3)^{-1}(q_n^f)^8(q_n^t)^4(q_n^u)^4(q_n^v)^4\log(n)^2\right)^{\frac{1}{2}}\right) \\ &= O(1)\cdot o(1) \quad \text{nach Voraussetzung} \end{aligned}$$

und daher

$$E[(\bar{h}(X))^2]=o(1).$$

Damit gilt analog zum Beweis von Lemma 10.2

$$E\left[\frac{1}{n}\sum_{i=1}^n Z_i\bar{h}(X_{i-1})\right]=0$$

und

$$\text{Var} \left(\frac{1}{n} \sum_{i=1}^n Z_i \bar{h}(X_{i-1}) \right) = o \left(\frac{1}{n} \right),$$

also

$$\left| \frac{1}{n} \sum_{i=1}^n Z_i \bar{h}(X_{i-1}) \right| = o_p \left(\frac{1}{\sqrt{n}} \right),$$

und somit (10.3.6).

Damit gilt (10.3.5), wenn

$$\sup_{s \in [0,1]} \left| \frac{1}{n} \sum_{j=1}^{\lfloor ns \rfloor} \hat{d}_{En}(X_{j-1}) \right| = o_p \left(\frac{1}{\sqrt{n}} \right), \quad (10.3.7)$$

wobei

$$\hat{d}_{En}(x) := \hat{d}_n(x) - \int \hat{d}_n(y) f_{X_0}(y) dy$$

wie im Beweis zu Lemma 10.2, und es bleibt zu zeigen, dass (10.3.7) gilt.

Auch hier gehört \hat{d}_{En} zu einer Funktionenklasse, was im Folgenden gezeigt wird.

Für die Funktion \hat{d}_{En} gilt wie im Beweis zu Lemma 10.2

$$\sup_{x \in \mathbb{R}} |\hat{d}_{En}(x)| \leq 2 \sup_{x \in I_n} |\hat{d}_n(x)|$$

und für die Ableitungen

$$\sup_{x \in \mathbb{R}} |\hat{d}'_{En}(x)| = \sup_{x \in I_n} |\hat{d}'_n(x)| \quad \text{und} \quad \sup_{x \in \mathbb{R}} |\hat{d}''_{En}(x)| = \sup_{x \in I_n} |\hat{d}''_n(x)|.$$

Weiter gilt

$$\begin{aligned} \hat{d}_n(x) &= \frac{w_n(x)t_n(x)}{f_{X_0}(x)} \frac{1}{nc_n} \sum_{i=1}^n K \left(\frac{x - X_{i-1}}{c_n} \right) Z_i v_n(X_{i-1}) u_n(X_{i-1}) \\ &\quad - \frac{w_n(x)t_n(x)v_n(x)}{f_{X_0}(x)} \frac{1}{nc_n} \sum_{i=1}^n K \left(\frac{x - X_{i-1}}{c_n} \right) Z_i u_n(X_{i-1}) \\ &= \frac{w_n(x)t_n(x)}{f_{X_0}(x)} \left(\underbrace{\frac{1}{nc_n} \sum_{i=1}^n K \left(\frac{x - X_{i-1}}{c_n} \right) Z_i v_n(X_{i-1}) u_n(X_{i-1})}_{=: B_1} - E[B_1] \right) \\ &\quad - \frac{w_n(x)t_n(x)v_n(x)}{f_{X_0}(x)} \left(\underbrace{\frac{1}{nc_n} \sum_{i=1}^n K \left(\frac{x - X_{i-1}}{c_n} \right) Z_i u_n(X_{i-1})}_{=: B_2} - E[B_2] \right), \end{aligned}$$

da nach Voraussetzung X_{i-1} und Z_i unabhängig sind und $E[Z_i] = 0$ für alle i .

Es gilt ebenfalls nach Voraussetzung $\sup_{x \in [a_n - Cc_n, b_n + Cc_n]} |v_n^{(\mu)}(x)| = O(q_n^v)$, sowie $\sup_{x \in [a_n - Cc_n, b_n + Cc_n]} |u_n^{(\mu)}(x)| = O(q_n^u)$, $\mu = 0, 1, 2, 3$, daher erfüllen die Funktionen $\frac{1}{q_n^v} v_n u_n$ und $\frac{1}{q_n^u} u_n$ die Bedingungen an die Funktionen g_n in Lemma 10.1.

Somit folgt, da $\sup_{x \in I_n} |t_n^{(\mu)}(x)| = O(q_n^t)$, $\sup_{x \in I_n} |w_n^{(\mu)}(x)| = O(1)$, $\sup_{x \in I_n} |f_{X_0}^{(\mu)}(x)| = O(1)$, $\mu = 0, 1, 2$, und $\frac{1}{\inf_{x \in I_n} f_{X_0}(x)} = O(q_n^f)$,

$$\sup_{x \in I_n} |\hat{d}_n(x)| = O_P \left(c_n^{-\frac{1}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} q_n^f q_n^t q_n^v q_n^u \right) = o_P(1)$$

$$\sup_{x \in I_n} |\hat{d}'_n(x)| = O_P \left(c_n^{-\frac{3}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} (q_n^f)^2 q_n^t q_n^v q_n^u \right) = o_P(1)$$

$$\sup_{x \in I_n} |\hat{d}''_n(x)| = O_P \left(c_n^{-\frac{5}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} (q_n^f)^3 q_n^t q_n^v q_n^u \right),$$

da $c_n^{-\frac{3}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} (q_n^f)^2 q_n^t q_n^v q_n^u = (n c_n^3)^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} (q_n^f)^2 q_n^t q_n^v q_n^u = o(1)$ nach Voraussetzung.

Für alle δ gilt außerdem

$$\begin{aligned} \sup_{x, y \in J_n, x \neq y} \frac{|\hat{d}'_n(x) - \hat{d}'_n(y)|}{|x - y|^\delta} &\leq \sup_{x, y \in J_n, |x - y| > c_n} \frac{|\hat{d}'_n(x) - \hat{d}'_n(y)|}{|x - y|^\delta} + \sup_{x, y \in J_n, 0 < |x - y| \leq c_n} \frac{|\hat{d}'_n(x) - \hat{d}'_n(y)|}{|x - y|^\delta} \\ &\leq 2 \sup_{x \in J_n} |\hat{d}'_n(x)| c_n^{-\delta} + \sup_{x \in J_n} |\hat{d}''_n(x)| c_n^{1-\delta} \\ &= O_P \left(c_n^{-\frac{3}{2}-\delta} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} (q_n^f)^3 q_n^t q_n^v q_n^u \right). \end{aligned}$$

