

HEINRICH-HERTZ-INSTITUT — BERLIN-CHARLOTTENBURG

Technischer Bericht Nr. 163

Wahrscheinlichkeitsdichte-Approximation mit Hilfe von
Momenten zur Zeichenerkennung und Prozeßbeschreibung

von

Dr.-Ing. Hans Dietrich Höhne



Berlin

1 9 7 3

HEINRICH - HERTZ - INSTITUT

Berlin-Charlottenburg

Technischer Bericht Nr. 163

Wahrscheinlichkeitsdichte-Approximation mit Hilfe von
Momenten zur Zeichenerkennung und Prozeßbeschreibung.

von

Dr.-Ing. H. D. Höhne



Berlin

1973

Wahrscheinlichkeitsdichte-Approximation mit Hilfe von
Momenten zur Zeichenerkennung und Prozeßbeschreibung.

Zusammenfassung:

Es wurde ein adaptives Verfahren zur Wahrscheinlichkeitsdichte-approximation entwickelt.

Mögliche Anwendungsgebiete sind Prozeßbeschreibung und Zeichenerkennung.

Als Momentenverfahren benötigt es keine Speicherung der Einzelwerte.

Es werden Äquivalenzklassen gebildet, repräsentiert durch Normalverteilungen. Jeder Äquivalenzklasse wird ein Satz von Momenten zugeordnet. Während des Datenempfanges werden die einlaufenden Zeichen durch Verändern des Momentensatzes der gewählten Äquivalenzklasse eingordnet.

Während der blockweisen Verarbeitung wird ein Test auf die einer jeden Klasse zugeordneten Momentensätze angewendet, um zu prüfen, ob Klassen zu zerlegen oder mit Nachbarklassen zusammenzulegen sind.

Dadurch können Klassenzahl und Klassengrenzen während des Datenempfanges adaptiv bestimmt werden.

Stehen nicht genügend Daten zur Verfügung, so kann der gleiche Datenblock mehrmalig eingelesen werden.

Eine Modifikation des Verfahrens zur Approximation von Mischverteilungen, bei denen Häufigkeiten in Klassen vorhanden sind, liegt ebenfalls vor.

HEINRICH - HERTZ - INSTITUT

Der Bearbeiter

Möhne

(Dr.-Ing. H. D. Möhne)

Der Abteilungsleiter

Der Institutsdirektor

Berger

Berger

(Prof. Dr.-Ing. Erich R. Berger) (Prof. Dr.-Ing. Erich R. Berger)

Inhaltsverzeichnis

	Seite
1. <u>Einleitung</u>	1
2. <u>Kurze Beschreibung des gewählten Algorithmus</u>	3
3. <u>Erstellung des Parametersatzes</u>	5
3.1 Nullpunktsmomente von Normal-, Dreieck- und Rechteckverteilung	5
3.1.1 Bezeichnungen	5
3.1.2 Nullpunktsmomente der $N(y, \mu, \sigma^2)$ - Verteilung	5
3.1.3 Nullpunktsmomente der Dreiecksverteilung	8
3.1.4 Nullpunktsmomente der Rechtecksverteilung	10
3.2 Zentrale Momente	11
3.2.1 Zusammenhang mit den Nullpunktsmomenten	11
3.2.2 Grundsätzliche Anwendbarkeit des Verfah- rens von K. Pearson auf mehr als zwei Ver- teilungen	12
3.3 Empirische Momente	15
3.3.1 Bezeichnungen	15
3.3.2 Zusammenhang der empirischen zentralen Mo- mente und der empirischen Nullpunktsmomente	16
3.3.3 Erwartungswert der empirischen Momente	17
3.3.3.1 Erwartungswert der Nullpunktsmomente	17
3.3.3.2 Erwartungswert der empirischen zentralen Momente	17
3.3.3.3 Herleitungen	19
3.3.3.4 Zerlegung in Summanden und zugehörige Erwartungswerte	21

3.3.3.5	Multiplikatoren der Erwartungswerte	21
3.3.3.6	Erwartungswert von m_5	26
3.3.3.7	Deutung der Ergebnisse	28
3.3.3.8	Biasfreie Schätzwerte der empirischen zentralen Momente	28
3.3.3.9	Rechengang für den biasfreien Schätzwert für μ_6	30
4.	<u>Zerlegungsverfahren nach K. Pearson</u>	32
4.1	Beschreibung des Zerlegungsverfahrens	33
4.2	Notwendige Bedingung dafür, daß zu p_2 eine Zerlegung gehört	36
4.3	Notwendige und hinreichende Bedingung dafür, daß zu p_2 eine Zerlegung gehört	37
4.4	Ausschließen von $p_2 = 0$ und $p_2 = -\mu_2$ als Wurzeln von Gleichung (21)	39
4.5	Numerische Betrachtung des Auftretens der Doppellösung	41
4.6	Schlußfolgerung für die praktische Anwendung	47
5.	<u>Signifikanztest für die Abweichung von der Normalität als Entscheidungskriterium für die Zerlegung</u>	49
5.1	Formfaktoren	49
5.2	Test auf Normalität	52

6.	<u>Nebenbedingungen</u>	53
6.1	Zweigipfligkeit	55
6.1.1	Ableitung der Gleichungen für die logarithmische Funktion und Parabel	58
6.1.2	Deutung der Gleichungen	61
6.1.3	Ersatzfunktion	63
6.1.4	Zahl der Schnittpunkte zwischen -1 und $+1$	64
6.1.4.1	Bestimmung der Schnittpunktsanzahl mit der Sturm'schen Kette	72
6.1.4.2	Einfache Bestimmung der Schnittpunkts- anzahl zwischen Ersatzfunktion und Parabel	78
6.2	Existenz der Bayesgrenze	82
7.	<u>Durchführung des Approximationsverfahrens</u>	85
7.1	Beschreibung des vorhandenen Programms	85
7.2	Noch nicht verwirklichte Möglichkeiten zur Verbesserung	90
7.2.1	Beschleunigung des Programms	90
7.2.2	Optimierung	91
8.	<u>Erzielte Ergebnisse</u>	92
8.1	Mischungen aus Normalverteilungen	92
8.2	Nicht-normale Mischungen	102
8.3	Zerlegung unter Nebenbedingungen	110
8.4	Approximation gemessener Wahrscheinlich- keitsdichten	113

9.	<u>Abschluß</u>	118
9.1	Leistungsfähigkeit und Grenzen des Verfahrens	118
9.2	Nebenergebnisse	119
9.3	Erweiterungen	120

Anhang I

Symmetrische Mischverteilungen	121
--------------------------------	-----

Anhang II

Sturm'sche Kette	130
------------------	-----

1. Einleitung

Ausgangspunkt der Untersuchung war die Frage, ob man beim Klassifizieren von Stichproben so verfahren könne, daß man die benötigte Klassenzahl und die Klassengrenzen adaptiv bestimmen könne, d.h., ohne vorher das Kurvenbild der empirischen Verteilung zu ermitteln.

Das ist vor allem wichtig, wenn man den Algorithmus zur Zeichen-erkennung verwendet. Jeder Klasse entspricht ein Merkmal von Taxemen, die weder nach Größe noch Anzahl bekannt zu sein brauchen (unsupervised learning). Solange man keine Vorinformation über die Verteilung hat, wäre man bei der üblichen Speicherung der Häufigkeit von Amplitudenklassen auf Vermutungen über die zu wählende Klassenzahl angewiesen und würde bei falscher Wahl entweder Information verschenken oder unnützen Aufwand treiben.

Dadurch scheidet eine größere Anzahl von Verfahren zur Trennung von Mischverteilungen aus, die an sich sonst Vorteile haben. Dazu gehört das Verfahren von Choi/Bulgren /1/, das die beobachtete mit der vermuteten Verteilung korreliert, und für die Zerlegung in Normalverteilungen das bereits 1936 von G. Doetsch /2/ angegebene, jedoch weitgehend unbeachtet gebliebene Verfahren der Gauß-Transformation, das innerhalb eines endlichen Intervalls die gesuchten Normal-Komponenten in einer empirischen Verteilungsfunktion mit Hilfe von FFT-Algorithmen zu bestimmen gestattet. Aus dem gleichen Grund wird auf die grafischen Verfahren nicht zurückgegriffen, von denen diejenigen von Weichselberger (1961) /3/ und von Bhattacharya (1967) /4/ genannt zu werden verdienen. Letzterer erwähnt u.a. das Zerlegungsverfahren von K. Pearson (1894) /5/ aufgrund der Momente und bemerkt: "The difficulties encountered with these methods increase at a tremendous rate as the number of components increases, and the general problem in which the number of components is unknown and may be more than two, does not seem to have been considered in the statistical literature."

Aus den vorher erwähnten Gründen, d.h., um den Speicherplatz klein und von der Anzahl der Beobachtungen unabhängig zu halten, soll hier für jede erkannte oder vermutete Komponente nur ein Satz von wenigen Parametern (hier sechs Momente) gespeichert werden. Daraus ergab sich, daß bei der Zerlegung auf ein Momentenverfahren zurückgegriffen werden mußte. Dasjenige von K. Pearson erschien am geeignetsten. Seine Bedeutung wird auch dadurch unterstrichen, daß Robertson und Fryer /6/ (1970) die Genauigkeit der Schätzwerte für die Komponenten bei der Zerlegung nach dem Verfahren nach K. Pearson abschätzen. Dabei weisen sie darauf hin, daß es Fälle gibt, in denen andere Verfahren versagen, so z.B. die Maximum likelyhood Methode bei der Schätzung der fünf Parameter aus einer Mischung zweier Normalverteilungen.

Der zitierte Einwand Bhattacharyas dürfte durch das vorgelegte Verfahren entkräftet sein.

Die Zerlegungen nach K. Pearson stellen Zwischenlösungen dar. Sie haben nur vorläufigen Charakter und werden laufend verbessert; die Frage nach der Zuverlässigkeit der Schätzwerte aus den Pearson-Verfahren ist daher für uns von untergeordneter Bedeutung.

Schließlich sei noch bemerkt, daß die Normalverteilung als Komponente nur stellvertretend für irgendeine andere Verteilungsart steht. Sie ist der Repräsentant der Äquivalenzklasse aller eingipfligen Komponenten mit gegebener Lage, Höhe und Breite des Maximums.

Ein auf Momenten basierendes Zerlegungsverfahren haben Rider /7/ (1961) und Blischke (1962) /8/ für Exponential- und Binomialverteilungen entwickelt.

Für die Verwendung von Normal-Komponenten spricht jedoch, daß sie häufig auftreten, z.B. bei Verkehrsmessungen in der Fernmeldetechnik (Rahko), bei Strommessungen an Halbleitern, in der Spektroskopie bei genetischen Untersuchungen, in der Fischerei-

kunde, um nur einige zu nennen.

Auch in der Regelungstechnik ist die genauere Approximation von Dichtefunktionen zur Beschreibung stochastischer Variabler (Regel- und Störgrößen) im nichtlinearen Regelkreis durch Normalverteilungen wegen der besonderen Rolle des Gauß-Markoffprozesses z.B. bei der Kalmanfilterung /9/ möglicherweise von Interesse.

2. Kurze Beschreibung des gewählten Algorithmus

Die Daten werden blockweise verarbeitet. Sie sollen entweder einem stationären Prozeß entstammen, es werden diesem Prozeß dann mehrere Blöcke entnommen, oder einem veränderlichem Prozeß, dann kann ein Datenblock z.B. in einen billigen Umlaufspeicher übernommen und daraus mehrmals abgelesen werden.

Aus den Eingangsdaten wird für jede erkannte bzw. vermutete Komponente ein Satz von 6 Parametern (Momenten) gebildet; zu Beginn der Approximation wird zunächst nur ein Momentensatz gebildet. Für jede Komponente wird nun getestet, ob sie signifikant von der Form der Normalverteilung abweicht. Jede nichtnormale Komponente wird nach K. Pearson in zwei normale Komponenten zerlegt, falls dies möglich ist. Diese Zerlegung kann unter bestimmten Nebenbedingungen verhindert oder rückgängig gemacht werden. Die Nebenbedingungen, unter denen man trotz Zerlegbarkeit wählen kann, nicht zu zerlegen, sind:

1. Unterschreiten einer Schwelle S für den Anteil einer Komponente an der Gesamtmischung (Mischungsfaktor) z.B. $S = 1 \%$
2. Fehlen einer Bayesgrenze zwischen beiden Komponenten der Zerlegung
3. Fehlen einer zweigipfligen Summe aus beiden Komponenten

Die gewonnenen Komponenten werden nach der Größe des Mittelwertes sortiert.

Nach der Zerlegung werden die Daten des nächsten abgelesenen Blockes auf die Komponenten verteilt, d.h. eine Meßgröße wird

zur Korrektur des Momentensatzes derjenigen Komponente verwendet, in deren "Bereich" sie fällt. Im Bereich sich überschneidender Komponenten wird sie anteilig nach Maßgabe der Dichten am Ort der Meßgrößen auf die vorhandenen Komponenten verteilt.

Im anschließenden Test auf Zusammenlegen werden die Parametersätze von je zwei benachbarten Komponenten daraufhin untersucht, ob sie zusammengenommen signifikant vom Typ der Normalverteilung abweichen. Zusätzlich kann auch geprüft werden, ob zwischen ihnen eine Bayesgrenze existiert oder ob ihre Summe zweigipflig ist. Je nach Ausfall des Tests werden die zwei benachbarten Komponenten wieder zusammengelegt.

Einlesen und Aufteilen eines neuen oder erneut des gleichen Datenblocke, falls zusammengelegt wurde, Test auf Zerlegen, Prüfung der Nebenbedingungen, Zerlegen, Sortieren der Komponenten, Einlesen eines weiteren oder erneut des gleichen Datenblockes, falls zerlegt wurde, Aufteilung, Test auf Zusammenlegen, Prüfung der Nebenbedingungen, Zusammenlegen werden so lange in einer Schleife hintereinander ausgeführt, bis keine Zerlegungen oder Zusammenlegungen mehr erfolgen, d.h. bis die gefundenen Komponenten endgültig sind.

Das Verfahren ist im Blockdiagramm S. 86 dargestellt.

In den folgenden Abschnitten des Berichtes werden die einzelnen Schritte einschließlich der dazu benötigten Formeln behandelt.

3. Erstellung des Parametersatzes

Die erste Stufe der Simulation umfaßt die Zerlegung einer Mischung von mehreren Normal-, Dreiecks- oder Rechteckskomponenten in zwei Normalkomponenten (Blockbild rechts oben, Ausgang b). Dazu sind die theoretischen zentralen Momente dieser Mischungen zu berechnen.

3.1 Nullpunktsmomente von Normal-, Dreieck- und Rechteckverteilung

3.1.1 Bezeichnungen:

X	stochastische Variable
$P\{X \leq x\} = F(x)$	Verteilungsfunktion
$\alpha_n = E\{X^n\}$	Nullpunktsmomente von X
$\alpha_0 = 1$	
$\alpha_1 = \mu$	Mittelwert
$\mu_n = E\{(X-\mu)^n\}$	zentrale Momente
$\mu_0 = 1$	
$\mu_1 = 0$	
$\mu_2 = \sigma^2$	Varianz

3.1.2 Nullpunktsmomente der $N(y, \mu, \sigma^2)$ -Verteilung:

Es sei $X \sim N(x, 0, 1)$ verteilt mit der Wahrscheinlichkeitsdichte (WD) $g(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$ und der charakteristischen Funktion

$$\Psi(t) = E\{e^{itX}\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{itu} e^{-\frac{1}{2}u^2} du = e^{-\frac{1}{2}t^2}$$

und $Y = \sigma X + \mu$ eine Transformation der stochastischen Variablen X , so daß $Y \sim N(y, \mu, \sigma^2)$ verteilt ist.

Die zu Y gehörige charakteristische Funktion ist:

$$\varphi(t) = E \{ e^{itY} \} = E \{ e^{it\sigma X} \} \cdot E \{ e^{it\mu} \} = \psi(\sigma \cdot t) \cdot e^{it\mu}$$

$$\varphi(t) = e^{-\frac{1}{2}\sigma^2 t^2} \cdot e^{it\mu} = e^{it\mu - \frac{1}{2}\sigma^2 t^2}$$

$$\varphi^{(n)}(t) = \varphi^{(n-1)}(t) [i\mu - \sigma^2 t] - (n-1) \sigma^2 \varphi^{(n-2)}(t)$$

was man durch Schluß von n auf $n+1$ beweist.

$$\varphi^{(n+1)}(0) = i\mu \varphi^{(n)}(0) + i^2 n \sigma^2 \varphi^{(n-1)}(0)$$

und mit $\alpha_n = i^{-n} \varphi^{(n)}(0)$

$$\boxed{\alpha_{n+1} = \mu \alpha_n + n \sigma^2 \alpha_{n-1}}$$

Damit erhält man

$$\alpha_0 = 1$$

$$\alpha_1 = \mu$$

$$\alpha_2 = \mu \cdot \mu + 1 \cdot \sigma^2 \cdot 1 = \mu^2 + \sigma^2$$

$$\alpha_3 = \mu^3 + 3 \mu \sigma^2$$

$$\alpha_4 = \mu^4 + 6\mu^2 \sigma^2 + 3 \sigma^4$$

$$\alpha_5 = \mu^5 + 10 \mu^3 \sigma^2 + 15 \mu \sigma^4 \tag{1}$$

$$\alpha_6 = \mu^6 + 15\mu^4 \sigma^2 + 45\mu^2 \sigma^4 + 15\sigma^6$$

Die Koeffizienten stimmen mit denjenigen der Hermitischen Polynome $He_n(x)$ überein. Die hierzugehörige Rekursionsformel findet man in /10/ S. 782

$$He_{n+1}(x) = x He_n(x) - n He_{n-1}(x)$$

Durch Vergleich

$$\boxed{\alpha_n = i^{-n} \sigma^n He_n \left(\frac{i\mu}{\sigma} \right)}$$

und unter Benutzung der Formel 22.3.1 in /10/ S. 775 ergibt sich schließlich:

$$\alpha_n = \sigma^n n! \sum_{m=0}^{\lfloor \frac{n}{2} \rfloor} \frac{\left(\frac{\mu}{\sigma}\right)^{n-2m}}{m! 2^m (n-2m)!}$$

3.1.3 Nullpunktsmomente der Dreiecksverteilung

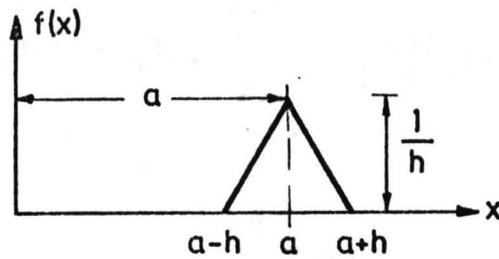


Bild 1 a Dreiecksverteilung

$$f(x) = \begin{cases} \frac{1}{h^2} [x - (a-h)] & \text{für } a-h \leq x \leq a \\ -\frac{1}{h^2} [x - (a+h)] & \text{für } a \leq x \leq a+h \\ 0 & \text{sonst} \end{cases}$$

$$\alpha_n = \int_{a-h}^{a+h} u^n f(u) du$$

$$= \frac{1}{h^2} \left[\frac{2a^{n+2}}{n+2} - \frac{(a-h)^{n+2}}{n+2} - \frac{(a+h)^{n+2}}{n+2} \right.$$

$$\left. - (a-h) \frac{a^{n+1}}{n+1} - (a+h) \frac{a^{n+1}}{n+1} + (a-h) \frac{(a-h)^{n+1}}{n+1} \right.$$

$$\left. + (a+h) \frac{(a+h)^{n+1}}{n+1} \right] \quad \text{und nach}$$

Umformung

$$\alpha_n = \frac{1}{h^2 (n+1) (n+2)} \left[(a+h)^{n+1} - 2a^{n+2} + (a-h)^{n+2} \right]$$

$$\text{mit } (a+h)^n = \sum_{i=0}^n \binom{n}{i} a^{n-i} h^i$$

$$\begin{aligned}
 & (a+h)^{n+2} - 2a^{n+2} + (a-h)^{n+2} = \\
 & = \sum_{i=1}^{n+2} \binom{n+2}{i} a^{n+2-i} h^i + \sum_{i=1}^{n+2} \binom{n+2}{i} (-1)^i a^{n+2-i} h^i \\
 & = h^2 \left[(n+1)(n+2) a^n + 2 \sum_{i=2,4,\dots}^n \binom{n+2}{i+2} a^{n-i} h^i \right]
 \end{aligned}$$

also

$$\alpha_n = \frac{2}{(n+1)(n+2)} \sum_{i=0}^{\lfloor \frac{n}{2} \rfloor} \binom{n+2}{2i+2} a^{n-2i} h^{2i}$$

$$\alpha_0 = 1$$

$$\alpha_1 = a$$

$$\alpha_2 = a^2 + \frac{h^2}{6}$$

$$\alpha_3 = a^3 + \frac{1}{2} ah^2$$

$$\alpha_4 = a^4 + a^2 h^2 + \frac{h^4}{15}$$

$$\alpha_5 = a^5 + \frac{5 a^3 h^2}{3} + \frac{ah^4}{3}$$

(2)

$$\alpha_6 = a^6 + \frac{5a^4 h^2}{2} + a^2 h^4 + \frac{h^6}{28}$$

3.1.4 Nullpunktsmomente der Rechteckverteilung

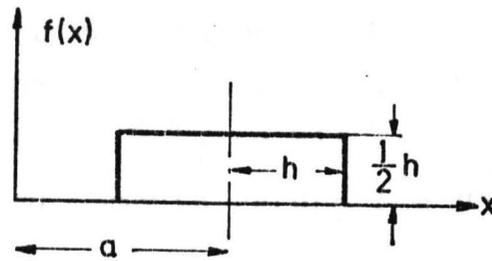


Bild 1 b Rechteckverteilung

$$\alpha_n = \int_{a-h}^{a+h} u^n \frac{1}{2h} dt$$

$$= \frac{1}{2h(n+1)} \left[(a+h)^{n+1} - (a-h)^{n+1} \right]$$

$$\alpha_n = \frac{1}{n+1} \sum_{i=0}^n (a+h)^i (a-h)^{n-i}$$

$$\alpha_n = \frac{1}{n+1} \sum_{k=0}^{[n/2]} \binom{n+1}{2k+1} a^{n-2k} h^{2k}$$

$$\alpha_0 = 1$$

$$\alpha_1 = a$$

$$\alpha_2 = a^2 + \frac{h^2}{3}$$

$$\alpha_3 = a^3 + ah^2$$

$$\alpha_4 = a^4 + 2a^2h^2 + \frac{h^4}{5}$$

$$\alpha_5 = a^5 + \frac{10}{3} a^3h^2 + ah^4$$

$$\alpha_6 = a^6 + 5a^4h^2 + 3a^2h^4 + \frac{h^6}{7}$$

(3)

3.2. Zentrale Momente

3.2.1 Zusammenhang mit den Nullpunktsmomenten.

Die zentralen Momente $\mu_n = E\{(X-\mu)^n\} = \int_{-\infty}^{+\infty} (X-\mu)^n f(x) dx$ sind unabhängig vom Skalennullpunkt. Wählt man diesen bei μ , setzt man also $\mu = 0$, so wird $\mu_n = E\{(x-\mu)^n\} = E\{x^n\} = \alpha_n$.

Die zentralen Momente werden für $\mu=0$ zu Nullpunktsmomenten.

Bei einer Mischung von i Komponenten mit den Mittelwerten

$$\alpha_{1(1)} = \int_{-\infty}^{+\infty} x dF_1(x)$$

$$\alpha_{1(2)} = \int_{-\infty}^{+\infty} x dF_2(x)$$

$$\alpha_{1(i)} = \int_{-\infty}^{+\infty} x dF_i(x)$$

und den Mischungsfaktoren

z_1, z_2, \dots, z_i

$$\text{mit } \sum_k z_k = 1$$

ist

$$\alpha_{1(\text{ges})} = \sum_k z_k \alpha_{1(k)}$$

Verschiebt man den Skalennullpunkt um diesen Wert, dann erhält man die neuen Mittelwerte γ :

$$\gamma_1 = \alpha_{1(1)} - \alpha_{1(\text{ges})}$$

$$\gamma_2 = \alpha_{1(2)} - \alpha_{1(\text{ges})}$$

$$\gamma_i = \alpha_{1(i)} - \alpha_{1(\text{ges})}$$

Die Momente um den neuen Nullpunkt sind nun auch zentrale Momente der Mischung:

$$\mu_0 = \sum_k z_k = 1$$

$$\mu_1 = \sum_k z_k \gamma_k = \sum_k z_k \alpha_{1(k)} - \alpha_{1(\text{ges})} \sum_k z_k$$

$$\mu_1 = \alpha_{1(\text{ges})} - \alpha_{1(\text{ges})} = 0$$

Auch die weiteren zentralen Momente erhält man aus den Nullpunktsmomenten der Komponenten, wobei nur zu beachten ist, daß der neue Skalennullpunkt gegenüber dem alten um $\alpha_1(\text{ges})$ verschoben ist. Für die Mischung zweier normalverteilter Komponenten ergibt sich:

$$\mu_0 = z_1 + z_2$$

$$\mu_1 = z_1 \gamma_1 + z_2 \gamma_2$$

$$\mu_2 = z_1 (\gamma_1^2 + \sigma_1^2) + z_2 (\gamma_2^2 + \sigma_2^2)$$

$$\mu_3 = z_1 (\gamma_1^3 + 3\gamma_1 \sigma_1^2) + z_2 (\gamma_2^3 + 3\gamma_2 \sigma_2^2) \quad (4)$$

...

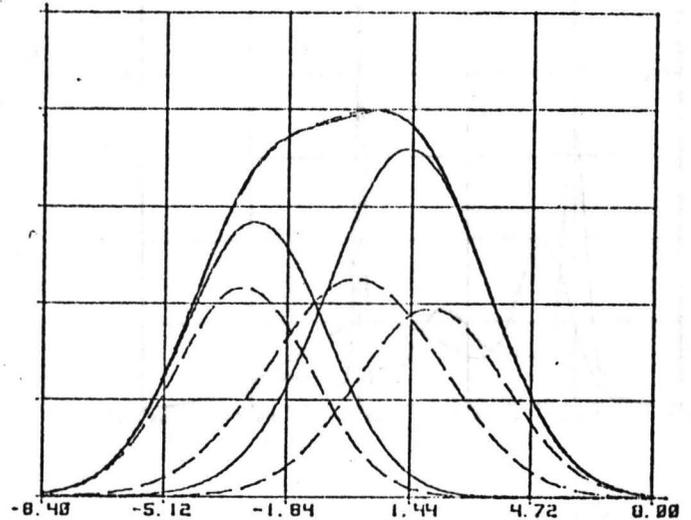
3.2.2 Grundsätzliche Anwendbarkeit des Verfahrens von Pearson

Am Ausgang b des Blockdiagrammes stehen die theoretischen zentralen Momente von Mischungen aus Normal-, Dreieck- und Rechteckverteilungen zur Verfügung. Es können nicht nur Mischungen von Komponenten eines Typs, sondern auch solche, die Komponenten mehrerer oder aller Typen enthalten, gebildet werden. Natürlich war die Frage, ob sich solche Mischungen überhaupt aufgrund der Momente in zwei Normalverteilungen trennen lassen, grundlegend für die weitere Fortführung der Aufgabe.

Bedenken hiergegen bestanden vor allem wegen der Aussage in der Literatur /10/S. 304, daß für die Anwendung des Zerlegungsverfahrens von K. Pearson sichergestellt sein müsse, daß die Gesamtheit eine Mischung von nur zwei homogenen Gruppen sei.

Zerlegung einer Mischung
von drei Normalvertei-
lungen aufgrund der
theoretischen Momente

Bild 2



EINGEGEBENE WERTE

(GAUSS-VERTEILUNG)

AA[1] =	.300000	AMY[1] =	2.000000	ASIGM[1] =	2.000000
AA[2] =	.400000	AMY[2] =	.000000	ASIGM[2] =	2.300000
AA[3] =	.300000	AMY[3] =	-3.000000	ASIGM[3] =	1.800000

AUS EINGABEWERTEN BERECHNETE ZENTRALE MOMENTE DER MISCHVERTEILUNG

MY1 =	.000000000000E 00	MY2 =	.809799999999E 01	MY3 =	.671999999999E-01
MY4 =	.163516379998E 03	MY5 =	.199453799985E 02	MY6 =	.498981496492E 04

LOESUNGSKOMPONENTEN

BA[1] =	.411029	BMY[1] =	-2.689724	BSIGM[1] =	1.877240
BA[2] =	.588971	BMY[2] =	1.367732	BSIGM[2] =	2.126809

AM1 =	-.109139364213E-10	AM2 =	.809799999998E 01	AM3 =	.671999999998E-01
AM4 =	.163516379998E 03	AM5 =	.199453799983E 02	AM6 =	.501442615858E 04

Im Bild 2 ist ein Beispiel für die Zerlegbarkeit einer Mischung aus drei Normalkomponenten zu sehen.

Die eingegebenen drei Normalkomponenten (strichliert) überlagern sich zu einer abgeplatteten eingipfligen Summe, die aufgrund der Momente in zwei Normalkomponenten zerlegt wird (durchgezogen). Die Summenkurve dieser zwei Normalkomponenten ist im Rahmen der Zeichengenaugigkeit deckungsgleich mit der Summenkurve aus den drei ursprünglichen Komponenten.

Zerlegung einer Mischung von fünf Normalverteilungen aufgrund der theoretischen Momente.

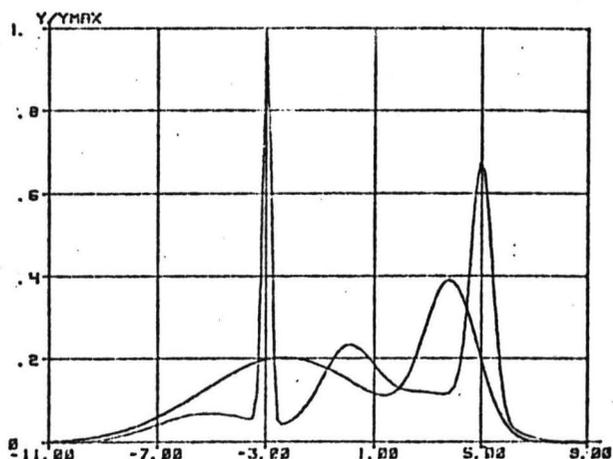


Bild 3

EINGEGEBENE WERTE

AA[1] =	.150000	AMY[1] =	=5.000000	ASIGM[1] =	2.000000
AA[2] =	.150000	AMY[2] =	=3.000000	ASIGM[2] =	.150000
AA[3] =	.200000	AMY[3] =	.000000	ASIGM[3] =	1.000000
AA[4] =	.250000	AMY[4] =	3.000000	ASIGM[4] =	2.000000
AA[5] =	.250000	AMY[5] =	5.000000	ASIGM[5] =	.400000

AUS EINGABEWERTEN BERECHNETE ZENTRALE MOMENTE DER MISCHVERTEILUNG

AMY1 =	-.727595761418E+11	AMY2 =	.148033750000E 02	AMY3 =	-.202704750013E 02
AMY4 =	.460012637807E 03	AMY5 =	-.188007513851E 04	AMY6 =	.231995416905E 05

Lösung:

A =	.631984345884E 00	CMY =	-.145993316459E 01	CSIGMA =	.299914706072E 01
1-A =	.368015654114E 00	DMY =	.468092834318E 01	DSIGMA =	.972580907975E 00

AUS ERGEBNISSEN ERRECHNETE ZENTRALE MOMENTE DER MISCHVERTEILUNG

AMY1 =	-.727595761418E+11	AMY2 =	.148033749997E 02	AMY3 =	-.202704750011E 02
AMY4 =	.460012637800E 03	AMY5 =	-.188007513848E 04	AMY6 =	.236283955643E 05

Auf das Problem der Zerlegbarkeit soll später noch näher eingegangen werden. Die Beispiele in Bild 2 und 3 entkräften grundsätzliche Einwände.

Der nächste Schritt in der Verwirklichung eines Klassifizierungsverfahrens, das ohne Speicherung der Zwischenwerte arbeiten soll, ist die Berechnung des für die Zerlegung benötigten Momentensatzes aus Realisierungen des betrachteten Prozesses. Die daraus zu berechnenden biasfreien empirischen Momente dienen als Schätzwerte für die tatsächlichen oder theoretischen Momente.

3.3 Empirische Momente

3.3.1 Bezeichnungen

X stochastische Variable

n Anzahl der Realisierungen

$$a_k = \frac{1}{n} \sum_{i=1}^n x_i^k$$

k-tes empirisches
Nullpunktsmoment

insbesondere:

$$a_0 = 1$$

$$a_1 = M$$

empirischer Mittelwert

$$m_k = \frac{1}{n} \sum_{i=1}^n (x_i - M)^k$$

k-tes empirisches zentrales
Moment mit Bias

insbesondere:

$$m_0 = 1$$

$$m_1 = 0$$

$$m_2 = s^2$$

empirische Varianz mit Bias

a_k und m_k sind ihrerseits stochastische Variable.

3.3.2 Zusammenhang der empirischen zentralen Momente und der empirischen Nullpunktsmomente:

$$m_0 = 1$$

$$m_1 = \frac{1}{n} \sum_{i=1}^n (X_i - M) = \frac{1}{n} \sum X_i - \frac{1}{n} M = M - M = 0$$

$$m_2 = \frac{1}{n} \sum (X_i - M)^2 = \frac{1}{n} \sum X_i^2 - \frac{2M}{n} \sum X_i + M^2 = a_2 - M^2$$

$$m_3 = a_3 - 3a_2 M + 2 M^3$$

$$m_4 = a_4 - 4a_3 M + 6 a_2 M^2 - 3M^4$$

$$m_5 = a_5 - 5a_4 M + 10a_3 M^2 - 10a_2 M^3 + 4M^5$$

$$m_6 = a_6 - 6a_5 M + 15a_4 M^2 - 20 \cdot a_3 M^3 + 15a_2 M^4 - 5M^6$$

allgemein

$$m_k = \sum_{j=0}^k \binom{k}{j} (-1)^{k-j} a_j M^{k-j} \quad (5)$$

vergl. /10/ S. 928

3.3.3 Erwartungswert der empirischen Momente

Der Erwartungswert von a_k oder m_k wird unter der Voraussetzung berechnet, daß die X_i zu einer Grundgesamtheit mit der Verteilungsfunktion F gehören. F ist in unserem Fall die Mischverteilung. Die Voraussetzung einer gemeinsamen Grundgesamtheit für alle X_i ist jedoch nur formal erfüllt, da es ja gerade das Ziel des Verfahrens ist, diese in mehrere Komponentenverteilungen aufzulösen.

3.3.3.1 Erwartungswert der Nullpunktsmomente

$$E \{a_k\} = E \left\{ \frac{1}{n} \sum_{i=1}^n X_i^k \right\} = \frac{1}{n} \sum_{i=1}^n E \{X_i^k\}$$

$$E \{a_k\} = \frac{1}{n} n \cdot \alpha_k = \alpha_k$$

$E \{a_k\} = \alpha_k$

Die empirischen Nullpunktsmomente sind biasfrei. (Man nennt diese Eigenschaft seltener auch Erwartungstreue)

3.3.3.2 Erwartungswert der empirischen zentralen Momente

Die Bestimmung der Erwartungswerte der empirischen zentralen Momente erfolgt mit dem Ziel, Formeln für biasfreie empirische zentrale Momente aufzustellen. Die Ableitungen sind langwierig, jedoch nicht prinzipiell schwierig. Grundsätzlich werden die zentralen Momente aus den Nullpunktsmomenten gemäß Abschnitt 3.3.2 berechnet, da sich letztere fortlaufend mit den Realisierungen des beobachteten Prozesses bilden lassen.

Ebenso wie die zentralen sind auch die empirischen Momente

$$m_k = \frac{1}{n} \sum_{i=1}^n (X_i - M)^k$$

unabhängig vom Skalennullpunkt. Wählt man ihn wieder so, daß gilt

$$E \{X_i\} = \mu = 0$$

für alle $i=1\dots n$

und sind die X_i, X_j, X_k paarweise unkorreliert, dann gilt auch

$$E \{X_i \cdot X_j \cdot X_k \cdot \dots\} = 0,$$

sobald mindestens ein Index verschieden von allen anderen ist.

(Es ist übrigens keine Unabhängigkeit der X_i, X_j usw. erforderlich, sondern nur die Gültigkeit der schwächeren Forderung

$$E \{X_i \cdot X_j\} = E \{X_i\} \cdot E \{X_j\} \quad \forall i, j)$$

Für die Berechnung der Erwartungswerte der $m_2 \dots m_6$ gehen wir von den in Abschnitt 3.3.2 abgeleiteten Formeln (5) aus:

Wir benötigen die Erwartungswerte der Ausdrücke

$$a_k \cdot M^l \quad \text{mit} \quad \begin{array}{l} k = 1, 2 \dots 5 \\ l = 1, 2 \dots 5 \end{array}$$

Da $a_0 = 1$

und $a_1 = M$ ist, läßt sich

z.B. M^6 durch a_1 und M^5 ausdrücken:

$$M^6 = a_1 M^5$$

während für M^5 die Möglichkeiten bestehen:

$$M^5 = a_0 M^5$$

oder $M^5 = a \cdot M^4$

3.3.3.3 Ableitung der Erwartungswerte $E \{a_k M^{l-1}\}$ mit $l = 2 \dots 6$

Zunächst verwandelt man in

$$a_k \cdot M^{l-1} = \frac{1}{n^l} \sum_{i=1}^n x_i^k \left(\sum_{i=1}^n x_i \right)^{l-1}$$

die Produkte von Summen in Mehrfachsummen

$$a_k M^{l-1} = \frac{1}{n^l} \sum_{i_1=1}^n \sum_{i_2=1}^n \dots \sum_{i_l=1}^n x_{i_1}^k x_{i_2} \dots x_{i_l}$$

Die Zahl der sich so ergebenden Summanden ist n^l (Variation von n Elementen zur l -ten Klasse mit Wiederholung).

