

$p$ -Werte und alternative Schranken von  
CUSUM-Tests

Achim Zeileis  
September 2000

Diplomarbeit

Betreuer: Prof. Dr. Walter Krämer

Fachbereich Statistik  
Universität Dortmund

# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>1</b>
<b>2</b>	<b>Methoden</b>	<b>4</b>
2.1	Strukturbruchtests im Linearen Regressionsmodell . . . . .	4
2.1.1	Standard-CUSUM-Test . . . . .	6
2.1.2	OLS-basierter CUSUM-Test . . . . .	10
2.2	$p$ -Werte . . . . .	13
2.2.1	Definition von $p$ -Werten . . . . .	13
2.2.2	Erwartete $p$ -Werte . . . . .	14
2.3	Software . . . . .	15
<b>3</b>	<b><math>p</math>-Werte für CUSUM-Tests</b>	<b>16</b>
3.1	$p$ -Wert des Standard-CUSUM-Tests . . . . .	16
3.2	$p$ -Wert des OLS-basierten CUSUM-Tests . . . . .	20
<b>4</b>	<b>Alternative Schranken für CUSUM-Tests</b>	<b>24</b>
4.1	Alternative Schranken für den Standard-CUSUM-Test . . . . .	25
4.2	Alternative Schranken für den OLS-basierten CUSUM-Test . . . . .	32

<i>INHALTSVERZEICHNIS</i>	2
<b>5 Simulation erwarteter <math>p</math>-Werte</b>	<b>38</b>
<b>6 Zusammenfassung</b>	<b>45</b>
Literaturverzeichnis . . . . .	48

# Abbildungsverzeichnis

2.1	Standard-CUSUM-Test . . . . .	10
2.2	OLS-basierter CUSUM-Test . . . . .	13
3.1	$p$ -Wert des Standard-CUSUM-Tests . . . . .	19
3.2	$p$ -Wert des OLS-basierten CUSUM-Tests . . . . .	22
4.1	Ergebnisse von Methode 1 und 2 . . . . .	29
4.2	Schranken des Standard-CUSUM-Tests . . . . .	30
4.3	Standard-CUSUM-Test mit alternativen Schranken . . . . .	31
4.4	Schranken des OLS-basierten CUSUM-Tests . . . . .	35
4.5	OLS-basierter CUSUM-Test mit linearen und alternativen Schranken . . . . .	36

# Tabellenverzeichnis

4.1	Kritische Werte des alternativen Standard-CUSUM-Tests . . .	28
4.2	Kritische Werte des alternativen OLS-basierten CUSUM-Tests	34
5.1	Simulation erwarteter $p$ -Werte des Standard-CUSUM-Tests . .	40
5.2	Simulation erwarteter $p$ -Werte des alternativen Standard-CUSUM- Tests . . . . .	41
5.3	Simulation erwarteter $p$ -Werte des OLS-basierten CUSUM-Tests	43
5.4	Simulation erwarteter $p$ -Werte des alternativen OLS-basierten CUSUM-Tests . . . . .	44

# Kapitel 1

## Einleitung

Im linearen Regressionsmodell wird die Abhängigkeit einer Variablen  $y$  von einem oder mehreren Regressoren  $x_1, \dots, x_k$  untersucht. Dabei wird nicht nur davon ausgegangen, daß der wahre Zusammenhang linear ist, sondern auch, daß dieser Zusammenhang sich nicht ändert. Das heißt vor allem, daß der Vektor der Regressionskoeffizienten  $\beta$  im Modell  $y = X\beta + u$  konstant ist. Um die für das lineare Regressionsmodell entwickelten Verfahren und Tests anwenden zu können, kann es deshalb sinnvoll sein, diese Annahme durch einen Test auf Strukturbruch in den Regressionskoeffizienten zu überprüfen. Es kann aber auch inhaltlich von Interesse sein, einen Wechsel der Beziehung zwischen den beobachteten Größen nachzuweisen; gerade in der Ökonometrie ist dies eine häufige Fragestellung. Untersucht man beispielsweise die Abhängigkeit des Konsums vom Einkommen, so kann es von Interesse sein, herauszufinden ob sich dieser Zusammenhang aufgrund eines politischen Ereignisses - wie einem Regierungswechsel oder einer Änderung der Steuergesetzgebung - geändert hat. Geht man davon aus, daß der wahre Zu-

sammenhang linear ist, so kann man dies in das lineare Regressionsmodell einbetten, indem man in  $y$  die zeitlich geordneten Beobachtungen des Konsums und im Regressor  $x$  die zugehörigen Beobachtungen des Einkommens zusammenfaßt. Würde dann ein Test einen Strukturbruch in den Regressionskoeffizienten aufdecken, so hieße dies, daß sich der Anteil des Einkommens, der für Konsum aufgewendet wird, über die Zeit geändert hat.

Auch in der Bio- oder Technometrie gibt es viele Situationen, in denen Strukturbrüche untersucht werden; alle diese Situationen haben aber gemeinsam, daß die Beobachtungen einer natürlichen Ordnung unterliegen: beispielsweise nach Größe, Alter oder, wie hier betrachtet, nach Zeit.

Einer der ersten Tests für diese Fragestellung wurde von [Chow \(1960\)](#) vorgeschlagen. Der Chow-Test geht von einem einzigen bekannten Bruchpunkt aus, paßt an die Daten vor und nach dem Bruchpunkt jeweils eine Kleinst-Quadrat-Gerade an und vergleicht mit Hilfe einer  $F$ -Teststatistik dieses Modell mit dem gewöhnlichen ohne Strukturbruch. Der Nachteil dabei ist, daß der Bruchpunkt bekannt sein muß, deshalb gibt es Modifikationen dieses Tests, die das Supremum ([Andrews, 1993](#)) oder den Durchschnitt ([Hansen, 1992](#)) der  $F$ -Statistiken für ein ganzes Intervall von möglichen Bruchpunkten betrachten.

Einen anderen Ansatz wählten [Brown, Durbin und Evans \(1975\)](#), die den CUSUM-Test als Möglichkeit zur Überprüfung der Konstanz von Regressionsbeziehungen über die Zeit vorschlugen. Dessen Teststatistik beinhaltet die kumulative Summe der rekursiven 1-Schritt-Prognosefehler. Als alternative Version dazu wurde von [Ploberger und Krämer \(1992\)](#) ein CUSUM-Test vorgeschlagen, der nicht auf diesen rekursiven sondern auf den herkömmlichen

OLS-Residuen basiert. Weitere Verfahren finden sich bei [Krämer und Sonnberger \(1986\)](#), die in ihrem Buch „The Linear Regression Model under Test“ unter anderem einen Überblick über verschiedene Strukturbruchtests geben. In dieser Diplomarbeit werden die Eigenschaften, d.h. insbesondere die Ablehnwahrscheinlichkeiten, des Standard-CUSUM-Tests sowie des OLS-basierten CUSUM-Tests untersucht. Nach der Definition der beiden Tests sowie von (erwarteten)  $p$ -Werten in Kapitel 2 wird eine (sehr gute) Approximation an die asymptotische  $p$ -Wert-Funktion der beiden Tests hergeleitet. In Kapitel 4 werden alternative Schranken als Modifikation der beiden Tests vorgeschlagen und schließlich in Kapitel 5 mit Hilfe von Simulationen erwarteter  $p$ -Werte verglichen. Dabei wird deutlich, daß der OLS-basierte CUSUM-Test durch die alternativen Schranken sowohl frühe wie auch späte Strukturbrüche besser aufdecken kann, während diese Schwäche beim Standard-CUSUM-Test durch die alternativen Schranken nicht behoben werden kann. In Kapitel 6 werden die wichtigsten Ergebnisse noch einmal zusammengefaßt.



# Kapitel 2

## Methoden

### 2.1 Strukturbruchtests im Linearen Regressionsmodell

Die in dieser Arbeit betrachteten Strukturbruchtests basieren auf einem Regressionsmodell der folgenden Form:

$$y_t = x_t^\top \beta_t + u_t \quad \text{für } t = 1, \dots, n, \quad (2.1)$$

wobei:

- $y_t$  die abhängige Variable der Beobachtung zum Zeitpunkt  $t = 1, \dots, n$ ,
- $x_t = (1, x_{t2}, \dots, x_{tk})^\top$  der  $k \times 1$ -Regressorvektor der  $t$ -ten Beobachtung ( $t = 1, \dots, n$ ),
- $\beta_t$  der Vektor der unbekanntenen Regressionskoeffizienten ( $t = 1, \dots, n$ ), der nicht konstant über die Zeit sein muß,

- $u_t$  der Fehlerterm der  $t$ -ten Beobachtung ( $t = 1, \dots, n$ ) ist.

Dabei wird angenommen, daß die Fehlerterme  $u_t$  unabhängig identisch verteilt sind mit Erwartungswert 0 und Varianz  $\sigma^2$ . Die Regressoren  $x_t$  sind nichtstochastisch mit  $\|x_t\| = O(1)$ , und es existiert eine endliche reguläre Matrix  $Q$ , so daß für  $n \rightarrow \infty$  gilt:

$$\frac{1}{n} \sum_{t=1}^n x_t x_t^\top \longrightarrow Q.$$

Mit dieser Annahme wird ein Trend in den Regressoren ausgeschlossen. In ökonometrischen Anwendungen kann es auch von Interesse sein, verzögerte Beobachtungen  $y_{t-j}$  ( $j \geq 1$ ) als Regressoren in das Modell mit einzuschließen. Für den CUSUM-Test, der im nächsten Abschnitt vorgestellt wird, betrachten [Krämer, Ploberger und Alt \(1988\)](#), im folgenden KPA, deshalb stationäre dynamische lineare Regressionsmodelle, die aber zu denselben asymptotischen Eigenschaften, d.h. insbesondere denselben Ablehnwahrscheinlichkeiten, beim Standard-CUSUM-Test führen wie bei nichtstochastischen Regressoren, weshalb hier nur letztere betrachtet werden.

Mit Hilfe eines Strukturbruchtests soll nun getestet werden, ob der Vektor der Regressionskoeffizienten konstant über die Zeit ist ( $H_0$ ) oder nicht. Die Nullhypothese lautet also:

$$H_0 : \beta_1 = \dots = \beta_n = \beta.$$

Gemäß [Brown, Durbin und Evans \(1975\)](#), im folgenden BDE, wird die Alternative zunächst nicht genauer spezifiziert, d.h. es wird kein genaues Muster einer möglichen Variation der Koeffizienten festgelegt. Einige spezielle Typen von Strukturbrüchen werden in späteren Kapiteln noch diskutiert werden;

insbesondere wenn sich der Koeffizientenvektor nur an einem bestimmten Bruchpunkt ändert. Ein Strukturbruch in den Varianzen der Fehlergrößen wird durch die Annahme gleicher Varianzen ausgeschlossen.

Unter der Nullhypothese kann das Regressionsmodell auch in der bekannten Matrixnotation geschrieben werden:

$$y = X\beta + u, \quad (2.2)$$

wobei:

- $y = (y_1, \dots, y_n)^\top$  den Vektor der abhängigen Variablen,
- $X = \begin{pmatrix} 1 & x_{12} & \cdots & x_{1k} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n2} & \cdots & x_{nk} \end{pmatrix}$  die  $(n \times k)$ -Regressormatrix,
- $u = (u_1, \dots, u_n)^\top$  den Vektor der Fehlerterme bezeichnet.

