Zur Schätzung von Varianzen in der Changepoint-Analyse

Diplomarbeit

vorgelegt von Markus Riedle

angefertigt bei Professor Dr. Josef Steinebach

Fachbereich Mathematik

und Informatik

Philipps-Universität Marburg

Januar 2000

Inhaltsverzeichnis

0	Einl	eitung und Zusammenfassung	2
1	Loka	ationsmodell mit Invarianzprinzipien	5
	1.1	Das Modell	5
	1.2	Der Standard-Schätzer	8
	1.3	Die kombinierten Schätzer	13
	1.4	Der 2-stufige Schätzer	21
2	Lineares Modell		38
	2.1	Das Modell	38
	2.2	Die kombinierten Schätzer	47
	2.3	Der Minimumschätzer	53
	2.4	Der Alternativschätzer	64
3	Simulationen		69
	3.1	Grundlagen der Simulationen	69
	3.2	Erwartungstreue einiger Varianzschätzer	70
	3.3	Verteilungsparameter der Kleinste-Quadrate-Schätzer	71
	3.4	Simulation mit normalverteilten Fehlern und äquidistanten Beobachtungen .	73
	3.5	Simulation mit normalverteilten Fehlern und nicht äquidistanten Beobachtungen	82
	3.6	Simulationen bei verschiedenen Stichprobenumfängen	84
	3.7	Simulation mit exponentialverteilten Fehlern	88
	3.8	Die Varianzschätzer im Vergleich	90
A	Anh	ang	94
Li	Literaturverzeichnis		

Einleitung und Zusammenfassung 0

Die Changepoint-Analyse entstand ursprünglich aus Problemstellungen der Qualitätskontrolle. Im Laufe ihrer Entwicklung traten immer weitere Anwendungen hinzu, so dass sie sich heute über viele Bereiche der Statistik erstreckt, wie Regressionsanalyse, Zeitreihenanalyse und Statistik stochastischer Prozesse. Allen Anwendungen gemeinsam ist die Frage nach Homogenität der akquirierten¹ Daten, also ob diesen Daten strukturelle Änderungen zu unbekannten Zeitpunkten unterliegen. Als Beispiel dienen soll eine Folge $X_1, \ldots, X_n, n \in \mathbb{N}$. von Zufallsvariablen auf einem Wahrscheinlichkeitsraum (Ω, \mathcal{A}, P) , deren Realisationen etwa die chronologischen, diskreten Ergebnisse einer physikalischen, medizinischen oder sonstigen Untersuchung repräsentieren. So kann die konstante Energiezufuhr in einer physikalischen Messung ab den Zeitpunkten t_1, t_2, \ldots für ein verändertes Verhalten von Teilchen verantwortlich sein oder das Uberschreiten der Inkubationszeit t_0 ein verändertes Krankheitsbild verursachen. Beschränkt man sich auf das Vorliegen einer einzigen solchen Änderung, dann ist dies die Frage nach einem Zeitpunkt $t = \theta n, \ \theta \in (0,1)$, dem sogenannten Changepoint, so dass sich die gemeinsame Verteilung der ersten [t] Beobachtungen² $X_1, \ldots, X_{[t]}$ von der Verteilung der letzten n - [t] Beobachtungen $X_{[t]+1}, \ldots, X_n$ unterscheidet. Die stochastische Präzisierung, in welchem Sinn sich diese Verteilungen unterscheiden, hängt von der Modellannahme der statistischen Untersuchung ab.

Eine Entscheidungsfunktion S_n , die in Abhängigkeit von der Anzahl n der Beobachtungen definiert ist, soll aufgrund der Daten X_1, \ldots, X_n Aufschluss darüber geben, ob

- eine homogene Verteilung vorliegt (Nullhypothese),
- ein Changepoint existiert (Alternative).

Vielen Verfahren der Changepoint-Analyse entsprechend wird im Folgenden davon ausgegangen, dass niedrige Werte der Funktion S_n für eine Bestätigung der Nullhypothese und hohe Werte für ein Verwerfen der Nullhypothese sprechen.

Die Beobachtungen X_1, \ldots, X_n unterliegen Zufallsschwankungen, die sich auf die Entscheidungsfunktion S_n übertragen. Um unabhängig von der Größenordnung dieser Schwankungen einen Bereich bestimmen zu können, in dem mit hoher Wahrscheinlichkeit die Werte einer Statistik bei homogener Verteilung der Daten liegen, ist eine geeignete Normierung der Funktion S_n erforderlich. Geht man von einer identischen Varianz σ^2 der Daten X_1, \ldots, X_n aus, so ist diese ein geeignetes Maß für die zufälligen Schwankungen und man erhält als normierte Statistik $S_n/g(\sigma^2)$, wobei $g: \mathbb{R}^*_+ \to \mathbb{R}^*_+$ eine passende, monoton wachsende Funktion sei.

Unter der Nullhypothese lässt sich für die Statistik $S_n/q(\sigma^2)$ im Allgemeinen keine Verteilung angeben. Um dennoch eine Entscheidung bei zumindest asymptotischer Einhaltung eines vorgegebenes Fehlerniveaus treffen zu können, ist die Kenntnis der Grenzverteilung der Statistik $S_n/g(\sigma^2)$ für $n \to \infty$ erforderlich. Die Nachweise solcher Grenzverteilungen basieren oft auf Verwendung schwacher und starker Invarianzprinzipien, die auf die sogenannte Ungarische Schule³ zurückgehen. Auch bei Benutzung der Invarianzprinzipien ergibt sich die Notwendigkeit einer geeigneten Normierung $1/q(\sigma^2)$ der Entscheidungsfunktion S_n .

¹Diese Arbeit beschränkt sich auf die sogenannte a posteriori Changepoint-Analyse im Gegensatz zu der sequentiellen Changepoint-Analyse.

²[t] sei der ganzzahlige Anteil von t. ³Die Anfänge sind in [KoMaTu75], [KoMaTu76] und [CsRe75a], [CsRe75b] zu finden.

Im Allgemeinen ist die Varianz σ^2 unbekannt, so dass aufgrund der beobachteten Daten X_1, \ldots, X_n diese mittels einer Schätzfunktion $\hat{\sigma}_n^2$ geschätzt werden muss. Als Entscheidungsfunktion ergibt sich $S_n/g(\hat{\sigma}_n^2)$, wobei zwei Forderungen zu stellen sind:

1. Unter der Nullhypothese sollen die geschätzten Werte $\hat{\sigma}_n^2$ möglichst gut dem wahren Wert σ^2 entsprechen, damit sich das asymptotische Verhalten der Statistik $S_n/g(\sigma^2)$ bei Ersetzen des unbekannten Parameters σ^2 durch die Schätzung $\hat{\sigma}_n^2$ nicht ändert. Mathematisch ist dafür meistens die Forderung nach konsistenter Schätzung unter der Annahme des Vorliegens der Nullhypothese ausreichend:

$$\left|\sigma^2 - \hat{\sigma}_n^2\right| \xrightarrow{P} 0, \quad n \to \infty.$$
 (0.1)

2. Unter der Alternative erwartet man von einer guten Entscheidungsfunktion S_n , dass diese bei größer werdender Beobachtungsanzahl n wachsende Werte mit hoher Wahrscheinlichkeit annimmt, um so für eine deutlichere Verwerfung der Nullhypothese zu sprechen, also $S_n \xrightarrow{P} \infty$ für $n \to \infty$. Das Ersetzen des unbekannten Parameters σ^2 durch die Schätzung $\hat{\sigma}_n^2$ birgt unter der Alternative die Gefahr, dass $g(\hat{\sigma}_n^2)$ ebenfalls wie die Funktion S_n wachsende Werte bei größer werdender Beobachtungsanzahl n annimmt. Kompensieren sich diese beiden Konvergenzen, so nimmt die Statistik $S_n/g(\hat{\sigma}_n^2)$ auch unter der Alternative "kleine Werte" an, die für keine Verwerfung der Nullhypothese mehr sprechen. Zur Vermeidung dieses Verhaltens fordert man die asymptotische Konsistenz⁴ der Statistik unter der Annahme des Vorliegens der Alternative:

$$\frac{1}{g(\hat{\sigma}_n^2)} S_n \xrightarrow{P} \infty, \quad n \to \infty.$$
(0.2)

In der Literatur⁵ konzentrieren sich viele Methoden der Varianzschätzung auf das Vorliegen der Nullhypothese. Die Konstruktion der Schätzer geschieht dann unter diesem Aspekt oder es werden Varianzschätzer vorgeschlagen, ohne die oben genannten Forderungen nachzuweisen. Varianzschätzer, die gemäß der Nullhypothese konstruiert sind, erfüllen sicherlich die erste der oben gestellten Forderungen. Unter der Alternative nehmen viele dieser Varianzschätzer in einem stochastischen Sinn größer werdende Werte bei wachsender Beobachtungsgröße nan. Da jedoch die Entscheidungsfunkion S_n unter der Alternative oft mit einer sehr hohen Rate stochastisch gegen Unendlich konvergiert, wird Forderung 2) ebenfalls eingehalten, falls die Varianzschätzer eben nicht "zu schnell" wachsen.

Diese Arbeit zeigt, dass durch das Einbeziehen eines eventuell vorliegenden Changepoints in die Varianzschätzung sich Schätzmethoden entwickeln lassen, ohne die oben gestellten Forderungen zu verletzen. Dabei stellt sich heraus, dass diese Schätzer unter der Alternative oft ein besseres Verhalten als die üblichen Schätzer aufweisen, da sie bei größer werdender Beobachtungsanzahl n zumindest "langsamer" wachsende Werte annehmen oder sogar wie unter der Nullhypothese gemäß (0.1) konsistente Schätzungen liefern. Dadurch ist von Statistiken, in denen der unbekannte Parameter σ^2 durch solche Schätzer ersetzt wird, unter der Alternative

⁴Mit asymptotischer Konsistenz einer Statistik wird in dieser Arbeit das Erfüllen der Forderung (0.2) unter allen zugelassenen Alternativen verstanden, da die Nullhypothesen für große Werte der betrachteten Statistiken verworfen werden.

⁵vgl. [CsHo97].

eine "schnellere" Konvergenz gegen Unendlich im Sinne von (0.2) zu erwarten. Entsprechend wird die Nullhypothese deutlicher verworfen.

Aus der Vielzahl der in der Changepoint-Analyse behandelten Modelle werden zwei herausgegriffen. Bei den formulierten Testproblematiken ergibt sich die Notwendigkeit, die den Modellen zugrunde liegenden Varianzen zu schätzen. Es werden jeweils für die beiden Modelle mehrere Varianzschätzer vorgestellt, die hier neu entwickelt oder in der Literatur vorgeschlagen werden. Das Verhalten der Varianzschätzer wird unter den in den Modellen formulierten Nullhypothesen und Alternativen betrachtet, um das Einhalten der oben gestellten Forderungen nachzuweisen.

Das erste Modell⁶ ist ein sehr allgemeiner Ansatz in der Changepoint-Analyse. Es setzt bestimmte Invarianzprinzipien voraus und fasst damit verschiedene Modelle der Changepoint-Analyse zusammen, in denen jeweils der Nachweis der Grenzverteilung der Statistik gerade auf solchen Invarianzprinzipien beruht.

Das zweite Modell⁷ behandelt eine Changepoint-Problematik in der Regressions-Analyse. Aufgrund der Vorgabe der Punkte, an denen die Beobachtungen gemacht werden, sind viele Zugänge der Changepoint-Analyse für lineare Modelle nicht anwendbar. Deshalb wird zu Beginn des zweiten Kapitels ein Verfahren ausführlich behandelt, das in der Literatur selten untersucht wird. Im letzten Kapitel werden die Schätzer dieses Modells in Simulationen betrachtet.

⁶vorgestellt in [HoSt98].

⁷vorgestellt in [ChWh92].

1 Lokationsmodell mit Invarianzprinzipien

Die Nachweise der asymptotischen Verteilungen von Teststatistiken in der Changepoint-Analyse basieren oft auf Invarianzprinzipien. Dies fassen Horváth und Steinebach [HoSt98] in einem allgemeinen Modell durch die Voraussetzung zusammen, dass die zugrunde liegenden Prozesse bestimmten Invarianzprinzipien genügen. Das Modell sowie die Teststatistik werden in Abschnitt 1.1 vorgestellt. Für die Varianz, mittels der die Teststatistik normiert wird, betrachten wir in Abschnitt 1.2 einen Varianzschätzer¹ sowie dessen Konvergenzraten unter der Nullhypothese und unter der Alternative. In Abschnitt 1.3 wird eine Familie von Schätzfunktionen für die unbekannte Varianz vorgestellt. Es zeigt sich, dass bei Ersetzen der unbekannten Varianz in der Teststatistik durch die Familie der Schätzfunktionen die asymptotische Konsistenz der Teststatistik unter schwächeren Voraussetzungen als bei Benutzung des Schätzers aus Abschnitt 1.2 gewährleistet ist. In Abschnitt 1.4 wird ein Varianzschätzer entwickelt, wobei das Intervall, in dem der eventuell vorhandene Changepoint liegt, geschätzt wird. Anders als in [Kühn99] werden für die Schätzung der Lage des Changepoints nur diskrete Beobachtungen benötigt. Mittels dieser Schätzung wird ein Varianzschätzer definiert und unter der Nullhypothese sowie unter der Alternative betrachtet.

1.1 Das Modell

Wie in der Einleitung dieser Arbeit bereits dargestellt, werden in der Changepoint-Analyse oft Folgen X_1, \ldots, X_N , $N \in \mathbb{N}$, von Zufallsvariablen auf einem Wahrscheinlichkeitsraum (Ω, \mathcal{A}, P) betrachtet, deren Verhalten sich zu einem nicht bekannten Zeitpunkt ändert. So geht man etwa von konstanten Erwartungswerten der X_i vor und nach dem Zeitpunkt T^* der Änderung, dem sogenannten Changepoint, aus, also $EX_1 = \ldots = EX_{[T^*]} \neq EX_{[T^*]+1} \ldots = EX_N$. Teststatistiken, mit denen aufgrund von Beobachtungen eine Entscheidung getroffen wird, ob solch eine Änderung in der Verteilung tatsächlich vorliegt, basieren dabei in vielen Ansätzen auf Schätzern des sich ändernden Parameters.

Das von Horváth und Steinebach [HoSt98] angegebene Modell verallgemeinert diesen Ansatz:

Definition 1.1

Sei $\{Z_T(t) : 0 \leq t \leq T\}$ für festes T > 0 ein reeller stochastischer Prozess auf einem Wahrscheinlichkeitsraum (Ω, \mathcal{A}, P) , der dem folgenden Modell genügt:

$$Z_T(t) := \begin{cases} a_T t + b_T Y_T(t) & : & 0 \le t \le T^*, \\ Z_T(T^*) + a_T^*(t - T^*) + b_T^* Y_T^*(t - T^*) & : & T^* < t \le T, \end{cases}$$

wobei a_T , a_T^* , b_T , b_T^* und $T^* \in [0, T]$ unbekannte reelle Parameter sind. Weiter seien $\{Y_T(t) : 0 \le t \le T^*\}$ und $\{Y_T^*(t) : 0 \le t \le T - T^*\}$ reelle stochastische Prozesse auf (Ω, \mathcal{A}, P) .

Die oben genannte Folge von Zufallsvariablen X_1, \ldots, X_N kann als Beispiel einer Anwendung des Modells der Definition 1.1 dienen. Dazu seien jeweils identische Erwartungswerte und Varianzen der Zufallsvariablen der Folgen $X_1, \ldots, X_{[T^*]}$ und $X_{[T^*]+1}, \ldots, X_N$ vorausgesetzt. Die Parameter a_T , a_T^* und b_T^2 , b_T^{*2} des Modells der Definition 1.1 seien diesen Erwartungswerten und Varianzen gleichgesetzt. Stellen $Y_T(t)$ und $Y_T^*(t)$ die Summen der stan-

¹vorgeschlagen in [HoSt98].

dardisierten Zufallsvariablen X_1, \ldots, X_N dar, also $Y_T(t) = \sum_{i=1}^{[t]} (X_i - a_T)/b_T$ und $Y_T^*(t) = \sum_{i=[T^*]+1}^{[t]} (X_i - a_T^*)/b_T^*$, so entspricht $Z_T([t])$ der Summe $\sum_{i=1}^{[t]} X_i$ für $0 \le t \le T := N$. In diesem Kapitel wird das Testproblem des sich ändernden Parameters a_T des Prozesses $\{Z_T(t): 0 \le t \le T\}$ gemäß Definition 1.1 betrachtet. Eine gleichzeitige Änderung des Parameters b_T ist ebenfalls erlaubt. Die Testproblematik formuliert sich folgendermaßen:

$$\begin{aligned} & \mathbf{H}_{0}: \quad T^{*} = T \\ & \mathbf{H}_{A}: \quad 0 < T^{*} < T \text{ und } a_{T} \neq a_{T}^{*}. \end{aligned}$$

Dazu sei der Prozess $\{Z_T(t): 0 \leq t \leq T\}$ für festes T > 0 und festes $N \in \mathbb{N}$ an den äquidistanten Zeitpunkten $t_i = t_{i,N} = i\frac{T}{N}, 1 \leq i \leq N$, beobachtet worden und die Nullhypothese werde für große Werte von

$$M_{T} := \frac{1}{\sqrt{Tb_{T}^{2}}} \max_{1 \le k \le N} |Z_{k}^{*}|, \qquad (1.2)$$

wobei $Z_{k}^{*} := \sum_{i=1}^{k} R_{i} - \frac{k}{N} \sum_{i=1}^{N} R_{i}, \quad 1 \le k \le N,$
und $R_{i} := Z_{T} \left(i \frac{T}{N} \right) - Z_{T} \left((i-1) \frac{T}{N} \right), \quad 1 \le i \le N, \text{ mit } Z_{T}(0) = 0,$

verworfen. Vorausgesetzt sei $b_T \neq 0$.

Der Parameter b_T , der in vielen Anwendungen des Modells der Definition 1.1 einer Standardabweichung von zugrunde liegenden Zufallsvariablen entspricht, ist im Allgemeinen unbekannt. In den folgenden Abschnitten werden Schätzer für diesen Parameter vorgestellt und ihr Verhalten unter der Nullhypothese und unter der Alternative gemäß (1.1) sowie das Verhalten der Statistik M_T bei Ersetzen des Parameters b_T^2 durch diese Schätzer betrachtet. Analog kann die ebenfalls in [HoSt98] vorgeschlagene Teststatistik für das Problem des sich ändernden Parameters b_T untersucht werden.

Bei den asymptotischen Betrachtungen dieses Kapitels untersuchen wir die Familie der Prozesse $\{Z_T(t) : 0 \le t \le T\}$ gemäß Definition 1.1 für $T \to \infty$. Vorausgesetzt wird, dass die Prozesse $\{Z_T(t) : 0 \le t \le T\}$ für alle T > 0 auf einem gemeinsamen Wahrscheinlichkeitsraum (Ω, \mathcal{A}, P) definiert sind und die Parameter a_T, a_T^*, b_T und b_T^* für $T \to \infty$ beschränkt sind.

Im Hinblick auf die asymptotischen Untersuchungen wird noch die Alternative des Testproblems (1.1) nach der Abhängigkeit der Parameter a_T und a_T^* von T unterschieden. Man spricht von einem

fixed change, falls H_A und $\Delta_T := a_T - a_T^* = \text{konstant}$ für alle T > 0 erfüllt ist, local change, falls H_A und $\Delta_T := a_T - a_T^* \to 0$ für $T \to \infty$ erfüllt ist.

Da der Prozess $\{Z_T(t) : 0 \leq t \leq T\}$ nur an diskreten Zeitpunkten beobachtet wird, ist unter der Alternative der letzte Beobachtungspunkt vor dem Changepoint in den folgenden Untersuchungen von Bedeutung. In diesem Kapitel sei

$$k^* := k_T^* := \left[\frac{T^*N}{T}\right].$$

Weiter werden die (stochastischen) Landau-Symbole o_p und O_P in diesem Kapitel für $T \to \infty$ benutzt, sofern nicht anders kenntlich gemacht.

Um eine asymptotische Verteilung der Statistik M_T aus (1.2) zu erhalten, ist die gemeinsame Herangehensweise in verschiedenen konkreten Anwendungen des Modells 1.1, eine Approximation der Prozesse $\{Y_T(t)\}$ durch Wiener-Prozesse $\{W_T(t)\}$, die auf demselben Wahrscheinlichkeitsraum (Ω, \mathcal{A}, P) definiert sind, nachzuweisen. Dem allgemeinen Ansatz des Modells 1.1 folgend, werden solche Invarianzprinzipien wie in [HoSt98] für die Prozesse $\{Y_T(t)\}$ und $\{Y_T^*(t)\}$ vorausgesetzt:

Es existiert $\alpha \in [0, \frac{1}{2})$ und es exisitieren für alle T > 0 unabhängige Wiener-Prozesse $\{W_T(t) : 0 \le t \le T^*\}$ und $\{W_T^*(t): 0 \le t \le T - T^*\}$ auf (Ω, \mathcal{A}, P) , so dass gilt:

$$\sup_{0 \le t \le T^*} |Y_T(t) - W_T(t)| = O_P(T^{\alpha}), \quad T \to \infty,$$
(1.3)

$$\sup_{0 \le t \le T - T^*} |Y_T^*(t) - W_T^*(t)| = O_P(T^{\alpha}), \quad T \to \infty.$$
(1.4)

Im Allgemeinen müssen die Prozesse $\{Z_T(t) : 0 \leq t \leq T\}$ zuerst auf einem geeigneten Wahrscheinlichkeitsraum redefiniert werden, auf dem die Wiener-Prozesse definiert sind. In dieser Arbeit wird dies vorausgesetzt. Des Weiteren muss die Messbarkeit der Suprema in (1.3) und (1.4) gewährleistet sein, etwa durch die Forderung nach Separabilität der Prozesse² $\{Y_T(t)\}\$ und $\{Y_T^*(t)\}$. Da jedoch im Folgenden die Invarianzprinzipien nur an den diskreten Beobachtungspunkten ausgenutzt werden, erübrigt sich dies.

Die Definition 1.1 erlaubt eine Abhängigkeit der Verteilung der Prozesse $\{Y_T(t)\}$ und $\{Y_T^*(t)\}$ von T. Bei einem konkreten Nachweis der Invarianzprinzipien muss jedoch meistens eine von T unabhängige Verteilung vorausgesetzt werden, damit die Approximation gemäß (1.3) und (1.4) tatsächlich gleichmäßig in T gewährleistet ist³.

Das bisher betrachtete Beispiel der Folge X_1, \ldots, X_N und die Interpretation der Zufallsvariablen Y_T und Y_T^* als Summe von diesen genügt ohne weiteren Voraussetzungen nicht den Invarianzprinzipien (1.3) und (1.4). Beispiele von linearen Prozessen oder Partialsummenprozessen, die der Definition 1.1 entsprechen sowie die Invarianzprinzipien (1.3) und (1.4)erfüllen, sind in [HoSt98] und [Kühn99] aufgeführt.

Von besonderem Interesse ist das Modell der Definition 1.1, wenn die Prozesse $\{Y_T(t)\}$ und $\{Y_T^*(t)\}$ nicht direkt beobachtbar sind, also diese nicht unbedingt einer Summe von Zufallsvariablen entsprechen. Als Beispiel aus der Praxis sei ein Schadenanzahlprozess in der Versicherungsmathematik genannt. Der Nachweis, dass solch ein Erneuerungsprozess den Invarianzprinzipien (1.3) und (1.4) genügt, findet sich [CsHoSt87].

Mittels des Invarianzprinzips (1.3) lässt sich die asymptotische Verteilung der Statistik M_T aus (1.2) unter der Nullhypothese bestimmen. Damit ist es möglich, aufgrund der Beobachtungen $Z_T(iT/N)$, $1 < i \leq N$, Entscheidungen bezüglich des Testproblems (1.1) zu treffen und asymptotisch ein vorgegebenes Fehlerniveau einzuhalten. Die Verteilung von $\sup_{0 \le t \le 1} |B(t)|$ findet sich in [ShWe86].

²Eine separable Modifikation existiert auf jeden Fall. ³vgl. [CsHo93].

Satz 1.2

 $\{Z_T(t)\}\$ sei wie in Definition 1.1 für alle T > 0 und genüge dem Invarianzprinzip (1.3). Weiter sei $N = N(T) \rightarrow \infty$ für $T \rightarrow \infty$ erfüllt. Dann gilt unter H_0 :

$$M_T \xrightarrow{\mathcal{D}} \sup_{0 \le t \le 1} |B(t)|, \quad T \to \infty,$$

wobei $\{B(t): 0 \le t \le 1\}$ eine Standard-Brownsche-Brücke ist.

Beweis: siehe [HoSt98].

Bei wachsender Stichprobengröße N erwartet man unter der Alternative von der Statistik M_T wachsende Werte, die für eine deutlichere Verwerfung der Nullhypothese sprechen. Diese asymptotische Konsistenz gewährleistet der folgende Satz.

Satz 1.3

 $\{Z_T(t)\}\$ sei wie in Definition 1.1 mit $T^* = \theta_T T$, $\theta_T \in (0, 1)$, für alle T > 0 und genüge den Invarianzprinzipien (1.3) und (1.4).

Weiter sei $N = N(T) \to \infty$ für $T \to \infty$ sowie $|\Delta_T| \theta_T (1 - \theta_T) \sqrt{T} \to \infty$ für $T \to \infty$ erfüllt. Dann gilt unter H_A mit $1 \le k^* \le N - 1$:

$$\frac{1}{|\Delta_T|\theta_T(1-\theta_T)T}|Z_{k^*}^*| \xrightarrow{P} 1, \quad T \to \infty.$$

Beweis: siehe [HoSt98].

Unter den Voraussetzungen des Satzes 1.3 ergibt sich die asymptotische Konsistenz des Tests M_T bei bekannten Parametern b_T aus

$$M_T \ge \left(\frac{\theta_T^2 (1-\theta_T)^2 \Delta_T^2 T}{b_T^2}\right)^{1/2} \frac{1}{\theta_T (1-\theta_T) |\Delta_T| T} |Z_{k^*}^*| \xrightarrow{P} \infty, \quad T \to \infty.$$

Bei Ersetzen des Parameters b_T^2 durch einen beliebigen Schätzer s_T^2 reicht es unter der Alternative aus, $s_T^2 = o_P \left(\theta_T^2 (1 - \theta_T)^2 \Delta_T^2 T\right)$ zu fordern, um die asymptotische Konsistenz des Tests gewährleisten zu können.

1.2 Der Standard-Schätzer

Betrachtet man den Prozess $\{Z_T(t) : 0 \le t \le T\}$ gemäß Definition 1.1 unter der Nullhypothese, so gilt die Darstellung $Z_T(t) = a_T t + b_T Y_T(t)$. Als Schätzer⁴ bieten sich an bei festem T > 0 für die unbekannten Parameter

$$a_T: \qquad \hat{a}_N := \hat{a}_{N,T} := \frac{1}{T} \sum_{i=1}^N \left(Z_T \left(i \frac{T}{N} \right) - Z_T \left((i-1) \frac{T}{N} \right) \right), \tag{1.5}$$

$$b_T^2: \qquad \hat{b}_N^2 := \hat{b}_{N,T}^2 := \frac{1}{T} \sum_{i=1}^N \left(Z_T\left(i\frac{T}{N}\right) - Z_T\left((i-1)\frac{T}{N}\right) - \frac{T}{N}\hat{a}_N \right)^2.$$
(1.6)

⁴vgl. hier wie auch im Folgenden [HoSt98].

Dies sind sicherlich messbare Funktionen auf (Ω, \mathcal{A}) . Außerdem sind von diesen Funktionen konsistente Schätzungen unter der Nullhypothese zu erwarten. Der folgende Satz gibt die Raten dazu an.

Satz 1.4

 $\{Z_T(t)\}\$ sei wie in Definition 1.1 für alle T > 0 und genüge dem Invarianzprinzip (1.3). Weiter sei $N = N(T) \rightarrow \infty$ für $T \rightarrow \infty$ erfüllt. Dann gilt unter H_0 :

a)
$$\hat{a}_N - a_T = O_P(\frac{1}{\sqrt{T}}), \quad T \to \infty.$$

b) falls weiter $N = o(T^{1-2\alpha})$ für $T \to \infty$ gilt, so folgt:

$$\hat{b}_N^2 - b_T^2 = O_P\left(\max\left\{\frac{1}{\sqrt{N}}, \sqrt{\frac{N}{T}T^{2\alpha}}\right\}\right), \quad T \to \infty$$

Beweis:

a) Da die Verteilung von $T^{-1/2}W_T(T)$ unabhängig von T ist, folgt

$$\hat{a}_{N} = \frac{1}{T}Z_{T}(T)$$

$$= a_{T} + b_{T}\frac{Y_{T}(T)}{T}$$

$$= a_{T} + b_{T}\frac{W_{T}(T)}{T} + O_{P}(T^{\alpha-1})$$

$$= a_{T} + O_{P}(T^{-1/2}) + O_{P}(T^{\alpha-1})$$

$$= a_{T} + O_{P}(T^{-1/2}).$$

b) Ohne Einschränkung sei $a_T = 0$. Betrachte ansonsten im Folgenden $\left(\left(R_i - a_T \frac{T}{N}\right) - \frac{T}{N} \left(\hat{a}_N - a_T\right)^2\right)$ als Summanden. Dann gilt mit a)

$$\hat{b}_{N}^{2} = \frac{1}{T} \sum_{i=1}^{N} \left(R_{i} - \frac{T}{N} \hat{a}_{N} \right)^{2}$$

$$= \frac{1}{T} \sum_{i=1}^{N} R_{i}^{2} - \frac{T}{N} \hat{a}_{N}^{2}$$

$$= \frac{1}{T} \sum_{i=1}^{N} R_{i}^{2} + O_{P} \left(\frac{1}{N} \right). \qquad (1.7)$$

Sei $W_i = W_T(iT/N) - W_T((i-1)T/N)$ für $1 \le i \le N$. Dann sind $\sqrt{\frac{N}{T}}W_i$ für $1 \le i \le N$ unabhängig und identisch verteilt, wobei $\sqrt{\frac{N}{T}}W_i \stackrel{\mathcal{D}}{=} N(0,1)$. Damit liefert der zentrale Grenzwertsatz

$$\frac{1}{T}\sum_{i=1}^{N}W_i^2 - 1 = \frac{1}{N}\sum_{i=1}^{N}\frac{N}{T}W_i^2 - 1 = O_P\left(\frac{1}{\sqrt{N}}\right).$$

Eine Anwendung der Markov-Ungleichung ergibt für beliebiges C > 0

$$P\left(\sum_{i=1}^{N} |W_i| \ge C\sqrt{NT}\right) = P\left(\sum_{i=1}^{N} \sqrt{\frac{N}{T}} |W_i| \ge CN\right) \le \frac{2}{C}$$

it gilt $\sum_{i=1}^{N} |W_i| = O_P\left(\sqrt{NT}\right)$.

und dam i=1

Sei weiter $\eta_i = R_i - b_T W_i$ für $1 \le i \le N$. Dann gilt wegen des Invarianzprinzips (1.3) und der Annahme $a_T = 0$

$$\max_{1 \le i \le N} |\eta_i| = O_P(T^{\alpha}).$$

Somit folgt

$$\begin{aligned} \frac{1}{T} \sum_{i=1}^{N} R_i^2 &= \frac{1}{T} \sum_{i=1}^{N} (b_T W_i + \eta_i)^2 \\ &= b_T^2 \frac{1}{T} \sum_{i=1}^{N} W_i^2 + b_T \frac{2}{T} \sum_{i=1}^{N} W_i \eta_i + \frac{1}{T} \sum_{i=1}^{N} \eta_i^2 \\ &= b_T^2 + O_P(N^{-1/2}) + O_P(T^{\alpha - 1/2} N^{1/2}) + O_P(NT^{2\alpha - 1}) \end{aligned}$$

und damit ergibt sich mit (1.7) die Behauptung b).

Mittels des Schätzers \hat{b}_N^2 lässt sich jetzt der unbekannte Parameter b_T^2 in der Statistik M_T aus (1.2) ersetzen. Man erhält als Statistik für festes T > 0

$$\hat{M}_T := \frac{1}{\sqrt{T\hat{b}_N^2}} \max_{1 \le k \le N} |Z_k^*| \,. \tag{1.8}$$

Für die asymptotischen Untersuchungen der Statistik \hat{M}_T unter der Nullhypothese reicht es aus, dass $\hat{b}_N^2 \ge c/2$ für ein c>0 auf Mengen von beliebig hohen Wahrscheinlichkeiten für T groß genug erfüllt ist. Dies wird aufgrund der Konsistenz des Schätzers \hat{b}_N^2 gewährleistet, falls $b_T^2 \geq c > 0$ für alle T > 0zusätzlich zu den Voraussetzungen des Satzes 1.4 gefordert wird. Weiter impliziert die Konsistenz des Schätzers \hat{b}_N^2 das Beibehalten der Grenzverteilung der Statistik M_T unter der Nullhypothese bei Ersetzen des Parameters b_T^2 durch den Schätzer.

Korollar 1.5

 $\{Z_T(t)\}\$ sei wie in Definition 1.1 für alle T > 0 und genüge dem Invarianzprinzip (1.3). Weiter sei $N = N(T) \to \infty$ für $T \to \infty$ und $N = o(T^{1-2\alpha})$ für $T \to \infty$ sowie $b_T^2 \ge c$ für alle T > 0 bei einem festen c > 0 erfüllt. Dann gilt unter H_0 :

$$\hat{M}_T \xrightarrow{\mathcal{D}} \sup_{0 \le t \le 1} |B(t)|, \quad T \to \infty,$$

wobei $\{B(t) : 0 \le t \le 1\}$ eine Standard-Brownsche-Brücke ist.

Beweis: Folgt direkt mit den Sätzen 1.2 und 1.4 sowie dem Lemma von Slutsky aus

$$\left| M_T - \frac{1}{\sqrt{T\hat{b}_N^2}} \max_{1 \le k \le N} |Z_k^*| \right| = \frac{1}{\sqrt{T}} \max_{1 \le k \le N} |Z_k^*| \left| \frac{1}{\sqrt{\hat{b}_N^2}} - \frac{1}{\sqrt{b_T^2}} \right| \xrightarrow{P} 0, \quad T \to \infty.$$

Unter der Alternative sind von dem Schätzer \hat{b}_N^2 aus (1.6) keine konsistenten Schätzungen mehr zu erwarten. Um die Konsistenz des Tests \hat{M}_T aus (1.8) zu gewährleisten, reicht es aus, dass der Schätzer \hat{b}_N^2 nicht "zu schnell" gegen Unendlich strebt bei wachsender Beobachtungsdauer T.

Satz 1.6

 $\{Z_T(t)\}\$ sei wie in Definition 1.1 mit $T^* = \theta_T T$, $\theta_T \in (0, 1)$, für alle T > 0 und genüge den Invarianzprinzipien (1.3) und (1.4). Weiter sei $N = N(T) \to \infty$ für $T \to \infty$ und $N = O(T^{1-2\alpha})$ für $T \to \infty$ erfüllt. Dann gilt unter H_A mit $1 \le k^* \le N - 1$:

$$\hat{b}_N^2 = O_P\left(\frac{T}{N}\right)O_P\left(\theta_T a_T^2 + (1-\theta_T)a_T^{*2} + \frac{1}{N} + \sqrt{\frac{N}{T}}\right), \quad T \to \infty$$

Beweis: Für den Parameter a_T gilt unter der Alternative:

$$\hat{a}_{N} = \frac{1}{T}Z_{T}(T)
= \frac{1}{T}Z_{T}(T^{*}) + \frac{1}{T}(Z_{T}(T) - Z_{T}(T^{*}))
= a_{T}\frac{T^{*}}{T} + a_{T}^{*}\frac{T - T^{*}}{T} + b_{T}\frac{1}{T}W_{T}(T^{*}) + b_{T}^{*}\frac{1}{T}W_{T}^{*}(T - T^{*}) + O_{P}(T^{\alpha - 1})
= \theta_{T}a_{T} + (1 - \theta_{T})a_{T}^{*} + O_{P}(T^{-1/2}).$$
(1.9)

Sei $W_i = W_T(iT/N) - W_T((i-1)T/N)$ für $1 \le i \le k^*$. Dann sind $\sqrt{\frac{N}{T}}W_i$ für $1 \le i \le k^*$ unabhängig und identisch verteilt, wobei $\sqrt{\frac{N}{T}}W_i \stackrel{\mathcal{D}}{=} N(0,1)$. Eine Anwendung der Markov-Ungleichung ergibt für beliebiges C > 0

$$P\left(\sum_{i=1}^{k^*} |W_i| \ge Ck^* \sqrt{\frac{T}{N}}\right) = P\left(\sum_{i=1}^{k^*} \sqrt{\frac{N}{T}} |W_i| \ge k^*C\right) \le \frac{2}{C},$$

und damit gilt
$$\sum_{i=1}^{k^*} |W_i| = O_P\left(k^* \sqrt{\frac{T}{N}}\right).$$

Mittels einer weiteren Anwendung der Markov-Ungleichung erhält man $\sum_{i=1}^{k^*} W_i^2 = O_P\left(k^* \frac{T}{N}\right)$. Weiter sei $\eta_i = R_i - b_T W_i - \frac{T}{N} a_T$ für $1 \le i \le k^*$. Dann gilt wegen des Invarianzprinzips (1.3)

$$\max_{1 \le i \le k^*} |\eta_i| = O_P(T^{\alpha}).$$

Für die Summe der quadratischen Zuwächse bis zum Changepoint folgt somit

$$\begin{split} &\frac{1}{T}\sum_{i=1}^{k^*} R_i^2 \\ &= \frac{1}{T}\sum_{i=1}^{k^*} \left(\frac{T}{N}a_T + b_T W_i + \eta_i\right)^2 \\ &= \frac{1}{T}\left(k^* a_T^2 \frac{T^2}{N^2} + 2b_T a_T \frac{T}{N}\sum_{i=1}^{k^*} W_i + b_T^2 \sum_{i=1}^{k^*} W_i^2 + 2\sum_{i=1}^{k^*} \left(\frac{T}{N}a_T + b_T W_i\right) \eta_i + \sum_{i=1}^{k^*} \eta_i^2\right) \\ &= \frac{k^* T}{N^2} \left(a_T^2 + O_P\left(\sqrt{\frac{N}{T}}\right) + O_P\left(\frac{N}{T}\right) + O_P\left(\frac{N}{T^{1-\alpha}}\right) + O_P\left(\frac{N^{3/2}}{T^{3/2-\alpha}}\right) + O_P\left(\frac{N^2}{T^{2-2\alpha}}\right)\right) \\ &= \frac{k^* T}{N^2} \left(a_T^2 + O_P\left(\sqrt{\frac{N}{T}}\right)\right). \end{split}$$

Analog folgt für die Summe der quadratischen Zuwächse nach dem Changepoint

$$\frac{1}{T} \sum_{i=k^*+2}^{N} R_i^2 = \frac{(N-k^*)T}{N^2} \left(a_T^{*2} + O_P\left(\sqrt{\frac{N}{T}}\right) \right).$$

Für den gemischten Summanden R_{k^*+1} , mit einem Anteil sowohl vor dem Changepoint als auch danach, gilt $R_{k^*+1} = O_P(T/N)$. Somit folgt mit (1.9) die Behauptung aus

$$\hat{b}_{N}^{2} = \frac{1}{T} \sum_{i=1}^{k^{*}} R_{i}^{2} + \frac{1}{T} \sum_{i=k^{*}+1}^{N} R_{i}^{2} - \frac{T}{N} \hat{a}_{N}^{2}$$

$$= \frac{T}{N} \left(\theta_{T} a_{T}^{2} + (1 - \theta_{T}) a_{T}^{*2} + O_{P} \left(\frac{1}{N} \right) + O_{P} \left(\sqrt{\frac{N}{T}} \right)$$

$$+ O_{P} \left(\max \left\{ \theta_{T}^{2} a_{T}^{2}, (1 - \theta_{T})^{2} a_{T}^{*2}, \frac{1}{T} \right\} \right) \right)$$

$$= O_{P} \left(\frac{T}{N} \right) O_{P} \left(\theta_{T} a_{T}^{2} + (1 - \theta_{T})^{2} a_{T}^{*2} + \frac{1}{N} + \sqrt{\frac{N}{T}} \right).$$

Das Resultat des vorangegangenen Satzes deckt auch Spezialfälle von Alternativen ab, wie etwa das Vorliegen eines "frühen" oder "späten" Changepoints, d.h. $\theta_T \to 0$ oder $\theta_T \to 1$ für $T \to \infty$. Unter bestimmten Voraussetzungen besitzt dann der Schätzer \hat{b}_N^2 nach Satz 1.6 bessere Konvergenzraten als $O_P(T/N)$. Liegen jedoch keine weiteren Kenntnisse über die Parameter a_T und a_T^* sowie über die Lage des Changepoints vor, so reduziert sich die Aussage des Satzes auf $\hat{b}_N^2 = O_P(T/N)$.

