# Zur Beschreibung und Analyse unscharfer Daten

Reinhard Viertl

*Technische Universität Wien*

## 1 Einleitung

Das Resultat einer Messung einer kontinuierlichen Größe ist keine exakte Zahl, sondern mehr oder weniger unscharf. Diese Unschärfe ist verschieden von Fehlern und kann mittels einer sogenannten unscharfen Zahl (fuzzy number) beschrieben werden. Im Folgenden werden unscharfe Zahlen mathematisch präzisiert und die Verallgemeinerung spezieller Verknüpfungen reeller Zahlen auf den Fall unscharfer Zahlen dargestellt.

## 2 Unscharfe Zahlen

Eine exakte Zahl $x_0 \in \mathbb{R}$ ist in eineindeutiger Weise durch die *Indikatorfunktion* $I_{\{x_0\}}(\cdot)$, also eine reelle Funktion, deren Funktionswerte folgendermaßen definiert sind,

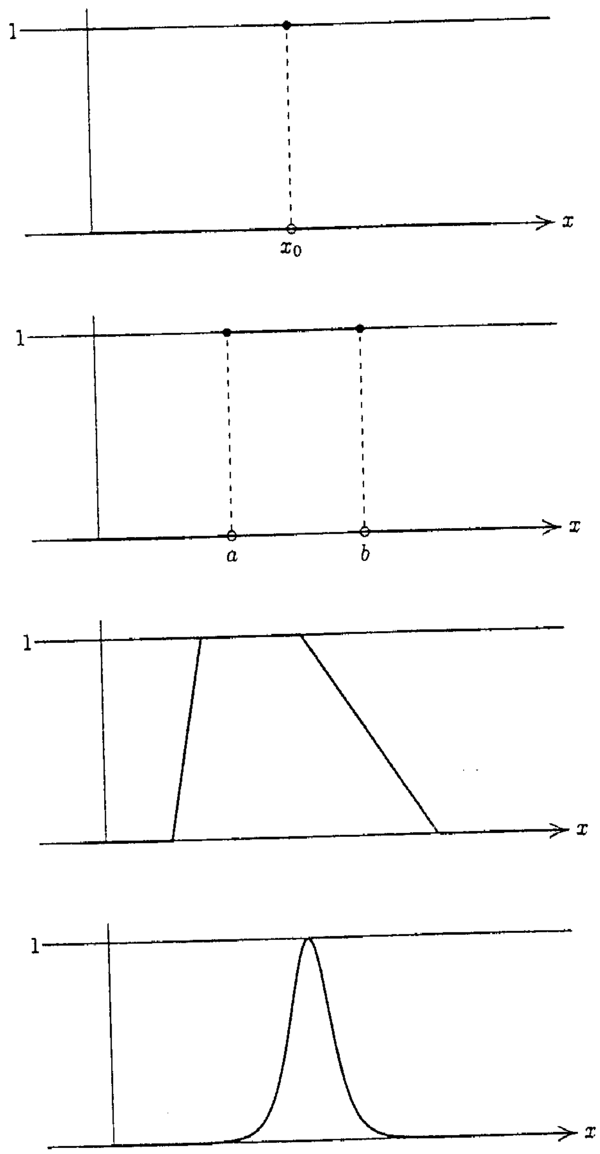
$$I_{\{x_0\}}(x) = \begin{cases} 1 & \text{für} \quad x = x_0 \\ 0 & \text{für} \quad x \neq x_0, \end{cases}$$

bestimmt. In analoger Weise ist die Indikatorfunktion $I_{[a,b]}(\cdot)$ eines Intervalles $[a,b]$ durch ihre Funktionswerte folgendermaßen bestimmt

$$I_{[a,b]}(x) = \begin{cases} 1 & \text{für} \quad x \in [a,b] \\ 0 & \text{für} \quad x \notin [a,b]\,. \end{cases}$$

Allgemein ist die Indikatorfunktion $I_A(\cdot)$ einer nichtleeren Teilmenge $A \subseteq \mathbb{R}$

