

# Die Maximum-Entropie-Methode zur Bestimmung von Mischanteilen

Christine Duller

IFAS - Institut für Angewandte Statistik  
Universität Linz

**Zusammenfassung:** Die Anforderungen an die Daten, Realisierungen von unabhängig und identisch verteilten Zufallsvariablen zu sein, sind Voraussetzung für eine Vielzahl statistischer Ansätze. In der Praxis erfüllen Daten nur sehr selten diese Bedingungen, meistens unterscheiden sich die Meßobjekte aber in Dimensionen, die bei der Datengewinnung nicht beachtet werden bzw. nicht beachtet werden konnten. Damit stellt sich folgende Aufgabe: Ausgangspunkt ist eine beobachtete Verteilung. Über die Struktur der Verteilung der Komponenten wird eine Hypothese formuliert. Ziel ist es, die Verteilung des unbeobachteten Parameters zu erfassen, also die Mischanteile festzustellen. Die Lösung erfolgt über eine Regularisierung mittels Maximum-Entropie-Methode, die anhand einiger theoretischer Überlegungen vorgestellt wird. Anschließend wird an einem Beispiel veranschaulicht, daß die so gefundene Lösung tatsächlich die Anforderungen einer Verteilung erfüllt.

**Abstract:** For several statistical methods realizations of identically, independently distributed random variables are required. But in reality most of the data do not satisfy this condition. More often than not data differ in dimensions, that are not observable. In this case we have a problem of mixed distributions. This problem can be specified as follows: Starting point is an observed distribution. Moreover we have a hypothesis about the distribution of the mixed components. The goal is to calculate the mixing proportions. Solving this problem with regression method or with Tikhonov regularization method was not very successful, because the solutions didn't fulfil the conditions of a distribution (to be positive and normed). This paper shows the solution using Maximum Entropy as a regularization method. With this method the conditions of a distribution are fulfilled.

**Schlüsselwörter:** Dekomposition, Mischverteilungen, Maximum-Entropie-Methode.

## 1 Problemstellung

Viele Methoden in der Statistik setzen die sogenannte „i.i.d.-Bedingung“ voraus, also die Anforderung an die Daten, Realisierungen von unabhängig und identisch verteilten Zufallsvariablen zu sein (independently identically distributed). Diese erleichtern die theoretische Betrachtung bzw. ermöglichen diese erst. Indessen erfüllen die Daten in der Praxis nur sehr selten diese Bedingungen. Meistens unterscheiden sich die Meßobjekte in Dimensionen, die bei der Datengewinnung nicht beachtet wurden bzw. gar nicht beachtet werden konnten, da sie a-priori unbekannt waren. Als einleuchtende Beispiele seien

verschiedene soziologische Untersuchungen erwähnt, bei denen es nicht möglich ist, von vornherein alle Einflußgrößen zu kennen und zu messen. Es werden also Variablen auftreten, die selbst nicht erhoben, oft nicht einmal als solche erkannt werden, die aber dennoch einen starken Einfluß auf die beobachteten Meßgrößen ausüben. Welche Auswirkungen solche unbeachtete „latente“ Variablen haben können, zeigt folgendes anschauliches Beispiel.

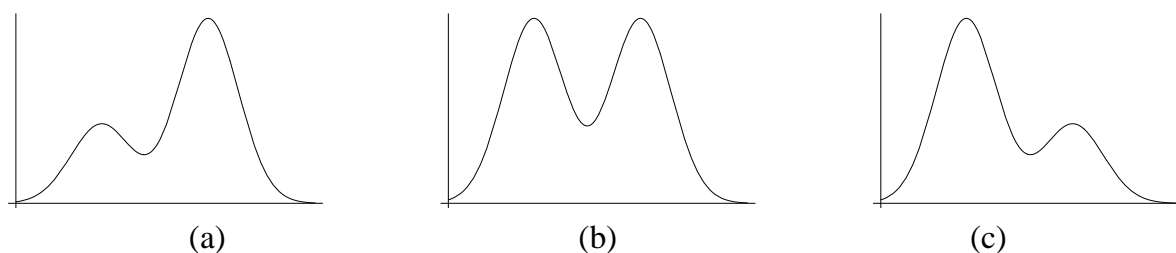


Abbildung 1: Drei mögliche Verteilungen der Körpergröße, wenn man bei der Untersuchung das Geschlecht der Personen nicht berücksichtigt

Gehen wir davon aus, daß die Körpergröße  $X$  eine vom Geschlecht abhängige normalverteilte Zufallsvariable mit fester Varianz ist, also:

$$X_w \sim N(\mu, \sigma^2) \quad \text{für weibliche Personen}$$

$$X_m \sim N(\mu + \Delta, \sigma^2), \Delta > 0 \quad \text{für männliche Personen}$$

Es ist unmittelbar einsichtig, daß die verschiedenen Ergebnisse dadurch entstanden sind, daß der Frauenanteil in den Untersuchungsgesamtheiten unterschiedlich hoch war. In (a) war der Frauenanteil deutlich niedriger als der Männeranteil, in (b) waren der Frauenanteil und der Männeranteil gleich groß und in (c) war der Frauenanteil deutlich höher als der Männeranteil.

Es liegen drei Mischungen zweier Normalverteilungen mit unterschiedlichen Mischanteilen vor. Eine Mischung ist das Ergebnis der Hintereinanderausführung von zwei Zufallsexperimenten, wobei der Ausgang des zweiten vom Ausgang des ersten Experimentes abhängt. Für das Beispiel mit der Körpergröße wäre das erste Zufallsexperiment die Auswahl einer Person aus der Untersuchungsgesamtheit, das zweite Experiment wäre das Abmessen der Körpergröße, wobei das Ergebnis dieser Messung vom Ergebnis des ersten Experimentes (Geschlecht) abhängig ist.

Damit stellt sich folgende grundlegende Aufgabe:

Ausgangspunkt ist eine beobachtete Verteilung. Über die Struktur der Verteilung der Komponenten gibt es eine Hypothese (bei der Körpergröße also z. B. Normalverteilungsannahme). Ziel ist es, die Verteilung des unbeobachteten Parameters zu erfassen, also die Mischanteile festzustellen. Man bezeichnet dies auch als die Dekomposition von Mischverteilungen.

Wir bezeichnen mit

$w(x)$  die Dichte der beobachteten Verteilung, mit

$f(x|\pi)$  die Dichte der Verteilung der Komponenten, darüber besteht eine Hypothese, und mit

$h(\pi)$  die Dichte der unbeobachteten Parameterverteilung, die errechnet werden soll.