Damit gilt für alle $\delta > 0$ mit $c_n^{-\frac{3}{2}-\delta} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} (q_n^f)^3 q_n^t q_n^v q_n^u = o(1)$

$$P \left(\hat{d}_{E_n} \notin C_1^{1+\delta}(I_n)^n \right) = o(1)$$

für die Funktionenklasse

$$C_1^{1+\delta}(I_n)^n := \left\{ d \in C_1^{1+\delta}(I_n), \sup_{x \in I_n} |d(x)| \leq z_n, \int d(y) f_{X_0}(y) dy = 0 \right\},$$

wobei $z_n := c_n^{-\frac{1}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} q_n^f q_n^t q_n^v q_n^u \log(n)^{\frac{1}{2}}$ und

$$C_1^{1+\delta}(I_n) := \left\{ d : I_n \rightarrow \mathbb{R} \left| \max \left(\sup_{x \in I_n} |d(x)|, \sup_{x \in I_n} |d'(x)| \right) + \sup_{x, y \in I_n, x \neq y} \frac{|d'(x) - d'(y)|}{|x - y|^\delta} \leq 1 \right. \right\}$$

sei.

Somit gilt (10.3.7), und damit die Behauptung des Satzes, wenn

$$\sup_{s \in [0, 1]} \sup_{d \in C_1^{1+\delta}(I_n)^n} \left| \frac{1}{n} \sum_{j=1}^{\lfloor ns \rfloor} d(X_{j-1}) \right| = o_p \left(\frac{1}{\sqrt{n}} \right).$$

Dies folgt mit den Bedingungen an die Bandbreite analog zum Beweis von Lemma 10.2.

□

Lemma 10.4 Es seien die Voraussetzungen von Lemma 3.4 (S. 35) und Lemma 10.1 (S. 142) gegeben.

Anmerkung: Damit müssen auch die Voraussetzungen von Lemma 3.3 (S. 21) erfüllt sein.

Außerdem sei der Prozess $(X_i)_{i \in \mathbb{Z}}$ α -mischend mit $\alpha(n) = O(\bar{\alpha}^{-n})$ für ein $\bar{\alpha}$ mit $1 < \bar{\alpha} < \infty$.

Zusätzlich seien die Beobachtungen X_0, \dots, X_n identisch verteilt und es sei $q_n^f := \frac{1}{\inf_{x \in I_n} f_{X_{i-1}}(x)}$ für alle i .

Anmerkung: Durch die Voraussetzungen von Lemma 3.4 gilt für die Dichte der Beobachtungen auch $\sup_{x \in \mathbb{R}} f_{X_0}(x) < \infty$, und damit auch $(q_n^f)^{-1} = O(1)$.

Die Gewichtsfunktion w_n sei zweimal differenzierbar und es gelte $\sup_{x \in \mathbb{R}} |w_n^{(\mu)}(x)| < \infty$ für $\mu = 1, 2$.

Für die Bandbreite c_n gelte $nc_n^5 = O(1)$ und $(nc_n^3)^{-1}(q_n^f)^4 \log(n)^2 = o(1)$ und es gebe ein δ , so dass $c_n^{-\frac{3}{2}-\delta} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} (q_n^f)^3 = o(1)$ und $(b_n - a_n)(q_n^f)^{-\frac{1}{1+\delta}+1} \log(n)^2 c_n^{-\frac{2}{1+\delta}-\frac{5}{2}} n^{-1} = o(1)$.

Anmerkung: Wenn $c_n \sim n^{-\frac{1}{4}} \log(n)^{r_1}$ für ein $|r_1| < \infty$, $(b_n - a_n) = O(\log(n)^{r_2})$ für ein $|r_2| < \infty$ und $q_n^f = O(\log(n)^{r_3})$ für ein $|r_3| < \infty$, gilt dies für alle $\delta \in (\frac{1}{3}, \frac{1}{2})$.

Dann gilt für zweimal differenzierbare Funktionen $t_n : \mathbb{R} \rightarrow \mathbb{R}$ mit $\sup_{x \in I_n} |t_n^{(\mu)}(x)| = O(1)$, $\mu = 0, 1, 2$ und dreimal differenzierbare Funktionen $v_n : \mathbb{R} \rightarrow \mathbb{R}$ und $u_n : \mathbb{R} \rightarrow \mathbb{R}$ mit $\sup_{x \in [a_n - Cc_n, b_n + Cc_n]} |v_n^{(\mu)}(x)| = O(1)$ und $\sup_{x \in [a_n - Cc_n, b_n + Cc_n]} |u_n^{(\mu)}(x)| = O(1)$, $\mu = 0, 1, 2, 3$

$$\begin{aligned} \sup_{s \in [0,1]} & \left| \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^{\lfloor ns \rfloor} \bar{w}_{j \lfloor ns \rfloor} t_n(X_{j-1}) \left(\frac{\frac{1}{c_n} K \left(\frac{X_{j-1} - X_{i-1}}{c_n} \right) (v_n(X_{i-1}) - v_n(X_{j-1})) u_n(X_{i-1})}{\frac{1}{nc_n} \sum_{k=1}^n K \left(\frac{X_{j-1} - X_{k-1}}{c_n} \right)} \right) \right| \\ & = O_P \left(\left(c_n^{\frac{1}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} + c_n^2 \right) q_n^f \right). \end{aligned}$$

Beweis Analog zum Beweis von Lemma 10.3 kann der zufällige Nenner ersetzt werden und es gilt

$$\begin{aligned} \sup_{s \in [0,1]} & \left| \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^{\lfloor ns \rfloor} \bar{w}_{j \lfloor ns \rfloor} t_n(X_{j-1}) \cdot \left(\frac{\frac{1}{c_n} K \left(\frac{X_{j-1} - X_{i-1}}{c_n} \right) (v_n(X_{i-1}) - v_n(X_{j-1})) u_n(X_{i-1})}{\frac{1}{nc_n} \sum_{k=1}^n K \left(\frac{X_{j-1} - X_{k-1}}{c_n} \right)} \right) \right| \\ & = \sup_{s \in [0,1]} \left| \frac{1}{n} \sum_{i=1}^n \frac{1}{n} \sum_{j=1}^{\lfloor ns \rfloor} w_{nj} t_n(X_{j-1}) \left(\frac{\frac{1}{c_n} K \left(\frac{X_{j-1} - X_{i-1}}{c_n} \right) (v_n(X_{i-1}) - v_n(X_{j-1})) u_n(X_{i-1})}{f_{X_0}(X_{j-1})} \right) \right| \\ & + O(q_n^f) \cdot O_P \left(c_n^{-\frac{1}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} + c_n^2 \right) \cdot O(c_n) \end{aligned} \quad (10.4.1)$$

$$\begin{aligned}
&= \sup_{s \in [0,1]} \left| \frac{1}{n} \sum_{i=1}^n \frac{1}{n} \sum_{j=1}^{\lfloor ns \rfloor} w_{nj} t_n(X_{j-1}) \left(\frac{\frac{1}{c_n} K \left(\frac{X_{j-1} - X_{i-1}}{c_n} \right) (v_n(X_{i-1}) - v_n(X_{j-1})) u_n(X_{i-1})}{f_{X_0}(X_{j-1})} \right) \right| \\
&\quad + O_P \left(c_n^{\frac{1}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} q_n^f + c_n^3 q_n^f \right). \tag{10.4.2}
\end{aligned}$$

