

ZUR AUSWAHL APPROXIMIERENDER DICHTEN,  
DIE LINEARKOMBINATIONEN ORTHOGONALER  
FUNKTIONEN SIND

8-Wochen-Arbeit im Rahmen der Prüfung für Diplom-  
Kaufleute an der Universität Göttingen

vorgelegt: 18.09.1987  
von: Achim Lewandowski  
aus: Uslar

# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>4</b>
<b>2</b>	<b>Grundlegende Begriffe der Modellauswahl</b>	<b>5</b>
2.1	Grundmodell, approximierende Modelle, Diskrepanzen . . . . .	5
2.2	Diskrepanz durch Approximation, Gesamtdiskrepanz . . . . .	6
2.3	Erwartete Gesamtdiskrepanz, Kriterium . . . . .	6
2.4	Empirische Diskrepanz, Minimum-Diskrepanz-Schätzer . . . . .	7
<b>3</b>	<b>Linearkombinationen von orthogonalen Funktionen als approximierende Dichten</b>	<b>7</b>
3.1	Eigenschaften von orthogonalen Funktionensystemen . . . . .	7
3.2	Herleitung von Kriterien . . . . .	8
3.3	Probleme bei der praktischen Anwendung . . . . .	10
<b>4</b>	<b>Ein Test zur Auswahl von Modellen</b>	<b>12</b>
4.1	Grundidee des Tests . . . . .	12
4.2	Entwicklung des Tests . . . . .	12
<b>5</b>	<b>Der Test für verschiedene orthogonale Funktionensysteme</b>	<b>15</b>
5.1	Gram-Charlier-Typ A . . . . .	15
5.2	Laguerre . . . . .	18
5.3	Legendre . . . . .	20
<b>6</b>	<b>Anwendungen</b>	<b>22</b>
6.1	Beispiel 1: Entrindungszeiten . . . . .	22
6.2	Beispiel 2: Höhen eines Baumbestandes . . . . .	26
6.3	Beispiel 3: Volumina eines Baumbestandes (Laguerre) . . . . .	30
6.4	Beispiel 4: Volumina (Legendre) . . . . .	34
6.5	Beispiel 5: Glucosekonzentrationen . . . . .	38
6.6	Anmerkungen . . . . .	42
<b>7</b>	<b>Zusammenfassung</b>	<b>43</b>
<b>8</b>	<b>Anhang</b>	<b>44</b>
8.1	Hermite-Polynome und ihre Quadrate . . . . .	44
8.2	Laguerre-Polynome und ihre Quadrate . . . . .	45
8.3	Legendre-Polynome und ihre Quadrate . . . . .	46
8.4	Daten . . . . .	47
8.4.1	Entrindungszeiten . . . . .	47
8.4.2	Höhen eines Bestandes . . . . .	47
8.4.3	Volumina eines Bestandes . . . . .	47
8.4.4	Glucosekonzentrationen . . . . .	47
8.5	Programmlisting - "ORTHO" . . . . .	48

<b>9 Danksagung</b>	<b>59</b>
<b>Literatur</b>	<b>60</b>

# 1 Einleitung

In verschiedensten Bereichen der Wirtschaft, in den Naturwissenschaften und in weiteren angewandten Disziplinen erhält man Daten als Ergebnis von Umfragen, Untersuchungen oder Messungen. Bei der Auswertung wird besonderes Gewicht auf das Erkennen der Struktur der Daten gelegt. Man ist daran interessiert, eine formale Darstellung zu finden, die die Verteilung der Beobachtungen möglichst gut erklärt.

Diese Überlegungen führen zur Aufstellung eines Modells. Die in der vorliegenden Arbeit genannten Verfahren lassen sich anwenden, falls man die Daten als Realisationen unabhängiger, identisch verteilter stetiger Zufallsvariablen ansehen kann. Ein Modell ist dann durch die Angabe der gemeinsamen Dichte- bzw. Verteilungsfunktion der Zufallsvariablen determiniert.

Abgesehen von den Fällen, bei denen die Verteilungsfunktion bekannt und leicht zu handhaben ist, steht man normalerweise vor dem Problem, ein komplexes Modell durch ein einfacheres zu ersetzen oder, falls nur unzureichende Informationen über die Zufallsvariablen zur Verfügung stehen, überhaupt erst einmal ein Modell anzupassen.

In beiden Situationen benötigt man Verfahren für die Entscheidung zwischen mehreren möglichen approximierenden Modellen. Die Modellauswahl, ein Teilgebiet der angewandten Statistik, gibt hierzu Hilfen.

Durch die Benutzung von orthogonalen Funktionensystemen kann in diversen Fällen der Prozeß der Auswahl entscheidend verkürzt und effektiver gestaltet werden. Ziel der Arbeit soll es sein, neben den Funktionsweisen Voraussetzungen für die Anwendbarkeit der genannten Methoden zu erarbeiten. In Kapitel 2 werden die wichtigsten Begriffe der Modellauswahl erläutert und grundlegende Methoden genannt, wie sie für das Verständnis der nachfolgenden Abschnitte erforderlich sind.

Durch die Benutzung von orthogonalen Funktionensystemen, dargestellt in 3., erreicht man eine schnelle Entscheidung zwischen einer großen Anzahl von potentiellen Modellen mit relativ geringem Rechenaufwand.

In Kapitel 4 wird ein Test zur Modellauswahl hergeleitet, der auf orthogonalen Funktionen basiert und eine verfeinerte Beurteilung der Qualität des ausgewählten Modells erlaubt.

An konkreten Beispielen werden in Kapitel 6 die Anwendungsbereiche von in 5. genannten Systemen demonstriert und gleichzeitig Vor- und Nachteile aufgezeigt.

Kapitel 7 schließt mit einer Zusammenfassung der erhaltenen Resultate ab.

Der Großteil der Terminologie und die Intentionen der Kapitel 2 bis 4 sind von Linhart ([1], [2]), bzw. Linhart und Zucchini [3] übernommen.

## 2 Grundlegende Begriffe der Modellauswahl

### 2.1 Grundmodell, approximierende Modelle, Diskrepanzen

Die in dieser Arbeit entwickelten Methoden lassen sich unter der Voraussetzung anwenden, daß man die zu untersuchenden Daten als Realisationen von identisch verteilten, unabhängigen stetigen Zufallsvariablen ansehen kann. Ein Modell ist durch die Angabe der gemeinsamen Verteilungs- bzw. Dichtefunktion der Zufallsvariablen vollständig beschrieben.

Für die gegebenen Beobachtungen wird ein Grundmodell postuliert. Die zugehörige Dichtefunktion beschreibt den tatsächlichen theoretischen Hintergrund für die Entstehung der Daten. Normalerweise ist höchstens die Modellfamilie, der das Grundmodell angehört, bekannt. Je mehr Informationen über die Zufallsvariablen vorliegen, umso enger läßt sich diese Familie eingrenzen.

Häufig ist es sinnvoll, auf die exakte Form der Verteilungsfunktion zu verzichten. Besitzt die Grundmodellfamilie viele freie Parameter, die aus einer Stichprobe geschätzt werden sollen, so entstehen, falls eine relativ geringe Anzahl von Beobachtungen vorliegt, Schätzfehler, die das Resultat unbrauchbar machen. Deshalb wählt man approximierende Modelle mit einfacherem Aufbau.

Wenn über die Grundmodellfamilie nur schwache Aussagen gemacht werden können, ist man ebenfalls in besonderem Maße an der Konstruktion eines approximierenden Modells zu einer gegebenen Stichprobe interessiert.

Für die Beurteilung einer solchen Approximation benötigt man ein Maß, das mit Diskrepanz bezeichnet wird. Ist  $F$  die gemeinsame Verteilungsfunktion aller Zufallsvariablen im völlig festgelegten Grundmodell und  $G_\theta$  die entsprechende Verteilungsfunktion der approximierenden Modellfamilie, so wird die Diskrepanz mit  $\Delta(G_\theta, F)$  oder nur mit  $\Delta(\theta)$  bezeichnet.  $\theta$  ist hier der Vektor der freien Parameter. Die Diskrepanz ist ein Funktional mit der Eigenschaft

$$\Delta(G_\theta, F) \geq \Delta(F, F) \quad \forall G_\theta . \quad (1)$$

Gute Approximationen sollten durch niedrige, schlechte durch hohe Werte angezeigt werden. In der Praxis werden verschiedene Diskrepanzen benutzt. Jede Diskrepanz gewichtet die Anpassung über den Definitionsbereich der Verteilungsfunktion anders, je nach Intention des Anwenders wird von Fall zu Fall eine geeignete Diskrepanz gewählt. Die in dieser Arbeit verwendete Diskrepanz hat die Form:

$$\Delta(G, F) = \int_{-\infty}^{+\infty} ((f(x) - g(x))^2 w(x) dx, \quad w(x) > 0 \quad (2)$$

Mit  $f$  und  $g$  werden die zugehörigen Dichtefunktionen bezeichnet, man sieht leicht, daß die Bedingung aus (2) erfüllt ist.

## 2.2 Diskrepanz durch Approximation, Gesamtdiskrepanz

Ist das Grundmodell völlig festgelegt, und hat man eine approximierende Modellfamilie  $G_\theta$ ,  $\theta \in \Theta$ , gewählt, so bezeichnet man mit  $\Delta(\theta_0)$  die Diskrepanz durch Approximation, wenn für  $\theta_0$  gilt :

$$\theta_0 = \arg \min \Delta(\theta) : \theta \in \Theta \quad (3)$$

Es soll angenommen werden, daß  $\theta_0$  existiert und eindeutig ist;  $\Theta$  bezeichnet die Menge der möglichen Parametervektoren.

Das zugehörige Modell  $G_{\theta_0}$  nähert von allen  $G_\theta$ ,  $\theta \in \Theta$ , bezüglich der gewählten Diskrepanz das Grundmodell am besten an. Enthält die approximierende Familie das Grundmodell, so gilt mit den Bezeichnungen aus 2.1:

$$\Delta(\theta_0) = \Delta(F, F) \quad (4)$$

In diesem Fall ist  $G_{\theta_0}$  gerade das Grundmodell. Besitzt eine approximierende Modellfamilie viele Parameter, so sollte sich diese Flexibilität in einer niedrigen Diskrepanz durch Approximation ausdrücken.

In der Praxis ist das völlig festgelegte Grundmodell selten bekannt, eine Berechnung von  $\theta_0$  ist nicht möglich. Es steht nur eine Stichprobe der Größe  $n$  zur Verfügung, mit deren Hilfe ein Schätzer  $\hat{\theta}$  für  $\theta_0$  angegeben werden kann. Mit

$$\Delta(G_{\theta_0}, G_{\hat{\theta}}) \quad (5)$$

wird die Diskrepanz durch Schätzung bezeichnet und mit

$$\Delta(F, G_{\hat{\theta}}) \text{ bzw. } \Delta(\hat{\theta}) \quad (6)$$

die Gesamtdiskrepanz.

Bei realistischen Anwendungen ist die Gesamtdiskrepanz die entscheidende Größe. Sie tendiert dazu, bei wachsender Parameterzahl zunächst durch den Einfluß der geringeren Diskrepanz durch Approximation zu sinken, dann aber aufgrund der zunehmenden Schätzfehler anzusteigen.

## 2.3 Erwartete Gesamtdiskrepanz, Kriterium

Nimmt man wiederholt gleichgroße Stichproben aus der-selben Grundgesamtheit, so erhält man im Prinzip verschiedene Schätzer  $\hat{\theta}$  und somit auch unterschiedliche Gesamtdiskrepanzen  $\Delta(\hat{\theta})$ .  $\Delta(\hat{\theta})$  ist daher Zufallsvariable und besitzt eine Verteilung unter dem Grundmodell. Soll zwischen mehreren approximierenden Modellfamilien aufgrund einer Größe entschieden werden, so kann man die Familie wählen, für die  $E_F \Delta(\hat{\theta})$  minimiert wird.  $E_F \Delta(\hat{\theta})$  wird mit erwarteter (Gesamt-) Diskrepanz bezeichnet.

Die genaue Verteilung von  $\hat{\theta}$  ist in der Praxis nicht gegeben oder schwer zu handhaben,  $E_F \Delta(\hat{\theta})$  läßt sich daher nur mit Mühe oder überhaupt nicht exakt ermitteln. Kann  $E_F \Delta(\hat{\theta})$  aus einer Stichprobe geschätzt werden, so sollte mit Hilfe dieses Schätzers  $\hat{E}_F \Delta(\hat{\theta})$ , des Kriteriums, die geeignetste approximierende Modellfamilie ausgewählt werden.

## 2.4 Empirische Diskrepanz, Minimum-Diskrepanz-Schätzer

Eine natürliche Möglichkeit,  $\theta_0$  zu schätzen, erhält man, indem die Verteilungsfunktion  $F$  durch die empirische Verteilungsfunktion  $F_n$  ersetzt wird. Sind die  $x_i$  ( $i = 1, 2, \dots, n$ ) die vorliegenden Beobachtungen, so gilt:

$$F_n(x) = \text{card}\{i : x_i \leq x\}/n \quad (7)$$

$\Delta_n(\theta) = \Delta(F_n, G_\theta)$ , die empirische Diskrepanz, ist somit ein Schätzer für  $\Delta(\theta)$ .

$$\hat{\theta} = \arg \min\{\Delta_n(\theta) : \theta \in \Theta\} \quad (8)$$

heißt Minimum-Diskrepanz-Schätzer von  $\theta_0$ .

## 3 Linearkombinationen von orthogonalen Funktionen als approximierende Dichten

### 3.1 Eigenschaften von orthogonalen Funktionensystemen

Die Darstellung von Dichten als Linearkombinationen von orthogonalen Funktionen ist in der Statistik seit längerem ein gebräuchliches Mittel. Für die Modellauswahl bringt die Verwendung von orthogonalen Funktionen den entscheidenden Vorteil, daß das Kriterium in Terme zerlegt werden kann, die den einzelnen Parametern direkt zurechenbar sind.

Sei die Dichte eines Elementes einer approximierenden Familie gegeben:

$$g_\theta(x) = \sum_{i \in A} \theta_i \phi_i(x) \quad (9)$$

wobei  $\phi(x)$ ,  $i = 0, 1, \dots$  eine beliebige Funktionenfolge und  $A$  eine endliche Menge von nichtnegativen, ganzen Zahlen sind. Die Zahl der freien Parameter ist identisch mit  $\text{card}(A)$ . Unter der Annahme, daß z.B. die Familien betrachtet werden, die  $\phi_0, \phi_1, \dots, \phi_k$  ( $k \in \mathbb{N}_0$ ) oder Teilmengen hiervon benutzen, so müßte man einen Vergleich zwischen  $2^{k+1} - 1$  approximierenden Modellfamilien durchführen und ebenso viele Kriterien berechnen. Es ist einleuchtend, daß mit steigender Parameterzahl der Rechen- und Arbeitsaufwand unverhältnismäßig ansteigt.

Verwendet man hingegen orthogonale Funktionen  $\phi_i$ , so ist es möglich, für jeden Parameter den zugehörigen Anteil am Kriterium zu errechnen. Aufgrund dieses Anteils kann, wie weiter unten gezeigt wird, über die Aufnahme von  $\phi_i$  in das approximierende Modell entschieden werden. In der Tat steigt der Rechenaufwand mit wachsender Parameterzahl nur noch linear an.

Möchte man die Vorteile der Orthogonalität nutzen, so kann ein Orthogonalsystem von Funktionen nicht mit jeder beliebigen Diskrepanz verknüpft werden, es gibt zu jedem System eine zugeordnete Diskrepanz. Für die Anwendung

werden deshalb die Funktionensysteme so ausgesucht, daß eine Approximation mit wenigen Parametern möglich erscheint und die zugehörige Diskrepanz sinnvoll gewählt ist.

Sei nun eine Dichtefunktion gegeben, deren Werte auf dem Intervall  $[a, b]$  nichtnegativ sind,  $a \in \mathbb{R} \cup \{-\infty\}$  und  $b \in \mathbb{R} \cup \{+\infty\}$ . Die approximierenden Dichten besitzen die Gestalt

$$g_\theta(x) = \sum_{i \in A} \theta_i \phi_i(x), \quad (10)$$

wobei die  $\phi_i(x)$  orthogonale Funktionen bzgl. einer Gewichtungsfunktion  $w(x) > 0$  sind:

$$\int_a^b \phi_i(x) \phi_j(x) w(x) dx = \delta_{ij} K(i), \quad \forall i, j \in A. \quad (11)$$

Hier ist  $K(i) > 0$  eine Konstante und  $\delta_{ij}$  das Kronecker-Symbol. Bei den in Kapitel 5 behandelten Funktionensystemen ist  $\phi_0 = 1/w(x)$ , und das Integral über  $\theta_0 \phi_0$  beträgt eins. Durch Einsetzen von  $i = 0$  und  $j > 0$  in (11) kann man folgern, daß das Integral über  $g_\theta(x)$  tatsächlich eins ist. Die Frage, ob die andere Bedingung für eine Dichtefunktion,  $g_\theta(x) \geq 0$ , immer erfüllt ist, wird in Kapitel 6 beantwortet.

Bemerkung: In der Literatur ist auch das Arbeiten mit *Orthonormalen*-systemen üblich. Man erhält es aus obigem Orthogonalsystem durch die Transformationen

$$\phi_i^*(x) = \phi_i(x) (w(x)/K(i))^{1/2} \quad \forall i \in A. \quad (12)$$

Mit diesen Bezeichnungen gilt:

$$\int_a^b \phi_i^*(x) \phi_j^*(x) dx = \delta_{ij} \quad \forall i, j \in A, \quad (13)$$

d.h.  $w^*(x) = 1$  und  $K^*(i) = 1$  für  $i \in A$ .