Für den Erwartungswert eines einzelnen Summanden

$$E \left\{ x_{i_1}^k x_{i_2} \dots x_{i_l} \right\}$$

ist es wichtig zu wissen, welche Beziehungen zwischen den $i_1, i_2 \dots i_l$ hinsichtlich der Gleichheit oder Verschiedenheit bestehen, da nach Abschnitt 3.3.3.2 ein Summand, in dem ein Index von allen anderen abweicht, den Erwartungswert Null hat und bei der Betrachtung des Erwartungswertes der Summe außer Betracht bleiben kann. Der erste Index nimmt jedoch insfern eine Sonderstellung ein, als er zum Faktor $x_{i_1}^k$ gehört. Für $k \neq 1$ braucht $E \{x_{i_1}^k\}$ nicht Null zu sein, vielmehr ist allgemein

$$E \left\{ x_i^k \right\} = \mu_k$$

Der erste Index kann also einzeln vorkommen.

Recht übersichtlich ist es, sich die Beziehungen zwischen den Indices an einem 1-Eck zu verdeutlichen. Für $l = 3$ kommt z.B.:

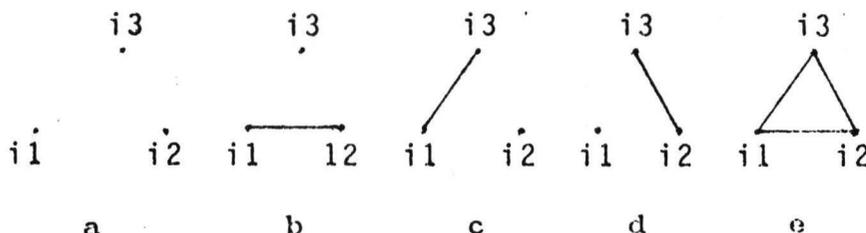


Bild 4 Kombinatorisches Schema der Beziehungen zwischen drei Indices.

Der Darstellung im 1-Eck entspricht die Zerlegung der Zahl 1 in ganzzahlige Summanden

$$\begin{aligned} 1 &= 1 + 1 + 1 \\ 1 &= 2 + 1 \\ 1 &= 3, \end{aligned}$$

wobei jeder Summand für eine Gruppe gleicher Indices steht. Es bedeutet

- 1 + 1 + 1 3 Gruppen von verschiedenen Indices mit je einem Glied (Bild 4 a))
- 2 + 1 2 Gruppen von verschiedenen Indices, von denen eine zwei, die andere ein Glied enthält. Dabei gibt es $\binom{3}{2} = 3$ Kombinationen. Sie sind in Bild 4 b) - 4 d) dargestellt. Gleiche Indices werden durch Strich verbunden gekennzeichnet.
- 3 1 Gruppe mit 3 gleichen Elementen (Bild 4 e)

Es ergeben sich also hier fünf Möglichkeiten der Zerlegung von $l = 3, 15$ für $l = 4$ und bereits 52 für $l = 5$. Wir verzichten daher auf die Ableitung einer allgemeinen Formel und begnügen uns mit der Ableitung von $E \{ a_k M^5 \}$ als größtem hier benötigten Ausdruck, die wir vorführen werden. Für kleinere l werden nur die Ergebnisse angegeben. Summanden, deren Erwartungswert 0 ist, werden ab sofort nicht mehr aufgeführt.

3.3.3.4 Zerlegung von l=6 und zugehörige Erwartungswerte $\neq 0$ der Summanden

$$\begin{aligned}
 6 & : E \left\{ X_i^k X_i X_i X_i X_i X_i \right\} = \mu_k + 5 \\
 4+2 & : E \left\{ X_i^k X_i X_i X_i X_j X_j \right\} = \mu_{k+3} \mu_2 \\
 3+3 & : E \left\{ X_i^k X_i X_i X_j X_j X_j \right\} = \mu_{k+2} \mu_3 \\
 2+2+2 & : E \left\{ X_i^k X_i X_j X_j X_m X_m \right\} = \mu_{k+1} \mu_2^2 \\
 1+5 & : E \left\{ X_i^k X_j X_j X_j X_j X_j \right\} = \mu_k \mu_5 \\
 1+2+3 & : E \left\{ X_i^k X_j X_j X_m X_m X_m \right\} = \mu_k \mu_2 \mu_3 \\
 1+3+2 & : E \left\{ X_i^k X_j X_j X_j X_m X_m \right\} = \mu_k \mu_3 \mu_2
 \end{aligned}$$

Alle verschiedenen Erwartungswerte $\neq 0$ des Produktes $X_i^k X_j X_m X_n X_p X_q$ der stochastischen Variablen X und alle zugehörigen nicht äquivalenten Indexanordnungen.

3.3.3.5 Multiplikatoren der Erwartungswerte aus 3.3.3.4

Der Faktor x_i^k trägt stets den Index i . Er kann jeden der n Werte annehmen. Alle Erwartungswerte kommen n-mal vor. Die folgenden Betrachtungen beziehen sich auf ein beliebiges, dann aber festgehaltenes i . Der Index i hat also einen Wert, die Indices j und m können n-1 Werte annehmen.

Es ist dann festzustellen:

1. Wieviel Kombinationen der Gruppen gleicher Indices i, j, m sich bei n-1 Elementen bilden lassen
2. Wieviel Permutationen mit Wiederholung bei fünf Faktoren möglich sind.

Kombinationen und Permutationen sind zu multiplizieren und bilden die Multiplikatoren für die Erwartungswerte der Summanden bei festgehaltenem i .

1.) Multiplikator für μ_{k+5}

$$E \left\{ X_i^k X_i X_i X_i X_i X_i \right\} = \mu_{k+5}$$

Alle Indices sind gleich i und fest.
Multiplikator für μ_{k+5} ist 1.

2.) Multiplikator für $\mu_{k+3} \mu_2$

$$E \left\{ X_i^k X_i X_i X_i X_i X_i \right\} = \mu_{k+3} \mu_2$$

Hier ist der Index j frei wählbar.

Kombinationen $C_1^{n-1} = \binom{n-1}{1} = (n-1)$

Unter den 5 gleichberechtigten und daher permutierbaren Indices sind 3 mal i und 2 mal j Permutationen $p_5^{3,2} = \frac{5!}{3!2!} = 10$

Multiplikator: $10(n-1)$

3.) Multiplikator für $\mu_{k+2} \mu_3$

$$E \left\{ X_i^k X_i X_i X_j X_j X_j \right\} = \mu_{k+2} \mu_3$$

Kombination $C_1^{n-1} = (n-1)$

Permutationen $p_5^{2,3} = 10$

Multiplikator $10(n-1)$

4.)

$$E \left\{ X_i^k X_i X_j X_j X_j X_j \right\} = \mu_{k+1} \mu_4$$

$$\text{Kombinationen } C_1^{n-1} = (n-1)$$

$$\text{Permutationen } P_5^{1,4} = 5$$

$$\text{Multiplikator : } 5(n-1)$$

$$5.) \quad E \left\{ X_i^k X_i X_j X_j X_m X_m \right\} = \mu_{k+1} \mu_2^2$$

frei wählbar j und m , also zwei Indices

$$\text{Kombinationen } C_2^{n-1} = \binom{n-1}{2} = \frac{(n-1)(n-2)}{2}$$

$$\text{Permutationen } P_5^{1,2,2} = \frac{5!}{1!2!2!} = 30$$

$$\text{Multiplikator: } 15 (n-1)(n-2)$$

$$6.) \quad E \left\{ X_i^k X_j X_j X_j X_j X_j \right\} = \mu_k \mu_5$$

$$\text{Kombinationen } C_1^{n-1} = (n-1)$$

$$\text{Permutationen } P_5^5 = 1$$

$$\text{Multiplikator : } (n-1)$$

$$7.) E \{ X_i X_j X_j X_m X_m X_m \} = \mu_k \mu_2 \mu_3$$

$$\text{Kombinationen } C_2^{n-1} = \frac{(n-1)(n-2)}{2}$$

$$\text{Permutationen } p_{5}^{2,3} = 10$$

$$\text{Multiplikator: } 5(n-1)(n-2)$$

$$8.) E \{ X_i^k X_j X_j X_j X_m X_m \} = \mu_k \mu_3 \mu_2$$

Dieser Summand ist analog dem Summanden unter 7) und erhält den gleichen Multiplikator. Der Ausdruck

$$\mu_k \mu_2 \mu_3 \text{ ist also mit } 10(n-1)(n-2)$$

zu multiplizieren. Trotzdem sind 7) und 8) verschiedene Fälle z.B. kann

$$X_1^k X_2 X_2 X_3 X_3 X_3 \text{ durch keine Permutation in}$$

$$X_1^k X_2 X_2 X_2 X_3 X_3 \text{ überführt werden.}$$

Das Ergebnis der Ableitung lautet also:

$$\begin{aligned} E \left\{ a_k M^5 \right\} = & \frac{1}{n^6} n \left[\mu_{k+5} + 10 (n-1) \mu_{k+3} \mu_2 \right. \\ & + 10(n-1) \mu_{k+2} \mu_3 + 5(n-1) \mu_{k+1} \mu_4 \\ & + 15(n-1) (n-2) \mu_{k+1} \mu_2^2 + (n-1) \mu_k \mu_5 \\ & \left. + 10(n-1) (n-2) \mu_k \mu_2 \mu_3 \right] \end{aligned}$$

Des weiteren sind:

$$\begin{aligned} E \left\{ a_k M^4 \right\} = & \frac{1}{n^4} \left[\mu_{k+4} + 6(n-1) \mu_{k+2} \mu_2 \right. \\ & + 4(n-1) \mu_{k+1} \mu_3 + (n-1) \mu_k \mu_4 + \\ & \left. + 3(n-1) (n-2) \mu_k \mu_2^2 \right] \end{aligned}$$

$$E \left\{ a_k M^3 \right\} = \frac{1}{n^3} \left[\mu_{k+3} + 3(n-1) \mu_{k+1} \mu_2 + (n-1) \mu_k \mu_3 \right]$$

$$E \left\{ a_k M^2 \right\} = \frac{1}{n^2} \left[\mu_{k+2} + (n-1) \mu_k \mu_2 \right]$$

$$E \left\{ a_k M \right\} = \frac{1}{n} \mu_{k+1}$$

(6)

Wir führen weiter die Berechnung des Erwartungswertes $E\{m_5\}$ vor und teilen für die übrigen empirischen zentralen Momente die Erwartungswerte mit.

3.3.3.6 Erwartungswert von m_5 :

Wie in Abschnitt 3.3.2 gezeigt, ist

$$m_5 = a_5 - 5a_4 M + 10a_3 M^2 - 10a_2 M^3 + 4M^5$$

$$E\{m_5\} = E\{a_5\} - 5E\{a_4 M\} + 10E\{a_3 M^2\} - 10E\{a_2 M^3\} + 4E\{M^5\}$$

mit $\mu_0=1$ und $\mu_1=0$ wird nach den vorstehenden Formeln

$$E\{a_5\} = \mu_5 \quad ; \quad E\{a_4 M\} = \frac{\mu_5}{n}$$

$$E\{a_3 M^2\} = \frac{1}{n^2} [\mu_5 + (n-1) \mu_2 \mu_3]$$

$$E\{a_2 M^3\} = \frac{1}{n^3} [\mu_5 + 4(n-1) \mu_2 \mu_3]$$

$$\begin{aligned} E\{M^5\} &= E\{a_1 M^4\} = \frac{1}{n^4} [\mu_5 + 6(n-1) \mu_3 \mu_2 + 4(n-1) \mu_2 \mu_3] \\ &= \frac{1}{n^4} [\mu_5 + 10(n-1) \mu_2 \mu_3] \end{aligned}$$

Die Formel für $E \{a_k M^5\}$ wurde hier nicht benötigt, obgleich man auch $M^5 = a_0 M^5$ schreiben und sie zur Kontrolle benutzen kann. Benötigt wird sie zur Berechnung von $E \{m_6\}$.

Nach kurzer Zwischenrechnung erhält man

$$E \{m_5\} = \frac{(n-1)(n-2)}{n^4} \left[(n^2 - 2n + 2) \mu_5 + 10(n-2) \mu_2 \mu_3 \right]$$

Die Ergebnisse für die Erwartungswerte aller zentralen Momente lauten:

$$E \{m_0\} = 1$$

$$E \{m_1\} = 0$$

$$E \{m_2\} = \frac{n-1}{n} \mu_2$$

$$E \{m_3\} = \frac{(n-1)(n-2)}{n^2} \mu_3$$

$$E \{m_4\} = \frac{n-1}{n^3} \left[(n^2 - 3n + 3) \mu_4 + 3(2n-3) \mu_2^2 \right]$$

$$E \{m_5\} = \frac{(n-1)(n-2)}{n^4} \left[(n^2 - 2n + 2) \mu_5 + 10(n-2) \mu_2 \mu_3 \right]$$

$$E \{m_6\} = \frac{(n-1)}{n^5} \left[(n^4 - 5n^3 + 10n^2 - 10n + 5) \mu_6 \right.$$

$$+ 15(n^3 - 4n^2 + 7n - 5) \mu_2 \mu_4 - 10(2n^2 - 6n + 5) \mu_3^2$$

$$\left. + 15(n-2)(3n-5) \mu_2^3 \right]$$

(7)

3.3.3.7 Deutung der Ergebnisse.

Die Erwartungswerte der empirischen zentralen Momente $m_2 \dots m_6$ stimmen nur für $n \rightarrow \infty$ mit den theoretischen zentralen Momenten $\mu_2 \dots \mu_6$ überein. Diese Eigenschaft nennt man nach R.A. Fischer Konsistenz der Schätzung.

Jedoch erhält man bei festgehaltener Stichprobengröße n systematische Fehler (m_2 und m_3 sind z.B. systematisch von zu kleinem Betrag). Man spricht hier vom Bias der Schätzung.

Man wird also die Schätzwerte für die Momente korrigieren, um den Bias zu entfernen und eine biasfreie konsistente Schätzung zu erhalten.

Nicht geprüft wurde, ob diese Schätzungen auch effizient d.h. von minimaler Varianz sind. Die an m_2 und m_3 anzubringenden Korrekturen bestehen lediglich darin, die Faktoren $\frac{n}{n-1}$ bzw. $\frac{n^2}{(n-1)(n-2)}$ vor die m_2 und m_3 zu setzen, bei den höheren Momenten gibt es jedoch mehrere Möglichkeiten, den Bias zu eliminieren. Die Vorgehensweise wird an der Erstellung der Formel für das fünfte Moment gezeigt. Die Formeln bis zum 4. Moment finden sich ohne Ableitung auch bei Cramér /12/ S. 352 und stimmen mit den selbst abgeleiteten überein, allerdings findet sich auch bei Cramer kein Hinweis auf die Effizienz. Dieser Frage konnte mit Rücksicht auf die Gesamtaufgabe nicht weiter nachgegangen werden.

3.3.3.8 Biasfreie Schätzwerte der empirischen zentralen Momente

Es war

$$E\{m_5\} = \frac{(n-1)(n-2)}{n^4} \left[(n^2 - 2n + 2) \mu_5 + 10(n-2) \mu_2 \mu_3 \right]$$

Es liegt nun nahe, das Glied $\mu_2 \mu_3$ durch ein solches, das aus dem Erwartungswert $E\{m_2 m_3\}$ gewonnen wurde, zu kompensieren.

$$m_2 m_3 = a_2 a_3 + 5a_2 M^3 - 3a_2^2 M - a_3 M^2 - 2M^5$$

$$E \{ a_2 a_3 \} = \frac{1}{n^2} E \left\{ \sum_i \sum_j x_i^2 x_j^3 \right\} = \frac{1}{n} \left[\mu_5 + (n-1) \mu_2 \mu_3 \right]$$

$$E \{ a_2^2 M \} = \frac{1}{n^3} E \left\{ \sum_i \sum_j \sum_k x_i^2 x_j^2 x_k \right\}$$

$$= \frac{1}{n^2} \left[\mu_5 + 2(n-1) \mu_2 \mu_3 \right]$$

Die Erwartungswerte der übrigen Summanden lassen sich aus den Formeln (6) ablesen.

Also wird:

$$E \{ m_2 m_3 \} = \frac{1}{n} \left[\mu_5 + (n-1) \mu_2 \mu_3 \right] + \frac{5}{n^3} \left[\mu_5 +$$

$$4(n-1) \mu_2 \mu_3 \right] - \frac{3}{n^2} \left[\mu_5 + 2(n-1) \mu_2 \mu_3 \right]$$

$$- \frac{1}{n^2} \left[\mu_5 + (n-1) \mu_2 \mu_3 \right] - \frac{2}{n^4} \left[\mu_5 + 10(n-1) \mu_2 \mu_3 \right]$$

$$E \{ m_2 m_3 \} = \frac{(n-1)(n-2)}{n^4} \left[(n-1) \mu_5 + (n^2 - 5n + 10) \mu_2 \mu_3 \right]$$

Mithilfe dieser Formel kann man nun den Term $\mu_2 \mu_3$ in $E \{ m_5 \}$ Formeln (7) durch $E \{ m_2 m_3 \}$ ausdrücken.

Nach elementarer Zwischenrechnung folgt:

$$\frac{n^2 \left[(n^2 - 5n + 10) E \{ m_5 \} - 10(n-2) E \{ m_2 m_3 \} \right]}{(n-1)(n-2)(n-3)(n-4)} = \mu_5$$

Daher hat der Ausdruck

$$M_5 = \frac{n^2 \left[(n^2 - 5n + 10) m_5 - 10(n-2) m_2 m_3 \right]}{(n-1)(n-2)(n-3)(n-4)}$$

den Erwartungswert μ_5 .

$$E \{ M_5 \} = \mu_5$$

Das biasfreie empirische zentrale Moment M_5 ist ein biasfreier Schätzwert für μ_5 .

3.3.3.9 Rechengang für den biasfreien Schätzwert für μ_6

Wie aus Formel (7) zu ersehen, enthält der Erwartungswert von m_6 außer μ_6 noch Ausdrücke mit μ_2 , μ_4 , μ_3^2 und μ_2^3 . Um sie zu eliminieren, benötigt man drei weitere Gleichungen, die man zweckmäßig durch Bilden der Erwartungswerte von m_2 , m_4 , m_3^2 und m_2^3 aufstellt. Diese Gleichungen sind ähnlich wie diejenige für den Erwartungswert $E \{ m_6 \}$ aufgebaut und erhalten ebenfalls nur Ausdrücke in μ_6 , μ_2 , μ_4 , μ_3^2 und μ_2^3 . Man erhält ein lineares Gleichungssystem, das man nach μ_6 auflöst, wodurch μ_2 , μ_4 , μ_3^2 und μ_2^3 eliminiert werden.

Der biasfreie Schätzwert für μ_6 setzt sich also aus den Erwartungswerten $E \{ m_6 \}$, $E \{ m_2 m_4 \}$, $E \{ m_3^2 \}$ und $E \{ m_2^3 \}$ multipliziert mit Ausdrücken aus Potenzen von n zusammen.

Die Erwartungswerte $E \{ m_2 m_4 \}$, $E \{ m_3^2 \}$ und $E \{ m_2^3 \}$ wurden von mir berechnet. Die letztgenannte Umformung zur Erstellung von M_6 wurde aus Aufwandsgründen nicht mehr durchgeführt. M_6 wird nur für den Symmetriefall mit $m_3=0$ und $m_5=0$ benötigt und ist daher praktisch entbehrlich, ggf. macht man keinen großen Fehler, wenn man mit m_6 anstelle von M_6 rechnet.

Es zeigt sich deutlich, daß abgesehen von den einfachsten Fällen, die Erstellung der Formeln für M_k sehr aufwendig wird. Von R.A. Fisher wurden die Kumulanten (Semi-Invarianten) eingeführt, für die sich entsprechende Berechnungen beträchtlich vereinfachen lassen. Als Literatur zitiert Cramér /12/ S. 352. Trotzdem erschien es

Wissenschaftszentrum
Berlin

nicht gerechtfertigt, für das anschließende Zerlegungsverfahren die Pearsonvariante von Charlier/13/ zu verwenden. Die Berechnung der biasfreien Schätzwerte für $\mu_2 \dots \mu_5$, der biasfreien empirischen zentralen Momente $M_2 \dots M_5$ konnte mit dem Formelmanipuliersystem Fordecad auf der IBM 360/67 bewältigt werden:

$$M_2 = \frac{n}{n-1} m_2 = \frac{1}{n-1} \sum (X_1 - M)^2 = s^2$$

$$M_3 = \frac{n^2}{(n-1)(n-2)} m_3$$

$$M_4 = \frac{n \left[(n^2 - 2n + 3) m_4 - 3(2n-3) m_2^2 \right]}{(n-1)(n-2)(n-3)}$$

$$M_5 = \frac{n^2 \left[(n^2 - 5n + 10) m_5 - 10(n-2) m_2 m_3 \right]}{(n-1)(n-2)(n-3)(n-4)} \quad (8)$$

4. Zerlegungsverfahren nach K. Pearson

Mit den theoretischen zentralen- oder den empirischen biasfreien Momenten stehen Ausgangsparameter für das in dieser Forschungsaufgabe entwickelte Wahrscheinlichkeitsdichteapproximationsverfahren zur Verfügung. Wichtigster Bestandteil dieses WD-Approximationsalgorithmus ist ein als "Zerlegung" bezeichnetes Verfahren von K. Pearson. Es soll kurz geschildert, und eine Reihe seiner Eigenschaften sollen untersucht werden.

Grundlegend für das gesamte W.D.-Approximationsverfahren ist die Möglichkeit, unter gewissen praktisch häufig gegebenen Bedingungen zu einer empirischen Verteilung mit den Momenten $M_2 \dots M_5$ eine Mischung von zwei normalen Komponenten mit gleichen Momenten $\mu_2 = M_2 \dots \mu_5 = M_5$ zu finden. Den Zusammenhang dieses Problems mit einem Wärmeleitungsproblem zeigt G. Doetsch /2/. Aus seiner Analogie erkennt man auch, daß die Bedingungen nicht immer erfüllt sein müssen.

K. Pearson hat die Aufgabe, aus $\mu_2 \dots \mu_5$ die Parameter der Komponenten zu berechnen, 1894 gelöst /5/. Sein Verfahren führt auf eine Gl. neunten Grades, deren sämtliche reelle Wurzeln er in Betracht zieht. Jede von ihnen kann auf zwei Komponenten führen. Wir bezeichnen in diesem Fall die Komponenten als Zerlegung der vorgegebenen empirischen Verteilung mit $M_2 \dots M_5$.

Es gibt entweder keine, eine oder mehrere Zerlegungen einer empirischen Verteilung, wie bereits K. Pearson feststellte. Die Mehrdeutigkeit kann im Zusammenhang gesehen werden mit der Mehrdeutigkeit der Lösung des modifizierten Momentenproblems nach Hamburger. (Schmetterer, 1966, S. 144 /14/).

In diesem Bericht wird diskutiert werden, welche Lösung im Falle der Mehrdeutigkeit zu wählen ist.

Der Vorschlag von K. Pearson, dazu das sechste Moment M_6 heranzuziehen, hat sich nicht bewährt - jedoch hatte Pearson selbst einschränkend hinzugefügt daß im Fall von zwei Lösungen, deren

sechste Momente oft wesentlich dichter benachbart sind als eines von ihnen mit dem sechsten Moment der empirischen Ausgangsverteilung.

1971 hat E.R. Berger den Bereich, in dem zur Lösung brauchbare Wurzeln der Pearsongleichung gesucht werden müssen, durch eine Ungleichung sehr eingengt /46/.

4.1 Beschreibung des Zerlegungsverfahrens

Es werden folgende Bezeichnungen eingeführt.

Z_1, Z_2	Mischungsfaktoren
γ_1, γ_2	Mittelwerte der Komponenten
σ_1, σ_2	Standardabweichungen der Komponenten.

$$\gamma_1 + \gamma_2 = p_1 \quad (9)$$

$$\gamma_1 \cdot \gamma_2 = p_2 \quad (10)$$

$$p_1 \cdot p_2 = p_3 \quad (11)$$

$$\lambda_4 = (9 \mu_2^2 - 3 \mu_4); \quad \lambda_5 = 10 \mu_2 \mu_3 - 3 \mu_5 \quad (12)$$

Der Skalennullpunkt wird so gewählt, daß der Mittelwert $\mu_1=0$ ist. Für $\mu_0 - \mu_6$ gelten dann in Fortführung der Gl'n (4)

$$\mu_0 = Z_1 + Z_2 = 1 \quad (13)$$

$$\mu_1 = Z_1 \gamma_1 + Z_2 \gamma_2 = 0 \quad (14)$$

$$\mu_2 = Z_1 (\gamma_1^2 + \sigma_1^2) + Z_2 (\gamma_2^2 + \sigma_2^2) \quad (15)$$

$$\mu_3 = Z_1 (\gamma_1^3 + 3\gamma_1 \sigma_1^2) + Z_2 (\gamma_2^3 + 3\gamma_2 \sigma_2^2) \quad (16)$$

$$\mu_4 = Z_1 (\gamma_1^4 + 6\gamma_1^2 \sigma_1^2 + 3\sigma_1^4) + Z_2 (\gamma_2^4 + 6\gamma_2^2 \sigma_2^2 + 3\sigma_2^4) \quad (17)$$

$$\mu_5 = Z_1 (\gamma_1^5 + 10\gamma_1^3 \sigma_1^2 + 15\gamma_1 \sigma_1^4) + Z_2 (\gamma_2^5 + 10\gamma_2^3 \sigma_2^2 + 15\gamma_2 \sigma_2^4) \quad (18)$$

$$\begin{aligned} \mu_6 = & Z_1 (\gamma_1^6 + 15\gamma_1^4 \sigma_1^2 + 45\gamma_1^2 \sigma_1^4 + 15\sigma_1^6) \\ & + Z_2 (\gamma_2^6 + 15\gamma_2^4 \sigma_2^2 + 45\gamma_2^2 \sigma_2^4 + 15\sigma_2^6) \quad (19) \end{aligned}$$

Aus den Gl'n (13) und (14) folgt:

$$\begin{aligned} Z_1 &= - \frac{\gamma_2}{\gamma_1 - \gamma_2} \\ Z_2 &= \frac{\gamma_1}{\gamma_1 - \gamma_2} \quad (20) \end{aligned}$$

Es soll gelten $0 \leq Z_1 \leq 1$; $0 \leq Z_2 \leq 1$

d.h. nur additive Mischungen werden betrachtet.

($Z_{1,2} < 0$ bedeuten "aussterben" einer Population oder dergl. und soll hier außer Betracht bleiben.)

Mit (14) folgt: γ_1 und γ_2 müssen verschiedenes Vorzeichen haben, $\gamma_2 \leq 0$, $\gamma_1 \geq 0$ (22). Dies wiederum bedeutet: $p_2 = \gamma_1 \cdot \gamma_2 < 0$. (23)

Da gleichzeitig γ_1 und γ_2 reell sind, muß dies auch für p_2 gelten.

Aus den Gl'n für die Momente der Mischung zweier Normalverteilungen Gl'n (13)-(18) stellt K. Pearson eine Gl. 9ten Grades in p_2 auf, die wir in Zukunft auch kurz die "Pearsongl." nennen wollen:

$$\begin{aligned}
 & 24 p_2^9 - 28 \lambda_4 p_2^7 + 36 \mu_3^2 p_2^6 - (24 \mu_3 \lambda_5 - 10 \lambda_4^2) p_2^5 \\
 & (148 \mu_3^2 \lambda_4 + 2 \lambda_5^2) p_2^4 + (288 \mu_3^4 - 12 \lambda_4 \lambda_5 \mu_3 - \lambda_4^3) p_2^3 \\
 & + (24 \mu_3^3 \lambda_5 - 7 \mu_3^2 \lambda_4^2) p_2^2 + 32 \mu_3^4 \lambda_4 p_2 - 24 \mu_3^6 = 0 \quad (21)
 \end{aligned}$$

Von dieser Gleichung interessieren nach (22) nur die negativen, reellen Wurzeln.

K. Pearson gibt weiterhin folgende Gl'n an:

$$p_3 = \frac{2 \mu_3^3 - 2 \mu_3 \lambda_4 p_2 - \lambda_5 p_2^2 - 8 \mu_3 p_2^3}{4 \mu_3^2 - \lambda_4 p_2 + 2 p_2^3}$$

und damit wegen (11) $p_1 = \frac{p_3}{p_2}$ (23)

Schließlich findet man γ_1 und γ_2 als Lösungen der quadratischen Gleichung

$$\gamma^2 - p_1 \gamma + p_2 = 0 \quad (24)$$

daraus ergeben sich mit (20) Z_1 und Z_2 , während σ_1^2 und σ_2^2 aus (15) und (16) abgeleitet werden:

$$\begin{aligned}
 \sigma_1^2 &= \mu_2 - \frac{1}{3} \frac{\mu_3}{\gamma_2} - \frac{1}{3} p_1 \gamma_1 + p_2 \\
 \sigma_2^2 &= \mu_2 - \frac{1}{3} \frac{\mu_3}{\gamma_1} - \frac{1}{3} p_1 \gamma_2 + p_2 \quad (25)
 \end{aligned}$$

4.2 Notwendige Bedingung dafür, daß zu p_2 eine Zerlegung gehört

K. Pearson hat alle reellen Wurzeln seiner Gl. neunten Grades auf die Möglichkeit einer Zerlegung untersucht. Im vorigen Abschnitt haben wir dann gezeigt, daß p_2 negativ und reell sein muß, damit die Zerlegung als Addition zweier Komponenten erscheint.

E.R. Berger hat 1971 eine untere Schranke für p_2 gefunden. p_2 muß notwendig oberhalb $-\mu_2$ liegen. Die Ableitung ist wie folgt: Die Pearsongl. (21) ist gedacht für unsymmetrische Mischverteilungen, dann wird aus (22) $\gamma_2 < 0 < \gamma_1$. Aus den Gl. (25)

wird durch Multiplizieren mit γ_2 bzw. γ_1 :

$$\begin{aligned} \gamma_2 \sigma_1^2 &= (\mu_2 + p_2) \gamma_2 - \frac{1}{3} (\mu_3 + p_3) \\ \gamma_1 \sigma_2^2 &= (\mu_2 + p_2) \gamma_1 - \frac{1}{3} (\mu_3 + p_3) \quad (26) \end{aligned}$$

σ_1^2 und σ_2^2 tatsächlicher Komponenten müssen pos. und reell sein, wegen (22) wird:

$$\gamma_2 \sigma_1^2 < 0, \quad \gamma_1 \sigma_2^2 > 0.$$

Da in beiden Gleichungen (26) $-\frac{1}{3} (\mu_3 + p_3)$ enthalten ist, gilt auch

$$(\mu_2 + p_2) \gamma_2 < (\mu_2 + p_2) \gamma_1$$

und da $\gamma_2 < \gamma_1$

notwendig

$$\mu_2 + p_2 > 0$$

$$p_2 > -\mu_2$$

$$\boxed{-\mu_2 < p_2 < 0} \quad (27)$$

Diese Bedingung ist äußerst nützlich, da von (21) nur die reellen Wurzeln im Bereich $-\mu_2$ bis 0 gesucht zu werden brauchen. Die Anzahl dieser Wurzeln kann mit Hilfe der Sturm'schen Kette bestimmt

werden. Obgleich auf dieses Verfahren noch eingegangen werden soll, sei hier bereits ein Ergebnis vorweggenommen:

1. Bei allen sehr zahlreich gerechneten Beispielen gab es stets zwei oder keine Wurzeln, die Bergers Bedingung (27) erfüllten.
2. Die Gl. von K. Pearson hat keine weiteren negativen Wurzeln.

Dabei wurde keine Voraussetzung über die Verteilungen gemacht, aus denen die Momente berechnet wurden. Um die Bedingungen 1. und 2. zu beweisen, könnte man daran denken, das Verfahren der Sturmschen Kette mit allgemeinen Zahlen durchzuführen. Nimmt man jedoch beliebige Zahlen anstelle von Momenten, dann gelten die obigen Feststellungen 1. u. 2. nicht. Es muß also den Beziehungen zwischen Momenten Rechnung getragen werden. Eine Beweisführung ist daher sehr aufwendig, wenn sie überhaupt gelingt. Die Bedeutung des Beweises rechtfertigt solchen Aufwand nicht.

4.3 Notwendige und hinreichende Bedingung dafür, daß zu p_2 eine Zerlegung gehört.

Die Bedingung (27) ist notwendig, jedoch nicht hinreichend, d.h. falls sich zwischen $-\mu_2$ und 0 zwei Wurzeln finden, können beide, eine oder auch keine zu einer Zerlegung führen. Im ersten Fall liegt dann eine zweideutige Lösung vor.

Daß die Bedingung (27) nicht hinreichend ist, liegt daran, daß sie erfüllt sein kann, ohne daß die Voraussetzung $\gamma_2 \sigma_1^2 < 0$, $\gamma_1 \sigma_2^2 > 0$, unter der sie abgeleitet wurde, erfüllt sein muß. Bezieht man die Gl'n (26) voll mit ein, dann wird

$$\begin{aligned} 3 \gamma_2 \sigma_1^2 &= 3 (\mu_2 + p_2) \gamma_2 - (\mu_3 + p_3) < 0 \\ 3 \gamma_1 \sigma_2^2 &= 3 (\mu_2 + p_2) \gamma_1 - (\mu_3 + p_3) > 0 \end{aligned} \quad (28)$$

also:

$$\boxed{3(\mu_2 + p_2)\gamma_2 < \mu_3 + p_3 < 3(\mu_2 + p_2)\gamma_1} \quad (29)$$

Diese Bedingung ist hinreichend für die Existenz einer zu p_2 gehörigen Zerlegung mit positiven Varianzen.

Sonderfall $\sigma_1 = \sigma_2 = \sigma \neq 0$

Wir gehen aus von Gl'n (25):

$$\sigma^2 = \sigma_1^2 = (\mu_2 + p_2) - \frac{\mu_3 + p_3}{3\gamma_2}$$

$$\sigma^2 = \sigma_2^2 = (\mu_2 + p_2) - \frac{\mu_3 + p_3}{3\gamma_1}$$

$$(\mu_3 + p_3)\gamma_1 = (\mu_3 + p_3)\gamma_2$$

Sind γ_1 und $\gamma_2 \neq 0$ (unsymmetrischer Fall), dann haben sie verschiedenes Vorzeichen, daher muß $\mu_3 + p_3 = 0$ sein; sind sie aber (beide) Null, so ist dies ebenfalls notwendig.

$$\mu_3 + p_3 = 0 \quad (30)$$

Nur in diesem Fall ist die Bedingung (27) zugleich notwendig und hinreichend.

Es wäre wünschenswert gewesen, bereits am Momentensatz $M_2 \dots M_5$ zu erkennen, ob Gl (21) zu Zerlegungen führt. Abgesehen von dem Fall, daß sich aus der Anwendung des Verfahrens der Sturm'schen Kette ergibt, daß in $(-\mu_2, 0)$ keine Wurzel vorliegt, ist dies nicht gelungen.

Es soll nun versucht werden zu zeigen, wie es dazu kommt, daß eine Wurzel nicht zu einer Zerlegung führt. Daraus ergibt sich ein Zugang zum Problem der zwei Lösungen einer Zerlegungsaufgabe und eine heuristische Regel für die Auswahl einer der Lösungen. Schließlich läßt sich in diesem Zusammenhang beweisen, daß im interessierenden unsymmetrischen Fall $p_2 = 0$ und $p_2 = -\mu_2$ selbst keine Lösungen der Pearsongl. (21) sein können, womit eine wichtige Voraussetzung für die Anwendung des Verfahrens der Sturmschen Kette im Intervall $(-\mu_2, 0)$ gegeben ist. Wir beginnen mit dem letzten Punkt.

4.4 Ausschließen von $p_2 = 0$ und $p_2 = -\mu_2$ als Wurzeln von Gleichung (21):

Im folgenden sei $F(x)$ eine unsymmetrische Mischverteilung mit $M_2 \neq 0, M_3 \neq 0, M_5 \neq 0$ bzw. $\mu_2 \neq 0, \mu_3 \neq 0, \mu_5 \neq 0$. Die Pearsongleichung (21) ist für diesen Fall vorgesehen. Die Lösung des Zerlegungsproblems im Symmetriefall wird im Anhang behandelt. In den Formeln von K. Pearson werden die Momente $\mu_2 \dots \mu_5$ bei der praktischen Zerlegung durch $M_2 \dots M_5$ ersetzt. $M_3 = 0$ bzw. $\mu_3 = 0$ hieße, in Gl. (21) verschwindet das absolute Glied: $-24\mu_3^6$.

Damit gäbe es mindestens eine Wurzel $p_2 = 0$.

Indem wir den Symmetriefall ausschließen, fordern wir:

$$p_2 \neq 0,$$

wegen (10) gilt dann auch

$$\gamma_1 \neq 0, \gamma_2 \neq 0 \quad (31)$$

Wäre nun $p_2 = -\mu_2$ dann käme aus Gl.(29)

$$0 < \mu_3 + p_3 < 0 \quad (32)$$

$0 < 0$ ist ein Widerspruch, also muß gelten:

$$p_2 \neq -\mu_2 \quad (33)$$

Würden die $<$ Zeichen in (31) durch \leq ersetzt, dann stände anstelle von (32) :

$$0 \leq \mu_3 + p_3 \leq 0$$

Das wäre nur möglich, wenn:

$$\mu_3 + p_3 = 0 \quad \text{wäre.}$$

Gleichungen (28) würden dann

$$3 \gamma_2 \sigma_1^2 = 3 \cdot 0 \cdot \gamma_2 - 0 = 0$$

$$3 \gamma_1 \sigma_2^2 = 3 \cdot 0 \cdot \gamma_1 - 0 = 0 \quad (34)$$

Wir setzen voraus

$$\sigma_1^2 \neq 0, \quad \sigma_2^2 \neq 0 \quad (35)$$

da es sinnlos wäre, mit Komponenten verschwindender Varianz zu arbeiten;

$$\gamma_1 \neq 0, \quad \gamma_2 \neq 0 \quad \text{galt wegen (31).}$$

Damit entsteht in den Gl'n (34) ein Widerspruch.

Ergebnis: Es gibt keine Wurzeln $p_2 = 0$ und $p_2 = -\mu_2$ der Pearsongl. (21), die zu Zerlegungen führen können. Die Werte 0 und $-\mu_2$ können als Grenzen im Verfahren nach Sturm verwendet werden.

