Das Diskrepanzprinzip in der nichtparametrischen Kurvenschätzung

Dissertation

zur Erlangung des Grades
eines Doktors der Naturwissenschaften
der Technischen Universität Dortmund

technische universität
dortmund

Der Fakultät Statistik
der Technischen Universität Dortmund

vorgelegt von

Thoralf Mildenberger

Dortmund, Dezember 2010

1. Gutachterin: Prof. Dr. Ursula Gather
2. Gutachter: Prof. Dr. Roland Fried
3. Gutachter: Prof. Dr. Holger Dette
Tag der mündlichen Prüfung: 24. Januar 2011

Inhaltsverzeichnis

Tabellenverzeichnis

1. Einleitung

Lange Zeit waren die meisten statistischen Modelle parametrisch, d.h. die Form einer Wahrscheinlichkeitsverteilung oder einer Regressionsfunktion war bis auf einige freie Parameter vorgegeben. Solche Modelle sind mathematisch relativ einfach zu handhaben, und ihr Einsatz ist insbesondere in solchen Fällen sinnvoll, in denen die zugrundeliegende Sachproblematik bereits weitgehend erforscht ist – zu denken ist hier z.B. an Fragestellungen aus der Physik, in denen bestimmte Gesetze bereits bekannt sind und nur noch die für den interessierenden Fall relevanten Werte von Konstanten zu bestimmen sind. Auch wird zur Schätzung der (meistens wenigen) Modellparameter nur eine vergleichsweise kleine Anzahl Beobachtungen benötigt, und in vielen Fällen ist der rechnerische Aufwand gering. Wenn über das zugrundeliegende Sachproblem jedoch nur wenig bekannt ist, sind parametrische Modelle häufig zu unflexibel und können unter Umständen zur Modellierung völlig ungeeignet sein. Dies hat zur Entwicklung nichtparametrischer Verfahren geführt, die inzwischen ein wichtiges Teilgebiet der Statistik darstellen. Zum einen gab es große mathematische Fortschritte beim Verständnis der Methoden, zum anderen sind aufgrund der Entwicklungen der Computertechnik viele aufwändige Verfahren erst heute praktisch anwendbar. Die Anwendung statistischer Methoden in immer weiteren Gebieten der Wissenschaft macht die Entwicklung flexibler Methoden zudem immer dringlicher. Auch sind nichtparametrische Methoden zur explorativen Datenanalyse und zur Visualisierung von Daten gut geeignet.

In dieser Arbeit sollen zwei der wichtigsten Probleme der nichtparametrischen Statistik betrachtet werden:

- In der *nichtparametrischen Dichteschätzung* ist es das Ziel, für eine Stichprobe die Dichte der Verteilung, aus der die Beobachtungen stammen, möglichst gut zu rekonstruieren.

- In der *nichtparametrischen Regression* soll der Zusammenhang zwischen einer oder mehreren Einflussvariablen und einer abhängigen Variablen modelliert werden.

In beiden Fällen soll eine Funktion geschätzt werden, die nicht durch nur endlich viele Parameter bestimmt ist, obwohl unter Umständen gewisse Annahmen über Glattheit, Monotonieverhalten oder den Träger der Funktion getroffen werden können. Jedoch stellt sich immer das Problem, dass ein Element einer unendlichdimensionalen Funktionenklasse auf Basis von nur endlich vielen Beobachtungen identifiziert werden soll. Die meisten Verfahren der nichtparametrischen Kurvenschätzung hängen von mindestens einem Glättungs- oder Regularisierungsparameter ab, dessen Wahl das Ergebnis entscheidend beeinflusst. In der Regel wird in irgendeinem Sinn die Glattheit oder Komplexität der geschätzten

Kurve gesteuert. Wird zu viel geglättet, so verschwindet der Einfluß der beobachteten Daten fast völlig, und man erhält nur noch eine sehr flache Kurve, die keine Information mehr enthält. Wird zu wenig geglättet, so passt sich die Kurve zu stark den – als zum Teil zufällig vorausgesetzten – konkret vorliegenden Beobachtungen an. Im Extremfall werden z.B. in der Regression die Beobachtungen interpoliert, und man erhält ebenfalls keine Information mehr über den zugrundeliegenden Zusammenhang. Zwar lässt sich oft der optimale Wert des Glättungsparameters theoretisch herleiten, jedoch hängt er meist von der gesuchten Funktion ab. Für die praktische Anwendung müssen Verfahren benutzt werden, die solche Glättungsparameter aufgrund der vorliegenden Daten wählen. Neben der Untersuchung des Verhaltens nichtparametrischer Kurvenschätzer bei optimaler Wahl der Glättungsparameter gehört daher die Konstruktion datengesteuerter Parameterwahlmethoden zu den wichtigsten Problemen der nichtparametrischen Kurvenschätzung.

Im Laufe der Jahre wurden zahlreiche Methoden vorgeschlagen, von denen die meisten in zwei Klassen fallen. Zum einen wird häufig versucht, die Güte des Schätzers (meistens gemessen durch eine Verlust- oder Riskofunktion) für verschiedene Werte des Glättungsparameters aus den Daten zu schätzen. Dies kann zum einen durch die Korrektur naiver Schätzer des Risikos durch Strafterme oder durch Resampling-Verfahren wie z.B. Kreuzvalidierung geschehen. Dann wird ein Wert gewählt, der ein gutes Ergebnis verspricht. Zum anderen wird häufig versucht, in einem expliziten Ausdruck für den optimalen Wert des Glättungsparameters die von der gesuchten Funktion abhängenden Größen aus den Daten zu schätzen und diese Schätzungen dann einzusetzen (sogenannte *Plug-In-Verfahren*).

In dieser Arbeit soll nun ein weiterer Ansatz zur Glättungsparameterwahl näher untersucht werden, der in einem anderen Gebiet der Mathematik, nämlich der Theorie der inversen bzw. schlecht gestellten Probleme, sehr weit verbreitet, in der Statistik jedoch relativ unbekannt ist. Die Grundidee ist, maximal möglich zu glätten unter einer Nebenbedingung bezüglich der Anpassung an die Daten. Letztere kann im Falle der Dichteschätzung über den Abstand zwischen der aus der Dichteschätzung durch Integration erhaltenen Schätzung für die Verteilungsfunktion und der empirischen Verteilungsfunktion formuliert werden, während in der Regression meist die Größe der Residuen gemessen wird. Dabei soll die Bedingung an die Anpassung ein Überglätten und die Wahl der unter dieser Bedingung maximalen Glättung ein Unterglätten verhindern.

Das zweite Kapitel gibt eine Einführung in die Thematik. Zunächst werden die beiden wichtigsten Probleme der nichtparametrischen Kurvenschätzung, die nichtparametrische Dichteschätzung und die nichtparametrische Regression, vorgestellt und die Glättungsparameterwahl diskutiert. Anschließend wird in einem Exkurs auf die Theorie der Regularisierung inverser bzw. schlecht gestellter Probleme eingegangen. Während einige Methoden der Regularisierungs- bzw. Glättungsparameterwahl in beiden Feldern verbreitet sind, wird das 1966 von Morozov eingeführte *Diskrepanzprinzip* (vgl. [Mor66]) zwar in fast jedem Lehrbuch über inverse Probleme vorgestellt, ist aber in der Statistik so gut wie unbekannt. Die wenigen Ausnahmen werden im Anschluss vorgestellt, insbesondere die im Kontext des maschinellen Lernens von Vapnik eingeführte Interpretation der nichtparametrischen Statistik als inverses Problem (vgl. z.B. [Vap98] und [Vap00]) sowie der von Davies vorgeschlagene *Data-Features-* bzw. *Data-Approximation-Ansatz* (vgl. vor allem [Dav95], [Dav03], [Dav08]), der zwar von einem anderen philosophischen Hintergund ausgeht, aber praktisch zu ähnlichen Methoden führt.

Das dritte und das vierte Kapitel sind der genaueren Untersuchung von auf dem Diskrepanzprinzip basierenden Methoden zur Glättungsparameterwahl in der nichtparametrischen Dichteschätzung gewidmet. Im dritten Kapitel wird die Bandbreitenwahl für Kernschätzer untersucht – in diesem Kontext wurden bereits von verschiedenen Autoren unabhängig voneinander Varianten des Diskrepanzprinzips vorgeschlagen und untersucht. Diese bereits vorliegenden Resultate werden in dieser Arbeit vereinheitlicht und erweitert. Außerdem wird gezeigt, dass für bestimmte Klassen von Dichten die Bandbreitenwahl mittels eines Diskrepanzprinzips zu einem inkonsistenten Schätzer führt. Zusätzlich werden auch einige neue Varianten, die auf der Normalverteilung als Referenzdichte basieren, vorgeschlagen. Da die theoretischen Ergebnisse asymptotischer Natur sind, werden die verschiedenen Versionen des Diskrepanzprinzips in einer umfangreichen Simulationsstudie erstmals sowohl untereinander als auch mit Standardmethoden der Bandbreitenwahl verglichen. Während sich einige Erfahrungen aus den bereits in der Literatur beschriebenen, kleineren Simulationsstudien bestätigen lassen, zeigen sich auch einige überraschende Ergebnisse. Insbesondere scheinen die asympotischen Resultate für Stichprobenumfänge bis mindestens $n = 2500$ nahezu irrelevant zu sein. Es zeigt sich jedoch auch, dass einige Versionen des Diskrepanzprinzips gute Resultate liefern.

Im vierten Kapitel wird die Anwendung des Diskrepanzprinzips auf die Wahl der Anzahl der Bins in regulären Histogrammen diskutiert. Hier liegen in der Literatur bisher noch keine theoretischen Ergebnisse vor. Es werden einige den Resultaten aus Kapitel 3 analoge Ergebnisse hergeleitet. Es zeigt sich unter anderem, dass auch die Wahl der Anzahl der Bins mit Hilfe des Diskrepanzprinzips für bestimmte Dichten zu inkonsistenten Schätzern führt. Bei der Dichteschätzung mit regulären Histogrammen gibt es jedoch auch einige wichtige Unterschiede zu Kernschätzern: Zum einen ist der zu wählende Glättungsparameter diskret, zum anderen kann die gesuchte Dichte selbst die Form eines regulären Histogramms besitzen. In diesem Fall wird gezeigt, dass bestimmte Versionen des Diskrepanzprinzips asymptotisch die korrekte Anzahl Bins wählen. Im Anschluss an diese theoretischen Ergebnisse werden auch hier verschiedene Versionen des Diskrepanzprinzips in einer umfangreichen Simulationsstudie mit Standardmethoden der Binanzahlwahl verglichen. Es zeigt sich auch in diesem Fall, dass einige Versionen des Diskrepanzprinzips gute Ergebnisse liefern.

Kapitel 5 behandelt einige Aspekte der Verwendung von Diskrepanzprinzipien in der nichtparametrischen Regression am Beispiel des Nadaraya-Watson-Kernschätzers sowie der kubischen Glättungssplines. In der Literatur wurden für diese Methoden bereits auf der Residuenquadratsumme basierende Versionen des Diskrepanzprinzips vorgestellt. Auch das im Rahmen des Data-Approximation-Ansatzes häufig verwendete Multiresolutionskriterium stellt ein auf einer speziellen Norm im $\mathbb{R}^n$ basierendes Diskrepanzprinzip dar. Es werden im weiteren Verlauf des Kapitels vor allem geometrische Eigenschaften dieser Norm, die nicht nur die Größe von Residuen beurteilt, sondern auch unerwünschte Muster erkennen kann, diskutiert. Diese Ergebnisse (Kapitel 5.2) wurden bereits separat in [Mil08] veröffentlicht. In einer Simulationsstudie werden verschiedene auf der euklidischen Norm der Residuen oder dem Multiresolutionskriterium basierende Varianten des Diskrepanzprinzips mit bekannten Standardmethoden verglichen.

Abschließend werden im sechsten Kapitel die Resultate dieser Arbeit zusammengefasst und diskutiert. Außerdem wird auf weitere offene Fragen eingegangen.

Im Anhang finden sich Details zu den verwendeten mathematischen Hilfsmitteln sowie zu den Simulationen.

Alle Simulationen in dieser Arbeit wurden mit dem Open-Source-Statistikpaket `R` [R D] durchgeführt. Für einige Formelmanipulationen, wie z.B. die Berechnung höherer Ableitungen der verwendeten Kerne, wurde außerdem das ebenfalls frei verfügbare Computeralgebrasystem `Sage` verwendet [S+10].

2. Nichtparametrische Kurvenschätzung, Regularisierung und das Diskrepanzprinzip

2.1. Nichtparametrische Kurvenschätzung, Verlust und Risiko

In dieser Arbeit werden die beiden einfachsten Probleme der nichtparametrischen Kurvenschätzung, die nichtparametrische Dichteschätzung und die nichtparametrische Regression, betrachtet. Diese sollen im Folgenden kurz vorgestellt werden.

Seien $X_1, \ldots, X_n$ unabhängig identisch verteilt gemäß einer Verteilung P, die eine Lebesgue-Dichte $f : \mathbb{R} \longrightarrow [0, \infty)$ besitze. Das Problem der *nichtparametrischen Dichteschätzung* besteht darin, f aus den Beobachtungen zu rekonstruieren. Dazu soll ein *Schätzer* $\hat{f}$ konstruiert werden, so dass für jedes $x \in \mathbb{R}$ $\hat{f}(x) := \hat{f}(x, X_1, \ldots, X_n)$ eine messbare Funktion von $\mathbf{X} = (X_1, \ldots, X_n)$ ist. Es wird vorausgesetzt, dass $f \in \mathcal{F}$ für eine geeignete Funktionenklasse $\mathcal{F}$ gilt. Im Gegensatz zum parametrischen Fall, in dem die Elemente von $\mathcal{F}$ durch einen endlichdimensionalen Parameter $\theta \in \mathbb{R}^k$ indiziert werden können, sind hier vor allem „große“ Klassen $\mathcal{F}$ von Interesse, insbesondere Teilmengen von unendlichdimensionalen Funktionenräumen.

Die *nichtparametrische Regression* beschäftigt sich mit der Modellierung des Zusammenhangs zwischen (mindestens) einer Einfluss- und einer Zielvariablen. Hier liegen n unabhängige Paare von Beobachtungen $(X_1, Y_1), \ldots, (X_n, Y_n)$ vor, wobei vorausgesetzt wird, dass zwischen X_i und Y_i ein Zusammenhang der Form

$$Y_i = f(X_i) + \varepsilon_i$$

mit $X_i \in [0, 1]$ besteht. Die X_i können dabei stochastisch oder deterministisch sein (z.B. $X_i = \frac{i}{n}$). Die ε_i sind unabhängige Zufallsvariablen mit $\mathbb{E}(\varepsilon_i) = 0$. Analog zum Fall der Dichteschätzung wird angenommen, dass $f \in \mathcal{F}$ ist für eine Funktionenklasse $\mathcal{F}$, die wieder – im Gegensatz zum parametrischen Fall – als „groß“ vorausgesetzt wird. Ein Schätzer $\hat{f}$ ist eine messbare Abbildung $\hat{f} : [0, 1] \times [0, 1]^n \times \mathbb{R}^n \longrightarrow \mathbb{R}$ mit $\hat{f}(x) := \hat{f}(x, X_1, \ldots, X_n, Y_1, \ldots, Y_n)$.

Die Güte eines Schätzers kann – zumindest theoretisch – durch eine *Verlustfunktion* gemessen werden. Zum Vergleich wahrer und geschätzter Kurven werden meist Metriken oder Normen auf den entsprechenden Funktionenräumen verwendet. Wichtige, sowohl für die Dichteschätzung als auch für die Regression relevante Beispiele sind die p-Normen d_p:

$$d_p(f,g) := \|f-g\|_p = \left(\int_{\mathbb{R}} |f(x)-g(x)|^p dx\right)^{1/p} \qquad (1 \leq p < \infty), \tag{2.1}$$

$$d_\infty(f,g) := \|f-g\|_\infty = \text{ess sup}_{\mathbb{R}}|f(x)-g(x)|, \tag{2.2}$$

wobei ess sup das essentielle Supremum bezeichnet.

Wünschenswert wäre ein Schätzer, der den Abstand zwischen der wahren und der aus dem vorliegenden Datensatz geschätzten Kurve minimiert. Dies ist jedoch schwierig, da $d(f,\hat{f}(x,X_1,\ldots,X_n))$ bzw. $d(f,\hat{f}(x,X_1,\ldots,X_n,Y_1,\ldots,Y_n))$ nicht nur vom unbekannten f abhängen, sondern auch Zufallsvariablen sind. In der Regel wird daher nicht versucht, den Verlust für die vorliegenden Beobachtungen zu minimieren, sondern das *Risiko*, d.h. den erwarteten Verlust

$$R(f,\hat{f}) := \mathbb{E}_P(d(f,\hat{f}(\cdot,X_1,\ldots,X_n)))$$

im Dichteschätzungsfall und

$$R(f,\hat{f}) := \mathbb{E}_P(d(f,\hat{f}(\cdot,X_1,\ldots,X_n,Y_1,\ldots,Y_n)))$$

im Regressionsfall. Hierbei wird über alle möglichen Realisationen der Stichprobe gemittelt, d.h. das Risiko hängt nicht mehr von der konkret beobachteten Stichprobe ab, wohl aber von der unbekannten Funktion f.

In anderen Fällen ist man nicht an der gesamten Funktion f interessiert, sondern nur an Werten von f an einem oder mehreren Punkten oder anderen Funktionalen $T(f)$. In diesen Fällen werden andere Verlustfunktionen benutzt und es ergeben sich andere optimale Schätzverfahren. In dieser Arbeit werden nur die Verlustfunktionen d_1, d_2 und d_∞ betrachtet.

Liegt das gesuchte Objekt f in einer „großen“, d.h. unendlichdimensionalen Funktionenklasse, so kann man nicht erwarten, f auf Basis von nur endlich vielen Beobachtungen vollständig zu rekonstruieren. Geschätzt werden kann immer nur eine einfache Approximation, meist eine geglättete Version von f. In der Regel hängen nichtparametrische Kurvenschätzer noch von mindestens einem Glättungs- oder Regularisierungsparameter ab. Man erhält so eine ganze Familie möglicher Rekonstruktionen, von denen dann eine ausgewählt werden muß.

Gegenstand dieser Arbeit ist die Untersuchung eines speziellen Ansatzes – des Diskrepanzprinzips – zur Wahl dieser Glättungsparameter. Die datengesteuerte Wahl des Glättungsparameters soll dabei soweit wie möglich von der Frage der Wahl des Schätzers getrennt werden. Daher werden jeweils zwei der bekanntesten Schätzer ausgewählt: Kerndichteschätzer und reguläre Histogramme für die Dichteschätzung sowie Nadaraya-Watson-Kernschätzer und kubische Glättungssplines für die Regression. Diese werden jeweils für eindimensionale Beobachtungen in der einfachsten Version, in der nur ein Glättungsparameter zu wählen ist, betrachtet. In diesen Fällen steht jeweils zum Vergleich eine Vielzahl gut funktionierender Parameterwahlmethoden zur Verfügung. Modifikationen dieser

Schätzer (z.B. mit lokalisierten Glättungsparametern) werden ebenso ausgeklammert wie z.B. Wavelet-Methoden oder Methoden für multivariate Beobachtungen. Es ist jedoch zu erwarten, dass die Verwendung des Diskrepanzprinzips in diesen Fällen ähnliche Vor- und Nachteile aufweist. Für andere Schätzer wie z.B. die Taut-String-Methode (vgl. [DK01] und [DK04]) sind bisher nur Varianten des Diskrepanzprinzips zur Parameterwahl vorgeschlagen worden.

Das Problem der Glättungsparameterwahl soll nun an einem einfachen Beispiel (vgl. [Ros56]) erläutert werden: Der Dichteschätzung an einem festen Punkt mittels Differenzenquotienten. Dies ist ein Spezialfall der in Kapitel 3 ausführlich behandelten Kerndichteschätzung bei Verwendung des Rechteckskerns (3.13). Es sei f die dreimal differenzierbare Dichte einer Verteilung mit Verteilungsfunktion F. Weiter seien $X_1, \dots, X_n$ unabhängige Beobachtungen aus dieser Verteilung und

$$F_n(x) := \frac{1}{n} \sum_{i=1}^{n} \mathbb{I}(X_i \leq x)$$

die zugehörige empirische Verteilungsfunktion. Gesucht ist für ein festes x_0 eine Schätzung $f(x_0)$, wobei die Güte einer Schätzung $\hat{f}(x_0)$ durch den mittleren quadratischen Fehler (MSE)

$$\mathbb{E}((\hat{f}(x_0) - f(x_0))^2)$$

beurteilt werden soll. Die natürliche Schätzung F_n für F lässt sich nicht differenzieren, jedoch können zentrale Differenzenquotienten betrachtet werden. So erhält man für jedes h einen Schätzer

$$\hat{f}_h(x_0) := \frac{F_n(x_0 + h) - F_n(x_0 - h)}{2h}.$$

Eine Aufspaltung des MSE in eine Varianz- und eine Biaskomponente ergibt unter Verwendung der Taylorformel den folgenden Ausdruck für $n \longrightarrow \infty, h \longrightarrow 0$ (vgl. [Ros56], S. 834):

$$\begin{aligned}
\mathbb{E}((\hat{f}(x_0) - f(x_0))^2) &= \mathbb{E}((\hat{f}(x_0) - \mathbb{E}\hat{f}(x_0) + \mathbb{E}\hat{f}(x_0) - f(x_0))^2) \\
&= \mathbb{E}(\hat{f}(x_0) - \mathbb{E}\hat{f}(x_0))^2 + (f(x_0) - \mathbb{E}\hat{f}(x_0))^2 \\
&= \left(\frac{1}{2hn} f(x_0) + o((hn)^{-1}) \right) + \left(\frac{h^4}{36} |f^{(2)}(x_0)|^2 + o(h^4) \right).
\end{aligned}$$

Der erste Summand ist dabei die asymptotische Varianz, der zweite Summand der asymptotische quadrierte Bias. Der Bias hängt dabei nur von h ab und wird klein, wenn h klein wird, während die Varianz klein wird, wenn hn groß wird. Damit die Schätzung bezüglich des gewählten Kriteriums konsistent ist, sind für datenunabhängig gewähltes h die Bedingungen

$$h \longrightarrow 0, nh \longrightarrow \infty$$

notwendig und hinreichend. Als optimale Wahl von h ergibt sich (vgl. [Ros56])

$$h = \left(\frac{9}{2} \frac{f(x_0)}{|f^{(2)}(x_0)|^2} \right)^{-1/5} n^{-1/5}.$$

Hier zeigt sich bereits das Problem, das im Folgenden bei allen betrachteten Methoden eine Rolle spielt: Selbst wenn ein expliziter Ausdruck für die optimale Wahl des Glättungsparameters h existiert, so

- ist die optimale Wahl nur unter – in der Regel nicht empirisch überprüfbaren – Bedingungen (meistens die Glattheit betreffend) an f gültig,
- gilt die explizite Darstellung häufig nur asymptotisch und
- hängt die optimale Wahl von der gesuchten Dichte f ab.

Aufgrund des letzten Punktes ist die optimale Bandbreite in der Praxis nicht implementierbar. Vorausgesetzt, f ist hinreichend glatt, würde die deterministische Wahl $h = cn^{1/5}$ für ein beliebiges $c > 0$ zumindest die optimale Konvergenzrate garantieren. Allerdings kann der MSE hier um einen unbekannten, beliebig großen Faktor größer sein als im optimalen Fall, was insbesondere bei kleinem n zu einem sehr schlechten Verhalten des Schätzers führen kann.

Einen Ausweg bieten datengesteuerte Methoden zur Wahl des Glättungsparameters, von denen in der Literatur zahlreiche vorgeschlagen werden. Die meisten davon fallen in eine von zwei großen Gruppen. In der ersten Gruppe wird versucht, die unbekannten, von f abhängenden Konstanten in einem expliziten Ausdruck für die optimale Wahl des Glättungsparameters entweder direkt aus den Daten zu schätzen oder durch entsprechende Werte für ein parametrisches Referenzmodell zu ersetzen. In der zweiten Gruppe wird versucht, für verschiedene Wahlen des Parameters das Risiko möglichst unverzerrt zu schätzen, und dann den Parameter zu wählen, für den die Schätzung des Risikos am kleinsten ist.

In dieser Arbeit wird eine dritte Gruppe untersucht, die auf dem sogenannten *Diskrepanzprinzip* beruht. Dieses wird bei der Regularisierung inverser Probleme häufig benutzt, hat aber in der Statistik bislang kaum Beachtung gefunden. Zur Erläuterung ist zunächst ein kleiner Exkurs über inverse bzw. schlecht gestellte Probleme notwendig.

2.2. Inverse Probleme

Die Theorie der Lösung sogenannter *inverser* oder *schlechtgestellter Probleme* stellt ein großes Teilgebiet der angewandten Mathematik dar, das zunächst nicht mit der Statistik verbunden und näher bei der Funktionalanalysis oder auch der numerischen Mathematik angesiedelt ist. Die Darstellung folgt im wesentlichen der klassischen Monographie von Tikhonov und Arsenin [TA77]. Tikhonov entwickelte zum einen als erster systematisch Regularisierungsverfahren zur approximativen Lösung schlecht gestellter Probleme, zum anderen beschränken sich die meisten neueren Lehrbücher und Monographien (z.B. [EHN00]) auf Probleme in Hilberträumen, für die wesentlich einfachere und effektivere Ansätze existieren, die aber für die hier zu behandelnden Fragestellungen weniger relevant sind.

Viele inverse Probleme stammen aus physikalischen Anwendungen. Meist geht es darum, aus beobachteten „Wirkungen" auf die zugrundeliegende „Ursache" zu schließen. Da oft sehr unterschiedliche „Ursachen“ sehr ähnliche „Wirkungen“ zur Folge haben, sind Probleme dieses Typs in der Regel erheblich schwieriger zu lösen als direkte Probleme, bei denen von der „Ursache“ auf die „Wirkung“ zu schließen ist.

Definition 2.1. Es seien $\mathcal{U}$ und $\mathcal{V}$ zwei metrische Räume, $A : \mathcal{U} \longrightarrow \mathcal{V}$ ein Operator. Gesucht werde für ein gegebenes $y \in \mathcal{V}$ eine Lösung der Gleichung

$$Ax = y. \tag{2.3}$$

Das Problem (2.3) heißt *gut gestellt im Sinne Hadamards (well posed in the sense of Hadamard)*, falls

1. zu jedem $y \in \mathcal{V}$ mindestens eine Lösung $x \in \mathcal{U}$ existiert mit $Ax = y$;
2. zu jedem $y \in \mathcal{V}$ höchstens eine Lösung $x \in \mathcal{U}$ existiert mit $Ax = y$;
3. die Lösung stabil ist, d.h. für jedes $\varepsilon > 0$ existiert ein $\delta > 0$, so dass zu zwei rechten Seiten y_1 und y_2 der Gleichung (2.3) $d_{\mathcal{V}}(y_1, y_2) < \delta$ für die zugehörigen Lösungen x_1 und x_2 $d_{\mathcal{U}}(x_1, x_2) < \varepsilon$ impliziert.

Das Problem heißt *schlecht gestellt im Sinne Hadamards (ill posed in the sense of Hadamard)*, falls eine der drei Bedingungen verletzt ist.

Bedingungen 1 und 2 sind algebraischer Natur und betreffen die Existenz und Eindeutigkeit von Lösungen bei exakt gegebener rechter Seite von (2.3). Existenz und Eindeutigkeit können stets durch Einschränkungen des Bild- oder Urbildraumes erzwungen werden, was aber im Hinblick auf das gegebene Sachproblem nicht immer sinnvoll ist. Die dritte Bedingung ist topologischer Art. Sind die ersten beiden Bedingungen erfüllt, so besitzt der Operator A eine Inverse A^{-1}. Die dritte Bedingung entspricht dann der Stetigkeit von A^{-1}, was in allen Fällen relevant ist, in denen die rechte Seite in irgendeiner Form fehlerbehaftet ist, weil sie z.B. durch Messungen empirisch erhoben wurde oder numerisch nur mit einer beschränkten Genauigkeit auszuwerten ist. Die Stetigkeit lässt sich mathematisch stets durch Übergang zu einer stärkeren Topologie auf $\mathcal{U}$ oder zu einer schwächeren Topologie auf $\mathcal{V}$ erzwingen. Jedoch sind die Topologien auf $\mathcal{U}$ und $\mathcal{V}$ in der Regel durch das Sachproblem (interessierende Aspekte der Lösung bzw. Art der Messungenauigkeit) vorgegeben. Ist Bedingung 3 verletzt, können geringe Abweichungen in der rechten Seite von (2.3) zu erheblichen Abweichungen in der zugehörigen Lösung führen. Eine lediglich approximativ bekannte rechte Seite kann außerdem außerhalb des Bildbereichs von A liegen, wenn Bedingung 1 verletzt ist.

Das wohl einfachste und dem Problem der nichtparametrischen Dichteschätzung am engsten verwandte schlecht gestellte Problem ist das der numerischen Differentiation einer nur approximativ bekannten Funktion (vgl. [TA77, Kap. 1.3]). Seien dazu $\mathcal{U} = C(\mathbb{R})$ und $\mathcal{V} = C^1(\mathbb{R})$ die Räume der stetigen bzw. einmal stetig differenzierbaren Funktionen, beide versehen mit der Supremumsnorm d_∞. Der Integraloperator $A : C(\mathbb{R}) \longrightarrow C^1(\mathbb{R})$ sei definiert duch $Ax = \int x$. Betrachte nun zu einer Funktion $y \in C^1(\mathbb{R})$ die Folge

$$y_n(t) := y(t) + \frac{1}{n}\sin(n^2 t).$$

Dann gilt

$$\lim_{n\to\infty} d_\infty(y_n, y) = 0,$$

aber mit $x = A^{-1}y = y'$ und $x_n = A^{-1}y_n = y' + n\sin(n^2 t)$ gilt

$$\lim_{n\to\infty} d_\infty(x_n, x) = \infty.$$

D.h. selbst mit einem geringen Fehler behaftete rechte Seiten von (2.3) können beliebig große Fehler in der Lösung zur Folge haben. Selbstverständlich verschwindet das Problem, wenn z.B. die Metrik auf $C^1(\mathbb{R})$ als $d'(x,z) = d_\infty(x,z) + d_\infty(x',z')$ gewählt wird. Dies ist aber von der Problemstellung her in der Regel nicht sinnvoll, da diese Metrik gerade von der gesuchten Ableitung abhängt.

Um schlecht gestellte Probleme trotzdem zumindest näherungsweise lösen zu können, werden Regularisierungsverfahren benutzt. Die Grundidee ist, statt (2.3) ein modifiziertes, gut gestelltes Ersatzproblem zu lösen. Ausgangspunkt ist die Beobachtung, dass die Lösung von (2.3), wenn sie existiert und die rechte Seite y exakt gegeben ist, die Bedingung

$$x = \operatorname*{argmin}_{z\in\mathcal{U}} d^2_{\mathcal{V}}(Az, y) \tag{2.4}$$

erfüllt. Sind Bedingungen 1 und 2 von Definition 2.1 erfüllt, so entspricht die Lösung von (2.4) genau

$$x = A^{-1}y.$$

Ist Bedingung 3 verletzt, so kann diese Lösung bei Vorliegen einer gestörten rechten Seite $\tilde{y}$ unbrauchbar sein. Tikhonov ([TA77], Kapitel 2) schlug vor, in diesem Fall statt (2.4) für ein $\lambda > 0$ eine regularisierte Lösung $\tilde{x}_\lambda$ zu bestimmen als

$$\tilde{x}_\lambda = \operatorname*{argmin}_{z\in\mathcal{U}} d^2_{\mathcal{V}}(Az, \tilde{y}) + \lambda\Omega(z). \tag{2.5}$$

Das Stabilisierungsfunktional $\Omega : \mathcal{U} \supset D(\Omega) \longrightarrow \mathbb{R}^+$ soll dabei folgende Bedingungen erfüllen (vgl. [TA77], Kapitel 2.2):

1. Die exakte Lösung x gehört zum Definitionsbereich $D(\Omega)$ von Ω.
2. Für jedes $c > 0$ ist die Menge $\mathcal{M}_c = \{z : \Omega(z) \leq c\}$ kompakt bezüglich der Metrik auf $\mathcal{U}$.

Die Addition des Stabilisierungsfunktionals in (2.5) erzwingt die Wahl von approximativen Lösungen „einfacher Gestalt“. Vorwissen über x, z.B. Glattheitseigenschaften, kann in die Konstruktion von Ω einfließen. Die Lösung des Minimierungsproblems lässt sich auffassen als Kompromiss zwischen der naiven Invertierung von A, die sehr instabil ist (bzw. gar nicht möglich, wenn die gestörte rechte Seite außerhalb des Bildes von A liegt) und der Auswahl von Lösungen mit kleinem Wert von Ω. Der Regularisierungsparameter λ steuert dabei die Gewichtung dieser beiden entgegengesetzten Ziele. Ähnlich wie in der nichtparametrischen Kurvenschätzung muss regularisiert werden, um die Lösung zu stabilisieren, es darf aber auch nicht zu stark regularisiert werden, da sonst nur noch das Stabilisierungsfunktional Ω minimiert wird. Es ist klar, dass λ in Abhängigkeit von der Größe der Störung der rechten Seite $\tilde{y}$ gewählt werden sollte: Je kleiner der Fehler, desto weniger muss regularisiert werden. Die übliche asymptotische Betrachtungsweise in diesem Gebiet besteht darin, das Verhalten der Lösungen von (2.5) zu betrachten, wenn die Größe des Fehlers in der rechten Seite von (2.3) gegen 0 konvergiert. Kann z.B. angenommen werden, dass $d_{\mathcal{V}}(y, \tilde{y}) \leq \delta$ für ein $\delta > 0$, so kann gezeigt werden (vgl. z.B. Theorem 1 in Kapitel A1.3.2 in [Vap98]), dass für $\delta \longrightarrow 0$ die Bedingungen

$$\lambda(\delta) \longrightarrow 0, \quad \lim_{\delta \longrightarrow 0} \frac{\delta^2}{\lambda(\delta)} < \infty$$

hinreichend sind für die Konvergenz der regularisierten Lösungen gegen die exakte Lösung, d.h.

$$d_{\mathcal{U}}(\tilde{x}_{\lambda(\delta)}, x) \longrightarrow 0 \text{ für } \delta \longrightarrow 0.$$

Ähnlich wie in der nichtparametrischen Kurvenschätzung sind solche Resultate zur praktischen Wahl des Regularisierungsparameters nur bedingt hilfreich. Es existieren zahlreiche Vorschläge, wie λ in der Praxis gewählt werden kann, von denen einige den Glättungsparameterwahlverfahren in der nichtparametrischen Statistik sehr ähnlich sind, vgl. z.B. [BL10] für eine umfassende Übersicht.

Eines der ältesten Verfahren, um λ in (2.5) zu wählen, ist das erstmals von 1966 von Morozov ([Mor66]) vorgeschlagene *Diskrepanzprinzip* (engl. *discrepancy principle*, auch *residual principle* oder *error principle*, wobei unter diesen Bezeichnungen teilweise auch andere Methoden verstanden werden) ([Mor68]):

> Nevertheless we can discuss the following proposition, which we have called the error principle, which reestablishes the role of the error in the approximate solution of a large group of problems (including ill-posed ones): the magnitude of the error must be in agreement with the accuracy of the assignment of the input data of the problem. Undoubtedly such an agreement may be more or less necessary and may be effected in various ways depending on the character of the original problem and the 'quality' of the input data.

Die Idee ist, dass es bei bekanntem Fehler in der rechten Seite, also z.B.

$$d_{\mathcal{V}}(y, \tilde{y}) \leq \delta$$

nicht sinnvoll ist zu verlangen, dass eine regularisierte Lösung $\tilde{x}$

$$d_{\mathcal{V}}(A\tilde{x}, \tilde{y}) < \delta$$

erfüllt. Das *Diskrepanzprinzip* besagt nun, dass der Regularisierungsparameter λ so zu wählen ist, dass für die Lösung $\tilde{x}_\lambda$ von (2.5)

$$d_{\mathcal{V}}(A\tilde{x}_\lambda, \tilde{y}) = c\delta \tag{2.6}$$

gilt für eine fest gewählte Konstante $c > 1$, vgl. [Mor66]. Zum Teil wird aber auch, wie in Kapitel II.6 von [TA77], $c = 1$ gewählt. Es existieren zahlreiche Varianten des Diskrepanzprinzips. Desweiteren kann es auch analog für andere Regularisierungsverfahren umgesetzt werden, insbesondere eignet es sich auch als Abbruchkriterium für iterative Verfahren. In diesem Fall werden iterativ Lösungen bestimmt (beginnend mit sehr regulären Löungen), bis eine Lösung zum ersten Mal (2.6) (mit „$\leq$“ anstelle von „$=$“) erfüllt. Diverse Versionen des Diskrepanzprinzips werden in praktisch allen Lehrbüchern und Monographien über inverse und schlecht gestellte Probleme vorgestellt, von denen hier nur [TA77] und [EHN00] genannt seien. Eine aktuelle Übersicht und ein praktischer Vergleich mit anderen Methoden der Parameterwahl finden sich außerdem in [BL10].

Im Folgenden sollen unter „Diskrepanzprinzip“ sowohl das allgemeine Prinzip („*das* Diskrepanzprinzip“) als auch jeweils die konkrete Implementierung („*ein* Diskrepanzprinzip“) verstanden werden.

2.3. Das Diskrepanzprinzip in der nichtparametrischen Statistik

Seit langem werden auch in der Statistik inverse Probleme behandelt, vornehmlich inverse Regressionsprobleme, bei denen für feste oder zufällige Designpunkte $t_1, \dots, t_n$ Beobachtungen von

$$\tilde{y}(t_i) = (Af)(t_i) + \varepsilon_t$$

vorliegen und f bestimmt werden soll. Der Operator A ist dabei oft ein Faltungsoperator bzw. etwas allgemeiner ist $Af = y$ eine Fredholm-Integralgleichung 1. Art. Dies ist ein direktes statistisches Analogon zu (2.3), wobei nun die rechte Seite nur in endlich vielen Designpunkten beobachtet wird und mit einem stochastischen Fehler ε_t behaftet ist. Für einen Überblick über Probleme dieses Typs vgl. [BHMR07] und die dort angegebene Literatur bzw. aus Bayesianischer Sicht die Monographie [KS05].

In der Dichteschätzung ist das Entfaltungsproblem wichtig, bei dem aus Beobachtungen von $Z = X+Y$ die Dichte von X rekonstruiert werden soll, wobei X und Y unabhängig sind und Y eine Störvariable mit meist als bekannt vorausgesetzter Dichte ist. Dies führt direkt auf Probleme vom Typ (2.3), vgl. vor allem die Monographie [Mei09]. Ähnliche Probleme treten z.B. bei intervallzensierten Daten auf, vgl. [Gro96]. Weiter treten statistische inverse Probleme in der Ökonometrie auf, z.B. bei der Regression mit Instrumentvariablen, vgl. Kapitel 5 in [Hor09] für eine Übersicht.

Obwohl schon die einfachsten Probleme der nichtparametrischen Kurvenschätzung – also die nichtparametrische Dichteschätzung und die nichtparametrische Regression – Regularisierung erfordern, werden sie selbst in der Regel *nicht* als inverse bzw. schlecht gestellte Probleme angesehen. Exemplarisch sei hier der Anfang der „Lectures on Inverse Problems“ von Groeneboom zitiert ([Gro96], S. 72):

> I will not try to define formally what an inverse problem is, but important aspects of what I call an inverse problem are:

> 1. We only have indirect information about the things we really would like to observe.
> 2. Estimators of the distribution function of the variables of interest do not (and cannot) have the usual $\sqrt{n}$-rate of convergence. In this sense the interval censoring problem, defined below, is an inverse problem, but in contrast the right-, left- or double censoring problems are not inverse problems and indeed have a completely different asymptotic theory.

Auch wenn gelegentlich ein Bezug zwischen statistischen Problemen – auch solchen, die nicht direkt statistische Varianten bekannter deterministischer inverser Probleme sind – und der Theorie inverser Probleme hergestellt wird (vgl. z.B. [Tit85] oder [BL06]), ist die explizite oder implizite Behandlung der nichtparametrischen Kurvenschätzung *als* inverses Problem selten. Ausnahmen bilden die Ansätze von Vapnik („Statistical Learning Theory") und Davies („Data Features" bzw. „Data Approximation"), auf die im Folgenden eingegangen wird. Beide Ansätze wurden über einen längeren Zeitraum und – besonders was die praktische Umsetzung in anwendbare Methoden betrifft – in Zusammenarbeit mit verschiedenen Co-Autoren entwickelt. In beiden Fällen kann die knappe Darstellung dem Ansatz kaum gerecht werden, weshalb nur die für den Kontext der vorliegenden Arbeit interessierenden Aspekte herausgearbeitet werden sollen, insbesondere der Bezug zu inversen Problemen und die Verwendung des Diskrepanzprinzips. Weitergehende, oft für den jeweiligen Ansatz zentralere Aspekte, wie z.B. die unterschiedlichen philosophischen Ausgangspunkte, werden dabei weitgehend ausgeklammert.

Die von Vapnik in den 1960er Jahren begründete und sukzessive mit verschiedenen Co-Autoren weiterentwickelte *Theorie des statististischen Lernens (Statistical Learning Theory,* im Folgenden kurz *SLT)* hat zu einer Vielzahl neuer Konzepte und Methoden geführt. Neben den grundlegenden Beiträgen zur Theorie der empirischen Prozesse (Vapnik-Cervonenkis-Klassen) sind wohl die zunächst für Klassifikationsprobleme entwickelten Support Vector Machines (SVM) am bekanntesten. Eine umfassende Darstellung des Ansatzes findet sich in den Monographien [Vap98] und [Vap00]. Auch wenn die für diese Arbeit relevanten Methoden nur einen Nebenaspekt der SLT darstellen, ist die Verbindung der nichtparametrischen Statistik mit der Theorie inverser bzw. schlecht gestellter Probleme für diesen Ansatz grundlegend. Insbesondere die nichtparametrische Dichteschätzung wird bereits in [VS78] als Lösung einer schlechtgestellten Integralgleichung aufgefasst, später wurden verschiedene bekannte Dichteschätzer als Spezialfall der Tikhonov-Regularisierung erkannt und gleichzeitig das Diskrepanzprinzip als Methode zur Glättungsparameterwahl vorgeschlagen [AV89, VMS92]. Auch die nichtparametrische Regression steht in engem Zusammenhang mit Integralgleichungen, insbesondere wenn die bedingte Verteilung bei gegebener Einflussvariable geschätzt werden soll. Ist nur die Schätzung der Regressionsfunktion das Ziel, ergeben sich aus dieser Formulierung jedoch keine praktischen Ansätze. Der folgende Abriss basiert vor allem auf [VS78], [AV89] und den Monographien [Vap98] und [Vap00], jeweils vor allem Kapitel 7.

Das Problem der Dichteschätzung ist ein Spezialfall des numerischen Differenzierens. Die

Dichte f ist Lösung der Integralgleichung

$$\int \mathbb{I}((x-t) \geq 0) f(t) dt = F(x). \tag{2.7}$$

Die Verteilungsfunktion F ist dabei unbekannt, mit der empirischen Verteilungsfunktion F_n steht aber eine gute Schätzung zur Verfügung. Weiterhin existieren zahlreiche Resultate über das stochastische Verhalten der Differenz $F_n - F$ der geschätzten und der wahren rechten Seite von (2.7). Ist nun $\mathcal{U}$ ein Raum, der die gesuchte Dichte enthält mit einer geeigneten Metrik (z.B. $L_1(\mathbb{R})$ mit d_1, $L_2(\mathbb{R})$ mit d_2 oder $C_b(\mathbb{R})$ mit d_∞) und $\mathcal{V}$ der mit der Supremumsnorm d_∞ versehene Raum der rechtsseitig stetigen beschränkten Funktionen auf $\mathbb{R}$, so ist das Problem (2.7) ein Spezialfall von (2.3), wobei der Operator A durch $Af := \mathbb{I}(\cdot \geq 0) * f$ gegeben ist. Dieses Problem ist im Sinne Hadamards schlecht gestellt, da Bedingungen 2 und 3 von (2.1) verletzt sind. Zum einen ist die als Approximation für F benutzte empirische Verteilungsfunktion kein Integral einer Lebesgue-Dichte (d.h. F_n liegt nicht im Bild von A), zum anderen ist das Problem nicht stabil, d.h. selbst geringe Störungen der wahren rechten Seite F können zu sehr unterschiedlichen Lösungen von (2.7) führen.

Dies stellt einen Spezialfall des oben behandelten numerischen Differenzierens dar. Im Unterschied zum deterministischen Fall liegt allerdings keine obere Schranke für $d_\infty(F_n, F)$ vor, sondern $d_\infty(F_n, F)$ ist eine Zufallsvariable, deren Verteilung jedoch bekannt ist. Auf dieser Grundlage schlugen Vapnik und Stefanyuk in [VS78] die Lösung des Dichteschätzungsproblems mit Hilfe der für inverse Probleme entwickelten Tikhonov-Regularisierung vor und kontrastierten dies mit Kerndichteschätzern. Aidu und Vapnik zeigten in [AV89], dass sich viele der bekannten nichtparametrischen Dichteschätzer als Tikhonov-Regularisierungen von (2.7) oder des etwas allgemeineren Problems

$$B \int \mathbb{I}((x-t) \geq 0) f(t) dt = BF(x) \tag{2.8}$$

für einen nicht singulären Operator B auffassen lassen. Je nach Wahl der zugrundeliegenden Räume, ihrer Metriken, des Operators B und des Regularisierungsfunktionals erhält man so Kerndichteschätzer, bestimmte Orthogonalreihenschätzer sowie bestimmte Splineschätzer. Zur Wahl des Regularsierungsparameters wurde in der selben Arbeit das Diskrepanzprinzip vorgeschlagen. Satz 3.7 der vorliegenden Arbeit ist eine leichte Verallgemeinerung des Satzes aus Kapitel 3 von [AV89], der die Verwendung des Diskrepanzprinzips zur Wahl der Bandbreite für Kerndichteschätzer behandelt. In der Simulationsstudie [Mar89] wurden Dichteschätzer mit durch das Diskrepanzprinzip gewählten Glättungsparametern mit anderen Dichteschätzern verglichen. In [VMS92] werden theoretische Resultate zu Projektionsdichteschätzern, deren Glättungsparameter mit Hilfe eines auf dem Cramér-von Mises-Tests basierenden Diskrepanzprinzips gewählt werden, hergeleitet.

Das Problem der bedingten Dichteschätzung, das eine Verallgemeinerung der nichtparametrischen Regression darstellt, lässt sich außerdem als Lösung der Integralgleichung

$$\int_{-\infty}^{y} \int_{-\infty}^{x} f(y'|x') dF(x') dy' = F(y,x)$$

formulieren (vgl. Kapitel 7.1.3 in [Vap00]), wobei statt F nur die gemeinsame empirische Verteilungsfunktion F_n der Beobachtungen $(x_1, y_1), \ldots, (x_n, y_n)$ gegeben ist. Die bedingte Dichte $f(y|x)$ enthält jedoch viel mehr Information als die in der Regel nur gesuchte Regressionsfunktion $r(x) = \int yf(y|x)dy$. Daher ist dieser Ansatz zur Lösung des Problems der nichtparametrischen Regression weniger geeignet.

Im Gegensatz dazu hat der *Data Features*- bzw. *Data Approximation*-Ansatz (im folgenden kurz: DA) von Davies seine Wurzeln in der robusten Statistik bzw. explorativen Datenanalyse. Die grundlegende Philosophie wird in [Dav95], [Dav02], [Dav03] und [Dav08] entwickelt, Anwendungen speziell auf die nichtparametrische Regression bzw. Dichteschätzung finden sich u.a. in [DK01], [DK04], [DM08] und [DKM09].

Ausgangspunkt ist ein vorliegender Datensatz. Herkömmliche statistische Methoden, seien sie frequentistisch oder Bayesianisch, basieren darauf, dass ein statistisches Modell postuliert wird, das die Daten erzeugt haben soll. Davon ausgehend wird dann versucht, interessierende Parameter möglichst optimal zu schätzen, Tests durchzuführen usw. DA versucht, ohne die Annahme eines wahren datengenerierenden Modells auszukommen, da die Existenz eines solchen nicht empirisch verifiziert werden kann. Statistische Verfahren werden als Prozeduren oder Algorithmen aufgefasst, die im Idealfall auf möglichst viele Datensätze sinnvoll angewendet werden können. Dies erfordert eine gute Kenntnis der Stärken und Schwächen solcher Methoden in möglichst vielen verschiedenen Situationen. Neben der Anwendung auf reale Datensätze sollte dabei auch das Verhalten der Prozedur in Fällen untersucht werden, in denen die Daten aus einem bekannten Modell stammen. Dies kann sowohl durch Simulationen wie auch durch theoretische (z.B. asymptotische) Resultate geschehen. Diese Modelle werden im Rahmen von DA auch als *test beds* bezeichnet. Unter bestimmten, eventuell sehr restriktiven, Annahmen bewiesene Resultate sollen dabei Eigenschaften der Prozedur kenntlich machen. Das Verfahren soll jedoch ausdrücklich nicht nur dann anwendbar sein, wenn diese – in der Praxis meist nicht überprüfbaren – Voraussetzungen erfüllt sind.

Da stochastische Modelle im Rahmen von DA als Konstrukte betrachtet werden, denen nicht unbedingt eine außermathematische Realität zukommt (auch nicht im Sinne einer Idealisierung oder Approximation), ist es auch nicht Ziel eines statistischen Verfahrens, das Modell, das die Daten erzeugt *hat*, einige seiner Parameter oder einen Bereich, in dem diese liegen, zu identifizieren. Zu einem vorliegenden Datensatz soll stattdessen ein *adäquates* Modell gefunden werden, wobei „adäquat“ bedeutet, dass das Modell die Daten erzeugt haben *könnte* bzw. die tatsächlich beobachteten Daten nicht von aus dem Modell simulierten Daten zu unterscheiden sind. Prinzipiell kann kein Modell ausgeschlossen werden, das die Daten erzeugt haben könnte. Sind bestimmte Kenngrößen gesucht, so kommen alle Werte in Frage, die sich aus adäquaten Modellen ergeben.

Gibt es adäquate Modelle verschiedener Komplexität, so ist ein Modell geringster Komplexität zu wählen. Diese Situation liegt insbesondere in den in dieser Arbeit betrachteten Problemen der nichtparametrischen Kurvenschätzung vor. Um zu praktisch einsetzbaren Prozeduren zu gelangen, sind die Konzepte der „Adäquatheit“ und der „Komplexität“ zu operationalisieren, in dem für beide formelle und praktisch überprüfbare Kriterien angegeben werden. Hierbei kann dann Wissen über das zugrundeliegende Sachproblem und den

Vorgang der Datenerhebung einfließen.

Die Formalisierung der *Adäquatheit* geschieht über die Definition sogenannter *Data Features*:

Definition 2.2. [Dav95, Sec. 4.1] Sei $\mathcal{X}$ eine Menge möglicher Stichproben und $\mathcal{P}$ eine Familie von Wahrscheinlichkeitsmaßen auf $\mathcal{X}$. Eine Abbildung $\mathcal{F} : \mathcal{X} \times \mathcal{P} \longrightarrow \{0, 1\}$ heißt *Data Feature*. Gilt für alle $P \in \mathcal{P}$ für eine Zufallsvariable X mit Werten in $\mathcal{X}$, die gemäß P verteilt ist

$$P(\mathcal{F}(X, P) = 1) \geq \alpha,$$

so heißt $\mathcal{F}$ ein α-Feature.

Für ein betrachtetes Modell P und einen Datensatz x lässt sich ohne Bezug auf ein datenerzeugendes Modell entscheiden, ob das Feature vorliegt ($\mathcal{F}(X, P) = 1$) oder nicht ($\mathcal{F}(X, P) = 0$). Insbesondere kann x auch deterministisch oder durch Simulationen erzeugt worden sein. Für die nichtparametrische Kurvenschätzung sind insbesondere solche Features relevant, die einen Abstand zwischen dem Datensatz und der geschätzten Kurve messen. Im Falle der Dichteschätzung wird z.B. in Section 4.2 von [Dav95] das Feature

$$\mathcal{F}(x, P^n) = \mathbb{I}(d_\infty(F_n, F) \leq c_\alpha)$$

vorgeschlagen, wobei P ein Wahrscheinlichkeitsmaß auf $\mathbb{R}$ mit zugehöriger stetiger Verteilungsfunktion F, $P^n := P \otimes \cdots \otimes P$, F_n die aus $x = (x_1, \ldots, x_n)$ gebildete empirische Verteilungsfunktion und c_α das $(1 - \alpha)$-Quantil der Verteilung von $d_\infty(F, F_n)$ ist, die ja nicht von F abhängt. Es können auch andere Abstände, wie z.B. die Kuiper-Metrik, bzw. Tests verwendet werden. Es ist aber wichtig, dass ein direkter Vergleich zwischen P und x möglich ist, was bei Abständen, die auf Dichten basieren, nicht der Fall ist. Im Folgenden soll der Einfachheit halber das Produktmaß P^n mit der Verteilungsfunktion der einzelnen Komponenten F identifiziert werden.

Die Menge aller Modelle, für die für einen vorliegenden Datensatz das interessierende Feature vorliegt, wird auch als *Adäquatheitsregion* bezeichnet. Im Falle der Dichteschätzung kann das z.B. die Menge

$$\mathcal{A} = \{G \mid G \text{ absolut stetige Verteilungsfunktion und } d_\infty(F_n, G) \leq F_{KS,n}^{-1}(1 - \alpha)\} \quad (2.9)$$

sein, wobei $F_{KS,n}^{-1}$ die Quantilsfunktion von $d_\infty(F_n, F)$ bezeichnet. Im einfachsten Modell der nichtparametrischen Regression

$$y_i = f(t_i) + \varepsilon_i$$

mit unabhängigen $N(0, \sigma^2)$-Fehlern $\varepsilon_1, \ldots, \varepsilon_n$ und festen Designpunkten $t_1, \ldots, t_n$ können Features bzw. die resultierende Adäquatheitsregion über Eigenschaften des Residualvektors definiert werden. Ein Modell wird dann mit der Regressionsfunktion f identifiziert, bzw. mit dem Vektor $(f(t_1), \ldots, f(t_n))$ der Auswertungen von f, der der Einfachheit halber ebenfalls mit f bezeichnet werden soll. Die Adäquatheitsregion kann dann zum Beispiel die Form

$$\mathcal{A} = \{g \mid \|y - g\| < \sigma c_{n,\alpha}\} \tag{2.10}$$

haben. Meistens wird die Adäquatheitsregion durch das sogenannte Multiresolutionskriterium definiert (vgl. vor allem [DK01], [DM08] und [DKM09]). Wie in Kapitel 6 dieser Arbeit gezeigt wird, entspricht dies jedoch genau (2.10) mit der Wahl einer speziellen Norm (vgl. auch [Mil08]). Da σ in der Regel unbekannt ist, wird eine Schätzung verwendet.

Ein adäquates Modell wird im Rahmen von DA auch als *Approximation* bezeichnet, womit aber entgegen dem herkömmlichen Sprachgebrauch gemeint ist, dass das Modell die Daten approximiert und nicht den die Daten generierenden Zufallsmechanismus, dessen Existenz ja explizit nicht unterstellt wird.

Die Wahl eines adäquaten Modells vermeidet im Falle der Kurvenschätzung eine zu große Verzerrung. Sie verhindert aber natürlich nicht die Überanpassung. Dies wird durch die Forderung, das Modell geringster Komplexität zu wählen, erzielt. Die *Komplexität* eines Modells kann z.B. durch Anzahl der freien Parameter gemessen werden. Im Falle der nichtparametrischen Kurvenschätzung sind die Modellklassen unendlichdimensional und als Komplexitätsmaß werden häufig die Anzahl der Extremwerte einer Funktion oder Glattheitsmaße, die z.B. auf Ableitungen basieren, verwendet. Für solche Komplexitätsmaße lassen sich in der Regel ohne weiteres Vorwissen nur einseitige Konfidenzintervalle angeben, ([Don88]):

> The quantity of interest is the complexity of a system – size, norm or number of components. This should make the phenomenon intuitively understandable. Empirical data can usually invalidate simple models, i.e. prove that a system possesses at least a certain degree of complexity. However, data can not usually rule out very complex models which differ from simpler ones in ways that are not detectable given the quantity and quality of data at hand. In short, measures of complexity usually admit of empirical lower but not upper bounds.

Idealerweise würde im Rahmen von DA nun unter allen Modellen, die für die vorliegenden Daten adäquat sind, ein Modell mit der geringsten Komplexität gewählt. Dies entspricht der Minimierung eines Komplexitätsfunktionals unter der Nebenbedingung, dass das Modell in der Adäquatheitsregion liegt. In vielen Fällen ist dies nicht praktikabel, weil entweder keine Lösungsmethode für das Optimierungsproblem bekannt oder eine exakte Lösung rechnerisch zu aufwändig ist. Die meisten der vorgeschlagenen Methoden basieren auf einem Algorithmus, der beginnend mit einem einfachen Modell solange Modelle aufsteigender Komplexität generiert, bis eines innerhalb der Adäquatheitsregion liegt, vgl. Abschnitt 9 in [Dav08]. Damit wird – unter anderem Namen – das Diskrepanzprinzip implementiert.

Es wurden bislang verschiedene datenapproximative Prozeduren vorgeschlagen. Schon die Beispiele 8-10 in Abschnitt 5 von [Dav95] befassen sich mit der nichtparametrischen

Dichteschätzung. Gesucht wird jeweils die einfachste Dichte, die einen gegeben Datensatz erzeugt haben könnte. Einfachheit wird gemessen durch die Anzahl der Modalwerte (Bsp. 8) oder die Anzahl der Intervalle auf denen die Dichte konvex bzw. konkav ist (Bsp. 9). Die Adäquatheit wird definiert durch eine Bedingung vom Typ (2.9) unter Verwendung der Kuiper-Metrik, teilweise kombiniert mit weiteren Data Features, die nicht mit dem Diskrepanzprinzip im Zusammenhang stehen. Beispiel 10 behandelt die Kerndichteschätzung. Dabei wird die Komplexität durch die Bandbreite gemessen, Adäquatheit durch eine Kuiper-Bedingung sowie ein weiteres sogenanntes Extremwertfeature. Bis auf letzteres entspricht dies den in Kapitel 3 dieser Arbeit behandelten Ansätzen zur Bandbreitenwahl. Davies bemerkt auch, dass die optimale Bandbreite bzgl. des L_2-Risikos mit dem Kuiper-Feature unverträglich ist.

Auch der Taut-String-Ansatz zur Dichteschätzung in [DK04] misst die Komplexität durch die Anzahl der Modalwerte. Adäquatheit wird über Bedingungen vom Typ (2.9) für die Kolmogorov-Metrik oder Kuiper-Metriken (auch sogenannte Kuiper-Metriken höherer Ordnung) gemessen. Zusätzlich wird noch eine Adäquatheitsregion vorgeschlagen, die über Differenzen von Kuiper-Metriken verschiedener Ordnungen definiert wird. Die Grundidee basiert auf einer von Hartigan und Hartigan in [HH85] vorgeschlagenen Teststatistik für Hypothesen über die Modalität von Dichten. Es wird ein Kolmogorov-Schlauch um die empirische Verteilungsfunktion gelegt, und dann die Funktion mit der kleinsten Bogenlänge innerhalb dieses Schlauches bestimmt. In der Version von Davies und Kovac wird deren Ableitung dann als Dichteschätzung verwendet. Die Schlauchbreite ist hier der Glättungsparameter. Bei Verwendung des Kolmogorov-Abstandes kann dieser direkt einem festen Quantil entsprechend gewählt werden, für die verschiedenen Kuiper-Kriterien wird mit einem breiten Schlauch begonnen und die Schlauchbreite so lange verringert, bis das Kriterium erfüllt ist. Es werden auch Varianten vorgeschlagen, in denen der Schlauch lokal verengt wird. Damit ist der Glättungsparameter nicht mehr eindimensional, die Methode der Schlauchbreitenwahl kann aber als eine Art verallgemeinertes Diskrepanzprinzip angesehen werden. In der selben Arbeit wird auch eine Variante zur Schätzung von Spektraldichten vorgestellt. In der Simulationsstudie [DGNW09] wird außerdem ein reguläres Histogramm mit der minimalen Anzahl Bins benutzt, für die das in [DK04] vorgeschlagene, auf Differenzen von Kuiper-Metriken basierende, Kriterium erfüllt ist. Diskrepanzprinzipien für die Wahl der Binanzahl in regulären Histogrammen werden in der vorliegenden Arbeit in Kapitel 4 untersucht.

Weiter existieren datenapproximative Ansätze für die nichtparametrische Regression. In Beispiel 10 von [Dav95] sowie in [DK01] werden Regressionsfunktionen mit kleinstmöglicher Anzahl lokaler Maxima gesucht, wobei die Adäquatheit über die Vorzeichen der Residuen definiert wird. Die meisten anderen Prozeduren benutzen Adäquatheitsregionen vom Typ (2.10) unter Verwendung der sogenannten Multiresolutionsnorm. Die Version der Taut-String-Methode für den Regressionfall, die in [DK01] vorgeschlagen wird und einer Splinemethode von Mammen und van de Geer (vgl. [MvdG97]) ähnelt, verwendet als Komplexitätsmaß ebenfalls die Anzahl Modalwerte. Der eigentliche Glättungsparameter ist jedoch wiederum die Schlauchbreite. Eine weitere Methode basiert auf kubischen Glättungssplines [DM08], wobei die Komplexität über die Norm der zweiten Ableitung definiert wird. Auf diese Methode wird in Kapitel 5 dieser Arbeit noch näher eingegangen. Weitere Varianten finden sich in [DKM09]. Die meisten der genannten Regressionsmethoden sind räumlich adaptiv, weshalb die Glättungsparameter nicht mehr nur eindimensional sind. Dies macht die resultierenden Algorithmen schwierig zu analysieren.

Außerdem wird das Diskrepanzprinzip explizit oder implizit noch in einigen weiteren Arbeiten verwendet, die nicht den oben ausführlich besprochenen Ansätzen des Maschinellen Lernens bzw. der Datenapproximation zuzuordnen sind. Für Kerndichteschätzer schlagen Eggermont und LaRiccia eine Variante des Diskrepanzprinzips vor, die unter Standardvoraussetzungen eine optimale Konvergenzrate erzielt (vgl. [EL96] und Kapitel 7.6 in [EL01]). Diese Version wird in der vorliegenden Arbeit in Kapitel 3 näher behandelt. Dieselben Autoren übertragen ihre Methode weiterhin in Kapitel 7.7 von [EL01] auf einen bestimmten Penalized-Maximum-Likelihood-Schätzer. Hjort und Walker zeigen in einem anderen Kontext, dass ein Kerndichteschätzer mit optimaler Bandbreite asymptotisch außerhalb jedes Kolmogorov-Schlauches liegt [HW01], was Implikationen für mit Hilfe des Diskrepanzprinzips gewählte Bandbreiten hat. Im Kontext der Bild- bzw. Clusteranalyse schlagen Frederix und Pauwels [FP04] einen Algorithmus zur Dichteschätzung vor, der ein Glattheitsmaß unter einer Nebenbedingung an den Abstand zwischen empirischer und geschätzter Verteilungsfunktion minimiert. Die Nebenbedingung entspricht einem Kolmogorov-Smirnov- bzw. Cramér-von Mises-Test zum Niveau 0.5. Die Autoren liefern jedoch keine theoretische Analyse. In [AGG02] werden Dichteschätzung und Entfaltungsprobleme mit bezug auf die Theorie inverser Probleme in einem einheitlichen Rahmen behandelt, jedoch ohne Benutzung des Diskrepanzprinzips.

In der Regression wird das Diskrepanzprinzip vor allem im Kontext von Glättungssplines untersucht: Bereits Reinsch schlägt in [Rei67] ein auf der Residuenquadratsumme basierendes Diskrepanzprinzip vor, das u.a. von Wahba in [Wah75] näher untersucht wird. Dieser Ansatz wird in der vorliegenden Arbeit in Kapitel 5 diskutiert. Eine ähnliche Methode wird auch in der Bildverarbeitung benutzt, vgl. z.B. [GK92]. Mächler untersucht eine spezielle Klasse von Penalized-Likelihood-Regressionsschätzern und schlägt vor, den größten Glättungsparameter zu wählen, für den die Autokorrelationsfunktion der Residuen keine relevante Struktur mehr aufweist [Mäc95]. Diese Methode hat eine gewisse Ähnlichkeit mit dem Multiresolutionskriterium, basiert aber nicht auf einer Norm des Residualvektors.

3. Dichteschätzung I: Kerndichteschätzung

3.1. Einleitung und Definitionen

Seien $X_1, \ldots, X_n$ unabhängig identisch verteilt gemäß einer Verteilung P mit Lebesgue-Dichte f und zugehöriger Verteilungsfunktion F. Bezeichne weiter

$$F_n(x) := \frac{1}{n} \sum_{i=1}^{n} \mathbb{I}(X_i \leq x)$$

die empirische Verteilungsfunktion. Während F sehr einfach durch F_n geschätzt werden kann (sogar mit einer $\sqrt{n}$-Konvergenzrate bzgl. d_∞), ist die Schätzung von f ein schwierigeres Problem. Neben Histogrammen, die in Kapitel 4 betrachtet werden, stellen Kerndichteschätzer den wohl populärsten Ansatz der nichtparametrischen Dichteschätzung dar. Sie basieren auf der Idee, die Ableitung der durch Faltung mit einer Kernfunktion K geglätteten empirischen Verteilungsfunktion als Schätzung der Dichte zu verwenden und wurden erstmals von Rosenblatt [Ros56] und Parzen [Par62] eingeführt. Es existieren zahlreiche Modifikationen und Erweiterungen. Einführungen in die Kerndichteschätzung und Zusammenfassungen theoretischer Resultate finden sich unter anderem in [PR83], [WJ95] und [Tsy09]. Das Integral des Kerndichteschätzers kann auch zur Schätzung der Verteilungsfunktion benutzt werden, vgl. z.B. [Yam73] oder Kapitel 9 in [PR83].

Definition 3.1. Eine Funktion $K : \mathbb{R} \longrightarrow \mathbb{R}$ heißt *Kern* der Ordnung ℓ für $\ell \geq 2, \ell \in \mathbb{N}$, falls

$$u^j K(u) \in L_1(\mathbb{R}) \qquad (j = 0, \ldots, \ell), \tag{3.1}$$

$$\int K(u) du = 1, \tag{3.2}$$

$$\int u^j K(u) du = 0 \qquad (j = 1, \ldots, \ell - 1), \tag{3.3}$$

$$\int u^\ell K(u) du \in \mathbb{R} \setminus \{0\}. \tag{3.4}$$

Nichtnegative Kerne der Ordnung 2 entsprechen genau Wahrscheinlichkeitsdichten von stetigen Verteilungen mit Erwartungswert 0 und endlicher Varianz, während für allgemeines ℓ K bzw. $\mathbb{K}$ Dichte bzw. Verteilungsfunktion eines signierten Maßes sind (vgl. Definition

A.1). Kerne der Ordnung ℓ für $\ell > 2$ sind notwendig für einige x negativ, ebenso die resultierenden Dichteschätzer. Kerne höherer Ordnung werden trotzdem zur Konstruktion von Kerndichteschätzern verwendet, da sie bei hinreichender Glattheit von f zu einem geringeren Risiko führen können, vgl. zum Beispiel die Diskussion in Kapitel 1 von [Tsy09].

Definition 3.2. Ein Kern heißt *symmetrisch*, falls

$$K(-u) = K(u) \qquad (x \in \mathbb{R}). \tag{3.5}$$

Meistens werden zur Dichteschätzung symmetrische Kerne verwendet. Kerne ungerader Ordnung können wegen (3.4) nicht symmetrisch sein. Weiter werden skalierte Versionen des Kernes und seines Integrals benötigt:

Definition 3.3. Für einen Kern K und $h > 0$ werden definiert:

$$K_h(u) := \frac{1}{h} K\left(\frac{u}{h}\right), \tag{3.6}$$

$$\mathbb{K}(u) = \int_{-\infty}^{u} K(t)dt, \tag{3.7}$$

$$\mathbb{K}_h(u) = \mathbb{K}\left(\frac{u}{h}\right). \tag{3.8}$$

Definition 3.4. Für Zufallsvariablen $X_1, \ldots, X_n$, einen Kern K der Ordnung $\ell \in \mathbb{N}$ und eine *Bandbreite* $h > 0$ heißt die Funktion $x \longrightarrow \hat{f}_h(x)$ mit

$$\hat{f}_h(x) := \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{x - X_i}{h}\right) = \frac{1}{n} \sum_{i=1}^{n} K_h(x - X_i) \tag{3.9}$$

Kerndichteschätzer (auch: *Parzen-Rosenblatt-Schätzer*). Ein entsprechender *Kernschätzer der Verteilungsfunktion* ergibt sich als

$$\hat{F}_n^h(x) := \int_{-\infty}^{x} \hat{f}_h(t)dt = \frac{1}{n} \sum_{i=1}^{n} \mathbb{K}_h(x - X_i) = (F_n * K_h)(x). \tag{3.10}$$

Zwei der am weitesten verbreiteten Kerne sind der *Gaußkern*

$$K^{\mathcal{N}}(x) := \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) \tag{3.11}$$

und der *Epanechnikov-Kern*

$$K^{\mathcal{E}}(x) := \frac{3}{4}(1 - x^2)\mathbb{I}(|x| \leq 1). \tag{3.12}$$

Diese Kerne werden auch in der Simulationsstudie in Kapitel 3.4 benutzt. Beide sind Wahrscheinlichkeitsdichten. Wichtige Eigenschaften sind in der Tabelle B.1 im Anhang zusammengestellt. Außerdem wird an einigen Stellen der Rechteckskern

$$K^{\mathcal{U}}(x) := \frac{1}{2}\mathbb{I}(|x| \leq 1) \tag{3.13}$$

benutzt. Einige theoretische Betrachtungen sind für diesen Kern besonders einfach. Für die praktische Anwendung ist er jedoch weniger geeignet, da die resultierenden Kernschätzer nicht stetig sind.

Die Güte eines Dichteschätzers $\hat{f}$ kann mit Hilfe verschiedener Verluste gemessen werden. Die größte Rolle spielen die p-Normen (2.1) bzw. (2.2). Mit L_1 und L_2 werden im Folgenden die Räume der integrierbaren bzw. quadratintegrierbaren reellen Funktionen bezeichnet, mit L_∞ der Raum der messbaren essentiell beschränkten reellen Funktionen, versehen jeweils mit den Normen d_1, d_2 bzw. d_∞.

Am häufigsten wird in der Literatur $d_2(\hat{f}, f)$ benutzt. Dies liegt hauptsächlich daran, dass sich dieser Verlust bzw. das zugehörige L_2-Risiko $\mathbb{E}_f(d_2^2(\hat{f}, f))$, das auch als *Mean Integrated Squared Error* bzw. *MISE* bezeichnet wird, mathematisch einfach handhaben lässt. Es gibt jedoch Dichten, deren Quadrat nicht integrierbar ist. Sie liegen daher nicht in L_2, während jede Dichte in L_1 liegt. Da die Metrik d_1 außerdem noch über weitere interessante Eigenschaften verfügt (z.B. Invarianz unter monotonen Transformationen), wird sie oft als das natürliche Gütemaß für Dichteschätzer angesehen (vgl. [DG85]). Jedoch ist die theoretische Untersuchung wesentlich komplizierter. In den Simulationen in dieser Arbeit werden die Risiken bezüglich d_1, d_2 und – soweit sinnvoll – d_∞ betrachtet. Ein bekannter Satz von Devroye gibt notwendige und hinreichende Bedingungen für die Konsistenz bezüglich d_1 von Kerndichteschätzern (vgl. Satz A.6 im Anhang).

Für das L_2-Risiko der Schätzung von f bei Verwendung eines Kerns K der Ordnung ℓ gilt mit $f_h = K_h * f$

$$\begin{aligned}\mathbb{E}(d_2^2(\hat{f}_h, f)) &= \mathbb{E}(d_2^2(\hat{f}_h, f_h)) + \mathbb{E}(d_2^2(f_h, f)) \\ &= \frac{1}{nh}\|K\|_2^2 + h^{2\ell}\frac{k_\ell^2}{\ell!^2}\|f^\ell\|_2^2 + o\left(\frac{1}{nh} + h^{2\ell}\right),\end{aligned}$$

vgl. Theorem 2.1.7 in [PR83]. Die Summe der beiden dominierenden Terme wird auch als AMISE *(Asymptotic Mean Integrated Square Error)* bezeichnet. Der Term $\frac{1}{nh}\|K\|_2^2$ stellt dabei die asymptotische Varianz dar, der Term $h^{2\ell}\frac{k_\ell^2}{\ell!}\|f^\ell\|_2^2$ die asymptotische quadrierte Verzerrung (*squared bias*). Die Varianz hängt asymptotisch nicht von f ab und wird geringer, je größer h ist, während der Bias-Term von f, aber nicht von n abhängt und kleiner wird für kleines h. Das Risko besteht also hier aus einer stochastischen und einer systematischen Komponente, die so auszubalancieren sind, dass das Risiko möglichst klein wird. Die asymptotisch optimale Bandbreite ergibt sich als

$$h_{L_2} = \left(\frac{\ell!^2\|K\|_2^2}{2\ell k_\ell^2\|f^{(\ell)}\|_2^2}\right)^{\frac{1}{2\ell+1}} n^{-\frac{1}{2\ell+1}}, \tag{3.14}$$

vgl. [WJ95], S.33. Ähnliche Überlegungen lassen sich auch für den L_1-Verlust anstellen, für den sich jedoch nur obere Schranken herleiten lassen. Für einen Kern der Ordnung $\ell = 2$ wird in der Literatur folgende asymptotische Wahl der Bandbreite vorgeschlagen:

$$h_{L_1} = \left(\frac{\|K\|_2^2\|\sqrt{f}\|_1^2}{2\pi k_2^2\|f^{(2)}\|_1^2}\right)^{\frac{1}{5}} n^{-\frac{1}{5}}, \tag{3.15}$$

vgl. Kapitel 5 in [DG85] bzw. [EL01], S. 298.

Die Bandbreiten (3.14) und (3.15) hängen von der gesuchten Dichte f über deren ℓ-te Ableitung ab, die natürlich in Anwendungen nicht bekannt ist. Eine Möglichkeit, h datengesteuert zu wählen, besteht darin, von f abhängige Konstanten in expliziten Ausdrücken wie (3.14) und (3.15) aus den Daten zu schätzen (sogenannte *Plug-In-Regeln*). Dies kann über eine Referenzdichte geschehen, meist eine $N(\mu, \sigma^2)$-Verteilung, wobei σ^2 dann aus den Daten geschätzt wird. Eine andere Möglichkeit ist es, einen weiteren nichtparametrischen Schätzer für die gesuchte Ableitung zu verwenden, für den sich allerdings wieder das Problem der Glättungsparameterwahl stellt. Es können auch in mehreren Schritten nacheinander höhere Ableitungen von f geschätzt werden, wobei im ersten Schritt meistens wiederum eine $N(\mu, \sigma^2)$-Verteilung als Referenzdichte verwendet wird. Eine Übersicht über Verfahren dieses Typs findet man z.B. in Kapitel 3 von [WJ95].

Eine andere wichtige Klasse von Methoden zur Bandbreitenwahl versucht, für verschiedene Werte von h das Risiko möglichst gut aus den Daten zu schätzen. Verwendet wird dann das h, das diese Risikoschätzung minimiert. Die populärste dieser Methoden ist die Leave-One-Out-L_2-Kreuzvaliderung, die auf der Zerlegung des L_2-Risikos

$$\mathbb{E}(d_2^2(\hat{f}_h, f)) = \mathbb{E} \int \hat{f}_h(x)^2 dx - 2\mathbb{E} \int \hat{f}_h(x) f(x) dx + \int f(x)^2 dx \tag{3.16}$$

basiert. Der dritte Term auf der rechten Seite hängt dabei nicht von h ab und kann daher bei der Minimierung ignoriert werden. Der erste Term hängt nur von der Schätzung ab, und kann einfach berechnet werden. Als Schätzung des zweiten Terms schlagen Rudemo [Rud82] und Bowman [Bow84] die Verwendung von

$$-\frac{2}{n} \sum_{i=1}^{n} \hat{f}_{h,-i}(X_i)$$

vor, wobei

$$\hat{f}_{h,-i}(x) = \frac{1}{n-1} \sum_{j \neq i} K_h(X_j - x).$$

Nun kann (3.16) (bis auf den dritten Term) für jedes h geschätzt werden. Gewählt wird das h, für das dieser Ausdruck minimal wird. Unter schwachen Voraussetzungen konvergieren die so gewählten Bandbreiten für $n \longrightarrow \infty$ gegen die optimale Bandbreite (3.23). Zur Verbesserung hinsichtlich der Variabilität und Konvergenzgeschwindigkeit existieren zahlreiche Modifikationen und Weiterentwicklungen der Leave-One-Out-L_2-Kreuzvaliderung, vgl. Kapitel 3 in [WJ95] für eine Diskussion und einen Überblick der wichtigsten Ansätze.

Beide Typen von Parameterwahlstrategien versuchen, das Risiko, d.h. einen Abstand zwischen der geschätzten Dichte $\hat{f}$ und der wahren Dichte f, zu minimieren. Die im Folgenden betrachteten Strategien basieren dagegen auf einem Abstand zwischen der empirischen und der geschätzten Verteilungsfunktion, also auf einem direkten Vergleich der Schätzung mit den Daten. Unter *Diskrepanzprinzipien zur Bandbreitenwahl in der Kerndichteschätzung* sollen Strategien vom folgenden Typ verstanden werden:

1. Wähle h so, dass
$$d(F_n, \hat{F}_n^h) = s(n), \tag{3.17}$$
oder

2. Wähle h maximal so, dass
$$d(F_n, \hat{F}_n^h) \leq s(n). \tag{3.18}$$

Dabei ist d ein geeigneter Abstand zwischen (evtl. zu signierten Maßen gehörigen) Verteilungsfunktionen und $s : \mathbb{N} \longrightarrow \mathbb{R}^+$ eine Funktion, die eine Schranke in Abhängigkeit von n liefert. Da die Konvergenz der gewählten Bandbreite gegen 0 für die Konsistenz des Kerndichteschätzers notwendig ist, kommen nur Schrankenfunktionen $s(n) = o(1)$ in Frage. Häufig besitzt (3.17) eine eindeutige Lösung. In diesem Fall sind die Strategien (3.17) und (3.18) äquivalent.

Im Folgenden wird der Abstand zwischen zwei Verteilungsfunktionen F und G meist gemessen durch die *Kolmogorov-Metrik*
$$d_\infty(F, G) := \|F - G\|_\infty$$
oder durch die *Kuiper-Metrik k-ter Ordnung*
$$d_{kuip,k}(F, G) := \sup_{a_1 \leq b_1 \leq a_2 \leq b_2 \leq \cdots \leq a_k \leq b_k} \sum_{i=1}^{k} |(F(b_i) - F(a_i)) - (G(b_i) - G(a_i))|.$$

Kuiper-Metriken erster Ordnung werden schon seit langem für Goodness-of-Fit-Tests benutzt (vgl. Kapitel 5 in [Dur73]), während die Kuiper-Metriken höherer Ordnung zuerst in [DK04] vorgeschlagen werden. Diese Metriken sind äquivalent und es ergeben sich folgende Abschätzungen:
$$\begin{aligned}
d_{kuip,k}(F, G) &= \sup \sum_{i=1}^{k} |(F(b_i) - F(a_i)) - (G(b_i) - G(a_i))| \\
&\leq k \sup |(F(b) - F(a)) - (G(b) - G(a))| \\
&= k d_{kuip,1}(F, G) \\
&\leq 2k \sup |F(a) - G(a)| \\
&= 2k d_\infty(F, G)
\end{aligned}$$
sowie
$$\begin{aligned}
d_{kuip,k}(F, G) &= \sup \sum_{i=1}^{k} |(F(b_i) - F(a_i)) - (G(b_i) - G(a_i))| \\
&\geq \sup |(F(b) - F(a)) - (G(b) - G(a))| \\
&= d_{kuip,1}(F, G) \\
&\geq \sup |F(a) - G(a)| \\
&= d_\infty(F, G).
\end{aligned}$$

Diese Abschätzungen ermöglichen eine weitgehend parallele Behandlung der Kolmogorov- und Kuiper-Metriken.

In den Simulationen wird weiterhin ein auf der Teststatistik des Cramér-von Mises-Tests basierender Abstand benutzt:

$$d_{CvM}(F,G) = \int_{-\infty}^{\infty} (F(t) - G(t))^2 dG(t). \tag{3.19}$$

Der Abstand d_{CvM} ist nicht symmetrisch und daher keine Metrik. Ist F stetige Verteilungsfunktion und F_n die empirische Verteilungsfunktion aus einer Stichprobe vom Umfang n aus der zugehörigen Verteilung, so hängen die Verteilungen von $d_\infty(F_n, F)$, $d_{kuip,k}(F_n, F)$ und $d_{CvM}(F_n, F)$ nicht von F ab. Außerdem sind die asymptotischen Verteilungen von $\sqrt{n} d_\infty(F_n, F)$, $\sqrt{n} d_{kuip,1}(F_n, F)$ und $n d_{CvM}(F_n, F)$ bekannt, vgl. Satz A.3. Daher eignen sich diese Abstände für die Konstruktion von Goodness-of-Fit-Tests und nichtparametrischen Konfidenzintervallen für die Verteilungsfunktion. Sie eignen sich daher auch für einen direkten Vergleich der Schätzung mit den Daten und damit zur Implementierung des Diskrepanzprinzips. Weiter sind $d_\infty(F_n, G)$, $d_{kuip,1}(F_n, G)$ und $d_{CvM}(F_n, G)$ für eine beliebige Verteilungsfunktion G eines Wahrscheinlichkeitsmaßes einfach auszuwerten, vgl. Satz A.3. Letzteres gilt jedoch nicht, wenn G die (dann nicht mehr unbedingt monotone) Verteilungsfunktion eines signierten Maßes ist. Diese Situation tritt bei Verwendung von Kernen der Ordnung $\ell > 2$ auf.

3.2. Analytische Resultate

Nun soll das Verhalten von Kernschätzern untersucht werden, deren Bandbreite mit Hilfe des Diskrepanzprinzips gewählt wurde. Sei dazu K ein Kern der Ordnung ℓ, $k_\ell := \int u^\ell K(u) du$ und $\tilde{k} = \|K\|_1$. Weiterhin sei $\mathbb{K}$ die zu K gehörige (nicht notwendig monotone) Verteilungsfunktion und

$$\kappa_0 := \sup_{x \in \mathbb{R}} |\mathbb{K}(x) - F_0(x)|,$$

wobei $F_0(x) := \mathbb{I}(x \geq 0)$ die Verteilungsfunktion zur Dirac-Verteilung in 0 bezeichne. Ist K eine Wahrscheinlichkeitsdichte, so ist $\kappa_0 = \max\{\mathbb{K}(0), 1 - \mathbb{K}(0)\}$. Ist K darüberhinaus symmetrisch, so gilt $\kappa_0 = \mathbb{K}(0) = 1/2$. Im Folgenden sei stets $d = d_\infty$ oder $d = d_{kuip,k}$ sowie

$$c_d = \begin{cases} 1, & d = d_\infty \\ 2k, & d = d_{kuip,k} \end{cases}.$$

Auf d_{CvM} basierende Versionen des Diskrepanzprinzips lassen sich mit den im Folgenden benutzten Techniken nicht ohne weiteres analysieren, da keine Dreiecksungleichung zur Verfügung steht. Sie werden jedoch in der Simulationsstudie in Kapitel 3.4 berücksichtigt.

Zunächst wird die Existenz einer Lösung von (3.17) gezeigt. Sei dazu $X_1, \ldots, X_n$ eine unabhängig identisch verteilte Stichprobe aus einer Verteilung auf $\mathbb{R}$ mit stetiger Verteilungsfunktion F. Die Existenz einer Dichte wird an dieser Stelle nicht benötigt, d.h. die

Verteilung darf einen singulären Anteil besitzen, jedoch keine Atome. Für festes n ist die Funktion $h \longrightarrow d(F_n, \hat{F}_n^h)$ ein Pfad eines stochastischen Prozesses mit Indexmenge $(0, \infty)$. Lösungen von (3.17) entsprechen genau denjenigen Werten von h, an denen der Pfad eine horizontale Gerade schneidet. Der folgende Satz zeigt, dass fast alle Pfade stetig sind und – unter schwachen Bedingungen an s – den Wert $s(n)$ für mindestens ein h annehmen müssen. Eine analoge Aussage wurde von Eggermont und LaRiccia [EL96, EL01] für den Spezialfall eines symmetrischen, nichtnegativen Kerns der Ordnung 2 und $d = d_\infty$ bewiesen.

Satz 3.1. *Sei F_n eine empirische Verteilungsfunktion und $\hat{F}_n^h$ wie in (3.10). Dann gilt fast sicher:*

1. *$d(F_n, \hat{F}_n^h)$ ist stetig in h.*
2. *$\lim_{h \to 0} d_\infty(F_n, \hat{F}_n^h) = \frac{\kappa_0}{n}$.*
3. *$\liminf_{h \to 0} d_{kuip,k}(F_n, \hat{F}_n^h) \leq 2k\frac{\kappa_0}{n}$.*
4. *$\lim_{h \to \infty} d_\infty(F_n, \hat{F}_n^h) = \kappa_0$.*
5. *$\limsup_{h \to \infty} d_{kuip,k}(F_n, \hat{F}_n^h) \geq \kappa_0$.*

Beweis. 1. Nach Lemma A.1 gilt:

$$\begin{aligned}
|d(F_n, \hat{F}_n^h) - d(F_n, \hat{F}_n^{h'})| &\leq d(\hat{F}_n^h, \hat{F}_n^{h'}) + d(F_n, F_n) \\
&\leq c_d d_\infty(\hat{F}_n^h, \hat{F}_n^{h'}) \\
&= c_d \|(K_h - K_{h'}) * F_n\|_\infty \\
&\leq c_d \|K_h - K_{h'}\|_1 \|F_n\|_\infty \\
&= c_d \|K_h - K_{h'}\|_1 \overset{h' \to h}{\longrightarrow} 0.
\end{aligned}$$

Der letzte Grenzübergang gilt nach Lemma A.2.
2. Wegen $K \in L_1(\mathbb{R})$ existiert für jedes $\varepsilon > 0$ ein M, so dass

$$\int_{-\infty}^{-M} |K(x)|dx + \int_M^\infty |K(x)|dx \leq \varepsilon.$$

Sei nun $\delta := \frac{1}{3}\min_{2 \leq j \leq n}(X_{(j)} - X_{(j-1)})$. Mit Wahrscheinlichkeit 1 ist $\delta > 0$. Für $h \leq \frac{\delta}{M}$ gilt nun

$$\int_{-\infty}^{-\delta} |K_h(x)|dx + \int_\delta^\infty |K_h(x)|dx \leq \varepsilon$$

und in jedem der (disjunkten) Intervalle $I_j := [X_{(j)} - \delta, X_{(j)} + \delta]$ gilt für alle $x \in I_j$

$$\left| \left| F_n(x) - \hat{F}_n^h(x) \right| - \left| \frac{1}{n} \left(\mathbb{I}(x \geq X_i) - K\left(\frac{x - X_i}{h} \right) \right) \right| \right| \leq \frac{n}{n}\varepsilon.$$

Für $x \in \mathbb{R} \setminus \bigcup I_j$ ist $|F_n(x) - \hat{F}_n^h(x)| \leq \varepsilon$. Insgesamt ist damit

$$\frac{\kappa_0}{n} - \varepsilon \leq d_\infty(F_n, \hat{F}_n^h) \leq \frac{\kappa_0}{n} + \varepsilon.$$

Es folgt $\lim_{h \to 0} d_\infty(F_n, \hat{F}_n^h) = \frac{\kappa_0}{n}$.
3. $\liminf_{h \to 0} d_{kuip,k}(F_n, \hat{F}_n^h) \leq 2k \liminf_{h \to 0} d_\infty(F_n, \hat{F}_n^h) \stackrel{2.}{=} 2k \frac{\kappa_0}{n}$.
4. Für $x \in [x_{(1)}, x_{(n)}]$ gilt $\lim_{h \to \infty} \hat{F}_n^h(x) = \mathbb{K}(0)$ und damit $\lim_{h \to \infty} d_\infty(F_n, \hat{F}_n^h) = \kappa_0$.
5. $\limsup_{h \to \infty} d_{kuip,k}(F_n, \hat{F}_n^h) \geq \limsup_{h \to \infty} d_\infty(F_n, \hat{F}_n^h) = \kappa_0$. □

Damit ist gezeigt, dass für Schrankenfunktionen $s(n)$ mit $s(n) = o(1)$ und $n^{-1} = o(s(n))$ zumindest für hinreichend großes n stets eine Lösung $h_{s,n}$ von

$$d(F_n, \hat{F}_n^h) = s(n)$$

existiert. Diese Bedingungen werden von den in der Literatur bislang vorgeschlagenen Wahlen von $s(n)$ erfüllt.

In Abbildung 3.1 sind jeweils zwei Realisationen für $n = 10$ und $n = 100$ zu sehen. Die Stichproben stammen aus einer Standardnormalverteilung, und es wird der Gaußkern verwendet. Eingezeichnet sind jeweils zwei horizontale Linien, die den Wahlen $s(n) = 0.6n^{-1/2}$ (vgl. Kapitel 7.9 in [Vap98]) bzw. $s(n) = 0.35n^{-2/5}$ (vgl. [EL96]) entsprechen. Es ist auch zu sehen, dass die Funktion $h \longrightarrow d(F_n, \hat{F}_n^h)$ nicht immer monoton sein muss. Die Eindeutigkeit von h kann also nicht garantiert werden. Bei größerem n existieren jedoch selten mehrere Lösungen.

Nun soll gezeigt werden, dass für $s(n) = o(1)$ der Glättungsparameter $h_{s,n}$ stets gegen 0 konvergiert und dass $\hat{F}_n^{h_{s,n}}$ stark konsistent für F ist. Dabei wird wieder nur die Stetigkeit von F vorausgesetzt, nicht die Existenz einer Dichte. Im Folgenden wird häufig die Funktion

$$F_h := F * K_h$$

benötigt. Ist K_h eine Wahrscheinlichkeitsdichte, so ist F_h Verteilungsfunktion zu einem Wahrscheinlichkeitsmaß, ansonsten Verteilungsfunktion zu einem signierten Maß (vgl. Definition A.1). Wichtig sind außerdem die folgenden Abschätzungen:

Satz 3.2. *Es gilt fast sicher*

1. $d(F_n, F) = O\left(\sqrt{\frac{\log\log n}{n}}\right)$ *und*

2. $d(\hat{F}_n^h, F_h) \leq c_d \tilde{k} d_\infty(F_n, F) = O\left(\sqrt{\frac{\log\log n}{n}}\right)$ *gleichmäßig in h.*

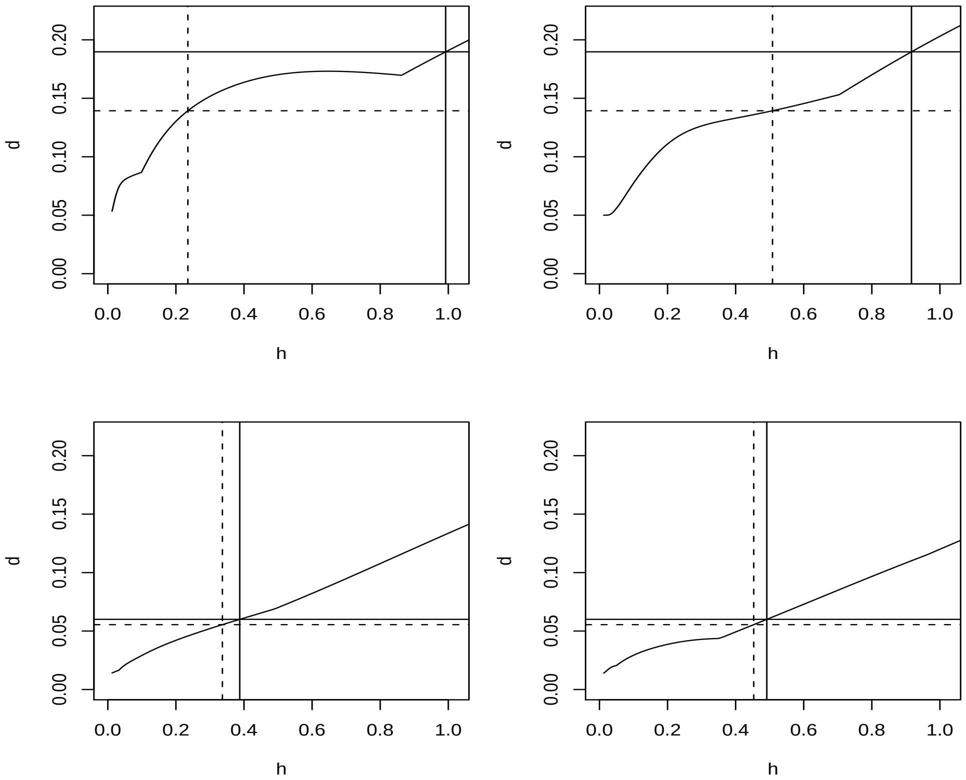

Abbildung 3.1.: Lösungen von $d_\infty(\hat{F}_n^h, F_n) = s(n)$ bei Verwendung von $K^{\mathcal{N}}$ für $X_1, \ldots, X_n \sim N(0,1)$. Oben: $n = 10$, unten: $n = 100$. Durchgezogene Linien: $s(n) = 0.6n^{-1/2}$, gestrichelte Linien: $s(n) = 0.35n^{-2/5}$.

Beweis. 1. $d(F_n, F) \leq c_d d_\infty(F_n, F) = O\left(\sqrt{\frac{\log\log n}{n}}\right)$ fast sicher nach Satz A.4.

2. Es gilt fast sicher:

$$\begin{aligned} d(\hat{F}_n^h, F_h) &\leq c_d d_\infty(\hat{F}_n^h, F_h) \\ &= c_d \|K_h * (F - F_n)\|_\infty \\ &= c_d \|K\|_1 \|F - F_n\|_\infty \\ &= c_d \tilde{k} \|F - F_n\|_\infty \\ &= O\left(\sqrt{\frac{\log\log n}{n}}\right). \end{aligned}$$

□

Im folgenden Satz wird gezeigt, dass die durch das Diskrepanzprinzip gewählten Bandbreiten gegen 0 konvergieren und zumindest die so erhaltene Schätzung der Verteilungsfunktion konsistent ist für F. Der Beweis des zweiten Teils von Satz 3.3 basiert auf der selben Idee wie der Beweis von Satz 3 in [Yam73]. Da hier jedoch Kerne zugelassen sind, die auch negative Werte annehmen können, wird ein zusätzliches Straffheitsargument benötigt.

Satz 3.3. *Es sei F eine stetige Verteilungsfunktion, F_n und $\hat{F}_n^h$ wie oben und $s(n) = o(1)$. Der Glättungsparameter $h_{s,n}$ werde als Lösung von*

$$d(F_n, \hat{F}_n^h) = s(n)$$

gewählt. Dann gilt:

1. $d(F, \hat{F}_n^{h_{s,n}}) \xrightarrow{f.s.} 0$ *sowie*

2. $h_{s,n} \xrightarrow{f.s.} 0$.

Beweis. 1. Es gilt fast sicher

$$\begin{aligned} d(F, F_{h_{s,n}}) &\leq d(F, F_n) + d(F_n, \hat{F}_n^{h_{s,n}}) + d(\hat{F}_n^{h_{s,n}}, F_{h_{s,n}}) \\ &= O\left(\sqrt{\frac{\log\log n}{n}}\right) + s(n) + O\left(\sqrt{\frac{\log\log n}{n}}\right) \\ &= o(1), \end{aligned}$$

und damit

$$d(F, \hat{F}_n^{h_{s,n}}) \leq d(F, F_{h_{s,n}}) + d(F_{h_{s,n}}, \hat{F}_n^{h_{s,n}}) = o(1).$$

2. Nach Teil 1 gilt fast sicher $d_\infty(F, F_{h_{s,n}}) \leq d(F, F_{h_{s,n}}) \longrightarrow 0$; zu zeigen ist noch, dass dies $h_{s,n} \xrightarrow{f.s.} 0$ impliziert. Sei dazu $h_n := h_{s,n}$ die Folge der gewählten Bandbreiten, P das zu F gehörige Wahrscheinlichkeitsmaß und μ_{h_n} das (signierte) Maß mit Lebesgue-Dichte K_{h_n}. Um Satz A.5 verwenden zu können, muss für jedes $\varepsilon > 0$ ein $M \in \mathbb{R}$ existieren, so dass für alle $n \in \mathbb{N}$

$$|P * \mu_{h_n}|(\mathbb{R} \setminus [-M, M]) \leq \varepsilon \tag{3.20}$$

gilt. Sei nun (3.20) nicht erfüllt, d.h. es existiere ein $\varepsilon > 0$, so dass für alle M

$$\sup_n |P * \mu_{h_n}|(\mathbb{R} \setminus [-M, M]) > \varepsilon. \tag{3.21}$$

Angenommen, $h^* := \sup_n h_n < \infty$. Dann existiert wegen $|K| \in L_1(\mathbb{R})$ zu jedem $\eta > 0$ ein m_η, so dass

$$|\mu_1|([-m_\eta, m_\eta]) \geq k - \eta$$

und damit

$$|\mu_{h_n}|([-h^* m_\eta, h^* m_\eta]) \geq k - \eta$$

für alle n. Da P ein Wahrscheinlichkeitsmaß ist, existiert ebenfalls ein q_η, so dass

$$P([-q_\eta, q_\eta]) \geq 1 - \eta$$

ist. Wird nun $\eta = \frac{1}{2}(k + 1 - \sqrt{k^2 + 2k - 4\varepsilon + 1})$ gewählt, so gilt mit

$$I := [-(h^* m_\eta + q_\eta), h^* m_\eta + q_\eta]$$

$$\begin{aligned}
|P * \mu_{h_n}|(I) &= P * |\mu_{h_n}|(I) \\
&= (P\,|_{[-q_\eta, q_\eta]} + P\,|_{[-q_\eta, q_\eta]^C}) * (|\mu_{h_n}|\,|_{[-h^* m_\eta, h^* m_\eta]} + |\mu_{h_n}|\,|_{[-h^* m_\eta, h^* m_\eta]^C})(I) \\
&\geq (P\,|_{[-q_\eta, q_\eta]} * |\mu_{h_n}|\,|_{[-h^* m_\eta, h^* m_\eta]})(I) \\
&\geq (1-\eta)(k-\eta) \\
&= 1 - \varepsilon
\end{aligned}$$

im Widerspruch zu (3.21). Damit ist klar, dass die Folge $(h_n)_n$ unbeschränkt sein muss, d.h. es existiert eine Teilfolge $(h_{n_m})_m$ mit $h_{n_m} \longrightarrow \infty$. Dies steht aber im Widerspruch zu

$$d_\infty(F, F_{h_n}) \longrightarrow 0.$$

Damit ist gezeigt, dass (3.20) erfüllt ist.
Seien nun $\hat{P}$, $\hat{K}$ und $\hat{K}_{h_n}$ die Fouriertransformierten von P, K und K_{h_n}. Mit Satz A.5 folgt, dass $\hat{P}\hat{K}_{h_n}(t) \longrightarrow \hat{P}(t)$ für alle $t \in \mathbb{R}$ gilt. Da wegen der Stetigkeit der Fouriertransformierten $\hat{P} > 0$ auf einem Intervall $[-\varepsilon, \varepsilon]$ gilt, muss $\hat{K}_{h_n}(t) = \hat{K}(h_n t) \longrightarrow 1$ gelten für alle $t \in [-\varepsilon, \varepsilon]$. Da aber $\hat{K}$ nach Satz A.1 wegen $\int u^\ell K(u) du \neq 0$ auf keinem Intervall um 0 identisch 1 sein kann, erzwingt dies $h_n \longrightarrow 0$. □

Damit ist gezeigt, dass durch das Diskrepanzprinzip gewählte Bandbreiten $h_{s,n}$ stets gegen 0 konvergieren und der Kernschätzer der Verteilungsfunktion konsistent für die Verteilungsfunktion ist. Um Aussagen über Konsistenz und mögliche Optimalität des Kerndichteschätzers $\hat{f}_{s,n}$ zu gewinnen, sind zusätzliche Annahmen an f, K und s notwendig. Insbesondere wird im Weiteren die Existenz der Dichte f vorausgesetzt. Zunächst wird gezeigt, dass Kernschätzer mit durch das Diskrepanzprinzip gewählter Bandbreite $h_{s,n}$ inkonsistent sein können. Für einige Dichten mit Polstellen konvergiert die Bandbreite zu schnell gegen 0:

Satz 3.4. *Sei $K = K^{\mathcal{U}}$ oder $K = K^{\mathcal{E}}$ und $0 < \varepsilon < 1/2$ so, dass $n^\varepsilon s(n) = o(1)$ ist. Sei weiter*

$$f(x) := \begin{cases} \varepsilon x^{-(1-\varepsilon)} & 0 < x \leq 1 \\ 0 & \textit{sonst} \end{cases}$$

und F_n die empirische Verteilungsfunktion einer unabhängig identisch verteilten Stichprobe aus einer Verteilung mit Dichte f. Ist $h_{s,n}$ nun die Lösung von

$$d(F_n, \hat{F}_n^h) = s(n),$$

so gilt:

1. $nh_{s,n} \longrightarrow 0$ *fast sicher und*

2. $\liminf_{n\to\infty} d_1(\hat{f}_{h_{s,n}}, f) \geq 1$ *fast sicher.*

Beweis. 1. Die zu f gehörige Verteilungsfunktion ist

$$F(x) := \begin{cases} 0 & x \leq 0 \\ x^\varepsilon & 0 < x \leq 1 \\ 1 & x > 1 \end{cases},$$

und wegen

$$(F * K_h)(t) = \int F(x) K_h(t-x) dx$$

ist für $K = K^{\mathcal{U}}$

$$(F * K_h^{\mathcal{U}})(h) = \frac{1}{2h} \int_0^{2h} x^\varepsilon dx = \frac{2^\varepsilon}{1+\varepsilon} h^\varepsilon.$$

Damit gilt:

$$|F(h) - (F * K_h^{\mathcal{U}})(h)| = \underbrace{\left(1 - \frac{2^\varepsilon}{1+\varepsilon}\right)}_{=:c_{\mathcal{U}}>0} h^\varepsilon.$$

Für $K = K^{\mathcal{E}}$ ist

$$(F * K_h^{\mathcal{E}})(h) = \frac{3}{4h} \int_0^{2h} \left(1 - \left(\frac{x-h}{h}\right)^2\right) x^\varepsilon dx = \frac{3 \cdot 2^{(\varepsilon+1)}}{\varepsilon^2 + 5\varepsilon + 6} h^\varepsilon,$$

und damit

$$|F(h) - (F * K_h^{\mathcal{E}})(h)| = \underbrace{\left(1 - \frac{3 \cdot 2^{(\varepsilon+1)}}{\varepsilon^2 + 5\varepsilon + 6}\right)}_{=:c_{\mathcal{E}}>0} h^\varepsilon.$$

Für $K = K^{\mathcal{U}}$ sei nun $c = c_{\mathcal{U}}$ und für $K = K^{\mathcal{E}}$ sei $c = c_{\mathcal{E}}$. Es folgt fast sicher

$$\begin{aligned} cn^\varepsilon h_{s,n}^\varepsilon &\leq n^\varepsilon d_\infty(F, F_{h_{s,n}}) \\ &\leq n^\varepsilon \left(d_\infty(F, F_n) + d(F_n, \hat{F}_n^{h_{s,n}}) + d_\infty(F_h, \hat{F}_n^{h_{s,n}}) \right) \\ &= n^\varepsilon O(n^{-1/2} \sqrt{\log \log n}) + n^\varepsilon s(n) \\ &= o(1), \end{aligned}$$

und damit $nh_{s,n} = o(1)$.