Anmerkung: In Zeile (10.4.1) gilt die Konvergenzrate $O(c_n)$ (anstatt $O_P(c_n q_n^f)$ im Beweis zu Lemma 10.3), da durch das Wegfallen von Z_i der Term $\sup_{x \in I_n} \left(\frac{1}{n} \sum_{i=1}^n \frac{\frac{1}{c_n} K \left(\frac{x - X_{i-1}}{c_n} \right) |X_{i-1} - x|}{\frac{1}{nc_n} \sum_{k=1}^n K \left(\frac{x - X_{k-1}}{c_n} \right)} \right)$ direkt durch Cc_n abgeschätzt werden kann.

Sei

$$\begin{aligned}
\hat{d}_n(x) : &= \frac{w_n(x) t_n(x)}{f_{X_0}(x)} \frac{1}{nc_n} \sum_{i=1}^n K \left(\frac{x - X_{i-1}}{c_n} \right) (v_n(X_{i-1}) - v_n(x)) u_n(X_{i-1}) \\
&\quad - \frac{w_n(x) t_n(x)}{f_{X_0}(x)} E \left[\frac{1}{nc_n} \sum_{i=1}^n K \left(\frac{x - X_{i-1}}{c_n} \right) (v_n(X_{i-1}) - v_n(x)) u_n(X_{i-1}) \right]
\end{aligned}$$

und

$$\hat{d}_{En}(x) := \hat{d}_n(x) - \int \hat{d}_n(y) f_{X_0}(y) dy.$$

Damit ist

$$\begin{aligned}
&\sup_{s \in [0,1]} \left| \frac{1}{n} \sum_{i=1}^n \frac{1}{n} \sum_{j=1}^{\lfloor ns \rfloor} w_{nj} t_n(X_{j-1}) \left(\frac{\frac{1}{c_n} K \left(\frac{X_{j-1} - X_{i-1}}{c_n} \right) (v_n(X_{i-1}) - v_n(X_{j-1})) u_n(X_{i-1})}{f_{X_0}(X_{j-1})} \right) \right| \\
&= \sup_{s \in [0,1]} \left| \frac{1}{n} \sum_{j=1}^{\lfloor ns \rfloor} \hat{d}_{En}(X_{j-1}) \right. \\
&\quad + \frac{1}{n} \sum_{i=1}^n \frac{1}{nc_n} \sum_{j=1}^{\lfloor ns \rfloor} \int \frac{w_n(y) t_n(y)}{f_{X_0}(y)} (v_n(X_{i-1}) - v_n(y)) K \left(\frac{X_{i-1} - y}{c_n} \right) u_n(X_{i-1}) f_{X_0}(y) dy \\
&\quad + \frac{1}{nc_n} \sum_{j=1}^{\lfloor ns \rfloor} \frac{w_n(X_{j-1}) t_n(X_{j-1})}{f_{X_0}(X_{j-1})} \int (v_n(x) - v_n(X_{j-1})) K \left(\frac{x - X_{j-1}}{c_n} \right) u_n(x) f_{X_0}(x) dx \\
&\quad \left. - \frac{1}{nc_n} \sum_{j=1}^{\lfloor ns \rfloor} \int \frac{w_n(y) t_n(y)}{f_{X_0}(y)} \int (v_n(x) - v_n(y)) K \left(\frac{x - y}{c_n} \right) u_n(x) f_{X_0}(x) dx f_{X_0}(y) dy \right|. \tag{10.4.3}
\end{aligned}$$

Weiter gilt

$$\begin{aligned}
&\sup_{s \in [0,1]} \left| \frac{1}{n} \sum_{i=1}^n \frac{1}{nc_n} \sum_{j=1}^{\lfloor ns \rfloor} \int \frac{w_n(y) t_n(y)}{f_{X_0}(y)} (v_n(X_{i-1}) - v_n(y)) K \left(\frac{X_{i-1} - y}{c_n} \right) u_n(X_{i-1}) f_{X_0}(y) dy \right| \\
&\leq \sup_{x \in \mathbb{R}} \sup_{s \in [0,1]} \left| \frac{\lfloor ns \rfloor}{n} \int \frac{1}{c_n} w_n(y) t_n(y) (v_n(x) - v_n(y)) K \left(\frac{x - y}{c_n} \right) u_n(x) dy \right| \\
&= \sup_{x \in \mathbb{R}} \left| \int \frac{1}{c_n} w_n(y) t_n(y) (v_n(x) - v_n(y)) K \left(\frac{x - y}{c_n} \right) u_n(x) dy \right|
\end{aligned}$$

$$\begin{aligned}
&\leq \sup_{x \in \mathbb{R}} \left| u_n(x) w_n(x) t_n(x) \int \frac{1}{c_n} (v_n(x) - v_n(y)) K \left(\frac{x-y}{c_n} \right) dy \right| \\
&\quad + \sup_{x \in \mathbb{R}} \left| u_n(x) \int \frac{1}{c_n} (v_n(x) - v_n(y)) (w_n t_n)'(\xi_{n,1}(x, y)) (y-x) K \left(\frac{x-y}{c_n} \right) dy \right| \\
&\quad \text{Taylorentwicklung von } w_n t_n \text{ in } x \text{ mit } \xi_{n,1}(x, y) \text{ zwischen } x \text{ und } y \\
&\leq \sup_{x \in \mathbb{R}} \left| u_n(x) w_n(x) t_n(x) v_n'(x) \int \frac{1}{c_n} (y-x) K \left(\frac{x-y}{c_n} \right) dy \right| \\
&\quad + \sup_{x \in \mathbb{R}} \left| u_n(x) w_n(x) t_n(x) \int \frac{1}{c_n} v_n''(\xi_{n,2}(x, y)) (y-x)^2 K \left(\frac{x-y}{c_n} \right) dy \right| \\
&\quad + \sup_{x \in \mathbb{R}} \left| u_n(x) \int \frac{1}{c_n} v_n'(\xi_{n,3}(x, y)) (w_n t_n)'(\xi_{n,1}(x, y)) (y-x)^2 K \left(\frac{x-y}{c_n} \right) dy \right| \\
&\quad \text{Taylorentwicklung von } v_n \text{ in } x \text{ mit } \xi_{n,2}(x, y) \text{ und } \xi_{n,3}(x, y) \text{ zwischen } x \text{ und } y \\
&\leq \underbrace{\sup_{x \in I_n} |u_n(x)| \sup_{x \in I_n} |t_n(x)| \sup_{x \in I_n} |v_n'(x)| c_n}_{=O(1) \text{ nach Voraussetzung}} \sup_{x \in \mathbb{R}} \left| \underbrace{\int \frac{1}{c_n} \frac{y-x}{c_n} K \left(\frac{x-y}{c_n} \right) dy}_{=\int z K(z) dz = 0} \right| \\
&\quad + \underbrace{\sup_{x \in I_n} |u_n(x)| \sup_{x \in I_n} |t_n(x)| \sup_{x \in [a_n - Cc_n, b_n + Cc_n]} |v_n''(x)| c_n^2}_{=O(1) \text{ nach Voraussetzung}} \int z^2 K(z) dz \\
&\quad + \underbrace{\sup_{x \in [a_n - Cc_n, b_n + Cc_n]} |u_n(x)| \sup_{x \in I_n} |(w_n t_n)'(x)| \sup_{x \in [a_n - 2Cc_n, b_n + 2Cc_n]} |v_n'(x)| c_n^2}_{=O(1) \text{ nach Voraussetzung}} \underbrace{\int z^2 K(z) dz}_{\leq 2CC^2 \bar{K} < \infty} \\
&= O(c_n^2),
\end{aligned}$$