4.5 Numerische Betrachtung des Auftretens der Doppellösung.

Eine Doppellösung tritt auf, wenn beide Wurzeln p_2 im Intervall $(-\mu_2, 0)$ und die zugehörigen Werte p_1, γ_1, γ_2 usw. die hinreichende Bedingung (29) erfüllen. Geht man bei der Simulation von Momenten aus der Mischung zweier Normalverteilungen gemäß (15) bis (18) aus, so zeigt sich, daß (29) dann von beiden Wurzeln p_2 befriedigt wird, wenn der Abstand der Mittelwerte γ_1 und γ_2 kleiner oder wenigstens nicht viel größer ist als die kleinere der Standardabweichungen σ_1, σ_2 . Um diese Angabe zu präzisieren, sollte man die Pearsongl. mit normierten Werten z.B. mit

$$\frac{\gamma_1 - \gamma_2}{\sigma_1 \sigma_2} ; \quad \gamma_2 < \gamma_1 \quad \text{ableiten}$$

Das war im Rahmen der Gesamtaufgabe zeitlich nicht möglich. Wir werden uns mit einem Zahlenbeispiel begnügen, an dem sich alles Wesentliche zeigen läßt.

Wir wollen folgende Bezeichnungen einführen:

Bei dem Zahlenbeispiel gehen wir von einer Mischung aus, deren theoretische zentrale Momente gemäß (15) - (18) berechnet werden. Wir nennen diese Mischung im weiteren einfach "Ausgangsverteilung". Bei der Lösung der Zerlegungsaufgabe wird stets als ein Ergebnis die Ausgangsverteilung zurückgewonnen. Diese Lösung wollen wir als Hauptlösung bezeichnen; daneben gibt es im Falle der Zweideutigkeit eine "Nebenlösung", das ist eine Lösung, die in den Momenten $\mu_2 \dots \mu_5$ mit der Hauptlösung übereinstimmt, nicht jedoch in allen höheren Momenten.

Es sei nun eine zweideutige Zerlegung gegeben. Wie das Experiment zeigt, vollzieht sich der Übergang zur eindeutigen Zerlegung so, daß mit kleiner werdendem σ_1 und σ_2 gegenüber $\gamma_1 - \gamma_2$ eine Komponente der Nebenlösung immer schmaler und höher wird, zur Spitze entartet und schließlich eine negative Varianz bekommt, womit die gesamte Nebenlösung nicht mehr als Zerlegung gelten kann, die Zerlegung ist also eindeutig geworden.

Solange die Nebenlösung existiert, gilt Gl. (29) auch für ihre Werte, die durch ' gekennzeichnet seien:

$$3(\mu_2 + p_2') \gamma_2' < \mu_3 + p_3' < 3(\mu_2 + p_2') \gamma_1'$$

Mit σ_1 und σ_2 der Ausgangsverteilung verkleinert man μ_2 , ohne daß sich die anderen Werte stark ändern. Dadurch verkleinert sich $\mu_2 + p_2'$ und die Grenzen $3(\mu_2 + p_2') \gamma_2'$ und $3(\mu_2 + p_2') \gamma_1'$ rücken näher aneinander. Da $\mu_3 + p_3' \neq 0$ (außer wenn $\sigma_1 = \sigma_2$), kommt der Punkt $\mu_3 + p_3'$ schließlich außerhalb der soeben genannten Grenzen zu liegen. Die hinreichende Bedingung (29) ist verletzt (obgleich die notwendige (27) - $\mu_2 < p_2 < 0$ noch gilt), die Nebenlösung ist verschwunden.

Wir betrachten ein Zahlenbeispiel und wählen als Ausgangsverteilung die Mischung:

Mischungsfaktoren	Mittelwerte	Standardabweichungen
$z_1 = 0,7$	$my_1 = -2$	$\sigma_1 = 2,099$
$z_2 = 0,3$	$my_2 = +2$	$\sigma_2 = 2,099$

Die Mittelwerte wurden mit my bezeichnet, um Verwechslungen mit empirischen M oder theoretischen μ Momenten oder mit auf Mittelwert der Mischung bezogenen Mittelwerten γ zu vermeiden. Der Fall $\sigma_1 = \sigma_2 = 0$ wurde gewählt, weil er sich einfacher rechnet und alles erforderliche zeigt. Für die Nebenlösung ist natürlich $\sigma_1' \neq \sigma_2'$.

Wir finden mit (14) als reduzierte Mittelwerte der Ausgangsverteilung

$$\gamma_2 = -1,2 \quad ; \quad \gamma_1 = 2,8$$

und daraus mit (9) (10) (11)

$$p_1 = \gamma_1 + \gamma_2 = 1,6$$

$$p_2 = \gamma_1 \gamma_2 = -3,36$$

$$p_3 = p_1 p_2 = -5,376$$

Für μ_2 und μ_3 findet man durch Einsetzen von (9), (10) und (11) in (15) und (16)

$$\mu_2 = -p_2 + \frac{\gamma_1 \sigma_2^2 - \gamma_2 \sigma_1^2}{\sigma_2^2 - \sigma_1^2} \quad (36)$$

$$\mu_3 = -p_3 + 3 p_2 \frac{\sigma_2^2 - \sigma_1^2}{\gamma_1 - \gamma_2} \quad (37)$$

hieraus wird für $\sigma_1 = \sigma_2$

$$u_2 = -p_2 + \sigma^2 = 7,765801$$

$$u_3 = -p_3 = 5,376$$

(Gl. (37) enthält Gl. (30) als Sonderfall)

Löst man Persons Gl. (21) auf, so gewinnt man zur Hauptlösung gehörig $p_2 = -3,36$ zurück.

Daneben findet man

$$p_2' = -1,07213676115$$

dazu

$$p_3' = 2,5064946424$$

$$p_1' = 2,3378497349$$

$$\gamma_2 = -2,7305016689$$

$$\gamma_1 = 0,3926519340$$

$$3(\mu_2 + p_2') \gamma_2' = -54,8311841259$$

$$\mu_3 + p_3' = 7,8824946424$$

$$3(\mu_2 + p_2') \gamma_1' = 7,8898406267$$

Die Bedingung (29) lautet

$$-54,83 \dots < 7,8824946424 < 7,8848406268$$

sie ist noch gültig, obgleich erkennbar wird, daß rechts fast das Gleichheitszeichen gilt.

Die zu diesem Zahlenbeispiel gehörigen Summenverteilungen sind im Bild 5 zu sehen. Die Nebenlösung trägt die Spitze.

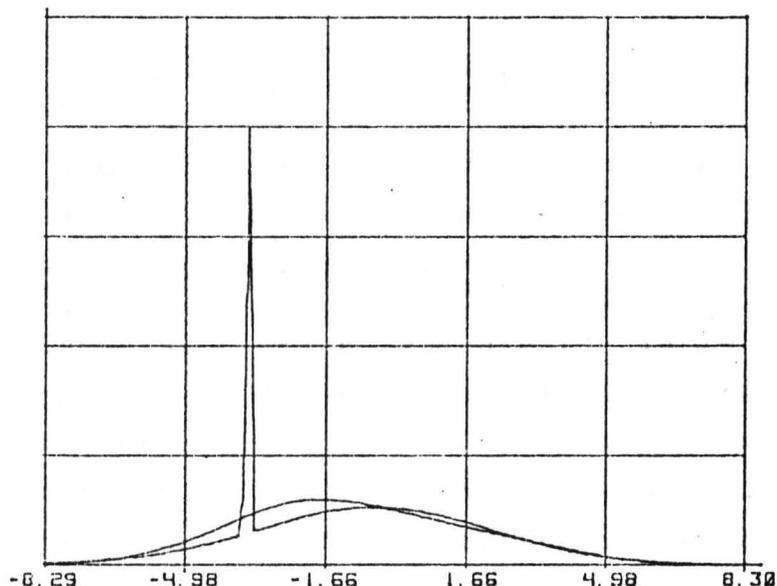


Bild 5 Haupt- und Nebenlösung einer zweideutigen Zerlegung

Die beiden Zerlegungen und die daraus reproduzierten Momente sind: (Erklärung der Parameter S. 46)

```

LSG. -.107213676115E 01
A = .125722901890E 00   CMY= *.353050166894E 01   CSI= .446270390596E-01
1-A= .874277098112E 00   DMY= *.407348065986E 00   DSI= .276693710947E 01
AM1= *.363797880709E-11   AM2= .776580099997E 01   AM3= .537599999961E 01
AM4= .166945395510E 03   AM5= .286745141745E 03   AM6= .629509018657E 04
    
```

```

LSG. -.336000000010E 01
A = .699999999906E 00   CMY= *.200000000029E 01   CSI= .209899999990E 01
1-A= .300000000096E 00   DMY= .199999999937E 01   DSI= .209900000016E 01
AM1= -.109139364213E-10   AM2= .776580099994E 01   AM3= .537599999935E 01
AM4= .166945395510E 03   AM5= .286745141737E 03   AM6= .567898347637E 04
    
```

Die Momente $\mu_2 \dots \mu_5$ stimmen bei beiden Lösungen überein. In den höheren Momenten weichen die Lösungen voneinander ab.

Als zweites Beispiel geben wir die Parameter von folgenden normalen Komponenten ein:

$$\begin{array}{lll}
 z_1 = 0,3 & my_1 = -2 & \sigma_1 = 6 \\
 z_2 = 0,7 & my_2 = +2 & \sigma_2 = 6
 \end{array}$$

Wegen $\sigma_1 = \sigma_2$ muß gelten $\mu_3 = -p_3$. Da die Mittelwerte mit denjenigen vom Beispiel 1 übereinstimmen, während die Mischungsfaktoren lediglich gegenüber Beispiel 1 vertauscht sind, was einer Umkehrung der x-Richtung entspricht und zur Vorzeichenumkehr der ungeraden Momente führt, müssen $p_2 = \gamma_1 \gamma_2$ und

$|p_3| = |(\gamma_1 \gamma_2)(\gamma_1 + \gamma_2)|$ in Beispiel 1 und 2 übereinstimmen.

Dies gilt mindestens für die Werte p_2 und p_3 , die zu den eingegebenen Komponenten gehören. Abweichungen in μ_3 sind auf Rundungsfehler zurückzuführen.

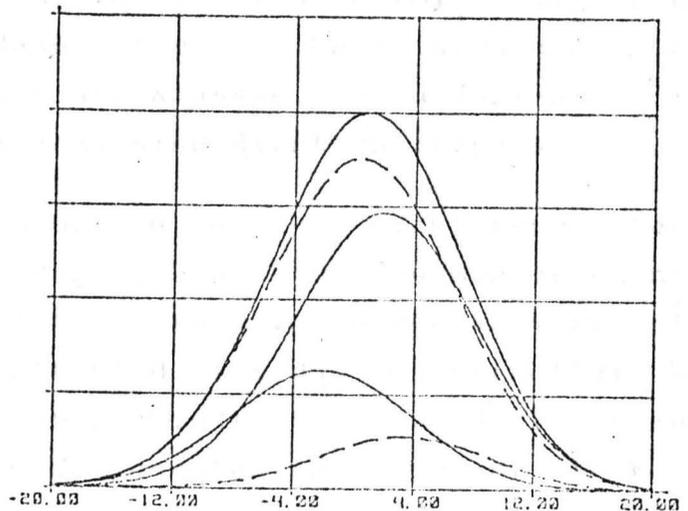
Interessanterweise stimmen jedoch beide Wurzeln p_2 und p_2' und bis auf die letzten Stellen auch alle Koeffizienten der Pearson-Gleichung in beiden Beispielen überein. Die Abweichungen werden voraussichtlich auf Rechenungenauigkeiten zurückzuführen sein. Die Übereinstimmung der Pearsongleichung bei bis auf μ_3 voneinander abweichenden Momentensätzen läßt sich sicher leicht exakt beweisen, jedoch besteht dafür keine Notwendigkeit, da die Ausgangsvoraussetzung $\sigma_1 = \sigma_2$ praktisch nie erfüllt ist und es hier nur darauf ankam, zu zeigen, daß bei festgehaltenem p_2, p_3 (und μ_3) sowie der γ_1, γ_2 Existenz oder Verschwinden der Nebenlösung nur von der Größe von μ_2 abhing. Für $\sigma_1 \neq \sigma_2$ wird diese Aussage nur näherungsweise gelten.

Die Nebenlösung hat die gleichen Mittelwerte und gegenüber Beispiel 1 zahlenmäßig gleiche (nur vertauschte) Mischungsfaktoren. Nur die Standardabweichungen sind gegenüber Beispiel 1 verändert. Beide Lösungen unterscheiden sich im σ_1/σ_2 -Verhältnis ihrer Komponenten nicht mehr stark. Wie Bild 6 zeigt, decken sich die Summenkurven aus beiden Zerlegungen.

Zweideutige Zerlegung
mit deckungsgleichen
Summenkurven.

— Hauptlösung
- - Nebenlösung

Bild 6



Hauptlösung

```
LSG. -.336000000028E 01
K,EG
A   . .299399999719E 00   CMY. . .200000000196E 01   CSI. .599999999980E 01
1-A . .700000000277E 00   DMY. .199999999925E 01   DSI. .600000000006E 01

AM1. .727595761418E-11   AM2. .393599999996E 02   AM3. .537600000389E 01
AM4. .463365119994E 04   AM5. .198524928117E 04   AM6. .906683105263E 06
```

Nebenlösung

```
LSG. -.107213676104E 01
A   . .874277098112E 00   CMY. .407348056023E 00   CSI. .626499321370E 01
1-A . .123722901390E 00   DMY. .353050166876E 01   DSI. .562104885282E 01

AM1. .363797280709E-11   AM2. .393599999996E 02   AM3. .537600000249E 01
AM4. .463365119994E 04   AM5. .198524928084E 04   AM6. .907299211967E 06
```

Parameter zu Bild 6

A, 1 - A	Mischungsfaktoren
CMY,DMY	Mittelwerte
CSI,DSI	Standardabweichungen
AM1...AM 6	Zentrale Momente
LSG	Wurzel der Pearsongleichung.

Dies ist ein sehr praktisches Ergebnis:

Liegt schon eine Doppellösung vor - ist also $(my_1 - my_2)$ nicht groß gegen σ_1 und σ_2 - dann ist dieser Fall für die Zeichen-erkennung ohnehin uninteressant. Klassen, deren Varianzen nicht klein sind gegen ihren Abstand, sind nicht zu trennen.

Ist dann zusätzlich σ_1/σ_2 in beiden Lösungen nicht sehr unterschiedlich, dann unterscheiden sich auch die Summenkurven nicht und es ist für die Approximation einer vorgegebenen Kurve gleichgültig, welche Lösung gewählt wird. Ist das σ -Verhältnis bei-der Lösungen jedoch stark unterschiedlich, so wird diejenige mit zwei sehr unterschiedlichen Komponenten in Bezug auf die Varianz die Nebenlösung sein, die instabil ist und deshalb verworfen werden sollte.

Schließlich sei als drittes Beispiel ein Fall gewählt, bei dem μ_2 gegen Fall 1 verkleinert wird, die übrigen Werte stimmen überein.

$$\begin{array}{lll} z_1 = 0,7 & my_1 = -2 & \sigma_1 = 2,098 \\ z_2 = 0,3 & my_2 = +2 & \sigma_2 = 2,098 \end{array}$$

LSS. ..107213676111E 01
KEINE ZERLEGUNG MÖGLICH

LSS. ..335999999994E 01

LOESUNGSKOMPONENTEN

BA[1]:	.700000	BMV[1]:	-2.000000	BSIGM[1]:	2.098000
BA[2]:	.300000	BMV[2]:	2.000000	BSIGM[2]:	2.098000

MY1:	.000000000000E 00	MY2:	.776160400000E 01	MY3:	.537599999991E 01
MY4:	.166749889957E 03	MY5:	.286519511033E 03	MY6:	.566847958371E 04

Die Standardabweichungen sind also um 0,5 ‰ verkleinert und desgleichen $\mu_2 = -p_2 + \sigma^2$.

Wieder stimmen die Koeffizienten der Pearsongleichung und damit auch p_2 und p_2' in jetzt allen drei Beispielen überein

$$\sigma_2'^2 = \mu_2 + p_2' - \frac{1}{3} \left(\frac{u_3 + p_3'}{\gamma_1'} \right)$$

enthält als einzige gegen Beispiel 1 veränderte Größen μ_2 . Ihre Verkleinerung um 0,5 ‰ ist dafür verantwortlich, daß die Nebenlösung verschwindet.

4.6 Schlußfolgerung für die praktische Anwendung

Ginge es darum, die "richtige" der beiden in Bild 5 dargestellten Lösung zu wählen und hätte man z.B. außer $\mu_2 - \mu_5$ noch μ_6 zur Verfügung, so könnte man durch Vergleich von μ_6 diese Aufgabe lösen.

Die empirischen Momente $M_2 \dots M_6$ sind hingegen mit Zufallsschwankungen behaftet. Da schon kleinste Schwankungen bei μ_2 (z.B. 0,5 ‰) die Lösung mit der Spitze zum Verschwinden bringen,

ist es besser, eine solche Lösung als instabil zu betrachten und zugunsten der gegen Schwankungen unempfindlichen Lösung zu verwerfen.

Liegen zwei Lösungen vor, so wird also diejenige verworfen, für die das Verhältnis zwischen den Varianzen ihrer Komponenten den größeren Wert annimmt, unabhängig davon, ob sie das sechste Moment besser reproduziert als die andere Lösung und ob sie zweigipflig ist und andere nicht.

Daß dieses Entscheidungskriterium zweckmäßig ist, hat sich bei Approximationsversuchen an vorgegebenen empirischen Verteilungen durch eine Mischung zahlreicher Normalverteilungen gezeigt.

Wählt man bei den hin und wieder auftretenden zweideutigen Lösungen jedoch z.B. diejenige, deren sechstes Moment das empirische sechste besser reproduziert, so approximiert der Algorithmus schlechter. Es wurden Spitzenbildungen beobachtet.

5. Signifikanztest für die Abweichung von der Normalität als Entscheidungskriterium für die Zerlegung

Der Gedanke, eine empirisch gegebene Klasse dann in zwei zu zerlegen, wenn die Verteilung der in ihr enthaltenen Daten signifikant von der Form der Normalverteilung abweicht, ist grundlegend für das vorliegende WD-Approximationsverfahren. Dabei bleibt es ohne prinzipielle Bedeutung, daß eine Zerlegung in Normalverteilungen vorgenommen wird und sich der Signifikanztest daher auch auf diesen Typ bezieht, ebensogut hätte ein anderer Verteilungstyp verwendet werden können.

Beim vorliegenden WD-Approximationsverfahren wird jede Klasse durch einen Satz von fünf zentralen Momenten repräsentiert. Dieser Momentensatz muß notwendig dem Signifikanztest zugrunde liegen.

5.1 Formfaktoren

In einer Stichprobe vom Umfang n

seien

$$m_k = \frac{1}{n} \sum_{i=1}^n (X_i - M)^k \quad \text{mit} \quad M = \frac{1}{n} \sum_{i=1}^n X_i$$

die biasbehafteten zentralen Stichprobenmomente und

$$e_1 = \frac{m_3}{m_2^{3/2}} \quad ; \quad e_2 = \frac{m_4}{m_2^2} - 3 \quad (38)$$

die empirischen Koeffizienten der Schiefe und des Exzesses der Stichprobe.

Näherungsweise gelten für beliebige Verteilungen die Erwartungswerte

$$E \{e_1\} = \frac{\mu_3}{\sigma^3} = \xi_1$$

$$E \{e_2\} = \frac{\mu_4}{\sigma^4} - 3 = \xi_2$$

und die Varianzen:

$$\text{Var} \{e_1\} =$$

$$(4 \mu_2^2 \mu_6 - 12 \mu_2 \mu_3 \mu_5 - 24 \mu_2^3 \mu_4 + 9 \mu_3^2 \mu_4 + 35 \mu_2^2 \mu_3^2 + 36 \mu_2^5) / 4 \mu_2^5 n$$

$$\text{Var} \{e_2\} =$$

$$(\mu_2^2 \mu_8 - 4 \mu_2 \mu_4 \mu_6 - 8 \mu_2^2 \mu_3 \mu_5 + 4 \mu_4^3 - \mu_2^2 \mu_4^2 + 16 \mu_2 \mu_3^2 \mu_4 + 16 \mu_2^3 \mu_3^2) / \mu_2^6 n$$

/177 S. 257

Bei Normalverteilungen ergibt das:

$$E \{e_1\} = 0 ; \quad E \{e_2\} = 0$$

$$\text{Var} \{e_1\} = \frac{6}{n} ; \quad \text{Var} \{e_2\} = \frac{24}{n} \quad (39)$$

Genauer kommt für die Normalverteilung

$$E \{e_1\} = 0 ; \quad E \{e_2\} = -\frac{6}{n+1}$$

$$\text{Var} \{e_1\} = \frac{6(n-2)}{(n+1)(n+3)} ; \quad \text{Var} \{e_2\} = \frac{24 n(n-2)(n-3)}{(n+1)^2(n+3)(n+5)}$$

/177 S. 387

Führt man die biasfreien Schätzwerte der Kumulanten ein

$$K_2 = \frac{n}{n-1} m_2$$

$$K_3 = \frac{n^2}{(n-1)(n-2)} m_3$$

$$K_4 = \frac{n^2}{(n-1)(n-2)(n-3)} \left[(n+1)m_4 - 3(n-1)m_2^2 \right],$$

dann kann man damit biasfreie empirische Formfaktoren E_1 und E_2

$$E_1 = \frac{K_3}{K_2^{3/2}} = \frac{\sqrt{n(n-1)}}{n-2} e_1$$

$$E_2 = \frac{K_4}{K_2^2} = \frac{n-1}{(n-2)(n-3)} \left[(n+1)e_2 + 6 \right]$$

ausschreiben.

Setzt man darin für e_1 und e_2 aus (38) ein und ersetzt dann wieder für m_2, m_3, m_4 aus (8) durch M_2, M_3, M_4 , so erhält man nach Zwischenrechnung:

$$E_1 = \frac{M_3}{M_2^{3/2}} \quad (40)$$

$$E_2 = \frac{n}{n^2 - 2n + 3} \left[(n+1) \frac{M_4}{M_2^2} - 3(n-1) \right] \quad (41)$$

Diese biasfreien empirischen Formfaktoren haben die Erwartungswerte 0:

$$E \{E_1\} = 0$$

$$E \{E_2\} = 0$$

und die Varianzen

$$\text{Var} \{E_1\} = \frac{6n(n-1)}{(n-2)(n+1)(n+3)} \quad (42)$$

$$\text{Var} \{E_2\} = \frac{24n(n-1)^2}{(n-2)(n-3)(n+3)(n+5)} \quad (43)$$

Für große n gehen (42) und (43) über in die asymptotischen Formeln (39).

5.2 Test auf Normalität

Beide Formafaktoren E_1 und E_2 sind Zufallsvariable und für nicht zu kleine Stichproben asymptotisch normal verteilt. Man kann also unter Benutzung von $\text{Var} \{E_1\}$ bzw. $\text{Var} \{E_2\}$ ein Intervall um $E \{E_1\} = 0$ bzw. $E \{E_2\} = 0$ konstruieren, außerhalb dessen das aus dem zu prüfenden Momentensatz berechnete E_1 oder E_2 nur mit einer Wahrscheinlichkeit (Irrtumswahrscheinlichkeit, Signifikanzniveau) zu liegen kommt. In weiteren Ausführungen wird nur E_1 erwähnt, obgleich der Formfaktor E_2 in gleicher Weise für einen Test verwendet wird.

Man stellt die Nullhypothese H_0 auf: " E_1 weicht nur zufällig von $E \{E_1\}$ ab", die man zugunsten der Alternativhypothese H_1 : " E_1 weicht systematisch von $E \{E_1\}$ ab" verwirft, wenn das nach Formel (40) aus der Messung berechnete E_1 außerhalb der Schranken $\pm S$ zu liegen kommt. Die Schranke S berechnet man unter Verwendung von (42) und von Tabellen des Fehlerintegrals:

$$\frac{\alpha}{2} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-t} e^{-\frac{1}{2}\tau^2} d\tau \quad (44)$$

$$\bar{-} S = \bar{+} t \sqrt{\text{Var} \{E_1\}} \quad (45)$$

Praktisch verläuft der Test so:

Wahl eines α z.B. $\alpha = 0,1$

Berechnung des zugehörigen t nach Tabelle zu (44) .

Untersuchung einer Stichprobe vom Umfang n .

Berechnung von $\text{Var} \{E_1\}$ nach (42) .

Berechnung von S nach (45) .

Berechnung von E_1 nach (40) .

Annahme oder Verwerfen der Nullhypothese .

Die Alternativhypothese H_1 :

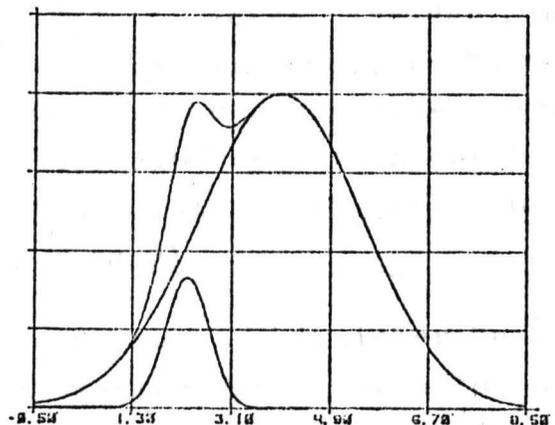
" E_1 weicht signifikant von $E \{E_1\}$ ab" deutet man so : "Die durch den Momentensatz repräsentierte Komponente ist nicht normalverteilt und sollte also zerlegt oder nicht wieder vereinigt werden."

Da $S = S(\alpha)$ ist und man in der Wahl des Signifikanzniveaus frei ist, liegt in der Entscheidung zugunsten H_0 oder H_1 , wie bei jedem Signifikanztest, ein gewisses Maß an Willkür. Es ist Sache des Gesamtalgorithmus, empirisch zu ermitteln, welches Signifikanzniveau zweckmäßig ist.

6. Nebenbedingungen

Zweigipflige Verteilung
ohne Bayesgrenze

Bild 7



EINGEGEBENE WERTE

(GAUSS-VERTEILUNG)

AA[1]•	•900000	AMY[1]•	4.000000	ASIGM[1]•	1.500000
AA[2]•	•100000	AMY[2]•	2.200000	ASIGM[2]•	•400000

Das Kriterium für die Steuerung des Zerlegens bzw. Zusammenlegens von Klassen ist die signifikante Abweichung eines Formfaktors von seinem Erwartungswert.

Unter Umständen ist es zweckmäßig, daneben weitere Forderungen zu stellen. Faßt man z.B. als Anwendung des Verfahrens die Zeichenerkennung ins Auge und betrachtet eine durch ihre empirischen Momente gekennzeichnete Komponente nur als Repräsentanten für ein Merkmal (Klasse), dann erscheint es wünschenswert, trotz signifikanter Abweichung des Formfaktors von seinem Erwartungswert Null eine Zerlegung der die Klasse repräsentierenden Komponente in zwei nur dann vorzunehmen, wenn deren Summenkurve zweigipflig ist.

Die Zweigipfligkeit tritt als Nebenbedingung auf.

Am praktischen Beispiel wird auf Vor- und Nachteile dieser Nebenbedingung einzugehen sein. Dichtekurven aus wenig Beobachtungen enthalten starke Schwankungen. Der Algorithmus könnte aufgrund eines Einbruches in der empirischen Verteilung einer Klasse, der sich in den Momenten und als signifikante Abweichung des Formfaktors von Null niederschlägt, zu einer Zerlegung gelangen, deren eine Komponente als kleine Spitze auf der Flanke der zweiten Komponente reitet. Auf diese Weise wäre der Einbruch in der empirischen Verteilung approximiert, die Summenkurve wäre sogar zweigipflig; dennoch kann die Dichte einer solchen kleinen Spitze ständig unterhalb derjenigen der Hauptkomponente liegen. Eine solche Klasse ist für die Zeichenerkennung uninteressant. (Vergl. Bild 7)

Es muß jedoch zugegeben werden, daß trotz zahlreicher Versuche mit normalverteilten Pseudozufallszahlen diese Situation nicht zu simulieren war. Dies spricht für die Stabilität des Zerlegungsverfahrens, wie später ausgeführt wird.

Man entscheidet sich z.B. beim Bayesklassifikator für eine Klasse, deren (bedingte) Wahrscheinlichkeit unter der Bedingung, daß eine Beobachtung X gemacht wurde, am größten ist.

Damit man sich nicht nur für eine Klasse entscheidet, müssen Grenzen existieren, an denen sich die Dichtekurven der Klassen schneiden. Man nennt sie Bayesgrenzen. Im Fall, daß eine Kurve unterhalb der anderen verläuft, existieren keine Bayesgrenzen. Die Existenz von Bayesgrenzen ist eine weitere interessante Nebenbedingung.

6.1. Zweigipfligkeit.

Es wäre äußerst praktisch, könnte man aus dem Momentensatz $M_2 \dots M_5$ bzw. $\mu_2 \dots \mu_5$ bereits erkennen, ob nach einer Zerlegung aufgrund dieser Momente aus den Komponenten eine ein- oder zweigipflige Dichtekurve als Summe resultiert. Nach Maßgabe der Nebenbedingung könnte man die Zerlegung dann ggf. unterlassen. Eine solche Bedingung kann es jedoch nicht geben, denn zu einem Momentensatz können zwei Zerlegungen gehören, von denen eine zweigipflig und die andere eingipflig sein kann.

Daß es auch bei Kenntnis der Parameter nicht trivial ist zu entscheiden, ob eine Mischung von zwei normalen Verteilungen eine ein- oder zweigipflige Dichtekurve hat, zeigt eine Reihe von Veröffentlichungen.

Zu nennen ist eine Betrachtung von G. Doetsch /2/ 1936, S. 314-315 an Komponenten mit gleichen Varianzen. Die Bedingung für die Extremwerte führt auf eine transzendente Gleichung. Ihre Lösung läßt sich nach Logarithmieren als Schnittpunktaufgabe einer logarithmischen Funktion mit einer Geraden auffassen:

Schnittpunkte
 der logarith-
 mischen Funktion
 $f_1(x)$ mit den zu
 Geraden entar-
 teten Parabeln
 $f_2(x)$; f_{21} , f_{22} ,
 f_{23} für drei
 Sätze von Ein-
 gangsparametern.
 (Erklärung von
 Z_2, R, D durch
 Gleichungen
 (46), (53) und
 (52) S. 58)

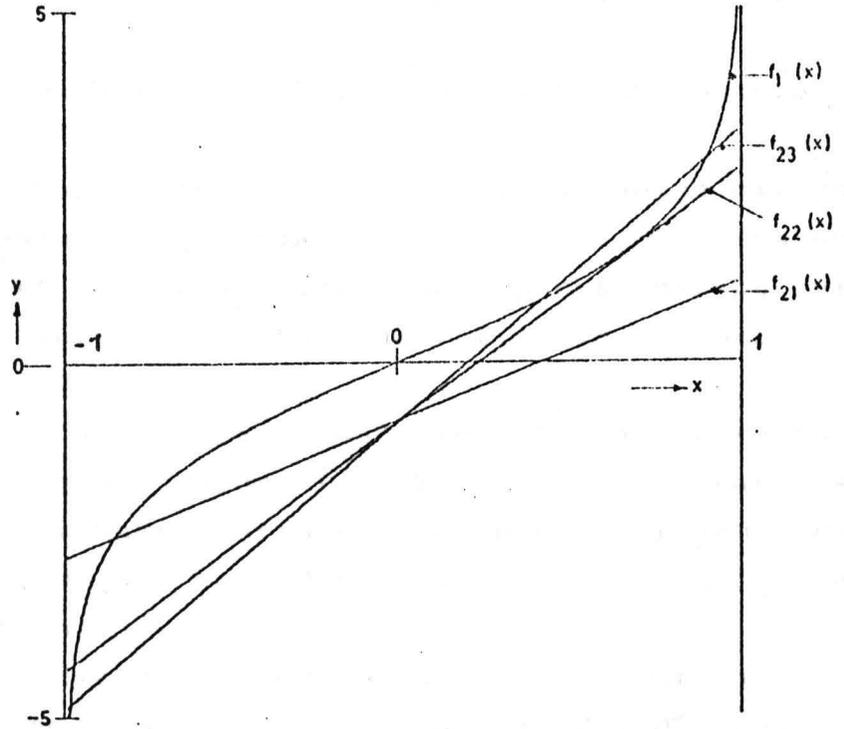


Bild 8

$$f_1(x) = \ln \frac{1+x}{1-x} \quad (64)$$

$$f_2(x) = \frac{d^2}{2r} \left((1+x)^2 - r^2(1-x)^2 \right) + \ln \frac{r^3 z_2}{1-z_2} \quad (65)$$

z2 = .3
R = 1.
D = 1.
F21 = 2. X - .84729786

z2 = .3
R = 1.
D = 1.35
F22 = 3.645 X - .84729786

z2 = .3
R = 1.
D = 1.45
F23 = 4.205 X - .84729786

Ergibt sich ein Schnittpunkt, so ist die zugehörige Mischverteilung eingipflig, bei drei Schnittpunkten ist sie zweigipflig.

Die weiteren Möglichkeiten sind bei G. Doetsch l.c. besprochen.

Es lag nahe, diese Betrachtung auf Mischverteilungen mit unterschiedlichen Varianzen der Komponenten zu erweitern. Dies wurde von mir durchgeführt und ergab eine Schnittpunktaufgabe zwischen o.g. logar. Kurve und einer quadratischen Parabel.

Nachträglich stellte sich heraus, daß diese Aufgabe bereits von Behbodian /16/ gelöst worden ist. Er gibt ein iteratives Verfahren zur Schnittpunktsbestimmung an, das in diesem Bericht nicht interessiert, da nur nach der Zahl nicht nach der Lage der Gipfel gefragt ist.

Behbodian reduziert die Parameter der Mischung, nämlich Mittelwerte γ_1, γ_2 , Standardabweichungen σ_1, σ_2 und Mischungsfaktoren z_1, z_2 auf drei wesentliche.

Halber Abstand der Mittelwerte bezogen auf das geometrische Mittel der Standardabweichungen:

$$d = \frac{|\gamma_1 - \gamma_2|}{2 \sqrt{\sigma_1 \sigma_2}}, \text{ Verhältnis der Standardabweichungen } r = \frac{\sigma_1}{\sigma_2} \text{ mit } \sigma_1 \leq \sigma_2 \text{ und Mischungsfaktoren } (1-z_2), z_2$$

Seine Variablentransformation kann man durch Einführen einer auf d bezogenen Variablen noch verbessern. Dadurch ergibt sich eine noch zweckmäßigere Darstellung der Schnittpunktskurven. Die logarithmische Kurve verändert im Gegensatz zu Behbodian nicht mehr Maßstab und Lage. Außerdem wird die Ableitung der Endgleichung einfacher, sodaß man eine lückenhafte Darstellung und einen in Behbodians Gleichung (15) enthaltenen Fehler vermeiden kann.

6.1.1 Ableitung der Gleichungen für die logarithmische Funktion und Parabel.

$$\text{Es sei } h(x) = z_1 h_1(x) + z_2 h_2(x) \quad (46)$$

$$\text{mit } z_1 + z_2 = 1$$

$$\text{und } n_1(x) = \frac{1}{\sqrt{2\pi} \sigma_1} e^{-\frac{1}{2} \left(\frac{x - \gamma_1}{\sigma_1} \right)^2} \quad (47)$$

$$h_2(x) = \frac{1}{\sqrt{2\pi} \sigma_2} e^{-\frac{1}{2} \left(\frac{x - \gamma_2}{\sigma_2} \right)^2}$$

Ohne Beschränkung der Allgemeinheit sei:

$$\sigma_1 \leq \sigma_2 \quad (48)$$

Weiterhin gelte:

$$\varepsilon = \text{sign}(\gamma_2 - \gamma_1) \quad (49)$$

$$\bar{\gamma} = (\gamma_1 + \gamma_2) / 2 \quad (50)$$

$$\bar{\sigma} = \sqrt{\sigma_1 \sigma_2} \quad (51)$$

$$d = \varepsilon (\gamma_2 - \gamma_1) / 2 \bar{\sigma} \quad (52)$$

$$r = \sigma_1 / \sigma_2 \quad (53)$$

Nun werde die Transformation eingeführt:

$$y = \varepsilon (x - \bar{\gamma}) / \bar{\sigma} \quad (54)$$

mit (50)

$$\frac{\gamma_1}{2} + \frac{\gamma_2}{2} = \bar{\gamma}$$

und (52)

$$-\frac{\gamma_1}{2} + \frac{\gamma_2}{2} = \frac{d\bar{\sigma}}{\epsilon}$$

$$\gamma_1 = \bar{y} - \frac{d\bar{\sigma}}{\epsilon} \quad (55)$$

$$\gamma_2 = \bar{y} + \frac{d\bar{\sigma}}{\epsilon}$$

aus (54)

$$x = \bar{y} + \frac{\bar{\sigma}}{\epsilon} y$$

und aus (55) für \bar{y}

$$x = \frac{\bar{\sigma}}{\epsilon} (y + d) + \gamma_1 \quad \text{oder}$$

$$x = \frac{\bar{\sigma}}{\epsilon} (y - d) + \gamma_2 \quad (56)$$

Die Exponenten in (47) werden:

$$-\frac{1}{2} \left(\frac{x - \gamma_1}{\sigma_1} \right)^2 = -\frac{1}{2\sigma_1^2} \left[\frac{\bar{\sigma}}{\epsilon} (y + d) + \mu_1 - \mu_1 \right]^2$$

mit (49) $\epsilon^2 = 1$ und (51), (53) $\frac{\sigma^2}{\sigma_1^2} = \frac{1}{r}$:

$$-\frac{1}{2} \left(\frac{x - \gamma_1}{\sigma_1} \right)^2 = -\frac{1}{2r} (y + d)^2$$

$$-\frac{1}{2} \left(\frac{x - \gamma_2}{\sigma_2} \right)^2 = -\frac{1}{2} r (y - d)^2 \quad (57)$$

weiter:

$$\begin{aligned}
 h_1(x) &= h_1(x(y)) = h_1(y) = \frac{1}{\sigma \sqrt{2\pi r}} e^{-\frac{1}{2r}(y+d)^2} = \frac{1}{\sigma} g_1(y) \\
 h_2(x) &= h_2(y) = \frac{1}{\sigma \sqrt{2\pi/r}} e^{-\frac{1}{2r}(y-d)^2} = \frac{1}{\sigma} g_2(y)
 \end{aligned} \tag{58}$$

Durch Einführen der neuen Variablen

$$z = \frac{y}{d} \tag{60}$$

wird aus (58)

$$\begin{aligned}
 h_1(z) &= \frac{1}{\sigma \sqrt{2\pi r}} e^{-\frac{d^2}{2r}(1+z)^2} \\
 h_2(z) &= \frac{1}{\sigma \sqrt{2\pi/r}} e^{-\frac{d^2 r}{2}(1-z)^2}
 \end{aligned} \tag{61}$$

Führt man (61) in (46) ein, so gewinnt man eine Gleichung $h(x) = h(x(z)) = h(z)$. Leitet man sie nach z ab und setzt $dh/dz = 0$, so entsteht eine Bestimmungsgleichung für die Abszissen der Extremwerte z_e . Es sei jedoch gestattet anstelle von z_e weiterhin z zu schreiben. Nach wenigen elementaren Umformungen findet man:

$$\frac{d^2}{2r} \left[(1+z)^2 - r^2(1-z)^2 \right] + \ln \frac{z_2 r^3}{1-z_2} = \ln \frac{1+z}{1-z} \tag{62}$$

Aus (56) und (60) gewinnt man:

$$x = \frac{\sigma_1 d}{\epsilon \sqrt{r}} (z+1) + y_1 = \frac{\sigma_2 d \sqrt{r}}{\epsilon} (z-1) + y_2, \quad (63)$$

womit man in z gewonnene Schnittpunktwerte auf x umrechnen kann.