Dabei wird zusätzlich noch vorausgesetzt, daß  $X$  vollen Spaltenrang hat, damit sichergestellt wird, daß keine linearen Abhängigkeiten zwischen den Regressoren vorliegen. Auf Grundlage dieses Modells werden nun der Standard-CUSUM-Test sowie der OLS-basierte CUSUM-Test definiert. Die Teststatistiken basieren auf der kumulativen Summe (CUSUM = cumulative sum) der rekursiven Residuen bzw. der OLS-Residuen.

### 2.1.1 Standard-CUSUM-Test

Die Idee des Standard-CUSUM-Tests ist es, die Tatsache auszunutzen, daß man jede Beobachtung aus den vergangenen Beobachtungen nur gut prognostizieren kann, wenn kein Strukturbruch vorliegt. Genauer gesagt basiert der

Test auf den rekursiven Residuen  $\tilde{u}_t$ , die für  $t = k+1, \dots, n$  wie folgt definiert sind:

$$\tilde{u}_t = \frac{y_t - x_t^\top \beta^{(t-1)}}{\sqrt{1 + x_t^\top \left( X^{(t-1)\top} X^{(t-1)} \right)^{-1} x_t}}. \quad (2.3)$$

Im Zähler steht dabei die Differenz zwischen der  $t$ -ten Beobachtung und der KQ-Prognose dieses Wertes aus den ersten  $(t-1)$  Werten (das Superskript in Klammern gibt an, daß die entsprechende Größe nur von den Werten abhängt, deren Index nicht größer ist). Die rekursiven Residuen sind also standardisierte 1-Schritt-Prognosefehler und haben deshalb unter  $H_0$  immer den Erwartungswert 0. Wenn aber die Regressionskoeffizienten nur bis zum Zeitpunkt  $t = t_0$  konstant sind und sich danach ändern, so werden die rekursiven Residuen auch nur bis zum Punkt  $t_0$  einen Erwartungswert von 0 haben und von da an im allgemeinen einen davon abweichenden. Im Nenner steht auf der rechten Seite von (2.3) eine Standardisierung, die dafür sorgt, daß die  $\tilde{u}_t$  unter der Nullhypothese die Varianz  $\sigma^2$  haben.

Wenn nun also der Pfad der CUSUM-Größe

$$W^{(t)} = \frac{1}{\tilde{\sigma}} \sum_{r=k+1}^t \tilde{u}_r \quad (2.4)$$

zu weit von seinem Mittelwert  $E[W^{(t)}]=0$  abweicht, so spricht dies gegen die Hypothese, daß kein Strukturbruch vorliegt. Der Skalenparameter wird dabei durch

$$\tilde{\sigma} = \sqrt{\frac{1}{n-k-1} \sum_{t=k+1}^n (\tilde{u}_t - \bar{\tilde{u}})^2}$$

geschätzt, wobei das Mittel der rekursiven Residuen  $\bar{\tilde{u}}$  einbezogen werden sollte, da es selbst unter der Hypothese nicht gleich 0 sein muß.

Der Standard-CUSUM-Test verwirft die Nullhypothese, wenn  $W^{(t)}$  zu weit

von seinem Mittel 0 abweicht, d.h. wenn er die Schranke  $c(t)$  aus Gleichung (2.5) für ein  $t \in \{k+1, \dots, n\}$  überschreitet bzw.  $-c(t)$  unterschreitet.

$$c(t) = \lambda\sqrt{n-k} + 2\lambda\frac{t-k}{n-k} \quad (2.5)$$

Der Parameter  $\lambda$  entscheidet dabei über das Niveau  $\alpha$  des Tests. Um dies bestimmen zu können, muß die Verteilung von  $W^{(t)}$  ermittelt werden. BDE tun dies eher heuristisch: unter der zusätzlichen Annahme, daß die Fehlerterme normalverteilt sind, und wenn man  $\tilde{\sigma}$  durch  $\sigma$  ersetzt, gilt nämlich:

$$\begin{aligned} \text{E} [W^{(t)}] &= 0, \\ \text{VAR} [W^{(t)}] &= t - k, \\ \text{COV} [W^{(t)}, W^{(s)}] &= \min(t, s) - k. \end{aligned} \quad (2.6)$$

Ersetzt man dann  $W^{(t)}$  durch eine Standard-Brownsche Bewegung  $B(t)$  mit Startpunkt  $k$  und Kovarianzfunktion (2.6) (gemäß Billingsley (1999)), so könnte man  $\lambda$  durch die Gleichung

$$\text{P} (|B(t)| \geq c(t) \text{ für ein } k+1 \leq t \leq n) = \alpha \quad (2.7)$$

bestimmen. Der zugehörige Test hätte dann (ungefähr) das Niveau  $\alpha$ . BDE betrachten nur das Zweifache der Wahrscheinlichkeit, daß  $B(t)$  die Gerade  $c(t)$  überschreitet, da man die Wahrscheinlichkeit, daß die Brownsche Bewegung, bzw. der CUSUM-Pfad, beide Geraden schneidet, vernachlässigen kann (jedenfalls zu üblichen Niveaus von  $\alpha \leq 0.1$ ).

Daß diese anschauliche Motivation auch gerechtfertigt ist, zeigen KPA. Dafür sei die CUSUM-Größe noch einmal leicht verändert und standardisiert auf das Intervall  $[0,1]$  angegeben:

$$W_n(t) = \frac{1}{\tilde{\sigma}\sqrt{n-k}} \sum_{i=k+1}^{\lfloor k+t(n-k) \rfloor} \tilde{u}_i \quad (0 \leq t \leq 1). \quad (2.8)$$

Dabei ist zu beachten, daß sich die Bedeutung von  $t$  ändert:  $t$  nimmt nun nicht mehr Werte von 1 bis  $n$  an, sondern vielmehr Werte aus dem Intervall  $[0,1]$ . Um aber die Interpretation von  $t$  als „Zeit“ zu wahren, wird  $t$  als Bezeichnung beibehalten. Die Schranke (2.5) hat dann die Form  $c(t) = \lambda + 2\lambda t$ .

Dann gilt gerade, daß  $W_n(t)$  für  $n \rightarrow \infty$  in Verteilung gegen eine Standard-Brownsche Bewegung konvergiert:

$$W_n(t) \xrightarrow{d} B(t). \quad (2.9)$$

Dies zeigt, daß die Bestimmung der kritischen Werte aus Gleichung (2.7) durch die Asymptotik gerechtfertigt ist. Ist also (2.7) für ein bestimmtes Wertepaar  $\lambda$  und  $\alpha$  erfüllt, so liefert folgende Testprozedur asymptotisch einen Niveau- $\alpha$ -Test: Lehne  $H_0$  ab, falls

$$S = \sup_{0 \leq t \leq 1} \left| \frac{W_n(t)}{1 + 2t} \right| \quad (2.10)$$

größer ist als der kritische Wert  $\lambda$ . Wie man dieses Ergebnis verfeinern sowie  $p$ -Werte für den Standard-CUSUM-Test bestimmen kann, wird in Kapitel 3 gezeigt.

Um zu veranschaulichen wie der Standard-CUSUM-Test arbeitet, zeigt Abbildung 2.1 den CUSUM-Pfad eines simulierten Strukturbruchs sowie die zugehörigen Schranken zum Niveau  $\alpha = 0.01$ . Die Vorgehensweise, die bei der Simulation verwendet wurde, ist in Kapitel 5 noch genauer beschrieben. Die simulierten Daten entstammen einem linearen Regressionsmodell mit einem Strukturbruch nach der Hälfte der 1000 Beobachtungen. In der Grafik ist deutlich zu sehen, daß der CUSUM-Pfad auch etwa nach der Hälfte der Beobachtungen beginnt, sich von dem Mittelwert 0 fortzubewegen. Aber erst

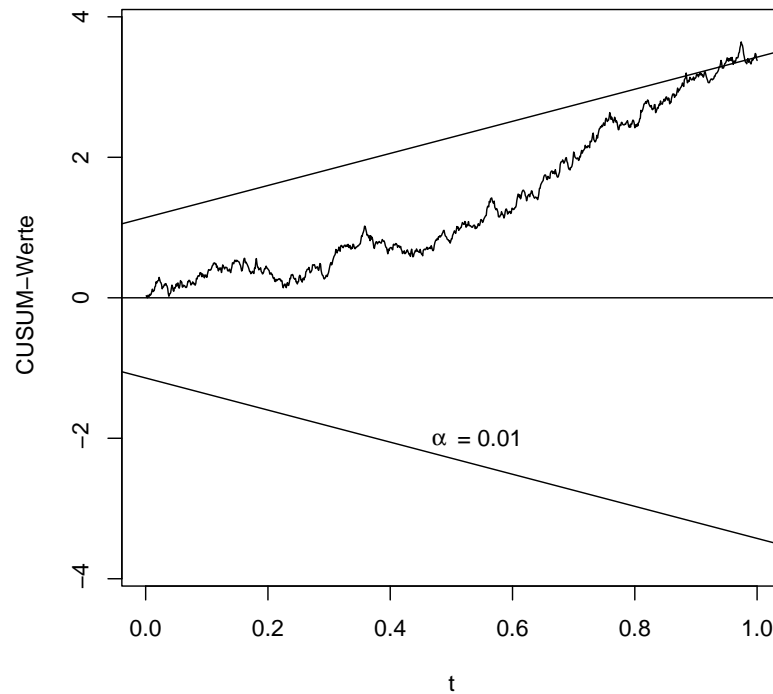


Abbildung 2.1: Standard-CUSUM-Test

gegen Ende überschreitet der Pfad schließlich die Grenze des kritischen Bereiches, so daß die Nullhypothese „Es liegt kein Strukturbruch vor“ zum Niveau 0.01 verworfen werden kann.

### 2.1.2 OLS-basierter CUSUM-Test

Die Grundidee beim OLS-basierten CUSUM-Test ist der des Standard-CUSUM-Tests sehr ähnlich. Der Unterschied liegt darin, daß beim OLS-basierten

CUSUM-Test anstatt der rekursiven Residuen einfach die OLS-Residuen verwendet werden. Diese haben genau wie die rekursiven Residuen unter  $H_0$  einen Erwartungswert von 0, jedoch ist bei ihnen auch immer die Summe 0, unabhängig davon wie stark der vorliegende Strukturbruch ist. Während der Standard CUSUM-Pfad bis zum Bruchpunkt  $t_0$  nahe 0 sein sollte und sich erst von da an von seinem Mittelwert entfernt, wird der OLS-basierte CUSUM-Pfad bis etwa zum Punkt  $t_0$  sich vom Mittel 0 entfernen und dann zur 0 zurückkehren.