Abbildung 1: *Beispiele charakterisierender Funktionen*




folgendermaßen durch ihre Funktionswerte definiert

$$I_A(x) = \begin{cases} 1 & \text{für} \quad x \in A \\ 0 & \text{für} \quad x \in \mathbb{R} \setminus A \ . \end{cases}$$

Es war die Idee von L.A. Zadeh, Mengen, die nicht scharf abgegrenzt sind, durch eine Verallgemeinerung von Indikatorfunktionen zu beschreiben. Dabei sind als Funktionswerte auch Zahlen zwischen 0 und 1 zulässig. Die so entstehenden Funktionen werden *Zugehörigkeitsfunktionen* genannt. Eine Zugehörigkeitsfunktion $\mu(\cdot)$ einer *unscharfen Teilmenge* $A^*$ einer gegebenen Menge $M$ ist eine Funktion

$$\mu : M \longrightarrow [0, 1].$$

*Unscharfe Zahlen* $x^*$ sind spezielle unscharfe Teilmengen der Menge $\mathbb{R}$ der reellen Zahlen, für deren Zugehörigkeitsfunktionen $\xi(\cdot)$ folgendes gefordert wird:

(1) $\exists x_0 \in \mathbb{R}\colon\ \xi(x_0) = 1$

(2) $\forall \alpha \in (0, 1]$ ist der sogenannte *$\alpha$-Schnitt* $C_\alpha(x^*) := \{x \in \mathbb{R}\colon\ \xi(x) \geq \alpha\}$ ein abgeschlossenes und beschränktes Intervall $[a_\alpha, b_\alpha]$.

Solche Zugehörigkeitsfunktionen, welche obige Bedingungen (1) und (2) erfüllen, nennt man *charakterisierende Funktionen.*

**Bemerkung** : Eine unscharfe Zahl $x^*$ ist durch ihre charakterisierende Funktion bestimmt.

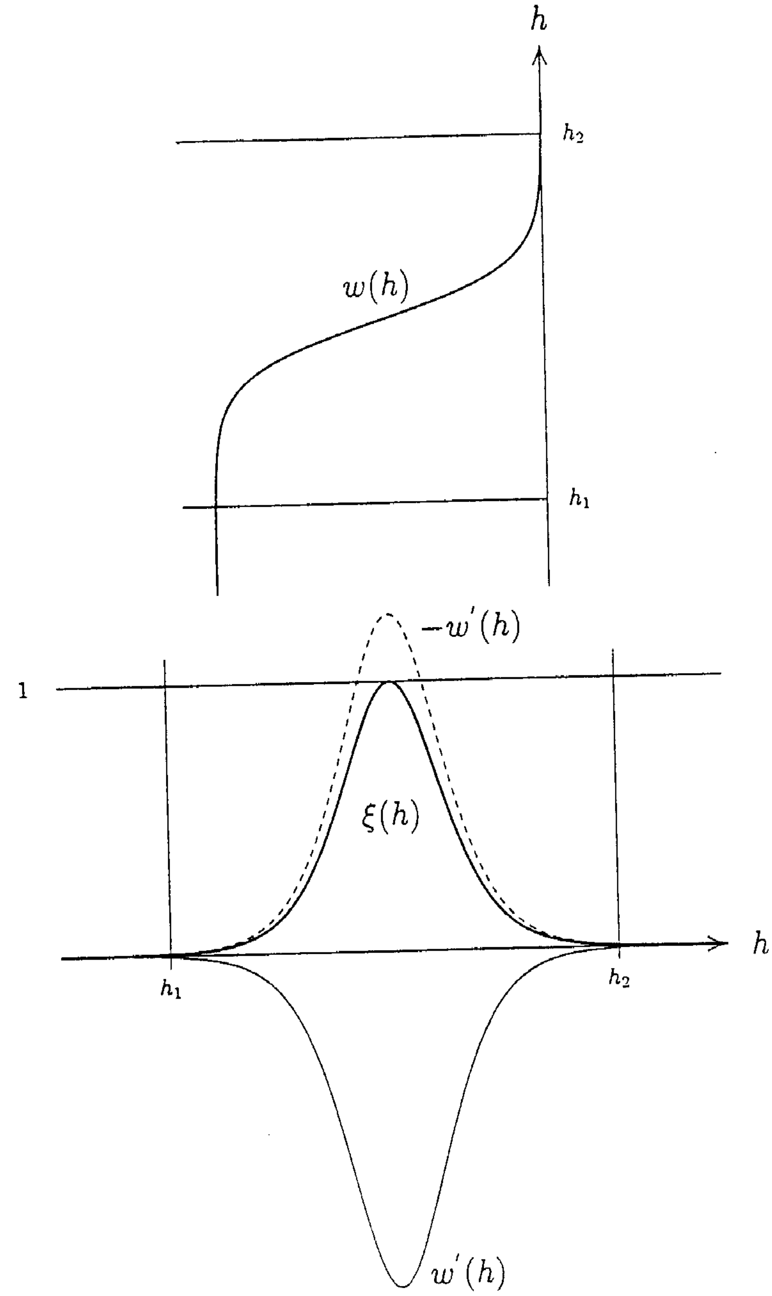
Beispiele von charakterisierenden Funktionen unscharfer Zahlen sind in Abbildung 1 dargestellt.