und damit auch

$$\begin{aligned}
&\sup_{s \in [0,1]} \left| \frac{1}{nc_n} \sum_{j=1}^{\lfloor ns \rfloor} \int \frac{w_n(y) t_n(y)}{f_{X_0}(y)} \int (v_n(x) - v_n(y)) K \left(\frac{x-y}{c_n} \right) u_n(x) f_{X_0}(x) dx f_{X_0}(y) dy \right| \\
&= \left| \frac{1}{c_n} \iint w_n(y) t_n(y) (v_n(x) - v_n(y)) K \left(\frac{x-y}{c_n} \right) u_n(x) f_{X_0}(x) dx dy \right| \\
&= O(c_n^2)
\end{aligned}$$

sowie analog dazu

$$\begin{aligned}
&\sup_{s \in [0,1]} \left| \frac{1}{nc_n} \sum_{j=1}^{\lfloor ns \rfloor} \frac{w_n(X_{j-1}) t_n(X_{j-1})}{f_{X_0}(X_{j-1})} \int (v_n(x) - v_n(X_{j-1})) K \left(\frac{x - X_{j-1}}{c_n} \right) u_n(x) f_{X_0}(x) dx \right| \\
&\leq \sup_{y \in I_n} |t_n(y)| \frac{1}{\inf_{y \in I_n} f_{X_0}(y)} \sup_{y \in I_n} \left| \int (v_n(x) - v_n(y)) K \left(\frac{x-y}{c_n} \right) u_n(x) f_{X_0}(x) dx \right| \\
&\leq \sup_{y \in I_n} |t_n(y)| \frac{1}{\inf_{y \in I_n} f_{X_0}(y)} \\
&\quad \cdot \left(\sup_{y \in I_n} |u_n(y)| \sup_{y \in \mathbb{R}} |f_{X_0}(y)| \sup_{y \in I_n} |v_n'(y)| c_n \int z K(z) dz \right. \\
&\quad \left. + \sup_{y \in I_n} |u_n(y)| \sup_{y \in \mathbb{R}} |f_{X_0}(y)| \sup_{y \in [a_n - Cc_n, b_n + Cc_n]} |v_n''(y)| c_n^2 \int z^2 K(z) dz \right)
\end{aligned}$$

$$\begin{aligned}
& + \sup_{y \in [a_n - Cc_n, b_n + Cc_n]} |(u_n f_{X_0})'(y)| \sup_{y \in [a_n - Cc_n, b_n + Cc_n]} |v'_n(y)| c_n^2 \int z^2 K(z) dz \\
= & O\left(c_n^2 q_n^f\right).
\end{aligned}$$

Mit diesen Ergebnissen folgt aus (10.4.3)

$$\begin{aligned}
& \sup_{s \in [0,1]} \left| \frac{1}{n} \sum_{i=1}^n \frac{1}{n} \sum_{j=1}^{\lfloor ns \rfloor} w_{nj} t_n(X_{j-1}) \left(\frac{\frac{1}{c_n} K\left(\frac{X_{j-1} - X_{i-1}}{c_n}\right) (v_n(X_{i-1}) - v_n(X_{j-1})) u_n(X_{i-1})}{f_{X_0}(X_{j-1})} \right) \right| \\
= & \sup_{s \in [0,1]} \left| \frac{1}{n} \sum_{j=1}^{\lfloor ns \rfloor} \hat{d}_{En}(X_{j-1}) \right| + O(c_n^2 q_n^f). \tag{10.4.4}
\end{aligned}$$

Auch hier gehört \hat{d}_{En} zu einer Funktionenklasse, was im Folgenden gezeigt wird.

Für die Funktion \hat{d}_{En} gilt wie im Beweis zu Lemma 10.2 und Lemma 10.3

$$\sup_{x \in \mathbb{R}} |\hat{d}_{En}(x)| \leq 2 \sup_{x \in I_n} |\hat{d}_n(x)|$$

und für die Ableitungen

$$\sup_{x \in \mathbb{R}} |\hat{d}'_{En}(x)| = \sup_{x \in I_n} |\hat{d}'_n(x)| \quad \text{und} \quad \sup_{x \in \mathbb{R}} |\hat{d}''_{En}(x)| = \sup_{x \in I_n} |\hat{d}''_n(x)|.$$

Weiter gilt

$$\begin{aligned}
\hat{d}_n(x) = & \frac{w_n(x) t_n(x)}{f_{X_0}(x)} \left(\underbrace{\frac{1}{nc_n} \sum_{i=1}^n K\left(\frac{x - X_{i-1}}{c_n}\right) v_n(X_{i-1}) u_n(X_{i-1}) - E[B_1]}_{=: B_1} \right) \\
& - \frac{w_n(x) t_n(x) v_n(x)}{f_{X_0}(x)} \left(\underbrace{\frac{1}{nc_n} \sum_{i=1}^n K\left(\frac{x - X_{i-1}}{c_n}\right) u_n(X_{i-1}) - E[B_2]}_{=: B_2} \right).
\end{aligned}$$

Es ist nach Voraussetzung $\sup_{x \in [a_n - Cc_n, b_n + Cc_n]} |v_n^{(\mu)}(x)| = O(1)$, sowie

$\sup_{x \in [a_n - Cc_n, b_n + Cc_n]} |u_n^{(\mu)}(x)| = O(1)$, $\mu = 0, 1, 2, 3$, daher erfüllen die Funktionen $v_n u_n$ und u_n die Bedingungen an die Funktionen g_n in Lemma 10.1.

