Regionen von Alternativen mit hoher Güte für Anpassungstests

I n a u g u r a l - D i s s e r t a t i o n

zur
Erlangung des Doktorgrades der
Mathematisch-Naturwissenschaftlichen Fakultät
der Heinrich-Heine-Universität Düsseldorf

vorgelegt von

Hülya Ünlü

aus Duisburg

Juli 2008

Aus dem Mathematischen Institut
der Heinrich-Heine-Universität Düsseldorf

Gedruckt mit der Genehmigung der
Mathematisch-Naturwissenschaftlichen Fakultät der
Heinrich-Heine-Universität Düsseldorf

Referent: Prof. Dr. A. Janssen

Koreferent: Prof. Dr. W. Bischoff

Tag der mündlichen Prüfung: 26.Juni 2008

Sevdiklerime

Zusammenfassung

Jedes statistische Computerpaket weist eine große Menge von nichtparametrischen Testverfahren für einseitige oder zweiseitige Testprobleme auf, welche durch stochastisch geordnete Alternativen gegeben sind. Häufig verwendete Anpassungstests sind Tests vom Kolmogoroff-Smirnoff-Typ oder Integraltests. Dabei ist der Vergleich der asymptotischen Gütefunktionen nichtparametrischer Tests ausschlaggebend für die Auswahl konkurrierender Tests. Nun ist es bekannt, dass die Auswahl eines nichtparametrischen Tests a priori eine endlichdimensionale Region von Alternativen mit hoher Güte festlegt. Für Alternativen aus dem orthogonalen Komplement ist die Gütefunktion dagegen flach, vergleiche JANSSEN [16], [17]. Das Problem ist, wie man Unterräume von Alternativen mit hoher Güte für einen konkreten Test findet. Typischerweise können nichtparametrische Gütefunktionen nicht explizit berechnet werden, da in der Nichtparametrik der Alternativenraum von unendlicher Dimension ist.

Es wird vorgeschlagen, das Problem für verschiedene Anpassungstests mit der Methode konkaver Majoranten zu lösen. Diese Idee beruht auf eine Serie von Arbeiten von BISCHOFF et al. Die Güte der Tests in Richtung eines Signals S_g, gegeben durch eine Alternative $g \in L_2([0,1], \lambda\!\!\lambda_{|[0,1]})$, lassen sich dann durch die Güte der konkaven Majoranten von S_g abschätzen.

In den ersten beiden Kapiteln werden die im Rahmen dieser Arbeit wesentlichen Hilfsmittel bereitgestellt. Dabei werden unter anderem zwei wichtige Regressionsmodelle vorgestellt, nämlich der Signalprozess der Brownschen Brücke und der Signalprozess der Brownschen Bewegung. Diese stellen Gauß-Shift-Experimente dar und treten in der Nichtparametrik als Limesexperimente auf.

Im dritten Kapitel wird für verschiedene einseitige Anpassungstests die Gütefunktion mit Hilfe einer Projektionsmethode abgeschätzt. Im weiteren Verlauf zeigt sich, dass Projektionen von Tangenten g auf Kegeln zu Signalen führen, die kleinste konkave Majoranten zu S_g sind. Dabei werden insbesondere Alternativen aus Räumen betrachtet, die von Rademacherfunktionen erzeugt werden. Mit Hilfe dieser Räume können für einseitige Anpassungstests Regionen von Alternativen angegeben werden, wo die Gütefunktion flach ist.

Im vierten Kapitel werden für zweiseitige Anpassungstests wieder mit Hilfe von Rademacherfunktionen obere Schranken für die Gütefunktion angegeben, die sich um einen additiven Term von den Schranken im einseitigen Fall unterscheiden.

Zum Schluss dieser Arbeit erfolgt die statistische Interpretation der Räume von Alternativen mit hoher Güte. Diese gehören zu den ungünstigsten Richtungen einer Klasse von statistischen Funktionalen, welche Linearkombinationen von Quantilfunktionen sind.

Abstract

Every statistical computerpackage shows a lot of non-parametric test procedures for one-sided or two-sided testproblems, which are given by stochastically ordered alternatives. Widespread goodness-of-fit tests are Kolmogorov-Smirnov type tests or integral tests. Thereby the comparision between the asymptotic power function of non-parametric tests is the determining factor for the choice of competing tests. Now it is known that the choice of a non-parametric test fixes a priori a finite dimensional region of alternatives with high power. For alternatives belonging to its orthogonal complement the power function is flat, see JANSSEN [16], [17]. The problem is finding the subspaces of alternatives with high power for a concrete test. Typically, non-parametric power functions can not be calculated explicitly because the region of alternatives has infinite dimension in non-parametric.

In this thesis it is proposed to treat the problem for various goodness-of-fit tests with a method of concave majorants. This idea is based upon a series of works of BISCHOFF et al. The power of tests towards a signal S_g, given by an alternative $g \in L_2([0,1], \lambda\!\!\lambda_{|[0,1]})$, can be estimated by the power of a concave majorants of S_g.

The first two chapters allocate the basic means within this thesis. In doing so there are two important regression modells introduced: the signal process for the Brownian bridge model and the signal process for the Brownian motion model. These bring out Gaussian shift and appear as limit experiments in the non-parametric.

In the third chapter the power function is going to be evaluated via a projection method for various one-sided goodness-of-fit tests. In the following course it is shown that projections of tangents g on cones lead to signals, which are the smallest concave majorants of S_g. In the process regions of alternatives, spanned by Rademacherfunctions, are especially examined. Via these spaces, regions of alternatives for one-sided goodness-of-fit tests can be named, where the power function is flat.

In the fourth chapter via Rademacherfunctions, upper bounds for the power functions of two-sided goodness-of-fit tests are named, which differ from the bounds in a one-sided case about an additional term.

In the end of this thesis the statistical interpretation of the regions of alternatives with high power is carried out. These alternatives belong to least favorable directions of a class of statistical functionals which are linear combinations of quantile functions.

Inhaltsverzeichnis

Einleitung

In dieser Arbeit wird die nichtparametrische Gütefunktion von Anpassungstests diskutiert. Es geht um das Bestimmen von Alternativen mit hoher Güte.
Es ist bekannt, dass die Auswahl eines nichtparametrischen Tests a priori eine endlich-dimensionale Region von Alternativen mit hoher Güte festlegt. Für Alternativen aus dem orthogonalen Komplement ist die Gütefunktion flach, vgl NEUHAUS [32], MILBRODT und STRASSER [30], JANSSEN [16], [17] sowie LEHMANN und ROMANO [28]. Nun ist das Problem, wie man die Unterräume von Alternativen mit hoher Güte für einen konkreten Test findet. Typischerweise können nichtparametrische Gütefunktionen nicht explizit berechnet werden. Nur für bestimmte Tests können Räume und Richtungen mit hoher Güte näherungsweise bestimmt werden, vgl. JANSSEN [13], HÁJEK et al. [10] und RAHNENFÜHRER [36]. Obwohl moderne Computer eine umfangreiche Monte-Carlo-Simulation erlauben, ist es ein sehr schwieriges Problem, Alternativen mit maximaler Güte zu finden. Dies ist ein Optimierungsproblem für unendlich-dimensionale Räume von Alternativen.
Es wird vorgeschlagen, das Problem für verschiedene Tests zu lösen, indem Räume von endlicher Dimension mit hoher Güte bestimmt werden. Diese Räume werden durch die Rademacherfunktionen beschrieben, welche ein Orthonormalsystem im Hilbertraum $L_2(\lambda\!\!\lambda_{|[0,1]})$ darstellen. Nachdem die Dimension von Alternativen reduziert ist, kann eine Monte-Carlo-Simulation der exakten Gütefunktion erfolgen. Die Lösungen sind dann näherungsweise bestimmt, wobei der approximative Fehler kontrolliert werden kann.
Der asymptotische Aufbau ist durch das folgende Einstichprobenproblem motiviert. Seien $X_1, ..., X_n$ reelle i.i.d. Zufallsvariable mit stetiger Verteilungsfunktion F und sei F_0 eine weitere stetige Verteilungsfunktion. Wir testen stochastisch geordnete Hypothesen, nämlich

$$H_1 : F \leq F_0 \quad \text{gegen} \quad K_1 : F \geq F_0, F \neq F_0 \tag{1}$$

oder das zweiseitige Testproblem

$$H_2 : F = F_0 \text{ gegen } K_2 : F \neq F_0.$$

Um unser bereits oben beschriebenes Problem zu lösen, nehmen wir an, dass eine stetige Verteilungsfunktion F_0 sowie komplexe Räume von Alternativen von unendlicher Dimension gegeben sind. Unser Ziel ist die Klassifizierung nichtparametrischer Test durch deren nichtparametrische Gütefunktionen. Dadurch erfolgt eine Reduktion auf Räume von Alternativen mit endlicher Dimension. An dieser Stelle tritt das bekannte Vierpunkteprogramm von LeCam [27] zur asymptotischen Behandlung von Entscheidungsproblemen ein. Zum endlichen Stichprobenumfang n liege ein Testproblem für ein statistisches Experiment E_n zugrunde.

- Man zeigt, dass die Loglikelihoodprozesse der Folge von Experimenten $(E_n)_n$ asymptotisch durch den Loglikelihoodprozess des Limesexperiments G ersetzt werden können, d.h. die Folge $(E_n)_n$ konvergiert schwach gegen den Gauß-Shift G.
- Man löst die Entscheidungsprobleme für G.
- Man zeigt, dass die Risikoschranken für Entscheidungsprobleme in G asymptotische Risikoschranken für die betrachtete Folge von Experimenten $(E_n)_n$ darstellen.
- Die optimalen statistischen Verfahren vom Gauß-Shift G werden dann auf $(E_n)_n$ übertragen.

Die vorliegende Arbeit besteht aus vier Kapiteln. Die Ergebnisse des Forschungprojekts werden teilweise in der Zeitschrift Journal of Statistical Planning and Inference, JANSSEN und ÜNLÜ [24], im August 2008 publiziert.
Im ersten Kapitel werden zwei wichtige Regressionsmodelle der Statistik vorgestellt. Zum einen wird das als Signalerkennungsproblem mit Brownscher Brücke B_0 als Rauschterm bezeichnete Experiment

$$E = (C_0, \mathcal{B}(C_0), \{\mathcal{L}(B_0 + S_g | B^0) : g \in H_0\}). \tag{2}$$

behandelt. Dabei seien $C_0 := \{f \in C([0,1]) : f(0) = f(1) = 0\}$, $B_0(\cdot)$ das kanonische Auswertungsfunktional auf C_0, d.h. $B_0(\omega)(t) := \omega(t)$ für $0 \leq t \leq 1$ und B^0 das Maß auf C_0, so dass der Prozess $(B_0(t))_{0 \leq t \leq 1}$ eine Brownsche Brücke ist. Die Abbildung

$$S : H_0 \to C_0, \quad S(g)(t) := S_g(t) := \int_0^t g(u)\ d\lambda\!\!\lambda_{|[0,1]}(u),\ 0 \leq t \leq 1$$

bezeichne das Signal und der Parameterraum sei gegeben durch den Hilbertraum

$$H_0 := L_2^0(0,1) := \{g \in L_2([0,1], \mathcal{B}([0,1]), \lambda\!\!\lambda_{|[0,1]}) : \int_0^1 g(u)\ d\lambda\!\!\lambda_{|[0,1]}(u) = 0\}.$$

Zum anderen wird das als Signalerkennungsproblem mit Brownscher Bewegung B als Rauschterm bezeichnete Experiment

$$F = (C, \mathcal{B}(C), \{\mathcal{L}(B + S_h | W_0) : h \in H\}), \tag{3}$$

behandelt, wobei $C := \{f \in C([0,1]) : f(0) = 0\}$, $B(\cdot)$ das kanonische Auswertungsfunktional auf C sei und W_0 das Maß auf C, so dass $(B(t))_{0 \leq t \leq 1}$ eine Brownsche Bewegung ist. Die Abbildung $S : H \to C$ sei gegeben wie oben und der Parameterraum sei der Hilbertraum

$$H := L_2(0,1) := L_2([0,1], \mathcal{B}([0,1]), \lambda\!\!\lambda_{|[0,1]}).$$

Nach einer Einführung in die Theorie der Gauß-Shift-Experimente, bei der wichtige Definitionen und Sätze für das Arbeiten mit solchen Experimenten zusammengefaßt werden, erfolgt die Identifizierung der Signalerkennungsprobleme (2) und (3) als Gauß-Shift-Experimente. Diese treten in der Nichtparametrik als Limesexperimente auf, worauf im zweiten Kapitel eingegangen wird. Der sogennante zentrale Prozess zum Gauß-Shift (2) ist das stochastische Integral nach der Brownschen Brücke

$$L_0(g) := \int g\ dB_0, \quad g \in L_2^0(0,1)$$

bzw. das stochastische Integral nach der Brownschen Bewegung

$$L(h) := \int h\ dB, \quad h \in L_2(0,1)$$

für den Gauß-Shift (3). Trivialerweise liegen die beiden Regressionsmodelle in derselben Äquivalenzklasse der Gauß-Shift-Experimente, da ein isometrischer Isomorphismus zwischen den Parameterräumen $H_0 = L_2^0(0,1)$ und $H = L_2(0,1)$ besteht.

Im zweiten Kapitel wird die L_2-Differenzierbarkeit wiederholt und grundlegende Begriffe wie Tangentialräume eingeführt. Darauf aufbauend wird mit Hilfe des Master-Modells eine $L_2(P_0)$-differenzierbare Kurve $(P_t)_{t \in \mathbb{R}}$ mit der Fußverteilung $P_0 = \lambda\!\!\lambda_{|[0,1]}$ konstruiert, so dass die zugehörigen Experimente lokal asymptotisch normal sind mit dem Regressionsmodell, das die Brownsche Bewegung als Rauschterm verwendet, als Limesexperiment. Dabei bedeutet Lokale Asymptotische Normalität Konvergenz der Loglikelihoodprozesse. Für die

Ableitung des Dichtequotienten $g \in L_2^0([0,1], \lambda\!\!\lambda_{|[0,1]})$ tritt der Kolmogoroff-Smirnoff-Test

$$\varphi_{KS} = \mathbb{1}_{\{\sup_{0\le t\le 1} B_0(t)+\int_0^t g(u)\ du-\rho(t)\ge 0\}}$$

als Limestest auf, wobei $\rho : [0,1] \to \mathbb{R}$ die Überschreitungsfunktion ist, d.h.

$$\lim_{n\to\infty} E_{P^n_{g/\sqrt{n}}}(\varphi^n_{KS}) = E_{P_0}(\varphi_{KS})$$

mit $\varphi^n_{KS} = \mathbb{1}_{\{\sup_{0\le t\le 1}(\sqrt{n}\ (F_n(t)-F_0(t))-\rho(t))\ge 0\}}$. Es ist nämlich nach dem Satz von Donsker bekannt, dass der empirische Prozess $\sqrt{n}\ (F_n(t) - F_0(t))$ unter lokalen Alternativen $P^n_{g/\sqrt{n}}$ gegen das Signalerkennungsproblem mit Brownscher Brücke als Rauschterm konvergiert:

$$\sqrt{n}\ (F_n(t) - F_0(t)) \xrightarrow[P^n_{g/\sqrt{n}}]{\mathcal{D}} Y(t) = B_0(F_0(t)) + \int_0^{F_0(t)} g(u)\ du \text{ für } n \to \infty$$

(vgl. STRASSER [41, Example 82.23] und VAN DER VAART [43]). Dabei bezeichnet F_n die empirische Verteilungsfunktion und F_0 die Verteilungsfunktion von P_0.
Die Parametrisierung der Wahrscheinlichkeitsmaße durch Tangenten $g \in L_2^0(0,1)$ führt zu einer Beschreibung der Hypothesen (1) durch Signale

$$H_1 = \{g \in L_2^0(0,1) : S_g(t) \le 0\} \quad \text{gegen} \quad K_1 = \{g \in L_2^0(0,1) : S_g \ge 0, S_g \ne 0\}$$

mit $S_g(t) := \int_0^t g(u)\ du$ für $g \in L_2^0(0,1)$ und $0 \le t \le 1$. Man testet also, ob das beobachtete verrauschte Signal von einem Parameter $g \in L_2^0(0,1)$ mit $S_g(t) \le 0$ für alle $0 \le t \le 1$ stammt, oder ob es von einem Parameter g mit $S_g(t) \ge 0$ für alle $0 \le t \le 1$ herrührt.

Im dritten Kapitel werden Testverfahren für Signalerkennungsprobleme angegeben, wie z.B. der Test vom Kolmogoroff-Smirnoff-Typ. Für solche Tests werden endlichdimensionale Räume von Alternativen angegeben, auf deren orhogonalem Komplement die Güte flach ist. Diese Räume werden durch die Rademacherfunktionen erzeugt, welche ein Orthonormalsystem für den Hilbertraum $L_2([0,1], \mathcal{B}([0,1]), \lambda\!\!\lambda_{|[0,1]})$ darstellen. Die Wahl der Rademacherfunktionen ist dadurch gerechtfertigt, dass konkave Signale für Alternativen, welche nach den Rademacherfunktionen entwickelt sind, schön handzuhaben sind. Die Güte der einseitigen Signalerkennungstests für solche Alternativen lassen sich dann einfach abschätzen. Der technische Teil dieses Kapitels wird durch die Arbeiten von BISCHOFF et al. beeinflußt. In einer Serie von Artikeln wie BISCHOFF und HASHORVA [4] und BISCHOFF et al. [5], [6], [7] und [8] wurden große Abweichungen für Gütefunktionen von Signalerkennungstests studiert.

Im vierten Kapitel wird die Gütefunktion zweiseitiger Tests für Signalerkennungsprobleme diskutiert. Das Testproblem ist dann gegeben durch

$$H_2 = \{g \in L_2(0,1) : S_g = 0\} \quad \text{gegen} \quad K_2 = \{g \in L_2(0,1) : S_g \neq 0\}.$$

Auch hier wird die Güte für Alternativen abgeschätzt, die in den von den Rademacherfunktionen erzeugten Räumen liegen.

Im fünften und letzten Kapitel erfolgt die statistische Interpretation. Zunächst werden anhand konkreter Beispiele die Bedeutung der konkaven Signale und deren Tangenten für parametrische Teilmodelle diskutiert. Diese entsprechen bestimmten Lokationsmodellen mit streng unimodalen Dichten.
Danach werden endlichdimensionale, von Rademacherfunktionen erzeugte Räume betrachtet. Die endlichen Linearkombinationen h lassen sich durch statistische Funktionale und deren Gradienten interpretieren. Es wird nämlich gezeigt, dass zu jedem solchen Element h eine endliche Linearkombination von Quantilfunktionen existiert, so dass deren Gradient gerade h ist. Dies bedeutet, dass ein Signalerkennungstest, welcher eine hohe Güte auf einer Kugel von Alternativen in h besitzt, für die Erkennung von Abweichungen, die von diesem Funktional herrühren, am empfindlichsten ist.

Zuletzt möchte ich mich an dieser Stelle bei Prof. Dr. Arnold Janssen für seine hervorragende Betreuung, Unterstützung und Motivation bedanken. Ohne ihn wäre diese Arbeit definitiv nicht entstanden. Desweiteren gilt mein Dank auch Prof. Dr. W. Bischoff für die Übernahme und Erstellung des Gutachtens.
Mein besonderer Dank gilt auch den Professoren und den Mitarbeitern des Lehrstuhls für Mathematische Statistik und Wahrscheinlichkeitstheorie am Mathematischen Institut der HHU Düsseldorf für die Unterstützung und großartige Arbeitsatmosphäre.
Ich danke auch meiner Familie und meinen Freunden für die Unterstützung, ganz besonders meiner Schwester Hasret Ünlü, die mich jahrelang sowohl in meiner akademischen Laufbahn als auch in sämtlichen Bereichen meines Lebens ausnahmslos begleitet, unterstützt und mir beigestanden haben. Ohne sie wären sehr viel mehr Dinge nicht möglich gewesen.

Kapitel 1

Ein Signalerkennungsproblem

Zu Anfang dieser Arbeit wird das Signalerkennungsproblem mit Brownscher Brücke und das Signalerkennungsproblem mit Brownscher Bewegung als Rauschterm vorgestellt. Nach einer Einführung in die Theorie der Gauß-Shift-Experimente werden diese Signalerkennungsprobleme als solche Experimente identifiziert.
Dieser Teil der Arbeit ist im Wesentlichen der Dissertation von KUNZ [26], der sich in seiner Arbeit mit den Signalerkennungsproblemen beschäftigt hat, sowie dem Buch von STRASSER [41], §§68-69 entnommen, das die Theorie der Gauß-Shift-Experimente ausführlich darstellt.

1.1 Das Signalerkennungsproblem mit Brownscher Brücke und Brownscher Bewegung als Rauschterm

Eine Motivation für die Beschäftigung mit dem Signalerkennungsproblem ergibt sich folgendermaßen:

Man stelle sich folgende (physikalische) Situation vor:
Ein Sender emittiert ein absolutstetiges Signal $s(\cdot)$, das sich über einen Zeitraum $[0, 1]$ erstreckt. Das Signal wird additiv von einem (zufälligen) Rauschterm $\omega \in C_0 := \{f \in C([0, 1] : f(0) = f(1) = 0\}$ überlagert. Das die Auswahl des Rauschterms steuernde Wahrscheinlichkeitsmaß B^0 ist so gewählt, dass $B_0(t)(\omega) := \omega(t)$ für $0 \leq t \leq 1$ eine Brownsche Brücke zum Start in 0 ist. Der Empfänger registriert das gestörte Signal

$$X_0(t) = s(t) + B_0(t), \quad 0 \leq t \leq 1.$$

Mithilfe des verrauschten Signals $X_0(t)$ werden dann Aussagen über Eigenschaften des Signals gemacht.

Es sei an die Definition der Brownsche Brücke erinnert:

Definition 1.1 (Brownsche Brücke)
Ein reeller Prozess $(B_0(t))_{0\le t\le 1}$ heißt eine Brownsche Brücke, falls folgende Bedingungen gelten:
(i) $B_0(0) = B_0(1) = 0$ fast sicher,
(ii) $(B_0(t))_{0\le t\le 1}$ ist ein zentrierter Gauß-Prozess mit fast sicher stetigen Pfaden,
(iii) $Cov(B_0(t_1), B_0(t_2)) = \min\{t_1, t_2\} - t_1 t_2$ für $0 \le t_1, t_2 \le 1$.

Definition 1.2 (Signalerkennungsproblem mit Brownscher Brücke als Rauschterm)
Seien $H_0 := L_2^0(0,1) := \{g \in L_2([0,1], \mathcal{B}([0,1]), \lambda\!\!\lambda_{|[0,1]}) : \int_0^1 g(u)\ d\lambda\!\!\lambda_{|[0,1]}(u) = 0\}$ und $C_0 := \{f \in C([0,1]) : f(0) = f(1) = 0\}$. Die Abbildung $S : H_0 \to C_0$ sei definiert durch $S(g)(t) := S_g(t) := \int_0^t g(u)\ d\lambda\!\!\lambda_{|[0,1]}(u)$, $0 \le t \le 1$. Es bezeichne für $0 \le t \le 1$

$$\begin{aligned} B_0(t) : C_0 &\to \mathbb{R} \\ \omega &\mapsto \omega(t) \end{aligned}$$

das kanonische Auswertungsfunktional auf C_0. B^0 bezeichne das Maß auf C_0, so dass $(B_0(t))_{0\le t\le 1}$ eine Brownsche Brücke ist. Dann heißt das Experiment

$$E = (C_0, \mathcal{B}(C_0), \{\mathcal{L}(B_0 + S_g | B^0) : g \in H_0\})$$

Signalerkennungsproblem (SEP) mit Brownscher Brücke als Rauschterm.

In einer ganz ähnlichen Situation tritt das folgende Experiment als Limesexperiment auf.

Definition 1.3 (Signalerkennungsproblem mit Brownscher Bewegung als Rauschterm)
Seien $H := L_2(0,1) := L_2([0,1], \mathcal{B}([0,1]), \lambda\!\!\lambda_{|[0,1]})$ und $C := \{f \in C([0,1]) : f(0) = 0\}$. Die Abbildung $S : H \to C$ sei definiert durch $S(h)(t) := S_h(t) := \int_0^t h(u)\ d\lambda\!\!\lambda(u)$, $0 \le t \le 1$. Es bezeichne für $0 \le t \le 1$

$$\begin{aligned} B(t) : C &\to \mathbb{R} \\ \omega &\mapsto \omega(t) \end{aligned}$$

das kanonische Auswertungsfunktional auf C. W_0 bezeichne das Wiener-Maß auf C, so dass $(B(t))_{0\leq t\leq 1}$ eine Brownsche Bewegung ist. Dann heißt das Experiment

$$F = (C, \mathcal{B}(C), \{\mathcal{L}(B + S_h|W_0) : h \in H\})$$

Signalerkennungsproblem (SEP) mit Brownscher Bewegung als Rauschterm.

Auch hier sei an die Definition der Brownschen Bewegung erinnert:

Definition 1.4 (Brownsche Bewegung)
Ein stochastischer Prozess $(B(t))_{t\geq 0}$ heißt reelle Brownsche Bewegung bzgl. einer stochastischen Basis $(\Omega, \mathcal{F}, P, (\mathcal{F}_t)_{t\geq 0})$, falls folgende Bedingungen erfüllt sind:
(i) $B(0) = 0$ fast sicher,
(ii) $(B(t))_{t\geq 0}$ besitzt fast sicher stetige Pfade und ist an $(\mathcal{F}_t)_{t\geq 0}$ adaptiert,
(iii) $(B(t))_{t\geq 0}$ besitzt unabhängige Zuwächse, d.h. $B(t) - B(s)$ ist unabhängig von $\mathcal{F}_s$ für $0 \leq s < t$,
(iv) $\mathcal{L}(B(t) - B(s)|P) = N(0, t - s)$ für $0 \leq s < t$.

1.2 Einführung in die Theorie der Gauß-Shift-Experimente

In diesem Abschnitt werden wichtige Sätze und Definitionen zur Theorie der Gauß-Shift-Experimente bereitgestellt.
Seien im Folgenden H stets ein reeller, separabler Hilbertraum, T ein nicht leerer, beliebiger Parameterraum.
Zunächst wird die Definition äquivalenter Experimente aus dem gemeinsamen Buch von JANSSEN, MILBRODT und STRASSER [23] angegeben, die für das Verständnis der Definition der Gauß-Shift-Experimente notwendig ist.