Ein wesentliches Problem ist die Frage, wie man die charakterisierende Funktion einer unscharfen Messung erhält. Dies ist vom konkreten Meßvorgang abhängig. Ein Beispiel ist die Beobachtung des Wasserstandes eines Flusses zu einer bestimmten Zeit an einem bestimmten Ort. Dazu dient eine Meßlatte, die verschieden naß ist. Man kann hier die Feuchtigkeitsintensität $w(h)$ der Meßlatte heranziehen. (vgl. Abb. 2, oberes Diagramm).

Um die charakterisierende Funktion $\xi(h)$ der unscharfen Höhe $h^*$ des Wasserstandes zu erhalten, differenziert man die Funktion $w(h)$, multipliziert das Resultat $w'(h)$ mit $(-1)$ und dividiert diese Funktion $-w'(h)$ durch ihr Maximum und erhält so die charakterisierende Funktion des unscharfen Wasserstandes, symbolisch geschrieben

$$\xi(h) = \frac{-w'(h)}{\max\limits_{h_1 \leq h \leq h_2} \{-w'(h)\}} \qquad \text{für} \quad h \in \mathbb{R}\,.$$

Abbildung 2: *Charakterisierende Funktion eines unscharfen Wasserstandes*



Für zweidimensionale Größen, wie z. B. die Lage eines Punktes auf einem Radarschirm, benötigt man das Konzept eines *unscharfen Vektors* $\underline{x}^*$.Unscharfe Vektoren sind durch zugehörige *vektorcharakterisierende Funktionen* $\xi_{\underline{x}^*}(\cdot,\cdot)$ beschrieben. Dies sind Funktionen von 2 reellen Variablen $x$ und $y$ mit Werten im Intervall $[0,1]$, die folgende Eigenschaften erfüllen:

(v1) $\exists (x_0, y_0) \in \mathbb{R}^2 : \ \xi(x_0, y_0) = 1$

(v2) $\forall \alpha \in (0,1]$ ist der sogenannte Schnitt $C_\alpha(\underline{x}^*) := \left\{(x,y) \in \mathbb{R}^2 : \ \xi_{\underline{x}^*}(x,y) \geq \alpha\right\}$ eine abgeschlossene, beschränkte und sternförmige Teilmenge des $\mathbb{R}^2$.