Damit folgt, da $\sup_{x \in I_n} |t_n^{(\mu)}(x)| = O(1)$, $\sup_{x \in I_n} |w_n^{(\mu)}(x)| = O(1)$, $\sup_{x \in I_n} |f_{X_0}^{(\mu)}(x)| = O(1)$, $\mu = 0, 1, 2$, und $\frac{1}{\inf_{x \in I_n} f_{X_0}(x)} = O(q_n^f)$,

$$\begin{aligned}
\sup_{x \in I_n} |\hat{d}_n(x)| & = O_P \left(c_n^{-\frac{1}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} q_n^f \right) = o_P(1) \\
\sup_{x \in I_n} |\hat{d}'_n(x)| & = O_P \left(c_n^{-\frac{3}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} (q_n^f)^2 \right) = o_P(1)
\end{aligned}$$

$$\sup_{x \in I_n} |\hat{d}_n''(x)| = O_P \left(c_n^{-\frac{5}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} (q_n^f)^3 \right),$$

da $c_n^{-\frac{3}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} (q_n^f)^2 = (nc_n^3)^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} (q_n^f)^2 = o(1)$ nach Voraussetzung.

Für alle δ gilt außerdem

$$\begin{aligned} \sup_{x, y \in J_n, x \neq y} \frac{|\hat{d}_n'(x) - \hat{d}_n'(y)|}{|x - y|^\delta} &\leq \sup_{x, y \in J_n, |x - y| > c_n} \frac{|\hat{d}_n'(x) - \hat{d}_n'(y)|}{|x - y|^\delta} + \sup_{x, y \in J_n, 0 < |x - y| \leq c_n} \frac{|\hat{d}_n'(x) - \hat{d}_n'(y)|}{|x - y|^\delta} \\ &\leq 2 \sup_{x \in J_n} |\hat{d}_n'(x)| c_n^{-\delta} + \sup_{x \in J_n} |\hat{d}_n''(x)| c_n^{1-\delta} \\ &= O_P \left(c_n^{-\frac{3}{2}-\delta} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} (q_n^f)^3 \right). \end{aligned}$$

Somit gilt für alle $\delta > 0$ mit $c_n^{-\frac{3}{2}-\delta} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} (q_n^f)^3 = o(1)$

$$P \left(\hat{d}_{En} \notin C_1^{1+\delta}(I_n)^n \right) = o(1)$$

für die Funktionenklasse

$$C_1^{1+\delta}(I_n)^n := \left\{ d \in C_1^{1+\delta}(I_n), \sup_{x \in I_n} |d(x)| \leq z_n, \int d(y) f_{X_0}(y) dy = 0 \right\},$$

wobei $z_n := c_n^{-\frac{1}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} q_n^f \log(n)^{\frac{1}{2}}$ und

$$C_1^{1+\delta}(I_n) := \left\{ d : I_n \rightarrow \mathbb{R} \left| \max \left(\sup_{x \in I_n} |d(x)|, \sup_{x \in I_n} |d'(x)| \right) + \sup_{x, y \in I_n, x \neq y} \frac{|d'(x) - d'(y)|}{|x - y|^\delta} \leq 1 \right. \right\}$$

sei.

Somit kann die Konvergenzrate von $\sup_{s \in [0, 1]} \left| \frac{1}{n} \sum_{j=1}^{\lfloor ns \rfloor} \hat{d}_{En}(X_{j-1}) \right|$ durch die Konvergenzrate von

$$\sup_{s \in [0, 1]} \sup_{d \in C_1^{1+\delta}(I_n)^n} \left| \frac{1}{n} \sum_{j=1}^{\lfloor ns \rfloor} d(X_{j-1}) \right| \tag{10.4.5}$$

abgeschätzt werden.

Wie im Beweis zu Lemma 10.2 folgt

$$\sup_{s \in [0, 1]} \sup_{d \in C_1^{1+\delta}(I_n)^n} \left| \frac{1}{n} \sum_{j=1}^{\lfloor ns \rfloor} d(X_{j-1}) \right| = \max_{1 \leq h \leq M_{n1}} \max_{1 \leq k \leq M_{n2}} \left| \frac{1}{n} \sum_{j=1}^{\lfloor ns \rfloor} d_k(X_{j-1}) \right| + O \left(\bar{\epsilon}_n + z_n \frac{1}{n} \right), \tag{10.4.6}$$

wobei M_{n1} und M_{n2} wie im Beweis zu Lemma 10.2 definiert sind.

Sei $\bar{\epsilon}_n = c_n^2 q_n^f$. Dann ist $z_n \frac{1}{n} = O(\bar{\epsilon}_n)$, da $n^{-\frac{3}{2}} c_n^{-\frac{1}{2}} \log(n) q_n^f = c_n^2 q_n^f \cdot c_n^2 (nc_n^3)^{-\frac{3}{2}} \log(n) = \bar{\epsilon}_n \cdot c_n^2 \left((nc_n^3)^{-1} \log(n)^{\frac{2}{3}} \right)^{\frac{3}{2}} = \bar{\epsilon}_n \cdot o(1) \cdot o(1)$ nach Voraussetzung. Damit kann in (10.4.6) $O(\bar{\epsilon}_n + z_n \frac{1}{n})$ durch $O(\bar{\epsilon}_n)$ ersetzt werden.

Mit Hilfssatz 10.2.A und weiter analog zum Vorgehen im Beweis zu Lemma 10.2 kann auf Grund der Annahmen an die Bandbreite c_n gezeigt werden, dass

$$\max_{1 \leq h \leq M_{n1}} \max_{1 \leq k \leq M_{n2}} \left| \frac{1}{n} \sum_{j=1}^{\lfloor ns_h \rfloor} d_k(X_{j-1}) \right| = O_P(\bar{\epsilon}_n).$$

Aus dieser Gleichung und aus (10.4.2), (10.4.4), (10.4.5) und (10.4.6) folgt

$$\begin{aligned} & \sup_{s \in [0,1]} \left| \frac{\sum_{k=1}^{\lfloor ns \rfloor} w_{nk}}{n} \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^{\lfloor ns \rfloor} \bar{w}_{j \lfloor ns \rfloor} t_n(X_{j-1}) \right. \\ & \qquad \qquad \qquad \left. \cdot \left(\frac{\frac{1}{c_n} K \left(\frac{X_{j-1} - X_{i-1}}{c_n} \right) (v_n(X_{i-1}) - v_n(X_{j-1})) u_n(X_{i-1})}{\frac{1}{nc_n} \sum_{k=1}^n K \left(\frac{X_{j-1} - X_{k-1}}{c_n} \right)} \right) \right| \\ &= O_P \left(\left(c_n^{\frac{1}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} + c_n^3 \right) q_n^f \right) + O_P(c_n^2 q_n^f) \\ &= O_P \left(\left(c_n^{\frac{1}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}} + c_n^2 \right) q_n^f \right). \end{aligned}$$

□

11 Ausblick

In diesem Kapitel wird ein Ausblick über mögliche weiterführende Forschung auf dem Gebiet dieser Arbeit gegeben.