Definition 1.5 (Äquivalente Experimente)
Zwei Experimente $E = (\Omega_1, \mathcal{A}_1, \{P_t : t \in T\})$ und $F = (\Omega_2, \mathcal{A}_2, \{Q_t : t \in T\})$ heißen äquivalent, i. Z. $E \sim F$, falls

$$\mathcal{L}\left(\left(\frac{dP_s}{dP_t}\right)_{s\in T}\Big|P_t\right) = \mathcal{L}\left(\left(\frac{dQ_s}{dQ_t}\right)_{s\in T}\Big|Q_t\right) \ \forall t \in T.$$

Interpretation: Die Äquivalenz beinhaltet die Gleichheit der Verteilungen der Likelihoodquotienten der Experimente E und F. Alle auf dem Likelihoodprozess beruhenden statistischen Eigenschaften von E und F sind gleich, z.B. das Verhalten der

Neyman-Pearson-Tests. Deshalb sagt man, dass E und F "gleich viel Informationen" besitzen, vgl. JANSSEN [14].

Zuerst werden Gauß-Shift-Experimente auf endlichdimensionalen Hilberträumen definiert. Dazu seien $(H, \langle\cdot,\cdot\rangle)$ ein endlichdimensionaler Hilbertraum und $\mathcal{B}(H)$ seine Borel-σ-Algebra. N_H bezeichne das Standard-Gauß-Maß auf $(H, \mathcal{B}(H))$ und ε_h das Einpunktmaß in $h \in H$.

Definition 1.6 (Gauß-Shift-Experiment)
Ein Gauß-Shift-Experiment auf einem endlichdimensionalen Hilbertraum $(H, \langle\cdot,\cdot\rangle)$ *ist ein Experiment* $E = (\Omega, \mathcal{A}, \{P_h : h \in H\})$*, welches äquivalent zu dem Experiment* $(H, \mathcal{B}(H), \{N_H * \varepsilon_h : h \in H\})$ *ist.*

Von größerem Interesse für diese Arbeit sind aber Gauß-Shift-Experimente auf unendlichdimensionalen Hilberträumen.

Definition 1.7 (Gauß-Shift-Experiment)
Ein Gauß-Shift-Experiment auf einem separablen Hilbertraum $(H, \langle\cdot,\cdot\rangle)$ *ist ein Experiment* $E = (\Omega, \mathcal{A}, \{P_h : h \in H\})$*, dessen Restriktion* $E_L = (\Omega, \mathcal{A}, \{P_h : h \in L\})$ *auf jeden endlichdimensionalen linearen Teilraum* $L \subseteq H$ *äquivalent zu* $(L, \mathcal{B}(L), \{N_L * \varepsilon_h : h \in L\})$ *ist.*

Unendlichdimensionale Gauß-Shift-Experimente spielen eine wichtige Rolle in der asymptotischen Statistik nichtparametrischer Probleme. Sie treten nämlich als Limesexperimente auf.

Definition 1.8 (Linearer Prozess)
Seien $(H, \langle\cdot,\cdot\rangle)$ *ein separabler Hilbertraum und* $(\Omega, \mathcal{A})$ *ein Messraum. Eine lineare Funktion* $Z : H \to \mathcal{L}(\Omega, \mathcal{A})$ *heißt ein Linearer Prozess, wobei* $\mathcal{L}(\Omega, \mathcal{A})$ *die Menge aller reellen* $\mathcal{A} - \mathcal{B}(\mathbb{R})$*-messbaren Funktionen bezeichnet.*

Definition 1.9 (Standard-Gauß-Prozess)
Ein stochastischer Prozess $L = (\Omega, \mathcal{A}, P, (L(h))_{h \in H})$ *auf einem Hilbertraum* $(H, \langle\cdot,\cdot\rangle)$ *heißt ein Standard-Gauß-Prozess, falls* L *ein zentrierter Gauß-Prozess ist mit der Kovarianzstruktur*

$$Cov(L(h_1), L(h_2)) = \langle h_1, h_2\rangle \ \forall h_1, h_2 \in H.$$

Zwischen den beiden Begriffen Gauß-Shift und Standard-Gauß-Prozess besteht ein enger Zusammenhang, den der folgende Satz beschreibt:

Satz 1.10 (**Girsanov-Formel**)
Ein Experiment $E = (\Omega, \mathcal{A}, \{P_h : h \in H\})$ ist genau dann ein Gauß-Shift, wenn ein stochastischer Prozess $(L(h))_{h \in H}$ existiert, der ein Standard-Gauß-Prozess unter P_0 ist und für den gilt

$$\frac{dP_h}{dP_0} = \exp\left(L(h) - \frac{1}{2}\|h\|^2\right) \ \ \forall h \in H. \tag{1.1}$$

Beweis. Vgl. STRASSER [41, Theorem 69.4]. ■

Aufgrund der zentralen Bedeutung dieses Prozesses definiert man:

Definition 1.11 (**Zentraler Prozess**)
Sei E ein Gauß-Shift. Der Standard-Gauß-Prozess $(L(h))_{h \in H}$, der im Sinne von Satz 1.10 mit E zusammenhängt, heißt Zentraler Prozess von E.

Bemerkung 1.12
Die Existenz von Gauß-Shift-Experimenten ist immer gegeben, da man ausgehend von einem Standard-Gauß-Prozess die Dichtequotienten wie in (1.1) definieren kann:
Seien $(H, \langle \cdot, \cdot \rangle)$ ein Hilbertraum und $(L(h))_{h \in H}$ ein Standard-Gauß-Prozess auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$. Setzt man $P_0 := P$ und definiert man für jedes $h \in H$

$$f_h := \exp\left(L(h) - \frac{\|h\|^2}{2}\right),$$

so erhält man

$$\begin{aligned} \int_\Omega f_h \, dP_0 &= \int_{\mathbb{R}} \exp\left(x - \frac{\|h\|^2}{2}\right) \, d\mathcal{L}(L(h)|P_0)(x) \\ &= \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi\|h\|^2}} \exp\left(x - \frac{\|h\|^2}{2}\right) \exp\left(-\frac{x^2}{2\|h\|^2}\right) \, dx \\ &= \int_{\mathbb{R}} dN(\|h\|^2, \|h\|^2)(x) = 1. \end{aligned}$$

Definiert man für jedes $h \in H$ das eindeutig bestimmte Wahrscheinlichkeitsmaß P_h durch $P_h := f_h P_0$, so ist $(\Omega, \mathcal{A}, \{P_h : h \in H\})$ ein Gauß-Shift. □

Satz 1.13
Sei $(H, \langle \cdot, \cdot \rangle)$ ein Hilbertraum. $L : H \to L_2(\Omega, \mathcal{A}, P)$ ist genau dann ein Standard-Gauß-Prozess, wenn L linear ist und $\mathcal{L}(L(h)|P) = N(0, \|h\|^2)$ für alle $h \in H$.

Beweis.
(1) Sei L ein Standard-Gauß-Prozess. Dann ist nach Definition $\mathcal{L}(L(h)|P) = N(0, \|h\|^2)$ für alle $h \in H$. Um die Linearität zu zeigen, wählt man beliebige $(\alpha_1, ..., \alpha_n) \in \mathbb{R}^n$ und $(h_1, ..., h_n) \in H^n, n \in \mathbb{N}$, und erhält

$$\begin{aligned}
\|\sum_{i=1}^n \alpha_i L(h_i)\|_{L_2}^2 &= Var_P(\sum_{i=1}^n \alpha_i L(h_i)) = \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j Cov_P(L(h_i), L(h_j)) \\
&\overset{\text{Definition1.9}}{=} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j \langle h_i, h_j \rangle = \langle \sum_{i=1}^n \alpha_i h_i, \sum_{j=1}^n \alpha_j h_j \rangle \\
&= \|\sum_{i=1}^n \alpha_i h_i\|^2.
\end{aligned}$$

Aus $\sum_{i=1}^n \alpha_i h_i = 0$ folgt also $\sum_{i=1}^n \alpha_i L(h_i) = 0$ P-f.s. Insbesondere für $\alpha_n := -1$ und $h_n := \sum_{i=1}^{n-1} \alpha_i h_i$ ergibt sich dann $\sum_{i=1}^{n-1} \alpha_i L(h_i) = L(\sum_{i=1}^{n-1} \alpha_i h_i)$ P-f.s.
(2) Sei L linear und $\mathcal{L}(L(h)|P) = N(0, \|h\|^2)$ für alle $h \in H$. Dann hat $\mathcal{L}(L(h)|P)$ die Fouriertransformierte $\varphi_{L(h)}(t) = \exp(-\frac{1}{2}\|h\|^2 t^2)$ für $t \in \mathbb{R}$. Man wähle wieder beliebige $\alpha := (\alpha_1, ..., \alpha_n) \in \mathbb{R}^n$ und $(h_1, ..., h_n) \in H^n, n \in \mathbb{N}$. Definiere $\tilde{L} := (L(h_1), ..., L(h_n))$. Dann erhält man

$$\begin{aligned}
\varphi_{\tilde{L}}(\alpha) &= E\left(\exp\left(i\sum_{i=1}^n \alpha_i L(h_i)\right)\right) \overset{L \text{ linear}}{=} E\left(\exp\left(iL\left(\sum_{i=1}^n \alpha_i h_i\right)\right)\right) \\
&= \varphi_{L(\sum_{i=1}^n \alpha_i h_i)}(1) = \exp\left(-\frac{\|\sum_{i=1}^n \alpha_i h_i\|^2}{2}\right) \\
&= \exp\left(-\frac{1}{2}\sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j \langle h_i, h_j \rangle\right).
\end{aligned}$$

Mit dem Eindeutigkeitssatz für Fouriertransformierte folgt hieraus, dass $(L(h_1), ..., L(h_n))$ eine gemeinsame Normalverteilung besitzt mit Erwartungswertvektor 0 und Kovarianzmatrix $(\langle h_i, h_j \rangle)_{1 \le i,j \le n}$. ∎

Satz 1.14
Sei $E = (\Omega, \mathcal{A}, \{P_h : h \in H\})$ Gauß-Shift mit dem zentralen Prozess $(L(h))_{h \in H}$. Dann gilt

$$\mathcal{L}(L(h_1)|P_{h_2}) = N(\langle h_1, h_2 \rangle, \|h_1\|^2) \quad \forall h_1, h_2 \in H.$$

Beweis. Die Behauptung folgt aufgrund der Translationsinvarianz von Gauß-Shift-Experimenten, vgl. STRASSER [41, Theorem 69.10]. ∎

1.3 Das Signalerkennungsproblem als Gauß-Shift

Das Signalerkennungsproblem $E = (C_0, \mathcal{B}(C_0), \{\mathcal{L}(B_0 + S_g | B^0) : g \in H_0\})$ mit dem Parameterraum $H_0 = L_2^0([0,1], \lambda\!\!\lambda_{|[0,1]})$ wird nun als Gauß-Shift identifiziert:

Satz 1.15
Seien $H_0 = L_2^0([0,1], \lambda\!\!\lambda_{|[0,1]}), C_0 = \{f \in C([0,1]) : f(0) = f(1) = 0\}$ und $P_0 := \mathcal{L}(S|N_{H_0})$. Dann existiert ein eindeutig bestimmter Standard-Gauß-Prozess $(L_0(g))_{g \in H_0}$ auf $(C_0, \mathcal{B}(C_0), P_0)$, so dass die Dichteformel (1.1) für $P_g := P_0 * \varepsilon_{S_g}$ gilt und

$$E = (C_0, \mathcal{B}(C_0), \{P_g : g \in H_0\})$$

ein Gauß-Shift mit Parameterraum H_0 und zentralem Prozeß $(L_0(g))_{g \in H_0}$ ist.

Beweis. Vgl. STRASSER [41, Theorem 70.4, 70.6 und 70.8]. ■

Nach Satz 1.15 ist also E ein Gauß-Shift. Sein zentraler Prozess $(L(g))_{g \in H_0}$ ist nach STRASSER [41, Example 70.7(2)] gegeben durch

$$L_0(g) := \int g \; dB_0 = \int_0^1 g(t) \; dB_0(t), \quad g \in H_0.$$

Dieser Prozess ist bekannt als das stochastische Itô-Integral nach der Brownschen Brücke.

Das SEP $F = (C, \mathcal{B}(C), \{\mathcal{L}(B + S_h | W_0) : h \in H\})$ mit $H = L_2([0,1], \lambda\!\!\lambda_{|[0,1]})$ als Parameterraum ist ebenfalls ein Gauß-Shift:

Satz 1.16
Seien $H = L_2([0,1], \lambda\!\!\lambda_{|[0,1]})$, $C = \{f \in C([0,1]) : f(0) = 0\}$ und $Q_0 := \mathcal{L}(S|N_H)$. Dann existiert ein eindeutig bestimmter Standard-Gauß-Prozess $(L(h))_{h \in H}$ auf $(C, \mathcal{B}(C), Q_0)$, so dass die Dichteformel (1.1) für $Q_h := Q_0 * \varepsilon_{S_h}$ gilt und

$$F = (C, \mathcal{B}(C), \{Q_h : h \in H\})$$

ein Gauß-Shift mit Parameterraum H und zentralem Prozess $(L(h))_{h \in H}$ ist.

Beweis. Vgl. STRASSER [41, Theorem 70.3, 70.6 und 70.8]. ■

Nach Satz 1.16 ist F ebenso ein Gauß-Shift. Sein zentraler Prozess $(L(h))_{h \in H}$ ist nach STRASSER [41, Example 70.7(1)] gegeben durch

$$L(h) := \int h \; dB = \int_0^1 h(t) \; dB(t), \quad h \in H.$$

Dieser Prozess ist das stochastische Itô-Integral nach der Brownschen Bewegung oder kurz, Wiener Integral.
Für eine ausführliche Darstellung der Theorie stochastischer Integrale wird auf IRLE [12] und KARATZAS/SHREVE [25] verwiesen.

Man beachte, dass das SEP mit Brownscher Bewegung als Rauschterm verwandt ist mit dem SEP mit Brownscher Brücke als Rauschterm, siehe Abschnitt 1.4.

Bemerkung 1.17
Das Maß B^0 auf C_0 ist nach Definition 1.2 so gewählt, dass $(B_0(t))_{0\leq t\leq 1}$ eine Brownsche Brücke ist. Nach dem Satz von Girsanov [25, Section 3.5.A,Theorem 5.1] bleibt der um S_g geshiftete Prozess $B_0 - S_g$ unter P_g eine Brownsche Brücke, d.h. $\mathcal{L}(B_0 - S_g|P_g) = \mathcal{L}(B_0|P_0)$ für $g \in L_2^0(0,1)$. Analog bleibt im Fall der Brownschen Bewegung (vgl. Definition 1.3) der Prozess $B - S_h$ unter Q_h eine Brownsche Bewegung, d.h. $\mathcal{L}(B - S_h|Q_h) = \mathcal{L}(B|Q_0)$ für $h \in L_2(0,1)$. □

Seien $G = (\Omega, \mathcal{A}, \{P_h : h \in H\})$ ein Gauß-Shift mit dem Hilbertraum $H \subseteq L_2(0,1)$ als Parameterraum und F_0 eine stetige Verteilungsfunktion[1] auf $\mathbb{R}$. Dann kann man einen Signalprozess allgemeiner definieren.

Definition 1.18 (Signalprozess)
Ein Prozess $Y(t) : (\Omega, \mathcal{A}) \to \mathbb{R}, t \in \mathbb{R}$, heißt Signalprozess bzgl. einem zugrunde liegenden Gauß-Shift $G = (\Omega, \mathcal{A}, \{P_h : h \in H\})$, falls folgende Bedingungen erfüllt sind.
(a) Es existiert ein fester zentrierter Gauß-Prozess $(Z(s))_{s\in[0,1]}$ mit stetigen Pfaden, so dass

$$(Y(t))_{t\in\mathbb{R}} \stackrel{\mathcal{D}}{=} (Z(F_0(t)))_{t\in\mathbb{R}}$$

unter P_0.
(b) Für jedes $h \in H$ erhält man Gleichheit in Verteilung

$$\begin{aligned}(Y(t))_{t\in\mathbb{R}} &\stackrel{\mathcal{D}}{=} \left(Z(F_0(t)) + \int_{-\infty}^{t} h \circ F_0(x)\ dF_0(x)\right)_{t\in\mathbb{R}} \\ &= \left(Z(F_0(t)) + \int_0^{F_0(t)} h(u)\ du\right)_{t\in\mathbb{R}}, \qquad (1.2)\end{aligned}$$

falls $Y(t) : (\Omega, \mathcal{A}, P_h) \to \mathbb{R}$.

[1] Für Regresssionsmodelle ergeben sich analog verwendete SEP. Dabei übernimmt F_0 die Rolle eines asymptotischen Designs, vgl. BISCHOFF [3, Section 3].

Man beschränkt sich häufig auf die Gleichverteilung $F_0(x) = x, x \in [0,1]$, denn allgemeinere Modelle mit stetiger Verteilungsfunktion F_0 haben dasselbe statistische Verhalten bis auf Quantiltransformationen.

Bemerkung 1.19
(a) Im Fall $Z(t) = B_0(t)$ erhält man für (1.2) den Signalprozess der Brownschen Brücke.
(b) Im Fall $Z(t) = B(t)$ erhält man für (1.2) den Signalprozess der Brownschen Bewegung.

1.4 Äquivalenz der Signalprozesse

Man betrachte wieder das Signalerkennungsmodell mit Brownscher Brücke als Rauschterm

$$X_0(t) = B_0(t) + \int_0^t g(u)\ du \quad \text{für} \quad g \in L_2^0(0,1), 0 \le t \le 1 \tag{1.3}$$

und das Signalerkennungsmodell mit Brownscher Bewegung als Rauschterm

$$X(t) = B(t) + \int_0^t h(u)\ du \quad \text{für} \quad h \in L_2(0,1), 0 \le t \le 1. \tag{1.4}$$

Die in (1.3) und (1.4) gegebenen Regressionsmodelle sind äquivalent im Sinne von Definition 1.5. Durch die Abbildung

$$\begin{aligned} R \quad &: \quad L_2^0(0,1) \to L_2(0,1) \\ R(g) \quad &: \quad t \mapsto g(t) + \frac{1}{1-t} \int_0^t g(u)\ du \end{aligned}$$

wird nämlich ein isometrischer Isomorphismus definiert, vgl. EFRON und JOHNSTONE [9], RITOV und WELLNER [37]. Dabei wird $R(g)$ **Hazardratenableitung** genannt und tritt bei der Analyse von Überlebenszeiten auf, vgl. JANSSEN und MILBRODT [22] sowie JANSSEN [14]. Die zugehörige Umkehrabbildung R^{-1} ist gegeben durch

$$\begin{aligned} R^{-1} \quad &: \quad L_2(0,1) \to L_2^0(0,1) \\ R^{-1}(h) \quad &: \quad t \mapsto h(t) - \int_0^t \frac{h(u)}{1-u}\ du, \end{aligned}$$

was im Folgenden gezeigt wird. Dazu seien $h \in L_2(0,1)$ und $0 < t < 1$. Mit dem Satz von

Fubini folgt

$$
\begin{aligned}
& R(R^{-1}(h(t))) \\
= \;& R^{-1}(h(t)) + \frac{\int_0^t R^{-1}(h)(u)\;du}{1-t} \\
= \;& h(t) - \int_0^t \frac{h(u)}{1-u}\;du + \frac{\int_0^t h(u)}{1-t}\;du - \frac{1}{1-t}\int_0^t \left(\int_0^u \frac{h(v)}{1-v}\;dv\right)\;du \\
= \;& h(t) - \int_0^t \frac{h(u)}{1-u}\;du + \frac{\int_0^t h(u)}{1-t}\;du - \frac{1}{1-t}\int_0^1 \int_0^1 \mathbb{1}_{(0,t)}(u)\mathbb{1}_{(0,u)}(v)\frac{h(v)}{1-v}\;dv\;du \\
= \;& h(t) - \int_0^t \frac{h(u)}{1-u}\;du + \frac{\int_0^t h(u)}{1-t}\;du - \frac{1}{1-t}\int_0^1 \frac{h(v)}{1-v}\left(\int_0^1 \mathbb{1}_{(0,t)\cap(v,1))}(u)\;du\right)\;dv \\
= \;& h(t) - \int_0^t \frac{h(u)}{1-u}\;du + \frac{\int_0^t h(u)}{1-t}\;du - \frac{1}{1-t}\int_0^1 \frac{h(v)}{1-v}\;(\max(t,v)-v)\;dv \\
= \;& h(t) - \int_0^t \frac{h(u)}{1-u}\;du + \frac{\int_0^t h(u)}{1-t}\;du - \frac{1}{1-t}\int_0^t \frac{h(v)}{1-v}(t-v)\;dv \\
= \;& h(t) - \frac{1}{1-t}\int_0^t \underbrace{\frac{h(u)(1-t) - h(u)(1-u) + h(u)(t-u)}{1-u}}_{=0}\;du \\
= \;& h(t).
\end{aligned}
$$

Analog rechnet man nach, dass $R^{-1}(R(g)) = g$ ist für alle $g \in L_2^0(0,1), 0 < t < 1$:

$$
\begin{aligned}
& R^{-1}(R(g(t))) \\
= \;& R(g(t)) - \int_0^t \frac{R(g(u))}{1-u}\;du \\
= \;& g(t) + \frac{\int_0^t g(u)\;du}{1-t} - \int_0^t \frac{g(u)}{1-u}\;du - \int_0^t \frac{\int_0^u g(v)\;dv}{(1-u)^2}\;du \\
= \;& g(t) + \frac{\int_0^t g(u)\;du}{1-t} - \int_0^t \frac{g(u)}{1-u}\;du - \int_0^1 g(v)\left(\int_0^1 \mathbb{1}_{(v,\max(t,v))}\frac{1}{(1-u)^2}\;du\right)\;dv \\
= \;& g(t) + \frac{\int_0^t g(u)\;du}{1-t} - \int_0^t \frac{g(u)}{1-u}\;du - \int_0^1 g(v)\left(\frac{1}{1-\max(t,v)} - \frac{1}{1-v}\right)\;dv \\
= \;& g(t) + \frac{\int_0^t g(u)\;du}{1-t} - \int_0^t \frac{g(u)}{1-u}\;du - \int_0^t g(v)\left(\frac{1}{1-t} - \frac{1}{1-v}\right)\;dv \\
= \;& g(t).
\end{aligned}
$$

Ferner kann mit Hilfe der Doob-Meyer Zerlegung von B_0 gegeben durch

$$B_0(t) \overset{\mathcal{D}}{=} B(t) - \int_0^t \frac{B_0(s)}{1-s}\ ds,\ \ 0 \le t \le 1,$$

wobei B wieder eine Brownsche Bewegung ist (vgl. KARATZAS und SHREVE [25][Section 5.6.B]), eine Transformation zwischen den Regressionsmodellen (1.3) und (1.4) hergestellt werden.

Satz 1.20
Gegeben seien die Tangente $g \in L_2^0(0,1)$ und die Hazardratenableitung $h := R(g) \in L_2(0,1)$. Dann erhält man eine Gleichheit in Verteilung

$$X(t) \overset{\mathcal{D}}{=} X_0(t) + \int_0^t \frac{X_0(s)}{1-s}\ ds,\ \ 0 \le t \le 1,$$

wobei X_0 und X die Prozesse aus (1.3) bzw. (1.4) sind.

Beweis. Vgl. JANSSEN und KUNZ [19].
Es ist

$$\begin{aligned}
X_0(t) + \int_0^t \frac{X_0(s)}{1-s}\ ds &= B_0(t) + \int_0^t g(u)\ du + \int_0^t \frac{B_0(s) + \int_0^s g(u)\ du}{1-s}\ ds \\
&= B_0(t) + \int_0^t \frac{B_0(s)}{1-s}\ ds + \int_0^t g(u)\ du + \int_0^t \frac{\int_0^s g(u)\ du}{1-s}\ ds \\
&\overset{\mathcal{D}}{=} B(t) + \int_0^t R(g)(s)\ ds \\
&= X(t).
\end{aligned}$$

■

Bemerkung 1.21
Im Falle der Gleichverteilung $F_0(x) = x, x \in [0,1]$ genügt das Studium der Signalprozesse mit der Parametermenge $[0,1]$. Bei einer beliebigen stetigen Verteilungsfunktion F_0 kann der Parameterraum $\mathbb{R}$ betrachtet werden. □

Kapitel 2

Bedeutung der Limesexperimente für asymptotische Testprobleme

In diesem Kapitel wird eine $L_2(P_0)$-differenzierbare Kurve $(P_t)_{t\in\mathbb{R}}$ von Wahrscheinlichkeitsmaßen mit der Fußverteilung $P_0 = \lambda\!\!\lambda_{|[0,1]}$ konstruiert, so dass die zugehörige Experimentenfolge schwach gegen das Signalerkennungsproblem mit der Brownschen Brücke als Rauschterm konvergiert. Ferner wird gezeigt, dass zu jeder Tangente $g \in L_2^0([0,1], \lambda\!\!\lambda_{|[0,1]})$ eine Testfolge existiert, die als Limestest den Kolmogoroff-Smirnoff-Typ Test besitzt. Es wird zum großen Teil mit STRASSER [41], §75, §80 und OSTROVSKI [33] gearbeitet.
Zunächst werden die L_2-Differenzierbarkeit wiederholt und Tangentialräume eingeführt. Eine ausführliche Darstellung der Theorie zur L_2-Differenzierbarkeit und Tangentialräume sowie Beispiele findet man in WITTING [44] und PFANZAGL [35].

2.1 L_2-Differenzierbarkeit und Tangentialräume

Seien $(\Omega, \mathcal{A})$ ein Messraum, $\mathcal{M}_1(\Omega, \mathcal{A})$ die Menge aller Wahrscheinlichkeitsmaße auf $(\Omega, \mathcal{A})$ und $\mathcal{P} \subset \mathcal{M}_1(\Omega, \mathcal{A}), \mathcal{P} \neq \emptyset$ eine nichtparametrische Familie von Wahrscheinlichkeitsmaßen.