Das folgende Korollar stellt die Voraussetzungen zusammen, um die asymptotische Konsistenz der Statistik \hat{M}_T aus (1.8) zu gewährleisten. Der negative Einfluss des Vorliegens eines frühen oder späten Changepoints auf das Konvergenzverhalten der Teststatistik lässt sich unter Umständen durch die geringere Rate kompensieren, mit der in diesen Fällen der Schätzer \hat{b}_N^2 stochastisch gegen Unendlich strebt.

Korollar 1.7

da

 $\{Z_T(t)\}\$ sei wie in Definition 1.1 mit $T^* = \theta_T T$, $\theta_T \in (0, 1)$ für alle T > 0 und genüge den Invarianzprinzipien (1.3) und (1.4). Weiter sei $N = N(T) \to \infty$ für $T \to \infty$ und $N = O(T^{1-2\alpha})$ für $T \to \infty$ sowie $|\Delta_T|\theta_T(1-\theta_T)\sqrt{T} \to \infty$ für $T \to \infty$ erfüllt. Außerdem gelte

$$\frac{\theta_T a_T^2 + (1 - \theta_T) a_T^{*2} + 1/N + \sqrt{N/T}}{\theta_T^2 (1 - \theta_T)^2 \Delta_T^2 N} \to 0, \quad T \to \infty.$$
(1.10)

Dann gilt unter H_A mit $1 \le k^* \le N - 1$:

$$\hat{M}_T = \frac{1}{\sqrt{T\hat{b}_N^2}} \max_{1 \le k \le N} |Z_k^*| \xrightarrow{P} \infty, \quad T \to \infty.$$

Beweis: Eine Anwendung des Satzes 1.3 ergibt

$$\frac{1}{\sqrt{T\hat{b}_N^2}} |Z_{k^*}^*| = \left(\frac{\theta_T^2 (1-\theta_T)^2 \Delta_T^2 T}{\hat{b}_N^2}\right)^{1/2} \frac{1}{\theta_T (1-\theta_T) |\Delta_T| T} |Z_{k^*}^*| \xrightarrow{P} \infty,$$
$$\frac{\hat{b}_N^2}{\theta_T^2 (1-\theta_T)^2 \Delta_T^2 T} = o_P(1) \text{ nach Satz 1.6.} \square$$

Schließt man einen frühen oder späten Changepoint in Korollar 1.7 durch $\theta_T \in (\epsilon, 1 - \epsilon)$, $0 < \epsilon < 1/2$, aus, so lässt sich die Voraussetzung (1.10) sinnvoll durch $\Delta_T^2 N \to \infty$ ersetzen, falls keine Kenntnisse über die Parameter a_T und a_T^* vorliegen. Die für die Anwendung des Satzes 1.3 notwendige Voraussetzung $|\Delta_T|\theta_T(1-\theta_T)\sqrt{T}\to\infty$ für $T\to\infty$ in Korollar 1.7 wird dadurch impliziert.

1.3 Die kombinierten Schätzer

Die Verteilung des Prozesses $\{Z_T(t) : 0 \leq t \leq T\}$ unter der Nullhypothese steht bei der Herleitung der Schätzer des letzten Abschnitts deutlich im Vordergrund. In diesem Abschnitt wird versucht, das eventuelle Vorliegen der Alternative in die Konstruktion der Schätzer miteinzubeziehen, ohne jedoch die Konsistenz der Schätzer unter der Nullhypothese zu verlieren. Dazu setzen wir für festes T > 0 kurzzeitig den Changepoint $T^* = \theta_T T$, $\theta_T \in (0, 1)$, als bekannt voraus und erhalten zwei Mengen $\{Z_T(T/N), \ldots, Z_T(k^*T/N)\}$ und $\{Z_T((k^* + 1)T/N), \ldots, Z_T(T)\}$ von Beobachtungen, wobei $k^* = [\theta_T N]$. Auf diesen Mengen lassen sich unter der Alternative Schätzer für die Parameter der Verteilungen gemäß Definition 1.1 angeben, nach denen die jeweiligen Beobachtungen verteilt sind. Da jedoch der Changepoint T^* und damit k^* unbekannt sind, nehmen wir obige Aufteilung der Beobachtungen für jedes $1 \leq k \leq N$ vor und erhalten für festes k und T > 0 als Schätzer für

$$a_T: \qquad \hat{a}_k := \hat{a}_{k,T} := \frac{N}{kT} \sum_{i=1}^k R_i, \quad 1 \le k \le N,$$
(1.11)

$$a_T^*: \quad \check{a}_k := \check{a}_{k,T} := \frac{1}{N-k-1} \frac{N}{T} \sum_{i=k+2}^N R_i, \quad -1 \le k \le N-2.$$
 (1.12)

Die Schätzer \check{a}_k für k = -1 oder k = 0 entsprechen den Fällen, dass entweder gar kein Changepoint existiert oder vor der ersten Beobachtung liegt und treten hier auf, weil eine einheitliche Definition gewählt wurde. Das Auslassen des Zuwachses R_{k+1} in der Definition der Schätzer \hat{a}_k und \check{a}_k wird bei den Betrachtungen unter der Alternative begründet.

Da diese Funktionen messbar sind, erhält man zwei Familien $\{\hat{a}_k : 1 \leq k \leq N\}$ und $\{\check{a}_k : -1 \leq k \leq N-2\}$ von Schätzern. Von den Schätzern \check{a}_k sind unter der Nullhypothese Schätzungen des Parameters a_T zu erwarten.

Bei konsequentem Weiterbegehen dieses Weges erhält man ebenso zwei Familien von Schätzern für die Parameter b_T^2 und b_T^{*2} . Für die Problemstellung, den unbekannten Parameter b_T^2 in der Statistik M_T aus (1.2) zu ersetzen, erscheint dies aber nicht unbedingt geeignet. Denn wie die Schätzer \hat{a}_k aus (1.11) würden die Schätzer des Parameters b_T^2 für kleine k auf einer zeitlich kurzen Beobachtung des Prozesses $\{Z_T(t) : 0 \le t \le T\}$ basieren. Unter der Nullhypothese ist es jedoch möglich, die Beobachtungen des gesamten Beobachtungszeitraumes bis T in die Schätzer für den Parameter b_T^{*2} liefern unter der Nullhypothese Schätzungen für b_T^2 . Deshalb kombinieren wir die Schätzer für b_T^2 und b_T^{*2} der beiden Mengen von Beobachtungen und erhalten als Schätzer für b_T^2 bei festem $1 \le k \le N$ und T > 0

$$\hat{b}_{k}^{2} := \hat{b}_{k,T}^{2} := \begin{cases} \frac{1}{T} \left(\sum_{i=1}^{k} \left(R_{i} - \frac{T}{N} \hat{a}_{k} \right)^{2} + \sum_{i=k+2}^{N} \left(R_{i} - \frac{T}{N} \check{a}_{k} \right)^{2} \right) & : \quad 1 \le k \le N-2, \\ \frac{1}{T} \sum_{i=1}^{N} \left(R_{i} - \frac{T}{N} \hat{a}_{N} \right)^{2} & : \quad k \in \{N-1,N\}. \end{cases}$$
(1.13)

Wie oben erhält man also eine Familie $\{\hat{b}_k^2 : 1 \leq k \leq N\}$ von Schätzern. Die heuristische Herleitung begründet der folgende Satz, in dem die Schätzer sogar in einem gleichmäßigen Sinn unter der Nullhypothese als konsistent nachgewiesen werden.

Satz 1.8

 $\{Z_T(t)\}\$ sei wie in Definition 1.1 für alle T > 0 und genüge dem Invarianzprinzip (1.3). Weiter sei $N = N(T) \rightarrow \infty$ für $T \rightarrow \infty$ erfüllt. Dann gilt unter H_0 :

a)
$$\max_{1 \le k \le N} |\hat{a}_k - a_T| = O_P\left(\max\left\{\sqrt{\frac{N}{T}}, NT^{\alpha - 1}\right\}\right), \quad T \to \infty,$$

und
$$\max_{-1 \le k < N-2} |\check{a}_k - a_T| = O_P\left(\max\left\{\sqrt{\frac{N}{T}}, NT^{\alpha - 1}\right\}\right), \quad T \to \infty$$

b) falls zusätzlich N und T/N monoton wachsend in T sind sowie

 $\begin{aligned} \frac{\log N}{\log \log T} &\to \infty \ f\ddot{u}r \ T \to \infty \quad und \quad \frac{N \log N}{T^{1-2\alpha}} \to 0 \ f\ddot{u}r \ T \to \infty \ gilt, \\ dann \ folgt \ auch: \\ & \max_{1 \le k \le N} \left| \hat{b}_k^2 - b_T^2 \right| = o_P(1), \quad T \to \infty. \end{aligned}$

Beweis:

a) Mit dem Gesetz vom iterierten Logarithmus⁵ für Wiener-Prozesse folgt

$$\begin{aligned} \max_{1 \le k \le N} |\hat{a}_k - a_T| &= |b_T| \max_{1 \le k \le N} \left| \frac{N}{kT} Y_T \left(k \frac{T}{N} \right) \right| \\ &\stackrel{\mathcal{D}}{=} |b_T| \max_{1 \le k \le N} \left| \sqrt{\frac{N}{T}} \frac{1}{k} W_T(k) \right| + O_P \left(NT^{\alpha - 1} \right) \\ &= |b_T| \sqrt{\frac{N}{T}} \max_{1 \le k \le N} \left| \sqrt{\frac{\log \log k}{k}} \frac{W_T(k)}{\sqrt{k \log \log k}} \right| + O_P \left(NT^{\alpha - 1} \right) \\ &= O_P \left(\sqrt{\frac{N}{T}} \right) + O_P \left(NT^{\alpha - 1} \right). \end{aligned}$$

Analog ergibt sich aufgrund der unabhängigen Zuwächse eines Wiener-Prozesses

$$\max_{-1 \le k < N-2} |\check{a}_k - a_T| = O_P\left(\sqrt{\frac{N}{T}}\right) + O_P\left(NT^{\alpha-1}\right)$$

und somit folgt Behauptung a).

b) Ohne Einschränkung sei $a_T = 0$. Dann erhält man mit denselben Methoden wie in Teil a)

$$\begin{aligned} \max_{1 \le k \le N} \left| \frac{\sqrt{kT}}{N} \hat{a}_k \right| &= |b_T| \max_{1 \le k \le N} \left| \frac{1}{\sqrt{kT}} W_T \left(k \frac{T}{N} \right) \right| + O_P \left(T^{\alpha - 1/2} \right) \\ &\stackrel{\mathcal{D}}{=} |b_T| \max_{1 \le k \le N} \left| \sqrt{\frac{\log \log k}{N}} \frac{W_T(k)}{\sqrt{k \log \log k}} \right| + O_P \left(T^{\alpha - 1/2} \right) \\ &= O_P \left(\sqrt{\frac{\log \log N}{N}} \right) + O_P \left(T^{\alpha - 1/2} \right) \\ &= o_P(1). \end{aligned}$$

Damit gilt

$$\max_{1 \le k \le N} \left| \frac{1}{T} \sum_{i=1}^{k} \left(R_i - \frac{T}{N} \hat{a}_k \right)^2 - \frac{1}{T} \sum_{i=1}^{k} R_i^2 \right| = \max_{1 \le k \le N} \left| k \frac{T}{N^2} \hat{a}_k^2 \right| = o_P(1).$$

Und auf gleiche Weise folgt für die Summe nach dem Changepoint

$$\max_{1 \le k \le N-2} \left| \frac{1}{T} \sum_{i=k+2}^{N} \left(R_i - \frac{T}{N} \check{a}_k \right)^2 - \frac{1}{T} \sum_{i=k+2}^{N} R_i^2 \right| = o_P(1).$$

⁵In dieser Arbeit wird log als $\max\{\log, 1\}$ benutzt.

Dann folgt für festes $k_0 \in \mathbb{N}$

$$\begin{split} \max_{1 \le k \le N-2} \left| \hat{b}_{k}^{2} - b_{T}^{2} \right| \\ &\leq \max_{1 \le k \le N-2} \left| \frac{1}{T} \sum_{i=1}^{k} \left(R_{i} - \frac{T}{N} \hat{a}_{k} \right)^{2} - \frac{k}{N} b_{T}^{2} \right| \\ &+ \max_{1 \le k \le N-2} \left| \frac{1}{T} \sum_{i=k+2}^{N} \left(R_{i} - \frac{T}{N} \check{a}_{k} \right)^{2} - \frac{N-k}{N} b_{T}^{2} \right| \\ &= \max_{1 \le k \le N-2} \left| \frac{1}{T} \sum_{i=1}^{k} R_{i}^{2} - \frac{k}{N} b_{T}^{2} \right| + \max_{1 \le k \le N-2} \left| \frac{1}{T} \sum_{i=k+2}^{N} R_{i}^{2} - \frac{N-k}{N} b_{T}^{2} \right| + o_{P}(1) \\ &= \max_{k_{0} \le k \le N-2} \left| \frac{N}{kT} \sum_{i=1}^{k} R_{i}^{2} - b_{T}^{2} \right| + \max_{1 \le k \le N-k_{0}} \left| \frac{N}{(N-k-1)T} \sum_{i=k+2}^{N} R_{i}^{2} - b_{T}^{2} \right| + o_{P}(1), \end{split}$$

$$(1.14)$$

da für festes k_0 sicherlich gilt

$$\max_{1 \le k < k_0} \left| \frac{1}{T} \sum_{i=1}^k R_i^2 - \frac{k}{N} b_T^2 \right| = o_P(1), \quad T \to \infty,$$
$$\max_{N-k_0 < k \le N-2} \left| \frac{1}{T} \sum_{i=k+2}^N R_i^2 - \frac{N-k}{N} b_T^2 \right| = o_P(1), \quad T \to \infty$$

•

Sei im Folgenden $W_i = W_T(iT/N) - W_T((i-1)T/N)$ für $1 \le i \le N$. Dann sind $\sqrt{\frac{N}{T}}W_i$ für $1 \le i \le N$ unabhängig und identisch verteilt, wobei $\sqrt{\frac{N}{T}}W_i \stackrel{\mathcal{D}}{=} N(0,1)$. Nach Satz 1.2.1 in [CsRe81] gilt $\max_{1\le i\le N} |W_i| = O_P\left(\sqrt{\frac{T}{N}\log N}\right)$.

Sei weiter $\eta_i = R_i - b_T W_i$ für $1 \le i \le N$. Dann gilt wegen des Invarianzprinzips (1.3) und der Annahme $a_T = 0 \max_{1 \le i \le N} |\eta_i| = O_P(T^{\alpha}).$

Damit folgt

$$\max_{k_{0} \leq k \leq N} \left| \frac{N}{kT} \sum_{i=1}^{k} R_{i}^{2} - b_{T}^{2} \frac{N}{kT} \sum_{i=1}^{k} W_{i}^{2} \right|
= \max_{k_{0} \leq k \leq N} \left| \frac{N}{kT} \sum_{i=1}^{k} (b_{T} W_{i} + \eta_{i})^{2} - b_{T}^{2} \frac{N}{kT} \sum_{i=1}^{k} W_{i}^{2} \right|
\leq |b_{T}| \max_{k_{0} \leq k \leq N} \left| \frac{2N}{kT} \sum_{i=1}^{k} \eta_{i} W_{i} \right| + \max_{k_{0} \leq k \leq N} \left| \frac{N}{kT} \sum_{i=1}^{k} \eta_{i}^{2} \right|
= O_{P} \left(\sqrt{\frac{N}{T} T^{2\alpha} \log N} \right) + O_{P} \left(\frac{N}{T} T^{2\alpha} \right)
= o_{P}(1).$$
(1.15)

Und analog ergibt sich

$$\max_{1 \le k \le N-k_0} \left| \frac{N}{(N-k-1)T} \sum_{i=k+2}^{N} R_i^2 - b_T^2 \frac{N}{(N-k-1)T} \sum_{i=k+2}^{N} W_i^2 \right| = o_P(1).$$
(1.16)

Für den letzten Schritt des Beweises seien beliebige $\epsilon, \delta > 0$ vorgegeben und es sei $B^2 > 0$ so gewählt, dass nach Voraussetzung $b_T^2 \leq B^2$ für alle T > 0 erfüllt ist.

Weiter seien $X_i \stackrel{\mathcal{D}}{=} N(0,1)$ für $i \in \mathbb{N}$ und unabhängig. Das starke Gesetz der großen Zahlen ergibt

$$\exists k_1 \in \mathbb{N}: \qquad P\left(\sup_{k_1 \leq k} \left| \frac{1}{k} \sum_{i=1}^k X_i^2 - 1 \right| \geq \frac{\delta}{B^2} \right) \leq \frac{\epsilon}{2}.$$

Damit folgt für alle $N \ge k_1$

$$P\left(b_T^2 \max_{k_1 \le k \le N} \left| \frac{N}{kT} \sum_{i=1}^k W_i^2 - 1 \right| \ge \delta \right) \le \frac{\epsilon}{2}, \quad \forall N \ge k_1,$$
(1.17)

und ebenso für die Summe nach dem Changepoint

$$\exists k_2 \in \mathbb{N}:$$

$$P\left(b_T^2 \max_{1 \le k \le N-k_2} \left| \frac{N}{(N-k-1)T} \sum_{i=k+2}^N W_i^2 - 1 \right| \ge \delta \right) \le \frac{\epsilon}{2}, \quad \forall N \ge k_2.$$

$$(1.18)$$

Da in (1.14) k_0 beliebig war, ergibt sich daraus und mit (1.15) bis (1.18) sowie Satz 1.4 die Behauptung.

Anders als im vorigen Abschnitt erhält man eine Familie $\{\hat{b}_k^2 : 1 \le k \le N\}$ von Schätzern für den Paramater b_T^2 . Da jedoch die Statistik M_T aus (1.2) auf der von k abhängigen Differenz Z_k^* basiert, ersetzen wir für jedes k den unbekannten Parameter b_T^2 durch den Schätzer \hat{b}_k^2 und erhalten so für festes T > 0

$$\tilde{M}_T := \frac{1}{\sqrt{T}} \max_{1 \le k \le N} \frac{1}{\sqrt{\hat{b}_k^2}} |Z_k^*| \,. \tag{1.19}$$

Dabei gilt $\min_{1 \le k \le N} \hat{b}_k^2 \ge c/2$ auf Mengen von beliebig hohen Wahrscheinlichkeiten für T groß genug, falls $\tilde{b}_T^2 \ge c > 0$ für alle T > 0 sowie die Voraussetzungen des Satzes 1.8.b) erfüllt sind. Mit den Sätzen 1.2 und 1.8 sowie dem Lemma von Slutsky ergibt sich dann die Grenzverteilung der Statistik \tilde{M}_T gemäß Satz 1.2 unter der Nullhypothese.

Aufgrund der Kombination der Schätzer für die Parameter b_T^2 und b_T^{*2} , wie in der Einleitung dieses Abschnitts dargelegt, liefert im Allgemeinen keiner der Schätzer \hat{b}_k^2 für $1 \le k \le N$ unter der Alternative konsistente Schätzungen für den Parameter b_T^2 . Doch gerade wegen dieser Kombination der beiden Schätzer ist unter der Alternative mit einem Changepoint T^* und $k^* = [NT^*/T]$ von den beiden Summen in der ersten Zeile von (1.13) für $k = k^*$ zu erwarten, dass diese jeweils für $T \to \infty$ entweder stochastisch konvergent oder zumindest stochastisch beschränkt (= $O_P(1)$) sind. Da k^* unbekannt ist, besitzt der Wert des eventuell vorhandenen Grenzwertes des Schätzers $\hat{b}_{k^*}^2$ keine Relevanz. Vielmehr ist für die Statistik \tilde{M}_T von Bedeutung, dass der Schätzer $\hat{b}_{k^*}^2$ unter der Alternative auf jeden Fall stochastisch beschränkt ist. Diesen Sachverhalt gibt der folgende Satz unter Alternativen mit einem beliebigen Changepoint $T^* = \theta_T T$ wieder, wobei $\theta_T \in (0, 1)$. Wegen der Konstruktion des Schätzers $\hat{b}_{k^*}^2$ ist im Gegensatz zu Satz 1.6 keine Abhängigkeit von θ_T bzw. $1 - \theta_T$ zu erwarten.

Satz 1.9

 $\{Z_T(t)\}\$ sei wie in Definition 1.1 mit $T^* = \theta_T T$, $\theta_T \in (0, 1)$, für alle T > 0 und genüge den Invarianzprinzipien (1.3) und (1.4).

Weiter sei $N = N(T) \rightarrow \infty$ für $T \rightarrow \infty$ sowie $N = O(T^{1-2\alpha})$ für $T \rightarrow \infty$ erfüllt. Dann gilt unter H_A mit $1 \le k^* \le N - 1$:

$$\hat{b}_{k^*}^2 = O_P(1), \quad T \to \infty.$$

Beweis: Ohne Einschränkung sei $1 \le k^* \le N - 2$. Betrachte anderenfalls im Folgenden nur \hat{a}_{k^*} und Summen, die bis k^* aufsummieren.

1. z.z.:
$$\hat{a}_{k^*} = a_T + O_P\left(\sqrt{\frac{N}{k^*T}}\right)$$
 und $\check{a}_{k^*} = a_T^* + O_P\left(\sqrt{\frac{N}{(N-k^*)T}}\right)$
 $\hat{a}_{k^*} = \frac{N}{k^*T}Z_T\left(k^*\frac{T}{N}\right)$
 $= a_T + b_T\frac{N}{k^*T}\left(W_T\left(k^*\frac{T}{N}\right) + O_P(T^{\alpha})\right)$
 $= a_T + O_P\left(\sqrt{\frac{N}{k^*T}}\right).$

Analog folgt mittels des Invarianzprinzips (1.4) die Behauptung für \check{a}_{k^*} .

2. z.z.: $\hat{b}_{k^*}^2 = O_P(1)$.

Sei $W_i = W_T(iT/N) - W_T((i-1)T/N)$ für $1 \le i \le k^*$. Dann sind $\sqrt{\frac{N}{T}}W_i$ für $1 \le i \le k^*$ unabhängig und identisch verteilt, wobei $\sqrt{\frac{N}{T}}W_i \stackrel{\mathcal{D}}{=} N(0,1)$. Damit ergibt eine Anwendung der Markov-Ungleichung

$$\sum_{i=1}^{k^*} |W_i| = O_P\left(k^*\sqrt{\frac{T}{N}}\right).$$

Analog ergibt sich

$$\sum_{i=1}^{k^*} W_i^2 = O_P\left(k^* \frac{T}{N}\right).$$
(1.20)

Sei weiter $\eta_i = R_i - b_T W_i - \frac{T}{N} a_T$ für $1 \le i \le k^*$. Dann gilt wegen des Invarianzprinzips (1.3)

$$\max_{1 \le i \le k^*} |\eta_i| = O_P(T^{\alpha})$$

Somit folgt für die Summe

$$\sum_{i=1}^{k^*} \left| R_i - \frac{T}{N} a_T \right| = \sum_{i=1}^{k^*} \left| b_T W_i + \eta_i \right|$$
$$= O_P \left(k^* \sqrt{\frac{T}{N}} \right) + O_P \left(k^* T^{\alpha} \right)$$
$$= O_P \left(k^* \sqrt{\frac{T}{N}} \right).$$

Damit erhält man nun mit Teil 1)

$$\frac{1}{T} \sum_{i=1}^{k^*} \left| \left(R_i - \frac{T}{N} \hat{a}_{k^*} \right)^2 - \left(R_i - \frac{T}{N} a_T \right)^2 \right| \\
\leq \frac{1}{T} \left(2\frac{T}{N} \left| \hat{a}_{k^*} - a_T \right| \sum_{i=1}^{k^*} \left| R_i - \frac{T}{N} a_T \right| + \frac{T^2}{N^2} \sum_{i=1}^{k^*} \left(\hat{a}_{k^*} - a_T \right)^2 \right) \\
= \frac{1}{N} O_P \left(\sqrt{\frac{N}{k^*T}} \right) O_P \left(k^* \sqrt{\frac{T}{N}} \right) + \frac{k^*T}{N^2} O_P \left(\frac{N}{k^*T} \right) \\
= O_P \left(\frac{\sqrt{k^*}}{N} \right).$$
(1.21)

Analog folgt

$$\frac{1}{T} \sum_{i=k^*+2}^{N} \left| \left(R_i - \frac{T}{N} \check{a}_{k^*} \right)^2 - \left(R_i - \frac{T}{N} a_T^* \right)^2 \right| = O_P \left(\frac{\sqrt{N-k^*}}{N} \right).$$
(1.22)

Für die Summe der quadratischen Zuwächse gilt mit (1.20)

$$\sum_{i=1}^{k^{*}} \left(R_{i} - \frac{T}{N} a_{T} \right)^{2} = \sum_{i=1}^{k^{*}} (b_{T} W_{i} + \eta_{i})^{2}$$

$$= b_{T}^{2} \sum_{i=1}^{k^{*}} W_{i}^{2} + 2b_{T} \sum_{i=1}^{k^{*}} W_{i} \eta_{i} + \sum_{i=1}^{k^{*}} \eta_{i}^{2}$$

$$= O_{P} \left(k^{*} \frac{T}{N} \right) + O_{P} \left(k^{*} T^{\alpha} \sqrt{\frac{T}{N}} \right) + O_{P} \left(k^{*} T^{2\alpha} \right)$$

$$= O_{P} \left(k^{*} \frac{T}{N} \right). \qquad (1.23)$$

Und analog ergibt sich

$$\sum_{i=k^*+2}^{N} \left(R_i - \frac{T}{N} a_T^* \right)^2 = O_P \left((N - k^*) \frac{T}{N} \right).$$
(1.24)

Somit gilt schließlich mit (1.21) und (1.23)

$$\frac{1}{T} \sum_{i=1}^{k^*} \left(R_i - \frac{T}{N} \hat{a}_{k^*} \right)^2 \\
\leq \frac{1}{T} \sum_{i=1}^{k^*} \left| \left(R_i - \frac{T}{N} \hat{a}_{k^*} \right)^2 - \left(R_i - \frac{T}{N} a_T \right)^2 \right| + \frac{1}{T} \sum_{i=1}^{k^*} \left(R_i - \frac{T}{N} a_T \right)^2 \\
= O_P \left(\frac{\sqrt{k^*}}{N} \right) + O_P \left(\frac{k^*}{N} \right) \\
= O_P \left(\frac{k^*}{N} \right).$$
(1.25)

Für die Summe nach dem Changepoint folgt mit (1.22) und (1.24)

$$\frac{1}{T}\sum_{i=k^*+2}^N \left(R_i - \frac{T}{N}\check{a}_{k^*}\right)^2 = O_P\left(\frac{N-k^*}{N}\right).$$

Daraus erhält man mit (1.25) die Behauptung.

Der Beweis zeigt auch den Grund des Auslassens des Zuwachses R_{k^*+1} in den Definitionen der Schätzer. Ansonsten würde der Schätzer \check{a}_{k^*} und die zweite Summe des Schätzers $\hat{b}_{k^*}^2$ auf einem falschen "Anfangswert" beruhen. Bei Alternativen mit einem Changepoint T^* , der für alle T > 0 mit der Beobachtungsstelle k^*T/N zusammenfällt, erübrigt sich dies.

Nach dem vorangegangen Satz ist $\hat{b}_{k^*}^2$ unter der Alternative für $k^* = [\theta_T N]$ stochastisch beschränkt (= $O_P(1)$). Für die Statistik \tilde{M}_T aus (1.19) liegt damit eine ausgezeichnete Situation vor. Denn bereits unter den Voraussetzungen des Satzes 1.3 und nur der zusätzlichen Forderung $N = O(T^{1-2\alpha})$ ist bereits die asymptotische Konsistenz der Statistik \tilde{M}_T gewährleistet, also

$$\tilde{M}_T \xrightarrow{P} \infty, \quad T \to \infty.$$

In der folgenden Diskussion, in der kein Anspruch auf mathematische Vollständigkeit erhoben wird, beschränken wir uns auf den am häufigsten vorkommenden Fall eines Changepoints $T^* = \theta_T T$ und $\theta_T \in (\epsilon, 1 - \epsilon)$, $0 < \epsilon < 1/2$, und gehen von keinen weiteren Kenntnissen über das Verhalten der Parameter a_T und a_T^* aus. Im Vergleich mit der Statistik \hat{M}_T aus (1.8) des vorangegangen Abschnitts ist zu erwarten, dass die hier vorliegende Statistik \tilde{M}_T "schneller" hohe Werte unter der Alternative bei wachsender Beobachtungsdauer T annimmt. Denn der Faktor $\Delta_T^2 N$, der unter diesen Voraussetzungen im Beweis des Korollars 1.7 die stochastische Konvergenz gegen Unendlich erzwingt, entspricht bei Übertragung auf den hier vorliegenden Fall der Statistik \tilde{M}_T dem Faktor $\Delta_T^2 T$. Dies ist sogar analog zu dem Fall der Statistik M_T aus (1.2), in der keine Schätzung des Parameters b_T^2 erfolgt, wie die Diskussion nach Satz 1.3 zeigt.

1.4 Der 2-stufige Schätzer

Ausgehend von der bekannten Lage des Changepoints unterteilten wir im letzten Abschnitt die Beobachtungen in zwei Mengen und erhielten so unter der Alternative Schätzer für die Parameter des Modells der Definition 1.1. Unter der Alternative ist die optimale Stelle der Aufteilung an der letzten Beobachtung vor dem Changepoint und unter der Nullhypothese an der letzten Beobachtung, so dass gar keine Aufteilung erfolgt. Mittels solch einer Aufteilung könnte man "optimale" Schätzer für die Parameter angeben.

Da jedoch unter der Alternative der Changepoint nicht bekannt ist, versuchen wir in diesem Abschnitt mittels eines Schätzers \hat{k}^* die optimale Stelle k^* der Aufteilung der Beobachtungen in zwei Mengen zu erhalten. Mittels dieser Schätzung \hat{k}^* definieren wir die Schätzer für die Parameter a_T und b_T^2 des Modells der Definition 1.1. Dazu seien bei festem T > 0 folgende Schätzer für

$$k^*: \qquad \hat{k}^* := \hat{k}^*_T := \operatorname{argmax}\{|Z^*_k|; k = 1, \dots N\}, \qquad (1.26)$$

$$a_T: \qquad \hat{a}_{\hat{k}^*} := \hat{a}_{\hat{k}^*,T} := \frac{N}{\hat{k}^*T} \sum_{i=1}^{k^*} \left(Z_T \left(i \frac{T}{N} \right) - Z_T \left((i-1) \frac{T}{N} \right) \right), \qquad (1.27)$$

$$b_T^2: \qquad \hat{b}_{\hat{k}^*}^2 := \hat{b}_{\hat{k}^*,T}^2 := \frac{N}{\hat{k}^*T} \sum_{i=1}^{\hat{k}^*} \left(Z_T\left(i\frac{T}{N}\right) - Z_T\left((i-1)\frac{T}{N}\right) - \frac{T}{N}\hat{a}_{\hat{k}^*} \right)^2 \qquad (1.28)$$

gegeben. Analog lassen sich Schätzer für die Parameter a_T^* und b_T^{*2} angeben, auf deren Behandlung verzichtet wird, da im Vordergrund das Ersetzen des Parameters b_T^2 in der Statistik M_T aus (1.2) steht. Die in diesem Abschnitt gewonnenen Resultate übertragen sich auf die Schätzer für die Parameter a_T^* und b_T^{*2} .

Unter der Nullhypothese erreicht man nicht, dass der Schätzer \hat{k}^* "dicht" an dem optimalen Wert N liegt, damit die Beobachtungen des Prozesses $\{Z_T(t) : 0 \le t \le T\}$ über die gesamte Beobachtungszeit T in die Schätzungen miteingehen. Doch reicht unter der Nullhypothese für eine konsistente Schätzung der Parameter a_T und b_T^2 mittels der Schätzer in (1.27) und (1.28) aus, dass die Werte von \hat{k}^* zumindest proportional zu der Anzahl N der Beobachtungen bei wachsender Beobachtungsdauer T zunehmen. Dazu das folgende Lemma:

Lemma 1.10

 $\{Z_T(t)\}\$ sei wie in Definition 1.1 für alle T > 0 und genüge dem Invarianzprinzip (1.3). Weiter sei $N = N(T) \rightarrow \infty$ für $T \rightarrow \infty$ erfüllt. Dann gilt unter H_0 :

$$\frac{\hat{k}^{*}}{N} \xrightarrow{P} \arg\max\{|B(t)|; 0 \le t \le 1\}, \quad T \to \infty,$$

wobei $\{B(t); 0 \le t \le 1\}$ eine Standard-Brownsche-Brücke ist.

Das Lemma zeigt, dass \hat{k}^* kein konsistenter Schätzer für N unter der Nullhypothese ist. Jedoch impliziert die Eigenschaft argmax $\{|B(t)| : 0 \le t \le 1\} \ne 0$ P-f.s. die stochastische Konvergenz des reziproken Quotienten N/\hat{k}^* . Für den proportionalen Fehler gilt somit die Aussage

$$\frac{N}{\hat{k}^*} = O_P(1), \quad T \to \infty.$$
(1.29)

Beweis von Lemma 1.10:

Wegen der Stetigkeit des Prozesses $\{B(t) : 0 \le t \le 1\}$ und der Kompaktheit des Trägers wird

$$t_0 := \operatorname{argmax}\{|B(t)| : 0 \le t \le 1\}$$

angenommen. Nach Satz 3.1 in [Fe95] gilt, dass t_0 P-f.s. eindeutig ist⁶. Da $\{Z_{[Ns]}^*: 0 \le s \le 1\}$ nur endlich viele Werte annimmt, wird auch

$$s_T := \operatorname{argmax}\{(Tb_T^2)^{-1/2} | Z^*_{[Ns]} | : 0 \le s \le 1\}$$

angenommen.

Wegen der P-f.s. Eindeutigkeit von t_0 gilt

$$\begin{aligned} \forall \, c > 0 : \quad P\left(|B(t_0)| - \sup_{\substack{0 \le t \le t_0 - c \\ t_0 + c \le t \le 1}} |B(t)| > 0\right) &= 1 \\ \Rightarrow \ 1 \ = \ P\left(\bigcup_{n=1}^{\infty} \left\{ |B(t_0)| - \sup_{\substack{0 \le t \le t_0 - c \\ t_0 + c \le t \le 1}} |B(t)| > \frac{1}{n} \right\} \right) \\ &= \ \lim_{n \to \infty} P\left(|B(t_0)| - \sup_{\substack{0 \le t \le t_0 - c \\ t_0 + c \le t \le 1}} |B(t)| > \frac{1}{n} \right). \end{aligned}$$

Für vorgegebene $\epsilon, c > 0$ erhält man somit ein geeignetes $\delta > 0$, so dass

$$P\left(|B(t_0)| \le \sup_{\substack{0 \le t \le t_0 - c \\ t_0 + c \le t \le 1}} |B(t)| + \delta\right) \le \frac{\epsilon}{2}.$$

Damit folgt

$$P(|s_{T} - t_{0}| \ge c) \le P\left(\sup_{\substack{0 \le t \le t_{0} - c \\ t_{0} + c \le t \le 1}} |B(t)| \ge |B(s_{T})|\right)$$

$$\le P\left(\sup_{\substack{0 \le t \le t_{0} - c \\ t_{0} + c \le t \le 1}} |B(t)| \ge |B(s_{T})|, \quad |B(t_{0})| > \sup_{\substack{0 \le t \le t_{0} - c \\ t_{0} + c \le t \le 1}} |B(t)| + \delta\right) + \frac{\epsilon}{2}$$

$$\le P(|B(t_{0})| > |B(s_{T})| + \delta) + \frac{\epsilon}{2}.$$
(1.30)

Sei $B_T(x) := \frac{1}{\sqrt{T}} (W_T(xT) - xW_T(T)), \quad 0 \le x \le 1,$ und $P_T(x) := \frac{1}{\sqrt{Tb_T^2}} Z^*_{[Nx]}, \quad 0 \le x \le 1.$

Dann gilt wegen des Invarianzprinzips (1.3) nach [HoSt98]

$$\exists T_0 > 0: \quad P\left(\sup_{0 \le x \le 1} \left| \frac{1}{\sqrt{Tb_T^2}} Z^*_{[Nx]} - B_T(x) \right| > \frac{1}{2}\delta \right) \le \frac{\epsilon}{2}, \quad \forall T \ge T_0.$$

⁶vgl. auch Bemerkung 3.1 und Satz 3.2 in [Fe95].

Damit folgt mittels (1.30) für $T \ge T_0$

$$P(|s_{T} - t_{0}| > c)$$

$$\leq P(|B_{T}(t_{0})| > |B_{T}(s_{T})| + \delta) + \frac{\epsilon}{2}$$

$$= P(|B_{T}(t_{0})| - |P_{T}(t_{0})| + |P_{T}(s_{T})| - |B_{T}(s_{T})| + |P_{T}(t_{0})| - |P_{T}(s_{T})| > \delta) + \frac{\epsilon}{2}$$

$$\leq P\left(2\sup_{0 \le x \le 1} |B_{T}(x) - P_{T}(x)| + |P_{T}(t_{0})| - |P_{T}(s_{T})| > \delta\right) + \frac{\epsilon}{2}$$

$$\leq P(|P_{T}(t_{0})| - |P_{T}(s_{T})| > 0) + \epsilon$$

$$= \epsilon.$$

Da $\hat{k}^* = Ns_T$, folgt die Behauptung.

Der folgende Nachweis der Konsistenz der Schätzer $\hat{a}_{\hat{k}*}$ und $\hat{b}_{\hat{k}*}^2$ unter der Nullhypothese gleicht dem analogen Beweis des vorigen Abschnitts. Jedoch erfolgt hier die Beobachtungsdauer des Prozesses $\{Z_T(t): 0 \le t \le T\}$ bis zu der "zufälligen" Zeit $\hat{k}*T/N$. Der auftretende Unterschied, dass das Argument des Prozesses $\{Y_T(t)\}$ nicht mehr deterministisch ist, verhindert nicht die Benutzung des Invarianzprinzips (1.3), da

$$\left|Y_T\left(\hat{k}^*\frac{T}{N}\right) - W_T\left(\hat{k}^*\frac{T}{N}\right)\right| \le \sup_{0\le t\le T} |Y_T(t) - W_T(t)| = O_P(T^{\alpha}).$$

Weiter wird die gültige Abschätzung

$$\frac{1}{\sqrt{T}} \left| W_T\left(\hat{k}^* \frac{T}{N}\right) \right| \le \sup_{0 \le x \le 1} \left| \frac{1}{\sqrt{T}} W_T(xT) \right|$$

benutzt, obwohl $\{W_T(\hat{k}^*t): t \ge 0\}$ kein Wiener-Prozess sein muss. Wir gelangen zu der Konsistenzaussage der Schätzer unter der Nullhypothese.

Satz 1.11

 $\{Z_T(t)\}\$ sei wie in Definition 1.1 für alle T > 0 und genüge dem Invarianzprinzip (1.3). Weiter sei $N = N(T) \rightarrow \infty$ für $T \rightarrow \infty$ erfüllt. Dann gilt unter H_0 :

a)
$$\hat{a}_{\hat{k}^*} - a_T = O_P(\frac{1}{\sqrt{T}}), \quad T \to \infty.$$

b) falls zusätzlich N und T/N monoton wachsend in T sind sowie

 $\frac{\log N}{\log \log T} \to \infty \quad f\ddot{u}r \ T \to \infty \quad und \quad \frac{N\log N}{T^{1-2\alpha}} \to 0 \quad f\ddot{u}r \ T \to \infty \quad gilt,$ dann folgt auch:

$$\hat{b}_{\hat{k}^*}^2 - b_T^2 = o_P(1), \quad T \to \infty.$$

Beweis:

a) Mit Invarianzprinzip (1.3) und (1.29) folgt

$$\hat{a}_{\hat{k}^*} - a_T = b_T \frac{N}{\hat{k}^* T} Y_T \left(\hat{k}^* \frac{T}{N} \right)$$

$$= b_T \frac{N}{\hat{k}^* T} \left(W_T \left(\hat{k}^* \frac{T}{N} \right) + O_P \left(T^{\alpha} \right) \right)$$

$$= O_P \left(\frac{1}{\sqrt{T}} \right) + O_P \left(T^{\alpha - 1} \right)$$

$$= O_P \left(\frac{1}{\sqrt{T}} \right). \qquad (1.31)$$

b) Ohne Einschränkung sei $a_T = 0$. Dann gilt mit (1.31)

$$\hat{b}_{\hat{k}^{*}}^{2} = \frac{N}{\hat{k}^{*}T} \sum_{i=1}^{\hat{k}^{*}} \left(R_{i} - \frac{T}{N} \hat{a}_{\hat{k}^{*}} \right)^{2}$$

$$= \frac{N}{\hat{k}^{*}T} \sum_{i=1}^{\hat{k}^{*}} R_{i}^{2} - \frac{T}{N} \hat{a}_{\hat{k}^{*}}^{2}$$

$$= \frac{N}{\hat{k}^{*}T} \sum_{i=1}^{\hat{k}^{*}} R_{i}^{2} + o_{P}(1). \qquad (1.32)$$

Seien im Folgenden $\epsilon,\delta>0$ beliebig vorgegeben.