Zum einen ist sicher von Interesse, die Ergebnisse dieser Arbeit auf autoregressive Modelle der Form

$$X_j = m(X_{j-1}, \dots, X_{j-d}) + \sigma(X_{j-1}, \dots, X_{j-d})\epsilon_j$$

für $d > 1$ auszuweiten.

Das Problem, das dabei auftritt, ist, dass die Konvergenzraten von $|\hat{m} - m|$ und $|\hat{\sigma} - \sigma|$ von d abhängen. Diese Konvergenzraten sind in Lemma 3.4 (S. 35) angegeben und basieren auf der Konvergenzrate in Lemma 3.3 (S. 21). Diese Konvergenzrate ist (für $d = 1$) $O_P\left(c_n^{-\frac{1}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}}\right)$ (siehe Lemma 3.3) und für $d \geq 1$

$$O_P\left(c_n^{-\frac{d}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}}\right), \quad (11.0.1)$$

was in Hansen [2008] für stationäre Modelle gezeigt wird.

Im Beweis zu Satz 4.2 (S. 49) wird benötigt, dass $|\hat{m} - m|^2$ und $|\hat{\sigma} - \sigma|^2$ von der Rate $o_p\left(\frac{1}{\sqrt{n}}\right)$ sind, und dass $c_n^2 = o(n^{-\frac{1}{2}})$. Mit (11.0.1) bedeutet das also, dass gelten muss

$$\left(c_n^{-\frac{d}{2}} n^{-\frac{1}{2}} \log(n)^{\frac{1}{2}}\right)^2 = c_n^{-d} n^{-1} \log(n) = o(n^{-\frac{1}{2}}),$$

also $nc_n^{2d} \xrightarrow{n \rightarrow \infty} \infty$, und $nc_n^4 = o(1)$, was für $d \geq 2$ zu einem Widerspruch führt.

Dieses Problem wurde im Fall unabhängiger Beobachtungen in Neumeyer and van Keilegom [2010] dadurch gelöst, dass ein lokal polynomieller Schätzer vom Grad p anstatt eines Nadaraya-Watson-Schätzers verwendet wurde. Mit einem lokal polynomiellen Schätzer vom Grad p verändert sich die Bedingung $nc_n^4 = o(1)$ zu $nc_n^{2p+2} = o(1)$, so dass sie mit geeignetem p nicht mehr im Widerspruch zu $nc_n^{2d} \xrightarrow{n \rightarrow \infty} \infty$ steht.

Eine weitere Möglichkeit das Modell dieser Arbeit zu erweitern, die zu dem selben Problem führt, ist Modelle der Form

$$X_j = m(X_{j-1}, Z_j) + \sigma(X_{j-1}, Z_j)\epsilon_j$$

zu betrachten, mit (möglicherweise mehrdimensionalen) Kovariablen (Z_j) .

Eine zweite interessante Erweiterung der Ergebnisse dieser Arbeit, ist die Untersuchung des sogenannten "Online-Monitoring" für den Change-Point-Test.

Dieses Verfahren basiert darauf, dass nicht auf Basis von n Beobachtungen getestet wird, ob in diesen ein Change-Point vorliegt, sondern ab einem bestimmten Zeitpunkt für jede neu hinzukommende Beobachtung getestet wird, ob die Nullhypothese keines Change-Points noch gilt. Für parametrische Modelle wird ein solches Verfahren in Hlávka et al. [2010] vorgestellt (siehe Abschnitt 5.5, S. 81).

Analog zu Hlávka et al. [2010] ist die Idee für ein solches “Online-Monitoring” die folgende: Es wird davon ausgegangen, dass ein (historisches) Trainingsdatenset zur Verfügung steht, in dem kein Change-Point vorliegt. Auf Basis dieses Trainingsdatensets werden die unbekannte Regressions- und Skalierungsfunktion geschätzt. Für jede neu hinzukommende Beobachtung wird dann getestet, ob sich die Verteilung der “neuen” Innovationen von der der Innovationen aus dem Trainingsdatenset unterscheidet, indem für jedes $k = 1, 2, \dots$ die Differenz

$$\sup_{t \in \mathbb{R}} \left| \sum_{j=1}^T \bar{w}'_{jT} I\{\hat{\epsilon}_j \leq t\} - \sum_{j=T+1}^{T+k} \bar{w}'_{jk} I\{\hat{\epsilon}_j \leq t\} \right| \quad (11.0.2)$$

betrachtet wird, wobei X_1, \dots, X_T das Trainingsdatenset bezeichnet und $\bar{w}'_{jT}, \bar{w}'_{jk}$ geeignete Gewichte. Dieses Verfahren wird so lange durchgeführt, bis die Nullhypothese abgelehnt wird, also bis zu dem k , bei dem die auf (11.0.2) basierende Teststatistik den kritischen Wert des Tests zum ersten Mal überschreitet.

Ein drittes interessantes Gebiet für die weiterführende Forschung ist die Untersuchung von weiteren Hypothesentests für autoregressive Modelle. In Kapitel 7 (S. 104) wird hierfür bereits ein Einblick gegeben.

Zusätzliche interessante Hypothesentests sind z.B. Tests zur Überprüfung der Modellannahmen, wie Tests auf Unabhängigkeit der Innovationen $\epsilon_1, \dots, \epsilon_n$ oder Tests auf Unabhängigkeit der Innovationen und der vorherigen Beobachtungen, ϵ_j und X_{j-1}, X_{j-2}, \dots

Eine weitere mögliche Ausweitung des Modells besteht darin, mehrdimensionale Zeitreihen zu betrachten, insbesondere mehrdimensionale Innovationen zuzulassen. Solche Modelle treten z.B. auf, wenn mehrere Finanzzeitreihen gemeinsam modelliert werden, um die Abhängigkeit von Ausfallrisiken zu untersuchen.

Ein Hypothesentest könnte in diesem Fall die Komponenten der Innovationen auf Unabhängigkeit testen, oder die Verteilungen der Komponenten der Innovationen auf Gleichheit.