2. Sei λ das Lebesgue-Maß auf $\mathbb{R}$. Es ist fast sicher

$$\begin{aligned}\lambda\{\hat{f}_{h_{s,n}} > 0\} &= \lambda\{\sum_{i=1}^{n} K_{h_{s,n}}(x - X_i) > 0\}\\ &= \lambda\{\sum_{i=1}^{n} \mathbb{I}(|x - X_i| \leq h_{s,n}) > 0\}\\ &\leq \sum_{i=1}^{n} \lambda\{\mathbb{I}(|x - X_i| \leq h_{s,n}) > 0\}\\ &\leq 2nh_{s,n} = o(1)\end{aligned}$$

nach Teil 1 des Satzes. Damit folgt weiter fast sicher

$$\begin{aligned}\liminf_{n\to\infty} d_1(\hat{f}_{h_{s,n}}, f) &= \liminf_{n\to\infty} \int |\hat{f}_{h_{s,n}}(x) - f(x)|dx\\ &\geq \liminf_{n\to\infty} \int_{[0,1]\cap\{\hat{f}_{h_{s,n}}=0\}} |\hat{f}_{h_{s,n}}(x) - f(x)|dx\\ &\geq \liminf_{n\to\infty} \int_{[0,1]\cap\{\hat{f}_{h_{s,n}}=0\}} f(x)dx\\ &\geq \liminf_{n\to\infty} \int_{2nh_{s,n}}^{1} f(x)dx\\ &= 1,\end{aligned}$$

wobei im vorletzten Schritt die Monotonie von f benutzt wird. □

In Satz 3.4 wird die Inkonsistenz nur für den Rechteckskern $K^{\mathcal{U}}$ und den Epanechnikov-Kern $K^{\mathcal{E}}$ bewiesen, da sich für diese eine einfache untere Abschätzung von $d_\infty(F, F_h)$ herleiten lässt. Der Beweis des zweiten Teils des Satzes zeigt insbesondere, dass für diese beiden Kerne mit kompaktem Träger das Lebesgue-Maß des Trägers von $\hat{f}_{h_{s,n}}$ gegen 0 konvergiert, d.h. der Dichteschätzer ist immer stärker um die einzelnen Beobachtungen konzentriert und verschwindet auf immer größeren Teilen des Trägers von f. Abbildung 3.2 zeigt, dass derselbe Effekt auch bei Verwendung anderer Kerne auftritt. Zu sehen sind Kerndichteschätzer unter Verwendung des Gaußkerns für vier Stichproben der Größe $n = 100$ aus der in Satz 3.4 gegeben Verteilung mit $\varepsilon = 0.32$. Die Bandbreite wurde jeweils als Lösung von $d_\infty(\hat{F}_n^h, F_n) = 0.6n^{-1/2}$ gewählt. Es ist deutlich zu sehen, dass die gewählten Bandbreiten zu klein sind, insbesondere im Bild links unten. Es zeigt sich auch in der in Kapitel 3.4 beschriebenen Simulationsstudie, dass das Diskrepanzprinzip für Dichten mit unendlichen Peaks nicht gut funktioniert, wobei allerdings die herangezogenen Vergleichsmethoden ähnliche Probleme haben. Ist f wie in Satz 3.4 mit $\varepsilon < \frac{1}{4}$, so ist $\lim_{x\to 0} xf(x) = \infty$. In [Dev89] wurde gezeigt, dass ein Dichteschätzer mit einem symmetrischen unimodalen beschränkten Kern und mittels L_2-Kreuzvalidierung gewählter Bandbreite in diesem Fall inkonsistent ist. Die kreuzvalidierte Bandbreite konvergiert dann ebenfalls zu schnell gegen 0.

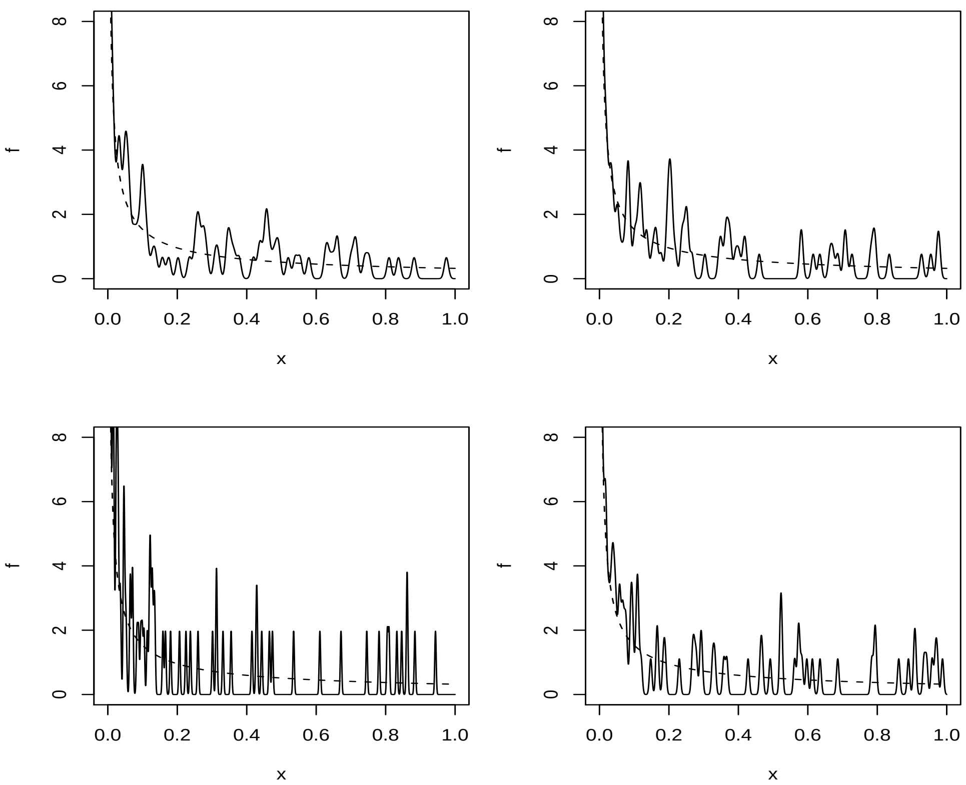

Abbildung 3.2.: Vier Realisationen von $\hat{f}_h$ bei Verwendung des Gaußkerns mit h Lösung von $d_\infty(\hat{F}_n^h, F_n) = 0.6n^{-1/2}$. Die Stichproben stammen aus der Dichte aus Satz 3.4 mit $\varepsilon = 0.32$ und $n = 100$.

Auch in Fällen, in denen $\hat{f}_{h_{s,n}}$ konsistent ist, interessiert die Geschwindigkeit, mit der $h_{s,n}$ gegen 0 konvergiert. Dazu werden Aussagen über das Verhalten von $F - F_h$ benötigt. Dazu wird vorausgesetzt, dass f in einem Sobolev-Raum

$$W^{\ell,p} := \{f : f, f^{(1)}, \dots, f^{(\ell)} \in L_p(\mathbb{R})\}$$

liegt. Der folgende Satz wurde von Eggermont und LaRiccia für den Fall eines nichtnegativen symmetrischen Kerns der Ordnung $\ell = 2$ und $d = d_\infty$ bewiesen. Der hier gegebene Beweis für den allgemeineren Fall lehnt sich eng an den Originalbeweis [EL96, EL01] an.

Satz 3.5. *Sei $f \in W^{\ell,1}(\mathbb{R})$ und K ein Kern der Ordnung ℓ. Dann gilt:*

1. $F_h(x) - F(x) = \frac{(-1)^p}{\ell!} k_\ell f^{(\ell-1)}(x) h^\ell + o(h^\ell)$ *gleichmäßig in* $x \in \mathbb{R}$.

2. $d(F_h, F) = \frac{1}{\ell!} k_\ell d(f^{(\ell-1)}, 0) h^\ell + o(h^\ell)$.

Beweis. 1. Wegen $f \in W^{\ell,1}(\mathbb{R})$ ist

$$\begin{aligned} F(y) - F(x) = (y-x)f(x) + \frac{1}{2}(y-x)^2 f'(x) + \\ \cdots + \frac{(y-x)^\ell}{\ell!} f^{(\ell-1)}(x) + \int_x^y \frac{(y-t)^\ell}{\ell!} f^{(\ell)}(t)dt \end{aligned}$$

und damit, da K Kern der Ordnung ℓ ist,

$$\begin{aligned} F_h(x) - F(x) &= \int K_h(x-y)(F(y)-F(x))dy \\ &= \int K_h(x-y)\left(\frac{1}{\ell!}(y-x)^\ell f^{(\ell-1)}(x) + \int_x^y \frac{(y-t)^\ell}{\ell!} f^{(\ell)}(t)dt\right) dy \\ &= \frac{(-1)^p}{\ell!} k_\ell f^{(\ell-1)}(x) h^\ell + \frac{1}{\ell!} \int K_h(x-y) \int_x^y (y-t)^\ell f^{(\ell)}(t)dtdy. \end{aligned}$$

Es bleibt zu zeigen, dass

$$\left| \frac{1}{\ell!} \int K_h(x-y) \int_x^y (y-t)^\ell f^{(\ell)}(t)dtdy \right| = o(h^\ell)$$

gleichmäßig in x ist. Es gilt:

$$\begin{aligned} \left| \frac{1}{\ell!} \int K_h(x-y) \int_x^y (y-t)^\ell f^{(\ell)}(t)dtdy \right| &\leq \frac{1}{\ell!} \int |K_h(x-y)| \int_x^y |y-t|^\ell |f^{(\ell)}(t)|dtdy \\ &\leq \frac{1}{\ell!} \int |x-y|^\ell |K_h(x-y)| \int_x^y |f^{(\ell)}(t)|dtdy \\ &= \frac{1}{\ell!} \left(R_1(x) + R_2(x)\right) \end{aligned}$$

mit

$$\begin{aligned} R_1(x) &:= \int_{|x-y|>h^{1/\ell}} |x-y|^\ell |K_h(x-y)| \int_x^y |f^{(\ell)}(t)|dtdy \\ R_2(x) &:= \int_{|x-y|<h^{1/\ell}} |x-y|^\ell |K_h(x-y)| \int_x^y |f^{(\ell)}(t)|dtdy. \end{aligned}$$

Es wird zunächst R_1 betrachtet:

$$\begin{aligned} R_1(x) &= \int_{|x-y|>h^{1/\ell}} |x-y|^\ell |K_h(x-y)| \int_x^y |f^{(\ell)}(t)|dtdy \\ &\leq \|f^{(\ell)}\|_1 \int_{|x-y|>h^{1/\ell}} |x-y|^\ell |K_h(x-y)|dy \\ &= \|f^{(\ell)}\|_1 \int_{|z|>h^{1/\ell-1}} h^\ell |z|^\ell |K(z)|dy = o(h^\ell), \end{aligned}$$

da $z^\ell K(z) \in L_1(\mathbb{R})$. Es sei

$$\omega(f;h) = \sup_{|x-y|<h^{1/\ell}} \int_y^x |f^{(\ell)}(z)|dz.$$

Wegen $f^{(\ell)} \in L_1(\mathbb{R})$ ist dann die Funktion $g(t) := \int_{-\infty}^t |f^{(\ell)}(x)|dx$ absolut stetig und damit gleichmäßig stetig, d.h. $\lim_{h\to 0} \omega(f;h) = 0$. Damit folgt:

$$\begin{aligned}
R_2(x) &\leq \omega(f;h) \int_{|x-y|<h^{1/\ell}} |x-y|^\ell |K_h(x-y)|dy \\
&\leq \omega(f;h) \int |x-y|^\ell |K_h(x-y)|dy \\
&= \omega(f;h) h^\ell \int |z|^\ell |K(z)|dy = o(h^\ell)
\end{aligned}$$

gleichmäßig in x und damit $(R_1(x) + R_2(x)) = o(h^\ell)$ gleichmäßig in x.
2. Sei zunächst $d = d_\infty$. Dann ist

$$\begin{aligned}
d(F_h, f) &= \sup_x |F_h(x) - F(x)| \\
&= \sup_x \left| \frac{(-1)^p}{\ell!} k_\ell f^{(\ell-1)}(x) h^\ell + o(h^\ell) \right| \\
&= \frac{1}{\ell!} k_\ell \sup_x |f^{(\ell-1)}(x)| h^\ell + o(h^\ell) \\
&= \frac{1}{\ell!} k_\ell d(f^{(\ell-1)}, 0) h^\ell + o(h^\ell)
\end{aligned}$$

und analog für $d = d_{kuip,k}$

$$\begin{aligned}
d(F_h, f) &= \sup_{a_1 \leq b_1 \leq a_2 \leq b_2 \leq \cdots \leq a_k \leq b_k} \sum_{i=1}^k |(F(b_i) - F_h(b_i)) + (F_h(a_i) - F(a_i))| \\
&= \sup_{a_1 \leq b_1 \leq a_2 \leq b_2 \leq \cdots \leq a_k \leq b_k} \frac{1}{\ell!} k_\ell \sum_{i=1}^k |f^{(\ell-1)}(b_i) - f^{(\ell-1)}(a_i)| h^\ell + o(h^\ell) \\
&= \frac{1}{\ell!} k_\ell d(f^{(\ell-1)}, 0) h^\ell + o(h^\ell).
\end{aligned}$$

□

Die in Satz 3.5 gegebenen Approximationen basieren auf Taylorentwicklungen und gelten daher nur für hinreichend kleines h. Wird die Bandbreite h jedoch durch das Diskrepanzprinzip gewählt, so impliziert Satz 3.3 für $n \to \infty$ fast sicher $h \to 0$. Damit sind Terme der Ordnung $o(1)$ für $h \to 0$ in diesem Fall auch für $n \to \infty$ von der Ordnung $o(1)$. Dies wird im Folgenden häufig benutzt, um konkretere Aussagen über das Verhalten der gewählten

Bandbreite treffen zu können; aus der Literatur bekannte Sätze scheinen an dieser Stelle eine Lücke aufzuweisen, vgl. z.B. [EL96].

Die folgenden Sätze 3.6-3.8 behandeln verschiedene Fälle, in denen $s(n)$ im Vergleich mit $d(F_n, F)$ schneller, genauso schnell oder langsamer gegen 0 konvergiert.

Satz 3.6. *Sei $f \in W^{\ell,1}$, K Kern der Ordnung ℓ und $s(n) = O\left(\sqrt{\frac{\log\log n}{n}}\right)$. Dann gilt fast sicher*

$$h_{s,n} = O(n^{-\frac{1}{2\ell}}(\log\log n)^{\frac{1}{2\ell}}).$$

Beweis. Nach Satz 3.5 gilt fast sicher

$$\begin{aligned}
\frac{1}{\ell!}k_\ell\|f^{(\ell-1)}\|_\infty {h_{s,n}}^\ell + o({h_{s,n}}^\ell) &= d_\infty(F_{h_{s,n}}, F) \\
&\leq d_\infty(F, F_n) + d(F_n, \hat{F}_n^{h_{s,n}}) + d_\infty(\hat{F}_n^{h_{s,n}}, F_{h_{s,n}}) \\
&= O\left(\sqrt{\frac{\log\log n}{n}}\right) + s(n) \\
&= O\left(\sqrt{\frac{\log\log n}{n}}\right).
\end{aligned}$$

Weiter ist

$$\frac{1}{\ell!}k_\ell\|f^{(\ell-1)}\|_\infty {h_{s,n}}^\ell + o({h_{s,n}}^\ell) = \frac{1}{\ell!}k_\ell\|f^{(\ell-1)}\|_\infty {h_{s,n}}^\ell(1 + o(1)),$$

wobei der zweite Term in der Klammer nicht nur von der Ordnung $o(1)$ für $h_{s,n} \to 0$, sondern auch von der Ordnung $o(1)$ für $n \to \infty$ ist, da nach Satz 3.3 $n \to \infty$ fast sicher $h_{s,n} \to 0$ impliziert. Es ist also

$$\begin{aligned}
&& \frac{1}{\ell!}k_\ell\|f^{(\ell-1)}\|_\infty {h_{s,n}}^\ell(1 + o(1)) &= O\left(\sqrt{\frac{\log\log n}{n}}\right) \\
&\Rightarrow & {h_{s,n}}^\ell &= O\left(\sqrt{\frac{\log\log n}{n}}\right)(1 + o(1))^{-1} \\
&\Rightarrow & {h_{s,n}}^\ell &= O\left(\sqrt{\frac{\log\log n}{n}}\right)(1 + o(1)) \\
&\Rightarrow & {h_{s,n}}^\ell &= O\left(\sqrt{\frac{\log\log n}{n}}\right) \\
&\Rightarrow & h_{s,n} &= O(n^{-\frac{1}{2\ell}}(\log\log n)^{\frac{1}{2\ell}}).
\end{aligned}$$

□

Satz 3.6 schließt Schrankenfunktionen der Form $s(n) = c\sqrt{\frac{\log\log n}{2n}}$ ein. In diesem Fall sind $s(n)$ und $d(F_n, F)$ von der selben Größenordnung. Wird die Konstante hinreichend groß gewählt, sind jedoch präzisere Resultate möglich. Aidu und Vapnik behandeln die Schrankenfunktion $s(n) = (1 + \tilde{k} + \varepsilon)\sqrt{\frac{\log\log n}{2n}}$ für Kerne der Ordnung ℓ und $d = d_\infty$. Der Beweis des folgenden Satzes folgt dem Originalbeweis in [AV89], es wird hier jedoch zusätzlich der Fall $d = d_{kuip,k}$ mit berücksichtigt.

Satz 3.7. *Sei $f \in W^{\ell,1}$, K Kern der Ordnung ℓ und $s(n) = c_d(\tilde{k} + 1 + \varepsilon)\sqrt{\frac{\log\log n}{2n}}$. Dann gilt fast sicher*

$$h_{s,n} \asymp \left(\frac{\log\log n}{n}\right)^{\frac{1}{2\ell}}.$$

Beweis. Es gilt nach Satz A.4 für $\eta > 0$ fast sicher für hinreichend großes n

$$\begin{aligned} c_d(\tilde{k} + 1 + \varepsilon)\sqrt{\frac{\log\log n}{2n}} &= d(F_n, \hat{F}_n^{h_{s,n}}) \\ &\leq d(F_n, F) + d(F, F_{h_{s,n}}) + d(F_{h_{s,n}}, \hat{F}_n^{h_{s,n}}) \\ &\leq c_d(\tilde{k} + 1)\sqrt{\frac{\log\log n}{2n}} + \eta + d(F, F_{h_{s,n}}) \end{aligned}$$

und damit

$$d(F, F_{h_{s,n}}) \geq c_d\varepsilon\sqrt{\frac{\log\log n}{2n}} - \eta.$$

Ebenfalls gilt fast sicher für hinreichend großes n

$$\begin{aligned} d(F, F_{h_{s,n}}) &\leq d(F, F_n) + d(F_n, \hat{F}_n^{h_{s,n}}) + d(\hat{F}_n^{h_{s,n}}, F_{h_{s,n}}) \\ &\leq c_d(2\tilde{k} + 2 + \varepsilon)\sqrt{\frac{\log\log n}{2n}} + \eta \end{aligned}$$

und damit nach Satz 3.5:

$$c_d\varepsilon\sqrt{\frac{\log\log n}{2n}} - \eta \leq \frac{k_\ell}{\ell!} d(f^{(\ell-1)}, 0){h_{s,n}}^{\ell}(1 + o(1)) \leq c_d(2\tilde{k} + 2 + \varepsilon)\sqrt{\frac{\log\log n}{2n}} + \eta$$

und damit $h_{s,n} \asymp \left(\frac{\log\log n}{n}\right)^{\frac{1}{2\ell}}$. □

Wesentlich präzisere Aussagen lassen sich für den Fall treffen, dass $s(n)$ langsamer gegen 0 konvergiert als $d(F_n, F)$. In diesem Fall kann das asymptotische Verhalten der Bandbreite exakt bestimmt werden. Der folgende Satz wurde von Eggermont und LaRiccia für symmetrische Kerne der Ordnung 2 und $d = d_\infty$ bewiesen. Der hier vorgestellte Beweis für den allgemeinen Fall lehnt sich eng an den Originalbeweis [EL96, EL01] an.

Satz 3.8. *Sei $f \in W^{\ell,1}$, K Kern der Ordnung ℓ und $s(n) = cn^{-\gamma}$ für $c > 0$ und $0 < \gamma < 1/2$. Dann gilt fast sicher:*

$$h_{s,n} = \left(\frac{c\ell!}{k_\ell d(f^{(\ell-1)}, 0)} \right)^{\frac{1}{\ell}} n^{-\frac{\gamma}{\ell}} \left(1 + o(1)\right). \tag{3.22}$$

Beweis. Nach der Vierecksungleichung A.1 gilt:

$$|d(F_n, \hat{F}_n^{h_{s,n}}) - d(F, F_{h_{s,n}})| \leq d(F_{h_{s,n}}, \hat{F}_n^{h_{s,n}}) + d(F, F_n) = O\left(\sqrt{\frac{\log\log n}{n}} \right).$$

Damit gilt:

$$\begin{aligned}
& \frac{1}{\ell!} k_\ell d(f^{(\ell-1)}, 0) {h_{s,n}}^\ell + o({h_{s,n}}^\ell) = cn^{-\gamma} + O\left(\sqrt{\frac{\log\log n}{n}} \right) \\
\Rightarrow \quad & \frac{1}{\ell!} k_\ell d(f^{(\ell-1)}, 0) {h_{s,n}}^\ell (1 + o(1)) = cn^{-\gamma} + O\left(\sqrt{\frac{\log\log n}{n}} \right) \\
\Rightarrow \quad & \frac{1}{\ell!} k_\ell d(f^{(\ell-1)}, 0) {h_{s,n}}^\ell = cn^{-\gamma} + o(n^{-\gamma}) + O\left(\sqrt{\frac{\log\log n}{n}} \right) + o\left(\sqrt{\frac{\log\log n}{n}} \right) \\
\Rightarrow \quad & {h_{s,n}}^\ell = \frac{c\ell!}{k_\ell d(f^{(\ell-1)}, 0)} n^{-\gamma} \left(1 + o(1)\right) \\
\Rightarrow \quad & h_{s,n} = \left(\frac{c\ell!}{k_\ell d(f^{(\ell-1)}, 0)} \right)^{\frac{1}{\ell}} n^{-\frac{\gamma}{\ell}} \left(1 + o(1)\right)^{\frac{1}{\ell}} \\
\Rightarrow \quad & h_{s,n} = \left(\frac{c\ell!}{k_\ell d(f^{(\ell-1)}, 0)} \right)^{\frac{1}{\ell}} n^{-\frac{\gamma}{\ell}} \left(1 + o(1)\right).
\end{aligned}$$

□

3.3. Varianten des Diskrepanzprinzips

Die Sätze 3.6-3.8 umfassen fast alle der bisher in der Literatur vorgeschlagenen Varianten des Diskrepanzprinzips für Kerndichteschätzer. Viele Fragen zum asymptotischen Verhalten der Methoden können dabei mit Hilfe dieser Resultate beantwortet werden. Satz 3.8 ermöglicht aber auch die Konstruktion neuer Methoden, die das Diskrepanzprinzip mit auf Referenzdichten basierenden Ansätzen verbinden. Weitere Aufschlüsse über das – zum Teil konträre – Verhalten der Methoden für kleine bis mittlere Stichprobenumfänge liefert die umfangreiche Simulationstudie, deren Ergebnisse im nächsten Abschnitt vorgestellt werden.

Die möglichen Schrankenfunktionen können nach der Rate, mit der sie für $n \longrightarrow \infty$ gegen 0 konvergieren, in drei Klassen eingeteilt werden:

1. $s(n) = o\left(\sqrt{\frac{\log\log n}{n}}\right)$
2. $s(n) \asymp \sqrt{\frac{\log\log n}{n}}$
3. $\sqrt{\frac{\log\log n}{n}} = o(s(n))$.

Vetreter aller drei Klassen sind in der Literatur vorgeschlagen worden. Diese sollen nun einzeln diskutiert werden. Dabei bezeichnen F_{KS}, $F_{kuip,k}$ und F_{CvM} die asymptotischen Verteilungsfunktionen von $\sqrt{n}d_\infty(F_n, F)$, $\sqrt{n}d_{kuip,k}(F_n, F)$ und $nd_{CvM}(F_n, F)$, vgl. Satz A.3. Weiter seien F_{KS}^{-1}, $F_{kuip,k}^{-1}$ und F_{CvM}^{-1} die zugehörigen Quantilsfunktionen. Das Diskrepanzprinzip kann natürlich auch auf Basis exakter Quantile implementiert werden, was aber höchstens für sehr kleine Stichproben zu anderen Ergebnissen führt.

1. Fall

Die einfachste Variante des Diskrepanzprinzips basiert auf Goodness-of-Fit-Tests zu einem unabhängig von n gewählten Niveau. Im „Data Approximation“-Ansatz von Davies wird vorgeschlagen, für die vorliegenden Daten das einfachste Modell zu wählen, das die Daten erzeugt haben könnte (vgl. [Dav08]). Letzteres wird formalisiert durch ein Kriterium, das sicherstellen soll, dass aus dem gewählten Modell simulierte Daten so aussehen wie die tatsächlich beobachteten. Im Fall von Kerndichteschätzern würde ein entsprechender Algorithmus ausgehend von einem sehr großen Startwert so lange die Bandbreite eines Kerndichteschätzers verringern, bis ein Goodness-of-Fit-Test die Hypothese, dass die Daten aus der durch das Integral des Kernschätzers gegeben Verteilungsfunktion stammen, nicht mehr ablehnt. Werden die asymptotischen Verteilungen des Kolmogorov-Smirnov- bzw. Kuiper-Tests benutzt, so ergibt sich ein Diskrepanzprinzip mit $d = d_\infty$ bzw. $d = d_{kuip,k}$ und

$$s(n) = cn^{-1/2}.$$

Die Konstante c wird dabei so gewählt, dass sie einem geeigneten Quantil der asymptotischen Verteilung von $\sqrt{n}d(F_n, F)$ entspricht.

Der „Data Approximation“-Ansatz legt hier extreme Quantile (95%, 99%) nahe. In Beispiel 10 in [Dav95] wird ein auf dem 98%-Quantil der Kuiper-Metrik basierendes Diskrepanzprinzip für Kerndichteschätzer vorgeschlagen, was allerdings noch mit einem weiteren Kriterium, dem sogenannten Extremwertfeature, kombiniert wird. Vapnik hingegen schlägt (für den Kolmogorov-Abstand) den Median (vgl. Kapitel 7.5.1 in [Vap00]) oder sogar den Modalwert vor. Der Modalwert der asymptotischen Verteilung von $\sqrt{n}d_\infty(F_n, F)$ liegt ungefähr bei 0.74, in [Vap98] wird $c = 0.6$ vorgeschlagen, in [Mar89] $c = 0.7$ oder $c = 0.5$. Es ist $F_{KS}(0.6) = 0.14$, d.h. die geschätzte Verteilungsfunktion soll in einem 14%-Konfidenzintervall liegen. Für die Rate, mit der die Bandbreite $h_{s,n}$ gegen 0 konvergiert, macht die Wahl des Quantils jedoch keinen Unterschied, solange ein festes Quantil gewählt wird. Satz 3.6 zeigt, dass für $f \in W^{\ell,1}$ und K Kern der Ordnung ℓ fast sicher

$$h_{s,n} = O(n^{-\frac{1}{2\ell}}(\log\log n)^{\frac{1}{2\ell}})$$

gilt, d.h. die Bandbreite konvergiert schneller gegen 0 als die nach verschiedenen Kriterien in diesem Fall optimalen Bandbreiten, für die $h \asymp n^{-\frac{1}{2\ell+1}}$ gelten sollte.

Das Problem liegt darin, dass die nichtparametrische Dichteschätzung ein schlecht gestelltes Problem ist, das Regularisierung erfordert. Auch mit optimal gewählter Bandbreite h wird eigentlich nicht die Dichte f, sondern eine geglättete Version $f * K_h$ geschätzt. Ein Kolmogorov-Smirnov-Test mit fest gewähltem Niveau entdeckt jedoch bei hinreichend großem n diesen Unterschied, und die zu schnell wachsende Macht erzwingt die Wahl eines zu kleinen Glättungsparameters. Asymptotisch liegt die geschätzte Verteilungsfunktion zu nahe an der empirischen Verteilungsfunktion, d.h. es kommt zur Unterglättung. In Satz 3.4 wurde bereits ein Beispiel gegeben, in dem die gewählte Bandbreite so schnell gegen 0 konvergieren kann, dass der resultierende Schätzer sogar inkonsistent ist; dieses Phänomen kann aber auch bei Diskrepanzprinzipien mit Ratenkorrektur (Fall 3) auftreten. In der Simulationsstudie in Kapitel 3.4 zeigt sich aber, dass insbesondere die auf extremen Quantilen basierenden Diskrepanzprinzipien auch noch für Stichprobenumfänge $n = 2500$ zu stark glätten. Die von Vapnik vorgeschlagene Version mit $c = 0.6$ funktioniert jedoch recht gut.

2. Fall

Das Gesetz vom iterierten Logarithmus für die empirische Verteilungsfunktion (vgl. Satz A.3) kommt einer deterministischen Schranke für $d(F_n, F)$ wohl am nächsten. Die Verwendung von Diskrepanzprinzipien mit $d = d_\infty$ und Schrankenfunktion $s(n) = c\sqrt{\frac{\log\log n}{2n}}$ wird in [AV89] vorgeschlagen. Dort wird gezeigt, dass mit der Wahl $c = (1 + \tilde{k} + \varepsilon)$ für $\varepsilon > 0$ fast sicher

$$h_{s,n} \asymp \left(\frac{\log\log n}{n}\right)^{\frac{1}{2\ell}}$$

erreicht wird (vgl. auch Satz 3.7 der vorliegenden Arbeit), was zu einer fast optimalen Konvergenzrate des Dichteschätzers bezüglich des L_∞-Risikos führt. In der Simulationsstudie [Mar89] wird $c = 1$ gewählt, was zu kleineren Bandbreiten führt. Trotzdem wird für die dort betrachteten Stichprobenumfänge sehr stark geglättet, und die Methode schneidet nicht besonders gut ab. Die Wahl $c = 1$ erfüllt jedoch nicht die Voraussetzungen von Satz 3.7. Die Simulationsstudie in Kapitel 3.4 bestätigt im Wesentlichen die Resultate aus [Mar89].

Verschiedene Werte von c sind im Gegensatz zum Fall $s(n) = cn^{-1/2}$, wo sie direkt verschiedenen festen Testniveaus des Kolmogorov-Smirnov–Tests bzw. äquivalent dazu Konfidenzwahrscheinlichkeiten entsprechen, nicht einfach zu interpretieren. Sie entsprechen jedoch von n abhängigen, mit wachsendem n gegen 1 konvergierenden Konfidenzwahrscheinlichkeiten, wie man durch die Umformung

$$s(n) = c\sqrt{\frac{\log\log n}{2n}} = \left(c\sqrt{\frac{\log\log n}{2}}\right) n^{-1/2}$$

Tabelle 3.1.: Werte von $F_{KS}(\sqrt{n}s(n))$ für verschiedene Wahlen von $s(n)$.

n	$0.0684n^{-2/5}$	$0.1357n^{-2/5}$	$0.3n^{-2/5}$	$0.35n^{-2/5}$	$0.4n^{-2/5}$	$\sqrt{\frac{\log\log n}{2n}}$	$2.1\sqrt{\frac{\log\log n}{2n}}$
10	0.00000000	0.00000000	0.00116	0.00989	0.03838	0.20146	0.94821
50	0.00000000	0.00000000	0.01071	0.04841	0.12467	0.49729	0.99492
100	0.00000000	0.00000000	0.02249	0.08199	0.18361	0.57015	0.99751
250	0.00000000	0.00000000	0.05117	0.14640	0.28011	0.63993	0.99888
500	0.00000000	0.00000004	0.08597	0.21043	0.36383	0.67952	0.99933
1000	0.00000000	0.00000045	0.13384	0.28601	0.45277	0.71135	0.99958
2500	0.00000000	0.00000695	0.21737	0.39861	0.57139	0.74491	0.99976
5000	0.00000000	0.00003976	0.29393	0.48848	0.65689	0.76556	0.99983
10000	0.00000000	0.00017993	0.37883	0.57786	0.73506	0.78313	0.99988
100000	0.00000000	0.00719288	0.67090	0.82752	0.91848	0.82640	0.99996
1000000	0.00000055	0.06771299	0.88466	0.95882	0.98746	0.85529	0.99998

zeigt. Tabelle 3.1 enthält die enstprechenden Werte von $F_{KS}(\sqrt{n}s(n))$ für $c = 1$ und $c = 2.1$. Diese Berechnungen basieren auf der asymptotischen Verteilung der Kolmogorov-Smirnov-Teststatistik, was bei kleinem n zu Abweichungen von der tatsächlichen Wahrscheinlichkeit führen kann. Es zeigt sich, dass das Konfidenzniveau im Fall $c = 1$ langsam wächst: für $n = 50$ entspricht diese Wahl ungefähr dem 50%-Quantil. Für den Fall $c = 2.1$ ist bereits für $n = 50$ die Konfidenzwahrscheinlichkeit größer als 0.99, d.h. es ist zu erwarten, dass die Verwendung dieser Version zu sehr großen Bandbreiten führt.

3. Fall

Die Beobachtung, dass ein Diskrepanzprinzip, das auf festen Quantilen oder dem Gesetz vom iterierten Logarithmus basiert, unter Standardvoraussetzungen in der nichtparametrischen Dichteschätzung eine zu schnelle Konvergenz des Glättungsparameters gegen 0 zur Folge hat, motiviert Eggermont und LaRiccia zu einer ratenkorrigerten Version des Diskrepanzprinzips ([EL96, EL01]). Für einen symmetrischen Kern, der eine Wahrscheinlichkeitsdichte ist, wird vorgeschlagen, h so zu wählen, dass

$$d_\infty(F_n, \hat{F}_n^h) = 0.35n^{-2/5}.$$

Die Wahl des Exponenten führt dazu, dass die gewählte Bandbreite bei hinreichender Glattheit mit der „richtigen“ Geschwindigkeit $n^{-1/5}$ gegen 0 konvergiert. Satz 3.8 der vorliegenden Arbeit ist eine Verallgemeinerung des Hauptergebnisses der zitierten Arbeiten, die auch die Verwendung von $d = d_{kuip,k}$ und von Kernen höherer Ordnung erlaubt. Damit erhält man bei Verwendung eines Kerns der Ordnung ℓ bei der Wahl von $s(n) = cn^{-\gamma}$ mit $\gamma = \frac{\ell}{2\ell+1}$ die – bei hinreichender Glattheit von f – bezüglich der L_1- bzw. L_2-Norm optimale Größenordnung $h = \alpha n^{-\frac{1}{2\ell+1}}$. Die Konstante α hängt dabei außer von c von der unbekannten Dichte f ab und entspricht keiner der nach den genannten Kriterien optimalen Konstanten. Eggermont und LaRiccia wählen $c = 0.35$ auf Basis von Simulationen, mit dem Ziel, dass der L_1-Verlust für ungefähr unimodale Dichten mit nicht zu schweren Tails möglichst gering sein sollte. Dazu werden verschiedene Werte von c ausprobiert, wobei generell Werte zwischen 0.3 und 0.4 für akzeptabel befunden werden.

Die Umformung

$$s(n) = cn^{-2/5} = (cn^{1/10})n^{-1/2}$$

ermöglicht wieder eine Interpretation im Rahmen von Konfidenzwahrscheinlichkeiten, die von n abhängen. Tabelle 3.1 enthält die entsprechenden Werte von $F_{KS}(\sqrt{n}s(n))$ für $c = 0.3$, $c = 0.35$ und $c = 0.4$. Es fällt auf, dass die Konfidenzwahrscheinlichkeiten sowohl mit c als auch mit n stark variieren. Insbesondere liegt das Konfidenzniveau für die Wahl $c = 0.35$ noch bis $n = 5000$ unterhalb von 0.5.

Im Prinzip lassen sich durch Simulationen auch geeignete Konstanten für andere Klassen von Dichten, andere Abstandsmaße oder Kerne höherer Ordnung finden. Satz 3.8 erlaubt jedoch außerdem einen anderen Ansatz: Diskrepanzprinzipien, die auf Referenzdichten (z.B. der Normalverteilung) basieren. Klassische Plug-In-Schätzer mit Referenzdichten beruhen auf der Idee, in einem Ausdruck für die optimale Bandbreite unbekannte, von f abhängende Konstanten durch die entsprechenden Werte für die Referenzdichte zu ersetzen. Demgegenüber können auch in der Schrankenfunktion $s(n) = cn^{-\gamma}$ die Konstanten γ und c so gewählt werden, dass die asymptotisch durch das Diskrepanzprinzip gewählte Bandbreite der „optimalen“ entspricht. Dies soll nun am Beispiel der Normalverteilung demonstriert werden. Bezüglich der L_2-Norm ergibt sich die asymptotisch optimale Bandbreite zur Schätzung von f bei Verwendung eines Kerns K der Ordnung ℓ als

$$h_{opt} = \left(\frac{\ell!^2 \|K\|_2^2}{2\ell k_\ell^2 \|f^{(\ell)}\|_2^2} \right)^{\frac{1}{2\ell+1}} n^{-\frac{1}{2\ell+1}}, \tag{3.23}$$

vgl. [WJ95], S.33. Mit der Wahl $\gamma = \frac{\ell}{2\ell+1}$ besitzt die Bandbreite (3.22) die korrekte Größenordnung. Damit auch die korrekte Konstante resultiert, muss c so gewählt werden, dass

$$\left(\frac{c\ell!}{k_\ell d(f^{(\ell-1)}, 0)} \right)^{\frac{1}{\ell}} = \left(\frac{\ell!^2 \|K\|_2^2}{2\ell k_\ell^2 \|f^{(\ell)}\|_2^2} \right)^{\frac{1}{2\ell+1}}$$

gilt. Auflösen nach c liefert

$$c = \left(\frac{\|K\|_2^{2\ell} k_\ell}{(2\ell)^\ell \ell!} \right)^{\frac{1}{2\ell+1}} \frac{d(f^{(\ell-1)}, 0)}{\|f^{(\ell)}\|_2^{2\ell/(2\ell+1)}}. \tag{3.24}$$

Der erste Faktor hängt dabei nur vom verwendeten Kern ab und ist invariant unter Reskalierungen des Kerns, wie man leicht nachprüft. Für den Gauß- und den Epanechnikov-Kern $K^{\mathcal{N}}$ bzw. $K^{\mathcal{E}}$ ergeben sich nach Einsetzen der Konstanten (vgl. Tabelle B.1)

$$\left(\frac{\|K^{\mathcal{N}}\|_2^{2\ell} k_\ell^{\mathcal{N}}}{(2\ell)^\ell \ell!} \right)^{\frac{1}{2\ell+1}} = \left(\frac{1}{128\pi} \right)^{\frac{1}{5}} \approx 0.3014$$

bzw.

$$\left(\frac{\|K^{\mathcal{E}}\|_2^{2\ell}k_\ell^{\mathcal{E}}}{(2\ell)^\ell \ell!}\right)^{\frac{1}{2\ell+1}} = \frac{1}{2}\left(\frac{9}{125}\right)^{\frac{1}{5}} \approx 0.2954.$$

Der zweite Faktor auf der rechten Seite von (3.24) hängt ebenfalls nur von der Form von f ab und ändert sich nicht, wenn f durch $a^{-1}f(a^{-1}(\cdot - b))$ ersetzt wird. Im Folgenden wird nun als Referenzdichte die Normalverteilung benutzt, also $f = \phi$. Da ϕ' sein globales Maximum in -1 und sein globales Minimum in 1 annimmt, der Funktionswert an diesen Stellen jeweils betragsmäßig gleich $\frac{1}{\sqrt{2\pi e}}$ ist und keine weiteren lokalen Extrema existieren, ergeben sich

$$d_\infty(\phi', 0) = \frac{1}{\sqrt{2\pi e}}$$

bzw.

$$d_{kuip,k}(\phi', 0) = \frac{\min\{k+1, 4\}}{\sqrt{2\pi e}}.$$

Weiter ist $\|\phi^{(2)}\|_2^2 = \frac{3}{8\sqrt{\pi}}$, und damit ergeben sich bei Verwendung von $d = d_\infty$ die Werte $c = 0.1357$ für den Gaußkern und $c = 0.1331$ für den Epanechnikovkern. Beide Werte sind kleiner als die von Eggermont und LaRiccia unabhängig vom Kern vorgeschlagene Konstante 0.35. Bei Verwendung von $d = d_{kuip,1}$ ergeben sich $c = 0.2715$ für den Gaußkern und $c = 0.2661$ für den Epanechnikovkern.

Analog lässt sich eine ähnliche Wahl für den L_1-Verlust durchführen. Für einen Kern der Ordnung $\ell = 2$ wurde in der Literatur folgende asymptotische Wahl der Bandbreite vorgeschlagen (vgl. Kapitel 5 in [DG85] bzw. [EL01], S. 298):

$$h_{L_1} = \left(\frac{\|K\|_2^2\|\sqrt{f}\|_1^2}{2\pi k_2^2\|f^{(2)}\|_1^2}\right)^{\frac{1}{5}} n^{-\frac{1}{5}}.$$

Mit der Wahl $\gamma = \frac{2}{5}$ besitzt die Bandbreite (3.22) die korrekte Größenordnung. Damit auch die korrekte Konstante resultiert, muss ähnlich wie oben c so gewählt werden, dass

$$\left(\frac{2c}{k_2 d(f^{(1)}, 0)}\right)^{\frac{1}{2}} = \left(\frac{\|K\|_2^2\|\sqrt{f}\|_1^2}{2\pi k_2^2\|f^{(2)}\|_1^2}\right)^{\frac{1}{5}}$$

gilt. Auflösen nach c liefert

$$c = \left(\frac{2^{3/5}\|K\|_2^{4/5}k_2^{1/5}}{4\pi^{2/5}}\right)\frac{\|\sqrt{f}\|_1^{4/5}d(f^{(1)}, 0)}{\|f^{(2)}\|_1^{4/5}}. \tag{3.25}$$

Dabei hängt der erste Faktor nur von der Form des benutzten Kerns ab und ist invariant unter Reskalierungen. Ebenso hängt der zweite Faktor nur von der Form von f ab und ändert sich nicht, wenn f durch $a^{-1}f(a^{-1}(\cdot - b))$ ersetzt wird. Mit den entsprechenden Werten aus Tabelle B.1 erhält man

$$\left(\frac{2^{3/5}\|K^{\mathcal{N}}\|_2^{4/5}(k_2^{\mathcal{N}})^{1/5}}{4\pi^{2/5}}\right) = \frac{2^{1/5}}{4\pi^{3/5}} \approx 0.1445$$

bzw.

$$\left(\frac{2^{3/5}\|K^{\mathcal{E}}\|_2^{4/5}(k_2^{\mathcal{E}})^{1/5}}{4\pi^{2/5}}\right) = \frac{2^{3/5}3^{2/5}5^{2/5}}{20\pi^{2/5}} \approx 0.1416.$$

Mit $\|\sqrt{f}\|_1 = \pi^{1/4}2^{3/4}$ und $\|f^{(2)}\|_1 = \frac{2\sqrt{2}}{\sqrt{\pi e}}$ ergeben sich bei Verwendung von $d = d_\infty$ die Werte $c = 0.0684$ für den Gaußkern und $c = 0.067$ für den Epanechnikovkern. Bei Verwendung von $d = d_{kuip,1}$ ergeben sich $c = 0.1368$ für den Gaußkern und $c = 0.1341$ für den Epanechnikovkern.

Die Tabelle 3.1 zeigt in Abhängigkeit von n die sich ergebenden Konfidenzniveaus bei Verwendung der auf der Normalverteilung als Referenzdichte basierenden Schrankenfunktionen für $K = K^{\mathcal{N}}$ und $d = d_\infty$. Die Konfidenzniveaus sind auch für großes n noch extrem klein, d.h. die auf Basis dieser Schrankenfunktionen gewählten Bandbreiten werden extrem klein sein.

Abbildung 3.3 zeigt am Beispiel einer Stichprobe vom Umfang $n = 100$ aus einer Standardnormalverteilung die Ergebnisse für verschiedene Varianten des Diskrepanzprinzips. Zur Schätzung wird der Gaußkern verwendet. Links oben ist zum Vergleich die Schätzung bei Verwendung der bezüglich des L_2-Risikos asymptotisch optimalen Bandbreite (3.23) zu sehen. Die anderen drei Bilder zeigen die Ergebnisse für das Diskrepanzprinzip unter Verwendung der d_∞-Metrik. Rechts oben ist die von Vapnik vorgeschlagene Version abgebildet ($s(n) = 0.6n^{-1/2}$). Sie liefert ein recht gutes Ergebnis. Im Bild links unten ist das Resultat bei Verwendung des asymptotischen 0.95-Quantils der Kolmogorov-Smirnov-Teststatistik zu sehen. Hier wird eine deutlich zu große Bandbreite gewählt. Die Bandbreite im Bild rechts unten wurde als Lösung von $d_\infty(\hat{F}_n^h, F) = 0.1357n^{-2/5}$ gewählt. Wie oben gezeigt, erhält man mit dieser Methode für die Normalverteilung asymptotisch die optimale Bandbreite. Man sieht allerdings sehr deutlich, dass für kleines n eine viel zu kleine Bandbreite gewählt wird. Generell zeigen die Simulationen in Kapitel 3.4, dass die asymptotischen Resultate aus Kapitel 3.2 mindestens für Stichproben bis zum Umfang $n = 2500$ nicht aussagekräftig sind.

3.4. Simulationstudie

In den Achtziger- und Neunzigerjahren sind zahlreiche Simulationsstudien zum praktischen Vergleich verschiedener Bandbreitenwahlmethoden für Kerndichteschätzer durchgeführt worden. Unter den besonders umfassenden Studien sind hier vor allem [CCGM94] (mit Schwerpunkt auf dem L_2-Risiko) sowie [BD94] und [Dev97] im L_1-Kontext zu nennen. Dem Verfasser ist jedoch keine größere Vergleichsstudie bekannt, die Diskrepanzprinzipien einbezieht, auch wenn in [Dev97] die Variante von Eggermont und LaRiccia [EL96]

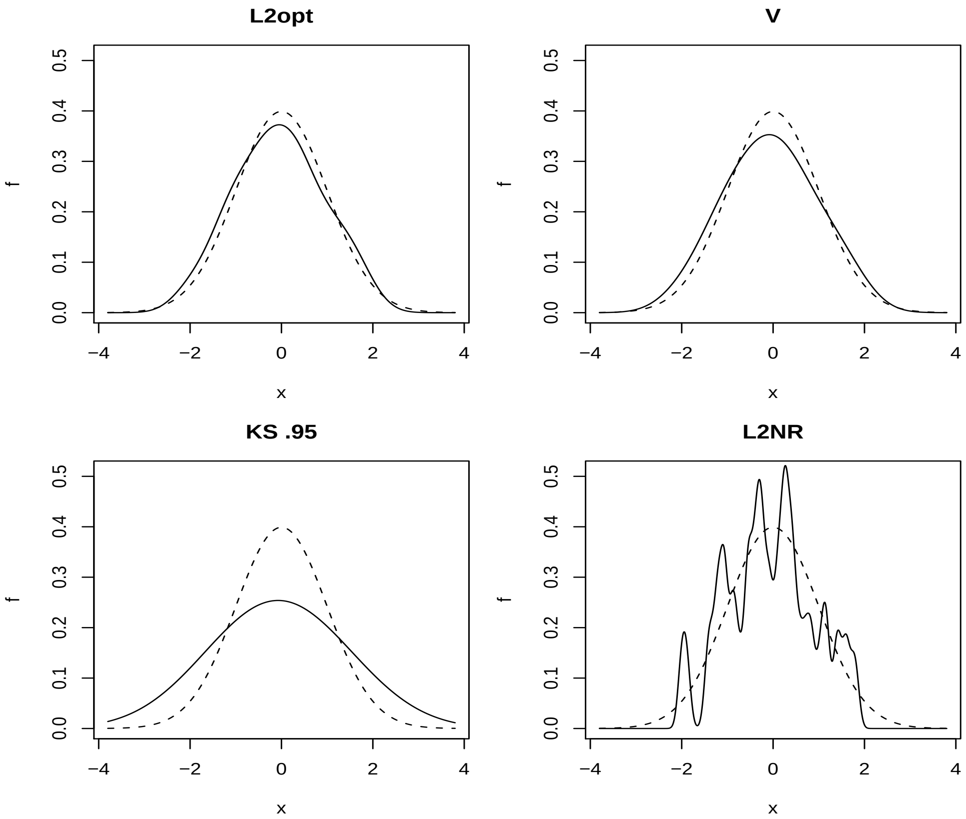

Abbildung 3.3.: Dichteschätzungen für eine $N(0,1)$-Stichprobe ($n = 100$) bei Verwendung des Gaußkerns. Als Bandbreiten wurde die optimale L_2-Bandbreite (3.23) (**L2opt**) bzw. Lösungen von $d_\infty(\hat{F}_n^h, F) = s(n)$ mit $s(n) = 0.6n^{-1/2}$ (**V**), $s(n) = 1.36n^{-1/2}$ (**KS .95**) sowie $s(n) = 0.1357n^{-2/5}$ (**L2NR**) gewählt.

erwähnt und als „interesting“ bezeichnet, aber nicht in die Studie aufgenommen wird. In der Literatur liegen bislang nur kleinere Simulationsstudien vor – im Rahmen der Arbeiten, in denen die Diskrepanzprinzipien vorgeschlagen werden oder direkt auf diese aufbauend [Mar89, EL96, EL01]. In einer umfangreichen Simulationsstudie werden daher im Folgenden erstmals die verschiedene Varianten des Diskrepanzprinzips für Kerndichteschätzer sowohl untereinander als auch mit anderen Methoden der Bandbreitenwahl in einer größeren Anzahl verschiedener Szenarien verglichen.

Es werden jeweils 100 Stichproben mit Umfängen 50, 100, 500, 1000 und 2500 aus den in [BD94] vorgestellten 28 Beispieldichten realisiert. Dazu wird das `R`-Paket `benchden` verwendet [MWT09]. Unter den Dichten finden sich sowohl aus der Statistik bekannte Standardverteilungen wie die Gleichverteilung, die Normalverteilung und die Exponentialverteilung als auch speziell konstruierte Dichten mit besonderen Eigenschaften. Enthalten sind uni- und multimodale Dichten, Dichten mit kompaktem Träger und solche mit schweren Tails, Mischungen von Normal- oder Gleichverteilungen, Dichten mit Polstellen oder nicht zu-

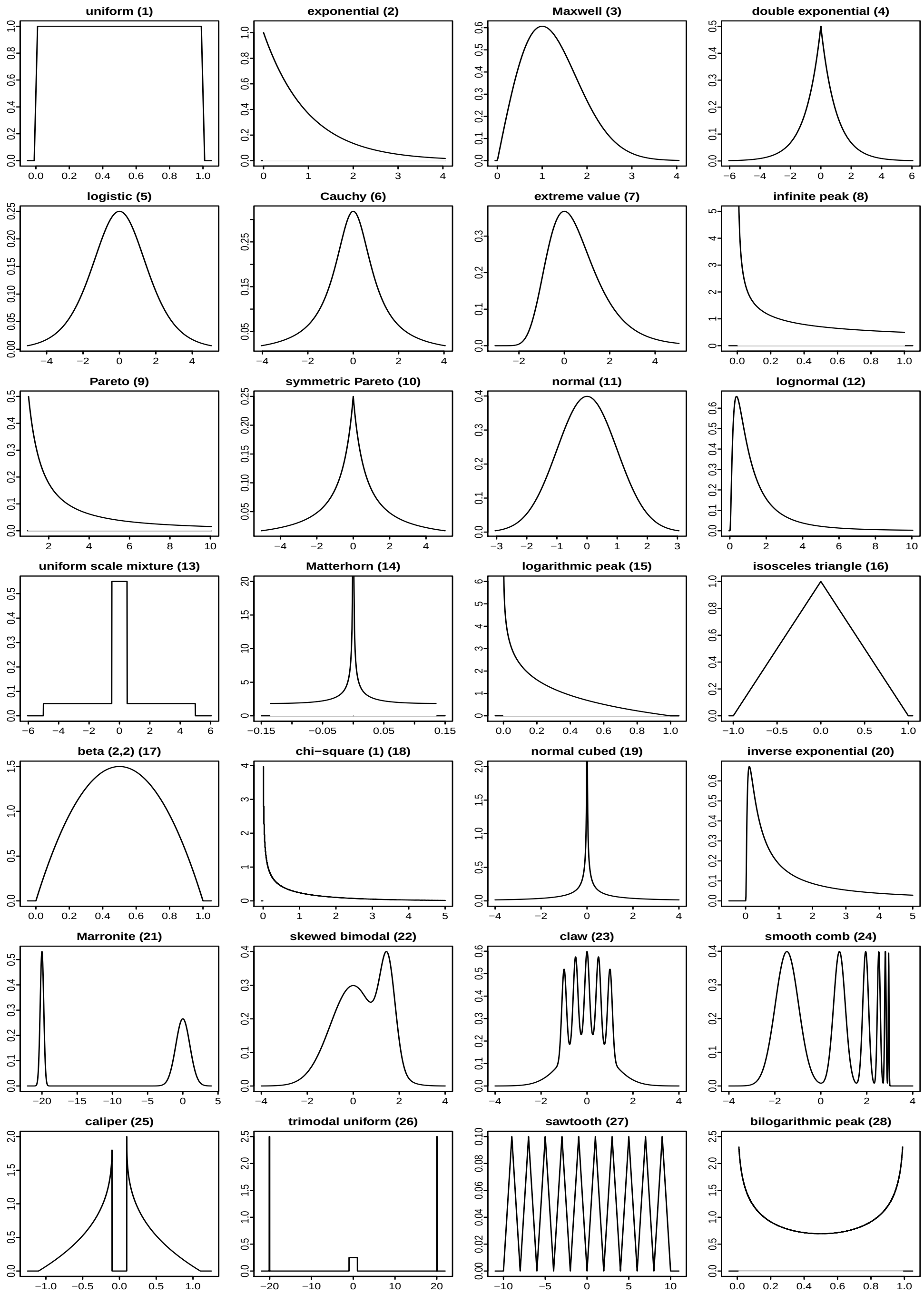

Abbildung 3.4.: Die verwendeten Dichten.

sammenhängendem Träger. Die Dichten sind in Abbildung 3.4 dargestellt, eine nähere Beschreibung gibt Anhang B.1.

Berechnet werden (soweit definiert) jeweils die Verluste bezüglich d_1, d_2^2 und d_∞. Die

Berechnung erfolgt ähnlich wie in Kapitel 5 von [RMG09] beschrieben durch numerische Integration mit Hilfe einer Trapezregel. Die arithmetischen Mittel über die 100 Simulationsläufe werden als Schätzungen der zugehörigen Risiken verwendet. Einige der Dichten liegen jedoch nicht in L_2 bzw. L_∞, weshalb der entsprechende Verlust nicht definiert ist. Desweiteren ist eine Betrachtung des L_∞-Risikos nur sinnvoll für stetige Dichten, da das bestmögliche Risiko ansonsten im Wesentlichen durch den höchsten Sprung bestimmt ist. Die Tabellen für das L_2-Risiko zeigen daher keine Ergebnisse für die Dichten 8, 14, 18 und 19, in den Tabellen für das L_∞-Risiko fehlen zusätzlich noch die Dichten 1, 2, 9, 13, 15, 25, 26 sowie 28. Zusätzlich werden die arithmetischen Mittel der gewählten Bandbreiten angegeben.

Es werden folgende Versionen des Diskrepanzprinzips verglichen:

- Auf dem Kolmogorov-Smirnov-Test basierende Versionen: h wird gewählt als Lösung von $d_\infty(F_n, \hat{F}_n^h) = cn^{-1/2}$ mit $c = 0.83, 1.22, 1.36, 1.62$. Dies entspricht dem Median, dem 90%-, dem 95% bzw. dem 99%-Quantil der asymptotischen Verteilung von $\sqrt{n} d_\infty(F_n, F)$. Diese Methoden werden in den Tabellen mit **KS .5**, **KS .9**, **KS .95** bzw. **KS .99** bezeichnet.

- Die von Vapnik vorgeschlagene Variante: h wird gewählt als Lösung von $d_\infty(F_n, \hat{F}_n^h) = cn^{-1/2}$ mit $c = 0.6$. In der Tabelle mit **V** bezeichnet.

- Die ratenkorrigierte Version von Eggermont und LaRiccia: h wird gewählt als Lösung von $d_\infty(F_n, \hat{F}_n^h) = 0.35n^{-2/5}$. In der Tabelle mit **E-LR** bezeichnet.

- Auf dem Gesetz vom Iterierten Logarithmus für den Kolmogorov-Smirnov-Abstand basierende Versionen: $d_\infty(F_n, \hat{F}_n^h) = c\sqrt{\frac{\log\log(n)}{2n}}$ für $c = 1$ und $c = 2.1$. Der Fall $c = 1$ wird nicht von Satz 3.7 abgedeckt, jedoch wurde diese Variante in der Simulationsstudie in [Mar89] verwendet. Der Wert $c = 2.1$ ergibt sich, wenn in Satz 3.7 $\varepsilon = 0.1$ gewählt wird. In den Tabellen mit **LIL** bzw. **LIL 2.1** bezeichnet.

- Auf der Kuiper-Metrik basierende Versionen: h wird als Lösung von $d_{kuip,1}(F_n, \hat{F}_n^h) = cn^{-1/2}$ gewählt mit $c = 1.22, 1.62, 1.75, 1.99$. Dies entspricht dem Median, dem 90%-, dem 95% bzw. dem 99%-Quantil der asymptotischen Verteilung von $\sqrt{n} d_{kuip,1}(F_n, F)$. Diese Methoden werden in den Tabellen mit **Kuip .5**, **Kuip .9**, **Kuip .95** bzw. **Kuip .99** bezeichnet.

- Auf dem Gesetz vom iterierten Logarithmus für den Kuiper-Abstand basierende Versionen: $d_{kuip,k}(F_n, \hat{F}_n^h) = c\sqrt{\frac{\log\log(n)}{2n}}$ für $c = 2$ und $c = 4.2$. Der Fall $c = 2$ wird nicht von Satz 3.7 abgedeckt, $c = 4.2$ ergibt sich, wenn in Satz 3.7 $\varepsilon = 0.1$ gewählt wird. In den Tabellen mit **kLIL** bzw. **kLIL 4.2** bezeichnet.

- Auf dem Cramér-von Mises-Test basierende Versionen: h wird gewählt als Lösung von $d_{CvM}(F_n, \hat{F}_n^h) = cn^{-1}$ mit $c = 0.12, 0.35, 0.46, 0.74$. Dies entspricht dem Median, dem 90%-, dem 95% bzw. dem 99%-Quantil der asymptotischen Verteilung von

$nd_{CvM}(F_n, F)$. Diese Methoden werden in den Tabellen mit **CvM .5**, **CvM .9**, **CvM .95** bzw. **CvM .99** bezeichnet. Die von Vapnik (vgl. z.B. Kapitel 7.9 in [Vap98]) vorgeschlagene Version mit $c = 0.05$ wird mit **CvM V** bezeichnet.

- Auf dem Gesetz vom iterierten Logarithmus für den Cramér-von Mises-Test basierende Version: h wird gewählt als Lösung von $d_{CvM}(F_n, \hat{F}_n^h) = \frac{2 \log\log(n)}{\pi^2 n}$ (vgl. Satz A.4). In den Tabellen mit **CvM LIL** bezeichnet.

- Die in Kapitel 3.3 vorgeschlagenen, auf der Normalverteilung als Referenzdichte basierenden Methoden. In den Tabellen bezeichnen **L2NR** und **L1NR** die L_2- und L_1-basierten Methoden für $d = d_\infty$ und **kL2NR** und **kL1NR** die entsprechenden Methoden für $d = d_{kuip,1}$.

In [EL96] wird zur Berechnung einer Lösung von $d(F_n, \hat{F}_n^h) = s(n)$ die Verwendung einer Sekantenmethode vorgeschlagen. Hier wird jedoch die eng verwandte Regula Falsi (vgl. Kapitel 5.3.1 in [Sch93]) verwendet, da sie sich als stabiler erweist. Wie in Kapitel 3.2 bemerkt, können gelegentlich mehrere Lösungen existieren. In [EL96] wird in diesem Fall die Verwendung der kleinsten Lösung vorgeschlagen, während der datenapproximative Ansatz die Verwendung der größten Lösung nahelegt. Das Problem eventuell existierender multipler Lösungen wird hier ignoriert, und es wird jeweils die erste gefundene Lösung verwendet. Benutzt wird der Gaußkern $K^{\mathcal{N}}$, zum Vergleich für einige Methoden auch der Epanechnikov-Kern $K^{\mathcal{E}}$. Auf die Implementierung für Kerne höherer Ordnung wird verzichtet, da sie zu nicht monotonen Verteilungsfunktionen führen, für die $d(F_n, \hat{F}_n^h)$ nicht mehr mit Hilfe der Formeln aus Satz A.3 effizient ausgewertet werden kann.

Zum Vergleich werden einige weit verbreitete Standardverfahren der Bandbreitenwahl herangezogen. Eine ausführliche Diskussion der Vor- und Nachteile der einzelnen Methoden findet man z.B. in Kapitel 3 von [WJ95]. Alle vier Methoden zielen auf die Minimierung des L_2-Risikos ab. Es existieren auch Methoden für L_1, diese sind jedoch weniger weit verbreitet und schwieriger zu implementieren.

- Leave-One-Out-L_2-Kreuzvalidierung unter Verwendung von Formel (13) für $p = 1$ in [CR08]. Die Bandbreite h_{L2CV} wird als Minimum der Funktion

$$L2CV(h) = \|K_h\|_2^2 + \frac{n-2}{n(n-1)} \sum_{i \neq j} (K * K)_h(X_i - X_j) - \frac{2}{n} \sum_{i \neq j} K_h(X_i - X_j)$$

gewählt. Dabei ist $(n-1)^{-1} L2CV(h)$ für jedes h bis auf eine additive Konstante, die nicht von h abhängt, ein unverzerrter Schätzer für $d_2^2(f, \hat{f}_n^h)$. Es ist bekannt, dass die Methode tendenziell sehr kleine Bandbreiten auswählt und eine hohe Variabilität aufweist (vgl. [WJ95], Kapitel 3.3). Zur praktischen Implementation wird $L2CV$ für 250 Werte auf einem Gitter ausgewertet. Dabei wird das größte lokale Minimum gewählt, das oft auch das globale Minimum ist. In den Tabellen mit **L2CV** bezeichnet.

- Die verzerrte Kreuzvalidierung (*biased cross-validation*) ist dagegen stabiler (vgl. [WJ95], Kapitel 3.4). Minimiert wird die Zielfunktion

$$BCV(h) = (nh)^{-1}\|K\|_2^2 + \frac{1}{4}h^4 k_2^2 n^{-2} \sum\sum_{i \neq j} (K_h'' * K_h'')(X_i - X_j).$$

  Da $BCV(h)$ ein globales Minimum in 0 besitzt, wird wie für **L2CV** das größte lokale Minimum auf einem Gitter gewählt. In den Tabellen mit **BCV** bezeichnet.

- Plug-In-Schätzer beruhen auf der Idee, in einem Ausdruck für die nach dem gewählten Kriterium optimale Bandbreite (z.B. (3.14) für das L_2-Risiko) die unbekannten Funktionale der gesuchten Dichte aus den Daten zu schätzen. Dies geschieht in mehreren Schritten, wobei im ersten Schritt die $N(0, \sigma^2)$-Verteilung als Referenzdichte benutzt wird. Hier wird die in Kapitel 3.6.1 von [WJ95] angegebene Version benutzt. Im ersten Schritt wird eine robuste Schätzung für σ verwendet:

$$\hat{\sigma} = 1.4826\text{mad}(X_1, \ldots, X_n),$$

  wobei

$$\text{mad}(X_1, \ldots, X_n) := \text{med}\{|X_1 - \text{med}(X_1, \ldots, X_n)|, \ldots, |X_n - \text{med}(X_1, \ldots, X_n)|\}.$$

  Nacheinander werden dann berechnet:

$$\begin{aligned}
\hat{\varphi}_8 &= \frac{105}{32\sqrt{\pi}\hat{\sigma}^9} \\
g_1 &= \left(\frac{-2K^{\mathcal{N}(6)}(0)}{k_2^{\mathcal{N}}\hat{\varphi}_8 n}\right)^{1/9} \\
\hat{\varphi}_6 &= \frac{1}{n^2}\sum_i\sum_j K_{g_1}^{\mathcal{N}(6)}(X_i - X_j) \\
g_2 &= \left(\frac{-2K^{\mathcal{N}(4)}(0)}{k_2^{\mathcal{N}}\hat{\varphi}_6 n}\right)^{1/7} \\
\hat{\varphi}_4 &= \frac{1}{n^2}\sum_i\sum_j K_{g_2}^{\mathcal{N}(4)}(X_i - X_j) \\
h_{PI}^{\mathcal{N}} &= \left(\frac{1}{2\sqrt{\pi}\hat{\varphi}_4 n}\right)^{1/5}.
\end{aligned}$$

  Da höhere Ableitungen des Kerns benötigt werden, wird für den Epanechnikov-Kern zunächst $h_{PI}^{\mathcal{N}}$ berechnet und dann $h_{PI}^{\mathcal{E}} = (900\pi)^{1/10} h_{PI}^{\mathcal{N}}$ verwendet. Der Faktor $(900\pi)^{1/10}$ ergibt sich aus den sogenannten kanonischen Bandbreiten für $K^{\mathcal{N}}$ bzw. $K^{\mathcal{E}}$, die es erlauben, Bandbreiten für verschiedene Kerne zu vergleichen (vgl. Tabelle 3.2 in [HMSW04]). In den Tabellen mit **PI** bezeichnet.

- Die geglättete Kreuzvalidierung (*smoothed cross-validation*) ist eine weitere Modifikation der L_2-Kreuzvalidierung, enthält aber auch Elemente eines Plug-In-Schätzers. Die hier verwendete und in den Tabellen als **SCV** bezeichnete Version ist die am Ende von Kapitel 3.7 von [WJ95] gegebene. Es werden nacheinander berechnet:

$$\begin{aligned}
\hat{\sigma} &= 1.4826\text{mad}(X_1, \dots, X_n) \\
g_1 &= \left(\frac{2}{7n}\right)^{1/9} \sqrt{2}\hat{\sigma} \\
g_2 &= \left(\frac{2}{11n}\right)^{1/13} \sqrt{2}\hat{\sigma} \\
\hat{\varphi}_6 &= \frac{1}{n^2} \sum_i \sum_j K_{g_1}^{\mathcal{N}(6)}(X_i - X_j) \\
\hat{\varphi}_{10} &= \frac{1}{n^2} \sum_i \sum_j K_{g_2}^{\mathcal{N}(10)}(X_i - X_j) \\
g_3 &= \left(\frac{-6}{\sqrt{2\pi}\hat{\varphi}_6 n}\right)^{1/7} \\
g_4 &= \left(\frac{-210}{\sqrt{2\pi}\hat{\varphi}_{10} n}\right)^{1/11} \\
\hat{\varphi}_4 &= \frac{1}{n^2} \sum_i \sum_j K_{g_3}^{\mathcal{N}(4)}(X_i - X_j) \\
\hat{\varphi}_8 &= \frac{1}{n^2} \sum_i \sum_j K_{g_4}^{\mathcal{N}(8)}(X_i - X_j).
\end{aligned}$$

  Nun sei für $h > 0$

$$g_h = \left(\frac{441}{64\pi}\right)^{1/18} (4\pi)^{-1/5} \hat{\varphi}_4^{-2/5} \hat{\varphi}_8^{-1/9} n^{-23/45} h^{-2}.$$

  Die gewählte Bandbreite ergibt sich nun als Minimum $h_{SCV}^{\mathcal{N}}$ von

$$SCV(h) = \frac{1}{2\sqrt{\pi}nh} + \sum_i \sum_j \left(K^{\mathcal{N}}_{\sqrt{2h^2+2g_h^2}} - 2K^{\mathcal{N}}_{\sqrt{h^2+2g_h^2}} + K^{\mathcal{N}}_{\sqrt{2g_h^2}} \right) (X_i - X_j)$$

  für den Gaußkern und $h_{SCV}^{\mathcal{E}} = (900\pi)^{1/10} h_{SCV}^{\mathcal{N}}$ für den Epanechnikov-Kern.

Zunächst sollen die verschiedenen auf dem Kolmogorov-Smirnov-Test basierenden Versionen des Diskrepanzprinzips mit den vier Referenzmethoden verglichen werden. Die Tabellen 3.2-3.3 zeigen die Ergebnisse für das L_1-Risiko für diese Methoden bei Verwendung eines Gaußkerns. Zur besseren Übersicht ist jeweils für jedes Simulationsszenario das beste

Ergebnis hervorgehoben, wobei aber oft mehrere Methoden gleich gut sind oder sehr nahe beieinander liegen. Es fällt zunächst auf, dass in vielen Fällen entweder eine der Referenzmethoden **L2CV**, **BCV**, **PI** oder **SCV** oder eines der beiden Diskrepanzprinzipien **E-LR** und **V** am besten abschneidet. Die Methode **KS .5** ist in den meisten Fällen etwas schlechter, während die auf extremen Quantilen basierenden Versionen **KS .9**, **KS .95** und **KS .99** mit nur wenigen Ausnahmen deutlich schlechtere Ergebnisse liefern.

Von den Referenzmethoden schneiden bei unimodalen Dichten (1-20) besonders **BCV** und **PI** gut ab. Für multimodale Dichten (21-28) ist **L2CV** häufig besser als alle anderen Methoden, wobei **SCV** mit Ausnahme von Dichte 26 sehr ähnliche Ergebnisse liefert. Die Ergebnisse für die beiden Methoden **E-LR** und **V** ähneln sich sehr. Sie gehören für unimodale Dichten oft zu den besten Methoden. Für multimodale Dichten schneiden sie fast durchgehend schlechter ab als **L2CV**, häufig aber besser als **BCV** und **PI**. In der kleinerern Simulationsstudie in Kapitel 8.3 von [EL01] wird **E-LR** mit einigen anderen Methoden verglichen, wobei sich ähnliche Resultate ergeben. Mit den in dieser Arbeit verwendeten multimodalen Dichten hat das Diskrepanzprinzip jedoch teilweise größere Schwierigkeiten als mit den dort verwendeten.

Auf dem Diskrepanzprinzip basierende Methoden, die den Median der Verteilung der Kolmogorov-Smirnov-Statistik (**KS .5**) oder extreme Quantile verwenden (**KS .9**, **KS .95** und **KS .99**), schneiden fast durchgehend schlecht ab. Zu den wenigen Ausnahmen gehören Dichten mit schweren Rändern und solche mit unendlichen Peaks. Der L_1-Abstand bestraft Abweichungen in den Tails relativ stark, weshalb eine starke Glättung nach diesem Kriterium vorteilhaft ist (vor allem Dichten 9 und 18-20). Dieser Effekt tritt selbst für die Exponential- und die Doppelexponentialverteilung auf (2 bzw. 4), wenn auch nicht so prononciert wie z.B. bei der Pareto-Verteilung (9). Das relativ gute Abschneiden bei einigen Dichten mit unendlichen Peaks (8, 14, 18, 19) ist hingegen erklärungsbedürftig, da Satz 3.4 zeigt, dass nach dem Diskrepanzprinzip gewählte Bandbreiten bei Dichten mit unendlichen Peaks zu schnell gegen 0 konvergieren können, so dass der resultierende Schätzer evtl. nicht mehr konsistent ist. Die genannten Dichten liegen allerdings nicht in L_2, weshalb es nicht verwundert, dass die auf ein geringes L_2-Risiko abzielenden Referenzmethoden auch nicht besonders gut funktionieren. Wird die Bandbreite jedoch nach dem Diskrepanzprinzip gewählt, führt die Verwendung eines größeren Quantils zu einer größeren Bandbreite, was das Problem etwas abmildert.

Für das L_2-Risiko (Tabellen 3.4 und 3.5) ergibt sich ein ähnliches Bild: Von den Referenzmethoden schneiden **BCV** und **PI** wieder für die unimodalen Dichten sehr gut ab, während **L2CV** und **SCV** (außer für Dichte 26) generell für die multimodalen Dichten besonders gut geeignet sind. Die Methoden **V** und **E-LR** verhalten sich auch nach diesem Kriterium untereinander sehr ähnlich und sind für unimodale Dichten häufig am besten oder nur unwesentlich schlechter als die beste der Referenzmethoden. Für multimodale Dichten sind sie schlechter geeignet als **L2CV** und **SCV** aber in der Regel besser als **BCV** und **PI**. Mit ganz wenigen Ausnahmen schneidet auch **KS .5** schlechter ab als **E-LR** und **V**. Da der L_2-Verlust Fehler in den Tails kaum bestraft (im Gegensatz zum L_1-Verlust) und die Dichten 8, 14, 18 und 19 nicht in L_2 liegen, bietet die Verwendung von **KS .9**, **KS .95** und **KS .99** bei keiner der Dichten einen Vorteil. Sehr ähnliche Ergebnisse ergeben sich für das L_∞-Risiko (vgl. Tabelle B.2 im Anhang).

Die Tabellen 3.6 und 3.7 zeigen die arithmetischen Mittel der jeweils gewählten Bandbreiten. In den meisten Fällen sind die mittels **SCV** gewählten Bandbreiten kleiner als die mittels **L2CV** gewählten, während die Verwendung von **BCV** und **PI** zu größeren Bandbreiten führt. Die durch **E-LR** und **V** gewählten Bandbreiten liegen häufig zwischen denen für **L2CV** und **BCV** bzw. **PI**. Wichtige Ausnahmen sind die Dichten 8, 14, 18 und 19, die unendliche Peaks besitzen. Hier wählen **SCV**, **PI** und besonders **BCV** im Vergleich zu den anderen Methoden relativ große Bandbreiten, was zu einem geringeren L_1-Risiko führt. Die auf dem Diskrepanzprinzip basierenden Methoden **KS .5**, **KS .9**, **KS .95** und **KS .99** liefern durchgehend sehr große Bandbreiten, besonders im Vergleich zu **L2CV**. Das schlechte Abschneiden dieser Varianten ist also auf Überglättung zurückzuführen, obwohl sie nach Satz 3.6 asymptotisch unterglätten. Dies zeigt, dass die asymptotischen Resultate aus Kapitel 3.2 selbst noch für den Stichprobenumfang $n = 2500$ kaum aussagekräftig sind.

Die Tabellen B.3-B.9 im Anhang zeigen die Simulationsergebnisse für das L_1-, L_2- und L_∞-Risiko sowie die gewählten Bandbreiten bei Verwendung des Epanechnikov-Kerns. Wie zu erwarten, ergeben sich im Vergleich zu den Ergebnissen für den Gaußkern kaum Unterschiede.

Nun soll untersucht werden, inwiefern die Verwendung anderer Goodness-of-Fit-Tests die Ergebnisse beeinflusst. Die Tabellen 3.8 und 3.9 zeigen die geschätzten L_1-Risiken für **Kuip .5**, **Kuip .9**, **Kuip .95**, **Kuip .99** sowie **CvM V**, **CvM .5**, **CvM .9**, **CvM .95** und **CvM .99**. Es fällt auf, dass die Ergebnisse für die auf der Kuiper-Metrik basierenden Methoden **Kuip .5**, **Kuip .9** und **Kuip .95** etwas besser sind als die für **KS .5**, **KS .9** und **KS .95**. Zum Teil schneidet **Kuip .99** sogar wesentlich besser ab als **KS .99**. Die Verwendung des Cramér-von Mises-Tests führt hingegen zu schlechten Ergebnissen – mit wenigen Ausnahmen sind die Ergebnisse für **CvM .5**, **CvM .9**, **CvM .95** und **CvM .99** schlechter als die für **KS .5**, **KS .9**, **KS .95** und **KS .99**, zum Teil sogar erheblich schlechter. Eine Ausnahme bildet **CvM V**: Diese Methode liefert sehr ähnliche Ergebnisse wie **V** und ist teilweise sogar besser. Ein ähnliches Bild ergibt sich für den L_2- bzw. L_∞-Verlust (vgl. Tabellen B.10-B.12 im Anhang). Ein Vergleich der gewählten Bandbreiten ergibt, dass die Verwendung der Kuiper-Metrik im Vergleich zur Kolmogorov-Metrik zu kleineren Werten führt. Die Verwendung des Cramér-von Mises-Tests führt in vielen Fällen, jedoch nicht durchgehend, zu größeren Bandbreiten, vor allem für größere n und extremere Quantile (vgl. Tabellen 3.10 und 3.11).

Die Tabellen 3.12-3.15 zeigen die L_1-Risiken bzw. die gewählten Bandbreiten für die restlichen untersuchten Varianten des Diskrepanzprinzips. Im Anhang finden sich entsprechende Tabellen für die L_2- und L_∞-Risiken (Tabellen B.13-B.15). Die auf dem Gesetz vom iterierten Logarithmus basierenden Methoden schneiden generell schlecht ab, besonders die auf der Kuiper-Metrik und dem Cramér-von-Mises-Test basierenden Versionen **kLIL**, **kLIL 2.1** und **CvM LIL**. Die auf dem Kolmogorov-Smirnov-Test basierende Version **LIL** wählt Bandbreiten, die ungefähr zwischen denen für **KS .5** und **KS .9** liegen, während die durch **LIL 2.1** gewählten Bandbreiten größer sind als die für **KS .99**. Beides ist auch nach Tabelle 3.1 zu erwarten. Es werden generell durch diese Methoden sehr große Bandbreiten gewählt, was nur für einige Dichten mit unendlichen Peaks (wie 8 und 14) einen gewissen

Vorteil hat. In der kleinen Simulationsstudie [Mar89] wird ebenfalls bemerkt, dass in den meisten Fällen **LIL** schlechter ist als **CvM V** bzw. die dort verwendete Variante von **V**.

Die auf der Normalverteilung als Referenzdichte basierenden Methoden **L2NR**, **L1NR**, **kL2NR** und **kL1NR** funktionieren schlecht. Die L_1-Methoden sind dabei schlechter als die L_2-Methoden, die auf $d_{kuip,k}$ basierenden Verfahren etwas besser als die für d_∞. Selbst **kL2NR** schneidet bezüglich des L_2-Risikos sogar für Normalverteilung durchgehend schlechter ab als **E-LR**, **V** und alle vier Referenzmethoden (teilweise mit Ausnahme von **SCV**). Generell wählen die auf der Normalverteilung basierenden Methoden sehr kleine Bandbreiten, was auch nach Tabelle 3.1 zu erwarten ist: Es wird ein extrem geringer Abstand zwischen der geschätzten und der empirischen Verteilungsfunktion erzwungen. Bemerkenswert ist, dass besonders für **L1NR** in einigen Fällen die Bandbreite für kleinere Werte von n extrem klein ist und dann zunächst ansteigt. Bei diesen Varianten können auch numerische Probleme auftreten, da die gesuchten Lösungen von $d(F_n, \hat{F}_n^h) = s(n)$ sehr nahe bei 0 liegen.

Insgesamt zeigt sich also, dass im Wesentlichen **E-LR**, **V** und **CvM V** auch im Vergleich zu den Referenzmethoden zu guten Ergebnissen führen. Dabei basiert **E-LR** auf einem Kolmogorov-Smirnov-Test mit von n abhängigem Niveau, während **V** auf einem Kolmogorov-Smirnov-Test zu einem Niveau von ca. 14% basiert. Beides lässt sich nicht anschaulich begründen. Intuitivere Wahlen wie **KS .95** oder **KS .99** schneiden hingegen schlecht ab, da zu stark geglättet wird, obwohl die Ergebnisse von Kapitel 3.2 zeigen, dass diese Methoden für hinreichend glatte Dichten asymptotisch zu wenig glätten.

Generell zeigen die Simulationen jedoch, dass die asymptotischen Resultate aus Kapitel 3.2 für die hier betrachteten Stichprobenumfänge (bis $n = 2500$) wenig relevant sind. Dies zeigt sich auch darin, dass die auf der Normalverteilung als Referenzdichte basierenden Methoden bei diesen Stichprobenumfängen nicht einmal für die Normalverteilung funktionieren. Umgekehrt zeigen die Simulationen aber auch, dass die gut funktionierenden Methoden **E-LR** und **V** bzw. **CvM V** auch für Dichten, die nicht glatt sind, meist akzeptable Ergebnisse liefern. Das Diskrepanzprinzip ist also in einigen Varianten durchaus auch praktisch zur Bandbreitenwahl für Kerndichteschätzer geeignet. Die Implementierung mit Hilfe der Regula Falsi ist zudem relativ einfach und das Verfahren liefert insbesondere bei größerem n teilweise erheblich schneller eine Bandbreite als z.B. die Kreuzvaliderung.

Tabelle 3.2.: Simulationsergebnisse für Dichteschätzung mit Gaußkern – L_1-Risiko für die Dichten 1-14.

Dichte	n	L2CV	BCV	PI	SCV	E-LR	V	KS .5	KS .9	KS .95	KS .99
1	50	0.2602	0.3382	0.2902	0.2478	**0.24**	0.2579	0.3681	0.5861	0.6666	0.8192
	100	0.2076	0.2784	0.2336	0.2047	**0.1968**	0.2046	0.2674	0.4217	0.4781	0.5821
	500	0.1291	0.1466	0.1384	0.1355	0.1214	**0.1192**	0.1342	0.185	0.2072	0.258
	1000	0.1001	0.1071	0.1084	0.1116	0.0955	**0.093**	0.1009	0.127	0.14	0.1682
	2500	0.0768	0.0741	0.0808	0.0886	0.0708	**0.0695**	0.0719	0.0854	0.092	0.1055
2	50	0.3564	0.3912	0.3609	0.3534	0.3344	**0.3311**	0.3605	0.4691	0.5184	0.6261
	100	0.2883	0.3132	0.2823	0.2884	0.2664	**0.264**	0.2712	0.3296	0.358	0.4212
	500	0.185	0.1649	0.1642	0.1958	0.1633	0.166	**0.1583**	0.1639	0.1698	0.1852
	1000	0.1563	0.1324	0.1317	0.1674	0.1334	0.139	0.1294	**0.1291**	0.131	0.1386
	2500	0.1188	0.0928	0.0948	0.1293	0.0979	0.1054	0.0959	**0.0915**	0.0921	0.0945
3	50	0.2172	0.2072	0.2094	0.258	0.2056	**0.2051**	0.2715	0.4804	0.5621	0.7151
	100	0.1689	**0.153**	0.1593	0.2127	0.1581	0.1591	0.1954	0.3204	0.3739	0.4758
	500	0.0871	**0.0786**	0.0822	0.1332	0.0819	0.0807	0.0934	0.1411	0.162	0.203
	1000	0.0651	**0.0608**	0.0632	0.1073	0.0631	0.0613	0.0686	0.098	0.1113	0.1387
	2500	0.0465	**0.0428**	0.0446	0.0829	0.0451	0.0429	0.0467	0.063	0.0705	0.0859
4	50	0.2655	**0.251**	0.2544	0.3285	0.2708	0.2592	0.3072	0.5043	0.584	0.7394
	100	0.2136	**0.2074**	0.2086	0.2756	0.2092	0.2085	0.2396	0.3614	0.4146	0.5183
	500	0.1214	**0.1138**	0.1145	0.1796	0.1149	0.1142	0.1253	0.1724	0.1942	0.2371
	1000	0.0927	0.0875	**0.0869**	0.1483	0.0878	**0.0869**	0.0931	0.1249	0.14	0.1699
	2500	0.0675	0.0622	**0.0615**	0.1167	0.0631	0.0618	0.0646	0.0831	0.092	0.1103
5	50	0.2297	**0.2074**	0.2148	0.2896	0.2226	0.2152	0.2721	0.4798	0.5647	0.727
	100	0.1799	**0.1602**	0.1678	0.2389	0.1729	0.1724	0.2088	0.3409	0.3964	0.5038
	500	0.0909	**0.0826**	0.086	0.1475	0.086	0.085	0.0984	0.1535	0.177	0.2224
	1000	0.0685	**0.0644**	0.067	0.1182	0.0677	0.0656	0.0749	0.1134	0.13	0.1618
	2500	0.049	**0.046**	0.048	0.0903	0.0499	0.0466	0.0525	0.0756	0.0857	0.1054
6	50	0.4105	0.3885	**0.3867**	0.4747	0.4016	0.3948	0.4398	0.6072	0.6733	0.8002
	100	0.3338	0.3238	**0.3217**	0.4105	0.3264	0.3236	0.354	0.4671	0.5148	0.6062
	500	0.2095	0.2004	**0.1998**	0.2889	0.2002	0.1999	0.2101	0.2567	0.2772	0.3162
	1000	0.166	0.1591	**0.159**	0.2462	0.1608	0.1594	0.1673	0.2011	0.2155	0.2436
	2500	0.124	0.124	**0.1234**	0.2002	0.1252	0.1235	0.127	0.1445	0.1526	0.1691
7	50	0.2387	0.2284	0.2297	0.2831	0.228	**0.2267**	0.2828	0.4746	0.553	0.7006
	100	0.1866	**0.1711**	0.177	0.2338	0.1774	0.1778	0.2108	0.3266	0.3751	0.4707
	500	0.0977	**0.0886**	0.0938	0.1465	0.0931	0.0919	0.1046	0.1499	0.1693	0.2077
	1000	0.0727	**0.0685**	0.0723	0.1179	0.0721	0.0696	0.0784	0.1083	0.1211	0.1463
	2500	0.0514	**0.0483**	0.0515	0.0904	0.0522	0.0488	0.0543	0.0722	0.0798	0.0948
8	50	0.4644	0.4741	0.4127	**0.3769**	0.4223	0.3975	0.3869	0.4803	0.5441	0.6424
	100	0.4341	0.4216	0.3488	**0.312**	0.3708	0.354	0.3136	0.3433	0.371	0.4403
	500	0.3863	0.3241	0.2411	0.2169	0.2862	0.3038	0.2458	0.2132	**0.2109**	0.2157
	1000	0.3771	0.2861	0.202	0.1814	0.2594	0.2933	0.2284	0.1854	0.1791	**0.1762**
	2500	0.3723	0.2541	0.167	0.1458	0.2291	0.2817	0.2146	0.1648	0.1558	**0.1452**
9	50	0.7549	0.787	0.7589	**0.7259**	0.7355	0.729	0.7265	0.7854	0.8121	0.8818
	100	0.6674	0.6865	0.6367	0.6256	0.6345	0.6279	**0.6139**	0.6344	0.6491	0.6854
	500	0.5347	0.5126	0.4552	0.4791	0.4831	0.4925	0.4598	0.4412	**0.4407**	0.4484
	1000	0.4893	0.4163	0.3872	0.4313	0.433	0.4485	0.4185	0.3889	0.383	**0.3814**
	2500	0.4627	0.3662	**0.3233**	0.4085	0.3987	0.4253	0.3902	0.3542	0.3468	0.3373
10	50	0.7114	0.6914	0.6825	0.7198	0.692	**0.6759**	0.6964	0.8089	0.8633	0.9728
	100	0.6457	0.6216	**0.6039**	0.6501	0.6174	0.611	0.6115	0.6798	0.7116	0.7772
	500	0.4882	0.4394	**0.434**	0.5336	0.4402	0.4434	0.4353	0.4485	0.4596	0.4841
	1000	0.4179	0.3839	**0.3696**	0.468	0.3752	0.3818	0.3701	0.3781	0.3845	0.3992
	2500	0.3425	0.3155	**0.2966**	0.405	0.2992	0.3081	0.2977	0.2994	0.3028	0.311
11	50	0.2181	**0.1955**	0.2055	0.2664	0.2093	0.2078	0.2763	0.4877	0.5716	0.7293
	100	0.1587	**0.1454**	0.1565	0.2113	0.1536	0.1557	0.2039	0.3483	0.4071	0.5165
	500	0.0851	**0.0775**	0.0823	0.1323	0.0816	0.0803	0.0958	0.1556	0.1805	0.2278
	1000	0.063	**0.0586**	0.0631	0.1068	0.0634	0.0602	0.0726	0.1137	0.1308	0.1635
	2500	0.0462	**0.0425**	0.0463	0.0806	0.0479	0.0431	0.0509	0.0762	0.0867	0.1068
12	50	0.368	0.3924	0.3656	0.3771	0.3536	**0.3513**	0.3831	0.5044	0.5593	0.6697
	100	0.2816	0.294	0.2784	0.3123	0.2643	**0.264**	0.2837	0.3554	0.389	0.4585
	500	0.1669	0.156	0.1584	0.2138	**0.1542**	0.1548	0.1568	0.1759	0.1854	0.2056
	1000	0.1306	0.1216	0.1229	0.181	**0.1208**	0.1218	0.1213	0.1313	0.1368	0.1501
	2500	0.0957	0.0892	0.0892	0.1444	**0.0879**	0.089	0.088	0.0931	0.0961	0.1024
13	50	0.4662	0.4858	0.458	0.4898	**0.4448**	0.4452	0.5069	0.6861	0.7537	0.8579
	100	0.3798	0.3871	0.3705	0.4036	**0.3594**	0.3597	0.3951	0.5146	0.5636	0.6629
	500	0.24	0.2382	0.2315	0.2604	0.2206	**0.22**	0.2294	0.2695	0.2886	0.3272
	1000	0.1993	0.1904	0.1907	0.2184	0.1802	**0.1797**	0.1844	0.2096	0.2216	0.2469
	2500	0.1519	0.1379	0.1439	0.171	**0.1351**	0.1356	0.1361	0.1495	0.1563	0.1704
14	50	1.4583	**0.6331**	0.6532	0.7381	1.4649	1.4187	1.063	0.7082	0.7055	0.746
	100	1.2251	**0.5767**	0.587	0.6724	1.41	1.5061	1.4356	0.7985	0.666	0.6202
	500	0.922	0.4873	**0.4733**	0.5355	1.0524	1.0023	1.1656	1.5574	1.6356	1.4465
	1000	0.9159	0.4629	**0.4359**	0.4887	0.9159	0.9159	0.9382	1.1267	1.3747	1.6487
	2500	0.9226	0.4325	**0.3909**	0.4324	0.9226	0.9226	0.9226	0.9226	0.9443	1.3606

Tabelle 3.3.: Simulationsergebnisse für Dichteschätzung mit Gaußkern – L_1-Risiko für die Dichten 15-28.

Dichte	n	L2CV	BCV	PI	SCV	E-LR	V	KS .5	KS .9	KS .95	KS .99
15	50	0.3676	0.3993	0.373	0.3471	0.3478	**0.3427**	0.3625	0.4625	0.5136	0.6211
	100	0.304	0.3427	0.3032	0.284	0.2803	**0.2754**	0.278	0.336	0.3652	0.4276
	500	0.1966	0.2061	0.1818	0.1818	0.1692	0.1731	**0.1628**	0.1689	0.175	0.189
	1000	0.1739	0.1485	0.1465	0.1532	0.1412	0.1487	0.1352	**0.132**	0.1352	0.1425
	2500	0.139	0.1037	0.1113	0.1206	0.1087	0.1192	0.1059	0.0988	**0.0987**	0.1005
16	50	0.2006	**0.181**	0.1904	0.2434	0.1944	0.197	0.2712	0.4837	0.5681	0.7294
	100	0.1547	**0.1428**	0.1505	0.2036	0.1485	0.1516	0.1963	0.3435	0.4016	0.5117
	500	0.0815	**0.0745**	0.0776	0.121	0.0787	0.0768	0.0932	0.1498	0.1749	0.2234
	1000	0.0624	**0.0585**	0.0611	0.1012	0.063	0.06	0.0703	0.1062	0.1219	0.1542
	2500	0.045	**0.042**	0.0432	0.0779	0.0459	0.0425	0.0482	0.0687	0.0782	0.0969
17	50	0.2062	0.1935	0.2002	0.2352	**0.1896**	0.1967	0.2822	0.497	0.5802	0.7389
	100	0.1544	0.1461	0.1496	0.1938	**0.1445**	0.1487	0.2064	0.3566	0.4131	0.5196
	500	0.0763	**0.0737**	0.0777	0.1172	0.0788	0.0753	0.099	0.1655	0.1918	0.2404
	1000	0.0604	**0.0569**	0.0599	0.0963	0.0628	0.0584	0.0719	0.1161	0.1349	0.1694
	2500	0.0432	**0.0408**	0.0426	0.0737	0.0456	0.0416	0.048	0.0704	0.0811	0.1014
18	50	0.6455	0.547	**0.5211**	0.5288	0.6148	0.573	0.522	0.5354	0.5647	0.6441
	100	0.6203	0.5015	0.4527	0.4535	0.5532	0.5283	0.4578	**0.4437**	0.4566	0.4941
	500	0.602	0.3825	0.3076	0.312	0.458	0.4863	0.3876	0.3146	0.3035	**0.2939**
	1000	0.5971	0.342	0.2591	0.2674	0.4193	0.4701	0.3698	0.2894	0.2735	**0.2575**
	2500	0.5853	0.2793	**0.2012**	0.2159	0.3802	0.4626	0.3565	0.268	0.2493	0.2251
19	50	1.1929	**0.7587**	0.7722	0.8501	1.0661	0.9752	0.8167	0.7794	0.8097	0.892
	100	1.2594	**0.6658**	0.6809	0.7802	1.0357	0.9813	0.7891	0.6742	0.674	0.7113
	500	1.4268	0.5154	**0.5147**	0.6133	0.9675	1.0276	0.794	0.6019	0.5655	0.5244
	1000	1.5123	0.4658	**0.4549**	0.5436	0.9467	1.0607	0.8238	0.608	0.5619	0.5041
	2500	1.5753	0.4132	**0.3827**	0.4643	0.94	1.107	0.8828	0.6319	0.576	0.5023
20	50	0.777	0.8316	0.7764	**0.733**	0.7518	0.739	0.742	0.8187	0.8554	0.9283
	100	0.69	0.7434	0.6851	**0.6513**	0.664	0.6559	0.655	0.68	0.6957	0.7312
	500	0.5742	0.5645	0.4901	0.5231	0.525	0.5321	0.506	0.4853	**0.4835**	0.4839
	1000	0.5423	0.4867	0.4312	0.4913	0.4751	0.4951	0.4565	0.4324	0.4289	**0.4253**
	2500	0.4479	0.4114	**0.3479**	0.4209	0.3927	0.4115	0.3858	0.3591	0.355	0.3513
21	50	**0.3497**	0.7172	0.5198	0.357	0.3645	0.3967	0.5639	0.987	1.1132	1.349
	100	0.2692	0.5929	0.4021	**0.2623**	0.284	0.2964	0.3909	0.6689	0.7818	0.9836
	500	**0.144**	0.1959	0.2075	0.1575	0.1574	0.1513	0.1826	0.2536	0.2857	0.3545
	1000	**0.1128**	0.1141	0.1533	0.1357	0.1241	0.1167	0.1364	0.1816	0.2004	0.2392
	2500	0.0747	**0.073**	0.0962	0.106	0.0852	0.0771	0.089	0.116	0.1268	0.1482
22	50	0.252	0.2675	0.2517	0.2568	**0.2324**	0.2363	0.3015	0.4798	0.556	0.7046
	100	0.1993	0.2294	0.2073	0.2116	**0.1873**	0.1917	0.2363	0.3471	0.3942	0.4899
	500	**0.1078**	0.1229	0.1198	0.1418	0.1143	0.11	0.1326	0.1812	0.1995	0.2329
	1000	0.0813	**0.0789**	0.09	0.1192	0.0883	0.0822	0.0982	0.1336	0.1473	0.1728
	2500	0.0569	**0.0555**	0.0619	0.0946	0.0636	0.0574	0.0667	0.0893	0.0981	0.1149
23	50	0.4178	0.4179	0.4174	**0.395**	0.4017	0.41	0.4613	0.5977	0.6566	0.7822
	100	0.3689	0.392	0.3907	**0.3333**	0.3704	0.3763	0.4106	0.4933	0.5278	0.5981
	500	0.175	0.3492	0.3448	**0.1715**	0.3096	0.2906	0.3424	0.3765	0.389	0.4134
	1000	0.1359	0.3367	0.3219	**0.1338**	0.2737	0.2354	0.3089	0.3482	0.3571	0.3737
	2500	**0.0933**	0.3183	0.2571	0.1021	0.2147	0.1642	0.2331	0.3058	0.3184	0.3341
24	50	**0.4387**	0.7967	0.6777	0.4518	0.5024	0.5684	0.7445	0.9365	0.9843	1.0719
	100	**0.3522**	0.7652	0.5847	0.359	0.4345	0.4629	0.6131	0.8169	0.8617	0.9315
	500	0.2107	0.6585	0.3701	**0.2064**	0.3161	0.2961	0.3865	0.5335	0.5796	0.6484
	1000	0.1729	0.5651	0.2978	**0.1669**	0.2652	0.2351	0.3073	0.4276	0.4675	0.5359
	2500	0.1254	0.2609	0.2211	**0.1218**	0.2029	0.1696	0.217	0.3081	0.3395	0.3936
25	50	0.4293	0.4519	0.4525	**0.3909**	0.4019	0.4215	0.5217	0.7048	0.7718	0.9004
	100	0.3547	0.4161	0.4051	**0.3341**	0.3408	0.3458	0.4251	0.5789	0.6263	0.709
	500	0.2156	0.3443	0.2798	**0.2031**	0.2071	0.2045	0.2221	0.2721	0.2949	0.4024
	1000	0.1766	0.3179	0.2233	0.1673	0.1662	**0.1635**	0.1731	0.2062	0.2208	0.2508
	2500	0.1359	0.2808	0.163	0.1289	0.1239	**0.122**	0.1256	0.143	0.1511	0.168
26	50	0.6031	1.5452	1.4592	1.2896	**0.6011**	0.6388	0.8176	1.3902	1.5359	1.7975
	100	**0.4806**	1.5385	1.432	1.1924	0.4836	0.499	0.6045	0.8909	1.0886	1.3891
	500	0.3029	1.4208	1.167	0.8244	0.2943	**0.2878**	0.3255	0.4186	0.4555	0.527
	1000	0.2512	1.4139	1.0886	0.6944	0.2391	**0.231**	0.2531	0.3131	0.3382	0.3869
	2500	0.1931	1.3348	0.9182	0.5269	0.1758	**0.1685**	0.1801	0.2143	0.2289	0.2586
27	50	0.5847	0.6614	0.6361	**0.5565**	0.5779	0.597	0.6874	0.8273	0.8775	0.9783
	100	**0.5016**	0.6277	0.5974	0.528	0.5531	0.5615	0.616	0.7278	0.7616	0.824
	500	**0.2083**	0.5467	0.5366	0.3248	0.5116	0.5004	0.5279	0.5628	0.5793	0.6159
	1000	**0.1608**	0.53	0.5256	0.2132	0.4899	0.4468	0.5109	0.5345	0.5436	0.5637
	2500	**0.1132**	0.5171	0.508	0.1223	0.4083	0.3088	0.4378	0.5065	0.514	0.5243
28	50	0.3596	0.5616	0.4486	**0.3215**	0.3336	0.3507	0.4462	0.646	0.7014	0.831
	100	0.2905	0.5183	0.367	**0.2642**	0.2708	0.2738	0.3276	0.4907	0.548	0.634
	500	0.186	0.3384	0.2267	0.1673	0.1666	**0.1657**	0.1749	0.2148	0.2348	0.2796
	1000	0.161	0.1978	0.181	0.1369	**0.1329**	0.1343	0.1359	0.1568	0.1666	0.1903
	2500	0.1303	0.1127	0.1396	0.1066	0.1011	0.1046	**0.1009**	0.1084	0.114	0.1262

Tabelle 3.4.: Simulationsergebnisse für Dichteschätzung mit Gaußkern – L_2-Risiko für die Dichten 1-13.