Die OLS-basierte CUSUM-Größe  $W_n^0(t)$  ist für  $t \in [0, 1]$  folgendermaßen definiert:

$$W_n^0(t) = \frac{1}{\hat{\sigma}\sqrt{n}} \sum_{i=1}^{\lfloor nt \rfloor} \hat{u}_i, \quad (2.11)$$

wobei über die OLS-Residuen

$$\hat{u}_t = y_t - x_t^\top (X^\top X)^{-1} X^\top y \quad (t = 1, \dots, n),$$

summiert wird und  $\hat{\sigma}$  die Schätzung der Standardabweichung der Fehlergrößen im klassischen Regressionsmodell ist:

$$\hat{\sigma} = \sqrt{\frac{1}{n-1} \sum_{t=1}^n \hat{u}_t^2}.$$

Die Testvorschrift lautet dann: lehne  $H_0$  ab, falls der OLS-CUSUM-Pfad die Schranke  $\lambda$  für irgendein  $t \in [0, 1]$  über- bzw.  $-\lambda$  unterschreitet. Unter  $H_0$  gilt dabei nach [Ploberger und Krämer \(1992\)](#), im folgenden PK, daß  $W_n^0(t)$  für  $n \rightarrow \infty$  in Verteilung gegen eine Standard Brownsche Brücke konvergiert:

$$W_n^0(t) \xrightarrow{d} B^0(t).$$



Dabei ist  $B^0(t)$  eine Standard-Brownsche Brücke (siehe auch Billingsley (1999)) auf dem Intervall  $[0,1]$  mit

$$\begin{aligned} \mathbb{E} [B^0(t)] &= 0, \\ \text{VAR} [B^0(t)] &= t(1-t), \\ \text{COV} [B^0(t), B^0(s)] &= \min(t, s) - ts. \end{aligned} \quad (2.12)$$

Die kritischen Werte  $\lambda$  können also mit denselben Argumenten wie im vorangegangenen Abschnitt aus der Gleichung

$$\mathbb{P} \left( |B^0(t)| \geq \lambda \quad \text{für ein } 0 \leq t \leq 1 \right) = \alpha \quad (2.13)$$

ermittelt werden. Ist Gleichung (2.13) für ein Wertepaar  $\lambda$  und  $\alpha$  erfüllt, so liefert wiederum folgende Testvorschrift einen asymptotischen Test zum Niveau  $\alpha$ : Lehne  $H_0$  ab, falls

$$S^0 = \sup_{0 \leq t \leq 1} |W_n^0(t)| \quad (2.14)$$

einen Wert größer als  $\lambda$  annimmt.

Auch hier soll veranschaulicht werden, wie der OLS-basierte CUSUM-Test arbeitet, und deshalb sind in Abbildung 2.2 der OLS-CUSUM-Pfad sowie die Schranken zum Niveau  $\alpha = 0.01$  abgetragen. Die zugrundeliegenden simulierten Daten sind dieselben wie im Beispiel zum Standard-CUSUM-Test. Auch hier kann die Nullhypothese „Es liegt kein Strukturbruch vor“ zum Niveau 0.01 verworfen werden. Auffällig ist jedoch, daß die Grenze des Ablehnungsbereich, anders als beim Standard-CUSUM-Test, genau beim Bruchpunkt  $t = 0.5$  überschritten wird. Diese Eigenschaft des OLS-basierten CUSUM-Tests ist entscheidend für die alternativen Schranken, die in Kapitel 4 vorgeschlagen werden.

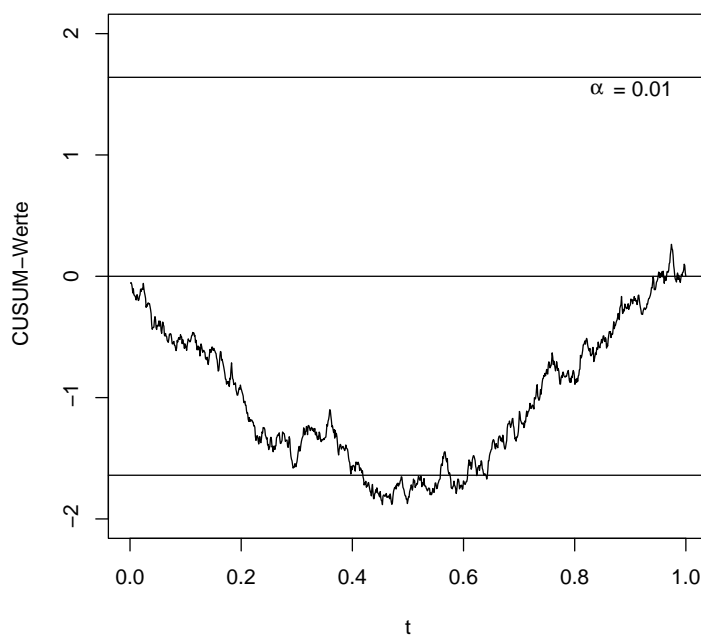


Abbildung 2.2: OLS-basierter CUSUM-Test

## 2.2 $p$ -Werte

### 2.2.1 Definition von $p$ -Werten

Ein  $p$ -Wert gibt bei einem Testergebnis an, wie wahrscheinlich es unter der Nullhypothese  $H_0$  ist, einen im Vergleich zur berechneten Teststatistik genauso extremen oder noch extremeren Wert zu beobachten. Das heißt, der  $p$ -Wert ist das Niveau, zu dem der Test die Nullhypothese gerade noch verwerfen würde. Ein kleiner  $p$ -Wert spricht gegen  $H_0$  (nach [Sachs \(1997\)](#)).

Ist die Teststatistik  $T$  unter  $H_0$  gemäß der stetigen Verteilung  $F_0$  verteilt und

wird die Nullhypothese zum Niveau  $\alpha$  verworfen, wenn  $T$  größer wird als  $c_\alpha$  (mit  $F_0(c_\alpha) = 1 - \alpha$ ), dann läßt sich der  $p$ -Wert folgendermaßen angeben: wird die Teststatistik  $T = t$  beobachtet, so berechnet sich der  $p$ -Wert gerade als

$$1 - F_0(t) = P_{H_0}(T \geq t). \quad (2.15)$$

### 2.2.2 Erwartete $p$ -Werte

Da der  $p$ -Wert von der Zufallsvariablen  $T$  abhängt, kann er selbst auch als Zufallsvariable aufgefaßt werden (gemäß [Sackrowitz und Samuel-Cahn \(1999\)](#)):

$$X = 1 - F_0(T).$$

Es ist bekannt, daß  $F_0(T)$  unter der Nullhypothese gleichverteilt ist auf dem Intervall  $[0, 1]$ ; deshalb ist es auch  $X$ . Damit ist klar, daß der erwartete  $p$ -Wert unter der Nullhypothese gerade 0.5 ist.

Unter einer speziellen Alternative  $\theta$ , wenn  $T$  gerade gemäß  $F_\theta$  verteilt ist, lautet der  $p$ -Wert per Definition:

$$P(T_0 \geq t | T = t), \quad (2.16)$$

wobei  $T_0$  der Nullverteilung folgt. Diese Darstellung erleichtert es, den erwarteten  $p$ -Wert, kurz *EPV* (expected  $p$  value), unter einer Alternative  $\theta$  anzugeben. Man erhält ihn nämlich, indem man in (2.16) einfach die Erwartung über  $T$  bildet:

$$EPV(\theta) = P(T_0 \geq T). \quad (2.17)$$

Hieraus ist auch noch einmal leicht zu sehen, daß für jede beliebige stetige Verteilung  $F_0$  der erwartete  $p$ -Wert genau 0.5 ist. Unter der Alternative  $\theta$

jedoch sollte die Verteilung  $F_\theta$  stochastisch größer sein als  $F_0$ , da die Teststatistik  $T$  unter der Alternative ja gerade größere Werte annehmen soll. Somit ist der erwartete  $p$ -Wert eine alternative Möglichkeit, die Güte eines Tests zu beschreiben, da er gerade  $1 - \text{erwartete Macht}$  (über alle möglichen Niveaus) unter der Alternative  $\theta$  beträgt.

Bei Güte-Simulationen haben erwartete  $p$ -Werte deshalb zwei Vorteile: zum einen sind sie unabhängig vom Niveau  $\alpha$  und zum anderen benötigt man keine kritischen Werte, um  $EPV(\theta)$  durch Monte-Carlo-Studien zu simulieren. Darüberhinaus können erwartete  $p$ -Werte eine Intuition darüber geben, welche Alternative wohl vorliegen oder besonders welche wohl nicht vorliegen könnte.

## 2.3 Software

Alle Berechnungen und Simulationen wurden mit R, einer unter GNU General Public Licence stehenden freien Implementation von S, durchgeführt (siehe z.B. [R Development Core Team \(2000\)](#)). Insbesondere wurden mit Hilfe der Funktion `rnorm` normalverteilte Zufallszahlen zur Simulation von Brownschen Bewegungen und Brownschen Brücken erzeugt. Diese wiederum wurden dann zur Simulation kritischer Werte und erwarteter  $p$ -Werte mittels Monte-Carlo-Methoden genutzt.

# Kapitel 3

## $p$ -Werte für CUSUM-Tests

Während in der Regel nur drei mögliche kritische Werte für den Standard-CUSUM-Test (und auch den OLS-basierten CUSUM-Test) zu den üblichen Niveaus 10%, 5% und 1% angegeben werden, soll in diesem Kapitel nun eine (sehr gute) Approximation an die asymptotische  $p$ -Wert-Funktion beider Tests hergeleitet werden. Das hat den Vorteil, da der Anwender nicht vorher eines der drei Niveaus des Tests festlegen muss und anschließend nur ein Ergebnis „ $H_0$  verwerfen“ oder „ $H_0$  nicht verwerfen“ erhält, sondern vielmehr das Niveau selbst wählen kann und darüberhinaus noch eine Vorstellung von der Art des Strukturbruchs bekommen kann. Speziell bei der Implementation in einem Statistik-Software-Paket ist dies von großem Vorteil.

### 3.1 $p$ -Wert des Standard-CUSUM-Tests

Wie in Kapitel 2 bereits angedeutet wurde, sind  $p$ -Werte der beiden untersuchten CUSUM-Tests eng verknüpft mit den Überschreitungswahrschein-

lichkeiten von Brownschen Bewegungen für Geraden.

Nach (2.15) kann der  $p$ -Wert eines Tests als eine Funktion der möglichen Beobachtungen der zugehörigen Teststatistik aufgefaßt werden. D.h. für die Teststatistik  $S$  des Standard-CUSUM-Tests aus Gleichung (2.10) ist der  $p$ -Wert folgendermaßen festgelegt:

$$P_{H_0}(S \geq s) = P_{H_0}(|W_n(t)| \geq s + 2st \quad \text{für ein } 0 \leq t \leq 1). \quad (3.1)$$

Nach KPA konvergiert die Wahrscheinlichkeit (3.1) für  $n \rightarrow \infty$  gegen die entsprechende Überschreitungswahrscheinlichkeit der Brownschen Bewegung, da ja die CUSUM-Größe in Verteilung gegen eine Brownsche Bewegung konvergiert (nach (2.9)). Der asymptotische  $p$ -Wert  $p(s)$  lautet also:

$$p(s) = P(|B(t)| \geq s + 2st \quad \text{für ein } 0 \leq t \leq 1). \quad (3.2)$$

Dies ist wiederum nichts anderes als die linke Seite von Gleichung (2.7) normiert auf das Intervall  $[0,1]$  und mit dem Parameter  $\lambda = s$  in der Schranke  $c(t)$  aus Gleichung (2.5). Mit anderen Worten ist der  $p$ -Wert nichts anderes als die Antwort auf die Frage, welches Niveau der Test mit dem kritischen Wert  $s$  hat. Um diese Frage zu beantworten, haben bereits BDE die Ergebnisse von Durbin (1971) genutzt, der Formeln für die Überschreitungswahrscheinlichkeiten von linearen Schranken für Brownsche Bewegungen angibt. Nach Lemma 3 bei Durbin gilt:

$$\begin{aligned} P(B(t) \geq at + b \quad \text{für ein } 0 \leq t \leq 1) = \\ 1 - \Phi(a + b) + \exp(-2ab) \Phi(a - b), \end{aligned} \quad (3.3)$$

wobei  $\Phi(\cdot)$  die Verteilungsfunktion der Standardnormalverteilung ist. Vernachlässigt man dann die Wahrscheinlichkeit, daß der Pfad einer Brownschen