Wie bereits in Kapitel 7 ausgeführt, ist für viele Hypothesentests die Verteilung der Teststatistik unbekannt, und daher die Anwendung von Bootstrap-Verfahren von Interesse (siehe Abschnitt 7.4, S. 109).

Somit ist die Untersuchung von Bootstrap-Verfahren für nichtparametrische Autoregressionsmodelle in den unterschiedlichen Testszenarien ebenfalls ein interessanter Punkt für die weiterführende Forschung.

Literaturverzeichnis

- M. G. Akritas and I. van Keilegom. Non-parametric Estimation of the Residual Distribution. *Scand. J. Statist.* 28, 549-567, 2001.
- E. Andreou and E. Ghysels. Structural Breaks in Financial Time Series. *T. G. Andersen (ed) et al, Handbook of Financial Time Series. Springer, 839-870, 2009.*
- A. Azzalini and A. Capitanio. Statistical applications of the multivariate skew-normal distribution. *J. Roy. Statist. Soc. Ser. B* 61, 579-602, 1999.
- J. Bai. Weak convergence of the sequential empirical processes of residuals in ARMA models. *Ann. Statist.* 22, 2051-2061, 1994.
- R. N. Bhattacharya and C. Lee. Ergodicity of Nonlinear First Order Autoregressive Models. *J. Theoret. Probab.* 8, 207-219, 1995.
- P. J. Bickel and M. J. Wichura. Convergence criteria for multiparameter stochastic processes and some applications. *Ann. Math. Statist.* 42, 1656-1670, 1971.
- P. Billingsley. *Convergence of Probability Measures.* John Wiley & Sons, 1968.
- T. Bollerslev. A Conditionally Heteroskedastic Time Series Model for Speculative Prices and Rates of Return. *The Review of Economics and Statistics* 69, 542-547, 1987.
- M. Csörgö, L. Horváth, and B. Szyszkowicz. Integral tests for suprema of Kiefer processes with application. *Statist. Decisions* 15, 365-377, 1997.
- H. Dette, N. Neumeyer, and I. van Keilegom. A new test for the parametric form of the variance function in nonparametric regression. *J. Roy. Statist. Soc. Ser. B* 69, 903-917, 2007.
- H. Dette, J. C. Pardo-Fernández, and I. van Keilegom. Goodness-of-Fit Tests for Multiplicative Models with Dependent Data. *Scand. J. Statist.* 36, 782-799, 2009.
- P. Doukhan. *Mixing, Properties and Examples.* Springer, 1994.
- J. Fan and Q. Yao. *Nonlinear Time Series.* Springer, 2005.
- J. Franke, J.-P. Kreiß, and E. Mammen. Bootstrap of kernel smoothing in nonlinear time series. *Bernoulli* 8, 1-37, 2002.
- D. A. Freedman. On tail probabilities for martingals. *Ann. Probab.* 3, 100-118, 1975.

- B. E. Hansen. Uniform Convergence Rates for Kernel Estimation with Dependent Data. *Econom. Theory* 24, 726-748, 2008.
- H. Heuser. *Lehrbuch der Analysis, Teil 1*. B. G. Teubner, 1988.
- Z. Hlávka, M. Hušková, C. Kirch, and S. G. Meintanis. Monitoring changes in the error distribution of autoregressive models based on fourier methods. *preprint* (http://mspcdip.mathematik.uni-karlsruhe.de/~ckirch/papers/pp_hhkm_06_12_10.pdf), 2010.
- M. Hušková and J. Antoch. Detection of structural changes in regression. *Tatra Mt. Math. Publ.* 26, 201-215, 2003.
- C. Kirch and J. Tadjuidje Kamgaing. Testing for parameter stability in nonlinear autoregressive models. *preprint* (http://mspcdip.mathematik.uni-karlsruhe.de/~ckirch/papers/pp_CPA_NN_offline.pdf), 2011.
- H. L. Koul. Asymptotics of some estimators and sequential residual empiricals in nonlinear time series. *Ann. Statist.* 24, 380-404, 1996.
- J.-P. Kreiß and G. Neuhaus. *Einführung in die Zeitreihenanalyse*. Springer, 2006.
- J.-P. Kreiß, M. H. Neumann, and Q. Yao. Bootstrap tests for simple structures in nonparametric time series regression. *Statistics And Its Interface* 1, 367-380, 2008.
- E. Liebscher. Strong convergence of sums of α -mixing random variables with applications to density estimation. *Stochastic Processes and their Applications* 65, 69-80, 1996.
- S. Ling. Weak convergence of the sequential empirical processes of residuals in nonstationary autoregressive models. *Ann. Statist.* 26, 741-754, 1998.
- U. U. Müller, A. Schick, and W. Wefelmeyer. Estimating the innovation distribution in nonparametric autoregression. *Probab. Theory Relat. Fields* 144, 53-77, 2009.
- M. H. Neumann and J.-P. Kreiß. Regression-Type Inference in Nonparametric Autoregression. *Ann. Statist.* 26, 1570-1613, 1998.
- M. H. Neumann and E. Paparoditis. Goodness-of-fit tests for markovian time series models: Central limit theory and bootstrap approximations. *Bernoulli* 14, 14-46, 2008.
- N. Neumeyer and H. Dette. Testing for symmetric error distribution in nonparametric regression models. *Statist. Sinica* 17, 775-795, 2007.
- N. Neumeyer and I. van Keilegom. Change-Point Tests for the Error Distribution in Non-parametric Regression. *Scand. J. Statist.* 36, 518-541, 2009.
- N. Neumeyer and I. van Keilegom. Estimating the error distribution in nonparametric multiple regression with applications to model testing. *J. Multivar. Anal.* 101, 1067-1078, 2010.
- N. Neumeyer, H. Dette, and E.-R. Nagel. Bootstrap tests for the error distribution in linear and nonparametric regression models. *Austr. New Zealand J. Statist.* 48, 129-156, 2006.

- E. S. Page. Continuous Inspection Schemes. *Biometrika* 41, 100-115, 1954.
- D. Picard. Testing and estimating change-points in time series. *Adv. Appl. Probab.* 17, 841-867, 1985.
- D. Pollard. *Empirical Processes: Theory and Applications*. NSF-CBMS Regional Conference Series in Probability and Statistics 2, Institute of Mathematical Statistics, 1990.
- E. Rio. The functional law of the iterated logarithm for stationary strongly mixing sequences. *Ann. Probab.* 23, 1188-1203, 1995.
- G. R. Shorack and J. A. Wellner. *Empirical processes with applications to statistics*. John Wiley & Sons, 1986.
- M. A. Stephens. Asymptotic Results for Goodness-of-Fit Statistics with Unknown Parameters. *Ann. Statist.* 4, 357-369, 1976.
- A. W. van der Vaart. *Asymptotic Statistics*. Cambridge University Press, 2000.
- A. W. van der Vaart and J. A. Wellner. *Weak Convergence and Empirical Processes*. Springer, 1996.
- I. van Keilegom, W. González Manteiga, and C. Sánchez Sellero. Goodness-of-fit tests in parametric regression based on the estimation of the error distribution. *Test* 17, 401-415, 2008.