Dichte	n	L2CV	BCV	PI	SCV	E-LR	V	KS .5	KS .9	KS .95	KS .99
1	50	0.0906	0.1049	0.0891	0.0909	**0.0749**	0.0793	0.113	0.185	0.2173	0.2873
	100	0.0662	0.084	0.0687	0.067	**0.056**	0.0584	0.0795	0.1277	0.1451	0.1816
	500	0.0291	0.0386	0.0358	0.0301	0.0283	**0.0271**	0.034	0.0527	0.0601	0.0769
	1000	**0.0185**	0.0257	0.0263	0.0207	0.0203	**0.0185**	0.023	0.0335	0.0382	0.0478
	2500	**0.0115**	0.0149	0.0182	0.0134	0.0133	0.0117	0.0141	0.0201	0.0226	0.0276
2	50	0.0692	0.0906	0.0774	0.0607	**0.0585**	0.061	0.0779	0.121	0.1384	0.1736
	100	0.0473	0.0708	0.0577	0.0416	**0.0413**	0.0422	0.0518	0.0785	0.0894	0.1118
	500	0.0193	0.028	0.0291	0.0189	0.0184	**0.018**	0.0208	0.0289	0.0324	0.0399
	1000	0.0137	0.0195	0.022	0.014	0.0134	**0.013**	0.0146	0.0196	0.0215	0.0262
	2500	0.0081	0.0106	0.015	0.0084	0.0082	**0.0077**	0.0085	0.0113	0.0124	0.0148
3	50	0.0229	**0.0176**	0.0181	0.0361	0.0195	0.0178	0.0273	0.0668	0.0852	0.123
	100	0.0154	**0.0106**	0.0114	0.0251	0.0123	0.0119	0.0157	0.0347	0.0443	0.0649
	500	0.0039	**0.0031**	0.0033	0.0095	0.0033	0.0032	0.0042	0.0085	0.0107	0.0156
	1000	0.0022	**0.0018**	0.002	0.006	0.002	0.0019	0.0024	0.0046	0.0057	0.0082
	2500	0.0011	**9e-04**	0.001	0.0036	0.0011	**9e-04**	0.0011	0.002	0.0025	0.0035
4	50	0.0166	**0.0157**	0.0161	0.0233	0.0168	0.0164	0.0255	0.0553	0.0677	0.0928
	100	0.0113	0.0113	0.0115	0.0163	**0.0108**	0.0112	0.0166	0.0338	0.0415	0.0572
	500	**0.0035**	**0.0035**	0.0038	0.0066	0.0038	**0.0035**	0.0053	0.0104	0.0127	0.0175
	1000	**0.002**	**0.002**	0.0022	0.0044	0.0023	**0.002**	0.003	0.006	0.0074	0.0103
	2500	0.0011	**0.001**	0.0012	0.0028	0.0013	**0.001**	0.0015	0.003	0.0037	0.0052
5	50	0.0087	**0.0065**	0.007	0.0144	0.0078	0.007	0.0107	0.0268	0.0345	0.0509
	100	0.0058	**0.004**	0.0044	0.01	0.005	0.0048	0.0066	0.0152	0.0194	0.0286
	500	0.0014	**0.0011**	0.0012	0.0036	0.0012	0.0012	0.0016	0.0036	0.0046	0.0068
	1000	7e-04	**6e-04**	7e-04	0.0022	7e-04	7e-04	9e-04	0.002	0.0025	0.0038
	2500	4e-04	**3e-04**	4e-04	0.0013	4e-04	**3e-04**	4e-04	9e-04	0.0012	0.0017
6	50	0.0144	0.013	**0.0126**	0.0181	0.013	0.0132	0.0204	0.0428	0.0519	0.069
	100	0.0083	0.0087	0.0084	0.0124	**0.0079**	0.0081	0.0128	0.0268	0.0329	0.045
	500	0.0026	**0.0024**	0.0027	0.0048	0.0027	0.0025	0.0039	0.008	0.01	0.0139
	1000	0.0015	**0.0014**	0.0016	0.0032	0.0017	0.0015	0.0022	0.0046	0.0056	0.0079
	2500	7e-04	**6e-04**	8e-04	0.0019	9e-04	7e-04	0.001	0.002	0.0025	0.0035
7	50	0.0139	**0.0111**	0.0114	0.0215	0.0121	0.0112	0.017	0.0404	0.0512	0.0729
	100	0.009	**0.0066**	0.0071	0.0149	0.0076	0.0074	0.0099	0.0218	0.0274	0.0395
	500	0.0024	**0.0018**	0.002	0.0056	0.002	0.002	0.0026	0.0053	0.0067	0.0097
	1000	0.0013	**0.0011**	0.0012	0.0035	0.0012	**0.0011**	0.0015	0.0029	0.0036	0.0051
	2500	6e-04	**5e-04**	6e-04	0.0021	6e-04	6e-04	7e-04	0.0013	0.0016	0.0022
9	50	0.0343	0.0603	0.0499	0.0342	**0.031**	0.0329	0.0417	0.0595	0.0663	0.0785
	100	0.0234	0.0551	0.0414	0.0241	**0.0223**	0.0229	0.0284	0.0412	0.0462	0.0555
	500	**0.01**	0.037	0.0256	0.0106	0.0103	**0.01**	0.0119	0.0167	0.0186	0.0227
	1000	**0.0067**	0.0235	0.0201	0.007	0.007	**0.0067**	0.0078	0.0109	0.0121	0.0146
	2500	0.0042	0.008	0.0146	0.0043	0.0045	**0.0041**	0.0047	0.0063	0.007	0.0085
10	50	0.0123	0.0177	0.015	**0.0114**	0.0119	0.0129	0.0189	0.0312	0.0357	0.0428
	100	0.0084	0.0149	0.0115	**0.0075**	0.0084	0.009	0.0133	0.0229	0.0262	0.0319
	500	**0.0028**	0.0046	0.0051	**0.0028**	0.0037	0.0033	0.005	0.0088	0.0103	0.0131
	1000	**0.0016**	0.0022	0.0033	0.0018	0.0025	0.0021	0.0032	0.0056	0.0065	0.0084
	2500	**8e-04**	0.001	0.0019	0.0011	0.0015	0.0011	0.0017	0.003	0.0036	0.0046
11	50	0.016	**0.0105**	0.0113	0.0239	0.0138	0.0121	0.0172	0.0429	0.0554	0.0814
	100	0.0083	**0.0058**	0.0065	0.0147	0.007	0.0068	0.0098	0.024	0.0312	0.0464
	500	0.0022	**0.0017**	0.0019	0.0056	0.0019	0.0018	0.0024	0.0055	0.0071	0.0107
	1000	0.0012	**0.001**	0.0011	0.0036	0.0011	**0.001**	0.0014	0.003	0.0038	0.0057
	2500	7e-04	**5e-04**	6e-04	0.002	6e-04	6e-04	7e-04	0.0014	0.0017	0.0025
12	50	0.0402	0.0514	0.0424	0.0371	**0.0354**	0.0362	0.0487	0.0846	0.0997	0.129
	100	0.0226	0.032	0.027	0.0248	**0.0202**	0.0209	0.0289	0.0501	0.0592	0.0776
	500	0.0071	0.008	0.0097	0.0105	0.0074	**0.007**	0.0092	0.0151	0.0177	0.023
	1000	**0.0042**	0.0044	0.006	0.0075	0.0046	**0.0042**	0.0053	0.0085	0.01	0.0132
	2500	**0.0022**	**0.0022**	0.0032	0.0047	0.0025	**0.0022**	0.0027	0.0043	0.0049	0.0064
13	50	0.0538	0.0667	0.0574	0.0543	**0.0498**	0.0524	0.0728	0.1292	0.1509	0.1892
	100	0.0381	0.0461	0.0413	0.0377	**0.0358**	0.0369	0.0478	0.0798	0.0943	0.1238
	500	**0.017**	0.0241	0.0227	**0.017**	0.0187	0.0179	0.0221	0.0311	0.0347	0.0423
	1000	**0.0118**	0.0175	0.018	0.0122	0.0143	0.0131	0.0162	0.0224	0.0247	0.0293
	2500	**0.0074**	0.0108	0.013	0.0077	0.0098	0.0084	0.0104	0.0144	0.0159	0.0187

Tabelle 3.5.: Simulationsergebnisse für Dichteschätzung mit Gaußkern – L_2-Risiko für die Dichten 15-28.

Dichte	n	L2CV	BCV	PI	SCV	E-LR	V	KS .5	KS .9	KS .95	KS .99
15	50	0.4127	0.5303	0.481	**0.3605**	0.3767	0.3872	0.4605	0.6218	0.6853	0.8025
	100	0.3058	0.471	0.4012	**0.269**	0.2782	0.2831	0.3373	0.4605	0.5065	0.5925
	500	0.1344	0.2879	0.2449	**0.1294**	0.1366	0.1319	0.1553	0.206	0.2261	0.2637
	1000	0.0995	0.1967	0.2004	**0.0952**	0.1019	0.0972	0.1112	0.146	0.1617	0.188
	2500	0.0664	0.1208	0.1574	**0.0642**	0.0705	0.0648	0.0732	0.0949	0.1036	0.1202
16	50	0.0374	**0.0232**	0.0249	0.0538	0.0301	0.027	0.0407	0.1007	0.1294	0.1893
	100	0.0197	**0.015**	0.0161	0.0372	0.0165	0.0164	0.0237	0.0574	0.0738	0.1088
	500	0.0058	**0.0044**	0.0048	0.013	0.0049	0.0048	0.0067	0.0147	0.0188	0.0276
	1000	0.0033	**0.0028**	0.003	0.0089	0.0032	0.0029	0.0041	0.0084	0.0105	0.0153
	2500	0.0018	**0.0015**	0.0016	0.0054	0.0019	**0.0015**	0.0021	0.0042	0.0052	0.0074
17	50	0.0678	**0.0444**	0.0472	0.0994	0.0555	0.0491	0.0694	0.1674	0.2156	0.3172
	100	0.0389	**0.0268**	0.0277	0.0639	0.0291	0.0288	0.0415	0.0952	0.1214	0.1781
	500	0.0088	**0.0078**	0.0084	0.023	0.0086	0.0081	0.012	0.0258	0.0325	0.0468
	1000	0.0058	**0.0049**	0.0053	0.0157	0.0058	0.0052	0.0071	0.0145	0.0182	0.026
	2500	0.003	**0.0026**	0.0028	0.0091	0.0031	0.0027	0.0035	0.0064	0.008	0.0113
20	50	0.0454	0.0973	0.0803	0.0532	**0.0441**	0.0471	0.0618	0.0932	0.1044	0.1233
	100	**0.0301**	0.0908	0.0674	0.036	0.0311	0.0324	0.0425	0.0646	0.0727	0.0877
	500	**0.0118**	0.0702	0.0395	0.0125	0.0126	0.012	0.0151	0.0232	0.0268	0.0332
	1000	0.0076	0.0525	0.0296	**0.0075**	0.008	**0.0075**	0.0091	0.014	0.0161	0.0202
	2500	0.0039	0.0226	0.0197	**0.0038**	0.0047	0.004	0.005	0.0075	0.0086	0.0108
21	50	0.0329	0.0993	0.068	**0.0324**	0.0373	0.0446	0.0744	0.138	0.157	0.188
	100	0.0196	0.0798	0.0483	**0.0176**	0.0241	0.027	0.0466	0.0906	0.1074	0.1372
	500	**0.0057**	0.0143	0.0164	0.0064	0.0081	0.0072	0.0124	0.0255	0.0313	0.0431
	1000	**0.0034**	0.0039	0.009	0.0047	0.005	0.0041	0.0067	0.014	0.0173	0.0241
	2500	0.0015	**0.0014**	0.0036	0.0029	0.0025	0.0018	0.0029	0.006	0.0074	0.0105
22	50	0.0203	0.0193	0.0177	0.0213	0.0166	**0.0164**	0.0226	0.0413	0.0503	0.07
	100	0.0134	0.0155	0.0132	0.0146	**0.0115**	0.0118	0.0162	0.027	0.0317	0.0421
	500	**0.0037**	0.005	0.0048	0.0062	0.0043	0.0039	0.006	0.0104	0.0121	0.0153
	1000	0.0022	**0.002**	0.0028	0.0044	0.0027	0.0022	0.0034	0.0062	0.0074	0.0096
	2500	**0.001**	**0.001**	0.0013	0.0028	0.0014	0.0011	0.0016	0.0029	0.0035	0.0047
23	50	0.0656	0.0599	**0.0598**	0.0612	0.0607	**0.0598**	0.0654	0.0903	0.1028	0.1296
	100	0.0532	0.0547	0.0546	**0.0444**	0.0525	0.0532	0.0566	0.0691	0.0754	0.0898
	500	0.012	0.047	0.0463	**0.0117**	0.0389	0.0345	0.0459	0.0503	0.0518	0.0551
	1000	0.0073	0.0452	0.0423	**0.007**	0.0315	0.0235	0.0394	0.0468	0.0479	0.0498
	2500	**0.0034**	0.042	0.0283	0.0039	0.0198	0.0114	0.0234	0.0393	0.042	0.0448
24	50	0.0588	0.1136	0.0943	**0.0574**	0.0651	0.0754	0.1049	0.1322	0.1384	0.1504
	100	**0.0403**	0.1074	0.0769	**0.0403**	0.0525	0.0569	0.0814	0.1151	0.1217	0.1311
	500	0.0162	0.0875	0.041	**0.0161**	0.0335	0.0307	0.0432	0.0654	0.0732	0.0856
	1000	**0.0111**	0.0717	0.0311	**0.0111**	0.0268	0.0227	0.0324	0.0488	0.0548	0.0654
	2500	**0.0058**	0.0277	0.0214	0.0059	0.0192	0.0148	0.0209	0.0323	0.0363	0.0437
25	50	0.1901	0.2125	0.212	**0.1634**	0.174	0.1894	0.228	0.2824	0.3078	0.3616
	100	0.1299	0.2024	0.1958	**0.1216**	0.1322	0.14	0.1941	0.2411	0.2548	0.2832
	500	0.0564	0.1776	0.1304	**0.0555**	0.0665	0.0632	0.0804	0.1239	0.1429	0.1916
	1000	0.0403	0.1653	0.096	**0.0401**	0.0512	0.0465	0.0582	0.0831	0.0941	0.1173
	2500	0.0255	0.1465	0.0656	**0.0254**	0.0354	0.03	0.0377	0.0524	0.0579	0.0687
26	50	**0.3527**	1.2861	1.2557	1.1613	0.3992	0.4822	0.8403	1.2587	1.3021	1.3428
	100	**0.2369**	1.2798	1.2333	1.09	0.2818	0.3087	0.5022	0.9687	1.1154	1.2626
	500	**0.1054**	1.2385	1.0943	0.7726	0.1398	0.131	0.1758	0.2949	0.3507	0.4687
	1000	**0.0751**	1.2309	1.0311	0.6916	0.1069	0.0947	0.1242	0.1868	0.2153	0.2777
	2500	**0.0476**	1.2002	0.8874	0.5543	0.0741	0.0607	0.0798	0.115	0.1282	0.1548
27	50	0.0213	0.0218	0.0213	0.0211	**0.0208**	**0.0208**	0.0221	0.0255	0.0271	0.0304
	100	**0.0173**	0.0208	0.0201	0.0191	0.0194	0.0195	0.0204	0.0227	0.0236	0.0253
	500	**0.0035**	0.0184	0.0181	0.0077	0.0172	0.0167	0.0179	0.0188	0.0192	0.02
	1000	**0.0021**	0.0178	0.0177	0.0035	0.0159	0.0135	0.0171	0.0179	0.0182	0.0187
	2500	**0.0011**	0.0173	0.0169	0.0012	0.0114	0.0068	0.013	0.0168	0.0171	0.0175
28	50	0.2669	0.4146	0.3488	**0.2383**	0.2509	0.2697	0.345	0.4564	0.4836	0.537
	100	0.1982	0.3892	0.294	**0.1822**	0.1962	0.2031	0.2598	0.3722	0.403	0.4487
	500	**0.0982**	0.2759	0.1951	0.0998	0.1145	0.1092	0.1341	0.1833	0.2015	0.2382
	1000	**0.0743**	0.1681	0.162	0.0755	0.0868	0.08	0.0984	0.1333	0.1453	0.1709
	2500	**0.0517**	0.0906	0.1296	0.0539	0.0638	0.0561	0.067	0.0893	0.0982	0.1141

Tabelle 3.6.: Simulationsergebnisse für Dichteschätzung mit Gaußkern – gewählte Bandbreiten für die Dichten 1-14.

Dichte	n	L2CV	BCV	PI	SCV	E-LR	V	KS .5	KS .9	KS .95	KS .99
1	50	0.109	0.1976	0.1746	0.066	0.1205	0.1512	0.2317	0.3688	0.4216	0.5296
	100	0.0809	0.1722	0.1446	0.0484	0.1051	0.1165	0.1724	0.2671	0.3015	0.3668
	500	0.0363	0.0946	0.0896	0.0228	0.0661	0.0605	0.0854	0.1278	0.1431	0.1715
	1000	0.0261	0.0681	0.0719	0.0162	0.0509	0.0433	0.061	0.091	0.1018	0.122
	2500	0.0168	0.0412	0.0531	0.0103	0.0366	0.0285	0.0398	0.0589	0.0658	0.0786
2	50	0.1682	0.3817	0.3103	0.123	0.1675	0.2068	0.323	0.5611	0.6637	0.8876
	100	0.1231	0.3085	0.2467	0.0867	0.1319	0.1455	0.2192	0.3582	0.4135	0.5261
	500	0.0523	0.1347	0.1459	0.0389	0.0732	0.067	0.0951	0.1452	0.164	0.2
	1000	0.0363	0.0954	0.1166	0.0277	0.0567	0.0483	0.0679	0.1021	0.1147	0.1386
	2500	0.023	0.0576	0.0866	0.0175	0.0389	0.0302	0.0423	0.0633	0.071	0.0853
3	50	0.294	0.3609	0.367	0.1388	0.2699	0.3373	0.5194	0.8327	0.9567	1.2135
	100	0.2459	0.3058	0.3184	0.1063	0.2557	0.2829	0.4113	0.6108	0.6841	0.8258
	500	0.1743	0.2004	0.2274	0.0569	0.2152	0.1984	0.2669	0.3636	0.3957	0.4537
	1000	0.1524	0.1684	0.1973	0.0438	0.1951	0.1715	0.2237	0.2971	0.3212	0.364
	2500	0.1184	0.1343	0.1612	0.0303	0.165	0.1374	0.1748	0.2276	0.2447	0.2749
4	50	0.4782	0.5668	0.5864	0.2251	0.427	0.5573	0.9155	1.6045	1.8869	2.5222
	100	0.4065	0.491	0.5038	0.1704	0.4115	0.4615	0.7021	1.1348	1.303	1.6405
	500	0.2377	0.2965	0.328	0.083	0.3216	0.2924	0.4148	0.6046	0.6724	0.7965
	1000	0.1907	0.2334	0.2704	0.0604	0.2823	0.2406	0.3348	0.4791	0.5289	0.6195
	2500	0.1424	0.1756	0.2131	0.0403	0.2345	0.1862	0.2527	0.3553	0.3896	0.4505
5	50	0.7887	0.8907	0.9503	0.359	0.7046	0.8873	1.3879	2.2669	2.6324	3.4379
	100	0.6583	0.7618	0.8218	0.2741	0.6725	0.7477	1.1169	1.7059	1.9231	2.357
	500	0.4693	0.521	0.5943	0.149	0.5645	0.518	0.7118	0.9968	1.0926	1.2658
	1000	0.4168	0.4413	0.5196	0.1154	0.5177	0.449	0.6031	0.8268	0.9002	1.0301
	2500	0.339	0.3615	0.4339	0.0821	0.4535	0.3671	0.4844	0.649	0.7019	0.7941
6	50	0.641	0.8324	0.796	0.314	0.6217	0.7835	1.2648	2.2883	2.7502	3.8312
	100	0.5186	0.6823	0.6666	0.2319	0.5449	0.6097	0.948	1.5599	1.8118	2.3455
	500	0.3326	0.3821	0.4468	0.1163	0.4417	0.4049	0.5612	0.8115	0.903	1.0707
	1000	0.2854	0.3121	0.3768	0.0866	0.395	0.3422	0.4617	0.6447	0.7072	0.8222
	2500	0.2256	0.2448	0.3022	0.0586	0.3293	0.2659	0.3522	0.4783	0.5205	0.5963
7	50	0.5093	0.6212	0.6238	0.2378	0.4548	0.574	0.8877	1.4448	1.6703	2.1416
	100	0.4275	0.517	0.5379	0.1811	0.4316	0.4793	0.7005	1.0541	1.1826	1.4365
	500	0.2982	0.333	0.3838	0.0969	0.3586	0.3303	0.4482	0.6168	0.6727	0.7742
	1000	0.2614	0.2813	0.3342	0.0748	0.3277	0.2872	0.3764	0.5027	0.5441	0.618
	2500	0.211	0.2282	0.2764	0.0527	0.2803	0.232	0.2975	0.3892	0.4188	0.4711
8	50	0.0388	0.1714	0.1241	0.0496	0.0368	0.0485	0.0875	0.1812	0.2238	0.3162
	100	0.0174	0.1478	0.0969	0.0345	0.021	0.0244	0.0447	0.0932	0.1148	0.1606
	500	0.0029	0.0986	0.0529	0.0145	0.0055	0.0047	0.0089	0.0189	0.0233	0.0329
	1000	0.0014	0.0847	0.0404	0.0099	0.003	0.0023	0.0043	0.0092	0.0114	0.0161
	2500	5e-04	0.0709	0.0284	0.006	0.0015	9e-04	0.0017	0.0036	0.0045	0.0063
9	50	0.4642	2.3242	1.5241	0.63	0.5022	0.635	1.0746	2.2312	2.8526	4.5026
	100	0.2894	1.9463	1.1541	0.4281	0.3578	0.3985	0.6324	1.1648	1.4111	1.9735
	500	0.1081	1.0397	0.6138	0.1763	0.1686	0.1532	0.2248	0.3629	0.418	0.5291
	1000	0.0736	0.6202	0.4711	0.121	0.1223	0.1031	0.1489	0.2341	0.2672	0.3324
	2500	0.0458	0.1748	0.3356	0.0744	0.0826	0.0637	0.0902	0.1377	0.1557	0.1903
10	50	1.1085	2.6323	2.0373	0.8337	1.2236	1.5975	2.9842	6.8096	9.0618	14.9373
	100	0.7317	2.1682	1.5886	0.5788	0.995	1.1347	1.9211	3.9009	4.8194	6.9161
	500	0.3649	0.7635	0.8667	0.2421	0.6504	0.5881	0.8711	1.3949	1.6045	2.023
	1000	0.2836	0.4636	0.6781	0.1685	0.5423	0.4582	0.6559	0.9947	1.1203	1.3703
	2500	0.2079	0.297	0.4997	0.1062	0.4268	0.3357	0.4622	0.6708	0.7471	0.8915
11	50	0.4791	0.538	0.5829	0.2201	0.4401	0.5558	0.8529	1.3604	1.5651	1.9983
	100	0.4236	0.4676	0.5159	0.1715	0.4166	0.4646	0.6894	1.0397	1.1696	1.4201
	500	0.2944	0.3281	0.3751	0.0938	0.3505	0.3225	0.4405	0.6167	0.6754	0.781
	1000	0.2571	0.2803	0.3289	0.0729	0.3225	0.279	0.376	0.5126	0.5572	0.6364
	2500	0.2145	0.2303	0.2764	0.0525	0.2837	0.2303	0.3028	0.4037	0.4361	0.4927
12	50	0.2601	0.4759	0.3951	0.1576	0.2478	0.3086	0.4712	0.7936	0.9335	1.2433
	100	0.1828	0.3566	0.3138	0.111	0.2135	0.2346	0.3386	0.5254	0.5994	0.7507
	500	0.1061	0.1565	0.1915	0.051	0.1493	0.139	0.1839	0.2572	0.2837	0.3337
	1000	0.0856	0.1156	0.1552	0.0367	0.1243	0.11	0.1427	0.195	0.2135	0.2477
	2500	0.0651	0.0836	0.119	0.0239	0.0997	0.0833	0.1058	0.1409	0.153	0.1751
13	50	0.2269	0.4391	0.3566	0.1481	0.2438	0.3036	0.4905	1.0583	1.4211	2.5257
	100	0.1656	0.3102	0.2785	0.103	0.2075	0.2298	0.343	0.5695	0.6747	0.9477
	500	0.0735	0.1803	0.1702	0.048	0.1275	0.1166	0.1657	0.2499	0.2806	0.3381
	1000	0.0498	0.1291	0.1376	0.0344	0.0991	0.0842	0.1192	0.1791	0.2007	0.2409
	2500	0.0323	0.0796	0.1022	0.0217	0.0711	0.0551	0.0773	0.1152	0.1288	0.1542
14	50	1e-04	0.0156	0.0106	0.0045	2e-04	6e-04	0.0047	0.0305	0.051	0.0885
	100	0	0.0131	0.0073	0.0028	0	0	2e-04	0.0038	0.0091	0.024
	500	0	0.0092	0.0034	0.001	0	0	0	0	0	0
	1000	0	0.008	0.0026	7e-04	0	0	0	0	0	0
	2500	0	0.0068	0.0017	4e-04	0	0	0	0	0	0

Tabelle 3.7.: Simulationsergebnisse für Dichteschätzung mit Gaußkern – gewählte Bandbreiten für die Dichten 15-28.

Dichte	n	L2CV	BCV	PI	SCV	E-LR	V	KS .5	KS .9	KS .95	KS .99
15	50	0.0478	0.1122	0.0927	0.0364	0.0445	0.0554	0.0892	0.1573	0.1862	0.2482
	100	0.0289	0.0945	0.0734	0.0257	0.0313	0.0349	0.0551	0.0953	0.1115	0.1445
	500	0.0094	0.0497	0.0402	0.0107	0.0141	0.0127	0.0191	0.0313	0.036	0.0453
	1000	0.0058	0.03	0.0313	0.0075	0.0101	0.0084	0.0124	0.0201	0.023	0.0288
	2500	0.0034	0.0157	0.0226	0.0046	0.0065	0.0048	0.0071	0.0113	0.0129	0.0161
16	50	0.2032	0.2319	0.2502	0.0937	0.1936	0.2442	0.3732	0.5881	0.6715	0.8482
	100	0.1768	0.2019	0.2184	0.0722	0.1858	0.2077	0.3033	0.4544	0.5081	0.6118
	500	0.1246	0.1417	0.1603	0.0397	0.1584	0.1447	0.201	0.279	0.3044	0.3494
	1000	0.1046	0.1218	0.139	0.0304	0.1447	0.1233	0.1697	0.2316	0.2511	0.2852
	2500	0.0823	0.0968	0.1135	0.021	0.1273	0.1009	0.1366	0.1847	0.1996	0.2252
17	50	0.1122	0.1367	0.1409	0.0529	0.1063	0.1344	0.2023	0.3147	0.3584	0.4497
	100	0.0964	0.1194	0.1234	0.0409	0.1042	0.1153	0.1663	0.2438	0.2717	0.3253
	500	0.0689	0.0799	0.0893	0.0222	0.0898	0.0827	0.1113	0.151	0.1642	0.1878
	1000	0.0581	0.0665	0.0769	0.0169	0.0822	0.0721	0.0941	0.125	0.1351	0.153
	2500	0.0464	0.0523	0.0625	0.0116	0.0696	0.0579	0.0738	0.0964	0.1038	0.1167
18	50	0.0614	0.3408	0.2374	0.0974	0.0571	0.0757	0.143	0.3167	0.4024	0.6055
	100	0.0286	0.2873	0.1807	0.0663	0.0335	0.0389	0.0725	0.155	0.193	0.277
	500	0.0043	0.2021	0.0978	0.0278	0.0083	0.0071	0.0133	0.0286	0.0356	0.0506
	1000	0.0021	0.1743	0.0742	0.0189	0.0048	0.0036	0.0067	0.0142	0.0176	0.025
	2500	8e-04	0.14	0.0519	0.0115	0.0023	0.0014	0.0026	0.0056	0.007	0.0099
19	50	0.0217	0.2834	0.1917	0.0801	0.0305	0.0485	0.1394	0.5814	0.839	1.6488
	100	0.0065	0.2223	0.1304	0.0491	0.0155	0.0196	0.0586	0.197	0.293	0.5348
	500	4e-04	0.1542	0.0654	0.0192	0.0031	0.0024	0.0068	0.0226	0.032	0.057
	1000	1e-04	0.1356	0.0496	0.0132	0.0015	9e-04	0.0026	0.0085	0.0119	0.0212
	2500	0	0.1145	0.0346	0.008	5e-04	3e-04	7e-04	0.0024	0.0034	0.0057
20	50	0.2673	1.7455	1.1243	0.4676	0.3071	0.3914	0.6869	1.5215	1.9876	3.2563
	100	0.1792	1.4435	0.8293	0.3093	0.2346	0.2606	0.413	0.7717	0.9428	1.3418
	500	0.0758	0.9324	0.4237	0.1229	0.1221	0.1121	0.1578	0.2447	0.2796	0.3503
	1000	0.0553	0.6567	0.3183	0.0826	0.0942	0.0816	0.1114	0.1652	0.186	0.2269
	2500	0.0398	0.2758	0.2223	0.0498	0.07	0.0567	0.0751	0.1068	0.1185	0.1409
21	50	0.2665	1.3691	0.7834	0.3222	0.3836	0.4911	0.8919	2.1653	2.8124	4.6449
	100	0.2125	0.9776	0.5619	0.205	0.3098	0.3456	0.5629	1.1603	1.4536	2.1324
	500	0.1299	0.2445	0.3016	0.0847	0.2048	0.1887	0.2588	0.3835	0.435	0.5439
	1000	0.1096	0.1286	0.2339	0.0586	0.1736	0.1515	0.2021	0.2845	0.3154	0.3754
	2500	0.0896	0.0949	0.1701	0.0366	0.1453	0.1203	0.1545	0.2068	0.2249	0.2583
22	50	0.416	0.684	0.6182	0.2346	0.4227	0.5296	0.8219	1.3268	1.5263	1.9403
	100	0.3378	0.6013	0.5236	0.1755	0.3945	0.4366	0.6416	0.9838	1.1098	1.3539
	500	0.2056	0.3095	0.3235	0.0818	0.3	0.2777	0.3721	0.5221	0.5758	0.6758
	1000	0.1716	0.1966	0.2625	0.0588	0.255	0.2243	0.2929	0.3997	0.4371	0.5064
	2500	0.1365	0.1485	0.2008	0.0384	0.2083	0.173	0.2214	0.2937	0.318	0.3621
23	50	0.3597	0.5106	0.5109	0.1924	0.3434	0.4411	0.6931	1.1018	1.2617	1.5997
	100	0.254	0.4404	0.436	0.1444	0.2953	0.3376	0.5286	0.8115	0.9127	1.1081
	500	0.0604	0.3007	0.2826	0.0701	0.1924	0.1684	0.2754	0.4235	0.4723	0.5604
	1000	0.0512	0.2676	0.2236	0.0496	0.1583	0.1285	0.2011	0.3163	0.3534	0.4198
	2500	0.0417	0.226	0.1483	0.0291	0.1207	0.0935	0.132	0.2028	0.2281	0.2738
24	50	0.216	1.1006	0.7523	0.2978	0.3958	0.5242	0.9437	1.7616	2.0599	2.664
	100	0.1553	0.9877	0.5737	0.2046	0.3384	0.3829	0.6318	1.1765	1.3646	1.7272
	500	0.0658	0.7218	0.3004	0.0832	0.2419	0.2197	0.3196	0.5029	0.5733	0.7003
	1000	0.0456	0.571	0.2307	0.057	0.1989	0.1665	0.2417	0.3694	0.4169	0.5062
	2500	0.0291	0.2149	0.1643	0.0349	0.1474	0.1143	0.1608	0.2458	0.2763	0.3317
25	50	0.1389	0.2757	0.2659	0.1	0.1232	0.183	0.3644	0.6249	0.7178	0.9114
	100	0.0555	0.2364	0.2085	0.0692	0.0852	0.0969	0.2303	0.4629	0.5254	0.6343
	500	0.0216	0.1738	0.1063	0.0277	0.044	0.0397	0.0596	0.1005	0.121	0.2463
	1000	0.0149	0.152	0.0776	0.0183	0.0338	0.0282	0.0415	0.0663	0.0761	0.0965
	2500	0.0091	0.1268	0.0522	0.0107	0.0229	0.0173	0.0252	0.0395	0.0449	0.0552
26	50	0.0289	6.9799	6.106	2.4433	0.0504	0.0686	0.2767	2.1468	3.8227	10.9467
	100	0.0216	6.6603	4.9093	1.7699	0.0407	0.0456	0.0796	0.4075	0.8353	1.9135
	500	0.0098	5.2276	1.8833	0.5558	0.0241	0.022	0.0317	0.0497	0.0573	0.0743
	1000	0.0069	4.6808	1.291	0.3482	0.019	0.0161	0.023	0.0348	0.0392	0.0479
	2500	0.0044	3.6548	0.7437	0.1752	0.0139	0.0107	0.0151	0.0226	0.0253	0.0304
27	50	2.215	3.836	3.471	1.3124	2.099	2.711	4.3733	7.1095	8.1403	10.2479
	100	1.3752	3.3436	2.8594	0.9566	1.8337	2.0811	3.2126	5.0837	5.7581	7.0415
	500	0.2128	1.9159	1.7493	0.4465	1.0918	0.9606	1.4907	2.3322	2.6353	3.1979
	1000	0.1799	1.5124	1.3987	0.3175	0.8603	0.6857	1.0679	1.6538	1.8672	2.2656
	2500	0.1433	1.2199	1.0283	0.2003	0.6001	0.4516	0.6639	1.0169	1.1477	1.3959
28	50	0.0729	0.2413	0.1782	0.0689	0.0821	0.1056	0.1785	0.3265	0.3853	0.503
	100	0.0419	0.2174	0.1404	0.0486	0.0629	0.0709	0.1154	0.2061	0.2429	0.316
	500	0.0146	0.1329	0.0787	0.0211	0.0302	0.027	0.042	0.0716	0.0832	0.1064
	1000	0.0089	0.0662	0.0615	0.0147	0.0214	0.0176	0.0269	0.0452	0.0524	0.0665
	2500	0.0051	0.0262	0.0441	0.0091	0.0139	0.0102	0.0154	0.0253	0.0291	0.0366

Tabelle 3.8.: Simulationsergebnisse für Dichteschätzung mit Gaußkern (alternative GoF-Tests) – L_1-Risiko für die Dichten 1-14.

Dichte	n	Kuip .5	Kuip .9	Kuip .95	Kuip .99	CvM V	CvM .5	CvM .9	CvM .95	CvM .99
1	50	0.3627	0.4803	0.5184	0.5884	0.2568	0.3832	0.6123	0.6899	0.8534
	100	0.2467	0.3258	0.3536	0.4038	0.2132	0.3056	0.4772	0.532	0.6445
	500	0.1239	0.1419	0.1495	0.1663	0.1332	0.1736	0.2721	0.3063	0.3661
	1000	0.0949	0.104	0.1078	0.1167	0.1063	0.1348	0.2032	0.2296	0.2858
	2500	0.0696	0.0727	0.0745	0.0789	0.08	0.0991	0.1431	0.1593	0.1956
2	50	0.3827	0.4609	0.4888	0.5455	0.3428	0.4013	0.5706	0.6398	0.7898
	100	0.2818	0.3261	0.3439	0.3804	0.2704	0.31	0.4304	0.4792	0.5828
	500	0.1583	0.1645	0.1689	0.1792	0.1589	0.1735	0.227	0.249	0.2962
	1000	0.1282	0.1295	0.131	0.1362	0.1288	0.1379	0.177	0.1935	0.2266
	2500	0.0931	0.092	0.0926	0.0947	0.0922	0.099	0.122	0.1309	0.1525
3	50	0.2735	0.3867	0.426	0.4988	0.1997	0.2568	0.452	0.5262	0.6856
	100	0.1867	0.2541	0.2787	0.3277	0.1527	0.1874	0.3178	0.3694	0.4786
	500	0.0855	0.1078	0.1175	0.1373	0.0792	0.0907	0.1432	0.1651	0.2129
	1000	0.0642	0.0771	0.0831	0.0957	0.0609	0.0677	0.1015	0.1164	0.1494
	2500	0.0441	0.0516	0.0552	0.0624	0.0429	0.0469	0.0676	0.0766	0.0973
4	50	0.2972	0.3949	0.4307	0.4981	0.2521	0.2844	0.4411	0.5061	0.65
	100	0.2272	0.2801	0.3016	0.3442	0.2056	0.2288	0.3286	0.3711	0.466
	500	0.1163	0.1329	0.1407	0.157	0.1128	0.12	0.1604	0.1783	0.2186
	1000	0.0874	0.0969	0.1018	0.1121	0.0864	0.0898	0.1163	0.1285	0.1561
	2500	0.0616	0.0662	0.0687	0.0745	0.0613	0.063	0.0785	0.086	0.1036
5	50	0.2657	0.37	0.4073	0.4785	0.2104	0.2508	0.4297	0.5009	0.6542
	100	0.1924	0.2523	0.2764	0.3238	0.1643	0.1872	0.3045	0.3526	0.4568
	500	0.0876	0.1082	0.1173	0.1359	0.0831	0.0906	0.1378	0.1579	0.2019
	1000	0.0669	0.0797	0.0856	0.0981	0.0646	0.0691	0.1003	0.1142	0.1448
	2500	0.0471	0.0544	0.0579	0.0652	0.0463	0.0483	0.0669	0.0753	0.0943
6	50	0.4352	0.5212	0.5507	0.6075	0.3863	0.427	0.5784	0.638	0.7678
	100	0.3391	0.3926	0.4125	0.4513	0.3204	0.3458	0.4513	0.4936	0.5859
	500	0.2013	0.2179	0.2255	0.2408	0.1987	0.2075	0.2522	0.2708	0.3115
	1000	0.1602	0.1713	0.1766	0.1875	0.1585	0.1643	0.197	0.2107	0.2407
	2500	0.1231	0.1285	0.1311	0.1368	0.1229	0.1259	0.1436	0.1514	0.1694
7	50	0.2859	0.3918	0.4296	0.501	0.2222	0.272	0.4546	0.5262	0.6814
	100	0.2029	0.2652	0.2885	0.3338	0.1711	0.2034	0.3237	0.3722	0.4772
	500	0.0964	0.1181	0.1273	0.1458	0.0893	0.1019	0.1509	0.171	0.2152
	1000	0.0727	0.0873	0.0933	0.1055	0.0683	0.0766	0.1102	0.1241	0.1545
	2500	0.0505	0.0594	0.0632	0.0707	0.0483	0.0535	0.0745	0.0831	0.1021
8	50	0.4083	0.4925	0.5295	0.596	0.3834	0.4398	0.6462	0.7054	0.8546
	100	0.3127	0.3527	0.3709	0.4083	0.3103	0.3568	0.5179	0.5755	0.6644
	500	0.2321	0.2133	0.2119	0.2136	0.2121	0.221	0.2831	0.3091	0.3679
	1000	0.2122	0.1844	0.18	0.1767	0.1772	0.1795	0.2166	0.2335	0.2726
	2500	0.1953	0.1624	0.1561	0.1476	0.1426	0.1423	0.1639	0.1733	0.1995
9	50	0.7351	0.7783	0.7919	0.8196	0.7211	0.765	0.8704	0.9171	1.0369
	100	0.6145	0.6322	0.6403	0.6563	0.6135	0.6335	0.7143	0.7533	0.8374
	500	0.4532	0.4417	0.4404	0.4431	0.4433	0.4451	0.4758	0.4904	0.5284
	1000	0.4094	0.3877	0.3841	0.3826	0.3869	0.3826	0.3964	0.4053	0.4268
	2500	0.3752	0.3505	0.3457	0.3391	0.3434	0.3274	0.3252	0.3294	0.3435
10	50	0.6926	0.7377	0.7585	0.8003	0.6756	0.6955	0.8057	0.8553	0.9715
	100	0.6086	0.6303	0.6408	0.6639	0.6051	0.6134	0.6806	0.7121	0.784
	500	0.439	0.4349	0.4358	0.4401	0.4398	0.4342	0.4552	0.4681	0.4991
	1000	0.3759	0.3696	0.3698	0.3727	0.3769	0.3693	0.3841	0.3924	0.4132
	2500	0.3033	0.297	0.2966	0.2971	0.303	0.2966	0.3037	0.3084	0.3205
11	50	0.2675	0.3743	0.4119	0.4827	0.1993	0.2439	0.4266	0.4989	0.6554
	100	0.1841	0.2531	0.2796	0.33	0.147	0.1793	0.3065	0.3557	0.4613
	500	0.0837	0.1067	0.1168	0.1372	0.0778	0.0871	0.1385	0.1596	0.2047
	1000	0.0622	0.0778	0.0846	0.0982	0.0587	0.0651	0.1	0.1146	0.1459
	2500	0.0442	0.0532	0.0572	0.0653	0.0423	0.0461	0.0671	0.076	0.0956
12	50	0.4022	0.4808	0.5103	0.5698	0.3537	0.4105	0.583	0.6491	0.7929
	100	0.2887	0.3361	0.355	0.3927	0.27	0.3079	0.4282	0.4761	0.5779
	500	0.1574	0.1703	0.1761	0.1884	0.1554	0.1694	0.2145	0.2335	0.2747
	1000	0.1218	0.1287	0.132	0.1395	0.1213	0.13	0.1627	0.1765	0.2076
	2500	0.0883	0.0919	0.0939	0.0979	0.0884	0.0942	0.1137	0.1216	0.141
13	50	0.5037	0.5904	0.6228	0.6851	0.4538	0.5347	0.7419	0.8002	0.871
	100	0.3823	0.4376	0.4581	0.4981	0.3682	0.4255	0.579	0.6386	0.7578
	500	0.2222	0.2362	0.2425	0.2558	0.2302	0.2602	0.3324	0.3591	0.416
	1000	0.1799	0.1874	0.1912	0.1994	0.1894	0.2136	0.2686	0.2885	0.3302
	2500	0.1348	0.137	0.1388	0.143	0.1433	0.1614	0.2003	0.2141	0.2431
14	50	1.09	0.8228	0.7606	0.7162	1.1721	0.8077	0.6706	0.6874	0.7155
	100	1.5522	1.2288	1.0856	0.8935	1.4383	0.9876	0.5996	0.5953	0.6492
	500	1.0573	1.2415	1.2258	1.3772	1.4041	1.6554	0.9258	0.7232	0.5261
	1000	0.9159	0.9809	1.1108	1.0857	1.1563	1.6532	1.3228	1.0652	0.6756
	2500	0.9226	0.9226	0.9226	0.9226	0.9898	1.1319	1.6482	1.788	1.2767

Tabelle 3.9.: Simulationsergebnisse für Dichteschätzung mit Gaußkern (alternative GoF-Tests) – L_1-Risiko für die Dichten 15-28.

Dichte	n	Kuip .5	Kuip .9	Kuip .95	Kuip .99	CvM V	CvM .5	CvM .9	CvM .95	CvM .99
15	50	0.3897	0.4625	0.4924	0.5524	0.3477	0.4102	0.5786	0.6433	0.7889
	100	0.2897	0.3364	0.3561	0.3948	0.2806	0.3284	0.4599	0.5084	0.6063
	500	0.1631	0.1704	0.1747	0.1846	0.1668	0.186	0.2444	0.2691	0.3181
	1000	0.1324	0.1337	0.1364	0.1414	0.1339	0.1483	0.1894	0.2049	0.2403
	2500	0.1016	0.0987	0.0989	0.1007	0.0996	0.1095	0.1393	0.1507	0.1734
16	50	0.258	0.3618	0.4007	0.4748	0.1838	0.2311	0.4132	0.4863	0.6469
	100	0.1759	0.2421	0.2682	0.3187	0.1438	0.1708	0.2909	0.3397	0.4458
	500	0.0806	0.1017	0.1109	0.129	0.0745	0.0818	0.1239	0.143	0.1876
	1000	0.0618	0.0743	0.0798	0.0917	0.0585	0.0628	0.0879	0.0999	0.1279
	2500	0.0432	0.0499	0.0528	0.0592	0.0419	0.0437	0.0566	0.0628	0.0781
17	50	0.2755	0.3865	0.4246	0.4969	0.1864	0.2537	0.4513	0.5266	0.6896
	100	0.1854	0.2626	0.2898	0.3398	0.1406	0.1853	0.3307	0.3821	0.4908
	500	0.0823	0.1092	0.12	0.1419	0.0731	0.0894	0.1525	0.1775	0.2287
	1000	0.0611	0.0761	0.0831	0.0972	0.0569	0.0662	0.1075	0.125	0.1618
	2500	0.0425	0.0495	0.0527	0.0596	0.041	0.0456	0.0681	0.0785	0.1012
18	50	0.5181	0.5399	0.5552	0.5985	0.5228	0.5247	0.6542	0.7167	0.8499
	100	0.4462	0.4464	0.4551	0.4749	0.4451	0.4497	0.5374	0.5734	0.6632
	500	0.3623	0.3124	0.3049	0.2967	0.3083	0.2923	0.3248	0.3406	0.3862
	1000	0.3413	0.2875	0.2761	0.2613	0.2651	0.2492	0.2622	0.2714	0.3014
	2500	0.3253	0.2641	0.251	0.2318	0.2178	0.1938	0.1946	0.2002	0.2167
19	50	0.8167	0.7674	0.764	0.7767	0.8904	0.784	0.7791	0.808	0.8812
	100	0.8307	0.7234	0.7046	0.679	0.8501	0.728	0.6662	0.6766	0.7246
	500	0.9199	0.7429	0.7038	0.6449	0.7752	0.6314	0.5281	0.5158	0.5092
	1000	0.9667	0.7827	0.7379	0.6717	0.7528	0.6024	0.4873	0.4704	0.4536
	2500	1.0395	0.8543	0.8029	0.7207	0.7235	0.5672	0.4433	0.4224	0.3961
20	50	0.7518	0.8112	0.8328	0.8691	0.7397	0.7838	0.9237	0.9723	1.0823
	100	0.6582	0.6739	0.6842	0.7044	0.6542	0.6786	0.7554	0.7918	0.8741
	500	0.5015	0.4862	0.4839	0.4824	0.4919	0.4825	0.5028	0.5165	0.5486
	1000	0.4581	0.4354	0.4319	0.428	0.4388	0.4269	0.4368	0.4446	0.4664
	2500	0.3798	0.3599	0.3566	0.3534	0.3597	0.3505	0.3485	0.352	0.3632
21	50	0.5266	0.7342	0.8089	0.9444	0.3902	0.542	0.9042	1.0293	1.3117
	100	0.3528	0.4667	0.5127	0.6039	0.3041	0.4057	0.6703	0.7627	0.9493
	500	0.1633	0.1946	0.2059	0.2281	0.1587	0.2014	0.3136	0.36	0.4604
	1000	0.1229	0.1427	0.15	0.1641	0.1218	0.1506	0.2229	0.2528	0.3215
	2500	0.0797	0.0915	0.0958	0.1042	0.0797	0.0972	0.1406	0.1579	0.1969
22	50	0.3104	0.4118	0.448	0.5171	0.2338	0.3042	0.5012	0.5758	0.7347
	100	0.2328	0.2938	0.3159	0.3591	0.1919	0.2462	0.3784	0.429	0.5378
	500	0.1255	0.1578	0.1691	0.189	0.1121	0.1444	0.2148	0.2373	0.2809
	1000	0.0936	0.118	0.1267	0.1431	0.0849	0.11	0.167	0.1862	0.2225
	2500	0.0631	0.0778	0.0833	0.094	0.0595	0.0752	0.1139	0.128	0.1567
23	50	0.4564	0.5255	0.5506	0.5995	0.4051	0.4546	0.5912	0.6455	0.775
	100	0.4006	0.4425	0.4573	0.4863	0.3768	0.4126	0.5008	0.5339	0.6084
	500	0.3243	0.3509	0.3567	0.3673	0.2773	0.3484	0.3944	0.409	0.4385
	1000	0.2684	0.3193	0.3278	0.3387	0.213	0.3038	0.3639	0.3756	0.3983
	2500	0.1829	0.2443	0.2613	0.286	0.1427	0.2091	0.3197	0.3364	0.3576
24	50	0.7275	0.8668	0.9001	0.9527	0.5334	0.71	0.9608	1.0153	1.1189
	100	0.5697	0.7024	0.7428	0.8047	0.4387	0.5847	0.8277	0.8899	0.9833
	500	0.332	0.4161	0.4425	0.4897	0.28	0.3623	0.5133	0.5613	0.6508
	1000	0.2572	0.3241	0.3457	0.3845	0.2282	0.2918	0.4116	0.45	0.5256
	2500	0.1813	0.2245	0.2395	0.2689	0.1721	0.2198	0.3091	0.338	0.3946
25	50	0.5101	0.6124	0.6436	0.7025	0.4169	0.4932	0.6666	0.7258	0.854
	100	0.4	0.4879	0.5137	0.5611	0.364	0.4209	0.5523	0.5952	0.6826
	500	0.2107	0.2317	0.2399	0.2565	0.2333	0.2777	0.3625	0.3918	0.451
	1000	0.1654	0.1773	0.1825	0.193	0.1886	0.2254	0.2968	0.3208	0.3707
	2500	0.1222	0.1267	0.1292	0.1345	0.1419	0.1687	0.2221	0.2401	0.2764
26	50	0.7593	0.9929	1.1041	1.2972	0.6672	0.8548	1.3918	1.596	1.8927
	100	0.5607	0.6663	0.7042	0.782	0.5244	0.6716	1.0638	1.211	1.4699
	500	0.3001	0.3403	0.3555	0.3849	0.3143	0.3765	0.5222	0.5795	0.7111
	1000	0.2367	0.2593	0.2684	0.2873	0.2552	0.2989	0.3984	0.436	0.5193
	2500	0.1702	0.1824	0.1876	0.1981	0.1897	0.22	0.2846	0.3077	0.3573
27	50	0.6895	0.7666	0.7894	0.832	0.599	0.7061	0.8505	0.8999	1.0108
	100	0.6027	0.668	0.6886	0.722	0.5694	0.6499	0.77	0.8034	0.8736
	500	0.5182	0.5336	0.5389	0.5504	0.4734	0.5464	0.6279	0.6577	0.7026
	1000	0.4817	0.516	0.5202	0.5267	0.3504	0.5149	0.5814	0.6017	0.6448
	2500	0.3423	0.449	0.469	0.4913	0.2243	0.353	0.5281	0.545	0.576
28	50	0.4523	0.5786	0.6102	0.6601	0.3737	0.5266	0.7395	0.812	0.9714
	100	0.3148	0.3879	0.417	0.4775	0.3042	0.4145	0.6292	0.6754	0.7716
	500	0.1678	0.181	0.1868	0.199	0.1859	0.2323	0.3506	0.3989	0.5052
	1000	0.1329	0.1376	0.1403	0.1474	0.148	0.1778	0.2582	0.2909	0.3616
	2500	0.1024	0.1011	0.1017	0.1037	0.1105	0.1319	0.1803	0.2001	0.24

Tabelle 3.10.: Simulationsergebnisse für Dichteschätzung mit Gaußkern (alternative GoF-Tests) – gewählte Bandbreiten für die Dichten 1-14.

Dichte	n	Kuip .5	Kuip .9	Kuip .95	Kuip .99	CvM V	CvM .5	CvM .9	CvM .95	CvM .99
1	50	0.2278	0.3018	0.3257	0.3702	0.1455	0.2367	0.3877	0.4393	0.5579
	100	0.158	0.2091	0.2257	0.2562	0.1256	0.1928	0.3013	0.336	0.4084
	500	0.0717	0.0942	0.1016	0.1151	0.0833	0.1197	0.1783	0.1966	0.2323
	1000	0.05	0.066	0.0712	0.0807	0.0687	0.0971	0.1431	0.1575	0.1855
	2500	0.0314	0.0415	0.0448	0.0508	0.0519	0.0724	0.1059	0.1165	0.1371
2	50	0.3758	0.5429	0.6011	0.7187	0.2534	0.4208	0.773	0.9164	1.2669
	100	0.248	0.3505	0.3855	0.452	0.2041	0.3175	0.5396	0.6226	0.8112
	500	0.1057	0.1462	0.1598	0.1853	0.119	0.173	0.2708	0.3042	0.3742
	1000	0.0768	0.1055	0.115	0.1325	0.0957	0.136	0.2082	0.2324	0.2825
	2500	0.0498	0.0681	0.0741	0.0853	0.0714	0.0995	0.149	0.1654	0.1988
3	50	0.5237	0.6985	0.7552	0.8616	0.32	0.4931	0.7935	0.9036	1.1626
	100	0.3923	0.5134	0.5509	0.6209	0.2762	0.4013	0.6079	0.6782	0.8305
	500	0.2393	0.3039	0.3233	0.3579	0.1926	0.2614	0.3674	0.4007	0.467
	1000	0.2006	0.251	0.2661	0.2926	0.1659	0.2204	0.3037	0.3294	0.3801
	2500	0.1572	0.1952	0.2063	0.2258	0.1343	0.1757	0.2386	0.2577	0.2948
4	50	0.8736	1.2339	1.3539	1.5818	0.5322	0.8212	1.3851	1.6041	2.1313
	100	0.6444	0.8693	0.9436	1.0826	0.4511	0.6579	1.0316	1.1659	1.4646
	500	0.3497	0.4583	0.4925	0.5541	0.2844	0.3896	0.5663	0.6244	0.7438
	1000	0.2788	0.3607	0.3861	0.4315	0.2324	0.3133	0.4474	0.4907	0.5781
	2500	0.2083	0.2675	0.2856	0.3179	0.1821	0.2412	0.3378	0.3685	0.4294
5	50	1.3633	1.8193	1.9701	2.2594	0.8394	1.2796	2.0595	2.352	3.0477
	100	1.0244	1.348	1.4515	1.6405	0.7134	1.03	1.5652	1.7505	2.156
	500	0.6101	0.7811	0.8318	0.9219	0.4871	0.6602	0.9311	1.0169	1.1893
	1000	0.5096	0.6431	0.6828	0.7532	0.4181	0.5547	0.7666	0.8325	0.963
	2500	0.4062	0.5066	0.5363	0.5883	0.3379	0.4414	0.6003	0.6489	0.7437
6	50	1.2365	1.7528	1.9293	2.2873	0.7606	1.1764	2.0881	2.4779	3.4923
	100	0.8548	1.1754	1.2809	1.4817	0.6193	0.9092	1.4791	1.6987	2.2155
	500	0.4783	0.6168	0.6601	0.7385	0.4032	0.5456	0.7911	0.8741	1.0487
	1000	0.3864	0.4919	0.5245	0.5826	0.3352	0.4448	0.6282	0.6883	0.8113
	2500	0.2942	0.3696	0.3923	0.4327	0.2636	0.3436	0.4733	0.5145	0.5971
7	50	0.8993	1.2162	1.3213	1.5224	0.5497	0.8486	1.391	1.5943	2.0775
	100	0.6675	0.8807	0.9485	1.0735	0.4703	0.6834	1.0482	1.1752	1.4538
	500	0.3983	0.5105	0.5443	0.6047	0.3237	0.4388	0.6201	0.6776	0.7938
	1000	0.3346	0.4213	0.4471	0.4927	0.2776	0.3684	0.5097	0.5537	0.6411
	2500	0.2643	0.3297	0.3491	0.383	0.2244	0.2932	0.3989	0.4313	0.4946
8	50	0.1149	0.1892	0.2157	0.27	0.0783	0.1504	0.3184	0.3813	0.517
	100	0.0571	0.0972	0.1126	0.1427	0.0542	0.1026	0.2111	0.2526	0.3416
	500	0.0113	0.0199	0.0232	0.0298	0.021	0.0382	0.078	0.0934	0.1274
	1000	0.0055	0.0096	0.0112	0.0146	0.0132	0.0242	0.0495	0.0593	0.0811
	2500	0.0022	0.0038	0.0045	0.0058	0.0072	0.0129	0.0266	0.032	0.0439
9	50	1.2826	2.1075	2.451	3.1721	0.9292	1.7208	4.2863	5.7323	10.6233
	100	0.7198	1.1108	1.2548	1.5496	0.6594	1.1324	2.3845	2.973	4.5776
	500	0.2552	0.363	0.4009	0.474	0.3228	0.492	0.8588	1.0021	1.3317
	1000	0.172	0.2416	0.2655	0.3107	0.2413	0.358	0.5963	0.6854	0.8839
	2500	0.1057	0.1456	0.159	0.1843	0.1658	0.2394	0.3833	0.4349	0.5458
10	50	2.8064	4.4434	5.1116	6.5306	1.7122	2.9521	6.6081	8.5705	14.8637
	100	1.6677	2.5385	2.867	3.5185	1.2504	2.0232	3.9556	4.8156	7.0656
	500	0.71	0.9661	1.0532	1.2199	0.6518	0.9409	1.5347	1.759	2.2659
	1000	0.5292	0.7065	0.7646	0.8713	0.5091	0.7136	1.1117	1.256	1.5713
	2500	0.3739	0.4855	0.5215	0.5878	0.3746	0.5111	0.7637	0.8514	1.0366
11	50	0.8313	1.0973	1.1839	1.3477	0.509	0.767	1.2196	1.3872	1.7826
	100	0.6306	0.8248	0.8869	0.9997	0.4373	0.6296	0.9471	1.0554	1.2902
	500	0.3816	0.4854	0.5164	0.5715	0.3036	0.4102	0.5753	0.627	0.7303
	1000	0.3183	0.4015	0.4263	0.4695	0.2602	0.3452	0.4757	0.5159	0.595
	2500	0.2536	0.3163	0.3347	0.3669	0.2114	0.2759	0.3743	0.4042	0.4623
12	50	0.5253	0.7357	0.8111	0.9647	0.3424	0.5472	0.9974	1.1892	1.6757
	100	0.3498	0.4788	0.522	0.6048	0.2717	0.4098	0.6848	0.7907	1.0383
	500	0.1848	0.2399	0.2578	0.2908	0.1716	0.2369	0.353	0.3928	0.4771
	1000	0.1437	0.1843	0.1971	0.2205	0.1409	0.1904	0.276	0.3046	0.3638
	2500	0.1073	0.1352	0.144	0.1597	0.1102	0.1458	0.2055	0.225	0.2643
13	50	0.4819	0.6989	0.7929	1.0237	0.3428	0.5578	1.2474	1.6116	2.5417
	100	0.3167	0.4248	0.4619	0.5353	0.2706	0.4029	0.7116	0.8618	1.2834
	500	0.1399	0.1845	0.199	0.2258	0.1659	0.2326	0.3461	0.3846	0.4679
	1000	0.0977	0.1295	0.1398	0.1588	0.1344	0.1863	0.2728	0.3006	0.3574
	2500	0.0619	0.0819	0.0884	0.1004	0.101	0.1385	0.201	0.2207	0.2599
14	50	0.0039	0.0131	0.0179	0.03	0.0017	0.0073	0.0324	0.0444	0.0728
	100	1e-04	6e-04	0.0011	0.0025	2e-04	0.0012	0.0104	0.0163	0.033
	500	0	0	0	0	0	0	2e-04	4e-04	0.0015
	1000	0	0	0	0	0	0	0	0	2e-04
	2500	0	0	0	0	0.0123	0	0	0	0

Tabelle 3.11.: Simulationsergebnisse für Dichteschätzung mit Gaußkern (alternative GoF-Tests) – gewählte Bandbreiten für die Dichten 15-28.

Dichte	n	Kuip .5	Kuip .9	Kuip .95	Kuip .99	CvM V	CvM .5	CvM .9	CvM .95	CvM .99
15	50	0.1078	0.1568	0.1743	0.2082	0.0717	0.122	0.2228	0.2618	0.3539
	100	0.0644	0.0957	0.1066	0.1274	0.0554	0.0909	0.1598	0.1845	0.239
	500	0.0223	0.0324	0.036	0.0429	0.0284	0.0436	0.0729	0.0831	0.1047
	1000	0.0147	0.0213	0.0236	0.0279	0.0215	0.0324	0.0532	0.0604	0.0755
	2500	0.0087	0.0125	0.0138	0.0162	0.0149	0.022	0.0355	0.0401	0.0497
16	50	0.3596	0.4712	0.5086	0.579	0.2176	0.3297	0.5207	0.5904	0.7538
	100	0.2744	0.3574	0.3837	0.4317	0.187	0.2704	0.4059	0.4515	0.5494
	500	0.1705	0.2176	0.2315	0.2558	0.1314	0.1781	0.25	0.2724	0.3168
	1000	0.1413	0.1803	0.1917	0.2117	0.1113	0.149	0.2063	0.2239	0.2582
	2500	0.1127	0.1426	0.1513	0.1664	0.0897	0.118	0.1611	0.1743	0.1997
17	50	0.1989	0.2586	0.2781	0.3149	0.1212	0.1852	0.2926	0.3311	0.4209
	100	0.1539	0.1973	0.211	0.2358	0.1064	0.1547	0.2314	0.2567	0.3109
	500	0.0954	0.1193	0.1264	0.139	0.0766	0.1037	0.1446	0.1572	0.1821
	1000	0.0797	0.0983	0.1039	0.1137	0.0658	0.0874	0.12	0.1299	0.1493
	2500	0.0622	0.076	0.0801	0.0874	0.0535	0.0699	0.0946	0.1021	0.1165
18	50	0.195	0.3257	0.3752	0.4862	0.1388	0.2711	0.6313	0.795	1.2102
	100	0.0946	0.1621	0.1871	0.2368	0.0932	0.1758	0.3763	0.4622	0.6728
	500	0.0167	0.03	0.035	0.0448	0.0322	0.0593	0.1251	0.1513	0.2109
	1000	0.0085	0.0148	0.0172	0.0223	0.0204	0.0372	0.0777	0.0938	0.1304
	2500	0.0034	0.0059	0.0069	0.009	0.0113	0.0204	0.0419	0.0504	0.0697
19	50	0.1319	0.2963	0.3749	0.561	0.0705	0.1803	0.6048	0.8311	1.4732
	100	0.0428	0.0982	0.1219	0.1754	0.0356	0.0858	0.2631	0.3555	0.6132
	500	0.0037	0.009	0.0112	0.0163	0.0073	0.0174	0.051	0.0673	0.1092
	1000	0.0014	0.0031	0.0039	0.0057	0.0036	0.0086	0.0256	0.0337	0.0545
	2500	3e-04	8e-04	0.001	0.0015	0.0014	0.0035	0.0103	0.0136	0.0219
20	50	0.8479	1.4498	1.6837	2.2402	0.6077	1.1887	3.2074	4.3831	8.2731
	100	0.4609	0.7197	0.822	1.0392	0.4204	0.7411	1.6538	2.101	3.3582
	500	0.1667	0.2303	0.2526	0.2971	0.1961	0.2987	0.5339	0.6291	0.8542
	1000	0.1171	0.1588	0.1731	0.1998	0.1466	0.2153	0.3631	0.42	0.5487
	2500	0.0813	0.1066	0.1149	0.1304	0.1057	0.1488	0.2345	0.2658	0.3343
21	50	0.8063	1.3061	1.5157	1.9739	0.4695	0.8316	1.8156	2.3193	4.223
	100	0.4813	0.7184	0.8122	1.0079	0.3622	0.5878	1.1448	1.3786	1.9714
	500	0.219	0.2817	0.3021	0.3411	0.2084	0.294	0.4786	0.5523	0.7215
	1000	0.1714	0.2152	0.2291	0.2546	0.1691	0.2302	0.3504	0.3957	0.4987
	2500	0.1303	0.1603	0.1694	0.1859	0.1307	0.1723	0.2467	0.2726	0.3281
22	50	0.8508	1.1549	1.2503	1.4287	0.5146	0.8323	1.3885	1.582	2.0286
	100	0.6239	0.8314	0.8986	1.0217	0.4414	0.676	1.0746	1.2042	1.4766
	500	0.3448	0.4504	0.4842	0.543	0.2901	0.4099	0.6213	0.6893	0.8226
	1000	0.2754	0.3545	0.3796	0.425	0.2408	0.331	0.4904	0.5428	0.6472
	2500	0.2062	0.2597	0.2764	0.3068	0.1883	0.2512	0.3594	0.3951	0.4678
23	50	0.679	0.899	0.9716	1.1081	0.4197	0.6733	1.086	1.2337	1.5794
	100	0.4866	0.6506	0.7003	0.7914	0.3382	0.5392	0.8355	0.9318	1.1373
	500	0.2185	0.3116	0.3381	0.3849	0.1538	0.2981	0.4926	0.5449	0.6447
	1000	0.1527	0.2192	0.2402	0.2771	0.1139	0.1928	0.381	0.4272	0.5106
	2500	0.1032	0.1392	0.152	0.1754	0.0826	0.1174	0.2314	0.2817	0.3626
24	50	0.8968	1.3995	1.5657	1.8576	0.4571	0.843	1.9061	2.2633	3.0064
	100	0.5514	0.8263	0.9336	1.1303	0.3493	0.5726	1.2109	1.4923	2.0484
	500	0.2595	0.3533	0.3846	0.4433	0.2007	0.2925	0.4734	0.5419	0.7028
	1000	0.1904	0.2587	0.2807	0.3213	0.158	0.2254	0.3508	0.3951	0.491
	2500	0.1262	0.1679	0.1819	0.2091	0.1168	0.1633	0.2465	0.2745	0.3328
25	50	0.3519	0.4988	0.5414	0.6215	0.1814	0.3343	0.5734	0.6535	0.8374
	100	0.1925	0.3393	0.376	0.44	0.1289	0.239	0.4296	0.4851	0.5987
	500	0.0489	0.0679	0.0745	0.0876	0.069	0.1052	0.1966	0.2349	0.3081
	1000	0.033	0.0453	0.0494	0.0573	0.0539	0.0792	0.1324	0.155	0.2098
	2500	0.0194	0.0265	0.0289	0.0334	0.0388	0.0557	0.087	0.0982	0.1236
26	50	0.1699	0.6094	0.8986	1.5876	0.0857	0.3083	1.859	4.3628	17.0524
	100	0.0636	0.1103	0.1385	0.2168	0.0537	0.112	0.7196	1.0839	2.2899
	500	0.026	0.0347	0.0377	0.0432	0.0293	0.0417	0.0732	0.0905	0.1589
	1000	0.0183	0.0245	0.0265	0.0302	0.0234	0.0323	0.0501	0.0574	0.0768
	2500	0.0118	0.0157	0.017	0.0193	0.0174	0.0237	0.0346	0.0384	0.0466
27	50	4.3593	5.8525	6.3283	7.2153	2.7834	4.6309	7.6119	8.6331	10.9809
	100	2.9606	4.0041	4.3362	4.9469	2.2579	3.7235	5.9305	6.6284	8.0791
	500	1.2082	1.6728	1.8178	2.0852	0.7686	1.9857	3.3898	3.7861	4.544
	1000	0.8096	1.1547	1.2567	1.4446	0.4986	1.1411	2.5694	2.91	3.5405
	2500	0.4951	0.6925	0.7585	0.8718	0.352	0.5093	1.4981	1.8598	2.4307
28	50	0.1849	0.2646	0.2917	0.3429	0.1279	0.2269	0.4241	0.4884	0.6296
	100	0.1073	0.1537	0.1699	0.2013	0.0979	0.1671	0.3097	0.3587	0.4539
	500	0.0336	0.0477	0.0526	0.0619	0.0518	0.0812	0.1407	0.1627	0.2104
	1000	0.021	0.0297	0.0327	0.0384	0.0387	0.0595	0.1008	0.1157	0.1478
	2500	0.0117	0.0164	0.018	0.0211	0.0266	0.0398	0.0657	0.0749	0.0942

Tabelle 3.12.: Simulationsergebnisse für Dichteschätzung mit Gaußkern (weitere Varianten) – L_1-Risiko für die Dichten 1-14.

Dichte	n	LIL	LIL 2.1	L2NR	L1NR	kLIL	kLIL 2.1	kL2NR	kL1NR	CvM LIL
1	50	0.3659	0.8835	0.6119	1.7675	0.4896	1.021	0.2686	0.4147	0.5525
	100	0.2834	0.666	0.3712	1.1764	0.3531	0.7476	0.2149	0.3507	0.4545
	500	0.1476	0.3335	0.1436	0.303	0.1603	0.3589	0.1252	0.1975	0.2792
	1000	0.1093	0.2268	0.1088	0.1895	0.1157	0.2458	0.1	0.1442	0.2134
	2500	0.0773	0.1393	0.0807	0.1148	0.0796	0.1457	0.0768	0.1015	0.1533
2	50	0.3597	0.6775	0.8087	1.7557	0.4676	0.9754	0.3995	0.5749	0.5217
	100	0.275	0.4805	0.5151	1.3086	0.3436	0.6989	0.3214	0.4896	0.411
	500	0.158	0.2177	0.2402	0.4501	0.1755	0.3241	0.1998	0.2991	0.2309
	1000	0.1281	0.1577	0.191	0.3042	0.1355	0.2356	0.1635	0.2352	0.1837
	2500	0.0922	0.1051	0.1417	0.2027	0.0951	0.1483	0.1194	0.1721	0.1275
3	50	0.2697	0.7794	0.6745	1.7622	0.3962	0.9498	0.2859	0.4635	0.3967
	100	0.2069	0.5585	0.4292	1.2316	0.2782	0.6845	0.2262	0.4067	0.2975
	500	0.1059	0.2646	0.1541	0.3637	0.1306	0.3248	0.1154	0.228	0.1474
	1000	0.0782	0.1873	0.0977	0.2265	0.0944	0.234	0.0794	0.164	0.1073
	2500	0.0533	0.1187	0.058	0.1271	0.0637	0.1515	0.0518	0.1009	0.0732
4	50	0.3056	0.8036	0.7999	1.7439	0.4034	0.9368	0.3481	0.5583	0.3939
	100	0.2499	0.6029	0.5032	1.3166	0.3012	0.6813	0.2813	0.4765	0.312
	500	0.1374	0.3032	0.1927	0.4556	0.1514	0.3283	0.1527	0.2814	0.1638
	1000	0.1034	0.2239	0.1317	0.2875	0.111	0.2382	0.1127	0.2063	0.121
	2500	0.0718	0.1492	0.0849	0.1688	0.0755	0.1578	0.0766	0.136	0.0831
5	50	0.2703	0.7937	0.7118	1.757	0.379	0.9297	0.3088	0.4994	0.3772
	100	0.2207	0.592	0.4664	1.2538	0.276	0.6747	0.2539	0.4426	0.2853
	500	0.113	0.2914	0.1742	0.4031	0.1296	0.3188	0.1304	0.2543	0.1416
	1000	0.0878	0.2168	0.1111	0.2543	0.0968	0.2328	0.0906	0.1834	0.1057
	2500	0.0621	0.1447	0.0664	0.1442	0.0664	0.1534	0.0588	0.1146	0.0721
6	50	0.4384	0.8543	0.8864	1.7771	0.5283	0.9791	0.4931	0.6885	0.5346
	100	0.3644	0.681	0.641	1.3538	0.4122	0.7542	0.4173	0.614	0.4346
	500	0.2227	0.3756	0.3046	0.5754	0.2356	0.3966	0.2502	0.4091	0.2557
	1000	0.1788	0.2931	0.22	0.4112	0.1863	0.3059	0.1939	0.3155	0.2023
	2500	0.1341	0.2038	0.152	0.2677	0.1378	0.2112	0.1408	0.2235	0.1484
7	50	0.2812	0.7631	0.7075	1.7537	0.4009	0.9456	0.3123	0.4933	0.4016
	100	0.2215	0.5495	0.4605	1.2501	0.2881	0.684	0.2474	0.4363	0.3046
	500	0.117	0.2664	0.1692	0.3907	0.1396	0.3259	0.1272	0.2482	0.1547
	1000	0.0887	0.1915	0.1076	0.2473	0.1042	0.2355	0.0882	0.1797	0.1156
	2500	0.062	0.126	0.0647	0.1394	0.072	0.156	0.0577	0.111	0.0799
8	50	0.3865	0.6838	1.0043	1.8032	0.5014	1.0085	0.5233	0.764	0.5993
	100	0.3118	0.5064	0.8556	1.5444	0.3706	0.735	0.5077	0.7962	0.4928
	500	0.2292	0.2353	0.6586	1.1908	0.212	0.3513	0.4317	0.787	0.2882
	1000	0.2055	0.1807	0.6159	1.1512	0.1769	0.2411	0.4016	0.7708	0.2225
	2500	0.1848	0.1405	0.5589	1.0638	0.1465	0.1643	0.3633	0.7367	0.1698
9	50	0.7261	0.9119	1.0883	1.7529	0.7815	1.1246	0.7874	0.9307	0.834
	100	0.6136	0.7268	0.9054	1.4539	0.6402	0.8756	0.7113	0.8825	0.6985
	500	0.45	0.4625	0.6238	0.8364	0.4412	0.5233	0.5576	0.6948	0.4786
	1000	0.4054	0.3859	0.5563	0.6948	0.3826	0.4174	0.4994	0.6096	0.3998
	2500	0.3703	0.3257	0.5111	0.6244	0.3381	0.3315	0.4592	0.5742	0.3276
10	50	0.6958	1.0202	1.0928	1.7156	0.7428	1.1262	0.7676	0.9484	0.7696
	100	0.6163	0.8381	0.9323	1.4371	0.6406	0.8948	0.7131	0.9079	0.6686
	500	0.436	0.5279	0.5867	0.8963	0.4383	0.5443	0.5303	0.6871	0.4576
	1000	0.37	0.4317	0.4732	0.7044	0.3723	0.441	0.4428	0.5746	0.3873
	2500	0.2968	0.3316	0.3676	0.5178	0.2973	0.3361	0.3515	0.4576	0.3066
11	50	0.2744	0.7952	0.7136	1.7595	0.3835	0.9332	0.2846	0.4823	0.373
	100	0.2174	0.6057	0.4338	1.2446	0.2791	0.6864	0.2293	0.414	0.2867
	500	0.1123	0.298	0.1493	0.384	0.1304	0.3248	0.1135	0.2345	0.1426
	1000	0.0871	0.2192	0.0993	0.235	0.0968	0.2355	0.0812	0.1662	0.1057
	2500	0.0618	0.1467	0.0603	0.1335	0.0667	0.1554	0.0539	0.1051	0.0726
12	50	0.3822	0.7185	0.8425	1.6473	0.4876	0.9749	0.4174	0.5962	0.5336
	100	0.2895	0.518	0.5631	1.3171	0.3546	0.703	0.3302	0.5374	0.4092
	500	0.1615	0.2407	0.2275	0.4991	0.1843	0.3287	0.1868	0.3218	0.218
	1000	0.124	0.177	0.1611	0.3195	0.1387	0.2438	0.1421	0.2361	0.1681
	2500	0.0896	0.1182	0.1074	0.1896	0.0986	0.1603	0.0988	0.1546	0.1185
13	50	0.5053	0.8879	0.932	1.7833	0.598	0.9505	0.5024	0.7061	0.6878
	100	0.4063	0.7457	0.6624	1.3547	0.4577	0.8104	0.4134	0.6345	0.5558
	500	0.2398	0.3881	0.2765	0.5816	0.2512	0.4108	0.2475	0.361	0.3376
	1000	0.1925	0.2949	0.2189	0.3727	0.1985	0.3092	0.2024	0.2821	0.2765
	2500	0.1412	0.202	0.1612	0.2307	0.1438	0.2091	0.1528	0.202	0.2089
14	50	1.0722	0.7725	1.0649	0.9271	0.8048	0.8677	1.29	1.1624	0.6682
	100	1.3757	0.6727	1.0097	0.8362	1.088	0.7134	1.0826	1.0016	0.6121
	500	1.2386	0.8642	0.76	0.5057	1.3632	0.7308	0.8394	0.5468	0.8811
	1000	1.1399	1.5475	0.5466	0.4406	1.0881	1.4342	0.6039	0.4406	1.2128
	2500	0.9226	1.3522	0.3943	0.3943	0.9226	1.4294	0.3943	0.3943	1.7299

Tabelle 3.13.: Simulationsergebnisse für Dichteschätzung mit Gaußkern (weitere Varianten) – L_1-Risiko für die Dichten 15-28.

Dichte	n	LIL	LIL 2.1	L2NR	L1NR	kLIL	kLIL 2.1	kL2NR	kL1NR	CvM LIL
15	50	0.3617	0.6697	0.7669	1.7741	0.4692	0.973	0.395	0.5461	0.5296
	100	0.282	0.4855	0.5253	1.3536	0.3557	0.7132	0.3327	0.5034	0.4403
	500	0.1625	0.2161	0.2653	0.482	0.1811	0.3341	0.2143	0.3198	0.2489
	1000	0.1316	0.1601	0.2182	0.3357	0.1409	0.2358	0.1774	0.2628	0.1954
	2500	0.1008	0.1108	0.1706	0.253	0.1011	0.1575	0.136	0.2023	0.1464
16	50	0.2694	0.7971	0.7057	1.781	0.3712	0.935	0.2625	0.462	0.3595
	100	0.2099	0.6012	0.4009	1.2817	0.2677	0.6797	0.2106	0.3866	0.2714
	500	0.1088	0.2942	0.1352	0.3489	0.1229	0.3187	0.1049	0.2088	0.1275
	1000	0.0824	0.2113	0.0937	0.2192	0.0904	0.2288	0.078	0.1535	0.0925
	2500	0.0567	0.1364	0.0585	0.1234	0.0603	0.1447	0.052	0.1001	0.0604
17	50	0.2802	0.8061	0.679	1.7717	0.3958	0.9463	0.2537	0.4363	0.395
	100	0.2218	0.6067	0.3699	1.1706	0.2893	0.6906	0.1997	0.358	0.3097
	500	0.1179	0.309	0.1308	0.3251	0.1344	0.3329	0.0992	0.1993	0.1574
	1000	0.0874	0.2273	0.0856	0.2014	0.0958	0.2416	0.0709	0.1406	0.1143
	2500	0.0571	0.1442	0.0521	0.1155	0.0608	0.1523	0.047	0.0914	0.0745
18	50	0.5224	0.6845	1.2024	1.7427	0.5428	0.9845	0.7276	0.9678	0.6113
	100	0.4518	0.5304	1.0592	1.602	0.4549	0.7241	0.7091	0.9969	0.523
	500	0.3548	0.2961	0.9048	1.3504	0.2987	0.3681	0.6649	1.008	0.328
	1000	0.3295	0.2482	0.8507	1.2932	0.2624	0.2737	0.6209	0.9982	0.2659
	2500	0.306	0.1991	0.8025	1.2413	0.2291	0.1943	0.574	0.9601	0.1977
19	50	0.8185	0.9283	1.6438	1.747	0.7658	1.0308	1.2968	1.5033	0.763
	100	0.7662	0.7617	1.627	1.7283	0.7049	0.8206	1.3791	1.6076	0.6656
	500	0.7119	0.5104	1.5905	1.7831	0.6628	0.5127	1.4551	1.7185	0.5243
	1000	0.7189	0.4599	1.5986	1.781	0.6777	0.4555	1.4974	1.7249	0.4795
	2500	0.7472	0.4241	1.5812	1.8075	0.7092	0.4155	1.5169	1.7754	0.4292
20	50	0.7417	0.962	1.1498	1.7394	0.8167	1.1594	0.8122	0.9706	0.8866
	100	0.6569	0.7678	0.9008	1.434	0.684	0.9114	0.7233	0.8802	0.7415
	500	0.4969	0.4946	0.6316	0.8777	0.4825	0.5446	0.5881	0.7017	0.5055
	1000	0.4432	0.43	0.5844	0.7319	0.4283	0.4594	0.5519	0.6433	0.4398
	2500	0.3697	0.3465	0.4794	0.5907	0.353	0.3564	0.4446	0.5468	0.3506
21	50	0.5598	1.4426	0.6906	1.7269	0.7522	1.5931	0.3557	0.4865	0.8078
	100	0.4154	1.1168	0.4403	1.2232	0.5118	1.2046	0.2814	0.4208	0.6318
	500	0.2029	0.4776	0.1817	0.4124	0.2206	0.5151	0.1534	0.257	0.3225
	1000	0.1528	0.3168	0.1325	0.2654	0.1626	0.3352	0.1187	0.1874	0.2345
	2500	0.1011	0.1949	0.0822	0.1501	0.1056	0.2036	0.0776	0.1202	0.1513
22	50	0.3001	0.7687	0.6798	1.7578	0.4205	0.9544	0.285	0.4491	0.4454
	100	0.2473	0.571	0.4301	1.2493	0.3155	0.7026	0.2317	0.4108	0.3584
	500	0.1474	0.2814	0.1429	0.3466	0.1826	0.3399	0.1166	0.2147	0.2193
	1000	0.1113	0.2142	0.1006	0.2228	0.1414	0.2648	0.0867	0.1585	0.1747
	2500	0.0768	0.148	0.0642	0.123	0.0957	0.1925	0.0594	0.0988	0.1227
23	50	0.4601	0.841	0.718	1.7314	0.5315	0.9806	0.41	0.5086	0.5523
	100	0.4188	0.6609	0.453	1.279	0.4571	0.7333	0.3299	0.4354	0.4876
	500	0.3539	0.451	0.1876	0.4078	0.3638	0.4667	0.1741	0.2546	0.3973
	1000	0.3304	0.4027	0.1393	0.2566	0.3378	0.4115	0.1362	0.1897	0.3686
	2500	0.2765	0.3568	0.0939	0.1508	0.2893	0.3616	0.0959	0.1226	0.3312
24	50	0.7401	1.1101	0.7072	1.7515	0.8753	1.2196	0.4312	0.5167	0.9114
	100	0.6412	0.9807	0.4561	1.2409	0.7422	1.0563	0.3497	0.4371	0.7983
	500	0.4357	0.7332	0.2087	0.3728	0.4744	0.7778	0.2093	0.2468	0.5229
	1000	0.3561	0.6239	0.1686	0.242	0.3806	0.6469	0.1722	0.1895	0.427
	2500	0.2596	0.488	0.1233	0.144	0.2737	0.5093	0.1262	0.1297	0.3273
25	50	0.5189	0.9552	0.7761	1.7991	0.62	1.0703	0.4103	0.544	0.6215
	100	0.4411	0.7776	0.4973	1.3047	0.5133	0.846	0.3567	0.4762	0.534
	500	0.2363	0.5373	0.2421	0.4061	0.2509	0.5644	0.2178	0.2871	0.3682
	1000	0.1845	0.3148	0.1902	0.2808	0.1919	0.3864	0.1775	0.228	0.3063
	2500	0.1326	0.2052	0.1413	0.1953	0.1354	0.2124	0.1348	0.174	0.2334
26	50	0.8142	1.8568	0.9922	1.7919	1.031	1.9123	0.6086	0.7664	1.2529
	100	0.6286	1.5344	0.7061	1.4176	0.7036	1.6238	0.4869	0.6736	0.9988
	500	0.3531	0.6386	0.3211	0.5924	0.3751	0.6647	0.2924	0.3967	0.5331
	1000	0.2738	0.4731	0.2529	0.3868	0.2853	0.4898	0.2379	0.3094	0.4131
	2500	0.1951	0.3198	0.1851	0.2526	0.1999	0.3293	0.1785	0.2257	0.299
27	50	0.6858	1.0237	0.6917	1.7832	0.7721	1.1306	0.5454	0.5254	0.8138
	100	0.6291	0.8765	0.4518	1.2991	0.6882	0.9355	0.45	0.4391	0.7561
	500	0.5376	0.6722	0.214	0.3512	0.5463	0.6917	0.2491	0.2346	0.6339
	1000	0.5216	0.6042	0.1717	0.2279	0.526	0.616	0.1964	0.1761	0.5897
	2500	0.4843	0.5458	0.1309	0.134	0.4938	0.5507	0.1451	0.1186	0.5393
28	50	0.4438	0.8928	0.6865	1.7489	0.5867	1.0499	0.3546	0.4721	0.6903
	100	0.3416	0.693	0.4387	1.2902	0.4164	0.7691	0.2946	0.417	0.6073
	500	0.185	0.3587	0.2183	0.384	0.1946	0.3903	0.1914	0.2653	0.3596
	1000	0.1417	0.2411	0.178	0.2763	0.1466	0.257	0.1604	0.2237	0.2709
	2500	0.1028	0.154	0.1378	0.2042	0.1042	0.1597	0.1281	0.1776	0.1928

Tabelle 3.14.: Simulationsergebnisse für Dichteschätzung mit Gaußkern (weitere Varianten) – gewählte Bandbreiten für die Dichten 1-14.

Dichte	n	LIL	LIL 2.1	L2NR	L1NR	kLIL	kLIL 2.1	kL2NR	kL1NR	CvM LIL
1	50	0.2303	0.5799	0.012	5e-04	0.3077	0.7013	0.0601	0.0244	0.349
	100	0.1831	0.4215	0.0148	0.0015	0.2254	0.477	0.0457	0.0161	0.2869
	500	0.099	0.2131	0.0203	0.0043	0.1107	0.2284	0.0296	0.0102	0.182
	1000	0.0728	0.1559	0.0178	0.0053	0.0798	0.1658	0.0225	0.0093	0.1489
	2500	0.0488	0.1034	0.0134	0.0058	0.0518	0.1078	0.0155	0.0076	0.1126
2	50	0.3207	0.9972	0.0195	9e-04	0.5569	1.8246	0.0977	0.0427	0.6718
	100	0.2338	0.6276	0.0248	0.0025	0.3848	1.0506	0.0702	0.0271	0.5066
	500	0.111	0.2555	0.0251	0.0067	0.1769	0.4161	0.0382	0.0156	0.2773
	1000	0.0812	0.1804	0.0207	0.0075	0.1307	0.2973	0.03	0.0128	0.2178
	2500	0.0522	0.1137	0.0144	0.0068	0.0871	0.1936	0.0216	0.0094	0.1593
3	50	0.5162	1.3354	0.0261	0.0011	0.7123	1.7196	0.1317	0.0533	0.7129
	100	0.4343	0.9483	0.0321	0.0032	0.5502	1.1552	0.1045	0.035	0.5793
	500	0.2997	0.5364	0.0472	0.0085	0.3468	0.6164	0.0797	0.0217	0.374
	1000	0.254	0.4327	0.056	0.0107	0.29	0.4952	0.0812	0.0203	0.314
	2500	0.2009	0.3294	0.0622	0.0135	0.2289	0.3782	0.0767	0.0215	0.2506
4	50	0.9086	2.8284	0.0406	0.0022	1.2627	3.5797	0.2251	0.0885	1.228
	100	0.7478	1.9394	0.0543	0.005	0.9423	2.2439	0.1753	0.0597	0.9778
	500	0.4763	0.9818	0.0776	0.0127	0.534	1.0528	0.1216	0.0345	0.5777
	1000	0.3932	0.7705	0.0825	0.0166	0.427	0.8098	0.1118	0.0331	0.4647
	2500	0.3026	0.565	0.0801	0.0205	0.323	0.589	0.0975	0.0313	0.3571
5	50	1.379	3.8245	0.0687	0.0031	1.8559	4.7418	0.3568	0.1407	1.8472
	100	1.1846	2.7435	0.0832	0.0083	1.4496	3.1369	0.2705	0.0914	1.4901
	500	0.8068	1.5192	0.1191	0.0214	0.8928	1.6198	0.2001	0.0547	0.9481
	1000	0.695	1.2387	0.1381	0.0267	0.7463	1.2978	0.1994	0.0512	0.793
	2500	0.566	0.9589	0.1537	0.0339	0.5964	0.9932	0.1904	0.0537	0.631
6	50	1.2554	4.398	0.0667	0.0033	1.7952	6.0496	0.3254	0.1367	1.8204
	100	1.0125	2.8547	0.0789	0.0081	1.279	3.4243	0.2369	0.0856	1.3932
	500	0.6425	1.3254	0.1077	0.02	0.7131	1.417	0.1762	0.0501	0.8073
	1000	0.5357	1.0167	0.1212	0.0241	0.5769	1.0679	0.1648	0.0481	0.6521
	2500	0.4137	0.738	0.1201	0.0295	0.439	0.767	0.145	0.046	0.4992
7	50	0.882	2.3675	0.045	0.002	1.2415	3.1707	0.2259	0.0936	1.2433
	100	0.7413	1.6572	0.0552	0.0056	0.9473	2.0755	0.1797	0.0604	0.9968
	500	0.5055	0.9205	0.0801	0.0147	0.5852	1.0663	0.1344	0.0369	0.6314
	1000	0.4285	0.7374	0.0955	0.0183	0.4883	0.8465	0.1354	0.0347	0.5273
	2500	0.3431	0.5648	0.1034	0.0231	0.3882	0.6457	0.1275	0.0366	0.4193
8	50	0.0867	0.3599	0.0042	2e-04	0.1955	0.6595	0.0208	0.0084	0.271
	100	0.0493	0.203	0.0032	4e-04	0.1124	0.4139	0.0098	0.0037	0.1947
	500	0.0117	0.0496	9e-04	2e-04	0.0276	0.1182	0.0021	6e-04	0.0809
	1000	0.006	0.0259	5e-04	1e-04	0.0143	0.0635	0.0012	3e-04	0.0533
	2500	0.0025	0.0108	2e-04	1e-04	0.0061	0.0269	5e-04	1e-04	0.0299
9	50	1.0655	5.4776	0.0712	0.0036	2.1874	16.4799	0.3082	0.1408	3.4208
	100	0.6827	2.5543	0.0721	0.0088	1.2521	5.5963	0.1882	0.079	2.1701
	500	0.267	0.7146	0.0557	0.018	0.4497	1.2906	0.0882	0.037	0.8861
	1000	0.1811	0.4535	0.0428	0.0169	0.3062	0.7991	0.0659	0.0288	0.6311
	2500	0.1122	0.2621	0.0301	0.0141	0.1884	0.4578	0.0467	0.0206	0.4154
10	50	2.9549	18.5923	0.1286	0.0078	4.6004	30.658	0.6532	0.2696	5.4006
	100	2.1006	9.1704	0.1536	0.0182	2.8611	11.7439	0.4436	0.1708	3.6356
	500	1.0325	2.7441	0.1724	0.037	1.1643	3.0282	0.2572	0.094	1.5776
	1000	0.7909	1.8392	0.1631	0.044	0.8607	1.9701	0.218	0.08	1.1683
	2500	0.5609	1.186	0.1519	0.0508	0.5983	1.2441	0.1808	0.0741	0.8184
11	50	0.8477	2.2063	0.038	0.0019	1.1184	2.7104	0.2228	0.0833	1.0974
	100	0.7286	1.6408	0.0485	0.0051	0.8858	1.8574	0.1686	0.0534	0.903
	500	0.4999	0.9313	0.0801	0.0125	0.5538	0.9882	0.1336	0.0326	0.5855
	1000	0.4326	0.7613	0.089	0.0163	0.4653	0.7966	0.1267	0.032	0.4918
	2500	0.3529	0.5932	0.0974	0.0202	0.3719	0.6139	0.1195	0.0324	0.3932
12	50	0.4682	1.3968	0.0278	0.0014	0.7536	2.5283	0.1397	0.0603	0.8647
	100	0.3585	0.8885	0.0335	0.0034	0.5212	1.3982	0.1052	0.0365	0.6433
	500	0.2076	0.4094	0.0477	0.0086	0.28	0.5778	0.0734	0.0224	0.3608
	1000	0.1635	0.306	0.0492	0.0112	0.2181	0.4275	0.0656	0.0216	0.2874
	2500	0.1227	0.2171	0.0469	0.014	0.1622	0.3004	0.0568	0.0213	0.2177
13	50	0.4866	3.1428	0.0306	0.0016	0.7203	4.936	0.1442	0.0644	1.0139
	100	0.3655	1.2861	0.0355	0.004	0.4612	1.7468	0.1027	0.0389	0.6582
	500	0.1927	0.4258	0.0421	0.0091	0.2171	0.4593	0.0579	0.023	0.3536
	1000	0.1426	0.309	0.035	0.0107	0.1569	0.3287	0.0438	0.0189	0.2838
	2500	0.0952	0.2036	0.0259	0.011	0.1023	0.2139	0.0304	0.0147	0.2134
14	50	0.0045	0.1075	2e-04	6e-04	0.014	0.1543	0	0	0.0242
	100	3e-04	0.0456	1e-04	9e-04	0.001	0.0732	0	1e-04	0.0084
	500	0	3e-04	0.001	0.0025	0	5e-04	5e-04	0.0023	2e-04
	1000	0	0	0.0021	0.0027	0	0	0.0018	0.0027	0
	2500	0	0	0.0027	0.0027	0	0	0.0027	0.0027	0

Tabelle 3.15.: Simulationsergebnisse für Dichteschätzung mit Gaußkern (weitere Varianten) – gewählte Bandbreiten für die Dichten 15-28.

Dichte	n	LIL	LIL 2.1	L2NR	L1NR	kLIL	kLIL 2.1	kL2NR	kL1NR	CvM LIL
15	50	0.0886	0.2781	0.006	3e-04	0.1609	0.4929	0.0269	0.0122	0.1948
	100	0.0592	0.174	0.006	6e-04	0.1064	0.3035	0.0161	0.0065	0.1498
	500	0.0229	0.0602	0.0043	0.0013	0.0406	0.1118	0.007	0.0029	0.0749
	1000	0.0153	0.0393	0.0032	0.0013	0.0275	0.0738	0.0052	0.0022	0.0561
	2500	0.0091	0.0226	0.0021	9e-04	0.0166	0.0432	0.0035	0.0015	0.0384
16	50	0.371	0.9323	0.0157	7e-04	0.4803	1.1281	0.0971	0.0351	0.4695
	100	0.3206	0.7014	0.0223	0.0019	0.3833	0.7854	0.0755	0.0238	0.3873
	500	0.2277	0.4124	0.036	0.0058	0.248	0.434	0.0568	0.0153	0.2544
	1000	0.1957	0.3388	0.0378	0.0072	0.2097	0.3543	0.0525	0.014	0.2133
	2500	0.1607	0.269	0.0387	0.0088	0.1688	0.2772	0.0485	0.0134	0.1694
17	50	0.2011	0.4932	0.0089	4e-04	0.2634	0.5968	0.0531	0.0198	0.2641
	100	0.1753	0.3714	0.0135	0.0013	0.2108	0.4187	0.0433	0.0142	0.221
	500	0.1248	0.2213	0.0191	0.0033	0.1349	0.233	0.0319	0.0084	0.1471
	1000	0.1069	0.1816	0.0226	0.0042	0.1127	0.1885	0.0325	0.0082	0.124
	2500	0.0849	0.1399	0.0249	0.0051	0.0885	0.144	0.0309	0.008	0.0994
18	50	0.1416	0.7104	0.0065	4e-04	0.3372	1.7724	0.0334	0.0138	0.5195
	100	0.0801	0.3595	0.0051	8e-04	0.1866	0.8367	0.0165	0.006	0.3438
	500	0.0176	0.0778	0.0014	4e-04	0.0415	0.1873	0.0032	0.001	0.1301
	1000	0.0093	0.0403	8e-04	2e-04	0.0218	0.0976	0.0018	5e-04	0.0839
	2500	0.0039	0.017	4e-04	1e-04	0.0093	0.0417	8e-04	2e-04	0.0471
19	50	0.1372	2.0528	0.0015	1e-04	0.3137	3.3998	0.0111	0.0039	0.4596
	100	0.0696	0.8237	8e-04	1e-04	0.1215	1.2037	0.0034	9e-04	0.2303
	500	0.0107	0.1109	2e-04	0	0.0145	0.1362	3e-04	1e-04	0.054
	1000	0.0044	0.0457	1e-04	0	0.0055	0.0535	1e-04	0	0.0286
	2500	0.0013	0.0134	0	0	0.0016	0.0153	0	0	0.0123
20	50	0.6806	4.0141	0.0417	0.0032	1.5031	12.7932	0.1852	0.0863	2.5069
	100	0.446	1.7646	0.0533	0.0067	0.8201	4.0886	0.128	0.0579	1.4937
	500	0.1843	0.4703	0.0452	0.0117	0.282	0.8271	0.0645	0.0285	0.5519
	1000	0.1319	0.3036	0.0384	0.0135	0.1971	0.5078	0.0515	0.0237	0.3853
	2500	0.09	0.1868	0.0313	0.0143	0.1329	0.2954	0.0413	0.0199	0.2539
21	50	0.8824	5.9054	0.0459	0.0025	1.3545	9.2776	0.2147	0.0962	1.5048
	100	0.6147	2.782	0.0572	0.0059	0.8104	3.3635	0.1533	0.0614	1.0557
	500	0.2968	0.7544	0.0666	0.0132	0.3281	0.8249	0.0947	0.0338	0.4926
	1000	0.2342	0.4917	0.0644	0.0158	0.252	0.5197	0.083	0.0317	0.368
	2500	0.1797	0.3254	0.063	0.0194	0.1885	0.3375	0.0723	0.0295	0.2628
22	50	0.8161	2.1385	0.0412	0.002	1.1781	2.8124	0.2233	0.0907	1.2443
	100	0.6801	1.5631	0.0515	0.0054	0.8974	1.9244	0.1681	0.0555	1.0208
	500	0.4207	0.823	0.0865	0.0149	0.5241	0.9934	0.1313	0.0378	0.6348
	1000	0.3355	0.6232	0.0875	0.0183	0.4205	0.776	0.1179	0.0357	0.5113
	2500	0.2567	0.4458	0.0874	0.024	0.3116	0.5634	0.1051	0.0367	0.3818
23	50	0.6888	1.7648	0.0301	0.0017	0.9166	2.2091	0.1616	0.0652	0.9772
	100	0.5621	1.2766	0.0384	0.0036	0.6994	1.4684	0.1109	0.041	0.7958
	500	0.3254	0.6846	0.0455	0.0091	0.3699	0.7337	0.0654	0.0231	0.5031
	1000	0.2489	0.5262	0.0456	0.0116	0.2735	0.5566	0.0581	0.0215	0.4
	2500	0.1658	0.3596	0.0443	0.0138	0.1791	0.3779	0.051	0.0203	0.2637
24	50	0.9297	2.9445	0.0416	0.0019	1.4402	3.8161	0.2	0.0863	1.6189
	100	0.6881	2.0356	0.0509	0.0051	0.9318	2.5472	0.1452	0.0555	1.1
	500	0.3764	0.8895	0.0633	0.013	0.4239	1.0194	0.0935	0.0321	0.4865
	1000	0.2915	0.6498	0.0609	0.0159	0.3172	0.6942	0.0771	0.0286	0.3682
	2500	0.2005	0.4411	0.0506	0.0181	0.2135	0.4684	0.0581	0.0246	0.264
25	50	0.3598	1.0032	0.0122	5e-04	0.5091	1.2206	0.0611	0.0261	0.5132
	100	0.2572	0.7283	0.0158	0.0016	0.3754	0.8269	0.0397	0.0169	0.406
	500	0.0714	0.4119	0.0134	0.004	0.0832	0.4457	0.0187	0.0084	0.204
	1000	0.0509	0.1497	0.011	0.0042	0.0565	0.231	0.0138	0.0067	0.141
	2500	0.0318	0.077	0.0077	0.0034	0.0341	0.0812	0.009	0.0044	0.094
26	50	0.2721	15.1441	0.0066	4e-04	0.714	22.8243	0.0273	0.0135	1.2539
	100	0.0908	3.2453	0.0078	9e-04	0.138	4.638	0.0196	0.0085	0.5852
	500	0.0372	0.114	0.0077	0.002	0.0414	0.1271	0.0105	0.0045	0.0762
	1000	0.0276	0.0653	0.0065	0.0023	0.0298	0.0692	0.008	0.0037	0.0529
	2500	0.0186	0.0404	0.0049	0.0022	0.0197	0.0419	0.0057	0.0028	0.037
27	50	4.3443	11.2387	0.1669	0.0071	5.9682	13.7698	0.9509	0.3402	6.8467
	100	3.4233	8.1184	0.198	0.0184	4.3302	9.336	0.5635	0.2075	5.6392
	500	1.7611	4.0221	0.2356	0.0546	1.9977	4.3463	0.3318	0.1298	3.4692
	1000	1.2955	2.9407	0.2327	0.0651	1.4259	3.1263	0.288	0.117	2.7082
	2500	0.831	1.8841	0.2189	0.0789	0.89	1.9845	0.2467	0.109	1.7351
28	50	0.1771	0.5569	0.0096	5e-04	0.2711	0.7064	0.0451	0.0198	0.3737
	100	0.1246	0.3781	0.011	0.0011	0.1696	0.4523	0.0295	0.012	0.2891
	500	0.051	0.1437	0.0084	0.0026	0.0588	0.1588	0.012	0.0054	0.1449
	1000	0.0337	0.0928	0.0062	0.0024	0.0378	0.1003	0.0081	0.0037	0.1066
	2500	0.0199	0.0525	0.0042	0.0018	0.0216	0.0557	0.005	0.0024	0.0714

4. Dichteschätzung II: Histogramme

4.1. Einführung und Definitionen

Es seien $X_1, \ldots, X_n$ unabhängig identisch verteilt gemäß einer Verteilung P, die eine Lebesgue-Dichte f besitze. Im Folgenden wird angenommen, dass die unbekannte Dichte f außerhalb eines bekannten, kompakten Intervalls – ohne Beschränkung der Allgemeinheit kann dies als $I := [0, 1]$ vorausgestzt werden – Null ist. Ziel ist es, einen Schätzer $\hat{f}$ für f aus den Beobachtungen zu konstruieren. Im Unterschied zu Kapitel 3 soll dies nun durch *Histogramme* geschehen, die eine spezielle Klasse stückweise konstanter Dichteschätzer bilden. Sie haben zwar weniger wünschenswerte Eigenschaften als Kernschätzer, spielen jedoch wegen ihrer Einfachheit und guten Interpretierbarkeit in der Praxis eine große Rolle.

Reguläre Histogramme basieren auf Partitionen $\{I_1, \ldots, I_d\}$ des Intervalls I mit D gleich langen Teilintervallen

$$I_i := \begin{cases} [0, \frac{1}{D}] & (i = 1) \\ (\frac{i-1}{D}, \frac{i}{D}] & (i = 2, \ldots, D) \end{cases}.$$

Diese Intervalle werden auch als *Bins* bezeichnet. Die Anzahl D der Bins bzw. die Binlänge D^{-1} ist ein Glättungsparameter und spielt eine ähnliche Rolle wie die Bandbreite in der Kerndichteschätzung.

Definition 4.1. Für $D \in \mathbb{N}$ sei

$$\mathcal{F}_D := \left\{ f \,\middle|\, f(x) = \sum_{i=1}^{D} \alpha_i \mathbb{I}(x \in I_i) \text{ fast überall}, \alpha_1, \ldots, \alpha_D \geq 0, \sum_{i=1}^{D} \alpha_i = D \right\}$$

die Menge der *Histogrammdichten mit D Bins.*

Für Beobachtungen $X_1, \ldots, X_n$ ist das *reguläre Histogramm mit D Bins* $\hat{f}_D$ definiert als der Maximum-Likelihood-Schätzer innerhalb der Funktionenklasse $\mathcal{F}_D$, d.h. als diejenige Funktion

$$\hat{f}_D := \text{argmax}_{g \in \mathcal{F}_D} L(g, X_1, \ldots, X_n) := \text{argmax}_{g \in \mathcal{F}_D} \sum_{i=1}^{n} \log(g(X_i)),$$

die in $\mathcal{F}_D$ die *Loglikelihood* maximiert. Es ergibt sich

$$\hat{f}_D(x) = \frac{D}{n} \sum_{j=1}^{D} N_j \mathbb{I}(x \in I_j), \tag{4.1}$$

wobei $N_j = \sum_{i=1}^n \mathbb{I}_{I_j}(X_i \in I_j)$ die Anzahl der Beobachtungen im j-ten Bin bezeichnet. Der Wert der maximierten Loglikelihood ist

$$L(\hat{f}_{\mathcal{I}}, X_1, \ldots, X_n) = \sum_{j=1}^{D} N_j \log N_j \; + n\log(D) - n\log(n). \tag{4.2}$$

Im Folgenden wird auch das Integral des Histogramms

$$\hat{F}_n^D(x) := \int_0^x \hat{f}_D(t)dt$$

als Schätzung der Verteilungsfunktion F benötigt. Es ergibt sich für $x \in I_i$

$$\hat{F}_n^D(x) = \begin{cases} DxF_n\left(\frac{1}{D}\right) & (i=1) \\ F_n\left(\frac{i-1}{D}\right) + \left(x - \frac{i-1}{D}\right) D \left(F_n\left(\frac{i}{D}\right) - F_n\left(\frac{i-1}{D}\right)\right) & (i = 2, \ldots, D) \end{cases}. \tag{4.3}$$

Mit Wahrscheilichkeit 1 hat F_n keine Sprungstelle in 0, und in diesem Fall ergibt sich $\hat{F}_n^D$ durch Verbinden der Punkte $\{(\frac{i}{D}, F_n(\frac{i}{D})) : i = 0, \ldots, D\}$.

Zur Konstruktion eines regulären Histogramms ist nur die Wahl der Zahl der Bins D nötig. Dieser Glättungsparameter spielt eine der Bandbreite in der Kerndichteschätzung analoge Rolle, ist jedoch diskret. Die Güte eines Histogramms kann mit denselben Verlust- bzw. Risikofunktionen wie im Fall der Dichteschätzung bewertet werden; je nach Gütekriterium und Voraussetzungen an f ergeben sich verschiedene optimale Wahlen für D. Eine Übersicht über theoretische Resultate zur Binwahl findet man z.B. in Kapitel 3 von [Sco92] oder in Kapitel 2.3 von [PR83].

Außer den hier betrachteten regulären Histogrammen auf einem kompakten Intervall existieren weitere Typen von Histogrammen. Ein Ansatz besteht z.B. darin, für einen fest gewählten Ankerpunkt $x_0 \in \mathbb{R}$ die komplette reelle Achse in gleich lange Intervalle zu partitionieren; die Grenzen zwischen den Bins sind dann gegeben durch $\{x_0 + hk | k \in \mathbb{Z}\}$, wobei dann die Binbreite h eine positive reelle Zahl ist (vgl. z.B. Kapitel 3 in [Sco92]). Dieser Fall ist der Kerndichteschätzung etwas ähnlicher, während der hier betrachtete Ansatz mit einem kompakten Intervall und diskretem Glättungsparameter Bezüge zu klassischen Modellwahlproblemen aufweist (vgl. [BR06] und Kapitel 7 in [Mas07]). Eine weitere Variante sind irreguläre Histogramme, bei denen die Bins verschieden groß sein dürfen. Sie sind zwar flexibler als reguläre Histogramme, ihre Verwendung ist dennoch nicht unbedingt von Vorteil, da nicht nur die Anzahl, sondern auch die jeweiligen Längen der Bins zu wählen sind, vgl. [RMG10] und die dort angegebene Literatur. In der Literatur werden verschiedentlich auch allgemeinere stückweise konstante Dichteschätzer, die sich nicht unbedingt als Maximum-Likelihood-Schätzer ergeben, als „Histogramme“ bezeichnet, so z.B. ein Spezialfall der in [AV89] vorgestellten Splineschätzer.