### 3.2 Herleitung von Kriterien

Für ein gegebenes System  $\{\phi_i(x)\}$  bestimmt die Wahl der Indexmenge  $A$  die approximierende Familie. Die Auswahl der Modellfamilie erfolgt auf der Basis der zu minimierenden Diskrepanz, die mit dem orthogonalen Funktionensystem verknüpft ist:

$$\Delta^+ F, G_\theta = \Delta^+(\theta) = \int_a^b (f(x) - g_\theta(x))^2 w(x) dx \quad (14)$$

Mit  $f$  wird die Dichtefunktion des Grundmodells bezeichnet. Der Ausdruck

$$U := \int_a^b f^2(x) w(x) dx \quad (15)$$



ist unabhängig von der Wahl von  $A$ , deshalb genügt es,

$$\Delta(\theta) = \Delta^+(\theta) - U \quad (16)$$

$$= \int_a^b \left( -2f(x) \sum_{i \in A} \theta_i \phi_i(x) + \left( \sum_{i \in A} \theta_i \phi_i(x) \right)^2 \right) w(x) dx \quad (17)$$

zu betrachten. Unter Ausnutzung der Orthogonalitätseigenschaften aus (11) läßt sich dieser Term umformen:

$$\begin{aligned} \Delta(\theta) &= - \sum_{i \in A} 2\theta_i \int_a^b \phi_i(x) w(x) dF(x) \\ &\quad + \int_a^b \left( \sum_{i \in A} \theta_i \phi_i(x) \right) \left( \sum_{j \in A} \theta_j \phi_j(x) \right) w(x) dx \\ &= - \sum_{i \in A} 2\theta_i \int_a^b \phi_i(x) w(x) dF(x) + \sum_{i \in A} \theta_i^2 K(i), \end{aligned} \quad (18)$$

deshalb folgt für  $\Delta(\theta)$ :

$$\Delta(\theta) = - \sum_{i \in A} \left( 2\theta_i \int_a^b \phi_i(x) w(x) dF(x) - \theta_i^2 K(i) \right). \quad (19)$$

$\Delta(\theta)$  nimmt daher sein Minimum, Existenz vorausgesetzt, an für

$$\theta_{0i} = 1/K(i) \int_a^b \phi_i(x) w(x) dF(x) \quad , \quad i \in A \quad (20)$$

Durch Einsetzen in (19) folgt:

$$\Delta(\theta) = \sum_{i \in A} \left( (\theta_i - \theta_{0i})^2 - \theta_{0i}^2 \right) K(i). \quad (21)$$

Ein erwartungstreuer Schätzer für  $\theta_{0i}$ , den man mit Hilfe einer Stichprobe  $x_1, x_2, \dots, x_n$  angeben kann, ist der Minimum-Diskrepanz-Schätzer

$$\begin{aligned} \hat{\theta}_i &= 1/K(i) \int_a^b \phi_i(x) w(x) dFn(x) \quad , \quad i \in A \\ &= \frac{1}{nK(i)} \sum_{j=1}^n \phi_i(x_j) w(x_j) \quad , \quad i \in A. \end{aligned} \quad (22)$$

Die Gesamtdiskrepanz beträgt in diesem Fall

$$\Delta(\hat{\theta}) = \sum_{i \in A} \left( (\hat{\theta}_i - \theta_{0i})^2 - \theta_{0i}^2 \right) K(i), \quad (23)$$

und demnach die erwartete Gesamtdiskrepanz:

$$E\Delta(\hat{\theta}) = \sum_{i \in A} \left( E(\hat{\theta}_i - \theta_{0i})^2 - \theta_{0i}^2 \right) K(i), \quad (24)$$

Da  $E\hat{\theta}_i = \theta_{0i}$  ist, folgt:

$$E\Delta(\hat{\theta}) = \sum_{i \in A} (Var \hat{\theta}_i - \theta_{0i}^2) K(i) \quad (25)$$

Nun ist  $E(\hat{\theta}_i^2 - Var \hat{\theta}_i) = E(\hat{\theta}_i^2) - Var \hat{\theta}_i = \theta_{0i}^2$ , d.h. man kann einen Schätzer für den Anteil eines Parameters  $\theta_i$  am Kriterium angeben:

$$(2Var \hat{\theta}_i - \hat{\theta}_i^2) K(i) \quad (26)$$

Hierbei ist  $Var \hat{\theta}_i$  ein unverzerrter Schätzer für  $Var \hat{\theta}_i$ .

Eine Funktion  $\phi_i$  wird zur Approximation der Dichtefunktion  $f$  aufgenommen, falls der zugeordnete Beitrag zum Kriterium negativ ist; da  $K(i) > 0$ , bedeutet dies, daß

$$\hat{\theta}_i^2 / Var \hat{\theta}_i > 2 \quad (27)$$

sein muß. Das gewählte  $A_*$  besteht dann aus sämtlichen  $i$ , für die (27) erfüllt ist. Die approximierende Dichte hat dann die Gestalt:

$$g_{\hat{\theta}}(x) = \sum_{i \in A_*} \hat{\theta}_i \phi_i(x) \quad (28)$$

### 3.3 Probleme bei der praktischen Anwendung

Es sei zunächst festgehalten, daß die eben genannte Methode nicht zu einer Auswahl zwischen Funktionensystemen führt, es werden nur Modelle innerhalb eines solchen Systems verglichen. Das Verfahren aus 3.2 bietet aber neben dem relativ geringen Rechenaufwand den Vorteil, daß die  $\hat{\theta}_i$  nicht von der gewählten Indexmenge  $A$  abhängen.

Größere Schwierigkeiten bereitet es, Aussagen über die Indexmenge  $I = \{0, 1, \dots, i_{max}\}$  herzuleiten, für deren Elemente  $i \in I$  die zugehörigen Terme  $\hat{\theta}_i^2 / Var \hat{\theta}_i$  berechnet werden sollen. Die in Kapitel 5 vorgestellten Funktionensysteme sind vollständig über ihren Definitionsbereich  $\mathcal{D}$  (Courant-Hilbert, [4]), d.h., daß für jede *bekannte* stetige Dichtefunktion  $f$  auf  $\mathcal{D}$  und jedes  $\epsilon > 0$   $i_{max}$  und  $\theta_i$ 's existieren, so daß gilt:

$$g_{\theta}(x) = \sum_{i=0}^{i_{max}} \theta_i \phi_i(x) \quad \text{und} \quad \Delta(F, G_{\theta}) \leq \Delta(F, F) + \epsilon \quad (29)$$

Die Diskrepanz durch Approximation läßt sich durch die Wahl von  $i_{max}$  dem prinzipiell möglichen Wert  $\Delta(F, F)$  beliebig genau nähern. In realistischen Anwendungen ist von vornherein nicht klar, wie man  $i_{max}$  zu wählen hat.

Kronmal und Tarter [5], bzw. Tarter und Kronmal [6] schlagen nun folgende Regeln vor:

a) Man bricht mit der Berechnung der Größen aus (27) ab, sobald  $t$  aufeinanderfolgende Terme die Bedingung aus (27) nicht erfüllen.

b) Jeder Ausdruck aus (27), der diesen Bedingungen genügt, wird zur Approximation aufgenommen (Tarter und Kronmal wählen eine Bedingung, die sich nur geringfügig von (27) unterscheidet).

Diggle und Hall ([7]) nehmen zu diesen Regeln Stellung: Die Schwierigkeiten,  $i_{max}$  zu bestimmen, werden bei a) auf die Herleitung des optimalen Werts von  $t$  umgewälzt, die mathematisch noch anspruchsvoller ist. Für komplizierte Dichtefunktionen mit vielen "Peaks" führt ein zu niedriger Wert von  $t$  zu Inkonsistenzen. Diggle und Hall räumen jedoch ein, daß etwa bei  $t = 1$  und symmetrischen, unimodalen Verteilungen zumindest beim Ansatz mit der trigonometrischen Reihe gute Ergebnisse zu erwarten sind.

Zu Regel b) führen sie einen Beweis an, daß im Falle der trigonometrischen Reihe die Wahrscheinlichkeit mehr als 0.15 beträgt, daß für unendlich viele  $i$ 's die Bedingung aus (27) erfüllt ist.

Ob sich diese Ergebnisse direkt auf die in Kapitel 5 dargestellten Funktionensysteme übertragen lassen, sei dahingestellt, fest steht nur, daß es keine allgemeingültige Formel für  $i_{max}$  gibt, die sämtliche Verläufe von Dichtefunktionen berücksichtigt.

Für den Praktiker bleibt deshalb nur das heuristische Verfahren, durch Betrachtung des Histogramms ein in Frage kommendes Funktionensystem zu wählen,  $i_{max}$  festzusetzen und anschließend die  $\theta_i$ 's und die Prüfgrößen  $\hat{\theta}_i^2 / \hat{Var} \hat{\theta}_i$  zu berechnen. Aus der Tendenz dieser Werte, unter Berücksichtigung der zur Verfügung stehenden Rechengenauigkeit, und der anschließenden Beurteilung der Approximation wird das genannte Vorgehen entweder nachträglich gerechtfertigt oder verworfen.

Hilfen für die Wahl der Funktionensysteme, die zu einer höheren Erfolgsquote bei der Approximation führen, werden in allgemeiner Art in Kapitel 5 und für konkrete Beispiele in Kapitel 6 gegeben.

## 4 Ein Test zur Auswahl von Modellen

### 4.1 Grundidee des Tests

Linhart [2] entscheidet zwischen Modellen nicht nur aufgrund eines Vergleichs absoluter Werte von Kriterien, sondern er entwickelt einen Test. Die Hypothese lautet, daß die erwartete Gesamtdiskrepanz einer bestimmten Modellfamilie größer als die einer anderen ist. Im Fall von orthogonalen Funktionensystemen wird eine Funktion  $\phi_i$  in das approximierende Modell aufgenommen, wenn der zugehörige Beitrag zur erwarteten Gesamtdiskrepanz negativ ist. Nach (25) ist dies gleichbedeutend mit

$$\theta_{0i}^2 / \text{Var}(\hat{\theta}_i) > 1 \quad (30)$$

Verwendet man die Hypothese

$$H_o : \quad \theta_{0i}^2 / \text{Var}(\hat{\theta}_i) = 1 \quad (31)$$

und benutzt einen einseitigen Ablehnungsbereich, rechts von eins gelegen, so ist dies gleichzeitig ein Test für die Hypothese

$$H_o : \quad \theta_{0i}^2 / \text{Var}(\hat{\theta}_i) < 1 \quad (32)$$

daß die Verwendung von  $\phi_i$  in der Modellfamilie zu einem Anstieg der erwarteten Diskrepanz führt. Bei diesem Test wird ein Term  $\hat{\theta}_i \phi_i$  zur Approximation benutzt, wenn die obenstehende Hypothese aus (31) verworfen werden kann.

In diesem Ansatz kommt der Wille zum Ausdruck, möglichst überschaubare Modelle zu verwenden, solange es keinen zwingenden Grund gibt, Modelle mit komplexerem Aufbau zu den gegebenen Daten aufzustellen. Ein Vorteil dieses Tests liegt darin, daß man das einfachere Modell nicht explizit angeben muß. Über die Aufnahme von  $\theta_i$  wird entschieden, ohne daß das genaue Aussehen des simpleren, bzw. des komplexeren Modells in das Testverfahren eingeht. Der Anwender hat deshalb theoretisch freie Hand bei der Auswahl der Reihenfolge der Tests für die einzelnen Funktionen  $\phi_i$ , obgleich er in der Praxis nach aufsteigenden Indizes vorgehen wird.

### 4.2 Entwicklung des Tests

Um einen Test herzuleiten, benötigt man neben Erwartung und (geschätzter) Varianz weitergehende Aussagen über die Verteilung von  $\hat{\theta}_i$ . Da in der Anwendung das völlig festgelegte Grundmodell nicht bekannt ist, besteht keine Möglichkeit, die genaue Verteilung zu bestimmen. Mittels Wahrscheinlichkeitstheorie kann jedoch eine asymptotische Verteilung von  $\hat{\theta}_i$  abgeleitet werden.

Die unabhängigen und identisch verteilten Zufallsvariablen  $\phi_i(X)w(X)/K(i)$  besitzen unter dem Grundmodell  $F$  den Erwartungswert  $\theta_{0i}$  und die Varianz

$$\sigma_i^2 = 1/K^2(i) \int_a^b \phi_i^2(x)w^2(x)dF(x) - \theta_{0i}^2 . \quad (33)$$

Für den Schätzer (siehe (22))

$$\begin{aligned}\hat{\theta}_i &= \frac{1}{nK(i)} \sum_{j=1}^n \phi_i(x_j)w(x_j) \\ &= \frac{1}{n} \sum_{j=1}^n \phi_i(x_j)w(x_j)/K(i)\end{aligned}\quad (34)$$

folgt unter Verwendung des Zentralen Grenzwertsatzes:

$$\hat{\theta}_i \text{ ist asymptotisch verteilt wie } N(\theta_{0i}, \sigma_i^2/n) \quad (35)$$

oder kürzer:

$$\hat{\theta}_i \sim AN(\theta_{0i}, \sigma_i^2/n) \quad (36)$$

Gilt

$$H_0 : \quad \theta_{0i} = \sigma_i^2/n, \quad (37)$$

so folgt:

$$\hat{\theta}_i \sim AN(\theta_{0i}, \theta_{0i}^2). \quad (38)$$

Unter  $H_0$  kann  $\hat{\theta}_i^2/(\sigma_i^2/n)$  deshalb als nichtzentral Chi-Quadrat-verteilt angesehen werden, mit einem Freiheitsgrad und Nichtzentralitätsparameter eins.  $\sigma_i^2$  kann durch  $\hat{\sigma}_i^2$  geschätzt werden:

$$\hat{\sigma}_i^2 = 1/(nK^2(i)) \sum_{j=1}^n \phi_i^2(x_j)w^2(x_j) - \hat{\theta}_i^2 \quad (39)$$

Einen approximierenden Test zu (37) erhält man nun, falls die Hypothese verworfen wird, wenn  $\hat{\theta}_i^2/(\hat{\sigma}_i^2/n)$  größer als der vorgegebene kritische Punkt von  $\chi^2(1, 1)$  ist. Ein Parameter  $\theta_i$  wird wiederum beim Verwerfen der Hypothese benutzt.

Nach (27) wird bei einem bloßen Vergleich von Kriterien ein Term  $\theta_i\phi_i$  in das Modell einbezogen, wenn

$$\hat{\theta}_i^2 / \widehat{Var} \hat{\theta}_i > 2, \quad (40)$$

wobei  $\widehat{Var} \hat{\theta}_i$  aus einem exakten Ausdruck für  $Var \hat{\theta}_i$  entwickelt wird.

Die Herleitung dieses Schätzers ist nicht immer einfach, es bietet sich daher an, auch in Kapitel 3  $Var \hat{\theta}_i$  durch die asymptotische Varianz von  $\hat{\theta}_i$  zu ersetzen:  $\theta_i\phi_i$  wird zur Approximation verwendet für

$$\hat{\theta}_i^2/(\hat{\sigma}_i^2/n) > 2 \quad (41)$$

Sei  $Z$  eine Zufallsvariable mit  $Z \sim \chi^2(1, 1)$ , dann ist  $\alpha = P(Z \geq 2) \approx 0.35$  die ungefähre Irrtumswahrscheinlichkeit der Hypothese (37) im oben angeführten Test; man muß berücksichtigen, daß  $\hat{\sigma}_i^2$  anstelle von  $\sigma_i^2$  verwendet wird.

$\alpha$	Kritischer Wert
0.01	11.1
0.05	7.0
0.10	5.2
0.35	2.0
0.50	1.1

Tabelle 1: Kritische Werte

Die Methode aus Kapitel 3 läßt sich daher als Spezialfall des Tests aus Kapitel 4 ansehen, deshalb genügt es, sich bei der Anwendung auf das eben genannte Testverfahren zu beschränken.

Will man sichergehen, daß die aufgenommenen Funktionen keinen Schaden anrichten, so verringert man  $\alpha$  und erhöht somit den kritischen Punkt der  $\chi^2(1, 1)$ -Verteilung. Gebräuchliche Irrtumswahrscheinlichkeiten und die zugehörigen kritischen Werte stehen in der folgenden Tabelle (ableitbar aus den Tabellen der Standard-Normalverteilung): Möchte man eher die charakteristischen Eigenschaften der vorliegenden Stichprobe erfassen, so ist es angebracht,  $\alpha$  heraufzusetzen, etwa  $\alpha = 0.50$ . Bei manchen Anwendungen kann dies jedoch zu einem komplizierten Modell führen, da relativ viele Parameter in das approximierende Modell aufgenommen werden (siehe auch Kapitel 6). Interessiert sich der Anwender für das grundlegende Phänomen der Daten, wird er  $\alpha$  verringern, um die stichproben-spezifischen Schwankungen auszufiltern.

Die in 3.3 genannten Probleme der Bestimmung von  $i_{max}$  lassen sich ohne größere Einschränkungen auf dieses Kapitel übertragen.