Im folgenden werden wir anstelle von z wieder x schreiben.

Auch die Festlegung zwischen y und x (54) lassen wir fallen.

Wir zerlegen (62) in

$$y_s = f_2(x) = f_1(x) \quad \text{Schnittpunktgleichung}$$

worin

$$y = f_1(x) = \ln \frac{1+x}{1-x} \quad (64)$$

$$y = f_2(x) = \frac{d^2}{2r} \left[(1+x)^2 - r^2(1-x)^2 \right] + \ln \frac{r^3 z_2}{1-z_2} \quad (65)$$

6.1.2 Deutung der Gleichungen

$$f_1(x) = \ln \frac{1+x}{1-x} \quad (64) \quad \text{existiert nur f\u00fcr}$$

$-1 < x < +1$, weil f\u00fcr alle anderen Werte ein Logarithmus aus einer negativen Zahl entst\u00fcnde.

Der Verlauf ist auf Bild 9-14 zu sehen. Die Funktion hat Pole bei $x_{p1} = -1$ und $x_{p2} = 1$ und eine Nullstelle bei $x_0 = 0$.

Ihre Steigung betr\u00e4gt :

$$f_1'(x) = \frac{2}{1-x^2} \quad \text{und ist im gesamten Bereich } -1 < x < +1$$

positiv.

$x_w = 0$ ist ein Wendepunkt.

$$f_2(x) = \frac{d^2}{2r} \left[(1+x)^2 - r^2(1-x)^2 \right] + \ln \frac{r^3 z_2}{1-z_2} \quad (65)$$

ist die Gleichung einer Parabel.

Nach einigen Umformungen kommt:

$$f_2(x) = \frac{d^2(1-r^2)}{2r} \left(x + \frac{1+r^2}{1-r^2} \right)^2 - \frac{2rd^2}{1-r^2} + \ln \frac{r^3 z_2}{1-z_2}$$

Für die Parameter z_2 , d , r gelten

$$0 < z_2 < 1$$

$$0 < r \leq 1$$

$$0 < d$$

(66)

Der Scheitelpunkt der Parabel hat die Koordinaten:

$$x_s(r) = - \frac{1+r^2}{1-r^2} \quad (67)$$

Läßt man r im Bereich (66) variieren, so findet man:

$$x_s(1) = -\infty \quad (\text{Vergl. Bild 13}),$$

die Parabel ist in eine Gerade übergegangen (Fall von Doetsch). Vergl. Bild 8

Da $r > 0$ gilt auch $x_s < -1$.

Die Parabel hat ihren Scheitel also stets links von -1 .

$$\text{Da } \frac{d^2(1-r^2)}{2r} > 0 \text{ für } 0 < d \text{ und } 0 < r < 1$$

ist die Parabel nach oben geöffnet.

Die Scheitelpunktsordinate $f_2(x_s) = -\frac{2rd^2}{1-r^2} + \ln \frac{r^3 z_2}{1-z_2}$

wird nur für sehr kleine z_2 (Anteil der Komponente mit größerer Standardabweichung) positiv.

Die Parabel hat mindestens einen und höchstens drei Schnittpunkte mit der logarithmischen Funktion im Bereich $-1 < x < 1$. Einem Schnittpunkt entspricht eine eingipflige, drei Schnittpunkten die zweigipflige Summenkurve.

6.1.3 Ersatzfunktion für $f_1(x)$.

Um die Zahl der Schnittpunkte abzuschätzen schlägt E.R. Berger vor, die transzendente Funktion (64) durch eine solche gebrochen rationale zu ersetzen, deren Steigung überall größer oder gleich derjenigen von $f_1(x)$ ist. /47/ Wenn diese Funktion $f_3(x)$ drei Schnittpunkte mit der Parabel hat, dann hat $f_1(x)$ sicher auch drei mit der Parabel, d.h., wenn wir eine mit Sicherheit zweigipflige Mischverteilung suchen, dann genügt es, eine Lösung von $f_3(x) = f_2(x)$ mit drei Wurzeln in $-1 < x < 1$ zu finden.

Berger schlägt als Ersatzfunktion vor:

$$f_3(x) = \frac{2x}{1-x^2} \quad (68)$$

$$f'_3(x) = \frac{2}{1-x^2} \cdot \frac{1+x^2}{1-x^2}$$

$$\frac{1+x^2}{1-x^2} \geq 1 \quad \text{für } -1 < x < 1$$

also

$$f'_3(x) \geq f'_2(x) \quad -1 < x < 1$$

Dies ist zugleich die einfachste gebrochen rationale Funktion mit Polen bei -1 und $+1$ einer Nullstelle bei 0 .

6.1.4 Zahl der Schnittpunkte zwischen -1 und +1 .

Wir bringen zunächst eine Reihe von Beispielen (Bild 9-14).
Eingegeben wurden die Daten z_2 , d und r .
Geplottet wurden die Funktionen $f_1(x)$, $f_2(x)$ und $f_3(x)$
zwischen -1 und +1.

In Bild 9 wurden $f_2(x)$ und $f_3(x)$ darüberhinaus im Bereich
-2 bis +2 eingezeichnet.

Auf die weiteren Angaben zu jedem Bild wird im folgenden ein-
gegangen.

Die Eingabedaten z_2 , r und d für Bild 9 stammen aus einer
Simulation mit 100 Zufallszahlen, bei der die obige Zerle-
gung entstand. Für die übrigen Beispiele wurden die Eingabe-
daten geeignet gewählt.

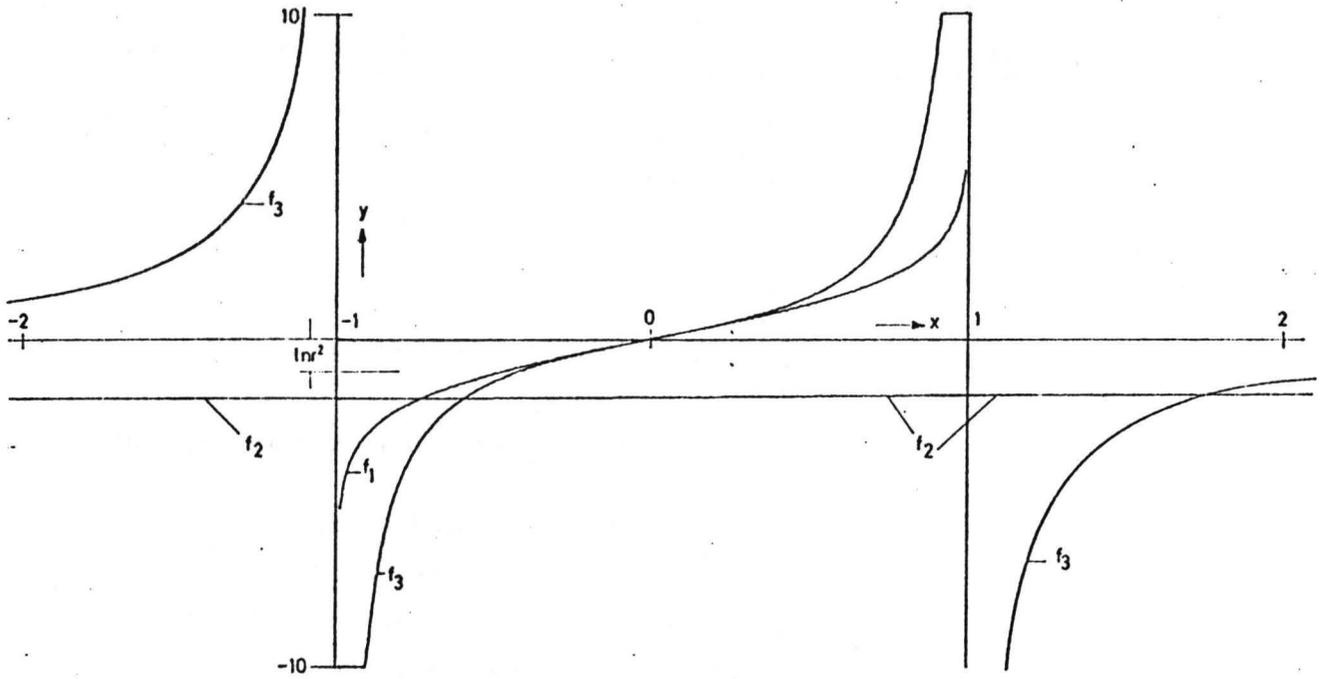


Bild 9

Eingipflige Mischung mit Bayesgrenzen

Logarithmische Funktion f_1 Gl. (64)

Ersatzfunktion f_3 Gl. (68) und Parabel f_2 Gl. (65)

$$f_1 = \ln \frac{1+x}{1-x} \quad f_3 = \frac{2x}{1-x^2} \cdot 2$$

f_2 siehe Rechnerausdruck.

```

J = .1622225      K = .62332579      Z2 = .394505
-----
F2 = .05662244 X + .01290769 X2 - 1.83356467
-----
F = 2.42022139 X - 2.30170769 X2 + .07308176 X3 + .01609139 X4 + 2.28581629
-----
K(-1) = 3          A = - 1          I = 5.20954063          H = - .94536871
-----
K(1) = 2          B = - 21.8678103       J = - 11.7596136
-----
ANZHL = 1         C = .38853628          T = 15.8233857
-----

```

Rechnerausdruck zu Bild 9

Erklärungen zum Rechnerausdruck

Z2, R, D nach Gl.n (46), (53), (52)

F Gl. (69)

K(-1) Zahl der Zeichenwechsel bei -1

K(1) Zahl der Zeichenwechsel bei +1

ANZHL Zahl der Schnittpunkte der Ersatzfunktion f_3 mit der Parabel f_2 zwischen -1 und +1
Siehe S. 76

A,B,C Bestandteile der Ungleichung (75). $A < B < C$ muß erfüllt sein, damit f_2 und f_3 keinen Schnittpunkt rechts von +1 haben, s. S. 81 .

T Gl. (76) Diskriminante

I,J Bestandteile der Diskriminante nach der Formel von Cayley-Boole, s. S. 82.

Wenn $T > 0$, haben f_2 und f_3 vier Schnittpunkte, sind zusätzlich $A < B < C$, so liegen drei davon zwischen -1 und +1, dann ist die Summe beider Gaußkomponenten sicher zweigipflig.

$H = \ln r^2$ Abstand der Parallelen zur x Achse. An den Schnittpunkten dieser Geraden mit der Parabel f_2 sind die Bayesgrenzen. Siehe auch S. 84 Gl. (79).

Gerade noch zwei-
 gipflige Mischung
 mit Bayesgrenzen.
 3 Schnittpunkte von
 f_2 mit f_1 , jedoch
 nur einer mit f_3
 in $(-1, +1)$.
 Schnittpunkt von f_2
 mit Gerade im Ab-
 stand $\ln r^2$ für
 Bayesgrenze.
 Vergl. Rechneraus-
 drucke.
 Erklärungen wie bei
 Bild 9

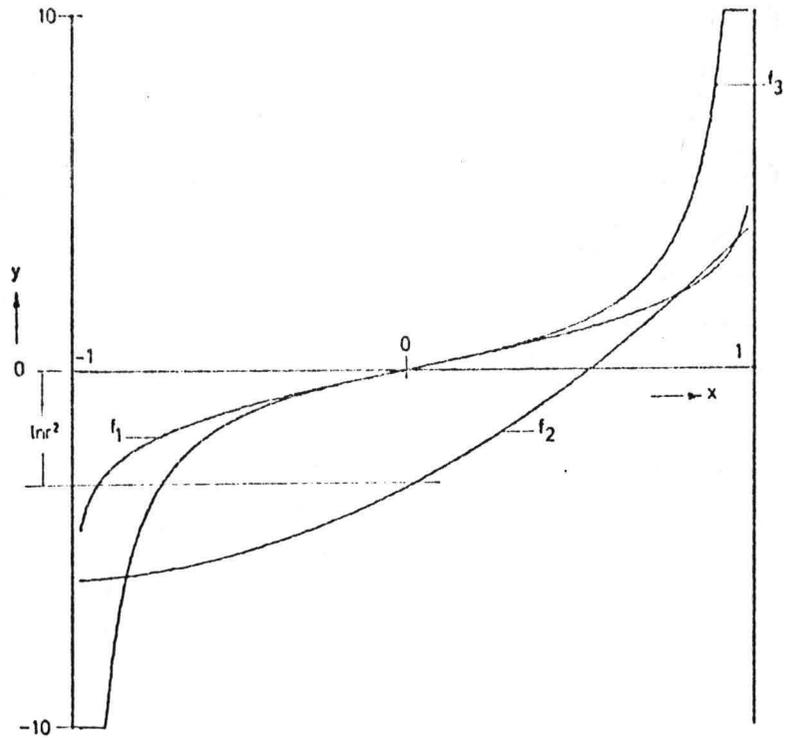


Bild 10

$J = .99$ $R = .2$ $ZZ = .3$

$$F_2 = 5.00052 X + 2.35224 X^2 - 3.32337159$$

$$F = - 1.238608 X - 2.27024463 X^2 + 2.038608 X^3 + .940896 X^4 + 1.32934863$$

$K(-1) = 2$

$A = - 1$

$I = 27.758435$

$II = - 3.21887582$

$K(1) = 1$

$B = - .57908495$

$J = 317.508468$

$ANZHL = 1$

$C = .04$

$T = - 573.044115$

Eingipflige Mischung
mit Bayesgrenzen.
Erklärungen zum
Rechnerausdruck
wie in Bild 9.

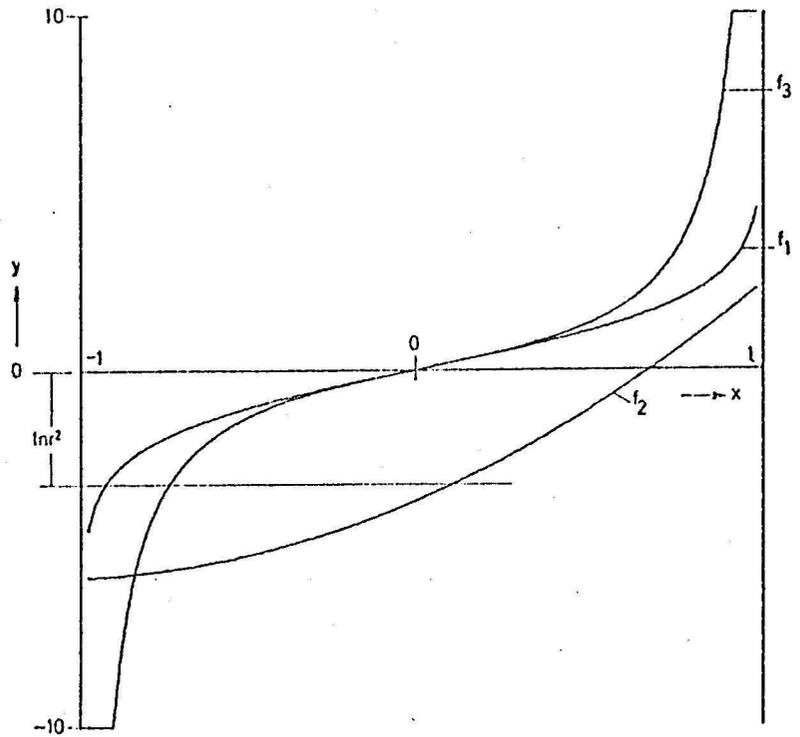


Bild 11

$\theta = .9 \quad \kappa = .2 \quad Z_2 = .3$

$F_2 = 4.212 X + 1.944 X^2 - 3.73161159$

$F = -.8343 X - 2.27024463 X^2 + 1.6348 X^3 + .7775 X^4 + 1.49264463$

$K(-1) = 2 \quad I = 23.5543034$

$K(1) = 1 \quad J = 266.695951$

$ANZHL = 1 \quad T = -693.314405$

Zweigipflige
Mischung mit
Bayesgrenzen.
Erklärungen zum
Rechnerausdruck
wie in Bild 9.

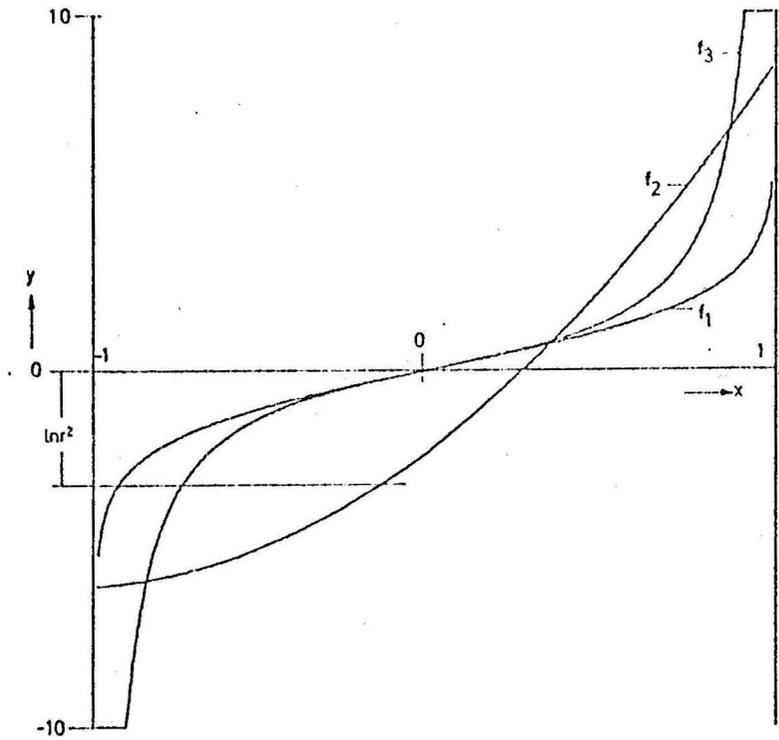


Bild 12

$D = 1.2$ $R = .2$ $Z2 = .3$

$F2 = 7.488 X + 3.456 X^2 - 2.21961159$

$F = -2.1952 X - 2.27024463 X^2 + 2.9952 X^3 + 1.3824 X^4 + .82724463$

$K(-1) = 3$

$I = 39.6074769$

$K(1) = 0$

$J = 437.797012$

$ANZHL = 3$

$T = 2196.33511$

Zweigipflige
Mischung ohne Bayes-
grenzen.

Diagramm zum Bei-
spiel Bild 7

3 Schnittpunkte

von f_1 und f_2 ,
daher zweigipflig.

Kein Schnittpunkt
der Parabel mit
Gerade im Abstand
 $\ln r^2$ daher keine
Bayesgrenze.

Erklärungen zum
Rechnerausdruck
wie in Bild 9.

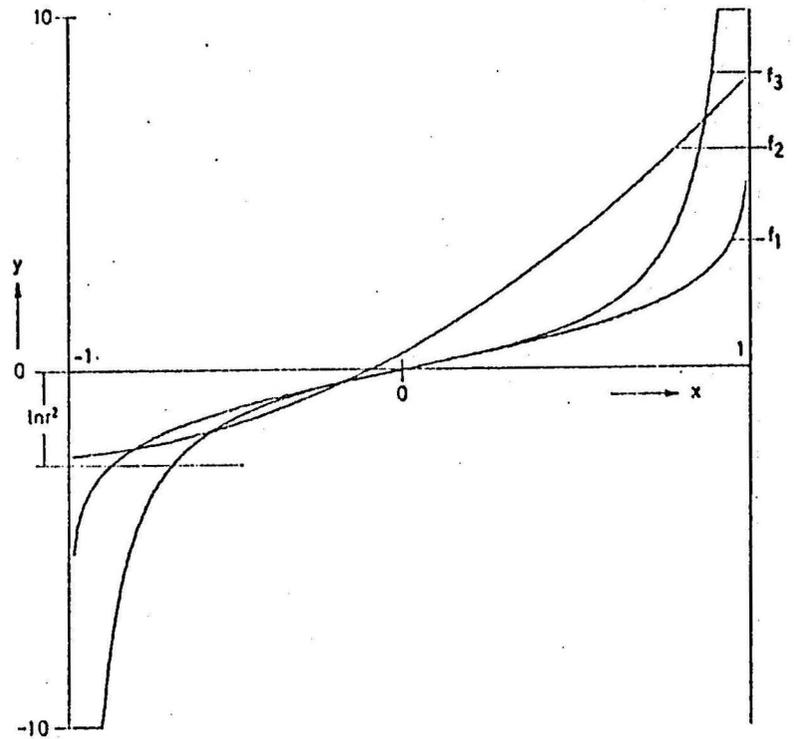


Bild 13

D = 1.162 R = .2667 Z2 = .9

$f_2 = 5.42283222 X + 2.35133603 X^2 + .58366807$

$F = -1.82577071 X - .94287409 X^2 + 2.89257071 X^3 + 1.25420264 X^4 - .31132854$

K(-1) = 3

A = -1

I = 12.0468951

II = -2.64326169

K(1) = 0

B = -.17457476

J = -30.4493335

ANZHL = 3

C = .07112889

T = 224.673693

Heinrich Heine
Institut

Gerade noch zwei-
 gipflige Mischung
 mit Bayesgrenzen.
 Drei Schnittpunkte
 von f_2 mit f_1
 ($x_{s1} \approx -1$, $x_{s2} \approx -0,6$,
 $x_{s3} \approx +1$).
 Nur ein Schnittpunkt
 von f_2 und f_3 in
 $(-1,1)$.
 Erklärungen zum
 Rechnerausdruck
 wie in Bild 9.

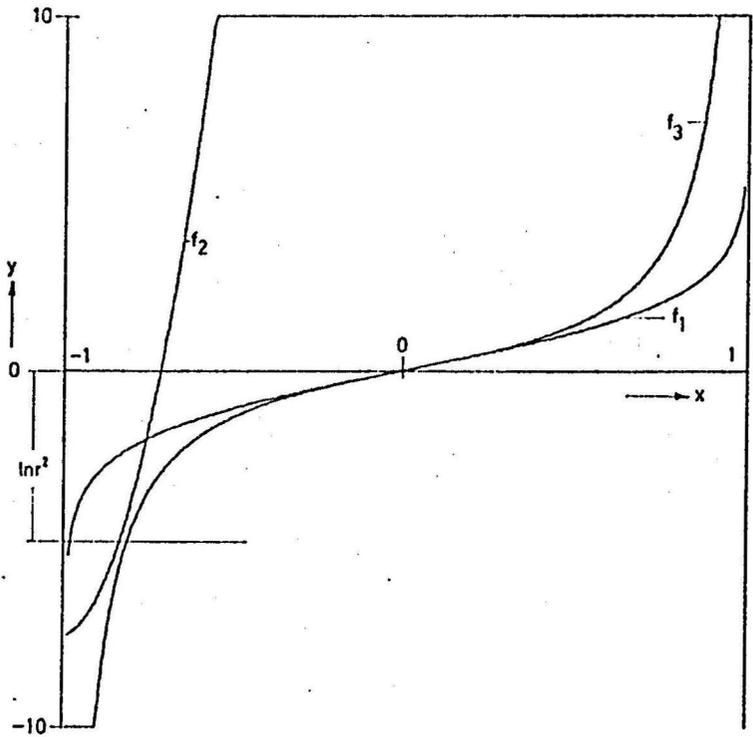


Bild 14

Z2 = .92 R = .09 D = 4.

F2 = 179.217777 X + 88.1633688 X² + 83.3873991

F = - 31.8992 X - .86066816 X² + 32.2592 X³ + 15.8704 X⁴ - 15.0097318

K(-1) = 2

A = - 1

I = 229.343389

H = - 4.81589121

K(1) = 1

B = - .01344794

J = - 8447.99164

ANZHL = 1

C = .0081

T = - 858155.192

Zweigipflige
Mischung zweier
Komponenten
gleicher Varianz
mit einer Bayes-
grenze
(Fall von
Doetsch).

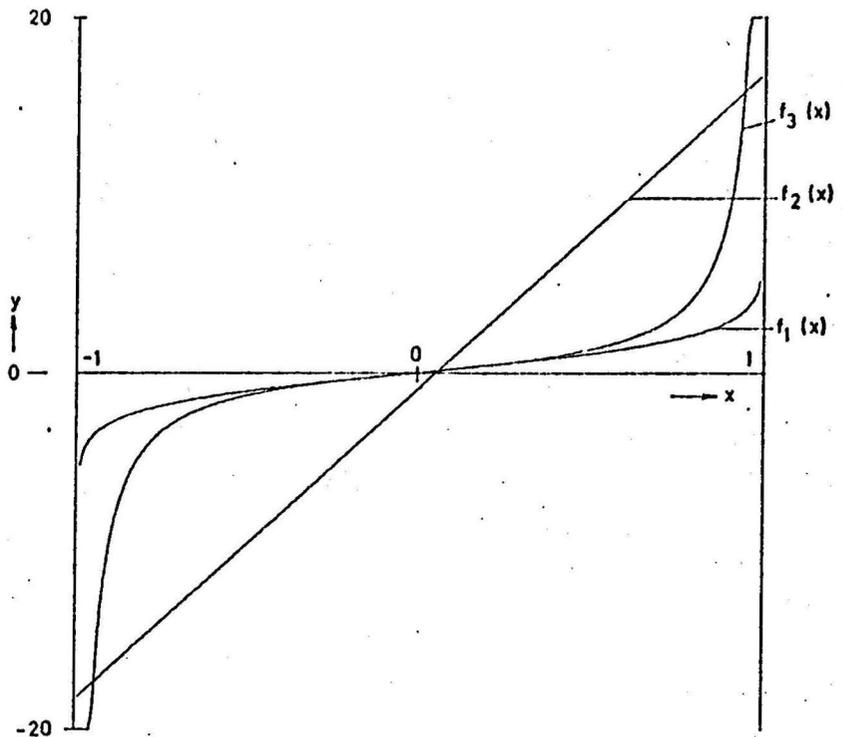


Bild 15

6.1.4.1 Bestimmung der Schnittpunktsanzahl mit der Sturm'schen Kette:

Das Verfahren der Sturm'schen Kette wird im Anhang behandelt, da es auch zur Optimierung der Lösung der Pearsongl. (21) benötigt wird.

Hier sei nur soviel gesagt:

Ausgehend von einem Polynom $f(x)$ und dessen Ableitung $f'(x)$ gewinnt man nach einem Euklidischen Teilverfahren Reste $R_i(x)$ (die im Gegensatz zu diesem negativ anzusetzen sind /17/ S. 155).

Die Folge der Funktionen $f(x), f'(x), R_1(x) \dots R_n = \text{const.}$ heißt die Sturm'sche Kette.

Die Zahl der Wurzeln von $f(x)$ zwischen zwei Grenzen ist dann gleich der Differenz der Zahl der Zeichenwechsel, die diese Kette an beiden Grenzen hat.

Die Grenzen selbst dürfen keine Wurzeln sein. Die Zahl der Wurzeln wird ohne Rücksicht auf deren Multiplizität bestimmt, d.h. Mehrfachwurzeln rechnen als eine.

$f(x)$ ist im allgemeinen ein Polynom. In der Literatur /18/ wird jedoch auch eine transzendente Funktion in der Form einer Potenzreihe mit einem modifizierten Sturm'schen Verfahren behandelt. Man bricht die Kette in diesem Fall ab, wenn man ein $R_m(x)$ gefunden hat, das im betrachteten Intervall sein Zeichen nicht ändert. Im vorliegenden Fall gelang das nicht, es wird mit der Ersatzfunktion $f_3(x)$ gearbeitet.

$$f_3(x) - f_2(x) = 0$$

also

$$\frac{2x}{(1-x^2)} - \frac{d^2}{2r} \left[(1+x)^2 - r^2(1-x)^2 \right] - \ln \frac{r^3 z_2}{1-z_2} = 0$$

werde mit $(1-x^2) (2r)$ multipliziert:

$$y(x) = -(1-x^2)d^2 \left[(1+x)^2 - r^2(1-x)^2 \right] - 2(1-x^2)r \ln \frac{r^3 z_2}{1-z_2} + 4rx$$

Diese Funktion ist vom vierten Grad und hat wegen der Pol-Nullstellenaufhebung bei ± 1 keine Wurzeln bei ± 1 . Daher kann das Verfahren der Sturm'schen Kette zwischen -1 und $+1$ angewandt werden. ± 1 können aber auch keine Schnittstellen von f_2 mit f_1 sein, wie die geometrische Anschauung zeigt.

(Vergl. dazu auch G. Doetsch /2/ S. 313 vorl. Textzeile und Behbodan /16/ S. 132 vorl. Zeile)

$y(x)$ hat also dort und nur dort Nullstellen, wo $f_3(x) - f_2(x) = 0$, wo also die Schnittstellen von f_3 und f_2 liegen.

Wir suchen jetzt die Nullstellen von

$$y(x) = a_4 x^4 + a_3 x^3 + a_2 x^2 + a_1 x + a_0 \quad (69)$$

mit den Koeffizienten:

$$a_4 = d^2(1-r^2)$$

$$a_3 = 2d^2(1+r^2)$$

$$a_2 = 2r \ln \frac{r^3 z_2}{1-z_2} \quad (70)$$

$$a_1 = 4r - 2d^2(1+r^2) = 4r - a_3$$

$$a_0 = -d^2(1-r^2) - 2r \ln \frac{r^3 z_2}{1-z_2} = -a_4 - a_2$$

zwischen den Grenzen -1 und $+1$.

Die Polynome der Sturm'schen Kette werden automatisch erstellt:

$$Y(1) = A_0 + X^3 A_3 + X^4 A_4 + A_1 X + A_2 X^2$$

$$Y(2) = A_1 + 3 X^2 A_3 + 4 X^3 A_4 + 2 A_2 X$$

$$Y(3) = 2 A_2 X A_3 + A_1 A_3 + 3 X^2 A_3^2 - 12 A_1 X A_4 - 16 A_0 A_4 - 8 A_2 X^2 A_4$$

$$\begin{aligned}
 y(1) &= f(x) \\
 y(2) &= f'(x) \\
 y(3) &= -R_1(x) \quad \text{usw.}
 \end{aligned}$$

Man erkennt, daß sich $y(3)$ noch durch a^4 dividieren läßt, was erlaubt ist, da wegen (70) a^4 nur positiv sein kann - eine Tatsache, die das allgemeine Programm zur Erstellung der Kette nicht berücksichtigen kann.

Nach Division von $Y(3)$, $Y(4)$, $Y(5)$ durch Potenzen von a^4 und Zahlenfaktoren kommt

$$\begin{aligned}
 Y(3) &= -12 A_1 X A^4 - 16 A_0 A^4 - 8 A_2 X^2 A^4 + 2 A_2 A_3 X + A_1 A_3 + 3 A_3^2 X^2 \\
 \hline
 Y(4) &= 28 A_1 A_2 A_3 X A^4 - 12 A_0 A_3^2 X A^4 - 8 A_2^3 X A^4 + 32 A_0 A_2 A_3 A^4 + 3 A_1^2 A_3 A^4 - 4 A_1 A_2^2 A^4 + 2 A_2^2 A_3^2 X \\
 &\quad - 6 A_1 A_3^3 X + 32 A_0 A_2 X A^4 - 36 A_1^2 X A^4 - 68 A_0 A_1 A^4 + A_1 A_2 A_3^2 - 9 A_0 A_3^3 \\
 \hline
 Y(5) &= -84 A_1^2 A_2^3 A_3^4 A^4 + 336 A_0 A_2^4 A_3^4 A^4 + 354 A_1^3 A_2 A_3^5 A^4 - 1584 A_0 A_1 A_2^2 A_3^5 A^4 + 2592 A_0^2 A_2 A_3^6 \\
 &\quad A^4 - 54 A_0 A_2^2 A_3^6 A^4 + 256 A_1^2 A_2^4 A_3^2 A^4 - 1024 A_0 A_2^5 A_3^2 A^4 - 1120 A_1^3 A_2^3 A_3^2 A^4 + 4992 A_0 A_1 A_2^3 \\
 &\quad A_3^2 A^4 + 1584 A_0 A_1^2 A_2 A_3^4 A^4 - 9792 A_0^2 A_2^2 A_3^4 A^4 - 243 A_1^4 A_3^4 A^4 - 1728 A_0^2 A_1 A_3^5 A^4 + 1152 \\
 &\quad A_1^3 A_2^3 A_3^3 A^4 - 5120 A_0 A_1 A_2^4 A_3^3 A^4 + 1296 A_1^4 A_2 A_3^2 A^4 - 7296 A_0 A_1^2 A_2^2 A_3^3 A^4 + 15360 A_0^2 A_2^2 A_3^3 \\
 &\quad A^4 + 9216 A_0^2 A_1 A_2 A_3^3 A^4 + 2304 A_0^3 A_3^4 A^4 - 256 A_1^2 A_2^5 A_3^3 A^4 + 1024 A_0 A_2^6 A_3^3 A^4 - 12288 A_0^2 A_1 A_2^2 \\
 &\quad A_3^4 A^4 - 12288 A_0^3 A_2 A_3^2 A^4 - 1728 A_1^4 A_2 A_3^4 A^4 + 9216 A_0 A_1^2 A_2^3 A_3^4 A^4 - 8192 A_0^2 A_2^4 A_3^4 A^4 + 16384 A_0^3 A_2 \\
 &\quad A_3^5 A^4 + 9 A_1^2 A_2^2 A_3^6 - 36 A_0 A_2^3 A_3^6 + 162 A_0 A_1 A_2 A_3^7 - 36 A_1^3 A_3^7 - 243 A_0^2 A_3^8
 \end{aligned}$$

(72)

Eine weitere Vereinfachung ließe sich erzielen, indem man $a^4 = 1$ setzt und $a^3 \dots a_0$ neu definiert. Die entstehenden Formeln entsprechen (71) und (72) mit $a^4=1$. Sie bleiben kompliziert genug, um eine andere Lösungsmöglichkeit ins Auge zu fassen.

Für das Beispiel in Bild 9 wurden die Zahlenwerte für z_2 , r und d eingesetzt und die Funktionswerte $Y(i)$ an den Grenzen -1 und $+1$ berechnet.

$$Y(1) = 2.42322139 X - 1.71263299 X^2 + .07338176 X^3 + .01699139 X^4 + 1.6965416$$

$$W(-1) = - 2.49330316$$

$$W(1) = 2.49330316$$

$$Y(2) = - 3.42526599 X + .21924529 X^2 + .06436558 X^3 + 2.42922139$$

$$W(-1) = 6.09936709$$

$$W(1) = - .72143372$$

$$Y(3) = - .01154817 X + .00380543 X^2 - .0041825$$

$$W(-1) = .01117116$$

$$W(1) = - .01192513$$

$$Y(4) = .30360474E-04 X - .41647501E-04$$

$$W(-1) = - .72007975E-04$$

$$W(1) = - .11287026E-04$$

$$Y(5) = .1135653E-10$$

$$W(-1) = .1135653E-10$$

$$W(1) = .1135653E-10$$

***>

An der Grenze -1 ergibt sich die Vorzeichenfolge (Vorzeichen von $w(-1)$ unter $Y(1) \dots Y(5)$) $- + + - + = 3$ Zeichenwechsel. An der Grenze $+1$ kommt $+ - - - + = 2$ Zeichenwechsel. Zwischen -1 und $+1$ geht ein Zeichenwechsel verloren. Daher hat das Polynom (69) eine Nullstelle zwischen -1 und $+1$ entsprechend einem Schnittpunkt in Bild 9 zwischen f_2 und f_3 .

Mit den im Anhang mitgeteilten Fordecalroutinen wurde für alle Beispiele in den Bildern 9 bis 14 die Sturm'sche Kette gebildet und die Zahl der Zeichenwechsel bestimmt. Die Ergebnisse bestätigen die Anschauung in den Bildern 9-14.

Die Restpolynome der Sturm'schen Kette sind schon beim Grad 4 des Ausgangspolynoms $y(x)$ recht umfangreich. Für die Pearsongl. 9ten Grades werden sie völlig unübersichtlich, abgesehen davon, daß der Rechenaufwand viel zu hoch wird.

Hier wurde unter Preisgabe einer gewissen Allgemeinheit ein anderer Weg der numerischen Berechnung gefunden. Auch im vorliegenden Fall ist es fraglich, ob der Rechenaufwand, nämlich die Koeffizienten $a^4 \dots a_0$, die Polynomwerte $Y(1) \dots Y(5)$ an den Grenzen -1 und $+1$ und dann die Differenz der Zeichenwechsel zu berechnen, gerechtfertigt ist.

Aufgrund der geometrischen Anschauung wurde eine erheblich einfachere Methode gefunden.

6.1.4.2 Einfache Bestimmung der Schnittpunktsanzahl
zwischen Ersatzfunktion und Parabel.

Bedingungen für die Zweigipfligkeit.

Behbodian /16/ S. 139 führt aus, daß man anhand der geometrischen Anschauung notwendige und hinreichende Bedingungen für die Eingipfligkeit bekäme, nur sei das Ergebnis recht kompliziert. Er teilt es jedoch leider weder mit, noch gibt er eine Literaturstelle an.

Er leitet eine hinreichende Bedingung für die Eingipfligkeit her Gl. (27) von der er behauptet, sie sei schärfer als die Eisenberges /19/ Gl.(4).

Wir sind an einer zumindest hinreichenden Bedingung für die Zweigipfligkeit interessiert. Aus der Negation einer hinreichenden Bedingung für die Eingipfligkeit folgt nur eine notwendige Bedingung für die Zweigipfligkeit.

Wir wollen auf Eisenberges und Behbodians Ungl'n trotzdem kurz eingehen, um zu zeigen, daß Behbodians Behauptung, seine Ungl. sei schärfer als die Eisenberges, für alle Werte von $0 < r < \frac{1}{4}\sqrt{11} \approx 0,829$, also für den größten Teilbereich von $0 < r \leq 1$ falsch ist.