Es gilt folgender Zusammenhang:

$$w(x) = \int_a^b f(x|\pi)h(\pi)d\pi$$

Zur Lösung wird von Sixtl (1993) eine Art Regressionsverfahren vorgeschlagen, welches den Vorteil hat, daß außer den Beobachtungen keine zusätzliche Informationen notwendig sind. Liegen für die Verteilung  $w(x)$  exakte Daten vor, können mit diesem Regressionsverfahren gute Ergebnisse erzielt werden. Standfest (1995) zeigte, daß die Ergebnisse im Fall von gestörten Daten (z.B. aufgrund von Meßfehlern) indessen unbrauchbar waren und versuchte das Verfahren mittels Tikhonov-Regularisierung zu verbessern. Allerdings führte auch das verbesserte Verfahren zu relativ unbefriedigenden Lösungen. Die Abweichungen zur tatsächlichen Lösung waren zwar relativ klein, aber die gefundene „Verteilung“ verletzte sowohl die Anforderung der Positivität  $h(\pi) \geq 0$  für  $\pi \in [a, b]$  als auch die Normierungsbedingung  $\int_a^b h(\pi)d\pi = 1$ .

Da aber der Ansatz eines Regressionsverfahrens in Kombination mit einem Regularisierungsverfahren an sich eine sehr gute Grundlage bildet, wird dieser prinzipielle Lösungsweg beibehalten. Um bessere Lösungen zu erzielen, wird anstelle der Tikhonov-Regularisierung eine Regularisierung mittels Maximum-Entropie-Methode vorgenommen.

## 2 Regularisierungsverfahren zur Lösung von Integralgleichungen

Das Gleichungssystem

$$w(x) = \int_a^b f(x|\pi)h(\pi)d\pi$$

kann diskretisiert in Matrixschreibweise  $w = Kh$  angeschrieben werden, mit

$$K = (K_{ji}) = f(x_j|\pi_i) \text{ und } h = h(\pi_i) = h_i \text{ mit } i = 1, \dots, n \text{ und } j = 1, \dots, m.$$

Die Lösung des Gleichungssystems gehört zu den schlecht gestellten oder zumindest zu den schlecht konditionierten Problemen. Man nennt ein Problem der Art  $Kh = w$  gut gestellt (oder korrekt gestellt), wenn die folgenden Bedingungen erfüllt sind:

1. Für jede Datenvorgabe  $w$  existiert eine Lösung  $h$ .
2. Diese Lösung  $h$  ist eindeutig.
3. Die Lösung  $h$  hängt stetig von den Daten  $w$  ab.

Sind die ersten beiden Bedingungen erfüllt, so folgt daraus die Existenz einer Umkehrfunktion. Aus der Linearität und der Stetigkeit des Operators  $K$  und dem Satz von

der inversen Abbildung folgt, daß die Umkehrfunktion ebenfalls stetig ist, das heißt im linearen Fall folgt Bedingung 3 aus den Bedingungen 1 und 2.

Die stetige Abhängigkeit der Lösung von den Daten ist eine notwendige, aber keine hinreichende Bedingung für die Stabilität der Lösung. Bei gut gestellten Problemen kann die relative Fehlerfortpflanzung mittels Konditionszahl gemessen werden, die definiert ist durch  $\text{cond}(K) = \|K\| \|K^{-1}\|$ .

Ist die Konditionszahl nicht zu groß, so wird das Problem als gut konditioniert bezeichnet und die Lösung ist stabil hinsichtlich kleiner Änderungen in den Daten, andernfalls liegt ein schlecht konditioniertes Problem vor. Anders als die Unterscheidung zwischen gut gestellten und schlecht gestellten Problemen ist also für die Unterscheidung zwischen gut konditionierten und schlecht konditionierten Problemen ein gewisser Ermessensspielraum vorhanden.

Bezeichnet  $\partial w$  eine Veränderung von  $w$ ,  $\partial h$  die zugehörige Veränderung in  $h$ , so bezeichnet  $\|\partial w\|/\|w\|$  den relativen Fehler der Daten und  $\|\partial h\|/\|h\|$  den relativen Fehler der zugehörigen Lösung. Durch die Konditionszahl besteht nun der Zusammenhang:

$$\frac{\|\partial h\|}{\|h\|} \leq \text{cond}(K) \cdot \frac{\|\partial w\|}{\|w\|}$$

In Worten ausgedrückt bedeutet dies, daß die Obergrenze des relativen Fehlers der Lösung mit der Konditionszahl steigt.

Sowohl schlecht gestellte als auch schlecht konditionierte Probleme können mit Regularisierungsverfahren approximativ gelöst werden, in denen die schlecht gestellte Aufgabe durch ähnliche, aber gut gestellte Probleme angenähert wird.

Im Fall der Tikhonov-Regularisierung wird beispielsweise das Minimum des Funktionals  $F(h) = \|Kh - w\|^2 + \alpha \|h\|^2$  aufgesucht.

Der Regularisierungsparameter  $\alpha \geq 0$  dient als Lagrangeparameter und übernimmt die Rolle eines Reglers zwischen Stabilität und Genauigkeit. Die Tikhonov-regularisierte Lösung für  $\alpha = 0$  geht in die Lösung des gewöhnlichen Regressionsverfahrens, also in die bestapproximierende Lösung, über. Diese Parameterwahl ist immer dann sinnvoll, wenn exakte Daten für die Mischung  $w(x)$  vorliegen. Im Normalfall liegen aber nur fehlerbehaftete Daten  $w_\delta$  vor, so daß ein Regularisierungsparameter  $\alpha$  gefunden werden muß, der einerseits groß genug ist, um Stabilität zu erreichen, andererseits aber klein genug ist, um das Problem nicht zu sehr abzuwandeln und dadurch Genauigkeit zu verlieren. In der Literatur gibt es eine Vielzahl von Methoden, die zur Wahl des Regularisierungsparameters vorgeschlagen werden.

Im folgenden wird immer von den Ergebnissen einer Tikhonov-Regularisierung mit geeigneter Parameterwahl ausgegangen.

### 3 Mathematische Grundlagen der Maximum-Entropie-Methode

#### 3.1 Anforderungen an ein Unsicherheits-Maß einer Wahrscheinlichkeitsverteilung

Die Wahrscheinlichkeiten  $\pi_1, \pi_2, \dots, \pi_n$  von  $n$  möglichen Ausgängen  $A_1, A_2, \dots, A_n$  eines Zufallsexperimentes bilden eine Verteilung

$$P = (\pi_1, \pi_2, \dots, \pi_n) \text{ mit } \pi_i \geq 0 \text{ für alle } i = 1, \dots, n \text{ und } \sum_{i=1}^n \pi_i = 1$$

Der tatsächliche Ausgang bei Durchführung des Experimentes ist demnach mit gewissen Unsicherheiten behaftet.