Nach Voraussetzung existiert ein $B^2 > 0$, so dass $b_T^2 \leq B^2$ für alle T > 0.

Sei $W_i = W_T(iT/N) - W_T((i-1)T/N)$ für $1 \le i \le N$. Dann sind $\sqrt{\frac{N}{T}}W_i$ für $1 \le i \le N$ unabhängig und identisch verteilt, wobei $\sqrt{\frac{N}{T}}W_i \stackrel{\mathcal{D}}{=} N(0,1)$.

Seien $X_i \stackrel{\mathcal{D}}{=} N(0,1)$ für $i \in \mathbb{N}$ und unabhängig. Das starke Gesetz der großen Zahlen ergibt

$$\exists k_0 \in \mathbb{N}: \quad P\left(\sup_{k \ge k_0} \left| \frac{1}{k} \sum_{i=1}^k X_i^2 - 1 \right| \ge \frac{\delta}{B^2} \right) \le \frac{\epsilon}{4},$$

und damit für alle $N \geq k_0$

$$P\left(b_T^2 \max_{k_0 \le k \le N} \left| \frac{N}{kT} \sum_{i=1}^k W_i^2 - 1 \right| \ge \delta \right) \le \frac{\epsilon}{4}.$$
(1.33)

Eine Anwendung des Satzes 1.2.1 aus [CsRe81] ergibt

$$\max_{1 \le i \le N} |W_i| = O_P\left(\sqrt{\frac{T}{N}\log N}\right).$$

Sei $\eta_i=R_i-b_TW_i$ für $1\leq i\leq N.$ Dann gilt wegen des Invarianz
prinzips (1.3) und der Annahme $a_T=0$

$$\max_{1 \le i \le N} |\eta_i| = O_P(T^{\alpha}).$$

Mittels der zwei letzten Aussagen folgt für die Summe der quadratischen Zuwächse

$$\frac{N}{\hat{k}^*T} \sum_{i=1}^{\hat{k}^*} R_i^2 = \frac{N}{\hat{k}^*T} \sum_{i=1}^{\hat{k}^*} (b_T W_i + \eta_i)^2$$

$$= b_T^2 \frac{N}{\hat{k}^*T} \sum_{i=1}^{\hat{k}^*} W_i^2 + b_T \frac{2N}{\hat{k}^*T} \sum_{i=1}^{\hat{k}^*} \eta_i W_i + \frac{N}{\hat{k}^*T} \sum_{i=1}^{\hat{k}^*} \eta_i^2$$

$$= b_T^2 \frac{N}{\hat{k}^*T} \sum_{i=1}^{\hat{k}^*} W_i^2 + O_P \left(\sqrt{N \log N} T^{\alpha - 1/2}\right) + O_P \left(N T^{2\alpha - 1}\right)$$

$$= b_T^2 \frac{N}{\hat{k}^*T} \sum_{i=1}^{\hat{k}^*} W_i^2 + o_P \left(1\right).$$
(1.34)

Weiter gilt wegen (1.29): $\exists C > 0 \quad \exists T_0 > 0 : P\left(\frac{N}{C} > \hat{k}^*\right) \leq \frac{\epsilon}{4} \quad \forall T \geq T_0.$ Für T groß genug, jedoch mindestens so, dass $N/C \geq k_0$ ist, folgt mit (1.32) bis (1.34):

$$\begin{split} P\left(\left|\hat{b}_{\hat{k}^*}^2 - b_T^2\right| \ge 3\delta\right) &= P\left(\left|\frac{N}{\hat{k}^*T}\sum_{i=1}^{\hat{k}^*}\left(R_i - \frac{T}{N}\hat{a}_{\hat{k}^*}\right)^2 - b_T^2\right| \ge 3\delta\right) \\ &\leq P\left(\left|\frac{N}{\hat{k}^*T}\sum_{i=1}^{\hat{k}^*}R_i^2 - b_T^2\right| \ge 2\delta\right) + \frac{\epsilon}{4} \\ &\leq P\left(\left|b_T^2\frac{N}{\hat{k}^*T}\sum_{i=1}^{\hat{k}^*}W_i^2 - b_T^2\right| \ge \delta\right) + \frac{2}{4}\epsilon \\ &\leq P\left(\left|b_T^2\frac{N}{\hat{k}^*T}\sum_{i=1}^{\hat{k}^*}W_i^2 - b_T^2\right| \ge \delta, \quad \frac{N}{C} \le \hat{k}^*\right) + \frac{3}{4}\epsilon \\ &\leq P\left(b_T^2\max_{k_0 \le k \le N}\left|\frac{N}{kT}\sum_{i=1}^{k}W_i^2 - 1\right| \ge \delta\right) + \frac{3}{4}\epsilon \\ &\leq \epsilon \end{split}$$

und man erhält die Behauptung des Satzes.

Mittels des Schätzers $\hat{b}_{\hat{k}^*}^2$ aus (1.28) lässt sich wie gewohnt der unbekannte Parameter b_T^2 in der Statistik M_T aus (1.2) ersetzen. Man erhält als Statistik für festes T > 0

$$\bar{M}_T := \frac{1}{\sqrt{T\hat{b}_{\hat{k}^*}^2}} \max_{1 \le k \le N} |Z_k^*| \,. \tag{1.35}$$

Dabei gilt $\hat{b}_{\hat{k}^*}^2 \ge c/2 > 0$ auf Mengen von beliebig hohen Wahrscheinlichkeiten für T groß genug, falls $b_T^2 \ge c > 0$ für alle T > 0 sowie die Voraussetzungen des Satzes 1.11.b) erfüllt sind.

Aufgrund der Konsistenz des Schätzers $\hat{b}_{\hat{k}^*}^2$ bleibt die Grenzverteilung der Statistik M_T gemäß Satz 1.2 für die Statistik \bar{M}_T unter der Nullhypothese bei Einhalten der Voraussetzungen des Satzes 1.11.b) erhalten.

Zur Untersuchung der Schätzer unter der Alternative wird eine Modifikation der Invarianzprinzipien (1.3) und (1.4) benötigt. Es wird wie bei Kühn [Kühn99] gefordert:

Es exisitiert $\alpha \in [0, \frac{1}{2})$ und $\delta \in \mathbb{R}^*_+$ sowie

es existieren für alle T > 0 unabhängige Wiener-Prozesse $\{W_T(t) : 0 \le t \le T^*\}$ und $\{W_T^*(t) : 0 \le t \le T - T^*\}$ auf (Ω, \mathcal{A}, P) , so dass gilt:

$$\sup_{\delta \le t \le T^*} \left| \frac{Y_T(t) - W_T(t)}{t^{\alpha}} \right| = O_P(1), \quad T \to \infty,$$
(1.36)

$$\sup_{5 \le t \le T - T^*} \left| \frac{Y_T^*(t) - W_T^*(t)}{t^{\alpha}} \right| = O_P(1), \quad T \to \infty.$$
(1.37)

Weiter werden mit denselben Parametern α und δ , symmetrisch zu den oben genannten, folgende Invarianzprinzipien gefordert:

Es existieren für alle T > 0 unabhängige Wiener-Prozesse $\{\tilde{W}_T(t) : 0 \le t \le T^*\}$ und $\{\tilde{W}_T^*(t) : 0 \le t \le T - T^*\}$ auf (Ω, \mathcal{A}, P) , so dass gilt:

$$\sup_{\delta \le t \le T^*} \left| \frac{Y_T(T^*) - Y_T(T^* - t) - \tilde{W}_T(t)}{t^{\alpha}} \right| = O_P(1), \quad T \to \infty,$$
(1.38)

$$\sup_{\delta \le t \le T - T^*} \left| \frac{Y_T^*(T - T^*) - Y_T^*(T - T^* - t) - \tilde{W}_T^*(t)}{t^{\alpha}} \right| = O_P(1), \quad T \to \infty.$$
(1.39)

Dabei seien die Wiener-Prozesse $\{W_T(t): 0 \le t \le T^*\}$ von $\{\tilde{W}_T^*(t): 0 \le t \le T - T^*\}$ sowie $\{W_T^*(t): 0 \le t \le T - T^*\}$ von $\{\tilde{W}_T(t): 0 \le t \le T^*\}$ für alle T > 0 unabhängig.

In der Nähe des Changepoints wird ein moderates Verhalten der Prozesse $\{Y_T(t)\}$ und $\{Y_T^*(t)\}$ verlangt:

$$\sup_{0 \le t \le \delta} |Y_T(T^*) - Y_T(T^* - t)| = O_P(1), \quad T \to \infty,$$
(1.40)

$$\sup_{0 \le t \le \delta} |Y_T^*(t)| = O_P(1), \quad T \to \infty.$$
(1.41)

Bei den Invarianzprinzipien (1.3) und (1.4) wird die Approximation durch die Wiener-Prozesse bis an den linken Rand der Intervalle $[0, T^*]$ und $[0, T - T^*]$ gefordert, dagegen ist dies bei den neuen Invarianzprinzipien (1.36) und (1.37) lediglich bis zu einem Randabstand $\delta > 0$ möglich. Wegen der diskreten Beobachtungen und $T/N \to \infty$ für $T \to \infty$ reicht eine Approximation bis zu einem Randabstand $\delta > 0$ für die Anwendungen des Invarianzprinzips (1.3) aus. Die Approximation gemäß Invarianzprinzip (1.4) bis zum Rand wird durch das Erfüllen des neu geforderten Invarianzprinzips (1.37) in Verbindung mit der Bedingung (1.41) impliziert. In diesem Sinne sind die hier betrachteten Invarianzprinzipien (1.36) und (1.37) eine Verschärfung der bisherigen (1.3) und (1.4). Die weiteren Invarianzprinzipien (1.38) und (1.39) verlangen eine analoge, unabhängige Approximation der umgekehrten Prozesse $Y_T(T^*) - Y_T(T^* - t)$ und $Y_T^*(T - T^*) - Y_T^*(T - T^* - t)$. Trotz der stärkeren Approximationsforderungen erfüllen die zu Beginn dieses Kapitels genannten Beispiele von Prozessen, die dem Modell der Definition 1.1 sowie den Invarianzprinzipien (1.3) und (1.4) genügen, auch die Invarianzprinzipien (1.36) bis (1.39). Die Nachweise sowie die konkreten Definitionen dieser Prozesse finden sich in [Kühn99].

Mittels der neuen Invarianzprinzipien lässt sich die schwache Konsistenz des Schätzers \hat{k}^* unter der Alternative nachweisen. Es werden hier Alternativen mit Changepoint $T^* = \theta T$ und festem $\theta \in (0, 1)$ sowie von T unabhängigen Parametern a_T und a_T^* betrachtet.

Satz 1.12

 $\{Z_T(t)\}\$ sei wie in Definition 1.1 mit $a_T = a$ und $a_T^* = a^*$ für alle T > 0 und genüge den Invarianzprinzipien (1.36) bis (1.39) sowie der Bedingung (1.40).

Weiter set $N = N(T) \rightarrow \infty$ für $T \rightarrow \infty$ erfüllt.

Dann gilt unter H_A mit $T^* = \theta T$, $\theta \in (0, 1)$:

$$\ddot{k}^* - k^* = O_P(1).$$

An dieser Stelle sei noch einmal darauf aufmerksam gemacht, dass für die Schätzung \hat{k}^* des Parameters k^* keine weiteren Daten außer an den äquidistanten Beobachtungspunkten $t_i = i\frac{T}{N}, 1 \leq i \leq N$, gefordert werden. Bei kontinuierlichen Beobachtungen dagegen erhält man mit

$$\hat{T}^* = \operatorname{argmax}\left\{ \left| Z_T(t) - \frac{t}{T} Z_T(T) \right| : t \in [\delta, T - \delta] \right\}$$

einen schwach konsistenten Schätzer⁷ für den Changepoint T^* . Damit lässt sich mittels

$$\tilde{k}^* = \left[\frac{\hat{T}^*N}{T}\right]$$

sogar ein Schätzer des Parameters k^* gewinnen, für den $P(\tilde{k}^* \in \{k^* - 1, k^*, k^* + 1\}) \to 1$ und $\max{\{\tilde{k}^*T/N - T^*, 0\}} = O_P(1)$ für $T \to \infty$ gilt, wobei $k^* = [NT^*/T]$. Dies erhält man jedoch nur auf Kosten von zusätzlichen Beobachtungen, die nicht weiter in die Varianzschätzung eingehen.

Der Beweis der schwachen Konsistenz des Schätzers \hat{k}^* erfolgt ähnlich wie in [Kühn99] für die schwache Konsistenz von \hat{T}^* , jedoch mit einigen kleinen technischen Erweiterungen, da in dem hier vorliegenden Fall die Schätzung \hat{k}^* nur auf diskreten Beobachtungen beruht und somit im Allgemeinen keine Beobachtung am tatsächlichen Changepoint vorliegt. Dazu werden folgende drei Lemmata benötigt, in denen $T^* = \theta T$ für festes $\theta \in (0, 1)$ sowie $k^* =$ $[\theta N]$ sei.

Lemma 1.13

a) Seien $\{Y_T(t): 0 \le t \le T^*\}$ für alle T > 0 stochastische Prozesse auf (Ω, \mathcal{A}, P) , die dem Invarianzprinzip (1.38) genügen. Dann gilt für beliebiges $u \in \mathbb{R}^+_+$:

$$\forall \epsilon > 0 \quad \exists t_0 = t_0(\epsilon) \ge \delta \quad \exists T_0 \ge 0 :$$

$$P\left(\sup_{0 \le t \le T^* - t_0} \frac{Y_T(T^*) - Y_T(t)}{T^* - t} \ge u\right) \le \epsilon, \quad \forall T \ge T_0.$$

⁷vgl. hierzu [Kühn99].

 b) Seien {Y_T^{*}(t) : 0 ≤ t ≤ T − T^{*}} für alle T > 0 stochastische Prozesse auf (Ω, A, P), die dem Invarianzprinzip (1.37) genügen. Dann gilt für beliebiges u ∈ ℝ^{*}₊:

$$\forall \epsilon > 0 \quad \exists t_0 = t_0(\epsilon) \ge \delta \quad \exists T_0 \ge 0 : \\ P\left(\sup_{t_0 \le t \le T - T^*} \frac{Y_T^*(t)}{t} \ge u\right) \le \epsilon, \quad \forall T \ge T_0.$$

Beweis:

b) Mit Invarianzprinzip (1.37) gilt

$$\sup_{\delta \le t \le T - T^*} \left| \frac{Y_T^*(t) - W_T^*(t)}{\sqrt{t \log \log t}} \right| = O_P(1)$$

und mit Hilfe des Gesetzes vom iterierten Logarithmus für Wiener-Prozesse folgt daraus für beliebiges $\epsilon>0$

$$\exists T_0, C > 0 : P\left(\sup_{\delta \le t \le T - T^*} \frac{Y_T^*(t)}{\sqrt{t \log \log t}} \ge C\right) \le \epsilon, \quad \forall T \ge T_0.$$

Wähle $t_0 > \delta$, so dass $\sqrt{\frac{\log \log t}{t}} \le \frac{u}{C}$ für alle $t \ge t_0$. Dann folgt die Behauptung für $T \ge T_0$ aus

$$P\left(\sup_{t_0 \le t \le T - T^*} \frac{Y_T^*(t)}{t} \ge u\right) \le P\left(\sup_{t_0 \le t \le T - T^*} \frac{Y_T^*(t)}{\sqrt{t \log \log t}} \ge C\right)$$
$$\le \epsilon.$$

a) Der Beweis erfolgt wie oben mit Invarianzprinzip (1.38).

Lemma 1.14

a) Seien {Y_T(t) : 0 ≤ t ≤ T*} für alle T > 0 stochastische Prozesse auf (Ω, A, P), die dem Invarianzprinzip (1.38) und der Bedingung (1.40) genügen. Weiter sei N = N(T) → ∞ für T → ∞ erfüllt und b_T sei für T → ∞ beschränkt. Dann gilt für beliebige u, v ∈ ℝ^{*}₊:

$$\begin{aligned} \forall \epsilon > 0 \quad \exists d \in \mathbb{N} \quad \exists T_0 \ge 0 : \\ P\left(b_T\left(Y_T\left(k\frac{T}{N}\right) - Y_T\left(k^*\frac{T}{N}\right)\right) \\ \ge u\left(T^* - k\frac{T}{N}\right) - v\left(T^* - k^*\frac{T}{N}\right), \quad k \in \{1, \dots, k^* - d\}\right) \le \epsilon, \quad T \ge T_0. \end{aligned}$$

b) Seien $\{Y_T(t): 0 \le t \le T^*\}$ für alle T > 0 stochastische Prozesse auf (Ω, \mathcal{A}, P) , die dem Invarianzprinzip (1.38) und der Bedingung (1.40) genügen. Ebenfalls seien $\{Y_T^*(t): 0 \le t \le T - T^*\}$ für alle T > 0 stochastische Prozesse auf (Ω, \mathcal{A}, P) , die dem Invarianzprinzip (1.37) genügen. Weiter sei $N = N(T) \rightarrow \infty$ für $T \rightarrow \infty$ erfüllt und b_T, b_T^* seien für $T \rightarrow \infty$ beschränkt. Dann gilt für beliebige $u, v \in \mathbb{R}_+^*$:

$$\forall \epsilon > 0 \quad \exists d \in \mathbb{N} \quad \exists T_0 \ge 0 : P\left(b_T\left(Y_T\left(T^*\right) - Y_T\left(k^*\frac{T}{N}\right)\right) + b_T^*Y_T^*\left(k\frac{T}{N} - T^*\right) \\ \ge u\left(k\frac{T}{N} - T^*\right) - v\left(T^* - k^*\frac{T}{N}\right), \quad k \in \{k^* + d, \dots, N\} \right) \le \epsilon \quad T \ge T_0.$$

Beweis:

b) Ohne Einschränkung sei $(k^* + 2) \frac{T}{N} \ge \delta$. 1. Fall: $T^* - k^* \frac{T}{N} > \delta > 0$. Sei $\epsilon > 0$ beliebig vorgegeben. Eine Anwendung des Invarianzprinzips (1.38) ergibt

$$\exists C = C(\epsilon) > 0 \quad \exists T_1 = T_1(\epsilon) > 0 :$$

$$P\left(b_T \frac{Y_T(T^*) - Y_T\left(k^* \frac{T}{N}\right)}{T^* - k^* \frac{T}{N}} \ge C, \quad T^* - k^* \frac{T}{N} > \delta\right) \le \frac{\epsilon}{2}, \quad \forall T \ge T_1.$$

Weiter folgt aus Lemma 1.13b)

$$\exists d_1 \in \mathbb{N} \quad \exists T_2 = T_2(\epsilon) > 0 :$$

$$P\left(b_T^* \max_{k^* + d_1 \le k \le N} \frac{Y_T^* \left(k \frac{T}{N} - T^*\right)}{k \frac{T}{N} - T^*} \ge \frac{1}{2}u\right) \le \frac{\epsilon}{2}, \quad \forall T \ge T_2$$

Ferner gilt sicherlich für alle T > 0 die Abschätzung

$$\exists d_2 \in \mathbb{N} : \frac{1}{2}u\left(k\frac{T}{N} - T^*\right) \ge (v+C)\left(T^* - k^*\frac{T}{N}\right) \quad \forall k \ge k^* + d_2$$

Somit erhält man insgesamt für $d = \max\{d_1, d_2\}$ und $T \ge T_0 = \max\{T_1, T_2\}$

$$\begin{split} P\left(b_{T}\left(Y_{T}\left(T^{*}\right)-Y_{T}\left(k^{*}\frac{T}{N}\right)\right)+b_{T}^{*}Y_{T}^{*}\left(k\frac{T}{N}-T^{*}\right)\\ &\geq u\left(k\frac{T}{N}-T^{*}\right)-v\left(T^{*}-k^{*}\frac{T}{N}\right), \quad k\in\{k^{*}+d,\ldots,N\}, \quad T^{*}-k^{*}\frac{T}{N}>\delta\right)\\ &\leq P\left(b_{T}^{*}Y_{T}^{*}\left(k\frac{T}{N}-T^{*}\right)\geq u\left(k\frac{T}{N}-T^{*}\right)-(v+C)\left(T^{*}-k^{*}\frac{T}{N}\right), \\ &\quad k\in\{k^{*}+d,\ldots,N\}\right)+\frac{\epsilon}{2}\\ &\leq P\left(b_{T}^{*}Y_{T}^{*}\left(k\frac{T}{N}-T^{*}\right)\geq\frac{1}{2}u\left(k\frac{T}{N}-T^{*}\right), \quad k\in\{k^{*}+d,\ldots,N\}\right)+\frac{\epsilon}{2}\\ &\leq P\left(b_{T}^{*}\max_{k^{*}+d\leq k\leq N}\frac{Y_{T}^{*}\left(k\frac{T}{N}-T^{*}\right)}{k\frac{T}{N}-T^{*}}\geq\frac{1}{2}u\right)+\frac{\epsilon}{2}\\ &\leq \epsilon. \end{split}$$

2. Fall: $T^* - k^* \frac{T}{N} \leq \delta$. Sei $\epsilon > 0$ beliebig vorgegeben.

In diesem Fall benutzt man statt des Invarianzprinzips (1.38) die Beschränktheit (1.40) und erhält:

$$\exists C = C(\epsilon) > 0 \quad \exists T_1 = T_1(\epsilon) > 0 :$$
$$P\left(b_T\left(Y_T(T^*) - Y_T\left(k^*\frac{T}{N}\right)\right) \ge C, \quad T^* - k^*\frac{T}{N} \le \delta\right) \le \frac{\epsilon}{2}, \quad \forall T \ge T_1.$$

Wie im ersten Fall gilt aufgrund des Lemmas 1.13b)

$$\exists d_1 \in \mathbb{N} \quad \exists T_2 = T_2(\epsilon) > 0 :$$

$$P\left(b_T^* \max_{k^* + d_1 \le k \le N} \frac{Y_T^* \left(k\frac{T}{N} - T^*\right)}{k\frac{T}{N} - T^*} \ge \frac{1}{2}u\right) \le \frac{\epsilon}{2} \quad \forall T \ge T_2$$

Hier benötigt man für alle T > 0 die Abschätzung

$$\exists d_2 \in \mathbb{N} : \quad \frac{1}{2}u\left(k\frac{T}{N} - T^*\right) \ge v\delta + C \quad \forall k \ge k^* + d_2.$$

Insgesamt ergibt sich so für $d = \max\{d_1, d_2\}$ und $T \ge T_0 = \max\{T_1, T_2\}$

$$P\left(b_{T}\left(Y_{T}\left(T^{*}\right)-Y_{T}\left(k^{*}\frac{T}{N}\right)\right)+b_{T}^{*}Y_{T}^{*}\left(k\frac{T}{N}-T^{*}\right)\right)$$

$$\geq u\left(k\frac{T}{N}-T^{*}\right)-v\left(T^{*}-k^{*}\frac{T}{N}\right), \quad k \in \{k^{*}+d,\ldots,N\}, \quad T^{*}-k^{*}\frac{T}{N} \leq \delta\right)$$

$$\leq P\left(b_{T}^{*}Y_{T}^{*}\left(k\frac{T}{N}-T^{*}\right)\geq u\left(k\frac{T}{N}-T^{*}\right)-\left(v\delta+C\right), \quad k \in \{k^{*}+d,\ldots,N\}\right)+\frac{\epsilon}{2}$$

$$\leq P\left(b_{T}^{*}Y_{T}^{*}\left(k\frac{T}{N}-T^{*}\right)\geq \frac{1}{2}u\left(k\frac{T}{N}-T^{*}\right), \quad k \in \{k^{*}+d,\ldots,N\}\right)+\frac{\epsilon}{2}$$

$$\leq P\left(b_{T}^{*}\max_{k^{*}+d \leq k \leq N}\frac{Y_{T}^{*}\left(k\frac{T}{N}-T^{*}\right)}{k\frac{T}{N}-T^{*}}\geq \frac{1}{2}u\right)+\frac{\epsilon}{2}$$

$$\leq \epsilon.$$

Beide Fälle zusammen ergeben die Aussage b).

a) Der Beweis erfolgt analog mit Lemma 1.13.a).

Für das letzte Lemma, das für den Beweis der Konsistenz des Schätzers \hat{k}^* notwendig ist, wird noch folgende Notation⁸ eingeführt:

$$\hat{k}^*_+ := \hat{k}^*_+(T) := \operatorname{argmax} \{ Z^*_k : k = 1, \dots, N \} , \hat{k}^*_- := \hat{k}^*_-(T) := \operatorname{argmin} \{ Z^*_k : k = 1, \dots, N \} .$$

⁸argmin{ $Z_k^*: k = 1, ..., N$ } := min{ $k: Z_k^* = \min_{1 \le l \le N} Z_l^*$ }.

Lemma 1.15

 $\{Z_T(t)\}\$ sei wie in Definition 1.1 mit $a_T = a$ und $a_T^* = a^*$ für alle T > 0 und genüge den Invarianzprinzipien (1.36) sowie (1.39). Dann gilt unter H_A und $T^* = \theta T$, $\theta \in (0, 1)$:

a) falls
$$a^* < a$$
: $P\left(\hat{k}^* = \hat{k}^*_+\right) \to 1, \quad T \to \infty.$
b) falls $a^* > a$: $P\left(\hat{k}^* = \hat{k}^*_-\right) \to 1, \quad T \to \infty.$

Beweis: siehe Lemma 1.2.6 in [Kühn99].

Beweis von Satz 1.12:

Wir betrachten den Fall $a^* < a$; der Beweis von $a^* > a$ erfolgt analog mit Abschätzungen in die andere Richtung.

1. Fall: $k^* + 1 \le \hat{k}^*$. Sei $k \ge k^* + 1$. Dann gilt

$$Z_{k}^{*} - Z_{k}^{**} = Z_{T} \left(k \frac{T}{N} \right) - \frac{k}{N} Z_{T}(T) - Z_{T} \left(k^{*} \frac{T}{N} \right) + \frac{k^{*}}{N} Z_{T}(T)$$

$$= \left(k \frac{T}{N} - k^{*} \frac{T}{N} \right) \left((a^{*} - a) \theta - \frac{b_{T} Y_{T} (T^{*}) + b_{T}^{*} Y_{T}^{*} (T - T^{*})}{T} \right)$$

$$- (a^{*} - a) \left(T^{*} - k^{*} \frac{T}{N} \right) + \tilde{Y}_{T}(k),$$
(1.42)

wobei $\tilde{Y}_T(k) := b_T^* Y_T^* \left(k \frac{T}{N} - T^* \right) + b_T \left(Y_T(T^*) - Y_T \left(k^* \frac{T}{N} \right) \right), \quad k^* + 1 \le k \le N.$

Sei $\epsilon > 0$ beliebig vorgegeben.

Wegen der Invarianzprinzipien (1.36) und (1.37) gilt

$$\exists T_1 \ge 0: P\left(\frac{1}{T}\left(b_T Y_T\left(T^*\right) + b_T^* Y_T^*\left(T - T^*\right)\right) \ge \frac{1}{2}\theta(a^* - a)\right) \ge 1 - \frac{\epsilon}{2} \quad \forall T \ge T_1. (1.43)$$

Damit erhält man mit (1.42)

$$P\left(Z_{k}^{*} - Z_{k}^{*} \leq \frac{1}{2}\theta\left(a^{*} - a\right)\left(k\frac{T}{N} - k^{*}\frac{T}{N}\right) - (a^{*} - a)\left(T^{*} - k^{*}\frac{T}{N}\right) + \tilde{Y}_{T}(k), \forall k \in \{k^{*} + 1, \dots, N\}\right) \geq 1 - \frac{\epsilon}{2} \quad \forall T \geq T_{1}.$$
(1.44)

Mit Lemma 1.14b) gilt

$$\exists d \in \mathbb{N} \quad \exists T_2 \ge 0:$$

$$P\left(\tilde{Y}_T(k) \ge -\frac{1}{2}\theta \left(a^* - a\right) \left(k\frac{T}{N} - k^*\frac{T}{N}\right) + (a^* - a) \left(T^* - k^*\frac{T}{N}\right), \ k \in \{k^* + d, \dots, N\}\right)$$

$$= P\left(\tilde{Y}_{T}(k) \geq -\frac{1}{2}\theta\left(a^{*}-a\right)\left(k\frac{T}{N}-T^{*}\right)-\left(a^{*}-a\right)\left(T^{*}-k^{*}\frac{T}{N}\right)\left(\frac{1}{2}\theta-1\right),$$

$$k \in \left\{k^{*}+d,\ldots,N\right\}\right)$$

$$\leq \frac{\epsilon}{2} \quad \forall T \geq T_{2}.$$
(1.45)

Damit ergibt sich mit (1.44) und (1.45) für $T \ge \max\{T_1, T_2\}$

$$P\left(\hat{k}_{+}^{*} < k^{*} + d\right)$$

$$\geq P\left(Z_{k}^{*} - Z_{k^{*}}^{*} < 0, \quad \forall k \in \{k^{*} + d, \dots, N\}\right)$$

$$\geq P\left(\frac{1}{2}\theta\left(a^{*} - a\right)\left(k\frac{T}{N} - k^{*}\frac{T}{N}\right) - (a^{*} - a)\left(T^{*} - k^{*}\frac{T}{N}\right) + \tilde{Y}_{T}(k) < 0,$$

$$\forall k \in \{k^{*} + d, \dots, N\}\right) - \frac{\epsilon}{2}$$

$$\geq 1 - \epsilon.$$

Daraus folgt

$$\max\{\hat{k}_{+}^{*} - k^{*}, 0\} = O_{P}(1)$$

und somit mit Lemma 1.15 auch

$$\max\{\hat{k}^* - k^*, 0\} = O_P(1).$$

2. Fall: $k^* \ge \hat{k}^*$.

Sei $k \leq k^*$. Dann erhält man wie im ersten Fall

$$Z_{k}^{*} - Z_{k^{*}}^{*} = \left(k^{*}\frac{T}{N} - k\frac{T}{N}\right) \left((a^{*} - a)\left(1 - \theta\right) + \frac{1}{T}\left(b_{T}Y_{T}(T^{*}) + b_{T}^{*}Y_{T}^{*}\left(T - T^{*}\right)\right)\right) - b_{T}\left(Y_{T}\left(k^{*}\frac{T}{N}\right) - Y_{T}\left(k\frac{T}{N}\right)\right).$$

Eine analoge Argumentation wie im ersten Fall und Benutzung des Lemmas 1.14a) ergibt

$$\min\{\hat{k}_{+}^{*} - k^{*}, 0\} = O_{P}(1)$$

und damit folgt mit Lemma 1.15

$$\min\{\hat{k}^* - k^*, 0\} = O_P(1).$$

Beide Fälle zusammen implizieren die Behauptung.

Mittels der schwachen Konsistenz des Schätzers \hat{k}^* nach Satz 1.12 lässt sich das Verhalten der Schätzer $\hat{a}_{\hat{k}^*}$ und $\hat{b}_{\hat{k}^*}^2$ unter der Alternative bestimmen. Zuerst wird eine Abschätzung der Differenz des Prozesses $\{Z_T(t)\}$ an der geschätzten Stelle \hat{k}^*T/N von der wahren Stelle k^*T/N der letzten Beobachtung vor dem Changepoint T^* benötigt.

Lemma 1.16

 $\{Z_T(t)\}\$ sei wie in Definition 1.1 mit $a_T = a$ und $a_T^* = a^*$ für alle T > 0 und genüge den Invarianzprinzipien (1.36) bis (1.39) sowie den Bedingungen (1.40) und (1.41). Weiter sei $N = N(T) \to \infty$ für $T \to \infty$ sowie $T/N \to \infty$ für $T \to \infty$ erfüllt. Dann gilt unter H_A und $T^* = \theta T$, $\theta \in (0, 1)$:

$$Z_T\left(\hat{k}^*\frac{T}{N}\right) = Z_T\left(k^*\frac{T}{N}\right) + O_P\left(\frac{T}{N}\right) + O_P\left(T^{\alpha}\right), \quad T \to \infty.$$

Beweis:

1. Fall: $\hat{k}^* \leq k^*$. Für $k^* - \hat{k}^* \leq C$ für ein C > 0 gilt

$$\left| W_T\left(k^* \frac{T}{N}\right) - W_T\left(\hat{k}^* \frac{T}{N}\right) \right| \le \max_{k^* - C \le i \le k^*} \left| W_T\left(k^* \frac{T}{N}\right) - W_T\left(i \frac{T}{N}\right) \right|$$

und damit folgt mit Invarianzprinzip (1.36) und Satz 1.12

$$Z_T\left(k^*\frac{T}{N}\right) - Z_T\left(\hat{k}^*\frac{T}{N}\right)$$

= $a\frac{T}{N}\left(k^* - \hat{k}^*\right) + b_T\left(Y_T\left(k^*\frac{T}{N}\right) - Y_T\left(\hat{k}^*\frac{T}{N}\right)\right)$
= $O_P\left(\frac{T}{N}\right) + b_T\left(W_T\left(k^*\frac{T}{N}\right) - W_T\left(\hat{k}^*\frac{T}{N}\right)\right) + O_P(T^{\alpha})$
= $O_P\left(\frac{T}{N}\right) + O_P(T^{\alpha})$.

2. Fall: $\hat{k}^* > k^*$.

Die Behauptung folgt analog mit den Invarianzprinzipien (1.36) und (1.37) sowie der Bedingung (1.41).

Lemma 1.17

 $\{Z_T(t)\}\$ sei wie in Definition 1.1 mit $a_T = a$ und $a_T^* = a^*$ für alle T > 0 und genüge den Invarianzprinzipien (1.36) bis (1.39) sowie den Bedingungen (1.40) und (1.41). Weiter sei $N = N(T) \rightarrow \infty$ für $T \rightarrow \infty$ sowie $T/N \rightarrow \infty$ für $T \rightarrow \infty$ erfüllt. Dann gilt unter H_A mit $T^* = \theta T$, $\theta \in (0, 1)$:

$$\hat{a}_{\hat{k}^*} - a = O_P\left(\max\left\{\frac{1}{N}, \frac{1}{\sqrt{T}}\right\}\right), \quad T \to \infty.$$

Beweis: Mit Lemma 1.16 folgt

$$\hat{a}_{\hat{k}^*} - a = \frac{N}{\hat{k}^* T} Z_T \left(\hat{k}^* \frac{T}{N} \right) - a$$

$$= \frac{N}{k^* + O_P(1)} \frac{1}{T} \left(Z_T \left(k^* \frac{T}{N} \right) + O_P \left(\frac{T}{N} \right) + O_P \left(T^{\alpha} \right) \right) - a$$

$$= O_P \left(\frac{1}{T} \right) W_T \left(k^* \frac{T}{N} \right) + O_P \left(\frac{1}{N} \right) + O_P \left(T^{\alpha-1} \right)$$

$$= O_P \left(\frac{1}{\sqrt{T}} \right) + O_P \left(\frac{1}{N} \right).$$

Satz 1.18

 $\{Z_T(t)\}\$ sei wie in Definition 1.1 mit $a_T = a$ und $a_T^* = a^*$ für alle T > 0 und genüge den Invarianzprinzipien (1.36) bis (1.39) sowie den Bedingungen (1.40) und (1.41). Weiter sei $N = N(T) \to \infty$ für $T \to \infty$ und $N = o(T^{1-2\alpha})$ für $T \to \infty$ erfüllt. Dann gilt unter H_A mit $T^* = \theta T$, $\theta \in (0, 1)$:

$$\hat{b}_{\hat{k}^*}^2 - b_T^2 = O_P\left(\max\left\{\sqrt{\frac{N}{T}}T^{2\alpha}, \frac{1}{\sqrt{N}}, \frac{T}{N^2}\right\}\right).$$

Beweis: In dem Beweis werden folgende Notationen benutzt:

$$b_{1}^{2} := b_{1,T}^{2} := \frac{N}{k^{*}T} \sum_{i=1}^{k^{*}} \left(Z_{T} \left(i \frac{T}{N} \right) - Z_{T} \left((i-1) \frac{T}{N} \right) - \frac{T}{N} \hat{a}_{\hat{k}^{*}} \right)^{2},$$

$$b_{2}^{2} := b_{2,T}^{2} := \frac{N}{k^{*}T} \sum_{i=1}^{k^{*}} \left(Z_{T} \left(i \frac{T}{N} \right) - Z_{T} \left((i-1) \frac{T}{N} \right) - \frac{T}{N} a \right)^{2},$$

$$W_{i} := W_{T} \left(i \frac{T}{N} \right) - W_{T} \left((i-1) \frac{T}{N} \right), \quad 1 \le i \le k^{*},$$

$$\eta_{i} := R_{i} - b_{T} W_{i} - \frac{T}{N} a, \quad 1 \le i \le k^{*}.$$

Aufgrund der Markov-Ungleichung gilt

$$\sum_{i=1}^{k^*} |W_i| = O_P\left(k^*\sqrt{\frac{T}{N}}\right) \tag{1.46}$$

und wegen des Invarianzprinzips (1.36) auch

$$\max_{1 \le i \le k^*} |\eta_i| = O_P(T^{\alpha}).$$
(1.47)

1. z.z:
$$b_2^2 - b_1^2 = O_P\left(\sqrt{\frac{T}{N^3}}\right) + O_P\left(\frac{T}{N^3}\right) + O_P\left(\frac{1}{\sqrt{N}}\right)$$
.
Mit (1.46) und (1.47) erhält man

$$\sum_{i=1}^{k^*} \left| R_i - \frac{T}{N} a \right| = \sum_{i=1}^{k^*} \left| b_T W_i + \eta_i \right|$$
$$= O_P \left(k^* \sqrt{\frac{T}{N}} \right) + O_P \left(k^* T^{\alpha} \right)$$
$$= O_P \left(k^* \sqrt{\frac{T}{N}} \right).$$

Damit folgt

$$|b_2^2 - b_1^2| = \frac{N}{k^*T} \left| \sum_{i=1}^{k^*} \left(R_i - \frac{T}{N} \hat{a}_{\hat{k}^*} \right)^2 - \left(R_i - \frac{T}{N} a \right)^2 \right|$$

$$\leq \frac{N}{k^*T} \left(2\frac{T}{N} \left| \hat{a}_{\hat{k}^*} - a \right| \sum_{i=1}^{k^*} \left| R_i - \frac{T}{N} a \right| + k^* \frac{T^2}{N^2} \left(\hat{a}_{\hat{k}^*} - a \right)^2 \right)$$

$$= O_P \left(\max\left\{ \frac{1}{N}, \frac{1}{\sqrt{T}} \right\} \right) O_P \left(\sqrt{\frac{T}{N}} \right) + \frac{T}{N} O_P \left(\max\left\{ \frac{1}{N^2}, \frac{1}{T} \right\} \right)$$

$$= O_P \left(\sqrt{\frac{T}{N^3}} \right) + O_P \left(\frac{T}{N^3} \right) + O_P \left(\frac{1}{\sqrt{N}} \right).$$
(1.48)

2. z.z.:
$$b_2^2 - b_T^2 = O_P\left(\frac{1}{\sqrt{N}}\right) + O_P\left(\sqrt{\frac{N}{T^{1-2\alpha}}}\right)$$
.
Mit dem zentralen Grenzwertsatz folgt

$$\frac{1}{k^*} \sum_{i=1}^{k^*} \frac{N}{T} W_i^2 - 1 = O_P\left(\frac{1}{\sqrt{N}}\right).$$
(1.49)

Somit ergibt sich die Behaupung aus

$$b_{2}^{2} = \frac{N}{k^{*}T} \sum_{i=1}^{k^{*}} \left(R_{i} - \frac{T}{N} a \right)^{2}$$

$$= \frac{N}{k^{*}T} \sum_{i=1}^{k^{*}} \left(b_{T} W_{i} + \eta_{i} \right)^{2}$$

$$= b_{T}^{2} \left(1 + O_{P} \left(\frac{1}{\sqrt{N}} \right) \right) + O_{P} \left(N^{1/2} T^{\alpha - 1/2} \right) + O_{P} \left(N T^{2\alpha - 1} \right).$$

$$3. \text{ z.z.: } \hat{b}_{\hat{k}^*}^2 - b_1^2 = O_P\left(\frac{T}{N^2}\right).$$

$$\left|\hat{b}_{\hat{k}^*}^2 - b_1^2\right|$$

$$= \frac{N}{\hat{k}^* k^* T} \left|k^* \sum_{i=1}^{\hat{k}^*} \left(R_i - \frac{T}{N} \hat{a}_{\hat{k}^*}\right)^2 - (k^* + O_P(1)) \sum_{i=1}^{k^*} \left(R_i - \frac{T}{N} \hat{a}_{\hat{k}^*}\right)^2\right|$$

$$\leq \begin{cases} \frac{N}{\hat{k}^* T} \sum_{i=k^*+1}^{\hat{k}^*} \left(R_i - \frac{T}{N} \hat{a}_{\hat{k}^*}\right)^2 + O_P(1) \frac{N}{\hat{k}^* k^* T} \sum_{i=1}^{k^*} \left(R_i - \frac{T}{N} \hat{a}_{\hat{k}^*}\right)^2 : k^* + 1 \le \hat{k}^*,$$

$$(1.50)$$

$$\frac{N}{\hat{k}^* T} \sum_{i=\hat{k}^*+1}^{k^*} \left(R_i - \frac{T}{N} \hat{a}_{\hat{k}^*}\right)^2 + O_P(1) \frac{N}{\hat{k}^* k^* T} \sum_{i=1}^{k^*} \left(R_i - \frac{T}{N} \hat{a}_{\hat{k}^*}\right)^2 : k^* \ge \hat{k}^* + 1.$$

Für den zweiten Summanden in (1.50) folgt in beiden Fällen mittels der Teile 1) und 2)

$$\frac{N}{k^* \hat{k}^* T} \sum_{i=1}^{k^*} \left(R_i - \frac{T}{N} \hat{a}_{\hat{k}^*} \right)^2 = O_P \left(\frac{1}{N} \right) \left(\left| b_1^2 - b_2^2 \right| + \left| b_2^2 \right| \right) \\ = O_P \left(\frac{T}{N^4} \right) + O_P \left(\frac{1}{N} \right).$$
(1.51)
Für die erste Summe der oberen Zeile in (1.50), also falls \hat{k}^* den wahren Wert k^* überschätzt, wird zuerst eine Abschätzung der einzelnen Summanden benötigt. Wegen Satz 1.12 reicht es aus, die Summanden auf der Menge $\{0 < \hat{k}^* - k^* \leq C\}$ für ein festes C > 0 zu betrachten.

$$\begin{split} \max_{k^* + 2 \le i \le \hat{k}^*} \left| R_i - \frac{T}{N} \hat{a}_{\hat{k}^*} \right| \\ & \leq \max_{k^* + 2 \le i \le k^* + C} \left| Z_T \left(i \frac{T}{N} \right) - Z_T \left((i-1) \frac{T}{N} \right) \right| + \frac{T}{N} \left| \hat{a}_{\hat{k}^*} \right| \\ & = \left| b_T^* \right| \max_{k^* + 2 \le i \le k^* + C} \left| W_T^* \left(i \frac{T}{N} - T^* \right) - W_T^* \left((i-1) \frac{T}{N} - T^* \right) \right| + O_P \left(\frac{T}{N} \right) + O_P \left(T^{\alpha} \right) \\ & = O_P \left(\frac{T}{N} \right). \end{split}$$

Für den gemischten Summanden R_{k^*+1} ergibt sich mittels (1.36), (1.37) sowie (1.41)

$$\begin{aligned} \left| R_{k^*+1} - \frac{T}{N} \hat{a}_{\hat{k}^*} \right| &\leq \left| Z_T \left(\left(k^* + 1 \right) \frac{T}{N} \right) - Z_T \left(k^* \frac{T}{N} \right) \right| + \frac{T}{N} \left| \hat{a}_{\hat{k}^*} \right| \\ &= O_P \left(\frac{T}{N} \right). \end{aligned}$$

Bei einem Unterschätzen von k^* durch den Schätzer \hat{k}^* reicht es wie oben nach Satz 1.12 aus, die Summanden auf der Menge $\{0 \le k^* - \hat{k}^* \le C\}$ für ein festes C > 0 zu betrachten.

$$\begin{aligned} \max_{\hat{k}^* \le i \le k^*} \left| R_i - \frac{T}{N} \hat{a}_{\hat{k}^*} \right| \\ &\le \max_{k^* - C \le i \le k^*} \left| R_i - \frac{T}{N} a \right| + \frac{T}{N} \left| a - \hat{a}_{\hat{k}^*} \right| \\ &= \left| b_T \right| \max_{k^* - C \le i \le k^*} \left| W_T \left(i \frac{T}{N} \right) - W_T \left((i-1) \frac{T}{N} \right) \right| + O_P \left(\frac{T}{N^2} \right) + O_P \left(\frac{\sqrt{T}}{N} \right) + O_P \left(T^{\alpha} \right) \\ &= O_P \left(\max \left\{ \sqrt{\frac{T}{N}}, \frac{T}{N^2} \right\} \right). \end{aligned}$$

Damit erhält man insgesamt für die ersten Summanden der beiden Fälle in (1.50)

$$\frac{N}{\hat{k}^*T} \sum_{i=k^*+1}^{k^*} \left(R_i - \frac{T}{N} \hat{a}_{\hat{k}^*} \right)^2 = \frac{N}{\hat{k}^*T} |\hat{k}^* - k^*| O_P\left(\frac{T^2}{N^2}\right) \\ = O_P\left(\frac{T}{N^2}\right), \quad \text{falls} \quad k^* + 1 \le \hat{k}^*,$$

$$\frac{N}{\hat{k}^*T} \sum_{i=\hat{k}^*+1}^{k^*} \left(R_i - \frac{T}{N} \hat{a}_{\hat{k}^*} \right)^2 = \frac{N}{\hat{k}^*T} |\hat{k}^* - k^*| O_P \left(\max\left\{\frac{T}{N}, \frac{T^2}{N^4}\right\} \right) \\ = O_P \left(\max\left\{\frac{1}{N}, \frac{T}{N^4}\right\} \right), \quad \text{falls} \quad k^* \ge \hat{k}^* + 1.$$

Damit folgt mit (1.50) und (1.51) die Behauptung 3.