Definition 2.1 (**L_2-Differenzierbarkeit**)
Es seien $\varepsilon > 0$ und $\mathcal{P} \subset \mathcal{M}_1(\Omega, \mathcal{A})$ eine nichtleere Teilmenge. Eine Abbildung $(-\varepsilon, \varepsilon) \to \mathcal{P}$, $t \mapsto P_t$ heißt eine $L_2(P_0)$-differenzierbare Kurve in $\mathcal{P}$ mit Tangente $g \in L_2(P_0)$, falls für $t \to 0$ gilt

$$\left\| 2\left(\left(\frac{dP_t}{dP_0}\right)^{\frac{1}{2}} - 1\right) - tg \right\|_{L_2(P_0)} = o(|t|) \quad \textit{und} \quad P_t\left(\left\{\frac{dP_t}{dP_0} = \infty\right\}\right) = o(t^2).$$

Bemerkung 2.2

(a) Die Tangente $g \in L_2(P_0)$ ist P_0-f.s. eindeutig bestimmt. Jede Version von g bezeichnet man deshalb als Tangente an die $L_2(P_0)$-differenzierbare Kurve $t \mapsto P_t$ in $\mathcal{P}$.

(b) Für die Tangente $g \in L_2(P_0)$ gilt

$$\int g \, dP_0 = 0,$$

vgl. WITTING [44, Satz 1.190, Hilfssatz 1.178]. □

Bemerkung 2.2(b) rechtfertigt folgende Definition.

Definition 2.3 (**Tangentenraum**)

Seien $P \in \mathcal{M}_1(\Omega, \mathcal{A})$ ein Wahrscheinlichkeitsmaß, $t \mapsto P_t$ eine $L_2(P_0)$-differenzierbare Kurve mit Tangente $g \in L_2(P_0)$ und $P_0 = P$. Der Vektorraum

$$L_2^0(P_0) = \left\{ h \in L_2(P) : \int h \, dP = 0 \right\}$$

wird als Tangentenraum bezeichnet.

Bezeichnung 2.4

Die Menge der Tangenten eines Wahrscheinlichkeitsmaßes $P \in \mathcal{P}$ bzgl. der Familie $\mathcal{P}$ ist gegeben durch

$$K(P, \mathcal{P}) := \{g \in L_2^0(P) : \text{es existiert eine } L_2(P_0)\text{-differenzierbare Kurve } t \mapsto P_t \text{ in } \mathcal{P} \\ \text{mit Tangente } g \text{ und } P_0 = P\}.$$

Definition 2.5 (**Tangentialraum**)

Der $L_2(P)$-Abschluss der linearen Hülle eines Tangentenkegels $K(P, \mathcal{P})$ ist gegeben durch

$$T(P, \mathcal{P}) := \overline{span K(P, \mathcal{P})}$$

und wird als Tangentialraum des Wahrscheinlichkeitsmaßes $P \in \mathcal{P}$ bzgl. der Familie $\mathcal{P}$ bezeichnet.

Der Tangentialraum spiegelt die lokale Struktur eines statistischen Experiments wieder. Für jedes Wahrscheinlichkeitsmaß $P \in \mathcal{P}$ ist der Tangentialraum $T(P, \mathcal{P})$ als abgeschlossener Unterraum von $L_2(P)$ wieder ein Hilbertraum.

2.2 Das Signalerkennungsproblem als Limesexperiment

Sei $\mathcal{P} \subset \mathcal{M}_1(\Omega, \mathcal{A})$ eine nichtparametrische Familie von Wahrscheinlichkeitsmaßen und $P_0 \in \mathcal{P}$. Um das Signalerkennungsproblem mit Brownscher Brücke als Rauschterm als Limesexperiment darzustellen, fehlt noch die Zuordnung zwischen den Wahrscheinlichkeitsmaßen aus $\mathcal{P}$ und den Tangenten aus $T(P_0, \mathcal{P})$. Deshalb wird eine semiparametrische Familie $\tilde{\mathcal{P}}(P_0)$ von Wahrscheinlichkeitsmaßen konstruiert, die an der Stelle P_0 den Tangentialraum $T(P_0, \mathcal{P})$ besitzt.

Satz 2.6 (Konstruktion $L_2(P_0)$-differenzierbarer Kurven)
Seien $P \in \mathcal{M}_1(\Omega, \mathcal{A})$ ein Wahrscheinlichkeitsmaß und $g \in L_2^0(P)$ eine Tangente. Dann existiert eine $L_2(P_0)$-differenzierbare Kurve

$$\mathbb{R} \to \mathcal{M}_1(\Omega, \mathcal{A}), \ t \to P_{tg}$$

mit Tangente g und $P_0 = P$. Außerdem gilt $P_{tg} \ll P_0$ für jedes $t \in \mathbb{R}$.

Beweis. Vgl. OSTROVSKI [33, Satz 4.5]. ■

Die Konstruktion solcher $L_2(P_0)$-differenzierbarer Kurven erfolgt mit Hilfe des Master-Modells aus JANSSEN [18]:
Man setze für $g \in L_2^0(P)$

$$f_g := \left(1 + \frac{g}{2}\right)^2 \ c(g)^{-1}, \quad c(g) := \int \left(1 + \frac{g}{2}\right)^2 \ dP_0 = 1 + \frac{1}{4} \int g^2 \ dP_0.$$

Dann wird durch

$$P_g := f_g P_0$$

ein Wahrscheinlichkeitsmaß auf $(\Omega, \mathcal{A})$ definiert, so dass die einparametrige Teilfamilie $t \mapsto P_{tg}$ an der Stelle P_0 die Tangente g hat. Insbesondere erhält man $P_0 = f_0 P = P$.

Definition 2.7 (Semiparametrische Familie)
Die semiparametrische Familie $\tilde{\mathcal{P}}(P_0)$ wird definiert durch

$$\tilde{\mathcal{P}}(P_0) := \{P_{tg} \in \mathcal{M}_1(\Omega, \mathcal{A}) : g \in T(P_0, \mathcal{P}), t \in \mathbb{R}\},$$

wobei $t \mapsto P_{tg}$ eine $L_2(P_0)$-differenzierbare Kurve ist mit Tangente g. Die Wahrscheinlichkeitsmaße P_{tg} sind durch den Dichtequotienten

$$\frac{dP_{tg}}{dP_0} = \left(1 + \frac{tg}{2}\right)^2 \ c(tg)^{-1}$$

gegeben mit $c(tg) := 1 + \frac{t^2}{4} \int g^2 \ dP_0$ für $g \in T(P_0, \mathcal{P})$ und $t \in \mathbb{R}$ (vgl. Satz 2.6).

Man beachte, dass die Gleicheit

$$\tilde{\mathcal{P}}(P_0) = \{P_g \in \mathcal{M}_1(\Omega, \mathcal{A}) : g \in T(P_0, \mathcal{P})\}$$

gilt, da $T(P_0, \mathcal{P})$ ein Vektorraum ist.

Nun werden grundlegende Begriffe in der Theorie der asymptotischen Statistik vorgestellt. Dabei wird LAN (Lokale Asymptotische Normalität) als schwache Konvergenz von Experimenten gegen einen Gauß-Shift neu definiert.

Definition 2.8 (**Schwache Konvergenz von Experimenten**)
Seien Θ eine nichtleere Menge und $(\Theta_n)_{n\in\mathbb{N}}$ eine Folge von Teilmengen von Θ mit $\Theta_n \uparrow \Theta$ für $n \to \infty$. Eine Folge von Experimenten $E_n = (\Omega_n, \mathcal{A}_n, \{P_{n,h} : h \in \Theta_n\}), n \in \mathbb{N}$, heißt schwach konvergent gegen ein Experiment $E = (\Omega, \mathcal{A}, \{P_h : h \in \Theta\})$, falls

$$\mathcal{L}\left(\left(\frac{dP_{n,h}}{dP_{n,s}}\right)_{h\in Z}\middle| P_{n,s}\right) \to \mathcal{L}\left(\left(\frac{dP_h}{dP_s}\right)_{h\in Z}\middle| P_s\right)$$

für jede endliche Teilmenge $Z \subset \Theta$ und für jedes $s \in \Theta$.

Definition 2.9 (**LAN**)
Seien H ein Hilbertraum und $(H_n)_{n\in\mathbb{N}}$ eine Folge von Teilmengen von H mit $H_n \uparrow H$ für $n \to \infty$. Eine Folge von Experimenten $E_n = (\Omega_n, \mathcal{A}_n, \{P_{n,h} : h \in H_n\}), n \in \mathbb{N}$, heißt Lokal Asymptotisch Normal (LAN), falls sie schwach gegen ein Gauß-Shift-Experiment $E = (\Omega, \mathcal{A}, \{P_h : h \in H\})$ konvergiert.

Satz 2.10 (**Hinreichendes und notwendiges Kriterium für LAN**)
Eine Folge von Experimenten $(E_n)_{n\in\mathbb{N}}$ ist genau dann lokal asymptotisch normal, wenn der stochastische Prozess $(L_n(h))_{h\in H_n}$, $n \in \mathbb{N}$, definiert durch

$$\frac{dP_{n,h}}{dP_{n,0}} = \exp\left(L_n(h) - \frac{\|h\|^2}{2}\right), \ h \in H_n, n \in \mathbb{N}$$

die folgenden Bedingungen erfüllt:

(1) $\mathcal{L}(L_n(h)|P_{n,0}) \xrightarrow[n\to\infty]{w} N(0, \|h\|^2)$ *für* $h \in H$

(2) $\alpha L_n(h_1) + \beta L_n(h_2) - L_n(\alpha h_1 + \beta h_2) \xrightarrow[n\to\infty]{} 0$ $P_{n,0}$*-stochastisch*

für $\alpha, \beta \in \mathbb{R}, h_1, h_2 \in H$.

Beweis. Vgl. STRASSER [41, Theorem 80.2] ■

Satz 2.11
Sei für $n \in \mathbb{N}$ *und* $g \in T(P_0, \mathcal{P})$

$$P_{n,g} := P^n_{\frac{1}{\sqrt{n}}g} \quad \textit{mit} \quad P_{\frac{1}{\sqrt{n}}g} \in \tilde{\mathcal{P}}(P_0).$$

Außerdem sei $E_n := (\Omega^n, \mathcal{A}^n, \{P_{n,g} : g \in T(P_0, \mathcal{P})\})$ *für* $n \in \mathbb{N}$. *Dann ist die Folge* $(E_n)_{n \in \mathbb{N}}$ *der Experimente lokal asymptotisch normal.*

Beweis. Vgl. OSTROVSKI [33, Satz 4.12]. ■

Bemerkung 2.12
Nach Satz 2.11 konvergiert die Folge $(E_n)_{n \in \mathbb{N}}$ mit $E_n := (\Omega^n, \mathcal{A}^n, \{P_{n,g} : g \in T(P_0, \mathcal{P})\})$ gegen einen Gauß-Shift auf $T(P_0, \mathcal{P}) \subset L_2^0(P_0)$, welcher ein Teilexperiment von unserem Signalerkennungsproblem mit Brownscher Brücke als Rauschterm ist.
Es gibt aber noch eine andere Konstruktion von Wahrscheinlichkeitsmaßen, wo das Signalerkennungsproblem als Limesexperiment der zugehörigen Experimente auftritt (vgl. STRASSER [41, Example 80.4]).
Seien $H_0 = L_2^0([0,1], \mathcal{B}([0,1]), \lambda\!\!\lambda_{|[0,1]})$ und $M := \{g \in H_0 : \int_0^1 g^2 \, d\lambda\!\!\lambda_{|[0,1]} \leq 4\}$. Für jedes $g \in M$ wird durch

$$\frac{dP_g}{d\lambda\!\!\lambda_{|[0,1]}} := \left(\frac{1}{2} g + \sqrt{1 - \frac{1}{4} \int_0^1 g^2 \, d\lambda\!\!\lambda_{|[0,1]}} \right)^2$$

ein Wahrscheinlichkeitsmaß P_g definiert mit $P_g \ll \lambda\!\!\lambda_{|[0,1]}$ und $P_0 = \lambda\!\!\lambda_{|[0,1]}$. Umgekehrt lässt jedes Wahrscheinlichkeitsmaß P mit $P \ll \lambda\!\!\lambda_{|[0,1]}$ eine Darstellung der Form $P = P_g$ zu, indem

$$g := 2 \left(\sqrt{\frac{dP}{d\lambda\!\!\lambda_{|[0,1]}}} - \int \sqrt{\frac{dP}{d\lambda\!\!\lambda_{|[0,1]}}} \, d\lambda\!\!\lambda_{|[0,1]} \right)$$

gesetzt wird. Also stimmen die Mengen von Verteilungen $\{P_{|\mathcal{B}([0,1])} : P \ll \lambda\!\!\lambda_{|[0,1]}\}$ und $\{P_g : g \in M\}$ überein. Damit erhält man eine Parametrisierung für die von $\lambda\!\!\lambda_{|[0,1]}$ dominierten Wahrscheinlichkeitsmaße auf $([0,1], \mathcal{B}([0,1]))$. In dem Beweis in STRASSER [41, Theorem 75.2] sieht man, dass $t \mapsto P_{tg}$ eine $L_2(P_0)$-differenzierbare Kurve ist mit Tangente g.
Seien $H_n := \{g \in H_0 : \frac{1}{\sqrt{n}} g \in M\}$ und $P_{n,g} := P^n_{\frac{1}{\sqrt{n}}g}$ für $n \in \mathbb{N}$. Man interessiert sich für das asymptotische Verhalten der Produktexperimente

$$E_n := ([0,1]^n, \mathcal{B}([0,1]^n), \{P_{n,g} : g \in H_n\}). \tag{2.1}$$

Es ist leicht zu sehen, dass $H_n \uparrow H_0$ gilt. Überdies folgt mit STRASSER [41, Theorem 75.8] durch Anwendung von Satz 2.10, dass die Folge $(E_n)_{n \in \mathbb{N}}$ der Experimente schwach gegen den Gauß-Shift auf H_0, d.h. gegen das Signalerkennungsproblem mit Brownscher Brücke als Rauschterm, konvergiert. $\square$

2.3 Der Kolmogoroff-Smirnoff-Test als Limestest

Im vorigen Abschnitt wurde dargestellt, wie das SEP mit Brownscher Brücke als Rauschterm als Limesexperiment einer geeigneten Folge von Experimenten auftritt. Nun wird der Kolmogoroff-Smirnoff-Test eingeführt und erläutert, wie dieser Test in der asymptotischen Statistik als Limestest für das Limesmodell auftritt. Vorher wird definiert, was man unter stochastisch geordneten Wahrscheinlichkeitmaßen versteht.

Definition 2.13 (Stochastisch größer)
Seien P und Q Wahrscheinlichkeitsmße auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ mit Verteilungsfunktionen F_P und F_Q. Dann heißt P stochastisch größer als Q (bzw. F_P heißt stochastisch größer als F_Q), falls gilt:

$$F_P(t) \leq F_Q(t) \ \forall t \in \mathbb{R} \quad \textit{und} \quad F_P(t) < F_Q(t) \quad \textit{für mindestens ein} \ \ t.$$

Analog wird stochastisch kleiner definiert.

Es wird folgendes nichtparametrisches Testproblem betrachtet. Dazu seien $X_1, ..., X_n$ reelle i.i.d. Zufallsvariable mit stetiger Verteilungsfunktion F und eine weitere stetige Verteilungsfunktion F_0 gegeben. Getestet werden stochastisch geordnete Hypothesen:

$$H_1 : F \leq F_0 \ \text{ gegen } \ K_1 : F \geq F_0, F \neq F_0, \tag{2.2}$$

oder man betrachtet das zweiseitige Testproblem

$$H_2 : F = F_0 \ \text{ gegen } \ K_2 : F \neq F_0. \tag{2.3}$$

Bei der Untersuchung nichtparametrischer Testverfahren betrachtet man, wie sich eine Familie von Tests $(\varphi_n)_{n\in\mathbb{N}}$, die in der Regel vom Stichprobenumfang abhängt, unter lokalen Alternativen verhält, wenn die zugrunde liegenden Verteilungen aus einer L_2-differenzierbaren Kurve von Verteilungen stammen. Von besonderem Interesse sind Kurven von Verteilungen $(P_s)_{|s|<\varepsilon}$, die folgende Eigenschaften erfüllen:

- Die Kurve hat die Fußverteilung $P_0 = \lambda\!\!\lambda_{|[0,1]}$ mit $P_s \ll P_0$ für jedes $|s| < \varepsilon$,

- P_s ist stochastisch größer als $\lambda\!\!\lambda_{|[0,1]}$ für alle $s < 0$ und P_s ist stochastisch kleiner als $\lambda\!\!\lambda_{|[0,1]}$ für alle $s > 0$.

Sei nun $(P_s)_{|s|<\varepsilon}$ eine Kurve von Verteilungen mit Fußverteilung $P_0 = \lambda\!\!\lambda_{|[0,1]}$ und $P_s \ll P_0$ für jedes $|s| < \varepsilon$. Die Kurve $(P_s)_{|s|<\varepsilon}$ sei $L_2(P_0)$-differenzierbar mit Tangente g. Folglich ist $(P_s)_{|s|<\varepsilon}$ auch $L_1(P_0)$-differenzierbar mit Ableitung g (vgl. WITTING [44, Satz 1.190]), d.h.

$$\lim_{s\to 0} \int_0^1 \left| \frac{1}{s} \left(\frac{dP_s}{dP_0} - 1 - sg \right) \right| dP_0 = 0.$$

Dann gilt für alle $t \in [0,1]$

$$\begin{aligned}
& \lim_{s\to 0} \left| \frac{1}{s} \int_0^t \left(\frac{dP_s}{dP_0} - 1 - sg \right) \right| dP_0 = 0 \\
\Longrightarrow \quad & \lim_{s\to 0} \frac{1}{s} \int_0^t \left(\frac{dP_s}{dP_0} - 1 - sg \right) \, dP_0 = 0 \\
\Longrightarrow \quad & \lim_{s\to 0} \frac{1}{s} \left(\int_0^t \frac{dP_s}{dP_0} \, dP_0 - \int_0^t 1 \, dP_0 \right) = \int_0^t g \, dP_0 \\
\Longrightarrow \quad & \lim_{s\to 0} \frac{1}{s} (F_{P_s}(t) - F_{P_0}(t)) = \int_0^t g \, dP_0.
\end{aligned}$$

Ist also $(P_{sg})_{|s|<\varepsilon}$ eine $L_2(P_0)$-differenzierbare Kurve von Verteilungen mit Tangente g (vgl. Abschnitt 2.2) und $S_g(t) := \int_0^t g \, d\lambda\!\!\lambda_{|[0,1]} \geq 0$ für alle $t \in [0,1]$, so ist P_{sg} stochastisch größer als P_0 für $s < 0$ und P_{sg} ist stochastisch kleiner als P_0 für $s > 0$ jeweils in einer ausreichend kleinen Umgebung von 0. Aufgrund dieser Überlegung führt die Parametrisierung der Wahrscheinlichkeitsmaße durch Tangenten $g \in L_2^0(0,1)$ zu einer Beschreibung der Hypothesen (2.2) durch Signale

$$H_1 = \{g \in L_2^0(0,1) : S_g \leq 0\} \quad \text{gegen} \quad K_1 = \{g \in L_2^0(0,1) : S_g \geq 0, S_g \neq 0\}$$

mit $S_g(t) := \int_0^t g(u) \, du$ für $g \in L_2^0(0,1)$ und $0 \leq t \leq 1$. Man testet also, ob das beobachtete verrauschte Signal von einem Parameter $g \in L_2^0(0,1)$ mit $S_g(t) \leq 0$ für alle $0 \leq t \leq 1$ stammt, oder ob es von einem Parameter g mit $S_g(t) \geq 0$ für alle $0 \leq t \leq 1$ herrührt.

Der einseitige Kolmogoroff-Smirnoff-Test ist auf das Prüfen von stochastisch-größer Verteilungen gegen stochastisch-kleiner Verteilungen zugeschnitten. Sei $(B_0(t))_{0\leq t\leq 1}$ eine Brownsche Brücke.

Definition 2.14 (**Test vom Kolmogoroff-Smirnoff-Typ**)
Der Test

$$\varphi_{KS} = \begin{cases} 1 & > \\ & \sup_{0\leq t\leq 1}(Y(t) - \rho(t)) \quad 0 \\ 0 & \leq \end{cases} \tag{2.4}$$

mit $Y(t) = B_0(t) + \int_0^t g(u)\ du$ *für* $0 \leq t \leq 1$ *und* $g \in L_2^0(0,1)$ *heißt einseitiger Test vom Kolmogoroff-Smirnoff-Typ. Im Folgenden wird auch die Bezeichnung KS-Test verwendet.*

Der folgende Satz stellt ein wichtiges Mittel zur Herleitung der oberen Schranken für die asymptotische Gütefunktion dar.

Satz 2.15 (**Hauptsatz der asymptotischen Testtheorie**)
Seien Θ *eine nichtleere Menge und* $(\Theta_n)_{n\in\mathbb{N}}$ *eine Folge von Teilmengen von* Θ *mit* $\Theta_n \uparrow \Theta$ *für* $n \to \infty$. *Eine Folge von Experimenten* $E_n = (\Omega_n, \mathcal{A}_n, \{P_{n,h} : h \in \Theta_n\}), n \in \mathbb{N}$, *konvergiere schwach gegen ein Limesexperiment* $E = (\Omega, \mathcal{A}, \{P_h : h \in \Theta\})$. *Außerdem seien* $H, K \subset \Theta$ *nichtleere Teilmengen mit* $H \cup K = \Theta$ *und* $H \cap K \neq \emptyset$. *Sei* $(\varphi_n)_{n\in\mathbb{N}}$ *eine Testfolge für das Testproblem* H *gegen* K. *Dann existiert zu jeder Teilfolge* $(\varphi_{n,j})_{j\in\mathbb{N}}$ *ein Test* φ *für* E, *so dass gilt:*

$$\liminf_{j\to\infty} \int \varphi_{n_j}\ dP_{n_j,h} \leq \int \varphi\ dP_h \quad \forall h \in K$$

und

$$\limsup_{j\to\infty} \int \varphi_{n_j}\ dP_{n_j,h} \geq \int \varphi\ dP_h \quad \forall h \in H.$$

Beweis. Vgl. STRASSER [42, Satz 11.8] ■

Beispiel 2.16
Sei

$$E_n := ([0,1]^n, \mathcal{B}([0,1]^n), \{P_{n,g} : g \in H_n\}).$$

das Produktexperiment (2.1) aus Bemerkung 2.12 mit $H_n = \{g \in L_2^0(0,1) : \frac{1}{\sqrt{n}}g \in M\}$ und $P_{n,g} := P^n_{\frac{1}{\sqrt{n}}g}$ für $n \in \mathbb{N}$, wobei $M := \{g \in L_2^0(0,1) : \int_0^1 g^2\ d\lambda\!\!\lambda_{|[0,1]} \leq 4\}$. Die Folge $(E_n)_{n\in\mathbb{N}}$ konvergiert schwach gegen das Signalerkennungsproblem mit Brownscher Brücke als Rauschterm

$$E = (C_0, \mathcal{B}(C_0), \{P_g : g \in L_2^0(0,1)\}).$$

Man führe für E_n den Test $\varphi_n := \varphi_{KS}^n$ ein, gegeben durch

$$\varphi_{KS}^n = \begin{cases} 1 & > \\ & \sup_{0\le t\le 1}(\sqrt{n}\ (F_n(t)-t)-\rho(t)) \qquad 0\ , \\ 0 & \le \end{cases}$$

wobei $\rho : [0,1] \to \mathbb{R}$ eine Überschreitungsfunktion ist. Dann erhält man als Limestest den einseitigen KS-Test (2.4). Es ist nämlich bekannt, dass der empirische Prozess $\sqrt{n}\,(F_n(t)-F_0(t))$ unter lokalen Alternativen $P^n_{g/\sqrt{n}}$ gegen das SEP mit Brownscher Brücke als Rauschterm konvergiert:

$$\sqrt{n}\ (F_n(t)-F_0(t)) \xrightarrow[P^n_{g/\sqrt{n}}]{\mathcal{D}} Y(t) = B_0(F_0(t)) + \int_0^{F_0(t)} g(u)\ du \text{ für } n \to \infty.$$

Daraus folgt für $g \in L_2^0(0,1)$

$$\lim_{n\to\infty} E_{P^n_{g/\sqrt{n}}}(\varphi_{KS}^n) = E_{P_g}(\varphi_{KS}),$$

vgl. STRASSER [41, Example 82.23], VAN DER VAART [43] sowie SHORACK und WELLNER [40]. □

Bemerkung 2.17

Es werden hier nur Familien $(P_t)_{t\in\mathbb{R}}$ mit $P_0 = \lambda\!\!\lambda_{|[0,1]}$ betrachtet. Hat aber für eine beliebige Kurve von Verteilungen die Fußverteilung P_0 eine stetige Verteilungsfunktion F_{P_0}, so kann der allgemeine Fall auf den Spezialfall $P_0 = \lambda\!\!\lambda_{|[0,1]}$ zurückgeführt werden, indem zu der Familie $\left(P_t^{F_{P_0}}\right)_{t\in\mathbb{R}}$ übergangen wird, da für eine reelle ZV Z mit stetiger Verteilungsfunktion F gilt

$$\left(P^Z\right)^F = \lambda\!\!\lambda_{|[0,1]}.$$

□

Kapitel 3

Die Güte einseitiger Signalerkennungstests

In diesem Kapitel wird die Gütefunktion einseitiger Signalerkennungstests untersucht. Es werden Räume von Alternativen angegeben, auf deren orthogonalem Komplement die Güte flach ist. Zunächst werden einige Testverfahren für das Signalerkennungsproblem vorgestellt. Dabei handelt es sich um einseitige Testprobleme.

3.1 Einige Anpassungstests für Signalerkennungsprozesse

Es sei gegeben ein Hilbertraum $(H, \langle \cdot, \cdot \rangle)$ mit $H \subseteq L_2([0,1], \lambda\!\!\lambda_{|[0,1]})$, der die Signale

$$S_g(t) = \int_0^t g(u)\ du, \quad g \in H,\ 0 \leq t \leq 1$$

parametrisiert. Das Testproblem ist gegeben durch

$$H_1 = \{g \in H : S_g \leq 0\} \ \text{ gegen } \ K_1 = \{g \in H : S_g \geq 0, S_g \neq 0\}. \tag{3.1}$$

Es werden folgende Bezeichnungen eingeführt, die in dieser Arbeit häufig verwendet werden:

Bezeichnung 3.1 (**und Definition**)
(a) Es bezeichne

$$V_{\geq} = V_{\geq}(H) := \{g \in H : S_g \geq 0\} \ \text{ und } \ V_{\leq} = V_{\leq}(H) := \{g \in H : S_g \leq 0\} \tag{3.2}$$

nichttriviale abgeschlossene konvexe Kegel in H.