Ein Maß für diese Unsicherheit soll folgenden Anforderungen genügen:

1. Es soll eine Funktion der Wahrscheinlichkeiten  $\pi_1, \pi_2, \dots, \pi_n$  sein. Damit läßt sich das Maß darstellen als

$$H = H_n(P) = H(\pi_1, \pi_2, \dots, \pi_n).$$

2. Es soll eine stetige Funktion sein, so daß kleine Veränderungen der  $\pi_i$  nur kleine Veränderungen in  $H$  nach sich ziehen.
3.  $H_n$  soll eine symmetrische Funktion bezüglich der Argumente sein, das heißt eine Umordnung der Ausgänge soll keine Veränderung des Maßes nach sich ziehen.

4. Das Maß soll sich nicht ändern, wenn ein unmöglicher Ausgang zum Experiment hinzugefügt wird, also

$$H_{n+1}(\pi_1, \pi_2, \dots, \pi_n, 0) = H_n(\pi_1, \pi_2, \dots, \pi_n).$$

5. Das Maß soll minimal und nach Möglichkeit gleich Null sein, falls keine Unsicherheit bezüglich des Ausganges besteht. Falls also ein möglicher Ausgang sicher eintritt, so soll das Maß gleich Null sein:

$$H_n(\pi_1, \pi_2, \dots, \pi_n) = 0, \text{ wenn } \pi_i = 1 \text{ und } \pi_j = 0 \text{ für alle } i, j = 1, \dots, n \text{ mit } i \neq j.$$

6. Bei maximaler Unsicherheit soll auch das Maß maximal sein. Dieser Fall tritt dann ein, wenn alle möglichen Ausgänge gleichwahrscheinlich sind, also

$$\pi_1 = \pi_2 = \dots = \pi_n = \frac{1}{n}.$$

7. Der Maximalwert von  $H$  soll mit der Anzahl der möglichen Ausgänge wachsen.

8. Für zwei unabhängige Wahrscheinlichkeitsverteilungen soll die Unsicherheit der gemeinsamen Verteilung die Summe der einzelnen Unsicherheitsmaße sein:

$$P = (\pi_1, \pi_2, \dots, \pi_n) \quad X = (\xi_1, \xi_2, \dots, \xi_m) \quad \text{mit} \quad \sum_{i=1}^n \pi_i = 1 \quad \text{und} \quad \sum_{j=1}^m \xi_j = 1$$

$$H_{nm}(PX) = H_n(P) + H_m(X)$$

Dabei sind die Ausgänge von  $PX$  gegeben durch  $A_i B_j$  mit den Wahrscheinlichkeiten  $\pi_i \cdot \xi_j$  für  $i = 1, \dots, n$  und  $j = 1, \dots, m$ , sofern  $A_1, A_2, \dots, A_n$  die Ausgänge von  $P$  und  $B_1, B_2, \dots, B_m$  die Ausgänge von  $X$  bezeichnen.

9. Die achte Anforderung läßt sich folgendermaßen verallgemeinern: Das Maß an Unsicherheit für eine zusammengesetzte Verteilung ist die Summe aus dem Maß der einen Verteilung und dem Erwartungswert des Maßes der durch die erste Verteilung bedingten anderen Verteilung:

$$H_{nm}(PX) = H_n(P) + \sum_{i=1}^n \pi_i H^{(i)}(X)$$

wobei  $H^{(i)}(X)$  das Maß an Unsicherheit von  $X$  gegeben  $A_i$  angibt.

### 3.2 Shannons Maß der Unsicherheit

Shannon schlug folgendes Maß für die Unsicherheit vor

$$S(\pi_1, \dots, \pi_n) = - \sum_{i=1}^n \pi_i \ln(\pi_i)$$

Shannons Maß wird durch die genannten Anforderungen an ein Unsicherheitsmaß charakterisiert. Als Shannon 1948 dieses Maß entwickelte und einen passenden Namen dafür suchte, riet ihm der Mathematiker John Van Neumann aus zwei Gründen zu dem Namen Entropie (siehe Kapur, 1989):

*“Firstly you have got the same expression  $-\sum \pi_i \ln(\pi_i)$  as is used for entropy in thermodynamics and secondly and more importantly, since even after one hundred years, nobody understands what entropy is, if you use the word entropy you will always win in an argument.”*

### 3.3 Das Maximum-Entropie-Prinzip

Das Problem, Wahrscheinlichkeiten zu erfassen in Fällen, wo nur wenig bzw. bis auf die Normierungsbedingungen überhaupt keine Information verfügbar ist, ist so alt wie die Wahrscheinlichkeitsrechnung selbst. Das bekannteste Beispiel ist wohl das folgende:

Ein Würfel mit  $n = 6$  Flächen wird  $N = 1000$ -mal geworfen, wobei die  $i$ -te Fläche gerade  $N_i$ -mal aufscheint. Gesucht sind die Erwartungswerte der relativen Häufigkeiten für das Auftreten der  $i$ -ten Fläche.

Bereits Jakob Bernoulli erklärte die sechs Wahrscheinlichkeiten der verschiedenen Augenzahlen eines Würfels für gleich groß, da „wegen der gleichen Gestalt aller Flächen und wegen des gleichmäßig verteilten Gewichts des Würfels kein Grund dafür vorhanden ist, daß eine Würfelseite leichter als eine andere fallen sollte“. Diese Argumentation entspricht dem „Prinzip der unzureichenden Ursache“, welches besagt, daß zwei Ereignissen dieselbe Wahrscheinlichkeit zugewiesen werden soll, falls es keinen offensichtlichen Grund gibt, von den beiden Ereignissen eines besonders hervorzuheben. Diese Annahme ist insofern plausibel, da jede andere Zuordnung von Wahrscheinlichkeiten dahingehend inkonsistent ist, daß man dann durch bloßes Umordnen der einzelnen Fälle ein

neues Problem erzeugen kann, über das zwar genau dieselbe Information wie über das ursprüngliche Problem vorliegt, bei dem die Ergebnisse aber mit anderen Wahrscheinlichkeiten bezeichnet werden.

Die Idee, gewisse Ereignisse a-priori als gleichwahrscheinlich anzunehmen, mag zwar in manchen Fällen hilfreich sein, sie kann aber auch in manchen Fällen zu Widersprüchen führen. Man benötigt daher eine Vorgehensweise, die es erlaubt, ausgehend von den verfügbaren Informationen eine Wahrscheinlichkeitsverteilung der anderen vorzuziehen.