## 5 Der Test für verschiedene orthogonale Funktionensysteme

### 5.1 Gram-Charlier-Typ A

Ein Funktionensystem, das häufig anwendbar ist, wird durch sukzessive Differentiation der Standard-Normalverteilungsdichte erzeugt. Mit

$$\alpha(x) = \frac{1}{(2\pi)^{1/2}} \exp(-x^2/2) \quad (42)$$

hat die approximierende Dichte die Form:

$$g_\theta(x) = \sum_{i \in A} \theta_i H_i(x) \alpha(x) \quad (43)$$

Kendall und Stuart [8] haben die Eigenschaften der Tschebyscheff-Hermite-Polynome  $H_i(x)$  explizit aufgeführt. Die Hermite-Polynome sind durch folgende Gleichungen rekursiv definiert:

$$\begin{aligned} H_0(x) &= 1, \\ H_1(x) &= x \text{ und} \\ H_i(x) &= xH_{i-1}(x) - (i-1)H_{i-2}(x) \quad , \quad i \geq 2 \end{aligned} \quad (44)$$

Die ersten neun Polynome sind im Anhang gegeben. Mit der Notation aus Kapitel 3 und der Berücksichtigung der Orthogonalitätseigenschaft

$$\int_{-\infty}^{+\infty} H_i(x) H_j(x) \alpha(x) dx = \delta_{ij} i! \quad , \quad i, j \in \mathbb{N}_0, \quad (45)$$

ergibt sich

$$\begin{aligned} a &= -\infty \quad b = +\infty \\ \phi_i(x) &= H_i(x) \alpha(x) \\ w(x) &= (\alpha(x))^{-1} \\ K(i) &= i! \quad \text{für } i \in \mathbb{N}_0. \end{aligned} \quad (46)$$

Ist  $F$  die Verteilungsfunktion, bzw.  $f$  die Dichtefunktion des Grundmodells, so lautet die Diskrepanz (vgl.(14)):

$$\Delta^+(\theta) = \int_{-\infty}^{+\infty} \left( f(x) - \sum_{i \in A} \theta_i H_i(x) \alpha(x) \right)^2 / \alpha(x) dx \quad (47)$$

Analog zu (19) verwendet man ersatzweise die Diskrepanz

$$\Delta(\theta) = - \sum_{i \in A} \left( 2\theta_i \int_{-\infty}^{+\infty} H_i(x) dF(x) - \sigma_i^2 i! \right). \quad (48)$$

Da  $w(x) = 1/\alpha(x)$ , wird die approximierende Dichte im Prinzip besonders für große Beträge von  $x$  die Dichte  $f$  gut annähern.

Der Wert von  $\Delta^+(\theta)$  ist abhängig von den Einheiten, in denen die Beobachtungen vorliegen. Ein sinnvoller Gebrauch der Diskrepanz ist normalerweise nur gewährleistet, falls Anpassungen an die standardisierte Dichte des Grundmodells, d.h. an die Dichte von  $(X - EX)/(VarX)^{1/2}$  betrachtet werden. Die standardisierten Beobachtungen  $z_i = (x_i - \bar{x})/s$  werden dann als Realisationen von unabhängigen Zufallsvariablen mit der standardisierten Dichte angesehen.

Seien von nun an  $f$  die standardisierte Dichte und  $z_i$  ( $i = 1, \dots, n$ ) die vorliegenden Beobachtungen der zugehörigen Verteilung. Nach (20) folgt für die "besten"  $\theta_i$ :

$$\theta_{0i} = 1/i! \int_{-\infty}^{+\infty} H_i(x) dF(x), \quad i \in A \quad (49)$$

Da  $H_0(x) = 1$ , ist  $\theta_{00}$  unabhängig von der genauen Gestalt von  $f$ :

$$\theta_{00} = 1 \quad (50)$$

$f$  liegt standardisiert vor, deshalb folgt ebenso:

$$\theta_{01} = \int_{-\infty}^{+\infty} x dF(x) = 0 \quad (51)$$

und

$$\theta_{01} = 1/2 \int_{-\infty}^{+\infty} (x^2 - 1) dF(x) = 0 \quad (52)$$

Jedes weitere  $\theta_{0i}$  ist nach (49) Linearkombination der Momente  $\mu_h$  ( $h = 0, 1, \dots, i$ ) der standardisierten Verteilung, deren Existenz bis zu genügend hohem Grad angenommen wird.

Aus Eigenschaft (50) folgt, daß  $\phi_0$  auf jeden Fall zur Approximation der Dichtefunktion  $f$  verwendet wird. Diese Basisfunktion ist bei der Gram-Charlier-Reihe vom Typ A gerade die Dichte der Standard-Normalverteilung. Wegen (50) wird deshalb ein Teilbereich der Anwendung determiniert: Kann aufgrund der vorliegenden Beobachtungen eine Normalverteilung oder eine ähnliche Verteilung angenommen werden, so ist es sinnvoll  $g_\theta$  wie in (43) anzusetzen. Genaugenommen sind bessere Resultate zu erwarten, falls  $f$  einer Linearkombination

$$g_\theta(x) = \sum a_i x^i \alpha(x), \quad a_i \in \mathbb{R} \quad (53)$$

ähnlich ist, doch Dichtefunktionen dieser Darstellung können recht unterschiedlich aussehen und sind etwa aufgrund eines Histogramms mangels Übung schwer zu erkennen.

Durch (50) wird die oben durchgeführte Standardisierung nachträglich gerechtfertigt. Varianz und Erwartungswert stimmen mit den entsprechenden zugeordneten Werten von  $\alpha(x)$  überein.



Analog zu (22) werden die Minimum-Diskrepanz-Schätzer hergeleitet:

$$\begin{aligned}\hat{\theta}_i &= 1/(i!n) \sum_{j=1}^n H_i(z_j) \\ &= 1/i! \sum_{j=1}^n (H_i(z_j)/n), \quad i \geq 3\end{aligned}\tag{54}$$

Seien  $m_h = 1/n \sum_{j=1}^n z_j^h$  die Stichprobenmomente, dann werden die  $\mu_h$  ( $h = 1, 2, \dots$ ) durch die  $m_h$  erwartungstreu geschätzt. Die ersten  $\theta_{0i}$  lauten (siehe (49) und Anhang):

$$\begin{aligned}\theta_{03} &= 1/6\mu_3 \\ \theta_{04} &= 1/24(\mu_4 - 6\mu_2 + 3) = 1/24(\mu_4 - 3) \\ \theta_{05} &= 1/120(\mu_5 - 10\mu_3 + 15\mu_1) = 1/120(\mu_5 - 10\mu_3) \\ &\text{usw.}\end{aligned}\tag{55}$$

Für die Schätzer  $\hat{\theta}_i$  erhält man deshalb aus (54):

$$\begin{aligned}\hat{\theta}_3 &= 1/6m_3 \\ \hat{\theta}_{04} &= 1/24(m_4 - 3) \\ \hat{\theta}_{05} &= 1/120(m_5 - 10m_3) \\ &\text{usw.}\end{aligned}\tag{57}$$

Um den Test aus Kapitel 4 anwenden zu können, sind Schätzer für die asymptotische Varianz  $\sigma_i^2$  erforderlich. Nach (39) folgt:

$$\begin{aligned}\hat{\sigma}_i^2 &= \frac{1}{(i!)^2 n} \sum_{j=1}^n H_i^2(z_j) - \hat{\theta}_i^2 \\ &= \frac{1}{(i!)^2} \sum_{j=1}^n H_i^2(z_j)/n - \hat{\theta}_i^2\end{aligned}\tag{59}$$

Für die Berechnung werden die Quadrate der Hermite-Polynome benötigt;  $H_0^2, H_1^2, \dots, H_8^2$  sind im Anhang aufgelistet.  $\mu_h$  wird wieder durch  $m_h$  geschätzt, die Schätzer haben folgendes Aussehen:

$$\begin{aligned}\hat{\sigma}_2^2 &= 1/36(m_6 - 6m_4 + 9m_2) - \hat{\theta}_3^2 \\ &= 1/36(m_6 - 6m_4 + 9) - \hat{\theta}_3^2 \\ \hat{\sigma}_4^2 &= 1/576(m_8 - 12m_6 + 42m_4 - 36m_2 + 9) - \hat{\theta}_4^2 \\ &= 1/576(m_8 - 12m_6 + 42m_4 - 27) - \hat{\theta}_4^2 \\ &\text{usw.}\end{aligned}\tag{60}$$

Die Prüfgröße des Tests wird dann für  $i \geq 3$  durch  $\hat{\theta}_i^2/(\hat{\sigma}_i^2/n)$  gegeben.

## 5.2 Laguerre

Auch das Funktionensystem, das mit Hilfe der Laguerre-Polynome gebildet wird, baut, wie später gezeigt wird, auf einer natürlichen Basisfunktion auf. Da die Vorgehensweise nahezu identisch wie im Falle der Hermite-Polynome ist, wird dieser Abschnitt ebenso wie 5.3 straffer gehalten.

Die approximierende Dichte lautet:

$$g_\theta(x) = \sum_{i \in A} \theta_i L_i(x) \exp(-x) . \quad (61)$$

Die Eigenschaften der Laguerre-Polynome  $L_i(x)$  werden z.B bei Courant-Hilbert [4] aufgeführt. Man kann eine Rekursionsformel für die  $L_i(x)$  angeben :

$$\begin{aligned} L_0(x) &= 1 \\ L_1(x) &= -x + 1 \text{ und} \\ L_i(x) &= (2i - 1 - x)L_{i-1}(x) - (i - 1)^2 L_{i-2}(x) \quad , \quad i \geq 2 \end{aligned} \quad (62)$$

Die ersten Laguerre-Polynome und ihre Quadrate sind im Anhang gegeben. Die Laguerre-Polynome besitzen die Orthogonalitätseigenschaft

$$\int_0^{+\infty} L_i(x) L_j(x) \exp(-x) dx = \delta_{ij} (i!)^2 \quad , \quad i, j \in \mathbb{N}_0. \quad (63)$$

Mit den Bezeichnungen aus Kapitel 3 folgt:

$$\begin{aligned} a &= 0, \quad b = +\infty \\ \phi_i(x) &= L_i(x) \exp(-x) \\ w(x) &= \exp(x) \\ K(i) &= (i!)^2 \quad \text{für } i \in \mathbb{N}_0. \end{aligned} \quad (64)$$

Die zugeordnete Diskrepanz lautet (vgl.(14)):

$$\Delta^+(\theta) = \int_0^{+\infty} \left( f(x) - \sum \theta_i L_i(x) \exp(-x) \right)^2 \exp(x) dx \quad (65)$$

Analog zu (19) wird auch

$$\Delta(\theta) = - \sum (2\theta_i \int_0^{+\infty} L_i(x) dF(x) - \theta_i^2 (i!)^2) \quad (66)$$

verwendet.  $w(x) = \exp(x)$ , d.h., die Anpassung wird im Prinzip mit wachsendem  $x$  besser.

Die Diskrepanz wird dann sinnvoll benutzt, wenn eine Anpassung an die standardisierte Dichte, d.h. an die Dichte von  $Z = (X - EX)/(Var X)^{1/2} + 1$  durchgeführt wird. Nach der Transformation beträgt die Erwartung eins, ebenso wie die Varianz.

Die transformierten Daten

$$z_i = (x_i - \bar{x})/s + 1 \quad (67)$$

werden als Realisationen von unabhängigen Zufallsvariablen angesehen. Nun kann in der Praxis der Fall auftreten, daß es eine Menge  $I^*$  gibt mit  $z_i < 0$  für  $i \in I^*$ . Sei  $Min := \min\{z_i\}$ , dann besteht die plausible Alternative darin,  $s$  bei der obigen Transformation durch  $(\bar{x} - Min)$  zu ersetzen, um sicherzustellen, daß sämtliche transformierten Beobachtungen im nichtnegativen Bereich liegen:

$$z_i = (x_i - \bar{x})/(x - Min) + 1. \quad (68)$$

Sei  $f$  nun die Dichte der transformierten Verteilung. Aus (20) erhält man:

$$\theta_{0i} = 1/(i!)^2 \int_0^{+\infty} L_i(x) dF(x) \quad , \quad i \in A. \quad (69)$$

Wiederum ergibt sich für  $\theta_{00}$  ein von  $f$  unabhängiger Wert:

$$\theta_{00} = 1 \quad (70)$$

Außerdem folgt für  $\theta_{01}$ , bzw.  $\theta_{02}$ :

$$\theta_{01} = \int_0^{+\infty} (-x + 1) dF(x) = 0 \quad (71)$$

und

$$\theta_{02} = 1/4 \int_0^{+\infty} (x^2 - 4x + 2) dF(x) = 0 \quad (72)$$

(71) und (72) besitzen nur dann Gültigkeit, falls die  $x_i$  nach (67) transformiert wurden. Jedes weitere  $\theta_{0i}$  (und  $\theta_{01}$ , bzw.  $\theta_{02}$ , wenn (68) angewendet wird) ist Linearkombination der Momente  $\mu_h$  der transformierten Verteilung.

Da  $L_0(x) = 1$ , lautet die Basisfunktion bei diesem Funktionensystem  $\exp(-x)$ . Deshalb wird man die Darstellung einer approximierenden Dichte wie in (61) vorziehen, wenn die Beobachtungen eine Exponentialfunktion oder eine ihr ähnliche Dichtefunktion vermuten lassen.

Die Minimum-Diskrepanz-Schätzer  $\hat{\theta}_i$ ,  $i \geq 3$  (bzw.  $i \geq 1$  bei Transformation nach (68)) haben die Form:

$$\hat{\theta}_i = 1/(i!)^2 \sum_{j=1}^n L_i(z_j)/n \quad , \quad \text{also} \quad (73)$$

$$\hat{\theta}_1 = -m_1 + 1 = 0$$

$$\hat{\theta}_2 = 1/4(m_2 - 2) \quad (\text{bei (68)})$$

$$\hat{\theta}_3 = 1/36(-m_3 + 9m_2 - 12)$$

$$\hat{\theta}_4 = 1/576(m_4 - 16m_3 + 72m_2 - 72)$$

usw.

Die Schätzer der asymptotischen Varianz  $\sigma_i^2$  lauten (vgl. (39)):

$$\hat{\sigma}_i^2 = 1/(i!)^4 \sum_{j=1}^n L_i^2(z_j)/n - \hat{\theta}_i^2 \quad , \quad i \geq 3 \quad (74)$$

bzw. für  $i \geq 2$ , falls nach (68) transformiert wurde.

### 5.3 Legendre

Wenn von vornherein bekannt ist, daß die Beobachtungen in einem beschränkten Intervall liegen, bietet sich das orthogonale Funktionensystem an, das von den Legendre-Polynomen erzeugt wird. Die approximierende Dichte hat in diesem Fall die Form:

$$g_\theta(x) = \sum_{i \in A} \theta_i P_i(x) \quad (75)$$

Die Legendre-Polynome  $P_i(x)$  sind rekursiv definierbar, Courant-Hilbert ([4]) haben diese und die weiteren Eigenschaften aufgeführt:

$$\begin{aligned} P_0(x) &= 1 \\ P_1(x) &= x \quad \text{und} \\ P_i(x) &= ((2i-1)/i)xP_{i-1}(x) - ((i-1)/i)P_{i-2}(x) \quad , \quad i \geq 2 \end{aligned} \quad (76)$$

Die ersten Polynome und ihre Quadrate sind im Anhang gegeben. Die Legendre-Polynome sind orthogonal bzgl. der Wichtungsfunktion  $w(x) = 1$ :

$$\int_{-1}^{+1} P_i(x)P_j(x)dx = \delta_{ij}2/(2i+1) \quad i, j \in A \quad (77)$$

Benutzt man die Bezeichnungen aus Kapitel 3, so folgt:

$$\begin{aligned} a &= -1, \quad b = +1 \\ \phi_i(x) &= P_i(x) \\ w(x) &= 1 \\ K(i) &= 2/(2i+1) \quad \text{für } i \in \mathbb{N}_0. \end{aligned} \quad (78)$$

Die verwendete Diskrepanz heißt wegen (14):

$$\Delta^+(\theta) = \int_{-1}^{+1} \left( f(x) - \sum_{i \in A} \theta_i P_i(x) \right)^2 dx . \quad (79)$$

Ersatzweise wird wie in (19)  $\Delta(\theta)$  benutzt:

$$\Delta(\theta) = - \sum_{i \in A} \left( 2\theta_i \int_{-1}^{+1} P_i(x)dF(x) - 2\theta_i^2/(2i+1) \right) \quad (80)$$

Die Legendre-Polynome besitzen nur über dem Intervall  $[-1, 1]$  die Orthogonalitätseigenschaft. Liegt ein Intervall  $[c, d]$  vor mit  $x_i \in [c, d]$ ,  $i = 1, 2, \dots, n$ , so ist es am einfachsten, eine Transformation so durchzuführen, daß  $[c, d]$  linear auf  $[-1, 1]$  abgebildet wird:

$$z_i = 2(x_i - c)/(d - c) - 1 \quad i = 1, 2, \dots, n \quad (81)$$

Sei  $f$  die Dichte der transformierten Verteilung, dann erhält man aus (20):

$$\theta_{0i} = (2i + 1)/2 \int_{-1}^{+1} P_i(x) dF(x) \quad , \quad i \in A . \quad (82)$$

$\theta_{00} = 1/2$ , d.h. die Basisfunktion lautet  $\theta_{00}P_0(x) = 1/2$ . Besonders übersichtliche Modelle wird man deshalb dann erhalten, falls die Daten nahezu gleichverteilt sind.