Mit den Behauptungen 1 - 3 ergibt sich die Aussage des Satzes.

Im Allgemeinen ist T/N^2 keine Nullfolge, so dass nach Satz 1.18 der Schätzer $\hat{b}_{\hat{k}^*}^2$ nicht als ein konsistenter Schätzer für b_T^2 unter der Alternative nachgewiesen ist. Doch falls die Invarianzprinzipien (1.36) bis (1.39) für $\alpha < 1/4$ erfüllt sind, gewährleistet die Wahl von $N = T^{1/2+\epsilon}$ für $0 < \epsilon < 1/2 - 2\alpha$ das Einhalten der Voraussetzung $N = o(T^{1-2\alpha})$ des Satzes 1.18, und der Schätzer $\hat{b}_{\hat{k}^*}^2$ liefert konsistente Schätzungen für den Parameter b_T^2 . Wie in dem hier betrachteten Fall, ist auch unter allgemeineren Alternativen mit einem Changepoint $T^* = \theta_T T$, $\theta_T \in (0, 1)$, und von T abhängigen Parametern a_T und a_T^* das Verhalten des Schätzers $\hat{b}_{\hat{k}^*}^2$ von der Konsistenzrate des Schätzers \hat{k}^* für $k^* = [\theta_T N]$ abhängig. Bei einem "local change", d.h. falls $\Delta_T = a_T - a_T^* \to 0$ für $T \to \infty$, verschlechtert sich die Konvergenzrate des Schätzers \hat{k}^* aufgrund der abnehmenden Unterscheidbarkeit der sich ändernden Parameter. Gleiches gilt unter Alternativen mit "frühen" oder "späten" Changepoints, d.h. $\theta_T \to 0$ oder $\theta_T \to 1$ für $T \to \infty$. In diesen Fällen ist für den Schätzer $\hat{b}_{\hat{k}^*}^2$ mit einem schlechteren Verhalten als unter den Voraussetzungen des Satzes 1.18 zu rechnen. Satz 1.18 betrachtet das Verhalten des Schätzers $\hat{b}_{\hat{k}^*}^*$ unter Alternativen, die häufig vor-

kommende Situationen beschreiben, jedoch von speziellerer Form als in den zwei vorigen Abschnitten sind. Um die Konsistenz der Teststatistik \overline{M}_T aus (1.35) unter der hier behandelten Alternative zu gewährleisten, reicht das Erfüllen der Voraussetzungen des Satzes 1.18 aus. Denn dann gilt $\hat{b}_{\hat{k}^*}^2 = o_P(T)$, so dass unter der Alternative mit Changepoint $T^* = \theta T$, $\theta \in (0, 1)$, nach Satz 1.3 folgt

$$\bar{M}_T^2 \xrightarrow{P} \infty, \quad T \to \infty,$$

was gerade der Konsistenzaussage der Teststatistik \bar{M}_T entspricht.

Mit dem Resultat der Konsistenz unter der Nullhypothese zu Beginn dieses Abschnitts ist $\hat{b}_{\hat{k}*}^2$ ein geeigneter Schätzer, um den unbekannten Parameter b_T^2 in der Statistik M_T aus (1.2) zu ersetzen. Die gewonnene Statistik \bar{M}_T aus (1.35) bietet gegenüber der Statistik \tilde{M}_T aus (1.19) des letzten Abschnitts, die auf einer Familie von Schätzern für den unbekannten Parameter b_T^2 beruht, den Vorteil einer schnellen Berechenbarkeit. Denn in statistischen Anwendungen ist max $\{|Z_k^*|: 1 \leq k \leq N\}$ zu bestimmen, so dass der Wert des Schätzers \hat{k}^* unmittelbar zur Verfügung steht. Die noch erforderlichen Rechnungen sind von minimalem Aufwand.

2 Lineares Modell

Den Betrachtungen dieses Kapitels liegt ein lineares Regressionsmodell der Changepoint-Analyse und eine damit verbundene Testproblematik zugrunde. Chu und White [ChWh92] geben dafür eine Teststatistik sowie Schätzer für die Varianzen der in dem Modell vorkommenden stochastischen Fehler an. In Abschnitt 2.1 wird dieses Modell und die Teststatistik unter stärkeren Voraussetzungen eingeführt, die jedoch für die Praxis ebenso relevant sind. Eine Familie von Varianzschätzern für das Modell wird unter der Nullhypothese und unter der Alternative in Abschnitt 2.2 untersucht. In den Abschnitten 2.3. und 2.4 stellen wir zwei neue, auf dieser Familie basierende Varianzschätzer¹ vor, die bessere Eigenschaften als die von Chu und White [ChWh92] unter allgemeineren Voraussetzungen vorgeschlagenen Schätzer besitzen.

Die für die Beweise dieses Kapitels notwendigen Konvergenzraten von deterministischen Folgen sind im Anhang A zusammengestellt. Simulationen zu den Ergebnissen dieses Kapitels findet sich im Kapitel "Simulationen".

2.1 Das Modell

Prozesse mit linearem Trend, deren Parameter einer Änderung zu einem Zeitpunkt unterliegen, gehören zu den klassischen Modellen, die in der Regressionsanalyse untersucht werden². Denn Änderungen von Einflüssen der realen Welt, z.B. ein Wechsel der Wirtschaftspolitik oder Technologiefortschritt, bewirken oft nur Änderungen der Parameter, ohne dass sich das Grundmodell als falsch erweist. Die Zeitpunkte dieser Änderungen der Parameter bezeichnet man als Changepoints.

In diesem Kapitel betrachten wir ein lineares Regressionsmodell mit drei unbekannten Regressionsparametern, in dem höchstens ein Changepoint vorliegen kann.

Definition 2.1

Seien $\{y_i : 1 \leq i \leq n\}$ für festes $n \in \mathbb{N}$ Zufallsvariablen auf einem Wahrscheinlichkeitsraum (Ω, \mathcal{A}, P) , die dem folgenden Modell genügen:

$$y_i := y_{i,n} := \begin{cases} a_n + b_n t_i + e_i & : & 1 \le i \le k^*, \\ a_n + b_n^* t_i + e_i & : & k^* + 1 \le i \le n, \end{cases}$$

wobei a_n , b_n und b_n^* unbekannte reelle Regressionsparameter sind und $k^* = k^*(n) \in \mathbb{N}$ ist. Dabei sind $e_1, e_2, \ldots e_n$ unabhängige, identisch verteilte Zufallsvariablen mit $\mathbb{E} e_i = 0$ und Varianz $0 < \sigma^2 := \operatorname{Var} e_i < \infty$.

Beobachtungspunkte seien $t_i := i^{\alpha}$ für $\alpha > 0$ und $1 \le i \le n$.

Aufgrund der beobachteten Daten y_1, \ldots, y_n soll eine Entscheidung getroffen werden, ob ein Changepoint k^* vorliegt oder nicht. Das Testproblem formuliert sich folgendermaßen:

$$\mathbf{H}_{0} : k^{*} = n
 \mathbf{H}_{A} : 1 \leq k^{*} < n \text{ und } b_{n} \neq b_{n}^{*}.$$
(2.1)

¹vorgeschlagen in [CsHo97].

²vgl. etwa [Seb77] oder [Hin71].

Das bedeutet, dass unter der Nullhypothese der Regressionsparameter b_n konstant bleibt, während dieser unter der Alternative zu dem Zeitpunkt k^* seinen Wert ändert.

In den asymptotischen Betrachtungen dieses Kapitels, in denen die Anzahl n der Beobachtungen wächst, wird die Familie von Prozessen $\{y_i : 1 \le i \le n\}$ gemäß Definition 2.1 für $n \in \mathbb{N}$ betrachtet. Dazu seien alle y_i für $i \in \mathbb{N}$ auf einem gemeinsamen Wahrscheinlichkeitsraum (Ω, \mathcal{A}, P) definiert und die Folge e_1, e_2, \ldots seien unabhängige und identisch verteilte Zufallsvariablen auf diesem Wahrscheinlichkeitsraum. Die Parameter a_n, b_n und b_n^* setzen wir als beschränkt voraus und benutzen die (stochastischen) Landau-Symbole o_p und O_P für $n \to \infty$, falls nicht anders gekennzeichnet.

Als weitere Notationen und Definitionen führen wir für festes $n \in \mathbb{N}$ ein:

$$Y_{k} := Y_{k,n} := \begin{pmatrix} y_{1} \\ \vdots \\ y_{k} \end{pmatrix}, \quad 1 \le k \le n, \qquad Y_{k}^{*} := Y_{k,n}^{*} := \begin{pmatrix} y_{k+1} \\ \vdots \\ y_{n} \end{pmatrix}, \quad 0 \le k \le n-1,$$
$$T_{k} := T_{k,n} := \begin{pmatrix} t_{1} & 1 \\ \vdots & \vdots \\ t_{k} & 1 \end{pmatrix}, \quad E_{k} := E_{k,n} := \begin{pmatrix} e_{1} \\ \vdots \\ e_{k} \end{pmatrix}, \quad 1 \le k \le n,$$
$$T_{k}^{*} := T_{k,n}^{*} := \begin{pmatrix} t_{k+1} & 1 \\ \vdots & \vdots \\ t_{n} & 1 \end{pmatrix}, \quad E_{k}^{*} := E_{k,n}^{*} := \begin{pmatrix} e_{k+1} \\ \vdots \\ e_{n} \end{pmatrix}, \quad 0 \le k \le n-1,$$
$$\gamma_{n} := \begin{pmatrix} b_{n} \\ a_{n} \end{pmatrix}, \quad \gamma_{n}^{*} := \begin{pmatrix} b_{n}^{*} \\ a_{n} \end{pmatrix}, \quad \delta_{n} := b_{n}^{*} - b_{n}, \quad \Delta_{n} := \begin{pmatrix} \delta_{n} \\ 0 \end{pmatrix}.$$

|| || bezeichne die euklidische Norm für Vektoren.

Da die Zufallsvariablen e_i für $1 \leq i \leq n$ die Schwankungen der Beobachtungen um die Gerade $a_n + b_n t_i$ bzw. $a_n + b_n^* t_i$ angeben, werden diese Fehler oder Residuen genannt. Weiter bezeichnet man $T_k^T T_k$ für $1 \leq k \leq n$ und $T_k^{*T} T_k^*$ für $0 \leq k \leq n - 1$ als Designmatrizen. Das Verhalten der Designmatrizen und deren Komponenten für $k \to \infty$ ist in vielen Zugängen der Changepoint-Analyse für das Testproblem (2.1) von grundlegender Bedeutung.

Für ein lineares Regressionsmodell mit beliebig vielen unbekannten Regressionsparametern und Fehlern e_i mit $E|e_i|^{2+\delta} < \infty$ für ein $\delta > 0$ geben Horváth [Ho95] sowie Csörgő und Horváth [CsHo97] eine Statistik und deren Grenzverteilung für das Testproblem (2.1) an. Überträgt man die dort gestellten Forderungen an das Konvergenzverhalten der Designmatrizen auf den hier vorliegenden Fall von nur einem sich unter der Alternative ändernden Parameter, so erhält man unter anderem:

$$\exists a \in \mathbb{R} \quad \exists \beta > 2 + 27/\min\{1, \delta\} : \frac{1}{k} \sum_{i=1}^{k} t_i^2 - a = o\left((\log k)^{-\beta}\right), \quad k \to \infty, \frac{1}{n-k} \sum_{i=k+1}^{n} t_i^2 - a = o\left((\log(n-k))^{-\beta}\right), \quad n-k \to \infty.$$
(2.2)

Sicherlich ist dies in dem hier vorliegenden Fall von $t_i = i^{\alpha}$ für $1 \le i \le n$ und $\alpha > 0$ nicht erfüllt. Auch die in Hušková [Hu96] gestellten Bedingungen, um Schätzer für die Regressi-

onsparameter eines linearen Regressionsmodells mit beliebig vielen unbekannten Regressionsparametern herzuleiten, resultieren bei Übertragung auf den Fall von nur drei unbekannten Parametern in ähnlichen Bedingungen wie (2.2).

Das von Chu und White [ChWh92] vorgestellte Verfahren für das Testproblem (2.1), das in diesem Kapitel betrachtet werden soll, kommt ebenfalls ohne eine Verteilungsannahme für die Fehler aus und und lässt die Designmatrizen $T_k^T T_k$ und $T_k^{*T} T_k^*$ gemäß Definition 2.1 zu. Dieser Test basiert auf einem Vergleich von Schätzern für die Regressionsparameter des linearen Modells. Denn bei bekanntem Changepoint $k = k^*$ bieten sich als Schätzer für die unbekannten Parameter a_n, b_n und b_n^* die Kleinste-Quadrate-Schätzer an, wenn die Grundgesamtheit der Beobachtungen in zwei Teile $\{y_i : 1 \le i \le k\}$ und $\{y_i : k+1 \le i \le n\}$ getrennt wird. Man erhält in Abhängigkeit von der Stichprobengröße n als Schätzer für

$$b_n: \qquad \hat{b}_k := \frac{k \sum_{i=1}^k t_i y_i - \left(\sum_{i=1}^k t_i\right) \left(\sum_{i=1}^k y_i\right)}{k \sum_{i=1}^k t_i^2 - \left(\sum_{i=1}^k t_i\right)^2}, \quad 2 \le k \le n,$$
(2.3)

$$b_n^*: \qquad \check{b}_k := \frac{(n-k)\sum_{i=k+1}^n t_i y_i - \left(\sum_{i=k+1}^n t_i\right) \left(\sum_{i=k+1}^n y_i\right)}{(n-k)\sum_{i=k+1}^n t_i^2 - \left(\sum_{i=k+1}^n t_i\right)^2}, \quad 0 \le k \le n-2,$$
(2.4)

$$a_n: \qquad \hat{a}_k := \frac{\left(\sum_{i=1}^k t_i^2\right) \left(\sum_{i=1}^k y_i\right) - \left(\sum_{i=1}^k t_i\right) \left(\sum_{i=1}^k t_i y_i\right)}{k \sum_{i=1}^k t_i^2 - \left(\sum_{i=1}^k t_i\right)^2}, \quad 2 \le k \le n,$$
(2.5)

$$a_{n}: \qquad \check{a}_{k} := \frac{\left(\sum_{i=k+1}^{n} t_{i}^{2}\right)\left(\sum_{i=k+1}^{n} y_{i}\right) - \left(\sum_{i=k+1}^{n} t_{i}\right)\left(\sum_{i=k+1}^{n} t_{i}y_{i}\right)}{\left(n-k\right)\sum_{i=k+1}^{n} t_{i}^{2} - \left(\sum_{i=k+1}^{n} t_{i}\right)^{2}}, \quad 0 \le k \le n-2.$$

$$(2.6)$$

Die erforderliche Eigenschaft der Messbarkeit ist für die Funktionen sicherlich gegeben. Für weitere Eigenschaften wie Erwartungstreue und Konsistenz sei auf das Kapitel "Simulationen" verwiesen.

Falls $\{e_i : 1 < i < n\}$ eine Familie von unabhängigen standard-normalverteilten Zufallsvariablen ist, so sind die Kleinste-Quadrate-Schätzer gerade die Maximum-Likelihood-Schätzer. Doch auch ohne diese Verteilungsannahme lassen sich die Schätzer (2.3) - (2.6) sowohl unter der Nullhypothese als auch unter der Alternative in der gewohnten Matrixnotation³ darstellen:

Lemma 2.2

Set $1 < k^* < n$ beliebig für festes $n \in \mathbb{N}$. Dann gilt:

a)
$$\hat{\gamma}_k := \begin{pmatrix} \hat{b}_k \\ \hat{a}_k \end{pmatrix} = (T_k^T T_k)^{-1} T_k^T Y_k, \quad 2 \le k \le n.$$

b) $\check{\gamma}_k := \begin{pmatrix} \check{b}_k \\ \check{a}_k \end{pmatrix} = (T_k^{*T} T_k^{*})^{-1} T_k^{*T} Y_k^{*}, \quad 0 \le k \le n-2.$

Beweis: Unter Beachtung von

$$(T_k^T T_k)^{-1} = \frac{1}{k \sum_{i=1}^k t_i^2 - \left(\sum_{i=1}^k t_i\right)^2} \begin{pmatrix} k & -\sum_{i=1}^k t_i \\ -\sum_{i=1}^k t_i & \sum_{i=1}^k t_i^2 \end{pmatrix}, \quad 2 \le k \le n,$$

³vgl. e 12 [T]

und
$$(T_k^{*T}T_k^*)^{-1} = \frac{1}{(n-k)\sum_{i=k+1}^n t_i^2 - (\sum_{i=k+1}^n t_i)^2} \begin{pmatrix} n-k & -\sum_{i=k+1}^n t_i \\ -\sum_{i=k+1}^n t_i & \sum_{i=k+1}^n t_i^2 \end{pmatrix},$$

 $0 \le k \le n-2,$

ergeben sich die Behauptungen mit üblicher Matrizenmultiplikation.

Die Determinanten der Designmatrizen $T_k^T T_k$ und $T_k^* T_k^*$ werden im Folgenden öfters auftreten, weshalb sie abgekürzt als

$$d_k := d_{k,n} := k \sum_{i=1}^k t_i^2 - \left(\sum_{i=1}^k t_i\right)^2, \quad 2 \le k \le n,$$

und $d_k^* := d_{k,n}^* := (n-k) \sum_{i=k+1}^n t_i^2 - \left(\sum_{i=k+1}^n t_i\right)^2, \quad 0 \le k \le n-2.$

bezeichnet werden sollen.

Ein direkter Nachweis zeigt die positive Definitheit der Matrizen $T_k^T T_k$ und $T_k^* T_k^*$ für alle $2 \le k \le n$ bzw. für alle $0 \le k \le n - 2$. Damit gilt

$$d_k > 0 \quad \forall \ 2 \le k \le n \quad \text{und} \quad d_k^* > 0 \quad \forall \ 0 \le k \le n-2$$

für beliebiges $\alpha > 0$.

Ein anderes Verfahren für das Testproblem (2.1), das wie der Test von Chu und White [ChWh92] auf einem Vergleich von Schätzern basiert, ist der von Hawkins [Ha89] vorgestellte Union-Intersection Test. Die asymptotische Verteilung der dort behandelten Statistik ist auch in [HoSh95] zu finden. Dieser Test beruht auf einem Vergleich der Schätzer $\hat{\gamma}_k$ und $\tilde{\gamma}_k$ der ersten k Beobachtungen und der letzten n - k Beobachtungen. Um jedoch eine Grenzverteilung der Statistik zu erhalten, wird wie in [Ho95] die Einhaltung der Bedingung (2.2) gefordert. Im Gegensatz dazu basiert das Verfahren von Chu und White [ChWh92] lediglich auf einem Vergleich der Kleinste-Quadrate-Schätzer \hat{b}_k der ersten k Beobachtungen mit dem wahren Parameter b_n . Da dieser im Allgemeinen unbekannt ist, wird er durch den Schätzer \hat{b}_n , dem Kleinste-Quadrate-Schätzer unter der Nullhypothese, ersetzt und man erhält als Teststatistik

$$L_n := \frac{\alpha}{\sigma(\alpha+1)\sqrt{(2\alpha+1)}} n^{\alpha+1/2} \max_{2 \le k \le n} |Q_k|$$
mit $Q_k := Q_{k,n} := \left(\frac{k}{n}\right)^{2\alpha+1} \left(\hat{b}_k - \hat{b}_n\right),$

$$(2.7)$$

wobei die zugrunde liegenden Beobachtungen y_1, \ldots, y_n der Definition 2.1 entsprechen. Dabei werden für große Werte von L_n die Nullhypothese H_0 des Testproblems (2.1) verworfen. In [ChWh92] wird der Test nur für Beobachtungen mit $\alpha = 1$ betrachtet. Es werden weder Unabhängigkeit noch eine identische Verteilung der Familie $\{e_i : 1 \le i \le n\}$ gefordert, sondern stattdessen die schwache Konvergenz im Sinne des funktionalen zentralen Grenzwertsatzes vorausgesetzt, d.h.

$$\frac{1}{\sqrt{\sigma_0^2 n}} \sum_{i=1}^{[nt]} e_i \stackrel{\mathcal{D}[0,1]}{\to} W(t), \quad n \to \infty,$$

$$0 < \sigma_0^2 := \lim_{n \to \infty} \frac{1}{n} E\left(\sum_{i=1}^n e_i\right)^2 < \infty,$$
(2.8)

wobei $\{W(t) : 0 \le t \le 1\}$ ein Wiener-Prozess sei. In dieser Arbeit schließen wir uns [CsHo97] an und setzen wie dort die Unabhängigkeit der Fehler voraus und lassen allgemeiner Beobachtungspunkte t_i mit $\alpha > 0$ zu.

Da in der Regel die Varianz σ^2 der identisch verteilten und unabhängigen Fehler e_i in der Statistik L_n aus (2.7) unbekannt ist, stellt sich auch hier das Problem einer geeigneten Schätzung. Unter der allgemeineren Voraussetzung (2.8) verweisen Chu and White [ChWh92] auf heteroskedastische und autokorrelations-konsistente Varianzschätzer⁴ für σ_0^2 , geben jedoch selbst eine schlechtere Konvergenzrate unter der Alternative H_A für diese an, als die im Folgenden betrachteten Schätzer aufweisen.

Um mittels der Statistik L_n aus (2.7) zu einem asymptotischen Fehlerniveau eine Entscheidung treffen zu können, benötigt man zumindest eine asymptotische Verteilung. Der folgende Beweis, in der angegebenen Quelle lediglich für $\alpha \ge 1$ ausgeführt, bedarf keiner Modifikation unter der schwächeren Voraussetzung $\alpha > 0$.

Satz 2.3

Set $\{y_i : 1 \leq i \leq n\}$ wie in Definition 2.1 für alle $n \in \mathbb{N}$. Dann gilt unter H_0 :

$$L_n \xrightarrow{\mathcal{D}} \sup_{0 \le t \le 1} |B(t)|, \quad n \to \infty,$$

wobei $\{B(t) : 0 \le t \le 1\}$ eine Standard-Brownsche-Brücke ist.

Beweis: siehe [CsHo97].

Bei der weiteren Untersuchung der Teststatistik (2.7) und der verschiedenen Varianzschätzer werden folgende Konvergenzaussagen immer wieder benötigt, weshalb sie an dieser Stelle in einem Lemma getrennt festgehalten werden.

Lemma 2.4

Sei $e_1, e_2...$ eine Folge von unabhängigen, identisch verteilten Zufallsvariablen mit $E e_1 = 0$ und $0 < Var e_1 = \sigma^2 < \infty$. Dann gilt für $t_i = i^{\alpha}$ mit $\alpha > 0$:

a)
$$\max_{1 \le k \le n} \left| \frac{1}{\sqrt{k \log \log k}} \sum_{i=1}^{k} e_i \right| = O_P(1).$$

b)
$$\max_{1 \le k \le n} \left| \frac{1}{k^{\alpha} \sqrt{k \log \log k}} \sum_{i=1} t_i e_i \right| = O_P(1).$$
⁴vgl. etwa [An91] oder [NeWe87].

$$c) \max_{1 \le k \le n} \left| \sum_{i=1}^{k} e_i \right| = O_P\left(\sqrt{n}\right).$$

$$d) \max_{1 \le k \le n} \left| \sum_{i=1}^{k} t_i e_i \right| = O_P\left(n^{1/2+\alpha}\right).$$

$$e) \max_{0 \le k \le n-1} \left| \sum_{i=k+1}^{n} e_i \right| = O_P(\sqrt{n}).$$

$$f) \max_{0 \le k \le n-1} \left| \sum_{i=k+1}^{n} t_i e_i \right| = O_P\left(n^{1/2+\alpha}\right).$$

Beweis:

- a) Folgt aus dem Gesetz vom iterierten Logarithmus.
- b) Mit dem Gesetz vom iterierten Logarithmus folgt:

$$\begin{aligned} \max_{1 \le k \le n} \left| \frac{1}{k^{\alpha} \sqrt{k \log \log k}} \sum_{i=1}^{k} i^{\alpha} e_{i} \right| \\ &= \max_{1 \le k \le n} \left| \frac{1}{k^{\alpha} \sqrt{k \log \log k}} \sum_{i=1}^{k} i^{\alpha} \left(S_{i} - S_{i-1} \right) \right| \quad , \text{ wobei } S_{i} = \sum_{j=1}^{i} e_{j}, \quad S_{0} = 0, \\ &\le \max_{1 \le k \le n} \left| \frac{1}{k^{\alpha} \sqrt{k \log \log k}} \sum_{i=1}^{k-1} \left(i^{\alpha} - (i+1)^{\alpha} \right) S_{i} \right| + \max_{1 \le k \le n} \left| \frac{1}{k^{\alpha} \sqrt{k \log \log k}} k^{\alpha} S_{k} \right| \\ &= O_{P} \left(\max_{1 \le k \le n} \left\{ \frac{1}{k^{\alpha}} \sum_{i=1}^{k-1} \left| i^{\alpha} - (i+1)^{\alpha} \right| \right\} \right) + O_{P}(1) \\ &= O_{P}(1). \end{aligned}$$

- c) und d) erhält man direkt aus der Ungleichung von Kolmogorov und Satz A.1.
- e) Mit der Ungleichung von Kolmogorov erhält man für beliebiges C > 0:

$$P\left(\max_{0 \le k \le n-1} \left| \sum_{i=k+1}^{n} e_i \right| \ge C\sqrt{n} \right) = P\left(\max_{0 \le k \le n-1} \left| \sum_{i=1}^{n-k} e_i \right| \ge C\sqrt{n} \right)$$
$$= P\left(\max_{1 \le l \le n} \left| \sum_{i=1}^{l} e_i \right| \ge C\sqrt{n} \right)$$
$$\le \frac{\sigma^2}{C^2}.$$

f) Auch hier erhält man für beliebiges C > 0 mit der Ungleichung von Kolmogorov:

$$P\left(\max_{0 \le k \le n-1} \left| \sum_{i=k+1}^{n} i^{\alpha} e_i \right| \ge C n^{1/2+\alpha} \right) = P\left(\max_{0 \le k \le n-1} \left| \sum_{i=1}^{n-k} (n-i+1)^{\alpha} e_i \right| \ge C n^{1/2+\alpha} \right)$$

$$= P\left(\max_{1 \le l \le n} \left| \sum_{i=1}^{l} (n-i+1)^{\alpha} e_i \right| \ge C n^{1/2+\alpha} \right)$$

$$\le \frac{1}{C^2 n^{1+2\alpha}} \operatorname{Var}\left(\sum_{i=1}^{n} (n-i+1)^{\alpha} e_i \right)$$

$$= \frac{\sigma^2}{C^2} \frac{1}{n^{1+2\alpha}} \sum_{i=1}^{n} i^{2\alpha},$$

woraus mit Satz A.1 die Behauptung folgt.

Nicht nur im Hinblick auf das Verhalten der Statistik L_n unter der Alternative, sondern auch auf das Verhalten der Varianzschätzer, die in den nächsten Abschnitten betrachtet werden, sind die Konvergenzraten der Kleinste-Quadrate-Schätzer (2.3) bis (2.6) unter der Alternative bedeutsam.

Lemma 2.5

Set $\{y_i : 1 \leq i \leq n\}$ wie in Definition 2.1 für alle $n \in \mathbb{N}$. Es gelte $\frac{k^*}{n} \to \lambda$ für $n \to \infty$ und $\lambda \in (0, 1)$ sowie $k = [\theta n]$ für $\theta \in (\lambda, 1]$. Dann gilt unter H_A mit $1 \leq k^* \leq n - 1$:

$$a) \ \hat{a}_{k} = a_{n} + \delta_{n}k^{\alpha} \frac{\alpha + 1}{\alpha^{2}} \left(\left(\frac{\lambda}{\theta} \right)^{1+2\alpha} - \left(\frac{\lambda}{\theta} \right)^{1+\alpha} \right) + o(\delta_{n}n^{\alpha}) + O_{P}\left(\frac{1}{\sqrt{n}} \right).$$

$$b) \ \hat{b}_{k} = b_{n} + \delta_{n} \frac{1}{\alpha^{2}} \left(\alpha^{2} \left(1 - \left(\frac{\lambda}{\theta} \right)^{1+2\alpha} \right) + (1+2\alpha) \left(\left(\frac{\lambda}{\theta} \right)^{1+\alpha} - \left(\frac{\lambda}{\theta} \right)^{1+2\alpha} \right) \right) \right)$$

$$+ o(\delta_{n}) + O_{P}\left(\frac{1}{n^{1/2+\alpha}} \right).$$

$$c) \ \hat{b}_{k} = b_{n}^{*} - \delta_{n} \frac{1}{\alpha^{2}} \left(\alpha^{2} \left(\frac{\lambda}{\theta} \right)^{1+2\alpha} + (1+2\alpha) \left(\left(\frac{\lambda}{\theta} \right)^{1+2\alpha} - \left(\frac{\lambda}{\theta} \right)^{1+\alpha} \right) \right) \right)$$

$$+ o(\delta_{n}) + O_{P}\left(\frac{1}{n^{1/2+\alpha}} \right).$$

Eine analoge Darstellung ließe sich auch für die Kleinste-Quadrate-Schätzer \check{a}_k und \check{b}_k ableiten. Da dies für das weitere Vorgehen nicht relevant ist, soll darauf verzichtet werden.

Betrachtet man den Fall $\delta_n \to 0$ für $n \to \infty$, so ist \hat{b}_n unter der Alternative sowohl für b_n als auch b_n^* ein konsistenter Schätzer. Da \hat{b}_{k^*} unter der Alternative mit Changepoint k^* ebenfalls ein konsistenter Schätzer ist, folgt, dass die Differenz von \hat{b}_{k^*} und \hat{b}_n stochastisch gegen Null konvergiert. Davon ausgehend, dass die größte Differenz von \hat{b}_k und \hat{b}_n für $k = k^*$ angenommen wird, gilt so für die Statistik L_n , dass bei einer großen Rate, mit der δ_n verschwindet, diese mit zunehmender Anzahl der Beobachtungen immer mehr zu einem extrem konservativen Verhalten tendiert.

Gilt darüber hinaus $\delta_n = o(n^{-\alpha})$, so sind auch die Kleinste-Quadrate-Schätzer für den Parameter a_n unter der Alternative konsistent. Von Varianzschätzern, die auf den Kleinste-Quadrate-Schätzern basieren, ist dann auch unter der Alternative eine gute Schätzung zu erwarten. Diese Überlegung greifen wir später noch einmal auf.

Beweis von Lemma 2.5:

Nach Voraussetzung existiert $n_0 \in \mathbb{N}$, so dass $k^* + 1 \leq [\theta n] = k$ für alle $n \geq n_0$ ist. Sei im Folgenden $n \geq n_0$. Teilt man die Summen an dem Changepoint k^* auf, so ergeben sich folgende Darstellungen:

$$\sum_{i=1}^{k} t_i y_i = a_n \sum_{i=1}^{k} t_i + b_n \sum_{i=1}^{k^*} t_i^2 + b_n^* \sum_{i=k^*+1}^{k} t_i^2 + \sum_{i=1}^{k} t_i e_i ,$$

$$\sum_{i=1}^{k} y_i = ka_n + b_n \sum_{i=1}^{k^*} t_i + b_n^* \sum_{i=k^*+1}^{k} t_i + \sum_{i=1}^{k} e_i.$$

a) Man erhält mit (2.3) bis (2.6) die Darstellungen

$$\begin{aligned} d_k \hat{a}_k &= a_n k \left(\sum_{i=1}^k t_i^2 \right) + b_n \left(\sum_{i=1}^k t_i \right) \left(\sum_{i=1}^k t_i^2 \right) + b_n^* \left(\sum_{i=k^*+1}^k t_i \right) \left(\sum_{i=1}^k t_i^2 \right) - a_n \left(\sum_{i=1}^k t_i \right)^2 \\ &- b_n \left(\sum_{i=1}^k t_i^2 \right) \left(\sum_{i=1}^k t_i \right) - b_n^* \left(\sum_{i=k^*+1}^k t_i^2 \right) \left(\sum_{i=1}^k t_i \right) \\ &+ \left(\sum_{i=1}^k t_i^2 \right) \left(\sum_{i=1}^k e_i \right) - \left(\sum_{i=1}^k t_i \right) \left(\sum_{i=1}^k t_i e_i \right) \\ &= a_n \left(k \sum_{i=1}^k t_i^2 - \left(\sum_{i=1}^k t_i \right)^2 \right) - \delta_n \left(\sum_{i=1}^k t_i \right) \left(\sum_{i=1}^k t_i^2 \right) + \delta_n \left(\sum_{i=1}^{k^*} t_i^2 \right) \left(\sum_{i=1}^k t_i \right) \\ &+ \left(\sum_{i=1}^k t_i^2 \right) \left(\sum_{i=1}^k e_i \right) - \left(\sum_{i=1}^k t_i \right) \left(\sum_{i=1}^k t_i e_i \right). \end{aligned}$$

Mit Lemma 2.4 und Satz A.1 folgt

$$\hat{a}_{n} = a_{n} + \left(\frac{(2\alpha+1)(\alpha+1)^{2}}{\alpha^{2}} + o(1)\right) \left(\delta_{n}k^{\alpha}\frac{1}{(2\alpha+1)(\alpha+1)} \times \left(\left(\frac{\lambda}{\theta}\right)^{1+2\alpha} - \left(\frac{\lambda}{\theta}\right)^{1+\alpha} + o(1)\right) + O_{P}\left(\frac{1}{\sqrt{n}}\right)\right)$$
$$= a_{n} + \delta_{n}k^{\alpha}\frac{\alpha+1}{\alpha^{2}} \left(\left(\frac{\lambda}{\theta}\right)^{1+2\alpha} - \left(\frac{\lambda}{\theta}\right)^{1+\alpha}\right) + o\left(\delta_{n}n^{\alpha}\right) + O_{P}\left(\frac{1}{\sqrt{n}}\right).$$

b) Analog gewinnt man die folgende Darstellung, mit der wie in a) die Behauptung folgt.

$$d_k \hat{b}_k = b_n \left(k \sum_{i=1}^k t_i^2 - \left(\sum_{i=1}^k t_i \right)^2 \right) + \delta_n k \sum_{i=k^*+1}^k t_i^2 - \delta_n \left(\sum_{i=1}^k t_i \right) \left(\sum_{i=k^*+1}^k t_i \right) + k \left(\sum_{i=1}^k t_i e_i \right) - \left(\sum_{i=1}^k t_i \right) \left(\sum_{i=1}^k e_i \right).$$

c) Folgt direkt aus Aussage b).

Oben wird angedeutet, dass für "rasch verschwindendes" δ_n bei wachsender Stichprobengröße *n* die Statistik L_n mit immer geringerer Wahrscheinlichkeit eine Entscheidung für die Alternative zulässt. Dies soll natürlich vermieden werden, falls tatsächlich die Alternative vorliegt. Dazu der folgende Satz.