Bewegung sowohl  $c(t)$  überschreiten wie auch  $-c(t)$  unterschreiten kann, so läßt sich eine erste Formel für die Wahrscheinlichkeit (3.2) angeben:

$$\begin{aligned} P(|B(t)| \geq s + 2st \quad \text{für ein } 0 \leq t \leq 1) = \\ 2(1 - \Phi(3s) + \exp(-4s^2) \Phi(s)). \end{aligned} \quad (3.4)$$

Jedoch ist diese Approximation nur für hinreichend große Werte von  $s$  zulässig; für kleine Werte von  $s$  kann (3.4) Werte größer als 1 annehmen (an der Stelle 0 sogar den Wert 2), was natürlich keinen Sinn macht. Deshalb kann sie vielmehr aufgefaßt werden als obere Schranke für  $p(s)$ :

$$p_O(s) = \min\{1, 2(1 - \Phi(3s) + \exp(-4s^2) \Phi(s))\}. \quad (3.5)$$

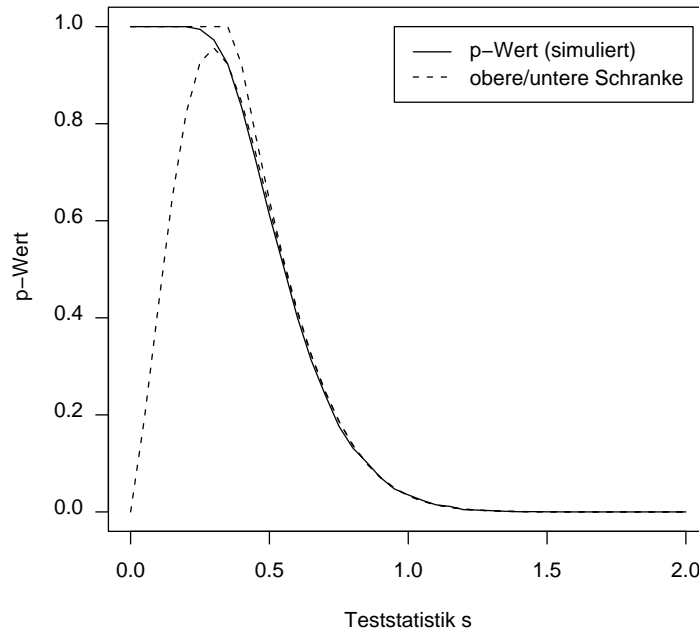
Um  $p(s)$  exakt angeben zu können, muß also noch die Wahrscheinlichkeit, daß  $B(t)$  beide Schranken schneidet, abgezogen werden. Doch auch diese Wahrscheinlichkeit kann nach Lemma 7 bei Durbin nur nach oben abgeschätzt werden:

$$\begin{aligned} P(B(t_1) \geq s + 2st_1 \wedge B(t_2) \leq -s - 2st_2 \quad \text{für jeweils ein } 0 \leq t_1, t_2 \leq 1) \leq \\ 2(\exp(-4s^2)(1 - \Phi(5s)) + \exp(-16s^2)(1 - \Phi(s))). \end{aligned} \quad (3.6)$$

Zieht man also die rechte Seite von (3.6) von der oberen Schranke (3.4) ab, so erhält man folglich eine untere Schranke für den gesuchten  $p$ -Wert  $p(s)$ :

$$\begin{aligned} p_U(s) = 2(1 - \Phi(3s) + \exp(-4s^2) (\Phi(s) + \Phi(5s) - 1) \\ - \exp(-16s^2) (1 - \Phi(s))). \end{aligned} \quad (3.7)$$

Diese Abschätzung hat den Nachteil, daß sie an der Stelle 0 den Wert 0 annimmt, während  $p(0)$  natürlich gerade 1 beträgt. In Abbildung 3.1 ist aber

Abbildung 3.1:  $p$ -Wert des Standard-CUSUM-Tests

deutlich zu sehen, daß beide Schranken sehr gute Approximationen für die Wahrscheinlichkeit (3.2) sind. Es ist ebenfalls leicht abzulesen, daß die untere Schranke  $p_U(s)$  näher bei den simulierten  $p$ -Werten liegt; jedenfalls bis zu ihrem Maximum, das etwa bei  $p_U(0.3) = 0.956$  liegt. Es können also  $p$ -Werte kleiner als 0.96 sehr gut approximiert werden; größere  $p$ -Werte könnte man auch durch Interpolation nähern, jedoch sind diese ohnehin wenig interessant. Für die Praxis relevant sind in aller Regel kleine  $p$ -Werte, und diese können mit beiden Schranken sehr gut approximiert werden. Die Differenz der beiden Schranken an der Stelle 0.85, dem kritischen Wert zum Niveau 0.1, beträgt



lediglich  $5 \cdot 10^{-6}$ .

Da beide Schranken für ein Statistik-Software-Paket gleichermaßen einfach zu berechnen sind, ist die untere Schranke wohl am ehesten zur Implementation zu empfehlen, da sie die größere Genauigkeit bietet. Für Werte kleiner als 0.3 kann dann linear interpoliert werden, so daß sich folgende Funktion ergibt:

$$\hat{p}(s) = \begin{cases} p_U(s) & s \geq 0.3 \\ 1 - 0.1465s & s < 0.3 \end{cases}. \quad (3.8)$$

### 3.2 *p*-Wert des OLS-basierten CUSUM-Tests

Da die Teststatistiken des OLS-basierten und des Standard-CUSUM-Tests, wie ihre asymptotischen Verteilungen, eng miteinander verwandt sind, verläuft auch die Argumentation bei der Herleitung der asymptotischen *p*-Werte bzw. der *p*-Wert-Funktion  $p^0(s)$  völlig analog. Der exakte *p*-Wert für die Teststatistik  $S^0$  aus (2.14) des OLS-basierten CUSUM-Tests ist dann folgendermaßen festgelegt:

$$P_{H_0}(S^0 \geq s) = P_{H_0}(|W_n^0(t)| \geq s \text{ für ein } 0 \leq t \leq 1). \quad (3.9)$$

Da die exakte Wahrscheinlichkeit (3.9) wiederum schwer zu evaluieren ist, wird hier ebenfalls die Asymptotik herangezogen. PK zeigen, daß sie für  $n \rightarrow \infty$  gegen die entsprechende Überschreitungswahrscheinlichkeit der Brownschen Brücke konvergiert:

$$\begin{aligned} p^0(s) &= P(|B^0(t)| \geq s \text{ für ein } 0 \leq t \leq 1) \\ &= P(|B(t)| \geq s \text{ für ein } 0 \leq t \leq 1 \mid B(1) = 0). \end{aligned} \quad (3.10)$$

Dies ist genau die linke Seite in Gleichung (2.13), die die Wahrscheinlichkeit angibt, daß der Pfad einer Brownschen Brücke die Schranke  $s$ , eine Parallele

zur x-Achse, überschreitet bzw. die Schranke  $-s$  unterschreitet. Auch hier ist der  $p$ -Wert  $p^0(s)$  die Antwort auf die Frage, zu welchem Niveau der Test bei realisierter Statistik  $S^0 = s$  die Nullhypothese gerade nicht verwirft.

Betrachtet man nun zunächst die Wahrscheinlichkeit, daß der Pfad einer Brownschen Brücke eine Gerade überschreitet (Lemma 4 bei Durbin)

$$P(B^0(t) \geq at + b \text{ für ein } 0 \leq t \leq 1) = \exp(-2b(a + b)) \quad (3.11)$$

so läßt sich - mit denselben Argumenten wie im vorigen Abschnitt - eine obere Schranke für die gesuchte  $p$ -Wert-Funktion  $p^0(s)$  angeben. Dabei ist, wie bereits erwähnt, die Gerade lediglich eine Parallele zur x-Achse, die somit eine Gerade mit Parametern  $a = 0$  und  $b = s$  ist.

$$p_O^0(s) = \min\{1, 2 \exp(-2s^2)\}. \quad (3.12)$$

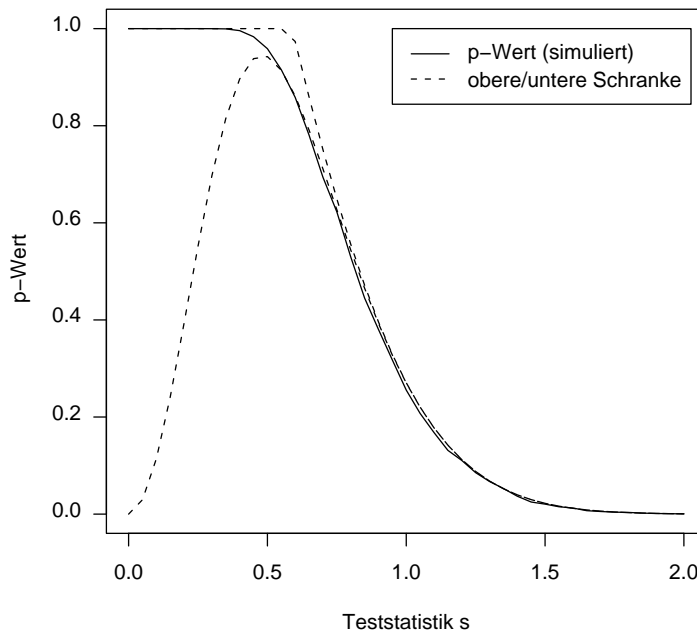
Um auch hier eine untere Schranke zu finden, muß wiederum die Wahrscheinlichkeit, daß der Pfad einer Brownschen Brücke sowohl  $s$  über- wie auch  $-s$  unterschreitet, abgeschätzt werden. Dafür gilt nach Lemma 6 bei Durbin folgende Ungleichung:

$$P(B^0(t_1) \geq s \wedge B^0(t_2) \leq -s \text{ für jeweils ein } 0 \leq t_1, t_2 \leq 1) \leq 2 \exp(-8s^2). \quad (3.13)$$

Hieraus läßt sich analog zur Vorgehensweise im letzten Abschnitt direkt eine untere Schranke für  $p^0(s)$  ableiten:

$$p_U^0(s) = 2(\exp(-2s^2) - \exp(-8s^2)). \quad (3.14)$$

Wie in Abbildung 3.2 zu sehen, haben auch diese Schranken sehr ähnliche Eigenschaften wie die des Standard-CUSUM-Tests. Auch hier muß die obere

Abbildung 3.2:  $p$ -Wert des OLS-basierten CUSUM-Tests

Schranke beim Wert 1 „abgeschnitten“ werden, da sie sonst gegen 2 gehen würde; die untere Schranke hingegen approximiert die simulierten  $p$ -Werte für große  $s$  zwar besser, für kleine  $s$  jedoch „stürzt sie ab“. Die untere Schranke nimmt das Maximum an der Stelle  $\sqrt{-\ln(0.25)/6} = 0.48$  mit dem Wert 0.94 an. Größere  $p$ -Werte könnten ebenfalls wieder durch Interpolation genähert werden, sind ohnehin aber von geringem Interesse. Die für die Praxis relevanten Werte werden von beiden Schranken ähnlich gut approximiert: an der Stelle 1.22, dem kritischen Wert zum Niveau 0.1, beträgt die Differenz  $1.3 \cdot 10^{-5}$ .