Zur Wahl von D für reguläre Histogramme sind zahlreiche Methoden vorgeschlagen worden. Die meisten basieren entweder auf der Darstellung des Histogramms als Maximum-Likelihood-Schätzer (sog. Penalized-Likelihood-Methoden), auf Kreuzvalidierung oder auf Plug-In-Methoden. Übersichten und Vergleiche findet man z.B. in [BR06] und in [DGNW09].

Im Unterschied zur Kerndichteschätzung kann die Dichteschätzung mit regulären Histogrammen als Approximation der unbekannten Dichte f durch niedrigdimensionale parametrische Modelle interpretiert werden. Interessant ist dabei der Fall, dass die gesuchte Dichte zu einem dieser Modelle gehört, d.h. dass f selbst eine Dichte vom Histogrammtyp ist. Im weiteren Verlauf des Kapitels soll auch untersucht werden, inwiefern in diesem Fall mit Hilfe des Diskrepanzprinzips die korrekte Anzahl Bins gewählt wird. Dazu werden einige Definitionen benötigt:

Definition 4.2. Es sei

$$\mathcal{F}_{hist} := \bigcup_{D=1}^{\infty} \mathcal{F}_D$$

die Menge aller *Histogrammdichten*. Weiter sei für eine Dichte f auf $[0,1]$

$$D^*(f) := \begin{cases} \min\{D : f \in \mathcal{F}_D\} & (f \in \mathcal{F}_{hist}) \\ \infty & (f \notin \mathcal{F}_{hist}) \end{cases}$$

die minimale Anzahl Bins von f.

Zu einer Dichte f auf $[0,1]$ wird für $D \in \mathbb{N}$ die Histogrammdichte f_D

$$f_D(x) := \sum_{i=1}^{D} \alpha_i \mathbb{I}(x \in I_i)$$

definiert mit

$$\alpha_i = \frac{1}{D} \int_{I_i} f(t)dt.$$

Die zugehörige linear interpolierte Verteilungsfunktion F_D ergibt sich als

$$F_D(x) = \int_0^x f_D(t)dt$$

bzw. als

$$F_D(x) = F\left(\frac{i-1}{D}\right) + \left(x - \frac{i-1}{D}\right) D \left(F\left(\frac{i}{D}\right) - F\left(\frac{i-1}{D}\right)\right)$$

für $x \in I_i$. Für $f \in \mathcal{F}_D$ ist $F_D = F$, ansonsten gilt $F_D(x) \neq F(x)$ für mindestens ein $x \in [0,1]$.

4.2. Analytische Resultate

Da der Glättungsparameter für reguläre Histogramme diskret ist, lassen sich nur Diskrepanzprinzipen vom Typ

$$\text{Wähle } D \text{ minimal so, dass } d(\hat{F}_n^D, F_n) \leq s(n) \tag{4.4}$$

realisieren. Für die Schrankenfunktion $s(n)$ gelte dabei wieder $s(n) = o(1)$. Algorithmisch ist diese Variante des Diskrepanzprinzips sehr einfach umzusetzen: Es werden beginnend mit $D = 1$ so lange reguläre Histogramme mit einer wachsenden Anzahl von Bins generiert, bis die Bedingung (4.4) erfüllt ist. Im Folgenden bezeichne $D_{s,n}$ die Lösung von (4.4). Während für Kerndichteschätzer schon einige Resultate zum Diskrepanzprinzip in der Literatur vorliegen, ist dies bei regulären Histogrammen nicht der Fall. In [DGNW09] wird eine (4.4) ähnliche Methode vorgeschlagen, bei der die Anzahl der Bins so lange erhöht wird, bis das Kuiper-Kriterium aus [DK04] erfüllt ist. Dabei werden jedoch – im Unterschied zu (4.4) – Differenzen von Kuiper-Metriken verschiedener Ordnungen benutzt, und es liegen bislang keine theoretischen Resultate für diese Methode vor.

Im Folgenden sei – wie in Kapitel 3 – stets $d = d_\infty$ oder $d = d_{kuip,k}$ und

$$c_d = \begin{cases} 1, & d = d_\infty \\ 2k, & d = d_{kuip,k} \end{cases}.$$

Zunächst soll die Existenz von $D_{s,n}$ gezeigt werden. Dazu wird die Existenz einer Dichte nicht benötigt, F muss lediglich stetig sein.

Satz 4.1. *Sei F_n die empirische Verteilungsfunktion und $\hat{F}_n^D$ wie in (4.3). Dann gilt fast sicher:*

$$\limsup_{D\to\infty} d(F_n, \hat{F}_n^D) \le \frac{c_d}{n}.$$

Beweis. Sei $\delta := \min_{2\le j\le n}(X_{(j)} - X_{(j-1)})$. Mit Wahrscheinlichkeit 1 ist $\delta > 0$, und für $D > \delta^{-1}$ enthält dann jedes Intervall $[\frac{k-1}{D}, \frac{k}{D}]$ für $k = 1, \dots, D$ höchstens eine Sprungstelle von F_n. Da F_n mit Wahrscheinlichkeit 1 keine Sprungstellen in 0 hat, gilt fast sicher $F_n(\frac{k}{D}) = \hat{F}_n^D(\frac{k}{D})$ für $k = 0, \dots, D$. Damit ist für $D > \delta^{-1}$

$$d_\infty(F_n, \hat{F}_n^D) \le \frac{1}{n}$$

und weiter

$$d(F_n, \hat{F}_n^D) \le c_d d_\infty(F_n, \hat{F}_n^D) \le \frac{c_d}{n}.$$

□

Damit ist die Existenz des Glättungsparameters für $n^{-1} = o(s(n))$ zumindest für hinreichend großes n gesichert. Abbildung 4.1 zeigt für jeweils zwei Stichproben aus einer $N(0,1)$-Verteilung mit $n = 10$ (oben) und $n = 100$ (unten) $d_\infty(F_n, \hat{F}_n^D)$ in Abhängigkeit von D. Der schwarze Punkt markiert den Wert von D, für den erstmals $d_\infty(F_n, \hat{F}_n^D) \le 0.6n^{-1/2}$ gilt. Dies entspricht der von Vapnik vorgeschlagenen Version des Diskrepanzprinzips, die in Kapitel 3 vorgestellt wurde. Man sieht, dass im Vergleich zur Situation für Kerndichteschätzer (vgl. Abbildung 3.1) $d_\infty(F_n, \hat{F}_n^D)$ stärker oszilliert.

Die folgende Abschätzung wird im weiteren Verlauf häufig benötigt:

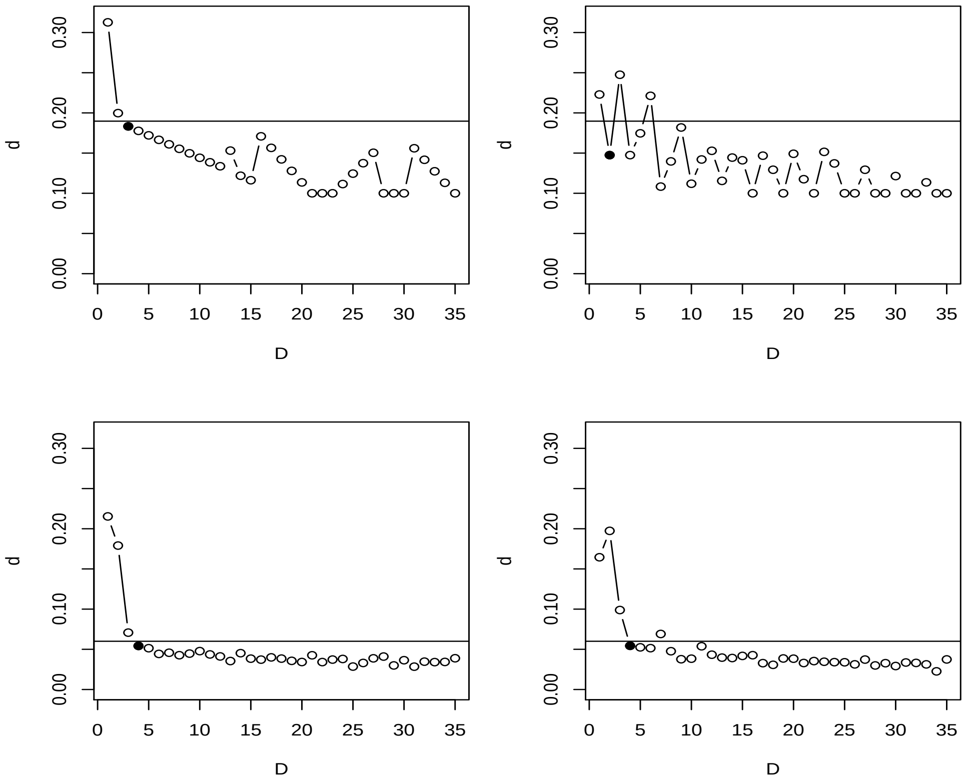

Abbildung 4.1.: $d_\infty(F_n, \hat{F}_n^D)$ in Abhängigkeit von D für $X_1, \ldots, X_n \sim N(0,1)$. Oben: $n = 10$, unten: $n = 100$. Der schwarz markierte Punkt ist jeweils das kleinste D, für das $d_\infty(F_n, \hat{F}_n^D) \leq 0.6n^{-1/2}$.

Satz 4.2. *Für $D \in \mathbb{N}$ gilt fast sicher*

$$d(\hat{F}_n^D, F_D) \leq c_d d_\infty(F_n, F) = O\left(\sqrt{\frac{\log\log n}{n}}\right)$$

gleichmäßig in D.

Beweis. Da $\hat{F}_n^D - F_D$ für jedes $k = 1, \ldots, D$ eingeschränkt auf das Intervall $[\frac{k-1}{D}, \frac{k}{D}]$ eine affin-lineare Funktion ist, gilt

$$d_\infty(\hat{F}_n^D, F_D) = \max_{i=0,\ldots,D} \left| \hat{F}_n^D\left(\frac{i}{D}\right) - F_D\left(\frac{i}{D}\right) \right|.$$

Da weiterhin für $k = 0, \ldots, D$ einerseits $F(\frac{k}{D}) = F_D(\frac{k}{D})$ gilt sowie fast sicher $F_n(\frac{k}{D}) =$

$\hat{F}_n^D(\frac{k}{D})$, folgt fast sicher:

$$\begin{aligned}
d(\hat{F}_n^D, F_D) &\leq c_d d_\infty(\hat{F}_n^D, F_D) \\
&= c_d \max_{i=0,\dots,D} \left| \hat{F}_n^D\left(\frac{i}{D}\right) - F_D\left(\frac{i}{D}\right) \right| \\
&= c_d \max_{i=0,\dots,D} \left| F_n\left(\frac{i}{D}\right) - F\left(\frac{i}{D}\right) \right| \\
&\leq c_d d_\infty(F_n, F) \\
&= O\left(\sqrt{\frac{\log\log n}{n}}\right).
\end{aligned}$$

□

Ähnlich wie im Fall der Kerndichteschätzung lässt sich zeigen, dass das integrierte Histogramm mit durch das Diskrepanzprinzip gewählter Binanzahl stets konsistent für F ist. Die Existenz einer Dichte wird dafür nicht benötigt, F darf jedoch keine Atome besitzen. Falls F nicht Integral einer Histogrammdichte ist, konvergiert außerdem die Anzahl der gewählten Bins gegen unendlich.

Satz 4.3. *1. Sei F stetige Verteilungsfunktion auf $[0,1]$, F_n und $\hat{F}_n^D$ wie oben und $s(n) = o(1)$. Wird $D_{s,n}$ als kleinste Lösung von*

$$d(F_n, \hat{F}_n^D) \leq s(n)$$

gewählt, so gilt fast sicher

$$d(\hat{F}_n^{D_{s,n}}, F) \longrightarrow 0.$$

2. Ist weiter $F \neq F_D$ für alle $D \in \mathbb{N}$, so gilt fast sicher $D_{s,n} \longrightarrow \infty$.

Beweis. 1. Es gilt fast sicher

$$\begin{aligned}
d(F, F_{D_{s,n}}) &\leq d(F, F_n) + d(F_n, \hat{F}_n^{D_{s,n}}) + d(\hat{F}_n^{D_{s,n}}, F_{D_{s,n}}) && (4.5) \\
&\leq O\left(\sqrt{\frac{\log\log n}{n}}\right) + s(n) + O\left(\sqrt{\frac{\log\log n}{n}}\right) && (4.6) \\
&= o(1), && (4.7)
\end{aligned}$$

und damit

$$d(F, \hat{F}_n^{D_{s,n}}) \leq d(F, F_{D_{s,n}}) + d(F_{D_{s,n}}, F_n^{D_{s,n}}) = o(1).$$

2. Sei $F \neq F_D$ für alle $D \in \mathbb{N}$. Sei weiter $D_n := D_{s,n}$ die Folge der nach dem Diskrepanzprinzip gewählten Binanzahlen. Konvergiert D_n nicht gegen ∞, so existiert eine beschränkte Teilfolge D_{m_n}, d.h. es existiert ein $\tilde{D}$ mit $D_{m_n} \leq \tilde{D}$ für alle n und damit ist $\liminf_{n\to\infty} d(F, F_{D_{m_n}}) \geq \min_{1 \leq D \leq \tilde{D}} d(F, F_D) > 0$ im Widerspruch zu (4.7). Also muss $D_{s,n} \longrightarrow \infty$ gelten. □

Die Bedingung $F \neq F_D$ ist für alle $D \in \mathbb{N}$ auch dann erfüllt, wenn F keine Dichte besitzt. Existiert jedoch eine Dichte f mit $f \in \mathcal{F}_{hist}$, so ist sie verletzt. Dieser Fall wird weiter unten in Satz 4.5 behandelt. Wenn f keine Histogrammdichte ist, kann $D_{s,n}$ – analog zur Kerndichteschätzung – zu schnell gegen unendlich konvergieren, so dass in diesem Fall $\hat{f}_n^{D_{s,n}}$ nicht für f konsistent ist.

Satz 4.4. *Sei* $0 < \varepsilon < 1/2$ *so, dass* $n^\varepsilon s(n) = o(1)$ *ist. Sei weiter*

$$f(x) := \begin{cases} \varepsilon x^{-(1-\varepsilon)} & 0 < x \leq 1 \\ 0 & \textit{sonst} \end{cases}$$

und F_n *die empirische Verteilungsfunktion einer unabhängig identisch verteilten Stichprobe aus einer Verteilung mit Dichte* f*. Ist* $D_{s,n}$ *nun die kleinste Lösung von*

$$d(F_n, \hat{F}_n^D) \leq s(n),$$

so gilt:

1. $\frac{n}{D_{s,n}} \longrightarrow 0$ *fast sicher und*

2. $\liminf_{n\to\infty} d_1(\hat{f}_{D_{s,n}}, f) \geq 1$ *fast sicher.*

Beweis. 1. Die zu f gehörige Verteilungsfunktion ist

$$F(x) := \begin{cases} 0 & x \leq 0 \\ x^\varepsilon & 0 < x \leq 1 \\ 1 & x > 1 \end{cases}.$$

Im Folgenden sei $h = D^{-1}$ die Länge der Bins. Für $h > 0$ gilt für $x \in [0, h]$

$$\begin{aligned} F(x) - \hat{F}^D(x) &= F(x) - x\frac{F(h)}{h} \\ &= x^\varepsilon - xh^{\varepsilon-1}. \end{aligned}$$

Diese Funktion ist konkav auf $[0, h]$ und ihre Ableitung

$$\varepsilon x^{\varepsilon-1} - h^{\varepsilon-1}$$

hat eine Nullstelle in $[0, h]$ bei $x_0 = h\varepsilon^{1/(1-\varepsilon)}$. Weiter ist

$$F(x_0) - \hat{F}^D(x_0) = \underbrace{\left(\varepsilon^{\varepsilon/(1-\varepsilon)} - \varepsilon^{1/(1-\varepsilon)}\right)}_{=:c>0} h^\varepsilon = c\frac{1}{D^\varepsilon}.$$

Damit ergibt sich fast sicher

$$
\begin{aligned}
c\left(\frac{n}{D_{s,n}}\right)^{\varepsilon} &\leq n^{\varepsilon} d_{\infty}(F, F_{D_{s,n}}) \\
&\leq n^{\varepsilon}\left(d_{\infty}(F, F_n) + d(F_n, \hat{F}_n^{D_{s,n}}) + d_{\infty}(F_h, \hat{F}_n^{D_{s,n}})\right) \\
&= n^{\varepsilon} O(n^{-1/2}\sqrt{\log\log n}) + n^{\varepsilon} s(n) \\
&= o(1),
\end{aligned}
$$

und damit $\frac{n}{D_{s,n}} = o(1)$.
2. Sei λ das Lebesgue-Maß auf $\mathbb{R}$. Nach Teil 1 ist klar, dass die Anzahl der gewählten Bins fast sicher schneller wächst als n. Damit gilt

$$
\lambda\{\hat{f}_{D_{s,n}} > 0\} \leq \frac{n}{D_{s,n}} = o(1),
$$

da die Gesamtlänge der nichtleeren Bins höchstens $\frac{n}{D_{s,n}}$ beträgt. Für den L_1-Verlust folgt damit fast sicher

$$
\begin{aligned}
\liminf_{n\to\infty} d_1(\hat{f}_{D_{s,n}}, f) &= \liminf_{n\to\infty} \int_{[0,1]} |\hat{f}_{D_{s,n}}(x) - f(x)| dx \\
&\geq \liminf_{n\to\infty} \int_{[0,1]\cap\{\hat{f}_{D_{s,n}}=0\}} |\hat{f}_{D_{s,n}}(x) - f(x)| dx \\
&\geq \liminf_{n\to\infty} \int_{[0,1]\cap\{\hat{f}_{D_{s,n}}=0\}} f(x) dx \\
&\geq \liminf_{n\to\infty} \int_{n/D_{s,n}}^{1} f(x) dx \\
&= 1,
\end{aligned}
$$

wobei im vorletzten Schritt die Monotonie von f benutzt wird. □

Der Satz zeigt, dass die Wahl der Binanzahl mit Hilfe des Diskrepanzprinzips für bestimmte Dichten zu inkonsistenten Schätzern führt. Ähnlich wie bei der Kerndichteschätzung können Polstellen der Dichte f extreme Unterglättung erzwingen. Abbildung 4.2 zeigt vier reguläre Histogramme für Stichproben vom Umfang $n = 50$ aus der Dichte aus Satz 4.4 für $\varepsilon = 0.32$. Die Anzahl D der Bins wurde kleinstmöglich gewählt, so dass $d_{\infty}(F_n, \hat{F}_n^D) \leq 0.6n^{-1/2}$.

Im Fall $D^*(f) < \infty$ werden bei Verwendung des Diskrepanzprinzips asymptotisch fast sicher mindestens $D^*(f)$ Bins gewählt. Für langsam genug fallende Schrankenfunktionen $s(n)$ ist das Diskrepanzprinzip sogar *modellwahlkonsistent*, d.h. wenn f eine Histogrammdichte ist, wird asymptotisch fast sicher die minimal korrekte Anzahl Bins gewählt:

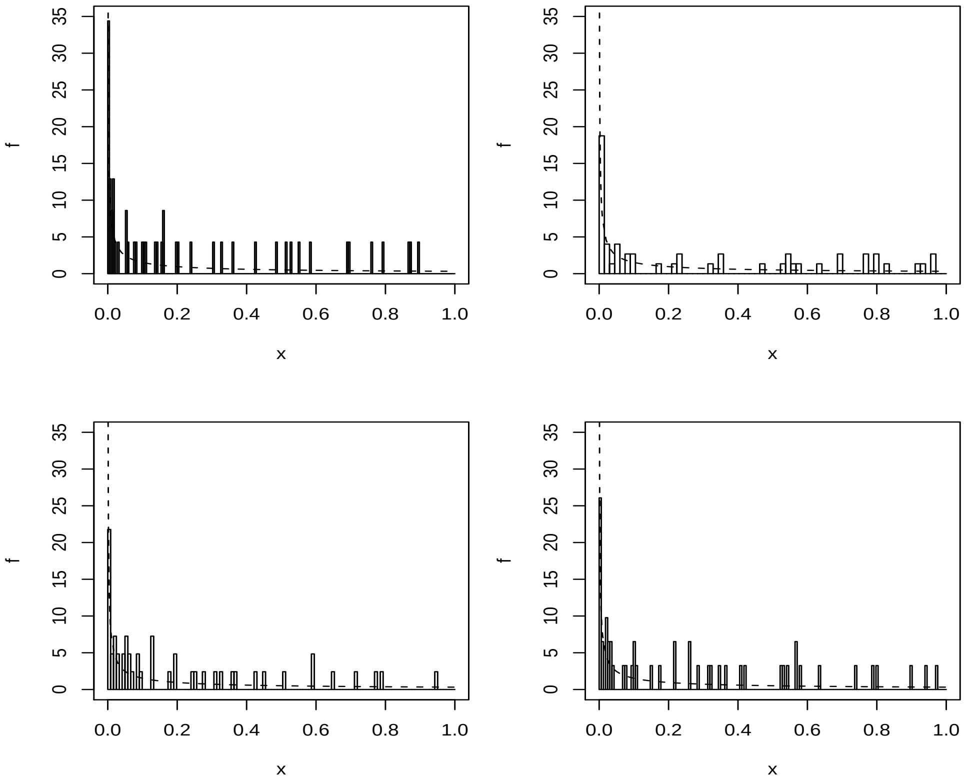

Abbildung 4.2.: Vier Realisationen von $\hat{f}_D$ mit D kleinste Lösung von $d_\infty(F_n, \hat{F}_n^D) \leq 0.6n^{-1/2}$. Die Stichproben stammen aus der Dichte aus Satz 4.4 mit $\varepsilon = 0.32$ und $n = 50$ (gestrichelte Linie).

Satz 4.5. *Sei $D^*(f) < \infty$. Dann gilt:*

1. $\liminf_{n\to\infty} D_{s,n} \geq D^*(f)$ *fast sicher.*
2. *Gilt außerdem* $\sqrt{\frac{\log\log n}{n}} = o(s(n))$, *so folgt*

$$D_{s,n} \overset{n\to\infty}{\longrightarrow} D^*(f) \text{ fast sicher.}$$

Beweis. 1. Es sei für $D = 1, \ldots, D^*(f)$

$$b_D := d(F, F_D).$$

Dann ist nach Definition von $D^*(f)$

$$\begin{aligned} b_D &> 0 \qquad\qquad (D = 1, \ldots, D^*(f) - 1), \\ b_{D^*(f)} &= 0. \end{aligned}$$

Sei zuächst $D \in \{1, \ldots, D^*(f) - 1\}$. Dann gilt fast sicher

$$\begin{aligned} b_D = d(F, F_D) &\le d(F, F_n) + d(F_n, \hat{F}_n^D) + d(\hat{F}_n^D, F^D) \\ &= d(F_n, \hat{F}_n^D) + O\left(\sqrt{\frac{\log\log n}{n}}\right). \end{aligned}$$

Für beliebiges $\varepsilon > 0$ existieren daher $n_{1,\varepsilon}, \ldots, n_{D^*(f)-1,\varepsilon}$, so dass

$$d(F_n, \hat{F}_n^D) \ge b_D - \varepsilon$$

für alle $D = 1, \ldots, D^*(f) - 1$ und $n \ge \max\{n_{1,\varepsilon}, \ldots, n_{D^*(f)-1,\varepsilon}\}$. Da $s(n) = o(1)$, gilt also für hinreichend großes n fast sicher

$$d(F_n, \hat{F}_n^D) \ge s(n),$$

d.h. für die gewählte Binanzahl gilt $D_{s,n} > D^*(f) - 1$.
2. Weiter gilt für $\sqrt{\frac{\log\log n}{n}} = o(s(n))$ fast sicher

$$\begin{aligned} d(F_n, \hat{F}_n^{D^*(f)}) &\le d(F_n, F) + \underbrace{d(F, F_{D^*(f)})}_{=0} + d(F_{D^*(f)}, \hat{F}_n^{D^*(f)}) \\ &= O\left(\sqrt{\frac{\log\log n}{n}}\right) \\ &= o(s(n)). \end{aligned}$$

Damit existiert fast sicher $n_{D^*(f)}$, so dass für $n \ge n_{D^*(f)}$

$$d(F_n, \hat{F}_n^{D^*(f)}) < s(n).$$

Für hinreichend großes n gilt damit fast sicher $D_{s,n} = D^*(f)$. □

Für schnell fallende Schrankenfunktionen wird asymptotisch nicht unbedingt die korrekte Binanzahl gewählt. Ist zum Beispiel $f = \mathbb{I}(0 \le x \le 1)$ die Dichte der Gleichverteilung auf $[0, 1]$ und damit $D^*(f) = 1$, so gilt mit $d = d_\infty$ und $s(n) = cn^{-1/2}$:

$$\begin{aligned} \lim_{n\to\infty} P(D_{s,n} > 1) &= \lim_{n\to\infty} P(d_\infty(F_n, \hat{F}_n^1) > cn^{-1/2}) \\ &= \lim_{n\to\infty} P(d_\infty(F_n, F) > cn^{-1/2}) \\ &= 1 - F_{KS}^{-1}(c) \in (0, 1). \end{aligned}$$

Im Fall $f \notin \mathcal{F}_{hist}$ ist eine ähnliche, allerdings etwas schwächere, Aussage wie in Satz 3.6 möglich:

Satz 4.6. *Sei f eine Dichte auf $[0,1]$ und es existiere ein Intervall $[a,b]$, so dass f auf $[a,b] \subseteq [0,1]$ stetig differenzierbar ist und $\min_{x\in[a,b]} |f'(x)| =: c > 0$. Weiter sei $s(n) = O((\frac{\log\log n}{n})^{1/2})$. Dann gilt für den gewählten Glättungsparameter fast sicher*

$$D_{s,n}^{-1} = O\left(\left(\frac{\log\log n}{n}\right)^{\frac{1}{4}}\right).$$

Beweis. Zunächst impliziert die Voraussetzung an die Ableitung, dass $f \notin \mathcal{F}_{hist}$ ist. Nach Satz 4.3 existiert fast sicher ein n_0, so dass für alle $n \geq n_0$ $D_{s,n} \geq \frac{1}{2(b-a)}$ ist, d.h. für $n \geq n_0$ enthält $[a,b]$ mindestens einen Bin. Sei nun $[a_n, b_n] \subseteq [a,b]$ ein solcher Bin mit $(b_n - a_n) = D_{s,n}^{-1}$. Es gilt unter Verwendung von Satz A.2 fast sicher

$$\begin{aligned}
\max_{x\in[a_n,b_n]} |F(x) - F_{D_{s,n}}(x)| &= \max_{x\in[a_n,b_n]} \left|(x-a_n)(x-b_n)\frac{f'(\xi)}{2}\right| \\
&\geq \frac{c}{2} \max_{x\in[a_n,b_n]} (x-a_n)(b_n-x) \\
&= \frac{c}{8D_{s,n}^2},
\end{aligned}$$

da $(x-a_n)(b_n-x)$ sein Maximum in $x_0 = \frac{a_n+b_n}{2}$ annimmt und $(x_0-a_n)(b_n-x_0) = \frac{1}{4D_{s,n}}$. Damit gilt weiter fast sicher:

$$\begin{aligned}
\frac{c}{8D_{s,n}^2} &\leq d_\infty(F, F_{D_{s,n}}) \\
&\leq d(F, F_{D_{s,n}}) \\
&\leq d(F, F_n) + d(F_n, \hat{F}_n^{D_{s,n}}) + d(\hat{F}_n^{D_{s,n}}, F_{D_{s,n}}) \\
&\leq s(n) + O\left(\sqrt{\frac{\log\log n}{n}}\right) \\
&= O\left(\sqrt{\frac{\log\log n}{n}}\right).
\end{aligned}$$

Damit ergibt sich fast sicher

$$D_{s,n}^{-1} = O\left(\left(\frac{\log\log n}{n}\right)^{\frac{1}{4}}\right).$$

□

Existiert also auf einem Intervall die Ableitung von f und ist diese von 0 weg beschränkt, so konvergiert die Bingröße mindestens so schnell wie $(\frac{\log\log n}{n})^{1/4}$ gegen 0 bzw. die Anzahl der Bins mindestens so schnell wie $(\frac{n}{\log\log n})^{1/4}$ gegen unendlich. Dies sind jedoch nur Abschätzungen der Konvergenzgeschwindigkeit, die Konvergenz kann tatsächlich wesentlich schneller sein. Unter den üblichen Glattheitsvoraussetzungen sind die optimalen Binanzahlen bzgl. d_p von der Form $D = cn^{1/3}$ für $p \in [1, \infty)$ und $D = c(\frac{n}{\log(n)})^{1/3}$ für $p = \infty$ (vgl. Kapitel 3 in [Sco92]). Die optimale Binanzahl konvergiert also etwas schneller gegen unendlich als die in Satz 4.6 hergeleitete Schranke.

4.3. Simulationsstudie

Auf dem Diskrepanzprinzip basierende Methoden zur Wahl der Binanzahl in regulären Histogrammen sollen nun in einer Simulationsstudie verglichen werden. Die verwendeten Dichten, Stichprobengrößen und Verlustfunktionen sind dieselben wie in Kapitel 3.4. Es werden ebenfalls jeweils 100 Simulationsläufe durchgeführt. Das Intervall $[X_{(1)}, X_{(n)}]$ wird dabei in D, $D \in \{1, 2, \ldots, n\}$, gleich lange Teilintervalle geteilt. Die maximale Anzahl der Bins wird aufgrund der Rechenzeit nach oben durch n begrenzt. Verglichen werden verschiedene Methoden zur Wahl von D, darunter folgende Versionen des Diskrepanzprinzips:

- Auf dem Kolmogorov-Smirnov-Test basierende Versionen: D wird gewählt als kleinste Lösung von $d_\infty(F_n, \hat{F}_n^D) \le cn^{-1/2}$ mit $c = 0.83, 1.22, 1.36, 1.62$. Dies entspricht dem Median, dem 90%-, dem 95%- bzw. dem 99%-Quantil der asymptotischen Verteilung von $\sqrt{n} d_\infty(F_n, F)$. Diese Methoden werden in den Tabellen mit **KS .5**, **KS .9**, **KS .95** bzw. **KS .99** bezeichnet.

- Die von Vapnik (für andere Dichteschätzer) vorgeschlagene Variante: D wird gewählt als kleinste Lösung von $d_\infty(F_n, \hat{F}_n^D) \le cn^{-1/2}$ mit $c = 0.6$. In der Tabelle mit **V** bezeichnet.

- Die Version von Eggermont und LaRiccia: D wird gewählt als kleinste Lösung von $d_\infty(F_n, \hat{F}_n^D) \le 0.35n^{-2/5}$. Die theoretische Rechtfertigung dieser Schrankenfunktion im Fall der Kerndichteschätzung lässt sich auf Histogramme nicht übertragen. Da diese Schrankenfunktion zum einen bei der Kerndichteschätzung sehr gute Ergebnisse liefert, zum anderen nach Satz 4.5 modellwahlkonsistent ist, wird sie auch für die Histogramme betrachtet und in den Tabellen mit **E-LR** bezeichnet.

- Auf der Kuiper-Metrik basierende Versionen: D wird gewählt als kleinste Lösung von $d_{kuip,1}(F_n, \hat{F}_n^D) \le cn^{-1/2}$ mit $c = 1.22, 1.62, 1.75, 1.99$. Dies entspricht dem Median, dem 90%-, dem 95%- bzw. dem 99%-Quantil der asymptotischen Verteilung von $\sqrt{n} d_{kuip,1}(F_n, F)$. Diese Methoden werden in den Tabellen mit **Kuip .5**, **Kuip .9**, **Kuip .95** bzw. **Kuip .99** bezeichnet.

- Auf dem Cramér-von Mises-Test basierende Versionen: D wird gewählt als kleinste Lösung von $d_{CvM}(F_n, \hat{F}_n^D) \le cn^{-1}$ mit $c = 0.12, 0.35, 0.46, 0.74$. Dies entspricht dem Median, dem 90%-, dem 95%- bzw. dem 99%-Quantil der asymptotischen Verteilung von $nd_{CvM}(F_n, F)$. Diese Methoden werden in den Tabellen mit **CvM .5**, **CvM**

.9, **CvM .95** bzw. **CvM .99** bezeichnet. Die von Vapnik (vgl. z.B. Kapitel 7.9 in [Vap98]) für andere Dichteschätzer vorgeschlagene Version mit $c = 0.05$ wird mit **CvM V** bezeichnet.

Existiert kein $D \in \{1, 2, \ldots n\}$, für das $d(F_n, \hat{F}_n^D) \leq s(n)$, so wird $D = n$ gewählt. Es werden wieder einige Standardmethoden zum Vergleich herangezogen. Ausgewählt werden Methoden, die in früheren Vergleichsstudien (vgl. [BR06] und [DGNW09]) gut abgeschnitten haben:

- Das von Birgé und Rozenholc vorgeschlagene Penalized-Likelihood-Verfahren (vgl. [BR06]). Die Anzahl der Bins D wird so gewählt, dass
$$L(\hat{f}_{\mathcal{I}}, x_1, \ldots, x_n) - (D - 1 + (\log D)^{2.5})$$
maximal wird. Dabei ist L die Loglikelihood (4.2). Der Strafterm ist asymptotisch äquivalent zu dem des klassischen Akaike-Informationskriteriums AIC (vgl. [Aka74]), führt jedoch bei kleinen Stichprobenumfängen zu besseren Ergebnissen. Daher wird statt des klassischen AIC diese Methode in die Simulationsstudie aufgenommen. Die Methode ist ursprünglich für das Hellinger-Risiko optimiert. In den Tabellen mit **BR** bezeichnet.

- Das Bayessche Informationskriterium BIC (vgl. [Sch78]). Die Anzahl der Bins D wird so gewählt, dass
$$L(\hat{f}_{\mathcal{I}}, x_1, \ldots, x_n) - \frac{1}{2} \log(n) D$$
maximal wird. Das Kriterium zielt darauf ab, asymptotisch das kleinste wahre Modell zu wählen, wenn das datenerzeugende Modell in der Menge der zur Auswahl stehenden Modelle enthalten ist. Für reguläre Histogramme entspricht dies dem Fall $f \in \mathcal{F}_{hist}$. Dies ist jedoch für das Problem der nichtparametrischen Dichteschätzung nicht realistisch, weshalb das Kriterium nur selten zur Histogrammkonstruktion verwendet wird. Da einige Varianten des Diskrepanzprinzips jedoch nach Satz 4.5 eine ähnliche Modellwahlkonsistenzeigenschaft besitzen, wird die Methode mit aufgenommen. Generell wird die Modellgröße stärker bestraft als beim AIC, was zu sehr sparsamen Modellen führt. In den Tabellen mit **BIC** bezeichnet.

- Leave-One-Out-L_2-Kreuzvalidierung. Gewählt wird das D, für das
$$\frac{D(n+1)}{n^2} \sum_{i=1}^{D} N_i^2 - 2D$$
maximal wird (vgl. Formel 2.8 in [Rud82]). In den Tabellen mit **L2CV** bezeichnet.

- Ein auf der Minimierung der stochastischen Komplexität basierendes Kriterium. Gewählt wird das D, für das
$$\frac{D^n (D-1)!}{(D+n-1)!} \prod_{i=1}^{D} N_i!$$

maximal wird (vgl. Formel 2.3 in [HH88]). In den Tabellen mit **SC** bezeichnet.

Verwendet werden jeweils die Implementierungen im `R`-Paket `histogram` [MRZ09].

Der Vergleich der auf dem Kolmogorov-Smirnov-Test basierenden Versionen des Diskrepanzprinzips ergibt ein ähnliches Bild wie bei den Kerndichteschätzern. Die Tabellen 4.1 und 4.2 zeigen die Simulationsergebnisse für das L_1-Risiko. Wie zu erwarten, liefern Histogramme oft schlechtere Ergebnisse als Kerndichteschätzer. Hierbei gibt es einige Ausnahmen, z.B. die Dichte 13, die einen kompakten Träger und hohe Sprünge hat. Ein Sonderfall ist auch die Gleichverteilung (Dichte 1), auf die weiter unten noch näher eingegangen wird. Im Gegensatz zur Bandbreite bei Kerndichteschätzern ist die Anzahl der Bins ein diskreter Parameter, der nur relativ wenige verschiedene Werte annehmen kann. Dies führt dazu, dass die Methoden oft ähnlichere Ergebnisse liefern als die verschiedenen Parameterwahlmethoden für Kernschätzer. Besonders bei Dichten mit schweren Tails (z.B. 6,9,10) versagen alle Methoden gleichermaßen. Ebenso liefert nur **BIC** etwas weniger schlechte Ergebnisse für Dichte 14, während alle Versionen des Diskrepanzprinzips versagen. Für beschränkte Dichten mit leichteren Tails schneiden wie bei den Kerndichteschätzern die Referenzmethoden **BR**, **BIC**, **L2CV** und **SC** gut ab, ähnlich die beiden auf dem Diskrepanzprinzip basierenden Methoden **E-LR** und **V**, von denen häufig sogar eine das beste Ergebnis liefert. Die auf dem Median oder extremen Quantilen der Kolmogorov-Smirnov-Teststatistik basierenden Methoden **KS .5**, **KS .9**, **KS .95** und **KS .99** funktionieren meist deutlich schlechter. Die geschätzten Risiken für Dichte 27 bei großem n sind für alle auf dem Diskrepanzprinzip basierenden Methoden wesentlich schlechter als die der Referenzmethoden. Die Ergebnisse für den L_2- bzw. L_∞-Verlust sind im Anhang abgedruckt (Tabellen B.16-B.18). Sie ergeben ein ähnliches Bild, wobei **V** zumindest für unimodale Dichten bezüglich des L_2-Risikos häufig am besten abschneidet, und in Bezug auf das L_∞-Risiko die Verwendung von Diskrepanzprinzipien mit extremen Quantilen vorteilhaft sein kann.

Betrachtet man nun die durchschnittliche Anzahl der gewählten Bins (Tabellen 4.3 und 4.4), so fällt auf, dass von den Referenzmethoden **L2CV** in der Regel die größte und **BIC** die kleinste Anzahl Bins wählt, während die Anzahlen für **SC** und **BR** dazwischen liegen. Die auf dem Diskrepanzprinzip basierenden Methoden **E-LR** und **V** verhalten sich wieder untereinander sehr ähnlich, wobei **E-LR** für kleinere Stichproben eine größere Anzahl Bins wählt und **V** für größere. Beide liegen für größere Stichproben meistens zwischen **BIC** und **L2CV**, bei kleineren Werten von n wählt insbesondere **E-LR** aber häufig sogar mehr Bins als die Referenzmethoden. Die sehr gleichförmige Struktur aus 10 Dreiecken in Dichte 27 kann natürlich nur für größeres n entdeckt werden. Die auf dem Diskrepanzprinzip basierenden Methoden wählen jedoch in diesem Fall noch für sehr große Stichproben eine geringe Anzahl Bins, so wählt z.B. **KS .99** für $n = 2500$ im Durchschnitt 1.28 Bins (in 83 von 100 Simulationsläufen wird nur ein Bin gewählt).

Ein Spezialfall, zu dem es bei Kerndichteschätzern kein Analogon gibt, ist die Gleichverteilung 1. Lässt man außer Acht, dass in der Simulationsstudie der Träger nicht als bekannt vorausgesetzt wird, ist das Histogramm mit einem Bin das wahre Modell, d.h. der Bias ist Null. Es ist gleichzeitig auch das optimale Modell, da die Erhöhung der Anzahl der

Bins den Bias nicht mehr veringern kann aber die Varianz erhöht. Hier sind also Methoden klar im Vorteil, die generell eine geringe Binanzahl wählen, so z.B. **KS .9**, **KS .95**, **KS .99** und die Referenzmethoden **BIC** und **SC**. Andererseits lehnt der Kolmogorov-Smirnov-Test zum Niveau α das wahre Modell mit Wahrscheinlichkeit α ab, d.h. auch asymptotisch wird mit positiver Wahrscheinlichkeit mehr als ein Bin gewählt. Wie in Satz 4.5 gezeigt, wird für **E-LR** asymptotisch mit Wahrscheinlichkeit 1 nur ein Bin gewählt, was sich in den Simulationen aber noch nicht zeigt. Sowohl **BIC** als auch **SC** wählen in allen 100 Simulationsläufen für $n = 2500$ das Histogramm mit einem Bin.

Die Verwendung der Kuiper-Metrik oder des Cramér-von Mises-Test (vgl. Tabellen 4.5 und 4.6) ergibt für die unimodalen Dichten (1-20) nur geringe Veränderungen des L_1-Risikos, während bei den multimodalen Dichten (21-28) teilweise **Kuip .5**, **Kuip .9**, **Kuip .95** und **Kuip .99** besser sind als **KS .5**, **KS .9**, **KS .95** und **KS .99**, wohingegen **CvM .5**, **CvM .9**, **CvM .95** und **CvM .99** schlechtere Ergebnisse liefern. Das Risiko der Methode **CvM V** ist fast identisch mit dem von **V**. Dasselbe Bild ergibt sich, wenn man die L_2- bzw. L_∞-Risiken betrachtet (Tabellen B.16-B.18 im Anhang). Die Tabellen B.22 und B.23 im Anhang enthalten die arithmetischen Mittel der gewählten Binanzahlen. Hier sieht man, dass bei der Verwendung der Kuiper-Metrik tendenziell eine etwas größere Zahl gewählt wird als bei Verwendung des Kolmogorov-Smirnov-Tests, bei Verwendung des Cramér-von Mises-Test eine etwas kleinere.

Insgesamt zeigt sich, dass auch für reguläre Histogramme einige Versionen des Diskrepanzprinzips eine durchaus praktikable Alternative zu den üblichen Binwahlmethoden darstellen. Das Diskrepanzprinzip für reguläre Histogramme ist zudem ein Abbruchkriterium, d.h. es müssen nur solange reguläre Histogramme mit wachsender Anzahl Bins konstruiert werden, bis (4.4) erfüllt ist. Für die vier Referenzmethoden in der Simulationsstudie sowie für die meisten Methoden ähnlichen Typs müssen jedoch zunächst Histogramme für alle in Frage kommenden Binanzahlen konstruiert werden, was unter Umständen einen erheblich größeren Rechenaufwand bedeutet.

Tabelle 4.1.: Simulationsergebnisse für reguläre Histogramme – L_1-Risiko für die Dichten 1-14.

Dichte	n	BR	BIC	L2CV	SC	E-LR	V	KS .5	KS .9	KS .95	KS .99
1	50	0.1072	0.0908	0.1979	0.3412	0.2942	0.2499	0.1621	0.0957	0.0877	**0.0857**
	100	0.0579	0.0468	0.1097	0.148	0.1844	0.1694	0.0962	0.0546	0.0442	**0.0429**
	500	0.0165	0.0114	0.0469	0.0114	0.0584	0.0659	0.0364	0.0162	0.0118	**0.0103**
	1000	0.0133	0.0053	0.0361	0.0053	0.0389	0.0487	0.0259	0.0103	0.0071	**0.005**
	2500	0.0063	**0.0016**	0.019	**0.0016**	0.0172	0.0293	0.0149	0.0041	0.0024	0.0024
2	50	0.3963	0.3963	0.4495	0.3933	0.4164	0.3862	**0.3485**	0.3723	0.3857	0.4142
	100	0.3075	0.3183	0.3274	0.2999	0.2799	0.2787	**0.269**	0.2965	0.3109	0.3367
	500	0.1874	0.2156	0.1786	0.2009	**0.1551**	0.1573	0.1617	0.1861	0.1956	0.2156
	1000	0.1528	0.1792	0.139	0.1692	0.1298	**0.1268**	0.134	0.1556	0.1631	0.1788
	2500	0.1102	0.1458	0.0991	0.1363	0.1004	**0.0934**	0.1024	0.121	0.1274	0.1394
3	50	0.3397	0.3438	0.3744	0.463	0.3593	0.3322	**0.323**	0.3477	0.3528	0.3795
	100	0.2863	0.2983	0.2991	0.2945	**0.2579**	0.2587	0.2712	0.3155	0.3297	0.3424
	500	0.1645	0.1767	0.1694	0.1671	0.1587	**0.1549**	0.1737	0.2056	0.2193	0.2484
	1000	**0.1279**	0.1407	0.1303	0.1326	0.1343	0.1295	0.1449	0.1717	0.1807	0.1987
	2500	**0.0932**	0.1061	0.1004	0.0972	0.1082	0.0991	0.1122	0.1335	0.1401	0.1545
4	50	0.426	0.4245	0.4467	0.4278	0.3992	**0.3868**	0.4025	0.4646	0.4889	0.5858
	100	0.3655	0.3738	0.3468	0.3496	**0.3115**	0.3141	0.3388	0.4019	0.4096	0.4464
	500	0.2272	0.2535	0.2048	0.234	0.1976	**0.1944**	0.2233	0.2699	0.2865	0.3133
	1000	0.1768	0.2159	0.1615	0.1981	0.1692	**0.1597**	0.1848	0.2258	0.2429	0.2647
	2500	0.1336	0.1728	**0.1173**	0.1566	0.1385	0.1241	0.1444	0.1742	0.1867	0.2068
5	50	0.3735	0.3778	0.4168	0.4359	0.3717	**0.357**	0.3575	0.4314	0.4642	0.5436
	100	0.3125	0.3212	0.3341	0.3073	**0.2921**	0.295	0.3139	0.3536	0.3682	0.3994
	500	0.1876	0.2088	0.187	0.1915	0.1813	**0.1755**	0.2027	0.2486	0.2688	0.3021
	1000	0.1442	0.1722	**0.1419**	0.1575	0.1544	0.1449	0.1677	0.2074	0.2198	0.2428
	2500	0.1093	0.1329	**0.1032**	0.1245	0.1294	0.1154	0.1358	0.1674	0.1788	0.1969
6	50	0.7781	0.8055	0.6061	0.6978	0.6046	**0.5992**	0.6063	0.6679	0.692	0.7443
	100	0.7221	0.7902	0.5262	0.6779	0.5259	**0.525**	0.539	0.5816	0.6009	0.6489
	500	0.6381	0.812	**0.407**	0.6618	0.4393	0.4376	0.45	0.4759	0.4873	0.5032
	1000	0.5881	0.8148	**0.3548**	0.6382	0.3907	0.3881	0.3935	0.4179	0.4278	0.4396
	2500	0.575	0.8465	**0.3182**	0.6394	0.3552	0.3523	0.356	0.3656	0.3694	0.3868
7	50	0.3826	0.3868	0.4192	0.4511	0.3752	**0.3446**	0.3552	0.3916	0.4088	0.4362
	100	0.3239	0.3278	0.3275	0.3175	0.2776	**0.2767**	0.298	0.362	0.38	0.4078
	500	0.1784	0.1892	0.1836	0.1811	0.18	**0.174**	0.1988	0.2275	0.2377	0.2566
	1000	0.1435	0.1591	**0.1413**	0.149	0.1481	0.1418	0.1622	0.2074	0.2157	0.2257
	2500	0.105	0.1228	**0.1044**	0.1144	0.1203	0.1105	0.1252	0.1504	0.1639	0.187
8	50	0.4589	0.4568	0.5856	0.6385	0.6406	0.5797	0.4823	0.4303	**0.4267**	0.4449
	100	0.3852	0.3862	0.5495	0.521	0.5631	0.5218	0.418	**0.3642**	0.3666	0.3678
	500	0.2728	0.2677	0.45	0.2904	0.3863	0.4152	0.3209	0.2616	0.254	**0.2491**
	1000	0.2274	0.2246	0.4262	0.2327	0.3415	0.3891	0.2974	0.2311	0.222	**0.2125**
	2500	0.1891	**0.1779**	0.4064	0.1841	0.2873	0.3598	0.269	0.2058	0.1933	0.1799
9	50	1.4708	1.4956	1.3463	1.3797	**1.3428**	**1.3428**	**1.3428**	1.3437	1.344	1.3481
	100	1.511	1.5665	1.4204	1.4453	**1.4188**	**1.4188**	**1.4188**	**1.4188**	**1.4188**	1.4196
	500	1.6605	1.7506	1.6333	1.6369	**1.6332**	**1.6332**	**1.6332**	**1.6332**	**1.6332**	**1.6332**
	1000	1.6926	1.7855	1.6714	1.6733	**1.6713**	**1.6713**	**1.6713**	**1.6713**	**1.6713**	**1.6713**
	2500	1.7577	1.8401	**1.7467**	1.7472	**1.7467**	**1.7467**	**1.7467**	**1.7467**	**1.7467**	**1.7467**
10	50	1.4033	1.4345	**1.2831**	1.3261	1.2987	1.2987	1.2958	1.3022	1.3024	1.3053
	100	1.4606	1.5231	**1.3644**	1.3932	1.3873	1.3873	1.3863	1.387	1.3874	1.3866
	500	1.621	1.7209	**1.5863**	1.5951	1.5988	1.5988	1.5988	1.5988	1.5988	1.5988
	1000	1.7025	1.7923	**1.6827**	1.6843	1.694	1.694	1.694	1.694	1.694	1.694
	2500	1.7538	1.837	**1.743**	1.7436	1.7516	1.7516	1.7516	1.7516	1.7516	1.7516
11	50	0.3627	0.367	0.4124	0.4737	0.3603	**0.3408**	0.3519	0.433	0.4567	0.4983
	100	0.2878	0.2936	0.3116	0.2987	0.2758	**0.2732**	0.2936	0.3287	0.3482	0.4143
	500	0.1699	0.1842	0.1767	0.1707	0.1726	**0.1689**	0.1905	0.2363	0.2543	0.2981
	1000	**0.1329**	0.153	0.1334	0.141	0.1494	0.1382	0.1637	0.2025	0.2144	0.2335
	2500	0.098	0.1181	**0.0961**	0.1068	0.1247	0.1085	0.1307	0.1614	0.1717	0.1935
12	50	0.4583	0.4576	0.4341	0.4386	0.4079	0.3809	**0.3805**	0.4052	0.4253	0.4698
	100	0.3573	0.3798	0.346	0.3478	0.3089	0.3045	**0.3016**	0.3337	0.3469	0.3793
	500	0.2518	0.2913	0.2495	0.2785	**0.2111**	0.2121	0.2189	0.2465	0.255	0.2699
	1000	0.2229	0.2611	0.2077	0.2503	0.1741	**0.1731**	0.1784	0.2123	0.2281	0.2431
	2500	0.1972	0.2303	0.1538	0.2247	0.1311	**0.1306**	0.1307	0.141	0.1469	0.1754
13	50	0.486	0.4833	0.5659	**0.466**	0.511	0.4862	0.4838	0.6007	0.6696	0.7746
	100	**0.3148**	0.3191	0.411	0.32	0.3807	0.3668	0.3367	0.4085	0.4533	0.554
	500	0.1636	0.1654	0.2176	**0.1629**	0.1875	0.1872	0.1819	0.1717	0.1703	0.1688
	1000	**0.1113**	0.1295	0.1413	0.1126	0.1207	0.1287	0.1259	0.1458	0.1453	0.1454
	2500	0.0684	0.0677	0.0785	**0.0674**	0.069	0.0796	0.0687	0.0937	0.1176	0.1271
14	50	0.7101	**0.7007**	0.9796	0.9523	0.982	0.9718	0.9025	0.7268	0.7024	0.7144
	100	0.6701	**0.63**	0.9521	0.9125	0.9487	0.9487	0.9368	0.8128	0.7221	0.6311
	500	0.626	**0.49**	0.8924	0.8524	0.8969	0.8969	0.8969	0.8969	0.8969	0.895
	1000	0.6104	**0.4465**	0.8755	0.8334	0.8784	0.8784	0.8784	0.8784	0.8784	0.8784
	2500	0.5833	**0.399**	0.856	0.8048	0.8602	0.8602	0.8602	0.8602	0.8602	0.8602

Tabelle 4.2.: Simulationsergebnisse für reguläre Histogramme – L_1-Risiko für die Dichten 15-28.

Dichte	n	BR	BIC	L2CV	SC	E-LR	V	KS .5	KS .9	KS .95	KS .99
15	50	0.3933	0.395	0.4412	0.4421	0.4229	0.394	**0.361**	0.3772	0.3875	0.4107
	100	0.3173	0.3283	0.3568	0.3275	0.3265	0.3063	**0.2877**	0.3055	0.3183	0.3468
	500	0.1901	0.2086	0.2058	0.1905	0.1732	0.1752	**0.168**	0.1801	0.1889	0.2051
	1000	0.1517	0.1699	0.1647	0.1552	0.1388	0.1423	**0.1363**	0.1457	0.152	0.1661
	2500	0.1114	0.1295	0.1263	0.1167	0.1024	0.1075	**0.1016**	0.1078	0.1124	0.1207
16	50	0.3414	0.3526	0.3587	0.4989	0.3317	**0.3085**	0.3137	0.3728	0.3787	0.3837
	100	0.2665	0.2855	0.2875	0.3151	**0.2574**	0.2577	0.261	0.3003	0.3392	0.3866
	500	0.1559	0.1666	0.1621	**0.1548**	0.163	0.1592	0.1846	0.2312	0.2448	0.2542
	1000	**0.1241**	0.1369	0.1283	0.1256	0.1412	0.1322	0.1578	0.1893	0.2132	0.2353
	2500	**0.0902**	0.1001	0.0931	0.0924	0.1234	0.1059	0.1292	0.1638	0.1712	0.1767
17	50	0.3031	0.2912	0.3448	0.4713	0.3406	0.315	0.2913	0.2845	0.2802	**0.2759**
	100	0.2686	0.2815	0.2847	0.3315	0.2415	**0.2407**	0.2471	0.2838	0.2925	0.2937
	500	0.1476	0.1536	0.1523	0.1481	0.1487	**0.1449**	0.1617	0.1894	0.1993	0.2193
	1000	**0.1185**	0.1239	0.1232	0.1193	0.1285	0.1233	0.1394	0.1679	0.1744	0.1827
	2500	0.0865	0.0912	0.0905	**0.0864**	0.1053	0.0957	0.1093	0.1299	0.1397	0.149
18	50	0.6231	0.6231	0.6834	0.6331	0.6845	0.6747	0.6303	0.5815	**0.576**	0.5895
	100	0.5334	0.5457	0.6068	0.5292	0.6085	0.6014	0.5691	0.5075	0.4984	**0.497**
	500	0.3684	0.4054	0.5173	0.3821	0.511	0.5217	0.4828	0.3948	0.3721	**0.3505**
	1000	**0.3096**	0.3511	0.49	0.3251	0.4806	0.4955	0.4541	0.362	0.3399	0.3122
	2500	**0.2441**	0.2883	0.4532	0.2687	0.4416	0.4599	0.4279	0.3327	0.3091	0.2784
19	50	0.8994	0.9002	0.8718	0.8916	0.8675	0.8676	0.8567	0.8386	**0.8384**	0.8651
	100	0.794	0.813	0.7761	0.7862	0.7811	0.7811	0.7756	0.7435	0.7415	**0.7403**
	500	0.5958	0.6585	0.5934	0.6101	0.602	0.602	0.602	0.6012	0.5949	**0.5722**
	1000	**0.5219**	0.5988	0.5297	0.5464	0.5427	0.5427	0.5427	0.5427	0.5424	0.534
	2500	**0.4363**	0.5165	0.4609	0.4695	0.4695	0.4695	0.4695	0.4695	0.4695	0.4695
20	50	1.5439	1.5645	1.4511	1.4786	**1.4489**	**1.4489**	**1.4489**	**1.4489**	1.4498	1.4501
	100	1.5632	1.6189	1.4924	1.5068	**1.4918**	**1.4918**	**1.4918**	**1.4918**	**1.4918**	**1.4918**
	500	1.6919	1.7704	1.6664	1.668	**1.6663**	**1.6663**	**1.6663**	**1.6663**	**1.6663**	**1.6663**
	1000	1.719	1.8003	**1.7008**	1.7029	**1.7008**	**1.7008**	**1.7008**	**1.7008**	**1.7008**	**1.7008**
	2500	1.7855	1.8581	1.7781	1.7783	**1.778**	**1.778**	**1.778**	**1.778**	**1.778**	**1.778**
21	50	0.6094	0.6391	0.4559	0.4975	**0.4218**	0.4301	0.4678	0.5962	0.6473	0.7486
	100	0.4145	0.4675	0.3906	0.4009	**0.3597**	0.3622	0.3947	0.4714	0.5038	0.5785
	500	0.2998	0.3604	0.2378	0.3412	0.2283	**0.2242**	0.2448	0.3149	0.3463	0.3978
	1000	0.2072	0.2978	0.1877	0.2643	0.1962	**0.185**	0.2073	0.2375	0.2529	0.2885
	2500	0.15	0.1946	**0.1342**	0.1842	0.1455	0.1377	0.1486	0.197	0.2037	0.2162
22	50	0.3613	0.3649	0.4069	0.4651	0.3661	0.3457	**0.3368**	0.3717	0.383	0.3899
	100	0.3019	0.3088	0.3237	0.3572	0.2788	**0.2773**	0.3006	0.3483	0.3672	0.3942
	500	0.1757	0.1848	0.1792	0.1761	0.1739	**0.1693**	0.1887	0.2303	0.2421	0.2681
	1000	**0.137**	0.1527	0.1398	0.14	0.146	0.1399	0.1557	0.1915	0.206	0.2259
	2500	**0.101**	0.1152	0.1024	0.1067	0.1177	0.1056	0.1209	0.1445	0.1522	0.1665
23	50	0.4901	0.4907	0.5382	0.5974	0.4825	**0.4808**	0.4859	0.5556	0.5745	0.6016
	100	0.4182	0.4264	0.4406	0.4292	**0.4091**	0.4138	0.4379	0.504	0.5212	0.5522
	500	0.2635	0.3383	**0.2634**	0.2692	0.2945	0.2708	0.3404	0.3712	0.3763	0.4005
	1000	0.2124	0.2688	**0.2086**	0.2235	0.237	0.2215	0.2868	0.3392	0.3512	0.3636
	2500	0.1615	0.1965	**0.1526**	0.1807	0.1989	0.1883	0.1988	0.2826	0.3026	0.3237
24	50	0.6017	0.6063	0.5957	0.6174	**0.5395**	0.5545	0.6142	0.6849	0.6946	0.7053
	100	**0.4709**	0.4964	0.4797	0.5017	0.4729	0.484	0.5507	0.6468	0.6668	0.6891
	500	0.2746	0.3127	0.2804	**0.272**	0.3423	0.3227	0.3886	0.465	0.483	0.5136
	1000	**0.2237**	0.2527	0.2342	0.2258	0.2926	0.2747	0.3187	0.3908	0.4053	0.4626
	2500	**0.1695**	0.193	0.1736	0.1731	0.2513	0.2062	0.2577	0.3075	0.3204	0.3557
25	50	0.5587	0.5616	0.5102	0.608	**0.469**	0.4855	0.5477	0.5984	0.609	0.614
	100	0.3864	0.4253	0.4034	0.4174	**0.3751**	0.3923	0.4435	0.547	0.5754	0.6066
	500	**0.2112**	0.2142	0.2328	0.2114	0.2237	0.2207	0.2403	0.3038	0.3229	0.4008
	1000	0.1753	0.1864	0.1848	0.1777	0.181	**0.1748**	0.1917	0.2413	0.2575	0.2832
	2500	**0.1243**	0.1533	0.1333	0.128	0.1359	0.129	0.1396	0.1794	0.1851	0.2121
26	50	1.0414	1.0414	1.1031	**1.0317**	1.1277	1.1277	1.1277	1.1277	1.1282	1.1658
	100	0.9848	1.0292	**0.9273**	0.9714	0.9629	0.9629	0.9629	0.9629	0.9629	0.9631
	500	0.5396	0.8523	**0.168**	0.4211	0.2064	0.2074	0.2357	0.3028	0.3311	0.3774
	1000	**0.1131**	0.6858	0.1228	0.117	0.1584	0.1555	0.1742	0.2228	0.2403	0.2778
	2500	**0.0669**	0.2528	0.0787	**0.0669**	0.1081	0.1027	0.1142	0.1503	0.1597	0.1839
27	50	0.5205	0.5163	0.5983	0.7265	0.5841	0.5627	0.5314	0.5169	0.515	**0.5131**
	100	0.5052	0.5003	0.5272	0.6088	0.531	0.5265	0.5093	0.5008	0.4987	**0.498**
	500	0.3181	0.4255	0.3201	**0.3156**	0.4783	0.4532	0.4925	0.4912	0.4909	0.4904
	1000	0.255	0.3026	0.2573	**0.2547**	0.429	0.377	0.4843	0.4919	0.4918	0.4917
	2500	**0.186**	0.2265	0.187	0.1897	0.3629	0.3235	0.378	0.4917	0.4938	0.4939
28	50	0.348	0.3434	0.4288	0.5639	0.407	0.3748	0.3413	0.3222	0.3215	**0.3171**
	100	0.3054	0.3101	0.3305	0.424	0.302	0.2916	**0.278**	0.3006	0.3033	0.3058
	500	0.1836	0.1977	0.1977	0.1839	**0.1648**	0.1685	0.169	0.1952	0.206	0.2269
	1000	0.1451	0.1568	0.1603	0.1472	**0.1336**	0.1352	0.1346	0.1555	0.1641	0.1832
	2500	0.1101	0.121	0.1225	0.112	0.1039	**0.1026**	0.1049	0.119	0.1253	0.1376

Tabelle 4.3.: Simulationsergebnisse für reguläre Histogramme – arithmetisches Mittel der Anzahl der gewählten Bins für Dichten 1-14.

Dichte	n	BR	BIC	L2CV	SC	E-LR	V	KS .5	KS .9	KS .95	KS .99
1	50	1.13	1.02	2.6	12.26	5.21	3.66	1.83	1.07	1.02	1
	100	1.12	1.04	2.3	9.84	4.48	3.88	1.77	1.1	1.02	1.01
	500	1.13	1.03	2.67	1.03	2.92	3.48	1.75	1.13	1.05	1.02
	1000	1.17	1.01	2.7	1.01	2.45	3.48	1.8	1.1	1.04	1.01
	2500	1.16	1	2.33	1	1.98	3.7	1.75	1.08	1.02	1.02
2	50	4.24	4.26	12.95	6.05	15.4	11.94	7.28	4.42	4	3.34
	100	5.76	5.46	15.53	6.34	13.66	12.47	8.38	5.69	5.23	4.55
	500	12.06	9.31	27.8	10.56	20.27	22.19	15.75	11.21	10.35	9.02
	1000	16.16	11.96	35.32	13.31	22.53	26.05	19.71	14.48	13.42	11.77
	2500	24.33	16.01	49.98	17.48	28.79	34.55	27.23	20.6	19.12	16.86
3	50	2.48	2.39	5.29	10.88	6.6	4.82	2.85	1.89	1.76	1.51
	100	3.92	3.61	7.65	5.87	6.63	5.93	4.05	2.59	2.26	1.98
	500	9.57	7.4	15.64	9.21	8.85	9.51	7.16	5.51	5.14	4.47
	1000	13.14	9.94	18.26	12.03	10.24	11.26	8.99	6.96	6.52	5.82
	2500	20.28	14.1	29.52	16.65	12.68	14.91	12.12	9.66	9.11	8.16
4	50	4.27	4.39	10.57	6.57	9.53	7.13	4.93	3.36	3.11	2.44
	100	5.65	5.26	13.28	6.82	10.69	9.91	6.34	4.27	4.02	3.49
	500	14.36	11.42	31.2	13.41	17.75	18.81	13.83	9.75	8.84	7.62
	1000	21.74	14.8	45.24	17.34	22.39	25.69	18.8	13.41	11.88	10.41
	2500	34.35	21.83	64.98	25	30.24	37.86	27.87	20.74	18.85	16.19
5	50	3.52	3.51	7.71	8.64	7.56	5.67	3.92	2.72	2.45	1.83
	100	4.98	4.67	10.23	6.33	8.13	7.2	4.92	3.55	3.33	3.1
	500	11.98	9.19	21.44	11.07	11.66	12.74	9.51	6.97	6.1	5.12
	1000	16.72	12.14	25.22	14.12	14	15.8	12.26	9.24	8.58	7.55
	2500	25.05	17.39	39.26	19.3	17.85	21.45	16.8	13.1	12.17	11
6	50	13.39	11.78	35.83	18	35.06	32.21	26.49	20.28	18.59	15.95
	100	26.14	20.63	75.19	31.72	70.74	67.13	56.9	45.25	42.3	36.68
	500	142.02	80.57	427.29	142.68	389.92	405.03	342.36	278.36	265.77	246.84
	1000	285.06	142.44	887.59	269.22	826.08	857.42	784.5	600.4	568.77	530.87
	2500	724.63	329.87	2337.04	655.81	2222.16	2297.93	2192.89	2007.57	1919.4	1658.34
7	50	3.03	2.99	7.18	9.04	8.2	5.96	3.53	2.37	2.11	1.85
	100	4.89	4.44	10.39	7.52	8.51	7.68	5.45	3.51	2.96	2.48
	500	12.21	10.33	19.92	11.54	11.31	12.33	9.39	7.3	6.89	6.37
	1000	16.81	12.78	25.74	14.75	14.06	15.63	12.15	8.9	8.31	7.69
	2500	25.87	17.96	41.19	20.3	18.14	20.98	17.23	13.63	12.45	10.82
8	50	3.97	3.85	17.13	18.75	28.05	21.57	10.58	4.72	3.64	2.47
	100	5.43	4.7	36.4	26.32	45.72	39.53	19.64	7.67	6.01	4
	500	17.62	9.71	142.17	28.72	110.92	132.03	62.99	28.06	22.36	15.24
	1000	28.64	13.84	276.29	32.69	178.99	245.48	121.6	54.14	43.02	29.26
	2500	64.16	24.1	667.65	53.05	319.77	544.08	270.53	125.09	99.54	68.66
9	50	25.47	20.58	49.61	43.89	50	50	50	49.78	49.42	48.62
	100	61.99	37.73	99.59	91.47	100	100	100	100	100	99.82
	500	415.62	152.51	499.94	493.2	500	500	500	500	500	500
	1000	867.48	272.99	999.92	992.46	1000	1000	1000	1000	1000	1000
	2500	2289.38	617.24	2499.98	2495.65	2500	2500	2500	2500	2500	2500
10	50	25.42	20.4	48.11	40.9	49.65	49.65	49.13	47.53	46.76	45.28
	100	59.47	36.6	98.42	86.84	100	100	99.8	99.5	99.43	98.72
	500	400.54	146.06	497.97	479.11	500	500	500	500	500	500
	1000	868.98	273.66	997.94	989.58	1000	1000	1000	1000	1000	1000
	2500	2267.35	617.42	2496.48	2484.3	2500	2500	2500	2500	2500	2500
11	50	3.36	3.36	6.96	10.95	6.33	4.91	3.31	2.12	1.88	1.47
	100	4.78	4.4	8.38	6.66	6.9	6.03	4.05	3.08	2.88	2.46
	500	9.9	7.97	15.66	9.34	8.96	9.6	7.29	5.44	4.96	3.97
	1000	13.37	9.95	18.17	11.84	10.01	11.5	8.85	6.75	6.33	5.67
	2500	19.97	13.61	26.82	15.98	12.2	14.95	11.51	8.97	8.41	7.41
12	50	5.91	5.83	15.5	7.38	18	14.13	10.07	7.09	6.36	5.43
	100	8.96	8.09	22.37	9.59	20.6	17.95	12.66	9.89	9.18	7.99
	500	19.83	15.88	104.36	16.9	62.97	71.19	38.81	19.94	18.97	17.58
	1000	26.77	20.87	183.83	22.1	104.53	120.8	84.92	37.28	26.76	22.77
	2500	41.3	30.05	324.23	31.18	176.61	209.57	166.84	119.91	107.29	74.31
13	50	7.57	7.47	20.82	12.47	17.19	14.09	8.83	5.79	4.89	3.37
	100	9.29	8.71	24.51	11.62	18.37	15.81	10.77	7.63	6.94	5.66
	500	12.14	9	40.51	13.09	21.13	23.52	13.92	9.98	9.43	9.02
	1000	19.2	13.09	34.88	18.27	21.48	25.5	16.95	9.63	9.25	9.06
	2500	20.43	19.78	30.23	19.98	20.99	24.69	20.79	15.59	11.64	9.46
14	50	8.82	8.22	46.13	39.27	49.97	48.8	40.4	16.46	9.46	4.26
	100	19.97	12.86	97.7	79.27	100	100	97.56	67.84	45.75	19.16
	500	152.24	37	497.16	417.71	500	500	500	500	500	498
	1000	327.91	61.32	997.91	851.56	1000	1000	1000	1000	1000	1000
	2500	835.47	125.1	2495.13	2106.5	2500	2500	2500	2500	2500	2500

Tabelle 4.4.: Simulationsergebnisse für reguläre Histogramme – arithmetisches Mittel der Anzahl der gewählten Bins für Dichten 15-28.

Dichte	n	BR	BIC	L2CV	SC	E-LR	V	KS .5	KS .9	KS .95	KS .99
15	50	3.18	3.22	8.93	7.28	12.04	9.55	5.47	3.18	2.81	2.3
	100	3.96	3.57	13.16	7.15	15.32	12.39	7.34	4.36	3.75	3.03
	500	8.72	6.57	26.23	9.02	18.31	20.55	13.58	8.73	7.7	6.31
	1000	13.17	8.65	34.25	11.58	22.5	27.56	18.36	11.77	10.4	8.62
	2500	20.05	11.99	54.87	15.34	30.43	40.61	27.6	17.93	15.96	13.12
16	50	2.43	2.22	4.52	12.47	5.6	4.21	2.76	1.56	1.3	1.09
	100	3.82	3.21	6.6	9.3	5.56	4.86	3.44	2.56	2.1	1.39
	500	7.9	6.44	11.27	8.19	6.62	7.12	5.19	3.95	3.49	3.1
	1000	10.31	8.06	13.75	10.06	7.31	8.24	6.24	4.81	4.42	3.96
	2500	16.92	11.36	18.34	14.01	8.12	10.11	7.6	5.79	5.33	5
17	50	1.73	1.42	3.5	11.8	5.25	3.79	2.17	1.24	1.11	1
	100	2.49	1.91	5.77	9.67	4.89	4.51	3.05	1.63	1.36	1.08
	500	7.41	5.92	9.38	7.56	6.08	6.64	4.89	3.77	3.55	3.05
	1000	10.21	7.77	12.19	9.76	6.57	7.49	5.75	4.4	4.17	3.92
	2500	14.84	10.87	17.09	13.45	7.71	9.13	7.32	5.81	5.32	4.92
18	50	7.42	7.22	35.14	12.99	46.12	42.7	32.32	17.06	13.38	9.23
	100	12.36	9.84	71.15	14.54	88.88	86.22	67.58	35.59	27.59	18.62
	500	39.65	21.45	412.84	30.72	448.63	467.75	385.73	196.83	154.8	104.8
	1000	65.4	31.28	872.94	45.07	905.13	970.24	784.74	398.59	317.62	219.73
	2500	135.07	52.34	2240.59	70.98	2245.82	2468.59	2082.4	1058.48	846.92	588.89
19	50	11.9	11.16	46.36	23.36	49.92	49.58	46.1	27.77	21.04	12.98
	100	22.87	17.18	95.53	35.73	100	100	97.79	70.74	57.23	38.28
	500	99.16	47.74	489.57	80.45	500	500	500	496.92	480.85	386.68
	1000	181.68	72.95	988.12	123.66	1000	1000	1000	1000	998.62	948.92
	2500	372.44	133.77	2484	219.26	2500	2500	2500	2500	2500	2500
20	50	27.96	22.41	49.8	45.03	50	50	50	50	49.87	49.78
	100	64.63	37.67	99.9	94.74	100	100	100	100	100	100
	500	414.3	151.03	499.97	496.76	500	500	500	500	500	500
	1000	875.74	277.08	999.98	992.51	1000	1000	1000	1000	1000	1000
	2500	2318.22	617.27	2499.95	2497.47	2500	2500	2500	2500	2500	2500
21	50	12.58	11.42	32.23	18.26	30	25.63	19.68	13.7	11.9	8.93
	100	19.51	16.68	49.9	20.69	34.64	30.74	20.88	16.47	15.14	12.95
	500	30.09	20.41	97.46	22.81	51.26	53.28	45.77	31.71	26.3	19.66
	1000	54.39	29.77	127.12	36.78	57.21	70.82	48.81	43.26	41.25	36.12
	2500	80.7	49.65	170.01	56.46	81.44	93.3	78.04	45.05	43.37	41.61
22	50	2.59	2.54	6.05	8.97	6.46	5.06	3.09	1.76	1.55	1.29
	100	4.65	3.99	9.35	13.69	6.89	6.36	4.24	2.7	2.37	1.95
	500	10.53	8.34	16.21	10.49	9.12	9.98	7.11	5.42	5.15	4.73
	1000	16.06	10.73	22.77	14.01	11.21	12.84	9.41	6.77	6.08	5.51
	2500	23.68	16.16	31.69	19.36	15.37	18.54	14.6	10.1	9.27	8.19
23	50	3.5	3.48	10.94	16.09	8.06	5.99	3.73	2.13	1.84	1.46
	100	5.14	4.81	16.51	11.12	9.4	7.99	4.63	3.22	3	2.66
	500	21.99	9.25	44.4	20.26	18.34	20.27	11.17	6.7	6.12	5.26
	1000	31.76	16.82	61.33	25.26	23.4	25.87	19.11	9.75	8.66	7.5
	2500	54.27	28.1	87.69	35.28	28.62	36.04	27.35	20.2	16.64	11.87
24	50	4.55	4.22	13.42	17.89	9.14	7.14	4.26	1.94	1.6	1.27
	100	10.03	8.06	17.14	22.03	9.99	8.94	5.98	3.06	2.47	1.71
	500	28.28	16.93	38.5	29.42	14.45	16.4	10.82	7.82	7.21	6.32
	1000	40.03	23.87	55.34	34.57	18.25	19.72	15.61	10.28	9.38	8.29
	2500	81.29	32.86	92.73	51.5	20.69	29.48	20.1	16.53	14.8	12.72
25	50	3	2.86	11.62	20.07	9.54	7.16	3.61	1.56	1.25	1.05
	100	9.15	7.21	17.35	20.15	11.06	10.07	6.41	2.92	2.36	1.63
	500	17.51	13.19	35	18.64	16.6	18.8	11.96	8.9	8.34	6.63
	1000	23.42	13.48	48.79	20.48	20.71	25.84	15.63	10.26	9.57	8.86
	2500	46.81	18.23	70.07	34.58	29.83	40.92	27.5	12.8	11.66	10.55
26	50	19.5	19.5	46.86	31.38	50	50	50	50	49.41	43.81
	100	40	19.93	99.22	60.27	100	100	100	100	100	99.31
	500	223.95	81.07	407.09	277.33	385.28	394.25	363.59	329.36	318.24	298.46
	1000	400.87	147.93	421.36	399.31	393.18	420.31	380.97	351.13	343.17	328.74
	2500	402	320.56	487.58	402	387.16	431.27	384.29	369.27	364.27	355.22
27	50	1.26	1.12	9.18	27.35	7.36	4.72	2.03	1.13	1.07	1.02
	100	1.85	1.13	20.67	40.97	6.92	5.65	2.31	1.27	1.09	1.02
	500	37.24	11.19	45.49	45.51	9.77	13.42	3.81	1.32	1.19	1.04
	1000	50.19	30.9	57.9	52.01	14.94	19.98	7.09	1.58	1.21	1.06
	2500	76.48	43.25	80.44	68.16	21.49	25.64	20.18	3.34	2.09	1.28
28	50	2.08	1.83	7.01	16.8	10.26	7.87	3.9	1.58	1.29	1.03
	100	2.94	2.22	8.66	18.37	10.15	8.76	4.67	2.05	1.71	1.29
	500	7.65	5.42	19.36	8.86	11.88	13.42	8.1	4.73	4.18	3.32
	1000	11.45	7.52	26.32	11.05	14.45	17.43	11.25	6.65	5.78	4.7
	2500	17.79	10.61	40.46	14.82	17.45	24.15	15.77	9.85	8.74	7.1

Tabelle 4.5.: Simulationsergebnisse für reguläre Histogramme (alternative GoF-Tests) – L_1-Risiko für die Dichten 1-14.

Dichte	n	Kuip .5	Kuip .9	Kuip .95	Kuip .99	CvM V	CvM .5	CvM .9	CvM .95	CvM .99
1	50	0.1729	0.093	0.091	0.0857	0.2634	0.1557	0.0921	0.0886	0.0857
	100	0.1033	0.0631	0.0527	0.0429	0.1654	0.097	0.0501	0.0459	0.041
	500	0.0399	0.0149	0.0111	0.0099	0.0649	0.0348	0.0138	0.0114	0.0103
	1000	0.0293	0.0081	0.0073	0.0053	0.0487	0.0276	0.0109	0.0096	0.0053
	2500	0.0165	0.0041	0.0027	0.002	0.0278	0.0147	0.004	0.0029	0.0024
2	50	0.3614	0.3927	0.4102	0.4344	0.3708	0.3405	0.3666	0.375	0.4094
	100	0.2865	0.3192	0.3351	0.3662	0.2682	0.2655	0.2966	0.3093	0.3332
	500	0.1735	0.2056	0.2157	0.2287	0.1527	0.1632	0.1919	0.2014	0.2201
	1000	0.147	0.1718	0.1786	0.1922	0.1266	0.1364	0.1613	0.1699	0.1852
	2500	0.1143	0.1336	0.14	0.1511	0.0941	0.1049	0.1263	0.1333	0.1471
3	50	0.316	0.3459	0.3589	0.3788	0.3236	0.3196	0.3457	0.3527	0.3791
	100	0.2622	0.2871	0.2971	0.3176	0.2579	0.2695	0.322	0.3317	0.3418
	500	0.1636	0.1844	0.1899	0.2001	0.1585	0.1814	0.2153	0.2298	0.2614
	1000	0.1345	0.1532	0.1601	0.1693	0.1323	0.1515	0.1879	0.1974	0.2167
	2500	0.1047	0.1217	0.1272	0.1324	0.1036	0.1226	0.1565	0.1644	0.1789
4	50	0.3877	0.4228	0.4378	0.4498	0.3829	0.3992	0.4533	0.4752	0.56
	100	0.3266	0.3655	0.3763	0.4008	0.31	0.3412	0.4067	0.4159	0.4448
	500	0.2039	0.2316	0.2411	0.2586	0.1956	0.2217	0.2731	0.2908	0.3253
	1000	0.1668	0.1922	0.1992	0.2118	0.1608	0.1849	0.2294	0.2457	0.2708
	2500	0.1295	0.1491	0.1541	0.1645	0.1248	0.1456	0.1801	0.193	0.2171
5	50	0.3496	0.3696	0.3957	0.4231	0.3546	0.3628	0.4243	0.4561	0.5521
	100	0.3008	0.3236	0.3367	0.3436	0.286	0.313	0.356	0.3673	0.3937
	500	0.1859	0.2114	0.2189	0.2311	0.1742	0.2017	0.2579	0.2767	0.3114
	1000	0.1525	0.1737	0.1834	0.1954	0.1462	0.1681	0.2137	0.2245	0.2596
	2500	0.12	0.1401	0.1458	0.1573	0.1135	0.1358	0.1727	0.1833	0.2029
6	50	0.5979	0.6336	0.6481	0.6853	0.5928	0.6066	0.6708	0.6945	0.7527
	100	0.5305	0.5519	0.5588	0.5791	0.5243	0.5402	0.5902	0.6107	0.6598
	500	0.4398	0.4534	0.4575	0.4688	0.4381	0.4482	0.477	0.4913	0.512
	1000	0.3894	0.3955	0.3996	0.4046	0.3882	0.3935	0.4186	0.4249	0.4409
	2500	0.3529	0.3568	0.3586	0.3609	0.352	0.3562	0.3691	0.3774	0.3945
7	50	0.3469	0.381	0.3916	0.4163	0.3484	0.3564	0.3993	0.4089	0.4319
	100	0.2861	0.3113	0.3259	0.3559	0.2804	0.2963	0.3703	0.3929	0.414
	500	0.184	0.2072	0.2134	0.2241	0.1738	0.2011	0.2461	0.2565	0.2905
	1000	0.1483	0.1712	0.1792	0.1985	0.143	0.1665	0.2193	0.2326	0.2517
	2500	0.1158	0.1345	0.1388	0.1476	0.1121	0.1335	0.1665	0.1765	0.2072
8	50	0.461	0.434	0.4413	0.4629	0.5127	0.4555	0.4258	0.4243	0.4348
	100	0.3923	0.3682	0.3688	0.3821	0.4353	0.3755	0.361	0.3614	0.3751
	500	0.2866	0.2583	0.252	0.252	0.3019	0.26	0.246	0.2488	0.2549
	1000	0.2539	0.2211	0.2154	0.2113	0.2611	0.224	0.2091	0.2089	0.2176
	2500	0.2231	0.1872	0.1816	0.1751	0.2158	0.1845	0.168	0.1686	0.1721
9	50	1.3435	1.3453	1.3473	1.3492	1.3428	1.3428	1.3443	1.3443	1.3472
	100	1.4188	1.4188	1.4195	1.4198	1.4188	1.4188	1.4188	1.4194	1.4195
	500	1.6332	1.6332	1.6332	1.6332	1.6332	1.6332	1.6332	1.6332	1.6332
	1000	1.6713	1.6713	1.6713	1.6713	1.6713	1.6713	1.6713	1.6713	1.6713
	2500	1.7467	1.7467	1.7467	1.7467	1.7467	1.7467	1.7467	1.7467	1.7467
10	50	1.2985	1.297	1.2981	1.3022	1.2987	1.2965	1.301	1.304	1.3097
	100	1.3873	1.3863	1.3864	1.3867	1.3873	1.3863	1.387	1.3858	1.3847
	500	1.5988	1.5988	1.5988	1.5988	1.5988	1.5988	1.5988	1.5988	1.5988
	1000	1.694	1.694	1.694	1.694	1.694	1.694	1.694	1.694	1.694
	2500	1.7516	1.7516	1.7516	1.7516	1.7516	1.7516	1.7516	1.7516	1.7516
11	50	0.3453	0.3636	0.3728	0.4061	0.3406	0.346	0.3983	0.4499	0.5039
	100	0.2798	0.308	0.3122	0.3252	0.266	0.2918	0.33	0.3393	0.38
	500	0.1764	0.2023	0.2147	0.223	0.1673	0.1927	0.2516	0.2703	0.3163
	1000	0.1468	0.1702	0.1787	0.1908	0.1377	0.1653	0.2047	0.2181	0.2554
	2500	0.1161	0.1352	0.1413	0.1517	0.1085	0.1318	0.1666	0.1773	0.1982
12	50	0.3959	0.4383	0.4553	0.4937	0.3794	0.3746	0.4062	0.4224	0.4681
	100	0.3118	0.3525	0.3718	0.4002	0.2986	0.3019	0.3379	0.3526	0.3803
	500	0.2115	0.2306	0.2403	0.2569	0.2112	0.2167	0.2493	0.2598	0.2777
	1000	0.1746	0.1836	0.1906	0.2045	0.1726	0.1833	0.2206	0.2306	0.2483
	2500	0.1311	0.1379	0.1412	0.149	0.1322	0.1416	0.1657	0.1755	0.2043
13	50	0.4511	0.5166	0.5235	0.5788	0.4743	0.4967	0.6887	0.7278	0.8459
	100	0.3286	0.3493	0.3669	0.4066	0.3517	0.3474	0.4855	0.5604	0.6672
	500	0.1769	0.1698	0.1687	0.1648	0.1853	0.1751	0.1692	0.1893	0.3236
	1000	0.1124	0.1328	0.1438	0.1462	0.1284	0.1451	0.1456	0.1454	0.1457
	2500	0.0687	0.0674	0.0681	0.0763	0.0704	0.111	0.1243	0.1282	0.1289
14	50	0.9482	0.8565	0.8214	0.7576	0.9443	0.8151	0.6705	0.6606	0.7017
	100	0.9487	0.9336	0.9232	0.8755	0.9436	0.8843	0.6367	0.5989	0.5897
	500	0.8969	0.8969	0.8969	0.8969	0.8969	0.8969	0.8652	0.7747	0.5883
	1000	0.8784	0.8784	0.8784	0.8784	0.8784	0.8784	0.8784	0.874	0.7626
	2500	0.8602	0.8602	0.8602	0.8602	0.8602	0.8602	0.8602	0.8602	0.8602

Tabelle 4.6.: Simulationsergebnisse für reguläre Histogramme (alternative GoF-Tests) – L_1-Risiko für die Dichten 15-28.

Dichte	n	Kuip .5	Kuip .9	Kuip .95	Kuip .99	CvM V	CvM .5	CvM .9	CvM .95	CvM .99
15	50	0.3758	0.3957	0.4053	0.426	0.3825	0.3554	0.3761	0.3832	0.4098
	100	0.2956	0.3302	0.3444	0.3708	0.2979	0.283	0.3059	0.319	0.3486
	500	0.1751	0.193	0.2035	0.2215	0.167	0.1708	0.1904	0.1987	0.2195
	1000	0.1394	0.1576	0.1632	0.1769	0.1352	0.1385	0.1575	0.1658	0.1829
	2500	0.1049	0.1133	0.1184	0.1287	0.1019	0.1046	0.1208	0.1268	0.1401
16	50	0.3043	0.3349	0.3458	0.3536	0.3065	0.3072	0.3633	0.3812	0.3845
	100	0.2551	0.2656	0.2671	0.281	0.2499	0.2579	0.2927	0.3193	0.3827
	500	0.1659	0.19	0.1979	0.2165	0.1569	0.1805	0.2382	0.2515	0.2564
	1000	0.1408	0.1663	0.1709	0.1771	0.1326	0.1546	0.1829	0.1989	0.2449
	2500	0.1121	0.1323	0.1372	0.1502	0.1059	0.1267	0.1667	0.1731	0.1758
17	50	0.2951	0.2861	0.2831	0.2787	0.3026	0.2911	0.2817	0.2791	0.2772
	100	0.239	0.2585	0.2636	0.2789	0.2393	0.2454	0.2804	0.2886	0.2952
	500	0.1506	0.1682	0.1731	0.1793	0.1482	0.1664	0.1978	0.2125	0.2213
	1000	0.1264	0.1415	0.1473	0.1565	0.1236	0.1441	0.1765	0.1795	0.1985
	2500	0.0989	0.1113	0.116	0.122	0.0974	0.1149	0.1452	0.1482	0.1765
18	50	0.6191	0.5827	0.5845	0.5953	0.6565	0.6098	0.5757	0.5693	0.582
	100	0.5433	0.4985	0.4986	0.5048	0.5842	0.5321	0.497	0.4907	0.503
	500	0.4296	0.3662	0.3569	0.346	0.4621	0.3911	0.3449	0.3405	0.3409
	1000	0.3929	0.3306	0.3191	0.3027	0.4083	0.3448	0.2986	0.2927	0.2889
	2500	0.3608	0.2954	0.2817	0.2613	0.3507	0.2902	0.2458	0.24	0.2342
19	50	0.868	0.852	0.8477	0.8484	0.8668	0.8447	0.8374	0.8438	0.8787
	100	0.7811	0.7703	0.7614	0.757	0.7796	0.7652	0.7334	0.7371	0.7586
	500	0.602	0.602	0.602	0.602	0.602	0.602	0.5765	0.5647	0.5555
	1000	0.5427	0.5427	0.5427	0.5427	0.5427	0.5427	0.5229	0.5093	0.4924
	2500	0.4695	0.4695	0.4695	0.4695	0.4695	0.4695	0.4618	0.4477	0.4216
20	50	1.4489	1.4503	1.4501	1.451	1.4489	1.4489	1.4492	1.4499	1.4501
	100	1.4918	1.4918	1.4918	1.4918	1.4918	1.4918	1.4918	1.4918	1.4918
	500	1.6663	1.6663	1.6663	1.6663	1.6663	1.6663	1.6663	1.6663	1.6663
	1000	1.7008	1.7008	1.7008	1.7008	1.7008	1.7008	1.7008	1.7008	1.7008
	2500	1.778	1.778	1.778	1.778	1.778	1.778	1.778	1.778	1.778
21	50	0.4681	0.5509	0.5961	0.6661	0.43	0.4943	0.6851	0.7519	0.8653
	100	0.3786	0.447	0.4664	0.5109	0.367	0.4176	0.5494	0.6012	0.7205
	500	0.2308	0.2567	0.2695	0.2938	0.2333	0.2688	0.3693	0.4037	0.4689
	1000	0.1929	0.2132	0.2171	0.2268	0.1994	0.224	0.282	0.3107	0.377
	2500	0.1409	0.1524	0.1625	0.1834	0.1454	0.1897	0.2201	0.2308	0.2574
22	50	0.3449	0.3613	0.3659	0.3823	0.3446	0.3424	0.3702	0.3785	0.3893
	100	0.2909	0.315	0.3236	0.3476	0.2824	0.3017	0.3621	0.3822	0.403
	500	0.1776	0.1973	0.206	0.2254	0.1724	0.1979	0.2589	0.2722	0.3016
	1000	0.1459	0.1608	0.1675	0.1822	0.1427	0.1621	0.2149	0.2314	0.2563
	2500	0.112	0.1296	0.1338	0.1401	0.1116	0.1307	0.1632	0.1766	0.207
23	50	0.4839	0.5146	0.529	0.5584	0.472	0.4986	0.5583	0.5724	0.6111
	100	0.4236	0.4536	0.4596	0.4851	0.4153	0.4465	0.5225	0.5338	0.5613
	500	0.3019	0.3515	0.3565	0.3695	0.2726	0.3451	0.3999	0.4078	0.4551
	1000	0.2246	0.3077	0.3259	0.3384	0.2193	0.2899	0.3597	0.3716	0.4149
	2500	0.1922	0.2032	0.2142	0.2579	0.1857	0.2028	0.2991	0.3209	0.3579
24	50	0.5879	0.6545	0.6753	0.6983	0.5542	0.6173	0.6991	0.7043	0.7061
	100	0.5114	0.5844	0.6089	0.6488	0.4753	0.5509	0.6654	0.686	0.7043
	500	0.3469	0.3948	0.4007	0.4323	0.3354	0.3959	0.497	0.5231	0.5594
	1000	0.2847	0.3328	0.3474	0.3768	0.2853	0.349	0.4182	0.4538	0.5109
	2500	0.2221	0.2611	0.2671	0.2811	0.2389	0.2774	0.3659	0.3726	0.4036
25	50	0.5193	0.573	0.5857	0.5999	0.4888	0.5493	0.6018	0.6118	0.614
	100	0.4111	0.4943	0.519	0.5403	0.4051	0.4763	0.5477	0.5642	0.6143
	500	0.2282	0.2557	0.2661	0.2886	0.234	0.2797	0.4226	0.4653	0.4951
	1000	0.1814	0.1995	0.2104	0.2389	0.2005	0.2145	0.3298	0.3642	0.4534
	2500	0.1309	0.1493	0.1583	0.1762	0.1477	0.1829	0.2172	0.2584	0.2971
26	50	1.1277	1.1277	1.1277	1.1277	1.1277	1.1277	1.1174	1.1346	1.4087
	100	0.9629	0.9629	0.9629	0.9629	0.9629	0.9629	0.962	0.9639	1
	500	0.2049	0.2447	0.2539	0.2755	0.2054	0.2586	0.3613	0.4007	0.4801
	1000	0.1543	0.1765	0.187	0.2012	0.1553	0.1885	0.2656	0.2919	0.35
	2500	0.1036	0.1174	0.1239	0.1358	0.1025	0.1252	0.174	0.1941	0.2289
27	50	0.5396	0.5184	0.5156	0.5133	0.5596	0.5306	0.5161	0.5153	0.5131
	100	0.5126	0.5027	0.5007	0.4988	0.5228	0.5073	0.4998	0.4989	0.4982
	500	0.4864	0.4925	0.4919	0.491	0.4309	0.4927	0.491	0.4905	0.4903
	1000	0.3934	0.4911	0.4916	0.4918	0.3656	0.4879	0.4918	0.4917	0.4916
	2500	0.3336	0.3755	0.4121	0.4836	0.3088	0.3692	0.494	0.4939	0.4938
28	50	0.3307	0.3283	0.327	0.3195	0.3564	0.327	0.3285	0.3236	0.319
	100	0.2798	0.2816	0.2821	0.2912	0.2737	0.2747	0.2965	0.3036	0.3051
	500	0.1633	0.173	0.177	0.189	0.1629	0.172	0.1993	0.2078	0.2269
	1000	0.1328	0.1362	0.1417	0.148	0.1331	0.1414	0.1655	0.1738	0.1908
	2500	0.1033	0.1064	0.1077	0.114	0.1039	0.1097	0.1301	0.1373	0.1505

5. Nichtparametrische Regression

5.1. Einführung und Definitionen

In diesem Kapitel wird die Verwendung des Diskrepanzprinzips in der nichtparametrischen Regression untersucht. Betrachtet wird der einfachste Fall der nichtparametrischen Regression, in dem n Beobachtungen

$$y_i = f(t_i) + \varepsilon_i \qquad (i = 1, ..., n) \tag{5.1}$$

vorliegen. Ziel ist die Schätzung von f durch einen geeigneten Schätzer $\hat{f}$. Die Designpunkte $0 \leq t_1 < \ldots < t_n \leq 1$ werden als fest und paarweise verschieden angenommen. Die Fehler $\varepsilon_1, \ldots, \varepsilon_n$ seien unabhängig identisch normalverteilt mit Varianz σ^2. In der Literatur werden häufig allgemeinere, auch heteroskedastische Fehler zugelassen. Die Varianz der Fehler ist in den meisten Anwendungen unbekannt und kann auf verschiedene Arten behandelt werden. Zum Beispiel kann σ auf der Basis von Residuen für eine Schätzung $\hat{f}$ für f geschätzt werden. Viele Methoden zur Glättungsparameterwahl, so auch die hier betrachteten Diskrepanzprinzipien, hängen jedoch von der Größe der Fehler ab. Deshalb wird σ hier als Störparameter betrachtet, der vor der Schätzung von f geschätzt werden muss. Ähnlich wie im Fall der Dichteschätzung kann die Güte eines Regressionsschätzers mittels Verlust- oder Risikofunktionen gemessen werden. Im Unterschied zur Dichteschätzung spielen in der Regression auch diskretisierte Versionen von Verlusten eine Rolle, bei denen Schätzung und wahre Funktion nur in den Designpunkten verglichen werden.

Bei der Untersuchung des Diskrepanzprinzips sollen wieder die Rolle der Parameterwahlmethode und die der betrachteten Klasse von Schätzern soweit wie möglich getrennt werden. Daher werden exemplarisch zwei der populärsten Methoden der nichtparametrischen Regression betrachtet: Nadaraya-Watson-Kernschätzer und Glättungssplines, beide jeweils mit globalem Glättungsparameter. Zum einen ist für diese Methoden klar, wie das Diskrepanzprinzip zu implementieren ist, zum anderen existieren weit verbreitete Vergleichsmethoden zur Parameterwahl, insbesondere Varianten der Kreuzvalidierung.

Im Folgenden werde das Modell (5.1) für äquidistante Designpunkte $t_i := (i-1)/n$, $i = 1, \ldots, n$, betrachtet. Sei K ein Kern wie in Kapitel 3 und $h > 0$ eine Bandbreite. Weiter sei $K_h(u) := \frac{1}{h} K\left(\frac{u}{h}\right)$. Der *Nadaraya-Watson-Kernschätzer* ist definiert durch

$$\hat{f}_h(x) := \begin{cases} \frac{\sum_{i=1}^n K_h(x-t_i) y_i}{\sum_{i=1}^n K_h(x-t_i)}, & \sum_{i=1}^n K_h(x-t_i) \neq 0 \\ 0 & \text{sonst} \end{cases}. \tag{5.2}$$

Dieser Schätzer wurde unabhängig von Nadaraya [Nad64] und Watson [Wat64] vorgeschlagen. Einen Überblick über Eigenschaften und Erweiterungen geben zum Beispiel [WJ95], [HMSW04] und [Tsy09]. Große Werte von h führen zu starker Glättung, während für $h \to 0$ die Schätzung die Beobachtungen interpoliert und zwischen den Beobachtungen verschwindet. Die optimale Wahl von h hängt von der verwendeten Verlustfunktion, der Stichprobengröße n, der unbekannten Funktion f und dem Rauschniveau σ ab. Während σ vergleichsweise einfach geschätzt werden kann (oder in manchen Anwendungen sogar bekannt ist), sorgt die Abhängigkeit von f dafür, dass eine optimale Bandbreite in der Praxis nicht implementierbar ist. Ähnlich wie im Fall der Dichteschätzung gibt es zahlreiche Methoden, den Glättungsparameter datengesteuert zu wählen. Viele, darunter die in der Simulationsstudie in Kapitel 5.3 verwendeten Vergleichsmethoden, versuchen, den *Average Squared Error*

$$ASE(h) = \frac{1}{n} \sum_{i=1}^{n} (f(t_i) - \hat{f}_h(t_i))^2, \tag{5.3}$$

eine diskrete Approximation für $d_2^2(f, \hat{f}_h)$, möglichst gut aus den Daten zu schätzen. Die Bandbreite h wird dann so gewählt, dass diese Schätzung minimal wird. Eine Übersicht über verschiedene Methoden der Glättungsparameterwahl findet man z.B. in Kapitel 4.3 von [HMSW04].

Für festes h ist der Nadaraya-Watson-Schätzer linear in den Daten: Mit den Gewichten

$$w_i^h(x) := \frac{K_h(x - t_i)}{\sum_{j=1}^{n} K_h(x - t_j)} \mathbb{I}(\sum_{j=1}^{n} K_h(x - t_j) \neq 0)$$

ergibt sich die Schätzung an einem Punkt $x \in [0, 1]$ als gewichtetes Mittel der Beobachtungen:

$$\hat{f}_h(x) = \sum_{i=1}^{n} w_i^h(x) y_i.$$

Im Folgenden bezeichne $y = (y_1, \ldots, y_n)$ den Beobachtungsvektor, $\hat{y}_h := (\hat{f}_h(t_1), \ldots, \hat{f}_h(t_n))$ den Vektor, dessen Komponenten die Auswertungen des Schätzers $\hat{f}_h$ in den Designpunkten enthalten und W_h die Gewichtsmatrix mit den Einträgen $w_{ij}^h := w_i(t_j)$. Dann gilt:

$$\hat{y}_h = W_h y. \tag{5.4}$$

Die Matrix W_h entspricht der sogenannten Hat-Matrix in der linearen Regression. Ihre Diagonale $w_h := (w_1^h, \ldots, w_n^h) := (w_{11}^h, \ldots, w_{nn}^h)$ spielt eine wichtige Rolle bei der datengesteuerten Wahl von h, da ihr i-ter Eintrag jeweils angibt, mit welchem Gewicht die i-te Beobachtung y_i in die Schätzung an der Stelle t_i eingeht.

Die zweite in diesem Kapitel betrachtete Klasse von nichtparametrischen Regressionsschätzern sind *kubische Glättungssplines*. Sie sind definiert als Lösung von

$$\hat{f}_\lambda = \underset{g \in W^{2,2}[0,1]}{\operatorname{argmin}} \sum_{i=1}^{n} (y_i - g(t_i))^2 + \lambda \int_0^1 (g^{(2)}(t))^2 dt, \tag{5.5}$$

vgl. z.B. [Rei67], [Wah90] und [GS94]. Die Lösung von (5.5) ist ein natürlicher Spline mit Knoten $t_1, \ldots, t_n$, also eine Funktion, die zweimal stetig differenzierbar ist und zwischen den Designpunkten jeweils mit einem Polynom zusammenfällt. Zusätzlich gilt $f^{(2)}(t_1) = f^{(2)}(t_n) = 0$. Der Glättungsparameter λ spielt eine ähnliche Rolle wie die Bandbreite h bei den Nadaraya-Watson-Kernschätzern. Für $\lambda = 0$ interpoliert der Glättungsspline die Daten, für $\lambda \to \infty$ konvergiert der Glättungsspline gegen die affin-lineare Kleinste-Quadrate-Schätzung. Für Glättungssplines wurden verschiedene Methoden zur Wahl von λ vorgeschlagen, einen Überblick gibt z.B. Kapitel 4 in [Wah90]. Kubische Glättungssplines sind für festes λ ebenfalls linear in den Daten, und es ergibt sich

$$\hat{y}_\lambda = W_\lambda y. \tag{5.6}$$

Eine explizite Darstellung der Matrix W_λ ist z.B. in Kapitel 2 von [GS94] zu finden. Wie im Fall der Nadaraya-Watson-Kernschätzer existieren viele Kriterien zur Wahl von λ, die die Diagonalelemente $w_\lambda := (w_1^\lambda, \ldots, w_n^\lambda)$ von W_λ zur Wahl von λ benutzen, so z.B. die Vergleichsmethoden in der Simulationsstudie in Kapitel 5.3. Zur schnellen Berechnung der Schätzung $\hat{y}_\lambda$ in den Designpunkten kann der von Reinsch vorgeschlagene Algorithmus verwendet werden, vgl. [Rei67] bzw. Kapitel 2.3.3 in [GS94].

Liegt eine Schätzung $\hat{f}_h$ bzw. $\hat{f}_\lambda$ vor, so wird der Vektor der Residuen als

$$r_h = (r_1^h, \ldots, r_n^h) := (\hat{f}_h(t_1) - y_1, \ldots, \hat{f}_h(t_n) - y_n)$$

bzw.

$$r_\lambda = (r_1^\lambda, \ldots, r_n^\lambda) := (\hat{f}_\lambda(t_1) - y_1, \ldots, \hat{f}_\lambda(t_n) - y_n)$$

definiert. Das *Diskrepanzprinzip für die nichtparametrische Regression* basiert auf einer Beurteilung des Residualvektors und lässt sich formulieren als

$$\text{Wähle } h \text{ maximal so, dass } \|r_h\| \leq \sigma s(n), \tag{5.7}$$

bzw.

$$\text{Wähle } \lambda \text{ maximal so, dass } \|r_\lambda\| \leq \sigma s(n), \tag{5.8}$$

wobei $\|\cdot\|$ eine Norm auf $\mathbb{R}^n$ ist und s eine geeignete Schrankenfunktion. Am weitesten verbreitet zur Messung der Größe von Residuen sind die ℓ_p-Normen, die durch

$$\|(r_1, \ldots, r_n)\|_p = \begin{cases} \left(\sum_{t=1}^n |r_t|^p\right)^{1/p} & (1 \leq p < \infty) \\ \max\{|r_1|, \ldots, |r_n|\} & (p = \infty) \end{cases}$$

definiert werden. Die bisher in der Literatur verwendeten Versionen des Diskrepanzprinzips basieren entweder auf der ℓ_2-Norm oder auf einer besonderen Norm, der Multiresolutionsnorm, auf die weiter unten eingangen wird. Zur Implementierung des Diskrepanzprinzips wird auch die Rauschvarianz σ benötigt, die in der Praxis häufig aus den Daten geschätzt werden muss.