Eine mögliche Vorgehensweise stellt die Maximum-Entropie-Methode dar, welche auf dem folgenden Prinzip beruht (Reiter, 1985):

*“Ist auf der Grundlage unzureichender Information aus einer Vielzahl von Wahrscheinlichkeitsverteilungen eine Verteilung auszuwählen, dann ist genau diejenige zu nehmen, welche die größte Entropie besitzt und mit der gesamten verfügbaren Information übereinstimmt.“*

Dieses Prinzip erscheint natürlich zunächst willkürlich, basiert aber auf einem bekannten Grenzverteilungssatz für die empirische Entropie,

$$\hat{S}_N = - \sum_{i=1}^n p_i \ln(p_i),$$

wobei die relativen Häufigkeiten  $p_i = \frac{N_i}{N}$  der Ereignisse  $A_i$  die Schätzer der Wahrscheinlichkeiten  $\pi_i$  sind.

Ist nämlich  $S_{max}$  die maximale Entropie des restringierten Optimierungsproblems

$$S(\pi_1, \dots, \pi_n) \rightarrow \text{Max!}$$

$$\sum_{i=1}^n \pi_i = 1$$

$$\pi_i \geq 0 \text{ für alle } i = 1, \dots, n$$

- gegebenenfalls unter einer zusätzlichen Nebenbedingung über die zugrundeliegende Verteilung  $P = (\pi_1, \dots, \pi_n)$  -, ist

$$\Delta S = \frac{\chi_{1-\alpha, n-1}^2}{2N}$$

und ist der Stichprobenumfang  $N$  hinreichend groß, dann ist die Wahrscheinlichkeit des Ereignisses  $S_{max} - \Delta S \leq \hat{S}_N \leq S_{max}$  annähernd gleich  $1 - \alpha$ .

Die genannte Optimierungsaufgabe läßt sich unter geeigneten Nebenbedingungen über  $P$  stets lösen, wobei zur Auffindung der optimalen Verteilung in der anwendungsorientierten Literatur in der Regel die Lagrange-Methode verwendet wird.

Liegen keine zusätzlichen Nebenbedingungen hinsichtlich  $P$  vor, so ist die Wahrscheinlichkeitsverteilung mit maximaler Entropie bekanntlich die Gleichverteilung  $P = (1/n, \dots, 1/n)$  und  $S_{max} = \ln(n)$ .

Für das Beispiel eines Würfels ist der Maximalwert  $S_{max}$  der Entropie somit  $S_{max} = \ln(6) = 1,7918$ . Da  $\chi_{1-0,05; 6-1}^2 = 11,071$ , liegt die empirische Entropie bei  $N = 1000$ -maligen Würfeln mit 95 %-iger Sicherheit im Bereich  $1,7862 \leq S \leq 1,7918$ .

Für den Fall, daß keine zusätzlichen Nebenbedingungen hinsichtlich  $P$  vorliegen, beruht der genannte Grenzverteilungssatz für die empirische Entropie ausschließlich auf

kombinatorischen Grundlagen und zählt lediglich die verschiedenen Möglichkeiten (Zustände) des Zufallsexperimentes ab. Daher kann man das Theorem auch dahingehend interpretieren, daß diejenigen Verteilungen mit höherer Entropie auf wesentlich mehr verschiedene Weisen realisiert werden können als solche mit niedrigerer. Dies bedeutet aber, daß es auch viel wahrscheinlicher ist, Zustände mit höherer Entropie zu beobachten, als solche mit niedrigerer Entropie.

Dazu folgendes Anschauungsbeispiel:

Ausprägung	$p_A(\mathbf{x})$	$p_B(\mathbf{x})$	$p_C(\mathbf{x})$
1	0,5	0,25	0
2	0,5	0,75	1
Entropie	0,693	0,562	$\sim 0$

Mögliche Realisationen für diese Verteilungen bei  $N = 4$

	$p_A(\mathbf{x})$	$p_B(\mathbf{x})$	$p_C(\mathbf{x})$
	1122	1222	2222
	1212	2122	
	1221	2212	
	2112	2221	
	2121		
	2211		
Anzahl der Realisationen	$\binom{4}{2} = 6$	$\binom{4}{1} = 4$	$\binom{4}{0} = 1$

### 3.4 Die Maximum-Entropie-Methode als Regularisierungsverfahren

Will man die Maximum-Entropie-Methode zur Lösung von Entfaltungsproblemen heranziehen, so hat man vor allem mit dem Problem fehlerbehafteter Daten zu kämpfen, darüber hinaus ist das Entfaltungsproblem im allgemeinen nicht eindeutig lösbar.

Es soll folgender Ansatz zur Lösung verwendet werden. Man sollte aus allen möglichen Lösungen, die mit der Residuenstatistik kompatibel sind, genau diejenige herausgreifen und als „wahr“ (= optimal) ansehen, die überall positiv ist, die Normierungsbedingung erfüllt und die möglichst glatt verläuft.

Ausgangspunkt des Entfaltungsproblems ist das Gleichungssystem

$$w = Kh \text{ bzw. } w = Kh + e$$

wobei  $w$  der gegebene Datenvektor ist,  $K$  eine bekannte Matrix,  $h$  der Lösungsvektor der Mischanteile und  $e$  der Fehlervektor, von dem wir annehmen, daß er normalverteilt ist ( $\mu = 0, \sigma^2$ ). Im Normalfall ist  $K$  eine unterbestimmte und schlecht konditionierte Matrix.

Dieses Problem wird gelöst mittels Regularisierung durch Minimieren des Problems

$$F_\alpha(h) = \|Kh - w\|^2 + \alpha\Omega(h)$$

Im Falle der Tikhonov-Regularisierung ist

$$\Omega(h) := \|h\|^2 = \sum_{i=1}^n h_i^2,$$



im Falle der Regularisierung mittels Maximum-Entropie-Methode ist hingegen

$$\Omega(h) := \sum_{i=1}^n h_i \ln(h_i)$$

Der Hauptnachteil dieses Ansatzes liegt darin, daß nun  $\Omega(h)$  keine quadratische Funktion mehr ist.

### 3.5 Ein Algorithmus zur Dekomposition

Ziel ist es also, folgendes Problem zu lösen:

$$Kh = w \quad \Leftrightarrow \quad \sum_{i=1}^n K_{ji} h_i = w_j \quad \text{für } j = 1, \dots, m$$

$$\text{mit } \sum_{i=1}^n h_i = 1 \text{ und } h_i \geq 0 \text{ für } i = 1, \dots, n$$

wobei

$w$  bekannter Datenvektor der Dimension  $m$

$h$  Vektor der unbekanntenen Parameterverteilung der Dimension  $n$

$K$  bekannte Matrix der Dimension  $(m \times n)$

Dadurch sind  $m + 1$  Nebenbedingungen zu erfüllen, wobei o.B.d.A. davon ausgegangen wird, daß die  $m + 1$  Nebenbedingungen nicht linear voneinander abhängig sind und  $m + 1 < n$  ist. In diesem Fall liegt keine eindeutige Lösung vor. Das Prinzip der maximalen Entropie sieht nun vor, jener Lösung den Vorzug zu geben, die zusätzlich zu den gegebenen Bedingungen die Entropie, gegeben durch

$$S(h_1, \dots, h_n) = - \sum_{i=1}^n h_i \ln(h_i),$$

maximiert, was gleichbedeutend ist damit, das Funktional

$$\sum_{i=1}^n h_i \ln(h_i)$$

unter Nebenbedingungen zu minimieren.