Da auch diese beiden Funktionen für ein Statistik-Software-Paket gleichermaßen einfach zu berechnen sind, ist  $p_U^0(s)$  wohl eher zur Implementation zu empfehlen, da sie die größere Genauigkeit bietet. Für Werte kleiner als 0.48 kann dann linear interpoliert werden, so daß sich folgende Funktion ergibt:

$$\hat{p}^0(s) = \begin{cases} p_U^0(s) & s \geq 0.48 \\ 1 - 0.1147s & s < 0.48 \end{cases} . \quad (3.15)$$

# Kapitel 4

## Alternative Schranken für CUSUM-Tests

In diesem Kapitel sollen alternative Schranken sowohl für den Standard-CUSUM wie auch für den OLS-basierten CUSUM-Test vorgeschlagen werden. Am Beispiel des Standard-CUSUM-Tests soll deshalb motiviert werden, warum dies sinnvoll ist.

Da Schranken für den Pfad einer Brownschen Bewegung gesucht werden, wäre es am naheliegendsten, eine Schranke zu betrachten, die in jedem Punkt proportional zur Standardabweichung dieses Pfades ist. Für eine Brownsche Bewegung mit Startpunkt 0 gilt, ähnlich wie in Gleichung (2.6), daß die Varianz gerade

$$\text{VAR}[B(t)] = t$$

beträgt. Deshalb wäre ein Ablehnungsbereich, der durch

$$b(t) = \lambda\sqrt{t} \tag{4.1}$$

begrenzt wird, wünschenswert. Da jedoch die Überschreitungswahrscheinlichkeit für Geraden leichter zu berechnen und insbesondere auch in geschlossener Form darstellbar ist, wählten BDE nicht die Funktion  $b(t)$  sondern vielmehr die Funktion  $c(t)$  aus (2.5), die für ein  $\lambda$  gerade die Tangente an  $b(t)$  im Punkt  $t = 0.5$  ist. So ist die verwendete lineare Schranke nach der Hälfte der Beobachtungen gerade die Tangente an die wünschenswerte Schranke.

Ähnliches gilt für den OLS-basierten CUSUM-Test. Die Parallele zur x-Achse  $\lambda$  ist im Punkt  $t = 0.5$  gerade die Tangente an die Funktion

$$d(t) = \lambda \sqrt{t(1-t)}, \quad (4.2)$$

die wiederum proportional zur Standardabweichung der Brownschen Brücke ist. Deren Varianz beträgt nämlich nach (2.12)

$$\text{VAR}[B^0(t)] = t(1-t).$$

## 4.1 Alternative Schranken für den Standard-CUSUM-Test

Wie in der Einleitung dieses Kapitels bereits erläutert, ist es wünschenswert und intuitiv, anstatt der üblichen linearen Schranken alternative Schranken zu betrachten, die proportional zur Standardabweichung des Pfades der Brownschen Bewegung sind. Die Testvorschrift dieses Tests lautet: lehne die Nullhypothese „Es liegt kein Strukturbruch vor“ ab, wenn der CUSUM-Pfad  $W_n(t)$  aus (2.8) die Schranke  $b(t)$  aus (4.1) über- bzw. die Schranke  $-b(t)$  unterschreitet. Völlig äquivalent hierzu ist es, die Nullhypothese zu verwerfen,

wenn die alternative Teststatistik

$$S_A = \sup_{\varepsilon \leq t \leq 1} \left| \frac{W_n(t)}{\sqrt{t}} \right|, \quad (4.3)$$

mit  $\varepsilon > 0$ , den kritischen Wert  $\lambda$  überschreitet.

Das Problem liegt nun darin, das zu jedem kritischen Wert  $\lambda$  gehörige Niveau  $\alpha$  zu bestimmen. Die zugehörigen Wertepaare sind, auch wieder asymptotisch, durch folgende Gleichung verknüpft:

$$\begin{aligned} \alpha &= P(S_A \geq \lambda) \\ &\stackrel{\text{asy}}{=} P(|B(t)| \geq b(t) \text{ für ein } \varepsilon \leq t \leq 1). \end{aligned} \quad (4.4)$$

Der Zeitpunkt  $t = 0$  wird ausgeschlossen, da die Bedingung  $B(0) \geq b(0) = 0$  ansonsten trivial erfüllt wäre. Es genügt aber auch nicht, allein diesen Punkt auszuschließen, da die Ablehnbedingung selbst dann asymptotisch immer erfüllt wäre. Also muß ein abgeschlossenes Intervall  $[\varepsilon, 1]$  mit  $\varepsilon > 0$  betrachtet werden; in dieser Arbeit wird  $\varepsilon = 0.001$  verwendet.

Um die Wertepaare von  $\lambda$  und  $\alpha$  zu ermitteln, für die Gleichung (4.4) erfüllt ist, werden nun zwei Methoden vorgestellt und angewendet:

**Methode 1:** Durch die kumulative Summe von  $n$  normalverteilten Zufallszahlen wird eine Brownsche Bewegung simuliert; anschließend wird überprüft, ob der Betrag dieser simulierten Brownschen Bewegung die Schranke  $b(t)$  (mit einem festen Parameter  $\lambda$ ) überschreitet oder nicht. Der Anteil der Überschreitungen bei  $k$  Wiederholungen dieser Prozedur liefert eine Schätzung für das zu  $\lambda$  gehörige Konfidenzniveau  $\alpha$ . Die Ergebnisse dieser Methode bei  $n = 5000$  und  $k = 5000$  wurden mit Hilfe der Funktion `smooth.spline` geglättet (durch Approximation durch ein Polynom). Die geglätteten Werte, die von den simulierten Wahrscheinlichkeiten nur einen

mittleren absoluten Abstand von  $2 \cdot 10^{-3}$  haben, sind in Tabelle 4.1 angegeben als  $\alpha_1$ .

**Methode 2:** Diese Methode ist eine Anwendung des Papiers von Wang und Pötzelberger (1997); diese betrachten Überschreitungswahrscheinlichkeiten von Brownschen Bewegungen für beliebige Schranken. Dabei wird die betrachtete Schranke durch eine stückweise lineare Funktion genähert. Für deren Überschreitungswahrscheinlichkeit geben Wang und Pötzelberger eine geschlossene Darstellung sowie einen Algorithmus zu deren Berechnung an. Bei der Durchführung dieser Methode wurde die Funktion  $b(t)$  an 128 äquidistanten Stellen im Intervall  $[0,1]$  ausgewertet und dazwischen linear interpoliert; für diese stückweise lineare Schranke wurde dann die angegebene Formel 200,000 Mal ausgewertet, was nach dem Algorithmus eine Schätzung für  $\alpha/2$  liefert. Denn auch hier wurde nur die Überschreitung einer einzigen Schranke betrachtet: die Approximation ist also wiederum nur für kleine  $\alpha$  gut und kann wieder für kleine  $\lambda$  den Wert 1 überschreiten. Vorteil dieser Methode ist allerdings, daß man eine Schätzung für die Standardabweichung der simulierten Wahrscheinlichkeiten erhält; diese lag für alle Ergebnisse  $\alpha_2$  dieser Methode in Tabelle 4.1 unterhalb von  $10^{-6}$ .

Beachtenswert ist, daß beide Methoden für große Werte von  $\lambda$  nahezu identische Werte liefern. Insbesondere sind die kritischen Werte zu den Niveaus 10%, 5% und 1% identisch:

$$\begin{aligned} \alpha &= 0.10, & \lambda &= 2.90, \\ \alpha &= 0.05, & \lambda &= 3.15, \\ \alpha &= 0.01, & \lambda &= 3.65. \end{aligned} \tag{4.5}$$



$\lambda$	1.00	1.05	1.10	1.15	1.20	1.25	1.30	1.35	1.40	1.45
$\alpha_1$	1.000	0.999	0.997	0.994	0.990	0.984	0.975	0.964	0.949	0.932
$\alpha_2$	1.676	1.636	1.594	1.548	1.502	1.454	1.402	1.348	1.294	1.238
$\lambda$	1.50	1.55	1.60	1.65	1.70	1.75	1.80	1.85	1.90	1.95
$\alpha_1$	0.912	0.889	0.864	0.836	0.806	0.773	0.739	0.703	0.666	0.627
$\alpha_2$	1.184	1.126	1.072	1.010	0.958	0.898	0.85	0.792	0.736	0.686
$\lambda$	2.00	2.05	2.10	2.15	2.20	2.25	2.30	2.35	2.40	2.45
$\alpha_1$	0.589	0.551	0.512	0.474	0.437	0.401	0.368	0.337	0.307	0.279
$\alpha_2$	0.636	0.588	0.542	0.502	0.458	0.416	0.382	0.346	0.312	0.284
$\lambda$	2.50	2.55	2.60	2.65	2.70	2.75	2.80	2.85	2.90	2.95
$\alpha_1$	0.253	0.228	0.205	0.183	0.163	0.145	0.129	0.114	0.100	0.088
$\alpha_2$	0.256	0.230	0.204	0.182	0.162	0.144	0.128	0.112	0.098	0.086
$\lambda$	3.00	3.05	3.10	3.15	3.20	3.25	3.30	3.35	3.40	3.45
$\alpha_1$	0.077	0.067	0.058	0.050	0.043	0.037	0.032	0.028	0.024	0.021
$\alpha_2$	0.076	0.066	0.058	0.050	0.042	0.038	0.032	0.026	0.024	0.020
$\lambda$	3.50	3.55	3.60	3.65	3.70	3.75	3.80	3.85	3.90	3.95
$\alpha_1$	0.018	0.015	0.012	0.010	0.009	0.007	0.006	0.005	0.004	0.004
$\alpha_2$	0.016	0.014	0.012	0.010	0.008	0.008	0.006	0.004	0.004	0.004
$\lambda$	4.00	4.05	4.10	4.15	4.20	4.25	4.30	4.35	4.40	4.45
$\alpha_1$	0.003	0.002	0.002	0.002	0.001	0.001	0.001	0.001	0.001	0.000
$\alpha_2$	0.002	0.002	0.002	0.002	0.002	0.002	0.000	0.000	0.000	0.000

Tabelle 4.1: Kritische Werte des alternativen Standard-CUSUM-Tests

Ein grafischer Vergleich der beiden Methoden in Abbildung 4.1 zeigt deutlich,

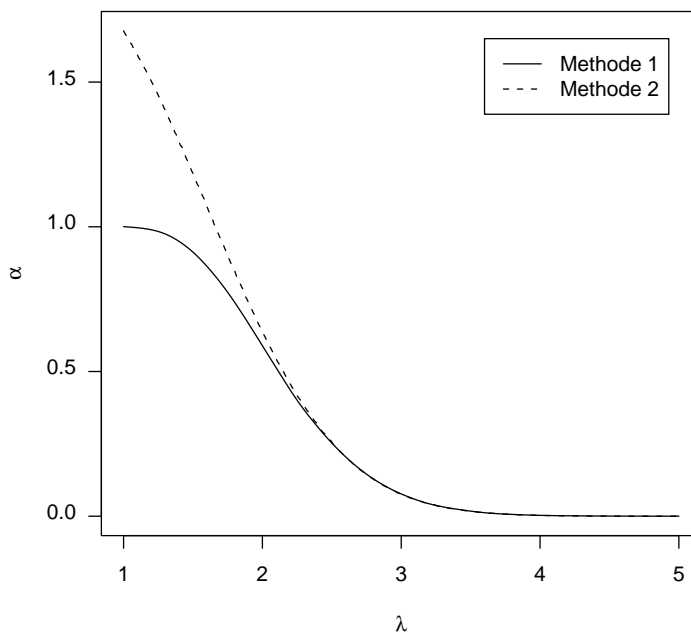


Abbildung 4.1: Ergebnisse von Methode 1 und 2

daß sich die Ergebnisse beider Simulationsmethoden nur für kleine  $\lambda$  unterscheiden. Dies gibt vor allem Vertrauen in die Ergebnisse von Methode 1: Bei Methode 1 ist es schwer zu evaluieren, wie zuverlässig die Ergebnisse einer solchen Simulation sind. Methode 2 hingegen liefert auch eine Schätzung für die Standardabweichung der simulierten Wahrscheinlichkeiten, die - wie bereits weiter oben gesagt - so gering ist, daß auch Wahrscheinlichkeiten  $\alpha_1$  als gute Approximation gelten können.