Auf der ℓ_2-Norm basierende Diskrepanzprinzipien werden vor allem für Glättungssplines verwendet. Bereits Reinsch [Rei67] schlägt vor, λ so zu wählen, dass

$$\|r_\lambda\|_2^2 = \sigma^2 S$$

für $n-\sqrt{2n} \le S \le n+\sqrt{2n}$ gilt. Dabei wird benutzt, dass (5.5) eine Lagrange-Formulierung des Problems

$$\int_0^1 (g^{(2)}(t))^2 dt \longrightarrow \min!$$

unter der Nebenbedingung

$$\sum_{i=1}^{n} (y_i - g(t_i))^2 \le \sigma^2 S$$

ist, vgl. [Rei67]. Für die Wahl $S = n$ entspricht dabei die rechte Seite der Nebenbedingung genau dem Erwartungswert von $\|(\varepsilon_1, \ldots, \varepsilon_n)\|_2^2$. Wahba zeigt, dass für den bezüglich des Average Mean Squared Error optimalen Wert S^* für periodische Glättungssplines

$$S^* = n(1 - k(1 + o(1)))$$

gilt mit $k = c(\|f^{(4)}\|_2^2/\sigma^2)^{1/9} n^{-8/9}$ und $c \approx 0.6$. Sie bemerkt weiter, dass sich ähnliche Ergebnisse auch für natürliche Glättungssplines ergeben, vgl. [Wah75], Kapitel 4.7 in [Wah90] sowie [HT87]. Damit führt die Wahl von $\|\cdot\| = \|\cdot\|_2$ und $s(n) = \sqrt{n}$ bzgl. des Average Mean Square Error zu zu großen Wahlen von λ, was durch die Simulationsstudie in [CW79] bestätigt wird.

Ein spezielles Diskrepanzprinzip, das sogenannte Multiresolutionskriterium, entstammt dem „Data Approximation“-Ansatz (vgl. Kapitel 2.3). Davies und Kovac schlagen in Verbindung mit ihrer Taut-String-Methode [DK01] vor, die einfachste Schätzung $\hat{f}$ zu wählen, für die die Residuen „aussehen wie weißes Rauschen“. Die Einfachheit bzw. Komplexität einer Schätzung kann zum Beispiel durch Glattheit gemessen werden. In der Taut-String-Methode wird mit der Anzahl lokaler Extrema von $\hat{f}$ ein diskretes Komplexitätskriterium verwendet. Dabei wird zur Definition von „Aussehen wie Rauschen“ ein objektives und automatisch überprüfbares Kriterium benötigt. Die Auswertung einer klassischen ℓ_p-Norm des Residualvektors würde zwar zur Ablehnung führen, wenn die Residuen groß sind, jedoch würden keine für das menschliche Auge in einem Residualplot deutlich sichtbaren Strukturen entdeckt. Das wichtigste Beispiel hierfür sind längere Intervalle, auf denen die meisten Residuen dasselbe Vorzeichen haben. Solche Muster sprechen für eine systematische Unter- oder Überschätzung des Signals f. Dies könnte erkannt werden, indem man die Residuen wiederum nichtparametrisch glättet und die resultierende Schätzung mit der Nullfunktion vergleicht, was aber die Wahl eines weiteren Glättungsparameters erfordert (vgl. z.B. [Har97], vor allem Kap. 7). Das in [DK01] vorgeschlagene Kriterium beurteilt Schätzungen bei verschiedenen „Auflösungen“ simultan: Die Residuen „sehen aus wie weißes Rauschen”, wenn für den Residualvektor $r = (r_1, \ldots, r_n)$

$$\max_{I \in \mathcal{I}} \frac{1}{\sqrt{|I|}} \left| \sum_{t \in I} r_t \right| \le \sigma s(n) \tag{5.9}$$

erfüllt ist, wobei $\mathcal{I} = \{I | I = \{i, i+1, \ldots, l\}, 1 \leq i \leq l \leq n\}$ das System aller diskreten Intervalle in $\{1, \ldots, n\}$ und $|I|$ die Anzahl der Designpunkte in I bezeichnet. Kleinere Systeme von Intervallen, z.B. nur die dyadischen Intervalle, können auch benutzt werden. Die Schrankenfunktion s sollte so gewählt werden, dass (5.9) mit hoher Wahrscheinlichkeit erfüllt ist, wenn die Daten tatsächlich aus dem Modell (5.1) mit $f = \hat{f}$ stammen.

Die Bedingung (5.9) wird üblicherweise als *Multiresolutionskriterium* bezeichnet. Da das Kriterium Summen von Residuen über Teilintervalle berücksichtigt, führen z.B. auch längere Folgen von Residuen moderater Größe mit dem selben Vorzeichen dazu, dass (5.9) groß wird. Das Kriterium (und die Wahl von $s(n)$) wurden ursprünglich durch das von Donoho und Johnstone [DJ94] vorgeschlagene Wavelet-Thresholding-Verfahren *VisuShrink* motiviert. Die Parallele wird deutlich, wenn $n = 2^j$ äquidistante Designpunkte, die Familie der dyadischen Intervalle und die Haar-Basis betrachtet werden. In der Hard-Thresholding-Wavelet-Regression wird f geschätzt, indem eine Wavelet-Transformation auf y angewandt und alle Wavelet-Koeffizienten auf Null gesetzt werden, die betragsmäßig kleiner als eine vorher definierte Schranke sind. Anschließend wird die Rücktransformation angewendet. Da für viele interessante Klassen von Regressionsfunktionen die relevante Information über f in wenigen Koeffizienten enthalten ist, werden nur diese verwendet und der Rest als durch das Rauschen verursacht angesehen. Da die Wavelet-Transformation linear ist, ergibt eine weitere Anwendung der Hard-Threshold-Methode mit derselben Schranke einen Nullvektor. Die Wahl der Schranke wird durch die Tatsache motiviert, dass wegen ihrer Orthogonalität die Wavelet-Transformation von Gaußschem weißen Rauschen wieder Gaußsches weißes Rauschen ergibt. Die Verwendung von (5.9) mit dem System der dyadischen Intervalle führt zu ähnlichen Ergebnissen, mit dem Unterschied, dass die Koeffizienten für die Skalierungsfunktion und nicht die Wavelet-Koeffizienten betrachtet werden, vgl. [DK01]. In Kapitel 5.2 wird eine etwas andere, auf Nadaraya-Watson-Kernschätzern basierende, Motivation gegeben. Dort wird gezeigt, dass sich (5.9) formulieren lässt als

$$\|r\|_{\mathsf{MR}} \leq \sigma s(n),$$

für die in Definition 5.1 definierte Multiresolutionsnorm $\| \cdot \|_{\mathsf{MR}}$. Damit entspricht das Multiresolutionskriterium also einem Diskrepanzprinzip vom Typ (5.7) bzw. (5.8). Die bisher benutzten Schrankenfunktionen sind vom Typ

$$s(n) = \sqrt{\tau \log(n)}$$

wobei τ in der Regel zwischen 2 und 3 gewählt wird, vgl. [DK01] bzw. [DM08]. Eine andere Möglichkeit besteht darin, τ in Abhängigkeit von n so zu wählen, dass (5.9) mit einer gegebenen Wahrscheinlichkeit α erfüllt ist, vgl. [DKM09].

Das Multiresolutionskriterium eignet sich besonders zur Wahl lokalisierter Glättungsparameter, da die Information, für welche Intervalle (5.9) verletzt ist, dazu benutzt werden kann, Glättungsparameter lokal anzupassen. Dieser Ansatz wurde bisher erfolgreich für Nadaraya-Watson-Kernschätzer und lokal-polynomiale Schätzer (vgl. [Mei04], [Mei06]) sowie für kubische Glättungssplines (vgl. [Mei04], [DM08] und [Mei06]) umgesetzt. Die Motivation im Sinne des „Data Approximation“-Ansatzes besteht dabei darin, unter allen

„adäquaten Approximationen“ (vgl. Kapitel 2.3) eine Funktion maximaler Glattheit zu finden. Wünschenswert wäre die Lösung des Problems

$$\int_0^1 (g^{(2)}(t))^2 dt \longrightarrow \min!$$

unter der Nebenbedingung

$$\|(g(t_1) - y_1, \dots, g(t_n) - y_n)\|_{\mathsf{MR}} \leq \sigma\sqrt{\tau \log(n)}.$$

Dies ist ein quadratisches Optimierungsproblem, das jedoch selbst bei Verwendung eines kleineren Intervallsystems für das Multiresolutionskriterium noch sehr groß und in vielen Fällen auch numerisch instabil ist, vgl. [DM08]. Daher werden iterativ kubische Glättungssplines mit lokalisiertem Glättungsparameter berechnet, bis das Multiresolutionskriterium erfüllt ist. Wird in diesem Ansatz jedoch ein globaler Glättungsparameter benutzt, entspricht dies genau einem Spline-Schätzer vom Typ (5.5), dessen Glättungsparameter mit Hilfe des Diskrepanzprinzips (5.8) für $\|\cdot\| = \|\cdot\|_{\mathsf{MR}}$ und $s(n) = \sqrt{\tau \log(n)}$ mit $\tau > 2$ gewählt wird.

In [DM08] wird außerdem gezeigt (Theorem 7.2), dass in diesem Fall für hinreichend glattes f und eine Folge δ_n mit $\delta_n^{-1} n^{-5/16} (\log(n))^{9/16} = o(1)$ die Konvergenzrate

$$\sup_{\delta_n < x < 1-\delta_n} |\hat{f}_{\lambda_{s,n}}(x) - f(x)| = O_P((\log n)^{7/32} n^{-11/32})$$

erreicht wird, wobei $\lambda_{s,n}$ den wie beschrieben gewählten Glättungsparameter bezeichnet.

Für diesen Ansatz existieren multivariate Verallgemeinerungen, insbesondere zur Bildverarbeitung mittels Thin-Plate-Splines (vgl. [Mei04] und [DM08]) bzw. mittels auf der Wärmeleitungsgleichung basierender Methoden (vgl. [HMS+10]). Außerdem wird das Kriterium (5.9) für stückweise konstante Kleinste-Quadrate-Regressionsschätzer verwendet, vgl. [BKL+09]. Auch zur sequentiellen Erkennung von Strukturbrüchen werden ähnliche Kriterien eingesetzt, vgl. z.B. [SV95]. In der Literatur werden außerdem einige vorwiegend asymptotische Eigenschaften von (5.9) behandelt, vgl. [DK01], [DM08], [BKL+09] und [Kab08]. Bernholt und Hofmeister entwickeln einen effizienten Algorithmus zur Entscheidung, ob (5.9) erfüllt ist (vgl. [BH06] und [BEH07]).

Es existieren außerdem weitere Verallgemeinerungen von (5.9), unter anderem robuste Varianten, die benutzt werden können, wenn das Rauschen eine Verteilung mit schweren Tails besitzt. Die Residuen werden in diesem Fall transformiert, z.B. in dem sie durch ihre Vorzeichen ± 1 ersetzt werden, die sich dann wie unabhängig identisch verteilte Rademacher-Zufallsvariablen verhalten sollten [Kov02]. Ein direkterer Ansatz, ebenfalls in [DK01] vorgestellt, besteht darin, (5.9) durch ein auf Runlängen basierendes Kriterium zu ersetzen, was jedoch zu einer völlig anderen Prozedur führt. Dümbgen und Kovac [DK09] stellen außerdem eine auf Score-Funktionen basierende Version für M-Schätzer mit Strafterm vor, wobei sich im Falle eines Kleinste-Quadrate-Schätzers wieder das Kriterium (5.9) ergibt.

5.2. Eine geometrische Interpretation des Multiresolutionskriteriums

Im Folgenden wird eine rein geometrische Interpretation von (5.9) für festes n vorgestellt. Es ist bekannt, dass die Menge der Residuen, die (5.9) erfüllen, ein Polyeder in $\mathbb{R}^n$ ist, da (5.9) einer endlichen Menge linearer Ungleichungen entspricht [DKM09]. Im weiteren Verlauf dieses Kapitels soll gezeigt werden, dass diese Menge eine Kugel in einer speziellen Norm – der *Multiresolutionsnorm* (MR-Norm) – ist. Weiter werden einige interessante Eigenschaften dieser Norm hergeleitet, die sie von den ℓ_p-Normen unterscheiden und als eine Art formalisierter Residualplot nützlich machen. In diesem Abschnitt sind nur die Residuen von Interesse, nicht die Funktion f. Daher wird nur der Vektor der Auswertung einer Schätzung $\hat{f}$ in den Designpunkten $(\hat{f}(t_1), \ldots, \hat{f}(t_n))$ betrachtet, der ebenfalls mit $\hat{f}$ bezeichnet werden soll. Die Gestalt des Schätzers $\hat{f}$ spielt für die folgenden Überlegungen keine Rolle. Weiter bezeichne $y := (y(t_1), \ldots, y(t_n))$ den Vektor der Beobachtungen und $r := (\hat{f}(t_1) - y(t_1), \ldots, \hat{f}(t_n) - y(t_n))$ den Vektor der Residuen.

Residuen $(r_1, \ldots, r_n)$ können mit Hilfe von Nadaraya-Watson-Kernschätzern mit der Nullfunktion verglichen werden. Betrachtet wird eine unimodale, symmetrische nichtnegative Kernfunktion $K : \mathbb{R} \longrightarrow \mathbb{R}^+$. Für eine Bandbreite $h > 0$ wird $K_h(\cdot) := h^{-1}K(h^{-1}\cdot)$ definiert. Statt eine einzige Bandbreite zu wählen, können auch alle Bandbreiten simultan betrachtet werden:

$$r_{t,h} := \begin{cases} \frac{\sum_{i=1}^n K_h(t_i-t) r_i}{\sqrt{\sum_{i=1}^n K_h^2(t_i-t)}}, & \text{falls } \sqrt{\sum_{i=1}^n K_h^2(t_i-t)} \neq 0 \\ 0, & \text{falls } \sqrt{\sum_{i=1}^n K_h^2(t_i-t)} = 0 \end{cases} \tag{5.10}$$

für alle $t \in [0,1], h > 0$. Dies entspricht dem Nadaraya-Watson-Kernschätzer (5.2) mit Bandbreite h im Punkt t, lediglich der Nenner wird zur Standardisierung anders gewählt. Werden die r_i als Realisierungen der ε_i angenommen, dann gilt $r_{t,h} \sim N(0, \sigma^2)$ für alle (t,h), für die $\sqrt{\sum_{i=1}^n K_h^2(t_i-t)}$ nicht Null ist. Nach der Cauchy-Schwarz-Ungleichung gilt für alle $(r_1, \ldots, r_n) \in \mathbb{R}^n$

$$|r_{t,h}| \leq \|(r_1, \ldots, r_n)\|_2, \tag{5.11}$$

was

$$\sup_{t,h} |r_{t,h}| \leq \|(r_1, \ldots, r_n)\|_2 < \infty$$

impliziert. Nun kann eine Norm auf $\mathbb{R}^n$ definiert werden:

Satz 5.1. *Für jede Kernfunktion K und feste Designpunkte $t_1, \ldots, t_n$ wie oben definiert*

$$\|(r_1, \ldots, r_n)\|_K := \sup_{t,h} |r_{t,h}| \tag{5.12}$$

eine Norm.

Beweis. Nichtnegativität, Homogenität und die Dreiecksungleichung sind offensichtlich. Zum Nachweis der positiven Definitheit sei mindestens eines der r_i nicht 0. Dann existiert immer eine Bandbreite h, für die die Beiträge aller anderen r_j zu $r_{t_i,h}$ beliebig klein sind und daher $|r_{t_i,h}| > 0$ ist. □

Da (5.12) als Supremum über eine überabzählbare Menge definiert ist, wird man in der Praxis ein endliches Gitter benutzen. Nichtnegativität, Homogenität und die Dreiecksungleichung bleiben dadurch erhalten. Falls das Gitter zu grob ist, kann aber die positive Definitheit verloren gehen.

Die Idee, Kernschätzer für eine Menge möglicher Bandbreiten zu benutzen, ist eng mit dem zuerst in der Bildverarbeitung vorgestellten Scale-Space-Ansatz verwandt. Die in der Statistik bekanntesten Methoden dieses Typs sind die sogenannten SiZer-Maps, die von Chaudhuri und Marron vorgeschlagen wurden, vgl. [CM99, CM00]. Die Definitionen sind allerdings etwas anders, und es werden nur Gaußkerne zugelassen. Dümbgen und Spokoiny [DS01] und Rufibach und Walther [RW10] benutzen ebenfalls ähnliche Kriterien, die auf Kernschätzern für verschiedene Skalen beruhen. Allerdings werden die zugehörigen Teststatistiken für jede Skala mit anderen kritischen Werten verglichen, die vom Niveau des Tests abhängen. Dies macht eine Interpretation im hier vorgestellten Rahmen unmöglich. Dette und Hetzler [DH07] schlagen eine auf Nadaraya-Watson-Kernschätzern für verschiedene Bandbreiten beuhende Teststatistik für Goodness-of-Fit-Tests vor; in diesem Fall liegen jedoch die Bandbreiten in einem von n abhängigen Intervall der Form $[an^{-1/5}, bn^{-1/5}]$ für Konstanten $a < b$.

Wird nun speziell der Rechteckskern $K^{\mathcal{U}}(x)$ (3.13) verwendet, hängen die Gewichte und der Nenner in (5.12) nur von der Anzahl der Designpunkte in einem Fenster der Gesamtlänge h um t ab:

$$\frac{\sum_{i=1}^{n} K_h^{\mathcal{U}}(t_i - t) r_i}{\sqrt{\sum_{i=1}^{n} (K_h^{\mathcal{U}}(t_i - t))^2}} = \frac{\sum_{\{i:|t-t_i| \leq h/2\}} r_i}{\sqrt{|\{i : |t - t_i| \leq h/2\}|}}. \tag{5.13}$$

In diesem Fall gibt es nur endlich viele $r_{t,h}$, eines für jedes *diskrete Intervall* $\{t_i, t_{i+1}, \dots, t_j\}$ mit $1 \leq i \leq j \leq n$. Das Supremum in (5.12) wird daher nur über eine endliche Menge gebildet und hängt von den Designpunkten nur über ihre Anordnung ab. Daher wird im Folgenden auf die Designpunkte nicht mehr Bezug genommen, und die Residuen werden nur noch als Vektor im $\mathbb{R}^n$ betrachtet. Damit kann auf $\mathbb{R}^n$ wie folgt eine Norm definiert werden:

Definition 5.1. Sei $n \in \mathbb{N}$ und $\mathcal{I} = \{I | I = \{i, i+1, \dots, l\}, 1 \leq i \leq l \leq n\}$. Die *Multiresolutionsnorm* (kurz *MR-Norm*) auf $\mathbb{R}^n$ ist definiert durch:

$$\|(x_1, \dots, x_n)\|_{\mathsf{MR}} := \|(x_1, \dots, x_n)\|_{K^{\mathcal{U}}} := \max_{I \in \mathcal{I}} \frac{1}{\sqrt{|I|}} \left| \sum_{t \in I} x_t \right|.$$

Da dies lediglich (5.12) für den Rechteckskern und beliebige Designpunkte ist, definiert $\| \cdot \|_{\mathsf{MR}}$ offensichtlich eine Norm.

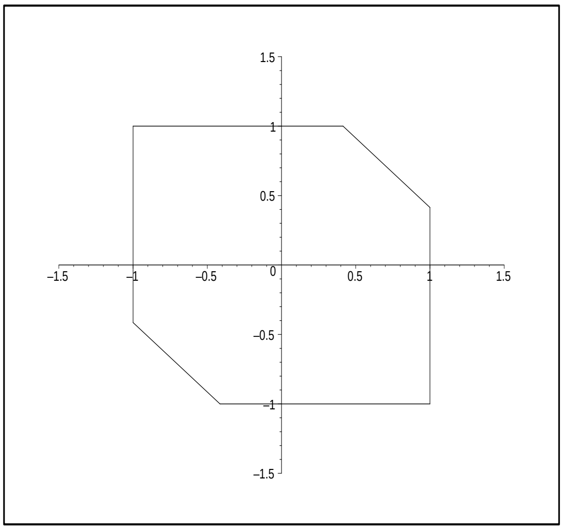

Abbildung 5.1.: Die Einheitskugel der Multiresolutionsnorm für $n = 2$.

Abbildung 5.1 zeigt die Einheitskugel von $\|\cdot\|_{\mathsf{MR}}$ für $n = 2$. Die Komponenten eines Vektors innerhalb der Einheitskugel dürfen größer sein, falls sie sich im Vorzeichen unterscheiden.

Eine Norm heißt strikt konvex, falls der Rand der Einheitskugel keine Streckensegmente enthält, d.h. für alle $x, y \in \mathbb{R}^n$ mit $\|x\| = \|y\| = 1$ impliziert $\|\frac{1}{2}(x+y)\| = 1$, dass $x = y$. Setzt man $x = (1, 0, \ldots, 0)$ und $y = (1, -1, 0, \ldots, 0)$, so ist $\|x\|_{\mathsf{MR}} = \|y\|_{\mathsf{MR}} = 1$, aber

$$\left\| \frac{1}{2}(x+y) \right\|_{\mathsf{MR}} = \left\| (1, -\frac{1}{2}, 0, \ldots, 0) \right\|_{\mathsf{MR}} = \max\left\{1, \frac{1}{2}, \frac{1}{2\sqrt{2}}\right\} = 1.$$

Damit ist klar, dass die Multiresolutionsnorm für $n \geq 2$ nicht strikt konvex ist. Da durch Skalarprodukte gegebene Normen stets strikt konvex sind, ist $(\mathbb{R}^n, \|\cdot\|_{\mathsf{MR}})$ kein Hilbertraum.

Weiter ist klar, dass die Residuen (5.9) genau dann erfüllen, wenn $\|r\|_{\mathsf{MR}} \leq \sigma s(n)$ gilt, d.h. wenn sie in einer Kugel um den Ursprung liegen bzw. – äquivalent dazu – wenn $\hat{f}$ in einer Kugel um die Daten y liegt. Im Gegensatz zu anderen Distanzmaßen zwischen Datenpunkten und geschätzten Werten kann die MR-Norm nicht als Schätzung eines Abstands zwischen f und $\hat{f}$ interpretiert werden, sie ist also kein empirisches Analogon zu einer (sinnvollen) Populationsgröße.

Die Nützlichkeit der Multiresolutionsnorm im Vergleich zu den ℓ_p-Normen ergibt sich nun aus dem Fehlen bestimmter Invarianzeigenschaften. Sei $\mathcal{S}_n$ die symmetrische Gruppe, d.h. die Gruppe der Permutationen von $\{1, \ldots, n\}$ mit

$$\pi x := (x_{\pi(1)}, \ldots, x_{\pi(n)}) \text{ für } \pi \in \mathcal{S}_n, x \in \mathbb{R}^n.$$

Desweiteren wird die Vorzeichengruppe $\mathcal{T}_n := \{-1, +1\}^n$ mit komponentenweiser Multiplikation betrachtet. Durch

$$sx := (s_1 x_1, \dots, s_n x_n) \text{ für } s \in \mathcal{T}_n, x \in \mathbb{R}^n$$

wird die Operation von $\mathcal{T}_n$ auf $\mathbb{R}^n$ definiert. Nach Definition von $\|\cdot\|_p$ ist direkt klar, dass $\|\pi x\|_p = \|x\|_p$ und $\|sx\|_p = \|x\|_p$ für alle $x \in \mathbb{R}^n, \pi \in \mathcal{S}_n, s \in \mathcal{T}_n$ und $p \in [1, \infty]$ gilt, da die p-Normen von den Komponenten nur über die Absolutbeträge abhängen. Die Multiresolutionsnorm eines Vektors ist dagegen im Allgemeinen nicht invariant unter diesen Transformationen. Betrachte z.B.:

$$\begin{aligned} \|(1,-1,1)\|_{\mathsf{MR}} &= 1 \\ \|(1,1,-1)\|_{\mathsf{MR}} &= \sqrt{2}. \end{aligned}$$

Dies ist ein Gegenbeispiel für beide Arten von Transformationen, da der zweite Vektor aus dem ersten sowohl durch Vertauschen der zweiten und dritten Komponente, als auch durch Vorzeichenwechsel in diesen Komponenten hervorgeht. Die Untergruppen von $\mathcal{S}_n$ und $\mathcal{T}_n$, unter denen $\|x\|_{\mathsf{MR}}$ invariant ist für alle $x \in \mathbb{R}^n$, bestehen jeweils nur aus zwei Elementen, wie nun gezeigt wird. Bezeichne id die Identität in $\mathcal{S}_n$ oder $\mathcal{T}_n$. Sei weiter $\rho \in \mathcal{S}_n$ die Permutation, die die Reihenfolge der Komponenten umkehrt, gegeben durch

$$\rho(x_1, \dots, x_n) = (x_n, \dots, x_1)$$

für alle $x \in \mathbb{R}^n$. Offensichtlich wird die Multiresolutionsnorm nicht geändert durch Umkehrung der Reihenfolge der Komponenten eines Vektors, da dieselben Summen betrachtet werden. Weiter bezeichne $\nu := (-1, \dots, -1)$ das Element von $\mathcal{T}_n$, das alle Vorzeichen gleichzeitig umkehrt. Die Umkehr aller Vorzeichen eines Vektors ist äquivalent zur Skalarmultiplikation des Vektors mit -1 und ändert daher die Multiresolutionsnorm eines Vektors nicht. Es gilt $\rho \circ \rho = \mathsf{id}$ und $\nu \circ \nu = \mathsf{id}$, so dass $\{\mathsf{id}, \rho\} \subset \mathcal{S}_n$ und $\{\mathsf{id}, \nu\} \subset \mathcal{T}_n$ Untergruppen sind. Invarianz unter diesen Untergruppen ist eine erwünschte Eigenschaft, da die Entscheidung, ob Residuen als Rauschen akzeptiert werden, sich nicht ändern soll, wenn sich die Anordnung umkehrt oder alle Vorzeichen gleichzeitig geändert werden. Es gibt keine weiteren Elemente in diesen Gruppen, die die Multiresolutionsnorm für alle $x \in \mathbb{R}^n$ unverändert lassen:

Satz 5.2. *1. Für $\pi \in \mathcal{S}_n$ gilt $\|\pi x\|_{\mathsf{MR}} = \|x\|_{\mathsf{MR}}$ für alle $x \in \mathbb{R}^n$ genau dann, wenn $\pi = \mathsf{id}$ oder $\pi = \rho$.*

2. Für $s \in \mathcal{T}_n$ gilt $\|sx\|_{\mathsf{MR}} = \|x\|_{\mathsf{MR}}$ für alle $x \in \mathbb{R}^n$ genau dann, wenn $s = \mathsf{id}$ oder $s = \nu$.

Beweis. 1. Die Invarianz der MR-Norm unter id und ρ ist nach der vorausgehenden Argumentation klar. Sei nun $\pi \in \mathcal{S}_n$ mit $\pi \neq \mathsf{id}$ und $\pi \neq \rho$. Da id und ρ die einzigen Permutationen sind, die die Anordnung der Komponenten nicht verändern, existieren $i, j \in 1, \dots, n$

mit $|j-i| \geq 2$ und $|\pi(j)-\pi(i)| = 1$. Für $x \in \mathbb{R}^n$, so dass $x_i = x_j = 1$ und alle anderen Komponenten 0 sind, ergibt sich

$$\|x\|_{\mathsf{MR}} = \max\{1, 2/\sqrt{|j-i|+1}\} = \begin{cases} \frac{2}{\sqrt{3}} & (|j-i| = 2) \\ 1 & (|j-i| > 2) \end{cases} .$$

Andererseits sind $x_{\pi(i)}$ und $x_{\pi(j)}$ benachbarte Komponenten in πx. Daher ist

$$\|\pi x\|_{\mathsf{MR}} = \max\{1, 2/\sqrt{2}\} = \sqrt{2},$$

und für jedes $\pi \notin \{\mathsf{id}, \rho\}$ existiert mindestens ein $x \in \mathbb{R}^n$, für das $\|\pi x\|_{\mathsf{MR}} \neq \|x\|_{\mathsf{MR}}$ gilt.
2. Die Invarianz der MR-Norm unter id und ν ist nach der vorausgehenden Argumentation klar. Sei $s \in \mathcal{T}_n$ mit $s \neq \mathsf{id}$ und $s \neq \nu$. In diesem Fall existiert ein $i \in \{1, \ldots, n-1\}$, so dass $s_i s_{i+1} = -1$. Sei nun x der Vektor mit $x_i = x_{i+1} = 1$ und $x_k = 0$ für $k \notin \{i, i+1\}$. Bei der Berechnung der Multiresolutionsnorm ergibt sich $\frac{1}{\sqrt{2}}(x_i + x_{i+1}) = \sqrt{2} > 1 = |x_i| = |x_{i+1}|$. Daraus folgt

$$\|x\|_{\mathsf{MR}} = \sqrt{2}.$$

In sx sind die beiden Komponenten x_i und x_{i+1} betragsmäßig gleich 1, haben aber entgegengesetzte Vorzeichen. Damit ist die Summe auf dem Intervall $\{i, i+1\}$ Null. Damit gilt

$$\|sx\|_{\mathsf{MR}} = 1.$$

Wie oben kann mindestens ein Vektor x mit $\|sx\|_{\mathsf{MR}} \neq \|x\|_{\mathsf{MR}}$ für jedes $s \neq \mathsf{id}$, $s \neq \nu$ konstruiert werden. □

Also lassen nur triviale Permutationen oder Vorzeichenwechsel die Multiresolutionsnorm invariant für alle $x \in \mathbb{R}^n$. Für eine feste Permutation oder einen festen Vorzeichenwechsel existieren natürlich immer Vektoren, für die sich die Norm nicht ändert; das einfachste Beispiel hierfür ist der Nullvektor. Es ist gerade das Fehlen der sonst üblichen Invarianzeigenschaften, das die Multiresolutionsnorm nützlich zur Erkennung von Mustern wie längeren Abfolgen großer Residuen oder Residuen desselben Vorzeichens macht.

Nun werden noch einige weitere Eigenschaften der Multiresolutionsnorm in Verbindung mit Vorzeichenmustern hergeleitet. Sei dazu $|x| = (|x_1|, \ldots, |x_n|)$.

Lemma 5.1. *Für $x \in \mathbb{R}^n$ gilt $\|x\|_{\mathsf{MR}} \leq \||x|\|_{\mathsf{MR}}$, und strikte Ungleichheit ist möglich.*

Beweis. Es gilt $\max_{I \in \mathcal{I}} \frac{1}{\sqrt{|I|}} \left|\sum_{t \in I} x_t\right| \leq \max_{I \in \mathcal{I}} \frac{1}{\sqrt{|I|}} \sum_{t \in I} |x_t|$. Ein Beispiel für strikte Ungleichheit ist $\|(1,-1)\|_{\mathsf{MR}} = 1 < \sqrt{2} = \|(1,1)\|_{\mathsf{MR}}$. □

Die Abhängigkeit von $\|x\|_{\mathsf{MR}}$ vom Vorzeichenmuster von x wird besonders deutlich, wenn Vektoren betrachtet werden, die nur aus Komponenten mit dem selben Absolutbetrag bestehen. Ein weiterer Grund, diesen Fall zu betrachten, ist das robuste Multiresolutionsvorzeichenkriterium, das in [Kov02] vorgeschlagen wird: Die Residuen werden durch ihre Vorzeichen ersetzt, d.h. (5.9) wird auf $(\mathrm{sign}(r_1), \ldots, \mathrm{sign}(r_n))$ angewendet statt auf den ursprünglichen Residuenvektor. Sind alle Residuen ungleich Null, entspricht dies der Ersetzung des Residuenvektors durch diejenige Ecke der Einheitskugel von $\|\cdot\|_\infty$, die im selben Orthanten liegt.

Satz 5.3. *Für die Menge aller $x = (x_1, \dots, x_n)$ mit $|x_1| = \dots = |x_n| =: m > 0$ gilt:*

1. *$\|x\|_{\mathsf{MR}}$ ist maximal genau dann, wenn alle Komponenten dasselbe Vorzeichen haben. Dann ist $\|x\|_{\mathsf{MR}} = \sqrt{n}m$.*
2. *$\|x\|_{\mathsf{MR}}$ ist minimal genau dann, wenn sich die Vorzeichen der Komponenten abwechseln. Dann ist $\|x\|_{\mathsf{MR}} = m$.*
3. *$\|x\|_{\mathsf{MR}} \geq \sqrt{\ell}m$, wobei ℓ die Länge des längsten Runs von Komponenten desselben Vorzeichens ist.*

Beweis. 1. Die Maximalität bei gleichen Vorzeichen ist nach Lemma 5.1 klar. Es ist ebenfalls klar, dass in diesem Fall $\frac{1}{\sqrt{|I|}}|\sum_{t\in I} x_t| = \sqrt{|I|}m$, so dass $\|x\|_{\mathsf{MR}} = \sqrt{n}m$. Existieren zwei Komponenten mit entgegengesetzten Vorzeichen, so heben sie sich in der Summation über das Intervall $I := \{1, \dots, n\}$ gegenseitig auf. Daher ist

$$\frac{1}{\sqrt{|I|}}\left|\sum_{t\in I} x_t\right| \leq \frac{n-2}{\sqrt{n}}m < \sqrt{n}m.$$

Ebenso ist klar, dass für jedes Teilintervall I mit $|I| < n$ $\frac{1}{\sqrt{|I|}}\left|\sum_{t\in I} x_t\right| \leq \sqrt{n-1}m$ gilt.
2. Zunächst sei x so, dass sich die Vorzeichen der Komponenten abwechseln. Für ein Intervall mit einer geraden Anzahl von Punkten heben sich die Komponenten auf. Hat das Intervall eine ungerade Länge, so bleibt bei der Summation gerade ein Term übrig. In diesem Fall ist $\frac{1}{\sqrt{|I|}}\left|\sum_{t\in I} x_t\right| = m/\sqrt{|I|}$. Dieser Wert wird maximal für Intervalle mit genau einem Punkt, und es ergibt sich $\|x\|_{\mathsf{MR}} = m$. Nimmt man hingegen an, dass mindestens zwei benachbarte Komponenten x_i und x_{i+1} mit dem selben Vorzeichen existieren, ergibt sich $\|x\|_{\mathsf{MR}} \geq \frac{1}{\sqrt{2}}|x_i + x_{i+1}| = \frac{2}{\sqrt{2}}m = \sqrt{2}m > m$.
3. Sei I ein Intervall mit Komponenten mit dem selben Vorzeichen. Dann ist
$\frac{1}{\sqrt{|I|}}|\sum_{t\in I} x_t| = \sqrt{|I|}m$. □

Nach Teil 1. und 2. von Satz 5.3 ist klar, dass die Ecken der Einheitskugeln der MR-Norm und $\|\cdot\|_\infty$ in denjenigen Orthanten des $\mathbb{R}^n$ übereinstimmen, in denen die Vorzeichen sich abwechseln, und am weitesten entfernt sind in den Orthanten, in denen alle Vorzeichen gleich sind. Durch vollständige Aufzählung aller Möglichkeiten erhält man, dass in Teil 3. von Satz 5.3 für $n \leq 6$ stets Gleichheit gilt, während die Ungleichung für $n \geq 7$ strikt sein kann: So ist z.B.

$$\|(1,1,1,-1,-1,-1,1)\|_{\mathsf{MR}} = \sqrt{3},$$

aber

$$\|(1,1,1,-1,1,1,1)\|_{\mathsf{MR}} = \frac{5}{\sqrt{7}} > \sqrt{3}.$$

Also können Residuenvektoren mit der selben maximalen Runlänge verschiedene MR-Normen haben. Dies bedeutet weiterhin, dass das Multiresolutionsvorzeichenkriterium in [Kov02] nicht äquivalent zu dem Runkriterium in [DK01] ist, das nur auf der Länge des längsten Runs basiert.

Eine weitere Eigenschaft der p-Normen ist die *Zerlegungseigenschaft*

$$\|(x_1, \dots, x_k, x_{k+1}, \dots, x_n)\|_p^p = \|(x_1, \dots, x_k, 0, \dots, 0)\|_p^p + \|(0, \dots, 0, x_{k+1}, \dots, x_n)\|_p^p$$

für $1 \leq p < \infty$ und

$$\|(x_1, \dots, x_k, x_{k+1}, \dots, x_n)\|_\infty = \max\{\|(x_1, \dots, x_k, 0, \dots, 0)\|_\infty, \|(0, \dots, 0, x_{k+1}, \dots, x_n)\|_\infty\},$$

d.h. wird ein Vektor in zwei Teilvektoren zerlegt, lässt sich die Norm des Gesamtvektors aus den Normen der Teilvektoren rekonstruieren. Durch ein einfaches Induktionsargument sieht man, dass dies auch für Partitionierungen in mehr als zwei Teilvektoren gilt. Diese Eigenschaft wird oft implizit benutzt und macht das Rechnen mit p-Normen sehr einfach. Die Multiresolutionsnorm besitzt diese Eigenschaft nicht, wie man einfach sieht: $\|(1,1)\|_{\mathsf{MR}} = \sqrt{2}$ und $\|(1,-1)\|_{\mathsf{MR}} = 1$, aber für die aus einer Komponente bestehenden Teilvektoren gilt: $\|(1,0)\|_{\mathsf{MR}} = \|(0,1)\|_{\mathsf{MR}} = \|(0,-1)\|_{\mathsf{MR}} = 1$. Daher kann keine Funktion existieren, die die Normen der Teilvektoren auf die Norm des Gesamtvektors abbildet.

Nachdem einige wichtige Unterschiede zwischen der MR-Norm und den p-Normen herausgestellt wurden, werden nun noch exakte Abschätzungen hergeleitet:

Satz 5.4. *Für $x \in \mathbb{R}^n$ gilt:*

$$\frac{1}{n^{1/p}}\|x\|_p \leq \|x\|_{\mathsf{MR}} \leq \|x\|_p \qquad (1 \leq p \leq 2), \tag{5.14}$$

$$\frac{1}{n^{1/p}}\|x\|_p \leq \|x\|_{\mathsf{MR}} \leq n^{1/2-1/p}\|x\|_p \qquad (2 < p < \infty), \tag{5.15}$$

$$\|x\|_\infty \leq \|x\|_{\mathsf{MR}} \leq \sqrt{n}\|x\|_\infty. \tag{5.16}$$

Alle Schranken sind scharf.

Beweis. 1. Untere Schranken: Der Fall $p = \infty$ ist trivial, da die Multiresolutionsnorm das Maximum über eine größere Menge ist.
Sei $p \in [1, \infty)$. Angenommen, es existiert $x \in \mathbb{R}^n$ so, dass $\|x\|_p =: m$ und

$$\|x\|_{\mathsf{MR}} < \frac{m}{n^{1/p}}$$

gilt. Da $\|x\|_\infty \leq \|x\|_{\mathsf{MR}}$, gilt $|x_t| < \frac{m}{n^{1/p}}$ für alle $t \in \{1, \dots, n\}$. Dies impliziert

$$\|x\|_p = \left(\sum_{t=1}^n |x_t|^p\right)^{1/p} < \left(n\frac{m^p}{n}\right)^{1/p} = m,$$

im Widerspruch zu $\|x\|_p = m$.
Um zu sehen, dass die Abschätzungen nicht verbessert werden können, betrachtet man Vektoren der Form $x = (a, -a, a, -a, \dots)$ für $a \in \mathbb{R}$. Dann ist $\|x\|_\infty = \|x\|_{\mathsf{MR}} = a$ und $\|x\|_p = n^{1/p}a$.
2. Obere Schranken: Es wird wieder zuerst der Fall $p = \infty$ betrachtet. Sei $\|x\|_\infty =: m$ fest. Es wird ein Vektor mit der größtmöglichen Multiresolutionsnorm konstruiert. Wegen Lemma 5.1 reicht es, Vektoren mit nichtnegativen Komponenten zu betrachten. Die Summe über ein Intervall I wird dann größer, wenn die Komponenten größer werden. Die Multiresolutionsnorm wird dann maximal, wenn jede Komponente gleich dem maximal möglichen Wert m ist. Dann ist

$$\|(m, \dots, m)\|_{\mathsf{MR}} = \max_{I \in \mathcal{I}} \frac{|I|}{\sqrt{|I|}} m = \max_{I \in \mathcal{I}} \sqrt{|I|} m = \sqrt{n} m = \sqrt{n} \|(m, \dots, m)\|_\infty.$$

Für $p \in [1, \infty)$ wird die Hölder-Ungleichung benutzt. Für $I \in \mathcal{I}$ ergibt sich:

$$\begin{aligned} \frac{1}{\sqrt{|I|}} \left| \sum_{t \in I} x_t \right| &\leq \frac{1}{\sqrt{|I|}} \sum_{t \in I} |x_t \cdot 1| && (5.17)\\ &\leq \frac{1}{\sqrt{|I|}} \left(\sum_{t \in I} |x_t|^p \right)^{1/p} \left(\sum_{t \in I} 1 \right)^{1-1/p} && (5.18)\\ &\leq \|x\|_p |I|^{1/2-1/p}. && (5.19) \end{aligned}$$

Für $p \in [1, 2]$ erhält man $|I|^{1/2-1/p} \leq 1$, und (5.19) ergibt die obere Schranke in (5.14). Um zu sehen, dass diese Schranke scharf ist, betrachtet man den Einheitsvektor $x = (1, 0, \dots, 0)$.
Für $p > 2$ erhält man $|I|^{1/2-1/p} \leq n^{1/2-1/p}$, und (5.19) ergibt die obere Schranke in (5.15). Gleichheit gilt hier wieder für Vektoren der Form $x = (m, \dots, m)$ mit $\|x\|_{\mathsf{MR}} = \sqrt{n} m$ und $\|x\|_p = n^{1/p} m$. □

Abbildung 5.2 zeigt die Einheitskugeln in der 1-, 2-, ∞- und MR-Norm für $n = 2$. Die Kugel in der 2-Norm berührt die MR-Kugel in $\pm(1/\sqrt{2}, 1/\sqrt{2})$ und Einheitskugeln für kleinere p sind vollständig in der MR-Kugel enthalten. Dies veranschaulicht die unterschiedliche Gestalt der oberen Schranken in (5.14) und (5.15) für die Fälle $p \leq 2$ und $p > 2$.

Abschließend sollen nun noch mögliche weitere, auf der hier vorgestellten geometrischen Interpretation basierende, Anwendungen des Multiresolutionskriteriums skizziert werden. Viele Regressionsschätzer sind als Minimierer eines Komplexitätsfunktionals unter der Nebenbedingung (5.9) definiert. Diese Nebenbedingung entspricht einer endlichen Menge linearer Ungleichungen. Für einige spezielle Komplexitätsfunktionale, wie z.B. die Totalvariation von $\hat{f}$, reduziert sich die Berechnung des Schätzers auf die Lösung eines linearen Optimierungsproblems. Im Rahmen der oben gegebenen geometrischen Interpretation bedeutet dies, dass über eine Kugel in der MR-Norm optimiert wird. In diesem Fall definiert

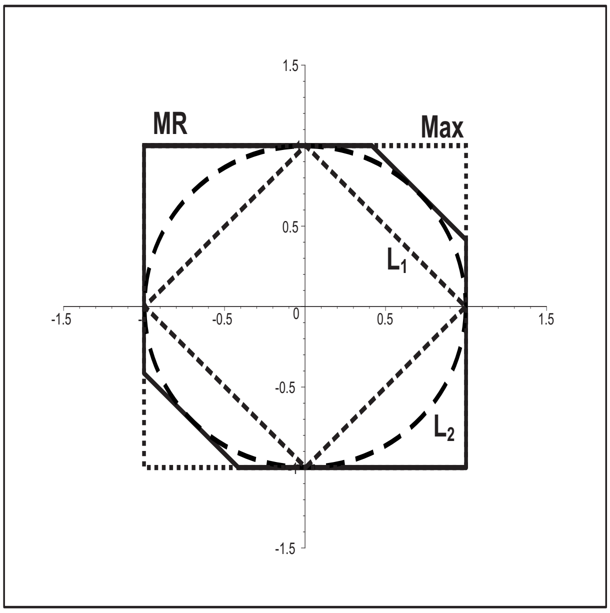

Abbildung 5.2.: Einheitskugeln in der MR-, 1-, 2- und ∞-Norm für $n = 2$.

die MR-Norm die Nebenbedingungen. Im Folgenden soll gezeigt werden, dass die lineare Approximation bezüglich der MR-Norm ebenfalls ein lineares Optimierungsproblem darstellt, in dem die MR-Norm die Rolle der Zielfunktion übernimmt. Desweiteren wird eine Anwendung auf Goodness-of-Fit-Tests für Lineare Modelle skizziert.

Betrachtet werde das Testproblem

$$H_0 : f = \sum_{i=1}^{k} \beta_i g_i \quad \text{vs.} \quad H_1 : f \neq \sum_{i=1}^{k} \beta_i g_i$$

im Regressionsmodel (5.1) für feste Funktionen $g_1, \ldots, g_k$. Zahlreiche auf einer Glättung der Residuen oder einem Vergleich einer parametrischen mit einer nichtparametrischen Schätzung basierende Tests wurden in der Literatur vorgeschlagen, vgl. zur Übersicht Kapitel 6 in [Har97]. Es bezeichne $X = (x_{ij}) = (g_j(t_i))$ die $n \times k$ Regressormatrix. Zu einem gegebenen Testniveau $\alpha \in (0, 1)$ sei $C_{n,1-\alpha}$ das $(1-\alpha)$-Quantil der linken Seite von (5.9) für $\sigma = 1$. Die Größe $C_{n,1-\alpha}$ kann durch Simulationen bestimmt werden. Wenn H_0 erfüllt ist für einen Vektor $\beta := (\beta_1, \ldots, \beta_k)^T$, liegt y mit Wahrscheinlichkeit $1-\alpha$ in einer MR-Kugel um $X\beta$ mit Radius $\sigma C_{n,1-\alpha}$. Dies legt folgende Testprozedur nahe: Zuerst wird die beste Approximation an y berechnet, d.h. ein Vektor $\hat{\beta} \in \mathbb{R}^k$, für den

$$\|X\hat{\beta} - y\|_{\mathsf{MR}} = \min_{\beta \in \mathbb{R}^k} \|X\beta - y\|_{\mathsf{MR}}.$$

Dies entspricht einer linearen Regression, wobei jedoch die übliche ℓ_2-Norm durch die MR-Norm ersetzt wird. H_0 wird genau dann abgelehnt, wenn $\|X\hat{\beta} - y\|_{\mathsf{MR}} > \sigma C_{n,1-\alpha}$.

Nun wird die Menge

$$B^* = \operatorname*{argmin}_{\beta \in \mathbb{R}^k} \|X\beta - y\|_{\mathsf{MR}}$$

charakterisiert. Da die Multiresolutionsnorm nicht strikt konvex ist, kann diese Menge aus mehr als einem Element bestehen, und da das Problem endlichdimensional ist, ist die Lösungsmenge nie leer. Analog zum Problem der Approximation bezüglich der ℓ_∞-Norm lässt sich das Problem mit Standardmethoden der linearen Optimierung lösen:

Satz 5.5. *Seien y, X, B^* wie oben. Dann besteht die Menge B^* aus allen $\beta \in \mathbb{R}^k$, die Lösungen des linearen Optimierungsproblems*

$$s \longrightarrow \min! \tag{5.20}$$

unter den Nebenbedingungen

$$s \geq 0 \tag{5.21}$$

$$\sqrt{|I|}s + \textstyle\sum_{t \in I}(X\beta)_t \geq \sum_{t \in I} y_t \quad \forall I \in \mathcal{I} \tag{5.22}$$

$$-\sqrt{|I|}s + \textstyle\sum_{t \in I}(X\beta)_t \leq \sum_{t \in I} y_t \quad \forall I \in \mathcal{I} \tag{5.23}$$

mit $\beta \in \mathbb{R}^k$ und $s \in \mathbb{R}$ sind.

Beweis. Sei $B_s := \{\beta \in \mathbb{R}^k \mid \|X\beta - y\|_{\mathsf{MR}} \leq s\}$. Es ist dann klar, dass $B^* = B_s$ für das kleinstmögliche $s \geq 0$, für das $B_s \neq \emptyset$. Das Minimum wird angenommen, da mindestens eine Bestapproximation existiert. Daraus ergibt sich die Zielfunktion und die Nebenbedingung (5.21). Die Bedingung $\|X\beta - y\|_{\mathsf{MR}} \leq s$ lässt sich durch elementare Umformungen in (5.22) und (5.23) transformieren. □

Da dieser Ansatz auf linearer Optimierung beruht, ist es möglich, z.B. Hypothesen über die Monotonie von f, die sich in Form linearer Nebenbedingungen formulieren lassen, zu testen.

5.3. Simulationsstudie

Im Folgenden werden verschiedene Versionen des Diskrepanzprinzips für die nichtparametrische Regression in einer Simulationsstudie verglichen. Dazu werden Datensätze aus acht verschiedenen Beispielfunktionen generiert. Neben den aus dem Wavelet-Kontext bekannten Beispielfunktionen von Donoho und Johnstone (vgl. [DJ94]) (hier Funktionen 4-7) sind dies eine etwas weniger stark oszillierende Version der Doppler-Funktion (Funktion 1, vgl. [RC00]), eine Sinusfunktion mit einem zusätzlichen Peak (Funktion 2, vgl. [Mei06]), eine stückweise polynomielle Funktion mit Sprungstelle (Funktion 3, vgl. [NS94]) und die Nullfunktion, die eine der Rechteckverteilung im Fall der Histogrammdichteschätzung analoge Rolle spielt und bereits in [DGW08] verwendet wurde (hier Funktion 8). Die Beispielfunktionen sind in Abbildung 5.3 abgebildet; die genauen Informationen und Quellenangaben finden sich in Anhang B.2. Für äquidistante Designpunkte werden jeweils 100 Datensätze vom Umfang $n = 50, 100, 1000$ erzeugt; für die Beispielfunktionen 3-7 geschieht dies mit

Hilfe des R-Pakets `waveband` [Bar10]. Zu den Funktionswerten werden jeweils unabhängig identisch verteilte Gaußfehler addiert. Fehler und Funktionswerte werden dabei für die Funktionen 1-7 so skaliert, dass sich die Signal-Rausch-Verhältnisse 3, 7 und 10 ergeben. Das *Signal-Rausch-Verhältnis* (*signal-to-noise-ratio*, in den Tabellen mit StN bezeichnet) ist dabei definiert als

$$\sqrt{\frac{\int_{[0,1]} f^2 - (\int_{[0,1]} f)^2}{\sigma^2}}.$$

Für Funktion 8 wird standardnormalverteiltes Rauschen verwendet.

Als Kern für den Nadaraya-Watson-Schätzer (5.2) wird dabei der Epanechnikov-Kern $K^{\mathcal{E}}$ (3.12) verwendet. Der Glättungsspline wird mit dem in Kapitel 2.3.3 von [GS94] gegebenen Algorithmus berechnet. Wie in [DM08] wird

$$\hat{\sigma} = \frac{1.4826}{\sqrt{2}} \operatorname*{med}_{i=2,\ldots,n} |y_i - y_{i-1}| \tag{5.24}$$

als Schätzung für σ verwendet. Mit Hilfe numerischer Integration wird jeweils der L_2-Verlust berechnet. Das arithmetische Mittel über die 100 Simulationsläufe dient als Schätzung für das zugehörige Risiko.

Verglichen werden folgende Versionen des Diskrepanzprinzips:

- Auf der ℓ_2-Norm basierende Varianten: Die Bandbreite h wird gewählt als größte Lösung von $\|r_h\|_2 \leq \hat{\sigma}\sqrt{F^{-1}_{\chi^2_n}(\alpha)}$ mit $\alpha = 0.5, 0.9, 0.95, 0.99$, wobei $F^{-1}_{\chi^2_n}$ die Quantilfunktion der χ^2-Verteilung mit n Freiheitsgraden bezeichnet. Die Wahl $\alpha = 0.5$ entspricht für die betrachteten Stichprobenumfänge fast der von Reinsch [Rei67] vorgeschlagenen Wahl. In den Tabellen als **Chi .5**, **Chi .9**, **Chi .95**, **Chi .99** bezeichnet.

- Auf festen Quantilen der MR-Norm basierende Varianten: h wird gewählt als größte Lösung von $\|r_h\|_{\mathsf{MR}} \leq \hat{\sigma} q_{\mathsf{MR},n,\alpha}$ mit $\alpha = 0.5, 0.9, 0.95, 0.99$. Dabei bezeichnet $q_{\mathsf{MR},n,\alpha}$ das α-Quantil von $\|(\varepsilon_1, \ldots, \varepsilon_n)\|_{\mathsf{MR}}$ für $\varepsilon_1, \ldots, \varepsilon_n$ unabhängig identisch standardnormalverteilt. Die Quantile werden auf Basis von Simulationen geschätzt. In den Tabellen als **MR .5**, **MR .9**, **MR .95**, **MR .99** bezeichnet.

- Auf von n abhängigen Quantilen der MR-Norm basierende Varianten: h wird gewählt als größte Lösung von $\|r_h\|_{\mathsf{MR}} \leq \hat{\sigma}\sqrt{\tau \log(n)}$ für $\tau = 2, 2.3, 2.5, 3$. In den Tabellen mit **MR 2**, **MR 2.3**, **MR 2.5**, **MR 3** bezeichnet.

Zum Vergleich werden folgende Referenzmethoden herangezogen:

- Kreuzvalidierung: Es wird diejenige Bandbreite h gewählt, die die Funktion

$$CV(h) := \frac{1}{n} \sum_{i=1}^{n} \frac{(y_i - \hat{f}_h(t_i))^2}{(1 - w_i^h)^2} \tag{5.25}$$

  minimiert, vgl. z.B. Formel (4.2.11) in [Wah90] (dort für Splines). In den Tabellen mit **CV** bezeichnet.

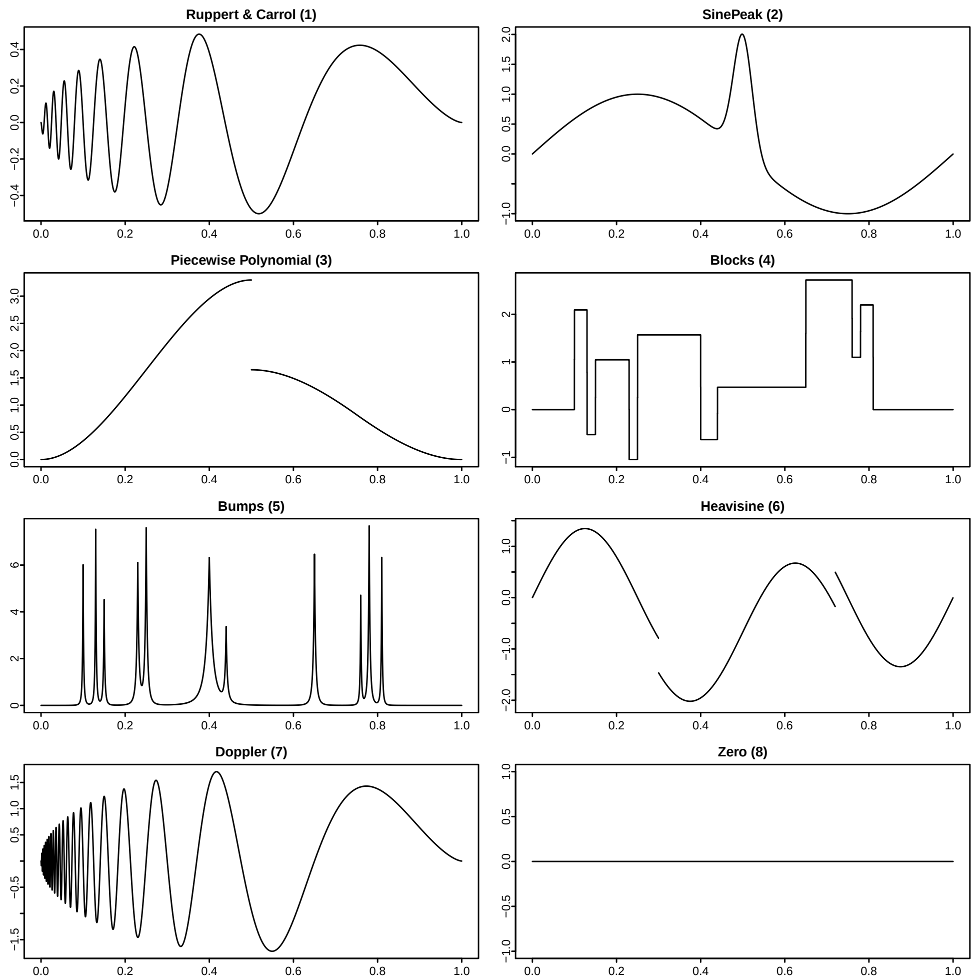

Abbildung 5.3.: Die verwendeten Signale.

- Generalisierte Kreuzvalidierung: Es wird diejenige Bandbreite h gewählt, die die Funktion

$$GCV(h) := \frac{1}{n}\sum_{i=1}^{n} \frac{(y_i - \hat{f}_h(t_i))^2}{(1 - \frac{1}{n}\sum_{j=1}^{n} w_j^h)^2} \tag{5.26}$$

  minimiert. Diese Methode wurde zuerst für Glättungssplines vorgeschlagen (vgl. Kapitel 4.3-4.4 in [Wah90]). Sie lässt sich aber auch für Nadaraya-Watson-Schätzer benutzen. In den Tabellen mit **GCV** bezeichnet.

- Mallow's C_p: Es wird diejenige Bandbreite h gewählt, die die Funktion

$$C_p(h) := \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{f}_h(t_i))^2 + \frac{2\hat{\sigma}^2}{n}\sum_{i=1}^{n} w_i^h \tag{5.27}$$

minimiert (vgl. [Tsy09, S. 63]). Das Kriterium (5.27) ist identisch mit dem AIC für den Fall normalverteilter Fehler mit bekannter Varianz; für den Fall unbekannter Varianz hat das AIC eine etwas andere Form, vgl. [Tsy09], S.71. Hier wird jedoch σ als Störparameter behandelt und mit Hilfe von (5.24) geschätzt. Diese Methode wird in den Tabellen mit **Cp** bezeichnet.

Die Kriterien sind durchgängig für Kernschätzer formuliert; für Glättungssplines ist jeweils lediglich h durch λ zu ersetzen. Mit Ausnahme von **CV** und **GCV** benötigen alle Methoden eine Schätzung für σ, für die (5.24) verwendet wird. Um zu untersuchen, wie groß der Einfluss der Schätzung auf das Resultat ist, werden zusätzlich auf dem wahren Wert für σ basierende Versionen betrachtet. Diese werden mit einem zusätzlichen Stern bezeichnet, also **Cp***, **Chi .5*** usw.

Die Ergebnisse dieser Simulationsstudie fallen wesentlich einheitlicher aus als die Ergebnisse für die Dichteschätzung. Die geschätzten Werte für das quadrierte L_2-Risiko für den Nadaraya-Watson-Kernschätzer finden sich in Tabellen 5.1-5.8. Zusätzlich sind (in Klammern) die arithmetischen Mittel der gewählten Bandbreiten angegeben.

Zunächst fällt auf, dass die verschiedenen Versionen des Diskrepanzprinzips größere Bandbreiten wählen als die Vergleichsmethoden **CV**, **GCV** und **Cp**. Dies führt zu schlechteren Ergebnissen bezüglich des L_2-Risikos. Die auf dem Multiresolutionskriterium basierenden Versionen **MR .5**, **MR .9**, **MR .95** und **MR .99** sind in den meisten Fällen besser als die entsprechenden auf dem ℓ_2-Abstand basierenden Methoden **Chi .5**, **Chi .9**, **Chi .95** und **Chi .99**. In einigen Fällen sind sie sogar erheblich besser, obwohl man zunächst vermuten könnte, dass ein ℓ_2-Diskrepanzprinzip bezüglich des L_2-Risikos besser funktioniert. Anscheinend führt die Fähigkeit des Multiresolutionskriteriums, Muster in den Residuen zu detektieren, aber dazu, dass die Bandbreiten stärker verkleinert werden und damit weniger stark überglättet wird. Obwohl das Multiresolutionskriterium besonders gut geeignet ist, um lokalisierte Glättungsparameter zu wählen, zeigt sich hier, dass eine Verwendung des Kriteriums auch bei der Wahl eines globalen Glättungsparameters gegenüber der ℓ_2-Norm Vorteile bietet. Grundsätzlich sind die auf kleineren Quantilen basierenden Versionen **Chi .5** und **MR .5** besser als die auf größeren Quantilen basierenden Versionen, auch wenn sie im Vergleich zu z.B. **CV** immer noch zu große Glättungsparameter wählen. Die auf mit n variierenden Quantilen basierenden Versionen **MR 2**, **MR 2.3**, **MR 2.5** und **MR 3** liegen zwischen **MR .5** und **MR .99** und glätten daher ebenfalls meistens zu stark.

Auffällig ist außerdem, dass die Schätzung für σ einen sehr großen Einfluss hat. Wird σ als bekannt vorausgesetzt, werden die auf dem Diskrepanzprinzip basierenden Methoden besser. Dies liegt hauptsächlich daran, dass (5.24) die Rauschvarianz σ überschätzt: Die Variation der Regressionsfunktion f verzerrt (5.24) nach oben, und dieser Effekt ist besonders stark, wenn n klein ist und wenn das Signal-Rausch-Verhältnis groß ist. Da σ vor der Trennung von Signal und Rauschen geschätzt werden muss, ist dies wohl unvermeidlich. Die unter Verwendung des wahren Werts für σ durch Varianten des Diskrepanzprinzips gewählten Bandbreiten sind daher kleiner als die auf Schätzungen für σ basierenden. Insbesondere **MR .5*** ist teilweise besser als die Kreuzvalidierungsmethode **CV**, die keine Schätzung für σ benötigt. Allerdings ist die Vergleichsmethode **Cp*** häufig noch besser.

Nur für wenige der Beispielfunktionen gibt es nennenswerte Abweichungen. Für Piecewise Polynomial (vgl. Tablle 5.3) sind für große Stichprobenumfänge und großes Signal-Rausch-Verhältnis die auf der ℓ_2-Norm basierenden Methoden besser als die auf dem Multiresolutionskriterium basierenden, und es zeigen sich teilweise auch keine Verbesserungen bei bekanntem σ. Für Blocks (Tabelle 5.4) und Bumps (Tabelle 5.5) schneiden **CV** und **GCV** teilweise sehr schlecht ab, wobei für die Bumps-Funktion für $n = 50$ alle Methoden nahe beieinander liegen. Für $n = 100$ sind die auf dem Diskrepanzprinzip basierenden Methoden jedoch deutlich besser als **CV**. Die Funktion Zero spielt eine ähnliche Rolle wie die Rechteckverteilung bei den regulären Histogrammen. Hierbei schneiden erwartungsgemäß diejenigen Methoden am besten ab, die die größten Bandbreiten wählen, wobei **MR .99** bei unbekanntem σ besser ist als **Chi .99**.

Die Ergebnisse für die Glättungssplines sind denen für den Nadaraya-Watson-Kernschätzer sehr ähnlich und befinden sich im Anhang (vgl. Tabellen B.24-B.31).

Insgesamt kann gesagt werden, dass die Verwendung des Diskrepanzprinzips bei unbekannter Rauschvarianz nicht unbedingt empfohlen werden kann. Bei bekannter Rauschvarianz kann das Diskrepanzprinzip jedoch durchaus attraktiv sein, zumal es als Abbruchkriterium unter Umständen wesentlich weniger Berechnungen von Schätzern für verschiedene Werte des Glättungsparameters erfordert als die Vergleichsmethoden, bei denen ein Minimum über verschiedene Werte des Glättungsparameters gebildet werden muss. Wird das Diskrepanzprinzip eingesetzt, so empfiehlt es sich, statt der ℓ_2-Norm das Multiresolutionskriterium zu verwenden, das fast immer bessere Ergebnisse liefert.

Tabelle 5.1.: Simulationsergebnisse für Kernschätzer: L_2-Risiko und gewählte Bandbreiten für Ruppert & Carroll.

StN		3			7			10	
n	50	100	1000	50	100	1000	50	100	1000
CV	0.0106	0.0058	9e-04	0.0052	0.0016	2e-04	0.0047	9e-04	1e-04
	(0.0287)	(0.0202)	(0.008)	(0.0201)	(0.0109)	(0.0053)	(0.0201)	(0.0101)	(0.0046)
GCV	0.0113	0.0058	9e-04	0.009	0.0028	2e-04	0.0087	0.0028	1e-04
	(0.0362)	(0.0208)	(0.008)	(0.0382)	(0.0183)	(0.0053)	(0.0389)	(0.0195)	(0.0046)
Cp	0.0122	0.0061	9e-04	0.009	0.0029	2e-04	0.0085	0.0023	1e-04
	(0.0431)	(0.0238)	(0.008)	(0.0386)	(0.0187)	(0.0055)	(0.0381)	(0.0174)	(0.0048)
Chi .5	0.0212	0.0104	0.0013	0.0142	0.0052	4e-04	0.0127	0.004	3e-04
	(0.0751)	(0.0412)	(0.0122)	(0.0559)	(0.0277)	(0.0085)	(0.0518)	(0.0244)	(0.0077)
Chi .9	0.0274	0.0129	0.0017	0.0176	0.0062	5e-04	0.0157	0.005	3e-04
	(0.0968)	(0.0499)	(0.0144)	(0.0669)	(0.0315)	(0.0093)	(0.0613)	(0.0277)	(0.0082)
Chi .95	0.0293	0.0137	0.0018	0.0187	0.0066	5e-04	0.0166	0.0052	3e-04
	(0.1039)	(0.0527)	(0.0151)	(0.0704)	(0.0326)	(0.0095)	(0.0642)	(0.0286)	(0.0084)
Chi .99	0.0334	0.0153	0.002	0.0209	0.0072	5e-04	0.0185	0.0058	3e-04
	(0.1191)	(0.0581)	(0.0162)	(0.0776)	(0.0348)	(0.0099)	(0.0702)	(0.0304)	(0.0086)
MR .5	0.0181	0.0082	0.001	0.0102	0.0036	3e-04	0.0083	0.0026	2e-04
	(0.0642)	(0.0331)	(0.0095)	(0.0426)	(0.0218)	(0.0065)	(0.0372)	(0.0189)	(0.0056)
MR .9	0.0259	0.0112	0.0012	0.0156	0.005	3e-04	0.0133	0.0037	2e-04
	(0.0916)	(0.0441)	(0.0115)	(0.0604)	(0.027)	(0.0073)	(0.0537)	(0.0231)	(0.0061)
MR .95	0.0288	0.0123	0.0012	0.0177	0.0055	3e-04	0.015	0.0041	2e-04
	(0.102)	(0.0479)	(0.0121)	(0.0672)	(0.0289)	(0.0075)	(0.0592)	(0.0246)	(0.0063)
MR .99	0.0355	0.0148	0.0014	0.0219	0.0067	3e-04	0.0192	0.0051	2e-04
	(0.1272)	(0.0566)	(0.0132)	(0.0811)	(0.033)	(0.008)	(0.0725)	(0.0282)	(0.0067)
MR 2	0.0183	0.0081	0.001	0.0103	0.0035	3e-04	0.0084	0.0026	2e-04
	(0.0648)	(0.0326)	(0.0091)	(0.0429)	(0.0216)	(0.0063)	(0.0374)	(0.0187)	(0.0055)
MR 2.3	0.0204	0.0091	0.0011	0.0117	0.004	3e-04	0.0096	0.003	2e-04
	(0.0724)	(0.0363)	(0.0103)	(0.0476)	(0.0235)	(0.0068)	(0.0418)	(0.0204)	(0.0058)
MR 2.5	0.0219	0.0097	0.0011	0.0127	0.0043	3e-04	0.0105	0.0032	2e-04
	(0.0774)	(0.0388)	(0.0109)	(0.0508)	(0.0246)	(0.007)	(0.0448)	(0.0213)	(0.006)
MR 3	0.0257	0.0115	0.0013	0.0154	0.0051	3e-04	0.0132	0.0038	2e-04
	(0.0909)	(0.045)	(0.0123)	(0.0599)	(0.0274)	(0.0076)	(0.0533)	(0.0234)	(0.0064)
Cp*	0.0097	0.0054	9e-04	0.0053	0.0015	2e-04	0.0047	0.001	1e-04
	(0.0234)	(0.014)	(0.0079)	(0.0206)	(0.0106)	(0.0053)	(0.0203)	(0.0103)	(0.0046)
Chi .5*	0.0106	0.006	0.0012	0.0056	0.0017	3e-04	0.0049	0.0011	2e-04
	(0.0352)	(0.0234)	(0.0114)	(0.0227)	(0.0127)	(0.0078)	(0.0216)	(0.0114)	(0.0065)
Chi .9*	0.0118	0.0069	0.0015	0.0058	0.0018	4e-04	0.005	0.0012	2e-04
	(0.0414)	(0.0277)	(0.0137)	(0.0235)	(0.0134)	(0.0086)	(0.0219)	(0.0117)	(0.0071)
Chi .95*	0.0123	0.0072	0.0016	0.0058	0.0019	4e-04	0.005	0.0012	2e-04
	(0.0433)	(0.0291)	(0.0143)	(0.0237)	(0.0137)	(0.0088)	(0.022)	(0.0118)	(0.0073)
Chi .99*	0.0131	0.0078	0.0018	0.0059	0.002	5e-04	0.005	0.0012	2e-04
	(0.0468)	(0.0317)	(0.0154)	(0.0242)	(0.0143)	(0.0092)	(0.0223)	(0.0119)	(0.0076)
MR .5*	0.0102	0.0059	0.001	0.0055	0.0016	3e-04	0.0048	0.0011	1e-04
	(0.0318)	(0.0222)	(0.0094)	(0.0217)	(0.0118)	(0.0063)	(0.021)	(0.0109)	(0.0053)
MR .9*	0.012	0.0071	0.0012	0.0056	0.0018	3e-04	0.0049	0.0011	2e-04
	(0.0415)	(0.0287)	(0.0114)	(0.0223)	(0.0128)	(0.0071)	(0.0213)	(0.0113)	(0.0059)
MR .95*	0.0129	0.0077	0.0012	0.0056	0.0018	3e-04	0.0049	0.0011	2e-04
	(0.0456)	(0.031)	(0.012)	(0.0226)	(0.0132)	(0.0073)	(0.0214)	(0.0114)	(0.0061)
MR .99*	0.015	0.0091	0.0014	0.0057	0.002	3e-04	0.0049	0.0012	2e-04
	(0.0536)	(0.0365)	(0.0131)	(0.0232)	(0.0146)	(0.0078)	(0.0217)	(0.0117)	(0.0064)
MR 2*	0.0102	0.0059	0.001	0.0055	0.0016	3e-04	0.0048	0.0011	1e-04
	(0.0319)	(0.022)	(0.0089)	(0.0217)	(0.0118)	(0.0061)	(0.021)	(0.0109)	(0.0052)
MR 2.3*	0.0107	0.0062	0.0011	0.0055	0.0017	3e-04	0.0048	0.0011	2e-04
	(0.0348)	(0.0242)	(0.0101)	(0.0219)	(0.0121)	(0.0066)	(0.0211)	(0.011)	(0.0055)
MR 2.5*	0.011	0.0065	0.0011	0.0055	0.0017	3e-04	0.0048	0.0011	2e-04
	(0.0367)	(0.0255)	(0.0108)	(0.022)	(0.0123)	(0.0069)	(0.0212)	(0.0111)	(0.0057)
MR 3*	0.0119	0.0072	0.0012	0.0056	0.0018	3e-04	0.0049	0.0011	2e-04
	(0.0413)	(0.0291)	(0.0121)	(0.0223)	(0.0129)	(0.0074)	(0.0213)	(0.0113)	(0.0061)

Tabelle 5.2.: Simulationsergebnisse für Kernschätzer: L_2-Risiko und gewählte Bandbreiten für SinePeak.

StN		3			7			10	
n	50	100	1000	50	100	1000	50	100	1000
CV	0.0352	0.019	0.003	0.0097	0.0051	8e-04	0.0076	0.0031	4e-04
	(0.0413)	(0.0322)	(0.0187)	(0.0308)	(0.0222)	(0.0132)	(0.0332)	(0.0175)	(0.0114)
GCV	0.0349	0.0195	0.003	0.0107	0.0051	8e-04	0.0084	0.0031	4e-04
	(0.0409)	(0.0317)	(0.0187)	(0.033)	(0.0221)	(0.0132)	(0.0345)	(0.018)	(0.0114)
Cp	0.0331	0.0191	0.003	0.0106	0.0051	8e-04	0.0075	0.0029	4e-04
	(0.0473)	(0.0343)	(0.0186)	(0.0357)	(0.0256)	(0.0133)	(0.0332)	(0.0221)	(0.0116)
Chi .5	0.0523	0.0293	0.0055	0.0175	0.0094	0.0013	0.0123	0.0059	7e-04
	(0.0789)	(0.0564)	(0.0276)	(0.0487)	(0.039)	(0.0197)	(0.0438)	(0.034)	(0.017)
Chi .9	0.072	0.0399	0.0082	0.023	0.012	0.0019	0.0154	0.0074	0.001
	(0.106)	(0.0709)	(0.0364)	(0.0558)	(0.0437)	(0.0239)	(0.0483)	(0.0373)	(0.0201)
Chi .95	0.0787	0.0435	0.0091	0.0248	0.0129	0.002	0.0164	0.0078	0.0011
	(0.116)	(0.0755)	(0.0385)	(0.058)	(0.0451)	(0.0249)	(0.0497)	(0.0383)	(0.0209)
Chi .99	0.0925	0.0508	0.0111	0.0281	0.0145	0.0024	0.0187	0.0088	0.0013
	(0.137)	(0.085)	(0.0424)	(0.0618)	(0.0475)	(0.0266)	(0.0526)	(0.0401)	(0.0222)
MR .5	0.0368	0.0201	0.0035	0.0105	0.0052	9e-04	0.0062	0.0029	5e-04
	(0.0568)	(0.0417)	(0.0202)	(0.0353)	(0.0271)	(0.0137)	(0.029)	(0.0229)	(0.0118)
MR .9	0.0452	0.0232	0.0033	0.0128	0.006	8e-04	0.0083	0.0034	4e-04
	(0.0697)	(0.0489)	(0.0242)	(0.0411)	(0.0307)	(0.0156)	(0.0358)	(0.0257)	(0.0132)
MR .95	0.0488	0.0247	0.0034	0.0136	0.0063	8e-04	0.0091	0.0036	5e-04
	(0.0746)	(0.0514)	(0.0251)	(0.0427)	(0.0317)	(0.0162)	(0.0376)	(0.0268)	(0.0136)
MR .99	0.0574	0.028	0.0037	0.0156	0.007	9e-04	0.0105	0.004	5e-04
	(0.0858)	(0.0564)	(0.0267)	(0.0461)	(0.0339)	(0.0171)	(0.0407)	(0.0287)	(0.0143)
MR 2	0.0369	0.02	0.0037	0.0106	0.0052	9e-04	0.0062	0.0029	5e-04
	(0.0571)	(0.0414)	(0.0192)	(0.0354)	(0.0269)	(0.0131)	(0.0291)	(0.0228)	(0.0114)
MR 2.3	0.0391	0.0209	0.0034	0.0113	0.0054	8e-04	0.0068	0.003	4e-04
	(0.0608)	(0.044)	(0.0219)	(0.0375)	(0.0283)	(0.0144)	(0.0311)	(0.0237)	(0.0124)
MR 2.5	0.0407	0.0216	0.0033	0.0117	0.0056	8e-04	0.0072	0.0031	4e-04
	(0.0632)	(0.0456)	(0.0232)	(0.0386)	(0.0291)	(0.0151)	(0.0325)	(0.0243)	(0.0128)
MR 3	0.045	0.0236	0.0035	0.0127	0.006	8e-04	0.0083	0.0034	5e-04
	(0.0694)	(0.0495)	(0.0255)	(0.041)	(0.0309)	(0.0164)	(0.0356)	(0.0259)	(0.0137)
Cp*	0.0333	0.0187	0.003	0.0087	0.0048	8e-04	0.0052	0.0029	4e-04
	(0.0383)	(0.0323)	(0.0187)	(0.0251)	(0.0228)	(0.0132)	(0.0223)	(0.0185)	(0.0114)
Chi .5*	0.038	0.0237	0.0048	0.0115	0.0066	0.0012	0.0063	0.0038	6e-04
	(0.0587)	(0.049)	(0.0266)	(0.039)	(0.033)	(0.0194)	(0.0304)	(0.0275)	(0.0162)
Chi .9*	0.0489	0.0321	0.0074	0.0134	0.0085	0.0017	0.0083	0.0046	9e-04
	(0.0745)	(0.062)	(0.036)	(0.043)	(0.0378)	(0.0236)	(0.0364)	(0.0307)	(0.0195)
Chi .95*	0.053	0.0352	0.0083	0.014	0.0092	0.0019	0.009	0.0048	0.001
	(0.0799)	(0.066)	(0.0383)	(0.044)	(0.0391)	(0.0246)	(0.0381)	(0.0315)	(0.0203)
Chi .99*	0.0618	0.0415	0.0104	0.0153	0.0103	0.0023	0.01	0.0053	0.0011
	(0.0913)	(0.0737)	(0.0422)	(0.0461)	(0.0414)	(0.0263)	(0.0403)	(0.0328)	(0.0216)
MR .5*	0.0336	0.0192	0.0035	0.0086	0.0048	9e-04	0.0053	0.0028	4e-04
	(0.0502)	(0.04)	(0.0203)	(0.027)	(0.0246)	(0.0134)	(0.0232)	(0.0212)	(0.0118)
MR .9*	0.0388	0.0219	0.0033	0.01	0.0055	8e-04	0.0054	0.0029	4e-04
	(0.0606)	(0.0472)	(0.0242)	(0.0343)	(0.0288)	(0.0158)	(0.0252)	(0.0227)	(0.0132)
MR .95*	0.0414	0.0232	0.0034	0.0107	0.0057	8e-04	0.0054	0.003	4e-04
	(0.0644)	(0.0496)	(0.025)	(0.0365)	(0.03)	(0.0162)	(0.0261)	(0.0234)	(0.0135)
MR .99*	0.047	0.0261	0.0037	0.0119	0.0061	9e-04	0.0061	0.0032	5e-04
	(0.0722)	(0.0541)	(0.0267)	(0.0399)	(0.0316)	(0.0171)	(0.0291)	(0.0251)	(0.0142)
MR 2*	0.0336	0.0191	0.0037	0.0086	0.0048	9e-04	0.0053	0.0028	5e-04
	(0.0504)	(0.0396)	(0.0191)	(0.0271)	(0.0245)	(0.0129)	(0.0232)	(0.0212)	(0.0114)
MR 2.3*	0.035	0.0199	0.0033	0.0088	0.005	8e-04	0.0053	0.0028	4e-04
	(0.0534)	(0.0424)	(0.0221)	(0.029)	(0.026)	(0.0142)	(0.0237)	(0.0217)	(0.0123)
MR 2.5*	0.036	0.0205	0.0033	0.0091	0.0052	8e-04	0.0053	0.0028	4e-04
	(0.0555)	(0.0442)	(0.0231)	(0.0304)	(0.027)	(0.0151)	(0.0241)	(0.022)	(0.0128)
MR 3*	0.0387	0.0222	0.0035	0.01	0.0055	8e-04	0.0054	0.0029	5e-04
	(0.0603)	(0.0478)	(0.0253)	(0.0341)	(0.0291)	(0.0163)	(0.0251)	(0.0228)	(0.0136)

Tabelle 5.3.: Simulationsergebnisse für Kernschätzer: L_2-Risiko und gewählte Bandbreiten für Piecewise Polynomial.

StN		3			7			10	
n	50	100	1000	50	100	1000	50	100	1000
CV	0.0505	0.0327	0.0097	0.0265	0.0164	0.0041	0.022	0.0125	0.0029
	(0.0736)	(0.0513)	(0.015)	(0.0453)	(0.0289)	(0.0069)	(0.0414)	(0.0229)	(0.005)
GCV	0.0511	0.0338	0.0097	0.026	0.0164	0.0041	0.022	0.0123	0.0029
	(0.0717)	(0.0496)	(0.015)	(0.0471)	(0.0284)	(0.0069)	(0.0409)	(0.0235)	(0.005)
Cp	0.0513	0.0339	0.0098	0.0258	0.0163	0.0041	0.0207	0.0122	0.0029
	(0.0722)	(0.0483)	(0.0144)	(0.0384)	(0.0216)	(0.0061)	(0.0324)	(0.0145)	(0.0041)
Chi .5	0.065	0.0391	0.0115	0.0287	0.0165	0.0042	0.0233	0.0123	0.003
	(0.1172)	(0.0776)	(0.0225)	(0.0594)	(0.0322)	(0.008)	(0.0502)	(0.0234)	(0.0054)
Chi .9	0.0924	0.0517	0.0141	0.0339	0.0185	0.0046	0.0269	0.0131	0.0031
	(0.1674)	(0.113)	(0.0351)	(0.0782)	(0.0419)	(0.0102)	(0.0639)	(0.0284)	(0.0064)
Chi .95	0.1023	0.0566	0.0151	0.0358	0.0192	0.0047	0.0282	0.0135	0.0032
	(0.1811)	(0.1241)	(0.0392)	(0.0839)	(0.045)	(0.0109)	(0.0681)	(0.0299)	(0.0067)
Chi .99	0.1228	0.0673	0.0175	0.0399	0.0209	0.005	0.031	0.0142	0.0033
	(0.2057)	(0.1448)	(0.0475)	(0.0953)	(0.0515)	(0.0123)	(0.0767)	(0.0331)	(0.0074)
MR .5	0.0553	0.0335	0.0122	0.0251	0.0164	0.0072	0.0203	0.0123	0.0053
	(0.0954)	(0.0492)	(0.0083)	(0.0366)	(0.0171)	(0.0021)	(0.028)	(0.0128)	(0.0013)
MR .9	0.0688	0.0356	0.0104	0.0268	0.0158	0.0062	0.021	0.0121	0.0049
	(0.1315)	(0.0724)	(0.0111)	(0.0505)	(0.0225)	(0.0026)	(0.0359)	(0.0147)	(0.0015)
MR .95	0.0742	0.0373	0.0101	0.0279	0.0158	0.0059	0.0215	0.012	0.0048
	(0.1425)	(0.0803)	(0.0122)	(0.0559)	(0.0247)	(0.0027)	(0.0393)	(0.0156)	(0.0015)
MR .99	0.0868	0.0422	0.0098	0.0304	0.0161	0.0053	0.0228	0.012	0.0044
	(0.1648)	(0.097)	(0.0146)	(0.067)	(0.0294)	(0.0031)	(0.0471)	(0.0178)	(0.0017)
MR 2	0.0555	0.0335	0.0127	0.0251	0.0164	0.0074	0.0203	0.0123	0.0054
	(0.0962)	(0.0484)	(0.0078)	(0.0368)	(0.0169)	(0.002)	(0.0282)	(0.0127)	(0.0013)
MR 2.3	0.0592	0.0337	0.0115	0.0253	0.0161	0.0068	0.0204	0.0122	0.0052
	(0.1082)	(0.0573)	(0.0091)	(0.0405)	(0.0186)	(0.0022)	(0.0302)	(0.0133)	(0.0013)
MR 2.5	0.0619	0.0341	0.0109	0.0256	0.0159	0.0065	0.0205	0.0122	0.0051
	(0.1156)	(0.0623)	(0.0101)	(0.0431)	(0.0198)	(0.0024)	(0.0316)	(0.0138)	(0.0014)
MR 3	0.0684	0.0359	0.01	0.0267	0.0158	0.0058	0.0209	0.012	0.0047
	(0.1307)	(0.0742)	(0.0126)	(0.05)	(0.023)	(0.0028)	(0.0356)	(0.0149)	(0.0016)
Cp*	0.0545	0.033	0.0097	0.0262	0.0171	0.0041	0.021	0.0125	0.0029
	(0.0623)	(0.0443)	(0.0146)	(0.0257)	(0.0169)	(0.006)	(0.0219)	(0.012)	(0.0041)
Chi .5*	0.0533	0.0325	0.0104	0.0251	0.0154	0.0042	0.0201	0.0118	0.003
	(0.0847)	(0.0579)	(0.0189)	(0.04)	(0.0257)	(0.0078)	(0.0284)	(0.0193)	(0.0054)
Chi .9*	0.0657	0.0405	0.0126	0.0264	0.0162	0.0045	0.0205	0.012	0.0031
	(0.1262)	(0.0927)	(0.0311)	(0.0509)	(0.0325)	(0.0098)	(0.0338)	(0.0225)	(0.0064)
Chi .95*	0.0716	0.0444	0.0136	0.0271	0.0167	0.0046	0.0207	0.0122	0.0032
	(0.1398)	(0.1038)	(0.0353)	(0.0544)	(0.0349)	(0.0105)	(0.0356)	(0.0236)	(0.0067)
Chi .99*	0.0859	0.0533	0.016	0.0288	0.0177	0.0049	0.0211	0.0125	0.0033
	(0.1658)	(0.1258)	(0.0437)	(0.0621)	(0.0397)	(0.0119)	(0.0393)	(0.0258)	(0.0074)
MR .5*	0.0519	0.0327	0.0121	0.0253	0.0168	0.0072	0.0204	0.0125	0.0053
	(0.0796)	(0.0441)	(0.0082)	(0.0296)	(0.0156)	(0.0021)	(0.0231)	(0.012)	(0.0013)
MR .9*	0.0592	0.0334	0.0104	0.0252	0.0159	0.0062	0.0201	0.0122	0.0049
	(0.1134)	(0.0672)	(0.011)	(0.0377)	(0.0198)	(0.0025)	(0.025)	(0.0134)	(0.0015)
MR .95*	0.0634	0.0351	0.01	0.0254	0.0157	0.0059	0.0201	0.0121	0.0048
	(0.1247)	(0.0751)	(0.0121)	(0.041)	(0.0214)	(0.0027)	(0.0258)	(0.014)	(0.0015)
MR .99*	0.0731	0.0394	0.0097	0.0263	0.0157	0.0054	0.0202	0.012	0.0044
	(0.1464)	(0.091)	(0.0144)	(0.0492)	(0.0256)	(0.0031)	(0.0284)	(0.0156)	(0.0017)
MR 2*	0.0519	0.0328	0.0127	0.0253	0.0168	0.0074	0.0204	0.0125	0.0054
	(0.0802)	(0.0431)	(0.0077)	(0.0297)	(0.0155)	(0.002)	(0.0231)	(0.012)	(0.0013)
MR 2.3*	0.0532	0.0322	0.0114	0.0252	0.0164	0.0068	0.0203	0.0124	0.0052
	(0.0898)	(0.0504)	(0.0091)	(0.0319)	(0.0169)	(0.0022)	(0.0236)	(0.0123)	(0.0014)
MR 2.5*	0.0545	0.0321	0.0108	0.0252	0.0162	0.0065	0.0203	0.0123	0.0051
	(0.097)	(0.056)	(0.01)	(0.0335)	(0.0178)	(0.0024)	(0.024)	(0.0126)	(0.0014)
MR 3*	0.0589	0.0337	0.01	0.0252	0.0158	0.0058	0.0201	0.0122	0.0047
	(0.1125)	(0.069)	(0.0124)	(0.0375)	(0.0201)	(0.0028)	(0.0249)	(0.0135)	(0.0016)

Tabelle 5.4.: Simulationsergebnisse für Kernschätzer: L_2-Risiko und gewählte Bandbreiten für Blocks.

StN		3			7			10	
n	50	100	1000	50	100	1000	50	100	1000
CV	0.3522	0.1382	0.0372	0.3208	0.0798	0.0163	0.3236	0.0735	0.0122
	(0.0677)	(0.0122)	(0.004)	(0.0602)	(0.0101)	(0.0021)	(0.0616)	(0.0101)	(0.0014)
GCV	0.3111	0.1489	0.0371	0.1395	0.1222	0.0163	0.1158	0.1177	0.0129
	(0.0565)	(0.0186)	(0.004)	(0.0233)	(0.0195)	(0.0021)	(0.02)	(0.0192)	(0.0019)
Cp	0.2018	0.1499	0.037	0.1255	0.0783	0.0162	0.117	0.0702	0.0115
	(0.0258)	(0.0196)	(0.0038)	(0.0205)	(0.0119)	(0.0021)	(0.0202)	(0.0107)	(0.0015)
Chi .5	0.2575	0.2101	0.0418	0.1406	0.1137	0.0194	0.1252	0.0784	0.0142
	(0.0365)	(0.0306)	(0.006)	(0.0226)	(0.0181)	(0.0031)	(0.0215)	(0.013)	(0.0023)
Chi .9	0.2971	0.2379	0.0449	0.1463	0.1224	0.0201	0.1276	0.0852	0.0146
	(0.0486)	(0.0358)	(0.0069)	(0.0233)	(0.0197)	(0.0033)	(0.0218)	(0.0141)	(0.0024)
Chi .95	0.3097	0.2464	0.0459	0.1482	0.1243	0.0203	0.1285	0.0878	0.0148
	(0.053)	(0.0376)	(0.0071)	(0.0236)	(0.0201)	(0.0033)	(0.0219)	(0.0144)	(0.0025)
Chi .99	0.3361	0.2638	0.048	0.1523	0.128	0.0207	0.1301	0.0932	0.0151
	(0.0626)	(0.0416)	(0.0076)	(0.0241)	(0.0207)	(0.0034)	(0.0221)	(0.0153)	(0.0025)
MR .5	0.231	0.1868	0.0375	0.1354	0.0884	0.0162	0.1226	0.0711	0.0113
	(0.0305)	(0.0266)	(0.004)	(0.022)	(0.0139)	(0.002)	(0.0211)	(0.0115)	(0.0014)
MR .9	0.2933	0.2249	0.0392	0.1425	0.1061	0.0164	0.1258	0.0746	0.0121
	(0.0491)	(0.0336)	(0.005)	(0.0228)	(0.0169)	(0.0022)	(0.0216)	(0.0123)	(0.0016)
MR .95	0.3196	0.2398	0.0402	0.1455	0.1113	0.0166	0.127	0.0767	0.0123
	(0.0589)	(0.0366)	(0.0054)	(0.0232)	(0.0178)	(0.0022)	(0.0217)	(0.0127)	(0.0017)
MR .99	0.3713	0.2712	0.043	0.1529	0.1209	0.0172	0.1298	0.0833	0.0128
	(0.079)	(0.044)	(0.0063)	(0.0242)	(0.0195)	(0.0024)	(0.0221)	(0.0138)	(0.0019)
MR 2	0.232	0.1854	0.0374	0.1355	0.0879	0.0162	0.1227	0.071	0.0112
	(0.0307)	(0.0263)	(0.0038)	(0.022)	(0.0138)	(0.0019)	(0.0212)	(0.0115)	(0.0014)
MR 2.3	0.2499	0.1982	0.0379	0.1374	0.0937	0.0163	0.1236	0.0718	0.0115
	(0.0352)	(0.0286)	(0.0043)	(0.0222)	(0.0147)	(0.002)	(0.0213)	(0.0118)	(0.0015)
MR 2.5	0.261	0.2064	0.0385	0.1387	0.0978	0.0163	0.1242	0.0726	0.0118
	(0.0383)	(0.03)	(0.0047)	(0.0224)	(0.0154)	(0.0021)	(0.0214)	(0.0119)	(0.0015)
MR 3	0.2901	0.2281	0.0406	0.1423	0.1072	0.0167	0.1257	0.0749	0.0124
	(0.0477)	(0.0342)	(0.0055)	(0.0228)	(0.0171)	(0.0023)	(0.0216)	(0.0124)	(0.0018)
Cp*	0.1806	0.1289	0.0368	0.1238	0.076	0.0162	0.1162	0.0715	0.0112
	(0.0211)	(0.0138)	(0.0037)	(0.0202)	(0.0106)	(0.0019)	(0.0201)	(0.0103)	(0.0014)
Chi .5*	0.2023	0.158	0.0397	0.1303	0.0809	0.0187	0.1202	0.0706	0.0139
	(0.0249)	(0.0216)	(0.0054)	(0.0213)	(0.0127)	(0.0029)	(0.0208)	(0.0115)	(0.0023)
Chi .9*	0.2141	0.1692	0.0422	0.1321	0.0854	0.0193	0.1211	0.0715	0.0142
	(0.0269)	(0.0237)	(0.0062)	(0.0215)	(0.0135)	(0.0031)	(0.0209)	(0.0118)	(0.0024)
Chi .95*	0.2185	0.1736	0.0431	0.1327	0.0872	0.0195	0.1214	0.0719	0.0144
	(0.0277)	(0.0244)	(0.0064)	(0.0216)	(0.0138)	(0.0031)	(0.021)	(0.0119)	(0.0024)
Chi .99*	0.2289	0.1834	0.0449	0.1339	0.0914	0.0198	0.122	0.0727	0.0146
	(0.0296)	(0.026)	(0.0069)	(0.0218)	(0.0144)	(0.0032)	(0.0211)	(0.012)	(0.0024)
MR .5*	0.1953	0.1552	0.0374	0.1284	0.0765	0.0162	0.1191	0.0698	0.0112
	(0.0238)	(0.0208)	(0.0039)	(0.021)	(0.0116)	(0.0019)	(0.0206)	(0.0109)	(0.0014)
MR .9*	0.2118	0.1744	0.0387	0.1308	0.08	0.0164	0.1204	0.0701	0.0119
	(0.0265)	(0.0245)	(0.0048)	(0.0214)	(0.0125)	(0.0021)	(0.0208)	(0.0112)	(0.0016)
MR .95*	0.2202	0.1825	0.0396	0.1317	0.0822	0.0165	0.1208	0.0704	0.0122
	(0.028)	(0.0259)	(0.0052)	(0.0215)	(0.0129)	(0.0022)	(0.0209)	(0.0114)	(0.0017)
MR .99*	0.244	0.2036	0.0421	0.1339	0.089	0.0169	0.1219	0.0712	0.0127
	(0.0326)	(0.0293)	(0.006)	(0.0218)	(0.014)	(0.0023)	(0.0211)	(0.0117)	(0.0019)
MR 2*	0.1956	0.1545	0.0374	0.1285	0.0764	0.0162	0.1191	0.0698	0.0111
	(0.0238)	(0.0206)	(0.0037)	(0.021)	(0.0116)	(0.0018)	(0.0206)	(0.0109)	(0.0013)
MR 2.3*	0.1996	0.1605	0.0377	0.1292	0.0773	0.0162	0.1195	0.0699	0.0114
	(0.0245)	(0.0219)	(0.0042)	(0.0211)	(0.0119)	(0.002)	(0.0207)	(0.011)	(0.0014)
MR 2.5*	0.2027	0.1648	0.0381	0.1296	0.078	0.0163	0.1198	0.0699	0.0116
	(0.025)	(0.0227)	(0.0045)	(0.0212)	(0.012)	(0.002)	(0.0207)	(0.0111)	(0.0015)
MR 3*	0.2113	0.1761	0.0399	0.1308	0.0804	0.0165	0.1204	0.0702	0.0123
	(0.0264)	(0.0248)	(0.0053)	(0.0214)	(0.0125)	(0.0022)	(0.0208)	(0.0113)	(0.0017)

Tabelle 5.5.: Simulationsergebnisse für Kernschätzer: L_2-Risiko und gewählte Bandbreiten für Bumps.

StN		3			7			10	
n	50	100	1000	50	100	1000	50	100	1000
CV	0.9139	0.785	0.0514	0.8794	0.7532	0.0154	0.8582	0.7535	0.0097
	(1.0276)	(0.2708)	(0.0022)	(0.6154)	(0.0598)	(0.0012)	(0.2931)	(0.06)	(0.0012)
GCV	0.9302	0.7158	0.0514	0.9014	0.2842	0.0222	0.8936	0.2445	0.019
	(1.3803)	(0.358)	(0.0022)	(1.0564)	(0.0135)	(0.002)	(0.9692)	(0.01)	(0.002)
Cp	0.8955	0.326	0.0523	0.8384	0.2526	0.0146	0.8307	0.2472	0.0081
	(0.0211)	(0.0104)	(0.0018)	(0.0202)	(0.0101)	(0.0011)	(0.0201)	(0.01)	(0.0011)
Chi .5	0.9235	0.4105	0.0607	0.8519	0.2806	0.0164	0.8402	0.2653	0.0103
	(0.0247)	(0.0118)	(0.0028)	(0.0213)	(0.0105)	(0.0014)	(0.0209)	(0.0103)	(0.0012)
Chi .9	0.9226	0.4292	0.0643	0.8554	0.2852	0.0168	0.8426	0.268	0.0104
	(0.0268)	(0.0122)	(0.003)	(0.0215)	(0.0106)	(0.0015)	(0.0211)	(0.0104)	(0.0012)
Chi .95	0.921	0.4351	0.0654	0.8565	0.2867	0.0169	0.8433	0.2691	0.0104
	(0.0277)	(0.0124)	(0.0031)	(0.0216)	(0.0106)	(0.0015)	(0.0211)	(0.0104)	(0.0012)
Chi .99	0.917	0.4466	0.0674	0.8586	0.2891	0.0172	0.8447	0.2705	0.0105
	(0.0297)	(0.0127)	(0.0032)	(0.0217)	(0.0106)	(0.0016)	(0.0212)	(0.0104)	(0.0012)
MR .5	0.9174	0.4079	0.0614	0.8485	0.2802	0.0142	0.8379	0.2651	0.0079
	(0.0234)	(0.0118)	(0.0012)	(0.021)	(0.0105)	(0.0011)	(0.0207)	(0.0103)	(0.001)
MR .9	0.9234	0.4503	0.0595	0.8536	0.2903	0.0142	0.8413	0.2714	0.0079
	(0.0258)	(0.0128)	(0.0013)	(0.0214)	(0.0106)	(0.0011)	(0.021)	(0.0104)	(0.001)
MR .95	0.9216	0.4663	0.0588	0.8556	0.2941	0.0143	0.8425	0.2736	0.008
	(0.0271)	(0.0133)	(0.0013)	(0.0215)	(0.0107)	(0.0011)	(0.0211)	(0.0105)	(0.001)
MR .99	0.9137	0.5013	0.0569	0.8594	0.3024	0.0143	0.8452	0.2786	0.008
	(0.0305)	(0.015)	(0.0014)	(0.0218)	(0.0108)	(0.0011)	(0.0212)	(0.0105)	(0.001)
MR 2	0.9176	0.4064	0.0617	0.8486	0.2798	0.0141	0.838	0.2649	0.0078
	(0.0234)	(0.0117)	(0.0012)	(0.021)	(0.0105)	(0.001)	(0.0207)	(0.0103)	(0.001)
MR 2.3	0.9206	0.4207	0.0608	0.8501	0.2835	0.0142	0.839	0.2671	0.0079
	(0.024)	(0.012)	(0.0012)	(0.0211)	(0.0105)	(0.0011)	(0.0208)	(0.0104)	(0.001)
MR 2.5	0.9221	0.4303	0.0602	0.8511	0.2855	0.0142	0.8397	0.2684	0.0079
	(0.0244)	(0.0123)	(0.0012)	(0.0212)	(0.0106)	(0.0011)	(0.0209)	(0.0104)	(0.001)
MR 3	0.9235	0.4537	0.0585	0.8535	0.2911	0.0143	0.8413	0.2719	0.008
	(0.0257)	(0.0129)	(0.0013)	(0.0214)	(0.0106)	(0.0011)	(0.021)	(0.0104)	(0.001)
Cp*	0.8901	0.3125	0.0574	0.8375	0.2521	0.0143	0.8297	0.2445	0.0078
	(0.0205)	(0.0102)	(0.0013)	(0.0201)	(0.0101)	(0.0011)	(0.02)	(0.01)	(0.001)
Chi .5*	0.9117	0.3608	0.0531	0.8453	0.2665	0.0157	0.835	0.2547	0.0095
	(0.0226)	(0.0109)	(0.0023)	(0.0208)	(0.0103)	(0.0013)	(0.0205)	(0.0102)	(0.0012)
Chi .9*	0.9177	0.3702	0.0548	0.8472	0.2702	0.0158	0.8361	0.2584	0.0096
	(0.0232)	(0.0111)	(0.0025)	(0.0209)	(0.0103)	(0.0013)	(0.0206)	(0.0102)	(0.0012)
Chi .95*	0.9196	0.3727	0.0554	0.8478	0.2703	0.0159	0.8363	0.2584	0.0096
	(0.0235)	(0.0111)	(0.0025)	(0.021)	(0.0103)	(0.0013)	(0.0206)	(0.0102)	(0.0012)
Chi .99*	0.9229	0.3786	0.0567	0.849	0.2716	0.0159	0.8367	0.2584	0.0097
	(0.0239)	(0.0112)	(0.0026)	(0.021)	(0.0104)	(0.0013)	(0.0206)	(0.0102)	(0.0012)
MR .5*	0.9061	0.3596	0.0623	0.8435	0.2668	0.0141	0.8337	0.2549	0.0078
	(0.022)	(0.0109)	(0.0012)	(0.0206)	(0.0103)	(0.001)	(0.0204)	(0.0102)	(0.001)
MR .9*	0.915	0.3801	0.0609	0.8462	0.2726	0.0141	0.8352	0.2584	0.0078
	(0.0229)	(0.0113)	(0.0012)	(0.0208)	(0.0104)	(0.0011)	(0.0205)	(0.0102)	(0.001)
MR .95*	0.9181	0.3882	0.0603	0.8472	0.2744	0.0142	0.836	0.2604	0.0078
	(0.0233)	(0.0114)	(0.0012)	(0.0209)	(0.0104)	(0.0011)	(0.0206)	(0.0103)	(0.001)
MR .99*	0.9244	0.406	0.0589	0.8494	0.2787	0.0142	0.8371	0.2622	0.0079
	(0.0242)	(0.0117)	(0.0013)	(0.0211)	(0.0105)	(0.0011)	(0.0207)	(0.0103)	(0.001)
MR 2*	0.9063	0.3588	0.0626	0.8436	0.2666	0.0141	0.8337	0.2547	0.0078
	(0.022)	(0.0109)	(0.0011)	(0.0206)	(0.0103)	(0.001)	(0.0204)	(0.0102)	(0.001)
MR 2.3*	0.9088	0.3655	0.0619	0.8444	0.269	0.0141	0.8342	0.2577	0.0078
	(0.0223)	(0.011)	(0.0012)	(0.0207)	(0.0103)	(0.001)	(0.0204)	(0.0102)	(0.001)
MR 2.5*	0.9106	0.3702	0.0614	0.8449	0.2702	0.0141	0.8347	0.2583	0.0078
	(0.0224)	(0.0111)	(0.0012)	(0.0207)	(0.0103)	(0.001)	(0.0205)	(0.0102)	(0.001)
MR 3*	0.9147	0.3819	0.0601	0.8462	0.2731	0.0142	0.8351	0.2586	0.0078
	(0.0229)	(0.0113)	(0.0012)	(0.0208)	(0.0104)	(0.0011)	(0.0205)	(0.0102)	(0.001)

Tabelle 5.6.: Simulationsergebnisse für Kernschätzer: L_2-Risiko und gewählte Bandbreiten für HeaviSine.