Die Legendre-Polynome eignen sich mit Einschränkungen auch zur Approximation von beliebigen Dichtefunktionen, eine hohe Rechengenauigkeit wird dabei vorausgesetzt. Hall ([9]) führt aus, daß die Legendre-Polynome der trigonometrischen Reihe dann überlegen sind, wenn bei der Dichtefunktion die Existenz von höheren Ableitungen vorausgesetzt werden darf; die differenzierte approximierende Dichte stellt dann zusätzlich einen guten Schätzer für die Ableitung der Dichtefunktion dar.

Schätzer von  $\theta_{0i}$  und  $\sigma_i^2$  leitet man wie in (22), bzw. (39) her:

$$\hat{\theta}_i = (2i + 1)/2 \sum_{j=1}^n P_i(z_j)/n \quad (83)$$

$$\hat{\sigma}_i^2 = ((2i + 1)/2)^2 \sum_{j=1}^n P_i^2(z_j)/n - \hat{\theta}_i^2 \quad (84)$$

für  $i \geq 1$ .

Die Prüfgröße des Tests ist wieder durch  $\hat{\theta}_i^2/(\hat{\sigma}_i^2/n)$  gegeben.

## 6 Anwendungen

### 6.1 Beispiel 1: Entrindungszeiten

In der Forstwirtschaft werden Maschinen zur Entrindung von Bäumen eingesetzt. Um eine Dispositionsgrundlage für den Einsatz von Arbeitern und Maschinen zu erhalten, wird die Verteilung der Entrindungszeiten (in Minuten) benötigt. Abbildung 1 zeigt ein Histogramm, das aufgrund von 1563 Zeitmessungen gebildet wurde (Häberle [10]).

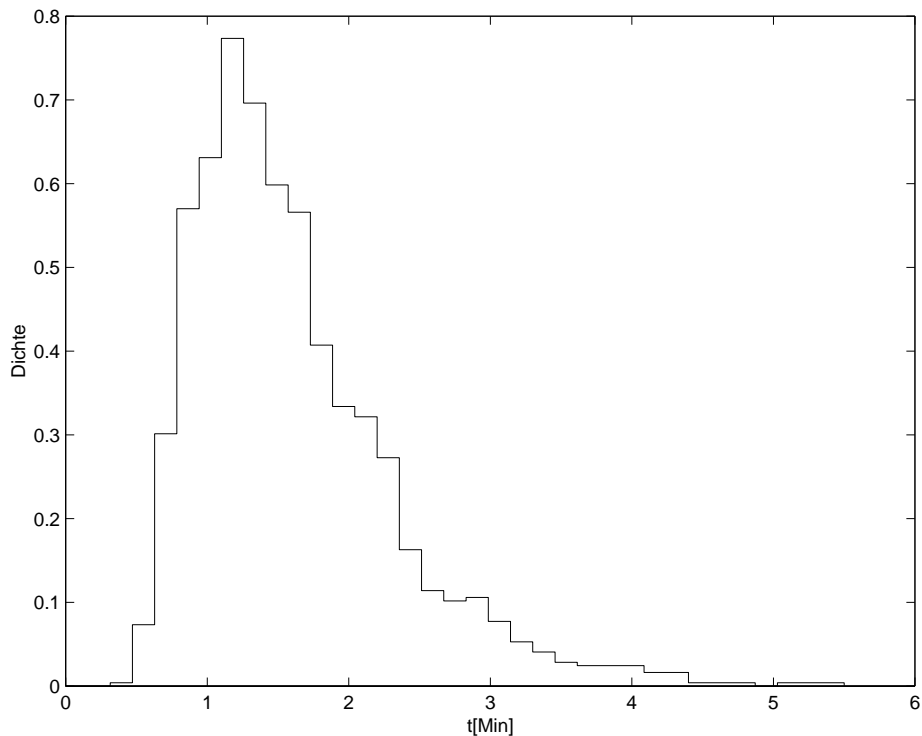


Abbildung 1: Entrindungszeiten

Die Form des Histogramms legt nahe, eine Approximation mit Hilfe der Gram-Charlier-Reihe Typ A an die *Logarithmen* der Originaldaten durchzuführen. Für die logarithmierten Entrindungszeiten erhält man:

$$\begin{aligned}\bar{x} &= 0.357212 \\ s &= 0.414371\end{aligned}\tag{85}$$

Betrachtet werden ab nun die standardisierten Beobachtungen  $z_i = (x_i - \bar{x})/s$ ,  $i = 1, \dots, 1563$ . Für die transformierten Daten errechnet man die Stichprobenmomente  $m_h$  (Tabelle 2).

$h$	$m_h$	$h$	$m_h$
1	$-3.612897 * 10^{-09}$	13	$+6.814980 * 10^{+03}$
2	$+1.000000 * 10^{+00}$	14	$+2.654495 * 10^{+04}$
3	$+1.983009 * 10^{-01}$	15	$+6.084432 * 10^{+04}$
4	$+2.805637 * 10^{+00}$	16	$+2.249924 * 10^{+05}$
5	$+1.574853 * 10^{+00}$	17	$+5.568614 * 10^{+05}$
6	$+1.217070 * 10^{+01}$	18	$+1.985629 * 10^{+06}$
7	$+1.197797 * 10^{+01}$	19	$+5.193480 * 10^{+06}$
8	$+6.802205 * 10^{+01}$	20	$+1.803738 * 10^{+07}$
9	$+9.491551 * 10^{+01}$	21	$+4.912936 * 10^{+07}$
10	$+4.485496 * 10^{+02}$	22	$+1.673044 * 10^{+08}$
11	$+7.882331 * 10^{+02}$	23	$+4.697571 * 10^{+08}$
12	$+3.314942 * 10^{+03}$	24	$+1.575701 * 10^{+09}$

Tabelle 2: Stichprobenmomente der transformierten Logarithmen der Entrindungszeiten

Sämtliche Werte wurden für die Tabelle auf sieben führende Stellen gerundet.  $m_1$  stimmt wegen Rechenungenauigkeiten nicht mit dem theoretischen Wert überein. Für die Schätzer von  $\theta_i$  und  $\sigma_i^2$ , bzw. für die Prüfgrößen  $\hat{\theta}_i^2/(\hat{\sigma}_i^2/n)$  wurden die Werte aus Tabelle 3 ermittelt.

$i$	$\hat{\theta}_i$	$\hat{\sigma}_i^2$	Prüfgröße
3	$+3.305016 * 10^{-02}$	$1.193764 * 10^{-01}$	14.30
4	$-8.098446 * 10^{-03}$	$2.217476 * 10^{-02}$	4.62
5	$-3.401308 * 10^{-03}$	$3.711217 * 10^{-03}$	4.87
6	$+1.196390 * 10^{-04}$	$5.956455 * 10^{-04}$	0.038
7	$-5.403499 * 10^{-05}$	$8.175190 * 10^{-05}$	0.056
8	$+3.537485 * 10^{-05}$	$9.145329 * 10^{-06}$	0.21
9	$+2.519765 * 10^{-05}$	$9.076034 * 10^{-07}$	1.09
10	$-7.330553 * 10^{-07}$	$9.485757 * 10^{-08}$	0.0089
11	$-1.305097 * 10^{-06}$	$7.813023 * 10^{-09}$	0.34
12	$-2.389651 * 10^{-07}$	$5.962660 * 10^{-10}$	0.15

Tabelle 3: Schätzer und Prüfgrößen. Bsp. Entrindungszeiten

Sei  $A_\alpha \subset \{0, 1, \dots, 12\}$  die zu gegebenem  $\alpha$  ausgewählte Indexmenge, dann folgt unter Zuhilfenahme von Tabelle 1 die in Tabelle 4 dargestellte Auswahl.

Die Wahl von  $i_{max} = 12$  scheint ausreichend für  $\alpha \leq 0.35$ , da die Prüfgrößen schnell abnehmen und das Histogramm keine "Peaks" besitzt, der kritische Wert zu  $\alpha = 0.50$  könnte vielleicht mit wachsendem  $i$  noch einmal überschritten werden. Es sollte klar sein, daß mit größerem  $\alpha$  auch  $i_{max}$  erhöht werden muß. Sei

$\alpha$	$A_\alpha$
0.01	$\{0, 3\}$
0.05	$\{0, 3\}$
0.10	$\{0, 3\}$
0.35	$\{0, 3, 4, 5\}$
0.50	$\{0, 3, 4, 5\}$

Tabelle 4: Ausgewählte Indexmengen. Bsp. Entrindungszeiten

$\alpha(x)$  wieder die Standard-Normalverteilungsdichte, dann sind die approximierenden Dichten von der Gestalt:

$$g_{\hat{\theta}, \alpha}(x) = \sum_{i \in A_\alpha} \hat{\theta}_i H_i(x) \alpha(x) \quad (86)$$

Für  $\alpha = 0.10$  wird z.B mit  $g_{\hat{\theta}, 0.10}(x) = \alpha(x)(1 + \hat{\theta}_3(x^3 - 3x))$  approximiert. Die gute Anpassung der approximierenden Dichten wird in Abb. 2 für  $\alpha = 0.10$ , bzw. in Abb. 3 für  $\alpha = 0.35$  demonstriert. Man sieht deutlich, etwa im Bereich zwischen  $-1$  und  $+1$ , wie in Abb. 3 die vorliegende Stichprobe bei größerem  $\alpha$  noch genauer wiedergegeben wird.

Im vorliegenden Fall war eine gute Anpassung durch die Gram-Charlier-Reihe vom Typ A zu erwarten. Das Histogramm zeigt einen ähnlichen Aufbau wie die Normalverteilungsdichte, die Verteilung ist eingipflig und nahezu symmetrisch. Sind diese Voraussetzungen gegeben, so liefert die Gram-Charlier-Reihe Typ A fast immer eine gute Approximation.

Manchmal kommt es vor, daß die approximierende Funktion zum Rand hin Werte unter null annimmt. Dieses passiert u.a. dann, wenn sich "steile" Passagen mit "flachen" abwechseln. Im vorliegenden Beispiel läuft das Histogramm zu den Rändern hin langsam aus, der obenstehende Effekt ist also nicht unbedingt zu erwarten.

Mit  $h_\alpha(x) = g_{\hat{\theta}, \alpha}((x - \bar{x})/s)/s$  wird die approximierende Dichte für die Logarithmen der Entrindungszeiten gegeben.



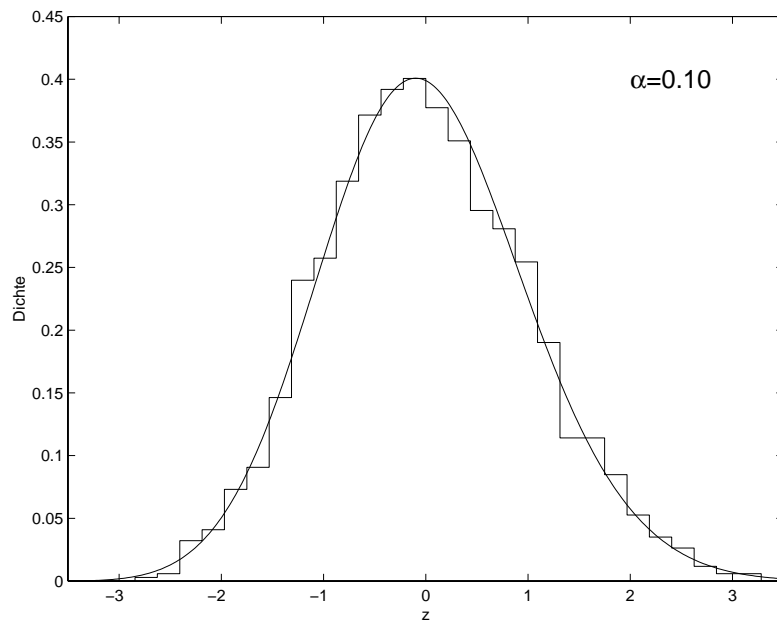


Abbildung 2: Transformierte Logarithmen der Entrindungszeiten (I)

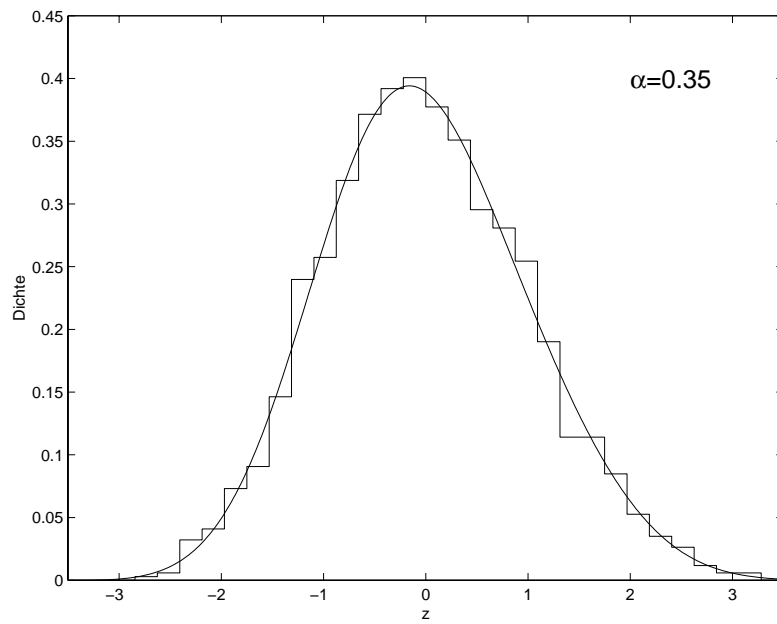


Abbildung 3: Transformierte Logarithmen der Entrindungszeiten (II)

## 6.2 Beispiel 2: Höhen eines Baumbestandes

Ebenfalls in der Forstwirtschaft tritt das Problem auf, Wachstumsgesetze für Bäume unter gegebenen Rahmenbedingungen aufzustellen. Dazu wird ein Bestand über Jahre hinweg in regelmäßigen Abständen vermessen. In diesem Beispiel soll die Höhenverteilung eines Bestandes zu einem festen Zeitpunkt approximiert werden.

Abbildung 4. zeigt ein Histogramm für die Höhen von 680 Bäumen in Metern (Saborowski [11]). Bei Betrachtung des Verlaufs erscheint eine Approximation mit Hilfe der Gram-Charlier-Reihe Typ A angebracht.

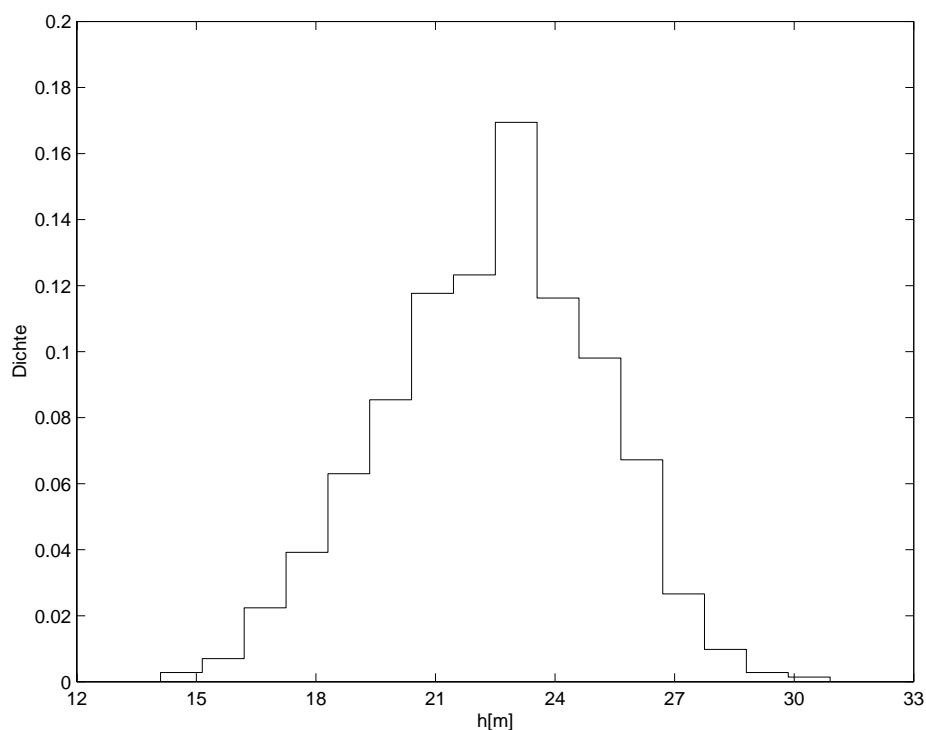


Abbildung 4: Höhenverteilung eines Baumbestandes

Aus den Daten werden  $\bar{x}$ , bzw. die Standardabweichung  $s$  berechnet:

$$\begin{aligned}\bar{x} &= 22.386 \\ s &= 2.6740\end{aligned}\tag{87}$$

Sei  $z_i = (x_i - \bar{x})/s$  für  $i = 1, \dots, 680$ . Die Stichprobenmomente sind in Tabelle 5 dargestellt. Die Schätzer für  $\theta_i$  bzw.  $\sigma_i^2$  und die Prüfgrößen sind in Tabelle 6 aufgelistet. Die Indextmengen aus Tabelle 7 werden deshalb den Irrtumswahrscheinlichkeiten zugeordnet.