Da nun die kritischen Werte bestimmt sind, stellt sich natürlich die Fra-

ge, welche Auswirkungen dies auf die Form des Ablehnungsbereichs hat. Als Antwort darauf zeigt Abbildung 4.2 sowohl die lineare Schranke  $c(t)$  des

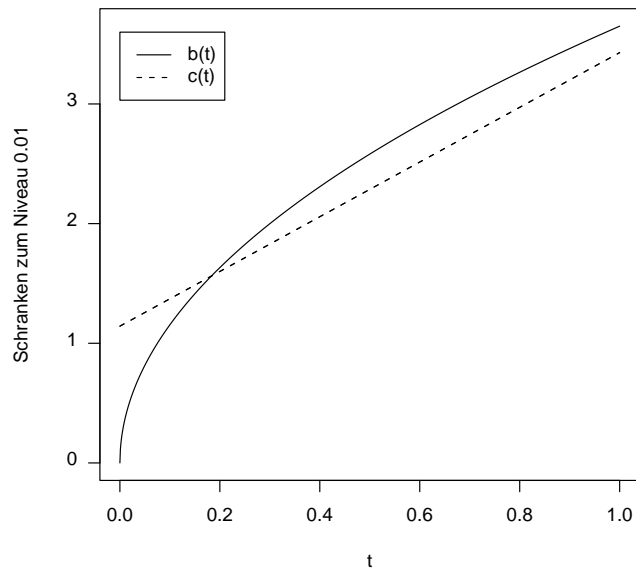


Abbildung 4.2: Schranken des Standard-CUSUM-Tests

Standard-CUSUM-Tests wie auch die neue Schranke  $b(t)$ , beide zum Niveau  $\alpha = 0.01$ . Die Form der entsprechenden Kurven zu anderen Niveaus ist sehr ähnlich, insbesondere liegt der Schnittpunkt der Kurven bei den üblichen kritischen Niveaus immer bei einem  $t \leq 0.2$ . Nach diesem Schnittpunkt liegt die neue Schranke über der linearen, was zur Folge hat, daß es mit diesem Ablehnungsbereich schwieriger wird, Strukturbrüche nach diesem Schnittpunkt zu erkennen. Der Vorteil, den man sich dadurch „erkauft“, ist aber eher gering: vor dem Schnittpunkt könnte man Strukturbrüche mit den al-

alternativen Schranken besser erkennen. Dabei bleibt aber zu beachten, daß der CUSUM-Pfad beim Bruchpunkt nur beginnt, vom Mittelwert 0 abzuwandern; bis er aber letztendlich den Ablehnungsbereich erreicht, kann aber bereits der Schnittpunkt passiert worden sein. Dies soll Abbildung 4.3 illustrieren. Darin ist der CUSUM-Pfad simulierter Daten zu sehen, die einen

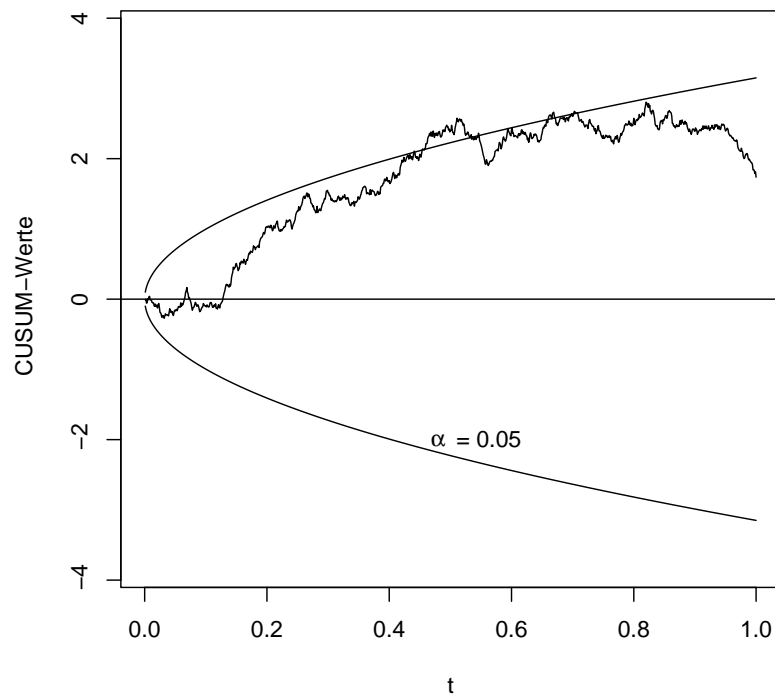


Abbildung 4.3: Standard-CUSUM-Test mit alternativen Schranken

Strukturbruch nach 10% der 1000 Beobachtungen enthalten. Ungefähr beim Bruchpunkt  $t = 0.1$  verläßt der CUSUM-Pfad den Mittelwert 0, jedoch überschreitet er erst wesentlich später die Schranke  $b(t)$ . Zu dem Punkt liegt die

Schranke aber bereits oberhalb der alten Funktion  $c(t)$ , so daß der herkömmliche Standard-CUSUM-Test die Nullhypothese erst recht verworfen hätte. Ein Vergleich der beiden Tests durch Simulation von erwarteten  $p$ -Werten folgt in Kapitel 5.

## 4.2 Alternative Schranken für den OLS-basierten CUSUM-Test

Auch für den OLS-basierten CUSUM-Test sollen nun alternative Schranken betrachtet werden, die proportional zur Standardabweichung der Brownschen Brücke sind. Die Testvorschrift des neuen Tests lautet dann: lehne  $H_0$  ab, wenn der OLS-basierte CUSUM-Pfad  $W_n^0(t)$  aus (2.11) die Schranke aus (4.2)

$$d(t) = \lambda \sqrt{t(1-t)}$$

überschreitet. Dazu äquivalent ist es wiederum, die Nullhypothese zu verwerfen, wenn die alternative Teststatistik

$$S_A^0 = \sup_{\varepsilon \leq t \leq 1-\varepsilon} \left| \frac{W_n^0(t)}{\sqrt{t(1-t)}} \right|. \quad (4.6)$$

den kritischen Wert  $\lambda$  überschreitet. Schwierig ist nun wieder die Bestimmung des zugehörigen Niveaus  $\alpha$ , das mit dem kritischen Wert asymptotisch über folgende Gleichung verknüpft ist:

$$\begin{aligned} \alpha &= P(S_A^0 \geq \lambda) \\ &\stackrel{\text{asy}}{=} P(|B^0(t)| \geq d(t) \text{ für ein } \varepsilon \leq t \leq 1-\varepsilon). \end{aligned} \quad (4.7)$$

Hier müssen beide Grenzen des Zeitintervalls ausgeschlossen werden, da die Bedingung  $B^0(t) \geq d(0) = d(1) = 0$  trivial erfüllt wäre. Mit den gleichen Argumenten wie beim Standard-CUSUM-Test wird deshalb das abgeschlossene Intervall  $[\varepsilon, 1 - \varepsilon]$  (hier mit  $\varepsilon = 0.001$ ) betrachtet.

Da sich die Methode 1 aus dem vorangegangenen Abschnitt als sehr gut für die Approximation der kritischen Werte bzw. Überschreitungswahrscheinlichkeiten herausgestellt hat, wird sie auch hier angewendet: es wurden in  $k = 5000$  Wiederholungen jeweils Brownsche Brücken durch die kumulative Summe von  $n = 5000$  mittelwertbereinigten normalverteilten Zufallszahlen simuliert. Die so ermittelten empirischen Überschreitungshäufigkeiten wurden wiederum geglättet. Die Ergebnisse sind aus Tabelle 4.2 abzulesen. Insbesondere ergeben sich zu den üblichen Konfidenzniveaus von 10%, 5% und 1% folgende kritische Werte:

$$\begin{aligned} \alpha = 0.10, \quad \lambda &= 3.13, \\ \alpha = 0.05, \quad \lambda &= 3.37, \\ \alpha = 0.01, \quad \lambda &= 3.83. \end{aligned} \tag{4.8}$$

Mit Hilfe dieser Werte können nun wieder die beiden verschiedenen Ablehnungsbereiche für den OLS-basierten CUSUM-Pfad verglichen werden. Abbildung 4.4 zeigt die beiden Schranken zum Niveau  $\alpha = 0.01$  im Vergleich. Dabei fällt sofort auf, daß, anders als bei den Schranken zum Standard-CUSUM-Pfad, sowohl am Anfang wie auch am Ende die neue Schranke unterhalb der alten liegt. Zwar muß sich dieser Vorteil an den Rändern auch wieder durch einen Nachteil in der Mitte „erkauft“ werden, wo die neue Schranke überhalb der alten verläuft, doch dieser Vorteil kann auch einen praktischen

$\lambda$	1.20	1.25	1.30	1.35	1.40	1.45	1.50	1.55	1.60	1.65
$\alpha$	1.000	1.000	1.000	0.999	0.997	0.994	0.990	0.985	0.977	0.967
$\lambda$	1.70	1.75	1.80	1.85	1.90	1.95	2.00	2.05	2.10	2.15
$\alpha$	0.954	0.938	0.919	0.896	0.871	0.843	0.812	0.778	0.743	0.705
$\lambda$	2.20	2.25	2.30	2.35	2.40	2.45	2.50	2.55	2.60	2.65
$\alpha$	0.666	0.625	0.585	0.544	0.504	0.464	0.426	0.389	0.353	0.320
$\lambda$	2.70	2.75	2.80	2.85	2.90	2.95	3.00	3.05	3.10	3.15
$\alpha$	0.288	0.258	0.230	0.205	0.182	0.161	0.142	0.124	0.109	0.095
$\lambda$	3.20	3.25	3.30	3.35	3.40	3.45	3.50	3.55	3.60	3.65
$\alpha$	0.082	0.071	0.062	0.053	0.046	0.039	0.034	0.029	0.025	0.021
$\lambda$	3.70	3.75	3.80	3.85	3.90	3.95	4.00	4.05	4.10	4.15
$\alpha$	0.017	0.014	0.011	0.009	0.008	0.006	0.005	0.005	0.004	0.003
$\lambda$	4.20	4.25	4.30	4.35	4.40	4.45	4.50	4.55	4.60	4.65
$\alpha$	0.003	0.002	0.002	0.001	0.001	0.001	0.001	0.001	0.000	0.000

Tabelle 4.2: Kritische Werte des alternativen OLS-basierten CUSUM-Tests

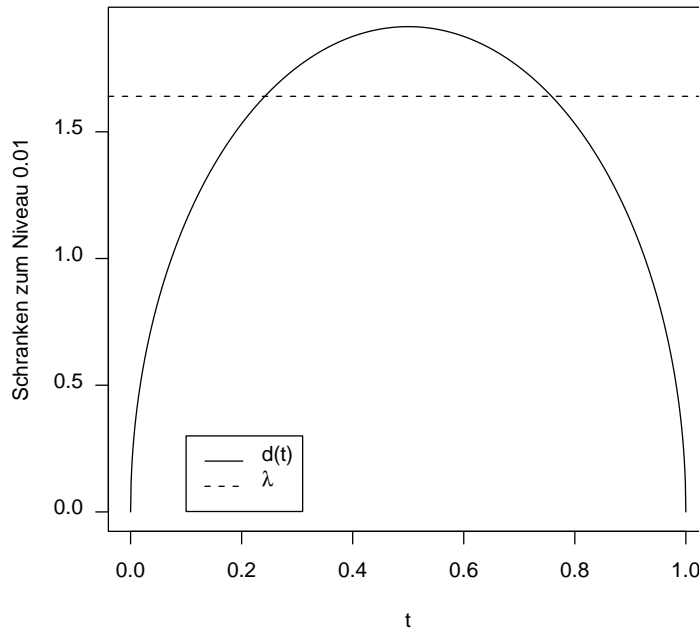


Abbildung 4.4: Schranken des OLS-basierten CUSUM-Tests

Wert haben, wie Abbildung 4.5 zeigt. Darin ist ebenfalls der OLS-basierte CUSUM-Pfad zu simulierten Daten mit einem Strukturbruch nach 10% der 1000 Beobachtungen zu sehen.