**Beispiel:** Um die verktorcharakterisierende Funktion $\xi(\cdot,\cdot)$ eines unscharfen Lagevektors aus einem Lichtpunkt zu erhalten, kann man die Helligkeitsintensität $h(x,y)$ heranziehen. Die Funktionswerte $\xi(x,y)$ von $\xi(\cdot,\cdot)$ erhält man durch

$$\xi(x,y) = \frac{h(x,y)}{\max\limits_{(x,y)\in\mathbb{R}^2} h(x,y)} \qquad \forall \quad (x,y) \in \mathbb{R}^2.$$

**Bemerkung:** Im Hinblick auf $n$-dimensionale unscharfe Größen ist folgende Bezeichnungsweise nützlich:

$$\underline{x} = (x,y) \in \mathbb{R}^2$$

allgemein

$$\underline{x} = (x_1, \cdots, x_n) \in \mathbb{R}^n$$

# 3 Unscharfe Funktionen

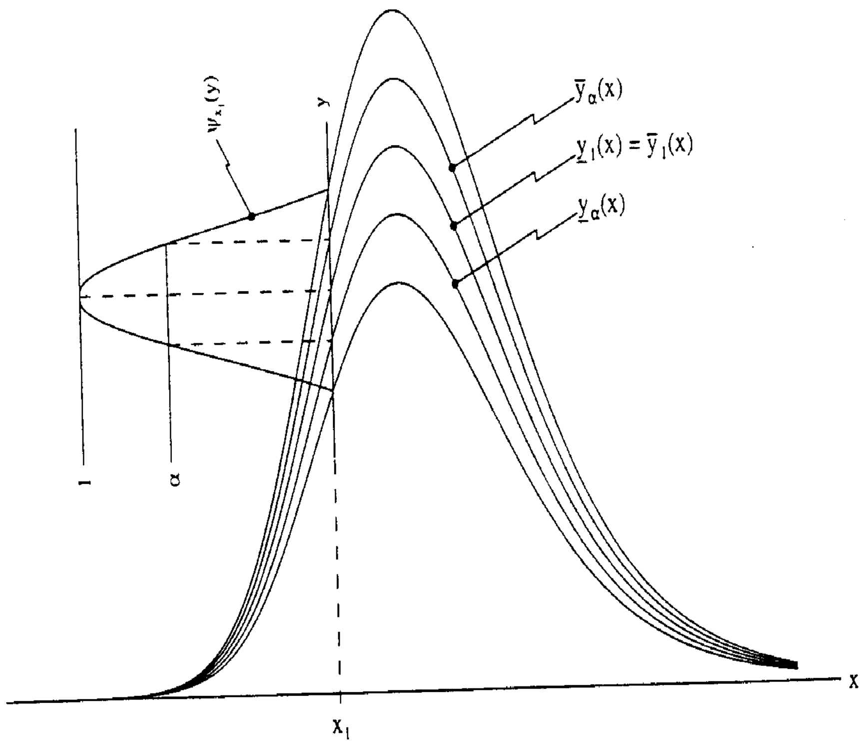
Bei der Beschreibung von Zeitverläufen von Messungen benötigt man das Konzept von Funktionen mit unscharfen Funktionswerten. In Verallgemeinerung von reellen Funktionen $y(x)$ wird jedem $x \in \mathbb{R}$ eine unscharfe Zahl $y^*(x)$ zugeordnet. Eine unscharfe reelle Funktion $y^*(\cdot)$ ist also durch eine Familie unscharfer Funktionswerte $y^*(x)$ mit entsprechenden charakterisierenden Funktionen $\psi_x(\cdot)$ bestimmt:

$$\left(y^*(x); \ x \in [a,b]\right)$$

Zur grafischen Darstellung unscharfer Funktionen dienen die sogenannten *$\alpha$-Niveaukurven* $\overline{y}_\alpha(\cdot)$ bzw. $\underline{y}_\alpha(\cdot)$, welche die jeweiligen Enden der $\alpha$-Schnitte

$\left[\underline{y}_\alpha(x), \overline{y}_\alpha(x)\right]$ der unscharfen Funktionswerte $y^*(x)$ verbindet. Für praktische Darstellungen werden einige $\alpha$-Werte (z. B. 5 solche) herangezogen. Ein Beispiel ist in Abb. 3 dargestellt.