Eisenberges hinreichende Bedingung für die Eingipfligkeit lautet:

$$(\gamma_2 - \gamma_1)^2 < \frac{27 \sigma_1^2 \sigma_2^2}{4(\sigma_1^2 + \sigma_2^2)} \quad /19/ \text{ Gl.}(4)$$

Worin γ_1, γ_2 die Mittelwerte und σ_1^2, σ_2^2 die Varianzen der Komponenten sind.

Mit (52)

$$(\gamma_2 - \gamma_1)^2 = 4d^2 \sigma_1 \sigma_2$$

und (53) $r = \frac{\sigma_1}{\sigma_2}$

wird Eisenberges Ungl. zu.

$$d^2 < r \cdot \frac{27}{16} \cdot \frac{1}{1+r^2}$$

Behbodians Ungl. lautet:

$$d^2 \leq r \quad /16/ \text{ Gl. (26)}$$

Der Faktor $\frac{27}{16} \cdot \frac{1}{1+r^2}$ ist für $r=1$ kleiner,

für $r = \frac{\sqrt{11}}{4}$ gleich und für $r = 0$ größer als 1, also

ist für $0 < r < \frac{\sqrt{11}}{4}$ Eisenbergers Ungleichung schärfer als die Behbodians.

Wir leiten zunächst eine eigene notwendige Bedingung für die Zweigipfligkeit ab, die mithilfe der Ersatzfunktion $f_3(x)$ zu einer hinreichenden Bedingung verschärft wird. Diese Bedingung leistet das gleiche, wie das Verfahren der Sturm'schen Kette im Zusammenhang mit der Ersatzfunktion, ist jedoch wesentlich einfacher:

Der Gedanke ist, die Diskriminante der Gleichung (69) dazu heranzuziehen zu entscheiden, ob die Funktionen f_2 und f_3 einen oder drei Schnittpunkte in $-1 < x < +1$ haben.

Wir betrachten in Bild 9 den Verlauf von $f_3(x)$. Die Parabel $f_2(x)$ ist nach oben geöffnet und hat ihren Scheitelwert links von -1 , wie bereits diskutiert. Sie hat also stets einen Schnittpunkt mit f_3 links von -1 . (Ausgenommen ist der Fall $r=1$ in dem sie zur Geraden entartet, in diesem Fall ergeben sich jedoch ein bis drei Schnittpunkte von f_2 und f_3 im Intervall $-1 < x < +1$). Ein bis drei Schnittpunkte können im Intervall $-1 < x < +1$ entstehen, schließlich können rechts von $+1$ weitere 2 Schnittpunkte oder

ein Berührungspunkt liegen. (Siehe z.B. Bild 9). Für $0 < r < 1$ hat Gl. (9) also mindestens zwei reelle Wurzeln.

Über die Diskriminante schreibt Dörrrie /20/ S. 70 :

"Bei positiver Diskriminante sind die Wurzeln der biquadratischen Gleichung alle reell oder alle komplex, je nachdem die Resolventenkoeffizienten E und F beide positiv sind oder nicht.

Bei verschwindender Diskriminante sind alle vier oder nur zwei Wurzeln reell, je nachdem E und F beide positiv sind oder nicht.

Bei negativer Diskriminante sind zwei Wurzeln reell, die beiden andern konjugiert komplex."

Aus der Tatsache der stets vorhandenen zwei reellen Wurzeln folgt, daß die Resolventenkoeffizienten E und F nicht betrachtet zu werden brauchen.

Ist die Diskriminante negativ, so hat f_2 nur zwei Schnittpunkte mit f_3 und davon nur einen in $-1 < x < +1$ und einen in $-\infty < x < -1$, ist sie jedoch positiv, so hat f_2 vier Schnittpunkte mit f_3 und zwar einen in $-\infty < x < -1$ und entweder einen in $-1 < x < +1$ und zwei in $1 < x < \infty$ (Bild 9) oder drei in $-1 < x < +1$ und keinen in $1 < x < \infty$ (Bild 12). (Von Berührungspunkten sei abgesehen, sie bringen keine prinzipiellen Schwierigkeiten)

Wenn aber drei Schnittpunkte zwischen f_2 und f_3 und damit notwendig auch drei zwischen f_2 und f_1 in $-1 < x < +1$ liegen sollen, so muß notwendig die Parabel f_2 eine Nullstelle in diesem Intervall haben.

So kommen wir zur notwendigen Bedingung für die Zweigipfligkeit:

Die Wurzeln von $f_2(x)$ sind :

Meinrich Herz-
Institut

$$x_{1,(2)} = - \frac{1+r^2}{1-r^2} (\pm) \sqrt{\left(\frac{1+r^2}{1-r^2}\right)^2 - \left(1 + \frac{2r}{d^2(1-r^2)} \ln \frac{r^3 z_2}{1-z_2}\right)} \quad (74)$$

Die größere Wurzel x_1 liegt im Intervall

$$-1 < x_1 < +1$$

$$\frac{1+r^2}{1-r^2} - 1 < + \sqrt{\quad} < \frac{1+r^2}{1-r^2} + 1$$

$$2 r^2 < (1-r^2) \sqrt{\quad} < 2$$

Da alle Zahlen positiv sind, kommt

$$4r^4 < 4r^2 - \frac{2r}{d^2} (1-r^2) \ln \frac{r^3 z_2}{1-z_2} < 4$$

und daraus schließlich:

$$\boxed{-1 < \frac{r}{2d^2} \ln \frac{r^3 z_2}{1-z_2} < r^2} \quad (75)$$

als notwendige Bedingung für die Zweigipfligkeit.

Hinreichend ist, daß zusätzlich die Diskriminante positiv ist

$$\boxed{T > 0} \quad (76)$$

Die Diskriminante wird nach der Formel von Cayley-Boole /20/ S. 67 berechnet:

$$T = \frac{1}{27} (4i^3 - j^2)$$

mit

$$i = a_2^2 + 12a_4a_0 - 3a_3a_1 \quad \text{und}$$

$$j = 2a_2^3 + 27a_4a_1^2 + 27a_0a_3^2 - 72a_4a_2a_0 - 9a_3a_2a_1$$

worin $a_0 \dots a_4$ die Koeffizienten (70) der Gleichung (69) sind.

Um auf Zweigipfligkeit zu entscheiden, werden die Bedingungen (75) und (76) nacheinander angewendet.

In den Beispielen Bild 9 - 14 sind verschiedene Fälle mit den zugehörigen Ungleichungen und Diskriminanten dargestellt. Es sind auch bewußt zwei Fälle konstruiert worden, in denen die hinreichende Bedingung nicht erfüllt ist und die Mischung trotzdem zweigipflig ist. In der Praxis tritt ein solcher Fall selten auf, und ihn bei der Suche nach zweigipfligen Lösungen zu verwerfen, stellt keine Beeinträchtigung des Verfahrens dar.

6.2. Existenz der Bayesgrenze

Die nach Bayes benannte Stelle gleicher Dichte beider Komponenten ist als weitere Nebenbedingung von Interesse. Bild 8 zeigt einen Fall, in dem keine Bayesgrenze existiert.

Wenn sie existieren soll, muß gelten:

$$z_1 h_1 = z_2 h_2 \quad (77)$$

Mit den Gleichungen (61) erhält man

$$\frac{z_1}{\sqrt{r}} \cdot e^{-\frac{d^2}{2r}(1+z)^2} = z_2 \sqrt{r} e^{-\frac{d^2}{2}r(1-z)^2}$$

$$\ln \frac{z_1}{z_2 r} = \frac{d^2}{2r} \left[(1+z)^2 - r^2 (1-z)^2 \right]$$

und mit $z_1 = 1-z_2$

$$\boxed{\frac{d^2}{2r} \left[(1+z)^2 - r^2(1-z)^2 \right] + \ln \frac{z_2 r}{1-z_2} = 0} \quad (78)$$

Setzt man in (78) wieder x anstelle von z und schreibt für die linke Seite $w(x)$, dann ist $w(x)$ eine Parabel, die sich von derjenigen $f_2(x)$ Gl. (65) nur um das additive Glied $\ln r^2$ unterscheidet:

$$w(x) + \ln r^2 = f_2(x)$$

Da $r \leq 1$ liegt $w(x)$ im allgemeinen oberhalb von $f_2(x)$. Man kann in Bild 9 bis 12 eine Parallele zur x -Achse im Abstand $|\ln r^2|$ unterhalb der x -Achse ziehen. Ihre Schnittpunkte mit f_2 geben die Lage der Bayesgrenzen an. In der Diskussion kann man der Diskussion von f_2 folgen.

Insbesondere ist bei $r=1$ stets ein Schnittpunkt der Geraden f_2 und w mit der x -Achse vorhanden. Der Wert der Scheitelpunktsabszisse x_s von f_2 gilt auch für w und bezeichnet die Abszisse des Berührungspunktes beider Komponenten, zu der symmetrisch die Nullstellen von $w(x)$ liegen, falls sie vorhanden sind, falls also eine Bayesgrenze existiert.

Für $r < 1$ gewinnt man die Bedingung für die Bayesgrenze also aus dem zur Diskriminante von (74) analogen Ausdruck:

$$\left(\frac{1+r^2}{1-r^2}\right)^2 - \left[1 + \frac{2r}{d^2(1-r^2)} \cdot \ln \frac{rz_2}{1-z_2}\right] > 0$$

nach Umformen:

$$\boxed{2r - \frac{1-r^2}{d^2} \cdot \ln \frac{rz_2}{1-z_2} > 0} \quad (79)$$

7. Durchführung des Approximationsverfahrens.

7.1 Beschreibung des vorhandenen Programms.

Der Ablauf des Zerlegungsverfahrens geht im wesentlichen aus dem Blockdiagramm 1 hervor.

Nicht eingezeichnet ist die Möglichkeit, empirisch gemessene und durch Häufigkeiten in Klassen beschriebene Wahrscheinlichkeitsdichten zu approximieren.

Dem Programm werden zunächst die Parameter zur Steuerung des Rechnungsablaufes eingegeben. Dazu gehören: Signifikanzniveaus für die Abweichung der Formfaktoren vom Typ der Normalverteilung, kleinster zulässiger Mischungsfaktor, Zusatzbedingungen, wie Existenz der Bayesgrenzen oder Zweigipfligkeit. Dann wird die Länge der zu bearbeitenden Datenblöcke vorgegeben.

Nach dem Einlesen eines Datenblockes und Berechnen der empirischen Momente werden aufeinanderfolgend Zerlegung, Zusammenlegung, weiteres Zerlegen usw. durchgeführt, wie auf S. 3 und 4 beschrieben. Diese Funktionen werden gesteuert durch einen Test auf Normalität, dessen Signifikanzniveau vorher eingegeben wurde (typisch: $\alpha = 10\%$).

Weiterhin ist die Unterdrückung kleiner Komponenten ab einer beliebig vorgebbaren Schranke S vorgesehen. Wenn nach Zerlegen eine Komponente den Anteil (Mischungsfaktor) $< S$ hat, wird die zugehörige Zerlegung rückgängig gemacht. Komponenten $< 2 \cdot S$ werden von vornherein von der Zerlegung ausgeschlossen, da nach Zerlegung mindestens eine Komponente $\leq S$ entstünde. (Typischer Wert für $S = 1\%$)

Eine weitere Wahlmöglichkeit des Programmes sieht Nebenbedingungen vor, die zusätzlich zur Zerlegbarkeit gefordert werden können: Existenz einer zweigipfligen Summenkurve aus zwei benachbarten Komponenten, Existenz von Bayesgrenzen. Wenn die Nebenbedingungen gestellt sind, wird eine Zerlegung rückgängig gemacht, falls sie nicht erfüllt sind, bzw. ein Zusammenlegen vorgenommen, obgleich die Formfaktoren dagegen sprechen.

Anfang

Eingabe der Parameter zur Steuerung des Rechnungsablaufes (Untere Grenze für den Mischungsfaktor, Irrtumswahrscheinlichkeit für den Formfaktor. Zusatzbedingung: Existenz der Bayesgrenze oder Zweigipfligkeit, Blocklängen)

Eingabe der Werte für die Simulation (Parameter für die Erzeugung einer Mischverteilung, Anzahl der Zufallszahlen, Blocklängen, Anfangszufallszahl)

Berechnung der theoretischen zentralen Momente aus den Simulationsparametern, ggf. direkte Eingabe von Momenten

Verarbeitg. von Meßdaten

nein

Erzeugung der Zufallszahlen nach den eingegebenen Parametern und den vorgeschriebenen Blocklängen.

b

c

Einlesen eines Blockes von Daten oder Zufallszahlen, Aufteilung auf die vorhandenen Klassen, Berechnung der Potenzsummen in jeder Klasse.

Wurden im letzten Durchlauf durch Zerlegung neue Klassen gebildet ?

nein

Berechnung der empirischen biasfreien zentralen Momente

a

Zerlegen

Zusammenlegen

nein

Ging dem Durchlauf ein solcher unmittelbar voran, in dem ein Zerlegungsversuch gemacht wurde ?

Läßt sich eine Klasse gemäß den Steuerparametern zerlegen ?

nein

b

Bestimmung der Gleichung nach Pearson und Berechnung der Wurzeln (Sonderlösung im Symmetriefall)

Zerlegung in Komponenten möglich ?

nein

Bestimmung der Komponenten und Prüfung auf Bayesgrenze und Zweigipfligkeit

Genügen die entstandenen Komponenten den Zusatzbedingungen, falls gestellt ?

nein

Abspeichern der Komponenten als Repräsentanten der neuen Klassen

Alle Klassen abgearbeitet ?

nein

Addition zweier aufeinanderfolgender Sätze der Potenzsummen und daraus Berechnung der biasfreien zentralen Momente

Lassen sich die Klassen gemäß den Steuerparametern zusammenlegen ?

nein

Abspeichern des Momentensatzes als Repräsentation der neuen Klasse

nein

Alle Klassen abgearbeitet ?

Sortieren der vorhandenen Klassen nach ihrem Mittelwert

Grafische Ausgabe möglich

nein

Gelang der Zerlegungsversuch in mindestens einer Klasse ?

Ende

folgender oder gleicher Block

gleicher Block

c

Im Programm sind zahlreiche Möglichkeiten zur Simulation vorgesehen. Es können anstelle der Daten a) Zufallszahlen nach eingegebenen Parametern erzeugt werden oder b) die theoretischen zentralen Momente der Mischung direkt aus diesen Parametern berechnet werden.

Es können Mischungen aus einer beliebigen Anzahl von Komponenten von Gleich-, Dreiecks- oder Normalverteilung erzeugt werden. Eine Mischverteilung kann Komponenten aller drei Typen enthalten.

Würden keine Zufallszahlen erzeugt, also die theoretischen zentralen Momente berechnet, dann kann natürlich nur einmal in zwei Komponenten zerlegt werden. Diese Version des Programms diene mehr grundsätzlichen Untersuchungen der Zerlegbarkeit.

Im Programm kann ein stationärer Prozeß simuliert werden, indem fortlaufend Pseudozufallszahlen nach dem vorgegebenen Gesetz gebildet und blockweise verarbeitet werden.

Die Zahl der benötigten Blöcke richtet sich nach dem Verlauf des Zerlegens und Zusammenlegens. Bei der Simulation des stationären Prozesses wird vor jedem Zerlegen und Zusammenlegen ein neuer Datenblock eingelesen. Beim mehrfachen Verarbeiten des gleichen Blockes wird dieser zu Beginn und **nachher** nur dann erneut eingelesen, wenn zerlegt wurde und dadurch neue Klassen entstanden. Nach einem Zusammenlegen benachbarter Komponenten ist ein erneutes Einlesen überflüssig, da sich die Momente der aus dem Zusammenlegen hervorgegangenen Komponente durch Addition derjenigen der Ausgangskomponenten ergeben.

Man könnte diese Steuerung auch für die fortlaufende Verarbeitung von Blöcken neuer Zufallszahlen verwenden, da nicht zu sehen ist, welchen Vorteil die Verwendung eines neuen Blockes nach dem Zusammenlegen hat.

Wenn ein Zerlegungsversuch erfolglos war, ist das Verfahren an seinem Ende angelangt.

Dies folgt daraus, daß sich Zerlegungs- und Zusammenlegungsversuche abwechseln. Gelingt bereits der erste Zerlegungsversuch nicht, so versteht sich von selbst, daß das Verfahren nicht weitergeführt werden kann. Mißlingt ein späterer Zerlegungsversuch, dann ist ihm ein Versuch zusammenzulegen vorangegangen. Hierbei sind bereits alle Möglichkeiten des Zusammenlegens benachbarter Komponenten abgeprüft worden. Ein dann auf den erfolglosen Zerlegungsversuch folgender Versuch zusammenzulegen würde keine neuen Gesichtspunkte erbringen.

Man kann dieses Abbruchkriterium natürlich außer Kraft setzen, wenn man einen langsam veränderlichen Prozeß beobachtet.

Die Frage nach der optimalen Länge der Blöcke ist offen geblieben. Sicherlich benötigt man am Anfang kürzere Blöcke (typisch 100 - 200 Zufallszahlen) während am Schluß des Verfahrens zur Steigerung der Genauigkeit längere Blöcke (500 - 2000 Zufallszahlen) angebracht sind. Es wurde daher vorgesehen, die Blocklänge durch manuellen Eingriff variabel zu gestalten. Da im Einzelfall nicht fest steht, wieviel Schritte das Verfahren benötigt, um zum Ende zu gelangen, konnte die Wahl der variablen Blocklänge bisher nicht automatisch erfolgen.

Anhand eines Beispiels wird gezeigt werden, weshalb abgesehen vom geringeren Aufwand kurze Blöcke am Anfang günstiger sind als lange.

Die Operationen während des Zerlegens bzw. Zusammenlegens sind aus dem Blockdiagramm 1 abzulesen. Einzelheiten sind in den entsprechenden Abschnitten dieses Berichtes beschrieben.

Zum Zerlegen sei noch bemerkt:

Bei der Lösung der Pearsongleichung neunten Grades wurde mit der Sturm'schen Kette bestimmt, ob Wurzeln zwischen $-\mu_2$ und 0 liegen. Die Gleichung neunten Grades wurde mit Hilfe eines im reellen arbeitenden Newton-Raphson-Verfahrens gelöst. Das eigens für diese Aufgabe geschriebene Programm (Werkauftrag Selim) vermeidet zahlreiche Fehlermöglichkeiten (Käfigbildung, waagerechte Tangente, etc.) und war bei gleicher Genauigkeit in zahlreichen Versuchen 2-3 mal so schnell wie die komplexe Newton-Raphson-Methode aus dem IBM-Scientific Subroutine Package (SSP). In einem Fall wurde sogar ein Verhältnis der Rechenzeiten von 1:50 gemessen. Die reelle Routine neigt zu Fehlern bei Doppelwurzeln oder bei im Vergleich zur Intervalllänge ($-\mu_2, 0$) sehr dicht beieinander liegenden Wurzeln. Der Fehler besteht im Übersehen der Wurzeln. Durch Vergleich der ermittelten mit der nach dem Sturm'schen Verfahren vorausgesagten Anzahl der Nullstellen werden die Fehler entdeckt. Automatisch wird dann die langsamere, im Fall der Doppelwurzeln zuverlässigere IBM-Routine aufgerufen.

Die IBM-Routine versagt u.U. jedoch auch, allerdings wurde beim Testen nur ein Fall beobachtet, in welchem die reelle Routine dann das richtige Ergebnis lieferte. Das Versagen von Wurzelroutinen in speziellen Fällen ist nichts Ungewöhnliches. Deshalb enthält das IBM-SSP verschiedene Verfahren (Graeffe, Bairstow, Nassita etc.). Drei wurden erprobt. Durch die durchgeführte Hintergrundschaltung von zwei Routinen erhält man eine hohe Zuverlässigkeit der Wurzelbestimmung.

Zum Zusammenlegen ist zu sagen, daß die geordneten Komponenten von links nach rechts, d.h. in aufsteigender Reihenfolge der Mittelwerte, auf Zusammenlegen geprüft werden und zwar, falls keine Nebenbedingungen bestehen, aufgrund des Signifikanztestes, sonst nach der gewählten Nebenbedingung.

Sind zwei Komponenten zusammengelegt worden, so wird die neu entstandene Komponente mit ihrer jetzigen Nachbarkomponente auf Zusammenlegen geprüft. Auf diese Weise kann es geschehen, daß erst z.B. Komponente 2 u. 3 und die jetzt als 2 bezeichnete Summenkomponente mit der als 3 bezeichneten Komponente 4 vereinigt wird.

(Rechnerausdruck: Zusammenlegen von 2 u. 3
Zusammenlegen von 2 u. 3)

Es wäre auch möglich, die Komponenten in absteigender Reihenfolge der Mittelwerte oder in mehreren Durchläufen auf Zusammenlegen zu prüfen. Obwohl theoretisch dabei andere Ergebnisse entstehen könnten, werden nur solche Klassen zusammengelegt, die nach dem Entscheidungskriterium äquivalent sind. Im allgemeinen dürften die verschiedenen Möglichkeiten des Durchprüfens zum gleichen Ergebnis führen. Dieser Frage wurde nicht weiter nachgegangen. Das gewählte Verfahren zeigt zufriedenstellende Resultate.

7.2 Noch nicht verwirklichte Möglichkeiten zur Verbesserung

7.2.1 Beschleunigung des Programms.

Im Programm wird häufig die Funktion e^{-x^2} benötigt. Versuche, sie durch Ausdrücke der Form $\frac{1}{1+x^2}$ zu ersetzen,

brachten keine merkliche Zeitersparnis. Würde man jedoch eine Wertotabelle $10\ 000 e^{-x^2}$ mit 500 bis 1000 Werten als ganze Zahlen speichern und alle mit e^{-x^2} zusammenhängenden Rechnungen im Integer-Bereich ganzzahlig durchführen, so würde man an der CAE ca. ein Zeitgewinn 1:20 erhalten.

(Fixpoint Addition 3.5 μ s Floatingpoint Addition 92 μ s
Multiplikation 7 μ s Multiplikation 147 μ s)

- 2) Nach dem Zerlegen einer Komponente und neuem Einlesen eines Blockes von Zufallszahlen werden die Momente neu berechnet. Hat man die gleichen Zufallszahlen und nur eine Ausgangskomponente, so genügt es, nach dem Zerlegen einen Momentensatz zu bilden. Den zweiten gewinnt man durch Differenzbildung des neuen mit dem alten Momentensatz,
- 3) Das Programm könnte in Assemblersprache geschrieben werden.

7.2.2 Optimierung

Es könnten Kriterien zur Steuerung der Blocklänge und des Signifikanzniveaus entwickelt werden. Das Signifikanzniveau α bestimmt, wie groß die Neigung des Verfahrens zum Zerlegen, bzw. Zusammenlegen ist. Bei kleinem α sind die Schranken, die der Formfaktor überschreiten muß, ehe auf Zerlegen entschieden wird, weit. Umgekehrt wird bei zu großem α nicht mehr zusammengelegt. Meist wurde $\alpha = 10\%$ gewählt. Dieser empirisch ermittelte Wert liefert Zerlegungen und Zusammenlegungen. Es wäre nun sicher sinnvoll, zu Beginn des Verfahrens $\alpha > 10\%$ und wenn kleine Komponenten auftreten $\alpha < 10\%$ zu machen. Dadurch würde die Neigung des Verfahrens, sehr kleine Komponenten weiter zu zerlegen gemindert und die Neigung, sie mit anderen zu vereinigen verstärkt. Vielleicht könnte man dann auf die Schwelle S für die kleinste zulässige Komponente verzichten.

8. Erzielte Ergebnisse

Wir zeigen jetzt einige mit dem Programmsystem erzielte Ergebnisse aus der Simulation und aus der Approximation gemessener Wahrscheinlichkeitsdichten.

8.1 Mischungen aus Normalverteilungen

Es ist wünschenswert, daß sich Mischungen aus Normalverteilungen wieder in solche zerlegen lassen. Natürlich ist es nicht selbstverständlich, daß es eine Mischung aus zwei Normalverteilungen gibt, deren Momente $\mu_2 \dots \mu_5$ mit denjenigen einer Mischung aus drei oder fünf Normalverteilungen übereinstimmen. Wir zeigten Beispiele, in denen dies der Fall ist (Bild 3 und 4) und werden solche bringen, bei denen es nicht der Fall ist. Bei letzteren würde der Zerlegungsversuch scheitern. Bei zwei Beispielen dieser Art, die bei der Simulation gefunden wurden, wurden Zufallszahlen nach dem gleichen Gesetz gemischt, nach dem zuvor die theoretischen Momente berechnet worden waren. Wenn man nun die Anzahl der Zufallszahlen klein genug machte, so wichen die aus den Zufallszahlen erzeugten empirischen Momente soweit von dem theoretischen ab, daß die Zerlegung möglich wurde. Dies ist ein wichtiger Grund dafür, mit variabler Blocklänge zu arbeiten. Kurze Blöcke zu Beginn der Zerlegung sind nicht nur sparsam im Aufwand, sondern auch günstig hinsichtlich der Zerlegungsneigung.

Im einzelnen ist zu den Beispielen zu sagen:

Bild 2 zeigt, daß die Approximation der Mischung aus drei Komponenten (strichliert) durch zwei Komponenten so gut ist, daß eine Weiterführung des Verfahrens nicht erfolgt. Die Momente für die Zerlegung wurden theoretisch berechnet.

Bild 3 zeigt ebenfalls eine Zerlegung, der die theoretischen Momente einer Mischung von fünf Normalverteilungen zugrunde liegt. Hier wäre eine Weiterführung des Verfahrens sinnvoll. Deshalb ist eine Mischung von 6000 Pseudozufallszahlen nach den Parametern der fünf Normalverteilungen im Bild hergestellt worden. (Bild 16)

Das Verfahren findet auch hier eine Zerlegung in zwei Komponenten und bricht dann ab. Das ist zweifellos nicht optimal. Dieses Beispiel zeigt am deutlichsten die Grenzen des Verfahrens. Es wird darauf noch abschließend einzugehen sein. Bereits hier sei jedoch gesagt, daß das Verfahren keinesfalls dazu neigt, nur zwei Komponenten zu finden, auch wenn dies nach einigen hier zu zeigenden Beispielen so scheinen sollte.

Relative Häufigkeiten von
6000 Zufallszahlen erzeugt
als Mischung von fünf Normalverteilungen und Approximation durch nur zwei Normalverteilungen.
Signifikanzniveau $\alpha = 10\%$.

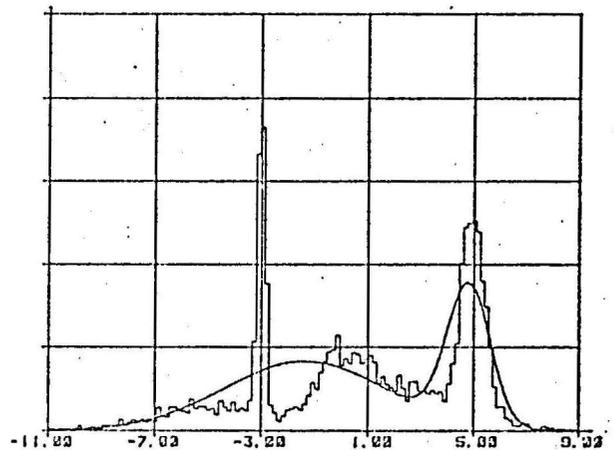


Bild 16

Rechnerausdruck:

AA[] BA[] Anteile, AMY[], BMY[] Mittelwerte,
ASIGM[], BSIGM[] Standardabweichungen.

EINGEGEBENE WERTE

(GAUSS-VERTEILUNG)

AA[1] =	.150000	AMY[1] =	-5.000000	ASIGM[1] =	2.000000
AA[2] =	.150000	AMY[2] =	-3.000000	ASIGM[2] =	.150000
AA[3] =	.200000	AMY[3] =	.000000	ASIGM[3] =	1.000000
AA[4] =	.250000	AMY[4] =	3.000000	ASIGM[4] =	2.000000
AA[5] =	.250000	AMY[5] =	5.000000	ASIGM[5] =	.400000

LSG = -.852557145484E 01
2-GIPFL.

LEESUNGSKOMPONENTEN

BA[1] =	.637681	BMY[1] =	-1.378265	BSIGM[1] =	3.039105
BA[2] =	.362319	BMY[2] =	4.696291	BSIGM[2] =	.932773

Wir geben jetzt zwei Beispiele für den Fall, daß die theoretisch aus einer Mischung von mehreren Normalverteilungen berechneten Momente keine Zerlegung in zwei Normalverteilungen zulassen, die Simulation mit Zufallszahlen jedoch zu einer Zerlegung führt.

Erfolgloser
Zerlegungsversuch einer
Mischung aus drei
Normalverteilungen aufgrund
der theoretischen
Momente.

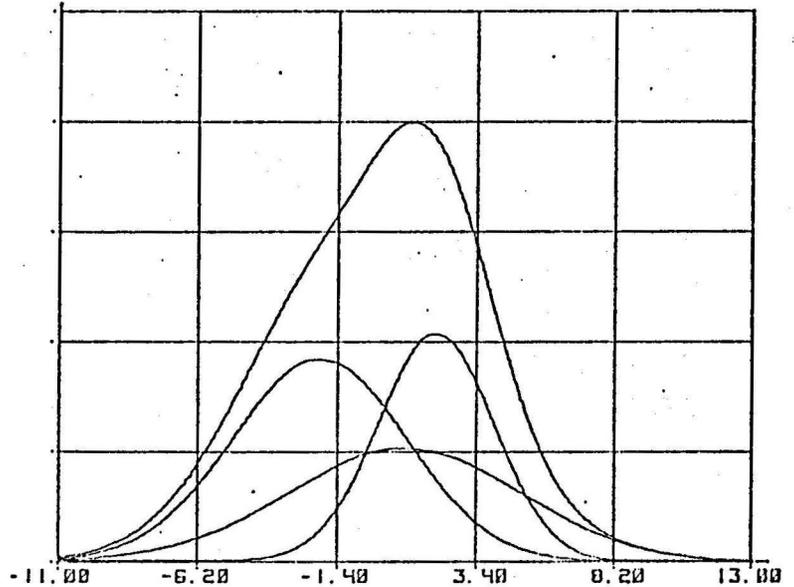


Bild 17

EINGEGEBENE WERTE

[GAUSS-VERTEILUNG]

AA(1) =	.400000	AMY(1) =	-2.000000	ASIGM(1) =	3.000000
AA(2) =	.300000	AMY(2) =	1.000000	ASIGM(2) =	4.000000
AA(3) =	.300000	AMY(3) =	2.000000	ASIGM(3) =	2.000000

AUS EINGABEWERTEN BERECHNETE ZENTRALE MOMENTE DER MISCHVERTEILUNG

MY1 =	-.363797820709E-11	MY2 =	.126899999999E 02	MY3 =	-.430800000054E 01
MY4 =	.498461699994E 03	MY5 =	-.718269600384E 02	MY6 =	.344834940441E 05

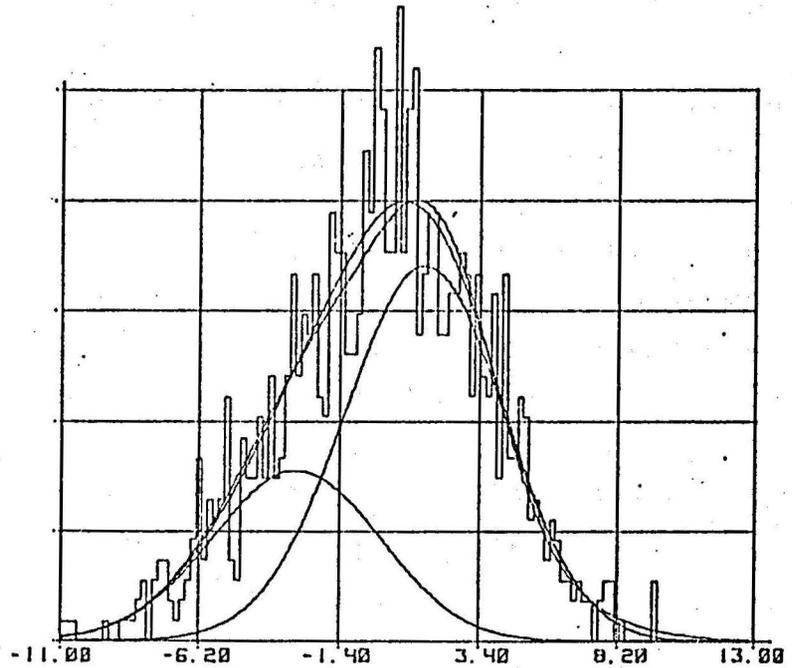
LSG: -.187821353985E 00
KEINE ZERLEGUNG MÖGLICH

LSG: -.513774456824E 00
KEINE ZERLEGUNG MÖGLICH

KEINE LÖSUNG

Zerlegung auf-
grund der empi-
rischen Momente.

Bild 18



EINGEGEBENE WERTE

(GAUSS-VERTEILUNG)

AA(1)#	.400000	AMY(1)#	-2.000000	ASIGM(1)#	3.000000
AA(2)#	.300000	AMY(2)#	1.000000	ASIGM(2)#	4.000000
AA(3)#	.300000	AMY(3)#	2.000000	ASIGM(3)#	2.000000

Simulation mit 1000 Zufallszahlen.
Signifikanzniveau $\alpha = 31,7 \% (1 \sigma)$

LSG. .414039615227E 01

LOESUNGSKOMPONENTEN

BA(1)#	.321961	BMY(1)#	-2.866504	BSIGM(1)#	2.862065
BA(2)#	.678039	BMY(2)#	1.438536	BSIGM(2)#	2.758174
AM1#	.727595761418E-11	AM2#	.119359087861E 02	AM3#	.808601163281E 01
AM4#	.408457066540E 03	AM5#	.737019096240E 03	AM6#	.226061709106E 05

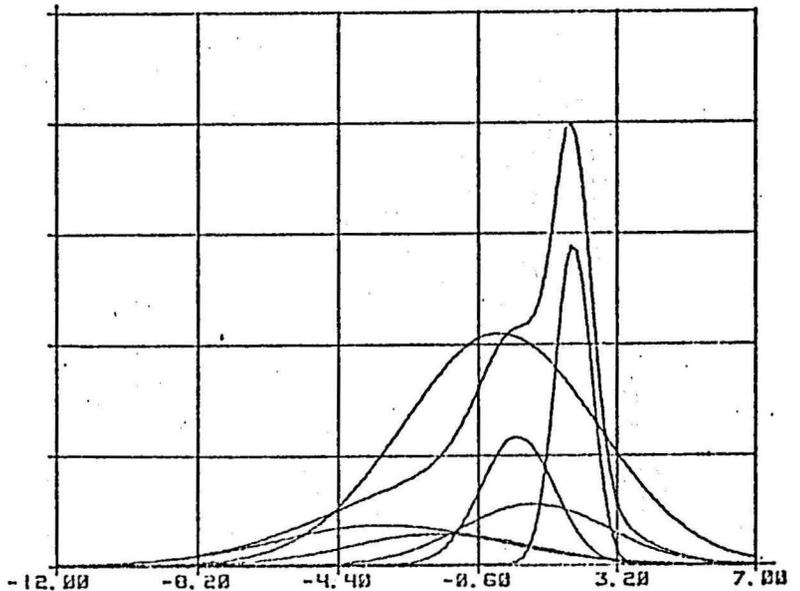
Bild 17 zeigt drei Normalverteilungen und die zugehörige Mischverteilung. Eine Zerlegung in zwei Normalverteilungen gelingt nicht.

Bild 18 zeigt erneut diese Mischverteilung, die empirische Häufigkeitsdichte einer Mischung von 1000 nach den drei Normalverteilungen erzeugten Zufallszahlen und die Komponenten, die aufgrund der Zerlegung der empirischen Momente gewonnen wurden. Die zu den zwei Komponenten gehörige Summenkurve der Verteilungsdichte stimmt, wie man sieht, sehr weitgehend mit derjenigen der Eingabewerte überein. Das Verfahren zerlegt nicht weiter.

Es ist zu bedenken, daß dem Zerlegungsverfahren nur die Momente der empirischen Häufigkeiten bekannt sind.

Erfolgloser Zerlegungsversuch einer Mischung aus fünf Normalverteilungen aufgrund der theoretischen Momente.

Bild 19



EINGEGEBENE WERTE

(GAUSS-VERTEILUNG)

AA(1)*	.200000	AMY(1)*	-3.000000	ASIGM(1)*	3.000000
AA(2)*	.130000	AMY(2)*	-1.500000	ASIGM(2)*	2.500000
AA(3)*	.210000	AMY(3)*	.500000	ASIGM(3)*	1.000000
AA(4)*	.200000	AMY(4)*	1.000000	ASIGM(4)*	2.000000
AA(5)*	.260000	AMY(5)*	2.000000	ASIGM(5)*	.500000

AUS EINGABEWERTEN BERECHNETE ZENTRALE MOMENTE DER MISCHVERTEILUNG

MY1*	.363797880709E-11	MY2*	.707159999997E 01	MY3*	-.209202210004E 02
MY4*	.212660364066E 03	MY5*	-.136592799451E 04	MY6*	.127663603036E 05

KEINE LÖSUNG

In Bild 19 ergibt die Mischung von fünf Komponenten theoretische Momente, die keine Zerlegung zulassen. Hier muß man bei der Erzeugung von Zufallszahlen bis auf 300 heruntergehen, ehe eine Zerlegung gelingt. (Bild 20)

Das Verfahren bleibt bei zwei Lösungskomponenten stehen, die man als Näherung für die empirische Dichteverteilung ansehen muß. Es wäre nun wünschenswert, das Verfahren mit größerer Blocklänge weiterlaufen zu lassen, doch war dies programmtechnisch z.Zt. nicht möglich, da die Blocklänge während eines Laufes z. Zt. nur im Vielfachen von 600 Zufallszahlen variabel ist und die Anfangszerlegung mit 600 Zufallszahlen nicht gelang.

Zerlegung aufgrund
der empirischen
Momente.

Simulation mit 300
Zufallszahlen $\alpha=10\%$.
AM2 ... AM5 sind die
verwendeten empirisch-
en biasfreien Momente.