Satz 2.6

Sei $\{y_i : 1 \le i \le n\}$ wie in Definition 2.1 für alle $n \in \mathbb{N}$. Weiter gelte $\frac{k^*}{n} \to \lambda$ für $n \to \infty$ und $\lambda \in (0, 1)$ wie auch $|\delta_n| n^{1/2+\alpha} \to \infty$ für $n \to \infty$. Dann gilt unter H_A mit $2 \le k^* \le n - 1$:

$$\frac{1}{|\delta_n|} |Q_{k^*}| \xrightarrow{P} \lambda^{1+2\alpha} \left| \lambda^{1+2\alpha} + \frac{1+2\alpha}{\alpha^2} \left(\lambda^{1+2\alpha} - \lambda^{1+\alpha} \right) - 1 \right| \neq 0, \quad n \to \infty$$

Beweis: Aus der Chebyshev-Ungleichung folgt

$$\frac{1}{\sqrt{k^*}} \sum_{i=1}^{k^*} e_i = O_P(1), \quad n \to \infty, \quad \text{und} \quad \frac{1}{k^{*1/2+\alpha}} \sum_{i=1}^{k^*} t_i e_i = O_P(1), \quad n \to \infty.$$

Ersetzt man in der Definition (2.3) des Schätzers \hat{b}_{k^*} die Beobachtung y_i durch $a_n + b_n t_i + e_i$, so erhält man mit Satz A.1:

$$\hat{b}_{k^*} - b_n = \frac{1}{d_{k^*}} \left(k^* \sum_{i=1}^{k^*} t_i e_i - \left(\sum_{i=1}^{k^*} t_i \right) \left(\sum_{i=1}^{k^*} e_i \right) \right) \\ = O_P \left(\frac{1}{n^{1/2 + \alpha}} \right).$$

Unter Beachtung von Lemma 2.5 folgt

$$\frac{1}{\delta_n}Q_{k^*} = \frac{1}{\delta_n} \left(\frac{k^*}{n}\right)^{1+2\alpha} \left(\left(\hat{b}_{k^*} - b_n\right) + \left(b_n - \hat{b}_n\right)\right) \\
= \frac{1}{\delta_n} \left(\frac{k^*}{n}\right)^{1+2\alpha} \left(-\delta_n \frac{1}{\alpha^2} \left(\alpha^2 \left(1 - \lambda^{1+2\alpha}\right) + \left(1 + 2\alpha\right) \left(\lambda^{1+\alpha} - \lambda^{1+2\alpha}\right)\right) \\
+ o(\delta_n) + O_P \left(\frac{1}{n^{1/2+\alpha}}\right)\right) \\
= \lambda^{2\alpha+1} \left(\lambda^{1+2\alpha} + \frac{1+2\alpha}{\alpha^2} \left(\lambda^{1+2\alpha} - \lambda^{1+\alpha}\right) - 1\right) + o_P(1)$$

und damit die Konvergenzaussage. Die zweite Aussage gilt offensichtlich.

Unter den Voraussetzungen des vorigen Satzes ist auch die asymptotische Konsistenz des Tests L_n gewährleistet, denn:

$$L_n \ge \frac{\alpha |\delta_n| n^{\alpha+1/2}}{\sigma(\alpha+1)\sqrt{2\alpha+1}} \left| \frac{1}{\delta_n} Q_{k^*} \right| \xrightarrow{P} \infty, \quad n \to \infty.$$

Bei Ersetzen der Varianz σ^2 durch einen beliebigen Schätzer s_n^2 reicht es unter der Alternative aus, $s_n = o_P(\delta_n n^{\alpha+1/2})$ zu fordern, um die Konsistenz der Teststatistik L_n zu erhalten.

2.2 Die kombinierten Schätzer

Um zu einer ersten Variante der Varianzschätzung zu gelangen, folgen wir der Motivation der Statistik L_n und teilen die Grundgesamtheit der Beobachtungen in die Teile $\{y_i : 1 \le i \le k\}$ und $\{y_i : k + 1 \le i \le n\}$ auf. Bei bekanntem Changepoint $k = k^*$ bietet sich dann als Varianzschätzer an:

$$\hat{\sigma}_{k}^{2} := \hat{\sigma}_{k,n}^{2} := \frac{1}{n} \left(\sum_{i=1}^{k} \left(y_{i} - \hat{b}_{k} t_{i} - \hat{a}_{k} \right)^{2} + \sum_{i=k+1}^{n} \left(y_{i} - \check{b}_{k} t_{i} - \check{a}_{k} \right)^{2} \right)$$

$$= \frac{1}{n} \left(\|Y_{k} - T_{k} \hat{\gamma}_{k}\|^{2} + \|Y_{k}^{*} - T_{k}^{*} \check{\gamma}_{k}\|^{2} \right), \quad 2 \le k \le n-2.$$

$$(2.9)$$

Da der Changepoint k^* unbekannt ist, geht man auf dieser Weise für jedes $2 \le k \le n-2$ vor und erhält so eine Familie $\{\hat{\sigma}_k^2 : 2 \le k \le n-2\}$ von Varianzschätzern, da jede dieser Funktionen messbar ist. Die Konsistenz bei wachsender Stichprobengröße *n* folgt für jedes $\hat{\sigma}_k^2$ mit $k = [\theta n], \theta \in (0, 1)$, unter der Nullhypothese leicht. Es gilt sogar:

Satz 2.7

Set $\{y_i : 1 \leq i \leq n\}$ wie in Definition 2.1 für alle $n \in \mathbb{N}$. Dann gilt unter H_0 :

$$\max_{2 \le k \le n-2} |\hat{\sigma}_k^2 - \sigma^2| = o_P(1).$$

Beweis: Mit Lemma 2.2 folgt für $2 \le k \le n$

$$||Y_{k} - T_{k}\hat{\gamma}_{k}||^{2} = Y_{k}^{T}Y_{k} - Y_{k}^{T}T_{k}(T_{k}^{T}T_{k})^{-1}T_{k}^{T}Y_{k}$$

$$= (T_{k}\gamma_{n} + E_{k})^{T}(I - T_{k}(T_{k}^{T}T_{k})^{-1}T_{k}^{T})(T_{k}\gamma_{n} + E_{k})$$

$$= \sum_{i=1}^{k} e_{i}^{2} - E_{k}^{T}T_{k}(T_{k}^{T}T_{k})^{-1}T_{k}^{T}E_{k}.$$
 (2.10)

Für den zweiten Term in der letzten Zeile gilt

$$E_{k}^{T}T_{k}(T_{k}^{T}T_{k})^{-1}T_{k}^{T}E_{k} = \frac{1}{d_{k}}\left(k\left(\sum_{i=1}^{k}t_{i}e_{i}\right)^{2} - 2\left(\sum_{i=1}^{k}t_{i}\right)\left(\sum_{i=1}^{k}e_{i}\right)\left(\sum_{i=1}^{k}t_{i}e_{i}\right) + \left(\sum_{i=1}^{k}t_{i}^{2}\right)\left(\sum_{i=1}^{k}e_{i}\right)^{2}\right),$$

und mit den Sätzen A.2 und A.4 sowie Lemma 2.4 folgt jeweils

$$\begin{aligned} \max_{2 \le k \le n} \left| \frac{1}{d_k} k \left(\sum_{i=1}^k t_i e_i \right)^2 \right| &= O_P(\log \log n), \\ \max_{2 \le k \le n} \left| \frac{1}{d_k} \left(\sum_{i=1}^k t_i \right) \left(\sum_{i=1}^k e_i \right) \left(\sum_{i=1}^k t_i e_i \right) \right| &= O_P(\log \log n), \\ \max_{2 \le k \le n} \left| \frac{1}{d_k} \left(\sum_{i=1}^k t_i^2 \right) \left(\sum_{i=1}^k e_i \right)^2 \right| &= O_P(\log \log n). \end{aligned}$$

Damit erhält man insgesamt

$$\max_{2 \le k \le n} \left| E_k^T T_k \left(T_k^T T_k \right)^{-1} T_k^T E_k \right| = O_P \left(\log \log n \right).$$

Für den ersten Summanden der letzten Zeile von (2.10) gilt nach dem starken Gesetz der großen Zahlen

$$\max_{1 \le k \le n} \left| \sum_{i=1}^{k} e_i^2 - k\sigma^2 \right| = o_P(n) ,$$

zusammengefasst folgt daraus:

$$\frac{1}{n} \max_{2 \le k \le n} \left| \|Y_k - T_k \hat{\gamma}_k\|^2 - k\sigma^2 \right| = o_P(1).$$
(2.11)

Für die zweite Summe des Schätzers $\hat{\sigma}_k^2$ in (2.9) ergibt sich ebenso

$$\|Y_k^* - T_k^* \check{\gamma}_k\|^2 = \sum_{i=k+1}^n e_i^2 - E_k^{*T} T_k^* \left(T_k^{*T} T_k^*\right)^{-1} T_k^{*T} E_k^*, \qquad (2.12)$$

wobei der zweite Summand hier die folgende Darstellung besitzt:

$$E_{k}^{*T}T_{k}^{*} \left(T_{k}^{*T}T_{k}^{*}\right)^{-1} T_{k}^{*T}E_{k}^{*}$$

$$= \frac{1}{d_{k}^{*}} \left(\left(n-k\right) \left(\sum_{i=k+1}^{n} t_{i}e_{i}\right)^{2} - 2\left(\sum_{i=k+1}^{n} t_{i}\right) \left(\sum_{i=k+1}^{n} e_{i}\right) \left(\sum_{i=k+1}^{n} t_{i}e_{i}\right) + \left(\sum_{i=k+1}^{n} t_{i}^{2}\right) \left(\sum_{i=k+1}^{n} e_{i}\right)^{2} \right).$$

Da am rechten Rand, also für k dicht bei n, die Konvergenzrate von d_k^* nicht mehr ausreicht, um wie für die erste Summe des Schätzers $\hat{\sigma}_k^2$ vorzugehen, betrachten wir zuerst unterhalb von

$$f(n) := n - 2 - \left[\frac{n}{\log \log n}\right]$$

das Maximum. Denn mit den Sätzen A.2 und A.4 und Lemma 2.4 gilt

$$\begin{aligned} \max_{0 \le k \le f(n)} \left| \frac{1}{d_k^*} (n-k) \left(\sum_{i=k+1}^n t_i e_i \right)^2 \right| &= O_P\left((\log \log n)^4 \right), \\ \max_{0 \le k \le f(n)} \left| \frac{1}{d_k^*} \left(\sum_{i=k+1}^n t_i \right) \left(\sum_{i=k+1}^n e_i \right) \left(\sum_{i=k+1}^n t_i e_i \right) \right| &= O_P\left((\log \log n)^4 \right), \\ \max_{0 \le k \le f(n)} \left| \frac{1}{d_k^*} \left(\sum_{i=k+1}^n t_i^2 \right) \left(\sum_{i=k+1}^n e_i \right)^2 \right| &= O_P\left((\log \log n)^4 \right), \end{aligned}$$

womit folgt:

$$\max_{0 \le k \le f(n)} \left| E_k^{*T} T_k^* \left(T_k^{*T} T_k^* \right)^{-1} T_k^{*T} E_k^* \right| = O_P \left(\left(\log \log n \right)^4 \right).$$

Aufgrund der identischen und unabhängigen Verteilung der Familie $\{e_i : 1 \le i \le n\}$ impliziert das starke Gesetz der großen Zahlen

$$\max_{0 \le k \le n-1} \left| \sum_{i=k+1}^{n} e_i^2 - (n-k)\sigma^2 \right| = o_P(n),$$
(2.13)

so dass man insgesamt erhält:

$$\frac{1}{n} \max_{0 \le k \le f(n)} \left| \|Y_k^* - T_k^* \check{\gamma}_k\|^2 - (n-k)\sigma^2 \right| = o_P(1).$$
(2.14)

Für $f(n) \le k \le n-2$ erhalten wir für den zweiten Summanden in (2.12)

$$\begin{split} E_k^{*T} T_k^* \left(T_k^{*T} T_k^*\right)^{-1} T_k^{*T} E_k^* \\ &= \frac{n-k}{d_k^*} \left(\left(\sum_{i=k+1}^n t_i e_i \right)^2 - 2 \left(\sum_{i=k+1}^n t_i \right) \left(\frac{1}{n-k} \sum_{i=k+1}^n e_i \right) \left(\sum_{i=k+1}^n t_i e_i \right) \right. \\ &+ \frac{1}{n-k} \left(\sum_{i=k+1}^n t_i^2 \right) \left(\sum_{i=k+1}^n e_i \right)^2 \right) \\ &= \frac{n-k}{d_k^*} \left(\left(\left(\sum_{i=k+1}^n t_i e_i \right)^2 - 2 \left(\sum_{i=k+1}^n t_i \right) \left(\frac{1}{n-k} \sum_{i=k+1}^n e_i \right) \left(\sum_{i=k+1}^n t_i e_i \right) \right. \\ &+ \left(\frac{1}{n-k} \sum_{i=k+1}^n t_i \right)^2 \left(\sum_{i=k+1}^n e_i \right)^2 \right) + \frac{1}{n-k} \left(\sum_{i=k+1}^n e_i \right)^2 \\ &= \left(\sum_{i=k+1}^n t_i^2 - \frac{1}{n-k} \left(\sum_{i=k+1}^n t_i \right)^2 \right)^{-1} \left(\sum_{i=k+1}^n t_i (e_i - \bar{e}_{k,n}) \right)^2 + \frac{1}{n-k} \left(\sum_{i=k+1}^n e_i \right)^2 , \\ &\qquad \text{wobei } \bar{e}_{k,n} = \frac{1}{n-k} \sum_{i=k+1}^n e_i . \end{split}$$

Da

$$\operatorname{Var}\left(\sum_{i=k+1}^{n} t_{i}\left(e_{i} - \bar{e}_{k,n}\right)\right) = \sigma^{2}\left(\sum_{i=k+1}^{n} t_{i}^{2} - \frac{1}{n-k}\left(\sum_{i=k+1}^{n} t_{i}\right)^{2}\right),$$

folgt mit der Chebyshev-Ungleichung für beliebiges C > 0

$$P\left(\max_{f(n)\leq k\leq n-2} \left| \frac{\sum_{i=k+1}^{n} t_i \left(e_i - \bar{e}_{k,n}\right)}{\left(\sum_{i=k+1}^{n} t_i^2 - \frac{1}{n-k} \left(\sum_{i=k+1}^{n} t_i\right)^2\right)^{1/2}} \right| \geq C\sqrt{\frac{n}{\log\log n}}\right)$$

$$\leq \sum_{k=f(n)}^{n-2} \frac{\log \log n}{nC^2} \sigma^2$$

$$\leq \frac{\sigma^2}{C^2}.$$

Damit gilt also

$$\max_{f(n) \le k \le n-2} \left| \frac{\left(\sum_{i=k+1}^{n} t_i \left(e_i - \bar{e}_{k,n}\right)\right)^2}{\sum_{i=k+1}^{n} t_i^2 - \frac{1}{n-k} \left(\sum_{i=k+1}^{n} t_i\right)^2} \right| = O_P\left(\frac{n}{\log\log n}\right).$$
(2.15)

Analog folgt

$$\max_{f(n) \le k \le n-2} \left| \frac{1}{n-k} \left(\sum_{i=k+1}^n e_i \right)^2 \right| = O_P \left(\frac{n}{\log \log n} \right)$$

und man erhält mit (2.13) und (2.15)

$$\frac{1}{n} \max_{f(n) \le k \le n-2} \left| \|Y_k^* - T_k^* \check{\gamma}_k\|^2 - (n-k)\sigma^2 \right| = o_P(1).$$
(2.16)

Insgesamt implizieren die Gleichungen (2.11), (2.14) und (2.16) die Behauptung des Satzes. □

Um mittels einer Modifikation der Statistik L_n aus (2.7) zu einer Statistik \hat{L}_n zu gelangen, in der für jedes $2 \le k \le n$ die unbekannte Varianz σ^2 geschätzt wird, benötigt man eine Erweiterung der bisher betrachteten Familie $\{\hat{\sigma}_k^2 : 2 \le k \le n-2\}$ von Schätzern. Da jedoch (2.11) im vorangeganen Beweis zeigt, dass

$$\frac{1}{n}\sum_{i=1}^{n}\left(y_{i}-\hat{b}_{n}t_{i}-\hat{a}_{n}\right)^{2}=\frac{1}{n}\|Y_{n}-T_{n}\hat{\gamma}_{n}\|^{2}$$

ebenfalls ein konsistenter Schätzer unter der Nullhypothese ist, erhalten wir mit

$$\hat{\sigma}_k^2 := \begin{cases} \hat{\sigma}_k^2 & : \ 2 \le k \le n-2\\ \frac{1}{n} \|Y_n - T_n \hat{\gamma}_n\|^2 & : \ k \in \{n-1,n\} \end{cases}$$

eine Familie $\{\hat{\sigma}_k^2: 2\leq k\leq n\}$ von konsistenten Schätzern und mit

$$\hat{L}_{n} := \frac{\alpha}{(\alpha+1)\sqrt{2\alpha+1}} n^{\alpha+1/2} \max_{2 \le k \le n} \frac{1}{\sqrt{\hat{\sigma}_{k}^{2}}} |Q_{k}|$$
(2.17)

die gewünschte Statistik. Aufgrund der eben geführten Diskussion und der Voraussetzung $\sigma^2 \neq 0$ gilt unter der Nullhypothese $\min_{2 \leq k \leq n} \hat{\sigma}_k^2 \geq c$ für ein c > 0 auf Mengen von beliebig hohen Wahrscheinlichkeiten für n groß genug. Dies reicht auch für die folgende asymptotische Betrachtung der Statistik \hat{L}_n aus.

Da aber sogar $\max_{2 \le k \le n} |\hat{\sigma}_k^2 - \sigma^2| = o_P(1)$ gilt, folgt aus dem Lemma von Slutsky auch die Grenzverteilung der modifizierten Statistik \hat{L}_n mittels Satz 2.3:

Korollar 2.8

Set $\{y_i : 1 \leq i \leq n\}$ wie in Definition 2.1 für alle $n \in \mathbb{N}$. Dann gilt unter H_0 :

$$\hat{L}_n \xrightarrow{\mathcal{D}} \sup_{0 \le t \le 1} |B(t)|, \quad n \to \infty,$$

wobei $\{B(t) : 0 \le t \le 1\}$ eine Standard-Brownsche-Brücke sei.

Im Allgemeinen kann man nicht erwarten, dass $\hat{\sigma}_k^2$ für alle $2 \leq k \leq n-2$ auch unter der Alternative H_A konsistente Schätzer sind. Denn für die Kleinste-Quadrate-Schätzer (2.3) bis (2.6) folgt nach Lemma 2.5 unter der Alternative H_A mit $k^*/n \to \lambda$ für $n \to \infty$, $\lambda \in (0, 1)$, und $k = [\theta n]$ mit einem $\theta > \lambda$:

$$\hat{a}_k = a_n + \delta_n n^{\alpha} c_1 + o_P(1), \quad n \to \infty, \quad c_1 \in \mathbb{R}, \\ \hat{b}_k = b_n + \delta_n c_2 + o_P(1), \quad n \to \infty, \quad c_2 \in \mathbb{R},$$
(2.18)

falls $|\delta_n|n^{\alpha} \to \infty$ für $n \to \infty$ gilt. Ein analoges Resultat erhält man für \check{a}_k und \check{b}_k mit einem $\theta < \lambda$. Daraus lassen sich Bedingungen formulieren, so dass die Summe der quadratischen Abweichungen, die über den Changepoint k^* hinweg aufsummiert, stochastisch gegen Unendlich konvergiert und damit $\hat{\sigma}_k^2$ keine konsistente Schätzungen mehr liefern kann.

Bei der Statistik \hat{L}_n aus (2.17) reicht es unter der Alternative H_A jedoch sicherlich aus, dass die Bedingung $\hat{\sigma}_k = o_P(\delta_n n^{\alpha+1/2}), n \to \infty$, und $|\delta_n| n^{\alpha+1/2} \to \infty, n \to \infty$, für ein $2 \le k \le n-2$ erfüllt ist, um nach Satz 2.6 die Konsistenz des Tests garantieren zu können. Der Motivation der Schätzer $\hat{\sigma}_k^2$ folgend, kann man sogar für $k = k^*$ einen konsistenten Schätzer erwarten, ohne jedoch auch unter der Alternative die Varianz schätzen zu können, da der Changepoint k^* unbekannt ist.

Satz 2.9

Set $\{y_i : 1 \leq i \leq n\}$ wie in Definition 2.1 für alle $n \in \mathbb{N}$. Weiter set $\frac{k^*}{n} \to \lambda$ für $n \to \infty$ und $\lambda \in (0, 1)$ erfüllt. Dann gilt unter H_A mit $2 \leq k^* \leq n - 2$:

$$\hat{\sigma}_{k^*}^2 - \sigma^2 = o_P(1).$$

Beweis: Der Beweis ist ähnlich zu dem Nachweis der Konsistenz des Schätzers unter der Nullhypothese. Mit Lemma 2.2 folgt

$$||Y_{k^*} - T_{k^*} \hat{\gamma}_{k^*}||^2 = Y_{k^*}^T Y_{k^*} - Y_{k^*}^T T_{k^*} (T_{k^*}^T T_{k^*})^{-1} T_{k^*}^T Y_{k^*}$$

$$= (T_{k^*} \gamma_n + E_{k^*})^T (I - T_{k^*} (T_{k^*}^T T_{k^*})^{-1} T_{k^*}^T) (T_{k^*} \gamma_n + E_{k^*})$$

$$= \sum_{i=1}^{k^*} e_i^2 - E_{k^*}^T T_{k^*} (T_{k^*}^T T_{k^*})^{-1} T_{k^*}^T E_{k^*}.$$
(2.19)

Mit der Ungleichung von Chebyshev und Satz A.2 erhält man

$$\frac{1}{\sqrt{k^*}} \sum_{i=1}^{k^*} e_i = O_P(1), \quad n \to \infty, \quad \text{und} \quad \frac{1}{k^{*1/2+\alpha}} \sum_{i=1}^{k^*} t_i e_i = O_P(1), \quad n \to \infty.$$

Damit folgt mit den Sätzen A.2 und A.4.a)

$$E_{k^*}^T T_{k^*} (T_{k^*}^T T_{k^*})^{-1} T_{k^*}^T E_{k^*}$$

$$= \frac{1}{d_{k^*}} \left(k^* \left(\sum_{i=1}^{k^*} t_i e_i \right)^2 - 2 \left(\sum_{i=1}^{k^*} t_i \right) \left(\sum_{i=1}^{k^*} e_i \right) \left(\sum_{i=1}^{k^*} t_i e_i \right) + \left(\sum_{i=1}^{k^*} t_i^2 \right) \left(\sum_{i=1}^{k^*} e_i \right)^2 \right)$$

$$= O_P(1).$$

Für den ersten Summanden der letzten Zeile in (2.19) gilt nach dem starken Gesetz der großen Zahlen

$$\left|\sum_{i=1}^{k^*} e_i^2 - k^* \sigma^2\right| = o(n) \text{ P-f.s.}, \quad n \to \infty,$$

da $k^* \to \infty$ für $n \to \infty$. Somit ergibt sich

$$\frac{1}{n} \left| \left\| Y_{k^*} - T_{k^*} \hat{\gamma}_{k^*} \right\|^2 - k^* \sigma^2 \right| = o_P(1).$$
(2.20)

Für die Summe nach dem Changepoint erhält man analog

$$||Y_{k^*}^* - T_{k^*}^* \check{\gamma}_{k^*}||^2 = \sum_{i=k^*+1}^n e_i^2 - E_{k^*}^{*T} T_{k^*}^* (T_{k^*}^{*T} T_{k^*}^*)^{-1} T_{k^*}^{*T} E_{k^*}^*.$$

Auch hier folgt mit der Ungleichung von Chebyshev und Satz A.2

$$\frac{1}{\sqrt{n}} \sum_{i=k^*+1}^n e_i = O_P(1), \quad n \to \infty, \quad \text{und} \quad \frac{1}{n^{1/2+\alpha}} \sum_{i=k^*+1}^n t_i e_i = O_P(1), \quad n \to \infty.$$

Da nach Voraussetzung $k^* \leq n(1-\epsilon)$ für ein $\epsilon>0$ und n groß genug gilt, folgt mit den Sätzen A.2 und A.4.c)

$$E_{k^*}^{*T} T_{k^*}^* (T_{k^*}^{*T} T_{k^*}^*)^{-1} T_{k^*}^{*T} E_{k^*}^*$$

$$= \frac{1}{d_{k^*}^*} \left((n - k^*) \left(\sum_{i=k^*+1}^n t_i e_i \right)^2 - 2 \left(\sum_{i=k^*+1}^n t_i \right) \left(\sum_{i=k^*+1}^n e_i \right) \left(\sum_{i=k^*+1}^n t_i e_i \right) \right)$$

$$+ \left(\sum_{i=k^*+1}^n t_i^2 \right) \left(\sum_{i=k^*+1}^n e_i \right)^2 \right)$$

$$= O_P(1).$$

Wegen der identischen Verteilung der Familie $\{e_i : 1 \leq i \leq n\}$ lässt sich wie oben mit $l = n - k^*$ schließen:

$$\left|\sum_{i=k^*+1}^n e_i^2 - (n-k^*)\sigma^2\right| \stackrel{\mathcal{D}}{=} \left|\sum_{i=1}^l e_i^2 - l\sigma^2\right| = o(n) \quad \text{P-f.s.}, \quad n \to \infty,$$

da auch hier $l \to \infty$ für $n \to \infty$. Somit folgt mit (2.20) die Behauptung aus:

$$\frac{1}{n} \left| \|Y_{k^*}^* - T_{k^*}^* \check{\gamma}_{k^*}\|^2 - (n - k^*)\sigma^2 \right| = o_P(1).$$

Wie bereits bemerkt, reicht die Konsistenz des Schätzers $\hat{\sigma}_{k^*}^2$ unter der Alternative mit Changepoint k^* aus, um auch die Konsistenz der Teststatistik \hat{L}_n gewährleisten zu können. Das folgende Korollar stellt nochmals die Voraussetzungen und die Vorgehensweise zusammen.

Korollar 2.10

Sei $\{y_i : 1 \leq i \leq n\}$ wie in Definition 2.1 für alle $n \in \mathbb{N}$. Weiter sei $\frac{k^*}{n} \to \lambda$ für $n \to \infty$ und $\lambda \in (0, 1)$ sowie $|\delta_n| n^{1/2+\alpha} \to \infty$ für $n \to \infty$ erfüllt. Dann gilt unter H_A mit $2 \leq k^* \leq n-2$:

$$\hat{L}_n = \frac{\alpha}{(\alpha+1)\sqrt{(2\alpha+1)}} n^{\alpha+1/2} \max_{2 \le k \le n} \left| \frac{1}{\sqrt{\hat{\sigma}_k^2}} Q_k \right| \xrightarrow{P} \infty, \quad n \to \infty.$$

Beweis: Mit Satz 2.9 gilt $\hat{\sigma}_{k^*}^2 = O_P(1)$ und damit auch $\hat{\sigma}_{k^*} = o_P(\delta_n n^{\alpha+1/2})$, womit die Behauptung aus Satz 2.6 folgt.

2.3 Der Minimumschätzer

Die Diskussion nach Korollar 2.8 des letzten Abschnitts legt die Vermutung nahe, dass unter der Alternative die Schätzer $\hat{\sigma}_k^2$ für $k \neq k^*$ und $k = [\theta n]$, $\theta \in (0, 1)$ bei wachsender Stichprobengröße *n* stochastisch gegen Unendlich konvergieren. Andererseits ist der Schätzer $\hat{\sigma}_{k^*}^2$ unter der Alternative, falls die Voraussetzungen von Satz 2.9 erfüllt sind, ein konsistenter Schätzer und somit stochastisch beschränkt (= $O_P(1)$). Damit ist es nahe liegend, über allen $\hat{\sigma}_k^2$, $2 \leq k \leq n-2$, zu minimieren, um asymptotisch den Varianzschätzer $\hat{\sigma}_{k^*}^2$ an der Stelle des Changepoints k^* zu erhalten. Die Bezeichnungen aus Abschnitt 2.2 benutzend, betrachten wir den Varianzschätzer

$$\tilde{\sigma}_{n}^{2} := \min_{2 \le k \le n-2} \{ \hat{\sigma}_{k}^{2} \} = \frac{1}{n} \min_{2 \le k \le n-2} \left\{ \sum_{i=1}^{k} \left(y_{i} - \hat{b}_{k} t_{i} - \hat{a}_{k} \right)^{2} + \sum_{i=k+1}^{n} \left(y_{i} - \check{b}_{k} t_{i} - \check{a}_{k} \right)^{2} \right\}$$
$$= \frac{1}{n} \min_{2 \le k \le n-2} \left\{ \|Y_{k} - T_{k} \hat{\gamma}_{k}\|^{2} + \|Y_{k}^{*} - T_{k}^{*} \check{\gamma}_{k}\|^{2} \right\}, \quad n \ge 2.$$
(2.21)

Aufgrund der abzählbaren Argumente des Minimums ist die Funktion $\tilde{\sigma}_n^2$ messbar. Da die Schätzer $\hat{\sigma}_k^2$ unter der Nullhypothese für alle $2 \le k \le n-2$ konsistent sind, überträgt sich dieses Ergebnis auf $\tilde{\sigma}_n^2$:

Satz 2.11

Set $\{y_i : 1 \leq i \leq n\}$ wie in Definition 2.1 für alle $n \in \mathbb{N}$. Dann gilt unter H_0 :

$$\left|\min_{2\le k\le n-2} \left\{ \hat{\sigma}_k^2 \right\} - \sigma^2 \right| = o_P(1)$$

Beweis: Die Behauptung folgt sofort mit Satz 2.7 aus der Ungleichung

$$|\min_{2 \le k \le n-2} \left\{ \hat{\sigma}_k^2 \right\} - \sigma^2 | \le \max_{2 \le k \le n-2} |\hat{\sigma}_k^2 - \sigma^2| = o_P(1).$$

Wie in der Einleitung dieses Abschnitts bemerkt, erhält man mit dem Minimumschätzer $\tilde{\sigma}_n^2$ einen Varianzschätzer, der sogar unter der Alternative konsistent ist. Diese Aussage beinhaltet der folgende Satz:

Satz 2.12

Sei $\{y_i : 1 \leq i \leq n\}$ wie in Definition 2.1 für alle $n \in \mathbb{N}$. Weiter sei $\frac{k^*}{n} \to \lambda$ für $n \to \infty$ und $\lambda \in (0, 1)$ sowie $|\delta_n|n^{\alpha-1/2} \to \infty$ für $n \to \infty$ erfüllt. Dann gilt unter H_A mit $2 \leq k^* \leq n-2$:

$$|\min_{2 \le k \le n-2} \{\hat{\sigma}_k^2\} - \sigma^2| = o_P(1).$$

Die grundlegende Idee des Beweises ist zu zeigen, dass unter der Alternative für $n \to \infty$

$$P\left(\min_{2\leq k\leq n-2}\left\{\|Y_k - T_k\hat{\gamma}_k\|^2 + \|Y_k^* - T_k^*\check{\gamma}_k\|^2\right\} = \|Y_{k^*} - T_{k^*}\hat{\gamma}_{k^*}\|^2 + \|Y_{k^*}^* - T_{k^*}^*\check{\gamma}_{k^*}\|^2\right) \to 1$$
(2.22)

gilt. Um dies zu erreichen, muss der unbeschränkte Anteil mit der größten Wachstumsrate von $||Y_k - T_k \hat{\gamma}_k||^2 + ||Y_k^* - T_k^* \check{\gamma}_k||^2$ für alle $k \neq k^*$ bei steigender Anzahl n von Beobachtungen schnell genug wachsen, damit das Minimum an der Stelle des Changepoints k^* asymptotisch angenommen wird. Dies spiegelt sich in der Voraussetzung $|\delta_n|n^{\alpha-1/2} \to \infty$ für $n \to \infty$ wieder, die für sich genommen seltsam anmutet, da man gerade bei schnell zunehmender Ununterscheidbarkeit der Verteilungen der Beobachtungen vor und nach dem Changepoint ein verbessertes Verhalten der Schätzer $\hat{\sigma}_k^2$ erwartet. Auf diese Vermutung werden wir nach dem Beweis des Satzes eingehen.

Es sei an dieser Stelle darauf hingewiesen, dass mittels (2.22) sich leicht Bedingungen finden lassen, um

$$\hat{k}_n := \min \left\{ k : \|Y_k - T_k \hat{\gamma}_k\|^2 + \|Y_k^* - T_k^* \check{\gamma}_k\|^2 \\ = \min_{2 \le l \le n-2} \left\{ \|Y_l - T_l \hat{\gamma}_l\|^2 + \|Y_l^* - T_l^* \check{\gamma}_l\|^2 \right\} \right\}$$

als einen konsistenten Schätzer für den Changepoint k* unter der Alternative nachzuweisen.

Der Beweis von Satz 2.12 erfolgt, indem die Beobachtungen Y_k und Y_k^* zerlegt werden in einen Anteil, der der Verteilung vor bzw. nach dem Changepoint unterliegt, und einen, der die Änderung nach bzw. vor dem Changepoint beschreibt. Dazu werden noch folgende Notationen eingeführt, bei denen auf eine Indizierung nach der Anzahl n der Beobachtungen verzichtet wird:

$$T_{k^*,k} := \begin{pmatrix} 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \\ t_{k^*+1} & 1 \\ \vdots & \vdots \\ t_k & 1 \end{pmatrix} \in \mathbb{R}^{k \times 2} \quad \text{für } k^* + 1 \le k \le n,$$
$$T_{k^*,k} := 0 \in \mathbb{R}^{k \times 2} \quad \text{für } 1 \le k \le k^*,$$

$$T_{k^*,k}^* := \begin{pmatrix} t_{k+1} & 1 \\ \vdots & \vdots \\ t_{k^*} & 1 \\ 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \end{pmatrix} \in \mathbb{R}^{(n-k)\times 2} \quad \text{für } 0 \le k \le k^* - 1,$$
$$T_{k^*,k}^* := 0 \in \mathbb{R}^{(n-k)\times 2} \quad \text{für } k^* \le k \le n - 1,$$
$$D_k := T_k (T_k^T T_k)^{-1} T_k^T, \quad 2 \le k \le n,$$
$$D_k^* := T_k^* (T_k^{*T} T_k^*)^{-1} T_k^{*T}, \quad 0 \le k \le n - 2,$$

wobei D_k und D_k^* symmetrische und idempotente Matrizen sind. Die beiden Summen der Varianzschätzer $\hat{\sigma}_k^2$ werden für $2 \le k \le n-2$ wieder als Normen von Vektoren dargestellt, die auf folgende Art zerlegt werden:

$$\begin{aligned} \|Y_{k} - T_{k}\hat{\gamma}_{k}\|^{2} &= \|Y_{k} - D_{k}Y_{k}\|^{2} \\ &= Y_{k}^{T}(I - D_{k})Y_{k} \\ &= (T_{k}\gamma_{n} + T_{k^{*},k}\Delta_{n} + E_{k})^{T}(I - D_{k})(T_{k}\gamma_{n} + T_{k^{*},k}\Delta_{n} + E_{k}) \\ &= (T_{k}\gamma_{n} + E_{k})^{T}(I - D_{k})(T_{k}\gamma_{n} + E_{k}) + (T_{k^{*},k}\Delta_{n})^{T}(I - D_{k})(T_{k^{*},k}\Delta_{n}) \\ &+ 2(T_{k}\gamma_{n} + E_{k})^{T}(I - D_{k})(T_{k^{*},k}\Delta_{n}) \\ &=: A_{k} + B_{k} + 2C_{k} \end{aligned}$$

und analog

$$\begin{aligned} \|Y_{k}^{*} - T_{k}^{*}\check{\gamma}_{k}\|^{2} &= (T_{k}^{*}\gamma_{n}^{*} + E_{k}^{*})^{T}(I - D_{k}^{*})(T_{k}^{*}\gamma_{n}^{*} + E_{k}^{*}) + (T_{k^{*},k}^{*}\Delta_{n})^{T}(I - D_{k}^{*})(T_{k^{*},k}^{*}\Delta_{n}) \\ &- 2(T_{k}^{*}\gamma_{n}^{*} + E_{k}^{*})^{T}(I - D_{k}^{*})(T_{k^{*},k}^{*}\Delta_{n}) \\ &=: A_{k}^{*} + B_{k}^{*} - 2C_{k}^{*}. \end{aligned}$$

Das Verhalten der einzelnen Summanden der Zerlegungen soll in den folgenden vier Lemmata untersucht werden.

Lemma 2.13

Set $\{y_i : 1 \leq i \leq n\}$ wie in Definition 2.1 für alle $n \in \mathbb{N}$. Dann gilt:

a)
$$\max_{2 \le k \le n} |A_k| = O_P(n).$$

b)
$$\max_{0 \le k \le n-2} |A_k^*| = O_p(n).$$

c) $\max_{2 \le k \le n-2} |A_k + A_k^* - n\sigma^2| = o_P(n).$

Beweis:

a) Sei 2 $\leq k \leq n$. Dann gilt

$$A_{k} = (T_{k}\gamma_{n} + E_{k})^{T}(I - D_{k})(T_{k}\gamma_{n} + E_{k})$$
$$= \sum_{i=1}^{k} e_{i}^{2} - E_{k}^{T}T_{k}(T_{k}^{T}T_{k})^{-1}T_{k}^{T}E_{k},$$

wobei der zweite Summand in der letzten Zeile folgende Darstellung besitzt

$$E_{k}^{T}T_{k}(T_{k}^{T}T_{k})^{-1}T_{k}^{T}E_{k} = \frac{1}{d_{k}}\left(k\left(\sum_{i=1}^{k}t_{i}e_{i}\right)^{2} - 2\left(\sum_{i=1}^{k}t_{i}\right)\left(\sum_{i=1}^{k}e_{i}\right)\left(\sum_{i=1}^{k}t_{i}e_{i}\right) + \left(\sum_{i=1}^{k}t_{i}^{2}\right)\left(\sum_{i=1}^{k}e_{i}\right)^{2}\right).$$

Nach dem starken Gesetz der großen Zahlen gilt für den ersten Summanden

$$\max_{1 \le k \le n} \left| \sum_{i=1}^{k} e_i^2 \right| = O_P(n),$$

während für die restlichen Summanden wie im Beweis zu Satz 2.7 gilt:

$$\begin{split} \max_{2 \le k \le n} \left| \frac{1}{d_k} k \left(\sum_{i=1}^k t_i e_i \right)^2 \right| &= O_P(\log \log n), \\ \max_{2 \le k \le n} \left| \frac{1}{d_k} \left(\sum_{i=1}^k t_i \right) \left(\sum_{i=1}^k e_i \right) \left(\sum_{i=1}^k t_i e_i \right) \right| &= O_P(\log \log n), \\ \max_{2 \le k \le n} \left| \frac{1}{d_k} \left(\sum_{i=1}^k t_i^2 \right) \left(\sum_{i=1}^k e_i \right)^2 \right| &= O_P(\log \log n). \end{split}$$

Damit folgt die Behauptung aus

$$\max_{2 \le k \le n} |A_k| = O_P(n) + O_P(\log \log n) = O_P(n).$$

b) Hier erhält man analog für $0 \leq k \leq n-2$ die Darstellung

$$\begin{aligned} A_k^* &= (T_k^* \gamma_n^* + E_k^*)^T (I - D_k^*) (T_k^* \gamma_n^* + E_k^*) \\ &= E_k^{*T} E_k^* - E_k^{*T} T_k^* (T_k^{*T} T_k^*)^{-1} T_k^{*T} E_k^* \\ &= \sum_{i=k+1}^n e_i^2 - \frac{1}{d_k^*} \left((n-k) \left(\sum_{i=k+1}^n t_i e_i \right)^2 - 2 \left(\sum_{i=k+1}^n t_i \right) \left(\sum_{i=k+1}^n e_i \right) \left(\sum_{i=k+1}^n t_i e_i \right) \right. \\ &+ \left(\sum_{i=k+1}^n t_i^2 \right) \left(\sum_{i=k+1}^n e_i \right)^2 \right). \end{aligned}$$

Aufgrund der identischen Verteilung der Familie $\{e_i : 1 \le i \le n\}$ folgt wie oben mittels des starken Gesetzes der großen Zahlen

$$\max_{0 \le k \le n-1} \left| \sum_{i=k+1}^n e_i^2 \right| = O_P(n).$$

Wie im Beweis zu Satz 2.7 gilt für $f(n) = n - 2 - \left[\frac{n}{\log\log n}\right]$

$$\max_{0 \le k \le f(n)} \left| \sum_{i=k+1}^{n} e_i^2 - E_k^{*T} T_k^* (T_k^{*T} T_k^*)^{-1} T_k^{*T} E_k^* \right| = O_P(n) + O_P\left((\log \log n)^4 \right)$$
$$= O_P(n)$$

und auch

$$\begin{split} \max_{f(n) \le k \le n-2} \left| \sum_{i=k+1}^{n} e_i^2 - E_k^{*T} T_k^* (T_k^{*T} T_k^*)^{-1} T_k^{*T} E_k^* \right| \\ &= \max_{f(n) \le k \le n-2} \left| \sum_{i=k+1}^{n} e_i^2 - \frac{\left(\sum_{i=k+1}^{n} t_i e_i - \frac{1}{n-k} \left(\sum_{i=k+1}^{n} t_i \right) \left(\sum_{i=k+1}^{n} e_i \right) \right)^2}{\sum_{i=k+1}^{n} t_i^2 - \frac{1}{n-k} \left(\sum_{i=k+1}^{n} t_i \right)^2} \right. \\ &- \frac{1}{n-k} \left(\sum_{i=k+1}^{n} e_i \right)^2 \right| \\ &= O_P(n) + O_P\left(\frac{n}{\log \log n} \right) \\ &= O_P(n). \end{split}$$

Aus den letzten beiden Gleichungen resultiert dann die Behauptung:

$$\max_{0 \le k \le n-2} |A_k^*| \le \max_{0 \le k \le f(n)} |A_k^*| + \max_{f(n) \le k \le n-2} |A_k^*| = O_P(n) .$$

c) Wie bereits in den beiden ersten Teilen des Lemmas folgt auch hier die Aussage analog zu Satz 2.7:

$$\frac{1}{n} \max_{2 \le k \le n-2} |A_k + A_k^* - n\sigma^2|$$

$$\leq \frac{1}{n} \max_{2 \le k \le n-2} \left| \sum_{i=1}^k e_i^2 - k\sigma^2 - E_k^T T_k (T_k^T T_k)^{-1} T_k^T E_k \right|$$

$$+ \frac{1}{n} \max_{2 \le k \le n-2} \left| \sum_{i=k+1}^n e_i^2 - (n-k)\sigma^2 - E_k^{*T} T_k^* (T_k^{*T} T_k^*)^{-1} T_k^{*T} E_k^* \right|$$

$$= o_P(1).$$

Lemma 2.14

Sei $\{y_i : 1 \leq i \leq n\}$ wie in Definition 2.1 für festes $n \in \mathbb{N}$. Dann gilt für $2 \leq k^* \leq n-2$:

- a) $B_k = 0 \ f\ddot{u}r \ 2 \le k \le k^* \ und \ B_k^* = 0 \ f\ddot{u}r \ k^* \le k \le n-2.$
- b) $B_k \ge B_{k-1} \ge \ldots \ge B_{k^*} = 0$ für $k^* + 1 \le k \le n$.

c)
$$B_k^* \ge B_{k+1}^* \ge \ldots \ge B_{k^*}^* = 0$$
 für $0 \le k \le k^* - 1$

Beweis:

- a) Ergibt sich aus der Definition von B_k und B_k^* .
- b) Sei $k^* + 1 \leq k \leq n$. Mit $P_k := T_k^T T_k$ folgt $T_{k^*,k}^T T_{k^*,k} = P_k - P_{k^*}$ und $T_{k^*,k}^T T_k = P_k - P_{k^*}$. Damit ergibt sich die Darstellung

$$B_{k} = (T_{k^{*},k}\Delta_{n})^{T}(I - T_{k}(T_{k}^{T}T_{k})^{-1}T_{k}^{T})(T_{k^{*},k}\Delta_{n})$$

= $\Delta_{n}^{T}(P_{k} - P_{k^{*}} - (P_{k} - P_{k^{*}})P_{k}^{-1}(P_{k} - P_{k^{*}}))\Delta_{n}$
= $\Delta_{n}^{T}P_{k^{*}}\Delta_{n} - (P_{k^{*}}\Delta_{n})^{T}P_{k}^{-1}(P_{k^{*}}\Delta_{n}).$ (2.23)

Weiter gelten für die quadratischen Formen P_k das Monotonieverhalten

 $v^T P_{k-1} v \leq v^T P_k v, \quad \forall v \in \mathbb{R}^2,$

woraus mit der positiven Definitheit⁵

$$v^T P_{k-1}^{-1} v \ge v^T P_k^{-1} v, \quad \forall v \in \mathbb{R}^2,$$

folgt und damit die Behauptung aus der Gleichung (2.23).

c) analog zu b).