Zusammenfassung

In dieser Arbeit wird das asymptotische Verhalten von sequentiellen empirischen Prozessen basierend auf nichtparametrisch geschätzten Residuen in heteroskedastischen autoregressiven Modellen untersucht und ein Test auf einen Change-Point (Strukturbruch) in der Verteilung der Innovationen hergeleitet.

Das betrachtete Modell ist eine Zeitreihe

$$X_j = m(X_{j-1}) + \sigma(X_{j-1}) \cdot \epsilon_j, \quad j \in \mathbb{Z}$$

mit unabhängigen Innovationen ϵ_j und geeignet glatten Funktionen m und σ .

Die Innovationen sind nicht beobachtbar und werden daher durch Residuen

$$\hat{\epsilon}_j = \frac{X_j - \hat{m}(X_{j-1})}{\hat{\sigma}(X_{j-1})}$$

mit nichtparametrischen Nadaraya-Watson-Schätzern \hat{m} und $\hat{\sigma}$ geschätzt. Basierend auf diesen Residuen werden sequentielle empirische Prozesse $\sum_{j=1}^{\lfloor ns \rfloor} w_j I\{\hat{\epsilon}_j \leq t\}$ mit $t \in \mathbb{R}$, $s \in [0, 1]$ und einer Gewichtsfunktion w_j , die den Prozess auf $[0, 1]$ normiert, gebildet.

Das Haupttheorem dieser Arbeit (Satz 4.2) gibt eine stochastische Entwicklung für den sequentiellen empirischen Prozess der nichtparametrisch geschätzten Residuen. Diese stochastische Entwicklung besteht aus einer Summe unabhängiger Zufallsvariablen und einem Fehlerterm, der mit schneller Rate gegen Null konvergiert. Daher können bekannte Resultate für unabhängige Zufallsvariablen verwendet werden, um schwache Konvergenz der Differenz zwischen sequentiellen empirischen Prozessen und der wahren Verteilungsfunktion gegen einen Gauß-Prozess zu beweisen. Dieses Ergebnis beinhaltet die schwache Konvergenz der Differenz zwischen nichtsequentiellen empirischen Prozessen und der wahren Verteilungsfunktion, daher können aus ihm, neben Change-Point-Tests, auch Goodness-of-fit-Tests abgeleitet werden.

Der beschriebene Ansatz ist vergleichbar zur Herleitung nichtparametrischer Change-Point-Tests, die für Regressionsmodelle existieren. Im autoregressiven Fall treten allerdings zwei zusätzliche Schwierigkeiten auf. Die Beobachtungen sind abhängig und der Träger des Prozesses kann nicht als kompakt angenommen werden, im Gegensatz zum Träger der Kovariablen in der Regression. Diese Probleme werden gelöst, indem die Annahme der Unabhängigkeit durch eine Mischungsbedingung ersetzt wird, und stochastische Gewichte verwendet werden, die den Wertebereich der Beobachtungen auf kompakte Intervalle beschränken, die gegen die reelle Zahlengerade konvergieren.

Ein Change-Point-Test für die Innovationenverteilung ist ein Test für die Nullhypothese keines Change-Points, was bedeutet dass die Innovationen $\epsilon_1, \dots, \epsilon_n$ i.i.d. sind, gegen die Alternative

$$\exists s_0 \in (0, 1) : \epsilon_1, \dots, \epsilon_{\lfloor ns_0 \rfloor} \text{ i.i.d. } \sim F, \quad \epsilon_{\lfloor ns_0 \rfloor + 1}, \dots, \epsilon_n \text{ i.i.d. } \sim G, \quad F \neq G.$$

Teststatistiken für solche Tests basieren auf der Differenz von geeignet angepassten sequentiellen empirischen Prozessen nichtparametrisch geschätzter Residuen. Schwache Konvergenz dieser Differenz kann aus der erwähnten stochastischen Entwicklung abgeleitet werden und es wird gezeigt, dass die selbe asymptotische Darstellung wie im Regressionsmodell gilt. Insbesondere ist die betrachtete Teststatistik verteilungsfrei und bekannte asymptotische Quantile können für den Test benutzt werden. Simulationen mehrerer Arten von Change-Points und verschiedener Modelle unterstützen dieses theoretische Resultat.

Danksagung

An dieser Stelle möchte ich den Menschen danken, die mich beim Schreiben dieser Arbeit unterstützt haben.

Zuallererst danke ich Prof. Dr. Natalie Neumeyer für die Anregung zu dieser Arbeit und die gute Betreuung. Auch in stressigen Zeiten und selbst während ihrer Elternzeit fand sie immer Zeit mir zuzuhören und zu helfen. Dafür danke ich ihr sehr.

Allen Mitarbeitern des Bereichs Mathematische Statistik und Stochastische Prozesse der Universität Hamburg danke ich für ihre Hilfsbereitschaft und die vielen anregenden Diskussionen. Besonders möchte ich an dieser Stelle meine Kollegin Jennifer Mylosz nennen, die immer ein offenes Ohr für mich hatte und mir eine wahre Freundin geworden ist.

Schließlich gilt mein ganz besonderer Dank meiner Familie für ihre große Unterstützung und Ermutigung. Meinen Eltern Gabriela und Dieter Stahl danke ich ganz herzlich dafür, dass sie mir ermöglicht haben zu studieren und mich immer mit Rat und Tat unterstützten. Meinem Mann Mario Selk danke ich für seinen unermüdlichen Zuspruch. Trotz der für ihn damit verbundenen Entbehrungen hat er mich stets darin bestärkt, meine Ziele zu verfolgen. Dafür danke ich ihm von ganzem Herzen.

Lebenslauf

Name:	Leonie Selk geb. Stahl
geboren:	20. Januar 1985 in Hamburg
1991-1995:	Besuch der Grundschule Scheeßeler Kehre, Hamburg
1995-2003:	Besuch des Alexander-von-Humboldt Gymnasiums, Hamburg
Juni 2003:	Schulabschluss Abitur
2003-2007:	Diplomstudium in Mathematik an der Universität Hamburg
September 2007:	Studienabschluss Diplom in Mathematik
2007-2011:	Promotionsstudium in Mathematik an der Universität Hamburg
2008-heute:	wissenschaftliche Mitarbeiterin im Department Mathematik der Universität Hamburg