Während die alten Schranken zum Niveau  $\alpha = 0.01$  keinen signifikanten Strukturbruch feststellen können, verwirft der alternative OLS-basierte CUSUM-Test zu demselben Niveau die Nullhypothese. Der Grund dafür liegt in der Eigenschaft des OLS-CUSUM-Pfades, etwa an der Stelle des Strukturbruchs seinen extremsten Wert anzunehmen; und nicht wie der Standard-CUSUM-Pfad am Bruchpunkt zu beginnen, den mittleren Wert zu verlas-



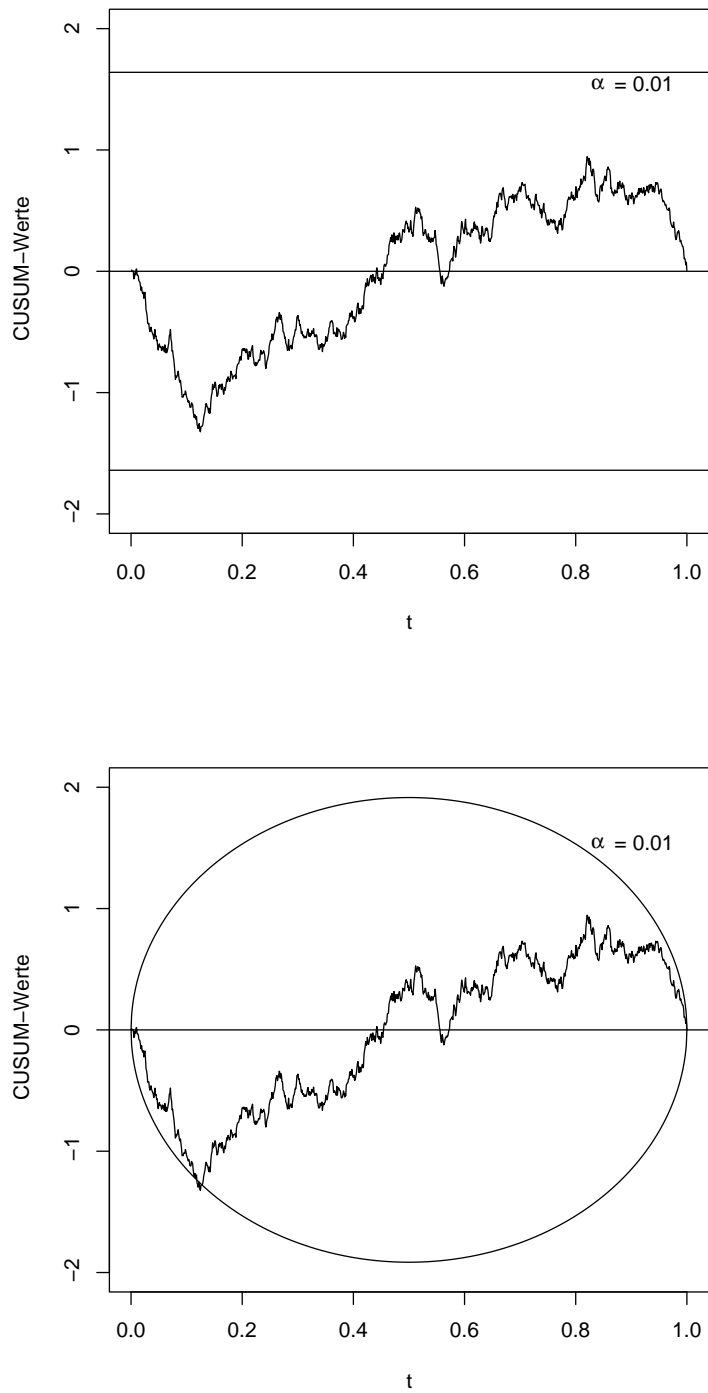


Abbildung 4.5: OLS-basierter CUSUM-Test mit linearen und alternativen Schranken

sen. So kann der neue OLS-basierte CUSUM-Test frühe Strukturbrüche wie in [Abbildung 4.5](#) und völlig symmetrisch auch späte Strukturbrüche besser erkennen als der herkömmliche Test.

Dies soll im folgenden Kapitel durch Simulationen belegt werden.

# Kapitel 5

## Simulation erwarteter $p$ -Werte

Die Vorgehensweise bei der Simulation von Strukturbrüchen ist dieselbe, wie sie [Ploberger und Krämer \(1992\)](#) verwenden. Diese betrachteten ein Modell mit  $k = 2$  Regressoren, in dem die Regressoren die Gestalt  $x_t = (1, (-1)^t)^\top$  haben und die Fehlerterme unabhängig identisch standardnormalverteilt sind. Um verschiedene Arten von Strukturbrüchen zu betrachten, wird der Zeitpunkt, die Intensität und der Winkel eines einzigen Strukturbruchs variiert, so daß die Regressionskoeffizienten folgende Form haben:

$$\beta_t = \begin{cases} \beta & \text{für } t \leq \lfloor qn \rfloor \\ \beta + \Delta\beta & \text{für } t > \lfloor qn \rfloor \end{cases}. \quad (5.1)$$

Dabei ist  $q$  der Anteil der Beobachtungen vor dem Strukturbruch und für die neuen Regressionskoeffizienten gilt:

$$\Delta\beta = \frac{g}{\sqrt{n}} \begin{pmatrix} \cos \psi \\ \sin \psi \end{pmatrix}. \quad (5.2)$$

So ist  $\psi$  der Winkel zwischen dem Bruch  $\Delta\beta$  und dem mittleren Regressor  $(1, 0)^\top$  und die Intensität des Bruches ist gegeben durch  $\|\Delta\beta\| = |g|\sqrt{n}$ .

Für die Simulation der erwarteten  $p$ -Werte wurde in  $k = 1000$  Wiederholungen jeweils eine Teststatistik basierend auf Daten ohne Strukturbruch und eine Teststatistik basierend auf denselben Daten (i.e. denselben Fehlertermen) mit einem Strukturbruch, festgelegt durch (5.1) und (5.2), erzeugt und jeweils ermittelt, ob die unter  $H_0$  erzeugte größer ist. Die empirische Wahrscheinlichkeit ist dann gemäß der Darstellung (2.17) eine Schätzung für den wahren erwarteten  $p$ -Wert. Hierfür benötigt man insbesondere keine kritischen Werte.

Um sowohl frühe wie auch späte Strukturbrüche zu betrachten, hat  $q$  in den Simulationen die Werte 0.1, 0.3, 0.5, 0.7 und 0.9, die Intensität nimmt Werte  $g = 4.8, 7.2, 9.6$  und  $12.0$  an und der Winkel variiert schließlich zwischen  $\psi = 0, 18, 36, 36, 54, 72$  und  $90$ .

Dabei ist, beispielsweise nach PK oder nach [Ploberger und Krämer \(1990\)](#), wohlbekannt, daß beide Tests (fast) keine Möglichkeit haben, Strukturbrüche mit einem Winkel von  $90$  zu erkennen. Der Standard-CUSUM-Test kann ebenfalls späte Strukturbrüche nur schwer aufdecken. Dies ist auch in [Tabelle 5.1](#) deutlich abzulesen: die erwarteten  $p$ -Werte in der letzten Spalte und untersten Zeile liegen alle nahe bei  $0.5$ , dem erwarteten  $p$ -Wert unter  $H_0$ , so daß es recht unwahrscheinlich ist, einen derartigen Strukturbruch mit Hilfe des Standard-CUSUM-Tests zu entdecken. Die besten, d.h. kleinsten,  $p$ -Werte hat der Test bei Strukturbrüchen nach  $30\%$  bzw.  $50\%$  der Daten bei einem Winkel zwischen  $0$  und  $36$ . Die Chancen einen solchen Strukturbruch zu offenbaren, steigen intuitiverweise mit der Intensität des Bruchs.

Vergleicht man nun aber diese Ergebnisse mit den analogen Werten für den alternativen Standard-CUSUM-Test in [Tabelle 5.2](#), so sieht man leicht, daß

$q$	$g$	$\psi$					
		0	18	36	54	72	90
0.1	4.8	0.356	0.367	0.401	0.478	0.484	0.469
	7.2	0.231	0.218	0.288	0.381	0.455	0.500
	9.6	0.140	0.155	0.184	0.289	0.448	0.497
	12.0	0.062	0.052	0.126	0.207	0.367	0.511
0.3	4.8	0.264	0.247	0.344	0.398	0.481	0.519
	7.2	0.105	0.116	0.197	0.305	0.423	0.489
	9.6	0.047	0.038	0.073	0.204	0.414	0.496
	12.0	0.007	0.014	0.036	0.093	0.320	0.485
0.5	4.8	0.263	0.300	0.362	0.383	0.479	0.513
	7.2	0.141	0.173	0.235	0.313	0.459	0.509
	9.6	0.051	0.069	0.135	0.221	0.425	0.502
	12.0	0.033	0.017	0.052	0.164	0.373	0.514
0.7	4.8	0.373	0.419	0.439	0.488	0.509	0.52
	7.2	0.318	0.304	0.342	0.420	0.492	0.530
	9.6	0.170	0.200	0.254	0.357	0.470	0.524
	12.0	0.104	0.130	0.172	0.288	0.452	0.524
0.9	4.8	0.511	0.533	0.472	0.499	0.486	0.522
	7.2	0.468	0.462	0.474	0.485	0.510	0.508
	9.6	0.482	0.453	0.507	0.471	0.493	0.524
	12.0	0.425	0.458	0.453	0.470	0.536	0.511

Tabelle 5.1: Simulation erwarteter  $p$ -Werte des Standard-CUSUM-Tests

$q$	$g$	$\psi$					
		0	18	36	54	72	90
0.1	4.8	0.400	0.409	0.417	0.501	0.510	0.504
	7.2	0.281	0.291	0.350	0.422	0.490	0.480
	9.6	0.185	0.188	0.245	0.322	0.446	0.515
	12.0	0.080	0.102	0.167	0.247	0.444	0.505
0.3	4.8	0.317	0.324	0.380	0.423	0.495	0.495
	7.2	0.154	0.186	0.265	0.360	0.498	0.532
	9.6	0.063	0.084	0.133	0.291	0.418	0.518
	12.0	0.010	0.015	0.057	0.189	0.404	0.511
0.5	4.8	0.328	0.396	0.416	0.448	0.507	0.504
	7.2	0.221	0.196	0.308	0.416	0.477	0.520
	9.6	0.101	0.103	0.159	0.313	0.483	0.532
	12.0	0.039	0.038	0.086	0.217	0.419	0.569
0.7	4.8	0.468	0.467	0.480	0.482	0.521	0.507
	7.2	0.360	0.371	0.383	0.468	0.523	0.519
	9.6	0.243	0.264	0.317	0.428	0.554	0.532
	12.0	0.165	0.176	0.241	0.386	0.501	0.550
0.9	4.8	0.513	0.505	0.484	0.536	0.466	0.535
	7.2	0.535	0.521	0.486	0.523	0.503	0.511
	9.6	0.493	0.507	0.494	0.514	0.525	0.543
	12.0	0.456	0.450	0.465	0.498	0.531	0.542

Tabelle 5.2: Simulation erwarteter  $p$ -Werte des alternativen Standard-CUSUM-Tests

letzterer keine Vorteile beim Nachweis von Strukturbrüchen bietet. Selbst bei frühen Strukturbrüchen ( $q = 0.1$ ) hat der neue Test eher höhere erwartete  $p$ -Werte. Bei späteren Brüchen ist, wie schon im vorigen Kapitel erläutert wurde, die Chance ohnehin geringer, diese zu entdecken.