Die Regularisierung mittels Maximum-Entropie-Methode ist daher nichts anderes als das Aufsuchen der maximalen Entropie unter Nebenbedingungen. Daher ist auch bei der Verwendung der Maximum-Entropie-Methode als Regularisierungsverfahren stets die Positivität und Normiertheit der Lösung garantiert.

Mittels Lagrangemultiplikatoren führt das restringierte Optimierungsproblem zu dem Optimierungsproblem

$$L(h, \lambda) = \sum_{i=1}^n h_i \ln(h_i) + \tilde{\lambda}_0 \left( \sum_{i=1}^n h_i - 1 \right) + \sum_{j=1}^m \lambda_j \left( \sum_{i=1}^n K_{ji} h_i - w_j \right) \Rightarrow \text{Min!}$$

Partielles Differenzieren führt auf

$$\frac{\partial L(h, \lambda)}{\partial h_i} = \ln(h_i) + \underbrace{1 + \tilde{\lambda}_0}_{=: \lambda_0} + \sum_{j=1}^m \lambda_j K_{ji} \quad i = 1, \dots, n$$

Durch Nullsetzen der  $n$  Ableitungen und Umformung ergeben sich die Gleichungen

$$h_i = \exp\left(-\sum_{j=1}^m \lambda_j K_{ji}\right) / \exp(\lambda_0) \quad i = 1, \dots, n$$

Aus der Bedingung  $\sum_{i=1}^n h_i = 1$  ergibt sich

$$\exp(\lambda_0) = \sum_{i=1}^n \exp\left(-\sum_{j=1}^m \lambda_j K_{ji}\right)$$

und somit

$$h_i = \frac{1}{\sum_{i=1}^n \exp\left(-\sum_{j=1}^m \lambda_j K_{ji}\right)} \exp\left(-\sum_{j=1}^m \lambda_j K_{ji}\right) \quad i = 1, \dots, n.$$

Unter Berücksichtigung von (1) ergibt sich somit

$$\sum_{i=1}^n (K_{ji} - w_j) \exp\left(-\sum_{j=1}^m \lambda_j K_{ji}\right) = 0 \quad j = 1, \dots, m$$

Dieses Gleichungssystem wird für die Simulationsstudie mittels Mathematica-Programm gelöst.

## 4 Simulation

Anhand eines konkreten Beispiels einer Mischung, wie sie bei Sixtl (1995) angegeben ist, sollen die Lösungsmethoden Regressionsverfahren, Tikhonov-Regularisierung mit geeigneter Parameterwahl und Regularisierung mittels Maximum-Entropie-Methode miteinander verglichen werden. Dazu greifen wir für die ersten beiden Methoden auf bereits vorhandenes Datenmaterial zurück und verweisen für Details auf Sixtl bzw. Standfest.

### 4.1 Exakte Daten

Für den Fall exakter Daten werden wir folgendermaßen vorgehen: Wir erzeugen eine exakte Mischverteilung  $w(x)$ , deren Komponenten  $f(x|\pi)$  eine Binomialverteilung aufweisen und deren Parameter  $\pi$  eine Beta-Verteilung über  $[0,1]$  aufweist. Diese Mischverteilung soll nun Ausgangspunkt für den Lösungsvergleich sein.

Damit ergibt sich als Verteilungsstruktur der Komponenten

$$f(x) = \binom{n}{x} \pi^x (1 - \pi)^{n-x} \quad x = 0, 1, 2, \dots, n$$

Die Mischanteile genügen einer Beta-Verteilung mit den Parametern  $a$  und  $b$  über dem Intervall  $[0, 1]$ , also

$$h(\pi) = \frac{1}{B(a, b)} \pi^{a-1} (1 - \pi)^{b-1} \quad \pi \in [0, 1]; a, b, > 0$$

wobei  $B(a, b) = \int_0^1 \pi^{a-1} (1 - \pi)^{b-1} d\pi$  die vollständige Betafunktion ist.

Damit ergibt sich die Mischverteilung aus

$$\begin{aligned} w(x) &= \int_0^1 f(x|\pi) h(\pi) d\pi = \binom{n}{x} \frac{1}{B(a, b)} \int_0^1 \pi^{x+a-1} (1 - \pi)^{n-x+b-1} d\pi \\ &= \binom{n}{x} \frac{B(x+a, n-x+b)}{B(a, b)} \end{aligned}$$

Wir folgen dem speziellen Beispiel von Sixtl und wählen  $n = 4$ ,  $a = 3$  und  $b = 5$ . Damit hat die exakte Mischung  $w(x)$  folgende Beschaffenheit:

Tabelle 1: Exakte Daten  $w(x)$

$x$	0	1	2	3	4	$\Sigma$
$w(x)$	$\frac{7}{33}$	$\frac{7}{22}$	$\frac{3}{11}$	$\frac{5}{33}$	$\frac{1}{22}$	1
	= 0,212121	= 0,318182	= 0,272727	= 0,151515	= 0,045455	

## 4.2 Ein Gütekriterium zur Beurteilung der Lösungen

Als Gütekriterium wird eine Art „Bestimmtheitsmaß“ herangezogen, welches folgendes Aussehen besitzt:

$$B^* = \frac{\|h\|^2 - \|h - \hat{h}\|^2}{\|h\|^2}$$

Für die Interpretation dieser Maßzahl sollte man die exakte Lösung  $h$  als Punkt in einem mehrdimensionalen (hier  $m$ -dimensionalen) Raum ansehen.

Liegt eine Näherungslösung  $\hat{h}$  „nahe“ beim gesuchten exakten Punkt  $h$ , so wird der quadrierte euklidische Abstand zwischen Näherungslösung und exakter Lösung sehr klein sein, und damit das Bestimmtheitsmaß nahe bei 1 liegen. Im Fall  $B^* = 1$  wurde die exakte Lösung gefunden.

Im folgenden wird anstelle der Dichte  $h(\pi)$  die Diskretisierung  $h_d(\pi)$  mit

$$h_d(\pi) := \int_{\pi_u}^{\pi_o} h(\pi) d\pi \quad \pi \in [\pi_u; \pi_o]$$

verwendet. Die für diese Diskretisierung gewählten Intervallgrenzen können aus der Tabelle 2 entnommen werden.