Abbildung 3: *$\alpha$-Niveaukurven einer unscharfen Funktion*




Ein wichtiges Beispiel unscharfer Funktionen sind Zeitverläufe von Schadstoffkonzentrationen in verschiedenen Medien.

Will man daraus Gesamtschadstoffmengen ermitteln, hat man das Problem der Integration von unscharfen Funktionen, d. h.

$$\int_{t_1}^{t_2} y^*(t)dt.$$

Methoden dazu existieren und ein Literaturhinweis ist [6]. Dabei werden $\alpha$-Niveaukurven verwendet, von denen vorausgesetzt wird, daß sie klassische integrierbare Funktionen sind.

# 4 Funktionen von unscharfen Argumenten

Bei der Analyse unscharfer Daten treten Funktionen von unscharfen Argumenten auf. Daher ist eine Verallgemeinerung klassischer, reeller Funktionen $f(\cdot)$ auf den Fall unscharfer Argumentwerte $x^*$ notwendig.

Im Fall einer reellen Variablen $x$ ist eine Unschärfe von $x^*$ durch eine charakterisierende Funktion $\xi(\cdot)$ gegeben. Der Funktionswert $f(x^*)$ wird auch eine unscharfe Größe $y^*$, deren charakterisierende Funktion $\psi(\cdot)$ mittels des sogenannten *Fortsetzungsprinzips* von L.A. Zadeh gegeben ist. Dieses definiert die Zugehörigkeitsfunktion $\psi(\cdot)$ folgendermaßen:

$$\psi(y) = \left\{ \begin{array}{lll} \sup\{\xi(x) : f(x) = y\} & \text{falls} & f^{-1}(\{y\}) \neq \emptyset \\ 0 & \text{falls} & f^{-1}(\{y\}) = \emptyset \end{array} \right\} \quad \forall\, y \in I\!R$$

**Bemerkung:** Die Funktion $\psi(\cdot)$ ist eine Zugehörigkeitsfunktion einer unscharfen Teilmenge von $I\!R$. Für allgemeine Funktionen $f(\cdot)$ folgt nicht zwingend, daß $\psi(\cdot)$ auch eine charakterisierende Funktion einer unscharfen Zahl ist. Für stetige Funktionen $f(\cdot)$ gilt folgender