(b) Für einen abgeschlossenen konvexen Kegel $V \subset H$ sei

$$\tilde{V} := \{g \in H : \langle g, v \rangle \leq 0 \ \forall v \in V\}$$

der polare Kegel zu V (vgl. zur Definition BARLOW et al. [1]).

Der polare Kegel $\tilde{V}$ ist ebenfalls abgeschlossen und konvex.
Für solche Hypothesen (3.2) werden nun nichtrandomisierte, einseitige Tests

$$\varphi : C(\mathbb{R}) \to \{0, 1\}$$

für den Signalprozess $(Y(t))_{t \in \mathbb{R}}$ auf der Pfadmenge der stetigen Funktionen $C(\mathbb{R})$, versehen mit der Supremumsnorm, eingeführt. Es sei daran erinnert, dass $Y(t) : (\Omega, \mathcal{A}) \to \mathbb{R}$ ein Signalprozess bzgl. eines Gauß-Shifts $(\Omega, \mathcal{A}, \{P_h : h \in H\})$ ist mit Gleichheit in Verteilung

$$(Y(t))_{t \in \mathbb{R}} \stackrel{\mathcal{D}}{=} \left(Z(F_0(t)) + \int_0^{F_0(t)} h(u) \ du \right)_{t \in \mathbb{R}}$$

unter P_h, wobei $(Z(s))_{0 \leq s \leq 1}$ ein zentrierter Gauß-Prozess mit f.s. stetigen Pfaden ist (vgl. Definition 1.18).

Definition 3.2 **(Einseitiger Signalerkennungstest)**
Ein Test $\varphi : C(\mathbb{R}) \to \{0, 1\}$ mit kritischem Bereich $C(\varphi) := \{x(\cdot) \in C(\mathbb{R}) : \varphi(x(\cdot)) = 1\}$ heißt ein einseitiger oberer Test oder einseitiger Signalerkennungstest für die Nullhypothese $V_{\leq}$, falls aus $x(\cdot) \in C(\varphi)$ folgt, dass $y(\cdot) \in C(\varphi)$ ist für alle $y(\cdot) \in C(\mathbb{R})$ mit $y(t) \geq x(t)$ für $t \in \mathbb{R}$.

Beispiel 3.3
Die folgenden Tests $\varphi : C(\mathbb{R}) \to \{0, 1\}$ sind einseitige Tests im Sinne von Definition 3.2. Sei $c > 0$ eine geeignete Konstante.
(a) **Einseitiger Test vom Kolmogoroff-Smirnoff-Typ** (vgl. Definition 2.4)
Für das Testproblem (3.1) wird häufig der einseitige Test vom Kolmogoroff-Smirnoff-Typ

$$\varphi_{\text{KS}}(Y(\cdot)) = \mathbb{1}_{\{\sup_{t \in \mathbb{R}}(Y(t) - \rho(t)) > 0\}}$$

benutzt. Dieser basiert auf einer messbaren Überschreitungsfunktion $\rho : \mathbb{R} \to \mathbb{R}$. Der KS-Test verwirft die Hypothese, wenn der beobachtete Pfad aus $C(\mathbb{R})$ die Funktion ρ an mindestens einer Stelle berührt oder überschreitet. Er ist ein spezieller Fall des Tests

$$\varphi(Y(\cdot)) = \mathbb{1}_{\{\sup_{t \in \mathbb{R}}(w(t) Y(t)) > c\}}$$

mit einer positiven Gewichtsfunktion w (vgl. BISCHOFF et al. [6]).
(b) Seien $I : C(\mathbb{R}) \to \mathbb{R}_+$ eine Seminorm mit $I(y(\cdot)) \geq I(x(\cdot))$ für alle stetigen Funktionen $y(t) \geq x(t) \geq 0, t \in \mathbb{R}$ und $Y^+ := \max(Y, 0)$. Dann ist durch

$$\varphi(Y(\cdot)) := \mathbb{1}_{\{I(Y(t)^+)>c\}}$$

ein einseitiger oberer Test gegeben.
(c) **Integral-Test**
Der Integral-Test ist definiert als

$$\varphi(Y(\cdot)) := \mathbb{1}_{\left\{\int q(Y(t)^+)d\mu(t)>c\right\}},$$

wobei $q : [0, \infty) \to \mathbb{R}_+$ eine nichtfallende Funktion ist und μ ein Maß auf $\mathbb{R}$ bezeichnet. Spezialfälle des Integral-Tests sind einseitige Versionen von Anderson-Darling-Tests sowie Cramér-von-Mises-Tests (vgl. SHORACK und WELLNER [40]). □

In dieser Arbeit stehen die Regressionsmodelle der Brownschen Brücke und der Brownschen Bewegung im Vordergrund. Deshalb wird bei der Betrachtung einseitiger Testprobleme häufig auf den Signalprozess der Brownschen Brücke

$$Y(t) = B_0(t) + \int_0^t g \; d\lambda\!\!\lambda_{|[0,1]}, \;\; g \in L_2^0(0,1), \;\; 0 \leq t \leq 1$$

und den Signalprozess der Brownschen Bewegung

$$Y(t) = B(t) + \int_0^t h \; d\lambda\!\!\lambda_{|[0,1]}, \;\; h \in L_2(0,1), \;\; 0 \leq t \leq 1$$

beschränkt.

3.2 Einhüllende Gütefunktionen für Gauß-Shift-Experimente

In diesem Abschnitt werden einhüllende Gütefunktionen für Gauß-Shift-Experimente angegeben, die durch Projektionen ausgedrückt werden. Dazu werden bekannte Eigenschaften für Gauß-Shift-Experimente aus Abschnitt 1.2 herangezogen. Ferner wird gezeigt, dass der Gradient eines einseitigen Signalerkennungstests in einem polaren Kegel liegt.
Zunächst werden beliebige reelle Hilberträume $(H, \langle \cdot, \cdot \rangle)$ betrachtet. Im Folgenden sei stets $E_h(\varphi) := E_{P_h}(\varphi)$ für jedes $h \in H$.

Hilfssatz 3.4
Sei $(\Omega, \mathcal{A}, \{P_h : h \in H\})$ ein Gauß-Shift mit dem zentralen Prozess $(L(h))_{h \in H}$ auf einem beliebigen Hilbertraum H. Dann ist für festes $h \in H$ die Familie $\{P_{th} : t \in \mathbb{R}\}$ eine Exponentialfamilie.

Beweis. Mit der Girsanov-Formel (1.1) erhält man für jedes $t \in \mathbb{R}$

$$\begin{aligned}\frac{dP_{th}}{dP_0} &= \exp\left(L(th) - \frac{\|th\|^2}{2}\right)\\ &= \exp\left(tL(h) - \frac{t^2\|h\|^2}{2}\right)\\ &= \exp\left(-\frac{t^2\|h\|^2}{2}\right)\exp(tL(h)).\end{aligned}$$

Damit folgt die Behauptung. ■

Satz 3.5
Sei $(\Omega, \mathcal{A}, \{P_h : h \in H\})$ ein Gauß-Shift mit dem zentralen Prozess $(L(h))_{h \in H}$ auf einem beliebigen Hilbertraum H. Sei η ein Test mit $E_0(\eta) = \alpha$. Dann gilt:
(a) Für jedes $h \in H$ existieren $a(h), b(h) \in \mathbb{R}$, so dass für $t \in \mathbb{R}$ gilt:

$$E_{th}(\eta) = E_0(\eta) + b(h)t + a(h)\frac{t^2}{2} + o(t^2).$$

(b) Es existiert ein $h_0 \in H$ mit $b(h) = \langle h, h_0\rangle = \int L(h)\eta \, dP_0$ für jedes $h \in H$.
(c) Es existiert ein selbstadjungierter Hilbert-Schmidt-Operator $T : H \to H$ mit

$$a(h) = \langle h, T(h)\rangle = \int L(h)^2(\eta - \alpha) \, dP_0.$$

Für ein geeignetes Orthonormalsystem $(h_i)_{i \in \mathbb{N}}$ gilt also $a(h) = \sum_{i=1}^{\infty} \lambda_i \langle h, h_i\rangle^2$ und $(\sum_{i=1}^{\infty} \lambda_i^2)^{\frac{1}{2}} < \infty$.

Beweis. (a) Die Behauptung folgt aus der Eigenschaft, dass $\{P_{th} : t \in \mathbb{R}\}$ nach Hilfsatz 3.4 für jedes $h \in H$ eine Exponentialfamilie ist.
(b) Aus Standardresultaten über Exponentialfamilien (vgl. WITTING [44]) folgt unmittelbar die Gleichheit $b(h) = \int L(h)\eta \, dP_0, h \in H$. Die Existenz des Elements $h_0 \in H$ folgt aus dem Darstellungssatz von Riesz, da $b : H \to \mathbb{R}$ ein stetig lineares Funktional ist. (Vgl. MEISE [29, Satz 11.9].)
(c) Vgl. JANSSEN [13, Theorem 2.1]. ■

Definition 3.6 (**Ableitung und Gradient eines Tests**)
Die im Satz 3.5(a) auftretenden Koeffizienten $b(h)$ bzw. $a(h)$ heißen die erste bzw. die zweite Ableitung des Tests η in Richtung $h \in H$. Das Element $h_0 \in H$ aus (b) heißt der Gradient des Tests η.

Lemma 3.7 (**Projektionslemma**)
(a) Seien H ein reeller Hilbertraum, $V \subset H$ ein abgeschlossener konvexer Kegel und $pV(h)$ bezeichne die eindeutig bestimmte Projektion von $h \in H$ in V. Eine orthogonale Zerlegung

$$h = pV(h) + p\tilde{V}(h), \;\; \langle pV(h), p\tilde{V}(h)\rangle = 0, \tag{3.3}$$

von $h \in H$ ist gegeben durch die Projektion $p\tilde{V}(h)$ in den polaren Kegel $\tilde{V} = \{g \in H : \langle g, v\rangle \leq 0 \; \forall v \in V\}$.
(b) Sei $h = h_1 + h_2$ eine weitere orthogonale Zerlegung mit $h_1 \in V, h_2 \in \tilde{V}$ und $h_1 \perp h_2$. Dann ist $h_1 = pV(h)$ und $h_2 = p\tilde{V}(h)$.

Beweis. Der Beweis folgt aus MOREAU [31], BARLOW et al. [1] sowie BEHNEN und NEUHAUS [2] und wird im Anhang ausgeführt. ■

Bemerkung 3.8
Sei $(\Omega, \mathcal{A}, \{P_h : h \in H\})$ ein Gauß-Shift-Experiment mit dem zentralen Prozess $(L(h))_{h\in H}$. Es gelten folgende Eigenschaften:
(a) Unter $P_g, g \in H$, ist der geshiftete Prozess

$$h \mapsto L(h) - \langle h, g\rangle$$

ein linearer zentrierter Gauß-Prozess mit der unveränderten Kovarianzstruktur

$$Cov_{P_g}(L(h_1), L(h_2)) = \langle h_1, h_2\rangle \; \forall h_1, h_2 \in H.$$

(b) Es bezeichne $\psi : \Omega \to [0,1]$ einen Test zum Niveau $\alpha \in (0,1)$ für das Testproblem

$$\{P_0\} \;\; \text{gegen} \;\; \{P_h\} \tag{3.4}$$

für ein $h \in H$. Dann gilt

$$\Phi(u_\alpha - \|h\|) \leq E_h(\psi) \leq \Phi(\|h\| - u_{1-\alpha}).$$

Neyman-Pearson Tests nehmen sogar die Schranken an. □

Beweis. (a) Vgl. Satz 1.13 und Satz 1.14.
(b) Unter Verwendung der Girsanov-Formel (1.1) läßt sich der Neyman-Pearson-Test ψ^* für das Testproblem (3.4) zum Niveau $\alpha \in (0,1)$ darstellen durch

$$\psi^* = \begin{cases} 1 & > \\ & \text{falls } L(h) \qquad \|h\| u_{1-\alpha}, \\ 0 & \leq \end{cases}$$

wobei $u_{1-\alpha}$ das $(1-\alpha)$-Quantil der Standardnormalverteilung ist. Damit folgt die obere Schranke

$$E_h(\psi) \leq E_h(\psi^*) = \Phi(\|h\| - u_{1-\alpha}).$$

Für die untere Schranke betrachte man den Neyman-Pearson-Test $\overline{\psi^*}$ für dasselbe Testproblem (3.4) zum Niveau $1-\alpha$. Setzt man $\overline{\psi} := 1 - \psi$, so folgt wegen $E_h(\overline{\psi}) \leq E_h(\overline{\psi^*}) = \Phi(\|h\| - u_\alpha)$ die Abschätzung

$$E_h(\psi) \geq 1 - \Phi(\|h\| - u_\alpha) = \Phi(u_\alpha - \|h\|).$$

■

Gegeben seien die abgeschlossenen konvexen Kegeln

$$V_{\leq} = \{g \in H : S_g \leq 0\} \quad \text{und} \quad V_{\geq} = \{g \in H : S_g \geq 0\}$$

aus Definition 3.1 mit den polaren Kegeln

$$\tilde{V}_{\leq} = \{g \in H : \langle g, v \rangle \leq 0 \ \forall v \in V_{\leq}\} \quad \text{bzw.} \quad \tilde{V}_{\geq} = \{g \in H : \langle g, v \rangle \leq 0 \ \forall v \in V_{\geq}\}.$$

Sei außerdem $(Y(t))_{t\in\mathbb{R}} = (Z(F_0(t)) + S_h(F_0(t)))_{t\in\mathbb{R}}$ ein Signalprozess bzgl. eines Gauß-Shift-Experiments $(\Omega, \mathcal{A}, \{P_h : h \in H\})$ (vgl. Definition 1.18).

Lemma 3.9
Sei $\varphi = \varphi(Y(\cdot))$ ein einseitiger oberer Test für den Signalprozess $(Y(t))_{t\in\mathbb{R}}$ zum Niveau $\alpha := E_0(\varphi), \alpha \in (0,1)$. Unter den Alternativen $P_h, h \in H$, erhält man folgende Ungleichungen für die einhüllende Gütefunktion.

(a) $\Phi(\Phi^{-1}(\alpha) - \|p\tilde{V}_{\geq}(h)\|) \ \leq \ E_h(\varphi) \ \leq \ \Phi(\|p\tilde{V}_{\leq}(h)\| - \Phi^{-1}(1-\alpha)),$

(b) $-\frac{\|p\tilde{V}_{\geq}(h)\|}{\sqrt{2\pi}} \ \leq \ E_h(\varphi) - \alpha \ \leq \ \frac{\|p\tilde{V}_{\leq}(h)\|}{\sqrt{2\pi}},$

(c) *Der Gradient h_0 des Tests φ ist von der Gestalt $h = \tilde{h} \circ F_0$ mit $\tilde{h} \in \tilde{V}_{\leq}$.*

Beweis. (a) Sei $h = h_1 + h_2$ die orthogonale Zerlegung (3.3) von $h \in H$ in $h_1 := pV_{\leq}(h)$ und $h_2 := p\tilde{V}_{\leq}(h)$. Wegen $h_1 \in V_{\leq}$ folgt $S_h = S_{h_1} + S_{h_2} \leq S_{h_2}$, d.h. die Signale sind geordnet. Dann ist $Z(F_0(\cdot)) + S_h(F_0(\cdot)) \leq Z(F_0(\cdot)) + S_{h_2}(F_0(\cdot))$. Somit erhält man

$$\begin{aligned} E_h(\varphi) &= P_0(\{Z(F_0(\cdot)) + S_h(F_0(\cdot)) \in C(\varphi)\}) \\ &\leq P_0(\{Z(F_0(\cdot)) + S_{h_2}(F_0(\cdot)) \in C(\varphi)\}) \\ &= E_{h_2}(\varphi) \end{aligned}$$

nach Definition 3.2 des einseitigen Signalerkennungstests.
Sei ψ der Neyman-Pearson-Test für $\{P_0\}$ gegen $\{P_{h_2}\}$ zum Niveau $\alpha \in (0,1)$. Mit Bemerkung 3.8 (b) folgt

$$E_h(\varphi) \leq E_{h_2}(\varphi) \leq E_{h_2}(\psi) = \Phi(\|h_2\| - u_{1-\alpha}) = \Phi(\|p\tilde{V}_{\leq}(h)\| - u_{1-\alpha}).$$

Für die untere Schranke betrachte man die orthogonale Zerlegung von $h = \overline{h}_1 + \overline{h}_2$ in $\overline{h}_1 := pV_{\geq}(h)$ und $\overline{h}_2 := p\tilde{V}_{\geq}(h)$. Wegen $\overline{h}_1 \in V_{\geq}$ folgt $S_h = S_{\overline{h}_1} + S_{\overline{h}_2} \geq S_{\overline{h}_2}$ und damit

$$E_h(\varphi) \geq E_{\overline{h}_2}(\varphi).$$

Nun betrachte man den Neyman-Pearson-Test $\overline{\psi}$ für P_0 gegen $P_{\overline{h}_2}$ zum Niveau $1-\alpha$. Wieder mit Bemerkung 3.8 (b) ergibt sich

$$E_h(\varphi) \geq E_{\overline{h}_2}(\varphi) \geq 1 - E_{\overline{h}_2}(\overline{\psi}) = 1 - \Phi(\|\overline{h}_2\| - u_\alpha) = \Phi(u_\alpha - \|p\tilde{V}_{\geq}(h)\|).$$

(b) Seien $h_2 := p\tilde{V}_{\leq}(h)$ und $\overline{h}_2 := p\tilde{V}_{\geq}(h)$ wie in Teil (a). Mit dem Mittelwertsatz der Differentialrechnung erhält man

$$\begin{aligned} E_h(\varphi) - \alpha &\overset{(a)}{\leq} \Phi(\|h_2\| - u_{1-\alpha}) - \alpha = \Phi(\|h_2\| - u_{1-\alpha}) - \Phi(-u_{1-\alpha}) \\ &= (\|h_2\| - u_{1-\alpha} - (-u_{1-\alpha}))\Phi'(\xi) = \|h_2\|\Phi'(\xi) \\ &\leq \|h_2\| \frac{1}{\sqrt{2\pi}} \end{aligned}$$

für einen Mittelwert $\xi \in (-u_{1-\alpha}, -u_{1-\alpha} + \|h_2\|)$. Analog

$$\begin{aligned} E_h(\varphi) - \alpha &\overset{(a)}{\geq} \Phi(u_\alpha - \|\overline{h}_2\|) - \alpha = \Phi(u_\alpha - \|\overline{h}_2\|) - \Phi(u_\alpha) \\ &= (u_\alpha - \|\overline{h}_2\| - u_\alpha)\Phi'(\overline{\xi}) = -\|\overline{h}_2\|\Phi'(\overline{\xi}) \\ &\geq -\|\overline{h}_2\| \frac{1}{\sqrt{2\pi}} \end{aligned}$$

für einen Mittelwert $\overline{\xi} \in (u_\alpha - \|\overline{h}_2\|, u_\alpha)$.
(c) Es genügt den Beweis für $F_0(x) := x,\ 0 \leq x \leq 1$ zu führen, vgl. Bemerkung 2.17. Nach Hilfsatz 3.4 ist bekannt, dass die Familie $\{P_{th} : t \in \mathbb{R}\}$ für jedes $h \in H$ eine Exponentialfamilie ist. Deshalb ist die Gütefunktion $t \mapsto E_{th}(\varphi)$ für jedes $h \in H$ beliebig oft differenzierbar auf $\mathbb{R}$ mit dem Gradienten h_0 des Tests φ, d.h.

$$\frac{d}{dt} E_{th}(\varphi)_{|t=0} = \langle h, h_0 \rangle \quad \forall h \in H, \tag{3.5}$$

vgl. Satz 3.5. Sei $h_0 = h_1 + h_2$ die orthogonale Zerlegung von h_0 in $h_1 := pV_{\leq}(h_0)$ und $h_2 := p\tilde{V}_{\leq}(h_0)$. Dann ist

$$\|h_0\|^2 = \|h_1\|^2 + \|h_2\|^2 \geq \|h_2\|^2. \tag{3.6}$$

Mit (3.5) folgt

$$\frac{d}{dt} E_{th_2}(\varphi)_{|t=0} = \langle h_2, h_0 \rangle = \langle h_2, h_1 \rangle + \langle h_2, h_2 \rangle = \|h_2\|^2.$$

Andererseits ist $E_{th_0}(\varphi) \leq E_{th_2}(\varphi)$ für $t \geq 0$, weil die Signale $S_{th_0} \leq S_{th_2}$ für $t \geq 0$ geordnet sind. Es ist

$$\frac{d}{dt} E_{th}(\varphi) = \lim_{s \downarrow 0} \frac{E_{(s+t)h}(\varphi) - E_{th}(\varphi)}{s}.$$

Damit erhält man

$$\begin{aligned} \|h_0\|^2 &\overset{(3.5)}{=} \frac{d}{dt} E_{th_0}(\varphi)_{|t=0} = \lim_{s \downarrow 0} \frac{E_{sh_0}(\varphi) - E_0(\varphi)}{s} = \lim_{s \downarrow 0} \frac{E_{sh_0}(\varphi) - \alpha}{s} \\ &\leq \lim_{s \downarrow 0} \frac{E_{sh_2}(\varphi) - \alpha}{s} = \frac{d}{dt} E_{th_2}(\varphi)_{|t=0} \\ &= \|h_2\|^2. \end{aligned} \tag{3.7}$$

Also folgt aus (3.6) und (3.7) die Gleichheit $\|h_0\|^2 = \|h_2\|^2$ und somit $h_0 = h_2 \in \tilde{V}_{\leq}$. ■

3.3 Konkave Signale für von Rademacherfunktionen erzeugten Räume

In diesem Abschnitt wird die Familie der Rademacherfunktionen vorgestellt. Diese bildet ein Orthonormalsystem für den Hilbertraum $L_2(0,1)$. Der Vorteil der Rademacherfunktionen ist, dass sich für jedes Signal einer Rademacherfunktion die kleinste konkave (nichtfallende) Majorante leicht angeben läßt, welche nichts anderes ist als das Signal zu der Projektion der Rademacherfunktion in den polaren Kegel $\tilde{V}^t_{\leq}$ bzw. $\tilde{V}^0_{\leq}$ (siehe (3.8)).

Somit wird eine explizite Darstellung der Schranken aus Lemma 3.9 für die Gütefunktion eines einseitigen Signalerkennungstest ermöglicht.

Im Vordergrund dieser Arbeit stehen die Signalerkennungsprozesse mit Brownscher Brücke bzw. mit Brownscher Bewegung als Rauschterm. Deshalb werden speziell die Hilberträume $L_2^0([0,1], \lambda\!\!\lambda_{|[0,1]})$ und $L_2([0,1], \lambda\!\!\lambda_{|[0,1]})$ betrachtet. Für diese werden die Kegel aus (3.2) neu eingeführt:

$$V_{\leq}^t := \{g \in L_2(0,1) : S_g \leq 0\} \quad \text{und} \quad V_{\leq}^0 := V_{\leq}^t \cap L_2^0(0,1). \tag{3.8}$$

Analog werden die Kegel $V_{\geq}^t$ und $V_{\geq}^0$ neu definiert.

Nun werden die Rademacherfunktionen vorgestellt.

Definition 3.10 (Rademacherfunktionen)
Die Rademacherfunktionen $(r_n)_{n \in \mathbb{N}_0}$ *sind ein System von Funktionen auf* $(0,1)$*, definiert durch*

$$\begin{aligned} r_0(u) &:= 1,\ r_1(u) := \mathbb{1}_{(0,\frac{1}{2}]}(u) - \mathbb{1}_{(\frac{1}{2},1)}(u), \\ r_{k+1}(u) &:= r_k(2u) \quad \textit{für} \quad 0 < u < \frac{1}{2}, k \in \mathbb{N}, \\ r_{k+1}\left(\frac{1}{2} + u\right) &:= r_{k+1}(u) \quad \textit{für} \quad 0 < u < \frac{1}{2}, k \in \mathbb{N}. \end{aligned}$$

Die Rademacherfunktionen bilden ein Orthonormalsystem für den Hilbertraum $L_2(0,1)$, welches nicht vollständig ist, vgl. SCHILLING [38]. Dies bedeutet, dass sich nicht jedes $h \in L_2(0,1)$ nach den Rademacherfunktionen entwickeln läßt. Deshalb beschränkt man sich auf Räume, die von den Rademacherfunktionen erzeugt werden. Dazu definiere man die Teilhilberträume

$$\tilde{H} := \overline{\operatorname{span}(r_i : i \in \mathbb{N}_0)} \quad \text{und} \quad \tilde{H}^0 := \tilde{H} \cap L_2^0(0,1).$$

Jedes $h \in \tilde{H}$ läßt sich dann darstellen als

$$\begin{aligned} h &= \sum_{i=0}^{\infty} \beta_i r_i \quad \text{mit} \quad \beta_i := \langle h, r_i \rangle \quad \text{und} \qquad (3.9) \\ \|h\|^2 &= \sum_{i=0}^{\infty} \beta_i^2 \quad \text{(Parsevalsche Gleichung).} \end{aligned}$$

Um zu den Hauptresultaten zu gelangen, sind folgende Aussagen über Signale und deren kleinste konkave (nichtfallende) Majoranten wichtig. Diese sind der gemeinsamen

Arbeit von BISCHOFF und HASHORVA [4] entnommen. Dabei beachte man, dass für $g \in L_2^0(0,1)$ ein $\tilde{g} \in L_2^0(0,1)$ existiert, so dass $S_{\tilde{g}}$ die kleinste konkave Majorante von S_g ist. Analog existiert für $h \in L_2(0,1)$ ein $\tilde{h} \in L_2(0,1)$, so dass $S_{\tilde{h}}$ die kleinste nichtfallende konkave Majorante von S_h ist.

Lemma 3.11
Sei $h \in L_2^0(0,1)$ mit dem zugehörigen Signal S_h und $S_{\tilde{h}}$ die kleinste konkave Majorante von S_h für ein $\tilde{h} \in L_2^0(0,1)$. Dann ist

$$\|\tilde{h}\|^2 = \min\{\|g\|^2 : S_g \geq S_h, g \in L_2^0(0,1)\}.$$

Beweis. Vgl. BISCHOFF und HASHORVA [4, Theorem 2.1]. ∎

Der Satz 3.11 impliziert unmittelbar das folgende Korollar.