$h$	$m_h$	$h$	$m_h$
1	$-3.560937 * 10^{-11}$	13	$-9.590529 * 10^{+02}$
2	$+1.000000 * 10^{+00}$	14	$+1.368942 * 10^{+04}$
3	$-1.405845 * 10^{-01}$	15	$-5.943395 * 10^{+03}$
4	$+2.737087 * 10^{+00}$	16	$+9.551500 * 10^{+04}$
5	$-8.333279 * 10^{-01}$	17	$-3.653186 * 10^{+04}$
6	$+1.106077 * 10^{+01}$	18	$+6.855705 * 10^{+05}$
7	$-4.608234 * 10^{+00}$	19	$-2.178017 * 10^{+05}$
8	$+5.577929 * 10^{+01}$	20	$+5.023754 * 10^{+06}$
9	$-2.630101 * 10^{+01}$	21	$-1.218301 * 10^{+06}$
10	$+3.228944 * 10^{+02}$	22	$+3.739149 * 10^{+07}$
11	$-1.565042 * 10^{+02}$	23	$-5.959202 * 10^{+06}$
12	$+2.041949 * 10^{+03}$	24	$+2.816661 * 10^{+08}$

Tabelle 5: Stichprobenmomente der standardisierten Höhen

$i$	$\hat{\theta}_i$	$\hat{\sigma}_i^2$	Prüfgröße
3	$-2.343076 * 10^{-02}$	$1.005136 * 10^{-01}$	3.71
4	$-1.095472 * 10^{-02}$	$1.899049 * 10^{-02}$	4.30
5	$+4.770979 * 10^{-03}$	$3.385795 * 10^{-03}$	4.57
6	$+6.213421 * 10^{-06}$	$5.999513 * 10^{-04}$	0.000044
7	$-3.709768 * 10^{-04}$	$7.177833 * 10^{-05}$	1.30
8	$+1.454817 * 10^{-04}$	$7.759277 * 10^{-06}$	1.85
9	$+4.778338 * 10^{-06}$	$9.733991 * 10^{-07}$	0.016
10	$-1.672980 * 10^{-05}$	$7.907267 * 10^{-08}$	2.41
11	$+1.684384 * 10^{-06}$	$7.159189 * 10^{-09}$	0.27
12	$+1.139698 * 10^{-06}$	$6.074256 * 10^{-10}$	1.45

Tabelle 6: Schätzer und Prüfgrößen. Bsp. Baumhöhen

$\alpha$	$A_\alpha$
0.01	{0}
0.05	{0}
0.10	{0}
0.35	{0, 3, 4, 5, 10}
0.50	{0, 3, 4, 5, 10, 12}

Tabelle 7: Ausgewählte Indexmengen. Bsp. Baumhöhen

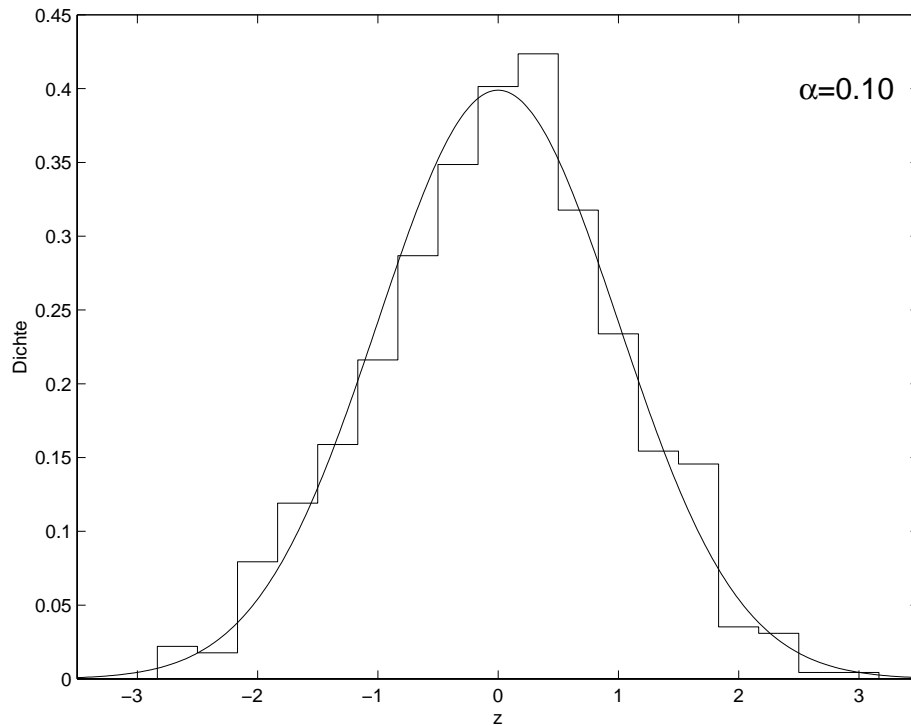


Abbildung 5: Anpassung an standardisierte Höhen (I)

Für  $\alpha = 0.50$  wird selbst für  $i = 20$  noch ein Term zur Approximation aufgenommen. Da in diesem Bereich schon deutliche Rundungsfehler auftreten, beschränkt man sich wieder auf  $\alpha \leq 0.35$ . Ist  $\alpha \leq 0.10$ , wird die Dichtefunktion der standardisierten Höhen durch die Standard-Normalverteilungsdichte angenähert.

In Abbildung 5 und 6 werden die Anpassungen bei  $\alpha = 0.10$  und  $\alpha = 0.35$  illustriert. Man sieht sehr schön, daß für  $\alpha = 0.35$  die approximierende Dichtefunktion links von Null flacher ausfällt und die gegebene Stichprobe genauer darstellt. In diesem Beispiel gerät die approximierende Dichtefunktion links von  $-3$  für  $\alpha = 0.35$  leicht unter die Achse. Die Approximation ist trotzdem brauchbar, falls man sich auf ein Teilintervall beschränkt.

Für die Originaldaten besitzt die approximierende Dichte zu einem bestimmten  $\alpha$  die Form  $g_{\hat{\theta},\alpha}((x - \bar{x})/s)/s$ .

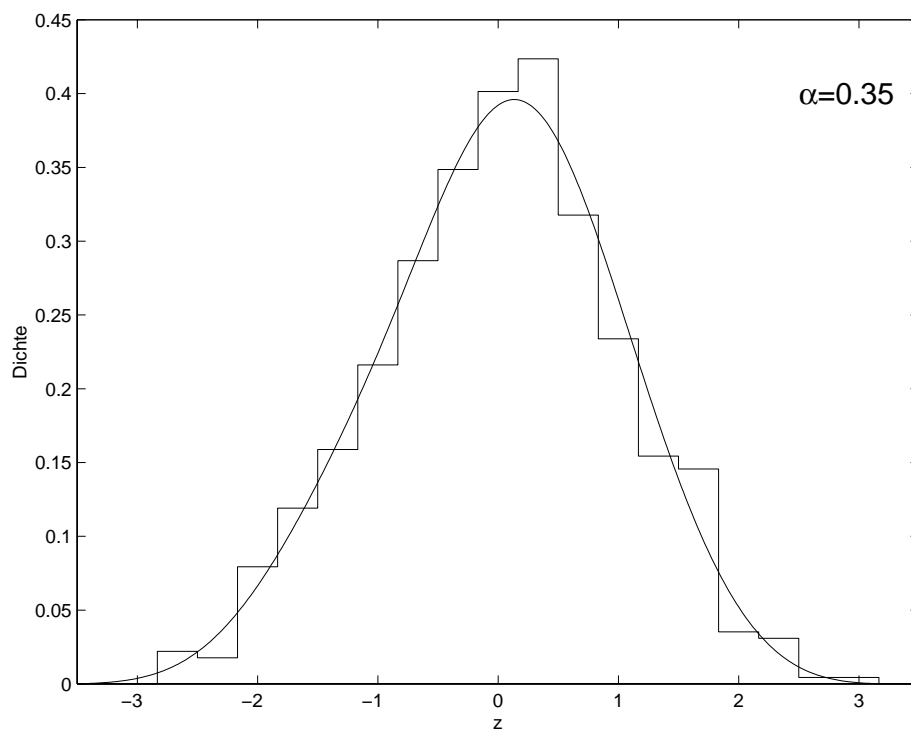


Abbildung 6: Anpassung an standardisierte Höhen (II)

### 6.3 Beispiel 3: Volumina eines Baumbestandes (Laguerre)

Für den Bestand aus 6.2 hat man näherungsweise die Volumina in Festmetern berechnet, indem zusätzlich zu den Baumhöhen in mehreren Höhen die Durchmesser bestimmt wurden. Die erhaltenen 680 Werte sind in Abb. 7 in einem Histogramm dargestellt.

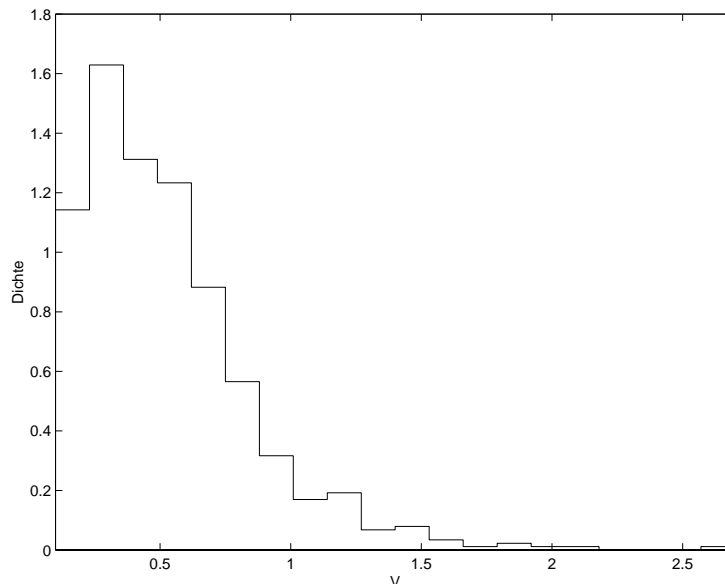


Abbildung 7: Histogramm zur Verteilung der Volumina

Die tendenziell abfallende Form des Histogramms läßt gute Approximationen bei Verwendung der Laguerre-Polynome erwarten. Transformiert man nach (67), so sind 75 der 680 transformierten Daten kleiner als null. Deshalb wird die Transformation nach (68) gewählt, bei der das Minimum der Originaldaten auf null abgebildet wird:

$$\begin{aligned}\bar{x} &= 0.531494 \\ Min &= 0.108 \\ z_i &= (x_i - \bar{x})/(\bar{x} - Min) + 1 \quad \text{für } i = 1, 2, \dots, 680\end{aligned}\quad (88)$$

Für die transformierten Beobachtungen werden die Stichprobenmomente wie in Tabelle 8 berechnet.

Aus den transformierten Daten erhält man die Schätzer und Prüfgrößen in Tabelle 9.

Die Stichprobenmomente wachsen schneller als die entsprechenden Größen der bisherigen Beispiele. Gleichzeitig besitzen die Laguerre-Polynome bei gleichem Grad höhere Koeffizienten als die Hermite-Polynome (siehe Anhang), im

$h$	$m_h$	$h$	$m_h$
1	$1.000000 * 10^{+00}$	11	$7.488636 * 10^{+05}$
2	$1.598450 * 10^{+00}$	12	$4.420316 * 10^{+06}$
3	$3.569925 * 10^{+00}$	13	$2.636208 * 10^{+07}$
4	$1.046850 * 10^{+01}$	14	$1.583421 * 10^{+08}$
5	$3.817184 * 10^{+01}$	15	$9.558219 * 10^{+08}$
6	$1.639706 * 10^{+02}$	16	$5.790168 * 10^{+09}$
7	$7.908303 * 10^{+02}$	17	$3.516448 * 10^{+10}$
8	$4.124993 * 10^{+03}$	18	$2.139497 * 10^{+11}$
9	$2.266052 * 10^{+04}$	19	$1.303461 * 10^{+12}$
10	$1.288043 * 10^{+05}$	20	$7.948943 * 10^{+12}$

Tabelle 8: Stichprobenmomente der transformierten Volumina (Laguerre)

$i$	$\hat{\theta}_i$	$\hat{\sigma}_i^2$	Prüfgröße
2	$-1.003875 * 10^{-01}$	$1.073036 * 10^{-01}$	63.86
3	$-3.288538 * 10^{-02}$	$8.739105 * 10^{-03}$	84.15
4	$-6.183848 * 10^{-03}$	$5.385738 * 10^{-04}$	48.28
5	$-7.898836 * 10^{-04}$	$1.852674 * 10^{-05}$	22.90
6	$-6.863625 * 10^{-05}$	$4.632284 * 10^{-07}$	6.92
7	$-3.393098 * 10^{-06}$	$9.883690 * 10^{-09}$	0.79
8	$+5.580851 * 10^{-08}$	$1.393852 * 10^{-10}$	0.015
9	$+3.241106 * 10^{-08}$	$1.398232 * 10^{-12}$	0.51
10	$+4.244758 * 10^{-09}$	$1.002135 * 10^{-14}$	1.22

Tabelle 9: Schätzer und Prüfgrößen. Bsp. Volumina (Laguerre)

Vergleich zu der Gram-Charlier-Reihe Typ A muß  $i_{max}$  kleiner gewählt werden, da sonst eine ausreichende Rechengenauigkeit nicht gewährleistet ist.

Die ausgewählten Indexmengen sind in Tabelle 10 gegeben.

Für  $\alpha = 0.50$  ist  $i_{max}$  zu klein, für die weiteren Irrtumswahrscheinlichkeiten ausreichend gewählt.

Ist  $\alpha$  gegeben, so heißt die approximierende Dichte:

$$g_{\hat{\theta},\alpha}(x) = \sum_{i \in A_\alpha} \hat{\theta}_i L_i(x) \exp(-x) \quad (89)$$

In Abb. 8 und Abb. 9. sind die approximierenden Dichten für  $\alpha = 0.05$  und  $\alpha = 0.10$  zusammen mit dem Histogramm gegeben. Die Dichte für die Originaldaten ist durch  $h_\alpha(x) = g_{\hat{\theta},\alpha}((x - \bar{x})/(\bar{x} - Min) + 1)/(\bar{x} - Min)$  gegeben.

Die Approximation unter Zuhilfenahme von Laguerre-Polynomen verspricht Erfolg, wenn die Dichtefunktion von der Tendenz her fallend ist. Ein kurzer

$\alpha$	$A_\alpha$
0.01	$\{0, 2, 3, 4, 5\}$
0.05	$\{0, 2, 3, 4, 5\}$
0.10	$\{0, 2, 3, 4, 5, 6\}$
0.35	$\{0, 2, 3, 4, 5, 6\}$
0.50	$\{0, 2, 3, 4, 5, 6, 8, 10\}$

Tabelle 10: Ausgewählte Indextmengen. Bsp. Volumina (Laguerre)

Anstieg zu Beginn stört die Güte der Annäherung in den meisten Fällen nicht. Falls hingegen die transformierte Dichtefunktion in der Nähe von Null nur geringe Funktionswerte aufweist und dann steil ansteigt, gerät die approximierende Dichte leicht unter die  $x$ -Achse.

Ein ähnliches Problem kann auftreten, wenn die Daten nach (67) transformiert werden: Im Falle, daß die Differenz zwischen Null und dem Minimum der transformierten Beobachtungen zu groß wird, verläuft die approximierende Funktion häufig unterhalb der  $x$ -Achse. Eine Transformation nach (68) auch unter diesen Voraussetzungen hat zumindest bei den betrachteten Datensätzen zu unbrauchbaren Resultaten geführt.



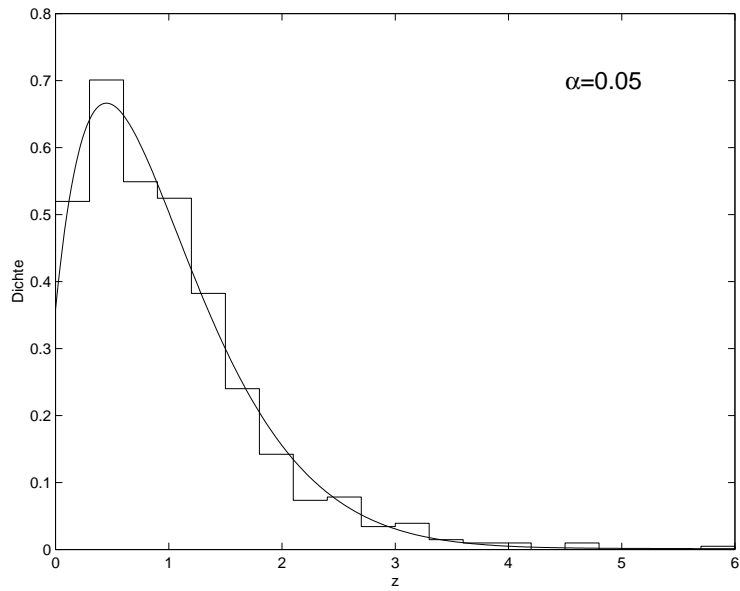


Abbildung 8: Anpassung an die transformierte Verteilung der Baumvolumina (I)

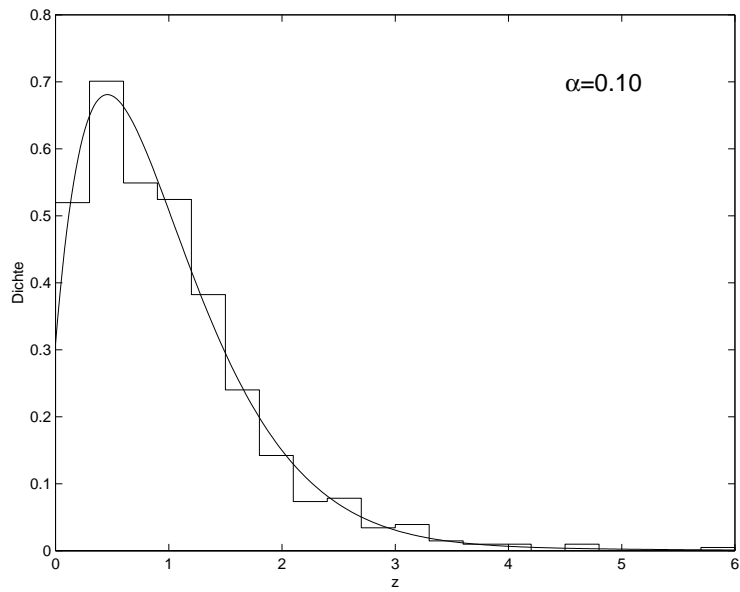


Abbildung 9: Anpassung an die transformierte Verteilung der Baumvolumina (II)

## 6.4 Beispiel 4: Volumina (Legendre)

Zu den Daten aus 6.3 lässt sich eine approximierende Dichtefunktion auch als Linearkombination von Legendre-Polynomen darstellen.