Lemma 2.15

Sei $\{y_i : 1 \le i \le n\}$ wie in Definition 2.1 für alle $n \in \mathbb{N}$. Es sei $\frac{k^*}{n} \to \lambda$ für $n \to \infty$ und $\lambda \in (0, 1)$ erfüllt. Dann gilt unter H_A mit $2 \le k^* \le n - 2$:

a) $\min_{k^*+1 \le k \le n} B_k = B_{k^*+1} = \delta_n^2 n^{2\alpha} (\lambda^{2\alpha} + o(1)).$

b)
$$\min_{0 \le k \le k^* - 1} B_k^* = B_{k^* - 1}^* = \delta_n^2 n^{2\alpha} (\lambda^{2\alpha} + o(1)).$$

c)
$$\max_{2 \le k \le n} |B_k| = O(\delta_n^2 n^{1+2\alpha})$$
.

d)
$$\max_{0 \le k \le n} |B_k^*| = O(\delta_n^2 n^{1+2\alpha}).$$

⁵siehe hierzu Theorem 20.1 und 33.1 in [Pra94].

Beweis:

a) Nach dem vorangegangenen Lemma gilt sicherlich $\min_{k^*+1 \le k \le n} B_k = B_{k^*+1}$. Weiter ergibt sich für ein beliebiges $k^* + 1 \le k \le n$:

$$B_{k} = (T_{k^{*},k}\Delta_{n})^{T}(I - T_{k}(T_{k}^{T}T_{k})^{-1}T_{k}^{T})(T_{k^{*},k}\Delta_{n})$$

$$= \delta_{n}^{2}\left(\sum_{i=k^{*}+1}^{k} t_{i}^{2} - \frac{1}{d_{k}}\left(k\left(\sum_{i=k^{*}+1}^{k} t_{i}^{2}\right)^{2} - 2\left(\sum_{i=k^{*}+1}^{k} t_{i}\right)\left(\sum_{i=1}^{k} t_{i}\right)\right)\left(\sum_{i=k^{*}+1}^{k} t_{i}^{2}\right)$$

$$+ \left(\sum_{i=k^{*}+1}^{k} t_{i}\right)^{2}\left(\sum_{i=1}^{k} t_{i}^{2}\right)\right). \qquad (2.24)$$

Damit erhält man die Konvergenzrate für B_{k^*+1} mit den Sätzen A.2 und A.4:

$$B_{k^*+1} = \delta_n^2 \left(t_{k^*+1}^2 - \frac{1}{d_{k^*+1}} \left((k^*+1) t_{k^*+1}^4 - 2t_{k^*+1}^3 \left(\sum_{i=1}^{k^*+1} t_i \right) + t_{k^*+1}^2 \left(\sum_{i=1}^{k^*+1} t_i^2 \right) \right) \right)$$

= $\delta_n^2 \left((k^*+1)^{2\alpha} + (k^*+1)^{2\alpha-1} O(1) \right)$
= $\delta_n^2 n^{2\alpha} \left(\lambda^{2\alpha} + o(1) \right).$

b) Hier gewinnt man für beliebiges $0 \leq k \leq k^* - 1$ die Darstellung

$$B_{k}^{*} = \delta_{n}^{2} \left(\sum_{i=k+1}^{k^{*}} t_{i}^{2} - \frac{1}{d_{k}^{*}} \left((n-k) \left(\sum_{i=k+1}^{k^{*}} t_{i}^{2} \right)^{2} - 2 \left(\sum_{i=k+1}^{n} t_{i} \right) \left(\sum_{i=k+1}^{k^{*}} t_{i} \right) + \left(\sum_{i=k+1}^{k^{*}} t_{i} \right)^{2} \left(\sum_{i=k+1}^{n} t_{i}^{2} \right) \right) \right).$$

Ebenfalls durch das vorangegangene Lemma ergibt sich $\min_{1 \le k \le k^*-1} B_k^* = B_{k^*-1}^*$. Weiter gilt nach Voraussetzung $k^* \le n(1-\epsilon)$ für ein $\epsilon > 0$, so dass mit Satz A.4.c)

$$\max_{2 \le k \le k^*} \left\{ \frac{n^{2+2\alpha}}{d_k^*} \right\} = O(1)$$

folgt. Damit erhält man wie in a)

$$\begin{split} B_{k^*-1}^* &= \delta_n^2 \left(t_{k^*}^2 - \frac{1}{d_{k^*-1}^*} \left(\left(n - k^* + 1 \right) t_{k^*}^4 - 2t_{k^*}^3 \left(\sum_{i=k^*}^n t_i \right) + t_{k^*}^2 \left(\sum_{i=k^*}^n t_i^2 \right) \right) \right) \\ &= \delta_n^2 \left(k^{*2\alpha} + O(1) k^{*2\alpha - 1} \right) \\ &= \delta_n^2 n^{2\alpha} \left(\lambda^{2\alpha} + o(1) \right) \end{split}$$

und damit die Behauptung.

c) Noch einmal das vorausgegange Lemma benutzend, erhält man:

$$\max_{2 \le k \le n} |B_k| = B_n.$$

Aus Satz A.2 gewinnt man folgende Konvergenzraten:

$$\begin{vmatrix} \sum_{i=k^*+1}^n t_i^2 \end{vmatrix} = O\left(n^{1+2\alpha}\right), \\ \left| \frac{n}{d_n} \left(\sum_{i=k^*+1}^n t_i^2 \right)^2 \right| = O\left(n^{1+2\alpha}\right), \\ \left| \frac{1}{d_n} \left(\sum_{i=k^*+1}^n t_i \right) \left(\sum_{i=1}^n t_i \right) \left(\sum_{i=k^*+1}^n t_i^2 \right) \right| = O\left(n^{1+2\alpha}\right), \\ \left| \frac{1}{d_n} \left(\sum_{i=k^*+1}^n t_i \right)^2 \left(\sum_{i=1}^n t_i^2 \right) \right| = O\left(n^{1+2\alpha}\right), \end{aligned}$$

und damit folgt aus (2.24) die Behauptung.

d) Ebenfalls mittels des vorigen Lemma erhält man

$$\max_{0 \le k \le n-2} |B_k^*| = B_0.$$

Unter Beachtung von Satz A.4.c) ergibt sich $n^{2+2\alpha}/d_0^* = O(1)$, so dass der verbleibende Beweis jetzt analog zum Beweis der Aussage c) geführt werden kann.

Lemma 2.16

Sei $\{y_i : 1 \le i \le n\}$ wie in Definition 2.1 für alle $n \in \mathbb{N}$. Es sei $\frac{k^*}{n} \to \lambda$ für $n \to \infty$ und $\lambda \in (0, 1)$ erfüllt. Dann gilt unter H_A mit $2 \le k^* \le n - 2$:

a)
$$\max_{2 \le k \le n} |C_k| = O_P(\delta_n n^{1/2 + \alpha}).$$

b)
$$\max_{0 \le k \le n-2} |C_k^*| = O_P(\delta_n n^{1/2+\alpha}).$$

Beweis:

a) Nach Definition gilt $C_k = 0$ für $2 \le k \le k^*$, so dass im Folgenden lediglich C_k für $k^* + 1 \le k \le n$ betrachtet werden muss. Dann gelangt man wieder durch einfaches Berechnen von C_k ans Ziel:

$$C_{k} = (T_{k}\gamma_{n} + E_{k})^{T}(I - \Delta_{k})(T_{k^{*},k}\Delta_{n})$$

$$= E_{k}^{T}(I - D_{k})T_{k^{*},k}\Delta_{n} + \gamma_{n}^{T}T_{k}^{T}(I - D_{k})T_{k^{*},k}\Delta_{n}$$

$$= E_{k}^{T}(I - T_{k}(T_{k}^{T}T_{k})^{-1}T_{k}^{T})T_{k^{*},k}\Delta_{n}$$

$$= \delta_{n}\left(\sum_{i=k^{*}+1}^{k} t_{i}e_{i} - \frac{1}{d_{k}}\left(k\left(\sum_{i=k^{*}+1}^{k} t_{i}^{2}\right)\left(\sum_{i=1}^{k} t_{i}e_{i}\right) - \left(\sum_{i=1}^{k} t_{i}\right)\left(\sum_{i=k^{*}+1}^{k} t_{i}^{2}\right)\left(\sum_{i=1}^{k} e_{i}\right) + \left(\sum_{i=k^{*}+1}^{k} t_{i}\right)\left(\sum_{i=1}^{k} t_{i}^{2}\right)\left(\sum_{i=1}^{k} e_{i}\right) - \left(\sum_{i=1}^{k} t_{i}\right)\left(\sum_{i=1}^{k} t_{i}e_{i}\right)\right)\right).$$

Für den ersten Summanden der letzten Zeile gilt aufgrund der identischen Verteilung der Familie $\{e_i : 1 \le i \le n\}$ und der Ungleichung von Kolmogorov für beliebiges C > 0

$$\begin{split} P\left(\max_{k^*+1 \le k \le n} \left| \sum_{i=k^*+1}^{k} i^{\alpha} e_i \right| \ge C n^{1/2+\alpha} \right) &= P\left(\max_{k^*+1 \le k \le n} \left| \sum_{i=1}^{k-k^*} (k^*+i)^{\alpha} e_i \right| \ge C n^{1/2+\alpha} \right) \\ &= P\left(\max_{1 \le l \le n-k^*} \left| \sum_{i=1}^{l} (k^*+i)^{\alpha} e_i \right| \ge C n^{1/2+\alpha} \right) \\ &\le \frac{1}{C^2 n^{1+2\alpha}} \operatorname{Var}\left(\sum_{i=1}^{n-k^*} (k^*+i)^{\alpha} e_i \right) \\ &= \frac{\sigma^2}{C^2} \frac{1}{n^{1+2\alpha}} \sum_{i=k^*+1}^{n} i^{2\alpha}, \end{split}$$

womit man mit Satz A.2 erhält:

$$\max_{k^*+1 \le k \le n} \left| \sum_{i=k^*+1}^k t_i e_i \right| = O_P\left(n^{1/2+\alpha} \right).$$

Die Raten der restlichen Summanden ergeben sich mit den Sätzen A.2, A.4.a) und Lemma 2.4 ebenfalls zu $O_P(n^{1/2+\alpha})$:

$$\max_{k^*+1 \le k \le n} \left| \frac{k}{d_k} \left(\sum_{i=k^*+1}^k t_i^2 \right) \left(\sum_{i=1}^k t_i e_i \right) \right| = O(1) O_P(n^{1/2+\alpha}) = O_P(n^{1/2+\alpha}),$$

$$\max_{k^*+1 \le k \le n} \left| \frac{1}{d_k} \left(\sum_{i=1}^k t_i \right) \left(\sum_{i=k^*+1}^k t_i^2 \right) \left(\sum_{i=1}^k e_i \right) \right| = O(n^{\alpha}) O_P(\sqrt{n}) = O_P(n^{1/2+\alpha}),$$

$$\max_{k^*+1 \le k \le n} \left| \frac{1}{d_k} \left(\sum_{i=k^*+1}^k t_i \right) \left(\sum_{i=1}^k t_i^2 \right) \left(\sum_{i=1}^k e_i \right) \right| = O(n^{\alpha}) O_P(\sqrt{n}) = O_P(n^{1/2+\alpha}),$$

$$\max_{k^*+1 \le k \le n} \left| \frac{1}{d_k} \left(\sum_{i=1}^k t_i \right) \left(\sum_{i=k^*+1}^k t_i \right) \left(\sum_{i=1}^k t_i e_i \right) \right| = O(1) O_P(n^{1/2+\alpha}) = O_P(n^{1/2+\alpha}),$$

woraus die Behauptung folgt.

b) Hier gilt $C_k^* = 0$ für $k^* \le k \le n-2$. Für $0 \le k \le k^* - 1$ erhält man die Darstellung

$$C_{k}^{*} = \delta_{n} \left(\sum_{i=k+1}^{k^{*}} t_{i} e_{i} - \frac{1}{d_{k}^{*}} \left((n-k) \left(\sum_{i=k+1}^{k^{*}} t_{i}^{2} \right) \left(\sum_{i=k+1}^{n} t_{i} e_{i} \right) - \left(\sum_{i=k+1}^{n} t_{i} \right) \left(\sum_{i=k+1}^{k^{*}} t_{i}^{2} \right) \left(\sum_{i=k+1}^{n} e_{i} \right) + \left(\sum_{i=k+1}^{k^{*}} t_{i} \right) \left(\sum_{i=k+1}^{n} t_{i}^{2} \right) \left(\sum_{i=k+1}^{n} e_{i} \right) - \left(\sum_{i=k+1}^{k^{*}} t_{i} \right) \left(\sum_{i=k+1}^{n} t_{i} \right) \left(\sum_{i=k+1}^{n} t_{i} \right) \left(\sum_{i=k+1}^{n} t_{i} e_{i} \right) \right) \right).$$

Auch hier erhält man mittels der Ungleichung von Kolmogorov für beliebiges C>0

$$\begin{split} P\left(\max_{0\le k\le k^*-1} \left|\sum_{i=k+1}^{k^*} i^{\alpha} e_i\right| \ge C n^{1/2+\alpha}\right) &= P\left(\max_{0\le k\le k^*-1} \left|\sum_{i=1}^{k^*-k} \left(k^*-i+1\right)^{\alpha} e_i\right| \ge C n^{1/2+\alpha}\right) \\ &= P\left(\max_{1\le l\le k^*} \left|\sum_{i=1}^{l} \left(k^*-i+1\right)^{\alpha} e_i\right| \ge C n^{1/2+\alpha}\right) \\ &\le \frac{1}{C^2 n^{1+2\alpha}} \operatorname{Var}\left(\sum_{i=1}^{k^*} \left(k^*-i+1\right)^{\alpha} e_i\right) \\ &= \frac{\sigma^2}{C^2} \frac{1}{n^{1+2\alpha}} \sum_{i=1}^{k^*} i^{2\alpha}, \end{split}$$

und damit wieder mit Satz A.2

$$\max_{0 \le k \le k^* - 1} \left| \sum_{i=k+1}^{k^*} t_i e_i \right| = O_P\left(n^{1/2 + \alpha} \right).$$

Da nach Voraussetzung $k^*-1 \le n\;(1-\epsilon)$ für ein $\epsilon>0$ und ngroß genug gilt, erhält man mit Satz A.4.c)

$$\max_{0 \le k \le k^* - 1} \left| \frac{n^{2+2\alpha}}{d_k^*} \right| = O(1).$$

Damit erhält man analog wie in a) für die restlichen Summanden die Raten

$$\begin{split} \max_{0 \le k \le k^* - 1} \left| \frac{n - k}{d_k^*} \left(\sum_{i=k+1}^{k^*} t_i^2 \right) \left(\sum_{i=k+1}^n t_i e_i \right) \right| &= O(1) O_P(n^{1/2 + \alpha}) = O_P(n^{1/2 + \alpha}), \\ \max_{0 \le k \le k^* - 1} \left| \frac{1}{d_k^*} \left(\sum_{i=k+1}^n t_i \right) \left(\sum_{i=k+1}^{k^*} t_i^2 \right) \left(\sum_{i=k+1}^n e_i \right) \right| &= O(n^{\alpha}) O_P(\sqrt{n}) = O_P(n^{1/2 + \alpha}), \\ \max_{0 \le k \le k^* - 1} \left| \frac{1}{d_k^*} \left(\sum_{i=k+1}^{k^*} t_i \right) \left(\sum_{i=k+1}^n t_i^2 \right) \left(\sum_{i=k+1}^n e_i \right) \right| &= O(n^{\alpha}) O_P(\sqrt{n}) = O_P(n^{1/2 + \alpha}), \\ \max_{0 \le k \le k^* - 1} \left| \frac{1}{d_k^*} \left(\sum_{i=k+1}^n t_i \right) \left(\sum_{i=k+1}^{k^*} t_i \right) \left(\sum_{i=k+1}^n t_i e_i \right) \right| &= O(1) O_P(n^{1/2 + \alpha}) = O_P(n^{1/2 + \alpha}), \end{split}$$

woraus ebenfalls die Behauptung folgt.

Beweis von Satz 2.12:

Unter Benutzung der Lemmata gilt

$$\min_{\substack{2 \le k \le k^* - 1 \\ k^* + 1 \le k \le n - 2}} \left\{ \hat{\sigma}_k^2 \right\} = \frac{1}{n} \min_{\substack{2 \le k \le k^* - 1 \\ k^* + 1 \le k \le n - 2}} \left\{ \|Y_k - T_k \hat{\gamma}_k\|^2 + \|Y_k^* - T_k^* \check{\gamma}_k\|^2 \right\} \\
= \frac{1}{n} \min_{\substack{2 \le k \le k^* - 1 \\ k^* + 1 \le k \le n - 2}} \left\{ A_k + B_k + 2C_k + A_k^* + B_k^* - 2C_k^* \right\} \\
\ge \frac{1}{n} \left(\min_{\substack{2 \le k \le k^* - 1 \\ k^* + 1 \le k \le n - 2}} \left\{ B_k + B_k^* \right\} - \max_{\substack{2 \le k \le k^* - 1 \\ k^* + 1 \le k \le n - 2}} \left\{ |A_k + A_k^* + 2C_k - 2C_k^*| \right\} \right)$$

$$\geq \frac{1}{n} \left(\min\{B_{k^*+1}, B_{k^*-1}^*\} - \max_{2 \le k \le n-2} \{|A_k + A_k^* + 2C_k - 2C_k^*|\} \right)$$

= $|\delta_n| n^{\alpha - 1/2} \left(|\delta_n| n^{\alpha - 1/2} (\lambda^{2\alpha} + o(1)) + O_P(1) \right)$
 $\stackrel{P}{\to} \infty , n \to \infty,$

da nach Voraussetzung $|\delta_n|n^{\alpha-1/2} \to \infty$ für $n \to \infty$. Damit folgt mit Satz 2.9 für beliebig vorgegebenes $\epsilon > 0$ und n groß genug sowie einem geeigneten C > 0

$$P\left(\min_{2\leq k\leq n-2} \left\{\hat{\sigma}_{k}^{2}\right\} \neq \hat{\sigma}_{k^{*}}^{2}\right) \leq P\left(\min_{2\leq k\leq n-2} \left\{\hat{\sigma}_{k}^{2}\right\} \neq \hat{\sigma}_{k^{*}}^{2}, \, \hat{\sigma}_{k^{*}}^{2} \leq C\right) + \frac{\epsilon}{2}$$
$$\leq \epsilon.$$

Mittels einer nochmaligen Anwendung des Satzes 2.9 folgt dann die Behauptung.

Im Gegensatz zu dem vorigen Abschnitt, in dem wir die Familie $\{\hat{\sigma}_k^2 : 2 \leq k \leq n-2\}$ betrachtet haben, erhalten wir hier also durch $\tilde{\sigma}_n^2$ einen von kunabhängigen Varianzschätzer, der bei Erfüllen der Voraussetzungen des Satztes 2.12 sowohl unter der Nullhypothese als auch unter der Alternative konsistent ist. Indem man jetzt die Varianz σ^2 in der Statistik L_n aus (2.7) durch die Schätzung $\tilde{\sigma}_n^2$ ersetzt, erhält man mit

$$\tilde{L}_n := \frac{\alpha}{\tilde{\sigma}_n(\alpha+1)\sqrt{2\alpha+1}} n^{\alpha+1/2} \max_{2 \le k \le n} |Q_k|$$
(2.25)

eine konsistente Statistik mittels Satz 2.6 unter den Voraussetzungen von Satz 2.12. Die Grenzverteilung für $n \to \infty$ von \tilde{L}_n unter der Nullhypothese ergibt sich unmittelbar aus den Sätzen 2.11 und 2.3. Die Diskussion für $\tilde{\sigma}_n^2 > c$ für ein c > 0 überträgt sich analog aus dem vorigen Abschnitt.

An dieser Stelle greifen wir die bereits zu Beginn dieses Abschnitts geäußerte Vermutung über das Verhalten der Schätzer $\hat{\sigma}_k^2$ für $2 \le k \le n-2$ bei "schnell verschwindendem" δ_n auf. Denn betrachtet man noch einmal die Aussagen des Lemmas 2.5 und die anschließende Diskussion, erkennt man die Konsistenz der Kleinste-Quadrate-Schätzer unter der Alternative, falls $\delta_n n^{\alpha} \to 0$ für $n \to \infty$. Diese Rate reicht sogar aus, damit sich die Konsistenz auf die Familie $\{\hat{\sigma}_k^2 : 2 \le k \le n-2\}$ überträgt:

Satz 2.17

Sei $\{y_i : 1 \le i \le n\}$ wie Definition 2.1 für alle $n \in \mathbb{N}$. Es sei $\frac{k^*}{n} \to \lambda$ für $n \to \infty$ und $\lambda \in (0, 1)$ sowie $\delta_n n^{\alpha} \to 0$ für $n \to \infty$ erfüllt. Dann gilt unter H_A mit $2 \le k^* \le n - 2$:

$$\max_{2\leq k\leq n-2} \left| \hat{\sigma}_k^2 - \sigma^2 \right| = o_P(1).$$

Beweis: Mit den Lemmata 2.13.c), 2.15.c) und d) folgt

$$\begin{aligned} \max_{2 \le k \le n-2} \left| \hat{\sigma}_k^2 - \sigma^2 \right| &= \frac{1}{n} \max_{2 \le k \le n-2} \left| \|Y_k - T_k \hat{\gamma}_k\|^2 + \|Y_k^* - T_k^* \check{\gamma}_k\|^2 - n\sigma^2 \right| \\ &\leq \frac{1}{n} \max_{2 \le k \le n-2} \left| A_k + A_k^* - n\sigma^2 \right| + \frac{1}{n} \max_{2 \le k \le n-2} \left| B_k + 2C_k + B_k^* - 2C_k^* \right| \\ &= o_P(1) + O_P\left(\delta_n^2 n^{2\alpha}\right) + O_P\left(\delta_n n^{\alpha - 1/2}\right) \\ &= o_P(1). \end{aligned}$$

Dieser Satz impliziert natürlich auch die Konsistenz des Schätzers $\tilde{\sigma}_n^2$ unter den angegebenen Voraussetzungen.

Da sich die beiden Bedingungen

- i) $\delta_n n^{\alpha} \to 0, \quad n \to \infty,$
- ii) $|\delta_n| n^{\alpha+1/2} \to \infty, \quad n \to \infty,$

nicht widersprechen, folgt ebenfalls aus dem vorausgegangenen Satz, dass unter den Voraussetzungen von Satz 2.6 und der Bedingung i) die Statistik \tilde{L}_n aus (2.25) asymptotisch konsistent ist. Obwohl die Kleinste-Quadrate-Schätzer bei Erfüllen der Bedingung i) unter der Alternative konsistent sind, nimmt die Statistik \tilde{L}_n bei zusätzlichem Einhalten der Bedingung ii) noch Werte an, die groß genug sind, um eine Entscheidung für die Alternative zu treffen.

2.4 Der Alternativschätzer

Das betrachtete Modell aus Definition 2.1 lässt keinen Wechsel in der Verteilung der Fehler e_i zu. Deshalb tritt mit $\tilde{\sigma}_n^2$ des letzten Abschnitts der seltene Fall in der Changepoint-Analyse ein, dass der Varianzschätzer sowohl unter der Nullhypothese als auch unter der Alternative konsistent ist. In manchen Anwendungen ist jedoch sicherlich der zu seiner Berechnung erforderliche quadratische Aufwand von Nachteil: so ist etwa bei Datenbanksystemen eine Echtzeitanalyse der Daten erwünscht. Folgt man unter diesem Gesichtspunkt der Motivation zu Beginn von Abschnitt 2.2 und teilt die Beobachtungen in zwei Mengen $\{y_i: 1 \le i \le k_0\}$ und $\{y_i : k_0 < i \leq n\}$ ein, erhält man zwei auf der Kleinste-Quadrate-Methode basierende Varianzschätzer. Beide sind sicherlich für $k_0 = [\theta n], \theta \in (0, 1)$, unter der Nullhypothese konsistent. Da die zugrunde liegenden Regressionsparameter der Beobachtungen einer der beiden Gruppen unter der Alternative konstant sind, bleibt die Konsistenz des Varianzschätzers dieser Gruppe erhalten, während mit den Überlegungen des letzten Abschnitts der andere Schätzer stochastisch gegen Unendlich konvergiert. Wählt man wieder den kleineren Wert der beiden Schätzer aus, so ist zu vermuten, dass sich auf diese Weise ein unter der Alternative konsistenter Schätzer ergibt. Hat man keine weiteren Informationen über die eventuelle Lage des Changepoints, wählt man $k_0 = n/2$ und erhält

$$\bar{\sigma}_{n}^{2} := \min\left\{\frac{2}{n}\sum_{i=1}^{[n/2]} \left(y_{i} - \hat{b}_{[n/2]}t_{i} - \hat{a}_{[n/2]}\right)^{2}, \frac{2}{n}\sum_{i=[n/2]+1}^{n} \left(y_{i} - \check{b}_{[n/2]}t_{i} - \check{a}_{[n/2]}\right)^{2}\right\}, \\
= \min\left\{\frac{2}{n}\|Y_{[n/2]} - T_{[n/2]}\hat{\gamma}_{[n/2]}\|^{2}, \frac{2}{n}\|Y_{[n/2]}^{*} - T_{[n/2]}^{*}\check{\gamma}_{[n/2]}\|^{2}\right\}, \\
=:\min\left\{\bar{\sigma}_{1}^{2}, \bar{\sigma}_{2}^{2}\right\}, \quad n \ge 4.$$
(2.26)

Da das Verhalten des Alternativschätzers $\bar{\sigma}_n^2$ unter der Nullhypothese und unter der Alternative maßgeblich von den Argumenten $\bar{\sigma}_1^2$ und $\bar{\sigma}_2^2$ des Minimums abhängt, werden diese Schätzer zuerst untersucht. Das folgende Lemma beinhaltet die Konsistenzaussage der beiden Schätzer $\bar{\sigma}_1^2$ und $\bar{\sigma}_2^2$ unter der Nullhypothese ($k^* = 0$ bzw. $k^* = n$ gesetzt) sowie unter der Alternative von einem der beiden, wobei dies von der Lage des Changepoints k^* abhängig ist.

Lemma 2.18

Sei $\{y_i : 1 \leq i \leq n\}$ wie in Definition 2.1 für alle $n \in \mathbb{N}$. Weiter sei $\frac{k^*}{n} \to \lambda$ für $n \to \infty$ und $\lambda \in [0, 1]$ gewährleistet.

a) Sei $\{r_n : n \in \mathbb{N}\}$ eine Folge mit $r_n \to \infty$ für $n \to \infty$ und $\exists n_0 > 0: r_n \leq k^* \quad \forall n \geq n_0$ sei erfüllt. Dann gilt:

$$\frac{1}{r_n} \sum_{i=1}^{[r_n]} (y_i - \hat{b}_{[r_n]} t_i - \hat{a}_{[r_n]})^2 = \sigma^2 + o_P(1).$$

b) Sei $\{s_n : n \in \mathbb{N}\}$ eine Folge mit $n - s_n \to \infty$ für $n \to \infty$. Weiter seien $\exists n_0 > 0: s_n \ge k^* \quad \forall n \ge n_0$ und $\exists \epsilon > 0: s_n \le n(1 - \epsilon) \quad \forall n \ge n_0 \text{ erfüllt.}$ Dann gilt:

$$\frac{1}{n-s_n}\sum_{i=[s_n]+1}^n (y_i - \check{b}_{[s_n]}t_i - \check{a}_{[s_n]})^2 = \sigma^2 + o_P(1).$$

Beweis:

a) Ohne Einschränkung sei $r_n \in \mathbb{N}.$ Mit Lemma 2.2 folgt für
 $n \geq n_0$

$$||Y_{r_n} - T_{r_n} \hat{\gamma}_{r_n}||^2 = Y_{r_n}^T Y_{r_n} - Y_{r_n}^T T_{r_n} (T_{r_n}^T T_{r_n})^{-1} T_{r_n}^T Y_{r_n}$$

$$= (T_{r_n} \gamma_n + E_{r_n})^T (I - T_{r_n} (T_{r_n}^T T_{r_n})^{-1} T_{r_n}^T) (T_{r_n} \gamma_n + E_{r_n})$$

$$= \sum_{i=1}^{r_n} e_i^2 - E_{r_n}^T T_{r_n} (T_{r_n}^T T_{r_n})^{-1} T_{r_n}^T E_{r_n}.$$
(2.27)

Für den zweiten Term in der letzten Zeile gilt

$$E_{r_n}^T T_{r_n} (T_{r_n}^T T_{r_n})^{-1} T_{r_n}^T E_{r_n} = \frac{1}{d_{r_n}} \left(r_n \left(\sum_{i=1}^{r_n} t_i e_i \right)^2 - 2 \left(\sum_{i=1}^{r_n} t_i \right) \left(\sum_{i=1}^{r_n} e_i \right) \left(\sum_{i=1}^{r_n} t_i e_i \right) + \left(\sum_{i=1}^{r_n} t_i^2 \right) \left(\sum_{i=1}^{r_n} e_i \right)^2 \right).$$

Mittels der Chebyshev-Ungleichung erhält man

$$\sum_{i=1}^{r_n} e_i = O_P\left(\sqrt{r_n}\right),$$
$$\sum_{i=1}^{r_n} t_i e_i = O_P\left(r_n^{1/2+\alpha}\right).$$

Damit folgt mit den Sätzen A.2 und A.4

$$\frac{1}{d_{r_n}} r_n \left(\sum_{i=1}^{r_n} t_i e_i\right)^2 = O_P(1),$$

$$\frac{1}{d_{r_n}} \left(\sum_{i=1}^{r_n} t_i\right) \left(\sum_{i=1}^{r_n} e_i\right) \left(\sum_{i=1}^{r_n} t_i e_i\right) = O_P(1),$$

$$\frac{1}{d_{r_n}} \left(\sum_{i=1}^{r_n} t_i^2\right) \left(\sum_{i=1}^{r_n} e_i\right)^2 = O_P(1),$$

und somit

$$E_{r_n}^T T_{r_n} (T_{r_n}^T T_{r_n})^{-1} T_{r_n}^T E_{r_n} = O_P(1).$$

Für den ersten Summanden in (2.27) gilt nach dem starken Gesetz der großen Zahlen

$$\left|\sum_{i=1}^{r_n} e_i^2 - r_n \sigma^2\right| = o(r_n) \text{ P-f.s.}, \quad n \to \infty.$$

Damit folgt

$$\frac{1}{r_n} \left| \|Y_{r_n} - T_{r_n} \hat{\gamma}_{r_n}\|^2 - r_n \sigma^2 \right| = o_P(1),$$

also die Behauptung.

b) Da man hier $\frac{n^{2+2\alpha}}{d_{s_n}^*} = O(1)$ mittels Satz A.4.c) erhält, folgt die Behauptung analog wie in Teil a).

Da nach diesem Lemma beide Argumente $\bar{\sigma}_1^2$ und $\bar{\sigma}_2^2$ des Minimums unter der Nullhypothese konsistent sind, erhält man unmittelbar auch die Konsistenz des Alternativschätzers $\bar{\sigma}_n^2$:

Satz 2.19

Set $\{y_i: 1\leq i\leq n\}$ wie in Definition 2.1 für alle $n\in\mathbb{N}$. Dann gilt unter H_0 :

$$\bar{\sigma}_n^2 = \sigma^2 + o_P(1).$$

Beweis: Mit Lemma 2.18 gilt

$$\begin{aligned} |\bar{\sigma}_{n}^{2} - \sigma^{2}| \} \\ &\leq \max\left\{ \left| \frac{2}{n} \sum_{i=1}^{[n/2]} \left(y_{i} - \hat{b}_{[n/2]} t_{i} - \hat{a}_{[n/2]} \right)^{2} - \sigma^{2} \right|, \left| \frac{2}{n} \sum_{i=[n/2]+1}^{n} \left(y_{i} - \check{b}_{[n/2]} t_{i} - \check{a}_{[n/2]} \right)^{2} - \sigma^{2} \right| \right\} \\ &= o_{P}(1). \end{aligned}$$

Für den Beweis der Konsistenz des Schätzers $\bar{\sigma}_n^2$ unter der Alternative wird auf die Ergebnisse des letzten Abschnitts zurückgegriffen, denn auch hier wird die stochastische Konvergenz gegen Unendlich von einem der beiden Varianzschätzers $\bar{\sigma}_1^2$ und $\bar{\sigma}_2^2$ gezeigt, falls der Changepoint k^* nicht [n/2] entspricht. Im zuletzt genannten Fall ergibt sich die Konsistenz aus Lemma 2.18.

Satz 2.20

Sei $\{y_i : 1 \leq i \leq n\}$ wie in Definition 2.1 für alle $n \in \mathbb{N}$. Weiter sei $\frac{k^*}{n} \to \lambda$ für $n \to \infty$ und $\lambda \in (0, 1)$ sowie $|\delta_n| n^{\alpha - 1/2} \to \infty$ für $n \to \infty$ erfüllt. Dann gilt unter H_A mit $2 \leq k^* \leq n - 2$:

$$\bar{\sigma}_n^2 = \sigma^2 + o_P(1).$$

Beweis: Zur Übersichtlichkeit der Notation sei n/2 im Sinne von [n/2] im folgenden Beweis als Index benutzt. Weiter werden die Bezeichnungen des letzten Abschnitts gebraucht.

1. Fall: sei $\lambda < \frac{1}{2}$.

Dann existiert $n_0 \in \mathbb{N}$: $k^* + 1 \leq \left[\frac{n}{2}\right] \quad \forall n \geq n_0$. So gilt wie im Beweis zu Satz 2.12 für $n \geq n_0$

$$\begin{split} \|Y_{n/2} - T_{n/2}\hat{\gamma}_{n/2}\|^2 \\ &= \|Y_{n/2} - D_{n/2}Y_{n/2}\|^2 \\ &= Y_{n/2}^T(I - D_{n/2})Y_{n/2} \\ &= (T_{n/2}\gamma_n + T_{k^*,n/2}\Delta_n + E_{n/2})^T(I - D_{n/2})(T_{n/2}\gamma_n + T_{k^*,n/2}\Delta_n + E_{n/2}) \\ &= (T_{n/2}\gamma_n + E_{n/2})^T(I - D_{n/2})(T_{n/2}\gamma_n + E_{n/2}) + (T_{k^*,n/2}\Delta_n)^T(I - D_{n/2})(T_{k^*,n/2}\Delta_n) \\ &+ 2(T_{n/2}\gamma_n + E_{n/2})^T(I - D_{n/2})(T_{k^*,n/2}\Delta_n) \\ &= A_{n/2} + B_{n/2} + 2C_{n/2}. \end{split}$$

Mit Lemmata 2.13 bis 2.16 folgt dann

$$\frac{2}{n} (A_{n/2} + B_{n/2} + 2C_{n/2}) \geq \frac{2}{n} \min_{k^* + 1 \leq k \leq n} \{A_k + B_k + 2C_k\} \\
\geq \frac{2}{n} \left(\min_{k^* + 1 \leq k \leq n} \{B_k\} - \max_{k^* + 1 \leq k \leq n} \{A_k + 2C_k\} \right) \\
= |\delta_n| n^{\alpha - 1/2} \left(|\delta_n| n^{\alpha - 1/2} (2\lambda^{2\alpha} + o(1)) + O_P(1) \right) \\
\xrightarrow{P}{\to} \infty, \quad n \to \infty.$$

Mit Lemma 2.18 folgt wegen $k^* \leq [n/2]$ die Konsistenz des Schätzers $\bar{\sigma}_2^2$ und damit auch $\bar{\sigma}_n^2 = \sigma^2 + o_P(1)$.

2. Fall: $\lambda > \frac{1}{2}$ Unter Beachtung von $k^* - 1 \ge [n/2]$ erfolgt der Beweis analog zum 1.Fall.

3. Fall: $\lambda = \frac{1}{2}$

Mit den beiden vorausgegangenen Fällen und Lemma 2.18 gilt für beliebige $\epsilon, \delta > 0$

$$P\left(\left|\bar{\sigma}_{n}^{2}-\sigma^{2}\right| \geq \delta\right)$$

$$\leq P\left(\left|\bar{\sigma}_{n}^{2}-\sigma^{2}\right| \geq \delta, \quad k^{*}+1 \leq \left[\frac{n}{2}\right]\right) + P\left(\left|\bar{\sigma}_{n}^{2}-\sigma^{2}\right| \geq \delta, \quad k^{*}-1 \geq \left[\frac{n}{2}\right]\right)$$

$$+ P\left(\left|\bar{\sigma}_{n}^{2}-\sigma^{2}\right| \geq \delta, \quad k^{*}=\left[\frac{n}{2}\right]\right)$$

$$\leq \epsilon.$$

Um jetzt noch zu einer Statistik zu gelangen, ersetzen wir analog zum vorigen Abschnitt die Varianz σ^2 in der Statistik L_n aus (2.7) durch die Schätzung $\bar{\sigma}_n^2$ und erhalten mit

$$\bar{L}_n := \frac{\alpha}{\bar{\sigma}_n(\alpha+1)\sqrt{2\alpha+1}} n^{\alpha+1/2} \max_{2 \le k \le n} |Q_k|$$
(2.28)

eine konsistente Statistik mittels Satz 2.6 unter den Voraussetzungen von Satz 2.20. Die Grenzverteilung von \bar{L}_n unter der Nullhypothese ergibt sich wieder unmittelbar aus den Sätzen 2.19 und 2.3. Dabei folgt auch hier, dass unter der Nullhypothese $\bar{\sigma}_n^2 > c$ für ein c > 0auf Mengen von beliebig hohen Wahrscheinlichkeiten für n groß genug gilt.