Korollar 3.12
Seien $g, h \in L_2^0(0,1)$ mit konkaven Signalen S_g und S_h. Es gelte $S_g \leq S_h$. Dann folgt

$$\|g\|^2 \leq \|h\|^2.$$

Beweis. Vgl. BISCHOFF und HASHORVA [4, Corollary 2.2]. ∎

Lemma 3.13
Sei $h \in L_2^0(0,1)$. Dann erhält man

$$\|h\|^2 = \|\tilde{h}\|^2 + \|h - \tilde{h}\|^2.$$

Beweis. Vgl. BISCHOFF und HASHORVA [4, Theorem 2.3.]. ∎

Nach Lemma 3.13 stehen also $\tilde{h}$ und $h - \tilde{h}$ orthogonal zueinander.

Analoge Ergebnisse erhält man auch für das Regressionsmodell der Brownschen Bewegung.

Lemma 3.14
Sei $h \in L_2(0,1)$ mit dem zugehörigen Signal S_h und $S_{\tilde{h}}$ die kleinste nichtfallende konkave Majorante von S_h für ein $\tilde{h} \in L_2(0,1)$. Dann ist

$$\|\tilde{h}\|^2 = \min\{\|g\|^2 : S_g \geq S_h, g \in L_2(0,1)\}.$$

Außerdem gilt

$$\|h\|^2 = \|\tilde{h}\|^2 + \|h - \tilde{h}\|^2.$$

Beweis. Vgl. BISCHOFF und HASHORVA [4]. ■

Bemerkung 3.15
(a) Sei $g \in L_2^0(0,1)$ eine beliebige nichtsteigende Funktion. Man erhält für $h \in V_{\leq}^0 = \{h \in L_2^0(0,1) : S_h \leq 0\}$ per partieller Integration

$$\begin{aligned}\int_0^1 h(u)(-g)(u)\ du &= \left(\int_0^u h(t)\ dt\right)(-g(u))\bigg|_0^1 - \int_0^1 \left(\int_0^u h(t)\ dt\right)\ d(-g)(u) \\ &= \underbrace{S_h(1)}_{=0}(-g(1-)) - \int_0^1 S_h(u)\ d(-g)(u).\end{aligned}$$

Dabei wird angenommen, dass die nichtfallende maßerzeugende Funktion $-g$ linksseitig stetig ist, vgl. SHIRYAEV [39]. Dann ist

$$\int_0^1 h(u)g(u)\ du = \int_0^1 S_h(u)\ d(-g)(u) \leq 0,$$

da $S_h \leq 0$. Folglich ist $g \in \tilde{V}_{\leq}^0 = \{v \in L_2^0(0,1) : \langle v, h\rangle \leq 0\ \forall h \in V_{\leq}^0\}$. Hat also die Tangente $g \in L_2^0(0,1)$ ein konkaves Signal S_g, so liegt g im polaren Kegel $\tilde{V}_{\leq}^0$.
(b) Sei $g \in L_2(0,1)$ eine beliebige nichtsteigende und nichtnegative Funktion. Man erhält für $h \in V_{\leq}^t = \{h \in L_2(0,1) : S_h \leq 0\}$ wieder per partieller Integration

$$\begin{aligned}\int_0^1 h(u)(-g)(u)\ du &= \left(\int_0^u h(t)\ dt\right)(-g(u))\bigg|_0^1 - \int_0^1 \left(\int_0^u h(t)\ dt\right)\ d(-g)(u) \\ &= S_h(1)(-g(1-)) - \int_0^1 S_h(u)\ d(-g)(u),\end{aligned}$$

wobei angenommen wird, dass die maßerzeugende Funktion $-g$ linksseitig stetig ist. Dann ist

$$\begin{aligned}\int_0^1 h(u)g(u)\ du &= -S_h(1)(-g(1-)) + \int_0^1 S_h(u)\ d(-g)(u) \\ &\leq \int_0^1 S_h(u)\ d(-g)(u) \\ &\leq 0,\end{aligned}$$

da g nichtnegativ ist und $S_h(u) \leq 0$ für jedes $0 \leq u \leq 1$. Folglich ist $g \in \tilde{V}_{\leq}^t = \{v \in L_2(0,1) : \langle v, h\rangle \leq 0\ \forall h \in V_{\leq}^t\}$. Hat also $g \in L_2(0,1)$ ein nichtfallendes konkaves Signal S_g, so liegt g im polaren Kegel $\tilde{V}_{\leq}^t$. □

Mit dem Projektionslemma 3.7 und Lemma 3.13 kommt man zum folgenden Ergebnis.

Lemma 3.16

(a) Sei $h \in L_2^0(0,1)$. Dann existiert eine orthogonale Zerlegung von $h = h_1 + h_2$ mit $h_1 \in V_{\leq}^0$ und $h_2 \in \tilde{V}_{\leq}^0$, so dass S_{h_2} die kleinste konkave Majorante von S_h ist.
(b) Sei $h \in L_2(0,1)$. Dann existiert eine orthogonale Zerlegung von $h = h_1 + h_2$ mit $h_1 \in V_{\leq}^t$ und $h_2 \in \tilde{V}_{\leq}^t$, so dass S_{h_2} die kleinste nichtfallende konkave Majorante von S_h ist.

Beweis. (a) Sei $h_2 \in L_2^0(0,1)$, so dass S_{h_2} die kleinste konkave Majorante von S_h ist. Setze $h_1 := h - h_2$. Dann ist $S_{h_1} = S_h - S_{h_2} \leq 0$. Also ist $h_1 \in V_{\leq}^0$. Da S_{h_2} ein konkaves Signal ist, liegt h_2 nach Bemerkung 3.15 im polaren Kegel $\tilde{V}_{\leq}^0$. Außerdem folgt mit Lemma 3.13, dass $h_2 \perp h_1 - h_2$ ist. Mit Lemma 3.7 ist diese orthogonale Zerlegung $h = h_1 + h_2$ eindeutig mit $h_1 = pV_{\leq}^0(h)$ und $h_2 = p\tilde{V}_{\leq}^0(h)$.
(b) Der Beweis entspricht im Wesentlichen dem Beweis zu Teil (a).
Sei $h_2 \in L_2(0,1)$, so dass S_{h_2} die kleinste nichtfallende konkave Majorante von S_h ist. Setze $h_1 := h - h_2$. Dann ist $h_1 \in V_{\leq}^t$. Da S_{h_2} ein nichtfallendes konkaves Signal ist, liegt h_2 nach Bemerkung 3.15 im polaren Kegel $\tilde{V}_{\leq}^t$. Mit Lemma 3.14 folgt außerdem $h_2 \perp h_1 - h_2$. Nach Lemma 3.7 ist diese orthogonale Zerlegung $h = h_1 + h_2$ eindeutig mit $h_1 = pV_{\leq}^t(h)$ und $h_2 = p\tilde{V}_{\leq}^0(h)$. ■

Nun folgen mit Hilfe von Lemma 3.16 wichtige Aussagen zu konkaven (nichtfallenden) Signalen.

Satz 3.17

(a) Sei $h = h_1 + h_2$ die orthogonale Zerlegung von $h \in L_2^0(0,1)$ in $h_1 := pV_{\leq}^0(h)$ und $h_2 := p\tilde{V}_{\leq}^0(h)$. Das Signal S_{h_2} ist die kleinste konkave Majorante von S_h. Überdies ist

$$\|p\tilde{V}_{\leq}^0(h)\| = \min\{\|g\| : g \in \tilde{V}_{\leq}^0, S_g \geq S_h\}. \tag{3.10}$$

(b) Der polare Kegel $\tilde{V}_{\leq}^0$ besteht aus allen Elementen $g \in L_2^0(0,1)$, die ein konkaves Signal S_g besitzen.
(c) Sei $h \in \tilde{H}^0 = \overline{span(r_i : i \in \mathbb{N})}$ mit $h = \sum_{i=1}^{\infty} \beta_i r_i$ wie in (3.9). Sei $\tilde{r}_i := r_i \mathbb{1}_{[2^{-i}, 1-2^{-i}]^c}$, für welche $\tilde{r}_i = p\tilde{V}_{\leq}^0(r_i)$ gilt, und $h_0 := \sum_{i \in \mathbb{N}, \beta_i > 0} \beta_i \tilde{r}_i$. Dann ist das Signal S_{h_0} konkav und man erhält $S_h \leq S_{h_2} \leq S_{h_0}$.

Beweis. (a) Aus Lemma 3.16(a) folgt, dass das Signal zu der Projektion $h_2 = p\tilde{V}_{\leq}^0(h)$ von h die kleinste konkave Majorante von S_h ist. Die Gleichheit in (3.10) folgt unmittelbar aus Lemma 3.11, da

$$\|p\tilde{V}_{\leq}^0(h)\| = \min\{\|g\| : g \in L_2^0(0,1), S_g \geq S_h\} = \min\{\|g\| : g \in \tilde{V}_{\leq}^0, S_g \geq S_h\}.$$

(b) Sei $g \in L_2^0(0,1)$ mit konkavem Signal S_g. Dann liegt g im polaren Kegel $\tilde{V}_{\le}^0$ nach Bemerkung 3.15(a). Sei umgekehrt $g \in \tilde{V}_{\le}^0$ mit der orthogonalen Zerlegung $g = g_1 + g_2$ in $g_1 = pV_{\le}^0(g)$ und $g_2 = p\tilde{V}_{\le}^0(g)$. Wegen der Eindeutigkeit der orthogonalen Zerlegung ist $g = g_2$. Somit stimmt S_g mit S_{g_2} überein und ist somit ebenfalls konkav nach Teil (a).
(c) Sei $h = \sum_{i=1}^{\infty} \beta_i r_i \in \tilde{H}^0$ mit $\beta_i = \langle h, r_i \rangle$. Das Signal $S_{\tilde{r}_i}$ ist die kleinste konkave Majorante von S_{r_i} und folglich nach Teil (a) $\tilde{r}_i = p\tilde{V}_{\le}^0(r_i)$. Damit folgt für $0 \le t \le 1$

$$\begin{aligned}
S_h(t) &= \int_0^t \sum_{i=1}^{\infty} \beta_i r_i \, d\lambda\!\!\lambda_{|[0,1]} = \lim_{n\to\infty} \sum_{i=1}^{n} \beta_i \underbrace{\int_0^t r_i \, d\lambda\!\!\lambda_{|[0,1]}}_{\ge 0} \\
&\le \lim_{n\to\infty} \sum_{i=1}^{n} \max(0,\beta_i) \int_0^t r_i \, d\lambda\!\!\lambda_{|[0,1]} = \lim_{n\to\infty} \sum_{\substack{i=1\\ \beta_i>0}}^{n} \beta_i \int_0^t r_i \, d\lambda\!\!\lambda_{|[0,1]} \\
&\le \lim_{n\to\infty} \sum_{\substack{i=1\\ \beta_i>0}}^{n} \beta_i \int_0^t \tilde{r}_i \, d\lambda\!\!\lambda_{|[0,1]} \\
&= S_{h_0}(t).
\end{aligned}$$

Somit ist das Signal S_{h_0} eine Majorante zu S_h. Da $S_{\tilde{r}_i}$ konkav ist, folgt, dass S_{h_0} als $L_2(\lambda\!\!\lambda_{|[0,1]})$-Limes von konkaven Signalen ebenfalls konkav ist (vgl. Satz A.1 aus dem Anhang). Da aber nach Teil (a) S_{h_2} die kleinste konkave Majorante zu S_h ist, folgt insgesamt

$$S_h \le S_{h_2} \le S_{h_0}.$$

■

Ein ähnliches Resultat erhält man für die Hilberträume $L_2(0,1)$ und $\tilde{H} = \overline{\mathrm{span}(r_i : i \in \mathbb{N}_0)}$.

Satz 3.18

(a) Sei $h = h_1 + h_2$ die orthogonale Zerlegung von $h \in L_2(0,1)$ in $h_1 := pV_{\le}^t(h)$ und $h_2 := p\tilde{V}_{\le}^t(h)$. Das Signal S_{h_2} ist die kleinste nichtfallende konkave Majorante von S_h. Überdies ist

$$\|p\tilde{V}_{\le}^t(h)\| = \min\{\|g\| : g \in \tilde{V}_{\le}^t, S_g \ge S_h\}. \tag{3.11}$$

(b) Der polare Kegel $\tilde{V}_{\le}^t$ besteht aus allen Elementen $g \in L_2(0,1)$, die ein nichtfallendes konkaves Signal S_g besitzen.
(c) Sei $h \in \tilde{H} = \overline{span(r_i : i \in \mathbb{N}_0)}$ mit $h = \sum_{i=0}^{\infty} \beta_i r_i$ wie in (3.9). Setze $\overline{r}_i := r_i \mathbb{1}_{[0,2^{-i}]}$, für

welche $\overline{r}_i = p\tilde{V}^t_{\leq}(r_i)$ gilt, und $h_0 := \sum_{i\in\mathbb{N}_0,\beta_i>0}\beta_i\overline{r}_i$. Dann ist das Signal S_{h_0} nichtfallend und konkav und man erhält $S_h \leq S_{h_2} \leq S_{h_0}$.

Beweis. Der Beweis entspricht im Wesentlichen dem Beweis zu Satz 3.17.
(a) Aus Lemma 3.16(b) folgt, dass das Signal S_{h_2} die kleinste nichtfallende konkave Majorante von S_h ist. Die Gleichheit in (3.11) folgt unmittelbar aus Lemma 3.14, da

$$\|p\tilde{V}^t_{\leq}(h)\| = \min\{\|g\| : g \in L_2(0,1), S_g \geq S_h\} = \min\{\|g\| : g \in \tilde{V}^t_{\leq}, S_g \geq S_h\}.$$

(b) Sei $g \in L_2(0,1)$ mit nichtfallendem konkavem Signal S_g. Dann liegt g im polaren Kegel $\tilde{V}^t_{\leq}$ nach Bemerkung 3.15(b). Sei umgekehrt $g \in \tilde{V}^t_{\leq}$ mit der orthogonalen Zerlegung $g = g_1 + g_2$ in $g_1 = pV^t_{\leq}(g)$ und $g_2 = p\tilde{V}^t_{\leq}(g)$. Wegen der Eindeutigkeit der orthogonalen Zerlegung ist $g = g_2$. Somit ist $S_g = S_{g_2}$ und S_g ist nichtfallend und konkav.
(c) Sei $h = \sum_{i=0}^{\infty}\beta_i r_i \in \tilde{H}$ mit $\beta_i = \langle h, r_i\rangle$. Das Signal $S_{\overline{r}_i}$ ist die kleinste konkave Majorante von S_{r_i} und folglich nach Teil (a) $\overline{r}_i = p\tilde{V}^t_{\leq}(r_i)$. Damit erhält man für $0 \leq t \leq 1$

$$\begin{aligned}
S_h(t) &= \int_0^t \sum_{i=0}^{\infty}\beta_i r_i \, d\lambda\!\!\lambda_{|[0,1]} = \lim_{n\to\infty}\sum_{i=0}^{n}\beta_i \int_0^t r_i \, d\lambda\!\!\lambda_{|[0,1]} \\
&\leq \lim_{n\to\infty}\sum_{\substack{i=0\\ \beta_i>0}}^{n}\beta_i \int_0^t r_i \, d\lambda\!\!\lambda_{|[0,1]} \\
&\leq \lim_{n\to\infty}\sum_{\substack{i=0\\ \beta_i>0}}^{n}\beta_i \int_0^t \overline{r}_i \, d\lambda\!\!\lambda_{|[0,1]} \\
&= S_{h_0}(t).
\end{aligned}$$

Also ist S_{h_0} eine konkave Majorante zu S_h. Da $S_{\overline{r}_i}$ nichtfallend und konkav ist, folgt, dass S_{h_0} als $L_2(\lambda\!\!\lambda_{|[0,1]})$-Limes ebenfalls nichtfallend und konkav ist (vgl. Satz A.1 aus dem Anhang). Da aber nach Teil (a) S_{h_2} die kleinste nichtfallende konkave Majorante zu S_h ist, folgt wieder insgesamt

$$S_h \leq S_{h_2} \leq S_{h_0}.$$

■

Aus Satz 3.17 und Satz 3.18 folgt:

Bemerkung 3.19
(a) Der polare Kegel $\tilde{V}^0_{\leq}$ ist die Menge aller nichtsteigenden Elemente $g \in L^0_2(0,1)$, welche ein konkaves Signal S_g besitzen.

(b) Der polare Kegel $\tilde{V}^t_{\leq}$ besteht aus allen nichtnegativen und nichtsteigenden Elementen $g \in L_2(0,1)$. Die zugehörigen Signale S_g sind dann nichtfallend und konkav. $\square$

3.4 Schranken für die Güte der Signalerkennungstests für von Rademacherfunktionen erzeugten Räume

In diesem Abschnitt werden Schranken für die Güte eines Signalerkennungstests (Definition 3.2) für die Teilhilberträume $\tilde{H} = \overline{\operatorname{span}(r_i : i \in \mathbb{N}_0)}$ und $\tilde{H}^0 = \tilde{H} \cap L_2^0(0,1)$ angegeben. Mit Hilfe der Sätze 3.17 und 3.18 kann nämlich die Güte für solche Regionen durch konkave (nichtfallende) Signale abgeschätzt werden. Man beschränkt sich auf Signalprozesse der Brownschen Brücke und der Brownschen Bewegung.

Der folgende Satz enthält einige der wichtigsten Resultate dieser Arbeit.

Satz 3.20

Sei φ ein einseitiger Signalerkennungstest zum Niveau $E_0(\varphi) := \alpha \in (0,1)$. Unter einer Alternative $h = \sum_{i=0}^{\infty} \beta_i r_i \in \tilde{H}$ erhält man folgende Ungleichungen für die Gütefunktion.
(a) Ist φ ein Test für den Signalprozess mit Brownscher Brücke als Rauschterm, so ist für $h \in \tilde{H}^0$

$$-\frac{1}{\sqrt{2\pi}} \sum_{\substack{i=1 \\ \beta_i<0}}^{\infty} |\beta_i| 2^{\frac{1-i}{2}} \leq E_h(\varphi) - \alpha \leq \frac{1}{\sqrt{2\pi}} \sum_{\substack{i=1 \\ \beta_i>0}}^{\infty} |\beta_i| 2^{\frac{1-i}{2}}.$$

(b) Ist φ ein Test für den Signalprozess mit Brownscher Bewegung als Rauschterm, so ist für $h \in \tilde{H}$

$$-\frac{1}{\sqrt{2\pi}} \sum_{\substack{i=0 \\ \beta_i<0}}^{\infty} |\beta_i| 2^{-\frac{i}{2}} \leq E_h(\varphi) - \alpha \leq \frac{1}{\sqrt{2\pi}} \sum_{\substack{i=0 \\ \beta_i>0}}^{\infty} |\beta_i| 2^{-\frac{i}{2}}.$$

Beweis. (a) Seien

$$\tilde{r}_i := r_i 1_{[2^{-i}, 1-2^{-i}]^c} \quad \text{und} \quad h_0 := \sum_{\substack{i=1 \\ \beta_i>0}}^{\infty} \beta_i \tilde{r}_i.$$

Nach Satz 3.17 ist $S_h \leq S_{h_0}$ und S_{h_0} konkav, d.h. $h_0 \in \tilde{V}^0_{\leq}$. Mit Lemma 3.9 folgt

$$E_h(\varphi) - \alpha \leq E_{h_0}(\varphi) - \alpha \leq \frac{\|p\tilde{V}^0_{\leq}(h_0)\|}{\sqrt{2\pi}} = \frac{\|h_0\|}{\sqrt{2\pi}}.$$

Wegen

$$\|h_0\| = \| \sum_{\substack{i=1 \\ \beta_i>0}}^{\infty} \beta_i \tilde{r}_i \| \le \sum_{\substack{i=1 \\ \beta_i>0}}^{\infty} \beta_i \|\tilde{r}_i\| = \sum_{\substack{i=1 \\ \beta_i>0}}^{\infty} \beta_i 2^{\frac{1-i}{2}}$$

erhält man die obere Schranke.

Nach Definition ist $S_{\tilde{r}_i}$ die kleinste konkave Majorante von S_{r_i}. Folglich ist $-S_{\tilde{r}_i}$ die größte konvexe Minorante von $-S_{r_i}$. Setzt man

$$g_0 := \sum_{\substack{i=1 \\ \beta_i<0}}^{\infty} \beta_i \tilde{r}_i,$$

so ist $S_{g_0} \le S_h$ und S_{g_0} ist als $L_2(\lambda\!\!\!\lambda_{|[0,1]})$-Limes konvexer Signale ebenfalls konvex. Deshalb ist $-g_0 \in \tilde{V}^0_{\le}$. Beachtet man, dass $p\tilde{V}_{\le}(-g) = -p\tilde{V}_{\ge}(g)$ für jedes $g \in L_2(0,1)$ gilt, so ergibt sich wieder mit Lemma 3.9

$$E_h(\varphi) - \alpha \ge E_{g_0}(\varphi) - \alpha \ge -\frac{\|p\tilde{V}^0_{\ge}(g_0)\|}{\sqrt{2\pi}} = -\frac{\|p\tilde{V}^0_{\le}(-g_0)\|}{\sqrt{2\pi}} = -\frac{\|g_0\|}{\sqrt{2\pi}}.$$

Aus der Abschätzung

$$\|g_0\| = \| \sum_{\substack{i=1 \\ \beta_i<0}}^{\infty} \beta_i \tilde{r}_i \| \le \sum_{\substack{i=1 \\ \beta_i<0}}^{\infty} |\beta_i| \|\tilde{r}_i\| = \sum_{\substack{i=1 \\ \beta_i<0}}^{\infty} |\beta_i| 2^{\frac{1-i}{2}}$$

folgt die untere Schranke.

(b) Der Beweis wird analog zum Beweis von Teil (a) geführt.

Seien

$$\overline{r}_i := r_i 1_{[0,2^{-i}]} \quad \text{und} \quad \overline{h}_0 := \sum_{\substack{i=0 \\ \beta_i>0}}^{\infty} \beta_i \overline{r}_i.$$

Nach Satz 3.18 ist $S_h \le S_{\overline{h}_0}$ und $S_{\overline{h}_0}$ nichtfallend und konkav, d.h. $\overline{h}_0 \in \tilde{V}^t_{\le}$. Mit Lemma 3.9 erhält man

$$E_h(\varphi) - \alpha \le E_{\overline{h}_0}(\varphi) - \alpha \le \frac{\|p\tilde{V}^t_{\le}(\overline{h}_0)\|}{\sqrt{2\pi}} = \frac{\|\overline{h}_0\|}{\sqrt{2\pi}}.$$

Wegen

$$\|\overline{h}_0\| = \| \sum_{\substack{i=0 \\ \beta_i>0}}^{\infty} \beta_i \overline{r}_i \| \le \sum_{\substack{i=0 \\ \beta_i>0}}^{\infty} \beta_i \|\overline{r}_i\| = \sum_{\substack{i=0 \\ \beta_i>0}}^{\infty} \beta_i 2^{-\frac{i}{2}}$$

folgt die obere Schranke.

Da $S_{\overline{r}_i}$ eine konkave Majorante von S_{r_i} ist, ist $-S_{\overline{r}_i}$ eine konvexe Minorante von $-S_{r_i}$.

Setzt man

$$\overline{g}_0 := \sum_{\substack{i=0 \\ \beta_i < 0}}^{\infty} \beta_i \overline{r}_i,$$

so ist $S_{\overline{g}_0} \leq S_h$ und $S_{\overline{g}_0}$ ist als $L_2(\lambda\!\!\lambda_{|[0,1]})$-Limes konvexer Signale ebenfalls konvex. Also ist $-\overline{g}_0 \in \tilde{V}^t_{\leq}$. Mit Lemma 3.9 ergibt sich

$$E_h(\varphi) - \alpha \geq E_{\overline{g}_0}(\varphi) - \alpha \geq -\frac{\|p\tilde{V}^t_{\geq}(\overline{g}_0)\|}{\sqrt{2\pi}} = -\frac{\|p\tilde{V}^t_{\leq}(-\overline{g}_0)\|}{\sqrt{2\pi}} = -\frac{\|\overline{g}_0\|}{\sqrt{2\pi}}.$$

Mit

$$\|\overline{g}_0\| = \|\sum_{\substack{i=0 \\ \beta_i < 0}}^{\infty} \beta_i \overline{r}_i\| \leq \sum_{\substack{i=0 \\ \beta_i < 0}}^{\infty} |\beta_i| \|\overline{r}_i\| = \sum_{\substack{i=1 \\ \beta_i < 0}}^{\infty} |\beta_i| 2^{-\frac{i}{2}}$$

erhält man die untere Schranke.

■

Der obige Satz 3.20 führt zum folgenden wichtigen Korollar.

Korollar 3.21
Sei φ ein einseitiger Signalerkennungstest zum Niveau $E_0(\varphi) := \alpha$.
(a) Für den Signalprozess mit Brownscher Brücke als Rauschterm bezeichne $W_k :=$ span $(r_1, ..., r_k)$. Dann ist

$$|E_h(\varphi) - \alpha| \leq \sqrt{\frac{1}{\pi}} \|h\| 2^{-\frac{k}{2}} \tag{3.12}$$

für jedes $h \in W_k^{\perp} \cap \tilde{H}^0$.
(b) Für den Signalprozess mit Brownscher Bewegung als Rauschterm erhält man

$$|E_h(\varphi) - \alpha| \leq \sqrt{\frac{1}{\pi}} \|h\| 2^{-\frac{k+1}{2}} \tag{3.13}$$

für jedes $h \in \tilde{W}_k^{\perp} \cap \tilde{H}$, wobei $\tilde{W}_k :=$span$(r_0, ..., r_k)$ ist.