Für die Daten gilt:

$$\min = 0.108 \quad \max = 2.698 \quad (90)$$

Es ist deshalb naheliegend (vgl. Abschnitt 5.3) das Intervall  $[0.1, 2.7]$  auf  $[-1, 1]$  abzubilden:

$$z_i = (x_i - 0.1)/1.3 - 1 \quad (91)$$

Man erhält dann folgende Stichprobenmomente aus Tabelle 11.

$h$	$m_h$	$h$	$m_h$
1	-0.6680814	11	-0.1262896
2	+0.5098419	12	+0.1175107
3	-0.3986966	13	-0.1046454
4	+0.3290339	14	+0.09891650
5	-0.2724968	15	-0.08854766
6	+0.2354740	16	+0.08485237
7	-0.2013424	17	-0.07618949
8	+0.1795623	18	+0.07390757
9	-0.1565919	19	-0.06645616
10	+0.1429892	20	+0.06518990

Tabelle 11: Stichprobenmomente der transformierten Volumina (Legendre)

Es ergeben sich die Schätzer und Prüfgrößen aus Tabelle 12.

$i$	$\hat{\theta}_i$	$\hat{\sigma}_i^2$	Prüfgröße
1	-1.002122	0.1428955	4778.94
2	+0.6619072	0.9716492	306.62
3	+0.01883230	1.850656	0.13
4	-0.4382288	2.315310	56.40
5	+0.4951654	2.778082	60.02
6	-0.2915928	3.558064	16.25
7	+0.1208535	4.256169	2.33
8	-0.01503172	4.706135	0.033
9	+0.06419770	5.212115	0.54
10	-0.08706580	5.685628	0.91

Tabelle 12: Schätzer und Prüfgrößen. Bsp. Volumina (Legendre)

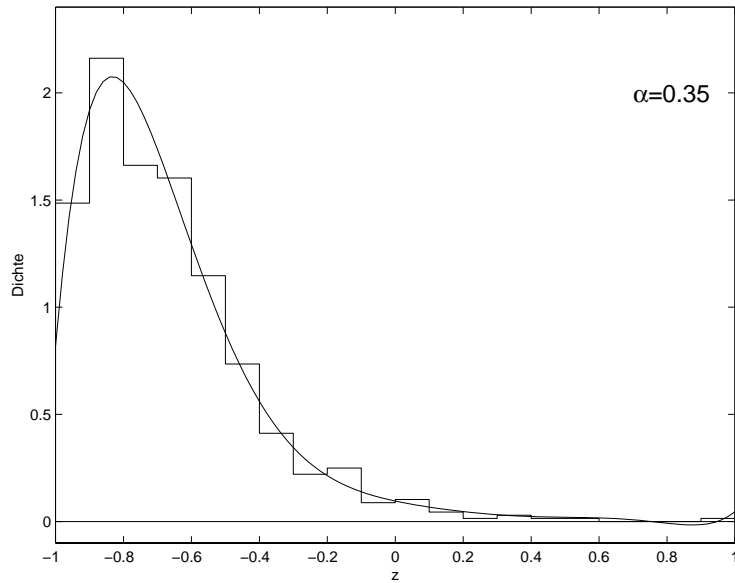


Abbildung 10: Anpassung an die transformierte Verteilung der Baumvolumina (Legendre)

Die den Irrtumswahrscheinlichkeiten zugeordneten Indextmengen sind in Tabelle 13 gegeben.

$\alpha$	$A_\alpha$
0.01	$\{0, 1, 2, 3, 4, 5, 6\}$
0.05	$\{0, 1, 2, 3, 4, 5, 6\}$
0.10	$\{0, 1, 2, 3, 4, 5, 6\}$
0.35	$\{0, 1, 2, 3, 4, 5, 6, 7\}$
0.50	$\{0, 1, 2, 3, 4, 5, 6, 7\}$

Tabelle 13: Ausgewählte Indextmengen. Bsp. Volumina (Legendre)

Die approximierende Dichte, abhängig von  $\alpha$ , ist dann durch

$$g_{\hat{\theta},\alpha}(x) = \sum_{i \in A_\alpha} \hat{\theta}_i P_i(x) \tag{92}$$

gegeben.

Betrachtet man die approximierende Dichte in Abbildung 10, so kann man mit der Anpassung auf dem Intervall  $[-1, 0.75]$  zufrieden sein, leider schwingt die Funktion für größere Werte unter die  $x$ -Achse (Abbildung 11 zeigt dies vergrößert).

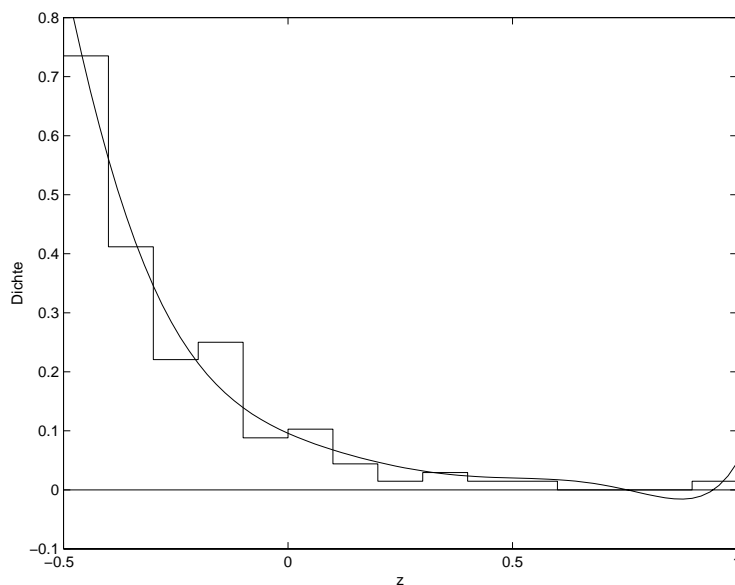


Abbildung 11: Ausschnitt: Anpassung an die transformierte Verteilung der Baumvolumina (Legendre)

Hier wird eine Schwäche der Legendre-Polynome deutlich. Liegt eine Dichtefunktion vor, deren Werte zu den Intervallgrenzen hin auf null zugehen, so tritt häufig das Phänomen auf, daß entweder die approximierende Dichte wie in diesem Beispiel Werte annimmt, die kleiner als null sind, und / oder es wird ein unnatürlicher Anstieg der Dichtefunktion an den Rändern beobachtet. Das genannte Problem läßt sich mindern, wenn man Maximum, bzw. Minimum der Daten als Intervallgrenzen ansieht.

Schwierigkeiten treten auch dann auf, falls die Dichtefunktion über ein größeres Teilintervall nahezu linear verläuft, aber ansonsten komplex aufgebaut ist, die approximierende Funktion schwingt dann um die imaginäre Dichtefunktion.

Betrachtet man die Ursprungsdaten des Beispiels, so läßt sich der Wert  $max = 2.698$  als Ausreißer ansehen. Das Maximum der übrigen Werte ist 2.105. Transformiert man das Intervall  $[0.1, 2.15]$  auf  $[-1, 1]$  und berechnet die Schätzer, so erhält man für  $\alpha = 0.35$  die in Abb. 12 und Abb. 13 illustrierte gute Annäherung.

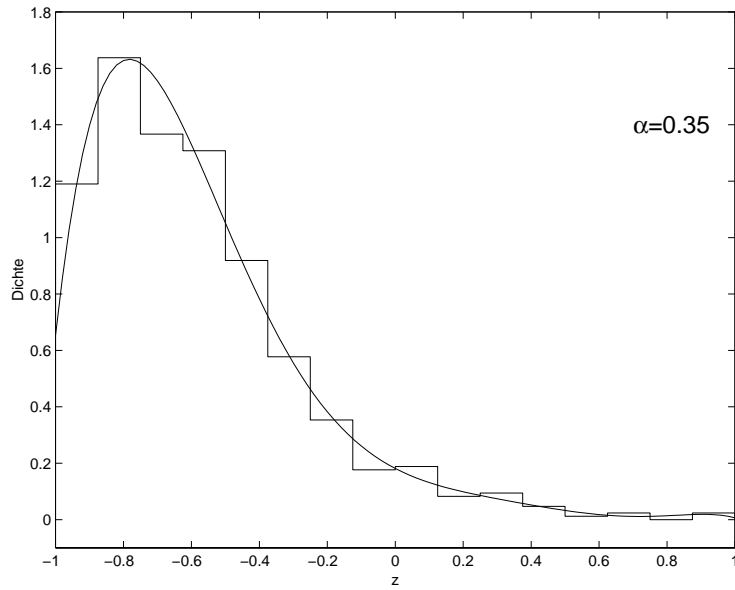


Abbildung 12: Anpassung an die transformierte Verteilung der Baumvolumina ohne Ausreißer (Legendre)

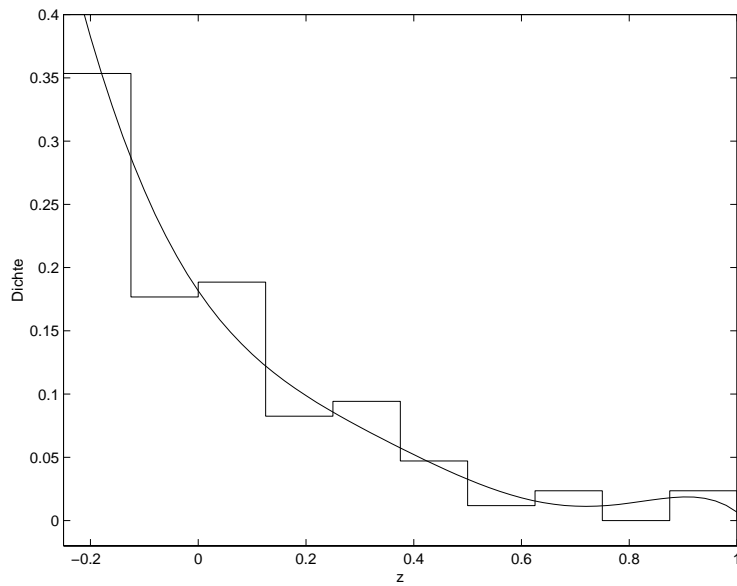


Abbildung 13: Ausschnitt: Anpassung an die transformierte Verteilung der Baumvolumina ohne Ausreißer (Legendre)

## 6.5 Beispiel 5: Glucosekonzentrationen

Anhand dieses Beispiels soll gezeigt werden, daß auf eine Beurteilung des Histogramms vor der Auswahl des Funktionensystems und der Anwendung des Tests nicht verzichtet werden kann.

Die Göttinger Risiko-, Inzidenz- und Prävalenzstudie (GRIPS) ist eine prospektive Studie zur Entdeckung von Risikofaktoren für Herzinfarkt. Im Frühjahr 1982 wurden diverse Daten bei Probanden von VW-Baunatal (40-60 Jahre, deutsch, männlich) erhoben. Gegeben ist eine Stichprobe von Glucosekonzentrationen in mg pro dl bei 1103 Probanden (Seidel [12]).

Betrachtet man das Histogramm der Originaldaten (Abb. 14), so fallen die vereinzelt Beobachtungen über 200 mg/dl auf. Inwiefern sich diese Ausreißer auf die Anpassungsgüte auswirken, soll nun untersucht werden.

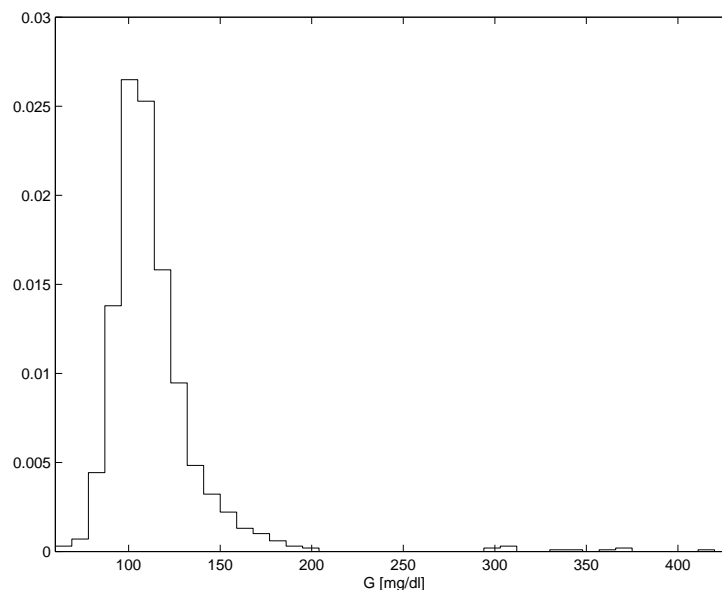


Abbildung 14: Glucosekonzentrationen

Zunächst kommt eine Anpassung mit Hilfe der Hermite-Polynome in Frage. Die approximierende Dichte für  $i_{max} = 10$  und  $\alpha = 0.05$  ist in Abb. 15 dargestellt. Offensichtlich ist der Einfluß der Ausreißer so groß, daß es zu einer unbrauchbaren Schätzung für die Dichtefunktion des Grundmodells kommt.

Aus (46) ist ersichtlich, daß die Güte der Anpassung besonders für große Beträge von  $x$  hoch gewichtet wird. Die Größe der Beträge der Ausreißer geht nach (54) dermaßen stark in die Schätzer für  $\theta_i$  ein, daß die Approximation für kleine Beträge von  $x$  nicht mehr gewährleistet ist.

Versucht man den Ansatz unter Verwendung der Legendre-Polynome, so erhält man für  $c = 61$ ,  $d = 417$  (vgl. 81),  $\alpha = 0.01$  und  $i_{max} = 12$  die in

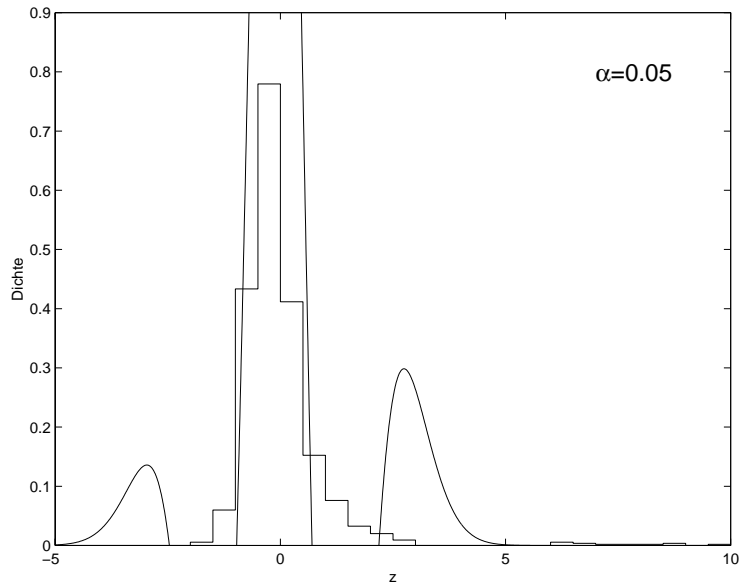


Abbildung 15: Anpassung an standardisierte Glucosekonzentrationen (Hermite)

Abb. 16 dargestellte Approximation. Das Problem liegt hier nicht bei der Wichtungsfunktion, sie ist konstant (78). Vielmehr ist es schwierig, mit Polynomen höheren Grades eine nahezu bei null verlaufende Dichtefunktion anzunähern. An den Rändern sind außerdem die typischen Schwingungen zu sehen, die bei der Verwendung von Legendre-Polynomen auftreten können.

Ein Ansatz unter Verwendung der Laguerre-Polynome kommt von vornherein nicht in Frage, da das Histogramm links nur geringe Höhen aufweist und so ein Verlauf der approximierenden Dichte unter die Achse nicht zu vermeiden ist.

Um trotzdem zu einer vernünftigen Approximation zu gelangen, vor allem für den Bereich zwischen 60 und 200 (Abb. 14), ist es sinnvoll, die Ausreißer vor der Analyse zu entfernen. Eliminiert man die elf Beobachtungen, die über 240 liegen, und führt die oben genannten Ansätze erneut durch, so gelangt man zu einer guten Approximation im Falle der Gram-Charlier-Reihe Typ A (Abb. 17) mit  $i_{max} = 12$  und  $\alpha = 0.35$ .

Die Annäherung unter der Verwendung der Legendre-Polynome weist zwar immer noch den Makel der Schwingungen an den Rändern auf, doch die Approximation (Abb. 18) innerhalb des Intervalls ist im Vergleich erheblich besser geworden ( $i_{max} = 14$ ,  $\alpha = 0.01$ ,  $c = 61$ ,  $d = 197$ ).