Für den Fall eines "schnell verschwindenden" δ_n , also $\delta_n = o(n^{-\alpha})$, überträgt sich die Diskussion am Ende des letzten Abschnitts analog.

3 Simulationen

In diesem Kapitel werden die in dem Kapitel "Lineares Modell" entwickelten Varianzschätzer in Simulationen betrachtet. In Abschnitt 3.1 werden das Zustandekommen der Simulationen sowie die Abbildungen erläutert. In den Abschnitten 3.2 und 3.3 werden Verteilungsparameter der Schätzer berechnet, wobei die Ergebnisse eine Normierung der bisher betrachteten Varianzschätzer nahe legen. In den Simulationen der Abschnitte 3.4 bis 3.7 wird vor allem das Verhalten der Familien der Kleinste-Quadrate-Schätzer sowie der Familie $\{\hat{\sigma}_k^2 : 2 \le k \le n-2\}$ von Varianzschätzern untersucht. Abschnitt 3.8 beschäftigt sich mit den einzelnen Schätzern $\tilde{\sigma}_n^2$ und $\bar{\sigma}_n^2$.

3.1 Grundlagen der Simulationen

Eine einzelne Simulation entsteht, indem die Beobachtungen $\{y_i : 1 \leq i \leq n\}$ zu den vorgegebenen Parametern a_n, b_n, b_n^*, k^* und zu der Verteilung der Fehler e_i gemäß Definition 2.1 durch einen Zufallsgenerator erzeugt werden. Mittels dieser Beobachtungen $\{y_i : 1 \leq i \leq n\}$ werden die Werte der Kleinste-Quadrate-Schätzer $\hat{a}_k, \check{a}_k, \hat{b}_k$ und \check{b}_k gemäß (2.3) bis (2.6) berechnet und mit diesen die Werte von normierten Versionen der Schätzers $\hat{\sigma}_k^2$ aus (2.9), des Minimumschätzers $\tilde{\sigma}_n^2 = \min\{\hat{\sigma}_k^2 : 2 \leq k \leq n-2\}$ aus (2.21) und des Alternativschätzers $\bar{\sigma}_n^2 = \min\{\bar{\sigma}_1^2, \bar{\sigma}_2^2\}$ aus (2.26) ermittelt. Da diese Funktionen natürlich Zufallsschwankungen unterliegen, werden S Durchläufe (hier: S=500) dieser Simulation durchgeführt und die Werte der Schätzer jeweils festgehalten. Aufgrund der Ergebnisse eines Schätzers \hat{x}_k wird dessen Erwartungswert und Standardabweichung geschätzt, wobei jeweils als Berechnung zugrunde gelegt wird:

$$\hat{\mu}(k) := \frac{1}{S} \sum_{i=1}^{S} \hat{x}_{i,k},$$
(3.1)

$$\hat{S}(k) := \sqrt{\frac{\sum_{i=1}^{S} \hat{x}_{i,k}^2 - \frac{1}{S} \left(\sum_{i=1}^{S} \hat{x}_{i,k}\right)^2}{(S-1)}}.$$
(3.2)

Dabei entspricht $\hat{x}_{i,k}$ dem Wert des Schätzers \hat{x}_k in der i-ten Simulation. Da die oben erwähnten Schätzer, die mit k parametrisiert sind, einer Familie von Schätzfunktionen angehören, gehen wir für alle k, für die jeweils ein Schätzer dieser Familie existiert, wie oben vor.

In den Abbildungen wird jeweils der Mittelwert $\hat{\mu}(k)$ eines Schätzers \hat{x}_k über der Beobachtungsstelle t_k aufgetragen. Zu diesem Mittelwert $\hat{\mu}(k)$ wird noch die geschätzte Standardabweichung $\hat{S}(k)$ addiert und subtrahiert. Bei gleicher Vorgehensweise für alle k, für die jeweils eine Schätzfunktion \hat{x}_k aus der dazugehörigen Familie definiert ist, erhält man so drei Graphen, wobei der mittlere die Mittelwerte und der obere und untere eine Art "Konfidenzintervall" für die Schätzer \hat{x}_k angeben, jeweils in Abhängigkeit von der Beobachtungsstelle t_k . Die Wahl der Beobachtungsstellen als Abszisse und der Sprachgebrauch "der Schätzer \hat{x}_k an der Stelle t_k " drücken den Sachverhalt aus, dass den Kleinste-Quadrate-Schätzern eine Summe zugrunde liegt, in die entweder die links oder rechts von dieser Beobachtungsstelle liegenden Beobachtungen eingehen. Analog bestehen die Varianzschätzer $\hat{\sigma}_k^2$ aus zwei Summen, in die jeweils die links und rechts liegenden Beobachtungen eingehen.

Werden die Werte der Standardabweichungen (3.2) für kleine oder dicht bei n liegende k unverhältnismäßig groß im Gegensatz zu den übrigen, werden diese Extremwerte in den Abbil-

dungen abgeschnitten, sofern dieses Verhalten nicht weiter von Bedeutung ist, oder gesondert betrachtet. Im anderen Fall von sehr kleinen Standardabweichungen relativ zu der vorliegenden Skalierung der Ordinate erscheinen die drei oben genannten Kurven als eine, meist etwas deutlicher hervorstechende Linie. Doch da bei den Betrachtungen nur der grobe Verlauf der Standardabweichungen um den Mittelwert von Interesse ist, reicht diese Darstellung aus. Im Wesentlichen werden vier verschiedene Simulationen in den Abschnitten 3.4 bis 3.7 betrachtet, jeweils unter einem besonderen Gesichtspunkt:

- 1. allgemeines Verhalten, insbesondere in Abhängigkeit der Lage des Changepoints (Simulation 3.3);
- 2. Einfluss der Verteilung der Beobachtungspunkte $t_i = i^{\alpha}$ (Simulation 3.4);
- 3. Konvergenzverhalten (Simulation 3.5);
- 4. Abhängigkeit von der Verteilung der Fehler e_i (Simulation 3.6).

Zu Beginn jedes Abschnitts werden in einer Definition die Parameter der dort behandelten Simulation festgehalten. Natürlich liegt das Interesse einer Simulation auf dem Verhalten bei Veränderung der modellbeschreibenden Parameter, so dass bei weiteren Betrachungen dieser Simulation jede Abweichung von den vorher definierten Parametern explizit angegeben wird. In der Definition der Simulation ist bereits der Wert des Changepoints mit angegeben, der sich nur auf Betrachtungen bezieht, in denen wir von Simulationen unter der Alternative sprechen.

3.2 Erwartungstreue einiger Varianzschätzer

In Kapitel 2 werden nur asymptotische Eigenschaften der Varianzschätzer untersucht, weshalb die Betrachtungen der zur Konsistenz ebenso wichtigen Eigenschaft der Erwartungstreue hier nachgeholt werden. Dazu das folgende Lemma:

Lemma 3.1

Seien $\{y_i : 1 \leq i \leq n\}$ wie in Definition 2.1 für festes $n \in \mathbb{N}$.

a) Dann gilt unter H_0 :

$$i) \quad \mathbf{E} \ \hat{\sigma}_k^2 = \frac{n-4}{n} \sigma^2, \quad 2 \le k \le n-2.$$

$$ii) \quad \mathbf{E} \ \bar{\sigma}_1^2 = \frac{2}{n} \left(\left[\frac{n}{2} \right] - 2 \right) \sigma^2, \quad n \ge 6.$$

$$iii) \quad \mathbf{E} \ \bar{\sigma}_2^2 = \frac{2}{n} \left(n - \left[\frac{n}{2} \right] - 2 \right) \sigma^2 \quad n \ge 4$$

b) Dann gilt unter der Alternative H_A mit $2 \le k^* \le n - 2$: $\mathrm{E} \,\hat{\sigma}_{k^*}^2 = \frac{n-4}{n} \sigma^2$.

Beweis zu a):

i) Der Beweis erfolgt analog zum Nachweis der Erwartungstreue des Varianzschätzers in gewöhnlichen linearen Modellen ohne Changepoint. Dazu seien:

 I_k die k-dimensionale Einheitsmatrix,

r der Rang einer Matrix und Sp die Spur einer Matrix. Mit Theorem 1 in 2.5. aus [Sea71] folgt für $2 \le k \le n-2$

$$E\left(\left\|Y_{k}-T_{k}\hat{\gamma}_{k}\right\|^{2}\right) = E\left(Y_{k}^{T}\left(I_{k}-T_{k}\left(T_{k}^{T}T_{k}\right)^{-1}T_{k}^{T}\right)Y_{k}\right)$$

$$= \operatorname{Sp}\left(\left(I_{k}-T_{k}\left(T_{k}^{T}T_{k}\right)^{-1}T_{k}^{T}\right)I_{k}\sigma^{2}\right)$$

$$+ (T_{k}\gamma_{n})^{T}\left(I_{k}-T_{k}\left(T_{k}^{T}T_{k}\right)^{-1}T_{k}^{T}\right)\sigma^{2}$$

$$= \operatorname{Sp}\left(I_{k}-T_{k}\left(T_{k}^{T}T_{k}\right)^{-1}T_{k}^{T}\right)\sigma^{2}$$

$$= (k-2)\sigma^{2}.$$

Analog ergibt sich $E\left(\|Y_k^* - T_k^*\check{\gamma}_k\|^2\right) = (n-k-2)\sigma^2$, woraus die Behauptung folgt.

ii) und iii) Ergeben sich aus den Berechnungen in a).

zu b): Folgt wie in a.i) für $k = k^*$.

Gemäß Lemma 3.1 werden die Varianzschätzer normiert. Dabei werden **dieselben Bezeich**nungen wie bisher beibehalten, die in dieser normierten Version nur Gültigkeit in diesem Kapitel "Simulationen" haben. Man erhält für $n \ge 5$:

$$\begin{aligned} \hat{\sigma}_k^2 &:= \frac{1}{n-4} \left(\sum_{i=1}^k \left(y_i - \hat{b}_k t_i - \hat{a}_k \right)^2 + \sum_{i=k+1}^n \left(y_i - \check{b}_k t_i - \check{a}_k \right)^2 \right), \quad 2 \le k \le n-2, \\ \tilde{\sigma}_n^2 &:= \frac{1}{n-4} \min_{2\le k\le n-2} \left\{ \sum_{i=1}^k \left(y_i - \hat{b}_k t_i - \hat{a}_k \right)^2 + \sum_{i=k+1}^n \left(y_i - \check{b}_k t_i - \check{a}_k \right)^2 \right\}, \\ \bar{\sigma}_1^2 &:= \frac{1}{[n/2]-2} \sum_{i=1}^{[n/2]} \left(y_i - \hat{b}_{[n/2]} t_i - \hat{a}_{[n/2]} \right)^2, \quad n \ge 6, \\ \bar{\sigma}_2^2 &:= \frac{1}{n-[n/2]-1} \sum_{i=[n/2]+1}^n \left(y_i - \check{b}_{[n/2]} t_i - \check{a}_{[n/2]} \right)^2, \\ \bar{\sigma}_n^2 &:= \min \left\{ \bar{\sigma}_1^2, \bar{\sigma}_2^2 \right\} \quad n \ge 6. \end{aligned}$$

Somit sind die Schätzer $\hat{\sigma}_k^2$, $\bar{\sigma}_1^2$ und $\bar{\sigma}_2^2$ erwartungstreu und damit zumindest die Argumente des Minimumsfunktionals der Schätzer $\tilde{\sigma}_n^2$ und $\bar{\sigma}_n^2$. Die Resultate des Kapitels 2 übertragen sich, da dort nur asymptotische Betrachtungen ausgeführt werden.

3.3 Verteilungsparameter der Kleinste-Quadrate-Schätzer

Da die Kleinste-Quadrate-Schätzer in die Varianzschätzungen maßgeblich miteingehen, ist deren Verhalten von grundlegender Bedeutung. Charakterisieren lässt sich ihr Verhalten zum Teil durch die Angabe von Erwartungswerten und Varianzen.
Lemma 3.2

Set $\{y_i : 1 \leq i \leq n\}$ wie in Definition 2.1 für festes $n \in \mathbb{N}$.

a) Dann gilt unter der Nullhypothese für die Kleinste-Quadrate-Schätzer:

$$i) \ \mathbf{E} \, \hat{a}_{k} = a_{n}, \quad \text{Var} \, \hat{a}_{k} = \frac{\sigma^{2}}{d_{k}} \sum_{i=1}^{k} t_{i}^{2}, \quad 2 \le k \le n,$$

$$ii) \ \mathbf{E} \, \check{a}_{k} = a_{n}, \quad \text{Var} \, \check{a}_{k} = \frac{\sigma^{2}}{d_{k}^{*}} \sum_{i=k+1}^{n} t_{i}^{2}, \quad 0 \le k \le n-2,$$

$$iii) \ \mathbf{E} \, \hat{b}_{k} = b_{n}, \quad \text{Var} \, \hat{b}_{k} = \frac{k}{d_{k}} \sigma^{2}, \quad 2 \le k \le n,$$

$$iv) \ \mathbf{E} \, \check{b}_{k} = b_{n}, \quad \text{Var} \, \check{b}_{k} = \frac{n-k}{d_{k}^{*}} \sigma^{2}, \quad 0 \le k \le n-2.$$

b) Dann gilt unter der Alternative für die Kleinste-Quadrate-Schätzer bei $2 \le k^* \le n-2$:

$$\begin{split} i) & \to \hat{a}_{k} = a_{n} + \frac{1}{d_{k}} \left(b_{n} \left(\sum_{i=1}^{k^{*}} t_{i} \sum_{i=1}^{k} t_{i}^{2} - \sum_{i=1}^{k^{*}} t_{i}^{2} \sum_{i=1}^{k} t_{i} \right) \\ & + b_{n}^{*} \left(\sum_{i=1}^{k} t_{i}^{2} \sum_{i=k^{*}+1}^{k} t_{i} - \sum_{i=1}^{k} t_{i} \sum_{i=k^{*}+1}^{k} t_{i}^{2} \right) \right), \quad k^{*} + 1 \le k \le n, \\ & \to \hat{a}_{k} = a_{n}, \quad 2 \le k \le k^{*}, \quad \text{Var } \hat{a}_{k} = \frac{\sigma^{2}}{d_{k}} \sum_{i=1}^{k} t_{i}^{2}, \quad 2 \le k \le n, \\ & ii) \quad \to \hat{a}_{k} = a_{n} + \frac{1}{d_{k}^{*}} \left(b_{n} \left(\sum_{i=k^{*}+1}^{n} t_{i} \sum_{i=k^{*}+1}^{n} t_{i}^{2} - \sum_{i=k^{*}+1}^{n} t_{i}^{2} \sum_{i=k^{*}+1}^{n} t_{i} \right) \\ & + b_{n}^{*} \left(\sum_{i=k^{*}+1}^{n} t_{i}^{2} \sum_{i=k^{*}+1}^{n} t_{i} - \sum_{i=k^{*}+1}^{n} t_{i} \sum_{i=k^{*}+1}^{n} t_{i}^{2} \right) \right), \quad 0 \le k \le k^{*} - 1, \\ & \to \hat{a}_{k} = a_{n}, \quad k^{*} \le k \le n - 2, \quad \text{Var } \tilde{a}_{k} = \frac{\sigma^{2}}{d_{k}^{*}} \sum_{i=k^{*}+1}^{n} t_{i}^{2}, \quad 0 \le k \le n - 2, \\ & iii) \quad \to \hat{b}_{k} = \frac{1}{d_{k}} \left(b_{n} \left(k \sum_{i=1}^{k^{*}} t_{i}^{2} - \sum_{i=1}^{k} t_{i} \sum_{i=k^{*}+1}^{k^{*}} t_{i} \right) \\ & \quad + b_{n}^{*} \left(k \sum_{i=k^{*}+1}^{k} t_{i}^{2} - \sum_{i=1}^{k} t_{i} \sum_{i=k^{*}+1}^{k} t_{i} \right) \\ & \quad + b_{n}^{*} \left(k \sum_{i=k^{*}+1}^{k^{*}} t_{i}^{2} - \sum_{i=k^{*}+1}^{k} t_{i} \sum_{i=k^{*}+1}^{k^{*}} t_{i} \right) \\ & \quad + b_{n}^{*} \left((n-k) \sum_{i=k^{*}+1}^{k} t_{i}^{2} - \sum_{i=k^{*}+1}^{n} t_{i} \sum_{i=k^{*}+1}^{n} t_{i} \right) \\ & \quad b_{k} = \frac{1}{d_{k}^{*}} \left(b_{n} \left((n-k) \sum_{i=k^{*}+1}^{k^{*}} t_{i}^{2} - \sum_{i=k^{*}+1}^{n} t_{i} \sum_{i=k^{*}+1}^{k^{*}} t_{i} \right) \\ & \quad + b_{n}^{*} \left((n-k) \sum_{i=k^{*}+1}^{n} t_{i}^{2} - \sum_{i=k^{*}+1}^{n} t_{i} \sum_{i=k^{*}+1}^{k^{*}} t_{i} \right) \\ & \quad b_{k} = b_{n}^{*}, \quad k^{*} \le k \le n - 2, \quad \text{Var } \tilde{b}_{k} = \frac{n-k}{d_{k}^{*}} \sigma^{2}, \quad 0 \le k \le n - 2. \end{aligned}$$

,

Beweis zu a):

- i) Nach Definition gilt $\hat{a}_k = \frac{1}{d_k} \sum_{i=1}^k y_i \left(\sum_{j=1}^k t_j^2 t_i \sum_{j=1}^k t_j \right)$, woraus mittels einfacher Rechnung die Behauptungen folgen.
- ii) bis iv) Die Beweise erfolgen wie oben.

zu b:

i) Hier benutzt man für $k^* + 1 \le k \le n$ die Darstellung

$$\hat{a}_{k} = \frac{1}{d_{k}} \left(\sum_{i=1}^{k^{*}} y_{i} \left(\sum_{j=1}^{k} t_{j}^{2} - t_{i} \sum_{j=1}^{k} t_{j} \right) + \sum_{i=k^{*}+1}^{k} y_{i} \left(\sum_{j=1}^{k} t_{j}^{2} - t_{i} \sum_{j=1}^{k} t_{j} \right) \right),$$

woraus sich $E\hat{a}_k$ und $Var\hat{a}_k$ für $k^* + 1 \le k \le n$ ergeben.

Die Werte des Erwartungswertes und der Varianz für $2 \le k \le k^*$ ergeben sich wie a).

ii) bis iv) Die Beweise folgen wie in i).

Das vorangegangene Lemma gibt Erwartungswerte und Varianzen der Kleinste-Quadrate-Schätzer vollständig an, weswegen auf die Schätzung dieser Werte mittels der Daten der Simulationen verzichtet werden könnte. Doch zur graphischen Veranschaulichung geben wir zum Teil den Verlauf dieser Funktionen trotzdem an.

3.4 Simulation mit normalverteilten Fehlern und äquidistanten Beobachtungen

Der in linearen Modellen am häufigsten vorkommende und auch wichtigste Fall von unabhängigen und normalverteilten Fehlern e_i sowie äquidistanten Beobachtungsstellen soll als Erstes behandelt werden. Dazu werden Simulationen unter der Nullhypothese und unter der Alternative gemäß den Parametern der folgenden Definition betrachtet. Im Anschluss daran wird das Verhalten unter der Alternative bei kleinerer Änderung δ_n des Trends sowie bei einem sehr frühen ($k^* = 10$) wie auch späten ($k^* = 90$) Changepoint diskutiert.

Simulation 3.3

Seien $\{y_i : 1 \le i \le n\}$ Beobachtungen, die dem folgenden Modell genügen:

$$y_i := \begin{cases} \mathbf{1} + \mathbf{2}t_i + e_i & : \quad 1 \le i \le k^*, \\ \mathbf{1} + 4t_i + e_i & : \quad k^* + 1 \le i \le n \end{cases}$$

wobei e_i , $1 \le i \le n$, unabhängige und normalverteilte Zufallsvariablen mit $Ee_i = 0$ und $Var e_i = 1$ sind.



Betrachten wir zunächst das Verhalten der Kleinste-Quadrate-Schätzer unter der Nullhypothese.

Abbildung 1: Die Schätzer \hat{a}_k und \check{a}_k unter der Nullhypothese

Bei den hier vorliegenden normalverteilten Fehlern e_i lässt sich sogar die Verteilung der Schätzer angeben. Denn in der Darstellung $\hat{a}_k = \frac{1}{d_k} \sum_{i=1}^k y_i \left(\sum_{j=1}^k t_j^2 - t_i \sum_{j=1}^k t_j \right)$ sind die einzelnen Summanden unabhängig und normalverteilt, so dass man

$$\hat{a}_k \stackrel{\mathcal{D}}{=} \mathcal{N}\left(a_n, \frac{\sigma^2}{d_k}\sum_{i=1}^k t_i^2\right), \ 2 \le k \le 100, \quad \check{a}_k \stackrel{\mathcal{D}}{=} \mathcal{N}\left(a_n, \frac{\sigma^2}{d_k^*}\sum_{i=k+1}^n t_i^2\right), \ 0 \le k \le 98, \quad (3.3)$$

mittels der analogen Darstellung für \check{a}_k erhält. Da weiter $d_k = d^*_{n-k}$, $2 \le k \le 100$, bei $\alpha = 1$ gilt, ist es offensichtlich, dass für kleine k die Abweichungen der Schätzer \check{a}_{n-k} von dem zu

schätzenden Wert $a_n = 1$ deutlich höher liegen als die der Schätzer \hat{a}_k . Diesen Sachverhalt illustriert Abbildung 1, die die geschätzten Mittelwerte und Standardabweichungen der Kleinste-Quadrate-Schätzer \hat{a}_k und \check{a}_k zeigt, wobei zunächst jeweils die ersten 18 Schätzer weggelassen wurden. Hier zeigt sich bereits, dass die Schätzer \hat{a}_k bereits bei wenigen miteinbezogenen Beobachtungen y_i sehr viel weniger von dem zu schätzenden Parameter $a_n = 1$ abweichen als \check{a}_k bei der entsprechenden Anzahl von miteingehenden Beobachtungen. Bei deutlich wenigen Beobachtungen y_i , die in die jeweiligen Schätzer eingehen, nimmt dieses asymmetrische Verhalten stark zu, wie Abbildung 2 zeigt, in der das Verhalten der ersten 18 Schätzer \hat{a}_k und \check{a}_k abgebildet ist.



Abbildung 2: Die Schätzer \hat{a}_k und \check{a}_k bei wenigen miteingehenden Beobachtungen

Wie für die Schätzer des Parameters a_n lassen sich hier auch die Verteilungen der Schätzer \hat{b}_k und \check{b}_k angeben:

$$\hat{b}_k \stackrel{\mathcal{D}}{=} \mathcal{N}\left(b_n, \frac{k}{d_k}\sigma^2\right), \quad 2 \le k \le 100, \quad \check{b}_k \stackrel{\mathcal{D}}{=} \mathcal{N}\left(b_n, \frac{n-k}{d_k^*}\sigma^2\right), \quad 0 \le k \le 98.$$
(3.4)

Wegen $d_k = d_{n-k}^*$ für $2 \le k \le 100$ verhalten sich Erwartungswerte und Varianzen dieser Schätzer symmetrisch und man erhält hier sogar $\hat{b}_k \stackrel{\mathcal{D}}{=} \check{b}_{n-k}$ für $2 \le k \le 100$. Dies wird in Abbildung 3 verdeutlicht, in der die geschätzten Erwartungswerte und Standardabweichungen der Schätzer \hat{b}_k und \check{b}_k abgebildet sind, wobei die ersten 18 Schätzer auch hier weggelassen werden. Außerdem zeigen (3.3) und (3.4), dass die Standardabweichungen der Schätzer für den Parameter b_n deutlich niedriger liegen als die der Schätzer für den Parameter a_n .

Damit gelangen wir schließlich zu Betrachtungen des Verhaltens **der Varianzschätzer** $\hat{\sigma}_{\mathbf{k}}^2$ **unter der Nullhypothese**. Da normalverteilte Fehler vorliegen, folgt wie in gewöhnlichen linearen Modellen ohne Changepoint für $2 \le k \le 98$

$$\hat{\sigma}_{k}^{2} = \frac{1}{n-4} \left(\sum_{i=1}^{k} \left(y_{i} - \hat{b}_{k} t_{i} - \hat{a}_{k} \right)^{2} + \sum_{i=k+1}^{n} \left(y_{i} - \check{b}_{k} t_{i} - \check{a}_{k} \right)^{2} \right)$$
$$\stackrel{\mathcal{D}}{=} \frac{1}{n-4} \left(\sigma^{2} \chi_{k-2}^{2} + \sigma^{2} \chi_{n-k-2}^{2} \right)$$
$$\stackrel{\mathcal{D}}{=} \frac{1}{n-4} \sigma^{2} \chi_{n-4}^{2}.$$



Abbildung 3: Die Schätzer \hat{b}_k und \check{b}_k unter der Nullhypothese

Somit sind die Schätzer $\hat{\sigma}_k^2$ für alle $2 \le k \le 98$ identisch verteilt, und es ergibt sich $E\hat{\sigma}_k^2 = \sigma^2$ und $Var\hat{\sigma}_k^2 = 2\sigma^4/(n-4)$. Es erübrigt sich eine Abbildung, da nur drei parallele Geraden zu erwarten sind. Eine Veranschaulichung findet sich trotzdem in Abbildung 13, wo ein Vergleich bei verschiedenen Beobachtungsgrößen stattfindet.

Auch unter der Alternative gemäß Simulation 3.3 folgt wie bereits unter der Nullhypothese, dass die Kleinste-Quadrate-Schätzer normalverteilt sind mit Erwartungswerten und Varianzen gemäß Lemma 3.2.b). Damit gilt, dass die Schätzer vor bzw. nach dem Changepoint wie unter der Nullhypothese verteilt sind, außer \check{b}_k , dessen Erwartungswert jetzt dem zu schätzenden Wert $b_n^* = 4$ entspricht. Begründet ist dies durch den offensichtlichen Sachverhalt, dass die Beobachtungen vor bzw. nach dem Changepoint der gleichen Verteilung wie unter der Nullhypothese bei Berücksichtigung des geänderten Trends unterliegen.

Abbildungen 4 und 5 stellen den Verlauf der geschätzten Erwartungswerte und Standardabweichungen der Schätzer dar, wobei der eingetragene Wert jeweils den Mittelwert des



Abbildung 4: Die Schätzer \hat{a}_k und \check{a}_k unter der Alternative

Schätzers an der Stelle des Changepoints $k^* = 60$ angibt. Deutlich erkennbar ist das Abweichen der Graphen nach Überschreiten des Changepoints von dem vorher angenäherten Wert. In diesem Bereich sind von den Funktionen keine konsistenten Schätzungen der Parameter mehr zu erwarten. Trotzdem nimmt die geschätzte Standardabweichung mit wachsender Anzahl von Beobachtungen ab, die in die Schätzungen eingehen. Die Begründung liefert Lemma 3.2, nach dem die Varianzen unter der Nullhypothese und unter der Alternative gleich sind. Eine weitere Diskussion der Kurven mittels Lemma 3.2 ist möglich, auch bietet sich Lemma 2.5 für eine leichtere Vorstellung der Graphen an, wobei dort natürlich keine expliziten Aussagen über Verteilungsparameter gemacht sind, sondern nur über asymptotische Näherungen. Betrachten wir jetzt die Schätzer $\hat{\sigma}_k^2$ unter der Alternative. Anders als bisher zeigt Abbildung 6 nur den Verlauf der geschätzten Erwartungswerte und Abbildung 7 davon getrennt den Verlauf der geschätzten Standardabweichungen aufgrund der Daten der Simulation. Die in den Abbildungen eingetragenen Werte geben jeweils den Wert an, den der Graph an der



Abbildung 5: Die Schätzer \hat{b}_k und \check{b}_k unter der Alternative

Stelle des Changepoints $k^* = 60$ annimmt.

Offensichtlich ist unter der Alternative lediglich $\hat{\sigma}_k^2$ für $k = k^*$ ein akzeptabler Schätzer, da der Graph der Erwartungswerte sich sehr steil von dem wahren Wert $\sigma^2 = 1$ bereits in der Nähe des Changepoints entfernt. Gleichzeitig ist dies eine Bekräftigung für den Minimumschätzer $\tilde{\sigma}_n^2 = \min{\{\hat{\sigma}_k^2 : 2 \le k \le 98\}}$, von dem zu erwarten ist, dass er dem Schätzer $\hat{\sigma}_{k^*}^2$ entspricht und somit einen Schätzer darzustellt, den man auch bei bekanntem Changepoint k^* gewählt hätte.

Weiterhin sprechen für den Minimumschätzer $\tilde{\sigma}_n^2$ auch die geringen Standardabweichungen der $\hat{\sigma}_k^2$ sowohl in der Nähe des Changepoints als auch entfernt davon. Denn im Gegensatz zu den extrem großen Erwartungswerten von $\hat{\sigma}_k^2$ für $k \neq k^*$ sind die Standardabweichungen sehr gering, so dass mit großer Wahrscheinlichkeit ausgeschlossen werden kann, dass das Minimum an einer anderen Stelle als $k = k^*$ angenommen wird, wo die Streuung ebenfalls den minimalen Wert erreicht. Tatsächlich wird das Minimum in der vorliegenden Simulation



Abbildung 7: Die Standardabweichung der Schätzer $\hat{\sigma}_k^2$

kein einziges Mal an einer anderen Stelle als an der des Changepoints $k^* = 60$ angenommen. Bei den zugrunde liegenden Parametern des Modells liegen fast "optimale" Daten vor: sowohl ein deutlicher Unterschied in den Parametern b_n und b_n^* als auch ein Changepoint nahe der Mitte der Beobachtungsstellen.

Betrachten wir deshalb im Folgenden Simulation 3.3 unter der Alternative mit veränderten Werten und setzen zuerst $\mathbf{b}_n^* = 2.2$. Die Kleinste-Quadrate-Schätzer sind sicherlich normalverteilt und liefern in dem Bereich vor bzw. nach dem Changepoint erwartungstreue Schätzungen der Parameter bei Varianzen gemäß den Formeln des Lemmas 3.2. In dem Bereich nach Überschreiten des Changepoints ergibt eine genauere Diskussion der Formeln des Lemmas 3.2, dass die Differenzen der Erwartungswerte von den zu schätzenden Parametern geringer ausfallen als bei $b_n^* = 4$. Obwohl in diesem Bereich die Schätzer keine geeigneten Funktionen mehr sind, um den jeweiligen Parameter zu schätzen, bedeutet dies, dass die Abweichungen der geschätzten Werte der Kleinste-Quadrate-Schätzer zu den wahren Parametern im Mittel geringer ausfallen als in der vorigen Simulation. Da die Varianzschätzer $\hat{\sigma}_k^2$ gerade Summen von quadratischen Abweichungen sind, nehmen diese ebenfalls kleinere Werte an, falls die "falschen" Schätzungen der Kleinste-Quadrate-Schätzer trotzdem weniger von den wahren Parametern abweichen.



Diesen Sachverhalt verdeutlicht Abbildung 8, in der die geschätzten Erwartungswerte und Standardabweichungen der Varianzschätzer $\hat{\sigma}_k^2$ dargestellt sind. Wie die geführte Diskussion begründet, sind die angenommenen Werte der Varianzschätzer für $k \neq k^*$ deutlich niedriger als in Abbildung 6. Doch trotz dieser niedrigen Werte liegt das Minimum der geschätzten Mittelwerte genau an der Stelle des Changepoints. Da die geschätzten Standardabweichungen der Varianzschätzer $\hat{\sigma}_k^2$ ebenfalls im Vergleich mit Abbildung 7 geringer sind, ist zu erwarten, dass das Minimum der $\{\hat{\sigma}_k^2 : 2 \leq k \leq 98\}$ mit großer Wahrscheinlichkeit an der Stelle des Changepoints angenommen wird. Die vorliegende Simulation bestätigt dies, in der kein einziges Mal das Minimum an einer anderen Stelle vorlag.

Für den nächsten Fall mit nicht "optimalen" Parametern betrachten wir wieder Simulationen von 3.3 unter der Alternative, jedoch mit einem sehr frühen wie auch sehr späten Changepoint, $\mathbf{k}^* = \mathbf{10}$ und $\mathbf{k}^* = \mathbf{90}$.

Der prinzipielle Verlauf der Erwartungswerte und Varianzen der Kleinste-Quadrate-Schätzer bleibt erhalten. Doch trotz der gleichen Abstände der jeweiligen Changepoints zu den Rändern des Intervalls [0, 100], in dem die Beobachtungsstellen t_i liegen, verhalten sich die Schätzer unterschiedlich. Die Differenzen der Erwartungswerte von den zu schätzenden Parametern, die die Kleinste-Quadrate-Schätzer nach Überschreiten des frühen Changepoints $k^* = 10$ aufweisen, liegen deutlich niedriger als etwa unter der Alternative mit einem Changepoint nahe der Mitte des Intervalls [0, 100].

Dagegen sind diese Differenzen der Kleinste-Quadrate-Schätzer bei einem späten Changepoint $k^* = 90$ deutlich höher. Mit derselben Begründung wie in der letzten Simulation für $b_n^* = 2.2$ überträgt sich das auf die Erwartungswerte der Varianzschätzer $\hat{\sigma}_k^2$ für $k \neq k^*$: diese sind bei der Simulation des frühen Changepoints $k^* = 10$ deutlich niedriger als bei der des späten



Abbildung 9: Die Schätzer $\hat{\sigma}_k^2$ unter der Alternative mit $k^* = 10$

Changepoints $k^* = 90$.

Die Abbildungen 9 und 10 zeigen Erwartungswerte und Standardabweichungen der Schätzer $\hat{\sigma}_k^2$ unter den Alternativen $k^* = 10$ und $k^* = 90$. Die eingetragenen Werte entsprechen den Mittelwerten an den Stellen der Changepoints. Aufgrund der hohen Erwartungswerte in Abbildung 10 erscheinen hier die Graphen als eine Kurve. In beiden Simulationen liegt das Minimum der geschätzten Erwartungswerte an den Stellen der jeweiligen Changepoints und weist Werte dicht an dem zu schätzenden Wert $\sigma^2 = 1$ auf. Da in beiden Simulationen die geschätzten Abweichungen der $\hat{\sigma}_k^2$ wieder relativ gering zu den Erwartungswerten sind, liegt mit großer Wahrscheinlichkeit das Minimum der $\{\hat{\sigma}_k^2: 2 \le k \le 98\}$ an den Stellen der Changepoints. In den vorliegenden Simulationen war dies sogar ohne Ausnahmen der Fall.



3.5 Simulation mit normalverteilten Fehlern und nicht äquidistanten Beobachtungen

Im Folgenden betrachten wir das Verhalten der Varianzschätzer $\hat{\sigma}_k^2$ bei nicht äquidistanten Beobachtungspunkten und fassen dabei die beiden Fälle $\alpha = 0.5$ und $\alpha = 2$ in der folgenden Simulation zusammen:

Simulation 3.4

Seien $\{y_i : 1 \le i \le n\}$ Beobachtungen, die dem folgenden Modell genügen:

$$y_i := \begin{cases} \mathbf{1} + \mathbf{2}t_i + e_i & : & 1 \le i \le k^*, \\ \mathbf{1} + \mathbf{4}t_i + e_i & : & k^* + 1 \le i \le n \end{cases}$$

wobei e_i , $1 \leq i \leq n$, unabhängige und normalverteilte Zufallsvariablen mit $Ee_i = 0$ und $Var e_i = 1$ sind.

Anzahl der Beobachtungen: n = 100.Beobachtungspunkte: $\alpha \in \{0.5, 2\}.$ Changepoint: $k^* = 60.$

Offensichtlich ist für $\alpha < 1$ die Dichte der Beobachtungsstellen nahe der letzten Beobachtungsstelle $t_n = n^{\alpha}$ am größten, dagegen bei $\alpha > 1$ nahe der ersten Beobachtungsstelle $t_1 = 1$. Die durch die unterschiedlichen Potenzen entstehenden verschiedenen Intervalle, in denen die Beobachtungsstellen liegen, lassen sich durch geeignete Multiplikation der Stützstellen t_i auf jeden beliebigen Bereich übertragen, dabei ist nur eine Reskalierung der Parameter b_n und b_n^* erforderlich.

Wie in der vorigen Simulation mit normalverteilten Fehlern sind auch hier die Kleinste-Quadrate-Schätzer unter der Nullhypothese normalverteilt mit den Parametern des Lemmas 3.2.a). Festzuhalten ist, dass die Varianzen der Kleinste-Quadrate-Schätzer für $\alpha = 2$ kleinere Werte aufweisen als für $\alpha = 0.5$. Begründet liegt dies in der Abhängigkeit der Formeln für die Varianzen aus Lemma 3.2 von der Verteilung der Beobachtungsstellen, also von α . Eine weitere Diskussion der Kleinste-Quadrate-Schätzer erübrigt sich, da für die Varianzschätzer $\hat{\sigma}_{\mathbf{k}}^2$ unter der Nullhypothese wie in Simulation 3.3 gilt, dass

$$\hat{\sigma}_k^2 \stackrel{\mathcal{D}}{=} \frac{1}{n-4} \sigma^2 \chi_{n-4}^2, \quad 2 \le k \le n-2.$$

Das bedeutet, dass die Varianzschätzer $\hat{\sigma}_k^2$ unabhängig von der Lage der Stützstellen sind und somit wie in Simulation 3.3 verteilt sind. Die gewohnte Veranschaulichung würde deshalb wieder lediglich drei parallele Geraden zeigen (vgl. auch Abbildung 13).

Auch **unter der Alternative sind die Kleinste-Quadrate-Schätzer** normalverteilt, jedoch gemäß den Parametern der Aussage b) des Lemmas 3.2. Die verschiedenen Verteilungen der Stützstellen bewirken unter der Alternative deutliche Unterschiede der Differenzen der Erwartungswerte zu den zu schätzenden Parametern der Kleinste-Quadrate-Schätzer nach Überschreiten des Changepoints $k^* = 60$. Für $\alpha = 0.5$ sind diese Differenzen deutlich kleiner als bei $\alpha = 2$.

Dieses unterschiedliche Verhalten der Erwartungswerte überträgt sich wie in der vorigen Simulation von verschiedenen Changepoints auf das Verhalten der **Varianzschätzer** $\hat{\sigma}_{\mathbf{k}}^2$ unter



Abbildung 11: Die Schätzer $\hat{\sigma}_k^2$ bei $\alpha = 0.5$ unter der Alternative

der Alternative. Dies zeigen die Abbildungen 11 und 12, die die geschätzten Mittelwerte und Standardabweichungen der Varianzschätzer $\hat{\sigma}_k^2$ darstellen. Die in die Graphiken eingetragenen Werte beziffern wieder den Mittelwert an den Stellen der Changepoints. Da die Differenzen der Erwartungswerte der Schätzer \hat{a}_k und \check{a}_k zu dem Parameter $a_n = 1$ bei $\alpha = 2$ unter der Alternative extrem große Werte nach Überschreiten des Changepoints aufweisen, erreichen die Erwartungswerte der $\hat{\sigma}_k^2$ in Abbildung 12 ebenfalls sehr große Werte, was im vorigen Abschnitt bereits begründet wurde. Das ausgezeichnete Verhalten des Minimumschätzers $\tilde{\sigma}_n^2$ in den vorigen Simulationen liegt auch hier vor, da das Minimum von $\{\hat{\sigma}_k^2: 2 \leq k \leq 98\}$ in beiden Simulationen immer an der Stelle des Changepoints angenommen wurde.



3.6 Simulationen bei verschiedenen Stichprobenumfängen

In diesem Abschnitt wird versucht, einen Einblick in das Konvergenzverhalten der Varianzschätzer $\hat{\sigma}_k^2$ zu gewinnen. Dazu wird Simulation 3.3 wiederholt, jedoch mit einer Anzahl von n = 50 und n = 200 Beobachtungen; zum Vergleich werden noch teilweise die aus Simulation 3.3 gewonnenen Graphiken mit einer Anzahl von n = 100 herangezogen.