StN		3			7			10	
n	50	100	1000	50	100	1000	50	100	1000
CV	0.043	0.0252	0.0059	0.0145	0.0088	0.0024	0.01	0.0059	0.0016
	(0.0563)	(0.0446)	(0.0215)	(0.0375)	(0.0284)	(0.0112)	(0.0303)	(0.0229)	(0.0079)
GCV	0.0444	0.0263	0.0059	0.0154	0.0092	0.0024	0.0101	0.0062	0.0016
	(0.0561)	(0.0446)	(0.0216)	(0.0319)	(0.0279)	(0.0112)	(0.0217)	(0.022)	(0.0079)
Cp	0.0428	0.0251	0.0059	0.0181	0.0091	0.0024	0.0159	0.0066	0.0016
	(0.0666)	(0.0477)	(0.0214)	(0.0598)	(0.0353)	(0.011)	(0.0595)	(0.0325)	(0.0078)
Chi .5	0.0822	0.0391	0.0086	0.0503	0.0167	0.0029	0.0503	0.0143	0.0019
	(0.1148)	(0.0796)	(0.033)	(0.1037)	(0.0609)	(0.0168)	(0.1052)	(0.0595)	(0.0121)
Chi .9	0.1196	0.0562	0.0117	0.066	0.0217	0.0035	0.0646	0.0178	0.0022
	(0.1401)	(0.1011)	(0.0486)	(0.1168)	(0.0709)	(0.0225)	(0.1168)	(0.067)	(0.0154)
Chi .95	0.1319	0.0621	0.013	0.0709	0.0233	0.0037	0.0691	0.0189	0.0023
	(0.147)	(0.1068)	(0.0529)	(0.1204)	(0.0737)	(0.0242)	(0.1201)	(0.069)	(0.0164)
Chi .99	0.1574	0.0742	0.0158	0.0809	0.0264	0.0042	0.078	0.021	0.0025
	(0.1599)	(0.1173)	(0.0606)	(0.1273)	(0.0786)	(0.0274)	(0.1263)	(0.0727)	(0.0184)
MR .5	0.077	0.0368	0.007	0.0409	0.0139	0.0026	0.0389	0.0107	0.0018
	(0.1123)	(0.0795)	(0.0272)	(0.0944)	(0.0539)	(0.0087)	(0.0941)	(0.0492)	(0.0054)
MR .9	0.1148	0.0532	0.0081	0.057	0.0193	0.0025	0.0529	0.0145	0.0017
	(0.1396)	(0.1016)	(0.0388)	(0.1099)	(0.0669)	(0.0118)	(0.1076)	(0.06)	(0.0069)
MR .95	0.1275	0.0593	0.0088	0.0624	0.0213	0.0025	0.058	0.016	0.0017
	(0.1469)	(0.1076)	(0.0429)	(0.1143)	(0.0708)	(0.0129)	(0.1118)	(0.0633)	(0.0074)
MR .99	0.156	0.0718	0.0107	0.0747	0.0255	0.0026	0.0688	0.019	0.0017
	(0.1614)	(0.1181)	(0.0505)	(0.1235)	(0.0778)	(0.0153)	(0.1201)	(0.0694)	(0.0086)
MR 2	0.0778	0.0362	0.0071	0.0412	0.0137	0.0026	0.0391	0.0105	0.0018
	(0.1129)	(0.0786)	(0.0236)	(0.0948)	(0.0533)	(0.0081)	(0.0943)	(0.0488)	(0.0051)
MR 2.3	0.0883	0.0416	0.0072	0.046	0.0155	0.0025	0.0433	0.0119	0.0017
	(0.1216)	(0.0876)	(0.0314)	(0.0999)	(0.0584)	(0.0097)	(0.0987)	(0.0529)	(0.0059)
MR 2.5	0.096	0.0453	0.0075	0.0493	0.0168	0.0025	0.046	0.0128	0.0017
	(0.1273)	(0.0926)	(0.0349)	(0.103)	(0.0615)	(0.0108)	(0.1014)	(0.0555)	(0.0064)
MR 3	0.1139	0.0545	0.0091	0.0566	0.0197	0.0025	0.0525	0.0148	0.0017
	(0.139)	(0.103)	(0.0443)	(0.1096)	(0.0678)	(0.0132)	(0.1072)	(0.0607)	(0.0076)
Cp*	0.0427	0.0239	0.0059	0.0137	0.0086	0.0024	0.0095	0.0058	0.0016
	(0.0581)	(0.046)	(0.0211)	(0.0368)	(0.0281)	(0.0108)	(0.0302)	(0.0225)	(0.0076)
Chi .5*	0.051	0.0295	0.0073	0.0168	0.0101	0.0026	0.0114	0.0067	0.0017
	(0.0833)	(0.0676)	(0.0294)	(0.0554)	(0.0416)	(0.0146)	(0.0454)	(0.0336)	(0.0104)
Chi .9*	0.0706	0.0428	0.0102	0.0206	0.0128	0.0031	0.0137	0.008	0.002
	(0.1083)	(0.0912)	(0.0469)	(0.0657)	(0.0518)	(0.0205)	(0.0531)	(0.0403)	(0.0136)
Chi .95*	0.0783	0.0479	0.0116	0.0221	0.0137	0.0033	0.0145	0.0084	0.0021
	(0.1153)	(0.0975)	(0.0517)	(0.0688)	(0.0547)	(0.0222)	(0.0555)	(0.0422)	(0.0146)
Chi .99*	0.0948	0.0586	0.0144	0.0253	0.0157	0.0037	0.0161	0.0093	0.0022
	(0.128)	(0.1085)	(0.0598)	(0.0746)	(0.0599)	(0.0256)	(0.0598)	(0.0457)	(0.0165)
MR .5*	0.0581	0.0327	0.0068	0.0176	0.0107	0.0026	0.0125	0.0068	0.0018
	(0.0943)	(0.0756)	(0.0268)	(0.0577)	(0.0435)	(0.0086)	(0.0488)	(0.033)	(0.0054)
MR .9*	0.0856	0.0477	0.008	0.0247	0.0142	0.0024	0.0161	0.0085	0.0017
	(0.1217)	(0.0978)	(0.0388)	(0.0733)	(0.0555)	(0.0115)	(0.0596)	(0.0419)	(0.0067)
MR .95*	0.0963	0.0532	0.0087	0.0272	0.0158	0.0025	0.0174	0.0092	0.0016
	(0.1296)	(0.1038)	(0.0429)	(0.0776)	(0.0598)	(0.0127)	(0.0626)	(0.0447)	(0.0072)
MR .99*	0.1191	0.0655	0.0106	0.0322	0.0189	0.0026	0.0204	0.011	0.0017
	(0.1437)	(0.1151)	(0.05)	(0.0851)	(0.0667)	(0.015)	(0.0687)	(0.0509)	(0.0084)
MR 2*	0.0586	0.0323	0.007	0.0177	0.0106	0.0026	0.0125	0.0067	0.0018
	(0.0949)	(0.0738)	(0.0235)	(0.0581)	(0.043)	(0.0079)	(0.049)	(0.0327)	(0.0051)
MR 2.3*	0.066	0.0373	0.007	0.0196	0.0116	0.0025	0.0137	0.0072	0.0017
	(0.1037)	(0.0839)	(0.0309)	(0.0631)	(0.0473)	(0.0096)	(0.0527)	(0.0357)	(0.0058)
MR 2.5*	0.0708	0.0406	0.0075	0.021	0.0124	0.0024	0.0144	0.0076	0.0017
	(0.1087)	(0.0889)	(0.0348)	(0.0662)	(0.0501)	(0.0105)	(0.0548)	(0.0377)	(0.0062)
MR 3*	0.0849	0.0489	0.009	0.0245	0.0145	0.0025	0.0161	0.0087	0.0016
	(0.1211)	(0.0992)	(0.0442)	(0.073)	(0.0564)	(0.013)	(0.0594)	(0.0426)	(0.0074)

Tabelle 5.7.: Simulationsergebnisse für Kernschätzer: L_2-Risiko und gewählte Bandbreiten für Doppler.

StN		3			7			10	
n	50	100	1000	50	100	1000	50	100	1000
CV	0.1999	0.1176	0.0241	0.133	0.0586	0.0074	0.1153	0.0521	0.0045
	(0.0342)	(0.0195)	(0.0045)	(0.0249)	(0.011)	(0.0024)	(0.021)	(0.0109)	(0.0021)
GCV	0.2105	0.1179	0.0241	0.181	0.0849	0.0074	0.1768	0.0817	0.0045
	(0.0404)	(0.0212)	(0.0045)	(0.0401)	(0.02)	(0.0024)	(0.04)	(0.02)	(0.0021)
Cp	0.209	0.1195	0.0242	0.1533	0.0704	0.0076	0.1417	0.0594	0.0046
	(0.0394)	(0.0217)	(0.0045)	(0.032)	(0.0157)	(0.0026)	(0.0298)	(0.0135)	(0.0022)
Chi .5	0.2919	0.1708	0.0305	0.2127	0.1	0.0117	0.1988	0.0886	0.0079
	(0.0637)	(0.036)	(0.0076)	(0.048)	(0.0234)	(0.0045)	(0.0454)	(0.0215)	(0.0037)
Chi .9	0.3556	0.2015	0.0343	0.2512	0.1144	0.0127	0.2322	0.0981	0.0085
	(0.0815)	(0.0439)	(0.0088)	(0.0576)	(0.0266)	(0.0048)	(0.0536)	(0.0236)	(0.0039)
Chi .95	0.3765	0.2112	0.0356	0.2631	0.1189	0.013	0.2431	0.1014	0.0087
	(0.0876)	(0.0464)	(0.0091)	(0.0606)	(0.0277)	(0.0049)	(0.0563)	(0.0243)	(0.004)
Chi .99	0.4195	0.2313	0.0381	0.2875	0.128	0.0136	0.265	0.1086	0.009
	(0.1005)	(0.0515)	(0.0098)	(0.0668)	(0.0297)	(0.0051)	(0.0617)	(0.0259)	(0.0041)
MR .5	0.2631	0.145	0.025	0.1547	0.069	0.0079	0.1412	0.059	0.0047
	(0.0555)	(0.0291)	(0.0054)	(0.0326)	(0.0153)	(0.0029)	(0.0298)	(0.0135)	(0.0023)
MR .9	0.3439	0.184	0.0273	0.2164	0.0885	0.0087	0.192	0.0725	0.0052
	(0.0783)	(0.0395)	(0.0065)	(0.0489)	(0.0205)	(0.0033)	(0.0436)	(0.0175)	(0.0026)
MR .95	0.3739	0.2004	0.0284	0.2434	0.0975	0.009	0.2168	0.0794	0.0054
	(0.0869)	(0.0437)	(0.007)	(0.0557)	(0.0227)	(0.0035)	(0.0498)	(0.0192)	(0.0027)
MR .99	0.4481	0.2322	0.0313	0.3072	0.121	0.0098	0.2742	0.0979	0.006
	(0.1096)	(0.0518)	(0.0079)	(0.0721)	(0.0281)	(0.0038)	(0.0641)	(0.0235)	(0.003)
MR 2	0.2644	0.1436	0.0248	0.1555	0.0685	0.0078	0.1419	0.0588	0.0047
	(0.0559)	(0.0287)	(0.0052)	(0.0328)	(0.0152)	(0.0028)	(0.03)	(0.0134)	(0.0023)
MR 2.3	0.289	0.1572	0.0256	0.1693	0.074	0.0082	0.1536	0.0616	0.0049
	(0.0629)	(0.0325)	(0.0057)	(0.0367)	(0.0168)	(0.003)	(0.0334)	(0.0144)	(0.0024)
MR 2.5	0.3057	0.1649	0.0264	0.182	0.0784	0.0084	0.1631	0.0643	0.005
	(0.0676)	(0.0345)	(0.0062)	(0.04)	(0.018)	(0.0032)	(0.036)	(0.0152)	(0.0025)
MR 3	0.342	0.1869	0.0287	0.2149	0.0904	0.0091	0.1904	0.074	0.0055
	(0.0778)	(0.0402)	(0.0071)	(0.0485)	(0.021)	(0.0035)	(0.0432)	(0.0178)	(0.0028)
Cp*	0.1754	0.1069	0.0241	0.119	0.0564	0.0074	0.1127	0.0499	0.0045
	(0.0222)	(0.0122)	(0.0042)	(0.0204)	(0.0104)	(0.0023)	(0.0202)	(0.0102)	(0.0018)
Chi .5*	0.1846	0.1141	0.0267	0.123	0.0595	0.0089	0.1152	0.0521	0.0054
	(0.0299)	(0.02)	(0.0063)	(0.022)	(0.0117)	(0.0034)	(0.0212)	(0.011)	(0.0027)
Chi .9*	0.1982	0.1222	0.0295	0.1244	0.0603	0.0095	0.1159	0.0526	0.0058
	(0.0359)	(0.0228)	(0.0074)	(0.0224)	(0.012)	(0.0037)	(0.0214)	(0.0111)	(0.0029)
Chi .95*	0.2029	0.1251	0.0304	0.1248	0.0605	0.0097	0.1161	0.0527	0.0059
	(0.0377)	(0.0237)	(0.0077)	(0.0226)	(0.0121)	(0.0038)	(0.0215)	(0.0112)	(0.003)
Chi .99*	0.2128	0.1316	0.0323	0.1256	0.061	0.01	0.1165	0.053	0.0061
	(0.0411)	(0.0255)	(0.0083)	(0.0229)	(0.0124)	(0.0039)	(0.0216)	(0.0112)	(0.0031)
MR .5*	0.1817	0.1094	0.0247	0.121	0.0576	0.0076	0.114	0.0509	0.0045
	(0.028)	(0.0173)	(0.0051)	(0.0213)	(0.0109)	(0.0025)	(0.0208)	(0.0106)	(0.0021)
MR .9*	0.1993	0.1231	0.0265	0.1223	0.0584	0.008	0.1147	0.0514	0.0046
	(0.0359)	(0.0228)	(0.0062)	(0.0217)	(0.0113)	(0.0029)	(0.021)	(0.0107)	(0.0022)
MR .95*	0.2071	0.129	0.0274	0.1228	0.0588	0.0082	0.115	0.0516	0.0047
	(0.0388)	(0.0246)	(0.0066)	(0.0219)	(0.0114)	(0.0031)	(0.0211)	(0.0108)	(0.0023)
MR .99*	0.2329	0.1466	0.0299	0.1239	0.0595	0.0088	0.1156	0.052	0.005
	(0.0469)	(0.0297)	(0.0075)	(0.0223)	(0.0117)	(0.0034)	(0.0213)	(0.0109)	(0.0025)
MR 2*	0.1819	0.1091	0.0245	0.1211	0.0576	0.0075	0.1141	0.0509	0.0045
	(0.0281)	(0.0171)	(0.0049)	(0.0213)	(0.0109)	(0.0024)	(0.0208)	(0.0106)	(0.002)
MR 2.3*	0.1858	0.1132	0.0252	0.1214	0.0579	0.0077	0.1142	0.0511	0.0045
	(0.0302)	(0.0192)	(0.0055)	(0.0214)	(0.011)	(0.0026)	(0.0209)	(0.0106)	(0.0021)
MR 2.5*	0.1888	0.1164	0.0258	0.1216	0.058	0.0078	0.1144	0.0512	0.0046
	(0.0316)	(0.0204)	(0.0058)	(0.0215)	(0.0111)	(0.0028)	(0.0209)	(0.0107)	(0.0022)
MR 3*	0.1988	0.1243	0.0277	0.1223	0.0585	0.0083	0.1147	0.0515	0.0047
	(0.0357)	(0.0231)	(0.0067)	(0.0217)	(0.0113)	(0.0031)	(0.021)	(0.0107)	(0.0023)

Tabelle 5.8.: Simulationsergebnisse für Kernschätzer: L_2-Risiko und gewählte Bandbreiten für Zero.

n	50	100	1000
CV	0.066	0.0226	0.0034
	(1.6178)	(2.1124)	(1.7344)
GCV	0.0763	0.0287	0.0034
	(1.6104)	(2.0792)	(1.7332)
Cp	0.0889	0.0399	0.0037
	(1.5269)	(1.9227)	(1.7323)
Chi .5	0.0701	0.055	0.0184
	(1.7196)	(1.4553)	(1.6353)
Chi .9	0.0346	0.0217	0.0065
	(2.4673)	(2.6038)	(2.3788)
Chi .95	0.0291	0.0192	0.0044
	(2.6435)	(2.6905)	(2.5689)
Chi .99	0.0236	0.0149	0.0024
	(2.8218)	(2.7557)	(2.7727)
MR .5	0.0645	0.0511	0.0133
	(1.7867)	(1.6956)	(1.8779)
MR .9	0.0262	0.0176	0.002
	(2.6236)	(2.661)	(2.7115)
MR .95	0.0238	0.0149	0.0015
	(2.7897)	(2.776)	(2.8422)
MR .99	0.0209	0.0114	0.0011
	(2.9521)	(2.9438)	(3)
MR 2	0.0636	0.053	0.0189
	(1.7901)	(1.6636)	(1.6493)
MR 2.3	0.0459	0.0321	0.0063
	(2.1523)	(2.0726)	(2.2303)
MR 2.5	0.0381	0.0249	0.0028
	(2.3142)	(2.2841)	(2.5465)
MR 3	0.0265	0.0165	0.0014
	(2.6178)	(2.6918)	(2.8931)
Cp*	0.058	0.0198	0.0035
	(1.7463)	(2.134)	(1.7571)
Chi .5*	0.0604	0.0432	0.0112
	(1.9193)	(1.7676)	(1.7122)
Chi .9*	0.0258	0.0139	0.0023
	(2.8132)	(2.666)	(2.7372)
Chi .95*	0.0227	0.0112	0.0016
	(2.8686)	(2.811)	(2.8543)
Chi .99*	0.0207	0.0104	0.0011
	(2.9764)	(3)	(2.9735)
MR .5*	0.0487	0.0375	0.0075
	(2.049)	(1.914)	(1.9215)
MR .9*	0.0223	0.0133	0.0017
	(2.8313)	(2.8493)	(2.8332)
MR .95*	0.0212	0.0122	0.0015
	(2.8997)	(2.9261)	(2.8888)
MR .99*	0.0206	0.0107	0.0011
	(3)	(2.9708)	(3)
MR 2*	0.0468	0.0392	0.0147
	(2.0937)	(1.8381)	(1.444)
MR 2.3*	0.0347	0.0246	0.0043
	(2.4864)	(2.3861)	(2.4273)
MR 2.5*	0.0279	0.0166	0.0024
	(2.6446)	(2.587)	(2.6756)
MR 3*	0.0224	0.013	0.0014
	(2.8173)	(2.8646)	(2.891)

6. Fazit und Ausblick

In dieser Arbeit wurde ein besonderes Verfahren zur Wahl von Glättungsparametern in der nichtparametischen Kurvenschätzung, das sogenannte Diskrepanzprinzip, untersucht. Dies geschah am Beispiel der zwei wichtigsten Probleme aus diesem Themenkreis, der nichtparametrischen Dichteschätzung und der nichtparametrischen Regression. Während das Diskrepanzprinzip in anderen Teilen der Mathematik eine sehr populäre Parameterwahlstrategie darstellt, ist es – trotz seiner Einfachheit – in der Statistik bislang kaum verwendet und nicht systematisch untersucht worden. Die bisheringen Verwendungen in der Statistik wurden in Kapitel 2 dargestellt, wo auch ein kurzer Einblick in die Verwendung des Diskrepanzprinzips zur Regularisierung inverser Probleme gegeben wurde.

Insbesondere im Fall der Kerndichteschätzung wurden in Kapitel 3 einige aus der Literatur bereits bekannte theoretische Resultate erweitert und durch neue ergänzt. Die bisherigen Vorschläge in der Literatur ließen sich so in einen einheitlichen Rahmen einordnen. Unter sehr schwachen Voraussetzungen konvergieren die mit Hilfe des Diskrepanzprinzips gewählten Bandbreiten stets gegen 0, so dass das Integral des Kerndichteschätzers konsistent ist für die Verteilungsfunktion. Allerdings kann die Bandbreite auch zu schnell gegen 0 konvergieren - so wurde z.B. gezeigt, dass Dichten existieren, für die auf dem Rechtecks- oder Epanechnikovkern basierende Dichteschätzer inkonsistent sein können, wenn die Bandbreite durch das Diskrepanzprinzip gewählt wird. Unter bestimmten Voraussetzungen an die gesuchte Dichte konnten außerdem präzisere Aussagen über die Konvergenzgeschwindigkeit der Bandbreite getroffen werden. Desweiteren wurden einige weitere neue Varianten des Diskrepanzprinzips untersucht, die asymptotisch die für eine bestimmte Referenzdichte (hier die Normalverteilung) optimale Bandbreite wählen. In einer umfangreichen Simulationsstudie wurden erstmalig die verschiedenen in der Literatur bisher existierenden Varianten des Diskrepanzprinzips sowohl untereinander als auch mit weit verbreiteten Standardmethoden verglichen. Hier zeigte sich, dass die besonders anschaulichen Varianten des Diskrepanzprinzips (z.B. diejenigen auf Basis des 95%- oder 99%-Quantils der Kolmogorov-Smirnov-Teststatistik) in den Simulationen fast durchgehend nicht gut abschneiden. Einige gut funktionierenden Versionen sind hingegen schwieriger zu interpretieren. Obwohl in den Simulationen Stichproben bis zu einer Größe von $n = 2500$ untersucht wurden, zeigten sich teilweise noch Effekte, die den asymtotischen Eigenschaften entgegengesetzt sind. Es ist also davon auszugehen, dass die asymptotischen Ergebnisse für praktisch relevante Stichprobengrößen nur bedingt aussagekräftig sind und sogar in die Irre führen können.

Im Fall der regulären Histogramme (Kapitel 4) ergaben sich ähnliche Resultate: Unter schwachen Voraussetzungen kann häufig die Konvergenz der Anzahl der Bins gegen unendlich mit einer gewissen Geschwindigkeit garantiert werden. Jedoch kann, ähnlich wie im Fall der Kerndichteschätzung, die Binanzahl für bestimmte Dichten so schnell wachsen, dass der resultierende Schätzer inkonsistent ist. Dichten, die selbst schon die Form eines regulären Histogramms haben, sind ein Spezialfall, der bei der Kerndichteschätzung nicht auftritt. Bestimmte Versionen des Diskrepanzprinzips sind modellwahlkonsistent, d.h. wenn die wahre Dichte ein reguläres Histogramm ist, wird asymptotisch fast sicher die minimale korrekte Anzahl Bins gewählt. Auch für die regulären Histogramme wurde eine umfangreiche Simulationsstudie durchgeführt, in der sich ähnliche Ergebnisse wie bei der Kerndichteschätzung zeigten.

In Kapitel 5 wurden einige Aspekte des Diskrepanzprinzips in der Regression am Beispiel von Nadaraya-Watson-Kernschätzern und kubischen Glättungssplines untersucht. Speziell für Glättungssplines wurden hier in der Literatur bereits früh auf der Residuenquadratsumme basierende Versionen des Diskrepanzprinzips vorgeschlagen. Das im Rahmen des „Data Approximation“-Ansatzes häufig benutzte Multiresolutionskriterium stellt ein auf einer besonderen Norm des Residualvektors basierendes Diskrepanzprinzip dar. Die geometrischen Eigenschaften dieser Norm wurden im weiteren Verlauf des Kapitels detailliert dargestellt. Insbesondere das Fehlen bestimmter Invarianzeigenschaften führt dazu, dass Muster in den Residuen entdeckt werden könnnen. Diese Ergebnisse sind auch für andere Verwendungen des Multiresolutionskriteriums interessant, z.B. für Tests. Beim abschließenden Vergleich der auf der Residuenquadratsumme bzw. dem Multiresolutionskriterium basierenden Diskrepanzprinzipien in einer Simulationsstudie ergab sich, dass diese gegenüber Standardmethoden der Glättungsparameterwahl nicht besonders gut abschneiden, zumindest wenn die Rauschvarianz nicht bekannt ist. Ist sie bekannt, funktionieren die Methoden etwas besser, können aber trotzdem nicht uneingeschränkt empfohlen werden. Die auf dem Multiresolutionskriterium basierenden Methoden sind jedoch klar den entsprechenden auf der Residuenquadratsumme basierenden Methoden überlegen.

Im Rahmen dieser Arbeit wurden nur die einfachsten nichtparametrischen Kurvenschätzer im univariaten Fall betrachtet. Hier ergeben sich zahlreiche Anknüpfungspunkte für zukünftige Verallgemeinerungen und Erweiterungen. Neben der Übertragung auf weitere Klassen von Schätzern, auf den Fall lokaliserter Glättungsparameter und auf multivariate Situationen sind auch für die hier betrachteten Schätzer noch einige weitergehende Fragen offen. So lassen sich z.B. auf dem Cramér-von Mises-Test basierende Versionen des Diskrepanzprinzips für die Dichteschätzung nicht ohne weiteres mit Hilfe der Methoden aus Kapitel 3 untersuchen, da der zugehörige Abstand keine Metrik ist. Aber auch für die hier näher untersuchten Versionen des Diskrepanzprinzips für Kerndichteschätzer besteht weiterer Forschungsbedarf. Während sich im Fall langsam fallender Schrankenfunktionen der Abstand zwischen der empirischen und der geschätzten Verteilungsfunktion im Wesentlichen wie der Abstand zwischen der theoretischen Verteilungsfunktion und ihrer geglätteten Version verhält, dominieren bei schnell fallender Schrankenfunktion die stochastischen Anteile. Hier könnten eventuell mit Hilfe speziellerer Methoden der Asymptotik präzisere Resultate erzielt werden. Im Fall der nichtparametrischen Regression sind ebenfalls noch viele Fragen, die das asymptotische Verhalten der gewählten Glättungsparameter bzw. der resultierenden Schätzer betreffen, offen.

Die Grundidee des Diskrepanzprinzips ist sehr einfach und lässt sich einem Anwender ohne Bezug auf Verlustfunktionen oder Asymptotik erklären. Auch sind die Verfahren in der Praxis z.B. im Vergleich zur Kreuzvaliderung oft schneller und in vielen Fällen einfach zu implementieren. Im Fall der Kerndichteschätzung genügt es, eine nichtlineare Gleichung zu lösen, was mit Hilfe der Regula Falsi schnell und effektiv geschehen kann. In den anderen betrachteten Fällen (Reguläre Histogramme, Nadaraya-Watson-Kernschätzer und Glättungssplines) ist das Diskrepanzprinzip ein Abbruchkriterium, d.h. es müssen nur solange in Frage kommende Werte des Glättungsparameters ausprobiert werden, bis das Kriterium erfüllt ist. Die meisten anderen Methoden zur Glättungsparemeterwahl erfodern die Minimierung einer Zielfunktion, die im Prinzip für alle in Frage kommenden Werte des Glätttungsparameters berechnet werden muss. Insgesamt ergibt sich, dass das Diskrepanzprinzip eine interessante – und bislang im statistischen Mainstream zu unrecht vernachlässigte – Alternative zu anderen Verfahren der Glättungsparameterwahl darstellt.

A. Mathematische Hilfsmittel

In diesem Anhang sind einige bekannte in dieser Arbeit benötigte Definitionen, Lemmata und Sätze zusammengestellt. Es handelt sich um Material, das sich in zahlreichen Lehrbüchern findet. Zusätzlich werden einige in der nichtparametrischen Kurvenschätzung häufig benutzte und oft nicht explizit ausgeführte Argumente und Techniken hier als Satz oder Lemma formuliert. Verwiesen wird jeweils auf die für den Kontext dieser Arbeit am besten geeignete Quelle, und es wurde kein Versuch unternommen, die beschriebenen Resultate bis zu ihrem historischen Ursprung zurückzuverfolgen. In einigen Fällen wurden Resultate gegenüber der zitierten Quelle kombiniert, verallgemeinert, spezialisiert oder auf andere Art leicht abgewandelt. Es wird dabei kein Anspruch auf Originalität, Allgemeinheit oder Vollständigkeit erhoben.

In den Sätzen A.1 und A.5 wird die Fouriertransformation einer Funktion bzw. eines Maßes wie in der Analysis üblich mit $\hat{f}$ bzw. $\hat{\mu}$ bezeichnet. Da Fouriertransformationen in dieser Arbeit nur im Beweis zum 2. Teil von Satz 3.3 verwendet werden, ist die Gefahr einer Verwechslung mit Schätzern gering. Daher wird darauf verzichtet, eine zusätzliche Notation für die Fouriertransformation einzuführen.

A.1. Einige Hilfsmittel aus der Analysis

Lemma A.1 (Vierecksungleichung)**.** *Sei E ein metrischer Raum mit Metrik d. Dann gilt für alle $w, x, y, z \in E$:*

$$|d(x,y) - d(w,z)| \leq d(x,z) + d(w,y).$$

Lemma A.2. *Sei $f \in L_1(\mathbb{R})$ und h,$h' > 0$. Dann gilt:*

$$\lim_{h' \to h} \left\| \frac{1}{h} f\left(\frac{\cdot}{h}\right) - \frac{1}{h'} f\left(\frac{\cdot}{h'}\right) \right\|_1 = 0.$$

Beweis. (vgl. Beweis zu Lemma 5 in [Dev83]). Wegen $f \in L_1(\mathbb{R})$ existiert ein $g \in C_c(\mathbb{R})$ mit $\int |g(x) - f(x)| dx \leq \frac{\varepsilon}{2}$. Dabei bezeichnet $C_c(\mathbb{R})$ den Raum der stetigen Funktionen auf $\mathbb{R}$ mit kompaktem Träger. Wähle nun m so, dass $g(x) = 0$ für $|x| > m$ und $M := \|g\|_\infty$. Dann gilt (mit Hilfe der Substitutionsregel):

$$\int \left| \frac{1}{h} f\left(\frac{x}{h}\right) - \frac{1}{h} g\left(\frac{x}{h}\right) \right| dx = \int \left| \frac{1}{h'} f\left(\frac{x}{h'}\right) - \frac{1}{h'} g\left(\frac{x}{h'}\right) \right| dx = \int |g(x) - f(x)| dx \leq \frac{\varepsilon}{2}.$$

Weiter ist

$$\begin{aligned}
&\int \left|\frac{1}{h} f\left(\frac{x}{h}\right) - \frac{1}{h'} f\left(\frac{x}{h'}\right)\right| dx \\
&= \int \left|\frac{1}{h} f\left(\frac{x}{h}\right) - \frac{1}{h} g\left(\frac{x}{h}\right) + \frac{1}{h} g\left(\frac{x}{h}\right) - \frac{1}{h'} g\left(\frac{x}{h'}\right) + \frac{1}{h'} g\left(\frac{x}{h'}\right) - \frac{1}{h'} f\left(\frac{x}{h'}\right)\right| dx \\
&\leq \int \left|\frac{1}{h} g\left(\frac{x}{h}\right) - \frac{1}{h'} g\left(\frac{x}{h'}\right)\right| dx + \varepsilon.
\end{aligned}$$

Wegen der Stetigkeit von g gilt $\lim_{h' \to h} \frac{1}{h} g\left(\frac{x}{h}\right) = \frac{1}{h'} g\left(\frac{x}{h'}\right)$ für alle $x \in \mathbb{R}$. Für jede Folge $(h_n)_{n \in \mathbb{N}} \subseteq [\min(h, h'), \max(h, h')]$ ist nun die Funktion

$$\frac{2M}{\min(h, h')} \mathbb{I}(-m \max(h, h') \leq x \leq m \max(h, h'))$$

eine integrierbare Majorante von

$$\left|\frac{1}{h} f\left(\frac{x}{h}\right) - \frac{1}{h_n} f\left(\frac{x}{h_n}\right)\right|.$$

Mit dem Satz von der majorisierten Konvergenz folgt damit die Behauptung. □

Satz A.1. *Es seien f und g in $L_1(\mathbb{R})$ mit $g(x) = -ix^k f(x)$. Dann ist die Fouriertransformierte $\hat{f}$ k mal stetig differenzierbar mit*

$$\hat{f}^{(k)}(t) = \hat{g}(t) = (-i)^{(k)} \widehat{x^k f}(t).$$

Insbesondere ist für $\int x^k f(x) dx \in \mathbb{R} \setminus \{0\}$ dann $\hat{f}^{(k)}(0) \neq 0$ und damit $\hat{f}$ auf keinem Intervall um 0 konstant.

Beweis. Folgt per Induktion aus Satz 9.2 f) in [Rud74]. □

Satz A.2. *Sei $g \in C^{(2)}([a, b])$. Für Stützstellen $a = \tau_1 < \cdots < \tau_n = b$ sei $\bar{g}$ die Funktion, die durch lineare Interpolation von g in den Punkten $\tau_1, \ldots, \tau_n$ entsteht. Dann ist für $\tau_i \leq x \leq \tau_{i+1}$*

$$g(x) - \bar{g}(x) = (x - \tau_i)(x - \tau_{i+1}) \frac{g^{(2)}(\xi)}{2}$$

für ein $\xi \in [\tau_i, \tau_{i+1}]$.

Beweis. Vgl. [dB01], S.31 und S. 6. □

A.2. Einige Hilfsmittel aus der Maß- und Wahrscheinlichkeitstheorie

Satz A.3. *Sei F eine stetige Verteilungsfunktion, F_n die empirische Verteilungsfunktion einer unabhängig identisch verteilten Stichprobe $X_1, \dots, X_n$ vom Umfang n aus der zu F gehörigen Verteilung sowie G eine weitere Verteilungsfunktion, die nicht notwendig gleich F ist. Dann gilt:*

1. $d_\infty(F_n, G) = \max\left(\max_{1\le i\le n}\left(\frac{i}{n} - G(X_i)\right), \max_{1\le i\le n}\left(G(X_i) - \frac{i-1}{n}\right)\right)$.
2. $d_{kuip,1}(F_n, G) = \max_{1\le i\le n}\left(\frac{i}{n} - G(X_i)\right) + \max_{1\le i\le n}\left(G(X_i) - \frac{i-1}{n}\right)$.
3. $d_{CvM}(F_n, G) = \frac{1}{n}\sum_{i=1}^n (G(X_i) - \frac{i-1/2}{n})^2 + \frac{12}{n^2}$.
4. $F_{KS}(x) := \lim_{n\to\infty} P(\sqrt{n} d_\infty(F_n, F) \le x) = \sum_{i\in\mathbb{Z}}(-1)^i \exp(-2i^2x^2)$.
5. $F_{kuip,1}(x) := \lim_{n\to\infty} P(\sqrt{n} d_{kuip,k}(F_n, F) \le x) = 1 - \sum_{i=1}^\infty (4i^2x^2 - 1)\exp(-2i^2x^2)$.
6. $$\begin{aligned} F_{CvM}(x) :&= \lim_{n\to\infty} P(n d_{CvM}(F_n, F) \le x) \\ &= 1 - \frac{1}{\pi}\sum_{i=1}^{\infty}(-1)^{i-1}\int_{(2i-1)^2\pi^2}^{(2i)^2\pi^2} t^{-1}\left(\frac{-\sqrt{t}}{\sin\sqrt{t}}\right)^{1/2}\exp\left(-\frac{xt}{2}\right)dt. \end{aligned}$$

Beweis.
1. vgl. [Dur73, S. 6].
2. vgl. [Dur73, S. 33f.].
3. vgl. Formel 4.1.7 in [Dur73, S. 27].
4. vgl. Formel 3.4.8 in [Dur73, S. 22].
5. vgl. Formel 5.3.3 in [Dur73, S. 36].
6. vgl. Formel 4.4.5 in [Dur73, S. 32].

□

Satz A.4. *Sei F eine stetige Verteilungsfunktion und F_n die auf einer Stichprobe vom Umfang n basierende empirische Verteilungsfunktion. Dann gilt fast sicher:*

1. $$\limsup_{n\to\infty}\sqrt{\frac{2n}{\log\log n}} d_\infty(F_n, F) = 1,$$
2. $$\limsup_{n\to\infty}\frac{n}{2\log\log n} d_{CvM}(F_n, F) = \frac{1}{\pi^2}.$$

Beweis. vgl. [Ser80], S.62 bzw. 64. □

Definition A.1. 1. Sind ν_1, ν_2 endliche Maße auf $(\mathbb{R}, \mathbb{B})$, so heißt die σ-additive Mengenfunktion $\mu := \nu_1 - \nu_2$ ein *signiertes Maß* auf $(\mathbb{R}, \mathbb{B})$.

2. Für ein signiertes Maß μ werden die (eindeutig definierten) positiven Maße μ^+ und μ^- mit $\mu = \mu^+ - \mu^-$ und $\mu^+(\emptyset) = \mu^-(\emptyset) = 0$ als *Positiv-* bzw. *Negativteil* von μ bezeichnet.

3. Das positive Maß $|\mu| := \mu^+ + \mu^-$ heißt das zu μ gehörige *Totalvariationsmaß.*

4. Eine Folge von signierten Maßen $\mu_1, \mu_2, \ldots$ heißt *schwach konvergent* gegen μ genau dann, wenn für alle stetigen beschränkten Funktionen f auf $\mathbb{R}$ gilt:

$$\int_{\mathbb{R}} f d\mu_n \longrightarrow \int_{\mathbb{R}} f d\mu.$$

Abkürzend wird die schwache Konvergenz der μ_n gegen μ auch mit

$$\mu_n \rightsquigarrow \mu$$

bezeichnet. Sind alle Maße Wahrscheinlichkeitsmaße, so entspricht die schwache Konvergenz der aus der Wahrscheinlichkeitsrechnung bekannten Verteilungskonvergenz.

Satz A.5. *Seien $\mu, \mu_1, \mu_2, \ldots$ endliche signierte Maße auf $(\mathbb{R}, \mathbb{B})$, die außerdem die folgende* Straffheitsbedingung *erfüllen: Für jedes $\varepsilon > 0$ existiert ein kompaktes Intervall J, so dass für alle $n \in \mathbb{N}$*

$$|\mu_n|(\mathbb{R} \setminus J) < \varepsilon. \tag{A.1}$$

Außerdem existiere $k \in \mathbb{R}$ mit $|\mu_n|(\mathbb{R}) \leq k$ für alle $n \in \mathbb{N}$ und $|\mu|(\mathbb{R}) \leq k$. Seien weiter $F, F_1, F_2, \ldots$ die zu $\mu, \mu_1, \mu_2, \ldots$ gehörigen Verteilungsfunktionen sowie $\hat{\mu}, \hat{\mu}_1, \hat{\mu}_2, \ldots$ die Fouriertransformierten der Maße. Dann gelten die folgenden Implikationen:

1. $d_\infty(F_n, F) \longrightarrow 0 \Longrightarrow \mu_n \rightsquigarrow \mu$

2. $\mu_n \rightsquigarrow \mu \Longrightarrow \hat{\mu}_n(t) \longrightarrow \hat{\mu}(t)$ *gleichmäßig auf kompakten Teilmengen von $\mathbb{R}$.*

Beweis. 1. Siehe Proposition 8.1.8 in [Bog07].

2. Siehe [Kat04], S.173f.

□

Satz A.6 (Devroye)**.** *Sei K Kern der Ordnung ℓ, $X_1, X_2, \ldots$ eine Folge von unabhängigen Zufallsvariablen mit Dichte f, h_n eine Folge von Bandbreiten und $\hat{f}_n := \frac{1}{n}\sum_{i=1}^n K_{h_n}(x - X_i)$. Die Bandbreiten können dabei deterministisch oder zufällig sein. Dann sind äquivalent:*

1. $d_1(\hat{f}_n, f) \longrightarrow 0$ *fast sicher*

2. $h_n \longrightarrow 0$ *fast sicher und* $nh_n \longrightarrow \infty$ *fast sicher.*

Beweis. vgl. Theorem 1.53 in Kapitel 4 von [EL01] bzw. [Dev83]. □

B. Details zu den Simulationen

B.1. Die Beispieldichten

Die folgende Aufstellung der von Berlinet und Devroye vorgeschlagenen und in der vorliegenden Arbeit verwendeten Beispieldichten wurde aus [BD94] entnommen und um einige weitere Eigenschaften ergänzt:

1. **Uniform:** Die Dichte einer Gleichverteilung auf $[0, 1]$.
2. **Exponential:** Die Dichte einer Exp(1)-Verteilung.
3. **Maxwell:** $f(x) = x \exp(-\frac{x^2}{2})$ auf $[0, \infty)$.
4. **Double Exponential:** $f(x) = \frac{1}{2} \exp(-|x|)$.
5. **Logistic:** $f(x) = \frac{\exp(-x)}{(1+\exp(-x))^2}$.
6. **Cauchy:** Die Dichte einer Cauchy(0,1)-Verteilung.
7. **Extreme value:** Die Dichte einer Extremwertverteilung mit Verteilungsfunktion $F(x) = \exp(-\exp(-x))$.
8. **Infinite peak:** $f(x) = (2\sqrt{x})^{-1}$ auf $(0, 1)$. Die Dichte liegt weder in L_2 noch in L_∞.
9. **Pareto:** Pareto-Verteilung mit Parameter 3/2: $f(x) = (2x^{3/2})^{-1}$ auf $[1, \infty)$.
10. **Symmetric Pareto:** $f(x) = (4(1 + |x|)^{3/2})^{-1}$.
11. **Normal:** Die Dichte einer N(0,1)-Verteilung.
12. **Lognormal:** $f(x) = (x\sqrt{2\pi})^{-1} \exp(-(\log x)^2/2)$ auf $[0, \infty)$.
13. **Uniform scale mixture:** Dichte einer Mischungsverteilung $\frac{1}{2}U[-\frac{1}{2}, \frac{1}{2}] + \frac{1}{2}U[-5, 5]$.
14. **Matterhorn:** Dichte von $S \exp(-2/U)$ mit $P(S = -1) = P(S = 1) = \frac{1}{2}$ und U gleichverteilt auf $[0, 1]$. Die Dichte ist gegeben durch $(|x|(\log(|x|))^2)^{-1}$ auf $[-e^{-2}, e^{-2}]$ und liegt weder in L_2 noch in L_∞. Durch die begrenzte Rechengenauigkeit können bei Simulationen in größeren Stichproben mehrere Realisationen den Wert 0 annehmen.

15. **Logarithmic peak:** Die Dichte von UV, wobei U und V unabhängig $U[0,1]$-verteilt sind. Die Dichte ist gegeben durch $f(x) = -\log(x)$ auf $(0,1)$ und liegt in L_2, aber nicht in L_∞.

16. **Isosceles triangle:** Dichte einer Dreiecksverteilung $f(x) = (1 - |x|)_+$.

17. **Beta (2,2):** Die Dichte einer Beta(2,2)-Verteilung.

18. **Chi-square (1):** Die χ^2-Verteilung mit 1 Freiheitsgrad. Die Dichte ist gegeben durch $(\sqrt{2\pi x})^{-1}\exp(-\frac{x}{2})$ für $x > 0$ und liegt weder in L_2 noch in L_∞.

19. **Normal cubed:** Die Dichte von N^3, wobei N standardnormalverteilt ist. Die Dichte ist gegeben durch $f(x) = \frac{\sqrt{2}}{6\sqrt{\pi}} x^{-2/3} \exp\left(-\frac{1}{2}x^{2/3}\right)$ und liegt weder in L_2 noch in L_∞.

20. **Inverse exponential:** Verteilung von E^{-2}, wobei E $Exp(1)$-verteilt ist. Die Dichte ist gegeben durch $f(x) = \frac{1}{2}x^{-3/2}\exp(-\frac{1}{\sqrt{x}})$ auf $[0,\infty)$.

21. **Marronite:** $f(x) = \frac{1}{3}\phi(\frac{x+20}{1/4}) + \frac{2}{3}\phi(x)$.

22. **Skewed bimodal:** $f(x) = \frac{3}{4}\phi(x) + \frac{1}{4}\phi(\frac{x-1.5}{1/3})$.

23. **Claw:** $f(x) = \frac{1}{2}\phi(x) + \frac{1}{10}\phi(\frac{x+1}{0.1}) + \frac{1}{10}\phi(\frac{x+0.5}{0.1}) + \frac{1}{10}\phi(\frac{x}{0.1}) + \frac{1}{10}\phi(\frac{x-0.5}{0.1}) + \frac{1}{10}\phi(\frac{x-1}{0.1})$.

24. **Smooth comb:** $f(x) = \frac{32}{63}\phi(\frac{x+31/21}{32/63}) + \frac{16}{63}\phi(\frac{x-17/21}{16/63}) + \frac{8}{63}\phi(\frac{x-41/21}{8/63}) + \frac{4}{63}\phi(\frac{x-53/21}{4/63}) + \frac{2}{63}\phi(\frac{x-59/21}{2/63}) + \frac{1}{63}\phi(\frac{x-62/21}{1/63})$.

25. **Caliper:** Die Dichte von $S(X + 0.1)$, wobei $P(S = -1) = P(S = 1) = \frac{1}{2}$ und X die Dichte $f(x) = 4(1 - x^{1/3})$ auf $[0,1]$ besitzt.

26. **Trimodal uniform:** Die Dichte einer Mischung aus drei Rechteckverteilungen mit disjunkten Trägern $\frac{1}{2}U[-1,1] + \frac{1}{4}U[-20.1,-20] + \frac{1}{4}U[20,20.1]$.

27. **Sawtooth:** Die Dichte von $N + X$, wobei N diskret gleichverteilt ist auf der Menge $\{-9,-7,-5,-3,-1,1,3,5,7,9\}$ und X gleichschenklig dreiecksverteilt auf $[-1,1]$ (wie in Nr. 16).

28. **Bilogarithmic peak:** Die Dichte ist gegeben durch $f(x) = -\frac{1}{2}\log(x(1-x))$ auf $[0,1]$ und liegt in L_2, aber nicht in L_∞.

Die Dichten sind vollständig im `R`-Zusatzpaket `benchden` [MWT09] implementiert.

B.2. Die Beispielfunktionen zur nichtparametrischen Regression

In der Simulationsstudie zur nichtparametrischen Regression werden die folgenden Beispielfunktionen verwendet:

1. **Ruppert & Carroll:** Entspricht Formel (10) in [RC00] für $j = 6$. Die Funktion ist gegeben durch
$$f_1(x) = \sqrt{x(1-x)} \sin\left(\frac{2\pi(1+0.125)}{x+0.125}\right)$$
und wird auch in [DM08] verwendet. Die Funktion ist der Doppler-Funktion sehr ähnlich, oszilliert aber etwas weniger.

2. **SinePeak:** Die Beispielfunktion aus dem R-Paket `mrsmooth` [Mei06]:
$$f_2(x) = \sin(2\pi x) + 2\exp(-1000(x-0.5)^2).$$

3. **Piecewise Polynomial:** Die Beispielfunktion aus [NS94]. Definiert als
$$f_3(x) = \begin{cases} 4x^2(3-4x) & (0 \leq x \leq 1/2) \\ \frac{4}{3}x(4x^2-10x+7) - \frac{3}{2} & (1/2 < x \leq 3/4) \\ \frac{16}{3}x(x-1)^2 & (3/4 < x \leq 1) \end{cases}.$$

4. **Blocks:** Eine der Testfunktionen aus [DJ94]. Definiert als
$$f_4(x) = \sum_{j=1}^{11} h_j \frac{1+\operatorname{sign}(x-t_j)}{2}$$
mit
$$\begin{aligned}(t_j) &= (0.1, 0.13, 0.15, 0.23, 0.25, 0.40, 0.44, 0.65, 0.76, 0.78, 0.81)\\ (h_j) &= (4, -5, 3, -4, 5, -4.2, 2.1, 4.3, -3.1, 2.1, -4.2).\end{aligned}$$

5. **Bumps:** Eine der Testfunktionen aus [DJ94]. Definiert als
$$f_5(x) = \sum_{j=1}^{11} \frac{h_j}{(1+|\frac{x-t_j}{w_j}|)^4},$$
wobei
$$\begin{aligned}(t_j) &= (0.1, 0.13, 0.15, 0.23, 0.25, 0.40, 0.44, 0.65, 0.76, 0.78, 0.81)\\ (h_j) &= (4, 5, 3, 4, 5, 4.2, 2.1, 4.3, 3.1, 5.1, 4.2)\\ (w_j) &= (0.005, 0.005, 0.006, 0.01, 0.01, 0.03, 0.01, 0.01, 0.005, 0.008, 0.005).\end{aligned}$$

6. **Heavisine:** Eine der Testfunktionen aus [DJ94]. Definiert als

$$f_6(x) = 4\sin 4\pi x - \operatorname{sign}(x - 0.3) - \operatorname{sign}(0.7 - x).$$

7. **Doppler:** Eine der Testfunktionen aus [DJ94]. Definiert als

$$f_7(x) = \sqrt{x(1-x)}\sin\left(\frac{2\pi(1+0.05)}{x+0.05}\right).$$

8. **Zero:** Die Nullfunktion $f_8(x) = 0$. Diese Funktion wird in [DGW08] verwendet.

B.3. Eigenschaften der verwendeten Kerne

Tabelle B.1.: Die verwendeten Kerne und einige ihrer Eigenschaften.

	Gaußkern	Epanechnikov-Kern
$K(x)$	$\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$	$\frac{3}{4}(1-x^2)\,(\lvert x\rvert \leq 1)$
$\mathbb{K}(x)$	$\Phi(x)$	$-\frac{1}{4}\,x^3 + \frac{3}{4}\,x + \frac{1}{2}\,(\lvert x\rvert \leq 1)$
k_2	1	$\frac{1}{5}$
$\lVert K\rVert_2^2$	$\frac{1}{2\sqrt{\pi}}$	$\frac{3}{5}$
$K * K(x)$	$\frac{1}{2\sqrt{\pi}}e^{-\frac{x^2}{4}}$	$-\frac{3}{160}\lvert x\rvert^5 + \frac{3}{8}\lvert x\rvert^3 - \frac{3}{4}x^2 + \frac{3}{5}\,(\lvert x\rvert \leq 2)$
$K_h'' * K_h''(x)$	$\frac{12h^4-12h^2x^2+x^4}{32\sqrt{\pi}h^9}e^{-\frac{x^2}{4h^2}}$	$-\frac{9x}{4h^6}\lvert x\rvert + \frac{9}{2h^5}\,(\lvert x\rvert \leq 2h)$
$K_h^{(4)}(x)$	$\frac{3h^4-6h^2x^2+x^4}{\sqrt{2\pi}h^9}e^{-\frac{x^2}{2h^2}}$	$0\ (x \notin \{-h, h\})$
$K_h^{(6)}(x)$	$-\frac{15h^6-45h^4x^2+15h^2x^4-x^6}{\sqrt{2\pi}h^{13}}e^{-\frac{x^2}{2h^2}}$	$0\ (x \notin \{-h, h\})$
$K_h^{(8)}(x)$	$\frac{105\,h^8-420\,h^6x^2+210\,h^4x^4-28\,h^2x^6+x^8}{\sqrt{2\pi}h^{17}}e^{-\frac{x^2}{2h^2}}$	$0\ (x \notin \{-h, h\})$
$K_h^{(10)}(x)$	$-\frac{945h^{10}-4725h^8x^2+3150h^6x^4-630h^4x^6+45h^2x^8-x^{10}}{\sqrt{2\pi}h^{21}}e^{-\frac{x^2}{2h^2}}$	$0\ (x \notin \{-h, h\})$

B.4. Weitere Simulationsergebnisse

Tabelle B.2.: Simulationsergebnisse für Dichteschätzung mit Gaußkern – L_∞-Risiko für die Dichten.

Dichte	n	L2CV	BCV	PI	SCV	E-LR	V	KS .5	KS .9	KS .95	KS .99
3	50	0.1722	**0.1431**	0.1465	0.2492	0.159	0.1461	0.1699	0.2382	0.2658	0.3174
	100	0.1536	**0.1215**	0.1257	0.2164	0.1311	0.1288	0.141	0.1839	0.2003	0.2337
	500	0.0869	**0.0804**	0.0872	0.1531	0.0845	0.0815	0.0983	0.1257	0.1342	0.1481
	1000	0.0706	**0.0666**	0.0754	0.1265	0.0747	0.0679	0.0839	0.1072	0.1144	0.1265
	2500	0.0542	**0.0526**	0.062	0.108	0.0634	0.0537	0.067	0.0855	0.0913	0.1012
4	50	0.1577	0.1648	0.1652	0.1632	**0.154**	0.1639	0.2176	0.2967	0.3188	0.3551
	100	**0.142**	0.1531	0.1539	0.1427	0.1423	0.1492	0.1886	0.2518	0.271	0.3028
	500	**0.0883**	0.1012	0.1081	0.0997	0.1069	0.0997	0.1298	0.1715	0.1846	0.2066
	1000	**0.0696**	0.0792	0.0889	0.0889	0.0921	0.0798	0.1071	0.1437	0.155	0.1741
	2500	**0.0542**	0.0613	0.0724	0.0747	0.0791	0.0639	0.0846	0.1141	0.1231	0.1383
5	50	0.064	**0.0524**	0.0545	0.0975	0.0628	0.0557	0.0653	0.1067	0.1206	0.1442
	100	0.0537	**0.0408**	0.0432	0.084	0.048	0.0463	0.0531	0.0815	0.0922	0.1111
	500	0.0277	**0.0229**	0.0237	0.0579	0.024	0.0242	0.027	0.0418	0.0477	0.0583
	1000	0.0211	**0.0186**	0.0189	0.0473	0.0192	0.0191	0.0211	0.032	0.0363	0.0443
	2500	0.0158	**0.014**	0.0144	0.0402	0.0149	0.0142	0.0156	0.0223	0.0252	0.0305
6	50	0.0981	0.0936	**0.0911**	0.1126	0.0919	0.0937	0.1249	0.1834	0.2002	0.2267
	100	0.0758	0.0794	0.0778	0.0975	**0.0734**	0.0757	0.1032	0.1513	0.1664	0.1907
	500	0.0461	**0.0441**	0.0487	0.07	0.0487	0.046	0.061	0.0904	0.1002	0.1167
	1000	0.0357	**0.0338**	0.0383	0.0603	0.0403	0.0361	0.0476	0.0705	0.0782	0.0916
	2500	0.0248	**0.0238**	0.0273	0.0501	0.03	0.025	0.0323	0.048	0.0536	0.0635
7	50	0.0978	**0.0807**	0.0826	0.1461	0.0932	0.0839	0.0998	0.1553	0.174	0.2057
	100	0.0815	**0.0631**	0.0658	0.1266	0.0716	0.0695	0.0774	0.116	0.1302	0.1555
	500	0.0442	**0.0363**	0.0378	0.0895	0.0383	0.0387	0.0419	0.0597	0.067	0.0807
	1000	0.0333	**0.0292**	0.0299	0.0738	0.0299	0.0298	0.0322	0.0444	0.0494	0.059
	2500	0.0252	**0.0224**	0.0227	0.0624	0.0229	0.0225	0.0235	0.0304	0.0335	0.0397
10	50	0.1289	0.1691	0.1562	**0.1196**	0.1316	0.1432	0.1729	0.2078	0.2164	0.2278
	100	0.1107	0.1635	0.1472	**0.1028**	0.1212	0.1277	0.1567	0.19	0.1982	0.2103
	500	0.0702	0.1073	0.1165	**0.0589**	0.0991	0.0934	0.1156	0.1424	0.1501	0.1624
	1000	0.0593	0.0823	0.1031	**0.044**	0.0903	0.0815	0.1005	0.124	0.1308	0.1421
	2500	0.0467	0.0617	0.0869	**0.0332**	0.0784	0.0668	0.0824	0.1023	0.1082	0.1182
11	50	0.1062	**0.0814**	0.0846	0.1562	0.1007	0.0888	0.0993	0.159	0.1805	0.2173
	100	0.0769	**0.0613**	0.0645	0.1281	0.0718	0.0691	0.0763	0.1202	0.1374	0.1672
	500	0.0446	**0.0358**	0.0364	0.092	0.0374	0.0383	0.0395	0.0588	0.0674	0.0834
	1000	0.0343	**0.0292**	0.0295	0.0786	0.0298	0.0303	0.0316	0.0449	0.0507	0.062
	2500	0.0264	**0.0225**	0.0228	0.0625	0.0231	0.0229	0.0238	0.0317	0.0353	0.0423
12	50	0.2835	0.3099	0.2939	**0.2475**	0.2674	0.2759	0.3066	0.3644	0.3872	0.4291
	100	0.2267	0.2679	0.2567	**0.208**	0.2261	0.2326	0.2636	0.3046	0.3209	0.3538
	500	**0.1399**	0.17	0.1916	0.1527	0.165	0.1576	0.187	0.2202	0.2292	0.2434
	1000	**0.1121**	0.1365	0.1685	0.1387	0.1438	0.1307	0.1589	0.1922	0.2013	0.2151
	2500	**0.0846**	0.1026	0.1389	0.1193	0.1196	0.1016	0.1259	0.1571	0.1661	0.1805
16	50	0.2869	**0.2173**	0.2268	0.3713	0.2558	0.2397	0.2876	0.4377	0.487	0.5712
	100	0.2121	**0.1864**	0.1939	0.3229	0.1945	0.1943	0.2359	0.3512	0.3899	0.4571
	500	0.1311	**0.1174**	0.1278	0.2162	0.1295	0.1234	0.1568	0.2186	0.2393	0.2754
	1000	0.1032	**0.0984**	0.1083	0.1856	0.1126	0.1013	0.1307	0.1811	0.197	0.2246
	2500	0.0825	**0.0786**	0.0873	0.1556	0.0977	0.0807	0.1049	0.1444	0.1566	0.1773
17	50	0.4754	**0.3692**	0.3846	0.6423	0.4241	0.388	0.4126	0.5147	0.5722	0.705
	100	0.3696	**0.3029**	0.3097	0.5639	0.3178	0.3149	0.3523	0.4368	0.4622	0.5179
	500	0.2053	**0.1956**	0.2084	0.3817	0.2103	0.2011	0.244	0.3069	0.3262	0.3582
	1000	0.1754	**0.1643**	0.1807	0.3285	0.1901	0.1736	0.211	0.2637	0.2801	0.3075
	2500	0.1328	**0.1264**	0.1452	0.273	0.1593	0.1364	0.1677	0.2107	0.2239	0.2464
20	50	**0.3786**	0.544	0.5076	0.43	0.3908	0.4073	0.4589	0.5375	0.5595	0.5929
	100	**0.3372**	0.5391	0.4876	0.3832	0.3565	0.3653	0.4116	0.4808	0.5015	0.5346
	500	**0.2479**	0.5012	0.4184	0.2826	0.2844	0.2783	0.3037	0.3467	0.3637	0.3931
	1000	**0.2105**	0.4445	0.38	0.2455	0.2558	0.2448	0.2674	0.2975	0.3096	0.3337
	2500	**0.1701**	0.3066	0.3306	0.1958	0.2277	0.2081	0.2337	0.2587	0.2657	0.2778
21	50	0.2122	0.4227	0.3592	**0.2077**	0.24	0.2733	0.3659	0.465	0.4818	0.5015
	100	0.165	0.3894	0.3139	**0.1474**	0.2011	0.2169	0.2994	0.4075	0.4336	0.4671
	500	**0.0925**	0.1364	0.1937	0.0972	0.1245	0.1144	0.1622	0.2395	0.2647	0.3073
	1000	**0.071**	0.0736	0.1456	0.0878	0.0995	0.0853	0.1206	0.1815	0.2019	0.2373
	2500	0.0495	**0.0481**	0.0921	0.0757	0.073	0.0571	0.0799	0.1215	0.1357	0.1612
22	50	0.1507	0.1513	0.1439	0.1467	**0.1347**	0.1368	0.1617	0.2007	0.2134	0.2371
	100	0.125	0.1432	0.1313	0.1287	**0.1179**	0.1218	0.1458	0.1782	0.188	0.205
	500	**0.0669**	0.0813	0.0809	0.0962	0.0753	0.0706	0.093	0.1254	0.1347	0.1492
	1000	0.0527	**0.05**	0.0627	0.0858	0.0611	0.0546	0.0711	0.0997	0.1087	0.1233
	2500	0.0364	**0.0345**	0.0424	0.0728	0.0445	0.0371	0.0479	0.0695	0.0767	0.0892
23	50	0.3365	0.2959	**0.2956**	0.3266	0.3091	0.2988	0.299	0.3226	0.3397	0.3773
	100	0.3125	**0.2827**	0.2838	0.2834	0.2861	0.2852	0.2867	0.2962	0.3015	0.3188
	500	0.1655	0.2521	0.2496	**0.1599**	0.2323	0.2235	0.2483	0.2642	0.268	0.2738
	1000	0.1302	0.243	0.2351	**0.128**	0.2104	0.1896	0.2286	0.2502	0.2547	0.2616
	2500	**0.0929**	0.2258	0.1907	0.1049	0.1665	0.1363	0.1771	0.2183	0.2263	0.2378
24	50	0.3231	0.3136	0.3133	**0.3116**	0.3129	0.3125	0.3137	0.3152	0.316	0.3178
	100	0.3089	0.3137	0.3126	**0.3054**	0.3093	0.3102	0.3128	0.3138	0.3143	0.3153
	500	**0.2759**	0.3139	0.3083	0.2837	0.3061	0.305	0.3089	0.3122	0.3128	0.3137
	1000	**0.2548**	0.3122	0.3068	0.2749	0.3047	0.3018	0.3072	0.3118	0.3126	0.3135
	2500	**0.2056**	0.2912	0.3013	0.2317	0.2989	0.2925	0.3008	0.3066	0.3079	0.3098
27	50	0.0779	0.0745	**0.074**	0.0763	0.0746	**0.074**	0.0743	0.0748	0.0751	0.0761
	100	0.074	0.0728	0.072	0.0716	**0.0705**	0.0706	0.0722	0.0735	0.0738	0.0741
	500	**0.042**	0.0669	0.0661	0.0502	0.0621	0.0612	0.0645	0.0683	0.0692	0.0703
	1000	**0.0343**	0.0641	0.0634	0.0388	0.0579	0.0545	0.0603	0.065	0.0661	0.0677
	2500	**0.0248**	0.0609	0.0586	0.0267	0.0488	0.0418	0.051	0.0583	0.0601	0.0625

Tabelle B.3.: Simulationsergebnisse für Dichteschätzung mit Epanechnikovkern – L_1-Risiko für die Dichten 1-14.

Dichte	n	L2CV	BCV	PI	SCV	E-LR	V	KS .5	KS .9	KS .95	KS .99
1	50	0.2957	0.3624	0.3292	**0.2679**	0.2784	0.2954	0.3792	0.5823	0.6602	0.7119
	100	0.2449	0.3134	0.2806	**0.2138**	0.2344	0.2445	0.2992	0.4243	0.4774	0.5792
	500	**0.1356**	0.2282	0.1838	0.137	0.1435	0.1384	0.1721	0.2247	0.2442	0.2822
	1000	0.1049	0.1988	0.1425	0.1122	0.109	**0.1022**	0.1198	0.1686	0.1836	0.2102
	2500	0.0786	0.1648	0.0997	0.0882	0.0785	**0.0743**	0.0819	0.1056	0.1167	0.1395
2	50	0.3703	0.4184	0.3902	0.3708	0.36	**0.3593**	0.385	0.4867	0.5347	0.6403
	100	0.2898	0.3516	0.3115	0.2981	0.2843	**0.2834**	0.2945	0.3553	0.3833	0.4422
	500	0.1848	0.2374	0.1851	0.1975	0.1705	0.1731	**0.1675**	0.1799	0.1883	0.2097
	1000	0.1545	0.2063	0.1503	0.1668	0.1381	0.1429	**0.1354**	0.1405	0.1471	0.1569
	2500	0.1165	0.1634	0.1065	0.1283	0.1005	0.1074	0.0988	**0.0972**	0.0988	0.1029
3	50	0.2143	0.2143	0.2165	0.2555	0.2127	**0.2113**	0.2702	0.4744	0.5568	0.6989
	100	0.1673	**0.1614**	0.1672	0.2096	0.1648	0.1657	0.199	0.3193	0.3711	0.4716
	500	0.0868	0.0869	0.0882	0.1305	0.0865	**0.0848**	0.0992	0.1455	0.1651	0.2037
	1000	0.0661	0.0674	0.0684	0.1052	0.0678	**0.0647**	0.0745	0.1049	0.118	0.1436
	2500	0.046	0.048	0.0477	0.0812	0.0483	**0.0449**	0.0505	0.0681	0.0757	0.0907
4	50	0.2623	**0.2549**	0.2591	0.3242	0.2758	0.2633	0.3131	0.5082	0.5865	0.7384
	100	**0.2088**	0.2095	0.2117	0.2716	0.2104	0.2096	0.2422	0.3646	0.4174	0.5201
	500	0.1195	0.1168	0.1153	0.1766	0.1148	**0.114**	0.126	0.1738	0.1955	0.2382
	1000	0.0908	0.0901	0.0875	0.1457	0.088	**0.087**	0.0935	0.1256	0.1406	0.1705
	2500	0.0656	0.0649	**0.0616**	0.1144	0.0629	0.0617	0.0645	0.083	0.092	0.1104
5	50	0.2161	**0.2072**	0.2163	0.2851	0.2248	0.2155	0.2703	0.478	0.5623	0.7198
	100	0.1703	**0.1583**	0.1675	0.2348	0.1736	0.1723	0.2066	0.3394	0.3948	0.5015
	500	0.0855	**0.0815**	0.0853	0.1444	0.0846	0.0839	0.0968	0.1514	0.1753	0.2207
	1000	0.0665	**0.0636**	0.0665	0.1158	0.0667	0.0648	0.0738	0.1122	0.1287	0.1605
	2500	0.047	**0.0456**	0.0474	0.0886	0.049	0.0459	0.0515	0.0748	0.0849	0.1047
6	50	0.3997	0.3961	**0.3942**	0.473	0.4084	0.4012	0.4493	0.6211	0.6873	0.814
	100	0.3313	0.3286	**0.3257**	0.408	0.3298	0.3265	0.3584	0.475	0.5227	0.6147
	500	0.2051	0.2052	0.2003	0.2869	0.2003	**0.1999**	0.2105	0.2582	0.279	0.3189
	1000	0.1648	0.1642	**0.1588**	0.2443	0.16	0.1591	0.1658	0.2028	0.2179	0.2464
	2500	**0.1209**	0.1295	0.1252	0.1986	0.1267	0.1252	0.1282	0.1444	0.1521	0.1683
7	50	0.2341	0.2285	0.2306	0.2782	0.2285	**0.2257**	0.2807	0.4749	0.5533	0.6996
	100	0.177	**0.1722**	0.1767	0.2297	0.1767	0.1768	0.2088	0.3261	0.3748	0.4712
	500	0.0937	0.0918	0.0928	0.1435	0.0918	**0.0909**	0.103	0.1486	0.1679	0.2063
	1000	0.0701	0.0707	0.0716	0.1156	0.0709	**0.0684**	0.0773	0.1076	0.1204	0.1455
	2500	0.0489	0.0511	0.051	0.0888	0.0514	**0.048**	0.0536	0.0717	0.0793	0.0943
8	50	0.4423	0.5304	0.4876	**0.408**	0.4483	0.4261	0.4356	0.5308	0.5728	0.6633
	100	0.3728	0.4933	0.4228	**0.3344**	0.3845	0.3696	0.3429	0.4057	0.4373	0.4965
	500	0.3448	0.4131	0.2829	**0.2268**	0.2936	0.3113	0.2539	0.2269	0.2272	0.2391
	1000	0.3532	0.3827	0.235	0.188	0.2659	0.2998	0.2349	0.1932	0.1875	**0.1871**
	2500	0.3447	0.3409	0.1918	**0.1513**	0.234	0.2864	0.2199	0.1697	0.161	0.1525
9	50	**0.7517**	0.8283	0.791	0.7665	0.7564	0.759	0.7679	0.8213	0.8483	0.9147
	100	0.6562	0.7416	0.6791	0.6451	0.6514	0.6476	**0.6419**	0.6722	0.6919	0.7403
	500	0.5356	0.5662	0.4872	0.4868	0.494	0.5039	0.4735	**0.4579**	0.4595	0.4704
	1000	0.4932	0.5047	0.406	0.4355	0.443	0.4574	0.4264	0.3992	**0.3946**	0.3949
	2500	0.4571	0.449	**0.3385**	0.4124	0.4082	0.4327	0.4021	0.3638	0.3543	0.3437
10	50	0.7111	0.7197	0.7065	0.7365	0.709	**0.703**	0.7233	0.8304	0.8859	0.9932
	100	0.647	0.6411	**0.6228**	0.6547	0.6323	0.6261	0.6287	0.6967	0.7306	0.797
	500	0.4909	0.473	**0.4437**	0.5339	0.4493	0.4548	0.4447	0.4598	0.4715	0.4937
	1000	0.4223	0.4128	**0.3768**	0.4687	0.3824	0.3894	0.3771	0.3866	0.3933	0.4073
	2500	0.3447	0.3378	**0.302**	0.4026	0.3056	0.3152	0.3039	0.3053	0.3086	0.3171
11	50	0.2036	**0.1946**	0.2062	0.2619	0.2104	0.2075	0.2706	0.482	0.5666	0.7155
	100	0.1546	**0.144**	0.1562	0.2069	0.1534	0.1549	0.1999	0.3427	0.4021	0.5128
	500	0.0806	**0.0765**	0.0812	0.1296	0.0801	0.079	0.0935	0.1526	0.1775	0.2248
	1000	0.0591	**0.0579**	0.0623	0.1048	0.0621	0.0591	0.0711	0.1118	0.1289	0.1616
	2500	0.0425	**0.0422**	0.0456	0.079	0.0468	**0.0422**	0.0499	0.0751	0.0857	0.1058
12	50	0.3703	0.4165	0.3855	0.3813	0.3662	**0.3623**	0.3972	0.5228	0.5768	0.6848
	100	0.2843	0.3327	0.2935	0.3111	0.2747	**0.2745**	0.2962	0.372	0.4054	0.477
	500	0.1648	0.2102	0.1662	0.2108	**0.1582**	0.1583	0.1624	0.1854	0.1972	0.2208
	1000	0.1305	0.1717	0.1286	0.1791	**0.1233**	0.1238	0.1248	0.1391	0.1464	0.1608
	2500	0.0954	0.1315	0.092	0.1416	**0.0888**	0.0894	0.0892	0.0967	0.1004	0.1085
13	50	0.4578	0.5063	0.472	0.4843	**0.45**	0.4508	0.5188	0.7191	0.7896	0.8857
	100	0.3794	0.4146	0.3807	0.4024	**0.3648**	0.3654	0.4001	0.5337	0.5897	0.6954
	500	0.2353	0.2691	0.2382	0.2585	0.2238	**0.223**	0.2333	0.274	0.2922	0.3306
	1000	0.1951	0.2307	0.1963	0.2163	0.1827	**0.1818**	0.1872	0.2136	0.2259	0.2509
	2500	0.15	0.1848	0.148	0.169	**0.1368**	0.1369	0.1379	0.152	0.1589	0.1734
14	50	1.1092	**0.6938**	0.698	0.7572	1.4884	1.4674	1.1149	0.7515	0.7394	0.7556
	100	1.2744	0.6359	**0.6207**	0.6801	1.3976	1.5121	1.4571	0.8399	0.7079	0.667
	500	0.922	0.5367	**0.4876**	0.5382	0.8794	0.8133	1.0177	1.5773	1.6383	1.4877
	1000	0.9159	0.5025	**0.4451**	0.4898	0.7224	0.7037	0.8334	1.0331	1.3029	1.6579
	2500	0.9226	0.4709	**0.3997**	0.4334	0.7014	0.7014	0.7014	0.7415	0.8829	0.9968

Tabelle B.4.: Simulationsergebnisse für Dichteschätzung mit Epanechnikovkern: L_1-Risiko für die Dichten 15-28.

Dichte	n	L2CV	BCV	PI	SCV	E-LR	V	KS .5	KS .9	KS .95	KS .99
15	50	0.3942	0.4383	0.4139	**0.3661**	0.3695	0.3688	0.3967	0.4885	0.5338	0.6336
	100	0.3135	0.3904	0.348	0.2969	0.2978	**0.2964**	0.3104	0.3739	0.4002	0.4548
	500	0.1924	0.2822	0.209	0.1852	0.1778	0.1815	**0.1729**	0.1849	0.1946	0.2143
	1000	0.1673	0.2493	0.1662	0.1549	0.1465	0.1532	**0.1421**	0.1424	0.1471	0.1568
	2500	0.1383	0.2083	0.1264	0.121	0.112	0.1217	0.1095	**0.1053**	0.1065	0.1101
16	50	0.1918	**0.189**	0.1982	0.2405	0.2028	0.2043	0.2683	0.4759	0.5603	0.6673
	100	0.1555	**0.1518**	0.1591	0.2013	0.155	0.1568	0.1991	0.3376	0.395	0.5056
	500	0.0822	**0.0812**	0.0851	0.1185	0.0851	0.0823	0.1005	0.1535	0.1766	0.2222
	1000	0.0643	**0.0642**	0.0676	0.0993	0.069	**0.0642**	0.078	0.1135	0.1281	0.1574
	2500	0.0458	0.0457	0.0473	0.0762	0.0509	**0.0453**	0.0538	0.0759	0.0852	0.1031
17	50	0.205	0.2052	0.2124	0.233	**0.2018**	0.206	0.2793	0.4879	0.5699	0.6527
	100	0.1559	0.1601	0.1626	0.1907	**0.1542**	0.1584	0.211	0.3525	0.4083	0.5135
	500	**0.0833**	0.0878	0.0894	0.1149	0.0892	0.085	0.1101	0.1722	0.1962	0.2413
	1000	**0.0649**	0.0692	0.0702	0.0946	0.0734	0.0668	0.084	0.1268	0.1436	0.1752
	2500	**0.0454**	0.0497	0.0494	0.0723	0.0542	0.0467	0.0576	0.0825	0.0929	0.1128
18	50	0.5775	0.6032	0.5653	0.5539	0.645	0.6014	**0.5477**	0.5839	0.6183	0.6954
	100	0.5444	0.5404	0.4931	**0.478**	0.5758	0.5538	0.4833	0.4803	0.4907	0.5242
	500	0.5499	0.4278	0.3414	0.3205	0.4713	0.5002	0.3991	0.3281	0.3187	**0.3136**
	1000	0.5574	0.3859	0.2874	0.2765	0.4297	0.4807	0.3796	0.3009	0.2866	**0.2693**
	2500	0.5747	0.3388	0.2228	**0.2202**	0.3887	0.4723	0.3648	0.2748	0.2558	0.2317
19	50	0.9861	**0.7994**	0.8053	0.8753	1.1085	1.0168	0.8553	0.8224	0.8457	0.9178
	100	1.0396	**0.7011**	0.7067	0.7982	1.0794	1.0242	0.8231	0.7103	0.7133	0.7494
	500	1.3534	0.5396	**0.5279**	0.6198	1.0092	1.0719	0.8287	0.6239	0.5862	0.5435
	1000	1.4602	0.4866	**0.4641**	0.549	0.9882	1.1048	0.8607	0.6328	0.5829	0.5204
	2500	1.5673	0.4295	**0.3926**	0.4695	0.9816	1.154	0.9222	0.66	0.6026	0.525
20	50	0.7783	0.8776	0.8209	**0.76**	0.7732	0.7651	0.7717	0.8595	0.8925	0.9644
	100	0.6906	0.7902	0.7279	0.6689	0.6959	0.6761	**0.6684**	0.7144	0.7327	0.7716
	500	0.5708	0.6176	0.5145	0.531	0.5378	0.5415	0.5225	0.5003	**0.4982**	0.5011
	1000	0.5425	0.5519	0.4498	0.4845	0.4742	0.4886	0.4613	0.4439	0.4422	**0.4415**
	2500	0.4603	0.4741	0.3631	0.4211	0.3999	0.4156	0.3924	0.3644	0.3604	**0.3563**
21	50	**0.3485**	0.7533	0.5518	0.3637	0.3657	0.3974	0.5584	1.0089	1.1641	1.4008
	100	0.2706	0.6599	0.4271	**0.2622**	0.2836	0.2959	0.3869	0.658	0.7722	1.0049
	500	**0.1431**	0.4885	0.2172	0.1545	0.156	0.1499	0.1816	0.2506	0.2805	0.3442
	1000	**0.1109**	0.4257	0.1578	0.1333	0.1226	0.1152	0.1351	0.18	0.1984	0.2355
	2500	**0.0741**	0.354	0.0972	0.1039	0.0841	0.076	0.0879	0.115	0.1259	0.147
22	50	0.2501	0.2805	0.265	0.2531	**0.237**	0.2412	0.3066	0.4742	0.5491	0.6838
	100	0.1947	0.2434	0.2194	0.2083	**0.1908**	0.1962	0.2432	0.3481	0.3925	0.4848
	500	**0.1066**	0.1655	0.124	0.139	0.115	0.1104	0.135	0.1871	0.2064	0.2406
	1000	**0.0786**	0.1355	0.0919	0.1169	0.0885	0.0819	0.0991	0.1366	0.151	0.1782
	2500	**0.0559**	0.1031	0.0622	0.0928	0.0634	0.0568	0.0667	0.0903	0.0995	0.1171
23	50	0.4126	0.4244	0.4244	0.4212	**0.4123**	0.415	0.462	0.5934	0.6525	0.7725
	100	0.382	0.3976	0.3976	**0.3552**	0.3805	0.384	0.4126	0.4916	0.525	0.5946
	500	0.3502	0.3521	0.3532	**0.1719**	0.3271	0.3018	0.3547	0.3787	0.3923	0.4137
	1000	0.3499	0.3499	0.3518	**0.1322**	0.2846	0.238	0.3313	0.3489	0.3548	0.3755
	2500	0.2038	0.3475	0.288	**0.1002**	0.2173	0.1632	0.2383	0.3331	0.3445	0.3456
24	50	0.5438	0.8393	0.7324	**0.473**	0.512	0.5827	0.7647	0.9473	0.988	1.0409
	100	0.3794	0.814	0.6343	**0.3712**	0.4403	0.4728	0.6313	0.8373	0.8798	0.9418
	500	0.2113	0.7155	0.3935	**0.2097**	0.3195	0.2994	0.3911	0.548	0.5987	0.6771
	1000	0.1737	0.658	0.3159	**0.1694**	0.2641	0.2357	0.308	0.4305	0.4746	0.5514
	2500	0.1268	0.582	0.2336	**0.1224**	0.2018	0.1689	0.2151	0.3039	0.3373	0.3939
25	50	0.4653	0.4737	0.4757	**0.405**	0.412	0.4348	0.5298	0.6982	0.7633	0.8552
	100	0.4027	0.4414	0.4332	**0.3399**	0.3469	0.353	0.4352	0.5782	0.623	0.7029
	500	0.2168	0.3737	0.3005	**0.2042**	0.2095	0.2066	0.2253	0.2777	0.3013	0.3832
	1000	0.1739	0.3461	0.2349	0.1676	0.1678	**0.1649**	0.1752	0.2092	0.2244	0.2558
	2500	0.1353	0.3057	0.1699	0.1285	0.1251	**0.1231**	0.1268	0.1447	0.153	0.1704
26	50	1.2196	1.5625	1.4969	1.3817	**0.6019**	0.6401	0.82	1.4301	1.6214	1.8354
	100	1.1882	1.5576	1.482	1.2789	**0.4829**	0.4979	0.6052	0.8866	1.0824	1.4197
	500	1.0624	1.4388	1.2723	0.8508	0.2965	**0.2906**	0.3253	0.4175	0.4548	0.5269
	1000	0.2545	1.4455	1.174	0.7082	0.2421	**0.2347**	0.2554	0.3126	0.3375	0.3863
	2500	0.1923	1.3751	0.9572	0.5399	0.18	**0.1728**	0.184	0.2162	0.23	0.2585
27	50	0.6478	0.6864	0.6709	**0.5709**	0.6111	0.6337	0.7018	0.8253	0.8725	0.9107
	100	0.6261	0.6613	0.6405	**0.5719**	0.5849	0.5996	0.6499	0.7326	0.7632	0.8225
	500	0.5965	0.6167	0.5874	**0.3529**	0.542	0.5381	0.5502	0.6037	0.6195	0.6449
	1000	0.5956	0.5978	0.5421	**0.2204**	0.5372	0.4661	0.5478	0.5672	0.5866	0.5993
	2500	0.5336	0.5912	0.5471	**0.123**	0.4222	0.3086	0.4631	0.5526	0.541	0.5384
28	50	0.4205	0.6061	0.5193	**0.3606**	0.3798	0.4077	0.5012	0.6684	0.7107	0.7715
	100	0.3256	0.5679	0.4516	**0.2913**	0.3049	0.3179	0.3975	0.5309	0.5792	0.6594
	500	0.1865	0.4673	0.2904	**0.1749**	0.1774	0.1751	0.1923	0.2574	0.2901	0.3464
	1000	0.1583	0.4267	0.224	0.142	**0.1415**	0.1416	0.1447	0.1763	0.1929	0.2276
	2500	0.1313	0.3751	0.164	0.1084	**0.1061**	0.1078	0.1069	0.118	0.1252	0.1417

Tabelle B.5.: Simulationsergebnisse für Dichteschätzung mit Epanechnikov-Kern – L_2-Risiko für die Dichten 1-13.

Dichte	n	L2CV	BCV	PI	SCV	E-LR	V	KS .5	KS .9	KS .95	KS .99
1	50	0.0958	0.1164	0.1041	0.0975	**0.0902**	0.0939	0.1215	0.1909	0.2222	0.2445
	100	0.0752	0.0973	0.0856	**0.0696**	**0.0696**	0.0729	0.0924	0.1351	0.1523	0.1878
	500	0.0316	0.0678	0.0514	**0.0307**	0.0362	0.0339	0.0471	0.0665	0.0736	0.0869
	1000	**0.021**	0.0585	0.0386	0.0211	0.0251	0.0219	0.0297	0.0478	0.0531	0.0626
	2500	**0.0124**	0.0481	0.025	0.0135	0.0162	0.0135	0.0177	0.0272	0.0313	0.0394
2	50	0.0757	0.1018	0.0885	**0.0672**	**0.0672**	0.07	0.0865	0.1291	0.1469	0.1824
	100	0.0509	0.0863	0.0687	**0.0455**	0.047	0.0484	0.0595	0.0878	0.0986	0.1201
	500	0.0214	0.0596	0.0371	0.0203	0.0208	**0.0202**	0.024	0.0345	0.039	0.0482
	1000	0.0148	0.0526	0.0289	0.0144	0.0151	**0.0143**	0.0167	0.0236	0.0272	0.0323
	2500	0.0085	0.0425	0.0195	0.0086	0.0092	**0.0084**	0.0096	0.0133	0.0149	0.0178
3	50	0.0197	**0.0182**	0.0187	0.0346	0.0203	0.0184	0.0277	0.0694	0.0891	0.126
	100	0.0133	**0.0113**	0.0119	0.0241	0.0128	0.0124	0.0161	0.036	0.0461	0.0678
	500	0.0036	0.0035	0.0036	0.009	0.0035	**0.0034**	0.0046	0.009	0.0113	0.0162
	1000	0.0021	0.0022	0.0023	0.0057	0.0022	**0.002**	0.0027	0.0051	0.0062	0.0088
	2500	**0.001**	0.0012	0.0012	0.0035	0.0012	**0.001**	0.0013	0.0023	0.0028	0.0039
4	50	0.0164	**0.0161**	0.0168	0.0224	0.0171	0.0166	0.0263	0.0576	0.0705	0.0963
	100	0.0109	0.0116	0.012	0.0156	**0.0108**	0.0112	0.0169	0.035	0.0431	0.0595
	500	0.0036	0.0043	0.004	0.0063	0.0038	**0.0035**	0.0054	0.0107	0.0132	0.0182
	1000	**0.0019**	0.0027	0.0023	0.0043	0.0024	0.002	0.0031	0.0062	0.0077	0.0107
	2500	**0.001**	0.0016	0.0012	0.0027	0.0014	**0.001**	0.0015	0.0031	0.0039	0.0054
5	50	0.0073	**0.0063**	0.0069	0.0138	0.0079	0.007	0.0107	0.0278	0.036	0.0528
	100	0.0049	**0.0039**	0.0044	0.0096	0.005	0.0048	0.0065	0.0156	0.0201	0.0298
	500	0.0012	**0.0011**	0.0012	0.0034	**0.0011**	**0.0011**	0.0015	0.0035	0.0046	0.007
	1000	7e-04	**6e-04**	7e-04	0.0021	7e-04	**6e-04**	9e-04	0.002	0.0025	0.0038
	2500	**3e-04**	**3e-04**	4e-04	0.0013	4e-04	**3e-04**	4e-04	9e-04	0.0011	0.0017
6	50	**0.013**	0.0135	0.0132	0.0173	0.0131	0.0134	0.0211	0.0449	0.0543	0.0719
	100	0.0082	0.0093	0.0087	0.0118	**0.0079**	0.0081	0.0131	0.0281	0.0345	0.0471
	500	**0.0025**	0.0034	0.0028	0.0046	0.0027	**0.0025**	0.0039	0.0082	0.0103	0.0145
	1000	**0.0015**	0.0021	0.0016	0.0031	0.0017	**0.0015**	0.0022	0.0047	0.0058	0.0082
	2500	**7e-04**	0.0011	8e-04	0.0018	9e-04	**7e-04**	0.001	0.002	0.0025	0.0036
7	50	0.0128	**0.0112**	0.0115	0.0206	0.0123	**0.0112**	0.0172	0.0423	0.0538	0.0765
	100	0.0075	**0.0068**	0.0071	0.0142	0.0076	0.0074	0.01	0.0226	0.0287	0.0416
	500	0.0022	**0.0019**	0.002	0.0053	0.002	**0.0019**	0.0025	0.0054	0.0068	0.0099
	1000	0.0012	0.0012	0.0012	0.0034	0.0012	**0.0011**	0.0014	0.0029	0.0036	0.0052
	2500	6e-04	6e-04	6e-04	0.002	6e-04	**5e-04**	7e-04	0.0013	0.0016	0.0023
9	50	0.0371	0.0647	0.0542	0.0375	**0.0331**	0.0352	0.0442	0.0623	0.0689	0.0809
	100	0.0251	0.0607	0.0463	0.0268	**0.0237**	0.0244	0.0307	0.0443	0.0496	0.0589
	500	**0.0106**	0.0501	0.0296	0.0119	0.0112	0.0108	0.013	0.0183	0.0206	0.0248
	1000	**0.0071**	0.0459	0.0229	0.0078	0.0076	0.0072	0.0085	0.012	0.0134	0.0161
	2500	**0.0044**	0.042	0.0172	0.0048	0.0049	**0.0044**	0.0052	0.0071	0.0079	0.0094
10	50	0.0132	0.0193	0.0165	**0.0118**	0.0123	0.0135	0.0197	0.0321	0.0365	0.0436
	100	0.0086	0.0169	0.0128	**0.0076**	0.0086	0.0093	0.0139	0.0237	0.027	0.0328
	500	**0.0028**	0.0112	0.0057	**0.0028**	0.0038	0.0034	0.0053	0.0091	0.0107	0.0136
	1000	**0.0016**	0.0092	0.0038	0.0017	0.0026	0.0021	0.0033	0.0058	0.0068	0.0088
	2500	**9e-04**	0.0072	0.0022	0.001	0.0016	0.0012	0.0018	0.0032	0.0037	0.0048
11	50	0.0127	**0.0102**	0.0112	0.0228	0.014	0.0121	0.0171	0.0444	0.0578	0.0836
	100	0.0074	**0.0056**	0.0064	0.014	0.007	0.0068	0.0097	0.0243	0.032	0.0483
	500	0.0019	**0.0016**	0.0018	0.0053	0.0018	0.0018	0.0023	0.0054	0.007	0.0108
	1000	0.001	**9e-04**	0.0011	0.0034	0.0011	0.001	0.0013	0.0029	0.0038	0.0057
	2500	**5e-04**	**5e-04**	6e-04	0.0019	6e-04	**5e-04**	7e-04	0.0014	0.0017	0.0025
12	50	0.04	0.0588	0.0481	**0.0374**	0.0378	0.0383	0.0527	0.0908	0.1062	0.1359
	100	0.0241	0.0435	0.0311	0.0243	**0.0217**	0.0227	0.0319	0.0549	0.0644	0.0839
	500	**0.0072**	0.0237	0.0114	0.0101	0.008	0.0075	0.0101	0.0171	0.0203	0.0265
	1000	**0.0043**	0.0181	0.0071	0.0072	0.0049	0.0044	0.0059	0.0099	0.0117	0.0154
	2500	**0.0022**	0.0125	0.0037	0.0044	0.0027	**0.0022**	0.0029	0.0048	0.0056	0.0074
13	50	0.0552	0.0718	0.0612	0.0546	**0.0514**	0.054	0.075	0.135	0.1573	0.1939
	100	0.0396	0.0538	0.0444	0.038	**0.0374**	0.0386	0.0495	0.083	0.0987	0.1299
	500	0.0179	0.032	0.0248	**0.0172**	0.0197	0.0188	0.0234	0.033	0.0367	0.0441
	1000	0.0125	0.0274	0.0198	**0.0123**	0.0151	0.0138	0.0172	0.0239	0.0264	0.0312
	2500	**0.0078**	0.0222	0.0143	**0.0078**	0.0104	0.0089	0.0111	0.0153	0.0169	0.02

Tabelle B.6.: Simulationsergebnisse für Dichteschätzung mit Epanechnikov-Kern – L_2-Risiko für die Dichten 15-28.

Dichte	n	L2CV	BCV	PI	SCV	E-LR	V	KS .5	KS .9	KS .95	KS .99
15	50	0.4774	0.5815	0.5376	**0.3958**	0.4056	0.4211	0.5059	0.6556	0.7159	0.8312
	100	0.3563	0.5378	0.4646	**0.2983**	0.3073	0.3161	0.3794	0.509	0.5511	0.6278
	500	0.1524	0.4221	0.2903	**0.1419**	0.1512	0.1465	0.1728	0.2327	0.2573	0.3023
	1000	0.1078	0.3847	0.2365	**0.104**	0.1123	0.1055	0.1244	0.1652	0.1825	0.2127
	2500	**0.0699**	0.3372	0.186	0.0702	0.0776	**0.0699**	0.0809	0.108	0.1186	0.1385
16	50	0.026	**0.0234**	0.0251	0.0515	0.0318	0.0278	0.0406	0.1041	0.1349	0.1763
	100	0.0171	**0.0152**	0.0164	0.0358	0.0168	0.0165	0.0237	0.0581	0.0754	0.1132
	500	0.0051	**0.0047**	0.0051	0.0125	0.0052	0.005	0.0071	0.0152	0.0192	0.0281
	1000	0.0031	**0.003**	0.0033	0.0085	0.0035	**0.003**	0.0044	0.0091	0.0113	0.016
	2500	0.0017	**0.0016**	0.0018	0.0051	0.0021	**0.0016**	0.0023	0.0046	0.0057	0.008
17	50	0.0573	**0.0467**	0.0493	0.0958	0.0591	0.0517	0.0703	0.1723	0.2235	0.2786
	100	0.0306	**0.0288**	0.0297	0.0612	0.0306	0.0303	0.0429	0.0975	0.1251	0.1854
	500	0.0092	0.0095	0.0098	0.022	0.0098	**0.0091**	0.0136	0.0275	0.0341	0.0484
	1000	**0.006**	0.0063	0.0065	0.0151	0.007	**0.006**	0.0087	0.0164	0.0199	0.0277
	2500	**0.003**	0.0034	0.0034	0.0088	0.004	0.0031	0.0045	0.008	0.0097	0.0131
20	50	0.0568	0.104	0.0879	0.0591	**0.0472**	0.0507	0.0665	0.0976	0.1087	0.127
	100	0.0339	0.0981	0.0752	0.0405	**0.0333**	0.0349	0.0459	0.0688	0.0771	0.0921
	500	**0.0122**	0.0832	0.0452	0.0142	0.0137	0.013	0.0168	0.0255	0.029	0.0359
	1000	**0.0078**	0.075	0.0345	0.0083	0.0088	0.0079	0.0102	0.0158	0.018	0.0226
	2500	0.0041	0.0665	0.0233	**0.004**	0.0051	0.0043	0.0055	0.0085	0.0097	0.0122
21	50	**0.0333**	0.1088	0.0775	0.0347	0.0385	0.0465	0.0775	0.1427	0.1625	0.1923
	100	0.0199	0.0951	0.0568	**0.0178**	0.0246	0.0277	0.0485	0.0935	0.1104	0.142
	500	**0.0057**	0.0695	0.0192	0.0061	0.0081	0.0071	0.0126	0.0266	0.0327	0.0449
	1000	**0.0034**	0.0597	0.0101	0.0046	0.005	0.004	0.0067	0.0144	0.0179	0.0251
	2500	**0.0015**	0.0477	0.0038	0.0028	0.0025	0.0018	0.0029	0.0061	0.0076	0.0108
22	50	0.019	0.0219	0.02	0.0204	**0.0174**	0.0175	0.0245	0.0425	0.0517	0.0704
	100	0.0124	0.0178	0.0151	0.014	**0.012**	0.0126	0.0177	0.0285	0.0331	0.0433
	500	**0.0036**	0.0094	0.0053	0.006	0.0044	0.004	0.0064	0.0116	0.0135	0.0171
	1000	**0.002**	0.0066	0.003	0.0042	0.0027	0.0022	0.0035	0.0067	0.0081	0.0107
	2500	**0.001**	0.004	0.0013	0.0027	0.0014	0.0011	0.0016	0.0031	0.0038	0.0051
23	50	**0.0602**	0.0611	0.0608	0.0692	0.0636	0.061	0.066	0.0915	0.1047	0.1319
	100	0.0538	0.0557	0.0557	**0.0502**	0.0552	0.0552	0.0572	0.0697	0.0762	0.0911
	500	0.047	0.0472	0.0483	**0.0117**	0.044	0.0377	0.0494	0.0511	0.0532	0.0557
	1000	0.0487	0.0487	0.0508	**0.0068**	0.0345	0.0242	0.0461	0.047	0.0473	0.0507
	2500	0.0237	0.0505	0.036	**0.0038**	0.0204	0.0113	0.0247	0.0474	0.0498	0.0482
24	50	0.0745	0.1231	0.105	**0.0617**	0.0678	0.0787	0.1101	0.1367	0.1411	0.1474
	100	0.0449	0.1182	0.0862	**0.043**	0.0546	0.0596	0.0849	0.1214	0.1279	0.1356
	500	**0.017**	0.0983	0.0458	**0.017**	0.0351	0.0323	0.0453	0.0686	0.0768	0.0909
	1000	**0.0118**	0.0872	0.0348	0.012	0.0278	0.0238	0.0337	0.0508	0.0573	0.0688
	2500	**0.0059**	0.0737	0.0242	0.0063	0.02	0.0154	0.0218	0.0332	0.0375	0.0453
25	50	0.2165	0.2201	0.2205	**0.1772**	0.1815	0.1972	0.2332	0.2846	0.3115	0.3523
	100	0.1812	0.214	0.2108	**0.1285**	0.137	0.1453	0.2008	0.2427	0.2564	0.2859
	500	0.0623	0.1964	0.1488	**0.0589**	0.07	0.0665	0.0839	0.1286	0.1491	0.1909
	1000	0.0426	0.186	0.105	**0.0424**	0.0539	0.0489	0.0614	0.0863	0.0971	0.1215
	2500	**0.0266**	0.1669	0.072	0.0267	0.0374	0.0317	0.0399	0.0555	0.0613	0.0722
26	50	0.8399	1.2941	1.2684	1.1904	**0.4049**	0.491	0.8575	1.2659	1.3081	1.345
	100	0.7998	1.2878	1.2528	1.1278	**0.2854**	0.3123	0.5155	0.9814	1.1247	1.2694
	500	0.6985	1.2526	1.1326	0.8038	0.1469	**0.1381**	0.1819	0.3003	0.3596	0.4839
	1000	**0.0787**	1.2484	1.0763	0.7121	0.1135	0.101	0.1314	0.1924	0.2199	0.2831
	2500	**0.0492**	1.2205	0.9385	0.5797	0.0803	0.0663	0.0861	0.1224	0.1359	0.162
27	50	0.0218	0.0225	0.022	**0.021**	0.0217	0.0216	0.0226	0.0258	0.0273	0.0285
	100	0.0207	0.0215	0.021	0.0216	**0.02**	0.0203	0.0212	0.0231	0.0239	0.0256
	500	0.0192	0.0199	0.0194	**0.0089**	0.0187	0.0188	0.0181	0.0195	0.02	0.0206
	1000	0.0192	0.0192	0.0175	**0.0038**	0.0188	0.0147	0.0189	0.0185	0.0192	0.0193
	2500	0.0171	0.0194	0.0186	**0.0013**	0.0121	0.0068	0.0144	0.0193	0.0181	0.0173
28	50	0.3139	0.4431	0.3916	**0.2725**	0.2867	0.3111	0.3793	0.4775	0.4997	0.5208
	100	0.2361	0.4187	0.3481	**0.2092**	0.2246	0.2382	0.306	0.3971	0.4256	0.4705
	500	**0.1078**	0.3601	0.2446	0.1111	0.1252	0.1187	0.1499	0.2158	0.2434	0.2834
	1000	**0.0789**	0.3366	0.1972	0.0843	0.0963	0.0887	0.1077	0.151	0.1682	0.2
	2500	**0.0556**	0.3062	0.1521	0.0591	0.0706	0.0606	0.0745	0.0992	0.1096	0.1287

Tabelle B.7.: Simulationsergebnisse für Dichteschätzung mit Epanechnikov-Kern – L_∞-Risiko für die Dichten.

Dichte	n	L2CV	BCV	PI	SCV	E-LR	V	KS .5	KS .9	KS .95	KS .99
3	50	0.1558	**0.1462**	0.1491	0.2398	0.1624	0.1484	0.174	0.2449	0.2716	0.3181
	100	0.1354	**0.1249**	0.1279	0.2107	0.1344	0.1316	0.1437	0.1912	0.2075	0.2405
	500	0.0845	0.0895	0.0905	0.1476	0.0862	**0.0825**	0.1014	0.1302	0.1393	0.1545
	1000	0.0716	0.0778	0.0789	0.1236	0.0774	**0.0699**	0.0871	0.1113	0.1187	0.1314
	2500	**0.0538**	0.0654	0.0647	0.1051	0.0657	0.0553	0.0694	0.089	0.095	0.1053
4	50	0.1618	0.1682	0.1711	0.1605	**0.1549**	0.1649	0.2194	0.2993	0.3211	0.3566
	100	0.1446	0.1568	0.158	**0.1392**	0.1418	0.1481	0.1898	0.2544	0.2737	0.3054
	500	**0.0932**	0.1198	0.1126	0.0985	0.1079	0.1004	0.1316	0.1739	0.1872	0.2095
	1000	**0.0705**	0.1044	0.0934	0.0872	0.0936	0.0812	0.1089	0.1463	0.1576	0.1768
	2500	**0.0567**	0.0898	0.076	0.0731	0.0806	0.0649	0.0862	0.1165	0.1258	0.1412
5	50	0.057	**0.0514**	0.0532	0.0937	0.0627	0.0554	0.0653	0.1083	0.1223	0.1451
	100	0.0478	**0.0402**	0.0429	0.0812	0.0485	0.0469	0.053	0.0826	0.0938	0.1129
	500	0.0242	**0.0223**	0.0233	0.0563	0.0234	0.0236	0.0268	0.042	0.0482	0.0594
	1000	0.0193	**0.0178**	0.0185	0.046	0.0188	0.0186	0.0207	0.0319	0.0365	0.045
	2500	0.0145	**0.0136**	0.0141	0.0394	0.0146	0.0138	0.0153	0.0223	0.0253	0.0308
6	50	**0.0904**	0.0957	0.0936	0.1077	0.0915	0.0932	0.1264	0.1853	0.2017	0.2279
	100	0.0765	0.0845	0.0809	0.094	**0.0735**	0.0755	0.1043	0.1534	0.1683	0.1924
	500	**0.045**	0.0576	0.05	0.0683	0.0485	0.0455	0.0615	0.0919	0.1021	0.119
	1000	0.0358	0.0472	0.0387	0.0592	0.0403	**0.0356**	0.0478	0.0718	0.0797	0.0936
	2500	**0.0247**	0.0347	0.0273	0.0489	0.0298	0.0249	0.0323	0.0488	0.0547	0.0649
7	50	0.0902	**0.0823**	0.0833	0.1402	0.0942	0.0841	0.1011	0.1585	0.1772	0.2079
	100	0.0704	**0.0634**	0.0658	0.1231	0.0726	0.0698	0.0783	0.1192	0.1336	0.1591
	500	0.041	**0.0366**	0.0374	0.0862	0.0377	0.0381	0.0414	0.0605	0.0683	0.0827
	1000	0.0312	**0.0291**	0.0297	0.0719	0.0296	0.0293	0.0322	0.0449	0.0502	0.0602
	2500	0.0228	0.0222	0.0224	0.0607	0.0226	**0.0221**	0.0232	0.0303	0.0336	0.0401
10	50	0.1372	0.1734	0.1615	**0.1233**	0.1327	0.1444	0.1738	0.2082	0.2167	0.228
	100	0.1186	0.1702	0.1528	**0.106**	0.122	0.1286	0.1576	0.1907	0.1988	0.2107
	500	0.0748	0.1536	0.1218	**0.0613**	0.1	0.0943	0.117	0.1435	0.1511	0.1633
	1000	0.0616	0.1455	0.1078	**0.0456**	0.0915	0.0826	0.1017	0.1255	0.1321	0.1433
	2500	0.0478	0.1355	0.091	**0.0338**	0.0796	0.068	0.0836	0.1036	0.1096	0.1195
11	50	0.0915	**0.0788**	0.0815	0.15	0.1001	0.0876	0.0986	0.1611	0.1833	0.218
	100	0.072	**0.06**	0.0633	0.1237	0.0716	0.068	0.0753	0.1214	0.1396	0.1703
	500	0.0393	**0.0348**	0.0356	0.0901	0.0366	0.0375	0.0387	0.0588	0.0677	0.0844
	1000	0.0301	**0.0278**	0.0286	0.076	0.0289	0.0295	0.0307	0.0444	0.0505	0.0622
	2500	0.023	**0.0217**	0.0222	0.0615	0.0225	0.0223	0.0232	0.0313	0.035	0.0421
12	50	0.2814	0.3219	0.3049	**0.2542**	0.2754	0.2831	0.3143	0.3694	0.3917	0.4328
	100	0.2376	0.2931	0.2692	**0.2089**	0.2348	0.2421	0.2728	0.312	0.3274	0.3591
	500	**0.1462**	0.246	0.204	0.1479	0.173	0.165	0.1961	0.2299	0.2392	0.2533
	1000	**0.1187**	0.2323	0.1804	0.1358	0.1509	0.1372	0.1668	0.2017	0.2111	0.2253
	2500	**0.0887**	0.2158	0.1486	0.1164	0.1255	0.1061	0.1322	0.1651	0.1746	0.1897
16	50	0.2278	**0.2132**	0.2241	0.361	0.2579	0.2393	0.2848	0.4438	0.4945	0.5528
	100	0.1926	**0.1847**	0.1939	0.312	0.1942	0.1933	0.2346	0.3528	0.3935	0.4642
	500	0.124	**0.1221**	0.1313	0.2135	0.1312	0.1253	0.1598	0.2221	0.2428	0.2788
	1000	0.1042	0.1051	0.1124	0.1817	0.1146	**0.1039**	0.1342	0.1864	0.2026	0.2302
	2500	**0.0822**	0.0859	0.0911	0.1526	0.1005	0.0826	0.1081	0.1494	0.1621	0.1833
17	50	0.423	**0.3771**	0.3913	0.6272	0.4417	0.3965	0.4285	0.5357	0.5919	0.6547
	100	0.3203	**0.3131**	0.319	0.5466	0.3232	0.3201	0.3645	0.4572	0.4842	0.5398
	500	**0.2032**	0.2129	0.2176	0.3735	0.2166	0.2058	0.2537	0.3187	0.3386	0.3727
	1000	**0.1734**	0.186	0.1888	0.3174	0.1976	0.1789	0.2194	0.274	0.2911	0.3197
	2500	**0.1343**	0.1537	0.1522	0.2658	0.1659	0.1413	0.1746	0.219	0.2328	0.2562
20	50	0.4202	0.5515	0.5172	0.4428	**0.395**	0.4121	0.4623	0.5399	0.5619	0.5948
	100	**0.3585**	0.5481	0.4994	0.3963	0.3615	0.37	0.416	0.4844	0.5046	0.5373
	500	**0.2578**	0.5234	0.4322	0.2937	0.2921	0.2859	0.3111	0.3527	0.3695	0.3985
	1000	**0.2179**	0.5073	0.3949	0.2575	0.2654	0.2543	0.2768	0.3053	0.3169	0.3404
	2500	**0.175**	0.4868	0.3458	0.2091	0.2376	0.2178	0.2434	0.2681	0.2747	0.2861
21	50	**0.2171**	0.4366	0.3801	0.2193	0.2449	0.2792	0.3701	0.4673	0.4834	0.5022
	100	0.1686	0.4209	0.337	**0.1505**	0.2034	0.2203	0.3037	0.4109	0.4367	0.4695
	500	0.0946	0.3774	0.2102	**0.0944**	0.1248	0.1138	0.1649	0.244	0.2692	0.3113
	1000	**0.0717**	0.3555	0.156	0.0853	0.0993	0.085	0.1212	0.1849	0.2059	0.2421
	2500	**0.0496**	0.3228	0.0964	0.0734	0.0731	0.0566	0.08	0.1233	0.1383	0.1648
22	50	0.1479	0.1625	0.1547	0.1415	**0.1368**	0.1431	0.1692	0.2027	0.2142	0.2347
	100	**0.1212**	0.1543	0.1421	0.1231	0.1215	0.1271	0.1528	0.1832	0.1916	0.2065
	500	**0.0661**	0.1204	0.0873	0.0951	0.0775	0.0721	0.0976	0.1324	0.1422	0.1568
	1000	**0.0506**	0.1048	0.0662	0.0836	0.0624	0.0548	0.0736	0.1048	0.1145	0.1301
	2500	**0.0359**	0.0834	0.0436	0.071	0.045	0.0373	0.0487	0.0722	0.0801	0.0936
23	50	0.3007	0.2999	**0.2971**	0.3393	0.3131	0.3007	0.3004	0.3222	0.3413	0.3787
	100	**0.2862**	0.2874	0.2873	0.2966	0.2925	0.2901	0.289	0.2956	0.3004	0.3199
	500	0.256	0.2561	0.2555	**0.1588**	0.2428	0.2307	0.2559	0.2689	0.2726	0.2762
	1000	0.2512	0.2512	0.2516	**0.1264**	0.2169	0.1918	0.2415	0.2536	0.2575	0.2667
	2500	0.1575	0.242	0.2085	**0.1021**	0.1689	0.1365	0.1809	0.2332	0.2405	0.2462
24	50	0.3187	0.3134	0.3147	**0.3126**	0.3162	0.3142	0.3142	0.3164	0.3175	0.3153
	100	0.3096	0.3149	0.3127	**0.3088**	0.3104	0.3121	0.3123	0.3135	0.3129	0.3159
	500	0.2863	0.313	0.3068	**0.2845**	0.3071	0.3087	0.3069	0.3147	0.3124	0.3112
	1000	**0.2713**	0.3121	0.3079	0.2905	0.3105	0.3057	0.3084	0.3119	0.3156	0.3165
	2500	**0.2119**	0.3142	0.3083	0.2503	0.2995	0.2882	0.3039	0.3075	0.3054	0.3074
27	50	0.0754	0.0751	0.0746	0.0759	0.0753	0.0748	0.0746	0.0748	0.0747	**0.0744**
	100	0.0728	0.0736	0.0729	0.0739	**0.0711**	0.0713	0.0729	0.0739	0.074	0.0741
	500	0.0687	0.07	0.0677	**0.0518**	0.064	0.0637	0.0652	0.0691	0.07	0.0709
	1000	0.0682	0.0684	0.0644	**0.0392**	0.0607	0.0558	0.0624	0.066	0.0673	0.0684
	2500	0.0628	0.0671	0.0612	**0.027**	0.0498	0.0415	0.0527	0.0607	0.0614	0.0632

Tabelle B.8.: Simulationsergebnisse für Dichteschätzung mit Epanechnikov-Kern – gewählte Bandbreiten für die Dichten 1-14.