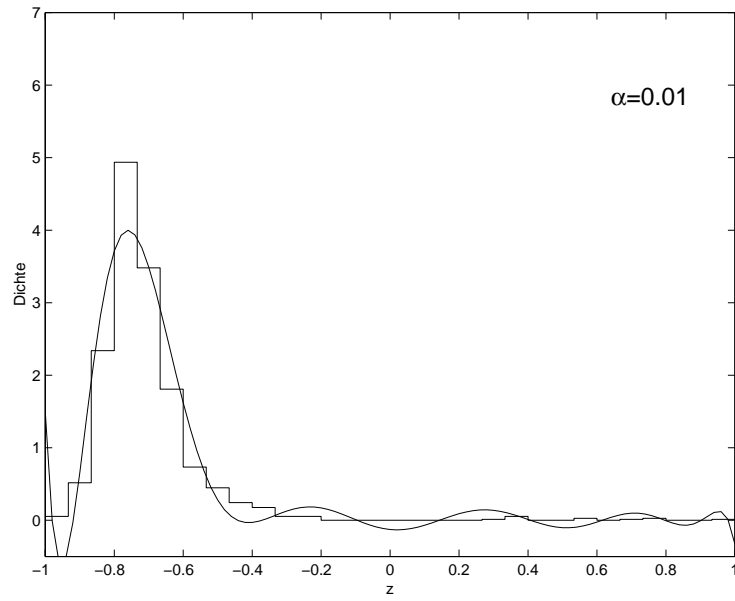


Abbildung 16: Anpassung an transformierte Glucosekonzentrationen (Legendre)

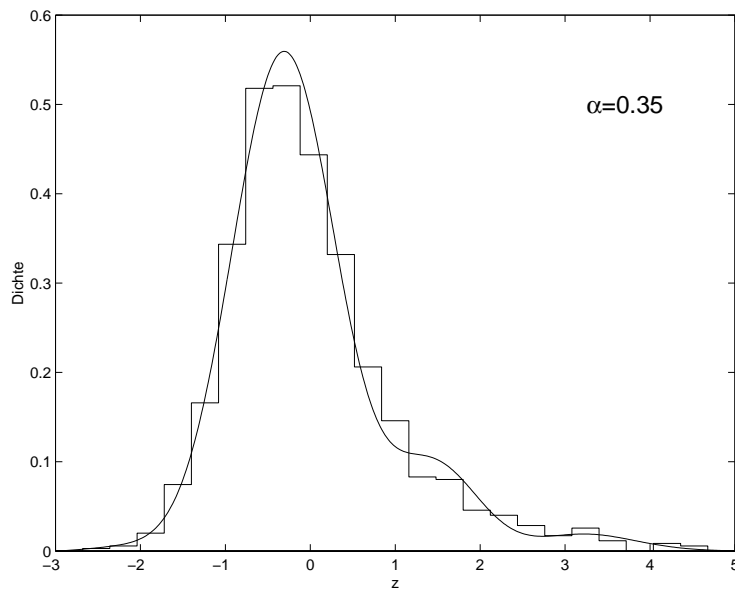


Abbildung 17: Anpassung an standardisierte Glucosekonzentrationen ohne Ausreißer (Hermite)



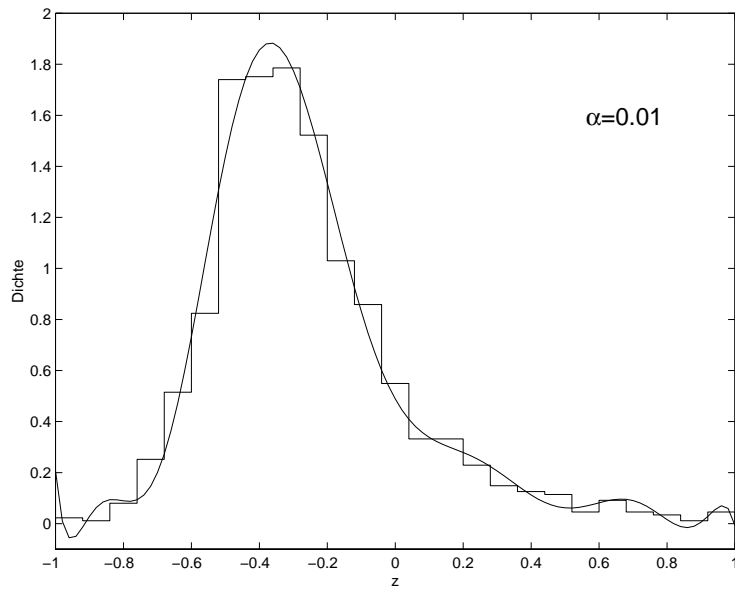


Abbildung 18: Anpassung an standardisierte Glucosekonzentrationen ohne Ausreisser (Legendre)

## 6.6 Anmerkungen

Die genannten Beispiele haben die gemeinsame Eigenschaft, daß jeweils sehr große Datensätze vorliegen. Steht nur eine geringe Anzahl von Beobachtungen zur Verfügung, so ist die Verwendung des Zentralen Grenzwertsatzes (vgl. Abschnitt 4.2) nicht gerechtfertigt, die Aussagekraft über Irrtumswahrscheinlichkeiten ist dann stark eingeschränkt. Die Approximationen können allerdings durchaus brauchbar sein.

Ein direkter Zusammenhang von  $i_{max}$  und der Stichprobengröße kann in allgemeiner Art nicht angegeben werden, der Verlauf der unbekanntes Dichtefunktion spielt hier die entscheidende Rolle. Je komplexer die Dichtefunktion ist, desto höher wird  $i_{max}$  bei gleicher Stichprobengröße gewählt.

## 7 Zusammenfassung

Es sollte deutlich geworden sein, daß der Test aus Kapitel 4 durchaus geeignet ist, ansprechende Approximationen zu liefern. Unentbehrlich für die Anwendung ist eine gewisse Erfahrung in der Auswahl der Funktionensysteme. Selbstverständlich decken die drei genannten Orthogonalensysteme nicht sämtliche Verteilungstypen ab, die Güte der Approximation muß in jedem Fall vom Anwender kontrolliert werden.

Die Berechnung der numerischen Größen erfolgte auf einem gewöhnlichen PC mit Hilfe des Programms "ORTHO", die Effizienz des Verfahrens kann noch erhöht werden, falls eine höhere Rechengenauigkeit vorausgesetzt werden darf, wie sie etwa für die Approximation von weniger glatten oder mehrgipfligen Dichtefunktionen erforderlich ist. Weiterhin ist die Tauglichkeit anderer Funktionensysteme für die Anwendung des Tests zu prüfen. Möglicherweise ist auch die Verallgemeinerung auf mehrdimensionale oder diskrete Verteilungen gegeben.

## 8 Anhang

### 8.1 Hermite-Polynome und ihre Quadrate

$$\begin{aligned}H_0(x) &= 1 \\H_1(x) &= x \\H_2(x) &= x^2 - 1 \\H_3(x) &= x^3 - 3x \\H_4(x) &= x^4 - 6x^2 + 3 \\H_5(x) &= x^5 - 10x^3 + 15x \\H_6(x) &= x^6 - 15x^4 + 45x^2 - 15 \\H_7(x) &= x^7 - 21x^5 + 105x^3 - 105x \\H_8(x) &= x^8 - 28x^6 + 210x^4 - 420x^2 + 105 \\&\dots\end{aligned}$$

$$\begin{aligned}H_0^2(x) &= 1 \\H_1^2(x) &= x^2 \\H_2^2(x) &= x^4 - 2x^2 + 1 \\H_3^2(x) &= x^6 - 6x^4 + 9x^2 \\H_4^2(x) &= x^8 - 12x^6 + 42x^4 - 36x^2 + 9 \\H_5^2(x) &= x^{10} - 20x^8 + 130x^6 - 300x^4 + 225x^2 \\H_6^2(x) &= x^{12} - 30x^{10} + 315x^8 - 1380x^6 + 2475x^4 - 1350x^2 + 225 \\H_7^2(x) &= x^{14} - 42x^{12} + 651x^{10} - 4620x^8 + 15435x^6 - 22050x^4 + 11025x^2 \\H_8^2(x) &= x^{16} - 56x^{14} + 1204x^{12} - 12600x^{10} + 67830x^8 \\&\quad - 182280x^6 + 220500x^4 - 88200x^2 + 11025 \\&\dots\end{aligned}$$

## 8.2 Laguerre-Polynome und ihre Quadrate

$$\begin{aligned}L_0(x) &= 1 \\L_1(x) &= -x + 1 \\L_2(x) &= x^2 - 4x + 2 \\L_3(x) &= -x^3 + 9x^2 - 18x + 6 \\L_4(x) &= x^4 - 16x^3 + 72x^2 - 96x + 24 \\L_5(x) &= -x^5 + 25x^4 - 200x^3 + 600x^2 - 600x + 120 \\L_6(x) &= x^6 - 36x^5 + 450x^4 - 2400x^3 + 5400x^2 - 4320x + 720 \\L_7(x) &= -x^7 + 49x^6 - 882x^5 + 7350x^4 - 29400x^3 + 52920x^2 - 35280x + 5040 \\L_8(x) &= x^8 - 64x^7 + 1568x^6 - 18816x^5 + 117600x^4 \\&\quad - 376320x^3 + 564480x^2 - 322560x + 40320 \\&\quad \dots\end{aligned}$$

$$\begin{aligned}L_0^2(x) &= 1 \\L_1^2(x) &= x^2 - 2x + 1 \\L_2^2(x) &= x^4 - 8x^3 + 20x^2 - 16x + 4 \\L_3^2(x) &= x^6 - 18x^5 + 117x^4 - 336x^3 + 432x^2 - 216x + 36 \\L_4^2(x) &= x^8 - 32x^7 + 400x^6 - 2496x^5 + 8304x^4 - 14592x^3 + 12672x^2 - 4608x + 576 \\L_5^2(x) &= x^{10} - 50x^9 + 1025x^8 - 11200x^7 + 71200x^6 \\&\quad - 270240x^5 + 606000x^4 - 768000x^3 + 504000x^2 \\&\quad - 144000x + 14400 \\&\quad \dots\end{aligned}$$

### 8.3 Legendre-Polynome und ihre Quadrate

$$\begin{aligned}P_0(x) &= 1 \\P_1(x) &= x \\P_2(x) &= 1.5x^2 - 0.5 \\P_3(x) &= 2.5x^3 - 1.5x \\P_4(x) &= 4.375x^4 - 3.75x^2 + 0.375 \\P_5(x) &= 7.875x^5 - 8.75x^3 + 1.875x \\P_6(x) &= 14.4375x^6 - 19.6875x^4 + 6.5625x^2 - 0.3125 \\P_7(x) &= 26.8125x^7 - 43.3125x^5 + 19.6875x^3 - 2.1875x \\&\dots\end{aligned}$$

$$\begin{aligned}P_{02}(x) &= 1 \\P_{12}(x) &= x^2 \\P_{22}(x) &= 2.25x^4 - 1.5x^2 + 0.25 \\P_{32}(x) &= 6.25x^6 - 7.5x^4 + 2.25x^2 \\P_{42}(x) &= 19.14063x^8 - 32.8125x^6 + 17.34375x^4 - 2.8125x^2 + 0.140625 \\P_{52}(x) &= 62.01563x^{10} - 137.8125x^8 + 106.0938x^6 - 32.8125x^4 + 3.515625x^2 \\&\dots\end{aligned}$$

Die Koeffizienten für  $P_4^2(x)$  und  $P_5^2(x)$  sind teilweise gerundet.

## **8.4 Daten**

**8.4.1 Entrindungszeiten**

**8.4.2 Höhen eines Bestandes**

**8.4.3 Volumina eines Bestandes**

**8.4.4 Glucosekonzentrationen**

## 8.5 Programmlisting - "ORTHO"

Für die Berechnung der Schätzer, Prüfgrößen und Koeffizienten, bzw. für den Ausdruck der approximierenden Dichtefunktion wurde das folgende selbstentworfene Programm benutzt. Es handelt sich um eine Arbeitsfassung, eine Optimierung z.B. bzgl. der Rechenzeit wurde noch nicht vorgenommen. Das Programm soll die Umsetzbarkeit des Tests aus Kapitel 4 unter Benutzung von EDV demonstrieren. Für die graphische Darstellung ist eine Auflösung von 640 x 400 Punkten nötig.

```
10 REM *****
20 REM * ORTHO Copyright by Achim Lewandowski ** 8/87 *
30 REM * Dieses Programm bestimmt approximierende *
40 REM * Dichten zu gegebenen Daten *
50 REM *****

60 REM ***Bei doppelter Genauigkeit muss ' entfernt werden 70
'DEFDBL A-J,0-R,V-Z

100 REM *****Vorbereitung
110 PI=ATN(1)*4
120 DIM PRUEF(20)
130 DIM RH(50)
140 DIM AKOEFF(20)
150 DIM ATHETA(20)

200 REM *****Einlesen der Datei
210 CLS:KEY OFF
220 PRINT "Laufwerk (ohne :) ?"
230 INPUT A$
240 PRINT "Name der Datei (ohne .dat) ?"
250 INPUT B$
260 C$=A$+"."+B$+".dat"
270 KEY OFF
280 PRINT "Einlesen der Daten in das Feld Beob ..."
290 OPEN C$ FOR INPUT AS #1
300 DIM BEOB(2000)
310 ANZAHL = 0
320 IF EOF(1) THEN GOTO 400
330 INPUT #1, BEOB(ANZAHL + 1)
340 ANZAHL = ANZAHL + 1
350 GOTO 320

400 REM *****Mittelwert,Varianz;Max,Min
410 CLOSE #1
420 LET AMIN = 1E+20
430 LET AMAX = -1E+20
```



```

440 ASUMME = 0
450 PRINT "Bestimmung von Mittelwert und Varianz"
460 FOR T= 1 TO ANZAHL
470 IF BEOB(T)< AMIN THEN AMIN = BEOB(T)
480 IF BEOB(T)> AMAX THEN AMAX = BEOB(T)
490 ASUMME = ASUMME + BEOB(T)
500 NEXT T
510 VARIANZ = 0
520 AMITTEL = ASUMME / ANZAHL
530 FOR T= 1 TO ANZAHL
540 VARIANZ = VARIANZ + ((BEOB(T)- AMITTEL)^2)/ANZAHL
550 NEXT T
560 PRINT "Anzahl = ";ANZAHL
570 PRINT "Mittelwert = ";AMITTEL
580 PRINT "Varianz = ";VARIANZ
590 PRINT "Min = ";AMIN
600 PRINT "Max = ";AMAX

700 REM *****Funktionenwahl
710 PRINT"Wahl der Funktionenfamilie : "
720 PRINT" 1: Hermite", "2: Laguerre", "3: Legendre", "4: nur
    Histogramm"
730 INPUT D$
740 IF (D$="1" OR D$="2" OR D$ ="3" OR D$="4") THEN GOTO 770
750 PRINT "Falsche Eingabe!!!"
760 GOTO 710
770 IF D$="4" THEN GOTO 800
780 PRINT"Maximaler Grad der Polynome (<= 20) ?"
790 INPUT MG

800 REM*****Datentransformation
810 DIM ATRB(2000)
820 CLS
830 PRINT "Daten werden geeignet transformiert..."
840 IF D$="1"THEN GOSUB 5000
850 IF D$="2" THEN GOSUB 5100
860 IF D$="3" THEN GOSUB 5300
870 IF D$="4" THEN GOSUB 5400
880 IF D$="4" THEN GOTO 3300
890 PRINT"Abspeicherung der transf. Daten unter tra(Dateiname)
    ?"
900 INPUT Z$
910 IF NOT Z$="j"THEN GOTO 1000
920 OPEN A$+":tra"+B$+".dat" FOR OUTPUT AS#4
930 FOR T= 1 TO ANZAHL
940 WRITE#4, ATRB(T)

```

```

950 NEXT T
960 CLOSE#4

1000 REM***** Berechnung der Polynomkoeffizienten
1010 DIM POLY(20,40)
1020 FOR T = 0 TO 20
1030 FOR U = 0 TO 40
1040 POLY(T,U) =0
1050 NEXT U
1060 NEXT T
1070 PRINT"Polynomkoeff. werden berechnet ..."
1080 IF D$="1" THEN GOSUB 6000
1090 IF D$="2" THEN GOSUB 6100
1100 IF D$="3" THEN GOSUB 6200

1200 REM*****Berechnung der Koeff. der quadr. Polynome
1210 PRINT"Koeff der quadr. Polynome werden berechnet"
1220 DIM QUPO(20,40)
1230 FOR K= 0 TO 20
1240 FOR L= 0 TO 40
1250 QUPO(K,L)=0
1260 NEXT L
1270 NEXT K
1280 FOR T= 0 TO MG
1290 FOR U =0 TO 2*T
1300 FOR M= 0 TO U
1310 QUPO(T,U)=QUPO(T,U) + POLY(T,M)*POLY(T,U-M)
1320 NEXT M
1330 NEXT U
1340 NEXT T

1400 REM *****Ausdruck der Koeffizienten
1410 CLS
1420 PRINT "Koeffizienten der Polynome"
1430 IF D$="1" THEN PRINT "Hermite"
1440 IF D$="2" THEN PRINT "Laguerre"
1450 IF D$="3" THEN PRINT "Legendre"
1460 FOR T=0 TO MG
1470 FOR U= 0 TO T
1480 PRINT POLY(T,U);" ";
1490 NEXT U
1500 PRINT
1510 NEXT T
1520 PRINT
1530 FOR T= 0 TO MG
1540 FOR U= 0 TO 2*T