Simulation 3.5

Seien $\{y_i : 1 \leq i \leq n\}$ Beobachtungen, die dem folgenden Modell genügen:

$$y_i := \begin{cases} 1 + 2t_i + e_i &: 1 \le i \le k^*, \\ 1 + 4t_i + e_i &: k^* + 1 \le i \le n \end{cases}$$

 $\begin{array}{ll} wobei \ e_i, \ 1 \leq i \leq n, \ unabhängige \ und \ normalverteilte \ Zufallsvariablen \ mit \ {\rm E} \ e_i = 0 \ und \\ {\rm Var} \ e_i = 1 \ sind. \\ {\it Anzahl \ der \ Beobachtungen: \ n \in \{50, 100, 200\}.} \\ {\it Beobachtungspunkte: } \qquad \alpha = 1. \\ {\it Changepoint: } \qquad k^* = 60. \end{array}$

Für die Kleinste-Quadrate-Schätzer unter der Nullhypothese folgt auch hier, dass diese normalverteilt sind gemäß den Verteilungsparametern aus Lemma 3.2.a). Daraus lassen sich leicht etwa die Konvergenzraten der Varianzen ermitteln. Eine weitere Behandlung dieser Schätzer erübrigt sich, denn für die Varianzschätzer $\hat{\sigma}_k^2$ unter der Nullhypothese gilt, wie bereits in Simulation 3.3 begründet, dass

$$\hat{\sigma}_k^2 \stackrel{\mathcal{D}}{=} \frac{1}{n-4} \sigma^2 \chi_{n-4}^2, \quad 2 \le k \le n-2.$$
(3.5)

Damit ist auch deren Verhalten unter der Nullhypothese vollständig bekannt; zur Illustration sind in Abbildung 13 die geschätzten Erwartungswerte und Standardabweichungen aufgrund der Simulationen abgebildet. Die geschätzten Werte entsprechen den Verteilungsparametern aus (3.5), also $E\hat{\sigma}_k^2 = \sigma^2$ und $Var\hat{\sigma}_k^2 = 2\sigma^4/(n-4)$. Damit folgt, dass die Varianzen der $\hat{\sigma}_k^2$ mit der Rate 1/n bei wachsender Anzahl n von Beobachtungen abnehmen.

Interessant sind die unterschiedlichen Abhängigkeiten der Erwartungswerte und damit des Verhaltens der Kleinste-Quadrate-Schätzer unter der Alternative von der Anzahl nder Beobachtungen. Lemma 2.5 zeigt, dass der bei wachsender Stichprobengröße n nichtverschwindende Anteil von \hat{a}_k für $k^* < k = [\theta n]$ ein Wachstum der Rate n^{α} bei festem θ und konstantem δ_n erfährt, dagegen bleibt dieser Anteil bei \hat{b}_k für solch ein k konstant. Analoges gilt auch für die Schätzer \check{a}_k und \check{b}_k (vgl. Bemerkung nach Lemma 2.5). Diesen Sachverhalt gibt Abbildung 14 wieder, in der die geschätzten Mittelwerte und Standardabweichungen der Schätzer \check{a}_k bei den verschiedenen Stichprobengrößen abgebildet sind. Die Schätzer sind normalverteilt.

Zusammenfassend gilt, dass die mit großer Wahrscheinlichkeit im Betrag angenommenen Werte der Schätzer \hat{a}_k und \check{a}_k mit der Rate *n* zunehmen, diese dagegen bei den Schätzern \hat{b}_k und \check{b}_k (fast) konstant bleiben bei wachsender Anzahl von Beobachtungen.







Abbildung 14: Die Schätzer \check{a}_k unter der Alternative



Abbildung 15: Die Schätzer $\hat{\sigma}_k^2$ unter der Alternative

Wie in den letzten Simulationen bereits beobachtet und begründet, üben die angenommenen Werte der Kleinste-Quadrate-Schätzer einen großen Einfluss auf **die Varianzschätzer** $\hat{\sigma}_k^2$ **unter der Alternative** aus. So ist bei dem Wachstum der im Betrag angenommenen Werte der Schätzer für den Parameter a_n auch bei den Varianzschätzern $\hat{\sigma}_k^2$ für $k \neq k^*$ eine Steigerung der angenommenen Werte bei wachsender Größe der Stichprobe zu erwarten. Abbildung 15 belegt dieses, in der die geschätzten Erwartungswerte und Standardabweichungen der Schätzer $\hat{\sigma}_k^2$ abgebildet sind. Die Zahlen in den Graphiken geben den Wert des Mittelwertes an der Stelle des Changepoints mit der geschätzten Standardabweichung in Klammern an. Die geschätzten Erwartungswerte wachsen ungefähr proportional zu n^2 , da die Schätzer \hat{a}_k und \check{a}_k quadratisch in die Varianzschätzer $\hat{\sigma}_k^2$ eingehen und die Abweichungen der Schätzer \hat{b}_k und \check{b}_k von den wahren Parametern $b_n = 2$ und $b_n^* = 4$ ungefähr konstant bei zunehmender Größe der Stichprobe bleiben.

Der Minimumschätzer $\tilde{\sigma}_n^2$ zeigt sogar bei der geringen Beobachtungsanzahl n = 50 wieder ein ausgezeichnetes Verhalten und entspricht in den vorliegenden Simulationen jeweils immer dem Schätzer $\hat{\sigma}_{k^*}^2$ an der Stelle des Changepoints.

3.7 Simulation mit exponentialverteilten Fehlern

In diesem Abschnitt untersuchen wir, ob die Varianzschätzer $\hat{\sigma}_k^2$ ein andersartiges Verhalten aufweisen, falls den Fehlern eine nicht symmetrische Verteilung zugrunde liegt. Dazu wird eine Expontialverteilung mit Parameter $\lambda = 1$ gewählt, die um den Erwartungswert verschoben wird, um eine zentrierte Verteilung zu erhalten. Diese Simulation dient nur dazu, exemplarisch die Abhängigkeit der Schätzer von der Verteilung der Fehler e_i zu untersuchen. Deshalb gehen wir auch auf keine weiteren Verteilungseigenschaften der Schätzer ein.

Simulation 3.6

Seien $\{y_i : 1 \le i \le n\}$ Beobachtungen, die dem folgenden Modell genügen:

$$y_i := \begin{cases} \mathbf{1} + \mathbf{2}t_i + e_i & : \quad 1 \le i \le k^*, \\ \mathbf{1} + \mathbf{4}t_i + e_i & : \quad k^* + 1 \le i \le n. \end{cases}$$

wobei e_i , $1 \le i \le n$, unabhängig und nach einer verschobenen Exponentialverteilung verteilt sind mit $E e_i = 0$ und $Var e_i = 1$.

Die Erwartungswerte und Varianzen der Kleinste-Quadrate-Schätzer unter der Nullhypothese sind in Aussage a) des Lemmas 3.2 gegeben. Jedoch gilt hier nicht mehr, dass die Schätzer gemäß diesen Parametern normalverteilt sind. Auch unterliegen die Varianzschätzer $\hat{\sigma}_k^2$ unter der Nullhypothese keiner χ^2 -Verteilung mehr. Trotzdem zeigt Abbildung 16, dass die geschätzten Erwartungswerte und Standardabweichungen der $\hat{\sigma}_k^2$ für alle $2 \leq k \leq n-2$ sich homogen über das gesamte Spektrum der Beobachtungstellen verhalten. Für den konstanten Mittelwert ist natürlich die Erwartungstreue der Schätzer $\hat{\sigma}_k^2$ für $2 \leq k \leq 98$ verantwortlich. Abbildung 16 legt nahe, dass ebenfalls die Standardabweichungen unabhängig von k denselben Wert annehmen. Dieser Wert der geschätzten Standardabweichungen liegt deutlich höher als bei normalverteilten Fehlern.



Abbildung 16: Die Schätzer $\hat{\sigma}_k^2$ unter der Nullhypothese

Erwartungswerte und Varianzen der Kleinste-Quadrate-Schätzer unter der Alternative sind ebenfalls durch Lemma 3.2 gegeben, doch genauso wenig ist hier von einer Normalverteilung auszugehen. Die Gleichheit der Erwartungswerte und Varianzen der Kleinste-Quadrate-Schätzer wie bei normalverteilten Fehlern lässt ein vergleichbares Verhalten wie in Simulation 3.3 der Varianzschätzer $\hat{\sigma}_k^2$ unter der Alternative vermuten. Die Abbildungen 17 und 18 stellen die geschätzten Erwartungswerte und Standardabweichungen der Varianzschätzter $\hat{\sigma}_k^2$ wieder getrennt dar, die eingetragenen Ergebnisse geben die Werte an der Stelle des Changepoints $k^* = 60$ an. Ein Vergleich der Abbildungen 17 und 18 mit den Abbildungen 6 und 7 zeigt, dass die geschätzten Erwartungswerte und Standardabweichungen der Schätzer $\hat{\sigma}_k^2$ für $k \neq k^*$ ungefähr dieselben Werte aufweisen. Für $k = k^*$ entspricht der



Abbildung 17: Die Schätzer $\hat{\sigma}_k^2$ unter der Alternative



Abbildung 18: Die Abweichung der Schätzer $\hat{\sigma}_k^2$ unter der Alternative

Mittelwert wieder einer guten Näherung des zu schätzenden Wertes $\sigma^2 = 1$, jedoch liegt hier die Standardabweichung höher als bei normalverteilten Fehlern und entspricht dem Wert der Standardabweichung unter der Nullhypothese (vgl. den folgenden Abschnitt). Aufgrund der geringen Standardabweichungen und der hohen Erwartungswerte der $\hat{\sigma}_k^2$ für $k \neq k^*$ ist zu erwarten, dass das Minimum der $\{\hat{\sigma}_k^2 : 2 \leq k \leq 98\}$ mit großer Wahrscheinlichkeit an der Stelle des Changepoints liegt. In der Tat wird in der vorliegenden Simulation das Minimum immer an der Stelle des Changepoints angenommen.

3.8 Die Varianzschätzer im Vergleich

Die Betrachtungen der letzten Abschnitte zeigen, dass unter der Alternative die Schätzer $\hat{\sigma}_k^2$ für $k \neq k^*$ im Mittel hohe Werte annehmen. Doch gerade dieses Verhalten spricht für den Minimumschätzer $\tilde{\sigma}_n^2 = \min\{\hat{\sigma}_k^2 : 2 \leq k \leq n-2\}$. Je größer diese Werte bei geringen Standardabweichungen sind, mit umso höherer Wahrscheinlichkeit wird das Minimum von $\{\hat{\sigma}_k^2 : 2 \leq k \leq n-2\}$ an der Stelle des Changepoints k^* angenommen. Zu erwarten ist dann, dass der Minimumschätzer $\tilde{\sigma}_n^2$ ähnliche Eigenschaften wie der Schätzer $\hat{\sigma}_{k^*}^2$ für festes k^* aufweist.

Ähnlich lassen sich von den im Mittel hohen Werten der Funktion $\hat{\sigma}_{[n/2]}^2$ der Familie $\{\hat{\sigma}_k^2: 2 \leq k \leq n-2\}$ in den Simulationen unter der Alternative auf Werte des Alternativeschätzers $\bar{\sigma}_n^2 = \min\{\bar{\sigma}_1^2, \bar{\sigma}_2^2\}$ nahe des zu schätzenden Wertes schließen.

Unter der Nullhypothese sind bei normalverteilten Fehlern gemäß der Herleitung in Simulation 3.3 die Schätzer $\hat{\sigma}_k^2$ für $2 \le k \le n-2$ identisch $\sigma^2/(n-4)\chi_{n-4}^2$ verteilt. Bei exponentialverteilten Fehlern lassen die nahe beieinander liegenden Werte der geschätzen Erwartungswerte und Standardabweichungen der $\hat{\sigma}_k^2$ in Simulation 3.6 ebenfalls eine identische Verteilung vermuten. Es gilt jedoch sicherlich nicht, dass trotz einer identischen Verteilung aufgrund der Erwartungstreue von $\hat{\sigma}_k^2$ für $2 \le k \le n-2$ der Minimumschätzer $\tilde{\sigma}_n^2 = \min{\{\hat{\sigma}_k^2: 2 \le k \le n-2\}}$ ebenfalls erwartungstreue Schätzungen liefert. Vielmehr sind unter der Nullhypothese von dem Minimumschätzer $\tilde{\sigma}_n^2$ Schätzungen zu erwarten, die im Mittel den wahren Wert σ^2 unterschätzen. Ohne Kenntnis der Verteilung der Fehler e_i lässt sich keine geeignete Normierung durchführen, um dies zu vermeiden.

Diese Diskussion überträgt sich auf den normierten Alternativschätzer $\bar{\sigma}_n^2 = \min\{\bar{\sigma}_1^2, \bar{\sigma}_2^2\}$ bezüglich der Argumente $\bar{\sigma}_1^2$ und $\bar{\sigma}_2^2$ des Minimums. Hier gilt bei normalverteilten Fehlern unter der Nullhypothese

$$\bar{\sigma}_1^2 \stackrel{\mathcal{D}}{=} \frac{\sigma^2}{[n/2] - 2} \chi^2_{[n/2] - 2} \quad \text{und} \quad \bar{\sigma}_2^2 \stackrel{\mathcal{D}}{=} \frac{\sigma^2}{n - [n/2] - 2} \chi^2_{n - [n/2] - 2} . \tag{3.6}$$

In der folgenden Tabelle sind die in den Simulationen ermittelten Werte zusammengefasst, wobei für einen Schätzer \hat{x} gilt:

$$\begin{aligned} \hat{\mu}(\hat{x}) &:= \frac{1}{S} \sum_{i=1}^{S} \hat{x}_i, \\ \hat{S}(\hat{x}) &:= \sqrt{\frac{\sum_{i=1}^{S} \hat{x}_i^2 - \frac{1}{S} \left(\sum_{i=1}^{S} \hat{x}_i\right)^2}{(S-1)}}, \\ m_i &:= \min\left\{m : \hat{\sigma}_{m,i}^2 := \min_{2 \le l \le n-2} \left\{\hat{\sigma}_{l,i}^2\right\}\right\}, \\ m &:= \frac{1}{S} \sum_{i=1}^{S} m_i. \end{aligned}$$

Dabei entspricht \hat{x}_i bzw. $\hat{\sigma}_{k,i}^2$ dem Wert des Schätzers \hat{x} bzw. $\hat{\sigma}_k^2$ in der i-ten Simulation. Weichen die Simulationen von den zu Beginn der Abschnitte 3.4 bis 3.7 gemachten Definitionen ab, so werden die damit nicht übereinstimmenden Parameter in der Tabelle gesondert aufgeführt. Auf die Werte von m unter den Nullhypothesen wird verzichtet, da diesen hier ohne weitere Untersuchungen keine besondere Aussagekraft für die Varianzschätzer zukommt.

Simulation	$\hat{\mu}(\tilde{\sigma}_n^2)$	m	$\hat{S}(\tilde{\sigma}_n^2)$	$\hat{\mu}(\bar{\sigma}_n^2)$	$\hat{S}(\bar{\sigma}_n^2)$
Simulation 3.3					
Nullhypothese:	0.951		0.141	0.898	0.150
Alternative: $k^* = 60 b_n^* = 2.2$	1.001	60	0.145	0.999	0.198
$k^* = 60 \ b_n^* = 4$	1.000	60	0.141	0.999	0.203
$k^* = 10 \ b_n^* = 4$	1.011	10	0.148	1.005	0.214
$k^* = 90 \ b_n^* = 4$	0.998	90	0.144	0.997	0.207
Simulation 3.4					
Nullhypothese: $\alpha = 0.5$	0.946		0.136	0.899	0.153
$\alpha = 2$	0.939		0.136	0.889	0.156
Alternative: $\alpha = 0.5$	1.003	60	0.146	0.998	0.212
$\alpha = 2$	1.002	60	0.141	1.003	0.207
Simulation 3.5					
Nullhypothese: $n = 50$	0.894		0.195	0.874	0.219
n = 200	0.972		0.099	0.932	0.112
Alternative : $n = 50$	0.995	30	0.209	0.999	0.306
n = 200	0.998	120	0.101	1.001	0.141
Simulation 3.6					
Nullhypothese:	0.926		0.262	0.800	0.263
Alternative :	0.989	60	0.279	0.967	0.363

Unter der Nullhypothese liegen in allen Simulationen die Mittelwerte der Schätzer $\tilde{\sigma}_n^2$ und $\bar{\sigma}_n^2$ unter den zu schätzenden Werten.

Die Resultate der Simulationen 3.3 und 3.5 zeigen eine Annäherung des Mittelwertes unter der Nullhypothese an den zu schätzenden Wert bei wachsender Stichprobengröße n, jedoch ist die dabei vorliegende Rate sehr gering. Die geschätzten Standardabweichungen des Schätzers $\tilde{\sigma}_n^2$ sind in den Simulationen 3.3 und 3.5 unter der Nullhypothese etwas geringer als die Standardabweichung $(2\sigma^4/(n-4))^{1/2}$ der Argumente $\hat{\sigma}_k^2$ des Minimumsfunktionals.

Anders dagegen sind die Standardabweichungen des Schätzers $\bar{\sigma}_n^2$ bei normalverteilten Fehlern unter der Nullhypothese deutlich niedriger als die Abweichungen $(2\sigma^4/(\frac{1}{2}n-2))^{1/2}$ der Argumente $\bar{\sigma}_1^2$ und $\bar{\sigma}_2^2$ des Minimums gemäß (3.6). Man beachte hier, dass sich aufgrund der Unabhängigkeit von $\bar{\sigma}_1^2$ und $\bar{\sigma}_2^2$ die Verteilung von $\bar{\sigma}_n^2$ leicht angeben lässt.

Die Feststellung in Simulation 3.4, dass die Lage der Beobachtungsstellen bei normalverteilten Fehlern unter der Nullhypothese keinen Einfluss auf die Verteilung der Schätzer $\hat{\sigma}_k^2$ für $2 \leq k \leq n-2$ hat, überträgt sich auf die Schätzer $\tilde{\sigma}_n^2$ und $\bar{\sigma}_n^2$. Dementsprechend sind die geschätzten Mittelwerte und Standardabweichungen der Simulation 3.4 unter der Nullhypothese ähnlich zu denen der Simulation 3.3.

Das schlechtere Verhalten der Familie $\{\hat{\sigma}_k^2 : 2 \le k \le n-2\}$ bei exponentialverteilten Fehlern als bei normalverteilten Fehlern setzt sich bei den Schätzern $\tilde{\sigma}_n^2$ und $\bar{\sigma}_n^2$ fort. Sowohl die Differenzen der geschätzten Mittelwerte von dem wahren Wert als auch die Standardabweichungen nehmen unter der Nullhypothese in Simulation 3.6 deutlich höhere Werte als in Simulation 3.3 an. Der Alternativschätzer liegt deutlicher als der Minimumschätzer unter dem zu schätzenden Wert.

Unter der Alternative wird in allen Simulationen das Minimum von $\{\hat{\sigma}_k^2 : 2 \leq k \leq n-2\}$ ohne Ausnahme an der Stelle des Changepoints k^* angenommen. Dies hat zur Folge, dass die Schätzungen der Erwartungswerte und Standardabweichungen des Minimumschätzers $\tilde{\sigma}_n^2$ aufgrund der Daten der Simulationen Schätzungen der Verteilungsparameter des Schätzers $\hat{\sigma}_{k^*}^2$ für festes k^* sind.

Bei normalverteilten Fehlern gilt deshalb unter der Alternative $\mu(\tilde{\sigma}_n^2) \approx \sigma^2$ und $\hat{S}(\tilde{\sigma}_n^2) \approx (2\sigma^4/(n-4))^{1/2}$, da $\hat{\sigma}_{k^*}^2 \stackrel{\mathcal{D}}{=} 2\sigma^2/(n-4)\chi_{n-4}^2$. In Simulation 3.5 und 3.3 mit $k^* = 60$ und $b_n^* = 4$ nehmen demzufolge die Raten der Standardabweichungen proportional zu $1/\sqrt{n}$ ab und die Mittelwerte liegen dicht an den zu schätzenden Werten.

Da die Lage der Beobachtungsstellen bei normalverteilten Fehlern keinen Einfluss auf die Verteilung des Schätzers $\hat{\sigma}_{k^*}^2$ ausübt, folgt wie oben, dass die geschätzten Werte des Minimumschätzers $\tilde{\sigma}_n^2$ in Simulation 3.4 den Parametern der Verteilung von $\hat{\sigma}_{k^*}^2$ gleichen.

Die geführte Diskussion für den Minimumschätzer $\tilde{\sigma}_n^2$ überträgt sich analog auf den Alternativschätzer $\bar{\sigma}_n^2$. Unter der Alternative wird in den Simulationen jeweils das Minimum an dem Wert der Funktion $\bar{\sigma}_j^2$, $j \in \{1, 2\}$, angenommen, die nicht über den Changepoint k^* aufsummiert. Diese Funktion $\bar{\sigma}_j^2$ ist bei normalverteilten Fehlern gemäß (3.6) verteilt.

Bei normalverteilten Fehlern sind die Unterschiede der geschätzten Standardabweichungen von $\tilde{\sigma}_n^2$ und $\bar{\sigma}_n^2$ auf die oben genannten Verteilungen zurückzuführen. Da in $\bar{\sigma}_1^2$ bzw. $\bar{\sigma}_2^2$ nur halb so viele Beobachtungen wie in $\hat{\sigma}_{k^*}^2$ eingehen, unterscheiden sich die Standardabweichungen von $\tilde{\sigma}_n^2$ und $\bar{\sigma}_n^2$ unter der Alternative um den Faktor $\sqrt{2}$.

Bei exponentialverteilten Fehlern liefern die Schätzer $\tilde{\sigma}_n^2$ und $\bar{\sigma}_n^2$ unter der Alternative im Mittel gute Schätzungen. Die Standardabweichungen entsprechen wie bei normalverteilten Fehlern der geschätzten Standardabweichung unter der Nullhypothese der Schätzer $\hat{\sigma}_k^2$ oder

 $\bar{\sigma}_1^2$ und $\bar{\sigma}_2^2$. Wie oben unterscheiden sich in Simulation 3.6 die Standardabweichungen der

Schätzer $\tilde{\sigma}_n^2$ und $\bar{\sigma}_n^2$ unter der Alternative etwa um den Faktor $\sqrt{2}$. Als Abschluss der Betrachtungen der Schätzer $\tilde{\sigma}_n^2$ und $\bar{\sigma}_n^2$ lässt sich festhalten, dass unter der Nullhypothese der Schätzer $\tilde{\sigma}_n^2$ den wahren Wert weniger unterschätzt als der Alternativschätzer $\bar{\sigma}_n^2$. Unter der Alternative liefern beide Funktionen im Mittel gute Schätzungen des zu schätzenden Wertes und zeigen sich sehr robust unter den verschiedenen Alternativen. Die Standardabweichungen liegen für den Alternativschätzer etwa um den Faktor $\sqrt{2}$ höher als die des Minimumschätzers $\tilde{\sigma}_n^2$.

In Anwendungen, in denen eine schnelle Berechenbarkeit gefordert ist, überwiegt eventuell der deutlich geringere Berechnungsaufwand des Alternativschätzers gegenüber dem besseren Verhalten des Minimumschätzers.

Bezugnehmend auf das ursprüngliche Ziel, die unbekannte Varianz in der Statistik L_n aus (2.7) durch eine Schätzung zu ersetzen, sei bemerkt, dass nach Chu und White [ChWh92] der Test L_n eher zu einem konservativen Verhalten neigt. Dagegen verursacht das Unterschätzen der Schätzer $\tilde{\sigma}_n^2$ und $\bar{\sigma}_n^2$ unter der Nullhypothese bei Ersetzen der unbekannten Varianz in der Statistik L_n ein progressives Verhalten.

A Anhang

In diesem Abschnitt werden die Konvergenzraten der deterministischen Summen, die im Kapitel "Lineares Modell" benutzt werden, sowie noch einige Aussagen über die dort vorkommenden Determinanten zusammengestellt.

Satz A.1

Set $\alpha > 0$. Dann gilt:

$$\frac{1}{n^{1+\alpha}}\sum_{i=1}^{n}i^{\alpha} \to \frac{1}{\alpha+1}, \quad n \to \infty.$$

Beweis:

$$\begin{aligned} \frac{1}{\alpha+1} &= \frac{1}{n^{1+\alpha}} \int_{1}^{n+1} (x-1)^{\alpha} dx &\leq \frac{1}{n^{1+\alpha}} \sum_{i=1}^{n} i^{\alpha} \\ &\leq \frac{1}{n^{1+\alpha}} \int_{1}^{n+1} x^{\alpha} dx = \frac{1}{n^{1+\alpha}} \left(\frac{(n+1)^{1+\alpha} - 1}{\alpha+1} \right). \quad \Box \end{aligned}$$

Satz A.2 Sei $\alpha > 0$. Dann gilt:

$$\begin{array}{l} a) & \max_{1 \le k \le n} \left| \frac{1}{k^{1+\alpha}} \sum_{i=1}^{k} i^{\alpha} \right| = O(1), \ n \to \infty. \\ b) & \max_{1 \le k \le n} \left| \frac{1}{k^{1+2\alpha}} \sum_{i=1}^{k} i^{2\alpha} \right| = O(1), \ n \to \infty. \\ c) & \max_{0 \le k \le n-1} \left| \frac{1}{n^{1+\alpha}} \sum_{i=k+1}^{n} i^{\alpha} \right| = O(1), \ n \to \infty. \\ d) & \max_{0 \le k \le n-1} \left| \frac{1}{n^{1+2\alpha}} \sum_{i=k+1}^{n} i^{2\alpha} \right| = O(1), \ n \to \infty. \end{array}$$

Beweis:

- a) und b) Der Beweis folgt jeweils direkt aus dem Beweis zu Satz A.1.
- c) Mit dem Mittelwertsatz ergibt sich

$$\begin{split} 0 &\leq \frac{1}{n^{1+\alpha}} \sum_{i=k+1}^{n} i^{\alpha} &\leq \frac{1}{n^{1+\alpha}} \int_{k+1}^{n+1} x^{\alpha} dx \\ &= \frac{1}{n^{1+\alpha}} \zeta^{\alpha} (n-k), \quad \zeta \in [k+1, n+1], \\ &= O(1) \quad \text{gleichmäßig für } k \in \{0, \dots, n-2\}. \end{split}$$

Sei im folgenden wie im Kapitel "Lineares Modell" für festes $\alpha > 0$

$$d_k := d_{k,n} := k \sum_{i=1}^k i^2 - \left(\sum_{i=1}^k i\right)^2, \quad 2 \le k \le n,$$

und
$$d_k^* := d_{k,n}^* := (n-k) \sum_{i=k+1}^n i^2 - \left(\sum_{i=k+1}^n i\right)^2, \quad 0 \le k \le n-2$$

Lemma A.3

Es gilt für $\alpha > 0$:

a) $d_k \leq d_{k+1} \quad \forall k \in \{2, \dots, n-1\}.$ b) $d_k^* \leq d_{k-1}^* \quad \forall k \in \{1, \dots, n-2\}.$

Beweis:

a) Sei
$$2 \le k \le n-1$$
 und $P_k = \begin{pmatrix} \sum_{i=1}^k i^2 & \sum_{i=1}^k i \\ \sum_{i=1}^k i & k \end{pmatrix}$, so dass Det $P_k = d_k$ gilt.

Ein direkter Nachweis zeigt die positive Definitheit von P_k für $k \ge 2$. Mittels einfacher Rechnung ergibt sich

$$v^T P_k v \le v^T P_{k+1} v, \quad \forall v \in \mathbb{R}^2.$$

Damit folgt dann¹

$$v^T P_k^{-1} v \ge v^T P_{k+1}^{-1} v, \quad \forall v \in \mathbb{R}^2,$$

so dass ein Einsetzen von v = (1,0) die Ungleichung

$$\frac{k}{d_k} \ge \frac{k+1}{d_{k+1}}$$

ergibt, woraus die Behauptung resultiert.

b) Erfolgt wie Teil a).

Satz A.4

Set $\alpha > 0$. Dann gilt:

a)
$$\max_{2 \le k \le n} \left| \frac{k^{2+2\alpha}}{k \sum_{i=1}^{k} i^{2\alpha} - \left(\sum_{i=1}^{k} i^{\alpha}\right)^2} \right| = O(1).$$

¹siehe hierzu Theorem 20.1 und 33.1 in [Pra94].

I		1
٩		

b)
$$f\ddot{u}r f(n) := n - 2 - \left[\frac{n}{\log\log n}\right]$$
:

$$\max_{\substack{0 \le k \le f(n) \\ 0 \le k \le n(1-\epsilon)}} \left| \frac{n^{2+2\alpha}}{(n-k)\sum_{i=k+1}^{n} i^{2\alpha} - \left(\sum_{i=k+1}^{n} i^{\alpha}\right)^{2}} \right| = O\left((\log\log n)^{4}\right).$$
c) $f\ddot{u}r \ alle \ \epsilon \in (0,1)$:

$$\max_{\substack{0 \le k \le n(1-\epsilon) \\ 0 \le k \le n(1-\epsilon)}} \left| \frac{n^{2+2\alpha}}{(n-k)\sum_{i=k+1}^{n} i^{2\alpha} - \left(\sum_{i=k+1}^{n} i^{\alpha}\right)^{2}} \right| = O(1).$$

Beweis:

a) Mit Satz A.1 ergibt sich

$$\frac{1}{k^{2+2\alpha}} \left(k \sum_{i=1}^{k} i^{2\alpha} - \left(\sum_{i=1}^{k} i^{\alpha} \right)^2 \right) \to \frac{\alpha^2}{(2\alpha+1)(\alpha+1)^2} \neq 0, \quad k \to \infty.$$

b) Die bereits in Satz A.2 benutzte Abschätzung ergibt mit einer Anwendung des Mittelwertsatzes

$$\sum_{i=k+1}^{n} i^{\alpha} \le \int_{k+1}^{n+1} x^{\alpha} dx = \frac{(n+1)^{\alpha+1} - (k+1)^{\alpha+1}}{\alpha+1} = \frac{n^{\alpha+1} - k^{\alpha+1}}{\alpha+1} + O\left(n^{\alpha}\right),$$

gleichmäßig für $k \in \{2, ..., n\}$. Damit gilt mit $t := \frac{k}{n}$

$$(n-k)\sum_{i=k+1}^{n} i^{2\alpha} - \left(\sum_{i=k+1}^{n} i^{\alpha}\right)^{2}$$

$$\geq \frac{(n-k)(n^{2\alpha+1}-k^{2\alpha+1})}{2\alpha+1} - \left(\frac{n^{\alpha+1}-k^{\alpha+1}}{\alpha+1} + O(n^{\alpha})\right)^{2}$$

$$= n^{2\alpha+2} \left(\frac{(1-t)(1-t^{2\alpha+1})}{2\alpha+1} - \left(\frac{1-t^{\alpha+1}}{\alpha+1} + O\left(\frac{1}{n}\right)\right)^{2}\right)$$

$$= n^{2\alpha+2} \left(\frac{1-t-t^{2\alpha+1}+t^{2\alpha+2}}{2\alpha+1} - \frac{1-2t^{\alpha+1}+t^{2\alpha+2}}{(\alpha+1)^{2}} + O\left(\frac{1}{n}\right)\frac{1-t^{\alpha+1}}{\alpha+1} + O\left(\frac{1}{n^{2}}\right)\right).$$
(1.1)

Setzen wir t := 1 - x, erhält man für die Potenzen von t für $x \to 0$ mittels der Taylorentwicklung die Darstellungen

$$\begin{split} t^{\alpha+1} &= 1 - (\alpha+1)x + \frac{(\alpha+1)\alpha}{2}x^2 - \frac{(\alpha+1)\alpha(\alpha-1)}{6}x^3 \\ &+ \frac{(\alpha+1)\alpha(\alpha-1)(\alpha-2)}{24}x^4 + O(x^5), \\ t^{2\alpha+1} &= 1 - (2\alpha+1)x + \frac{(2\alpha+1)2\alpha}{2}x^2 - \frac{(2\alpha+1)2\alpha(2\alpha-1)}{6}x^3 \\ &+ \frac{(2\alpha+1)2\alpha(2\alpha-1)(2\alpha-2)}{24}x^4 + O(x^5), \end{split}$$

$$t^{2\alpha+2} = 1 - (2\alpha+2)x + \frac{(2\alpha+2)(2\alpha+1)}{2}x^2 - \frac{(2\alpha+2)(2\alpha+1)(2\alpha)}{6}x^3 + \frac{(2\alpha+2)(2\alpha+1)(2\alpha)(2\alpha-1)}{24}x^4 + O(x^5).$$

Ersetzt man die Potenzen von t der ersten zwei Summanden in (1.1) durch die eben gewonnenen Darstellungen und sortiert nach den Potenzen von x, erhält man

$$\frac{1-t-t^{2\alpha+1}+t^{2\alpha+2}}{2\alpha+1} - \frac{1-2t^{\alpha+1}+t^{2\alpha+2}}{(\alpha+1)^2} = \frac{\alpha^2}{12}x^4 + O(x^5)$$
$$= \frac{\alpha^2}{12}(1+o(1))x^4, \quad x \to 0, \qquad (1.2)$$

da die Koeffizienten von x^{μ} für $\mu \in \{0, 1, 2, 3\}$ verschwinden.

Wegen des Lemmas A.3 verhält sich die Folge $\{d_k^*: 0 \le k \le f(n)\}$ monoton, so dass es ausreicht, d_k^* für k = f(n) zu betrachten, also

$$x = \frac{1}{n} \left(2 + \left[\frac{n}{\log \log n} \right] \right) \ge \frac{1}{\log \log n}.$$

Somit folgt mit (1.1) und (1.2) die Behauptung aus

$$\begin{aligned} \frac{1}{n^{2+2\alpha}} \left((n-k) \sum_{i=k+1}^{n} i^{2\alpha} - \left(\sum_{i=k+1}^{n} i^{\alpha} \right)^2 \right) &\geq \frac{\alpha^2}{12} \left(1 + o(1) \right) x^4 + O\left(\frac{x}{n} \right) + O\left(\frac{1}{n^2} \right) \\ &= x^4 \left(\frac{\alpha^2}{12} \left(1 + o(1) \right) + o(1) \right) \\ &\geq \frac{1}{\left(\log \log n \right)^4} \left(\frac{\alpha^2}{12} \left(1 + o(1) \right) \right). \end{aligned}$$

c) Wie in [HuSt99] werden folgende Summen durch Integrale gleichmäßig in $t \in (0, 1)$ approximiert:

$$\begin{split} \frac{1}{n} \sum_{i=0}^{[nt]-1} \left(1 - \frac{i}{n}\right)^{\alpha} &- \int_{0}^{t} (1 - x)^{\alpha} \, dx \\ &= \sum_{i=0}^{[nt]-1} \int_{\frac{i}{n}}^{\frac{i+1}{n}} \left(\left(1 - \frac{i}{n}\right)^{\alpha} - (1 - x)^{\alpha} \right) \, dx - \int_{\frac{[nt]}{n}}^{t} (1 - x)^{\alpha} \, dx \\ &\leq \sum_{i=0}^{[nt]-1} \alpha \int_{\frac{i}{n}}^{\frac{i+1}{n}} \left(\int_{\frac{i}{n}}^{x} (1 - y)^{\alpha - 1} \, dy \right) \, dx - (1 - t)^{\alpha} \left(t - \frac{[nt]}{t} \right) \\ &= \alpha \sum_{i=0}^{[nt]-1} \int_{\frac{i}{n}}^{\frac{i+1}{n}} (1 - y)^{\alpha - 1} \left(\frac{i+1}{n} - y \right) \, dy + (1 - t)^{\alpha} \, O\left(\frac{1}{n}\right) \\ &\leq \frac{1}{n} \left(1 - \left(1 - \frac{[nt]}{n} \right)^{\alpha} \right) + (1 - t)^{\alpha} \, O\left(\frac{1}{n}\right) \\ &= O\left(\frac{1}{n}\right), \end{split}$$

und ebenso

$$\frac{1}{n} \sum_{i=0}^{[nt]-1} \left(1 - \frac{i}{n}\right)^{2\alpha} - \int_0^t (1 - x)^{2\alpha} \, dx = O\left(\frac{1}{n}\right).$$

Mittels dieser beiden Approximationen erhält man schließlich mit l = n - k

$$(n-k)\sum_{i=k+1}^{n} i^{2\alpha} - \left(\sum_{i=k+1}^{n} i^{\alpha}\right)^{2}$$

= $l\sum_{i=0}^{l-1} (n-i)^{2\alpha} - \left(\sum_{i=0}^{l-1} (n-i)^{\alpha}\right)^{2}$
= $n^{2+2\alpha} \left(\frac{l}{n^{2}}\sum_{i=0}^{l-1} \left(1 - \frac{i}{n}\right)^{2\alpha} - \left(\frac{1}{n}\sum_{i=0}^{l-1} \left(1 - \frac{i}{n}\right)^{\alpha}\right)^{2}\right)$
= $n^{2+2\alpha} \left(t\int_{0}^{t} (1-x)^{2\alpha} dx - \left(\int_{0}^{t} (1-x)^{\alpha} dx\right)^{2} + O\left(\frac{1}{n}\right)\right).$

Da der Integrand an keiner Stelle konstant ist, folgt aus der Jensen-Ungleichung

$$g(t) := t \int_0^t (1-x)^{2\alpha} \, dx - \left(\int_0^t (1-x)^{\alpha} \, dx \right)^2 > 0 \quad \forall t \ge \epsilon,$$

was die Behauptung aufgrund der gleichmäßigen Approximation impliziert.

Danksagung

Nicht weil hier der übliche Platz ist, sondern um den Gegebenheiten gerecht zu werden: Besonderen Dank möchte ich meinem Betreuer Herrn Prof. Dr. J. Steinebach für seine Unterstützung und zahlreiche Anregungen während der Erstellung dieser Arbeit aussprechen.

Literatur

- [An91] Andrews, D. (1991). Heteroskedasticity and autocorrelation consistent covariance matrix estimation. Econometrica 59, 817-858.
- [Bill79] Billingsley, Patrick. (1979). Probability and Measure. Wiley, New York.
- [ChWh92] Chu, C.-S.J., White, H. (1992). A direct test for changing trend. J. Business Economic Statist. 10, 289 - 299.
- [CsHo93] Csörgő, M., Horváth, L. (1993). Weighted Approximations in Probability and Statistics. Wiley, Chichester.
- [CsHo97] Csörgő, M., Horváth, L. (1997). Limit Theorems in Change-Point Analysis. Wiley, Chichester.
- [CsHoSt87] Csörgő, M., Horváth, L., Steinebach, J. (1987). Invariance principles for renewal processes. Ann. Probab. 15, 1441-1460.
- [CsRe75a] Csörgő, M. und Révész, P. (1975). A new method to prove Strassen type laws of invariance principle I. Z. Wahrsch. verw. Geb. 31, 255-260.
- [CsRe75b] Csörgő, M. und Révész, P. (1975). A new method to prove Strassen type laws of invariance principle II. Z. Wahrsch. verw. Geb. 31, 261-269.
- [CsRe81] Csörgő, M. und Révész, P. (1981). Strong Approximations in Probability and Statistics. Academic Press, New York.
- [Fe95] Ferger, D. (1995). Change-Point Schätzer basierend auf gewichteten Empirischen Prozessen mit Anwendungen auf das Zweistichprobenproblem in allgemeinen meßbaren Räumen. Habilitationsschrift. Justus-Liebig-Universität Gießen.
- [Ha89] Hawkins, D.L. (1989). A u-i approach to retrospective testing for shifting parameters in a linear model. Commun. Stat., Theory Methods 18, 3117-3134.
- [Hin71] Hinkley, D.V. (1971). Inference in two-phase regression. J. Amer. Statist. Assoc. 66, 736 - 743.
- [Ho95] Horváth, L. (1995). Detecting changes in linear regressions. Statistics 26, 189-208.
- [HoSh95] Horváth, L., Shao, Q.-M. (1995). Limit theorems for the union-intersection test. J. Statist. Plann. Inference 44, 133-148.
- [HoSt98] Horváth, L. und Steinebach, J. (1998). Testing for changes in the mean or variance of a stochastic process under weak invariance. J. Statist. Plann. Inference (erscheint demnächst).
- [Hu96] Hušková, M. (1996). Estimation of a change in linear models. Statistics and Probability Letters 26, 13-24.
- [HuSt99] Hušková, M., Steinebach, J. (1999). Limit theorems for a class of tests of gradual changes. J. Statist. Plann. Inference (erscheint demnächst).

- [KoMaTu75] Komlós, J., Major, P., Tusnády, G. (1975). An approximation of partial sums of independent r.v.'s and the sample df. I. Z. Wahrsch. verw. Geb. 32, 111-131.
- [KoMaTu76] Komlós, J., Major, P., Tusnády, G. (1976). An approximation of partial sums of independent r.v.'s and the sample df. II. Z. Wahrsch. verw. Geb. 34, 33-58.
- [Kühn99] Kühn, C. (1999). Statistische Analyse von Changeparametern auf der Basis von Invarianzprinzipien. Diplomarbeit an der Philipps-Universität Marburg.
- [NeWe87] Newey, W., West, K. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, Econometrica 55, 703-708.
- [Pra94] Prasolov, V.V. (1994). Problems and Theorems in Linear Algebra. American Mathematical Society, Providence.
- [Ri87] Ripley, B.D. (1987). Stochastic Simulation. Wiley, New York.
- [Sea71] Searle, S.R. (1971). Linear Models. Wiley, New York.
- [Seb77] Seber, G.A.F. (1977). Linear Regression Analysis. Wiley, New York.
- [ShWe86] Shorack, G.R., Wellner, J.A. (1986). Empirical Processes with Applications to Statistics. Wiley, New York.

Hiermit bestätige ich, dass ich die vorliegende Arbeit selbstständig verfasst und keine anderen Hilfsmittel als die angegebenen benutzt habe.