Dichte	n	L2CV	BCV	PI	SCV	E-LR	V	KS .5	KS .9	KS .95	KS .99
1	50	0.291	0.4286	0.3865	0.1462	0.2499	0.318	0.4894	0.7759	0.881	0.9513
	100	0.2245	0.3744	0.3201	0.1071	0.2208	0.2458	0.3646	0.5658	0.6382	0.7736
	500	0.0954	0.2753	0.1983	0.0504	0.1404	0.1286	0.1813	0.2715	0.304	0.3643
	1000	0.0662	0.24	0.1591	0.036	0.1081	0.092	0.1295	0.1934	0.2164	0.2593
	2500	0.0378	0.2009	0.1176	0.0228	0.0778	0.0604	0.0845	0.1253	0.1399	0.167
2	50	0.5133	0.8386	0.687	0.2724	0.3511	0.433	0.6736	1.1594	1.3645	1.8077
	100	0.3326	0.7217	0.5462	0.192	0.2785	0.3071	0.46	0.7473	0.8605	1.0885
	500	0.1281	0.5261	0.3229	0.0862	0.1552	0.1421	0.2014	0.3065	0.3455	0.4205
	1000	0.0858	0.4619	0.2581	0.0613	0.1203	0.1026	0.1441	0.2163	0.2428	0.2928
	2500	0.0537	0.387	0.1918	0.0388	0.0826	0.0642	0.0899	0.1343	0.1505	0.1808
3	50	0.6916	0.7871	0.8125	0.3073	0.5635	0.7137	1.0963	1.724	1.9626	2.4033
	100	0.5907	0.6791	0.7048	0.2353	0.5435	0.603	0.8776	1.2894	1.4356	1.7145
	500	0.4136	0.4933	0.5034	0.126	0.4662	0.4291	0.5802	0.7866	0.8538	0.9744
	1000	0.3603	0.43	0.4368	0.0969	0.4271	0.3745	0.4897	0.6476	0.6985	0.7885
	2500	0.2797	0.3604	0.3568	0.0671	0.3626	0.3014	0.3843	0.4988	0.5355	0.6001
4	50	1.1244	1.2293	1.2982	0.4984	0.8871	1.1648	1.8987	3.2575	3.798	4.9836
	100	0.9428	1.0825	1.1153	0.3773	0.8694	0.9739	1.4767	2.3418	2.6727	3.3265
	500	0.5607	0.7719	0.726	0.1837	0.6868	0.6227	0.8849	1.2776	1.4156	1.6668
	1000	0.4348	0.6695	0.5986	0.1337	0.604	0.5144	0.7147	1.0161	1.1185	1.3052
	2500	0.3339	0.5601	0.4719	0.0892	0.504	0.3991	0.5424	0.759	0.8308	0.9578
5	50	1.7637	1.9358	2.1037	0.7947	1.482	1.8714	2.9223	4.6612	5.3513	6.8024
	100	1.5017	1.6698	1.8194	0.6068	1.427	1.588	2.3692	3.5754	4.0012	4.8309
	500	1.0825	1.2109	1.3156	0.3299	1.2196	1.1164	1.5413	2.1446	2.3458	2.7023
	1000	0.9524	1.0544	1.1502	0.2554	1.1269	0.9757	1.3134	1.7954	1.9508	2.2232
	2500	0.7782	0.883	0.9607	0.1818	0.9917	0.8005	1.06	1.4219	1.5364	1.7337
6	50	1.4977	1.8114	1.7621	0.695	1.305	1.6379	2.6057	4.5964	5.4756	7.5295
	100	1.1765	1.5675	1.4757	0.5135	1.1463	1.2833	1.9764	3.1878	3.6712	4.6976
	500	0.7839	1.1281	0.989	0.2576	0.9484	0.8701	1.201	1.7131	1.8965	2.2327
	1000	0.6655	0.9775	0.8341	0.1916	0.8544	0.7401	0.9956	1.3769	1.5053	1.7387
	2500	0.526	0.81	0.669	0.1298	0.7174	0.5782	0.767	1.035	1.1233	1.2798
7	50	1.1679	1.3545	1.381	0.5265	0.9465	1.21	1.866	2.9794	3.4146	4.2988
	100	1.0112	1.164	1.1908	0.4009	0.9155	1.0199	1.4905	2.2153	2.4711	2.969
	500	0.7009	0.8439	0.8497	0.2145	0.7763	0.7133	0.9729	1.3321	1.4481	1.657
	1000	0.6189	0.7354	0.74	0.1656	0.716	0.6255	0.8236	1.0964	1.1838	1.3379
	2500	0.497	0.6169	0.612	0.1166	0.6162	0.5073	0.6544	0.855	0.9186	1.03
8	50	0.1637	0.372	0.2747	0.1098	0.0769	0.1013	0.1812	0.3748	0.4628	0.6551
	100	0.0863	0.3256	0.2145	0.0764	0.0439	0.0509	0.093	0.1933	0.2381	0.3328
	500	0.009	0.238	0.1171	0.032	0.0115	0.0098	0.0183	0.0391	0.0484	0.0682
	1000	0.0037	0.2079	0.0894	0.0219	0.0063	0.0047	0.0088	0.0189	0.0235	0.0334
	2500	0.0014	0.1746	0.0628	0.0133	0.003	0.0019	0.0035	0.0075	0.0093	0.0131
9	50	1.4431	5.0494	3.374	1.3948	1.046	1.3163	2.2037	4.518	5.756	9.0189
	100	0.7951	4.3111	2.5551	0.9478	0.7489	0.8324	1.31	2.3882	2.8834	4.0079
	500	0.2592	3.0328	1.3589	0.3903	0.3545	0.3224	0.4716	0.7578	0.8711	1.0984
	1000	0.1677	2.6345	1.043	0.2678	0.2578	0.2177	0.3133	0.4909	0.5596	0.6944
	2500	0.1017	2.207	0.7429	0.1648	0.1747	0.135	0.1906	0.2901	0.3275	0.3995
10	50	2.9258	5.7141	4.5101	1.8457	2.5196	3.2603	5.9686	13.3414	17.6044	28.8788
	100	1.856	4.8973	3.5169	1.2813	2.045	2.3237	3.8882	7.7511	9.5429	13.5687
	500	0.8406	3.362	1.9188	0.536	1.354	1.2291	1.8015	2.8432	3.2576	4.0788
	1000	0.6251	2.8981	1.5011	0.373	1.1368	0.963	1.3673	2.0534	2.3021	2.7952
	2500	0.4492	2.4119	1.1062	0.2351	0.8983	0.7109	0.9712	1.3971	1.5505	1.8402
11	50	1.0876	1.1705	1.2905	0.4872	0.9267	1.1779	1.8013	2.8057	3.1951	3.9427
	100	0.9336	1.0223	1.1422	0.3796	0.8859	0.9901	1.4713	2.1842	2.4414	2.9256
	500	0.6757	0.7426	0.8303	0.2076	0.7572	0.6951	0.954	1.333	1.4565	1.6745
	1000	0.5863	0.6486	0.728	0.1614	0.7018	0.6054	0.8198	1.1162	1.2116	1.3791
	2500	0.4927	0.5442	0.6118	0.1161	0.6212	0.5022	0.6634	0.8861	0.9567	1.0788
12	50	0.6715	1.0608	0.8747	0.3489	0.5208	0.6434	0.9836	1.628	1.904	2.5149
	100	0.4632	0.9037	0.6947	0.2458	0.4494	0.4968	0.7129	1.0942	1.243	1.5438
	500	0.2476	0.6554	0.424	0.113	0.3212	0.2995	0.3939	0.5469	0.6016	0.7044
	1000	0.1923	0.574	0.3436	0.0811	0.2684	0.2382	0.3073	0.4171	0.4556	0.5267
	2500	0.1441	0.4827	0.2634	0.0528	0.2163	0.1814	0.2294	0.3035	0.329	0.3753
13	50	0.6205	0.9636	0.7895	0.3279	0.508	0.6356	1.0034	2.0838	2.7859	4.8656
	100	0.4263	0.7722	0.6164	0.228	0.4375	0.4835	0.7204	1.1575	1.3492	1.8554
	500	0.1843	0.5085	0.3767	0.1062	0.2708	0.2476	0.3518	0.5301	0.5951	0.716
	1000	0.1239	0.4418	0.3047	0.0761	0.2104	0.1786	0.2531	0.3804	0.4263	0.5116
	2500	0.0752	0.3661	0.2263	0.0481	0.1511	0.1172	0.1644	0.2446	0.2736	0.3275
14	50	0.0105	0.0339	0.0236	0.0099	4e-04	0.0011	0.0093	0.0607	0.1011	0.1676
	100	0.001	0.0285	0.0161	0.0062	0.0038	0.0026	4e-04	0.0073	0.018	0.0472
	500	0	0.0199	0.0076	0.0023	0.0394	0.0746	0.0054	0	0	1e-04
	1000	0	0.0174	0.0057	0.0016	0.1225	0.1347	0.0653	0	0	0
	2500	0	0.0147	0.0039	9e-04	0.1365	0.1365	0.1365	0.1119	0.0314	0.0014

Tabelle B.9.: Simulationsergebnisse für Dichteschätzung mit Epanechnikov-Kern – gewählte Bandbreiten für die Dichten 15-28.

Dichte	n	L2CV	BCV	PI	SCV	E-LR	V	KS .5	KS .9	KS .95	KS .99
15	50	0.1513	0.2455	0.2053	0.0806	0.0935	0.1159	0.1863	0.3267	0.3848	0.5082
	100	0.0966	0.2157	0.1624	0.0568	0.0655	0.0731	0.1154	0.1989	0.2326	0.3004
	500	0.0256	0.1547	0.0889	0.0238	0.0297	0.0268	0.0402	0.0655	0.0754	0.0948
	1000	0.0151	0.1346	0.0694	0.0165	0.0212	0.0176	0.0261	0.0421	0.0483	0.0603
	2500	0.0076	0.1123	0.05	0.0102	0.0136	0.0102	0.015	0.0238	0.0272	0.0338
16	50	0.4669	0.5055	0.5539	0.2075	0.4047	0.5154	0.7892	1.217	1.3745	1.5781
	100	0.3985	0.4415	0.4834	0.1599	0.3939	0.4402	0.6461	0.9576	1.0637	1.2646
	500	0.288	0.3225	0.3548	0.0879	0.3422	0.312	0.4364	0.6041	0.6582	0.7519
	1000	0.2459	0.2822	0.3077	0.0674	0.313	0.2654	0.37	0.5064	0.5485	0.6212
	2500	0.1926	0.2354	0.2512	0.0466	0.2774	0.2185	0.2982	0.4055	0.4388	0.4947
17	50	0.2611	0.298	0.312	0.1172	0.2223	0.2837	0.4289	0.6551	0.7382	0.8229
	100	0.2248	0.264	0.2731	0.0905	0.2215	0.246	0.3554	0.5172	0.5736	0.6788
	500	0.159	0.1917	0.1978	0.0492	0.1946	0.1789	0.2422	0.3279	0.3557	0.4053
	1000	0.1344	0.167	0.1703	0.0374	0.1798	0.1571	0.206	0.2727	0.2943	0.3324
	2500	0.1081	0.1397	0.1383	0.0258	0.1527	0.1265	0.162	0.2111	0.2269	0.2548
18	50	0.2895	0.7456	0.5256	0.2157	0.1183	0.1571	0.2965	0.6512	0.8247	1.2352
	100	0.1253	0.6396	0.4	0.1468	0.0696	0.0806	0.1505	0.3206	0.3985	0.57
	500	0.0137	0.4708	0.2164	0.0616	0.0172	0.0147	0.0275	0.0593	0.0737	0.1046
	1000	0.0057	0.4065	0.1642	0.0418	0.01	0.0075	0.0139	0.0294	0.0365	0.0516
	2500	0.0019	0.339	0.1149	0.0254	0.0047	0.0029	0.0055	0.0117	0.0145	0.0205
19	50	0.1618	0.614	0.4244	0.1773	0.0607	0.0949	0.2746	1.133	1.637	3.213
	100	0.0512	0.4821	0.2888	0.1087	0.0303	0.0385	0.115	0.384	0.5695	1.0394
	500	0.0015	0.3343	0.1447	0.0426	0.006	0.0046	0.0132	0.0442	0.0622	0.1109
	1000	4e-04	0.2938	0.1099	0.0291	0.0029	0.0017	0.005	0.0164	0.0232	0.0411
	2500	1e-04	0.248	0.0765	0.0177	0.001	5e-04	0.0013	0.0047	0.0065	0.0111
20	50	1.0738	3.7818	2.489	1.0351	0.6336	0.806	1.4084	3.08	4.0125	6.5346
	100	0.4875	3.1679	1.836	0.6847	0.4891	0.5414	0.8508	1.5776	1.9227	2.719
	500	0.1731	2.1909	0.9381	0.272	0.2574	0.2369	0.3312	0.509	0.5798	0.7236
	1000	0.1204	1.8653	0.7047	0.1828	0.1996	0.1734	0.2351	0.3461	0.3887	0.4722
	2500	0.0838	1.5287	0.4921	0.1103	0.1491	0.1213	0.1599	0.2258	0.25	0.2962
21	50	0.6151	2.9867	1.7343	0.7133	0.7856	0.9984	1.7809	4.2594	5.4783	8.8333
	100	0.4728	2.3077	1.244	0.4539	0.6425	0.7113	1.1251	2.3053	2.8843	4.2029
	500	0.294	1.5804	0.6676	0.1874	0.4371	0.4034	0.5465	0.7868	0.8831	1.0854
	1000	0.2494	1.3636	0.5179	0.1298	0.3734	0.327	0.4324	0.597	0.6569	0.772
	2500	0.1998	1.1386	0.3765	0.081	0.3157	0.2621	0.3352	0.4438	0.4806	0.5475
22	50	1.0732	1.4867	1.3685	0.5194	0.8803	1.1092	1.7277	2.7732	3.1637	3.8687
	100	0.8281	1.3236	1.1592	0.3884	0.8313	0.9237	1.3552	2.0809	2.3437	2.8356
	500	0.473	0.9735	0.7161	0.1811	0.6435	0.596	0.7957	1.1051	1.2172	1.4279
	1000	0.3915	0.8447	0.581	0.1302	0.5527	0.4861	0.6327	0.8526	0.9294	1.0724
	2500	0.3063	0.7036	0.4446	0.0849	0.4545	0.3782	0.4823	0.635	0.6854	0.7762
23	50	0.9459	1.1105	1.1311	0.426	0.7194	0.9317	1.4695	2.2994	2.6069	3.1971
	100	0.785	0.9636	0.9651	0.3197	0.6158	0.7102	1.1279	1.7214	1.9278	2.3158
	500	0.6626	0.6785	0.6257	0.1551	0.3831	0.3338	0.5854	0.9111	1.0156	1.2054
	1000	0.5884	0.5887	0.495	0.1097	0.3188	0.2616	0.4068	0.6817	0.7636	0.9039
	2500	0.266	0.4897	0.3283	0.0644	0.2494	0.1971	0.2708	0.4178	0.4758	0.5829
24	50	0.977	2.3846	1.6655	0.6592	0.8113	1.0628	1.8618	3.6954	4.3207	5.138
	100	0.462	2.14	1.2702	0.4529	0.6934	0.7864	1.2628	2.3865	2.8132	3.6214
	500	0.159	1.5722	0.6651	0.1841	0.5028	0.4569	0.6611	1.0223	1.1636	1.424
	1000	0.105	1.3594	0.5107	0.1262	0.4059	0.3425	0.4972	0.7535	0.8474	1.0305
	2500	0.0598	1.1166	0.3638	0.0772	0.3029	0.2357	0.329	0.4977	0.5631	0.6769
25	50	0.5246	0.5989	0.5888	0.2214	0.2405	0.3693	0.7601	1.3005	1.4757	1.7239
	100	0.3501	0.5122	0.4615	0.1533	0.1747	0.1964	0.4542	0.9767	1.1086	1.3214
	500	0.0605	0.3766	0.2353	0.0613	0.0923	0.0834	0.1248	0.2041	0.2377	0.4208
	1000	0.0369	0.3294	0.1717	0.0404	0.0712	0.0595	0.0873	0.139	0.1588	0.1973
	2500	0.0202	0.2747	0.1155	0.0237	0.0483	0.0364	0.0531	0.0831	0.0944	0.1161
26	50	13.2096	18.9889	13.5176	5.4089	0.102	0.1357	0.5537	4.1907	7.6725	21.7118
	100	14.0111	17.1404	10.8683	3.9182	0.0846	0.0941	0.1567	0.798	1.6496	3.7066
	500	10.4311	11.3264	4.1693	1.2305	0.0511	0.0466	0.0669	0.1029	0.1168	0.1476
	1000	0.0148	10.1417	2.858	0.7709	0.0403	0.0341	0.0486	0.0736	0.0827	0.0998
	2500	0.0098	7.9188	1.6464	0.388	0.0294	0.0227	0.0321	0.048	0.0537	0.0644
27	50	6.5211	8.3297	7.6842	2.9053	4.3864	5.6754	9.2486	14.9577	17.0014	18.6274
	100	5.7585	7.2977	6.3302	2.1177	3.8459	4.3757	6.8146	10.7814	12.1951	14.8497
	500	4.7394	5.4851	3.8726	0.9884	2.2374	1.9022	3.1673	4.9511	5.5949	6.7949
	1000	4.6954	4.7973	3.0963	0.703	1.7016	1.3195	2.2211	3.5129	3.9565	4.8118
	2500	2.8892	4.0203	2.2764	0.4434	1.1943	0.9295	1.3121	2.1017	2.4058	2.9746
28	50	0.2316	0.5229	0.3945	0.1526	0.1706	0.2189	0.3719	0.6848	0.8114	0.9629
	100	0.1364	0.471	0.3109	0.1076	0.1312	0.1481	0.2411	0.4299	0.5071	0.6628
	500	0.0365	0.3551	0.1743	0.0466	0.0631	0.0565	0.0878	0.1495	0.174	0.2221
	1000	0.0216	0.3117	0.1361	0.0325	0.0449	0.0368	0.0563	0.0945	0.1095	0.139
	2500	0.0115	0.26	0.0977	0.0201	0.0292	0.0215	0.0324	0.053	0.061	0.0767

Tabelle B.10.: Simulationsergebnisse für Dichteschätzung mit Gaußkern (alternative GoF-Tests) – L_2-Risiko für die Dichten 1-13.

Dichte	n	Kuip .5	Kuip .9	Kuip .95	Kuip .99	CvM V	CvM .5	CvM .9	CvM .95	CvM .99
1	50	0.1116	0.1475	0.16	0.1852	0.0794	0.118	0.1956	0.2276	0.3046
	100	0.0728	0.0988	0.1074	0.1222	0.0619	0.0924	0.1454	0.1638	0.2068
	500	0.0296	0.0371	0.04	0.0463	0.0336	0.0488	0.0815	0.0918	0.109
	1000	0.02	0.0244	0.0261	0.0296	0.0254	0.0363	0.0595	0.0682	0.0858
	2500	0.0122	0.0145	0.0155	0.0175	0.0179	0.0252	0.0403	0.0456	0.0573
2	50	0.0873	0.1175	0.1275	0.1471	0.0686	0.0952	0.1559	0.1777	0.2225
	100	0.0575	0.0771	0.0841	0.0973	0.0503	0.0706	0.1146	0.1302	0.1618
	500	0.0224	0.0291	0.0318	0.037	0.0243	0.0344	0.0559	0.0636	0.0791
	1000	0.0159	0.0201	0.0215	0.0249	0.0187	0.0258	0.0422	0.048	0.0591
	2500	0.0093	0.012	0.013	0.0148	0.0125	0.0174	0.0274	0.0309	0.0387
3	50	0.0277	0.0478	0.0555	0.0709	0.0174	0.025	0.0608	0.077	0.1155
	100	0.0145	0.024	0.0278	0.036	0.011	0.0146	0.0342	0.0434	0.0655
	500	0.0036	0.0054	0.0063	0.0082	0.0031	0.004	0.0088	0.0111	0.0169
	1000	0.0021	0.003	0.0034	0.0044	0.0018	0.0023	0.0049	0.0061	0.0093
	2500	0.001	0.0014	0.0016	0.002	9e-04	0.0012	0.0023	0.0029	0.0044
4	50	0.0239	0.0387	0.0441	0.0543	0.0153	0.022	0.0455	0.0555	0.0782
	100	0.0148	0.0224	0.0254	0.0314	0.0108	0.0151	0.0292	0.0352	0.0491
	500	0.0041	0.0062	0.007	0.0088	0.0033	0.0047	0.0091	0.011	0.0154
	1000	0.0023	0.0034	0.0039	0.0049	0.0019	0.0027	0.0052	0.0063	0.0089
	2500	0.0011	0.0017	0.0019	0.0024	0.001	0.0014	0.0027	0.0033	0.0046
5	50	0.0103	0.0177	0.0206	0.0266	0.0068	0.0092	0.0224	0.0285	0.0433
	100	0.0057	0.0092	0.0107	0.0139	0.0043	0.0054	0.0124	0.0158	0.0242
	500	0.0012	0.0019	0.0022	0.0029	0.0011	0.0013	0.0029	0.0037	0.0057
	1000	7e-04	0.001	0.0012	0.0015	6e-04	8e-04	0.0016	0.002	0.0031
	2500	3e-04	5e-04	5e-04	7e-04	3e-04	4e-04	7e-04	9e-04	0.0014
6	50	0.0198	0.0312	0.0352	0.0429	0.0124	0.0186	0.039	0.0471	0.0648
	100	0.011	0.0176	0.0201	0.0249	0.0079	0.0119	0.0249	0.0302	0.0423
	500	0.003	0.0046	0.0053	0.0066	0.0025	0.0036	0.0076	0.0093	0.0134
	1000	0.0017	0.0025	0.0029	0.0036	0.0014	0.0021	0.0043	0.0053	0.0077
	2500	8e-04	0.0011	0.0012	0.0015	7e-04	9e-04	0.0019	0.0024	0.0035
7	50	0.0174	0.0298	0.0345	0.044	0.0109	0.0158	0.0377	0.0474	0.0699
	100	0.0092	0.0151	0.0176	0.0226	0.0068	0.0092	0.0214	0.027	0.0403
	500	0.0022	0.0033	0.0038	0.005	0.0019	0.0024	0.0053	0.0068	0.0103
	1000	0.0012	0.0018	0.0021	0.0027	0.0011	0.0014	0.0029	0.0037	0.0056
	2500	6e-04	8e-04	0.001	0.0012	5e-04	7e-04	0.0014	0.0017	0.0026
9	50	0.0456	0.0576	0.0615	0.0684	0.0389	0.0527	0.0768	0.0842	0.0973
	100	0.0307	0.0401	0.0431	0.0486	0.0292	0.0406	0.0609	0.0672	0.0792
	500	0.0129	0.0167	0.018	0.0207	0.0153	0.0214	0.0329	0.0367	0.0446
	1000	0.0086	0.0112	0.012	0.0138	0.0112	0.0156	0.0243	0.0272	0.033
	2500	0.0051	0.0066	0.0072	0.0082	0.0075	0.0106	0.0164	0.0186	0.0229
10	50	0.0182	0.0245	0.0266	0.0305	0.0134	0.0189	0.031	0.035	0.0427
	100	0.0119	0.0167	0.0184	0.0213	0.0095	0.014	0.0232	0.0263	0.0325
	500	0.004	0.0057	0.0063	0.0075	0.0037	0.0055	0.0098	0.0114	0.0147
	1000	0.0024	0.0035	0.0039	0.0047	0.0023	0.0036	0.0065	0.0076	0.0099
	2500	0.0013	0.0018	0.002	0.0025	0.0013	0.002	0.0037	0.0043	0.0058
11	50	0.0163	0.0278	0.0325	0.0421	0.0115	0.014	0.0344	0.0444	0.0687
	100	0.0083	0.0139	0.0164	0.0217	0.0062	0.0078	0.019	0.0246	0.0382
	500	0.0019	0.0029	0.0033	0.0044	0.0017	0.002	0.0044	0.0056	0.0088
	1000	0.0011	0.0015	0.0018	0.0023	0.001	0.0011	0.0023	0.003	0.0046
	2500	6e-04	8e-04	8e-04	0.0011	5e-04	6e-04	0.0011	0.0014	0.0021
12	50	0.0547	0.078	0.0862	0.1024	0.038	0.0573	0.106	0.1236	0.1604
	100	0.0304	0.0446	0.0499	0.0601	0.0236	0.0366	0.0697	0.0822	0.1087
	500	0.0093	0.0135	0.0151	0.0184	0.0085	0.0132	0.0253	0.0301	0.0405
	1000	0.0054	0.0078	0.0087	0.0107	0.0053	0.0082	0.0162	0.0193	0.0264
	2500	0.0028	0.004	0.0044	0.0054	0.0029	0.0045	0.0088	0.0105	0.0145
13	50	0.0715	0.0988	0.109	0.1285	0.0559	0.0815	0.1459	0.1651	0.1955
	100	0.0446	0.0586	0.064	0.075	0.0404	0.0556	0.0993	0.1166	0.1508
	500	0.0197	0.0239	0.0254	0.0283	0.0222	0.0292	0.0434	0.049	0.0622
	1000	0.0142	0.0171	0.0182	0.0202	0.0177	0.0232	0.0332	0.0367	0.0446
	2500	0.009	0.0109	0.0115	0.0128	0.0129	0.0169	0.0241	0.0264	0.0311

Tabelle B.11.: Simulationsergebnisse für Dichteschätzung mit Gaußkern (alternative GoF-Tests) – L_2-Risiko für die Dichten 15-28.

Dichte	n	Kuip .5	Kuip .9	Kuip .95	Kuip .99	CvM V	CvM .5	CvM .9	CvM .95	CvM .99
15	50	0.5105	0.621	0.6595	0.7289	0.4221	0.546	0.7587	0.827	0.9723
	100	0.3675	0.4612	0.4925	0.549	0.3423	0.4481	0.6325	0.6875	0.7908
	500	0.1693	0.2109	0.2254	0.2534	0.1953	0.2565	0.3695	0.4077	0.4766
	1000	0.1206	0.1531	0.1648	0.1838	0.1548	0.2046	0.2933	0.3208	0.3771
	2500	0.0811	0.1009	0.1077	0.121	0.1138	0.1531	0.2208	0.2423	0.2826
16	50	0.0375	0.0635	0.0746	0.0974	0.0249	0.0315	0.0782	0.1011	0.1577
	100	0.0199	0.0328	0.0386	0.0507	0.0156	0.0189	0.0437	0.056	0.087
	500	0.0051	0.0078	0.009	0.0115	0.0045	0.0053	0.0108	0.0136	0.0209
	1000	0.0031	0.0045	0.0052	0.0066	0.0028	0.0032	0.0061	0.0076	0.0113
	2500	0.0016	0.0022	0.0025	0.0032	0.0015	0.0017	0.0029	0.0035	0.0052
17	50	0.0669	0.111	0.1291	0.1668	0.0474	0.0597	0.1433	0.184	0.2844
	100	0.036	0.0586	0.0681	0.0876	0.0269	0.0358	0.0839	0.1063	0.1616
	500	0.0091	0.0138	0.0159	0.0204	0.0077	0.0103	0.0228	0.0288	0.0431
	1000	0.0055	0.0078	0.0089	0.0112	0.005	0.0063	0.013	0.0163	0.0243
	2500	0.0028	0.0036	0.004	0.0049	0.0026	0.0032	0.0061	0.0076	0.0112
20	50	0.0687	0.0897	0.0961	0.1079	0.0581	0.0817	0.1216	0.1326	0.1511
	100	0.0457	0.0615	0.0667	0.0763	0.043	0.0627	0.0963	0.1064	0.1247
	500	0.016	0.0218	0.024	0.0283	0.0186	0.0285	0.048	0.0546	0.0674
	1000	0.0096	0.0134	0.0148	0.0175	0.0122	0.0191	0.0337	0.0387	0.0489
	2500	0.0054	0.0075	0.0083	0.0098	0.0074	0.0116	0.0208	0.0241	0.0309
21	50	0.0685	0.1004	0.1114	0.1316	0.0431	0.0718	0.126	0.1445	0.1844
	100	0.0392	0.0598	0.0671	0.081	0.0284	0.0497	0.0915	0.1049	0.1321
	500	0.009	0.0145	0.0166	0.0208	0.0083	0.0157	0.0362	0.0441	0.0601
	1000	0.0049	0.0076	0.0087	0.011	0.0047	0.0088	0.0212	0.0265	0.0381
	2500	0.002	0.0031	0.0036	0.0045	0.002	0.0037	0.0094	0.0119	0.018
22	50	0.0236	0.034	0.0378	0.0456	0.0159	0.023	0.0437	0.0527	0.0742
	100	0.0159	0.0219	0.0241	0.0284	0.0116	0.0172	0.0303	0.0354	0.0476
	500	0.0053	0.0082	0.0093	0.0111	0.0041	0.007	0.0136	0.0158	0.0201
	1000	0.003	0.0049	0.0056	0.007	0.0024	0.0043	0.0091	0.0108	0.0142
	2500	0.0014	0.0022	0.0025	0.0032	0.0012	0.002	0.0046	0.0057	0.0081
23	50	0.0645	0.076	0.0808	0.0907	0.0592	0.0643	0.0889	0.1004	0.1281
	100	0.0554	0.0607	0.063	0.0677	0.0535	0.0569	0.0704	0.0766	0.0919
	500	0.0423	0.0472	0.0479	0.0492	0.0314	0.0468	0.0526	0.0545	0.0587
	1000	0.0303	0.0418	0.0435	0.0454	0.019	0.0383	0.0487	0.0501	0.0529
	2500	0.0142	0.0256	0.0293	0.0348	0.0085	0.0187	0.0422	0.0451	0.0477
24	50	0.1023	0.123	0.1274	0.1343	0.0697	0.0995	0.1354	0.1425	0.1572
	100	0.0739	0.0969	0.1035	0.1133	0.0532	0.0763	0.1169	0.1257	0.1379
	500	0.0356	0.0474	0.0513	0.0584	0.0284	0.0399	0.0621	0.0699	0.0859
	1000	0.0258	0.0346	0.0374	0.0427	0.0217	0.0303	0.0466	0.0521	0.0637
	2500	0.0164	0.0219	0.0237	0.0274	0.0152	0.0213	0.0324	0.0361	0.0438
25	50	0.2255	0.2519	0.2614	0.2813	0.1936	0.2216	0.2683	0.2896	0.3414
	100	0.1861	0.219	0.225	0.2363	0.162	0.2018	0.2334	0.2449	0.2732
	500	0.0704	0.0886	0.0956	0.1101	0.0899	0.1287	0.1856	0.1959	0.2092
	1000	0.0505	0.0616	0.0656	0.0734	0.0701	0.0977	0.1514	0.1666	0.1891
	2500	0.032	0.039	0.0415	0.0461	0.0516	0.0692	0.1054	0.1183	0.1436
26	50	0.7495	1.0499	1.1268	1.2235	0.5487	0.9457	1.27	1.3161	1.3547
	100	0.4179	0.6209	0.6945	0.8355	0.3557	0.6345	1.1489	1.2132	1.2883
	500	0.1472	0.1926	0.2105	0.2478	0.1636	0.2374	0.461	0.5615	0.8009
	1000	0.1038	0.131	0.1406	0.1597	0.1266	0.1719	0.2936	0.3487	0.4831
	2500	0.0649	0.0826	0.0886	0.0996	0.0908	0.1203	0.18	0.2045	0.264
27	50	0.0222	0.0238	0.0244	0.0256	0.0208	0.0225	0.0262	0.0278	0.0316
	100	0.0202	0.0215	0.0219	0.0226	0.0196	0.0212	0.0238	0.0247	0.0268
	500	0.0176	0.018	0.0182	0.0185	0.0152	0.0184	0.0203	0.0209	0.0219
	1000	0.0155	0.0173	0.0175	0.0177	0.0087	0.0173	0.0191	0.0196	0.0206
	2500	0.0083	0.0136	0.0147	0.016	0.0038	0.0087	0.0177	0.0181	0.0189
28	50	0.35	0.421	0.4375	0.4634	0.2909	0.3937	0.5005	0.5294	0.5963
	100	0.2496	0.3071	0.3277	0.3663	0.2389	0.326	0.4462	0.47	0.5126
	500	0.12	0.1436	0.1515	0.1666	0.1501	0.1994	0.2879	0.3195	0.3787
	1000	0.0859	0.1034	0.1093	0.1209	0.1216	0.158	0.2272	0.2505	0.2966
	2500	0.0589	0.0695	0.0729	0.0797	0.0925	0.1207	0.1696	0.1859	0.2161

Tabelle B.12.: Simulationsergebnisse für Dichteschätzung mit Gaußkern (alternative GoF-Tests) – L_∞-Risiko für die Dichten.

Dichte	n	Kuip .5	Kuip .9	Kuip .95	Kuip .99	CvM M	CvM .5	CvM .9	CvM .95	CvM .99
3	50	0.17	0.2091	0.2219	0.2452	0.1451	0.1643	0.2296	0.2541	0.3075
	100	0.1373	0.1635	0.1715	0.1861	0.1225	0.1391	0.1829	0.1982	0.2346
	500	0.0905	0.1093	0.1148	0.1244	0.0788	0.0971	0.1271	0.1357	0.1512
	1000	0.0766	0.093	0.0978	0.1061	0.0662	0.083	0.1096	0.1171	0.131
	2500	0.0606	0.0742	0.0782	0.085	0.0524	0.0673	0.0894	0.0958	0.1076
4	50	0.2112	0.2601	0.2732	0.2949	0.1585	0.2046	0.2777	0.2981	0.3357
	100	0.1779	0.2154	0.2265	0.2454	0.1459	0.1813	0.2396	0.2566	0.2882
	500	0.1137	0.1401	0.1478	0.1612	0.0972	0.124	0.1642	0.176	0.1981
	1000	0.0912	0.1142	0.1209	0.1324	0.0776	0.1015	0.1365	0.1467	0.1659
	2500	0.0708	0.0891	0.0945	0.1039	0.0626	0.0813	0.1097	0.118	0.1335
5	50	0.0642	0.0866	0.0938	0.1066	0.0551	0.061	0.098	0.1104	0.1342
	100	0.049	0.0628	0.0682	0.0783	0.0431	0.0474	0.0745	0.0844	0.1035
	500	0.0242	0.0298	0.0323	0.0373	0.0236	0.0248	0.0378	0.0429	0.0536
	1000	0.019	0.0224	0.0242	0.0278	0.019	0.0193	0.0283	0.0321	0.0401
	2500	0.0141	0.0161	0.0171	0.0193	0.0143	0.0144	0.0198	0.0222	0.0275
6	50	0.1234	0.1577	0.1673	0.1838	0.0894	0.1196	0.176	0.1921	0.221
	100	0.0945	0.1229	0.1314	0.146	0.0748	0.0991	0.1462	0.1601	0.1858
	500	0.0517	0.0675	0.0728	0.082	0.0451	0.0592	0.0883	0.0974	0.115
	1000	0.0395	0.0513	0.0554	0.0628	0.0352	0.0454	0.0686	0.076	0.0904
	2500	0.0269	0.0342	0.037	0.0421	0.0246	0.0312	0.0474	0.0528	0.0636
7	50	0.1009	0.1336	0.1438	0.1619	0.0827	0.0964	0.1505	0.1681	0.202
	100	0.0745	0.0967	0.1045	0.1185	0.0651	0.0745	0.1157	0.1298	0.1574
	500	0.0385	0.0472	0.051	0.0583	0.0372	0.0405	0.0602	0.0679	0.0835
	1000	0.0301	0.0358	0.0383	0.0433	0.0295	0.0314	0.0451	0.0506	0.0623
	2500	0.0224	0.0253	0.0268	0.0298	0.0224	0.0232	0.0314	0.0349	0.0427
10	50	0.1701	0.1915	0.1973	0.2065	0.1479	0.1738	0.2073	0.2151	0.2274
	100	0.1487	0.1709	0.1767	0.1858	0.1341	0.1599	0.1911	0.1987	0.2114
	500	0.1039	0.1216	0.1266	0.135	0.0998	0.1206	0.1481	0.1554	0.1684
	1000	0.0891	0.1047	0.1091	0.1165	0.0874	0.1056	0.1306	0.1374	0.1498
	2500	0.0719	0.0849	0.0887	0.0951	0.0723	0.0878	0.1097	0.1158	0.1269
11	50	0.0962	0.1275	0.1383	0.1577	0.0877	0.0891	0.1422	0.162	0.2008
	100	0.071	0.0908	0.0992	0.1149	0.0653	0.0683	0.1077	0.1227	0.1528
	500	0.0366	0.0425	0.0457	0.0524	0.0377	0.037	0.0527	0.06	0.0756
	1000	0.0294	0.033	0.0351	0.0395	0.0304	0.0296	0.04	0.0448	0.0557
	2500	0.0225	0.0245	0.0256	0.028	0.0232	0.0227	0.0285	0.0315	0.0382
12	50	0.317	0.3543	0.3673	0.3912	0.2824	0.3206	0.3958	0.4215	0.4699
	100	0.2662	0.2954	0.3046	0.3226	0.2449	0.2799	0.3404	0.3619	0.4036
	500	0.1874	0.2136	0.2203	0.2313	0.1803	0.2123	0.2484	0.2591	0.2819
	1000	0.1598	0.1864	0.1932	0.2043	0.158	0.1897	0.224	0.2318	0.2469
	2500	0.1275	0.1526	0.1595	0.1708	0.1307	0.161	0.1967	0.2052	0.2188
16	50	0.2793	0.3578	0.3852	0.433	0.2267	0.2579	0.3936	0.4398	0.5299
	100	0.2166	0.2764	0.2972	0.3346	0.1872	0.2129	0.3142	0.3492	0.4182
	500	0.1347	0.1694	0.1804	0.2	0.1147	0.1385	0.1946	0.2129	0.249
	1000	0.11	0.1393	0.1485	0.1648	0.094	0.1148	0.1602	0.1745	0.2025
	2500	0.087	0.1099	0.117	0.1295	0.075	0.0902	0.125	0.1357	0.1564
17	50	0.4101	0.4605	0.4758	0.5125	0.3885	0.395	0.489	0.5346	0.6615
	100	0.3407	0.3902	0.4052	0.4302	0.3046	0.3388	0.4255	0.4479	0.4979
	500	0.2181	0.2571	0.2686	0.2886	0.1918	0.2315	0.2982	0.317	0.3512
	1000	0.1854	0.2182	0.2279	0.2448	0.164	0.1991	0.2559	0.2722	0.3021
	2500	0.1446	0.1721	0.1801	0.1939	0.1285	0.1601	0.2077	0.2212	0.2463
20	50	0.478	0.5283	0.5418	0.5654	0.4461	0.5114	0.5894	0.6068	0.6327
	100	0.4224	0.4714	0.4855	0.509	0.4129	0.4744	0.5503	0.5685	0.5978
	500	0.3083	0.3396	0.3506	0.3712	0.3223	0.3724	0.4456	0.4645	0.4973
	1000	0.2705	0.2943	0.3024	0.318	0.2871	0.3273	0.3968	0.4155	0.448
	2500	0.2395	0.2584	0.2634	0.2721	0.2577	0.2825	0.3375	0.3548	0.3857
21	50	0.3517	0.4184	0.4349	0.4589	0.2711	0.3655	0.4564	0.4732	0.4997
	100	0.2738	0.3403	0.3593	0.3899	0.2262	0.3138	0.4147	0.4347	0.4642
	500	0.1338	0.1783	0.192	0.2162	0.1265	0.1868	0.2842	0.3113	0.3569
	1000	0.0975	0.1306	0.1413	0.1603	0.0958	0.1422	0.2235	0.2483	0.2939
	2500	0.0626	0.0841	0.0914	0.1046	0.0625	0.0936	0.1526	0.1716	0.2089
22	50	0.1659	0.1907	0.1971	0.209	0.1327	0.1647	0.2066	0.2187	0.2428
	100	0.1452	0.1671	0.173	0.1831	0.1206	0.1508	0.1875	0.1969	0.2143
	500	0.0861	0.1115	0.1186	0.1298	0.0727	0.1022	0.1424	0.1516	0.1665
	1000	0.0663	0.088	0.0946	0.1059	0.0574	0.0817	0.1204	0.1303	0.1466
	2500	0.044	0.0591	0.0642	0.0734	0.0396	0.0566	0.0886	0.098	0.115
23	50	0.2984	0.3069	0.3111	0.3223	0.2979	0.2989	0.321	0.3359	0.3755
	100	0.2857	0.2902	0.2919	0.2952	0.2853	0.2874	0.298	0.3032	0.3218
	500	0.2394	0.253	0.256	0.2608	0.2174	0.2514	0.2699	0.2733	0.2788
	1000	0.2075	0.2341	0.2385	0.2446	0.1768	0.2257	0.2581	0.2627	0.2694
	2500	0.1477	0.1836	0.1931	0.207	0.1229	0.1631	0.2273	0.2394	0.2522
24	50	0.3135	0.3138	0.3143	0.315	0.3127	0.3141	0.3157	0.3164	0.3189
	100	0.3124	0.3136	0.3136	0.3135	0.3099	0.3126	0.3144	0.3148	0.3158
	500	0.3067	0.3099	0.3106	0.3115	0.3039	0.308	0.312	0.3126	0.3138
	1000	0.3042	0.3079	0.3089	0.3104	0.3009	0.3065	0.3114	0.3124	0.3134
	2500	0.2949	0.3016	0.303	0.3049	0.2929	0.3013	0.3066	0.3078	0.3099
27	50	0.0742	0.0745	0.0746	0.0747	0.0736	0.0745	0.075	0.0752	0.0764
	100	0.0718	0.0729	0.0731	0.0734	0.0711	0.0729	0.0739	0.074	0.0744
	500	0.0627	0.0656	0.0663	0.0674	0.0594	0.0671	0.0706	0.0711	0.0718
	1000	0.0571	0.0612	0.0621	0.0636	0.0482	0.0611	0.0686	0.0694	0.0704
	2500	0.0442	0.0519	0.0535	0.0559	0.0357	0.0449	0.0634	0.0655	0.0677

Tabelle B.13.: Simulationsergebnisse für Dichteschätzung mit Gaußkern (weitere Varianten) – L_2-Risiko für die Dichten 1-13.

Dichte	n	LIL	LIL 2.1	L2NR	L1NR	kLIL	kLIL 2.1	kL2NR	kL1NR	CvM LIL
1	50	0.1123	0.3197	0.7043	55.4636	0.1505	0.3933	0.1117	0.3122	0.1733
	100	0.0849	0.2156	0.2359	3.2733	0.1072	0.2518	0.0756	0.207	0.1381
	500	0.0393	0.0997	0.0338	0.1523	0.0441	0.1069	0.0268	0.0633	0.0837
	1000	0.0267	0.0673	0.02	0.0583	0.0293	0.0734	0.0178	0.0333	0.0629
	2500	0.0168	0.039	0.0118	0.0212	0.0178	0.0412	0.0112	0.0169	0.0437
2	50	0.0775	0.1893	0.4116	19.8054	0.1199	0.2756	0.0787	0.1766	0.1396
	100	0.0542	0.1309	0.1388	1.8643	0.0839	0.1959	0.051	0.1222	0.1082
	500	0.023	0.0526	0.0262	0.0981	0.0353	0.088	0.0197	0.0402	0.0573
	1000	0.0165	0.0348	0.017	0.0412	0.0245	0.0621	0.0138	0.0245	0.0446
	2500	0.0096	0.0203	0.0095	0.0178	0.0151	0.0372	0.0079	0.013	0.0296
3	50	0.027	0.1399	0.2981	23.5895	0.0496	0.1861	0.0463	0.1328	0.0496
	100	0.0172	0.0836	0.1119	1.5717	0.0278	0.1147	0.03	0.0987	0.0308
	500	0.0053	0.0242	0.0132	0.0769	0.0075	0.034	0.0072	0.0293	0.0092
	1000	0.0031	0.0135	0.0051	0.0286	0.0043	0.0196	0.0033	0.0146	0.0053
	2500	0.0015	0.0061	0.0018	0.0087	0.0021	0.0092	0.0014	0.0054	0.0027
4	50	0.0253	0.1033	0.2017	14.4429	0.04	0.1251	0.0289	0.0859	0.0385
	100	0.0181	0.0705	0.0637	1.0237	0.0254	0.0831	0.0177	0.0558	0.0269
	500	0.0067	0.0257	0.0079	0.0495	0.0082	0.029	0.0048	0.0173	0.0094
	1000	0.004	0.0161	0.0036	0.0183	0.0048	0.0177	0.0026	0.0093	0.0057
	2500	0.0021	0.0085	0.0015	0.0061	0.0025	0.0093	0.0012	0.0039	0.0031
5	50	0.0106	0.0579	0.1155	8.8945	0.0184	0.0727	0.0173	0.0509	0.0182
	100	0.0073	0.0369	0.0434	0.644	0.0107	0.0452	0.012	0.0385	0.0111
	500	0.002	0.011	0.0053	0.031	0.0026	0.0129	0.0029	0.0116	0.003
	1000	0.0012	0.0064	0.002	0.0115	0.0015	0.0073	0.0013	0.0058	0.0017
	2500	6e-04	0.003	7e-04	0.0035	7e-04	0.0034	5e-04	0.0022	8e-04
6	50	0.0202	0.0762	0.1498	7.1023	0.0322	0.0921	0.0219	0.0665	0.033
	100	0.0141	0.055	0.0473	0.5776	0.02	0.0647	0.0136	0.0405	0.0228
	500	0.005	0.0204	0.0057	0.0318	0.0062	0.0228	0.0035	0.0127	0.0079
	1000	0.003	0.0123	0.0025	0.0127	0.0036	0.0136	0.0018	0.0063	0.0047
	2500	0.0014	0.0059	0.001	0.0041	0.0016	0.0064	8e-04	0.0026	0.0022
7	50	0.0168	0.0824	0.1725	10.8581	0.0309	0.1113	0.0275	0.0767	0.031
	100	0.0109	0.0503	0.0657	0.8929	0.0175	0.07	0.0175	0.0577	0.0193
	500	0.0033	0.015	0.0078	0.0448	0.0046	0.0212	0.0043	0.0171	0.0056
	1000	0.0019	0.0084	0.003	0.0168	0.0026	0.0121	0.002	0.0086	0.0032
	2500	9e-04	0.0039	0.0011	0.0051	0.0013	0.0057	8e-04	0.0032	0.0016
9	50	0.0415	0.0834	0.1148	5.3322	0.0586	0.1044	0.035	0.0595	0.0707
	100	0.0297	0.0631	0.052	0.4978	0.0431	0.0841	0.0239	0.0465	0.0582
	500	0.0132	0.0286	0.0126	0.0356	0.0197	0.0437	0.01	0.0176	0.0337
	1000	0.0089	0.0193	0.0081	0.0178	0.0136	0.0307	0.0067	0.0107	0.0254
	2500	0.0053	0.0115	0.0049	0.009	0.0084	0.0195	0.0041	0.0065	0.0178
10	50	0.0188	0.0454	0.0728	2.6104	0.025	0.0503	0.0141	0.0317	0.0278
	100	0.0144	0.0363	0.0249	0.2706	0.0184	0.04	0.0095	0.0217	0.022
	500	0.0062	0.0174	0.004	0.0182	0.0071	0.0189	0.0029	0.007	0.0101
	1000	0.0041	0.0118	0.0019	0.0072	0.0046	0.0127	0.0016	0.0037	0.0069
	2500	0.0023	0.0069	9e-04	0.0025	0.0025	0.0073	8e-04	0.0016	0.0041
11	50	0.017	0.0928	0.2491	11.2064	0.0289	0.1175	0.0305	0.0895	0.0276
	100	0.0109	0.0603	0.0681	1.0776	0.0164	0.0736	0.0182	0.0615	0.017
	500	0.0031	0.0174	0.0074	0.0512	0.004	0.0203	0.0042	0.0184	0.0046
	1000	0.0019	0.0098	0.0032	0.0183	0.0022	0.0112	0.0021	0.0089	0.0026
	2500	0.001	0.0046	0.0011	0.0058	0.0011	0.0051	9e-04	0.0035	0.0013
12	50	0.0484	0.1415	0.3642	19.9617	0.0799	0.2043	0.0526	0.1346	0.0926
	100	0.0309	0.0932	0.104	1.3723	0.0498	0.1403	0.0299	0.0925	0.0647
	500	0.0108	0.032	0.0127	0.0762	0.0173	0.0537	0.0082	0.0271	0.0262
	1000	0.0064	0.0195	0.0059	0.027	0.0104	0.0347	0.0046	0.0137	0.0174
	2500	0.0033	0.0098	0.0025	0.0086	0.0055	0.0186	0.0022	0.0055	0.0098
13	50	0.0723	0.2032	0.2593	21.7966	0.1013	0.233	0.0582	0.1183	0.1294
	100	0.0507	0.1479	0.1022	1.4404	0.0639	0.1693	0.0408	0.0929	0.0924
	500	0.0248	0.0555	0.0186	0.0732	0.0274	0.0608	0.0164	0.029	0.0445
	1000	0.0185	0.0379	0.0123	0.03	0.02	0.0405	0.0116	0.0178	0.0346
	2500	0.0122	0.0244	0.0074	0.0119	0.013	0.0256	0.0073	0.0096	0.0255

Tabelle B.14.: Simulationsergebnisse für Dichteschätzung mit Gaußkern (weitere Varianten) – L_2-Risiko für die Dichten 15-28.

Dichte	n	LIL	LIL 2.1	L2NR	L1NR	kLIL	kLIL 2.1	kL2NR	kL1NR	CvM LIL
15	50	0.4589	0.8531	1.6314	90.8487	0.6299	1.149	0.4176	0.736	0.7047
	100	0.35	0.6619	0.6307	9.092	0.4919	0.8979	0.2985	0.5807	0.6091
	500	0.1704	0.321	0.1716	0.5038	0.2441	0.4976	0.1357	0.2307	0.3768
	1000	0.1232	0.2345	0.1196	0.2418	0.1819	0.3703	0.0978	0.1579	0.3041
	2500	0.0827	0.1563	0.0767	0.1391	0.1233	0.255	0.0648	0.0955	0.2344
16	50	0.0403	0.2158	0.7612	53.39	0.0662	0.2711	0.0687	0.293	0.0627
	100	0.0263	0.1403	0.166	2.3204	0.0385	0.1695	0.0406	0.1563	0.0391
	500	0.0087	0.0432	0.0168	0.1128	0.0106	0.0492	0.0099	0.0394	0.0113
	1000	0.0055	0.0251	0.0077	0.0438	0.0064	0.0284	0.0053	0.0212	0.0067
	2500	0.0029	0.0126	0.0031	0.0137	0.0033	0.0138	0.0024	0.0089	0.0033
17	50	0.0688	0.3632	0.9613	67.4055	0.1152	0.4622	0.1238	0.3715	0.1158
	100	0.046	0.2301	0.2607	4.0064	0.0679	0.2838	0.0749	0.2416	0.0756
	500	0.0154	0.0716	0.0294	0.1859	0.0188	0.0814	0.0168	0.0684	0.024
	1000	0.0095	0.0421	0.0128	0.0707	0.0109	0.0466	0.0085	0.0344	0.0142
	2500	0.0046	0.0197	0.0045	0.0232	0.0051	0.0216	0.0036	0.0145	0.007
20	50	0.0615	0.1308	0.214	9.2367	0.0913	0.1608	0.0474	0.0957	0.1123
	100	0.0448	0.0997	0.068	0.7264	0.0667	0.1319	0.0308	0.0617	0.0918
	500	0.0175	0.0433	0.0142	0.0565	0.0269	0.066	0.0117	0.0211	0.0494
	1000	0.0107	0.0282	0.0089	0.0242	0.0172	0.0458	0.0076	0.0138	0.0357
	2500	0.006	0.0157	0.0043	0.0086	0.01	0.0271	0.0038	0.0062	0.0229
21	50	0.0738	0.1989	0.1497	13.1943	0.103	0.2147	0.0345	0.0701	0.1119
	100	0.0511	0.1571	0.0569	0.6768	0.067	0.1701	0.0213	0.0512	0.0859
	500	0.016	0.0625	0.0088	0.0486	0.0193	0.0682	0.0062	0.0178	0.0378
	1000	0.0092	0.0373	0.0046	0.0193	0.0107	0.0404	0.0037	0.0094	0.0232
	2500	0.0042	0.0177	0.0018	0.006	0.0047	0.0191	0.0016	0.0038	0.0109
22	50	0.0225	0.0793	0.1849	23.5245	0.0349	0.1078	0.0286	0.0771	0.0375
	100	0.0172	0.0519	0.0724	1.0076	0.024	0.0694	0.0191	0.066	0.0283
	500	0.0073	0.0201	0.0066	0.0406	0.0105	0.0261	0.0043	0.015	0.014
	1000	0.0044	0.0134	0.0032	0.0163	0.0068	0.0184	0.0024	0.0081	0.0097
	2500	0.0021	0.0073	0.0013	0.0049	0.0034	0.0112	0.0011	0.0031	0.0053
23	50	0.0652	0.1421	0.2634	15.8634	0.0771	0.1724	0.0691	0.1186	0.0811
	100	0.0575	0.1033	0.0926	1.3617	0.0629	0.119	0.0446	0.0841	0.068
	500	0.0476	0.0607	0.0139	0.0699	0.0488	0.0633	0.012	0.0259	0.0529
	1000	0.044	0.0534	0.0075	0.0257	0.0453	0.0546	0.0074	0.0137	0.0492
	2500	0.0327	0.0477	0.0034	0.0088	0.0356	0.0482	0.0037	0.0057	0.0443
24	50	0.1043	0.1559	0.2043	14.7281	0.1241	0.1727	0.0578	0.0942	0.1291
	100	0.0863	0.1375	0.0748	1.0531	0.1034	0.1478	0.0398	0.0673	0.1124
	500	0.0503	0.1005	0.0158	0.0479	0.0561	0.108	0.017	0.0204	0.0636
	1000	0.0388	0.0808	0.0115	0.0197	0.0422	0.085	0.0127	0.0122	0.0488
	2500	0.0262	0.0575	0.0073	0.0069	0.028	0.0608	0.008	0.0058	0.0347
25	50	0.2273	0.3855	0.7362	74.7361	0.2541	0.4368	0.1698	0.3121	0.2537
	100	0.2016	0.3097	0.2507	3.59	0.2249	0.3378	0.1266	0.2267	0.2291
	500	0.0925	0.2279	0.0623	0.162	0.1052	0.2347	0.0552	0.0819	0.1879
	1000	0.0671	0.1624	0.0417	0.0779	0.0726	0.1892	0.0396	0.0533	0.1577
	2500	0.0445	0.0934	0.0255	0.0384	0.0468	0.0985	0.0251	0.0319	0.1134
26	50	0.8341	1.3506	1.3207	129.3661	1.0732	1.3575	0.3537	0.5672	1.226
	100	0.5486	1.3017	0.518	6.7115	0.6932	1.3191	0.2396	0.4504	1.1076
	500	0.2079	0.6693	0.1093	0.3292	0.235	0.7175	0.1006	0.1529	0.4797
	1000	0.1461	0.4066	0.0737	0.1446	0.1577	0.4338	0.0716	0.0961	0.3146
	2500	0.0965	0.2182	0.0465	0.0658	0.1014	0.2293	0.0467	0.0555	0.195
27	50	0.0221	0.0321	0.0514	12.1752	0.024	0.0359	0.0215	0.0243	0.0251
	100	0.0207	0.0269	0.0191	0.3062	0.0219	0.0289	0.015	0.0176	0.0234
	500	0.0182	0.0212	0.0037	0.0112	0.0184	0.0217	0.0048	0.0048	0.0204
	1000	0.0175	0.0197	0.0024	0.0046	0.0177	0.0199	0.003	0.0027	0.0193
	2500	0.0156	0.0182	0.0014	0.0016	0.0161	0.0183	0.0017	0.0012	0.018
28	50	0.3434	0.5624	0.9459	78.5397	0.4253	0.633	0.2623	0.423	0.4785
	100	0.2712	0.4786	0.3578	4.8619	0.3274	0.5118	0.1924	0.3246	0.4349
	500	0.149	0.2929	0.1065	0.2643	0.161	0.3139	0.0965	0.138	0.2941
	1000	0.1115	0.2138	0.0768	0.1409	0.1197	0.2264	0.0717	0.101	0.2366
	2500	0.077	0.1441	0.0522	0.082	0.0807	0.1501	0.0506	0.067	0.1802

Tabelle B.15.: Simulationsergebnisse für Dichteschätzung mit Gaußkern (weitere Varianten) – L_∞-Risiko für die Dichten.

Dichte	n	LIL	LIL 2.1	L2NR	L1NR	kLIL	kLIL 2.1	kL2NR	kL1NR	CvM LIL
3	50	0.1692	0.3379	1.2514	45.8771	0.2123	0.389	0.3029	0.6749	0.2116
	100	0.1457	0.2631	0.673	5.393	0.1713	0.3075	0.2451	0.6118	0.177
	500	0.1079	0.1653	0.194	0.6656	0.1213	0.1802	0.1267	0.3403	0.1288
	1000	0.0937	0.1439	0.1122	0.373	0.1053	0.1575	0.0829	0.2331	0.1126
	2500	0.0762	0.1178	0.0661	0.1892	0.086	0.1314	0.0551	0.14	0.0935
4	50	0.2166	0.3682	0.8154	24.7059	0.2633	0.3927	0.1927	0.4169	0.2606
	100	0.1961	0.325	0.3669	3.4932	0.2263	0.3436	0.1534	0.3324	0.2322
	500	0.1441	0.235	0.1126	0.3913	0.1569	0.2446	0.0865	0.1885	0.1666
	1000	0.1227	0.2021	0.0794	0.2269	0.1313	0.2087	0.0663	0.1445	0.1407
	2500	0.0993	0.1641	0.052	0.1273	0.1053	0.1692	0.0484	0.0946	0.1149
5	50	0.0649	0.153	0.4786	17.4359	0.0884	0.1696	0.1128	0.257	0.088
	100	0.0559	0.1251	0.2593	1.9841	0.0681	0.1374	0.0966	0.2363	0.0704
	500	0.0312	0.073	0.0765	0.2657	0.0356	0.0786	0.0498	0.1339	0.0388
	1000	0.0249	0.0572	0.0437	0.1485	0.0274	0.0608	0.0326	0.0919	0.0298
	2500	0.0184	0.0405	0.0262	0.0763	0.0196	0.0426	0.0218	0.0555	0.0213
6	50	0.1242	0.2364	0.6182	16.4592	0.1601	0.256	0.1328	0.3171	0.1625
	100	0.109	0.2075	0.277	2.0445	0.1313	0.2222	0.1053	0.2448	0.1401
	500	0.0707	0.1381	0.0786	0.2629	0.079	0.1449	0.0555	0.1385	0.0901
	1000	0.0567	0.1118	0.0518	0.1572	0.062	0.1167	0.0409	0.0974	0.0715
	2500	0.0397	0.0811	0.0316	0.0834	0.0429	0.0846	0.0278	0.0614	0.0508
7	50	0.0992	0.2178	0.7182	25.466	0.1361	0.25	0.1762	0.3882	0.1366
	100	0.0811	0.1743	0.3927	3.0612	0.1043	0.2029	0.142	0.356	0.1098
	500	0.0468	0.1	0.1136	0.3856	0.0559	0.118	0.0747	0.1983	0.0617
	1000	0.0365	0.0755	0.0659	0.2178	0.0428	0.0905	0.0491	0.1364	0.0472
	2500	0.0263	0.0522	0.0392	0.1104	0.0303	0.0635	0.0326	0.0822	0.0336
10	50	0.1724	0.2314	0.2839	5.9991	0.193	0.2376	0.1206	0.1635	0.2003
	100	0.1615	0.2183	0.1448	0.9177	0.1766	0.2241	0.0998	0.1323	0.1876
	500	0.1253	0.1773	0.0613	0.1399	0.1323	0.1818	0.0613	0.0756	0.1496
	1000	0.1109	0.158	0.0447	0.0826	0.1158	0.1616	0.0502	0.0555	0.1334
	2500	0.0925	0.1344	0.0383	0.0474	0.096	0.1371	0.0427	0.0383	0.1136
11	50	0.0988	0.2314	0.9971	23.6413	0.1302	0.2589	0.1917	0.4377	0.1269
	100	0.08	0.1895	0.4132	3.6759	0.099	0.2084	0.1496	0.3822	0.1015
	500	0.0443	0.106	0.1115	0.446	0.0501	0.1142	0.0752	0.2137	0.0541
	1000	0.036	0.0809	0.0715	0.2314	0.039	0.0863	0.0534	0.1437	0.0418
	2500	0.027	0.0564	0.0417	0.123	0.0285	0.0594	0.0349	0.0895	0.0303
12	50	0.3059	0.4461	1.2838	45.4605	0.3574	0.5193	0.3194	0.6391	0.376
	100	0.2684	0.3798	0.5906	5.0194	0.3045	0.4473	0.24	0.5389	0.3316
	500	0.1995	0.2638	0.1745	0.6401	0.2279	0.3116	0.1333	0.3015	0.2505
	1000	0.1737	0.2325	0.1169	0.3306	0.2032	0.2646	0.101	0.2088	0.2272
	2500	0.1419	0.202	0.0801	0.1799	0.1724	0.2284	0.0777	0.1347	0.2022
16	50	0.2861	0.6035	3.1788	104.248	0.3646	0.6646	0.449	1.4266	0.357
	100	0.2481	0.5071	1.0122	8.4092	0.2969	0.5481	0.3538	0.9726	0.2994
	500	0.1776	0.3246	0.2583	0.9501	0.1936	0.341	0.178	0.4597	0.1983
	1000	0.1518	0.2676	0.1704	0.5465	0.1632	0.2801	0.1306	0.3356	0.1659
	2500	0.1247	0.2126	0.1072	0.2908	0.1314	0.2192	0.0906	0.2188	0.1317
17	50	0.4118	0.7611	3.8462	119.1707	0.464	0.8716	0.763	1.784	0.4627
	100	0.3627	0.5881	1.6095	13.3053	0.405	0.6627	0.6297	1.5204	0.4152
	500	0.266	0.3974	0.4526	1.5841	0.2822	0.4095	0.3024	0.8101	0.302
	1000	0.2333	0.3472	0.2853	0.9189	0.2432	0.356	0.2159	0.567	0.2625
	2500	0.1893	0.2841	0.1684	0.501	0.196	0.2904	0.141	0.3683	0.2162
20	50	0.458	0.6044	0.8886	20.1674	0.5317	0.6447	0.3725	0.5022	0.5732
	100	0.42	0.557	0.4158	2.6489	0.4852	0.6083	0.3201	0.3959	0.5419
	500	0.3169	0.4305	0.2129	0.4488	0.3646	0.4938	0.2362	0.2322	0.4495
	1000	0.2795	0.3731	0.181	0.2705	0.3164	0.4386	0.2039	0.1888	0.4045
	2500	0.2473	0.3066	0.1462	0.1557	0.2734	0.369	0.1746	0.1293	0.3485
21	50	0.3643	0.507	0.631	20.2921	0.4226	0.5133	0.2172	0.3464	0.4401
	100	0.314	0.4832	0.3525	2.3598	0.359	0.4919	0.1696	0.3237	0.4049
	500	0.1882	0.3616	0.1203	0.4046	0.2084	0.3747	0.096	0.1959	0.2898
	1000	0.145	0.2907	0.0869	0.2328	0.1585	0.3013	0.0733	0.1469	0.2335
	2500	0.0997	0.2071	0.0544	0.1237	0.1067	0.2146	0.0504	0.0914	0.1646
22	50	0.1612	0.2469	0.7896	30.4811	0.1922	0.2749	0.1835	0.3965	0.1969
	100	0.1501	0.2181	0.4498	3.4583	0.1729	0.239	0.1584	0.4203	0.1833
	500	0.1045	0.166	0.0985	0.3446	0.1264	0.1816	0.0736	0.1732	0.1443
	1000	0.0829	0.1429	0.0691	0.2082	0.1049	0.1622	0.056	0.1316	0.1245
	2500	0.0582	0.11	0.0429	0.1041	0.0748	0.1332	0.0383	0.0782	0.0946
23	50	0.2988	0.3935	1.0697	31.5076	0.3078	0.4284	0.3771	0.5712	0.3124
	100	0.2875	0.3397	0.5528	4.8558	0.2919	0.3635	0.298	0.51	0.2964
	500	0.2545	0.2806	0.1869	0.587	0.2593	0.283	0.1637	0.2915	0.2706
	1000	0.2399	0.2702	0.1346	0.3159	0.2441	0.2721	0.1292	0.2066	0.26
	2500	0.2015	0.2517	0.0926	0.1804	0.2088	0.254	0.0929	0.1362	0.2355
24	50	0.3137	0.3191	0.83	34.6149	0.3139	0.3242	0.3294	0.4553	0.3152
	100	0.3132	0.3159	0.4459	3.5741	0.3136	0.3164	0.3079	0.4144	0.3143
	500	0.3105	0.3146	0.2769	0.4036	0.3113	0.3147	0.286	0.2537	0.3121
	1000	0.3093	0.3144	0.2757	0.2449	0.3102	0.3147	0.2839	0.2192	0.3118
	2500	0.3044	0.312	0.2666	0.1604	0.3051	0.3123	0.2753	0.1841	0.3073
27	50	0.0743	0.0767	0.2097	18.9079	0.0746	0.0784	0.0827	0.1083	0.0749
	100	0.0724	0.0745	0.1076	1.1084	0.0731	0.075	0.0714	0.0999	0.0738
	500	0.066	0.0713	0.0427	0.0957	0.0671	0.0716	0.0449	0.0518	0.0707
	1000	0.0624	0.0694	0.035	0.0578	0.0635	0.0697	0.0374	0.0395	0.0689
	2500	0.0551	0.0656	0.0277	0.0313	0.0563	0.0661	0.0293	0.0264	0.0649

Tabelle B.16.: Simulationsergebnisse für reguläre Histogramme – L_2-Risiko für die Dichten 1-13.

Dichte	n	BR	BIC	L2CV	SC	E-LR	V	KS .5	KS .9	KS .95	KS .99
1	50	0.0551	0.0475	0.1255	0.4201	0.164	0.1254	0.0746	0.0495	0.0466	**0.0458**
	100	0.0265	0.0235	0.0566	0.158	0.0729	0.0653	0.0356	0.0247	0.0219	**0.0216**
	500	0.0051	0.0044	0.0126	0.0044	0.0096	0.011	0.0068	0.0048	0.0044	**0.0043**
	1000	0.003	**0.0023**	0.0066	**0.0023**	0.0049	0.0059	0.0037	0.0026	0.0024	**0.0023**
	2500	0.0011	**8e-04**	0.0021	**8e-04**	0.0015	0.0023	0.0014	9e-04	**8e-04**	**8e-04**
2	50	0.0861	0.0861	0.1269	0.0902	0.0991	0.0854	**0.0691**	0.0781	0.0839	0.0947
	100	0.0552	0.059	0.0652	0.0525	0.0445	0.0439	**0.0425**	0.0525	0.0571	0.0661
	500	0.0206	0.0281	0.0176	0.0242	**0.013**	0.0132	0.0149	0.0207	0.0231	0.0281
	1000	0.0135	0.0197	0.0102	0.0173	0.0093	**0.0086**	0.0103	0.0146	0.0162	0.0196
	2500	0.0071	0.0133	0.0051	0.0115	0.0058	**0.0048**	0.0061	0.0089	0.01	0.0121
3	50	0.0561	0.0574	0.0796	0.145	0.0678	0.0562	**0.0495**	0.0555	0.0572	0.0654
	100	0.0408	0.0437	0.0498	0.047	0.0351	**0.0345**	0.0358	0.0457	0.0493	0.0526
	500	0.0144	0.0172	0.0159	0.015	0.0133	**0.0125**	0.0164	0.0239	0.0268	0.0328
	1000	**0.0086**	0.0109	0.0089	0.0094	0.0097	0.0089	0.0115	0.0168	0.0187	0.0228
	2500	**0.0046**	0.0062	0.0054	0.005	0.0065	0.0053	0.007	0.0101	0.0111	0.0136
4	50	0.0402	0.0399	0.0521	0.0421	0.0368	**0.0337**	0.0357	0.047	0.0518	0.0732
	100	0.0306	0.032	0.0294	0.0283	**0.0221**	0.0223	0.0256	0.0357	0.0368	0.0436
	500	0.013	0.0161	0.0101	0.0138	0.0096	**0.0092**	0.0123	0.0175	0.0196	0.0232
	1000	0.0079	0.0121	**0.0062**	0.0101	0.0073	0.0064	0.0088	0.0129	0.0146	0.017
	2500	0.0047	0.0082	**0.0033**	0.0067	0.0052	0.004	0.0057	0.0081	0.0092	0.0111
5	50	0.0217	0.0222	0.0324	0.0386	0.0233	0.0204	**0.0197**	0.0276	0.0315	0.0416
	100	0.0153	0.0161	0.0209	0.0153	0.0147	**0.0144**	0.0156	0.0182	0.0195	0.0231
	500	0.0058	0.0073	0.0063	0.0061	0.0055	**0.0051**	0.0068	0.0101	0.0114	0.0137
	1000	**0.0034**	0.005	**0.0034**	0.0042	0.004	0.0035	0.0047	0.0072	0.0081	0.0098
	2500	0.002	0.003	**0.0018**	0.0026	0.0028	0.0022	0.0031	0.0047	0.0054	0.0065
6	50	0.0654	0.0691	0.0387	0.0542	0.0375	**0.0368**	0.0386	0.0497	0.0539	0.0624
	100	0.0611	0.0705	0.0311	0.0553	0.0307	**0.0305**	0.0328	0.0404	0.0435	0.0515
	500	0.0546	0.0776	**0.0224**	0.0588	0.0269	0.0266	0.0279	0.0304	0.0318	0.034
	1000	0.0495	0.0796	**0.0189**	0.0574	0.0238	0.0235	0.0242	0.0255	0.0264	0.0276
	2500	0.0495	0.0848	**0.0175**	0.0592	0.0224	0.0222	0.0226	0.0234	0.0237	0.0246
7	50	0.0369	0.0375	0.0504	0.0657	0.0374	**0.0305**	0.0309	0.0364	0.0391	0.0442
	100	0.0275	0.0281	0.0312	0.028	0.0204	**0.0201**	0.0233	0.0326	0.0351	0.0394
	500	0.0084	0.0096	0.0092	0.0087	0.0086	**0.008**	0.0105	0.0145	0.0162	0.0192
	1000	0.0054	0.0069	**0.0053**	0.006	0.006	0.0054	0.0072	0.0113	0.0124	0.014
	2500	**0.0029**	0.0042	**0.0029**	0.0036	0.004	0.0033	0.0044	0.0064	0.0075	0.0094
9	50	0.1158	0.1169	0.1072	0.1103	**0.1069**	**0.1069**	**0.1069**	0.1071	0.1073	0.108
	100	0.1182	0.12	0.113	0.1152	**0.1129**	**0.1129**	**0.1129**	**0.1129**	**0.1129**	**0.1129**
	500	0.1231	0.1242	**0.1224**	0.1225	**0.1224**	**0.1224**	**0.1224**	**0.1224**	**0.1224**	**0.1224**
	1000	0.1235	0.1245	**0.1231**	0.1232	**0.1231**	**0.1231**	**0.1231**	**0.1231**	**0.1231**	**0.1231**
	2500	0.1243	0.1248	**0.1241**	**0.1241**	**0.1241**	**0.1241**	**0.1241**	**0.1241**	**0.1241**	**0.1241**
10	50	0.056	0.0569	**0.0513**	0.0534	0.0518	0.0518	0.0517	0.0525	0.0527	0.0531
	100	0.0583	0.0595	**0.0553**	0.0566	0.0559	0.0559	0.0559	0.056	0.056	0.056
	500	0.061	0.0619	**0.0604**	0.0607	0.0606	0.0606	0.0606	0.0606	0.0606	0.0606
	1000	0.0618	0.0622	**0.0616**	**0.0616**	**0.0616**	**0.0616**	**0.0616**	**0.0616**	**0.0616**	**0.0616**
	2500	**0.0621**	0.0624	**0.0621**	**0.0621**	**0.0621**	**0.0621**	**0.0621**	**0.0621**	**0.0621**	**0.0621**
11	50	0.0393	0.0401	0.0618	0.0919	0.0434	0.0371	**0.0359**	0.051	0.0556	0.0648
	100	0.0242	0.0251	0.0328	0.0285	0.023	**0.022**	0.0239	0.0294	0.0333	0.0467
	500	0.0085	0.0098	0.01	0.0085	0.0086	**0.0083**	0.0103	0.0157	0.018	0.0231
	1000	**0.0052**	0.0068	0.0055	0.0059	0.0065	0.0056	0.0076	0.0115	0.0128	0.0153
	2500	**0.0028**	0.0041	**0.0028**	0.0033	0.0045	0.0034	0.0049	0.0074	0.0083	0.0105
12	50	0.0677	0.0677	0.0692	0.0658	0.0564	0.0481	**0.0475**	0.0538	0.0594	0.0734
	100	0.0435	0.0499	0.0433	0.0407	0.0301	**0.0295**	0.0296	0.0373	0.041	0.0501
	500	0.026	0.0345	0.0207	0.0317	0.0169	**0.0163**	0.0196	0.0248	0.0263	0.0293
	1000	0.0234	0.03	0.0135	0.028	0.0127	**0.0115**	0.0147	0.0213	0.0237	0.0261
	2500	0.022	0.0261	0.0071	0.0254	0.0081	**0.0067**	0.0086	0.0125	0.014	0.0185
13	50	0.0816	0.0812	0.0956	**0.0676**	0.0751	0.0715	0.0769	0.1182	0.14	0.1699
	100	0.0446	0.047	0.0584	**0.0407**	0.0498	0.0503	0.0487	0.0754	0.088	0.1155
	500	0.0217	0.0262	**0.0176**	0.0206	0.0207	0.0191	0.0257	0.0259	0.0264	0.027
	1000	**0.0079**	0.0171	0.0086	0.0088	0.0092	0.0095	0.0137	0.0247	0.0248	0.0249
	2500	0.0032	0.0032	0.0035	**0.0031**	0.0032	0.0048	**0.0031**	0.0127	0.0211	0.0245

Tabelle B.17.: Simulationsergebnisse für reguläre Histogramme – L_2-Risiko für die Dichten 15-28.

Dichte	n	BR	BIC	L2CV	SC	E-LR	V	KS .5	KS .9	KS .95	KS .99
15	50	0.4587	0.4616	0.5688	0.5523	0.5213	0.4622	**0.4129**	0.445	0.4635	0.5039
	100	0.3447	0.361	0.3898	0.3485	0.321	0.2985	**0.2873**	0.3301	0.3525	0.3955
	500	0.1538	0.185	0.1304	0.153	0.109	**0.107**	0.118	0.1479	0.1616	0.1859
	1000	0.104	0.1372	0.0848	0.1126	0.0762	**0.0728**	0.0813	0.1061	0.1162	0.1357
	2500	0.0663	0.0966	0.0514	0.0789	0.0495	**0.0456**	0.0519	0.0688	0.0754	0.0879
16	50	0.0885	0.093	0.1179	0.2655	0.0955	0.0785	**0.0768**	0.1015	0.1037	0.1057
	100	0.0553	0.0632	0.0713	0.099	0.0531	**0.0524**	**0.0524**	0.0692	0.0867	0.1083
	500	0.0189	0.021	0.0221	**0.0187**	0.02	0.0192	0.0252	0.0385	0.0437	0.048
	1000	**0.0119**	0.0141	0.0134	0.0121	0.0147	0.0131	0.0182	0.0261	0.0325	0.039
	2500	**0.0065**	0.0075	0.0072	0.0066	0.0109	0.0082	0.012	0.019	0.0208	0.0224
17	50	0.1444	0.1275	0.2097	0.512	0.2006	0.1666	0.1291	0.1233	0.1185	**0.115**
	100	0.1085	0.117	0.1401	0.2258	**0.0935**	**0.0935**	0.0942	0.1189	0.1245	0.1254
	500	0.0336	0.037	0.0365	0.0338	0.0343	**0.0323**	0.0418	0.0603	0.0673	0.0828
	1000	**0.0218**	0.0241	0.0238	0.0222	0.0264	0.0238	0.0316	0.047	0.051	0.0563
	2500	0.0116	0.013	0.013	**0.0115**	0.0181	0.0145	0.0195	0.0286	0.0333	0.0379
20	50	0.1803	0.1813	0.1744	0.1767	**0.1742**	**0.1742**	**0.1742**	**0.1742**	0.1744	0.1745
	100	0.1819	0.1838	**0.1781**	0.1788	**0.1781**	**0.1781**	**0.1781**	**0.1781**	**0.1781**	**0.1781**
	500	0.1857	0.1867	**0.185**	0.1851	**0.185**	**0.185**	**0.185**	**0.185**	**0.185**	**0.185**
	1000	0.186	0.187	**0.1856**	0.1857	**0.1856**	**0.1856**	**0.1856**	**0.1856**	**0.1856**	**0.1856**
	2500	0.1869	0.1873	**0.1868**	**0.1868**	**0.1868**	**0.1868**	**0.1868**	**0.1868**	**0.1868**	**0.1868**
21	50	0.0943	0.0999	0.0597	0.0712	**0.0518**	0.0538	0.0645	0.0907	0.1004	0.1176
	100	0.0535	0.0658	0.0448	0.0502	**0.0387**	0.04	0.0494	0.0663	0.0729	0.0865
	500	0.0312	0.0437	**0.0166**	0.0395	0.0175	0.0167	0.0215	0.0373	0.0432	0.0523
	1000	0.0159	0.032	**0.0105**	0.026	0.0137	0.0121	0.0148	0.0217	0.0255	0.0337
	2500	0.0092	0.0149	**0.0053**	0.0142	0.0086	0.0072	0.0092	0.0136	0.0147	0.0174
22	50	0.0408	0.0415	0.0578	0.0821	0.044	0.0383	**0.0353**	0.0405	0.0427	0.0438
	100	0.0294	0.0305	0.0368	0.0543	0.025	**0.0244**	0.0287	0.0371	0.0404	0.0451
	500	0.0103	0.0116	0.0105	0.0104	0.01	**0.0095**	0.0111	0.0156	0.0174	0.0214
	1000	**0.0063**	0.0083	0.0066	0.0068	0.0073	0.0068	0.0081	0.0112	0.0125	0.0147
	2500	**0.0034**	0.0049	**0.0034**	0.004	0.005	0.004	0.0053	0.0072	0.0079	0.009
23	50	0.0886	0.0888	0.1329	0.1676	0.0937	0.0899	**0.0843**	0.1012	0.107	0.115
	100	**0.0657**	0.068	0.0849	0.0771	0.0665	0.0671	0.0698	0.0854	0.0895	0.0983
	500	0.0301	0.0458	**0.0296**	0.0315	0.0372	0.0321	0.0469	0.0525	0.0534	0.0593
	1000	0.0199	0.0319	**0.018**	0.0226	0.0255	0.0221	0.0363	0.0463	0.0489	0.0512
	2500	0.0115	0.0182	**0.0096**	0.0153	0.0186	0.0166	0.0185	0.0358	0.0394	0.0427
24	50	0.0953	0.0955	0.1146	0.1265	0.0867	**0.0857**	0.0954	0.1079	0.1094	0.1112
	100	0.0657	0.0702	0.0741	0.0829	**0.0653**	0.0668	0.0791	0.098	0.1023	0.1063
	500	0.026	0.0321	0.0271	**0.0254**	0.0374	0.0343	0.0443	0.0581	0.0614	0.0671
	1000	**0.0181**	0.0228	0.0196	0.0185	0.0292	0.0256	0.0337	0.0442	0.0464	0.0576
	2500	**0.0105**	0.015	0.0106	0.0119	0.0223	0.0166	0.0232	0.0325	0.034	0.0393
25	50	0.2672	0.2682	0.2794	0.3898	0.2265	**0.226**	0.2587	0.2802	0.2828	0.284
	100	0.1592	0.1831	0.1827	0.1945	**0.1576**	0.1699	0.2009	0.2535	0.2694	0.2859
	500	0.064	0.0702	0.0679	**0.0626**	0.0793	0.0757	0.0902	0.1272	0.1385	0.1812
	1000	0.0467	0.0576	**0.0447**	0.0491	0.0598	0.0549	0.064	0.0934	0.1018	0.1155
	2500	0.0265	0.0453	**0.0244**	0.0314	0.0428	0.035	0.0457	0.0564	0.0581	0.0748
26	50	1.2114	1.2114	1.1628	1.1852	**1.1565**	**1.1565**	**1.1565**	**1.1565**	1.158	1.1759
	100	1.1439	1.2022	**0.9922**	1.0934	0.9955	0.9955	0.9955	0.9955	0.9955	0.9972
	500	0.5747	1.0057	**0.0402**	0.4147	0.1158	0.1117	0.1585	0.2574	0.2913	0.3511
	1000	**0.023**	0.7977	0.0242	0.028	0.0852	0.0723	0.109	0.1773	0.2016	0.2463
	2500	**0.0075**	0.2596	0.0106	**0.0075**	0.058	0.0409	0.0657	0.1111	0.1262	0.1547
27	50	0.0189	0.0184	0.0292	0.0453	0.0258	0.0234	0.02	0.0185	0.0183	**0.0181**
	100	0.0176	0.017	0.0224	0.0326	0.0203	0.0198	0.0179	0.017	**0.0168**	**0.0168**
	500	**0.008**	0.013	0.0084	0.0081	0.016	0.0146	0.0165	0.0163	0.0162	0.0162
	1000	**0.0051**	0.0069	0.0054	0.0052	0.0131	0.0103	0.016	0.0163	0.0163	0.0163
	2500	**0.0028**	0.0038	**0.0028**	**0.0028**	0.0094	0.0075	0.0102	0.0163	0.0164	0.0164
28	50	0.2722	0.2685	0.4011	0.6725	0.3644	0.3212	0.2754	0.2539	0.2525	**0.2471**
	100	0.2118	0.2163	0.2468	0.4073	0.212	0.2012	**0.193**	0.2128	0.2144	0.2151
	500	0.0961	0.112	0.0927	0.0933	0.0791	**0.0787**	0.0879	0.1155	0.1245	0.1414
	1000	0.0645	0.0795	0.0626	0.0664	0.0558	**0.0536**	0.0614	0.0829	0.0914	0.1058
	2500	0.0409	0.0555	0.0377	0.0448	0.0397	**0.0347**	0.0418	0.0568	0.0621	0.0727

Tabelle B.18.: Simulationsergebnisse für reguläre Histogramme – L_∞-Risiko für die Dichten 1-14.

Dichte	n	BR	BIC	L2CV	SC	E-LR	V	KS .5	KS .9	KS .95	KS .99
3	50	0.3361	0.3375	0.3829	0.5268	0.3441	0.3208	**0.32**	0.3283	0.3306	0.3252
	100	0.316	0.324	0.3259	0.3232	0.2791	**0.2734**	0.2976	0.3367	0.3488	0.3576
	500	0.2104	0.2383	0.2022	0.215	0.2104	**0.2007**	0.2393	0.2821	0.2955	0.3192
	1000	0.1672	0.1967	**0.1606**	0.178	0.188	0.1769	0.2069	0.2473	0.2588	0.2803
	2500	**0.123**	0.1526	0.1251	0.1358	0.162	0.1426	0.1675	0.2012	0.2114	0.231
4	50	0.2185	0.2183	0.267	0.229	0.222	**0.2111**	0.2132	0.235	0.2421	0.2804
	100	0.195	0.1996	0.2077	0.195	0.1843	**0.1815**	0.1856	0.2124	0.2144	0.2283
	500	0.1449	0.153	0.1424	0.1477	0.1321	**0.1314**	0.1405	0.1561	0.1627	0.173
	1000	0.1196	0.1347	0.1173	0.128	0.1185	**0.1133**	0.1237	0.1358	0.1401	0.1493
	2500	0.1024	0.1199	**0.094**	0.1123	0.1053	0.0963	0.107	0.1167	0.1209	0.1268
5	50	0.1199	0.12	0.1514	0.1631	0.1296	0.1211	**0.1137**	0.1192	0.1243	0.1315
	100	**0.1069**	0.1087	0.1284	0.1089	0.1098	0.1083	0.1078	0.1095	0.1107	0.1146
	500	0.0727	0.0794	0.0766	0.0742	0.0731	**0.0702**	0.0767	0.0904	0.0949	0.1005
	1000	**0.0585**	0.0666	0.0595	0.0632	0.0621	0.0596	0.0658	0.0783	0.0816	0.0889
	2500	0.0456	0.0536	**0.0453**	0.0507	0.053	0.0481	0.0551	0.0628	0.0661	0.0712
6	50	0.1968	0.2021	0.1772	0.1816	0.1669	0.1651	**0.1623**	0.1739	0.1796	0.192
	100	0.1914	0.2059	0.1591	0.1828	0.1527	**0.1502**	0.1517	0.1618	0.1656	0.1788
	500	0.1826	0.2179	**0.1292**	0.1902	0.1387	0.1377	0.1388	0.1396	0.1421	0.1444
	1000	0.1746	0.2225	**0.1172**	0.1895	0.1304	0.1279	0.1327	0.1285	0.1307	0.1329
	2500	0.1753	0.2304	**0.111**	0.1932	0.1239	0.1213	0.1251	0.1309	0.131	0.1268
7	50	0.2087	0.2065	0.2338	0.2606	0.1986	**0.1842**	0.1922	0.1993	0.1964	0.1957
	100	0.1901	0.1928	0.1983	0.1886	**0.163**	0.1634	0.1755	0.1954	0.2006	0.2032
	500	0.1185	0.1281	0.121	0.1218	0.1203	**0.1178**	0.1306	0.1589	0.1668	0.1774
	1000	0.0997	0.1147	**0.095**	0.1066	0.1074	0.1009	0.1153	0.136	0.1449	0.1562
	2500	0.076	0.0939	**0.073**	0.086	0.0907	0.083	0.0938	0.1138	0.1189	0.1266
10	50	0.2406	0.2421	**0.2316**	0.2363	0.2328	0.2328	0.2326	0.235	0.2356	0.2367
	100	0.2448	0.2463	**0.2401**	0.2424	0.241	0.241	0.2409	0.2411	0.2412	0.2414
	500	0.2483	0.2493	**0.2475**	0.2478	0.2476	0.2476	0.2476	0.2476	0.2476	0.2476
	1000	0.2492	0.2497	**0.249**	**0.249**	**0.249**	**0.249**	**0.249**	**0.249**	**0.249**	**0.249**
	2500	0.2496	0.2499	**0.2495**	**0.2495**	**0.2495**	**0.2495**	**0.2495**	**0.2495**	**0.2495**	**0.2495**
11	50	0.2032	0.2041	0.2672	0.3352	0.2213	0.2033	**0.1964**	0.2065	0.2101	0.2147
	100	0.1699	0.1718	0.1987	0.1789	0.1665	**0.1623**	0.1672	0.178	0.1834	0.1964
	500	0.112	0.1173	0.1221	0.1115	0.112	**0.1103**	0.119	0.1385	0.1479	0.1639
	1000	**0.0921**	0.1007	0.0954	0.0958	0.0985	0.0935	0.1038	0.1192	0.1236	0.1352
	2500	**0.0698**	0.0784	0.0705	0.0735	0.083	0.0741	0.0852	0.0995	0.1034	0.1139
12	50	0.307	0.3066	0.3933	0.3302	0.3715	0.3388	0.3036	**0.2874**	0.2914	0.3035
	100	0.2993	0.3017	0.3847	0.3042	0.3467	0.3429	0.3142	0.2877	0.2829	**0.2819**
	500	0.411	0.3732	**0.3361**	0.3844	0.3949	0.378	0.421	0.4063	0.3972	0.3829
	1000	0.4568	0.4176	**0.2861**	0.4281	0.3687	0.3419	0.399	0.4444	0.4404	0.4291
	2500	0.4874	0.4577	**0.2168**	0.4635	0.3143	0.2791	0.328	0.3986	0.4187	0.4488
16	50	0.4589	0.4671	0.5485	0.916	0.4958	0.451	**0.4335**	0.4782	0.4796	0.4843
	100	0.3855	0.3958	0.4508	0.5209	0.3864	0.3825	**0.3781**	0.4044	0.4369	0.4712
	500	0.242	0.2478	0.2731	0.2429	0.242	**0.2419**	0.2604	0.2988	0.3214	0.344
	1000	**0.2002**	0.207	0.2216	0.2005	0.2068	0.2015	0.2211	0.2476	0.2654	0.2856
	2500	0.1609	**0.1556**	0.1705	0.1564	0.1737	0.1612	0.1787	0.2088	0.2183	0.2253
17	50	0.8927	0.8689	1.0038	1.7467	0.9587	0.8762	**0.8358**	0.8884	0.8754	0.8762
	100	0.7906	0.8397	0.8428	1.1358	**0.6983**	0.7145	0.7535	0.8627	0.885	0.9076
	500	0.4621	0.4882	0.4717	0.4623	0.4768	**0.4574**	0.5424	0.6454	0.6721	0.7425
	1000	**0.3686**	0.4022	0.3807	0.374	0.437	0.4072	0.4767	0.576	0.6004	0.6281
	2500	0.2826	0.3061	0.3005	**0.282**	0.3729	0.3309	0.3877	0.4638	0.4967	0.5268
20	50	0.6613	0.6629	**0.6523**	0.6561	**0.6523**	**0.6523**	**0.6523**	**0.6523**	0.6526	0.6528
	100	0.6652	0.6676	**0.66**	0.6609	**0.66**	**0.66**	**0.66**	**0.66**	**0.66**	**0.66**
	500	0.6701	0.6713	**0.6694**	0.6695	**0.6694**	**0.6694**	**0.6694**	**0.6694**	**0.6694**	**0.6694**
	1000	0.6705	0.6716	**0.67**	0.6701	**0.67**	**0.67**	**0.67**	**0.67**	**0.67**	**0.67**
	2500	0.6715	0.672	**0.6714**	**0.6714**	**0.6714**	**0.6714**	**0.6714**	**0.6714**	**0.6714**	**0.6714**
21	50	0.3541	0.3625	0.3456	0.3335	**0.3213**	0.3221	0.3279	0.3482	0.3628	0.3942
	100	0.3036	0.3144	0.3212	0.3038	**0.2938**	0.2956	0.3042	0.3186	0.327	0.3473
	500	0.2708	0.2752	**0.2185**	0.2728	0.2469	0.2435	0.2635	0.2902	0.2911	0.2946
	1000	0.2411	0.269	**0.1824**	0.2657	0.2326	0.2147	0.2392	0.2732	0.2835	0.2944
	2500	0.194	0.2438	**0.1336**	0.2351	0.1881	0.1711	0.1934	0.2295	0.2403	0.2598
22	50	0.2271	0.2271	0.2783	0.3272	0.2456	0.2315	0.2245	0.206	0.2069	**0.2043**
	100	0.224	0.2221	0.2414	0.2856	0.2071	**0.2035**	0.2251	0.2245	0.2223	0.22
	500	0.1668	0.1801	**0.1504**	0.1677	0.1677	0.1634	0.1728	0.195	0.2048	0.2251
	1000	0.1329	0.1634	**0.1246**	0.1438	0.1537	0.1463	0.1594	0.1767	0.1801	0.194
	2500	0.1043	0.1312	**0.0962**	0.1172	0.1319	0.1198	0.1354	0.1551	0.1626	0.1688
23	50	0.3575	**0.3568**	0.5361	0.6063	0.4179	0.3939	0.3598	0.3658	0.3645	0.3584
	100	**0.3307**	0.3337	0.4555	0.3948	0.3612	0.3573	0.3541	0.3652	0.3698	0.3769
	500	0.2837	**0.2772**	0.3115	0.2798	0.2888	0.2851	0.292	0.3179	0.317	0.3241
	1000	0.2527	0.2542	**0.2498**	0.2536	0.2615	0.2557	0.2729	0.2833	0.2998	0.3126
	2500	0.199	0.2336	**0.1905**	0.2211	0.2329	0.2184	0.234	0.2639	0.2607	0.2697
24	50	0.2817	0.2761	0.3901	0.4373	0.3258	0.2993	0.2789	0.2448	0.2355	**0.2246**
	100	0.2956	0.2886	0.3512	0.3709	0.2897	0.2815	0.2818	0.2562	0.253	**0.2399**
	500	0.272	0.2607	0.2846	0.2708	0.2732	0.2686	**0.2543**	0.2562	0.2624	0.2668
	1000	0.2712	0.2462	0.2683	0.2685	0.2576	0.2479	0.2704	0.242	**0.2372**	0.2516
	2500	0.2366	0.2671	**0.2324**	0.2612	0.2434	0.2502	0.244	0.2647	0.2723	0.267
27	50	0.0564	0.0547	0.0948	0.1466	0.0886	0.0784	0.0624	0.0549	0.0542	**0.0535**
	100	0.0566	0.0533	0.0999	0.1296	0.0752	0.0724	0.0598	0.054	0.0528	**0.0523**
	500	0.0628	0.0578	0.0674	0.0663	0.0637	0.0656	0.0574	0.052	0.0516	**0.051**
	1000	0.0514	0.0528	0.055	0.0519	0.0613	0.0584	0.0585	0.0518	0.0512	**0.0508**
	2500	0.0401	**0.0399**	0.0415	**0.0399**	0.0538	0.0486	0.0552	0.0529	0.0518	0.0508

Tabelle B.19.: Simulationsergebnisse für reguläre Histogramme (alternative GoF-Tests) – L_2-Risiko für die Dichten 1-14.

Dichte	n	Kuip .5	Kuip .9	Kuip .95	Kuip .99	CvM V	CvM .5	CvM .9	CvM .95	CvM .99
1	50	0.0818	0.0491	0.0483	0.0458	0.1279	0.0702	0.048	0.0471	0.0458
	100	0.0401	0.0276	0.025	0.0216	0.0606	0.0349	0.0232	0.0223	0.0212
	500	0.0079	0.0047	0.0044	0.0042	0.0108	0.0066	0.0045	0.0043	0.0043
	1000	0.0041	0.0025	0.0024	0.0023	0.0058	0.0036	0.0026	0.0026	0.0023
	2500	0.0015	9e-04	9e-04	8e-04	0.0021	0.0013	9e-04	9e-04	8e-04
2	50	0.0724	0.086	0.0927	0.1024	0.0768	0.0655	0.076	0.0798	0.0928
	100	0.0475	0.0598	0.0654	0.0764	0.041	0.0416	0.0522	0.0566	0.0649
	500	0.0176	0.0255	0.028	0.0317	0.0126	0.0153	0.0219	0.0244	0.0293
	1000	0.0127	0.018	0.0195	0.0226	0.0086	0.0107	0.0157	0.0175	0.0211
	2500	0.0079	0.0111	0.0122	0.0143	0.0049	0.0064	0.0097	0.0109	0.0135
3	50	0.0486	0.0558	0.0597	0.0656	0.054	0.0481	0.0547	0.057	0.0652
	100	0.0345	0.0394	0.0414	0.0463	0.0348	0.0353	0.0472	0.0496	0.0525
	500	0.0143	0.0187	0.0201	0.0225	0.0133	0.0179	0.026	0.029	0.0352
	1000	0.0098	0.0131	0.0143	0.0162	0.0093	0.0127	0.0204	0.0226	0.0269
	2500	0.006	0.0083	0.0092	0.01	0.0058	0.0084	0.0139	0.0155	0.019
4	50	0.0334	0.0387	0.0416	0.0439	0.0328	0.0349	0.0443	0.0489	0.0672
	100	0.0244	0.0299	0.0315	0.0355	0.0218	0.0264	0.0363	0.0377	0.0433
	500	0.0103	0.013	0.0141	0.0161	0.0093	0.012	0.0179	0.0201	0.0249
	1000	0.007	0.0095	0.0101	0.0114	0.0064	0.0087	0.0131	0.0149	0.0177
	2500	0.0045	0.006	0.0064	0.0072	0.0041	0.0057	0.0085	0.0097	0.012
5	50	0.0189	0.0204	0.0235	0.0265	0.0204	0.02	0.0266	0.0303	0.0425
	100	0.0146	0.0159	0.0168	0.0171	0.0135	0.0152	0.0182	0.0193	0.0222
	500	0.0057	0.0074	0.008	0.0089	0.0049	0.0067	0.0108	0.012	0.0141
	1000	0.0039	0.0051	0.0057	0.0064	0.0035	0.0047	0.0078	0.0086	0.0109
	2500	0.0024	0.0033	0.0036	0.0042	0.0021	0.0031	0.005	0.0057	0.007
6	50	0.0373	0.0436	0.0463	0.0524	0.0362	0.0386	0.0504	0.0542	0.0638
	100	0.0319	0.035	0.0365	0.0401	0.0305	0.0329	0.0416	0.045	0.0537
	500	0.027	0.0281	0.0283	0.0295	0.0267	0.0277	0.0304	0.0324	0.0353
	1000	0.0237	0.0244	0.0247	0.0247	0.0235	0.0242	0.0256	0.0261	0.0277
	2500	0.0222	0.0226	0.0228	0.0231	0.0222	0.0226	0.0235	0.024	0.0251
7	50	0.0307	0.035	0.0366	0.0412	0.0314	0.0314	0.0375	0.0389	0.0434
	100	0.0213	0.0257	0.0278	0.0321	0.0206	0.0232	0.034	0.0372	0.0403
	500	0.009	0.0113	0.0122	0.0139	0.008	0.0108	0.0177	0.0192	0.0242
	1000	0.006	0.008	0.0087	0.0104	0.0055	0.0076	0.013	0.0153	0.0183
	2500	0.0037	0.0051	0.0054	0.0062	0.0034	0.005	0.0079	0.0088	0.0113
9	50	0.107	0.1075	0.1079	0.1084	0.1069	0.1069	0.1073	0.1075	0.1081
	100	0.1129	0.1129	0.1129	0.113	0.1129	0.1129	0.1129	0.1129	0.113
	500	0.1224	0.1224	0.1224	0.1224	0.1224	0.1224	0.1224	0.1224	0.1224
	1000	0.1231	0.1231	0.1231	0.1231	0.1231	0.1231	0.1231	0.1231	0.1231
	2500	0.1241	0.1241	0.1241	0.1241	0.1241	0.1241	0.1241	0.1241	0.1241
10	50	0.0518	0.0519	0.0521	0.0524	0.0518	0.0518	0.0525	0.0527	0.0534
	100	0.0559	0.0559	0.0559	0.056	0.0559	0.0559	0.056	0.0559	0.0559
	500	0.0606	0.0606	0.0606	0.0606	0.0606	0.0606	0.0606	0.0606	0.0606
	1000	0.0616	0.0616	0.0616	0.0616	0.0616	0.0616	0.0616	0.0616	0.0616
	2500	0.0621	0.0621	0.0621	0.0621	0.0621	0.0621	0.0621	0.0621	0.0621
11	50	0.0359	0.0373	0.0393	0.0459	0.0371	0.0342	0.0438	0.0545	0.0657
	100	0.0223	0.0263	0.027	0.0291	0.0207	0.0238	0.0296	0.0313	0.0394
	500	0.009	0.0115	0.0129	0.014	0.0081	0.0105	0.0176	0.0199	0.025
	1000	0.0062	0.0082	0.0091	0.0103	0.0054	0.0078	0.0118	0.0134	0.0181
	2500	0.0039	0.0053	0.0057	0.0066	0.0034	0.005	0.0078	0.0089	0.0111
12	50	0.052	0.0636	0.0686	0.0813	0.0475	0.0442	0.054	0.0589	0.0725
	100	0.0322	0.0425	0.0479	0.0565	0.0281	0.0297	0.0385	0.0426	0.0505
	500	0.0179	0.0222	0.024	0.0269	0.0175	0.0201	0.0253	0.0271	0.031
	1000	0.0139	0.0172	0.0185	0.0209	0.0138	0.0172	0.0227	0.024	0.027
	2500	0.0087	0.0118	0.0129	0.0146	0.0097	0.013	0.018	0.0195	0.0227
13	50	0.0669	0.0924	0.0967	0.1132	0.0697	0.0829	0.1448	0.1567	0.1906
	100	0.0454	0.0581	0.0631	0.0748	0.0482	0.0545	0.097	0.1176	0.1475
	500	0.0218	0.0253	0.0258	0.0259	0.023	0.0258	0.0265	0.0325	0.068
	1000	0.0074	0.0186	0.0231	0.0246	0.0131	0.0227	0.0248	0.0249	0.0249
	2500	0.0031	0.003	0.0032	0.0063	0.0037	0.0178	0.0237	0.0246	0.0249

Tabelle B.20.: Simulationsergebnisse für reguläre Histogramme (alternative GoF-Tests) – L_2-Risiko für die Dichten 15-28.