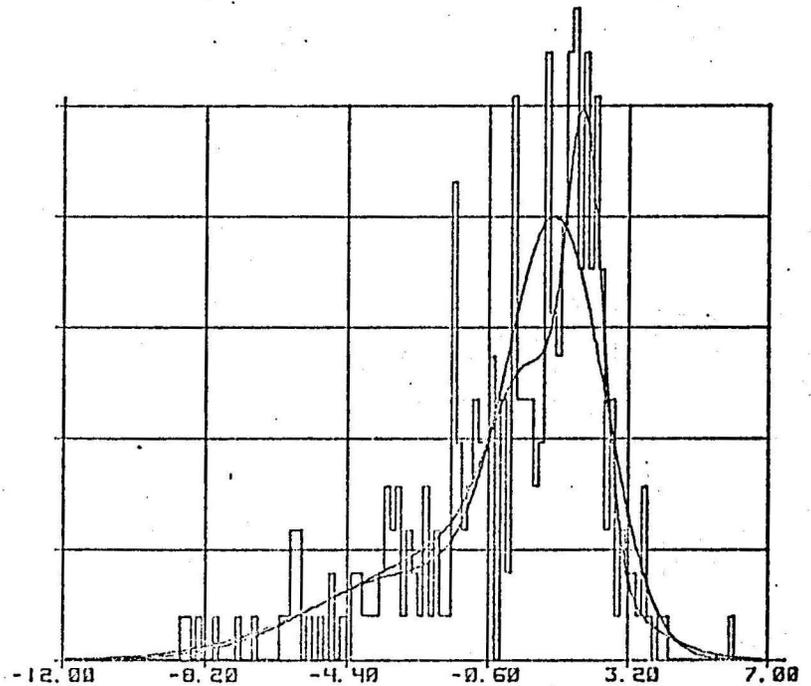


Bild 20

EINGEGEBENE WERTE

[GAUSS-VERTEILUNG]

AA[1] =	.200000	AMY[1] =	-3.000000	ASIGM[1] =	3.000000
AA[2] =	.130000	AMY[2] =	-1.500000	ASIGM[2] =	2.500000
AA[3] =	.210000	AMY[3] =	.500000	ASIGM[3] =	1.000000
AA[4] =	.200000	AMY[4] =	1.000000	ASIGM[4] =	2.000000
AA[5] =	.260000	AMY[5] =	2.000000	ASIGM[5] =	.500000

LSG: -.322220318820E 01

LOESUNGSKOMPONENTEN

BA[1] =	.266871	BMY[1] =	-2.844395	BSIGM[1] =	2.588549
BA[2] =	.733129	BMY[2] =	1.213822	BSIGM[2] =	1.313632
AM1 =	-.109139364213E-10	AM2 =	.627560135888E 01	AM3 =	-.179469156938E 02
AM4 =	.168291572035E 03	AM5 =	-.101524216277E 04	AM6 =	.846290497446E 04

Ein Beispiel bei dem 2000 Zufallszahlen als Mischung von drei Normalverteilungen erzeugt werden und je nach Zufallszahlensatz und Signifikanzniveau drei oder vier Komponentenverteilungen gefunden werden, soll den Abschluß der Approximation von Mischungen aus Normalverteilungen bilden:

In Bild 21 ist punktiert die Dichtefunktion zu sehen, nach der die gezeigten Zufallszahlen erzeugt wurden.

Signifikanzniveau $\alpha = 10\%$. Zunächst erfolgt eine Zerlegung in zwei (strichpunktiert) und dann in vier Komponenten (strichliert). Die einzelnen Komponenten sind mit dargestellt.

Zerlegung aufgrund der empirischen Momente von 2000 Zufallszahlen als Mischung von 3 Normalverteilungen.
Zufallszahlensatz 1, Signifikanzniveau $\alpha = 10\%$.
Zustand nach zweimaligem Zerlegen.

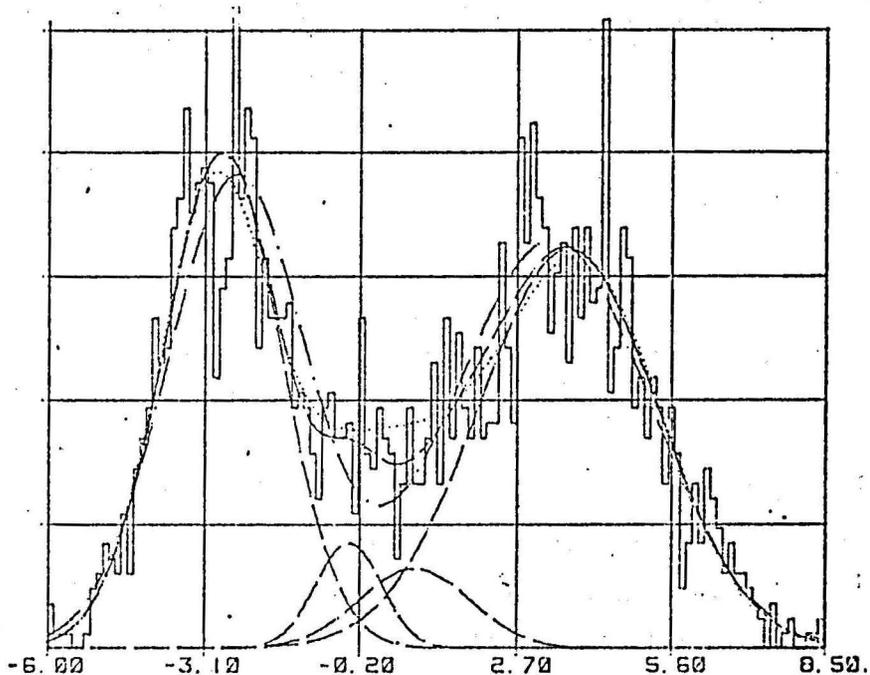


Bild 21

EINGEGEBENE WERTE

(GAUSS-VERTEILUNG)

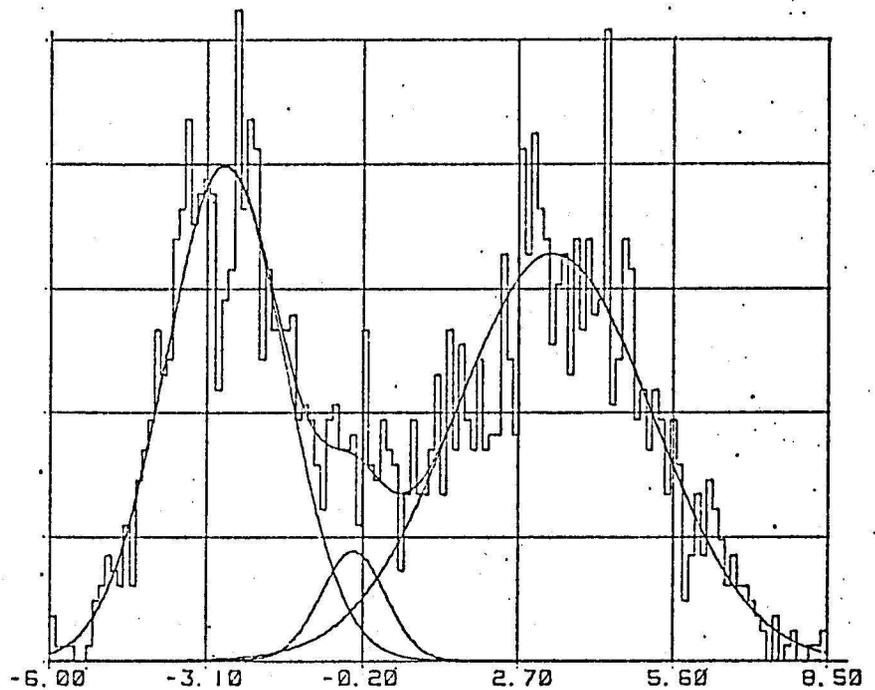
AA[1]#	.300000	AMY[1]#	-3.000000	ASIGM[1]#	1.000000
AA[2]#	.300000	AMY[2]#	.000000	ASIGM[2]#	2.000000
AA[3]#	.400000	AMY[3]#	4.000000	ASIGM[3]#	1.500000

LOESUNGSKOMPONENTEN

BA[1]#	.399651	BYM[1]#	-2.783463	BSIGM[1]#	1.093814
BA[2]#	.048057	BYM[2]#	-.382150	BSIGM[2]#	.616130
BA[3]#	.060848	BYM[3]#	.853718	BSIGM[3]#	1.032129
BA[4]#	.491443	BYM[4]#	3.676325	BSIGM[4]#	1.670766

Endergebnis von
Bild 21 nach
Zusammenlegen.

Bild 22



LOESUNGSKOMPONENTEN

BA(1)*	.398857	BMY(1)*	-2.784804	BSIGM(1)*	1.097490
BA(2)*	.049313	BMY(2)*	.374930	BSIGM(2)*	.610382
BA(3)*	.551830	BMY(3)*	3.365347	BSIGM(3)*	1.842867

Als nächster Schritt erfolgt ein Zusammenlegen der dritten und vierten Komponente. Die Entscheidung des Zusammenlegens fällt knapp. Der Betrag der Differenz $Dif\ 2 = -1,63$ liegt nur knapp unterhalb des für $\alpha = 10\%$ maßgeblichen Wertes 1,645.

Die Dichtefunktion Bild 22 stellt eine gute Annäherung an die empirischen Häufigkeiten dar. Die drei Komponenten sind ebenfalls gezeichnet.

Es ist noch ein anderer Satz Zufallszahlen nach dem gleichen Verteilungsgesetz erzeugt worden. Hier erfolgt bei $\alpha = 10\%$

die Zerlegung zunächst in zwei und dann durch Zerlegen der linken Komponente in drei Komponenten. Die empirischen Häufigkeiten und das Endergebnis sind in Bild 23 zu sehen. Wählt man jedoch $\alpha = 25\%$, so wird auch noch die rechte Komponente in zwei zerlegt, Bild 24. In beiden Fällen (Bild 23 und Bild 24) wird nicht zusammgelegt.

2000 Zufallszahlen als Mischung von 3 Normalverteilungen.
 Parameter wie Bild 21 und 22, nur andere Zufallszahlen.
 Ergebnis: Zerlegung in zwei und dann in drei Komponenten.

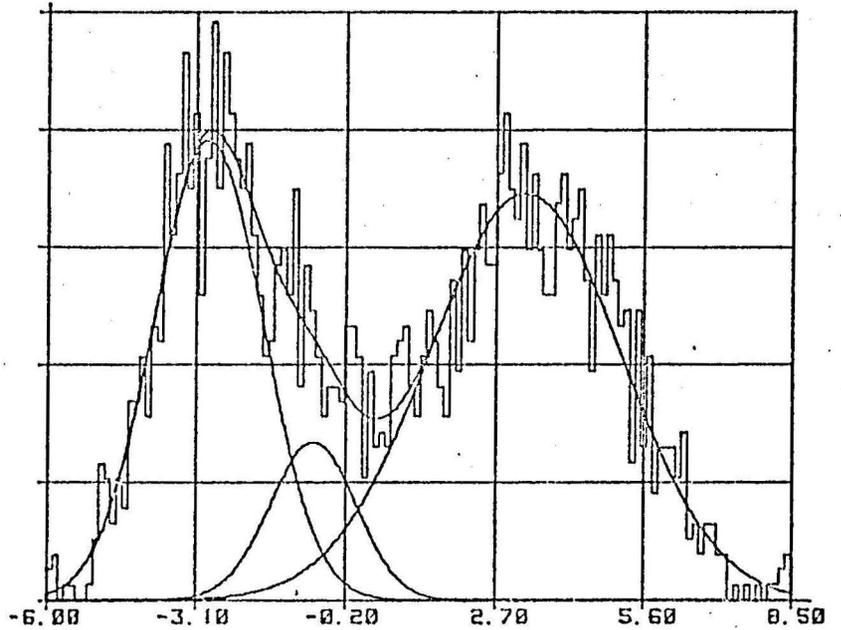


Bild 23

EINGEGEBENE WERTE

(GAUSS-VERTEILUNG)

AA[1] =	.300000	AMY[1] =	-3.000000	ASIGM[1] =	1.000000
AA[2] =	.300000	AMY[2] =	.000000	ASIGM[2] =	2.000000
AA[3] =	.400000	AMY[3] =	4.000000	ASIGM[3] =	1.500000

LOESUNGSKOMPONENTEN

BA[1] =	.352957	BMY[1] =	-2.876188	BSIGM[1] =	1.054139
BA[2] =	.095335	BMY[2] =	-.816826	BSIGM[2] =	.835517
BA[3] =	.551208	BMY[3] =	3.292713	BSIGM[3] =	1.862505

2000 Zufalls-
zahlen als Mischung
von 3 Normalvertei-
lungen.

Parameter wie in
Bild 21 und 22
außer Signifikanz-
niveau $\alpha = 25\%$.
Gleiche Zufalls-
zahlen wie in
Bild 23.

Ergebnis: Zerlegung
in zwei und dann
in vier Komponenten.

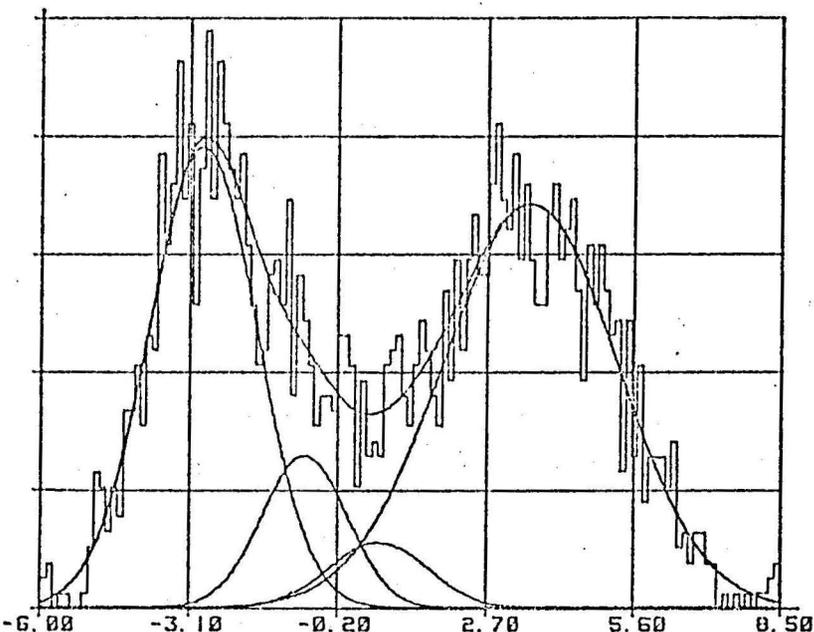


Bild 24

EINGEGEBENE WERTE					
[GAUSS-VERTEILUNG]					
AA[1]*	.300000	AMY[1]*	.3.000000	ASIGM[1]*	1.000000
AA[2]*	.300000	AMY[2]*	.000000	ASIGM[2]*	2.000000
AA[3]*	.400000	AMY[3]*	4.000000	ASIGM[3]*	1.500000
LOESUNGSKOMPONENTEN					
BA[1]*	.353635	BMY[1]*	.2.877508	BSIGM[1]*	1.050195
BA[2]*	.093250	BMY[2]*	.842749	BSIGM[2]*	.834526
BA[3]*	.047284	BMY[3]*	.612188	BSIGM[3]*	.956131
BA[4]*	.505831	BMY[4]*	3.536249	BSIGM[4]*	1.719529

8.2 Nicht-normale Mischungen

Das Verfahren findet häufig auch dann eine Zerlegung in Normalverteilungen, wenn das Verteilungsgesetz der Zufallszahlen nicht normal ist.

In Bild 25 ist eine Mischung von 6000 Zufallszahlen erzeugt worden. Der Mittelwert liegt bei 1, die Breite des Rechteckes ist 2. Die gewählten Parameter des Zerlegungsverfahrens sind: Irrtumswahrscheinlichkeit $\alpha = 25\%$ d.h. Schranke für die Varianz der Formfaktoren: 1,15; kleinste zulässige Komponente 1%.

6000 Zufallszahlen mit Rechteckverteilung.
 Signifikanzniveau $\alpha = 25\%$
 Kleinste zulässige Komponente $S = 1\%$
 Beginn der Zerlegung zunächst in zwei, dann in vier Komponenten.

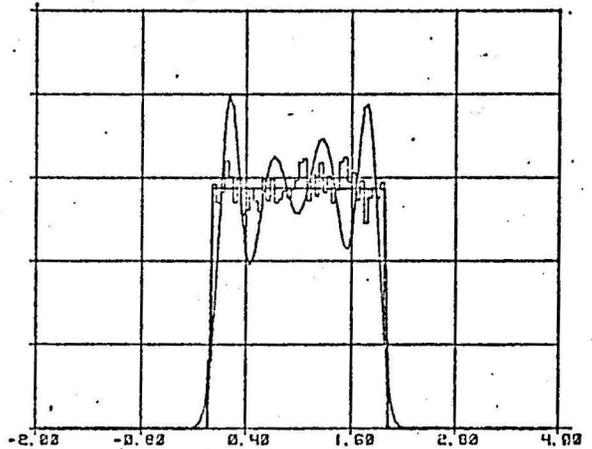


Bild 25

LOESUNGSKOMPONENTEN

BA[1] =	.486170	BMY[1] =	.487239	BSIGM[1] =	.272435
BA[2] =	.513830	BMY[2] =	1.497778	BSIGM[2] =	.277867

LOESUNGSKOMPONENTEN

BA[1] =	.193634	BMY[1] =	.204129	BSIGM[1] =	.116363
BA[2] =	.290152	BMY[2] =	.701951	BSIGM[2] =	.210200
BA[3] =	.315935	BMY[3] =	1.282439	BSIGM[3] =	.212954
BA[4] =	.200279	BMY[4] =	1.788096	BSIGM[4] =	.124152

Fortsetzung von Bild 25
Zerlegung in acht Komponenten.

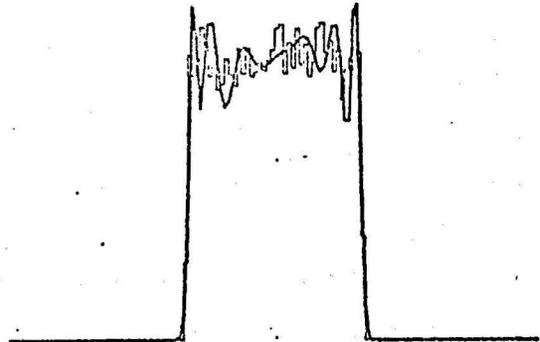


Bild 26

LOESUNGSKOMPONENTEN

BA[1] =	.055726	BMV[1] =	.074564	BSIGM[1] =	.039129
BA[2] =	.132078	BMV[2] =	.255138	BSIGM[2] =	.105883
BA[3] =	.212956	BMV[3] =	.612487	BSIGM[3] =	.176231
BA[4] =	.032989	BMV[4] =	.905201	BSIGM[4] =	.153067
BA[5] =	.129000	BMV[5] =	1.116667	BSIGM[5] =	.163832
BA[6] =	.193409	BMV[6] =	1.407140	BSIGM[6] =	.167981
BA[7] =	.121591	BMV[7] =	1.711715	BSIGM[7] =	.104832
BA[8] =	.072251	BMV[8] =	1.920302	BSIGM[8] =	.046557

Endergebnis der Zerlegung von 6000 rechteckverteilten Zufallszahlen mit Parametern wie in Bild 25.

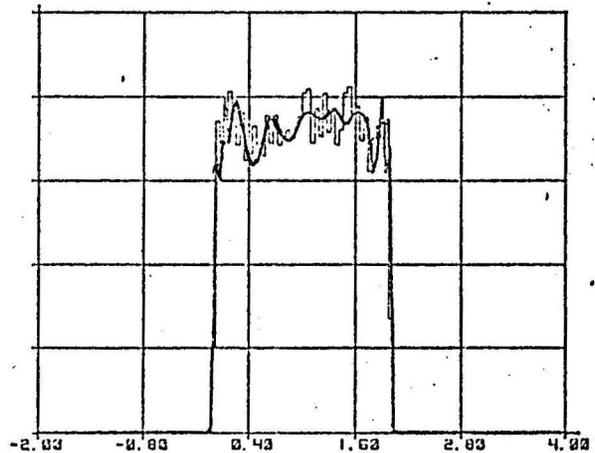


Bild 27

LOESUNGSKOMPONENTEN

BA[1] =	.016723	BMV[1] =	.027827	BSIGM[1] =	.014298
BA[2] =	.036694	BMV[2] =	.089389	BSIGM[2] =	.037335
BA[3] =	.076914	BMV[3] =	.198838	BSIGM[3] =	.081399
BA[4] =	.056496	BMV[4] =	.329600	BSIGM[4] =	.088074
BA[5] =	.212775	BMV[5] =	.609035	BSIGM[5] =	.177723
BA[6] =	.213970	BMV[6] =	1.032647	BSIGM[6] =	.186769
BA[7] =	.193424	BMV[7] =	1.411688	BSIGM[7] =	.169059
BA[8] =	.037868	BMV[8] =	1.611387	BSIGM[8] =	.077710
BA[9] =	.034943	BMV[9] =	1.757008	BSIGM[9] =	.086193
BA[10] =	.045277	BMV[10] =	1.896885	BSIGM[10] =	.041515
BA[11] =	.024916	BMV[11] =	1.969645	BSIGM[11] =	.016716

Das Verfahren zerlegt zunächst in zwei, in vier (Bild 25) und dann in 8 Komponenten (Bild 26). Nun werden die mittleren Komponenten 4 und 5 zusammengelegt. Von den jetzt vorhandenen 7 Komponenten werden die äußeren vier zerlegt (1,2,6,7), bei der Komponente 5 wird ein Zerlegungsversuch verworfen, da sich eine zu kleine Komponente ergibt. Das Verfahren endet bei 11 Komponenten, deren Summenfunktion in Bild 27 dargestellt ist und als brauchbare Näherung für die empirischen Häufigkeiten betrachtet werden kann.

In Bild 28 ist die Mischung von 6000 dreieckverteilten Zufallszahlen mit dem Mittelwert -3 und der Grundlinie 10 erzeugt worden. Das Verfahren wurde mit $\alpha = 10\%$ entsprechend einer Schranke von $1,645$ für die Varianzen der Formfaktoren gestartet.

Es werden zunächst zwei etwa gleich große Normalverteilungen gefunden, deren Mittelwerte $-4,34$ und $-1,36$ ziemlich symmetrisch zum empirischen Mittelwert der Zufallszahlen $-2,94$ liegen und die diese bereits recht gut approximieren (Bild 28).

6000 dreieckverteilte Zufallszahlen. Signifikanzniveau $\alpha = 10\%$. Beginn der Zerlegung zunächst in zwei Komponenten.

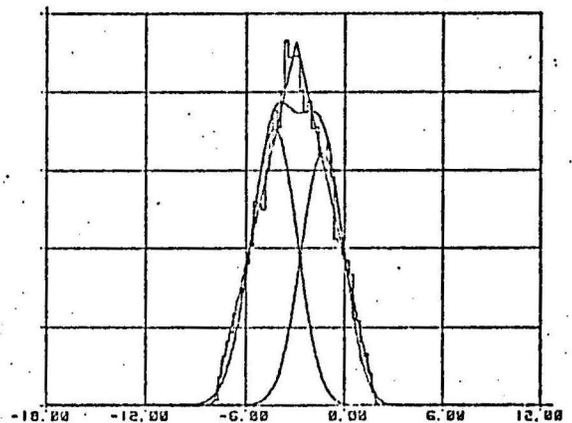


Bild 28

.....
LÖSUNGSKOMPONENTEN

BA(1) =	.537771	BM(1) =	-4.344917	BSIGM(1) =	1.403813
BA(2) =	.462229	BM(2) =	-1.363617	BSIGM(2) =	1.320824

Nach dem Versuch die erste dieser Komponenten zu zerlegen und der Zerlegung der zweiten Komponente in zwei kommt das Verfahren zum Ende (Bild 29).

Fortsetzung von Bild 28
und Endergebnis der Zerlegung von 6000 dreieck-
verteilten Zufallszahlen.

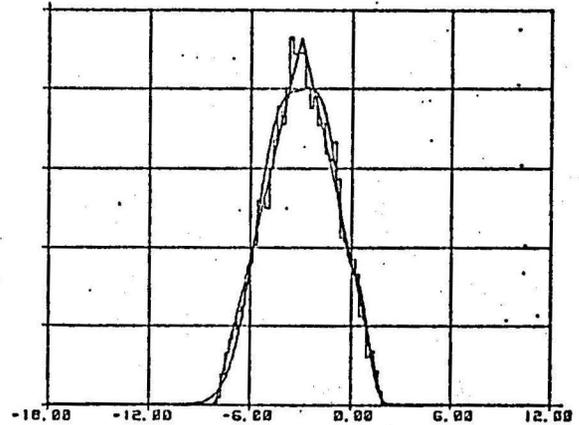


Bild 29

LOESUNGSKOMPONENTEN

BA[1] =	.539408	BMY[1] =	-4.305807	BSIGM[1] =	1.441029
BA[2] =	.411044	BMY[2] =	-1.639299	BSIGM[2] =	1.228669
BA[3] =	.049548	BMY[3] =	.596090	BSIGM[3] =	.539020

Es soll ein Beispiel gezeigt werden, bei dem eine Mischung aus drei Dreieckverteilungen mit insgesamt 2000 Zufallszahlen approximiert wird.

2000 Zufallszahlen als Mischung von drei Dreieckverteilungen.
 Signifikanzniveau $\alpha = 10\%$
 Beginn der Zerlegung in zwei (nicht dargestellt), drei (punktiert) und schließlich in 4 Normalverteilungen.
 Eingegebene Werte zur Erzeugung der Zufallszahlen. AA[] Mischungsfaktoren, AMY[] Mittelwerte, ASIGM[] halbe Grundlinien der Dreieckskomponenten.

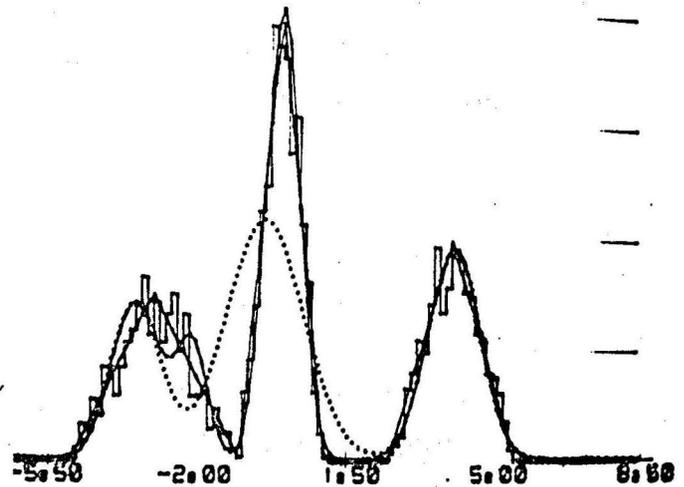


Bild 30

 EINGEGEBENE WERTE

 -- [DREIECK-VERTEILUNG]

AA[1]#	.300000	AMY[1]#	-3.000000	ASIGM[1]#	2.000000
AA[2]#	.400000	AMY[2]#	.000000	ASIGM[2]#	1.000000
AA[3]#	.300000	AMY[3]#	4.000000	ASIGM[3]#	1.500000

 LOESUNGSKOMPONENTEN

BA[1]#	.211777	BMY[1]#	-3.383793	BSIGM[1]#	.619059
BA[2]#	.085990	BMY[2]#	-2.128509	BSIGM[2]#	.410537
BA[3]#	.402232	BMY[3]#	.001161	BSIGM[3]#	.409613
BA[4]#	.300000	BMY[4]#	4.036761	BSIGM[4]#	.617675

Daten des Zerlegungsverfahrens:

Irrtumswahrscheinlichkeit $\alpha = 10 \%$

Untere Schranke für kleinste Komponente $S = 0$.

Das Ergebnis der Approximation ist in Bild 30 zu sehen.

Die empirischen Häufigkeiten bilden die Stufenkurve, durch die hindurch (besonders links deutlich zu sehen) die theoretische Dreiecksverteilung gezeichnet ist, nach der die Zufallszahlen erzeugt wurden.

Das Zerlegungsverfahren findet zunächst zwei Normalverteilungen (nicht dargestellt), von denen die rechte offenbar die rechte Dreiecksverteilung sehr gut annähert, die linke wird im nächsten Schritt erneut zerlegt, so daß jetzt eine Mischung aus drei Normalverteilungen vorliegt (punktiert). Im letzten Schritt zerlegt das Verfahren die mittlere der drei Verteilungen und kommt zum Schluß. Auf Bild 30 erkennt man, daß jetzt die rechte und die mittlere Komponente sehr gut angenähert sind, während die linke Komponente durch eine Mischung von zwei Verteilungen ebenfalls gut angenähert wird.

In einem anderen Beispiel wurden 500 Zufallszahlen nach drei Dreiecksverteilungen erzeugt.

Der Algorithmus wurde mit $\alpha = 10 \%$, $S = 0$ angewendet.

Zu den Bildern 31 - 34:

500 Zufallszahlen als Mischung von drei Dreiecksverteilungen. Eingegebene Werte zur Erzeugung der Zufallszahlen:

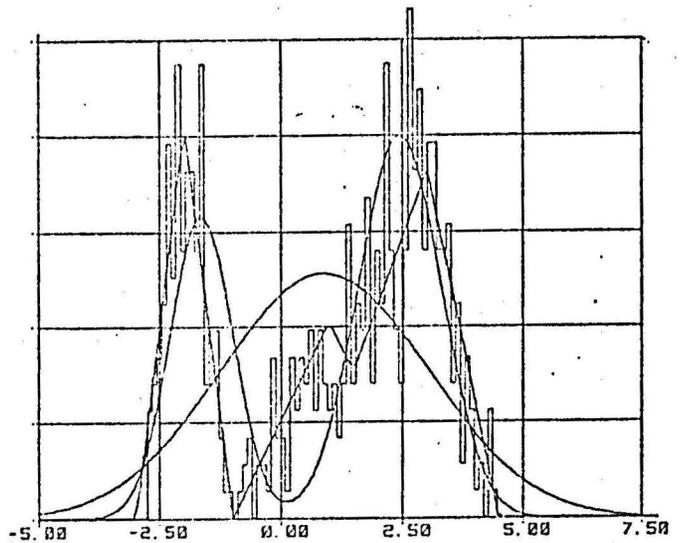
EINGEGEBENE WERTE

[DREIECK-VERTEILUNG]

AA[1]*	•300000	AMY[1]*	•2.000000	ASIGM[1]*	1.000000
AA[2]*	•300000	AMY[2]*	1.000000	ASIGM[2]*	2.000000
AA[3]*	•400000	AMY[3]*	3.000000	ASIGM[3]*	1.500000

Beginn der Zerlegung in zwei Normalverteilungen.

Bild 31



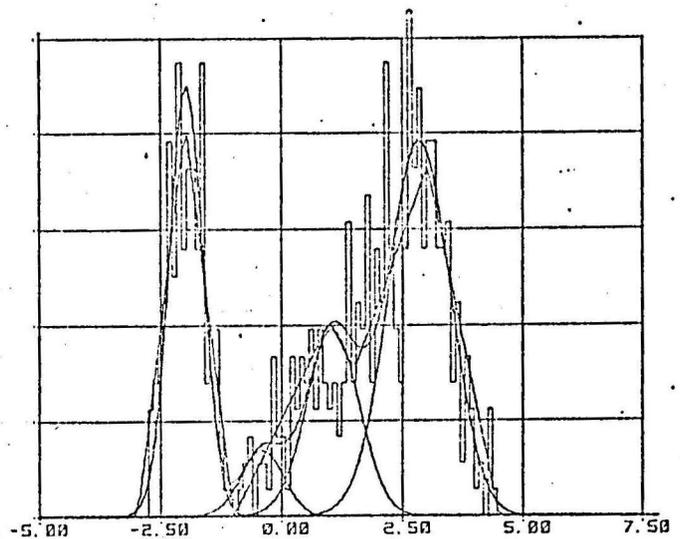
LÖSUNGSKOMPONENTEN

BA[1] =	.375179	BMU[1] =	-1.648673	BSIGM[1] =	.646340
BA[2] =	.624821	BMU[2] =	2.441062	BSIGM[2] =	.850725

Fortsetzung zu
Bild 31

Zerlegung in vier
Normalverteilungen.

Bild 32



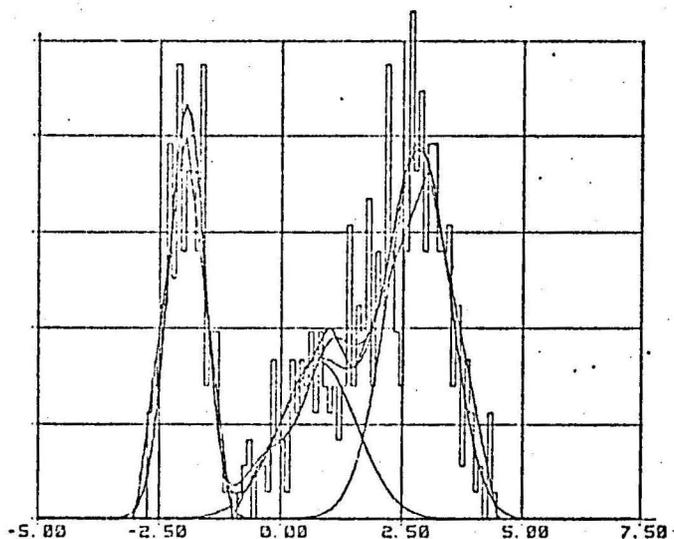
LÖSUNGSKOMPONENTEN

BA[1] =	.290058	BMU[1] =	-1.986072	BSIGM[1] =	.350728
BA[2] =	.056208	BMU[2] =	-.399043	BSIGM[2] =	.435018
BA[3] =	.195787	BMU[3] =	1.051271	BSIGM[3] =	.544283
BA[4] =	.457947	BMU[4] =	2.837365	BSIGM[4] =	.636158

Fortsetzung zu
Bild 32

Zusammenlegen von
zwei Komponenten.

Bild 33

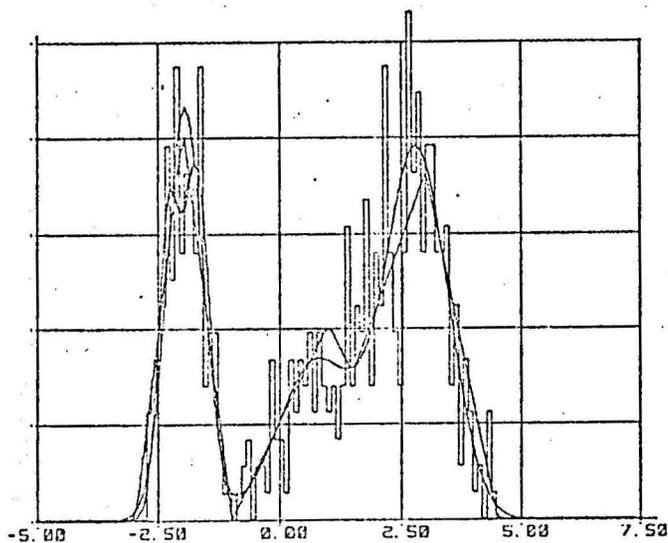


LOESUNGSKOMPONENTEN

BA[1] =	.295724	BMY[1] =	-1.965248	BSIGM[1] =	.372379
BA[2] =	.242450	BMY[2] =	.766663	BSIGM[2] =	.793405
BA[3] =	.461826	BMY[3] =	2.819188	BSIGM[3] =	.652180

Zerlegen einer
anderen Komponente.

Bild 34



LOESUNGSKOMPONENTEN

BA[1] =	.124791	BMY[1] =	-2.292270	BSIGM[1] =	.223770
BA[2] =	.172388	BMY[2] =	-1.722147	BSIGM[2] =	.263043
BA[3] =	.240022	BMY[3] =	.800667	BSIGM[3] =	.819754
BA[4] =	.462800	BMY[4] =	2.803446	BSIGM[4] =	.669930

In den Bildern 31-34 sind jeweils Häufigkeiten der Zufallszahlen, die theoretische Funktion nach der sie erzeugt wurden und die Ergebnisse der Zerlegung zu sehen, teilweise mit Komponenten. Das Vorstadium ist in jedes Bild mit aufgenommen.

Bild 31 zeigt die Zerlegung in zwei Normalverteilungen. Der Vorzustand - eine Normalverteilung - ist ebenfalls eingezeichnet. In Bild 32 ist zu erkennen, daß beide Komponenten erneut zerlegt wurden. Es sind jetzt vier Komponenten vorhanden. Die beiden mittleren Komponenten werden nun zusammengelegt (Bild 33). Die Summenkurve aus den vier Komponenten in Bild 32 ist mit eingezeichnet. Das Verfahren bricht jedoch erst ab, nachdem die linke Komponente nochmal in zwei zerlegt wurde (Bild 34). Auch hier ist die Summenkurve der drei Dichten aus Bild 33 mit eingezeichnet. Die letzte Zerlegung ist sicherlich auf den speziellen Zufallszahlensatz zurückzuführen, der im Bereich der linken Komponente ebenfalls zwei Spitzen und einen Einbruch hat.

8.3 Zerlegung unter Nebenbedingungen.

An den Zufallszahlen des letzten Beispiels wurde die Wirkung der Nebenbedingung "Zweigipfligkeit" bzw. "Existenz der Bayesgrenze" als Entscheidungskriterium erprobt. Die Bezeichnung "Nebenbedingung" ist nur für das Zerlegen korrekt. Hier muß neben der am Formfaktor erkennbaren Zerlegungswürdigkeit auch die Nebenbedingung erfüllt sein. Beim Zusammenlegen wird nur die "Nebenbedingung" abgefragt.