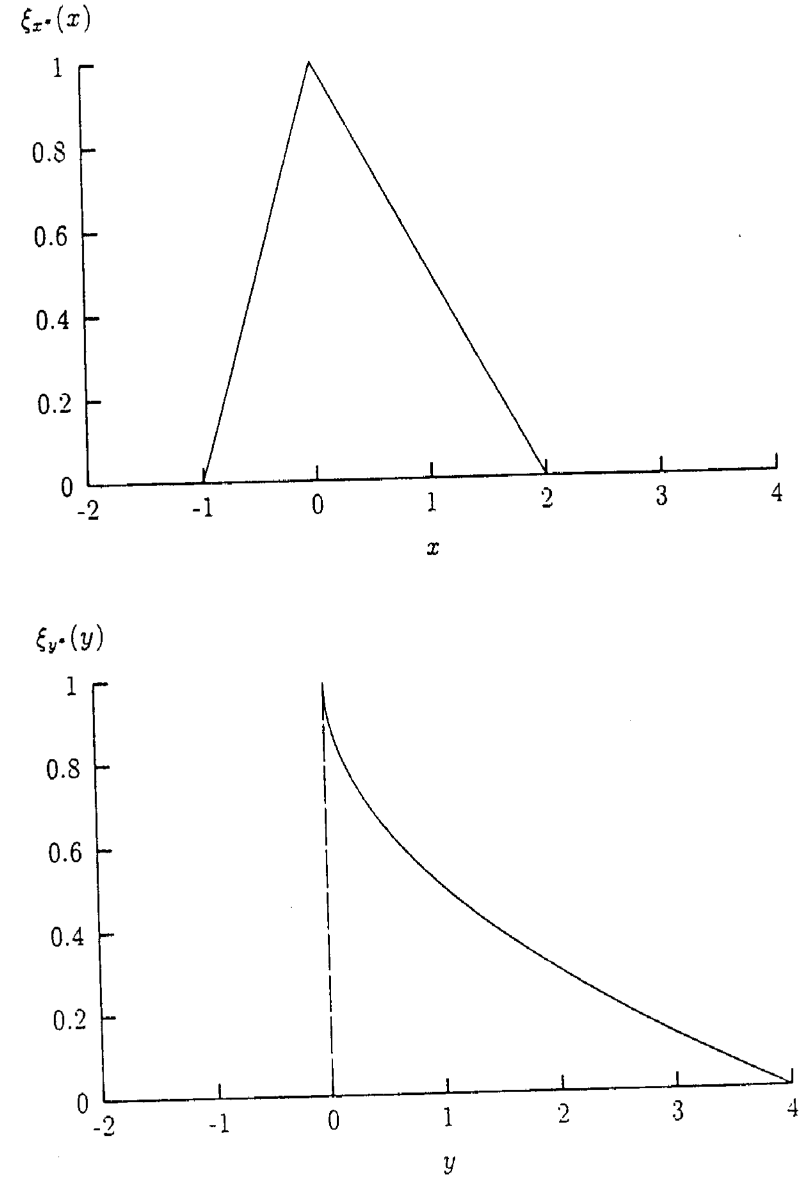
**Satz:** Ist $\underline{x}^*$ ein unscharfer Vektor mit vektorcharakterisierender Funktion $\xi(\cdot,\ldots,\cdot)$ und $f(\cdot,\ldots,\cdot)$ eine stetige, klassische reellwertige Funktion, so ist die oben beschriebene Funktion $\psi(\cdot)$ eine charakterisierende Funktion im Sinne von Abschnitt 1 und für die $\alpha$-Schnitte $C_\alpha\big(f(\underline{x}^*)\big)$ gilt:

$$C_\alpha\big(f(\underline{x}^*)\big) = \left[\min_{\underline{x}\in C_\alpha(\underline{x}^*)} f(\underline{x}),\ \max_{\underline{x}\in C_\alpha(\underline{x}^*)} f(\underline{x})\right] \qquad \forall \quad \alpha \in (0,1].$$

Den Beweis findet man in [4].

**Beispiel:** Die charakterisierende Funktion des Quadrates einer unscharfen Zahl ist in Abbildung 4 dargestellt. Der obere Teil der Abbildung zeigt die charakterisierende Funktion der unscharfen Zahl, der untere Teil die charakterisierende Funktion ihres Quadrates.