In dieser Tabelle sind auch die Ergebnisse des Regressionsverfahrens ( $\alpha = 0$ ) und mehrere Ergebnisse einer Tikhonov-Regularisierung mit vorgegebenen  $\alpha$  angeführt (vgl. Tabelle 2).

Tabelle 2: Lösungsvergleich für exakte Daten

Intervall	$h_d(\pi)$	$\hat{h}(\pi)$ geschätzt			
		$\alpha = 0$	$\alpha = 0.02$	$\alpha = 0,04$	$\alpha = 0,06$
[0 ; 0,05]	0,003757	-0,005865	0,019723	0,026798	0,029040
]0,05 ; 0,15]	0,070008	0,089989	0,096368	0,093086	0,088377
]0,15 ; 0,25]	0,169826	0,176436	0,143426	0,125848	0,113452
]0,25 ; 0,35]	0,224126	0,212292	0,163346	0,138978	0,123298
]0,35 ; 0,45]	0,215843	0,205830	0,158537	0,135942	0,120461
]0,45 ; 0,55]	0,163512	0,167614	0,137581	0,120293	0,107543
]0,55 ; 0,65]	0,097320	0,110497	0,105228	0,095672	0,087207
]0,65 ; 0,75]	0,042729	0,049626	0,067394	0,065806	0,062176
]0,75 ; 0,85]	0,011657	0,002436	0,030166	0,034512	0,035229
]0,85 ; 0,95]	0,001216	-0,011344	-0,000202	0,005691	0,009208
]0,95 ; 1]	0,000006	0,006702	-0,007309	-0,005952	-0,003980
$\sum$	1	1,004213	0,913258	0,836674	0,772011
$B^*$	1	0,992287	0,931609	0,874507	0,825551

Im Fall von exakten Daten liefert also das gewöhnliche Regressionsverfahren die Lösung mit dem größten Bestimmtheitsmaß. Diese Lösung ist aber dennoch unbefriedigend, weil sowohl die Anforderung der Positivität als auch der Normiertheit verletzt sind. Darüber hinaus liegt im Fall von  $\alpha = 0$  eine äußerst instabile Lösung vor. Bereits eine kleine Änderung des Parameters  $\alpha$  ruft eine starke Änderung der einzelnen Lösungselemente hervor. Die Stabilität für ein größeres  $\alpha$  ist weitaus besser. Anders ausgedrückt liefert ein kleines  $\alpha$  also eine genauere, aber instabilere Lösung, während mit einem größerem  $\alpha$  eine stabilere, aber ungenauere Lösung erzielt werden kann.

### 4.3 Gestörte Daten

Ausgangspunkt sind die exakten Daten von  $w(x)$  an den  $s$  Stützstellen  $x_1, \dots, x_s$  (in unserem Fall  $s = 5$ ). Durch die Verteilungsfunktion  $F(x_i)$  wird das Intervall  $[0, 1]$  in  $s$

Teilintervalle untergliedert:

$$F(x_i) = \sum_{k=1}^i w(x_k) \quad i = 1, \dots, s$$

Für eine Stichprobe vom Umfang  $N$  werden nun  $N$  im Intervall  $[0, 1]$  gleichverteilte Pseudozufallszahlen erzeugt und beobachtet, in welches der  $s$  Teilintervalle die Ergebnisse fallen. Aus den relativen Häufigkeiten  $p_i = N_i/N$ ,  $i = 1, \dots, s$  in den einzelnen Teilintervallen ergibt sich dann eine simulierte fehlerbehaftete Datenvorgabe als  $w_\delta(x_i) = p_i$ ,  $i = 1, \dots, s$ .

Je größer der Stichprobenumfang  $N$ , desto besser die Anpassung an die exakten Daten, desto geringer also der Fehler. In der Tabelle 3 ist eine mögliche Datenvorgabe angegeben, in der Tabelle 4 ist der Lösungsvergleich für diese Daten angeführt.

Tabelle 3: Mögliche Datenvorgabe für gestörte Daten

$x$	0	1	2	3	4
$w(x)$	0,212121	0,318182	0,272727	0,151515	0,045455
$w_\delta(x)$	0,24	0,38	0,18	0,16	0,04

Tabelle 4: Lösungsvergleich für gestörte Daten

Intervall	$h_d(\pi)$	$\hat{h}(\pi)$ geschätzt			
		$\alpha = 0$	$\alpha = 0,01$	$\alpha = 0,02$	$\alpha = 0,03$
[0 ; 0,05]	0,003757	-0,201930	0,017178	0,030809	0,034481
]0,05 ; 0,15]	0,070008	0,289620	0,133377	0,120340	0,110899
]0,15 ; 0,25]	0,169826	0,506640	0,183712	0,151384	0,132949
]0,25 ; 0,35]	0,224126	0,314580	0,176410	0,147507	0,129671
]0,35 ; 0,45]	0,215843	0,036000	0,142104	0,125644	0,112736
]0,45 ; 0,55]	0,163512	-0,127500	0,102538	0,098028	0,090696
]0,55 ; 0,65]	0,097320	-0,095280	0,070557	0,072271	0,068988
]0,65 ; 0,75]	0,042729	0,092340	0,050114	0,051371	0,049930
]0,75 ; 0,85]	0,011657	0,274068	0,036269	0,033707	0,032722
]0,85 ; 0,95]	0,001216	-0,167700	-0,015186	0,013042	0,013448
]0,95 ; 1]	0,000006	-0,169583	-0,009358	-0,005390	-0,003239
$\sum$	1	1,086655	0,918087	0,838713	0,773281
$B^*$	1	-2,547257	0,892883	0,853360	0,809308

Offensichtlich führt die Tikhonov-Regularisierung im Fall fehlerbehafteter Daten zu eindeutig besseren Ergebnissen als das Regressionsverfahren, wie aus dem Bestimmt-

heitsmaß zu entnehmen ist. Allerdings weisen auch diese „Verteilungen“ negative Werte auf und die Normiertheit ist ebenfalls verletzt.

Zusammenfassend läßt sich für dieses Simulationsbeispiel vorerst folgendes feststellen: Im Fall von exakten Daten liefert das Regressionsverfahren (bzw. Tikhonov-Regularisierung mit  $\alpha = 0$ ) die Lösung mit dem größten Bestimmtheitsmaß, im Fall gestörter Daten ist eine Tikhonov-Regularisierung mit einem geeigneten Regularisierungsparameter  $\alpha > 0$  auf jeden Fall vorzuziehen.

Der Regularisierungsparameter fungiert als Regler zwischen Stabilität (je größer  $\alpha$ , desto stabiler) und Genauigkeit (je kleiner  $\alpha$ , desto genauer die Lösung). Eines aber ist allen gefundenen Lösungen gemeinsam: Sie verletzen sowohl die Anforderungen an die Positivität als auch an die Normiertheit, so daß bei den gefundenen Lösungen auf keinen Fall von einer Verteilung gesprochen werden kann.