Beweis. (a) Sei $h \in W_k^{\perp} \cap \tilde{H}^0$, d.h. $h = \sum_{i=1}^{\infty} \beta_i r_i$ mit $\beta_1 = ... = \beta_k = 0$. Nach Satz 3.20 ist

$$|E_h(\varphi) - \alpha| \leq \frac{1}{\sqrt{2\pi}} \sum_{i=1}^{\infty} |\beta_i| 2^{\frac{1-i}{2}}.$$

Mit der Cauchy-Schwarz-Ungleichung erhält man

$$
\begin{aligned}
\frac{1}{\sqrt{2\pi}} \sum_{i=1}^{\infty} |\beta_i| 2^{\frac{1-i}{2}} &\leq \frac{1}{\sqrt{2\pi}} \left(\sum_{i=1}^{\infty} \beta_i^2 \right)^{\frac{1}{2}} \left(\sum_{i=1}^{\infty} \left(2^{\frac{1-i}{2}} \right)^2 \right)^{\frac{1}{2}} \\
&= \frac{1}{\sqrt{2\pi}} \left(\sum_{i=k+1}^{\infty} \beta_i^2 \right)^{\frac{1}{2}} \left(\sum_{i=k+1}^{\infty} 2^{1-i} \right)^{\frac{1}{2}} \\
&= \frac{1}{\sqrt{2\pi}} \|h\| 2^{\frac{1}{2}} 2^{-\frac{k}{2}} \\
&= \frac{1}{\sqrt{\pi}} \|h\| 2^{-\frac{k}{2}}.
\end{aligned}
$$

(b) Sei $h \in \tilde{W}_k^{\perp} \cap \tilde{H}$, d.h. $h = \sum_{i=0}^{\infty} \beta_i r_i$ mit $\beta_0 = \beta_1 = ... = \beta_k = 0$. Nach Satz 3.20 ist

$$
\begin{aligned}
|E_h(\varphi) - \alpha| &\leq \frac{1}{\sqrt{2\pi}} \sum_{i=0}^{\infty} |\beta_i| 2^{-\frac{i}{2}} \\
&\leq \frac{1}{\sqrt{2\pi}} \left(\sum_{i=k+1}^{\infty} \beta_i^2 \right)^{\frac{1}{2}} \left(\sum_{i=k+1}^{\infty} 2^{-i} \right)^{\frac{1}{2}} \\
&= \frac{1}{\sqrt{2\pi}} \|h\| 2^{-\frac{k}{2}} \\
&= \frac{1}{\sqrt{\pi}} \|h\| 2^{-\frac{k+1}{2}}.
\end{aligned}
$$

■

Wie man in Korollar 3.21 sieht, kann man für einen einseitigen Signalerkennungstests endlichdimensionale Räume $W_k = \text{span}\ (r_1, ..., r_k)$ bzw. $\tilde{W}_k = \text{span}(r_0, ..., r_k)$ von Alternativen angeben, auf deren orthogonalem Komplement die Güte beliebig klein werden kann. In seiner Arbeit aus dem Jahre 2000 [16] hat JANSSEN bereits gezeigt, dass auf dem orthogonalen Komplement von endlichdimensionalen Räumen von Alternativen die Gütefunktion flach ist.

Theorem 3.22

Sei φ ein Test zum Niveau $E_{P_0}(\varphi) = \alpha \in (0,1)$ für die Nullhypothese P_0 eines Gauß-Shift-Experiments $E = (\Omega, \mathcal{A}, \{P_h : h \in H\})$. Für jedes $\varepsilon > 0$ und $K > 0$ existiert ein linearer Unterraum $V \subset H$ von endlicher Dimension mit

$$\sup\{|E_{P_h}(\varphi) - \alpha| : h \in V^\perp, \|h\| \leq K\} \leq \varepsilon.$$

Außerdem besitzt die Dimension von V folgende, vom Test φ unabhängige obere Schranke:

$$dim(V) - 1 \leq \varepsilon^{-1}\alpha(1-\alpha)(\exp(K^2) - 1).$$

Beweis. Vgl. JANSSEN [16, Theorem 2.1]. ■

Bemerkung 3.23

Sei φ ein einseitiger Signalerkennungstest für die Brownsche Brücke zum Niveau $E_0(\varphi) = \alpha \in (0,1)$. Für $h \in W_k^\perp \cap \tilde{H}^0$ und $t \in \mathbb{R}$ gilt mit (3.12) die Abschätzung

$$E_{th}(\varphi) - \alpha \leq \sqrt{\frac{1}{\pi}}\, |t| \|h\| 2^{-\frac{k}{2}} = |t|\|h\| M_k$$

mit $M_k := \sqrt{\frac{1}{\pi}} 2^{-\frac{k}{2}}$. Sei h_0 der Gradient des Tests φ. Es gilt für jedes $h \in W_k^\perp \cap \tilde{H}^0$:

$$\begin{aligned} \langle h, h_0 \rangle &= \frac{d}{dt} E_{th}(\varphi)_{|t=0} = \lim_{s \downarrow 0} \frac{E_{sh}(\varphi) - E_0(\varphi)}{s} = \lim_{s \downarrow 0} \frac{E_{sh}(\varphi) - \alpha}{s} \\ &\leq \lim_{s \downarrow 0} \frac{s\|h\| M_k}{s} = \|h\| M_k. \end{aligned}$$

Analog erhält man für $-h \in W_k^\perp \cap \tilde{H}^0$:

$$\langle -h, h_0 \rangle \leq \|h\| M_k.$$

Also ergibt sich insgesamt

$$|\langle h, h_0 \rangle| \leq \|h\| M_k.$$

Ebenso erhält man, falls φ ein einseitiger Signalerkennungstest für die Brownsche Bewegung ist, für jedes $h \in \tilde{W}_k^\perp \cap \tilde{H}$ mit (3.13) die Abschätzung

$$|\langle h, h_0 \rangle| \leq \|h\| N_k,$$

wobei $N_k := \sqrt{\frac{1}{\pi}} 2^{-\frac{k+1}{2}}$ ist. □

Bemerkung 3.24

Sei φ ein nichtrandomisierter Test für einen Gauß-Shift $E = (\Omega, \mathcal{A}, \{P_h : h \in H\})$ auf einem beliebigen Hilbertraum $(H, \langle \cdot, \cdot \rangle)$ mit dem zentralen Prozess $(L(h))_{h \in H}$ (vgl. Satz 1.10). Für $g, h \in H$ gilt dann

$$
\begin{aligned}
E_{g+h}(\varphi) &= \int \varphi \exp\left(L(g+h) - \frac{\|g+h\|^2}{2}\right) \, dP_0 \\
&= \exp\left(-\frac{\|g+h\|^2}{2}\right) \int \varphi \exp(L(g)) \varphi \exp(L(h)) \, dP_0 \\
&\leq \exp\left(-\frac{\|g+h\|^2}{2}\right) \left(\int (\varphi \exp(L(g)))^p \, dP_0\right)^{\frac{1}{p}} \left(\int (\varphi \exp(L(h)))^q \, dP_0\right)^{\frac{1}{q}} \\
&= \exp\left(-\frac{\|g+h\|^2}{2}\right) \left(\int \varphi \exp(L(pg)) \, dP_0\right)^{\frac{1}{p}} \left(\int \varphi \exp(L(qh)) \, dP_0\right)^{\frac{1}{q}} \\
&= \left(\int \varphi \exp\left(L(pg) - \frac{\|pg\|^2}{2}\right) \, dP_0\right)^{\frac{1}{p}} \left(\int \varphi \exp\left(L(qh) - \frac{\|qh\|^2}{2}\right) \, dP_0\right)^{\frac{1}{q}} \\
&\quad \cdot \exp\left(\frac{\|pg\|^2}{2p} + \frac{\|qh\|^2}{2q} - \frac{\|g+h\|^2}{2}\right) \\
&= (E_{pg}(\varphi))^{\frac{1}{p}} (E_{qh}(\varphi))^{\frac{1}{q}} \exp\left(\frac{p\|g\|^2 + q\|h\|^2 - \|g+h\|^2}{2}\right)
\end{aligned}
$$

nach der Hölder-Ungleichung für $p, q \in (1, \infty)$ mit $\frac{1}{p} + \frac{1}{q} = 1$. □

Bemerkung 3.24 führt zur folgenden Anwendung von Korollar 3.21.

Anwendung 3.25

Sei φ ein einseitiger nichtrandomisierter oberer Test zum Niveau $E_0(\varphi) = \alpha$ für den Signalprozess der Brownschen Brücke. Die orthogonale Zerlegung von $h \in \tilde{H}^0 = \overline{\text{span}(r_i : i \in \mathbb{N})}$ sei gegeben durch $h = h_1 + h_2$ mit $h_1 = pW_k(h), h_2 = p(W_k^\perp \cap \tilde{H}^0)(h)$, wobei $W_k = \text{span}$ $(r_1, ..., r_k)$ ist. Sei $1 = \|h\|^2 = \|h_1\|^2 + \|h_2\|^2$. Weiterhin nehme man $\|h_1\| \in (0, 1)$ an und setze $p := \frac{1}{\|h_1\|}$. Sei q so gewählt, dass $\frac{1}{p} + \frac{1}{q} = 1$ gilt, also $q = \frac{p}{p-1} = \frac{1}{1-\|h_1\|}$. Es folgt dann

$p\|h_1\|^2 + q\|h_2\|^2 - \|h_1 + h_2\|^2 = 2\|h_1\|$. Man erhält mit Bemerkung 3.24

$$\begin{aligned} E_h(\varphi) &\leq (E_{ph_1}(\varphi))^{\frac{1}{p}} (E_{qh_2}(\varphi))^{\frac{1}{q}} \exp\left(\frac{p\|h_1\|^2 + q\|h_2\|^2 - \|h_1+h_2\|^2}{2}\right) \\ &\overset{(3.12)}{\leq} (E_{ph_1}(\varphi))^{\frac{1}{p}} \left(\sqrt{\frac{1}{\pi}}\, q\|h_2\|\; 2^{-\frac{k}{2}} + \alpha\right)^{\frac{1}{q}} \exp(\|h_1\|) \\ &= (E_{\frac{h_1}{\|h_1\|}}(\varphi))^{\|h_1\|} \left(\sqrt{\frac{1}{\pi}} \sqrt{\frac{1+\|h_1\|}{1-\|h_1\|}}\; 2^{-\frac{k}{2}} + \alpha\right)^{1-\|h_1\|} \exp(\|h_1\|). \end{aligned}$$

Sei φ ein einseitiger nichtrandomisierter oberer Test zum Niveau $E_0(\varphi) = \alpha$ für den Signalprozess der Brownschen Bewegung. Man erhält analog zum Fall der Brownschen Brücke die Abschätzung

$$E_h(\varphi) \leq (E_{\frac{h_1}{\|h_1\|}}(\varphi))^{\|h_1\|} \left(\sqrt{\frac{1}{\pi}} \sqrt{\frac{1+\|h_1\|}{1-\|h_1\|}}\; 2^{-\frac{k+1}{2}} + \alpha\right)^{1-\|h_1\|} \exp(\|h_1\|).$$

□

Als Nächstes wird gezeigt, dass die Güte eines einseitigen Signalerkennungstests auf Bällen von Alternativen ihr Maximum annimmt, und zwar für eine Alternative aus dem polaren Kegel. Dazu wird unter anderem die Normstetigkeit der Gütefunktionen für Gauß-Shift-Experimente benutzt.

Hilfssatz 3.26

Sei φ ein Test für einen Gauß-Shift $E = (\Omega, \mathcal{A}, \{P_h : h \in H\})$ auf einem beliebigen Hilbertraum $(H, \langle\cdot,\cdot\rangle)$. Dann folgt für jede Folge $h_n \to h$ in H die Konvergenz der Gütefunktion

$$E_{h_n}(\varphi) \to E_h(\varphi) \quad \textit{für} \quad n \to \infty.$$

Beweis. Es bezeichne für $g, h \in H$

$$d_1(P_g, P_h) := \|P_g - P_h\| := \frac{1}{2}\left\|\frac{dP_g}{dP_0} - \frac{dP_h}{dP_0}\right\|_{L_1(P_0)} = \frac{1}{2}\int \left|\frac{dP_g}{dP_0} - \frac{dP_h}{dP_0}\right| dP_0$$

die Norm der Totalvariation zwischen P_g und P_h und

$$d_2(P_g, P_h) := \frac{1}{\sqrt{2}}\left\| \left(\frac{dP_g}{dP_0}\right)^{\frac{1}{2}} - \left(\frac{dP_g}{dP_0}\right)^{\frac{1}{2}} \right\|_{L_2(P_0)} = \frac{1}{\sqrt{2}}\left[\int \left(\left(\frac{dP_g}{dP_0}\right)^{\frac{1}{2}} - \left(\frac{dP_h}{dP_0}\right)^{\frac{1}{2}}\right)^2\right]^{\frac{1}{2}}$$

die Hellingerdistanz zwischen P_g und P_h.

Nach STRASSER [41, Lemma 2.15] gilt die Abschätzung

$$d_1(P_g, P_h) \leq [d_2^2(P_g, P_h)(2 - d_2^2(P_g, P_h))]^{\frac{1}{2}}. \tag{3.14}$$

Weiterhin ist die Hellingerdistanz zwischen P_g und P_h nach STRASSER [41, Remark 69.8] gegeben durch

$$d_2^2(P_g, P_h) = 1 - \exp\left(-\frac{1}{8}\|g-h\|^2\right).$$

Aus $h_n \to h$ folgt

$$d_2^2(P_{h_n}, P_h) = 1 - \exp\left(-\frac{1}{8}\|h_n - h\|^2\right) \to 0 \quad \text{für} \quad n \to \infty$$

und somit wegen (3.14)

$$d_1(P_{h_n}, P_h) \to 0 \quad \text{für} \quad n \to \infty.$$

Also ergibt sich

$$\begin{aligned}
|E_{h_n}(\varphi) - E_h(\varphi)| &= \left| \int \varphi dP_{h_n} - \int \varphi dP_h \right| \\
&= \left| \int \varphi \frac{dP_{h_n}}{dP_0} dP_0 - \int \varphi \frac{dP_h}{dP_0} dP_0 \right| \\
&\leq \int |\varphi| \left| \frac{dP_{h_n}}{dP_0} - \frac{dP_h}{dP_0} \right| dP_0 \\
&\leq \int \left| \frac{dP_{h_n}}{dP_0} - \frac{dP_h}{dP_0} \right| dP_0 \\
&\to 0 \quad \text{für} \quad n \to \infty.
\end{aligned}$$

■

Sei an folgende Definitionen erinnert:

$$\tilde{H} := \overline{\operatorname{span}(r_i : i \in \mathbb{N}_0)} \quad \text{und} \quad \tilde{H^0} := \tilde{H} \cap L_2^0(0,1)$$

mit den abgeschlossenen konvexen Kegeln

$$V_{\leq}(\tilde{H}) := \{g \in \tilde{H} : S_g \leq 0\} \quad \text{und} \quad V_{\leq}(\tilde{H^0}) = V_{\leq}(\tilde{H}) \cap L_2^0(0,1)$$

und den zugehörigen polaren Kegeln $\tilde{V}_{\leq}(\tilde{H})$ bzw. $\tilde{V}_{\leq}(\tilde{H^0})$.

Satz 3.27

Sei φ ein einseitiger oberer Test für den Signalprozess der Brownschen Brücke (oder der Brownschen Bewegung). Der Test φ erreicht für jedes $K > 0$ die maximale Güte auf Bällen $\{h \in H : \|h\| \leq K\}$ von Alternativen mit $H = \tilde{H^0}$ (bzw. $H = \tilde{H}$). Auf diesen Bällen ist die Gütefunktion maximal für ein $h \in \tilde{V}_{\leq}(\tilde{H^0})$ (bzw. $h \in \tilde{V}_{\leq}(\tilde{H})$).

Beweis. Der Beweis wird für den Signalprozess der Brownschen Brücke ausgeführt. Sei $a := \sup\{E_h(\varphi) : h \in \tilde{H}^0, \|h\| \leq K\}$. Dann existiert eine Folge $(h_n)_{n\in\mathbb{N}} \in \tilde{H}^0$ mit $E_{h_n}(\varphi) \to a$ für $n \to \infty$ und $\|h_n\| \leq K$ für jedes $n \in N$. Sei $h_n = \sum_{i=1}^{\infty} \beta_{i,n} r_i$ für $n \in \mathbb{N}$ die Darstellung von h_n durch die Rademacherfunktionen (3.9). Wegen

$$\sum_{i\in\mathbb{N}} \beta_{i,n}^2 = \|h_n\|^2 \leq K^2, \tag{3.15}$$

folgt $|\beta_{i,n}| \leq K$. Ohne Beschränkung der Allgemeinheit sei $\beta_{i,n} \geq 0$ für jedes $i, n \in \mathbb{N}$. Ansonsten geht man zu $|\beta_{i,n}|$ über, ohne, dass sich die Norm verändert. Das Signal wird nicht kleiner. Durch eine diagonale Teilfolgenauswahl erhält man eine Teilfolge $\{\beta_{i,m}\} \subset \{\beta_{i,n}\}$, so dass $(\beta_{i,m})_{m\in\mathbb{N}}$ für jedes $i \in \mathbb{N}$ konvergiert mit $\beta_i := \lim_{m\to\infty} \beta_{i,m}$. Setze nun für festes $k \in \mathbb{N}$

$$g_k := \sum_{i=1}^{k} \beta_i r_i + \sum_{i=k+1}^{\infty} \beta_i \tilde{r}_i \text{ und } g := \sum_{i\in\mathbb{N}} \beta_i r_i \tag{3.16}$$

mit $\tilde{r}_i = r_i 1_{[2^{-i}, 1-2^{-i}]^c}$ aus Satz 3.17. Wegen (3.15) ist $\left(\sum_{i\in\mathbb{N}} \beta_{i,n}^2\right)^{1/2} \leq K$ und folglich auch $\left(\sum_{i=1}^{k} \beta_i^2\right)^{1/2} \leq K$ für jedes $k \in \mathbb{N}$ und somit $\|g\| \leq K$. Man betrachte nun für festes $k \in \mathbb{N}$

$$f_m := \sum_{i=1}^{k} \beta_{i,m} r_i + \sum_{i=k+1}^{\infty} \beta_{i,m} \tilde{r}_i.$$

Analog zum Beweis von Satz 3.17 erhält man für $0 \leq t \leq 1$

$$\begin{aligned}
S_{h_m}(t) &= \sum_{i=1}^{\infty} \beta_{i,m} \int_0^t r_i \; d\lambda\!\!\lambda_{|[0,1]} \\
&= \sum_{i=1}^{k} \beta_{i,m} \int_0^t r_i \; d\lambda\!\!\lambda_{|[0,1]} + \sum_{i=k+1}^{\infty} \beta_{i,m} \int_0^t r_i \; d\lambda\!\!\lambda_{|[0,1]} \\
&\leq \sum_{i=1}^{k} \beta_{i,m} \int_0^t r_i \; d\lambda\!\!\lambda_{|[0,1]} + \sum_{i=k+1}^{\infty} \beta_{i,m} \int_0^t \tilde{r}_i \; d\lambda\!\!\lambda_{|[0,1]} \\
&= \int_0^t \left(\sum_{i=1}^{k} \beta_{i,m} r_i + \sum_{i=k+1}^{\infty} \beta_{i,m} \tilde{r}_i \right) d\lambda\!\!\lambda_{|[0,1]} \\
&= S_{f_m}(t).
\end{aligned}$$

Folglich ist

$$E_{h_m}(\varphi) \leq E_{f_m}(\varphi). \tag{3.17}$$

Andererseits erhält man für festes $k \in \mathbb{N}$

$$\begin{aligned}
\|f_m - g_k\| &= \|\sum_{i=1}^{k} \beta_{i,m} r_i + \sum_{i=k+1}^{\infty} \beta_{i,m} \tilde{r}_i - \sum_{i=1}^{k} \beta_i r_i - \sum_{i=k+1}^{\infty} \beta_i \tilde{r}_i\| \\
&\leq \|\sum_{i=1}^{k} \beta_{i,m} r_i - \sum_{i=1}^{k} \beta_i r_i\| + \|\sum_{i=k+1}^{\infty} \beta_{i,m} \tilde{r}_i - \sum_{i=k+1}^{\infty} \beta_i \tilde{r}_i\| \\
&\leq \sum_{i=1}^{k} |\beta_{i,m} - \beta_i| \|r_i\| + \sum_{i=k+1}^{\infty} |\beta_{i,m} - \beta_i| \|\tilde{r}_i\| \\
&= \sum_{i=1}^{k} |\beta_{i,m} - \beta_i| + \sum_{i=k+1}^{\infty} |\beta_{i,m} - \beta_i| 2^{\frac{1-i}{2}} \\
&\to 0 \ \text{ für } \ m \to \infty. \qquad (3.18)
\end{aligned}$$

Für Gauß-Shift-Experimente ist nach Hilfsatz 3.26 bekannt, dass die Gütefunktionen $h \mapsto E_h(\varphi)$ stetig bzgl. der Norm des zugehörigen Hilbertraumes sind. Deshalb gilt mit (3.17) und (3.18)

$$a = \lim_{n \to \infty} E_{h_n}(\varphi) \leq \lim_{n \to \infty} E_{f_n}(\varphi) = E_{g_k}(\varphi).$$

Weiterhin folgt

$$\begin{aligned}
\|g_k - g\| &= \|\sum_{i=1}^{k} \beta_i r_i + \sum_{i=k+1}^{\infty} \beta_i \tilde{r}_i - \sum_{i \in \mathbb{N}} \beta_i r_i\| \\
&= \|\sum_{i=1}^{k} \beta_i r_i - \sum_{i=1}^{k} \beta_i r_i + \sum_{i=k+1}^{\infty} \beta_i \tilde{r}_i - \sum_{i=k+1}^{\infty} \beta_i r_i\| \\
&= \|\sum_{i=k+1}^{\infty} \beta_i r_i - \sum_{i=k+1}^{\infty} \beta_i \tilde{r}_i\| \\
&\leq \|\sum_{i=k+1}^{\infty} \beta_i r_i\| + \|\sum_{i=k+1}^{\infty} \beta_i \tilde{r}_i\| \\
&\leq \left(\sum_{i=k+1}^{\infty} \|\beta_i r_i\|^2 \right)^{\frac{1}{2}} + \sum_{i=k+1}^{\infty} |\beta_i| \|\tilde{r}_i\| \\
&= \left(\sum_{i=k+1}^{\infty} |\beta_i|^2 \right)^{\frac{1}{2}} + \sum_{i=k+1}^{\infty} |\beta_i| 2^{\frac{1-i}{2}} \\
&\to 0 \ \text{ für } \ k \to \infty.
\end{aligned}$$

Daraus folgt

$$a = \lim_{k\to\infty} E_{g_k}(\varphi) = E_g(\varphi),$$

da $g \in \tilde{H}^0$ (nach Definition (3.16)) mit $\|g\| \leq K$.
Es muss noch gezeigt werden, dass auf den Bällen $\{h \in \tilde{H}^0 : \|h\| \leq K\}$ die Gütefunktion von φ maximal ist für ein $h \in \tilde{V}_{\leq}(\tilde{H}^0)$. Betrachte dazu die orthogonale Zerlegung von $g = \overline{g}_1 + \overline{g}_2 \in \tilde{H}^0$ in $\overline{g}_1 = pV_{\leq}(\tilde{H}^0)(g)$ und $\overline{g}_2 = p\tilde{V}_{\leq}(\tilde{H}^0)(g)$. Nach Satz 3.17 ist $S_{\overline{g}_2}$ die kleinste konkave Majorante von S_g und somit

$$E_g(\varphi) \leq E_{\overline{g}_2}(\varphi).$$

Andererseits erhält man

$$E_{\overline{g}_2}(\varphi) \leq a = \sup\{E_h(\varphi) : h \in \tilde{H}^0, \|h\| \leq K\} = E_g(\varphi),$$

da $\|\overline{g}_2\|^2 = \|g\|^2 - \|\overline{g}_1\|^2 \leq \|g\|^2 \leq K^2$ gilt. Folglich ist

$$E_{\overline{g}_2}(\varphi) = E_g(\varphi) = a$$

mit $\overline{g}_2 \in \tilde{V}_{\leq}(\tilde{H}^0)$. ∎

Kapitel 4

Die Güte zweiseitiger Signalerkennungstests

In diesem Kapitel wird die Gütefunktion zweiseitiger Signalerkennungstests untersucht. Es wird das Testproblem aus (2.3)

$$H_2 : F = F_0 \ \text{ gegen } \ K_2 : F \neq F_0,$$

oder durch Signale ausgedrückt

$$H_2 = \{g \in H : S_g = 0\} \ \text{ gegen } \ K_2 = \{g \in H : S_g \neq 0\},$$

betrachtet mit dem Hilbertraum H.
Hierfür werden die Begriffe gespiegelter und symmetrischer Test eingeführt. Dazu sei $(Y(t))_{t\in\mathbb{R}}$ ein Signalprozess bzgl. eines Gauß-Shift-Experiments $(\Omega, \mathcal{A}, \{P_h : h \in H\})$ auf einem beliebigen Hilbertraum $(H, \langle\cdot,\cdot\rangle)$, vgl. Definition 1.18.

Definition 4.1 (Gespiegelter und symmetrischer Signalerkennungstest)
Gegeben sei ein einseitiger oberer Test $\varphi = \varphi(Y(\cdot)) : C(\mathbb{R}) \to \{0,1\}$ für den Signalprozess $(Y(t))_{t\in\mathbb{R}}$ auf der Pfadmenge der stetigen Funktionen $C(\mathbb{R})$ (vgl. Definition 3.2).
(a) Es bezeichne

$$\overline{\varphi}(y(\cdot)) := \varphi(-y(\cdot))$$

die Spiegelung des Tests φ.
(b) Ein nichtrandomisierter Test $\psi = \psi(Y(\cdot))$ heißt ein symmetrischer zweiseitiger Signalerkennungstest, falls ψ von der folgenden Gestalt ist:

$$\psi = \varphi + \overline{\varphi} - \varphi\overline{\varphi}$$

Für die Gütefunktionen ergeben sich folgende Relationen.

Bemerkung 4.2
Gegeben sei ein einseitiger oberer Test $\varphi = \varphi(Y(\cdot))$ mit $\overline{\varphi}$ und ψ zum Signalprozess $(Y(t))_{t\in\mathbb{R}}$ aus Definition 4.1. Dann gilt für jedes $h \in H$:
(a)

$$E_h(\psi) \leq E_h(\varphi) + E_h(\overline{\varphi}), \tag{4.1}$$

(b)

$$E_h(\overline{\varphi}) = E_{-h}(\varphi). \tag{4.2}$$

□

Beweis. (a) Die Behauptung folgt unmittelbar aus $A(\psi) = A(\varphi) \cap A(\overline{\varphi})$, wobei $A(\eta) := \{\eta = 0\}$ den Annahmebereich eines Tests η bezeichnet.
(b) Es ist aufgrund der Symmetrieeigenschaft der zentrierten Normalverteilung

$$\begin{aligned}
E_h(\varphi) &= P_0(\{Z(F_0(\cdot)) + S_h(F_0(\cdot)) \in C(\varphi)\}) \\
&= P_0(\{-Z(F_0(\cdot)) + S_h(F_0(\cdot)) \in C(\varphi)\}) \\
&= P_0(\{-Z(F_0(\cdot)) - S_{-h}(F_0(\cdot)) \in C(\varphi)\}) \\
&= P_0(\{-[Z(F_0(\cdot)) + S_{-h}(F_0(\cdot))] \in C(\varphi)\}) \\
&= P_0(\{Z(F_0(\cdot)) + S_{-h}(F_0(\cdot)) \in C(\overline{\varphi})\}) \\
&= E_{-h}(\overline{\varphi}).
\end{aligned}$$

■

Nach Bemerkung 4.2 muß für Alternativen aus dem Raum $W_k^{\perp} \cap \tilde{H}^0$ mit $W_k =\text{span}(r_1, \ldots, r_k)$ und $\tilde{H}^0 = \overline{\text{span}(r_i : i \in \mathbb{N})}$, wo die Güte der Tests φ und $\overline{\varphi}$ klein ist, auch der zweiseitige Signalerkennungstest ψ von geringer Güte sein. Beschränkt man sich auf den Signalprozess der Brownschen Brücke, so erhält man für den zweiseitigen Test ψ ähnliche obere Schranken wie in Satz 3.20 und Korollar 3.21 für den einseitigen Test φ mit einem additiven Term $a(\alpha)$. Wie man im nächsten Satz sehen wird, kann $a(\alpha)$ für ein hinreichend kleines Signifikanzniveau α sehr gering werden.
Zuvor wird ein Lemma formuliert, welches im Beweis zu Satz 4.4 verwendet wird.