500 Zufallszahlen,
gleiche Eingabe wie
in Bild 31, jedoch
Nebenbedingung
Zweigipfligkeit.
Ergebnis: 3 Kompo-
nenten

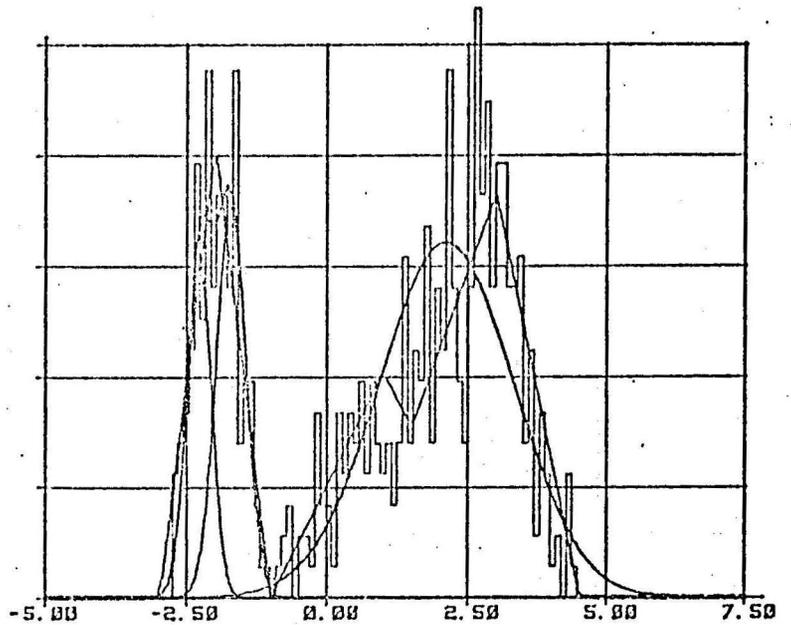


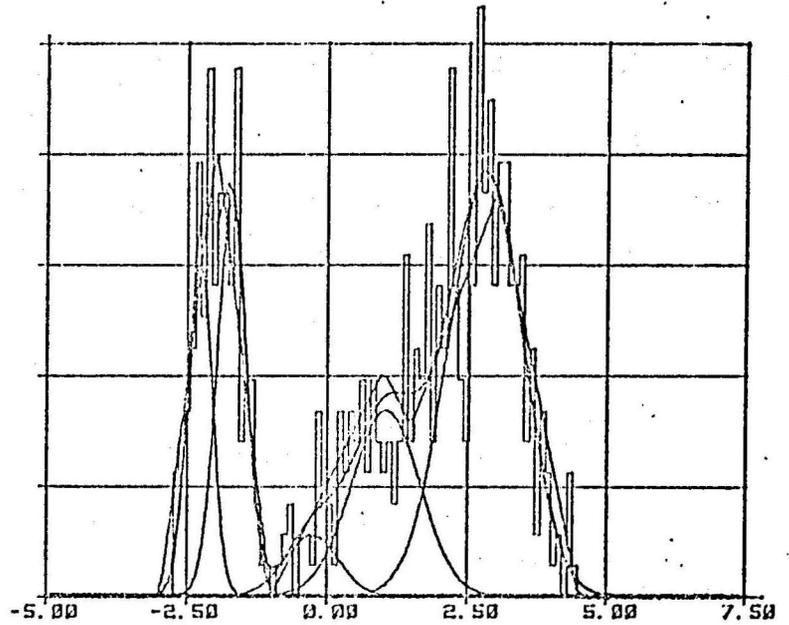
Bild 35

LÖSUNGSKOMPONENTEN					
BAC (1) *	.124791	BMY (1) *	2.292270	BSIGM (1) *	.223770
BAC (2) *	.172383	BMY (2) *	1.722147	BSIGM (2) *	.263043
BAC (3) *	.702321	BMY (3) *	2.119474	BSIGM (3) *	1.194960

Bild 35 zeigt das Endergebnis einer Approximation mit $\alpha = 10 \%$, $S = 0$ und Nebenbedingung: "Zweigipfligkeit". Es unterscheidet sich vom Endergebnis Bild 34 nur dadurch, daß die dritte und vierte Komponente zusammengelegt wurden. Der Ablauf stimmt ebenfalls völlig mit dem in Bild 31 - 34 überein, nur endet er erst einen Schritt später nach dem Zusammenlegen der dritten und vierten Komponente. Dies geschieht, weil die schwache Einsattelung nicht mehr als Zweigipfligkeit erkannt wird. Die einprogrammierte notwendige Bedingung für die Zweigipfligkeit ist nur bei deutlichen Einsattelungen erfüllt. Beim Blick auf die empirischen Häufigkeiten erscheint das erzielte Endergebnis gar nicht so ungünstig.

500 Zufallszahlen,
gleiche Eingabe wie
in Bild 31, jedoch
Nebenbedingung Bay-
esgrenze existiert.
Ergebnis: 5 Kompo-
nenten.

Bild 36



..... LOESUNGSKOMPONENTEN

BA[1] =	.124841	BMY[1] =	-2.292215	BSIGM[1] =	.223785
BA[2] =	.172059	BMY[2] =	-1.723834	BSIGM[2] =	.261415
BA[3] =	.048454	BMY[3] =	.266572	BSIGM[3] =	.475756
BA[4] =	.191459	BMY[4] =	1.046425	BSIGM[4] =	.617867
BA[5] =	.463188	BMY[5] =	2.810984	BSIGM[5] =	.661739

In Bild 36 ist das Ergebnis der Approximation des gleichen Datenmaterials wie bei Bild 31-34 und 35 mit $\alpha = 10\%$, $S = 0$ und Nebenbedingung: "Bayesgrenze existiert" zu sehen. Der Unterschied im Verlauf des Algorithmus mit und ohne Nebenbedingung besteht darin, daß durch sie das Zusammenlegen verhindert wurde, sodaß am Schluß 5 Komponenten vorhanden sind.

8.4 Approximation gemessener Wahrscheinlichkeitsdichten.

Die Daten entstammen Messungen von Herrn Prof. Dr. D. Wolf, Frankfurt /45/. Einige Ergebnisse der durchgeführten Zerlegungen sollen zur Illustration betrachtet werden:

GES. ZAHL. DER ZZ: 998800.

EMPIRISCHE NULLPUNKTSMOMENTE DER MISCHVERTEILUNG

A1 = .262126708864E 02	A2 = .724557549983E 03	A3 = .209422632980E 05
A4 = .628846232761E 06	A5 = .195177002756E 08	A6 = .623579117047E 09

EMPIRISCHE ZENTRALE MOMENTE DER MISCHVERTEILUNG MIT BIAS

MG1 = .000000000000E 00	MG2 = .374534349851E 02	MG3 = -.138344275951E 02
MG4 = .377826132584E 04	MG5 = -.342859545898E 04	MG6 = .537440687500E 06

BIASFREIE EMPIRISCHE ZENTRALE MOMENTE DER MISCHVERTEILUNG

MK4 = .377826803040E 04	MK2 = .374534724834E 02	MK3 = -.138344691484E 02
	MK5 = -.342860743496E 04	MK6 = .537441790844E 06

NACH DER SHEPPARDSCHEN KORREKTUR

MK4 = .375957046084E 04	MK2 = .373701391502E 02	MK3 = -.138344691484E 02
	MK5 = -.341707871066E 04	MK6 = .532735318638E 06

998 800 Amplitudenwerte in 512 Klassen eingeordnet. Approximation durch eine Normalverteilung.

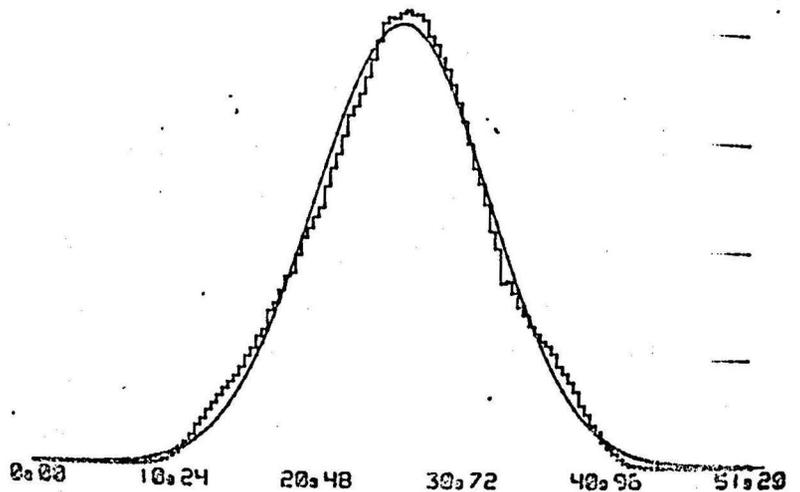


Bild 37

LOESUNGSKOMPONENTEN

BA[1] = 1.000000 BMY[1] = 26.212671 BSIGM[1] = 6.113112

Fortsetzung zu
Bild 37.

Approximation
durch zwei Nor-
malverteilungen.

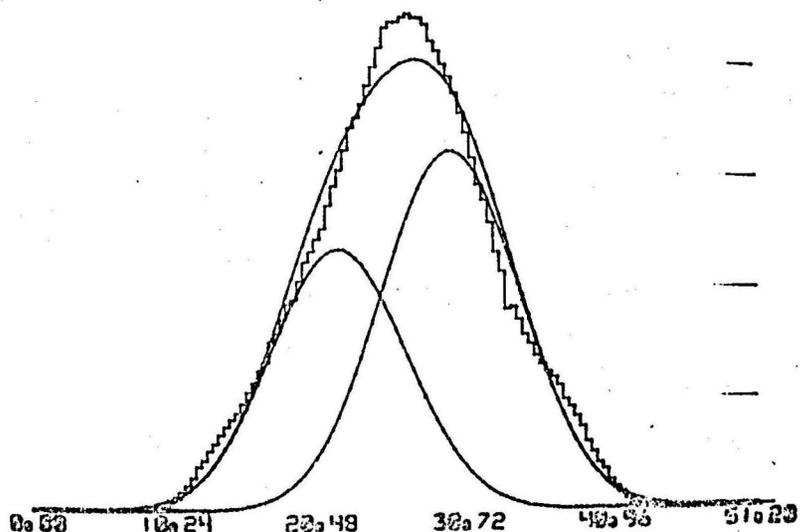


Bild 38

LOESUNGSKOMPONENTEN

BA(1)*	.414310	BMY(1)*	21.674567	BSIGM(1)*	4.743777
BA(2)*	.585690	BMY(2)*	29.422868	BSIGM(2)*	4.797199

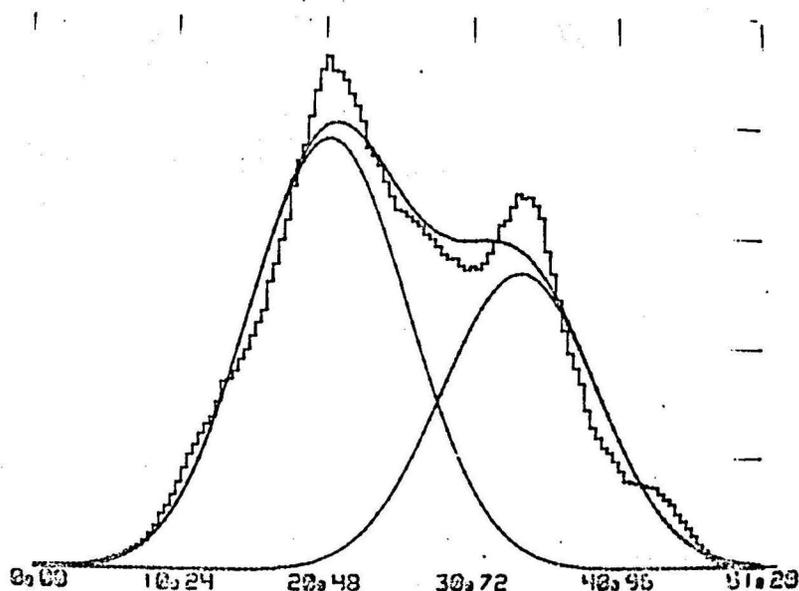
DNH* .276356793263E+01 DAH* .173944140561E+01

Das erste Beispiel zeigt eine Zerlegung in zwei Normalverteilungen. Es ist insofern bemerkenswert, als die Einzelverteilung Bild 37, die nur in zwei Momenten mit der gemessenen übereinstimmt, im Gegensatz zur Mischverteilung Bild 38, die in 5 Momenten übereinstimmt, eine bessere Anpassung liefert als letztere. Dies folgt durch Augenschein und auch nach dem Test von Kolmogoroff ($DAH < DNH$). Auf den Test wird im Rahmen dieses Berichtes nicht eingegangen. Dieser Fall ist eine Ausnahme.

Ein weiteres Beispiel der Zerlegung in zwei Komponenten zeigt Bild 39. Hier ist eine deutliche Verbesserung der Approximation mit zwei gegen eine Komponente zu sehen (DAH > DNH). Ob eine weitere Zerlegung durch Veränderung des Signifikanzniveaus möglich ist, wurde nicht überprüft. Dieses Beispiel ist insofern interessant, als hier die Nebenbedingung Zweigipfligkeit zu einem Verwerfen der Zerlegung geführt hätte, obgleich die empirische Funktion zwei deutliche Maxima besitzt.

999 998 Amplitudenwerte in 512 Klassen eingeordnet. Approximation durch zwei Normalverteilungen.

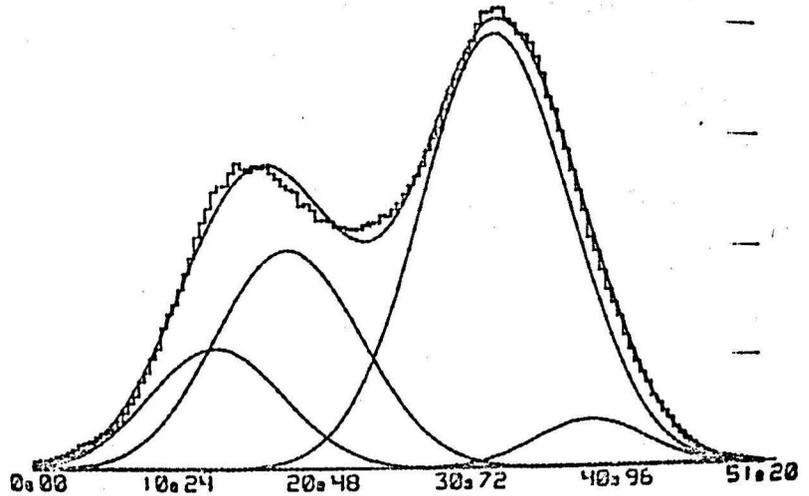
Bild 39



```
LOESUNGSKOMPONENTEN
-----
BA[ 1] = .594970      BMY[ 1] = 20.653141      BSIGM[ 1] = 5.483615
BA[ 2] = .405030      BMY[ 2] = 33.913333      BSIGM[ 2] = 5.405852
-----
DNH = .169509334936E-01      DAH = .619392590179E-01
```

998 041 Amplitudenwerte in 512 Klassen eingeordnet. Approximation durch vier Normalverteilungen.

Bild 40



... LOESUNGSKOMPONENTEN

BA(1) =	.138280	BMY(1) =	12.950937	BSIGM(1) =	4.667304
BA(2) =	.268979	BMY(2) =	18.025023	BSIGM(2) =	5.041034
BA(3) =	.547129	BMY(3) =	32.473299	BSIGM(3) =	5.198083
BA(4) =	.045611	BMY(4) =	38.991502	BSIGM(4) =	3.935587

Die meisten Verteilungsfunktionen ließen sich jedoch sehr gut annähern z.B., die folgende durch 4 Komponenten.

(Bild 40)

Ebensogut ist die Annäherung in Bild 41 durch 10 Komponenten. Hier ist die Verbesserung ($DNH < DAH$) gegenüber 8 Komponenten nur noch gering.

Dieses Beispiel wurde noch mit einer Programmversion ohne Schwelle S für die kleinste Komponente und ohne Sortieren der Komponenten nach dem Mittelwert gerechnet. Streicht man in Bild 41 z.B. einfach alle Komponenten mit einem Anteil $BA [i] < 0,5 \%$, so verbleiben sieben Komponenten, deren Mittelwerte außerdem richtig geordnet sind. Am Bild der Funktion dürfte sich nichts ändern.

999 999 Amplitudenwerte in 512 Klassen eingeordnet.
Approximation durch 10 (7) Normalverteilungen.

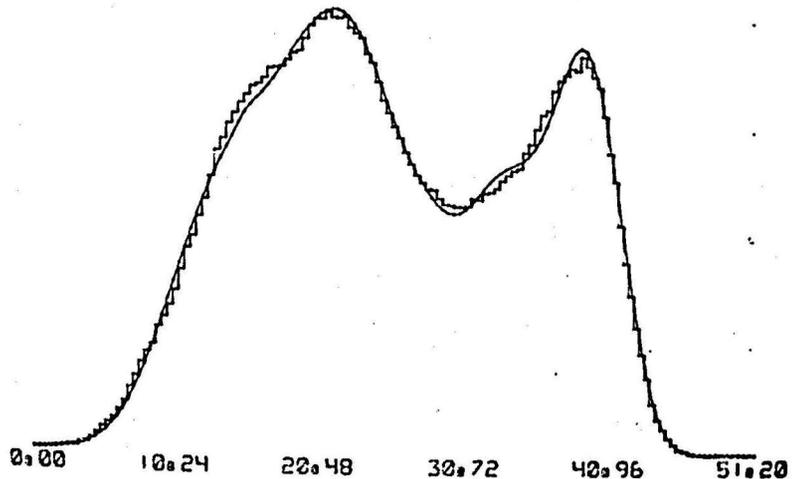


Bild 41

LOESUNGSKOMPONENTEN

BA[1]#	.142542	BMY[1]#	12.520966	BSIGM[1]#	3.421713
BA[2]#	.292837	BMY[2]#	19.499500	BSIGM[2]#	4.344391
BA[3]#	.001317	BMY[3]#	14.967146	BSIGM[3]#	.729148
BA[4]#	.002017	BMY[4]#	17.013415	BSIGM[4]#	2.303795
BA[5]#	.184265	BMY[5]#	24.177128	BSIGM[5]#	4.062094
BA[6]#	.008382	BMY[6]#	27.832022	BSIGM[6]#	1.611093
BA[7]#	.001095	BMY[7]#	27.404692	BSIGM[7]#	.445205
BA[8]#	.185981	BMY[8]#	33.525407	BSIGM[8]#	3.221282
BA[9]#	.162834	BMY[9]#	38.936504	BSIGM[9]#	2.184021
BA[10]#	.018733	BMY[10]#	41.978412	BSIGM[10]#	1.577955

DNH# .107555063913E+01

DAH# .110487381622E+01

9. Abschluß

Es ist ein heuristisches Verfahren zur Wahrscheinlichkeitsdichteapproximation entwickelt worden. Als Momentenverfahren benötigt es keine Speicherung der Einzelwerte. Die Darstellung einer Verteilung als Mischung - und nicht etwa durch eine Reihe - hat den Vorteil der leichteren Deutbarkeit. Bei der Zeichenerkennung stehen die Komponenten für Äquivalenzklassen.

Zur Steuerung des Verfahrens wird ein Signifikanztest verwendet /48/. Im Falle, daß das klassifizierte Beobachtungsmaterial gespeichert wird (Beispiele unter 8.4 S. 113 f.) wurde auch die Güte der Approximation in jedem Teilschritt mit einem Signifikanztest überprüft (Test von Kolmogoroff). Dadurch läßt sich erkennen, welche Zerlegungsschritte einen besonderen Zuwachs an Signifikanz liefern.

In einer neueren Veröffentlichung (April 1973) wurde das Prinzip, einen Signifikanztest zur Steuerung und Beurteilung der Klassifizierung heranzuziehen, ebenfalls beschrieben /15/. Es wurde vom Verfasser dieses Berichtes jedoch ohne Kenntnis der Arbeiten des Autors von /15/ benutzt. So finden sich Angaben über die Signifikanz der Approximation in den Ergebnissen, die im März 1973 dem Auftraggeber /45/ zugesandt wurden.

9.1 Leistungsfähigkeit und Grenzen des Verfahrens.

Die Aufgabe, in einer multimodalen Mischung auch alle enthaltenen Komponenten zu finden, wird nicht immer gelöst (Bild 16). Doch sind auch sehr gute Ergebnisse zu verzeichnen (Bild 30). Weitere Untersuchungen der Ursache dieses unterschiedlichen Verhaltens wären wünschenswert.

Bei der adaptiven Approximation von Wahrscheinlichkeitsdichten, insbesondere auch wenn sie nicht äußerlich erkennbar aus mehreren Anteilen bestehen, ist das Verfahren recht leistungsfähig. Bilder 25 - 29 und 37 - 41.

9.2 Nebenergebnisse

Die umfangreichen Untersuchungen führten zu einer Reihe von Nebenergebnissen. Dazu gehört eine Schranke für die Wurzeln der Gleichung neunten Grades von K. Pearson /46/, die Erkenntnis über das Entstehen einer Zweideutigkeit und eine heuristische Entscheidungsregel zur Beseitigung ihrer Auswirkung. Pearsons Untersuchungen über den Symmetriefall wurden erweitert, seine Ungleichungen zwischen den Formfaktoren ergänzt, eine Unterscheidung von 2 Symmetrietypen anhand der Formfaktoren vorgenommen.

Das Pearsonverfahren wurde weiter optimiert durch Anwendung der Sturm'schen Kette auf das Pearsonpolynom neunten Grades. Es wurde mit Fordecap-Programmen geklärt, daß man einen besonders einfachen numerischen Algorithmus erhält, wenn man voraussetzt, daß sich die Funktionen der Kette um 1 im Grad unterscheiden. Diese Bedingung ist nur bei einer Nullmenge von Fällen nicht erfüllt. Um auch hier kein fehlerhaftes Ergebnis zu erhalten, wird der Algorithmus in diesem Fall verlassen.

Gegenüber /16/ wurde eine Transformation verbessert. Dadurch benötigt man zur Bestimmung der Lage der Gipfel und derjenigen der Bayesgrenzen nur ein Diagramm. Eine dem Diagramm entsprechende numerische Behandlung der Frage nach Zweigipfligkeit und Existenz von Bayesgrenzen läßt sich zweckmäßig auch gemeinsam durchführen. Eine Ersatzfunktion zur Abschätzung der Gipfelzahl /47/ und eine Ungleichung in Verbindung mit dem Diskriminantenkriterium liefern enge notwendige Bedingungen für die Zweigipfligkeit.

9.3 Erweiterungen

Ein dynamisches Signifikanzniveau dürfte eine weitere Verbesserung des Verfahrens ermöglichen. (7.2.3 S. 91)

Das Grundprinzip des Approximationsverfahrens könnte bei anderen Verteilungstypen angewendet werden, wenn man

- a) ein Zerlegungsverfahren in zwei Komponenten,
- b) einen Signifikanztest über die Zugehörigkeit zum gewünschten Verteilungstyp entwickelt.

Zerlegungsverfahren aufgrund der Momente existieren bereits bei einigen Verteilungstypen /7/, /8/.

Durch eine geeignete -z. B. exponentielle- Alterung der Momente könnte man das Verfahren auch für einen zeitveränderlichen Prozeß anwenden. /48/

Anhang I

Symmetrische Mischverteilungen

Symmetrie - erkennbar am Verschwinden der ungeraden Momente - kommt bei empirischen Daten nur mit Wahrscheinlichkeit 0 vor.

Theoretisch sind bei einer Mischung von zwei Komponenten zwei Fälle denkbar:

Fall 1: $z_1 = z_2 = 0,5$, $\gamma_1 = -\gamma_2$, $\sigma_1 = \sigma_2$
gleichzeitig (sehr selten)

Fall 2:
 $\gamma_1 = \gamma_2 = 0$

Formt man die Gleichung für das dritte Moment (16) unter Verwendung von (14)

$$\mu_1 = z_1 \gamma_1 + z_2 \gamma_2 = 0$$

um, so kommt

$$\mu_3 = z_1 \gamma_1 \left[(\gamma_1^2 - \gamma_2^2) + 3(\sigma_1^2 - \sigma_2^2) \right] \quad (a1)$$

Schließt man den trivialen Fall $z_1 = 0$ (oder $z_2 = 0$) aus, dann kann μ_3 verschwinden, wenn

Fall 1: $\gamma_1^2 = \gamma_2^2$ und gleichzeitig $\sigma_1^2 = \sigma_2^2$

Da gelten soll $\gamma_2 \leq 0 \leq \gamma_1$ und $\sigma_1 > 0$, $\sigma_2 > 0$

folgt daraus

$\gamma_1 = -\gamma_2 = \gamma$ und $\sigma_1 = \sigma_2 = \sigma$ und wegen (20)

$z_1 = \frac{-\gamma_2}{\gamma_1 - \gamma_2}$, $z_2 = \frac{\gamma_1}{\gamma_1 - \gamma_2}$ auch

$z_1 = 0,5$ und $z_2 = 0,5$.

Die Momente lauten in diesem Fall

$$\begin{aligned}\mu_1 &= 0 \\ \mu_2 &= \gamma^2 + \sigma^2\end{aligned}\quad (\text{a2})$$

$$\begin{aligned}\mu_3 &= 0 \\ \mu_4 &= \gamma^4 + 6\gamma^2\sigma^2 + 3\sigma^4 \\ \mu_5 &= 0\end{aligned}\quad (\text{a3})$$

Fall 2:

$\gamma_1 = 0$ und $\gamma_2 = 0$ bedeutet, daß die Mittelwerte beider Komponenten gleich sind. Dies ist bei beliebigen Mischungsfaktoren

$$z_1 = \frac{0}{0} \quad \text{und} \quad z_2 = \frac{0}{0} \quad \text{möglich.}$$

Die Gleichungen für die Momente werden aus (13) bis (19) mit $\gamma_1 = \gamma_2 = 0$

$$\mu_0 = 1 \quad \text{definitionsgemäß}$$

$$\mu_1 = 0 \quad \text{definitionsgemäß}$$

$$\mu_2 = z_1\sigma_1^2 + z_2\sigma_2^2 = z_1(\sigma_1^2 - \sigma_2^2) + \sigma_2^2$$

$$\mu_3 = 0$$

$$\mu_4 = 3z_1\sigma_1^4 + 3z_2\sigma_2^4 = 3z_1(\sigma_1^4 - \sigma_2^4) + 3\sigma_2^4$$

$$\mu_5 = 0$$

$$\mu_6 = 15z_1\sigma_1^6 + 15z_2\sigma_2^6 \quad (\text{a4})$$

Berechnung der Parameter der Komponenten aus den Momenten:

Fall 1:

Es stehen 2 Gleichungen zur Bestimmung der zwei Unbekannten γ und σ zur Verfügung.

Fall 2:

In Fall 2 sind drei Unbekannte zu bestimmen: z_1, σ_1, σ_2 .
Dafür kommt man mit den Momenten $\mu_2 \dots \mu_5$ nicht aus, da nur zwei nicht identisch verschwinden.

Lösungen

Pearson schreibt:
$$\xi_1 = \frac{\mu_4 - 3 \mu_2^2}{3 \mu_2^2} \quad (a5)$$

$$\xi_2 = \frac{\mu_6 - 5 \mu_2 \mu_4}{5 \mu_2^3} \quad (a6)$$

und nennt ξ_1 Exzess und ξ_2 Defekt.

Er setzt voraus, daß bekannt ist, welcher Symmetriefall vorliegt (z.B. aus der Anschauung) und findet neben Lösungsformeln in beiden Fällen auch Bedingungen für ξ_1 und ξ_2 , die erfüllt sein müssen, damit eine Lösung existiert.

Wir haben diese Untersuchung etwas weitergeführt, insbesondere die Voraussetzung der Kenntnis des Typs der Symmetrie fallengelassen und eine Bedingung zwischen ξ_1 und ξ_2 hergeleitet, an der man erkennen kann, welcher Symmetriefall vorliegt.

Die von K. Pearson abgeleiteten Ungleichungen zwischen ξ_1 und ξ_2 im Fall 2 wurden (nach Beseitigung eines Vorzeichenfehlers) in ein Diagramm (Bild A1) eingetragen.

Schließlich wurde überprüft, wie sich die für den unsymmetrischen Fall abgeleitete Pearsongleichung (21) bei Symmetrie verhält.

Zerlegung im Symmetriefall 1:

Aus (a2) und (a3) folgt sofort

$$3 \mu_2^2 - \mu_4 = 2 \gamma^4 \quad (a7)$$

$$\gamma_{1,2} = \pm \sqrt[4]{\frac{3 \mu_2^2 - \mu_4}{2}} \quad (a8)$$

Aus Gl. (a2)

$$\sigma^2 = \mu_2 - \gamma^2 = \mu_2 - \sqrt{\frac{3\mu_2^2 - \mu_4}{2}} \quad (a9)$$

Durch Einsetzen von (a2) und (a7) in (a5) wird

$$-\xi_1 = \frac{+2\gamma^4}{3\gamma^4 + 6\gamma^2\sigma^2 + 3\sigma^4} \quad (a10)$$

Da $\sigma > 0 \rightarrow \xi_1 > -2/3$

weiter, da $\gamma^2 > 0 \rightarrow \xi_1 < 0$

(nur die letzte Bedingung findet sich sinngemäß bei K. Pearson /5/ S. 21)

Bedingung zwischen ξ_1 und ξ_2 im Symmetriefall 1 :

Mit (19) und $\gamma_1 = -\gamma_2 = \gamma$, $\sigma_1 = \sigma_2 = \sigma$

$z_1 = z_2 = 0,5$ wird:

$$\mu_6 = \gamma^6 + 15\gamma^4\sigma^2 + 45\gamma^2\sigma^4 + 15\sigma^6 \quad (a11)$$

ersetzt man in (a11) nach (a2)

$\sigma^2 = \mu_2 - \gamma^2$, so ergibt sich:

$$\mu_6 = 16\gamma^6 - 30\mu_2\gamma^4 + 15\mu_2^3 \quad (a12)$$

Nun folgt aus (a8) und (a5)

$$\gamma^2 = \mu_2 \sqrt{\frac{3(-\xi_1)}{2}}$$

und durch Einsetzen in (a12)

$$\mu_6 = 3 \mu_2^3 \left(8 (-\varepsilon_1) \sqrt{\frac{3(-\varepsilon_1)}{2}} - 15 (-\varepsilon_1) + 5 \right) \quad (\text{a13})$$

Nach (a6) ist

$$-\varepsilon_2 = \frac{\mu_4}{\mu_2^2} - \frac{\mu_6}{5 \mu_2^3}$$

worin mit (a5) $\frac{\mu_4}{\mu_2^2} = 3(\varepsilon_1 + 1)$

während $\frac{\mu_6}{5 \mu_2^3}$ mit (a13) nach kurzer Zwischenrechnung ergibt:

$$\boxed{\varepsilon_2 = 6(\varepsilon_1) \left(1 - \sqrt{0,96(-\varepsilon_1)} \right)} \quad (\text{a 14})$$

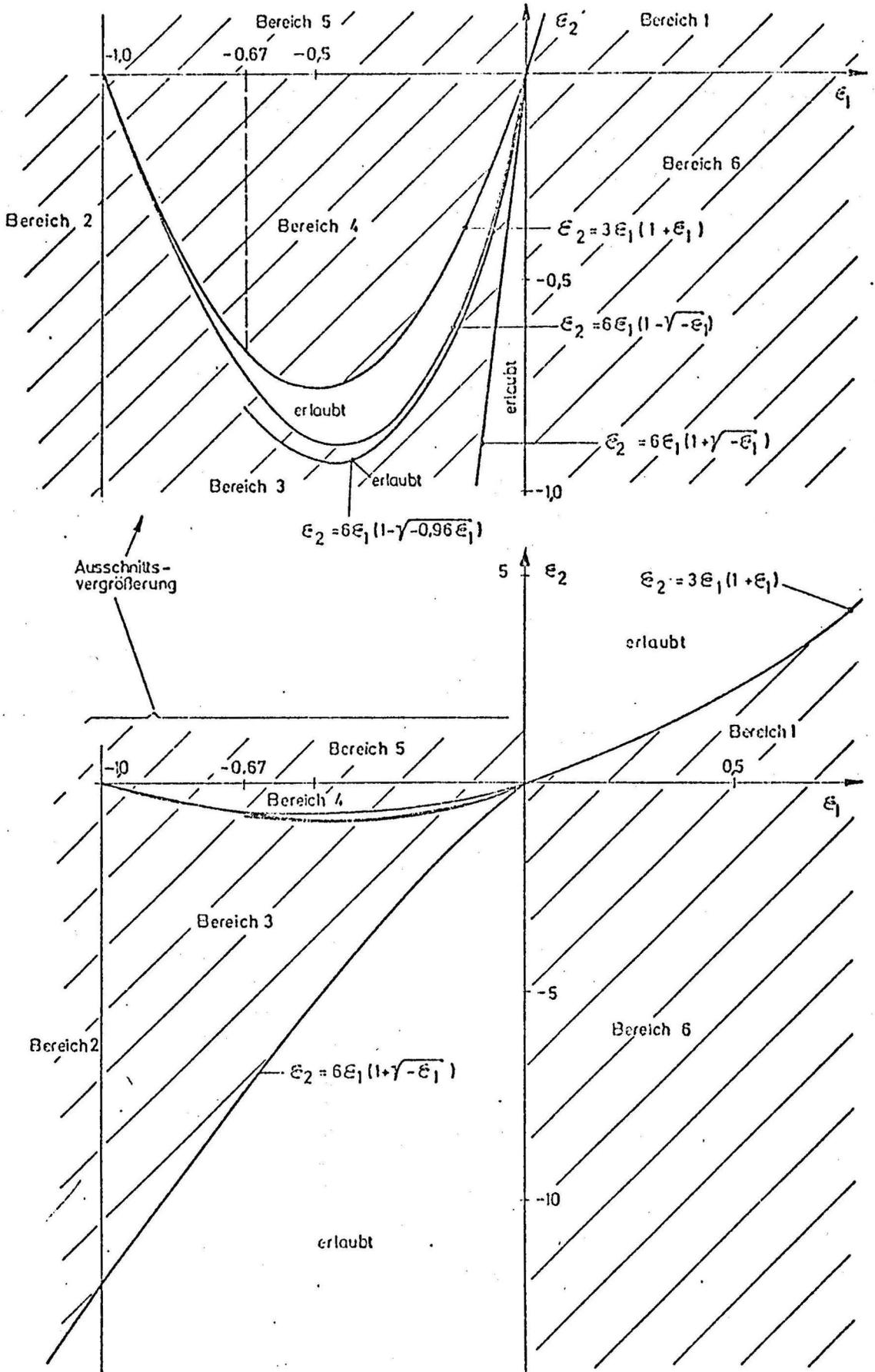


Bild A1: Abhängigkeit des Defektes ϵ_2 vom Exzess ϵ_1 bei symmetrischen Mischverteilungen

Im Diagramm Bild A1 erkennen wir, daß der Symmetriefall 1 auf einer Kurve $\xi_2 = f(\xi_1)$ vorliegt, während Symmetriefall 2 durch Bereiche $g_i(\xi_1) < \xi_2 < h_i(\xi_1)$ gekennzeichnet ist. Das heißt, die Ereignisse "Es liegt ein Symmetriefall 1 vor" bilden eine Nullmenge in der Zahl der Ereignisse "Es liegt ein Symmetriefall vor".

Bedingungen zwischen ξ_1 und ξ_2 im Symmetriefall 2 :

Es war $\gamma_1 = \gamma_2 = 0$

K. Pearson leitet her: /5/ (S. 24)

$$\sigma_{1,2}^2 = \mu_2 \frac{\xi_2 \pm \sqrt{(\xi_2 - 6\xi_1)^2 + 36\xi_1^3}}{6\xi_1}$$

$$z_1 = \frac{\mu_2 - \sigma_2^2}{\sigma_1^2 - \sigma_2^2} ; \quad z_2 = \frac{\sigma_1^2 - \mu_2}{\sigma_1^2 - \sigma_2^2}$$

Es müssen folgende Ungleichungen zwischen den ξ_1 und ξ_2 gelten; damit obige Lösungen existieren:

1) ξ_1 und ξ_2 müssen gleiches Vorzeichen haben.

2) Wenn $\xi_1 > 0, \xi_2 > 0$

dann muß gelten $\xi_2 > 3\xi_1(1 + \xi_1)$

3) Wenn $\xi_1 < 0, \xi_2 < 0$

dann muß sein $\xi_2 < 6\xi_1(1 + \sqrt{-\xi_1})$

Zusätzlich existiert eine Lösung, wenn $-1 < \xi_1 < 0$

und $6\xi_1(1 - \sqrt{-\xi_1}) \leq \xi_2 \leq 3\xi_1(1 + \xi_1) < 0$.

Dies ist vereinfacht die Aussage einer Reihe von Ungln., die K. Pearson aufstellte /5/ S. 25. Diese Ungln. wurden zur besseren Übersicht in das bereits erwähnte Bild eingetragen.

Verhalten der Gleichung (21) in den Symmetriefällen:

Die Symmetriefälle sind erkennbar am Verschwinden der Momente μ_3 und μ_5 . Das Pearsonverfahren für den unsymmetrischen Fall kann nicht angewendet werden, schon deshalb nicht, weil das Gl'n-System im Symmetriefall 2 unterbestimmt ist.

Für den Symmetriefall 1 reicht die Zahl der Bestimmungsgl'n aus und das Verfahren mit Gl. (21) liefert interessanterweise eine richtige und eine verkehrte, das Moment μ_4 nicht reproduzierende Lösung.

Wegen $\mu_3 = 0$, $\mu_5 = 0$, $\lambda_5 = 0$ wird Gl (21) nach Kürzen durch p_2^3 zu

$$24 p_2^6 - 28 \lambda_4 p_2^4 + 10 \lambda_4^2 p_2^2 - \lambda_4^3 = 0$$

Geben wir mit (a2) und (a3)

$$\begin{aligned} \mu_2 &= \tilde{\gamma}^2 + \tilde{\sigma}^2 \\ \mu_4 &= \tilde{\gamma}^4 + 6 \tilde{\gamma}^2 \tilde{\sigma}^2 + 3 \tilde{\sigma}^4 \\ \lambda_4 &= 6 \tilde{\gamma}^4 \end{aligned} \quad \text{ein,}$$

so wird $p_2^6 - 7 \tilde{\gamma}^4 p_2^4 + 15 \tilde{\gamma}^8 p_2^2 - 9 \tilde{\gamma}^{16} = 0$ und man errät leicht als Wurzeln

$$p_2^2 = \tilde{\gamma}^4 \quad (\text{a15})$$

$$p_2^2 = 3 \tilde{\gamma}^4 \quad (\text{Doppelwurzel}) \quad (\text{a16})$$

Dabei gehört zu (a15) die Hauptlösung nämlich $p_2 = -\tilde{\gamma}^2$

$$\begin{aligned} \gamma &= \sqrt{-p_2} = \sqrt{\tilde{\gamma}^2} = \pm \tilde{\gamma} \\ \sigma^2 &= \mu_2 + p_2 = \tilde{\gamma}^2 + \tilde{\sigma}^2 - \tilde{\gamma}^2 = \tilde{\sigma}^2 \end{aligned}$$

d.h. $\tilde{\gamma}$ und $\tilde{\sigma}^2$ werden durch die Lösung γ und σ^2 reproduziert.