#### 4.4 Vergleich der Lösungen

Für den ersten Vergleich gehen wir von dem Fall aus, daß exakte Daten vorliegen. Eine grafische Darstellung ermöglicht uns einen ersten Einblick in die erhaltenen Lösungen. Abbildung 2 zeigt den Vergleich zwischen exakter Lösung und der Lösung mittels Tikhonov-Regularisierung. Man sieht, daß die Tikhonov-regularisierte Lösung an den Rändern die Positivitätsanforderung nicht erfüllt. Über die Qualität der Lösung läßt sich keine genaue Aussage machen, dafür werden wir im nächsten Schritt das Bestimmtheitsmaß heranziehen.

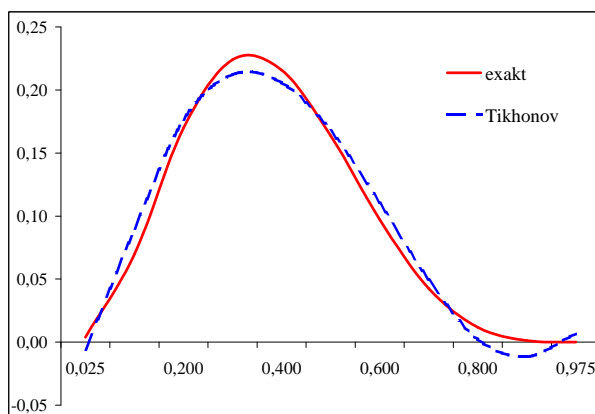


Abbildung 2: Vergleich exakte Lösung und Lösung mittels Tikhonov-Regularisierung

Abbildung 3 zeigt die exakte Lösung im Vergleich zur Lösung mittels Maximum-Entropie-Methode.

Vergleicht man nun die beiden Abbildungen, so fällt sofort auf, daß im Gegensatz zu vorher jetzt auch an den Rändern die Positivitätsbedingung erfüllt ist.

Welche der beiden regularisierten Lösungen bezüglich des Bestimmtheitsmaßes besser ist, sieht man im Zahlenvergleich in den Tabellen 5 und 6.

Aus den Ergebnissen ist erkennbar, daß im Fall der Tikhonov-Regularisierung die Positivitätsbedingung und die Normierungsbedingung verletzt sind, daß es sich also bei

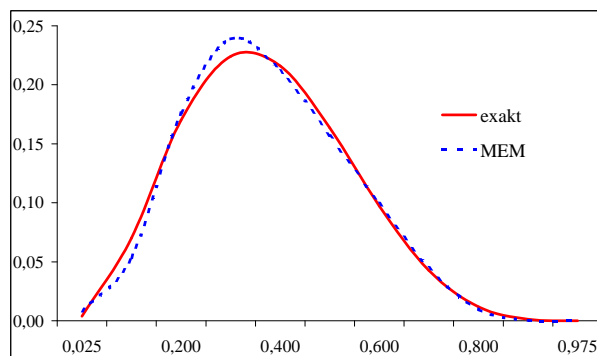


Abbildung 3: Vergleich exakte Lösung und Lösung mittels Maximum-Entropie-Methode

Tabelle 5: Lösungsvergleich für exakte Daten

		Intervall	$h_d(\pi)$	$\hat{h}(\pi)$	
				Tikhonov $\alpha = 0$	MEM
		[0 ; 0,05]	0,0038	-0,0059	0,0082
		]0,05 ; 0,15]	0,0700	0,0900	0,0534
		]0,15 ; 0,25]	0,1698	0,1764	0,1760
		]0,25 ; 0,35]	0,2241	0,2123	0,2390
		]0,35 ; 0,45]	0,2158	0,2058	0,2119
		]0,45 ; 0,55]	0,1635	0,1676	0,1567
		]0,55 ; 0,65]	0,0973	0,1105	0,0995
		]0,65 ; 0,75]	0,0427	0,0496	0,0450
		]0,75 ; 0,85]	0,0117	0,0024	0,0097
		]0,85 ; 0,95]	0,0012	-0,0113	0,0005
		]0,95 ; 1]	0,0000	0,0067	0,0000
		$\Sigma$	1,0000	1,0042	1,0000
			$B^* =$	0,9923	0,9963
$x$	$w(x)$				
0	0,212121				
1	0,318182				
2	0,272727				
3	0,151515				
4	0,045455				
$\Sigma$	1				

der gefundenen Lösung um keine Verteilung handelt. Die Maximum-Entropie-Methode hingegen liefert als Lösung eine Verteilung und das Bestimmtheitsmaß weist einen - wenn auch unwesentlich - höheren Wert auf. Im Fall von exakten Daten scheint also die Maximum-Entropie-Methode in jeder Hinsicht bessere Resultate zu liefern als die Tikhonov-Regularisierung.

Im nächsten Schritt soll nun dieser Vergleich auch für gestörte Daten angestellt werden. In unserem ersten Beispiel haben wir für die Erzeugung der gestörten Daten einen sehr niedrigen Stichprobenumfang gewählt, was natürlich zu stark fehlerbehafteten Daten führt.

Die Lösung mittels Tikhonov-Regularisierung verletzt auch hier wieder die Positi-

Tabelle 6: Lösungsvergleich für gestörte Daten,  $N = 50$ 

			Intervall	$h_d(\pi)$	$\hat{h}(\pi)$	
					Tikhonov	MEM
					$\alpha = 0,0212$	
			[0 ; 0,05]	0,0038	0,0339	0,1573
			]0,05 ; 0,15]	0,0700	0,0664	0,0038
			]0,15 ; 0,25]	0,1698	0,1963	0,0030
$x$	$w(x)$	$w_\delta(x)$	]0,25 ; 0,35]	0,2241	0,1358	0,0253
0	0,212	0,20	]0,35 ; 0,45]	0,2158	0,1638	0,2033
1	0,318	0,24	]0,45 ; 0,55]	0,1635	0,1676	0,3835
2	0,273	0,36	]0,55 ; 0,65]	0,0973	0,1431	0,1113
3	0,152	0,16	]0,65 ; 0,75]	0,0427	0,0944	0,0087
4	0,045	0,04	]0,75 ; 0,85]	0,0117	0,0342	0,0009
$\Sigma$	1	1	]0,85 ; 0,95]	0,0012	-0,0164	0,0014
			]0,95 ; 1]	0,0000	-0,0153	0,0014
			$\Sigma$	1,0000	0,9039	1,0000
				$B^* =$	0,8657	0,0778

vitäts- und Normierungsbedingung, was bei der Lösung mittels Maximum-Entropie-Methode nicht der Fall ist, aber das Bestimmtheitsmaß spricht in diesem Fall eindeutig zugunsten der Tikhonov-Regularisierung. Aber auch ohne Bestimmtheitsmaß ist erkennbar, daß die Lösung der Maximum-Entropie-Methode definitiv unbrauchbar ist. Da dieser Umstand nicht unwesentlich darauf zurückzuführen ist, daß der Stichprobenumfang (zu) klein bemessen ist, wollen wir uns vor weiteren Analysen ansehen, was passiert, wenn ein größerer Stichprobenumfang ( $N = 1000$ ) gewählt wird.