Abbildung 4: *Quadrat $y^* = (x^*)^2$ einer unscharfen Zahl $x^*$*



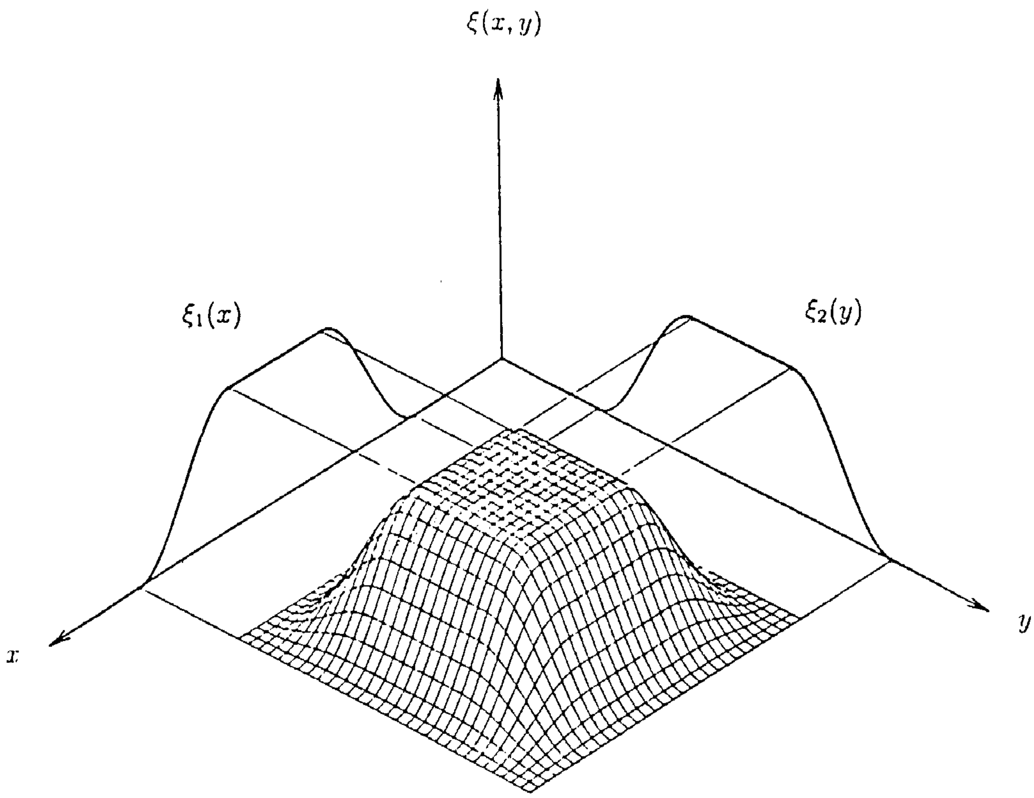
# 5 Addition unscharfer Zahlen

Um unscharfe Zahlen $x^*$ und $y^*$ zu addieren kann man das Fortsetzungsprinzip heranziehen. Die verallgemeinerte Addition $x^* \oplus y^*$ liefert als Resultat eine unscharfe Zahl $z^* = x^* \oplus y^*$, deren charakterisierende Funktion folgendermaßen gefunden werden kann: Um das Fortsetzungsprinzip anwenden zu können, müssen die unscharfen Zahlen $x^*$ und $y^*$ zuerst zu einem unscharfen Vektor $\underline{x}^* = (x, y)^*$ kombiniert werden. Die vektorcharakterisierende Funktion $\xi(\cdot, \cdot)$ des unscharfen Vektors $\underline{x}^*$ erhält man aus den charakterisierenden Funktionen $\xi_1(\cdot)$ von $x^*$ und $\xi_2(\cdot)$ von $y^*$ durch folgende Kombination:

$$\xi(x, y) = \min\{\xi_1(x), \xi_2(y)\} \qquad \forall \quad (x, y) \in \mathbb{R}^2$$

In Abbildung 5 ist die Kombination zweier unscharfer Zahlen zu einem unscharfen Vektor durch die charakterisierenden Funktionen und die erzeugte vektorcharakterisierende Funktion dargestellt.

Abbildung 5: *Kombination zweier unscharfer Zahlen*



Mit Hilfe der vektorcharakterisierenden Funktion von $\underline{x}^*$ kann durch die Anwendung des Fortsetzungsprinzips die charakterisierende Funktion $\zeta(\cdot)$ der unscharfen Summe $x^* \oplus y^*$ ermittelt werden:

$$\zeta(z) := \left\{ \begin{array}{ll} \sup\{\xi(x,y) : \; x+y=z\} & \text{falls } \exists (x,y): \; x+y=z \\ 0 & \text{sonst} \end{array} \right\}$$

**Bemerkung:** Es läßt sich zeigen, daß die so entstandene Funktion alle Eigenschaften einer charakterisierenden Funktion hat. Die Erweiterung anderer algebraischer Operationen für den Fall unscharfer Zahlen ist mit Hilfe des unscharfen kombinierten Vektors und des Fortsetzungsprinzips möglich. Eine konkrete Fragestellung zu dieser Aufgabe ist die Ermittlung der Fläche eines Rechteckes, dessen Seitenlängen unscharf sind.