Wählt man $p_2^2 = 3 \tilde{\gamma}^4$, so wird

$$\gamma = \pm \sqrt[4]{3} \tilde{\gamma} \quad \sigma^2 = \mu_2 - \sqrt{3} \tilde{\gamma}^2$$
$$\sigma^2 > 0 \quad \text{wenn} \quad \tilde{\sigma}^2 > (\sqrt{3}-1) \tilde{\gamma}^2$$

Bei dieser Lösung wird μ_4 nicht richtig reproduziert.

Zahlenbeispiel:

$$\tilde{\gamma} = 1, \quad \tilde{\sigma}^2 = 1.$$

$$\mu_2 = \tilde{\gamma}^2 + \tilde{\sigma}^2 = 1 + 1 = 2$$

$$\mu_4 = \tilde{\gamma}^4 + 6 \tilde{\gamma}^2 \tilde{\sigma}^2 + 3 \tilde{\sigma}^4 = 1 + 6 + 3 = 10$$

Zu $p_2^2 = 1$ gehört $\gamma = 1; \sigma^2 = 1$

Zu $p_2^2 = 3$ $\gamma = \sqrt[4]{3}$
 $\sigma^2 = 2 - \sqrt{3}$

$$\mu_2 = \gamma^2 + \sigma^2 = \sqrt{3} + 2 - \sqrt{3} = 2$$

$$\mu_4 = 3 + 6\sqrt{3}(2 - \sqrt{3}) + 3(4 + 4\sqrt{3} + 3) = 6 \neq 10$$

Die Nebenlösung verschwindet bei der geringsten Unsymmetrie. Anstelle der Doppelwurzel entsteht ein konjugiert komplexes Wurzelpaar.

Es empfiehlt sich daher nicht, die Symmetriefälle mit Gl (21) zu behandeln, sondern vorher anhand von $\mu_3; \xi_1$ und ξ_2 eine Fallunterscheidung durchzuführen.

Anhang II

Sturm'sche Kette.

Wie bereits auf S.72 gesagt kann man mit Hilfe der Sturm'schen Kette bestimmen, wieviele Nullstellen ein Polynom oder ggf. eine in eine unendliche Reihe entwickelte transzendente Funktion zwischen zwei Grenzen hat.

Die Sturm'sche Kette besteht aus der Funktion selbst, ihrer Abgeleiteten und den negativen Resten beim Euklidischen Teilungsverfahren.

$$f(x) = Q_1(x) f'(x) - R_2(x)$$

$$f'(x) = Q_2(x) R_2(x) - R_3(x)$$

$$R_1(x) = Q_3(x) R_3(x) - R_4(x) \quad (b1)$$

$$R_{n-2}(x) = Q_{n-1}(x) R_{n-1}(x) - R_n$$

Bei Polynomen gibt es schließlich ein R_n , das nicht mehr von x abhängt. Dies ist die Kette:

$$f(x), f'(x), R_2(x), R_3(x), \dots, R_n.$$

An den Grenzen $x=x_1$ und $x=x_2$ werden die Funktionswerte der Kette

$$f(x_1), f'(x_1), R_2(x_1) \dots R_n \quad (b2)$$

$$f(x_2), f'(x_2), R_2(x_2) \dots R_n$$

berechnet. Es muß gelten $f(x_1) \neq 0, f(x_2) \neq 0$.

Von den Funktionswerten der Kette interessieren nur die Vorzeichen. Ist $x_2 > x_1$, dann ist die Differenz der Zahl der Zeichenwechsel bei $x=x_1$ und derjenigen bei $x=x_2$ gleich der Zahl der reellen Wurzeln von $f(x)$ zwischen x_1 und x_2 , ohne Rücksicht auf deren Multiplizität. Mehrfachwurzeln werden

einfach gezählt. Als Zeichenwechsel gelten Übergänge von + nach - und umgekehrt, auch wenn dazwischen Funktionen den Wert 0 annehmen.

Das Verfahren führt bei der Bildung der Reste auf sehr große Zahlen, siehe auch /17/ S. 157. Man kann jedoch, um dies zu vermeiden, jeden Rest durch eine positive Konstante dividieren.

Die besten Resultate hinsichtlich des Zahlenbereiches erzielten wir mit einer Normierung des Koeffizienten der höchsten Potenz jeder Kettenfunktion auf ± 1 .

Wenn alle Wurzeln eines Polynoms vom Grade n reell sind, muß die Differenz der Zeichenwechsel n sein. Die Sturm'sche Kette besteht dann aus $n+1$ Funktionen, die sich um eins im Grad unterscheiden:

$$f(x), f'(x), R_2 \dots R_n$$

haben den Grad $n \quad n-1 \quad n-2 \dots 0$.

Die Koeffizienten der Glieder der höchsten Potenz müssen in diesem Fall alle gleiches Vorzeichen haben, denn sie bestimmen das Vorzeichen der Funktionen der Kette bei $-\infty$ und bei $+\infty$. Die Funktionen verlieren also ihre n Zeichenwechsel, wenn man x von $-\infty$ nach $+\infty$ laufen läßt.

Sind nicht alle Wurzeln reell, dann können sich weniger Funktionen in der Kette ergeben, dies muß aber keinesfalls der Fall sein. Siehe Beispiel (b 11).

Dies sei an einem Beispiel erläutert $n=3$

Gegeben ist

$$f(x) = x^3 + a_2x^2 + a_1x + a_0$$

Die Funktionen der Kette werden durch eine von mir in Fordecap geschriebene Prozedur erstellt und im Feld Y abgespeichert. $Y(1)$ enthält $f(x)$, $Y(2) = f'(x)$, $Y(3) = R_2(x)$ usw. (Auf die Reihenfolge der Größen in einer Funktion hat der Benutzer keinen Einfluß)

$$Y(1) = A_0 + X^2 A_2 + A_1 X + X^3$$

(b3)

$$Y(2) = A_1 + 2 X A_2 + 3 X^2$$

$$Y(3) = -9 A_0 + A_1 A_2 - 6 A_1 X + 2 X A_2^2$$

$$Y(4) = 162 A_0 A_1 A_2 + 9 A_1^2 A_2^2 - 36 A_0 A_2^3 - 36 A_1^3 - 243 A_0^2$$

Es entstehen also, da über den Wert der Koeffizienten $a_2 \dots a_0$ nichts ausgesagt ist, $n+1=4$ Funktionen.

Wir wollen nun sehen, was geschieht, wenn $a_1 = \frac{1}{3} a_2^2$ gesetzt wird.

`$z(3)=eval(y(3),a1,a2**2/3);`

$$Z(3) = -9 A_0 + 1/3 A_2^3$$

(b4)

`==>`

(Erläuterung: `$` ist der Druckbefehl für $z(3)$, `eval=evaluate` = ersetze in $Y(3)$ den Koeffizienten a_1 durch $\frac{1}{3} a_2^2$. \Rightarrow heißt "Befehl ausgeführt".)

Es verschwindet das lineare Glied in $Y(3)$ d.h. die Kette dürfte nur aus drei Funktionen bestehen. In diesem Fall kann $f(x)$ nicht drei reelle Wurzeln haben.

Jedoch zeigt sich, daß $Y(4)$ keinesfalls zu 0 wird.

`$z(4)=eval(y(4),a1,a2**2/3);`

$$Z(4) = 18 A_0 A_2^3 - 1/3 A_2^6 - 243 A_0^2$$

(b5)

Man sieht, daß $z(4) = -3(z(3))^2$ ist und daher stets ein negatives Vorzeichen hat, solange $z(3) \neq 0$ ist.

Daß dieses Glied tatsächlich nicht entsteht, zeigt man, indem man vor Erstellung der Kettenfunktionen a_1 durch $a_2^2/3$ ersetzt.

call sturml;

==>

call druck y;

$$Y(1) = A_0 + X^2 A_2 + 1/3 X A_2^2 + X^3$$

$$Y(2) = 2 X A_2 + 1/3 A_2^2 + 3 X^2$$

(b6)

$$Y(3) = -9 A_0 + 1/3 A_2^3$$

Ein zusätzliches absolutes Glied Y(4) verändert die Differenz der Zeichenwechselanzahlen an beiden Grenzen nicht.

Die Frage ist jedoch, ob Fehler bei dieser Differenz entstehen können, wenn andere, als absolute Glieder als Funktionen in der Kette zusätzlich vorhanden sind. Der Fall wurde an einem automatisch berechnetem Beispiel geklärt.

Das Beispiel lautet:

f=x**4+x**3+a2*x**2+x+1;

==>

call sturml;

==>

call druck y;

$$Y(1) = X + X^2 A_2 + X^3 + X^4 + 1$$

$$Y(2) = 2 X A_2 + 3 X^2 + 4 X^3 + 1$$

$$Y(3) = -12 X + 2 X A_2 - 8 X^2 A_2 + 3 X^3 - 15$$

$$Y(4) = 528 A_2 - 864 X + 960 X A_2 + 32 X A_2^2 - 128 X A_2^3 - 64 A_2^4 - 364$$

$$Y(5) = 746496 A_2 - 4458240 A_2^2 + 7833600 A_2^3 - 3256320 A_2^4 - 327630 A_2^5 + 262144 A_2^6$$

(b7)

Hierin setzen wir $A_2 = 0,375$ und erhalten

$$\text{ZZ}(1) = X + .375 X^2 + X^3 + X^4 + 1$$

$$\text{ZZ}(2) = .75 X + 3 X^2 + 4 X^3 + 1$$

$$\text{ZZ}(3) = - 11.25 X - 15$$

$$\text{ZZ}(4) = - 506.25 X - 675.$$

$$\text{ZZ}(5) = 0$$

(b8)

Man sieht

$$\text{zz}(4) = 45 \cdot z(3)$$

Bei direkter Berechnung

call sturml; call druck y;

==>

$$Y(1) = X + .375 X^2 + X^3 + X^4 + 1$$

$$Y(2) = .75 X + 3 X^2 + 4 X^3 + 1$$

$$Y(3) = - 11.25 X - 15$$

$$Y(4) = 66445.3125.$$

==>

(b9)

Nun berechnen wir mit beiden Ketten die Zahl der Nullstellen von $Y(1)$.

Wir haben dazu für x die Werte $-\infty$ und $+\infty$ einzusetzen. An diesen Stellen ist für jede Funktion der Kette nur das Glied mit der höchsten Potenz in x maßgeblich:

Aus (b8) folgt

$$x = -\infty: \quad + - + + 0 \quad 2 \text{ Zeichenwechsel}$$

$$x = +\infty: \quad + + - - 0 \quad 1 \text{ Zeichenwechsel}$$

Man würde auf eine Nullstelle schließen. Bei einer Funktion 4. Grades könnte dies nur eine Berührungsstelle sein.

Aus (b9) finden wir jedoch

$$x = -\infty : + - + + \quad 2 \text{ Zeichenwechsel}$$

$$x = +\infty : + + - + \quad 2 \text{ Zeichenwechsel}$$

Also hat die Funktion keine reellen Wurzeln, das mit (b8) gewonnene Ergebnis ist falsch.

Daß nicht etwa die Formeln (b7) fehlerhaft sind, wenn die Kette wirklich aus fünf Funktionen besteht, wurde überprüft.

Ein Beispiel für diesen Fall sei gegeben: $A2 = 2$ in (b7) eingesetzt findet sich in (b10), die direkte Berechnung in (b11)

$$\text{ZZZ(1)} = x + 2 \cdot x^2 + x^3 + x^4 + 1$$

$$\text{ZZZ(2)} = 4 \cdot x + 3 \cdot x^2 + 4 \cdot x^3 + 1$$

$$\text{ZZZ(3)} = -8 \cdot x - 13 \cdot x^2 - 15$$

$$\text{ZZZ(4)} = 160 \cdot x - 64$$

$$\text{ZZZ(5)} = 519168.$$

(b10)

==>
call sturml;

==>
call druck y;

$$Y(1) = x + 2 \cdot x^2 + x^3 + x^4 + 1$$

$$Y(2) = 4 \cdot x + 3 \cdot x^2 + 4 \cdot x^3 + 1$$

$$Y(3) = -8 \cdot x - 13 \cdot x^2 - 15$$

$$Y(4) = 160 \cdot x - 64.$$

$$Y(5) = 519168.$$

(b11)

==>

Die Anzahl der Zeichenwechsel ist:

$$x = -\infty : + - - - + \quad 2 \text{ Zeichenwechsel}$$

$$x = +\infty : + + - + + \quad 2 \text{ Zeichenwechsel}$$

Es sind keine reellen Wurzeln von $Y(1)$ vorhanden, obgleich Sturm'sche Kette die volle Anzahl von Funktionen, $n+1=5$, besitzt. (Dieser Fall war in zahlreichen numerischen Beispielen sehr häufig.)

Die Ursache für die fehlende Übereinstimmung der direkt erzeugten mit den durch Einsetzen in die allgemeine Form erzielten Kettenfunktionen im Sonderfall verschwindender Koeffizienzen ist eine in allgemeinen Koeffizienten formal ausgeführte Division durch Null. Folgerung: Die allgemeinen Formeln gelten nicht für Sonderfälle.

Man rechnet die Funktionen der Kette zweckmäßigerweise numerisch aus. Man kann dabei auf den häufigsten Fall abstellen, daß sie sich nur um die Einheit im Grad unterscheiden. Der Algorithmus läßt sich divisionsfrei gestalten. Die umstehend abgeleiteten Formeln täuschen Kompliziertheit vor. Er wird so einfach, daß er mit einem Kleinstrechner (Olivetti-Programma) durchgeführt werden kann. Eine geeignete Normierung verhindert, daß Koeffizienten zu groß oder zu klein werden. Falls die Voraussetzung, daß sich die Kettenfunktionen nur um eins im Grad unterscheiden, einmal verletzt sein sollte, erfolgt Fehlermeldung.

Algorithmus zur Berechnung der Koeffizienten der Sturm'schen Kette, für den Fall, daß sich die Funktionen im Grad um eins unterscheiden:

Man habe die Kette: $f(x), f'(x), R_2(x) \dots$ im Feld Y abgespeichert:

$$Y(1)=f(x), Y(2) = f'(x), Y(3) = R_2(x), \dots$$

Die Gleichungen (b1) lassen sich dann schreiben

$$Y(i) = Q_{i+1} \cdot Y(i+1) - Y(i+2)$$

Die nachfolgende Kettenfunktion $Y(i+2)$ entsteht also als negativ zu nehmender Divisionsrest aus den beiden vorhergehenden.

$$Y(i) : Y(i+1) = Q_{i+1} - \frac{Y(i+2)}{Y(i+1)}$$

Wenn sich die Kettenfunktionen alle um eins im Grad unterscheiden, dann kann man für ein beliebiges i schreiben:

$$Y(i) = a_0 x^n + a_1 x^{n-1} + a_2 x^{n-2} + \dots + a_n$$

$$Y(i+1) = b_1 x^{n-1} + b_2 x^{n-2} + \dots + b_n$$

$$Q_{i+1} = c_1 x + c_2$$

$$Y(i+2) = d_2 x^{n-2} + \dots + d_n$$

Darin ist $c_1 = \frac{a_0}{b_1}$ und

$$c_2 = \frac{a_1 - \frac{a_0 b_2}{b_1}}{b_1} = \frac{a_1}{b_1} - \frac{a_0 b_2}{b_1^2}$$

Q_{i+1} und damit auch der Rest $Y(i+2)$ enthalten keine Division mehr, wenn man den Divisor mit dem Quadrat des Koeffizienten der höchsten Potenz multipliziert.

Dies ist beim Sturm'schen Verfahren zulässig, da nur die Vorzeichen interessieren und eine Multiplikation mit einer positiven Zahl ungeändert läßt. Auf diese Weise gelangt man zu allgemeinen Formeln für die divisionsfreie Berechnung der $d_2, d_3 \dots d_n$ aus den $a_0, a_1, \dots a_n$ und $b_1, b_2, \dots b_n$. Die Formeln werden dadurch überhaupt erst übersichtlich.

$$b_1^2 a_0 x^n + b_1^2 a_1 x^{n-1} + b_1^2 a_2 x^{n-2} + b_1^2 a_3 x^{n-3} + \dots + b_1^2 a_n : (b_1 x^{n-1} + b_2 x^{n-2} + b_3 x^{n-3} + \dots + b_{n-1} x + b_n) = a_0 b_1 x + (a_1 b_1 - a_0 b_2) \quad c_1 \quad c_2$$

$$\therefore (b_1^2 a_1 - b_1 b_2 a_0) x^{n-1} + (b_1^2 a_2 - b_1 b_3 a_0) x^{n-2} + (b_1^2 a_3 - b_1 b_4 a_0) x^{n-3} + \dots + (b_1^2 a_{n-1} - b_1 b_n a_0) x + b_1^2 a_n$$

$$(b_1^2 a_1 - b_1 b_2 a_0) x^{n-1} + (b_2^2 a_1 - b_2 b_3 a_0) x^{n-2} + (b_3^2 a_1 - b_3 b_4 a_0) x^{n-3} + \dots + (b_n^2 a_1 - b_n b_2 a_0)$$

$$-r(i+2) = \underbrace{[b_1(a_2 b_1 - b_3 a_0) - b_2(b_3 a_1 - b_2 a_0)] x^{n-2}}_{-d_{n-1}} + \underbrace{[b_1(a_3 b_1 - a_0 b_4) - b_3(a_1 b_1 - a_0 b_2)] x^{n-3} + \dots + [b_1(a_{n-1} b_1 - a_0 b_n) - b_{n-1}(a_1 b_1 - a_0 b_2)] x}_{-d_n} + \underbrace{[b_i(a_n b_1) - b_n(a_1 b_1 - a_0 b_2)]}_{-d_n}$$

Die Koeffizienten $d_2 \dots d_n$ lassen sich leicht berechnen. Für jedes $Y(i+2)$ bestimmt man zunächst $c_2 = a_1 b_1 - a_0 b_2$, weiterhin sind a_0 und b_1 in allen Koeffizienten d enthalten. Um sie z.B. auf einem Kleinstrechner zu bestimmen, hat man nur a_0 , b_1 , a_1 und b_2 einzugeben und a_0 , b_1 und c_2 zu speichern. Für jedes $i = 2 \dots n$ sind dann a_i , b_i und b_{i+1} einzugeben. Mit vier Multiplikationen und zwei Subtraktionen wird der Koeffizient d_i gebildet. Anstelle von b_{n+1} ist 0 zu setzen.

Wird in einem $Y(i+2)$ mit $n > 2$ $d_2 = 0$, so würde beim nächsten Schritt der Gültigkeitsbereich des Algorithmus verlassen, da sich der Divisor vom Dividenden um mehr als eins im Grad unterscheidet. Dies ist nur dann unschädlich für die Differenz zwischen den Zeichenwechselanzahlen, wenn $n=3$ war, wenn also $Y(i+2)$ durch $d=0$ zum absoluten Glied wurde.

In allen Fällen $d=0$ bei $n > 3$ muß die Berechnung der Funktionen der Sturm'schen Kette abgebrochen werden. - Dieser Fall trat bei der praktischen Benutzung unter ca. 200 gerechneten Beispielen für die Pearsongleichung nur einmal auf.

Auch in diesem Fall hätte bei der später durchgeführten Normierung auf ± 1 für den Koeffizienten der höchsten Potenz für alle $Y(i)$ mit $i > 2$ das Unterschreiten der im Rechner möglichen kleinsten Zahl vermieden werden können. d_2 wurde also nicht echt, sondern nur fast Null.

Betrachtet man Funktionen $Y(1)$, deren Koeffizienten aus empirischen Daten stammen, dann ist der Fall, daß zwei Koeffizienten eine durch eine Gleichung zu beschreibende Relation erfüllen, nur eine Nullmenge unter allen Möglichkeiten. Nur unter der Bedingung, daß die Koeffizienten ganz bestimmte, durch eine Gleichung zu beschreibende Beziehungen untereinander haben, verschwindet aber d_2 .

Wir benutzen den Algorithmus zur Optimierung der Wurzelbestimmung bei der Pearsongleichung und sehen im Falle des Abbruches vor, nach allenreellen negativen Wurzeln der Gleichung 9ten Grades zu suchen. Dieser Fall wird jedoch fast sicher nicht eintreten.

Abschließend sollen in Fordecap geschriebene Prozeduren zur Berechnung der Sturm'schen Kette angegeben werden. Fordecap steht in Berlin, Pisa und Grenoble an der IBM 360/67 zur Verfügung und über Terminals allen Informatikern der BRD.

Sturm 1 : Procedure; ist vorgesehen für die Erstellung der Sturm'schen Kette mit allgemeinen Koeffizienten.

```
&BEG STACK
FORD
STURM1: PROCEDURE;
F=EXPAND(F);
G=DERIV(F,X);
Y(1)=F;
Y(2)=G;
I=3;

E3:  A=HIGHPOW(F,X);
      B=HIGHPOW(G,X);
      BB=COEFF(G,X**B);
      K=0;

E1:  AA=COEFF(F,X**A);
      F=EXPAND(F*BB);
      K=K+1;
      F=EXPAND(F-G*AA*X**(A-B));
      A=HIGHPOW(F,X);
      IF A .GE. B THEN GOTO E1;
      KK=(-1)**K;
      IF KK .EQ. 1 THEN GOTO E5;
      F=EXPAND(F*BB);

E5:  F=-F;
      Y(I)=F;

IF A .EQ. 0 THEN GOTO E2;
B=F;
F=G;
G=B;
I=I+1;
GOTO E3;

E2:  REMOB A,B,AA,BB,K,KK;

END;
```

Mit expand wird die Funktion F nach dem multinomialen und distributiven Gesetz entwickelt. In G wird die Abgeleitete von F nach X gebildet. Diese beiden Funktionen werden in Y(1) und in Y(2) gespeichert.

In A wird mit Highpow (F,X) der höchste Exponent von X in F festgehalten.

In der Sprache von Gleichung (b12) ist $A=n$, B ist beim ersten Durchlauf (Division v. $Y(1):Y(2) = F:G$) gleich $n-1$, später kann es sich von A um mehr als 1 unterscheiden. BB enthält den Koeffizienten der höchsten Potenz von G (also b1 im für Gleichung (b12) gültigen Fall).

Entsprechend wird in AA der Koeffizient der höchsten Potenz in X in F. Zur Vermeidung von Brüchen beim Quotienten und beim Rest wird der Dividend mit BB multipliziert.

Im Gegensatz zu (b12) erfolgt nicht die Multiplikation mit BB^2 , sondern mit $BB * * (A-B+1)$. Das ist in der Regel auch $BB * * 2$. Die Multiplikation erfolgt schrittweise, so oft eine Division erfolgt. Die Divisionen werden in K mitgezählt.

Später wird überprüft, ob K gerade oder ungerade ist:

$$KK = (-1) * * K$$

IF KK .EQUAL.1.THEN...

War die Anzahl ungerade, so erfolgt eine weitere Multiplikation mit BB. Auf diese Weise ist gesichert, daß der Dividend mit einer positiven Zahl multipliziert wurde.

Mit $F = \text{EXPAND}(F - G * AA * X * * (A-B))$; wird das erste Glied des Quotienten $AA * X * * (A-B)$ mit dem Divisor G multipliziert und vom Dividenden F subtrahiert. (Eigentlich wäre das erste Glied des Quotienten

$$\frac{AA}{BB} * X * * (A-B),$$

aber F wurde ja mit BB multipliziert.) Anschließend wird überprüft, ob der Rest von höherer oder gleicher Potenz in X ist, wie der Divisor. GE. = greater or equal. Gegebenenfalls wird erneut mit B1 multipliziert und durch $Y(i+1)$ dividiert.

Schließlich steht in F der negativ zu nehmende Rest

$$E5: F = -F;$$

Ist der Rest ein absolutes Glied: $A=0$, dann ist man am Schluß angelangt, die Objekte A,B,AA, ... werden entfernt (REMOB = remove objects), um Speicherplatz in der Freelist für weitere Operationen zu gewinnen.

Normalerweise wird der Rest jedoch zum Divisor und der vorherige Divisor zum Dividenden:

$$B = F;$$

$$F = G;$$

$$G = B;$$

Danach erfolgt ein neuer Durchlauf.

Dieses Programmbeispiel zeigt die Leistungsfähigkeit des Formac Formel-Manipulier-Systems.

Eine Anleitung zu seiner Benutzung findet man in einer Veröffentlichung von A. Laplace et.al. /21/ S. 26.

Ein dort mitgeteiltes Programm zur Bildung des größten gemeinsamen Teilers (GCD) von zwei Polynomen eignet sich nur für solche Polynome, deren Koeffizienten numerisch gegeben sind. Das vorliegende Programm hat, wie übrigens auch das Laplace'sche zur Bestimmung des GCD nach Beseitigung eines darin enthaltenen Fehlers, einen Nachteil: Bei Rechnung mit numerisch gegebenen Koeffizienten können infolge von Rechenungenauigkeiten bei der Differenzbildung kleinste Anteile der vorher höchsten Potenz übrig bleiben, wodurch ein falscher Grad des Restes vorgetäuscht wird und das Verfahren versagt. Dies wurde in einer numerischen Version:

S.: Procedure; ausgeschlossen in der nach jeder Division abgefragt wird, ob das höchste Glied im Rest von kleinerer Potenz ist als das Glied des Divisors, andernfalls wird es Null gesetzt.

Außerdem ist in dieser Version die besprochene Normierung auf ± 1 für die Koeffizienten der höchsten Potenz in X vorgesehen.

```
&BEG STACK  
FORD
```

```
S.: PROCEDURE;  
F=EXPAND(F);  
G=DERIV(F,X);
```

```
Y(1)=F;  
Y(2)=G;  
I=3;
```

```
E3: A=HIGHPOW(F,X);  
B=HIGHPOW(G,X);  
BB=COEFF(G,X**B);  
K=0;
```

```
E1: AA=COEFF(F,X**A);  
F=EXPAND(F*BB);  
K=K+1;  
F=EXPAND(F-G*AA*X**(A-B));  
C=HIGHPOW(F,X);
```

```
IF A .NE. C THEN GOTO E6;  
AAA=COEFF(F,X**A); F=EXPAND(F-AAA*X**A);  
C=HIGHPOW(F,X);
```

```
E6: A=C;  
IF A .GE. B THEN GOTO E1;  
KK=(-1)**K;  
IF KK .EQ. 1 THEN GOTO E5;  
F=EXPAND(F*BB);
```

```
E5: BBB=ABS(COEFF(F,X**A));
```

```
IF A .EQ. 0 THEN GOTO E2;  
F=EXPAND(-F/BBB);  
Y(1)=F;  
I = I+1;
```

```
S=F;  
F=G;  
G=S;  
GOTO E3;
```

```
E2: Y(1)=-F;  
REMOB A,B,AA,BB,K,KK,AAA,BBB,F,G;
```

```
END;
```

Von Interesse ist noch ein Programm zur Bestimmung der n
Zahl der Zeichenwechsel der Polynome $Y(1) \dots Y(i)$ an
den Grenzen X_1 und X_2 . (In der G.: Procedure wurde $X=-1$;
 $X_2 = +1$ gesetzt, da sie bei der Entscheidung über die
Gipfelzahl benutzt wurde).

```
G.:PROCEDURE;
L=1;
X1=-1;
X2=1;
E1:
  DO J=1 TO 1;
  Z(L,J)=EVAL(Y(J),X,X1);
  END;
L=L+1;
IF L .EQ. 3 THEN GOTO E2;
X1=X2;
GOTO E1;

E2:
  DO L=1 TO 2;
  K=0;
  A=LOP Z(L,1);

  DO J=2 TO 1;
  F=LOP Z(L,J);

  IF A .LT. 36 THEN GOTO E3;
  GOTO E10;

E3:
  IF F .LE. 36 THEN GOTO E20;
  K=K+1;
  A=F;
  GOTO E20;

E10:
  IF F .GE. 36 THEN GOTO E20;
  K=K+1;
  A=F;

E20:
  END;

K(L)=K;

END;
X1=-1;
$K(X1)=K(1); $K(X2)=K(2); $ANZHL=K(X1)-K(X2);

END;
```

In G.: Procedure werden mit $Z(L,Y) = \text{EVAL}(Y(J),X,X1)$; die Werte $X1$ (und später $X2$) für X in $Y(J)$ eingesetzt und diese Zahlen im Feld $Z(1,J)$ und $Z(2,Y)$ gespeichert.

Die Zahl der Zeichenwechsel kann in Fordecad mithilfe der Syntaxanalyse LOP (leading operator) festgestellt werden. Negative Zahlen haben $\text{LOP} < 36$, positive $\text{LOP} > 36$ und $0 \text{ LOP} = 36$. Aus den neun Möglichkeiten die zwischen den LOP's zweier Ausdrücke bestehen, gewinnt man durch Fallunterscheidung zwei, die einen Zeichenwechsel bedeuten.

	+		
+	0	nämlich	+ nach -
	-		
	+	und	- nach +
0	0		
	-		
	+	Beim Auftreten von 0 wird keine Ent-	
-	0	scheidung gefällt, sondern das vorherige	
	-	Zeichen gespeichert.	

Im Fortranprogramm kann man die Zahl der Zeichenwechsel nicht ganz so elegant mit der Signumfunktion bestimmen. Da $\text{sign}(0)$ nicht definiert ist, muß man diesen Fall ausschließen, z.B. indem man den vorhergehenden Ausdruck $\neq 0$ einsetzt.

Literaturverzeichnis:

- (1) Choi, K. und Bulgren, W.G.,
An Estimation Procedure for Mixtures of
Distributions.
(University of Missouri)
Journal of the Royal Statist. Society B 30 (1968)
S. 444 - 460
- (2) Doetsch, G.,
Zerlegung einer Funktion in Gaußsche Fehler-
kurven und zeitliche Zurückverfolgung eines
Temperaturzustands.
Mathematische Zeitschrift, 41, S. 283-318 (1936)
- (3) Weichselberger, K.,
Über ein graphisches Verfahren zur Trennung von
Mischverteilungen und zur Identifikation kupierter
Normalverteilungen bei großem Stichprobenumfang.
Metrika 4, 1961, S. 179-229.
- (4) Bhattacharya, C.G.,
A Simple Method of Resolution of a Distribution
into Gaussian Components.
Biometrics, März 1967
- (5) Pearson, K.,
On the Mathematical Theory of Evolution.
Philosophical Transactions of the Royal Society
of London, 1894 series A, vol. 185 pp. 71-110
Reprinted in "Karl Pearson's Early Statistical
Papers"
Cambridge, University Press 1956, pp. 1-40,
plates 1-5.

- (6) Robertson, C.A. und Fryer, J.G.
The bias and accuracy of moment estimators.
Biometrika (1970), 57,1, p. 57.
- (7) Rider, P.R.
The Method of Moments Applied to a Mixture
of Two Exponential Distributions.
Ann. Math. Statistik, Vol. 32,1 (1961)
- (8) Blischke, W.R.
Moment estimators for the parameters of
a mixture of two binominal distributions.
Ann. Math. Statist 33, (1962), 444-54.
- (9) Kalman, R.E.
New methods and results in linear filtering
and prediction theory.
Erschienen im: Fundamental Study of Adaptive
Control Systems,
R.E. Kalman et al., RIAS Division, The Martin
Company, Baltimore, Maryland, April 1962.
- (10) Abramowitz, M. und Stegun, I.A.
Handbook of Mathematical Functions.
Dover Publications, Inc. New York 1965.
- (11) Rao, C.R.,
Advanced Statistical Methods in Biometric
Research, Wiley, New York 1952, S. 304
- (12) Cramér, H.
Mathematical Methods of Statistics.
Princeton University Press, 12. Printing,
1971.

- (13) Charlier, C.V.L.
Researches into the Theory of Probability.
Kongl. Fysiografiska Sällskapetets Handlingar,
N.F. Bd. 16 Nr. 5, S. 1-51 (1906).
- (14) Schmetterer, L.,
Einführung in die mathematische Statistik.
Springer Verlag, Wien, New York 1966.
- (15) Paulus, E.,
Ein Algorithmus zur störsicheren Gruppierung
von Mustern.
Vortrag auf der Fachtagung "Cognitive Ver-
fahren und Systeme"
Hamburg 11.-13. April 1973 in
Lecture Notes in Economics and Mathematical
Systems, Bd. 83
Springer, Berlin, Heidelberg, New York 1973.
- (16) Behbodian, J.,
On the Modes of a Mixture of Two Normal
Distributions.
Technometrics Vol 12 Nr. 1 (Febr. 1970)
S. 131 - 138.
- (17) Bieberbach, L. und Bauer, G.,
Vorlesungen über Algebra
B.G. Teubner Verlag, Berlin, Leipzig 1928.
- (18) Stern,
Über die Auflösung der transzendenten
Gleichungen.
Crelles Journal, Bd. 33, S. 363.
- (19) Eisenberger, I.
Genesis of Bimodal Distributions
Technometrics, Vol 6, No.4, November 1964.

- (20) Dörrie, H.
Kubische und biquadratische Gleichungen.
Leibnitz-Verlag München 1948.
- (21) Laplace, A. und Siret, Y.
FORMac DEsk CALcULATOR.
University Grenoble, 1971.
- (22) Kühl, B.
Ein adaptives Wahrscheinlichkeitsdichte-
Approximationsverfahren mit Momenten.
Regelungstechnik und Prozeß-Datenverarbei-
tung Heft 8, 1972.
- (23) Berger, E.R. Aktenvermerk
über das dynamische Verhalten der adaptiven
Regelung mit exponentieller Alterung der
Beobachtungen.
SiemensZentral-Labor für NT Theoret. Aufgaben,
München 20.1.66.
- (24) Berger, E.R. Aktenvermerk
über die Berechnung der Regressionskoeffizien-
ten für eine adaptive Regelung.
Siemens Zentral-Labor für NT Theoret. Aufgaben,
München 19.1.66.
- (25) Burr, I.W.
A Useful Approximation to the Normal
Distribution Function, with Application
to Simulation.
Technometrics Vol. 9, Nr. 4 (Nov. 1967),
S. 647 - 51

- (26) Burrau, C.,
The half-invariants of the sum of two typical laws of errors, with an application to the problem of dissecting a frequency-curve into components.
Skandinavisk Aktuarietidskrift 1934 (82 646)
S. 1-54.
- (27) Cohen, A.C.,
Estimation in Mixtures of Two Normal Distributions.
Technometrics, Vol. 9, Nr. 1, Februar 1967,
S. 15 - 28.
- (28) Cooper, P.W.,
Statistical classification with quadratic forms.
Biometrika (1963), 50, 3 and 4, S. 439-448.
- (29) Day, N.E.,
Estimating the components of a mixture of normal distributions.
Biometrika (1969), Vol 56, 3 p. 463-474
- (30) Hasselblad, V.,
Estimation of Parameters for a Mixture of Normal Distributions.
- (31) Ifram, A.F.,
On Mixtures of Distributions with Applications to Estimation.
Journal of the American Statistical Association,
Vol 65, Nr. 330 (June 1970), S. 749-754.
- (32) John, S.,
On Analyzing Mixed Samples.
Journal of American Statistical Association,
June 1970, Vol 65, S. 755-762.

- (33) Kravčuk, M.,
Sur Quelques Approximations Dans le Problème
Des Moments pour les Fonctions de Deux Variables.
Comptes Rendus de l'Académie des Sciences de
l'URSS
Vol. XVII, Nr.6 (1937), S. 283-85.
- (34) Kravčuk, M.,
Sur le problème des moments généralisé.
Journal de l'Institut Mathématique de
l'Académie des Sciences de la RSS d'Ukraine,
1936 Nr. 4 S. 1-8.
- (35) Lonn, E.,
Beweis der Eindeutigkeit der Zerlegung einer
Intensitätskurve in ihre Komponenten.
Zeitschrift für Physik 75 (1932), S. 348-349.
- (36) Marriott, F.H.C.,
Practical Problems in a method of Cluster
Analysis.
Biometrics 27, 501-14 Sept. 1971.
- (37) Rao, C.,
The Utilization of Multiple Measurements in
Problems of Biological Classification.
J.R. Statist. Soc., 10 (1948), S. 159-203.
- (38) Rao, C.,
Resolution of a Mixed Series into Two
Gaussian Components.
Advanced Statistical Methods in Biometric
Research (1970) S. 300-304.

Heinrich Heertz-
Institut

- (39) Teicher, H.,
On the Mixture of Distributions.
Ann. Math. Stat., Vol. 31 (1960), S. 55-73.
- (40) Teicher, H.,
IDENTIFIABILITY of Mixtures.
Ann. Math. Statist., Vol 32 (1961), S. 244-48.
- (41) Teicher, H.,
IDENTIFIABILITY of Finite Mixtures.
Ann. Math. Statist., Vol. 34, S. 1265-1269
(1963).
- (42) Tchouproff, A.,
On the Mathematical Expectation of the Moments
of Frequency Distributions.
Biometrika Vol. XII, 1918/19, S. 140-295.
- (43) Weibull, W.,
A Statistical Distribution Function of Wide
Applicability.
Journal of Applied Mechanics, Sept. 1951,
S. 293 - 7.
- (44) Yakowitz, S.J.,
Unsupervised Learning and the Identification
of Finite Mixtures.
IEEE Transactions on information theory,
May 1970, S. 330-338.
- (45) Wolf, D. Korrespondenz vom 27.7.72 und Ergebnisse
der Amplitudenanalyse von Stromschwankungen
einer Diode im Relaxationsfall, mit der Bitte
um Bearbeitung.
Institut für angewandte Physik, Frankfurt/Main.

- (46) Berger, E.R., handschriftliche Notiz v. 9.6.71
- (47) Berger, E.R., mündliche Mitteilung August 71
- (48) Berger, E.R., Arbeitsanweisung v. 25.3.71

Herr Bernhard Kriener hat selbständig bei der Erstellung des Programmsystems zur Wahrscheinlichkeitsdichteapproximation mitgearbeitet und bei der Simulation und Auswertung von Meßdaten geholfen.

Herr Hany Selim und Herr Noffz haben spezielle Unterprogramme geschrieben.

Die Untersuchungen wurden von der Deutschen Forschungsgemeinschaft in Bad Godesberg (AZ Be 462/6 und Be 462/12, Sachbeihilfen an Herrn Prof. Dr. Berger) gefördert.

Aufgabenstellung und Grundgedanken des Verfahrens gehen auf Herrn Prof. Berger zurück.

Der Fachbereich 20 (Informatik) der Technischen Universität Berlin stellte die Rechenzeit zur Verfügung.

Allen Genannten gilt mein Dank.