```

```

1550 PRINT QUP0(T,U);" ";
1560 NEXT U
1570 PRINT
1580 NEXT T
1590 PRINT
1600 PRINT " Weiter mit RETURN"
1610 INPUT P$

1700 REM***** Berechnung der Momente
1710 CLS
1720 PRINT"Berechnung der Momente ..."
1730 DIM AMOMENT(40)
1740 FOR T= 0 TO 40
1750 AMOMENT(T)=0
1760 NEXT T
1770 LOCATE 2,2:PRINT ANZAHL
1780 FOR T%= 1 TO ANZAHL
1790 LOCATE 3,2:PRINT T%
1800 FOR L%= 1 TO 2*MG
1810 AMOMENT(L%)=AMOMENT(L%)+(ATRB(T%)^L%)/ANZAHL
1820 NEXT L%
1830 NEXT T%
1840 AMOMENT(0)=1
1850 BEEP
1860 PRINT"Ausgabe der Momente"
1870 FOR T= 0 TO 2*MG
1880 PRINT "m";T;" = ";AMOMENT(T)
1890 NEXT T
1900 PRINT "Weiter mit RETURN"
1910 INPUT P$

2000 REM *****Bestimmung der Wichtungsfaktoren
2010 PRINT"Berechnung der Wichtungsfaktoren"
2020 DIM W(20)
2030 IF D$="1" THEN GOSUB 6300
2040 IF D$="2" THEN GOSUB 6360
2050 IF D$="3" THEN GOSUB 6420

2100 REM *****THETA
2110 PRINT "Berechnung der besten Thetas"
2120 FOR T= 0 TO MG
2130 ATHETA(T)=0
2140 NEXT T
2150 FOR T= 0 TO MG
2160 FOR U= 0 TO T
2170 ATHETA(T) =ATHETA(T)+POLY(T,U)*AMOMENT(U)

```

```

2180 NEXT U
2190 ATHETA(T)=ATHETA(T)*W(T)
2200 NEXT T

2300 REM *****Sigmas
2310 PRINT "Schaetzer fuer Sigma"
2320 DIM ASIGMA1(20)
2330 ASIGMA1(0)=1
2340 FOR T= 1 TO 20
2350 ASIGMA1(T)=0
2360 NEXT T
2370 FOR T= 1 TO MG
2380 FOR U=0 TO 2*T
2390 ASIGMA1(T)=ASIGMA1(T) + QUP0(T,U)*AMOMENT(U)
2400 NEXT U
2410 ASIGMA1(T)= ASIGMA1(T)*(W(T)^2)
2420 NEXT T
2430 DIM ASIGMA2(20)
2440 FOR T= 0 TO MG
2450 ASIGMA2(T)=ASIGMA1(T)- (ATHETA(T)^2)
2460 NEXT T
2470 DIM PG(20)
2480 FOR T= 1 TO MG
2490 IF W(T) <> 0 THEN PG(T)=ATHETA(T)^2/ASIGMA2(T)*ANZAHL
2500 IF W(T)=0 THEN PG(T)=0
2510 NEXT T
2520 PG(0)=99999!

2600 REM *****Ausgabe der Sch"atzer
2610 CLS
2620 IF D$="1" THEN PRINT "Hermite"
2630 IF D$="2" THEN PRINT "Laguerre"
2640 IF D$="3" THEN PRINT "Legendre"
2650 PRINT"Grad", "Theta", "Sigma1", "Sigma2", "PG"
2660 PRINT
2670 FOR T=0 TO MG
2680 PRINT T, ATHETA(T), ASIGMA2(T), PG(T)
2690 NEXT T
2700 PRINT
2710 PRINT"Weiter mit RETURN .."
2720 INPUT P$
2730 CLS

2800 REM *****Kritische Werte
2810 DIM KRIT(4)
2820 KRIT(0) = 2

```

```

2830 KRIT(1) = 5.218
2840 KRIT(2) = 7
2850 KRIT(3) = 11.07
2860 KRIT(4)=1.1
2870 CLS
2880 IF D$="4" THEN GOTO 3300
2890 PRINT"Gewuenshtes Alpha eingeben ..."
2900 PRINT"0: Alpha = 0.35"
2910 PRINT"1: Alpha = 0.1"
2920 PRINT"2: Alpha = 0.05"
2930 PRINT"3: Alpha = 0.01"
2940 PRINT"4: Alpha = 0.5"
2950 INPUT GK
2960 IF NOT (GK=1 OR GK=2 OR GK= 3 OR GK = 0 OR GK=4) THEN GOTO
      2870
2970 FOR T= 0 TO MG
2980 IF PG(T) > KRIT(GK) THEN PRUEF(T)=1
2990 IF NOT PG(T)> KRIT(GK) THEN PRUEF(T)=0
3000 NEXT T

3100 REM *****Berechnung der Funktion
3110 CLS
3120 PRINT" Berechnung der Funktion"
3130 FOR T= 0 TO MG
3140 AKOEFF(T)=0
3150 NEXT T
3160 IF D$="1" THEN GOSUB 7000
3170 IF D$="2" THEN GOSUB 7000
3180 IF D$="3" THEN GOSUB 7000
3190 PRINT
3200 IF D$="1" THEN PRINT "exp(-x^2/2)/sqr(2PI)*";
3210 IF D$="2" THEN PRINT"exp(-x)*";
3220 PRINT"( ";
3230 FOR T= 0 TO MG
3240 PRINT AKOEFF(T);"x";T;" ";
3250 NEXT T
3260 PRINT ")"

3300 REM*****AUSDRUCK
3310 PRINT "Linke Grenze der Daten eingeben !"
3320 INPUT ALIN
3330 PRINT"Rechte Grenze eingeben"
3340 INPUT AREC
3350 PRINT "Markierungsdifferenzen horizontal?"
3360 INPUT DELTA
3370 PRINT "Anzahl der Intervalle "

```

```

3380 INPUT INTER
3390 LET ALPHA =(AREC -ALIN)/DELTA
3400 A= 560/INTER
3410 AUA=560/ALPHA

3500 REM *****Berechnung der H"aufigkeiten
3510 FOR T= 1 TO INTER
3520 RH(T)=0
3530 NEXT T
3540 FOR T= 1 TO ANZAHL
3550 ZEIGER =INT((ATRB(T) - ALIN)/(AREC - ALIN)*INTER)+1
3560 IF ZEIGER =INTER + 1 THEN ZEIGER =INTER
3570 IF (ZEIGER> 0 AND ZEIGER < INTER +1 ) THEN
    RH(ZEIGER)=RH(ZEIGER)+1
3580 NEXT T
3590 FOR T= 1 TO INTER
3600 RH(T)=RH(T)*INTER/(ANZAHL*(AREC - ALIN))
3610 NEXT T

3620 REM ****H"ochster Wert des Histogramms = 4/5 der H"ohe
3630 OBEN =0
3640 FOR T= 1 TO INTER
3650 IF RH(T)>OBEN THEN OBEN =RH(T)
3660 NEXT T
3670 FAKTOR = .4/OBEN
3680 FOR T= 1 TO INTER
3690 RH(T)=RH(T)*FAKTOR
3700 NEXT T
3710 SCREEN 3
3720 IF D$="4" THEN GOTO 3750
3730 PRINT"Nur Histogramm (der transf. Daten)?"
3740 INPUT Z$ 3750 CLS 3760 IF D$="4" THEN Z$="j"

3800 REM *****Ausdruck der Achsen und der Markierungen
3810 LINE(10,370) - (590,370)
3820 LINE(10,370) - (10,10)
3830 FOR T= 0 TO ALPHA
3840 LINE (20+T*AUA,370) - (20+T*AUA,375)
3850 NEXT T
3860 FOR T=0 TO INTER
3870 LINE (20+T*A,370) - (20+T*A,369)
3880 NEXT T
3890 IF OBEN > .5 THEN OBEN = OBEN/2
3900 NUM =INT(OBEN*10)
3910 FOR T=1 TO NUM+2
3920 LINE (10,370-T*28.8/OBEN) - (5,370- T*28.8/OBEN)

```

```

3930 LINE (10,370-(T-.5)*28.8/OBEN) - (8,370-(T-.5)*28.8/OBEN)
3940 NEXT T

4000 REM *****Ausdruck des Histogramms
4010 LINE (20,370)-(20,370-RH(1)*720)
4020 FOR T= 1 TO INTER -1
4030 LINE (20+(T-1)*A,-RH(T)*720+370) - (20+T*A,-RH(T)*720+370)
4040 LINE (20+T*A,-RH(T)*720+370)-(20+T*A,-RH(T+1)*720+370)
4050 NEXT T
4060 LINE(20 + (INTER - 1)*A,-RH(INTER)*720+370)-
      (20+INTER*A,-RH(INTER)*720 +370)
4070 LINE (20+INTER*A,-RH(INTER)*720+370)- (20+INTER*A,370)
4080 IF Z$="j" THEN GOTO 4370

4100 REM *****Ausdruck der approximierenden Funktion
4110 FOR T= 20 TO 580 STEP 1
4120 U=(T-20)/560*(AREC - ALIN) + ALIN
4130 F=0
4140 FOR S=0 TO MG
4150 F=F + AKOEFF(S)*U^S
4160 NEXT S
4170 IF D$="1" THEN F = F*EXP((-U^2)/2)/SQR(2*PI)*FAKTOR
4180 IF D$="2" THEN F=F *EXP(-U)*FAKTOR
4190 IF D$="3" THEN F=F*FAKTOR
4200 IF F>-.1 AND F<.5 THEN PSET(T,-F*720 + 370)
4210 NEXT T

4220 XL =462
4230 YU =85
4240 XR =548
4250 YO = 39
4260 LINE (XL,YU) - (XL,YO)
4270 LINE (XL,YO) - (XR,YO)
4280 LINE (XR,YO) - (XR,YU)
4290 LINE (XR,YU) - (XL,YU)
4300 LOCATE 4,60:PRINT"I = ";INTER
4310 LOCATE 5,60:PRINT" = ";
4320 IF GK=0 THEN PRINT"0.35"
4330 IF GK=1 THEN PRINT "0.10"
4340 IF GK=2 THEN PRINT "0.05"
4350 IF GK=3 THEN PRINT "0.01"
4360 IF GK=4 THEN PRINT "0.50"
4370 LOCATE 24,79
4380 INPUT Z$
4390 IF NOT Z$="q" THEN GOTO 2870
4400 END

```

```

5000 REM *****HERMITE
5010 FOR T= 1 TO ANZAHL
5020 ATRB(T)=(BEOB(T) - AMITTEL)/ SQR(VARIANZ)
5030 NEXT T
5040 PRINT"Mittelwert ist 0, Varianz = 1"
5050 RETURN

5100 REM *****Laguerre
5110 UNTER = 0
5120 AM =1
5130 FOR T= 1 TO ANZAHL
5140 IF (BEOB(T)-AMITTEL)/SQR(VARIANZ) + 1 < 0 THEN
    UNTER=UNTER+1
5150 NEXT T
5160 IF UNTER > 0 THEN PRINT UNTER;" von ";ANZAHL;" Werten <
    0":PRINT"TRANSFORMATION NACH (5.2.6) ERW"uNSCHT ?"
5170 INPUT A$
5180 IF NOT A$="j" THEN GOTO 5200
5190 AM =-SQR(VARIANZ)/(AMIN -AMITTEL)
5200 FOR T= 1 TO ANZAHL
5210 ATRB(T)=(BEOB(T)-AMITTEL)*AM/SQR(VARIANZ) + 1
5220 NEXT T
5230 PRINT "Mittelwert ist nun 1 , Varianz =";AM^2
5240 RETURN

5300 REM *****LEGENDRE
5310 PRINT"Geben Sie die untere Grenze der urspr. Daten an"
5320 INPUT AUG
5330 PRINT"obere Grenze ?"
5340 INPUT AOG
5350 FOR T= 1 TO ANZAHL
5360 ATRB(T)= 2*(BEOB(T) - AUG)/(AOG - AUG) -1
5370 NEXT T
5380 PRINT" Das Intervall ug;og wird auf -1;1 abgebildet"
5390 RETURN

5400 REM*****Umspeichern von Beob nach Atrb
5410 FOR T= 1 TO ANZAHL
5420 ATRB(T)=BEOB(T)
5430 NEXT T
5440 RETURN

6000 REM*****Koeffizienten Hermite
6010 POLY(0,0) =1
6020 POLY(1,1)=1

```



```

6030 FOR T= 2 TO MG
6040 POLY(T,0)=- (T-1)*POLY(T-2,0)
6050 FOR U= 1 TO T
6060 POLY(T,U)= POLY(T-1,U-1) - (T-1)* POLY(T-2,U)
6070 NEXT U
6080 NEXT T
6090 RETURN

6100 REM *****Koeffizienten Laguerre
6110 FOR T= 0 TO MG
6120 POLY(T,T) = (-1)^T
6130 FOR U= T-1 TO 0 STEP -1
6140 POLY(T,U)= -(U+1)^2/(T-U)*POLY(T,U+1)
6150 NEXT U
6160 NEXT T
6170 RETURN

6200 REM *****Koeffizienten Legendre
6210 POLY(0,0)=1
6220 POLY(1,1)=1
6230 FOR T= 2 TO MG
6240 POLY(T,0)=(1-T)/T*POLY(T-2,0)
6250 FOR U=1 TO T
6260 POLY(T,U)=(2*T-1)/T*POLY(T-1,U-1) - (T-1)/T*POLY(T-2,U)
6270 NEXT U
6280 NEXT T
6290 RETURN

6300 REM *****Wichtung Hermite
6310 W(0)=1
6320 FOR T=1 TO MG
6330 W(T)=W(T-1)/T
6340 NEXT T
6350 RETURN
6360 REM *****Wichtung Laguerre
6370 W(0)=1
6380 FOR T= 1 TO MG
6390 W(T)=W(T-1)/T^2
6400 NEXT T
6410 RETURN
6420 REM *****Wichtung Legendre
6430 FOR T= 0 TO MG
6440 W(T)=((2*T +1)/2)
6450 NEXT T
6460 RETURN

```

```
7000 REM *****Funktion Hermite/Laguerre/Legendre
7010 FOR U= 0 TO MG
7020 FOR T= 0 TO MG
7030 AKOEFF(U)=AKOEFF(U)+ ATHETA(T)*PRUEF(T)*POLY(T,U)
7040 NEXT T
7050 NEXT U
7060 RETURN
```

## 9 Danksagung

Für die Genehmigung zur Verwendung der Daten aus Anhang 8.4.1, bzw. für die technische Abwicklung möchte ich mich bei Prof. Dr. Siegfried Häberle und Herrn Axel Quentin vom Institut für Waldarbeit und Forstmaschinenkunde, Göttingen, bedanken.

Die Daten aus Anhang 8.4.2 und Anhang 8.4.3 stellte freundlicherweise Dr. Joachim Saborowski vom Institut für forstliche Biometrie und Informatik, Göttingen, zur Verfügung.

Schließlich danke ich Prof. Dr. Brunner, Abt. Medizinische Statistik, für die freundliche Beratung, bzw. Prof. Dr. Seidel und Dr. Cremer, Abt. Klinische Chemie der Universitätsklinik Göttingen, für die Erlaubnis zur Auswertung der Daten aus Anhang 8.4.4. Dipl.-Statist. Rainer Muche übernahm die Aufbereitung des Datensatzes.

Unterstützt wurde ich durch die Mitarbeiter des Instituts für Statistik und Ökonometrie, Göttingen, herausheben möchte ich die persönliche Betreuung durch Dipl.-Math. Martina Hochhaus.

## Literatur

- [1] LINHART, H. (1984). Modellauswahl und die Analyse von Kontingenztafeln. Vorlesungsarbeit, Institut für Statistik und Ökonometrie, Universität Göttingen.
- [2] LINHART, H. (1987). A test for the selection of models based on systems of orthogonal functions. Internal report, Institut für Statistik und Ökonometrie, Universität Göttingen.
- [3] LINHART, H. and ZUCCHINI, W. (1986). Model selection. Wiley, New York.
- [4] COURANT, R. and HILBERT, D. (1953). Methods of mathematical physics, Vol. 1. Interscience Publishers, Inc., New York.
- [5] KRONMAL, R.A., and TARTER, M.E. (1968). The estimation of probability densities and cumulatives by Fourier series methods. J. Amer. Statist. Ass. 63, 925- 952.
- [6] TARTER, M.E., and KRONMAL, R.A. (1976). An introduction to the implementation and theory of nonparametric density estimation. Amer. Statist. 30, 105-112.
- [7] DIGGLE, P.J., and HALL, P. (1986). On the selection of terms in an orthogonal series density estimator. J. Amer. Statist. Ass. 81, 230-233.
- [8] KENDALL, M.G. and STUART, A. (1979). The advanced theory of statistics, Vol. 1: Distribution theory, 4th edition. Griffin, London.
- [9] HALL, P. (1982). Comparison of two orthogonal series methods of estimating a density and its derivatives on an interval. Journal of Multivariate Analysis 12, 432-449.
- [10] HÄBERLE, S. (1987). Persönliche Mitteilung. Institut für Waldarbeit und Forstmaschinenkunde, Universität Göttingen.
- [11] SABOROWSKI, J. (1987). Persönliche Mitteilung. Institut für forstliche Biometrie und Informatik, Universität Göttingen.
- [12] SEIDEL, D. (1987). Persönliche Mitteilung. Universitäts- klinik Göttingen, Abt. Klinische Chemie.