Auch in unserem dritten Fall verletzt die Lösung mittels Tikhonov-Regularisierung wieder die Anforderungen an die Positivität und die Normiertheit, beide werden im Gegensatz dazu von der Maximum-Entropie-Lösung erfüllt. Das Bestimmtheitsmaß ist allerdings bei der Maximum-Entropie-Lösung etwas schlechter ausgefallen als bei der Tikhonov-Regularisierung. In diesem Fall scheint aber auch die Maximum-Entropie-Lösung noch brauchbar zu sein.

Bei der Analyse der Ergebnisse ist folgendes zu beachten: Das zweite Beispiel liefert aufgrund des geringen Stichprobenumfanges für die Maximum-Entropie-Methode unbrauchbare Ergebnisse, bei der Verwendung dieser Methode ist also auf jeden Fall auf einen ausreichenden Stichprobenumfang zu achten. Beim Vergleich der Ergebnisse ist zu berücksichtigen, daß wir die Lösung der Maximum-Entropie-Methode mit der Lösung einer Tikhonov-Regularisierung mit einer geeigneten Parameterwahl vergleichen. Genauer gesagt ist diese Parameterwahl optimal in Bezug auf das Bestimmtheitsmaß und damit ist im praxisrelevanten Fall diese Parameterwahl gar nicht möglich. Darüber hinaus haben wir im Fall der Tikhonov-Regularisierung - im Gegensatz zur Maximum-Entropie-Methode - immer die Erfüllung der Ungleichung  $\|w - w_\delta\| \leq \delta$  gefordert, weil dies ist für die Bestimmung eines geeigneten Regularisierungsparameters notwendig war (Standfest,



Tabelle 7: Lösungsvergleich für gestörte Daten,  $N = 1000$ 

			Intervalle	$h_d(\pi)$	$\hat{h}(\pi)$	
					Tikhonov	MEM
					$\alpha = 0,0027$	
			[0 ; 0,05]	0,0038	-0,0016	0,0243
			]0,05 ; 0,15]	0,0700	0,0769	0,0405
			]0,15 ; 0,25]	0,1698	0,1564	0,1038
$x$	$w(x)$	$w_\delta(x)$	]0,25 ; 0,35]	0,2241	0,1990	0,2233
0	0,212	0,197	]0,35 ; 0,45]	0,2158	0,2037	0,2893
1	0,318	0,312	]0,45 ; 0,55]	0,1635	0,1746	0,1978
2	0,273	0,285	]0,55 ; 0,65]	0,0973	0,1211	0,0761
3	0,152	0,154	]0,65 ; 0,75]	0,0427	0,0579	0,0214
4	0,045	0,052	]0,75 ; 0,85]	0,0117	0,0050	0,0070
$\Sigma$	1	1	]0,85 ; 0,95]	0,0012	-0,0124	0,0051
			]0,95 ; 1]	0,0000	0,0080	0,0115
			$\Sigma$	1,0000	0,9888	1,0000
				$B^* =$	0,9867	0,9212

1995). Anders ausgedrückt verwenden wir bei der Tikhonov-Regularisierung mehr Informationen, optimieren dieses Verfahren zusätzlich in Hinblick auf unser Bewertungskriterium und vergleichen dieses optimierte Verfahren mit der Maximum-Entropie-Methode, die mit weniger Information auskommt.

## 5 Resümee

Ausgangspunkt der Überlegungen war eine Wahrscheinlichkeitsfunktion, von der man annimmt, daß es sich um die Realisation einer Mischung handelt. Ziel ist die Erhebung der Mischanteile, die zu der beobachteten Verteilung geführt haben, oder anders ausgedrückt, die Berechnung der Parameterverteilung. Die Lösung unseres Problems sollte also nicht nur den beobachteten Daten entsprechen, sondern auch die Anforderungen an eine Verteilung, also Positivität und Normiertheit, erfüllen.

Bisherige Überlegungen von Sixtl und Standfest führten zu unbefriedigenden Lösungen, weil sie nicht alle Anforderungen erfüllen konnten. Mit der Maximum-Entropie-Methode konnte nun ein Verfahren gefunden werden, welches alle Anforderungen erfüllt. Aber auch diese Methode kann nicht immer alle Anforderungen zur vollsten Zufriedenheit erfüllen. Bei der Verwendung dieser Methode sollte man sich immer folgende grundsätzlichen Aussagen vor Auge halten:

- Die gefundenen Wahrscheinlichkeiten stellen die vorurteilsloseste und gleichförmigste Verteilung dar, die mit den gegebenen Nebenbedingungen vereinbar ist.

- Die Maximum-Entropie-Methode führt nur bei großem Stichprobenumfang zu brauchbaren Ergebnissen.
- Die Maximum-Entropie-Methode liefert in manchen Fällen bezüglich des Bestimmtheitsmaßes schlechtere Resultate als die Tikhonov-Regularisierung, aber hat gegenüber dieser folgende Vorteile:
  - Positivität und Normiertheit der Lösung
  - weniger aufwendig in der Berechnung
  - keine zusätzlichen Vorkenntnisse notwendig (z. B. über Datenfehler)

## Literatur

J.N. Kapur. *Maximum entropy models in science engineering*. John Wiley & Sons, New Delhi, 1989.

J. Reiter. *Bildverschärfung durch Lösung der Fredholmschen Integralgleichung 1. Art mittels der Maximum-Entropie-Methode mit astronomischen Anwendungen*. PhD thesis, Innsbruck, 1985.

F. Sixtl. *Der Mythos des Mittelwertes: Neue Methodenlehre der Statistik*. Oldenbourg, München, Wien, 1993.

T. Standfest. *Die Tikhonov-Regularisierung in Verbindung mit einem Regressionsverfahren zur Dekomposition von Mischverteilungen*. Master's thesis, Linz, 1995.

Adresse der Autorin:

Dipl.-Ing. Mag. Dr. Christine Duller  
IFAS - Institut für Angewandte Statistik  
Johannes Kepler Universität Linz  
Altenbergerstr. 69  
A-4040 Linz

Tel. +43 732 2468 / 9128

Fax +43 732 2468 / 9846

E-mail: christine.duller@jku.at

<http://www.ifas.uni-linz.ac.at>