Lemma 4.3

Seien $X_1, \ldots, X_n$ unabhängige Zufallsvariable und $f(x_1, \ldots, x_n)$ und $g(x_1, \ldots, x_n)$ reelle Funktionen, die in jedem Argument x_i, $1 \leq i \leq n$, nichtfallend sind. Dann ist

$$Cov(f(X_1, \ldots, X_n), g(X_1, \ldots, X_n)) \geq 0,$$

falls dieser existiert.

Beweis. Vgl. HÁJEK [10, Lemma 3.1] oder HÁJEK et al. [11, Lemma 7.5.2.1]. ■

Satz 4.4

Sei ψ ein symmetrischer zweiseitiger Test zum Niveau $E_0(\psi) = \alpha \in (0,1)$ für den Signalprozess der Brownschen Brücke, gegeben durch den einseitigen Signalerkennungstest $\varphi : C_0(\mathbb{R}) \to \{0,1\}$. Sei der Rand $\partial\{\varphi = 1\}$ des kritischen Bereichs des Tests φ eine Nullmenge, d.h. $P_0(\{B_0(F_0(\cdot)) \in \partial\{\varphi(Y(\cdot)) = 1\}\}) = 0$. Dann gelten

(a) Für jedes $h \in \tilde{H}^0$, gegeben durch $h = \sum_{i=1}^{\infty} \beta_i r_i$, erhält man

$$E_h(\psi) - \alpha \leq \frac{1}{\sqrt{2\pi}} \sum_{i=1}^{\infty} |\beta_i| 2^{\frac{1-i}{2}} + a(\alpha)$$

mit $a(\alpha) := 2(1 - \sqrt{1-\alpha}) - \alpha$.

(b) Für jedes $h \in W_k^\perp \cap \tilde{H}^0$, wobei $W_k^\perp$ den Orthogonalraum zu $W_k :=$ span $(r_1, ..., r_k)$ bzgl. $\tilde{H}^0$ bezeichnet, ist

$$E_h(\psi) - \alpha \leq \frac{1}{\sqrt{\pi}} \|h\| 2^{-\frac{k}{2}} + a(\alpha).$$

Beweis. (a) Sei $\alpha' := E_0(\varphi) \overset{(4.2)}{=} E_0(\overline{\varphi})$. Dann folgt für jedes $h \in \tilde{H}^0,\ h = \sum_{i \in \mathbb{N}} \beta_i r_i$ mit Satz 3.20

$$\begin{aligned}
E_h(\psi) - \alpha &\leq E_h(\varphi) + E_h(\overline{\varphi}) - \alpha \\
&\overset{(4.1)}{=} (E_h(\varphi) - \alpha') + (E_h(\overline{\varphi}) - \alpha') + (2\alpha' - \alpha) \\
&\overset{(4.2)}{=} (E_h(\varphi) - \alpha') + (E_{-h}(\varphi) - \alpha') + (2\alpha' - \alpha) \\
&\leq \frac{1}{\sqrt{2\pi}} \sum_{i \in \mathbb{N}} |\beta_i| 2^{\frac{1-i}{2}} + (2\alpha' - \alpha).
\end{aligned}$$

Dabei wird für h die Abschätzung für positive Koeffizienten und für $-h$ die Abschätzung für negative Koeffizienten gebraucht. Die Abschätzung $2\alpha' - \alpha \leq a(\alpha)$ folgt aus folgenden Überlegungen:

Sei $(X_i)_{i \in \mathbb{N}}$ eine Folge i.i.d. verteilter Zufallsvariablen mit gemeinsamer Verteilungsfunktion F_0. Es bezeichne $X_{1:n} \leq X_{2:n} \leq X_{3:n} \leq ...$ die Orderstatistik zu $X_1, ..., X_n$. Man erhält eine stetige Modifikation $\tilde{F}_n$ der empirischen Verteilungsfunktion $\hat{F}_n$ von $X_1, ..., X_n$, indem man

$$\tilde{F}_n(X_{i:n}) := \hat{F}_n(X_{i:n})$$

setzt und $\tilde{F}_n$ stetig und linear zwischen $X_{i:n}$ und $X_{i+1:n}$ definiert sowie $X_{0:n} := X_{1:n} - 1$ und $X_{n+1:n} := X_{n:n} + 1$ setzt, vgl. SHORACK und WELLNER [40].
Das Invarianzprinzip für empirische Prozesse liefert die schwache Konvergenz

$$\sqrt{n}(\tilde{F}_n - F_0) \ \xrightarrow[P_0]{\mathcal{D}} \ B_0 \circ F_0 \ \text{ für } \ n \to \infty \tag{4.3}$$

auf $C_0(\mathbb{R})$. Weiterhin gelten

$$\begin{aligned} \partial\{\varphi\overline{\varphi}(Y(\cdot)) = 1\} &= \partial\{\{\varphi(Y(\cdot)) = 1\} \cap \{\overline{\varphi}(Y(\cdot)) = 1\}\} \\ &\subseteq \partial\{\varphi(Y(\cdot)) = 1\} \cup \partial\{\overline{\varphi}(Y(\cdot)) = 1\} \end{aligned}$$

und

$$\begin{aligned} P_0(\{B_0(F_0(\cdot)) \in \partial\{\overline{\varphi}(Y(\cdot)) = 1\}) &= P_0(\{B_0(F_0(\cdot)) \in \partial\{\varphi(-Y(\cdot)) = 1\}) \\ &= P_0(\{B_0(F_0(\cdot)) \in -\partial\{\varphi(Y(\cdot)) = 1\}) \\ &= P_0(\{-B_0(F_0(\cdot)) \in \partial\{\varphi(Y(\cdot)) = 1\}) \\ &= P_0(\{B_0(F_0(\cdot)) \in \partial\{\varphi(Y(\cdot)) = 1\}) \\ &= 0 \end{aligned}$$

nach Voraussetzung, d.h. die fast sichere Stetigkeit von φ impliziert die fast sichere Stetigkeit von $\overline{\varphi}$. Daraus folgt

$$\begin{aligned} & P_0(\{B_0(F_0(\cdot)) \in \partial\{\varphi\overline{\varphi}(Y(\cdot)) = 1\}) \\ \leq\ & P_0(\{B_0(F_0(\cdot)) \in \partial\{\varphi(Y(\cdot)) = 1\}) + P_0(\{B_0(F_0(\cdot)) \in \partial\{\overline{\varphi}(Y(\cdot)) = 1\}) \\ =\ & 0. \end{aligned}$$

Damit erhält man auch die fast sichere Stetigkeit von $\varphi\overline{\varphi}$ auf $C_0(\mathbb{R})$ bzgl. der Supremumsnorm. Da $\varphi, \overline{\varphi}$ und $\varphi\overline{\varphi}$ zusätzlich beschränkt sind, folgt mit der schwachen Konvergenz des empirischen Prozesses (4.3) die Konvergenz

$$Cov(\varphi(\sqrt{n}(\tilde{F}_n - F_0)), \overline{\varphi}(\sqrt{n}(\tilde{F}_n - F_0))) \to Cov(\varphi(B_0 \circ F_0), \overline{\varphi}(B_0 \circ F_0)),$$

vgl. SHORACK und WELLNER [40]. Seien nun $x_1, ..., x_n$ Realisierungen von $X_1, ..., X_n$. Betrachtet man nur x_i und hält die übrigen Argumente $x_1, \dots, x_{i-1}, x_{i+1}, \dots x_n$ fest, so stellt man fest, dass mit steigendem x_i die Abbildung

$$x_i \mapsto \sqrt{n}(\tilde{F}_n - F_0)$$

fallend ist und somit die Abbildungen

$$x_i \mapsto \varphi(\sqrt{n}(\tilde{F}_n - F_0))$$

und

$$x_i \mapsto 1 - \overline{\varphi}(\sqrt{n}(\tilde{F}_n - F_0))$$

ebenfalls fallend sind nach Definition des einseitigen Signalerkennungstests (vgl. Definition (3.2)). Mit Lemma 4.3 folgt

$$\begin{aligned} & Cov(\varphi(B_0 \circ F_0), \overline{\varphi}(B_0 \circ F_0)) \leq 0 \\ \Longleftrightarrow \quad & E_0(\varphi\overline{\varphi}) - E_0(\varphi)E_0(\overline{\varphi}) \leq 0 \\ \Longleftrightarrow \quad & 2\alpha' - \alpha \leq (\alpha')^2 \\ \Longleftrightarrow \quad & \alpha' \leq 1 - \sqrt{(1-\alpha)}. \end{aligned}$$

Also

$$2\alpha' - \alpha \leq 2(1 - \sqrt{1-\alpha}) - \alpha = a(\alpha).$$

(b) Sei $h \in W_k^{\perp} \cap \tilde{H}^0$, $h = \sum_{i=1}^{\infty} \beta_i r_i$ mit $\beta_1 = ... = \beta_k = 0$. Dann folgt mit Teil (a)

$$\begin{aligned} E_h(\psi) - \alpha \quad & \leq \quad \frac{1}{\sqrt{2\pi}} \sum_{i=1}^{\infty} |\beta_i| 2^{\frac{1-i}{2}} + a(\alpha) \\ & \leq \quad \frac{1}{\sqrt{\pi}} \|h\| 2^{-\frac{k}{2}} + a(\alpha) \end{aligned}$$

(vgl. Beweis zu Korollar 3.21(a)). ■

Ähnliche Ergebnisse wie im Satz 4.4 erhält man auch für den Signalprozess der Brownschen Bewegung.

Kapitel 5

Die statistische Interpretation

In diesem Kapitel wird die Bedeutung der konkaven Signale und deren Tagenten für parametrische Teilfamilien, insbesondere für Lokationsfamilien erläutert. Danach erfolgt eine Interpretation nichtparametrischer statistischer Funktionale, die Linearkombinationen von Quantilfunktionen sind.

5.1 Bedeutung der konkaven Signale und deren Tangenten für parametrische Modelle

Dieser Abschnitt richtet sich nach dem Buch von HÁJEK et al. [11], welches die Theorie der Rangtests ausführlich darstellt.

Definition 5.1 (**Streng unimodal**)
Eine absolut stetige Dichtefunktion $f : (a, b) \to \mathbb{R}$ *mit* $-\infty \leq a < b \leq \infty$ *heißt streng unimodal, falls* $x \mapsto -\log(f(x))$ *eine konvexe Funktion ist, d.h., falls*

$$x \mapsto (-\log f(x))' = -\frac{f'(x)}{f(x)}$$

nichtfallend ist.

Lemma 5.2
Sei $g \in L_2^0(0,1), g \neq 0,$ *eine nichtsteigende Tangente. Für festes* $c \in (0,1)$ *existiert eine Verteilungsfunktion* F *mit* $F(0) = c$ *und eine streng unimodale Dichtefunktion* f*, welche gegeben ist durch die inverse Verteilungsfunktion*

$$F^{-1}(u) = \int_c^u \frac{1}{S_g(t)}\, dt, \quad 0 < u < 1. \tag{5.1}$$

Die zugehörige Lokationsfamilie $\vartheta \mapsto F * \varepsilon_{-\vartheta}$ mit der Dichte $x \mapsto f(x+\vartheta)$ hat die score erzeugende Funktion $g(u) = \frac{f'}{f} \circ F^{-1}(u)$.

Beweis. Da g eine nichtsteigende Tangente ist, besitzt sie ein konkaves Signal S_g. Zusammen mit $S_g \neq 0$ folgt aus der Konkavität $S_g(t) > 0$ für alle $t \in (0,1)$. Damit folgt

$$0 < S_g(t) = \int_0^t g(u)\, du = \int_t^1 (-g(u))\, du \quad \forall t \in (0,1).$$

Nach HÁJEK et al. [11, Lemma 2.2.4.6, S.19f.] erfüllt die durch (5.1) gegebene Verteilung die Eigenschaften $F(0) = c$ und $g(u) = \frac{f'}{f} \circ F^{-1}(u)$. ■

Bemerkung 5.3

Die durch (5.1) gegebene Verteilung ist sogar eindeutig.

Beispiel 5.4

Die folgenden nichtsteigenden Tangenten $g \in L_2^0(0,1)$ führen zu streng unimodalen Dichten f mit $g(u) = \frac{f'}{f} \circ F^{-1}(u)$.

(a) Seien $a > 0, b > 0$ und $p := \frac{a}{a+b}$. Die Tangente $g := \frac{1}{a}\mathbb{1}_{[0,p]} - \frac{1}{b}\mathbb{1}_{(p,1]}$ führt zu der Dichte

$$f(x) = \frac{1}{a+b}\left[\exp\left(\frac{x}{a}\right)\mathbb{1}_{(-\infty,0]}(x) + \exp\left(-\frac{x}{b}\right)\mathbb{1}_{(0,\infty)}(x)\right].$$

Es gilt $F(0) = p$ ist. Im Fall $a = b = 1$ erhält man die Dichte der Doppel-Exponentialverteilung.

(b) Die Tangente $g := \tilde{r}_i := r_i \mathbb{1}_{[2^{-i},1-2^{-i}]^c}$ (vgl. Satz 3.17) führt für $i \geq 2$ zu der achsensymmetrischen Dichte

$$f(x) = a\exp(x)\mathbb{1}_{(-\infty,-x_0)}(x) + a\mathbb{1}_{[-x_0,x_0]}(x) + a\exp(-x)\mathbb{1}_{(x_0,\infty)}(x),$$

wobei x_0 die Lösung der Gleichung $\exp(-x_0) = \frac{x_0}{2^{i-1}-1}$ ist und $a := \frac{2^{i-1}-1}{x_0 2^i}$ gesetzt wird. Es gilt $F(0) = \frac{1}{2}$ ist.

Beweis. (a) Es ist

$$f'(x) = \frac{1}{a+b}\left[\frac{1}{a}\exp\left(\frac{x}{a}\right)\mathbb{1}_{(-\infty,0]}(x) - \frac{1}{b}\exp\left(-\frac{x}{b}\right)\mathbb{1}_{(0,\infty)}(x)\right]$$

und damit

$$\frac{f'(x)}{f(x)} = \frac{1}{a}\mathbb{1}_{(-\infty,0]}(x) - \frac{1}{b}\mathbb{1}_{(0,\infty)}(x).$$

Weiterhin ergibt sich als Verteilungsfunktion F von f

$$F(t) = \frac{1}{a+b}\left[a\exp\left(\frac{t}{a}\right)\mathbb{1}_{(-\infty,0]}(t) + \left(a+b-b\exp\left(-\frac{t}{b}\right)\right)\mathbb{1}_{(0,\infty)}(t)\right].$$

Zur Ermittlung der inversen Verteilungsfunktion F^{-1} ist eine Fallunterscheidung notwendig. Man erhält für $t \leq 0$

$$\begin{aligned} u = F(t) &= \frac{a}{a+b}\exp\left(\frac{t}{a}\right) \\ \Longleftrightarrow \quad F^{-1}(u) = t &= a\log\left(\frac{a+b}{a}u\right) \end{aligned}$$

und für $t > 0$

$$\begin{aligned} u = F(t) &= \frac{1}{a+b}\left(a+b-b\exp\left(-\frac{t}{b}\right)\right) \\ \Longleftrightarrow \quad F^{-1}(u) = t &= -b\log\left(\frac{(a+b)(1-u)}{b}\right). \end{aligned}$$

Also ist F^{-1} von der Gestalt

$$F^{-1}(u) = a\log\left(\frac{a+b}{a}u\right)\mathbb{1}_{(0,p]}(u) - b\log\left(\frac{(a+b)(1-u)}{b}\right)\mathbb{1}_{(p,1)}(u).$$

Somit folgt

$$\frac{f'}{f}\circ F^{-1}(u) = \frac{1}{a} \Longleftrightarrow F^{-1}(u) \leq 0 \Longleftrightarrow u \leq p \Longleftrightarrow g(u) = \frac{1}{a}$$

und analog

$$\frac{f'}{f}\circ F^{-1}(u) = -\frac{1}{b} \Longleftrightarrow F^{-1}(u) > 0 \Longleftrightarrow u > p \Longleftrightarrow g(u) = -\frac{1}{b}.$$

(b) Es ist

$$f'(x) = a\exp(x)\mathbb{1}_{(-\infty,-x_0)}(x) - a\exp(-x)\mathbb{1}_{(x_0,\infty)}(x)$$

und damit

$$\frac{f'(x)}{f(x)} = \mathbb{1}_{(-\infty,-x_0)}(x) - \mathbb{1}_{(x_0,\infty)}(x).$$

Die Verteilungsfunktion F von f ergibt sich als

$$\begin{aligned} F(t) = {} & a \quad \exp(t)\mathbb{1}_{(-\infty,-x_0)}(t) + a(t+\exp(-x_0)+x_0)\mathbb{1}_{[-x_0,x_0]}(t) \\ + {} & a \quad (2x_0 - \exp(-t) + 2\exp(-x_0))\mathbb{1}_{(x_0,\infty)}(t). \end{aligned}$$

Die inverse Verteilungsfunktion F^{-1} ist dann von der Gestalt

$$\begin{aligned} F^{-1}(u) = {} & \log \quad \left(\frac{u}{a}\right)\mathbb{1}_{(0,2^{-i})}(u) + \left(\frac{u}{a} - \exp(-x_0) - x_0\right)\mathbb{1}_{[2^{-i},1-2^{-i}]}(u) \\ - {} & \log \quad \left(2x_0 + 2\exp(-x_0) - \frac{u}{a}\right)\mathbb{1}_{(1-2^{-i},1)}(u). \end{aligned}$$

Also ist

$$\frac{f'}{f} \circ F^{-1}(u) = -1 \iff F^{-1}(u) > x_0 \iff u > 1 - 2^{-i} \iff g(u) = -1,$$

$$\frac{f'}{f} \circ F^{-1}(u) = 0 \iff -x_0 \leq F^{-1}(u) \leq x_0 \iff 2^{-i} \leq u \leq 1 - 2^{-i} \iff g(u) = 0$$

und

$$\frac{f'}{f} \circ F^{-1}(u) = 1 \iff F^{-1}(u) < -x_0 \iff u < 2^{-i} \iff g(u) = 1.$$

■

5.2 Bedeutung der Rademachersignale

Wie man in Beispiel 5.4 sehen kann, führt die erste Rademacherfunkion $r_1 = \mathbb{1}_{[0,\frac{1}{2}]} - \mathbb{1}_{(\frac{1}{2},0]}$ zu einer Lokationsfamilie, die als streng unimodale Dichte die Dichte der Doppel-Exponentialverteilung besitzt. Die Signale S_{r_i} sind nicht konkav, falls $i > 1$ ist. Linearkombintionen von Rademacherfunktionen innnerhalb des Raumes $W_k = \text{span}(r_1, \ldots, r_k)$ können aber zu konkaven Signalen führen. Diese Linearkombinationen lassen sich als kanonische Gradienten von statistischen Funktionalen darstellen. Dazu sei an die Definition der L_2-Differenzierbarkeit und Tangentialräume aus Abschnitt 2.1 erinnert.
Seien $(\Omega, \mathcal{A})$ ein Messraum und $\mathcal{M}_1(\Omega, \mathcal{A})$ die Menge aller Wahrscheinlichkeitsmaße auf $(\Omega, \mathcal{A})$. Jede nichtleere Teilmenge $\mathcal{P} \subset \mathcal{M}_1(\Omega, \mathcal{A})$ heißt eine nichtparametrische Familie von Wahrscheinlichkeitsmaßen.

Definition 5.5 (Differenzierbares statistisches Funktional)
Ein statistisches Funktional $\kappa : \mathcal{P} \to \mathbb{R}$ heißt differenzierbar an der Stelle $P \in \mathcal{P}$ mit dem Gradienten $\dot{\kappa} = \dot{\kappa}(P) \in L_2^0(P)$, falls für jede $L_2(P_0)$-differenzierbare Kurve $t \mapsto P_t$ in $\mathcal{P}$ mit Tangente g und $P_0 = P$ gilt

$$\lim_{t \downarrow 0} \frac{1}{t}(\kappa(P_t) - \kappa(P_0)) = \int \dot{\kappa} g \; dP_0.$$

Satz 5.6
Ein statistisches Funktional $\kappa : \mathcal{P} \to \mathbb{R}$ sei differenzierbar an der Stelle $P \in \mathcal{P}$. Dann existiert ein eindeutiger Gradient $\tilde{\kappa} \in T(P, \mathcal{P})$, der die kleinste Norm innerhalb der Menge aller Gradienten von κ an der Stelle P besitzt.

Beweis. Vgl. JANSSEN [14, Seite 185]. ■

Definition 5.7 (Kanonischer Gradient)
Den Gradienten $\tilde{\kappa}$ aus Satz 5.6 bezeichnet man als kanonischen Gradienten von κ an der Stelle P.

Bemerkung 5.8
Sei $\kappa : \mathcal{P} \to \mathbb{R}$ an der Stelle $P \in \mathcal{P}$ differenzierbar mit dem Gradienten $\dot{\kappa} \in L_2^0(P)$. Der kanonische Gradient $\tilde{\kappa}$ von κ an der Stelle P ergibt sich dann als orthogonale Projektion von $\dot{\kappa}$ in den Tangentialraum $T(P, \mathcal{P})$, vgl. JANSSEN [14, Seite 186]. □

Sei nun $\mathcal{P}$ eine Menge von Verteilungen auf $\mathbb{R}$ mit positiver Lebesgue-Dichte auf ihrem Träger, welcher ein Intervall (möglicherweise unbeschränkt) sein soll. Man betrachte für diese Verteilungsklasse das Funktional

$$\kappa : \mathcal{P} \to \mathbb{R}, \quad F \mapsto -F^{-1}(\xi) \tag{5.2}$$

für ein $\xi \in (0,1)$. Nach JANSSEN und KUNZ [20] ist das obige Quantilfunktional (5.2) differenzierbar an der Stelle $F_0 \in \mathcal{P}$ mit einem kanonischen Gradienten $\tilde{\kappa}$, welcher proportional ist zu $h_\xi \circ F_0$ mit

$$h_\xi(u) = \frac{1}{\sqrt{\xi(1-\xi)}}(1_{[0,\xi)}(u) - \xi), \quad 0 < u < 1.$$

Die Methodik differenzierbarer Funktionale und derer Gradienten wurde von PFANZAGL und WEFELMEYER [34] eingeführt und wird ebenfalls in JANSSEN [15] sowie JANSSEN und KUNZ [20] diskutiert.
Betrachtet man nun für $k \in \mathbb{N}$ Linearkombinationen der Quantilfunktionen

$$F \mapsto \sum_{i=1}^{2^k-1} a_i F^{-1}\left(\frac{i}{2^k}\right),$$

so stellt man fest, dass Folgendes gilt (vgl. Satz 5.9):

$$\{g \circ F_0 : g \in W_k\} \subseteq \left\{\tilde{\kappa} : \kappa(F) = \sum_{i=1}^{2^k-1} a_i F^{-1}\left(\frac{i}{2^k}\right), a_i \in \mathbb{R}\right\},$$

wobei W_k =span$(r_1, \dots r_k)$.
Für die Anwendung sind die Konsequenzen von sehr großem Interesse. Nimmt man nämlich an, dass ein Signaltest an der Stelle $g \circ F_0$ mit $g \in W_k$ hohe Güte auf einer Kugel hat, oder kann sein Gradient durch $g \circ F_0$ approximiert werden, dann existiert eine Linearkombination von Quantilfunktionen, welche exakt diesen als kanonischen Gradienten hat. Der Signaltest ist dann für die Erkennung von Abweichungen, die von diesem Funktional herrühren, sehr empfindlich.

Satz 5.9
Für festes $k \in \mathbb{N}$ sei W_k =span$(r_1, \ldots, r_k)$. Dann gilt

$$\{g \circ F_0 : g \in W_k\} \subseteq V_k,$$

wobei V_k die Menge aller Gradienten $\tilde{\kappa}$ von in F_0 differenzierbaren Funktionalen κ darstellt mit $\kappa : \mathcal{P} \to \mathbb{R}, F \mapsto \sum_{i=1}^{2^k-1} a_i F^{-1}\left(\frac{i}{2^k}\right)$ für $a_i \in \mathbb{R}$.