Sieht man sich nun aber die entsprechenden Tabellen für den OLS-basierten CUSUM-Test an, so ist zu erkennen, daß der herkömmliche Test leichte Schwächen bei frühen respektive späten Strukturbrüchen hat, während er bei Strukturbrüchen für  $q$  zwischen 0.3 und 0.7 recht gute Eigenschaften hat (wie in Tabelle 5.3 abzulesen). Vergleicht man diese erwarteten  $p$ -Werte mit denen aus Tabelle 5.4 für den alternativen OLS-basierten CUSUM-Test, so sind die Chancen einen Bruch in der Mitte nachzuweisen, wie bereits im letzten Kapitel angedeutet, etwas geringer, aber dafür sind die erwarteten  $p$ -Werte an den Rändern niedriger. Da die OLS-Residuen immer von der Gesamtheit aller Beobachtungen abhängen, haben sie unter der Alternative günstigere Eigenschaften, um diesen Test zu konstruieren, der recht gleichmäßige Macht für Strukturbrüche zu verschiedenen Zeitpunkten hat. Ist also vor der Anwendung des Tests nicht klar, wo der Bruchpunkt (ungefähr) liegen könnte, so ist der alternative OLS-basierte CUSUM-Test eine gute Möglichkeit, einen Strukturbruch nachzuweisen.

$q$	$g$	$\psi$					
		0	18	36	54	72	90
0.1	4.8	0.395	0.415	0.419	0.489	0.458	0.505
	7.2	0.293	0.329	0.374	0.425	0.509	0.483
	9.6	0.201	0.207	0.267	0.389	0.484	0.526
	12.0	0.124	0.133	0.166	0.331	0.442	0.545
0.3	4.8	0.151	0.198	0.255	0.350	0.449	0.536
	7.2	0.045	0.060	0.121	0.221	0.435	0.513
	9.6	0.005	0.015	0.029	0.127	0.332	0.544
	12.0	0.002	0.001	0.008	0.055	0.308	0.519
0.5	4.8	0.118	0.144	0.224	0.338	0.468	0.509
	7.2	0.020	0.024	0.067	0.171	0.381	0.500
	9.6	0.004	0.006	0.016	0.065	0.294	0.508
	12.0	0.000	0.000	0.003	0.039	0.247	0.504
0.7	4.8	0.179	0.185	0.252	0.364	0.464	0.484
	7.2	0.038	0.051	0.119	0.226	0.398	0.512
	9.6	0.005	0.010	0.031	0.121	0.377	0.521
	12.0	0.000	0.000	0.005	0.056	0.299	0.512
0.9	4.8	0.399	0.431	0.421	0.443	0.488	0.482
	7.2	0.334	0.339	0.392	0.412	0.480	0.487
	9.6	0.211	0.204	0.268	0.384	0.488	0.507
	12.0	0.109	0.136	0.190	0.308	0.458	0.508

Tabelle 5.3: Simulation erwarteter  $p$ -Werte des OLS-basierten CUSUM-Tests



$q$	$g$	$\psi$					
		0	18	36	54	72	90
0.1	4.8	0.367	0.382	0.396	0.436	0.477	0.503
	7.2	0.256	0.246	0.308	0.409	0.457	0.494
	9.6	0.138	0.145	0.231	0.309	0.455	0.500
	12.0	0.054	0.071	0.118	0.260	0.395	0.494
0.3	4.8	0.227	0.243	0.300	0.393	0.456	0.507
	7.2	0.066	0.087	0.138	0.293	0.438	0.533
	9.6	0.023	0.012	0.037	0.156	0.390	0.539
	12.0	0.001	0.003	0.010	0.106	0.346	0.514
0.5	4.8	0.197	0.207	0.262	0.359	0.465	0.505
	7.2	0.049	0.075	0.114	0.271	0.422	0.513
	9.6	0.004	0.011	0.038	0.141	0.395	0.536
	12.0	0.001	0.001	0.007	0.045	0.338	0.520
0.7	4.8	0.224	0.250	0.303	0.415	0.485	0.522
	7.2	0.080	0.077	0.170	0.276	0.441	0.523
	9.6	0.009	0.016	0.039	0.184	0.394	0.527
	12.0	0.002	0.006	0.012	0.095	0.347	0.544
0.9	4.8	0.383	0.376	0.426	0.453	0.482	0.488
	7.2	0.229	0.264	0.291	0.397	0.480	0.488
	9.6	0.137	0.154	0.203	0.328	0.447	0.500
	12.0	0.060	0.074	0.129	0.229	0.404	0.494

Tabelle 5.4: Simulation erwarteter  $p$ -Werte des alternativen OLS-basierten CUSUM-Tests

# Kapitel 6

## Zusammenfassung

In dieser Diplomarbeit werden zunächst zwei Approximationen für die  $p$ -Wert-Funktionen der beiden CUSUM-Tests mit linearen Schranken hergeleitet. Es kann jeweils eine obere Schranke angegeben werden, die nichts anderes ist als das Zweifache der Überschreitungswahrscheinlichkeit einer Brownschen Bewegung, respektive einer Brownschen Brücke, für eine einzige lineare Schranke. Dabei wird also die Überschreitung beider Schranken vernachlässigt, so daß man die untere Schranke erhält, wenn man die Wahrscheinlichkeit für die Überschreitung beider Schranken abzieht. Dies ist nur eine untere Schranke, da noch häufigere Über- bzw. Unterschreitungen vernachlässigt werden. Beide Schranken ermöglichen die äußerst exakte Berechnung von für die Praxis relevanten  $p$ -Werten ( $p \leq 0.1$ ) und die untere Schranke eignet sich auch für die Approximation größerer  $p$ -Werte. Zuletzt wird auch eine Funktion angegeben, die sich besonders zur Implementation im Rahmen eines Software-Paketes eignet.

Im Anschluß daran werden alternative Schranken für beide Tests vorgeschla-

gen. Zunächst wird für den Standard-CUSUM-Test eine Schranke betrachtet, die proportional zur Standardabweichung der Brownschen Bewegung ist. Durch zwei verschiedene Simulationsmethoden werden (Schätzungen für) kritische Werte evaluiert. Allerdings stellt sich heraus, daß der Test mit den alternativen Schranken wegen des Verhaltens der rekursiven Residuen unter Hypothese und Alternative kaum Vorteile gegenüber der Standard-Version bieten kann. Da die Schranken (zu identischen Niveaus) sich nämlich nur einmal schneiden und die alternative Schranke vor diesem Schnittpunkt kleiner und anschließend überall größer ist als die herkömmliche, kann sie keine Vorteile bei späten Strukturbrüchen bieten. Und selbst bei frühen Strukturbrüchen braucht der CUSUM-Pfad zuviel „Zeit“ um die Schranke zu überschreiten, so daß die Vorteile der neuen Schranke zu Beginn des betrachteten Zeitintervalls nicht genutzt werden können.

Völlig analog werden alternative Schranken für den OLS-basierten CUSUM-Test vorgeschlagen, die nämlich proportional zur Standardabweichung der Brownschen Brücke sind. Mit Hilfe von Monte-Carlo-Methoden können auch hier kritische Werte evaluiert werden. Wegen der Eigenschaften der OLS-Residuen unter Hypothese und Alternative kann hier die Güte des OLS-basierten CUSUM-Tests bei frühen und späten Strukturbrüchen in der Tat verbessert werden. Der OLS-CUSUM-Pfad erreicht seinen extremsten Wert in der Nähe des Strukturbruchs, so daß die Vorteile der alternativen Schranke, die zu Beginn und zu Ende des Zeitintervalls unterhalb der herkömmlichen Schranke (zum selben Niveau) liegt, wirklich ausgenutzt werden können. Auch wenn die Güte bei Strukturbrüchen nach etwa der Hälfte der Beobachtungen etwas abnimmt, so ist das Ergebnis doch ein Test, der recht

gleichmäßige Ablehneigenschaften für verschiedene mögliche Bruchzeitpunkte hat. Sollte also einer der vier betrachteten Tests auf Daten angewendet werden, in denen ein Strukturbruch vermutet wird, dessen Zeitpunkt aber unbekannt ist (was in der Praxis häufig der Fall ist), so sollte am ehesten der alternative OLS-basierte CUSUM-Test verwendet werden.

# Literaturverzeichnis

- D. W. K. Andrews. Tests for parameter instability and structural change with unknown change point. *Econometrica*, 61:821–856, 1993.
- P. Billingsley. *Convergence of Probability Measures*. Wiley, New York, 2nd edition, 1999.
- R. L. Brown, J. Durbin, und J. M. Evans. Techniques for testing the constancy of regression relationships over time. *Journal of the Royal Statistical Society*, B 37:149–163, 1975.
- G. C. Chow. Tests of equality between sets of coefficients in two linear regressions. *Econometrica*, 28:591–605, 1960.
- J. Durbin. Boundary-crossing probabilities for the Brownian motion and poisson process and techniques for computing the power of the Kolmogorov-Smirnov test. *Journal of Applied Probability*, 8:431–453, 1971.
- B. E. Hansen. Tests for parameter instability in regressions with  $I(1)$  processes. *Journal of Business & Economic Statistics*, 10:321–335, 1992.
- W. Krämer, W. Ploberger, und R. Alt. Testing for structural change in dynamic models. *Econometrica*, 56(6):1355–1369, 1988.

- W. Krämer und H. Sonnberger. *The Linear Regression Model under Test*. Physica-Verlag, Heidelberg, 1986.
- W. Ploberger und W. Krämer. The local power of the CUSUM and CUSUM of squares test. *Econometric Theory*, 6:335–347, 1990.
- W. Ploberger und W. Krämer. The CUSUM test with OLS residuals. *Econometrica*, 60(2):271–285, 1992.
- R Development Core Team. An introduction to R, 2000. URL <http://cran.r-project.org>. Manual.
- L. Sachs. *Angewandte Statistik*. Springer, Berlin, 1997.
- H. Sackrowitz und E. Samuel-Cahn.  $p$  values as random variables – expected  $p$  values. *The American Statistician*, 53(4):326–331, 1999.
- L. Wang und K. Pötzelberger. Boundary crossing probability for Brownian motion and general boundaries. *Journal of Applied Probability*, 34:54–65, 1997.