Dichte	n	Kuip .5	Kuip .9	Kuip .95	Kuip .99	CvM V	CvM .5	CvM .9	CvM .95	CvM .99
15	50	0.4338	0.4749	0.4901	0.527	0.432	0.4034	0.4452	0.4571	0.5008
	100	0.3023	0.3663	0.39	0.4293	0.2873	0.2833	0.334	0.3559	0.398
	500	0.1313	0.1674	0.1816	0.2068	0.1093	0.1281	0.1646	0.1782	0.2055
	1000	0.0907	0.1222	0.1306	0.1496	0.0772	0.0922	0.125	0.1363	0.1591
	2500	0.0591	0.0777	0.0842	0.0972	0.0515	0.0637	0.0881	0.0954	0.112
16	50	0.0722	0.0848	0.0895	0.0915	0.0765	0.0733	0.0967	0.1044	0.1057
	100	0.0504	0.0539	0.0546	0.0609	0.0488	0.0508	0.0659	0.0776	0.1064
	500	0.0205	0.0266	0.0287	0.0336	0.0187	0.024	0.0415	0.0468	0.0491
	1000	0.0147	0.02	0.0211	0.0228	0.0131	0.0174	0.0243	0.0287	0.0432
	2500	0.0091	0.0125	0.0135	0.016	0.0082	0.0115	0.0197	0.0214	0.0222
17	50	0.1401	0.123	0.1213	0.1171	0.1532	0.1284	0.1196	0.1174	0.1159
	100	0.0888	0.1029	0.1056	0.115	0.091	0.0943	0.1175	0.1218	0.1264
	500	0.0353	0.0456	0.0489	0.0536	0.0334	0.0445	0.0662	0.0772	0.0846
	1000	0.0255	0.0328	0.0358	0.0405	0.0239	0.0337	0.0518	0.0541	0.0681
	2500	0.0157	0.0204	0.0224	0.025	0.015	0.0218	0.0359	0.0375	0.0525
20	50	0.1742	0.1745	0.1745	0.1746	0.1742	0.1742	0.1743	0.1744	0.1745
	100	0.1781	0.1781	0.1781	0.1781	0.1781	0.1781	0.1781	0.1781	0.1781
	500	0.185	0.185	0.185	0.185	0.185	0.185	0.185	0.185	0.185
	1000	0.1856	0.1856	0.1856	0.1856	0.1856	0.1856	0.1856	0.1856	0.1856
	2500	0.1868	0.1868	0.1868	0.1868	0.1868	0.1868	0.1868	0.1868	0.1868
21	50	0.0653	0.0829	0.0916	0.1043	0.0544	0.071	0.1071	0.1181	0.1354
	100	0.0454	0.0609	0.0652	0.0747	0.0413	0.0546	0.0816	0.0907	0.1107
	500	0.018	0.0239	0.0267	0.0323	0.0185	0.0275	0.0483	0.0542	0.0659
	1000	0.0133	0.0159	0.0168	0.0191	0.0141	0.0184	0.0334	0.04	0.0523
	2500	0.0077	0.0097	0.011	0.0129	0.0086	0.0131	0.0182	0.0207	0.0277
22	50	0.0371	0.0398	0.0404	0.0434	0.0383	0.0359	0.0406	0.0417	0.0438
	100	0.027	0.0311	0.0327	0.0371	0.0252	0.0289	0.0401	0.0436	0.0466
	500	0.0103	0.0118	0.0128	0.0149	0.0098	0.012	0.0203	0.0226	0.0278
	1000	0.0074	0.0087	0.0092	0.0104	0.0071	0.009	0.0138	0.0159	0.0193
	2500	0.0045	0.0061	0.0064	0.007	0.0046	0.0064	0.0092	0.0103	0.0127
23	50	0.085	0.0899	0.0939	0.1016	0.0857	0.0866	0.1013	0.1059	0.118
	100	0.0674	0.0731	0.0748	0.0808	0.067	0.0719	0.0896	0.0924	0.1013
	500	0.0393	0.0489	0.0498	0.0521	0.0327	0.0475	0.059	0.0609	0.0712
	1000	0.023	0.0404	0.0439	0.0461	0.0217	0.0372	0.0504	0.0528	0.0627
	2500	0.0172	0.0193	0.0218	0.0307	0.016	0.0194	0.0386	0.0423	0.0496
24	50	0.0914	0.1034	0.1067	0.11	0.0856	0.0957	0.1103	0.1111	0.1104
	100	0.072	0.0864	0.0914	0.0994	0.0648	0.0793	0.1019	0.1055	0.1087
	500	0.0379	0.0448	0.0455	0.0518	0.0359	0.0447	0.0638	0.0685	0.0759
	1000	0.0276	0.0357	0.0378	0.0418	0.0279	0.0381	0.0492	0.0559	0.0659
	2500	0.0187	0.0239	0.0251	0.028	0.0213	0.0275	0.0405	0.0405	0.0461
25	50	0.2467	0.2696	0.2748	0.281	0.2312	0.2592	0.2806	0.2844	0.2839
	100	0.1809	0.2231	0.2384	0.2507	0.1779	0.2194	0.2547	0.2632	0.2901
	500	0.0821	0.0976	0.1041	0.1169	0.0873	0.1126	0.1925	0.2129	0.2253
	1000	0.059	0.0674	0.0733	0.0904	0.0771	0.076	0.1433	0.1636	0.2063
	2500	0.0386	0.0505	0.053	0.0564	0.0527	0.0703	0.0768	0.1009	0.1237
26	50	1.1565	1.1565	1.1565	1.1565	1.1565	1.1565	1.1659	1.1846	1.2655
	100	0.9955	0.9955	0.9955	0.9955	0.9955	0.9955	0.9975	1.0019	1.0645
	500	0.1218	0.1752	0.1914	0.2215	0.1185	0.1928	0.332	0.3802	0.4802
	1000	0.0803	0.1177	0.13	0.1519	0.0813	0.1326	0.2321	0.2662	0.3379
	2500	0.0485	0.0719	0.0795	0.0935	0.0475	0.0814	0.1446	0.1668	0.2122
27	50	0.0209	0.0186	0.0183	0.0181	0.023	0.0199	0.0184	0.0183	0.0181
	100	0.0183	0.0172	0.017	0.0168	0.0194	0.0177	0.0169	0.0168	0.0168
	500	0.0164	0.0164	0.0164	0.0163	0.0135	0.0165	0.0162	0.0162	0.0162
	1000	0.0112	0.0164	0.0163	0.0163	0.0097	0.0162	0.0163	0.0163	0.0163
	2500	0.008	0.01	0.012	0.0159	0.0069	0.0097	0.0164	0.0164	0.0164
28	50	0.266	0.2567	0.2546	0.2485	0.2935	0.2618	0.2583	0.2533	0.2491
	100	0.1922	0.1967	0.1979	0.205	0.1888	0.1905	0.2103	0.2136	0.2148
	500	0.0803	0.0931	0.0981	0.1089	0.0793	0.0929	0.1194	0.1264	0.1419
	1000	0.056	0.0653	0.0697	0.0767	0.0572	0.0689	0.0916	0.0984	0.1127
	2500	0.0372	0.0444	0.0466	0.0521	0.0396	0.0485	0.0661	0.0721	0.0826

Tabelle B.21.: Simulationsergebnisse für reguläre Histogramme (alternative GoF-Tests) – L_∞-Risiko für die Dichten.

Dichte	n	Kuip .5	Kuip .9	Kuip .95	Kuip .99	CvM V	CvM .5	CvM .9	CvM .95	CvM .99
3	50	0.3147	0.3227	0.3242	0.3299	0.3164	0.3135	0.3325	0.3312	0.3237
	100	0.283	0.3154	0.325	0.3379	0.2796	0.2965	0.3414	0.35	0.3553
	500	0.2204	0.256	0.2641	0.2776	0.2066	0.2497	0.2941	0.3088	0.3313
	1000	0.191	0.2232	0.2309	0.2453	0.1822	0.2196	0.2691	0.2808	0.3022
	2500	0.1554	0.1834	0.1914	0.2	0.1512	0.1854	0.2338	0.2445	0.2684
4	50	0.2081	0.2173	0.2235	0.228	0.2091	0.2104	0.2285	0.2376	0.2691
	100	0.1849	0.1954	0.1981	0.2094	0.1802	0.1874	0.2127	0.2171	0.2283
	500	0.1342	0.1383	0.1432	0.1508	0.1302	0.1372	0.1569	0.1626	0.1788
	1000	0.1166	0.1261	0.1271	0.1299	0.1144	0.1225	0.134	0.1412	0.1512
	2500	0.1006	0.1093	0.1099	0.1139	0.0968	0.1072	0.1173	0.1226	0.1286
5	50	0.1144	0.1134	0.117	0.1188	0.121	0.114	0.1181	0.1217	0.1331
	100	0.1069	0.1076	0.1089	0.1087	0.1052	0.1068	0.1094	0.1102	0.1132
	500	0.0736	0.0798	0.0823	0.0865	0.0684	0.0766	0.0943	0.0966	0.101
	1000	0.0617	0.0682	0.071	0.0737	0.0595	0.0661	0.0815	0.085	0.0931
	2500	0.0494	0.0569	0.0583	0.06	0.0466	0.0552	0.0645	0.068	0.0755
6	50	0.1644	0.1685	0.171	0.1762	0.1635	0.1643	0.1743	0.1811	0.1945
	100	0.1512	0.1561	0.1578	0.1601	0.1489	0.1517	0.1623	0.1671	0.1834
	500	0.1391	0.1374	0.1365	0.1389	0.1382	0.1372	0.1402	0.1433	0.1476
	1000	0.129	0.1329	0.1327	0.1302	0.1284	0.1325	0.1287	0.1303	0.1332
	2500	0.1223	0.1255	0.1267	0.1288	0.1217	0.1267	0.1303	0.1286	0.1265
7	50	0.1882	0.1974	0.1965	0.1999	0.1871	0.1909	0.1975	0.197	0.1973
	100	0.1691	0.1876	0.1914	0.1966	0.1647	0.1798	0.1982	0.2019	0.2059
	500	0.1217	0.134	0.1405	0.1525	0.1174	0.1306	0.1735	0.1787	0.1866
	1000	0.1086	0.1193	0.1236	0.1308	0.1023	0.1182	0.1477	0.1642	0.1775
	2500	0.0865	0.1014	0.1054	0.1125	0.0837	0.1008	0.1227	0.1269	0.1379
10	50	0.2329	0.2333	0.2337	0.2347	0.2328	0.2328	0.2351	0.2356	0.2374
	100	0.241	0.2409	0.2409	0.241	0.241	0.2409	0.2411	0.2411	0.2412
	500	0.2476	0.2476	0.2476	0.2476	0.2476	0.2476	0.2476	0.2476	0.2476
	1000	0.249	0.249	0.249	0.249	0.249	0.249	0.249	0.249	0.249
	2500	0.2495	0.2495	0.2495	0.2495	0.2495	0.2495	0.2495	0.2495	0.2495
11	50	0.1978	0.1967	0.1995	0.2036	0.2029	0.1938	0.2024	0.2119	0.2136
	100	0.1648	0.1733	0.1751	0.1783	0.1592	0.1663	0.1786	0.1805	0.1895
	500	0.1145	0.1228	0.1286	0.133	0.109	0.1186	0.1473	0.1544	0.1691
	1000	0.097	0.107	0.111	0.1152	0.0912	0.1048	0.1204	0.1279	0.1473
	2500	0.0776	0.0878	0.0912	0.0953	0.0748	0.0852	0.1012	0.1062	0.1168
12	50	0.306	0.2973	0.3013	0.3136	0.3357	0.2941	0.2874	0.2915	0.3022
	100	0.3079	0.2843	0.2822	0.2833	0.3391	0.3115	0.2846	0.2801	0.2806
	500	0.4166	0.4237	0.4139	0.3968	0.4102	0.4357	0.4032	0.3924	0.3742
	1000	0.3949	0.4425	0.4535	0.46	0.3945	0.4453	0.4523	0.4416	0.4238
	2500	0.3306	0.3907	0.4059	0.4289	0.3504	0.4083	0.4692	0.4813	0.4773
16	50	0.4327	0.4437	0.4532	0.4573	0.4429	0.4261	0.4677	0.4795	0.4841
	100	0.373	0.383	0.381	0.3903	0.3669	0.3727	0.3985	0.42	0.4704
	500	0.2421	0.2624	0.269	0.2807	0.2377	0.2559	0.3146	0.3389	0.349
	1000	0.2067	0.2275	0.2328	0.2394	0.2031	0.216	0.2415	0.2538	0.308
	2500	0.1628	0.1804	0.184	0.1937	0.1594	0.1744	0.2115	0.222	0.2252
17	50	0.8411	0.8523	0.8704	0.8723	0.8492	0.8311	0.8712	0.877	0.8768
	100	0.7218	0.8059	0.8196	0.8569	0.7013	0.7592	0.8568	0.8744	0.9119
	500	0.4893	0.5658	0.5842	0.6192	0.465	0.559	0.6663	0.7151	0.7544
	1000	0.4296	0.4909	0.5122	0.5405	0.4133	0.4959	0.6037	0.6209	0.6743
	2500	0.345	0.3981	0.4149	0.4377	0.3372	0.4084	0.5151	0.5256	0.6087
20	50	0.6523	0.6527	0.6528	0.653	0.6523	0.6523	0.6524	0.6525	0.6528
	100	0.66	0.66	0.66	0.66	0.66	0.66	0.66	0.66	0.66
	500	0.6694	0.6694	0.6694	0.6694	0.6694	0.6694	0.6694	0.6694	0.6694
	1000	0.67	0.67	0.67	0.67	0.67	0.67	0.67	0.67	0.67
	2500	0.6714	0.6714	0.6714	0.6714	0.6714	0.6714	0.6714	0.6714	0.6714
21	50	0.3309	0.3388	0.3453	0.3722	0.3243	0.3336	0.379	0.4037	0.4423
	100	0.3055	0.3143	0.3158	0.3274	0.2993	0.3082	0.34	0.3555	0.3887
	500	0.2509	0.2731	0.2808	0.289	0.2517	0.2834	0.2925	0.2971	0.3172
	1000	0.226	0.2458	0.2521	0.2644	0.2341	0.2605	0.2978	0.3027	0.306
	2500	0.1776	0.1991	0.2081	0.222	0.1901	0.228	0.2636	0.2749	0.2945
22	50	0.2233	0.2135	0.2105	0.2072	0.2336	0.2207	0.2102	0.2063	0.2052
	100	0.2209	0.2241	0.2259	0.2218	0.2074	0.2264	0.2245	0.2242	0.2181
	500	0.1708	0.1732	0.1786	0.188	0.1674	0.1751	0.2148	0.2247	0.2352
	1000	0.1554	0.167	0.1693	0.1717	0.1541	0.1729	0.1869	0.1992	0.2154
	2500	0.1259	0.147	0.151	0.1552	0.1272	0.1545	0.176	0.1791	0.1791
23	50	0.3692	0.3636	0.3632	0.3584	0.3857	0.3642	0.3632	0.3647	0.3558
	100	0.3473	0.3552	0.3567	0.3602	0.3515	0.3509	0.3748	0.3764	0.3782
	500	0.2906	0.2963	0.3052	0.3193	0.2867	0.2941	0.3302	0.3287	0.3707
	1000	0.2578	0.2697	0.2702	0.2861	0.254	0.2778	0.3047	0.3107	0.3465
	2500	0.2214	0.2379	0.2419	0.2533	0.2171	0.2381	0.2538	0.2629	0.3163
24	50	0.2891	0.2615	0.2479	0.2298	0.3014	0.2751	0.2321	0.2261	0.2178
	100	0.2837	0.2766	0.2717	0.2589	0.2769	0.2796	0.2522	0.2416	0.2273
	500	0.2711	0.242	0.2375	0.2458	0.2715	0.2375	0.2656	0.2641	0.2552
	1000	0.2524	0.2723	0.2681	0.2451	0.2508	0.2653	0.24	0.2547	0.2614
	2500	0.2453	0.2452	0.247	0.2531	0.2515	0.2505	0.2534	0.2312	0.2319
27	50	0.0656	0.0558	0.0544	0.0535	0.0776	0.0619	0.0546	0.0543	0.0535
	100	0.0631	0.055	0.0538	0.0527	0.0698	0.0583	0.0534	0.0528	0.0524
	500	0.0623	0.0542	0.053	0.0517	0.0678	0.0551	0.0517	0.0512	0.0509
	1000	0.0602	0.056	0.0542	0.0518	0.0581	0.0585	0.051	0.0509	0.0508
	2500	0.0498	0.0557	0.057	0.0547	0.0467	0.0551	0.0515	0.0508	0.0505

Tabelle B.22.: Simulationsergebnisse für reguläre Histogramme (alternative GoF-Tests) – arithmetisches Mittel der Anzahl der gewählten Bins für Dichten 1-14.

Dichte	n	Kuip .5	Kuip .9	Kuip .95	Kuip .99	CvM V	CvM .5	CvM .9	CvM .95	CvM .99
1	50	2.13	1.07	1.04	1	3.69	1.75	1.06	1.02	1
	100	2.01	1.2	1.1	1.01	3.6	1.69	1.07	1.03	1
	500	2.02	1.13	1.04	1.03	3.38	1.66	1.1	1.04	1.02
	1000	1.98	1.09	1.06	1.01	3.37	1.76	1.11	1.08	1.01
	2500	1.97	1.09	1.03	1.01	3.26	1.64	1.08	1.04	1.02
2	50	5.93	3.94	3.55	3.12	10.42	6.64	4.34	4.01	3.35
	100	7.08	5.03	4.66	4.02	11.51	7.93	5.6	5.14	4.51
	500	13.2	9.74	9.12	8.32	19.23	14.58	10.57	9.78	8.64
	1000	16.53	12.53	11.86	10.77	23.9	18.25	13.52	12.6	11.11
	2500	22.67	17.95	16.83	15.29	32.24	25.13	19.03	17.63	15.55
3	50	3.43	2.16	1.93	1.6	4.83	2.87	1.87	1.75	1.51
	100	4.91	3.42	3.1	2.57	5.69	3.93	2.43	2.18	2
	500	8.23	6.48	6.15	5.69	9.14	6.73	5.17	4.86	4.2
	1000	10.1	8.1	7.68	7.09	10.75	8.29	6.17	5.82	5.26
	2500	13.45	10.87	10.28	9.74	13.85	10.76	8.03	7.57	6.78
4	50	5.61	3.96	3.62	3.33	7.12	4.97	3.37	3.13	2.52
	100	7.61	5.18	4.74	4.23	9.29	6.22	4.09	3.84	3.47
	500	16.39	12.38	11.58	10.23	18.13	13.52	9.41	8.5	7.22
	1000	22.71	17.45	16.21	14.66	24.73	18.24	12.73	11.49	9.98
	2500	33.98	26.33	24.86	22.45	36.31	27.21	19.56	17.75	15.07
5	50	4.21	3.34	3.03	2.7	5.5	3.75	2.74	2.46	1.73
	100	5.68	4.19	3.85	3.62	7.28	4.76	3.43	3.3	3.09
	500	11.04	8.81	8.23	7.64	12.42	9.35	6.41	5.76	4.82
	1000	14.18	11.54	10.82	10	15.52	12.05	8.76	8.12	6.75
	2500	19.86	16.16	15.33	14.04	21.75	16.7	12.54	11.72	10.41
6	50	28.55	23.41	22.12	20.04	32.21	26.49	20.09	18.68	15.74
	100	61.05	52.57	50.09	46.28	66.73	56.45	43.86	41.51	35.84
	500	387.11	328.57	309.27	289.37	400.31	341.11	276.31	262.5	240.34
	1000	842.55	763.38	739.15	678.45	854.13	765.28	605.86	572.74	521.35
	2500	2275.71	2175.89	2142.17	2089.41	2296.33	2160.22	1873.36	1753.32	1579.89
7	50	4.4	2.62	2.47	2.14	5.89	3.63	2.27	2.09	1.85
	100	6.27	4.7	4.28	3.58	7.35	5.22	3.3	2.76	2.33
	500	10.8	8.72	8.24	7.55	12.06	9.14	6.57	6.34	5.81
	1000	13.88	11.3	10.63	9.43	15.08	11.64	8.12	7.35	6.74
	2500	19.32	15.75	15.04	13.95	20.35	15.82	11.96	11.28	9.59
8	50	8.35	3.93	3.32	2.38	14.32	7.6	3.61	3.13	2.33
	100	13.59	6.4	5.25	3.66	21.84	11.24	5.27	4.43	3.31
	500	41.83	21.32	17.78	13.26	52.07	28.35	13.71	11.41	8.45
	1000	73.06	39.24	32.67	24.58	81.1	45.05	21.46	17.81	13.08
	2500	158.54	85.72	72.66	54.7	144.53	79.88	39.5	32.92	23.93
9	50	49.94	49.19	48.72	48.16	50	50	49.36	49.1	48.45
	100	100	100	99.92	99.72	100	100	100	99.85	99.69
	500	500	500	500	500	500	500	500	500	500
	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
	2500	2500	2500	2500	2500	2500	2500	2500	2500	2500
10	50	49.57	48.76	48.38	47.72	49.65	49.16	47.43	46.87	44.65
	100	100	99.8	99.73	99.57	100	99.8	99.5	99.34	98.36
	500	500	500	500	500	500	500	500	500	500
	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
	2500	2500	2500	2500	2500	2500	2500	2500	2500	2500
11	50	3.77	2.82	2.65	2.23	4.99	3.32	2.34	1.93	1.41
	100	4.87	3.53	3.34	3.09	6	4.06	3.01	2.93	2.63
	500	8.35	6.7	6.2	5.83	9.48	7.12	4.98	4.54	3.56
	1000	10.33	8.33	7.77	7.2	11.39	8.69	6.6	6.12	5.02
	2500	13.46	10.94	10.39	9.59	14.81	11.33	8.65	8.07	7.12
12	50	9.27	6.11	5.73	4.97	13.67	9.37	6.9	6.36	5.42
	100	12.39	8.92	8.2	7.38	17.4	12.19	9.51	8.88	7.84
	500	52.16	27.79	22.86	19.05	59.3	33.04	19.48	18.42	16.82
	1000	92.04	64.3	56.3	40.6	92.37	64.35	27.66	24.79	22.1
	2500	164.67	126.49	117.5	103.29	150.34	115.73	78.6	67.31	41.2
13	50	9.64	7.05	6.57	5.64	12.25	8.25	4.61	3.91	2.39
	100	11.54	8.55	8.03	7.39	13.65	9.6	6.55	5.6	4.22
	500	16.53	10.61	9.93	9.13	16.87	10.76	9.18	8.86	7.3
	1000	20.39	13.67	10.91	9.65	18.67	11.55	9.1	9.04	9.02
	2500	21.02	20.27	20.02	18.47	20.93	13.25	10.03	9.2	9
14	50	46.23	34.18	29.6	20.07	46.03	30.13	7.5	5.14	2.58
	100	100	96.6	94.28	82.9	99.01	85.94	22.57	13.45	6.62
	500	500	500	500	500	500	500	464.64	360.91	153.73
	1000	1000	1000	1000	1000	1000	1000	1000	989.59	733.66
	2500	2500	2500	2500	2500	2500	2500	2500	2500	2500

Tabelle B.23.: Simulationsergebnisse für reguläre Histogramme (alternative GoF-Tests) – arithmetisches Mittel der Anzahl der gewählten Bins für Dichten 15-28.

Dichte	n	Kuip .5	Kuip .9	Kuip .95	Kuip .99	CvM V	CvM .5	CvM .9	CvM .95	CvM .99
15	50	4.62	2.77	2.55	2.14	8.13	5.04	3.01	2.8	2.31
	100	6.06	3.61	3.2	2.71	9.85	6.33	3.93	3.5	2.93
	500	11.09	7.43	6.65	5.6	15.7	10.88	7.31	6.55	5.49
	1000	15.14	9.9	9.14	7.7	20.11	14.17	9.28	8.36	6.96
	2500	22.79	15.4	13.93	11.77	27.46	19.45	12.92	11.79	9.85
16	50	3.21	2.15	1.92	1.61	4.31	2.79	1.61	1.26	1.07
	100	3.94	3.09	2.96	2.76	4.91	3.35	2.62	2.29	1.43
	500	6.21	4.9	4.67	4.32	6.93	5.17	3.69	3.23	3.01
	1000	7.3	5.74	5.45	5.07	8.1	6.32	4.87	4.6	3.62
	2500	9.07	7.32	6.97	6.42	9.97	7.72	5.62	5.19	4.99
17	50	2.52	1.51	1.27	1.11	3.78	2.17	1.21	1.04	1.01
	100	3.58	2.42	2.16	1.64	4.4	3.01	1.76	1.44	1.05
	500	5.79	4.57	4.33	3.99	6.31	4.68	3.58	3.23	3
	1000	6.74	5.5	5.18	4.8	7.24	5.43	4.13	3.98	3.57
	2500	8.56	7.1	6.72	6.27	8.82	6.84	5.07	4.94	4.11
18	50	25.22	12.58	10.37	7.87	38.64	26.78	13.52	11.41	8.11
	100	52.27	25.78	21.39	15.18	75.74	50.86	23.93	19.77	14.51
	500	272.85	139.03	114.85	86.45	341.05	192.11	93.51	77.82	55.39
	1000	525.13	280.3	236.6	177.15	587.1	329.46	161.27	133.34	97.69
	2500	1325.7	716	607.53	459.15	1228.06	678.48	326.93	272.91	198.82
19	50	49.21	43.12	39.25	32.22	49.38	43.77	22.21	17.18	10.65
	100	99.93	95.44	91.2	80.64	99.98	94.43	55.04	42.5	25.94
	500	500	500	500	500	500	500	401.44	323.44	207.54
	1000	1000	1000	1000	1000	1000	1000	883.17	748.65	474.65
	2500	2500	2500	2500	2500	2500	2500	2387.15	2121.82	1402.27
20	50	50	49.81	49.78	49.67	50	50	49.95	49.9	49.76
	100	100	100	100	100	100	100	100	100	100
	500	500	500	500	500	500	500	500	500	500
	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
	2500	2500	2500	2500	2500	2500	2500	2500	2500	2500
21	50	20.06	14.85	13.15	10.61	25.03	17.63	10.32	8.37	5.71
	100	24.38	17.96	16.88	14.75	28.36	19.36	13.5	12.13	9.58
	500	49.53	42.6	39.86	34.3	48.6	41.51	23.11	19.53	15.46
	1000	62.25	47.04	46.07	44.47	53.59	44.83	37.89	33.66	24.3
	2500	88.36	74.93	68.23	54.49	81.28	48.74	41.19	40.2	38.12
22	50	3.39	2.01	1.77	1.46	4.89	2.93	1.78	1.57	1.34
	100	4.86	3.4	3.15	2.69	6.02	4.05	2.56	2.23	1.79
	500	8.44	6.51	6.17	5.5	9.42	6.59	4.88	4.65	4.27
	1000	11.14	8.74	8.15	7.08	11.68	8.62	5.86	5.45	4.91
	2500	16.72	12.58	11.69	10.56	16.56	11.9	8.6	7.96	6.28
23	50	4.03	2.94	2.62	2.07	5.67	3.49	2.08	1.87	1.34
	100	5.77	4.01	3.82	3.42	7.37	4.31	3.07	2.92	2.58
	500	17.1	8.98	8.19	7.02	19.9	9.59	5.49	4.94	3.89
	1000	23.76	15.71	12.78	10.28	24.73	18.67	7.45	6.8	5.67
	2500	32.87	26.5	25.62	22.89	34.87	26.46	16.49	12.15	8.39
24	50	4.84	2.74	2.23	1.41	7.08	3.78	1.5	1.33	1.12
	100	7.19	4.66	4.01	2.77	8.79	5.68	2.41	1.77	1.28
	500	13.58	9.74	9.3	8.55	14.29	9.66	6.65	6.08	5.4
	1000	18.92	14.03	13.15	10.98	19.06	13.14	9	7.95	6.33
	2500	26.29	19.88	19.59	18.66	23.62	19.07	12.36	10.64	9.43
25	50	4.69	2.32	1.85	1.38	6.46	3.28	1.42	1.16	1.04
	100	7.96	4.49	3.76	2.97	8.5	5	2.78	2.41	1.4
	500	13.82	10.4	9.83	9.1	12.49	9.3	5.8	4.41	3.68
	1000	19.63	12.5	11.65	10.19	15.77	10.84	7.92	7.28	4.81
	2500	32.04	22.15	19.13	13.53	21.17	14.13	10.3	9.24	8.5
26	50	50	50	50	50	50	50	45.34	38.62	17.19
	100	100	100	100	100	100	100	99.08	97.45	74.42
	500	375.66	356.72	351.15	341.21	378.82	350.87	304.71	288.88	256
	1000	390.31	370.55	366.59	359.41	389.79	365.68	333.29	322.18	298.88
	2500	393.58	381.84	379.42	374.96	393.9	378.81	358.38	351.29	336.61
27	50	2.51	1.24	1.11	1.02	4.62	2.04	1.1	1.08	1.02
	100	3.15	1.43	1.23	1.09	4.94	1.98	1.18	1.08	1.03
	500	8.23	2.19	1.73	1.25	16.11	2.65	1.23	1.07	1.01
	1000	17.82	4.72	2.92	1.61	19.98	7.36	1.17	1.08	1.04
	2500	24.51	19.66	16.27	6.68	27.2	19.9	1.86	1.29	1.04
28	50	4.49	2.46	2.12	1.49	6.41	3.6	1.6	1.28	1.06
	100	5.8	3.54	3.05	2.4	6.97	4.38	2.23	1.87	1.25
	500	10.26	6.99	6.32	5.25	10.26	6.84	4.37	4	3.29
	1000	13.85	9.7	8.79	7.45	12.88	8.83	5.71	5.16	4.24
	2500	19.75	14.12	12.96	11.06	16.96	12.13	7.98	7.12	5.97

Tabelle B.24.: Simulationsergebnisse für Splines: L_2-Risiko und gewählte $\lambda \cdot 10^4$ für Ruppert & Carroll.

StN		3			7			10	
n	50	100	1000	50	100	1000	50	100	1000
CV	0.0107	0.0057	9e-04	0.0053	0.0021	2e-04	0.0047	0.0016	1e-04
	(0.0174)	(0.0083)	(0.0014)	(0.0027)	(0.0019)	(3e-04)	(0.0021)	(0.0015)	(2e-04)
GCV	0.0106	0.0058	9e-04	0.0054	0.0019	2e-04	0.0048	0.0013	1e-04
	(0.0163)	(0.0073)	(0.0014)	(0.0036)	(0.0013)	(3e-04)	(0.0028)	(8e-04)	(2e-04)
Cp	0.0123	0.0061	9e-04	0.0077	0.0027	2e-04	0.007	0.0019	1e-04
	(0.0687)	(0.0145)	(0.0014)	(0.0252)	(0.0042)	(4e-04)	(0.0203)	(0.0023)	(2e-04)
Chi .5	0.0205	0.0105	0.0014	0.0139	0.0056	4e-04	0.0129	0.004	3e-04
	(0.8513)	(0.1464)	(0.0107)	(0.2055)	(0.0297)	(0.0024)	(0.1565)	(0.014)	(0.0015)
Chi .9	0.026	0.0131	0.0017	0.0171	0.0067	5e-04	0.0158	0.0048	3e-04
	(2.3908)	(0.3171)	(0.0193)	(0.4271)	(0.0495)	(0.0033)	(0.3121)	(0.0218)	(0.0019)
Chi .95	0.0278	0.0139	0.0018	0.0181	0.007	5e-04	0.0167	0.005	3e-04
	(3.2166)	(0.3931)	(0.0225)	(0.5262)	(0.057)	(0.0035)	(0.38)	(0.0247)	(0.002)
Chi .99	0.0315	0.0155	0.0021	0.0202	0.0077	6e-04	0.0186	0.0055	3e-04
	(5.7307)	(0.5857)	(0.0296)	(0.7802)	(0.0745)	(0.0041)	(0.5518)	(0.0313)	(0.0023)
MR .5	0.0166	0.0079	0.001	0.0097	0.0035	2e-04	0.0083	0.0023	1e-04
	(0.3597)	(0.0513)	(0.0032)	(0.0718)	(0.0086)	(7e-04)	(0.0419)	(0.0039)	(4e-04)
MR .9	0.024	0.0105	0.0011	0.0144	0.0049	3e-04	0.0131	0.0033	1e-04
	(1.5663)	(0.1494)	(0.0058)	(0.2669)	(0.0215)	(9e-04)	(0.1904)	(0.0086)	(5e-04)
MR .95	0.0263	0.0115	0.0012	0.0162	0.0054	3e-04	0.0148	0.0036	1e-04
	(2.2735)	(0.2037)	(0.0069)	(0.3844)	(0.0285)	(0.0011)	(0.2714)	(0.0111)	(6e-04)
MR .99	0.0317	0.0138	0.0013	0.02	0.0066	3e-04	0.0184	0.0045	2e-04
	(5.1341)	(0.3764)	(0.0097)	(0.7603)	(0.0487)	(0.0013)	(0.5478)	(0.0194)	(7e-04)
MR 2	0.0167	0.0078	0.001	0.0098	0.0034	2e-04	0.0084	0.0023	1e-04
	(0.3713)	(0.0494)	(0.0028)	(0.0738)	(0.0083)	(6e-04)	(0.0432)	(0.0038)	(4e-04)
MR 2.3	0.0186	0.0086	0.001	0.0111	0.0039	3e-04	0.0098	0.0026	1e-04
	(0.5703)	(0.0726)	(0.004)	(0.1168)	(0.0121)	(8e-04)	(0.0753)	(0.0051)	(5e-04)
MR 2.5	0.02	0.0092	0.0011	0.012	0.0042	3e-04	0.0107	0.0028	1e-04
	(0.7642)	(0.0914)	(0.0049)	(0.152)	(0.0148)	(9e-04)	(0.1018)	(0.006)	(5e-04)
MR 3	0.0238	0.0108	0.0012	0.0143	0.005	3e-04	0.013	0.0033	2e-04
	(1.511)	(0.1623)	(0.0073)	(0.2601)	(0.023)	(0.0011)	(0.1853)	(0.0091)	(6e-04)
Cp*	0.01	0.0052	9e-04	0.0051	0.0015	2e-04	0.0045	9e-04	1e-04
	(0.0057)	(0.0024)	(0.0014)	(8e-04)	(2e-04)	(3e-04)	(4e-04)	(1e-04)	(2e-04)
Chi .5*	0.0111	0.0061	0.0013	0.0056	0.0019	3e-04	0.0048	0.0011	2e-04
	(0.0353)	(0.014)	(0.0086)	(0.0052)	(0.0014)	(0.0016)	(0.0028)	(6e-04)	(9e-04)
Chi .9*	0.0124	0.0069	0.0016	0.0059	0.0021	4e-04	0.0049	0.0012	2e-04
	(0.065)	(0.0264)	(0.016)	(0.0069)	(0.0019)	(0.0022)	(0.0035)	(8e-04)	(0.0012)
Chi .95*	0.0128	0.0072	0.0017	0.0059	0.0021	4e-04	0.005	0.0012	2e-04
	(0.0777)	(0.0316)	(0.0187)	(0.0075)	(0.002)	(0.0024)	(0.0037)	(8e-04)	(0.0013)
Chi .99*	0.0137	0.0079	0.002	0.0061	0.0023	5e-04	0.0051	0.0013	3e-04
	(0.1092)	(0.0443)	(0.0249)	(0.0088)	(0.0024)	(0.0029)	(0.0042)	(9e-04)	(0.0015)
MR .5*	0.0105	0.0057	0.001	0.0053	0.0016	2e-04	0.0046	9e-04	1e-04
	(0.0236)	(0.0102)	(0.0031)	(0.0026)	(7e-04)	(6e-04)	(0.0015)	(3e-04)	(4e-04)
MR .9*	0.0119	0.0068	0.0011	0.0054	0.0018	3e-04	0.0047	0.001	1e-04
	(0.0587)	(0.0255)	(0.0057)	(0.0038)	(0.0011)	(9e-04)	(0.002)	(4e-04)	(5e-04)
MR .95*	0.0127	0.0073	0.0012	0.0055	0.0019	3e-04	0.0047	0.001	1e-04
	(0.0847)	(0.0347)	(0.0067)	(0.0043)	(0.0013)	(0.001)	(0.0022)	(5e-04)	(5e-04)
MR .99*	0.0146	0.0085	0.0013	0.0057	0.0021	3e-04	0.0048	0.0011	2e-04
	(0.1721)	(0.0628)	(0.0093)	(0.0055)	(0.002)	(0.0012)	(0.0027)	(6e-04)	(6e-04)
MR 2*	0.0105	0.0057	0.001	0.0053	0.0016	2e-04	0.0046	9e-04	1e-04
	(0.0241)	(0.0098)	(0.0026)	(0.0026)	(6e-04)	(6e-04)	(0.0015)	(3e-04)	(3e-04)
MR 2.3*	0.0109	0.006	0.001	0.0053	0.0017	2e-04	0.0046	9e-04	1e-04
	(0.032)	(0.0143)	(0.0039)	(0.003)	(8e-04)	(7e-04)	(0.0017)	(3e-04)	(4e-04)
MR 2.5*	0.0111	0.0062	0.0011	0.0053	0.0017	3e-04	0.0046	0.001	1e-04
	(0.0378)	(0.0172)	(0.0047)	(0.0032)	(9e-04)	(8e-04)	(0.0018)	(3e-04)	(4e-04)
MR 3*	0.0118	0.0069	0.0012	0.0054	0.0018	3e-04	0.0047	0.001	1e-04
	(0.0571)	(0.0273)	(0.007)	(0.0038)	(0.0012)	(0.001)	(0.002)	(4e-04)	(5e-04)

Tabelle B.25.: Simulationsergebnisse für Splines: L_2-Risiko und gewählte $\lambda \cdot 10^4$ für SinePeak.

StN	3			7			10		
n	50	100	1000	50	100	1000	50	100	1000
CV	0.0363	0.0191	0.0026	0.0092	0.0045	6e-04	0.0049	0.0024	3e-04
	(0.0578)	(0.0431)	(0.0672)	(0.0059)	(0.0086)	(0.0245)	(0.002)	(0.0056)	(0.0171)
GCV	0.0381	0.0191	0.0026	0.0095	0.0045	6e-04	0.0051	0.0025	3e-04
	(0.0527)	(0.0403)	(0.0672)	(0.005)	(0.0082)	(0.0245)	(0.0015)	(0.0053)	(0.0171)
Cp	0.0356	0.0195	0.0026	0.0092	0.0045	6e-04	0.0051	0.0025	3e-04
	(0.1363)	(0.0609)	(0.0681)	(0.0212)	(0.0154)	(0.0251)	(0.013)	(0.0106)	(0.0175)
Chi .5	0.0579	0.0343	0.0058	0.0194	0.01	0.0013	0.012	0.0061	7e-04
	(3.0521)	(1.0534)	(0.4496)	(0.1308)	(0.0957)	(0.1185)	(0.0661)	(0.0571)	(0.073)
Chi .9	0.0777	0.0459	0.0084	0.0248	0.0127	0.0019	0.0153	0.0076	0.001
	(13.255)	(3.7965)	(0.8958)	(0.2299)	(0.1448)	(0.1891)	(0.0996)	(0.0783)	(0.1119)
Chi .95	0.0845	0.0497	0.0094	0.0266	0.0136	0.002	0.0164	0.0081	0.0011
	(19.0917)	(5.4152)	(1.057)	(0.271)	(0.1619)	(0.2107)	(0.1119)	(0.0854)	(0.1234)
Chi .99	0.0986	0.0574	0.0113	0.0301	0.0154	0.0024	0.0185	0.0091	0.0012
	(35.7224)	(10.0838)	(1.4059)	(0.3726)	(0.1994)	(0.2523)	(0.1397)	(0.1)	(0.145)
MR .5	0.041	0.0226	0.0033	0.0117	0.0054	7e-04	0.0066	0.0031	4e-04
	(0.3127)	(0.1709)	(0.1475)	(0.0426)	(0.0306)	(0.0461)	(0.0233)	(0.0196)	(0.0318)
MR .9	0.0498	0.0268	0.0035	0.0147	0.0065	8e-04	0.0085	0.0037	4e-04
	(0.8454)	(0.3133)	(0.221)	(0.072)	(0.0449)	(0.0664)	(0.0367)	(0.0273)	(0.0429)
MR .95	0.0532	0.0285	0.0036	0.0159	0.0069	8e-04	0.0092	0.004	4e-04
	(1.221)	(0.3799)	(0.244)	(0.0848)	(0.0506)	(0.0712)	(0.0421)	(0.0303)	(0.046)
MR .99	0.0614	0.0322	0.004	0.0185	0.0079	9e-04	0.0109	0.0045	5e-04
	(3.341)	(0.5619)	(0.2946)	(0.1174)	(0.0636)	(0.0814)	(0.0556)	(0.037)	(0.052)
MR 2	0.0412	0.0225	0.0034	0.0117	0.0054	7e-04	0.0066	0.0031	4e-04
	(0.3189)	(0.1665)	(0.126)	(0.0431)	(0.0301)	(0.04)	(0.0236)	(0.0193)	(0.0283)
MR 2.3	0.0436	0.0238	0.0033	0.0126	0.0057	7e-04	0.0071	0.0033	4e-04
	(0.4282)	(0.2085)	(0.1737)	(0.0509)	(0.0348)	(0.0544)	(0.0273)	(0.0219)	(0.0367)
MR 2.5	0.0453	0.0247	0.0033	0.0131	0.006	7e-04	0.0075	0.0034	4e-04
	(0.5171)	(0.2387)	(0.1987)	(0.0564)	(0.0381)	(0.061)	(0.0298)	(0.0236)	(0.04)
MR 3	0.0496	0.0272	0.0037	0.0146	0.0066	8e-04	0.0084	0.0038	4e-04
	(0.8228)	(0.3267)	(0.2516)	(0.0712)	(0.0462)	(0.0727)	(0.0363)	(0.0279)	(0.0468)
Cp*	0.0337	0.0189	0.0026	0.0082	0.0044	6e-04	0.0043	0.0024	3e-04
	(0.0486)	(0.0402)	(0.0679)	(0.0085)	(0.0101)	(0.0249)	(0.0052)	(0.0064)	(0.0172)
Chi .5*	0.0422	0.0246	0.0045	0.0116	0.0068	0.0011	0.0065	0.0037	5e-04
	(0.3562)	(0.2433)	(0.3306)	(0.0417)	(0.048)	(0.1002)	(0.0229)	(0.0271)	(0.0591)
Chi .9*	0.0534	0.0327	0.0075	0.0142	0.0086	0.0017	0.0078	0.0046	8e-04
	(1.0851)	(0.5996)	(0.7484)	(0.064)	(0.0734)	(0.1715)	(0.0317)	(0.0382)	(0.0983)
Chi .95*	0.0574	0.0356	0.0086	0.015	0.0092	0.0019	0.0083	0.0049	9e-04
	(1.5048)	(0.773)	(0.8974)	(0.0722)	(0.0822)	(0.1927)	(0.0347)	(0.0418)	(0.1095)
Chi .99*	0.066	0.0416	0.0107	0.0168	0.0105	0.0022	0.0091	0.0055	0.0011
	(2.8362)	(1.2555)	(1.2168)	(0.0903)	(0.1008)	(0.2336)	(0.041)	(0.0492)	(0.1309)
MR .5*	0.0368	0.0209	0.0032	0.0093	0.0049	7e-04	0.0049	0.0027	4e-04
	(0.1591)	(0.1243)	(0.1509)	(0.022)	(0.0231)	(0.0457)	(0.0117)	(0.0138)	(0.0313)
MR .9*	0.0431	0.0243	0.0035	0.0108	0.0057	8e-04	0.0057	0.0031	4e-04
	(0.3431)	(0.2236)	(0.2207)	(0.0348)	(0.0338)	(0.0653)	(0.0173)	(0.0192)	(0.0423)
MR .95*	0.0455	0.0257	0.0036	0.0114	0.006	8e-04	0.006	0.0032	4e-04
	(0.4352)	(0.2662)	(0.2437)	(0.0397)	(0.0378)	(0.0705)	(0.0194)	(0.0212)	(0.0453)
MR .99*	0.0511	0.0288	0.004	0.0128	0.0067	9e-04	0.0067	0.0036	4e-04
	(0.7133)	(0.3755)	(0.2934)	(0.0519)	(0.0469)	(0.0807)	(0.0243)	(0.0257)	(0.0512)
MR 2*	0.0369	0.0208	0.0032	0.0093	0.0049	7e-04	0.0049	0.0027	4e-04
	(0.1617)	(0.1211)	(0.128)	(0.0223)	(0.0227)	(0.0405)	(0.0118)	(0.0136)	(0.0273)
MR 2.3*	0.0386	0.0218	0.0032	0.0097	0.0051	7e-04	0.0051	0.0028	4e-04
	(0.2062)	(0.151)	(0.1741)	(0.0259)	(0.0263)	(0.0531)	(0.0134)	(0.0154)	(0.0353)
MR 2.5*	0.0398	0.0226	0.0033	0.01	0.0053	7e-04	0.0052	0.0029	4e-04
	(0.2397)	(0.1724)	(0.1997)	(0.0283)	(0.0288)	(0.0602)	(0.0144)	(0.0166)	(0.0394)
MR 3 *	0.0429	0.0246	0.0037	0.0108	0.0057	8e-04	0.0056	0.0031	4e-04
	(0.3377)	(0.2328)	(0.2509)	(0.0344)	(0.0346)	(0.0719)	(0.0171)	(0.0196)	(0.0461)

Tabelle B.26.: Simulationsergebnisse für Splines: L_2-Risiko und gewählte $\lambda \cdot 10^4$ für Piecewise Polynomial.

StN		3			7			10	
n	50	100	1000	50	100	1000	50	100	1000
CV	0.054	0.0355	0.0095	0.0272	0.0162	0.004	0.0222	0.0126	0.0029
	(1.2573)	(0.5121)	(0.025)	(0.1295)	(0.035)	(0.001)	(0.0674)	(0.0126)	(3e-04)
GCV	0.0548	0.0353	0.0095	0.0269	0.0162	0.004	0.0225	0.0126	0.0029
	(1.1866)	(0.4561)	(0.0248)	(0.1211)	(0.0327)	(0.001)	(0.0652)	(0.0123)	(3e-04)
Cp	0.054	0.0355	0.0096	0.0259	0.0161	0.0041	0.0215	0.0127	0.0029
	(1.3264)	(0.4506)	(0.0211)	(0.0702)	(0.0138)	(7e-04)	(0.0244)	(0.0035)	(2e-04)
Chi .5	0.0681	0.0394	0.0112	0.0284	0.0164	0.0043	0.0232	0.0127	0.003
	(15.7912)	(5.8278)	(0.2971)	(0.4293)	(0.0808)	(0.0037)	(0.1476)	(0.0188)	(7e-04)
Chi .9	0.0938	0.0514	0.0136	0.0334	0.0182	0.0048	0.0261	0.0135	0.0032
	(49.8982)	(20.4642)	(1.4649)	(1.3278)	(0.2376)	(0.0091)	(0.3854)	(0.041)	(0.0014)
Chi .95	0.1033	0.0561	0.0146	0.0353	0.0189	0.0049	0.0272	0.0138	0.0033
	(65.3313)	(27.8575)	(2.1691)	(1.7793)	(0.3178)	(0.0116)	(0.5039)	(0.0512)	(0.0017)
Chi .99	0.1237	0.0662	0.0168	0.0393	0.0205	0.0053	0.0296	0.0145	0.0034
	(103.3319)	(46.6562)	(4.2421)	(2.975)	(0.5286)	(0.0182)	(0.8227)	(0.0775)	(0.0024)
MR .5	0.057	0.035	0.0122	0.0253	0.0163	0.0066	0.0211	0.0129	0.005
	(4.304)	(0.973)	(0.0031)	(0.0611)	(0.0054)	(0)	(0.017)	(0.0013)	(0)
MR .9	0.0699	0.0391	0.0103	0.0266	0.0158	0.0058	0.0214	0.0125	0.0046
	(13.3346)	(3.7995)	(0.0095)	(0.2238)	(0.0178)	(0)	(0.0464)	(0.0026)	(0)
MR .95	0.0753	0.0412	0.01	0.0276	0.0158	0.0056	0.0217	0.0123	0.0044
	(17.8202)	(5.35)	(0.0135)	(0.3387)	(0.0261)	(0)	(0.0643)	(0.0033)	(0)
MR .99	0.0871	0.0464	0.0097	0.0304	0.0161	0.0051	0.0229	0.0123	0.0041
	(30.0487)	(9.5988)	(0.0247)	(0.762)	(0.0517)	(1e-04)	(0.1318)	(0.0057)	(0)
MR 2	0.0572	0.0349	0.0129	0.0253	0.0164	0.0068	0.0211	0.013	0.0051
	(4.4239)	(0.9212)	(0.0023)	(0.0627)	(0.0052)	(0)	(0.0173)	(0.0013)	(0)
MR 2.3	0.0605	0.0356	0.0114	0.0255	0.0161	0.0063	0.0211	0.0128	0.0049
	(6.672)	(1.5398)	(0.0046)	(0.0913)	(0.0077)	(0)	(0.0232)	(0.0016)	(0)
MR 2.5	0.063	0.0367	0.0107	0.0257	0.0159	0.0061	0.0211	0.0126	0.0047
	(8.2937)	(2.2149)	(0.0067)	(0.1175)	(0.0102)	(0)	(0.0284)	(0.0019)	(0)
MR 3	0.0695	0.0395	0.0099	0.0266	0.0158	0.0055	0.0213	0.0124	0.0044
	(13.0568)	(4.0918)	(0.0149)	(0.2169)	(0.0194)	(0)	(0.0453)	(0.0027)	(0)
Cp*	0.0536	0.0337	0.0095	0.0271	0.0163	0.004	0.0233	0.0133	0.0029
	(0.7315)	(0.303)	(0.02)	(0.0098)	(0.0059)	(6e-04)	(0.0031)	(0.001)	(1e-04)
Chi .5*	0.0574	0.0361	0.0103	0.0249	0.0156	0.0042	0.0209	0.0121	0.0029
	(5.0822)	(2.0858)	(0.1008)	(0.0647)	(0.0301)	(0.0022)	(0.0217)	(0.0067)	(4e-04)
Chi .9*	0.0735	0.0454	0.0126	0.0262	0.0167	0.0046	0.0211	0.0124	0.0031
	(18.2287)	(9.2414)	(0.6492)	(0.1659)	(0.0765)	(0.0054)	(0.0407)	(0.0126)	(8e-04)
Chi .95*	0.0805	0.0496	0.0137	0.0269	0.0172	0.0047	0.0213	0.0125	0.0032
	(25.0859)	(13.1488)	(1.0301)	(0.2193)	(0.1003)	(0.007)	(0.0492)	(0.0152)	(9e-04)
Chi .99*	0.0963	0.059	0.016	0.0285	0.0183	0.005	0.0218	0.0129	0.0033
	(43.578)	(23.9762)	(2.2631)	(0.3703)	(0.1666)	(0.0112)	(0.0707)	(0.0218)	(0.0013)
MR .5*	0.0537	0.0337	0.012	0.0253	0.0165	0.0067	0.0218	0.0132	0.005
	(2.5779)	(0.6143)	(0.0027)	(0.0195)	(0.0037)	(0)	(0.006)	(9e-04)	(0)
MR .9*	0.0644	0.0373	0.0102	0.0249	0.0157	0.0058	0.0211	0.0127	0.0046
	(8.795)	(2.5313)	(0.0086)	(0.0515)	(0.009)	(0)	(0.011)	(0.0016)	(0)
MR .95*	0.0688	0.0395	0.0099	0.025	0.0156	0.0056	0.021	0.0125	0.0045
	(11.7627)	(3.8058)	(0.012)	(0.0703)	(0.0123)	(0)	(0.0136)	(0.002)	(0)
MR .99*	0.0794	0.0452	0.0096	0.0258	0.0156	0.0051	0.0209	0.0123	0.0042
	(20.1169)	(7.6154)	(0.0232)	(0.1368)	(0.0252)	(1e-04)	(0.0215)	(0.003)	(0)
MR 2*	0.0538	0.0337	0.0126	0.0253	0.0166	0.0069	0.0218	0.0133	0.0051
	(2.6425)	(0.582)	(0.002)	(0.0199)	(0.0036)	(0)	(0.006)	(9e-04)	(0)
MR 2.3*	0.0561	0.0343	0.0112	0.0251	0.0162	0.0064	0.0216	0.013	0.0049
	(3.8952)	(1.0148)	(0.004)	(0.0271)	(0.0049)	(0)	(0.0073)	(0.0011)	(0)
MR 2.5*	0.0581	0.0351	0.0106	0.025	0.016	0.0061	0.0214	0.0129	0.0047
	(5.0337)	(1.4225)	(0.006)	(0.0328)	(0.006)	(0)	(0.0082)	(0.0012)	(0)
MR 3*	0.0641	0.0378	0.0098	0.0249	0.0157	0.0055	0.0212	0.0126	0.0044
	(8.591)	(2.8001)	(0.0132)	(0.0504)	(0.0096)	(0)	(0.0109)	(0.0017)	(0)

Tabelle B.27.: Simulationsergebnisse für Splines: L_2-Risiko und gewählte $\lambda \cdot 10^4$ für Blocks.

StN	3			7			10		
n	50	100	1000	50	100	1000	50	100	1000
CV	0.357	0.1369	0.0374	0.3246	0.0793	0.0163	0.3199	0.0695	0.0119
	(0.5172)	(0.0012)	(1e-04)	(0.3455)	(1e-04)	(0)	(0.3207)	(0)	(0)
GCV	0.3592	0.1373	0.0374	0.3278	0.0793	0.0162	0.3233	0.0695	0.0119
	(0.5474)	(0.0011)	(1e-04)	(0.3732)	(0)	(0)	(0.3491)	(0)	(0)
Cp	0.2121	0.1506	0.0374	0.1074	0.0882	0.0162	0.0958	0.0763	0.012
	(0.0297)	(0.0043)	(1e-04)	(5e-04)	(8e-04)	(0)	(2e-04)	(4e-04)	(0)
Chi .5	0.2704	0.2194	0.0431	0.1405	0.1161	0.0196	0.1164	0.0944	0.0148
	(0.0981)	(0.0422)	(7e-04)	(0.0039)	(0.0037)	(0)	(0.002)	(0.0018)	(0)
Chi .9	0.3074	0.2455	0.0462	0.1509	0.1237	0.0203	0.1221	0.0986	0.0153
	(0.4381)	(0.0976)	(0.0012)	(0.0055)	(0.0048)	(0)	(0.0025)	(0.0022)	(0)
Chi .95	0.3197	0.2534	0.0473	0.1542	0.126	0.0206	0.1239	0.0999	0.0154
	(0.6598)	(0.1277)	(0.0014)	(0.0061)	(0.0052)	(1e-04)	(0.0027)	(0.0023)	(0)
Chi .99	0.3456	0.2691	0.0493	0.1608	0.1305	0.021	0.1274	0.1023	0.0157
	(1.3837)	(0.217)	(0.0019)	(0.0074)	(0.006)	(1e-04)	(0.003)	(0.0026)	(0)
MR .5	0.2487	0.1959	0.0381	0.1341	0.103	0.0162	0.113	0.0851	0.012
	(0.0941)	(0.0237)	(1e-04)	(0.0031)	(0.0021)	(0)	(0.0017)	(0.001)	(0)
MR .9	0.3024	0.2341	0.0403	0.1486	0.1136	0.0166	0.1214	0.0906	0.0123
	(0.6568)	(0.0839)	(4e-04)	(0.0051)	(0.0034)	(0)	(0.0024)	(0.0014)	(0)
MR .95	0.3223	0.2478	0.0414	0.154	0.1176	0.0169	0.1244	0.0927	0.0125
	(1.0292)	(0.1342)	(5e-04)	(0.006)	(0.004)	(0)	(0.0027)	(0.0016)	(0)
MR .99	0.3695	0.2774	0.0444	0.166	0.1266	0.0175	0.1311	0.0973	0.0128
	(2.6112)	(0.3194)	(9e-04)	(0.0085)	(0.0054)	(0)	(0.0035)	(0.0021)	(0)
MR 2	0.2496	0.1945	0.0379	0.1343	0.1026	0.0161	0.1132	0.0849	0.012
	(0.0972)	(0.0226)	(1e-04)	(0.0031)	(0.0021)	(0)	(0.0017)	(0.001)	(0)
MR 2.3	0.2645	0.2078	0.0387	0.1385	0.1062	0.0163	0.1156	0.0867	0.0121
	(0.1993)	(0.0349)	(2e-04)	(0.0036)	(0.0025)	(0)	(0.0019)	(0.0011)	(0)
MR 2.5	0.2752	0.2166	0.0394	0.1413	0.1085	0.0164	0.1172	0.088	0.0122
	(0.2804)	(0.0472)	(3e-04)	(0.004)	(0.0028)	(0)	(0.002)	(0.0012)	(0)
MR 3	0.3012	0.237	0.0418	0.1482	0.1145	0.0169	0.1211	0.0911	0.0125
	(0.643)	(0.0926)	(5e-04)	(0.005)	(0.0035)	(0)	(0.0024)	(0.0015)	(0)
Cp*	0.1844	0.1369	0.0371	0.1049	0.0822	0.016	0.0947	0.0726	0.0119
	(0.001)	(0.0015)	(1e-04)	(2e-04)	(3e-04)	(0)	(1e-04)	(2e-04)	(0)
Chi .5*	0.2149	0.1697	0.0412	0.119	0.0978	0.019	0.1039	0.0837	0.0144
	(0.008)	(0.0087)	(4e-04)	(0.0016)	(0.0016)	(0)	(9e-04)	(9e-04)	(0)
Chi .9*	0.229	0.1841	0.0438	0.123	0.1018	0.0196	0.1062	0.086	0.0148
	(0.0127)	(0.0137)	(7e-04)	(0.0019)	(0.002)	(0)	(0.0011)	(0.0011)	(0)
Chi .95*	0.2337	0.1886	0.0447	0.1243	0.103	0.0198	0.1069	0.0867	0.0149
	(0.0147)	(0.0156)	(9e-04)	(0.002)	(0.0021)	(0)	(0.0012)	(0.0011)	(0)
Chi .99*	0.2434	0.1977	0.0464	0.1268	0.1054	0.0202	0.1083	0.0881	0.0151
	(0.0196)	(0.02)	(0.0012)	(0.0023)	(0.0023)	(0)	(0.0013)	(0.0012)	(0)
MR .5*	0.2076	0.1626	0.0379	0.117	0.0919	0.0161	0.1028	0.0789	0.012
	(0.0063)	(0.0072)	(1e-04)	(0.0014)	(0.0011)	(0)	(8e-04)	(6e-04)	(0)
MR .9*	0.2275	0.185	0.0397	0.1232	0.0974	0.0165	0.1063	0.082	0.0122
	(0.0142)	(0.015)	(3e-04)	(0.0019)	(0.0016)	(0)	(0.0011)	(8e-04)	(0)
MR .95*	0.2354	0.1939	0.0408	0.1254	0.0995	0.0167	0.1076	0.0832	0.0124
	(0.0187)	(0.0195)	(4e-04)	(0.0021)	(0.0018)	(0)	(0.0012)	(9e-04)	(0)
MR .99*	0.2552	0.2139	0.0435	0.1305	0.1042	0.0172	0.1105	0.0857	0.0126
	(0.0441)	(0.0344)	(8e-04)	(0.0026)	(0.0022)	(0)	(0.0014)	(0.001)	(0)
MR 2*	0.208	0.1618	0.0378	0.1171	0.0917	0.016	0.1028	0.0788	0.0119
	(0.0064)	(0.007)	(1e-04)	(0.0014)	(0.0011)	(0)	(8e-04)	(6e-04)	(0)
MR 2.3*	0.2134	0.1691	0.0383	0.1189	0.0935	0.0162	0.1039	0.0799	0.012
	(0.0078)	(0.0091)	(2e-04)	(0.0015)	(0.0012)	(0)	(9e-04)	(6e-04)	(0)
MR 2.5*	0.2171	0.174	0.039	0.1201	0.0948	0.0163	0.1046	0.0806	0.0121
	(0.009)	(0.0107)	(2e-04)	(0.0016)	(0.0013)	(0)	(0.001)	(7e-04)	(0)
MR 3*	0.227	0.1869	0.0411	0.123	0.0979	0.0168	0.1063	0.0823	0.0124
	(0.0139)	(0.0158)	(5e-04)	(0.0019)	(0.0016)	(0)	(0.0011)	(8e-04)	(0)

Tabelle B.28.: Simulationsergebnisse für Splines: L_2-Risiko und gewählte $\lambda \cdot 10^4$ für Bumps.

StN	3			7			10		
n	50	100	1000	50	100	1000	50	100	1000
CV	0.9211	0.8215	0.0503	0.8657	0.7393	0.0152	0.8537	0.7273	0.0095
	(17832.1518)	(11313.2078)	(0)	(2955.9554)	(1479.5945)	(0)	(2216.9154)	(0.7169)	(0)
GCV	0.9512	0.8758	0.0503	0.9291	0.7975	0.0152	0.927	0.7962	0.0095
	(37983.6263)	(31823.8948)	(0)	(36945.3876)	(19950.6592)	(0)	(36945.3728)	(19950.6585)	(0)
Cp	1.0087	0.3676	0.0497	0.9637	0.2967	0.0147	0.9566	0.289	0.0091
	(0.001)	(1e-04)	(0)	(1e-04)	(0)	(0)	(1e-04)	(0)	(0)
Chi .5	0.9269	0.4094	0.062	0.9183	0.296	0.0168	0.9195	0.284	0.0095
	(0.0087)	(6e-04)	(0)	(0.0014)	(1e-04)	(0)	(0.001)	(1e-04)	(0)
Chi .9	0.9119	0.4247	0.0655	0.9124	0.2982	0.0173	0.9147	0.2847	0.0098
	(0.0155)	(9e-04)	(0)	(0.0018)	(2e-04)	(0)	(0.0012)	(1e-04)	(0)
Chi .95	0.908	0.4297	0.0666	0.9106	0.2989	0.0175	0.9134	0.2849	0.0098
	(0.0188)	(9e-04)	(0)	(0.0019)	(2e-04)	(0)	(0.0013)	(1e-04)	(0)
Chi .99	0.9011	0.4397	0.0688	0.9074	0.3004	0.0178	0.9109	0.2854	0.01
	(0.0275)	(0.0011)	(0)	(0.0021)	(2e-04)	(0)	(0.0014)	(1e-04)	(0)
MR .5	0.9346	0.4115	0.056	0.9195	0.2969	0.0156	0.9202	0.2843	0.0092
	(0.0069)	(7e-04)	(0)	(0.0014)	(1e-04)	(0)	(0.001)	(1e-04)	(0)
MR .9	0.9137	0.4467	0.0534	0.9097	0.3029	0.0152	0.9122	0.2864	0.009
	(0.0147)	(0.0014)	(0)	(0.002)	(2e-04)	(0)	(0.0013)	(1e-04)	(0)
MR .95	0.9077	0.4611	0.0527	0.9066	0.3055	0.015	0.9097	0.2874	0.0089
	(0.0192)	(0.0019)	(0)	(0.0022)	(2e-04)	(0)	(0.0014)	(1e-04)	(0)
MR .99	0.8971	0.4951	0.0514	0.9006	0.3115	0.0148	0.9048	0.2899	0.0088
	(0.0329)	(0.0045)	(0)	(0.0027)	(3e-04)	(0)	(0.0017)	(1e-04)	(0)
MR 2	0.9342	0.4104	0.0566	0.9193	0.2967	0.0157	0.92	0.2842	0.0093
	(0.007)	(7e-04)	(0)	(0.0014)	(1e-04)	(0)	(0.001)	(1e-04)	(0)
MR 2.3	0.9274	0.4217	0.0551	0.9162	0.2986	0.0154	0.9175	0.2848	0.0091
	(0.0088)	(8e-04)	(0)	(0.0015)	(2e-04)	(0)	(0.0011)	(1e-04)	(0)
MR 2.5	0.9232	0.4296	0.0542	0.9143	0.2999	0.0153	0.9159	0.2853	0.0091
	(0.0101)	(0.001)	(0)	(0.0017)	(2e-04)	(0)	(0.0011)	(1e-04)	(0)
MR 3	0.9142	0.4498	0.0525	0.9099	0.3034	0.015	0.9124	0.2866	0.0089
	(0.0144)	(0.0015)	(0)	(0.0019)	(2e-04)	(0)	(0.0013)	(1e-04)	(0)
Cp*	1.0349	0.3737	0.0508	0.9688	0.2985	0.0158	0.9601	0.29	0.0098
	(4e-04)	(0)	(0)	(1e-04)	(0)	(0)	(0)	(0)	(0)
Chi .5*	0.9561	0.3766	0.0544	0.932	0.2917	0.015	0.9332	0.2834	0.0087
	(0.0034)	(3e-04)	(0)	(9e-04)	(1e-04)	(0)	(5e-04)	(1e-04)	(0)
Chi .9*	0.944	0.3821	0.0565	0.9273	0.2921	0.0152	0.9298	0.2832	0.0088
	(0.0046)	(3e-04)	(0)	(0.001)	(1e-04)	(0)	(6e-04)	(1e-04)	(0)
Chi .95*	0.9405	0.3839	0.0571	0.9259	0.2922	0.0153	0.9288	0.2832	0.0088
	(0.005)	(3e-04)	(0)	(0.0011)	(1e-04)	(0)	(7e-04)	(1e-04)	(0)
Chi .99*	0.9338	0.3876	0.0585	0.9234	0.2925	0.0155	0.927	0.2832	0.0088
	(0.006)	(4e-04)	(0)	(0.0012)	(1e-04)	(0)	(7e-04)	(1e-04)	(0)
MR .5*	0.9595	0.3779	0.0579	0.9326	0.2918	0.0161	0.9334	0.2833	0.0095
	(0.0031)	(3e-04)	(0)	(8e-04)	(1e-04)	(0)	(5e-04)	(1e-04)	(0)
MR .9*	0.9407	0.392	0.0551	0.9245	0.2931	0.0156	0.9275	0.2832	0.0093
	(0.005)	(4e-04)	(0)	(0.0011)	(1e-04)	(0)	(7e-04)	(1e-04)	(0)
MR .95*	0.9348	0.3979	0.0543	0.922	0.2938	0.0155	0.9256	0.2833	0.0092
	(0.0059)	(5e-04)	(0)	(0.0012)	(1e-04)	(0)	(8e-04)	(1e-04)	(0)
MR .99*	0.9231	0.4118	0.0528	0.9169	0.2955	0.0152	0.9218	0.2837	0.0091
	(0.0082)	(6e-04)	(0)	(0.0015)	(1e-04)	(0)	(9e-04)	(1e-04)	(0)
MR 2*	0.9591	0.3775	0.0585	0.9324	0.2918	0.0161	0.9333	0.2833	0.0096
	(0.0032)	(3e-04)	(0)	(8e-04)	(1e-04)	(0)	(5e-04)	(1e-04)	(0)
MR 2.3*	0.9532	0.3819	0.0569	0.9299	0.2921	0.0159	0.9314	0.2832	0.0095
	(0.0037)	(3e-04)	(0)	(9e-04)	(1e-04)	(0)	(6e-04)	(1e-04)	(0)
MR 2.5*	0.9495	0.385	0.056	0.9283	0.2924	0.0158	0.9303	0.2832	0.0094
	(0.004)	(4e-04)	(0)	(0.001)	(1e-04)	(0)	(6e-04)	(1e-04)	(0)
MR 3*	0.9412	0.3933	0.054	0.9247	0.2933	0.0155	0.9276	0.2832	0.0092
	(0.005)	(4e-04)	(0)	(0.0011)	(1e-04)	(0)	(7e-04)	(1e-04)	(0)

Tabelle B.29.: Simulationsergebnisse für Splines: L_2-Risiko und gewählte $\lambda \cdot 10^4$ für HeaviSine.

StN	3			7			10		
n	50	100	1000	50	100	1000	50	100	1000
CV	0.0329	0.0194	0.0057	0.0123	0.0081	0.0023	0.0083	0.0056	0.0016
	(0.5365)	(0.7587)	(0.2763)	(0.133)	(0.1342)	(0.0077)	(0.0659)	(0.0412)	(0.0017)
GCV	0.0343	0.0196	0.0056	0.0125	0.0081	0.0023	0.0085	0.0056	0.0016
	(0.5917)	(0.7708)	(0.2739)	(0.1479)	(0.1362)	(0.0077)	(0.0724)	(0.0379)	(0.0017)
Cp	0.0321	0.0199	0.0057	0.0144	0.0083	0.0023	0.0126	0.0064	0.0016
	(0.8657)	(0.8985)	(0.2811)	(0.6529)	(0.2766)	(0.0068)	(0.6613)	(0.1724)	(0.0015)
Chi .5	0.0683	0.0366	0.0091	0.0515	0.0166	0.0026	0.0515	0.0142	0.0018
	(4.6363)	(5.0538)	(7.3858)	(4.1079)	(2.4135)	(0.061)	(4.2179)	(2.0894)	(0.0125)
Chi .9	0.1027	0.0563	0.0124	0.0679	0.0221	0.0031	0.066	0.0179	0.0021
	(7.7388)	(8.7794)	(15.9621)	(5.4164)	(3.5886)	(0.2288)	(5.3722)	(2.9319)	(0.034)
Chi .95	0.1138	0.0628	0.0138	0.073	0.0239	0.0033	0.0706	0.0191	0.0021
	(8.7641)	(9.9207)	(19.1861)	(5.824)	(3.935)	(0.3211)	(5.7296)	(3.1741)	(0.0444)
Chi .99	0.1366	0.0759	0.0167	0.0831	0.0273	0.0037	0.0796	0.0214	0.0023
	(11.0042)	(12.1765)	(25.8579)	(6.6335)	(4.6009)	(0.5759)	(6.4482)	(3.6406)	(0.0718)
MR .5	0.0634	0.032	0.0066	0.0373	0.012	0.0025	0.0347	0.0088	0.0018
	(4.1474)	(4.2477)	(1.0467)	(2.9318)	(1.2843)	(0.004)	(2.8376)	(0.7399)	(5e-04)
MR .9	0.098	0.0488	0.0069	0.0536	0.0165	0.0024	0.0494	0.0122	0.0017
	(7.2781)	(7.6239)	(3.0359)	(4.2788)	(2.4227)	(0.0124)	(4.0515)	(1.5981)	(0.0012)
MR .95	0.1108	0.0552	0.0073	0.0597	0.0181	0.0024	0.0548	0.0133	0.0017
	(8.4295)	(8.7315)	(4.068)	(4.7639)	(2.7837)	(0.0171)	(4.4815)	(1.8855)	(0.0016)
MR .99	0.1396	0.0683	0.0085	0.0732	0.0217	0.0025	0.0668	0.0159	0.0017
	(11.1664)	(10.9416)	(7.1126)	(5.838)	(3.5392)	(0.0325)	(5.4342)	(2.4822)	(0.0027)
MR 2	0.0639	0.0316	0.007	0.0376	0.0118	0.0026	0.035	0.0087	0.0018
	(4.1968)	(4.092)	(0.7465)	(2.9579)	(1.2488)	(0.0031)	(2.864)	(0.7113)	(5e-04)
MR 2.3	0.074	0.0371	0.0065	0.0425	0.0132	0.0024	0.0393	0.0098	0.0017
	(5.1534)	(5.3748)	(1.6831)	(3.3728)	(1.6264)	(0.006)	(3.2244)	(1.0002)	(7e-04)
MR 2.5	0.0808	0.0407	0.0067	0.0456	0.0143	0.0024	0.0421	0.0106	0.0017
	(5.7614)	(6.0994)	(2.3664)	(3.6275)	(1.8876)	(0.0088)	(3.4574)	(1.1921)	(9e-04)
MR 3	0.0971	0.0502	0.0074	0.0532	0.0169	0.0024	0.049	0.0124	0.0017
	(7.2019)	(7.8669)	(4.4041)	(4.2464)	(2.5025)	(0.0188)	(4.0195)	(1.6592)	(0.0017)
Cp*	0.0324	0.0192	0.0057	0.0119	0.008	0.0023	0.008	0.0054	0.0016
	(0.6553)	(0.7514)	(0.2566)	(0.1522)	(0.0987)	(0.0068)	(0.055)	(0.0229)	(0.0014)
Chi .5*	0.0475	0.027	0.0067	0.015	0.0089	0.0025	0.0097	0.0062	0.0018
	(2.6922)	(2.9497)	(2.2105)	(0.7041)	(0.4719)	(0.0438)	(0.2966)	(0.1443)	(0.0083)
Chi .9*	0.0722	0.0414	0.01	0.0198	0.0112	0.003	0.0119	0.0073	0.002
	(4.9848)	(6.2296)	(10.9717)	(1.2823)	(1.1285)	(0.1767)	(0.5859)	(0.3738)	(0.0236)
Chi .95*	0.0808	0.047	0.0115	0.0215	0.0121	0.0032	0.0128	0.0078	0.0021
	(5.7203)	(7.2718)	(14.5285)	(1.4647)	(1.3674)	(0.2541)	(0.6867)	(0.473)	(0.0313)
Chi .99*	0.0986	0.0584	0.0146	0.025	0.014	0.0037	0.0145	0.0087	0.0023
	(7.2218)	(9.2687)	(22.0291)	(1.8205)	(1.8619)	(0.477)	(0.8925)	(0.7017)	(0.052)
MR .5*	0.0493	0.0282	0.0064	0.0161	0.0088	0.0025	0.0097	0.0058	0.0018
	(2.9216)	(3.4608)	(0.7826)	(0.8237)	(0.3988)	(0.004)	(0.2984)	(0.0885)	(5e-04)
MR .9*	0.0789	0.0435	0.0068	0.0224	0.0116	0.0024	0.0131	0.0069	0.0017
	(5.591)	(6.7058)	(2.8252)	(1.5603)	(1.1913)	(0.0121)	(0.7125)	(0.283)	(0.0012)
MR .95*	0.0899	0.0491	0.0072	0.0246	0.0129	0.0024	0.0144	0.0074	0.0017
	(6.5183)	(7.7198)	(3.8693)	(1.7824)	(1.5599)	(0.017)	(0.8773)	(0.3958)	(0.0015)
MR .99*	0.1137	0.0609	0.0083	0.0299	0.0153	0.0025	0.0174	0.0088	0.0017
	(8.5495)	(9.7202)	(6.6533)	(2.2879)	(2.1694)	(0.032)	(1.2125)	(0.72)	(0.0026)
MR 2*	0.0498	0.0277	0.0066	0.0162	0.0088	0.0025	0.0098	0.0058	0.0018
	(2.9673)	(3.3383)	(0.5125)	(0.8383)	(0.3783)	(0.0031)	(0.3056)	(0.0835)	(4e-04)
MR 2.3*	0.0582	0.0326	0.0063	0.0181	0.0094	0.0024	0.0106	0.0061	0.0017
	(3.764)	(4.5787)	(1.2368)	(1.0749)	(0.5916)	(0.0059)	(0.411)	(0.1312)	(7e-04)
MR 2.5*	0.0637	0.0362	0.0065	0.0193	0.01	0.0024	0.0112	0.0063	0.0017
	(4.2601)	(5.3224)	(1.9942)	(1.2162)	(0.7551)	(0.0087)	(0.4869)	(0.1681)	(9e-04)
MR 3*	0.078	0.0446	0.0074	0.0222	0.0118	0.0024	0.013	0.007	0.0017
	(5.5186)	(6.9156)	(4.3464)	(1.5344)	(1.2629)	(0.0188)	(0.7014)	(0.3074)	(0.0017)

Tabelle B.30.: Simulationsergebnisse für Splines: L_2-Risiko und gewählte $\lambda \cdot 10^4$ für Doppler.

StN	3			7			10		
n	50	100	1000	50	100	1000	50	100	1000
CV	0.193	0.1182	0.0239	0.139	0.0597	0.0075	0.1323	0.0523	0.0044
	(0.0245)	(0.0073)	(2e-04)	(0.0109)	(7e-04)	(0)	(0.0095)	(5e-04)	(0)
GCV	0.1918	0.118	0.0239	0.1374	0.0599	0.0075	0.1307	0.0526	0.0044
	(0.0228)	(0.0068)	(2e-04)	(0.0101)	(7e-04)	(0)	(0.0088)	(6e-04)	(0)
Cp	0.2043	0.1227	0.024	0.1489	0.0675	0.0077	0.1407	0.0581	0.0047
	(0.0446)	(0.0112)	(2e-04)	(0.0174)	(0.0019)	(0)	(0.0144)	(0.0012)	(0)
Chi .5	0.2897	0.1796	0.0309	0.2154	0.102	0.0119	0.2028	0.0878	0.0079
	(0.4013)	(0.1047)	(0.0016)	(0.1159)	(0.0131)	(2e-04)	(0.0913)	(0.0078)	(1e-04)
Chi .9	0.3524	0.2127	0.0348	0.2524	0.1158	0.0129	0.2356	0.0986	0.0085
	(1.2202)	(0.2415)	(0.0028)	(0.2461)	(0.0223)	(2e-04)	(0.1849)	(0.0124)	(1e-04)
Chi .95	0.3732	0.2233	0.0361	0.2643	0.1201	0.0132	0.2462	0.102	0.0087
	(1.7011)	(0.3061)	(0.0033)	(0.3058)	(0.026)	(2e-04)	(0.2269)	(0.0142)	(1e-04)
Chi .99	0.4159	0.2446	0.0386	0.2887	0.1286	0.0137	0.2676	0.1087	0.009
	(3.3085)	(0.4752)	(0.0044)	(0.4639)	(0.0347)	(3e-04)	(0.3347)	(0.0182)	(1e-04)
MR .5	0.2477	0.1443	0.025	0.159	0.0719	0.0078	0.1456	0.06	0.0046
	(0.2086)	(0.037)	(3e-04)	(0.0289)	(0.0029)	(0)	(0.0186)	(0.0015)	(0)
MR .9	0.3334	0.183	0.0272	0.2181	0.0918	0.0083	0.1937	0.0747	0.0049
	(0.9631)	(0.1194)	(8e-04)	(0.151)	(0.0092)	(0)	(0.0918)	(0.0043)	(0)
MR .95	0.3653	0.1951	0.0282	0.2413	0.1005	0.0086	0.2206	0.0815	0.0051
	(1.6435)	(0.1604)	(9e-04)	(0.2295)	(0.0136)	(0)	(0.1591)	(0.0062)	(0)
MR .99	0.4349	0.2247	0.0307	0.288	0.1194	0.0092	0.2691	0.0963	0.0054
	(5.1441)	(0.3067)	(0.0015)	(0.4793)	(0.0268)	(1e-04)	(0.3469)	(0.0122)	(0)
MR 2	0.2489	0.1431	0.0247	0.1598	0.0713	0.0077	0.1462	0.0596	0.0046
	(0.2146)	(0.0353)	(3e-04)	(0.0298)	(0.0027)	(0)	(0.0192)	(0.0014)	(0)
MR 2.3	0.2753	0.1557	0.0255	0.1758	0.078	0.0079	0.1565	0.0639	0.0047
	(0.3587)	(0.0555)	(5e-04)	(0.0532)	(0.0044)	(0)	(0.0295)	(0.0021)	(0)
MR 2.5	0.2911	0.1643	0.0263	0.1874	0.082	0.0081	0.1661	0.0673	0.0048
	(0.4885)	(0.072)	(6e-04)	(0.0741)	(0.0056)	(0)	(0.0435)	(0.0027)	(0)
MR 3	0.3313	0.1857	0.0285	0.2166	0.0936	0.0087	0.1923	0.0761	0.0051
	(0.9348)	(0.1275)	(0.001)	(0.1466)	(0.01)	(0)	(0.0888)	(0.0046)	(0)
Cp*	0.18	0.1058	0.0238	0.1189	0.0573	0.0075	0.1109	0.0503	0.0044
	(0.0032)	(0.001)	(1e-04)	(4e-04)	(1e-04)	(0)	(2e-04)	(1e-04)	(0)
Chi .5*	0.19	0.1155	0.0275	0.1221	0.0593	0.0092	0.1126	0.0507	0.0056
	(0.0207)	(0.0065)	(8e-04)	(0.003)	(8e-04)	(1e-04)	(0.0016)	(4e-04)	(0)
Chi .9*	0.2021	0.1246	0.0304	0.1243	0.0608	0.0097	0.1136	0.0513	0.0059
	(0.0361)	(0.0116)	(0.0014)	(0.0039)	(0.001)	(1e-04)	(0.002)	(5e-04)	(0)
Chi .95*	0.2064	0.1278	0.0314	0.125	0.0614	0.0099	0.114	0.0515	0.006
	(0.0426)	(0.0137)	(0.0017)	(0.0042)	(0.001)	(1e-04)	(0.0022)	(5e-04)	(0)
Chi .99*	0.2158	0.1346	0.0334	0.1266	0.0624	0.0103	0.1147	0.0519	0.0062
	(0.0586)	(0.0189)	(0.0022)	(0.0048)	(0.0012)	(1e-04)	(0.0024)	(5e-04)	(0)
MR .5*	0.1815	0.1085	0.0246	0.1193	0.0565	0.0076	0.1111	0.0495	0.0044
	(0.012)	(0.0037)	(3e-04)	(0.0017)	(3e-04)	(0)	(0.001)	(2e-04)	(0)
MR .9*	0.1932	0.1218	0.0264	0.1208	0.057	0.0078	0.1119	0.0496	0.0045
	(0.0299)	(0.0111)	(6e-04)	(0.0025)	(4e-04)	(0)	(0.0014)	(2e-04)	(0)
MR .95*	0.1999	0.128	0.0273	0.1215	0.0573	0.008	0.1123	0.0497	0.0046
	(0.0414)	(0.0155)	(8e-04)	(0.0028)	(5e-04)	(0)	(0.0015)	(2e-04)	(0)
MR .99*	0.2215	0.145	0.0296	0.123	0.0583	0.0084	0.113	0.05	0.0048
	(0.0894)	(0.0328)	(0.0012)	(0.0034)	(6e-04)	(0)	(0.0018)	(3e-04)	(0)
MR 2*	0.1816	0.1082	0.0244	0.1193	0.0565	0.0076	0.1112	0.0495	0.0044
	(0.0123)	(0.0036)	(2e-04)	(0.0018)	(3e-04)	(0)	(0.001)	(2e-04)	(0)
MR 2.3*	0.1842	0.1115	0.025	0.1197	0.0566	0.0076	0.1114	0.0495	0.0045
	(0.016)	(0.0051)	(4e-04)	(0.002)	(3e-04)	(0)	(0.0011)	(2e-04)	(0)
MR 2.5*	0.1863	0.1144	0.0256	0.12	0.0567	0.0077	0.1115	0.0496	0.0045
	(0.019)	(0.0067)	(5e-04)	(0.0021)	(4e-04)	(0)	(0.0012)	(2e-04)	(0)
MR 3*	0.1928	0.1231	0.0276	0.1208	0.0571	0.008	0.1119	0.0497	0.0046
	(0.0292)	(0.012)	(8e-04)	(0.0025)	(4e-04)	(0)	(0.0014)	(2e-04)	(0)

Tabelle B.31.: Simulationsergebnisse für Splines: L_2-Risiko und gewählte $\lambda \cdot 10^4$ für Zero.

n	50	100	1000
CV	0.0742	0.0395	0.0034
	(40814.3914)	(39363.4433)	(53302.8752)
GCV	0.079	0.0417	0.0034
	(37330.8605)	(39196.1607)	(53151.9403)
Cp	0.0976	0.0544	0.0036
	(40430.3481)	(40274.5203)	(52689.8356)
Chi .5	0.0789	0.0621	0.0139
	(44568.1754)	(37346.7552)	(39942.7582)
Chi .9	0.0534	0.0366	0.005
	(65068.7857)	(62077.6167)	(62811.8179)
Chi .95	0.0504	0.0327	0.0038
	(67245.5337)	(62834.9491)	(68010.9252)
Chi .99	0.046	0.0274	0.0027
	(69522.0819)	(66650.6525)	(70935.0786)
MR .5	0.0872	0.0626	0.015
	(43147.9118)	(39239.7491)	(36352.5143)
MR .9	0.0465	0.0275	0.0039
	(66538.694)	(65194.6831)	(64436.2844)
MR .95	0.0443	0.0263	0.003
	(69492.1904)	(68848.8552)	(67530.4673)
MR .99	0.0427	0.0245	0.002
	(72596.3741)	(72412.8021)	(72413.2525)
MR 2	0.0862	0.0647	0.023
	(43317.0533)	(36287.617)	(29628.3946)
MR 2.3	0.0663	0.0475	0.0098
	(56306.6155)	(48942.7833)	(49521.7279)
MR 2.5	0.0606	0.0367	0.0056
	(60611.4461)	(54128.3627)	(57841.8004)
MR 3	0.0466	0.0271	0.0027
	(65336.5678)	(65436.2243)	(68738.4792)
Cp*	0.0754	0.0405	0.0034
	(43303.9511)	(43196.7573)	(53303.776)
Chi .5*	0.0664	0.051	0.0113
	(45988.2801)	(42537.4314)	(38587.861)
Chi .9*	0.0429	0.0269	0.0028
	(70945.2002)	(69458.8011)	(68001.7712)
Chi .95*	0.0417	0.0251	0.0022
	(72461.4947)	(70206.1132)	(70099.1728)
Chi .99*	0.0417	0.0234	0.0018
	(73890.561)	(73151.6676)	(73890.561)
MR .5*	0.0683	0.0541	0.0149
	(50328.3519)	(45920.7796)	(34649.4385)
MR .9*	0.0452	0.025	0.0036
	(69462.6623)	(67267.6493)	(65303.2324)
MR .95*	0.0429	0.0238	0.0028
	(71724.9112)	(70362.0277)	(67818.1885)
MR .99*	0.0417	0.0225	0.0018
	(73890.561)	(73074.3453)	(72427.4214)
MR 2*	0.0672	0.0557	0.0207
	(50380.1388)	(45891.9635)	(26145.197)
MR 2.3*	0.0566	0.0404	0.0077
	(56810.4538)	(52997.7411)	(50417.2059)
MR 2.5*	0.0523	0.0312	0.005
	(62947.6612)	(58927.6435)	(61732.6937)
MR 3*	0.0453	0.0246	0.0026
	(69462.0102)	(68752.4211)	(68836.7674)

Literaturverzeichnis

[AGG02] Abdous, B., S. Germain und N. Ghazzali: *A unified treatment of direct and indirect estimation of a probability density and its derivatives.* Statistics and Probability Letters, 56:239–250, 2002.

[Aka74] Akaike, H.: *A new look at the statistical model identification.* IEEE Transactions on Automatic Control, 19:716–723, 1974.

[AV89] Aidu, F.A. und V.N. Vapnik: *Estimation of probability density on the basis of the method of stochastic regularization.* Automation and Remote Control, 50:499–509, 1989.

[Bar10] Barber, S.: *waveband: Computes credible intervals for Bayesian wavelet shrinkage*, 2010. R package version 4.5.

[BD94] Berlinet, A. und L. Devroye: *A Comparison of kernel density estimates.* Publications de l'Institute de Statistique de L'Universite de Paris, 38:3–59, 1994.

[BEH07] Bernholt, T., F. Eisenbrand und T. Hofmeister: *A geometric framework for solving subsequence problems in computational biology efficiently.* In: *Proceedings of the 23rd Annual ACM Symposium on Computational Geometry (SoCG 2007)*, Seiten 310–318, 2007.

[BH06] Bernholt, T. und T. Hofmeister: *An algorithm for a generalized maximum subsequence problem.* In: Correa, J.R., A. Hevia und M. Kiwi (Herausgeber): *Latin 2006: Theoretical Informatics*, Seiten 178–189, Berlin, 2006. Springer Verlag.

[BHMR07] Bissantz, N., T. Hohage, A. Munk und F. Ruymgaart: *Convergence rates of general regularization methods for statistical inverse problems and applications.* SIAM Journal of Numerical Analysis, 45:2610–2636, 2007.

[BKL+09] Boysen, L., A. Kempe, V. Liebscher, A. Munk und O. Wittich: *Consistencies and rates of convergence of jump penalized least squares estimators.* The Annals of Statistics, 37:157–183, 2009.

[BL06] Bickel, P.J. und B. Li: *Regularization in Statistics.* Test, 15:271–344, 2006.

[BL10] BAUER, F. und M. LUKAS: *Comparing parameter choice methods for regularization of ill-posed problems.* `http://www.bmath.de/FBauerPublications.html`, 2010.

[Bog07] BOGACHEV, V.I.: *Measure theory. Volume 2.* Springer, New York, 2007.

[Bow84] BOWMAN, A.W.: *An alternative method of cross-validation for the smoothing of density estimates.* Biometrika, 71:353–360, 1984.

[BR06] BIRGÉ, L. und Y. ROZENHOLC: *How many bins should be put in a regular histogram.* ESAIM: Probability and Statistics, 10:24–45, 2006.

[CCGM94] CAO, R., A. CUEVAS und W. GONZÁLEZ MANTEIGA: *A comparative study of several smoothing methods in density estimation.* Computational Statistics and Data Analysis, 17:153–176, 1994.

[CM99] CHAUDHURI, P. und J.S. MARRON: *SiZer for exploration of structure in curves.* Journal of the American Statistical Association, 94:807–823, 1999.

[CM00] CHAUDHURI, P. und J.S. MARRON: *Scale space view of curve estimation.* Annals of Statistics, 28:408–428, 2000.

[CR08] CELISSE, A. und S. ROBIN: *Nonparametric density estimation by exact leave-p-out cross-validation.* Computational Statistics and Data Analysis, 52:2350–2368, 2008.

[CW79] CRAVEN, P. und G. WAHBA: *Smoothing noisy data with spline functions. Estimating the correct degree of smoothing by the method of genaeralized cross-validation.* Numerische Mathematik, 31:377–403, 1979.

[Dav95] DAVIES, P. L.: *Data features.* Statistica Neerlandica, 49:185–245, 1995.

[Dav02] DAVIES, P.L.: *Statistical procedures and robust statistics.* Estadística, 162-163:1–27, 2002.

[Dav03] DAVIES, P. L.: *Approximating data and statistical procedures – I. Approximating data.* Technical Report 7/2003, SFB 475, University of Dortmund, Dortmund, Germany, 2003.

[Dav08] DAVIES, P. L.: *Approximating data (with discussion).* Journal of the Korean Statistical Society, 37:191–240, 2008.

[dB01] BOOR, C. DE: *A practical guide to splines. Revised edition.* Springer, New York, 2001.

[Dev83] DEVROYE, L.: *The equivalence of weak, strong and complete convergence in L_1 for kernel density estimates.* Annals of Statistics, 11:896–904, 1983.

[Dev89] DEVROYE, L.: *On the non-consistency of the L_2-cross-validated kernel density estimate.* Statistics and Probality Letters, 8:425–433, 1989.

[Dev97] DEVROYE, L.: *Universal smoothing factor selection in density estimation: Theory and practice (with discussion).* Test, 6:223–320, 1997.

[DG85] DEVROYE, L. und L. GYÖRFI: *Nonparametric density estimation. The L_1 view.* Wiley, New York, 1985.

[DGNW09] DAVIES, P. L., U. GATHER, D. J. NORDMAN und H. WEINERT: *A comparison of automatic histogram constructions.* ESAIM: Probability and Statistics, 13:181–196, 2009.

[DGW08] DAVIES, P.L., U. GATHER und H. WEINERT: *Nonparametric regression as an example of model choice.* Communications in Statistics – Simulation and Computation, 37:274–289, 2008.

[DH07] DETTE, H. und B. HETZLER: *Specification tests indexed by bandwidths.* Sankhya, 69:28–54, 2007.

[DJ94] DONOHO, D. L. und I.M. JOHNSTONE: *Ideal spatial adaption by wavelet shrinkage.* Biometrika, 81:425–455, 1994.

[DK01] DAVIES, P.L. und A. KOVAC: *Local extremes, runs, strings and multiresolution (with discussion and rejoinder).* The Annals of Statistics, 29:1–65, 2001.

[DK04] DAVIES, P.L. und A. KOVAC: *Densities, spectral densities and modality.* The Annals of Statistics, 32:1093–1136, 2004.

[DK09] DÜMBGEN, L. und A. KOVAC: *Extensions of smoothing via taut strings.* Electronic Journal of Statistics, 3:41–57, 2009.

[DKM09] DAVIES, P.L., A. KOVAC und M. MEISE: *Nonparametric regression, confidence regions and regularization.* The Annals of Statistics, 37:2597–2625, 2009.

[DM08] DAVIES, P.L. und M. MEISE: *Approximating data with weighted smoothing splines.* Journal of Nonparametric Statistics, 20:207–228, 2008.

[Don88] DONOHO, D.L.: *One-sided inference about functionals of a density.* The Annals of Statistics, 16:1390–1420, 1988.

[DS01] DÜMBGEN, L. und V.G. SPOKOINY: *Multiscale testing of qualitative hypotheses.* The Annals of Statistics, 29:124–152, 2001.

[Dur73] DURBIN, J.: *Distribution theory for tests based on the sample distribution function.* SIAM, Philadelphia, 1973.

[EHN00] Engel, H.W., M. Hanke und A. Neubauer: *Regularization of inverse problems.* Kluwer Academic Publishers, Dordrecht, 2000.

[EL96] Eggermont, P.P.B. und V.N. LaRiccia: *A simple and effective bandwidth selector for kernel density estimation.* Scandinavian Journal of Statistics, 23:285–301, 1996.

[EL01] Eggermont, P.P.B. und V.N. LaRiccia: *Maximum penalized likelihood estimation. Volume I: Density estimation.* Springer, New York, 2001.

[FP04] Frederix, G. und E.J. Pauwels: *A statistically principled approach to histogram segmentation.* Technical Report PNA-E0401, Centrum voor Wiskunde en Informatica, Amsterdam, 2004.

[GK92] Galatsanos, N.P. und A.K. Katsaggelos: *Methods for choosing the regularization parameter and estimating the noise variance in image restoration and their relation.* IEEE Transactions on Image Processing, 1:322–336, 1992.

[Gro96] Groeneboom, P.: *Lectures on inverse problems.* In: *Lectures on Probability Theory and Statistics. Ecole d'Eté de Probabilités de Saint-Flour XXXII*, Seiten 67–164. Springer, Berlin, 1996.

[GS94] Green, P.J. und B.W. Silverman: *Nonparametric regression and generalized linear models.* Chapmann and Hall, Boca Raton, 1994.

[Har97] Hart, J.D.: *Nonparametric smoothing and lack-of-fit tests.* Springer, New York, 1997.

[HH85] Hartigan, J.A. und P.M. Hartigan: *The dip test of unimodality.* The Annals of Statistics, 13:70–84, 1985.

[HH88] Hall, P. und E.J. Hannan: *On stochastic complexity and nonparametric density estimation.* Biometrika, 75:705–714, 1988.

[HMS+10] Hotz, T., P. Marnitz, R. Stichtenoth, L. Davies, Z. Kabluchko und A. Munk: *Locally adaptive image denoising by a statistical multiresolution criterion.* `http://arxiv.org/abs/1001.5447`, 2010.

[HMSW04] Härdle, W., M. Müller, S. Sperlich und A. Werwatz: *Nonparametric and semiparametric models.* Springer, Berlin, 2004.

[Hor09] Horowitz, J.L.: *Semiparametric and nonparametric methods in econometrics.* Springer, New York, 2009.

[HT87] Hall, P. und D.M. Titterington: *Common structure of techniques for choosing the smoothing parameter in regression.* Journal of the Royal Statistical Society. Series B, 49:184–198, 1987.

[HW01] Hjort, N.L. und S.G. Walker: *A note on kernel density estimators with optimal bandwidths.* Statistics and Probability Letters, 54:153–159, 2001.

[Kab08] Kabluchko, Z.: *Extreme-value analysis of standardized gaussian increments.* `http://arxiv.org/abs/0706.1849v3`, 2008.

[Kat04] Katznelson, Y.: *An introduction to harmonic analysis. 3rd edition.* Cambridge University Press, Cambridge, 2004.

[Kov02] Kovac, A.: *Robust nonparametric regression and modality.* In: Dutter, R., P. Filzmoser, U. Gather und P.J. Rousseeuw (Herausgeber): *Developments in Robust Statistics*, Seiten 218–227. Physica-Verlag, 2002.

[KS05] Kaipio, J. und E. Somersalo: *Statistical and computational inverse problems.* Springer, New York, 2005.

[Mar89] Markovich, N.M.: *Experimental analysis of nonparametric density estimates and of methods for smoothing them.* Automation and Remote Control, 50:941–948, 1989.

[Mas07] Massart, P.: *Concentration inequalities and model selection.* Springer, Berlin, 2007.

[Mäc95] Mächler, M.B.: *Variational Solution of Penalized Likelihood Problems and Smooth Curve Estimation.* The Annals of Statistics, 23:1496–1517, 1995.

[Mei04] Meise, M.: *Residual based selection of smoothing parameters.* Doktorarbeit, Universität Duisburg-Essen, 2004.

[Mei06] Meise, M.: *mrsmooth: multiresolution based smoothing methods*, 2006. R package version 0.1-0.

[Mei09] Meister, A.: *Deconvolution Problems in Nonparametric Statistics.* Springer, Berlin, 2009.

[Mil08] Mildenberger, T.: *A geometric interpretation of the multiresolution criterion in nonparametric regression.* Journal of Nonparametric Statistics, 20:599–609, 2008.

[Mor66] Morozov, V. A.: *On the solution of functional equations by the method of regularization.* Soviet Mathematics, 7:414–417, 1966.

[Mor68] Morozov, V. A.: *The error principle in the solution of operational equations by the regularization method.* USSR Computational Mathematics and Mathematical Physics, 8:63–87, 1968.

[MRZ09] Mildenberger, T., Y. Rozenholc und D. Zasada.: *histogram: Construction of regular and irregular histograms with different options for automatic choice of bins*, 2009. R package version 0.0-23.

[MvdG97] Mammen, E. und S. van de Geer: *Locally adaptive regression splines.* The Annals of Statistics, 25(1):387–413, 1997.

[MWT09] Mildenberger, T., H. Weinert und S. Tiemeyer: *benchden: 28 benchmark densities from Berlinet/Devroye (1994)*, 2009. R package version 1.0.3.

[Nad64] Nadaraya, E.A.: *On estimating regression.* Theory of Probability and its Applications, 9:141–142, 1964.

[NS94] Nason, G.P und B. W. Silverman: *The discrete wavelet transform in S.* Journal of Computational and Graphical Statistics, 3:163–191, 1994.

[Par62] Parzen, E.: *On the estimation of a probability density and the mode.* The Annals of Mathematical Statistics, 33:1065–1076, 1962.

[PR83] Prakasa Rao, B.L.S.: *Nonparametric functional estimation.* Academic Press, Orlando, 1983.

[R D] R Development Core Team: *R: A language and environment for statistical computing version 11.1 (und frühere Versionen).*

[RC00] Ruppert, D. und R.J. Carroll: *Spatially-adpative penalties for spline fitting.* Australian and New Zealand Journal of Statistics, 42:205–223, 200.

[Rei67] Reinsch, C.M.: *Smoothing by spline functions.* Numerische Mathematik, 10:177–183, 1967.

[RMG09] Rozenholc, Y., T. Mildenberger und U. Gather: *Combining regular and irregular histograms by penalized likelihood.* Discussion Paper 31/2009, SFB 823, Technische Universität Dortmund, 2009.

[RMG10] Rozenholc, Y., T. Mildenberger und U. Gather: *Combining regular and irregular histograms by penalized likelihood.* Computational Statistics and Data Analysis, 54:3313–3323, 2010.

[Ros56] Rosenblatt, M.: *Remarks on some nonparametric estimates of a density function.* The Annals of Mathematical Statistics, 27:832–837, 1956.

[Rud74] Rudin, W.: *Real and complex analysis. Second edition.* McGraw-Hill, New York, 1974.

[Rud82] Rudemo, M.: *Empirical choice of histograms and kernel density estimators.* Scandinavian Journal of Statistics, 9:65–78, 1982.

[RW10] RUFIBACH, K. und G. WALTHER: *The block criterion for multiscale inference about a density, with applications to other multiscale problems.* Journal of Computational and Graphical Statistics, 19:175–190, 2010.

[S+10] STEIN, W.A. et al.: *Sage mathematics software version 4.4.4 (und frühere Versionen).* The Sage Development Team, 2010.

[Sch78] SCHWARZ, G.: *Estimating the dimension of a model.* The Annals of Statistics, 6:461–464, 1978.

[Sch93] SCHWARZ, H.R.: *Numerische Mathematik. 3., überarbeitete und erweiterte Auflage.* Teubner, Stuttgart, 1993.

[Sco92] SCOTT, D.W.: *Multivariate density estimation. Theory, practice, and visualization.* Wiley, New York, 1992.

[Ser80] SERFLING, R.J.: *Approximation theorems of mathematical statistics.* Wiley, New York, 1980.

[SV95] SIEGMUND, D. O. und E. S. VENKATRAMAN: *Using the generalized likelihood ratio statistic for sequential detection of a change-point.* The Annals of Statistics, 23:255–271, 1995.

[TA77] TIKHONOV, A.N. und V.Y. ARSENIN: *Solutions of ill-posed problems.* Winston and Sons, Washington, D.C., 1977.

[Tit85] TITTERINGTON, D.M.: *Common structure of smoothing techniques in statistics.* International Statistical Review, 53:141–170, 1985.

[Tsy09] TSYBAKOV, A.B.: *Introduction to nonparametric estimation.* Springer, New York, 2009.

[Vap98] VAPNIK, V.N.: *Statistical learning theory.* Wiley, New York, 1998.

[Vap00] VAPNIK, V.N.: *The nature of statistical learning theory. Second edition.* Springer, New York, 2000.

[VMS92] VAPNIK, V.N., N.M. MARKOVICH und A.R. STEFANYUK: *Rate of convergence in L_2 of the projection estimator of the distribution density.* Automation and Remote Control, 53:677–686, 1992.

[VS78] VAPNIK, V.N. und A.R. STEFANYUK: *Nonparametric methods of reconstructing the probability density.* Automation and Remote Control, 39:1127–1140, 1978.

[Wah75] WAHBA, G.: *Smoothing noisy data with spline functions.* Numerische Mathematik, 24:383–393, 1975.

[Wah90] WAHBA, G.: *Spline models for observational data.* SIAM, Philadelphia, 1990.

[Wat64] WATSON, G.S.: *Smooth regression analysis.* Sankhya Series A, 26:359–372, 1964.

[WJ95] WAND, M.P und M.C. JONES: *Kernel smoothing.* Chapmann and Hall, Boca Raton, 1995.

[Yam73] YAMAMOTO, H.: *Uniform convergence of an estimator of a distribution function.* Bulletin of Mathematical Statistics, 15:69–78, 1973.