Beweis. Sei $\kappa : \mathcal{P} \to \mathbb{R}, F \mapsto \sum_{i=1}^{2^k-1} a_i F^{-1}\left(\frac{i}{2^k}\right)$ ein differenzierbares Funktional in F_0 und sei ohne Einschränkung $F_0(x) = x, 0 \leq x \leq 1$. Dann ist der kanonische Gradient $\tilde{\kappa}$ von der Gestalt

$$\tilde{\kappa} = -\sum_{i=1}^{2^k-1} \frac{a_i}{\sqrt{\frac{i}{2^k}\left(1 - \frac{i}{2^k}\right)}} \left(\mathbb{1}_{[0, \frac{i}{2^k})} - \frac{i}{2^k}\right), \quad a_i \in \mathbb{R}.$$

Sei $g \in W_k$ mit $g = \sum_{i=1}^{k} \beta_i r_i$. Setze $U_k :=$ span$(u_j : 0 \leq j \leq 2^k)$ mit $u_j := \mathbb{1}_{[0, \frac{j}{2^k})}$. Betrachte die orthogonale Projektion π_0 von $L_2(0,1)$ auf $L_2^0(0,1)$ gegeben durch

$$\begin{aligned} \pi_0 : L_2(0,1) &\to L_2^0(0,1) \\ f &\mapsto f - \int_0^1 f(x)\, dx. \end{aligned}$$

Es ist

$$\pi_0(u_j) = u_j - \int_0^1 u_j \, d\lambda\!\!\lambda_{|[0,1]} = u_j - \frac{j}{2^k} = \frac{a_j}{\sqrt{\frac{i}{2^k}\left(1 - \frac{i}{2^k}\right)}} \left(\mathbb{1}_{[0, \frac{j}{2^k})} - \frac{j}{2^k}\right) \in V_k$$

mit $a_j := \sqrt{\frac{j}{2^k}\left(1 - \frac{j}{2^k}\right)}$ für $1 \leq j \leq 2^k$. Wegen der Linearität von π_0 folgt $\pi_0(U_k) \subseteq V_k$. Weiterhin ist für $1 \leq m \leq k$ und $1 \leq l \leq 2^m$

$$\mathbb{1}_{[\frac{l-1}{2^m}, \frac{l}{2^m})} = u_{l2^{k-m}} - u_{(l-1)2^{k-m}}.$$

Die Rademacherfunktionen r_j lassen sich darstellen als

$$r_j = \sum_{i=1}^{2^j} (-1)^{i+1} \mathbb{1}_{[\frac{i-1}{2^j}, \frac{i}{2^j})}, 1 \leq j \leq k.$$

Deshalb folgt $W_k \subseteq U_k$. Man erhält also insgesamt

$$W_k = \pi_0(W_k) \subseteq \pi_0(U_k) \subseteq V_k.$$

■

Beispiel 5.10

Sei F_0 die Gleichverteilung.

(a) Der Median $F \mapsto F^{-1}\left(\frac{1}{2}\right)$ hat den kanonischen Gradienten $\tilde{\kappa} = -r_1$.

(b) Für $k \in \mathbb{N}$ betrachte man das winsorisierte Mittel $F \mapsto \frac{1}{2}\left(F^{-1}\left(\frac{1}{2^k}\right) + F^{-1}\left(1 - \frac{1}{2^k}\right)\right)$.

Für $k = 2$ ergibt sich das robuste Lagemaß

$$F \mapsto \frac{1}{2}\left(F^{-1}\left(\frac{1}{4}\right) + F^{-1}\left(\frac{3}{4}\right)\right)$$

mit dem kanonischen Gradienten $\tilde{\kappa} = -\frac{1}{\sqrt{3}}r_1 - \frac{1}{\sqrt{3}}r_2$. $\square$

Anhang

Zur Erinnerung seien $H \subseteq L_2(0,1)$ ein Teilhilbertraum und

$$S_h(t) = \int_0^t h \, d\lambda\!\!\lambda_{|[0,1]} \quad \text{für} \quad h \in H, 0 \leq t \leq 1.$$

Im Folgenden wird für bestimmte Teilmengen K von H gezeigt, dass diese Kegel bilden. Dabei versteht man unter einer Kegel K eine Menge, für die gilt:

$$g, h \in K, \alpha, \beta \geq 0 \quad \Longrightarrow \quad \alpha g + \beta h \in K.$$

Satz A.1 Sei $K := \{h \in H : t \mapsto S_h(t) \text{ ist konkav}\}$. Dann ist K ein abgeschlossener, konvexer Kegel von H.

Beweis. Seien $h_1, h_2 \in H, \alpha, \beta \geq 0$ und $\lambda, t_1, t_2 \in [0,1]$. Mit der Konkavität von $t \mapsto S_{h_1}(t)$ und $t \mapsto S_{h_2}(t)$ erhält man

$$\begin{aligned}
&\lambda S_{\alpha h_1 + \beta h_2}(t_1) + (1-\lambda) S_{\alpha h_1 + \beta h_2}(t_2) \\
&= \lambda S_{\alpha h_1}(t_1) + (1-\lambda) S_{\alpha h_1}(t_2) + \lambda S_{\beta h_2}(t_1) + (1-\lambda) S_{\beta h_2}(t_2) \\
&\leq S_{\alpha h_1 + \beta h_2}(\lambda t_1 + (1-\lambda) t_2).
\end{aligned}$$

Somit ist $\alpha h_1 + \beta h_2 \in K$. Daraus folgt auch unmittelbar die Konvexität von K. Um die Abgeschlossenheit von K in H zu zeigen, sei $(h_n)_{n \in \mathbb{N}}$ eine Folge in K mit $\lim_{n \to \infty} h_n = h$ für ein $h \in H$. Aus $\lim_{n \to \infty} \|h_n - h\|_{L_2} = 0$ folgt $\lim_{n \to \infty} S_{h_n}(t) = S_h(t)$ für jedes $t \in [0,1]$, da L_2-Konvergenz die L_1-Konvergenz impliziert. Mit der Konkavität von $t \mapsto S_{h_n}(t)$ für jedes $n \in \mathbb{N}$ ergibt sich für $\lambda, t_1, t_2 \in [0,1]$

$$\begin{aligned}
\lambda S_h(t_1) + (1-\lambda) S_h(t_2) &= \lim_{n \to \infty} [\lambda S_{h_n}(t_1) + (1-\lambda) S_{h_n}(t_2)] \\
&\leq \lim_{n \to \infty} S_{h_n}(\lambda t_1 + (1-\lambda) t_2) \\
&= S_h(\lambda t_1 + (1-\lambda) t_2).
\end{aligned}$$

Also ist $t \mapsto S_h(t)$ konkav und somit $h \in K$. ∎

Bemerkung A.2 Genauso zeigt man für einen Teilhilbertraum $H \subseteq L_2(0,1)$, dass $V_\geq = V_\geq(H) := \{g \in H : S_g \geq 0\}$ und $V_\leq = V_\leq(H) := \{g \in H : S_g \leq 0\}$ abgeschlossene konvexe Kegel sind.

Satz A.3 Sei $K_h := \{g \in H : t \mapsto S_g(t)$ ist konkav, $S_g(t) \geq S_h(t)$ für jedes $t \in [0,1]\}$ für ein festes $h \in H$. Dann ist K_h eine abgeschlossene, konvexe Teilmenge von H.

Beweis. Um die Abgeschlossenheit von $\mathcal{K}_h$ nachzuweisen, betrachte man eine Folge $(g_n)_{n\in\mathbb{N}}$ in K_h mit $\lim_{n\to\infty} g_n = g$ für ein $g \in H$. Wie bereits im Beweis zu Satz A.2 gezeigt worden ist, ist $t \mapsto S_g(t)$ ebenfalls konkav. Wegen $S_{g_n}(t) \geq S_h(t)$ für jedes $n \in \mathbb{N}, t \in [0,1]$ und $\lim_{n\to\infty} S_{g_n}(t) = S_g(t)$ folgt

$$S_g(t) \geq S_h(t)\ \forall\ t \in [0,1].$$

Somit ist $g \in K_h$ und K_h abgeschlossen.
Um die Konvexität von $\mathcal{K}_h$ zu zeigen, wähle man $g_1, g_2 \in K_h$ und $\alpha \in [0,1]$. Dann folgt mit der Konkavität von $t \mapsto S_{g_1}(t)$ und $t \mapsto S_{g_2}(t)$ die Konkavität von $t \mapsto S_{\alpha g_1+(1-\alpha)g_2}(t)$ (vergleiche Beweis zu Satz A.2). Außerdem ist

$$\begin{aligned} S_{\alpha g_1+(1-\alpha)g_2}(t) &= \alpha S_{g_1}(t) + (1-\alpha)S_{g_2}(t) \\ &\geq \alpha S_h(t) + (1-\alpha)S_h(t) \\ &= S_h(t). \end{aligned}$$

Folglich ist K_h auch konvex. ■

Satz A.4 (und Definition) Sei $(H, \langle\cdot,\cdot\rangle)$ ein reeller Hilbertraum. Dann gelten:
(a) Sei V eine abgeschlossene konvexe Teilmenge von H. Dann existiert für jedes $h \in H$ ein eindeutig bestimmtes Element $pV(h) \in V$ mit

$$\|h - pV(h)\| = \inf_{v\in V} \|h - v\|. \tag{5.3}$$

Das Element $pV(h)$ wird als Projektion von h in V bezeichnet.
(b) Sei V ein abgeschlossener konvexer Kegel in H. Dann ist die Projektion $pV(h) \in V$ von $h \in H$ eindeutig bestimmt durch die folgenden Eigenschaften:

$$\langle pV(h), h\rangle = \|pV(h)\|^2 \tag{5.4}$$

und

$$\langle h, v\rangle \leq \langle pV(h), v\rangle \quad \forall v \in V. \tag{5.5}$$

Beweis. Vgl. BEHNEN und NEUHAUS [2, Proposition 3.2.1]. ■

Lemma A.5 [Projektionslemma 3.7]
(a) Seien H ein reeller Hilbertraum, $V \subset H$ ein abgeschlossener konvexer Kegel und $pV(h)$ bezeichne die eindeutig bestimmte Projektion von $h \in H$ in V. Eine orthogonale Zerlegung

$$h = pV(h) + p\tilde{V}(h), \ \langle pV(h), p\tilde{V}(h)\rangle = 0,$$

von $h \in H$ ist gegeben durch die Projektion $p\tilde{V}(h)$ in den polaren Kegel $\tilde{V} = \{g \in H : \langle g, v\rangle \leq 0 \ \forall v \in V\}$.
(b) Sei $h = h_1 + h_2$ eine weitere orthogonale Zerlegung mit $h_1 \in V, h_2 \in \tilde{V}$ und $h_1 \perp h_2$. Dann ist $h_1 = pV(h)$ und $h_2 = p\tilde{V}(h)$.

Beweis. (a) Seien $g_1 := pV(h)$ und $g_2 := h - g_1$. Mit Satz A.4 erhält man

$$\langle g_1, g_2\rangle = \langle g_1, h - g_1\rangle = \langle g_1, h\rangle - \langle g_1, g_1\rangle \overset{(5.4)}{=} \|g_1\|^2 - \|g_1\|^2 = 0$$

sowie

$$\langle g_2, v\rangle = \langle h - g_1, v\rangle = \langle h, v\rangle - \langle g_1, v\rangle \overset{(5.5)}{\leq} \langle g_1, v\rangle - \langle g_1, v\rangle = 0 \ \ \forall v \in V.$$

Also ist $g_2 \in \tilde{V}$ mit $g_1 \perp g_2$.
Sei nun $u \in V$ fest und $\tilde{v} \in \tilde{V}$ beliebig. Dann ist $\langle u, \tilde{v}\rangle \leq 0$ und folglich $u \in \tilde{\tilde{V}}$. Somit gilt $g_1 \in V \subseteq \tilde{\tilde{V}}$ und

$$\langle h, \tilde{v}\rangle = \langle g_1 + g_2, \tilde{v}\rangle = \langle g_1, \tilde{v}\rangle + \langle g_2, \tilde{v}\rangle \leq \langle g_2, \tilde{v}\rangle \ \forall \tilde{v} \in \tilde{V}$$

sowie

$$\langle g_2, h\rangle = \langle g_2, g_1 + g_2\rangle = \langle g_2, g_1\rangle + \langle g_2, g_2\rangle = \|g_2\|^2.$$

Da also g_2 die Eigenschaften (5.4) und (5.5) erfüllt, folgt

$$p\tilde{V}(h) = g_2.$$

(b) Seien $g_1 = pV(h)$ und $g_2 := p\tilde{V}(h)$. Es ist

$$\|g_2\| = \|h - g_1\| \overset{(5.3)}{\leq} \|h - h_1\| = \|h_2\|$$

und analog

$$\|g_1\| \leq \|h_1\|.$$

Daraus folgt

$$\|h_1\|^2+\|h_2\|^2 = \|h_1+h_2\|^2 = \|h\|^2 = \|g_1+g_2\|^2 = \|g_1\|^2+\|g_2\|^2 \leq \|h_1\|^2+\|g_2\|^2 \leq \|h_1\|^2+\|g_2\|^2.$$

Also ist

$$\|h_1\|^2 = \|g_1\|^2 \quad \text{und} \quad \|h_2\|^2 = \|g_2\|^2.$$

Damit erhält man

$$\|h-h_1\|^2 = \|h_2\|^2 = \|g_2\|^2 = \|h-g_1\|^2 \overset{(5.3)}{\leq} \|h-v\|^2 \quad \forall v \in V$$

und analog

$$\|h-h_2\|^2 \leq \|h-\tilde{v}\|^2 \quad \forall \tilde{v} \in \tilde{V}.$$

Also ist

$$h_1 = g_1 = pV(h) = \quad \text{und} \quad h_2 = g_2 = p\tilde{V}(h).$$

■

Symbol- und Abkürzungsverzeichnis

Im Folgenden werden im Rahmen dieser Arbeit verwendete Symbole und Abkürzungen erläutert.

Symbole

$B = (B(t))_{0 \le t \le 1}$	Prozess der Brownschen Bewegung
$B_0 = (B_0(t))_{0 \le t \le 1}$	Prozess der Brownschen Brücke
$C([0,1])$	Menge aller stetigen Funktionen $f : [0,1] \to \mathbb{R}$
$C(\mathbb{R})$	Menge aller stetigen Funktionen $f : \mathbb{R} \to \mathbb{R}$
C	$:= \{f \in C([0,1]) : f(0) = 0\}$
F_n	empirische Verteilungsfunktion zum Stichprobenumfang n
C_0	$:= \{f \in C([0,1]) : f(0) = f(1) = 0\}$
$g(t-)$	$:= \lim_{s \uparrow t} g(s)$
H	Hilbertraum (meistens Teilhilbertraum $H \subseteq L_2(0,1)$)
H_0	$L_2^0(0,1)$
$\tilde{H}$	$:= \overline{\mathrm{span}(r_i : i \in \mathbb{N}_0)}$
$\tilde{H}^0$	$:= \overline{\mathrm{span}(r_i : i \in \mathbb{N})}$
$L_2(\Omega, \mathcal{A}, P)$	$= \{X : \Omega \to \mathbb{R} \text{Zufallsvariable} : \lVert X \rVert = (\int X^2 \, dP)^{1/2} < \infty\}$
$L_2(0,1)$	$:= L_2([0,1], \mathcal{B}([0,1]), \lambda\!\!\lambda_{\vert[0,1]})$
$L_2^0(0,1)$	$:= L_2^0([0,1], \mathcal{B}([0,1]), \lambda\!\!\lambda_{\vert[0,1]}) := \{h \in L_2(0,1) : \int h \, d\lambda\!\!\lambda_{\vert[0,1]} = 0\}$
$(L(h))_{h \in H}$	zentraler Prozess zum Gauß-Shift $(\Omega, \mathcal{A}, \{P_h : h \in H\})$
$\overline{M}$	Abschluß der Menge M
$N(\mu, \sigma^2)$	Normalverteilung mit Erwartungswert μ und Varianz σ^2

$o(\cdot)$	Landau-Symbol
$pV(h)$	Projektion von $h \in H$ in den Kegel $V \subseteq H$
$(r_i)_{i \in \mathbb{N}_0}$	System der Rademacherfunktionen
S_g	Signalabbildung zu $g \in H$ mit $S_g(t) := \int_0^t g \; d\lambda\!\!\lambda_{\vert[0,1]}, t \in [0,1]$
$\mathrm{span}(M)$	lineare Hülle der Menge M
$T(P, \mathcal{P})$	Tangentialraum von $P \in \mathcal{P}$ bzgl. der Familie $\mathcal{P}$
u_α	α-Quantil der Standardnormalverteilung
V	abgeschlossener konvexer Kegel in H
$\tilde{V} = \tilde{V}(H)$	$:= \{g \in H : \langle g, v \rangle \leq 0 \; \forall v \in V\}$
$V_\leq = V_\leq(H)$	$:= \{g \in H : S_g \leq 0\} = \{g \in H : S_g(t) \leq 0$ für jedes $t \in [0,1]\}$
$V_\geq = V_\geq(H)$	$:= \{g \in H : S_g \leq 0\} = \{g \in H : S_g(t) \leq 0$ für jedes $t \in [0,1]\}$
$V_\leq^t$	$:= \{g \in L_2(0,1) : S_g \leq 0\}$
$V_\leq^0$	$:= V_\leq^t \cap L_2^0(0,1) = \{g \in L_2^0(0,1) : S_g \leq 0\}$
W_k	$:= \mathrm{span}(r_1, \ldots, r_k)$
$W_k^\perp$	Orthogonalraum zu W_k
$\tilde{W}_k$	$:= \mathrm{span}(r_0, \ldots, r_k)$
$Y = (Y(t))_{t \in \mathbb{R}}$	Signalprozess
$\mathbb{N}$	Menge der natürlichen Zahlen
$\mathbb{N}_0$	$:= \mathbb{N} \cup \{0\}$
$\mathbb{R}$	Menge der reellen Zahlen
$\mathbb{R}_+$	$:= \{x \in \mathbb{R} : x \geq 0\}$
$\mathcal{B}(\Omega)$	Borelsche σ-Algebra über Ω
$\mathcal{L}(X \vert P)$	Verteilung der Zufallsvariablen X unter dem Maß P
$\mathbb{1}_A$	Indikatorfunktion der Menge A
ε_a	Einpunktmaß in a
$\lambda\!\!\lambda_{\vert[0,1]}$	Lebesgue-Maß auf $[0,1]$
Φ	Verteilungsfunktion der Standardnormalverteilung
$\varphi = \varphi(Y(\cdot))$	einseitiger Signalerkennungstest
$\langle g, h \rangle$	$:= \int gh \; d\lambda\!\!\lambda_{\vert[0,1]}$ Skalarprodukt auf $L_2(0,1)$
$\Vert g \Vert$	$:= (\int g^2 \; d\lambda\!\!\lambda_{\vert[0,1]})^{1/2}$ Norm auf $L_2(0,1)$
$X \stackrel{\mathcal{D}}{=} Y$	Gleichheit in Verteilung

$X_n \xrightarrow{\mathcal{D}} X$	Kovergenz in Verteilung
$\blacksquare$	Ende einer Beweisführung
$\square$	Ende einer Bemerkung oder eines Beispiels

Abkürzungen

bzgl.	bezüglich
bzw.	beziehungsweise
d.h.	das heißt
et al.	et alii (und andere)
f.	folgende
f.s.	fast sicher
i.i.d.	unabhängig identisch verteilt
SEP	Signalerkennungsproblem
vgl.	vergleiche

Literaturverzeichnis

[1] BARLOW, R.E., BARTHOLOMEW, D.J., BREMMER, J.M., BRUNK, H.D., 1972. *Statistical Inference under Order Restrictions.* Wiley, New York.

[2] BEHNEN, K. und NEUHAUS, G., 1989. *Rank tests with estimated scores and their application.* Teubner, Stuttgart.

[3] BISCHOFF, W., 1998. *A functional central lilit theorem for regression models.* Annals of Statistics **26**, No.4, 1398-1410.

[4] BISCHOFF, W. und HASHORVA, E., 2005. *A lower bound for boundary crossing probabilities of Brownian bridge/motion with trend.* Statistics and Probability Letters. **74**, No.3, 265-271.

[5] BISCHOFF, W., HASHORVA, E., HÜSLER und J., MILLER, F., 2003a. *Exact asymptotics for boundary crossings of the Brownian bridge with trend with application to the Kolmogorov test.* Annals of the Institute of Statistical Mathematics **55**, No.4, 849-864.

[6] BISCHOFF, W., HASHORVA, E., HÜSLER und J., MILLER, F., 2003b. *Asymptotics of a boundary crossing probability of a Brownian bridge with general trend.* Methodology and Computing in Applied Probability **5**, No.3, 271-287.

[7] BISCHOFF, W., HASHORVA, E., HÜSLER und J., MILLER, F., 2004. *On the power of the Kolmogorov test to detect the trend of a Brownian bridge with applications to a change-point problem in regression models.* Statistics and Probability Letters **66**, No.2, 105-115.

[8] BISCHOFF, W., HASHORVA, E., HÜSLER und J., MILLER, F., 2005. *Analysis of a change-point regression problem in quality control by partial sums processes and Kolmorov type tests.* Metrica 62, No.1, 85-98.

[9] EFRON, B. und JOHNSTONE, I. M., 1990. *Fisher's information in terms of hazard rate.* Annals of Statistics **18**, No.1, 38-62.

[10] HÁJEK, J., 1968. *Asymptotic normality of simple linear rank statistics under alternatives.* Annals of Mathematical Statistics **39**, No.2, 325-346.

[11] HÁJEK, J., ŠIDÁK, Z., SEN, P. K., 1999. *Theory of rank tests.* 2.Auflage. Academic Press, San Diego.

Funktionalanalysis. 4.Auflage. Teubner Verlag, Wiesbaden.

[12] IRLE, A, 1998. *Finanzmathematik: Die Bewertung von Derivaten.* Teubner Verlag, Stuttgart.

[13] JANSSEN, A., 1995. *Principal component decomposition of non-parametric tests.* Probability Theory and Related Fields **101**, No.2, 193-209.

[14] JANSSEN, A., 1998. *Zur Asymptotik nichtparametrischer Tests.* Skripten zur Mathematischen Statistik Nr.29.

[15] JANSSEN, A., 1999. *Testing nonparametric statistical functionals with applications to rank tests.* Journal of Statistical Planning and Inference. **81**, No.1, 71-93.

[16] JANSSEN, A., 2000. *Global power functions of goodness of fit tests.* Annals of Statistics **28**, No.1, 239-253.

[17] JANSSEN, A., 2003. *Which power of goodness of fit tests can really be expected: Intermediate versus contiguous alternatives.* Statistics and Decisions **21**, No.4, 301-325.

[18] JANNSEN, A., 2004. *Asymptotic relative efficiency of tests at the boundary of regular statistical models.* Journal of Statistical Planning and Inference **126**, No.2, 461-477.

[19] JANSSEN, A., KUNZ, M., 2000. *On The equivalence of infinite dimensional Gaussian shift regression models.* Technical report.

[20] JANSSEN, A., KUNZ, M., 2002. *Global extrapolations for power functions of one-sided nonparametric tests.* Statistics and Decisions **20**, No.2, 153-176.

[21] JANSSEN, A., KUNZ, M., 2004. *Brownian type boundary crossing probabilities for piecewise linear boundary functions.* Communications in Statistics. Theory and Methods **33**, No.7, 1445-1464.

[22] JANSSEN, A. und MILBRODT, H, 1993. *Rényi type goodness of fit tests with adjusted principal direction of alternatives.* Scandinavian Journal of Statistics **20**, No.3, 177-194.

[23] JANSSEN, A., MILBRODT, H. und STRASSER, H., 1985. *Infinitely Divisible Statistical Experiments.* Lecture Notes in Statistics, Vol.27. Springer Verlag, Belin/Heidelberg.

[24] JANSSEN, A., ÜNLÜ, H., 2008. *Regions of alternatives with high and low power for goodness-of-fit tests.* To appear in Journal of Statistical Planning and Inference, Vol. 138, No.8, 2526-2543.

[25] KARATZAS, I., SHREVE, S.E., 1988. *Brownian Motion and Stochastic Calculus*, Springer Verlag, New York.

[26] KUNZ, M., 1999. *Anwendung der Theorie von Gauß-Shift-Experimenten auf den Kolmogorov-Smirnov-Test und das einseitige Boundary-Crossing-Problem.* Dissertation, Heinrich-Heine-Universität Düsseldorf.

[27] LE CAM, L., 1986. *Asymptotic Methods in Statistical Decision Theory* Springer Verlag, New York

[28] LEHMANN, E.L., ROMANO, J.P., 2005. *Testing Statistical Hypotheses.* Third Edition. Springer Verlag, New York.

[29] MEISE, R. und VOGT, D., 1992. *Einführung in die Funktionalanalysis.* Vieweg Verlag, Braunschweig/Wiesbaden.

[30] MILBRODT, H. und STRASSER, H., 1990. *On the asymptotic power of the two-sided Kolmogorov-Smirnov test.* Journal of Statistical Planning and Inference **26**, No.1, 1-23.

[31] MOREAU, J.J., 1962. *Décomposition orthogonale d'un espace hilbertien selon deux cônes mutuellement pollaire.* Comptes Rendus de l'Acadmie des Sciences. Paris. French. Journal **255**, 238-240.

[32] NEUHAUS, G., 1976. *Asymptotic power properties of the Cramér-von Mises test under contiguous alternatives.* Journal of Multivariate Analysis **6**, 95-110.

[33] OSTROVSKI, V., 2006. *Testen statistischer Funktionale für Zweistichprobenprobleme.* Dissertation, Heinrich-Heine-Universität Düsseldorf.

[34] PFANZAGL, J. und WEFELMEYER, W., 1982. *Contributions to a general asymptotic statistical theory.* Lecture Notes in Statistics 13. Springer Verlag, Berlin.

[35] PFANZAGL, J. und WEFELMEYER, W., 1985. *Asymptotic Expansions for General Statistical Models.* Lecture Notes in Statistics 31. Springer Verlag, Berlin.

[36] RAHNENFÜHRER, J., 2003. *On preferences of general two-sided tests with applications to Kolmogorov-Smirnov-type tests.* Statistics and Decisions **21**, No. 2, 149-170.

[37] RITOV, Y. und WELLNER, J., 1988. *Censoring, martingales, and the Cox model.* Statistical inference from stochastic processes. Contemporary Mathematics **80**, 191-219.

[38] SCHILLING, R.L., 2005. *Measures, integrals and martingales.* Cambridge University Press.

[39] SHIRYAEV, A.N., 1984. Probability. Springer Verlag, New York.

[40] SHORACK, G.R. und WELLNER, J.A., 1986. *Empirical Processes with Applications to Statistics.* Wiley, New York.

[41] STRASSER, H., 1985. *Mathematical Theory of Statistics.* De Gruyter Studies in Mathematics.

[42] STRASSER, H., 1985. *Einführung in die lokale asymptotische Theorie der Statistik.* Bayreuther Mathematische Schriften, Heft 19.

[43] VAN DER VAART, A.W., 1998. *Asymptotic Statistics.* Cambridge University Press.

[44] WITTING, H., 1985. *Mathematische Statistik I.* Teubner Verlag, Stuttgart.

Erklärung

Die hier vorgelegte Dissertation habe ich eigenständig und ohne unerlaubte Hilfe angefertigt. Die Dissertation wurde in der vorgelegten oder in ähnlicher Form noch bei keiner anderen Institution eingereicht. Ich habe bisher keine erfolglosen Promotionsversuche unternommen.

Hülya Ünlü Düsseldorf, den 04.07.2008