# 6 Zur statistischen Analyse unscharfer Daten

Auch Beobachtungen von kontinuierlichen stochastischen Größen $X$, wie beispielsweise Lebensdauern, sind meist unscharf (nicht zu verwechseln mit Fehlern). Daher ist die Adaption statistischer Methoden für den Fall unscharfer Daten notwendig.

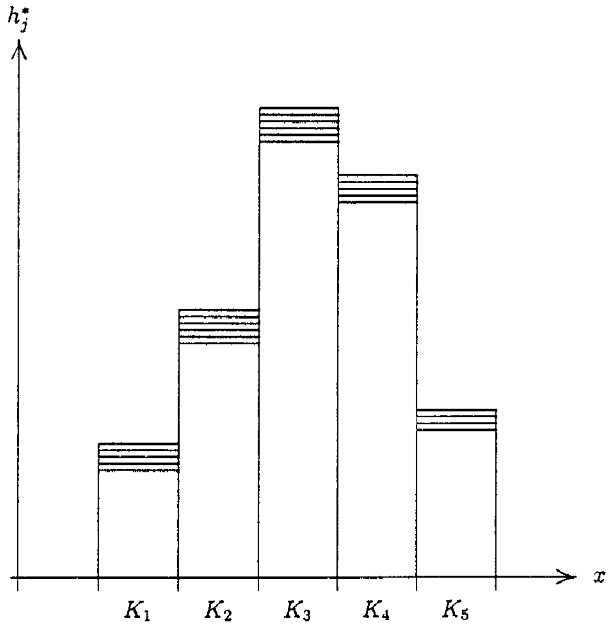
## 6.1 Beschreibende Statistik

Bereits in der beschreibenden Statistik, z. B. bei Histogrammen, ist die Unschärfe von Daten zu berücksichtigen, sollen nicht unrealistische Resultate entstehen.

Für unscharfe Beobachtungen kann unter Umständen nicht entschieden werden, in welcher Klasse $K_j$ eines zu berechnenden Histogrammes eine solche unscharfe Beobachtung liegt. Aus diesem Grunde wird die Höhe eines Histogrammbalkens eine unscharfe Zahl $h_j^*$, deren charakterisierende Funktion aus den charakterisierenden Funktionen der Beobachtungen ermittelt wird. Grafisch dargestellt erhält man sogenannte *Fuzzy Histogramme*. Ein Beispiel ist in Abbildung 6 dargestellt.

Die Ränder der schraffierten Bereiche der Histogrammhöhen entsprechen der Anzahl jener Beobachtungen, die sicher in der jeweiligen Klasse $K_j$ liegen. bzw. der Anzahl jener Beobachtungen, die nicht sicher außerhalb von $K_j$ liegen. Zur genauen Bestimmung der charakterisierenden Funktion der unscharfen Höhen vgl. die Arbeit [3].

Abbildung 6: *Fuzzy Histogramm*




## 6.2 Schließende Statistik

Eine der wichtigsten Aufgaben der schließenden Statistik ist die Schätzung von Parametern $\theta$ in stochastischen Modellen $X \sim W_\theta$, $\theta \in \Theta$.

Betrachtet man die Schätzung des Erwartungswertes $\mathbb{E}X$ einer eindimensionalen stochastischen Größe $X$ auf Grundlage einer konkreten Stichprobe $x_1, \cdots, x_n$ so ist die im statistischen Sinn beste Schätzung für $\mathbb{E}X$ das sogenannte *Stichprobenmittel*

$$\overline{x}_n = \frac{x_1 + \cdots + x_n}{n}.$$

Allgemein sind Schätzungen Funktionen $\vartheta(x_1, \cdots, x_n)$ von Stichproben.

Im Fall unscharfer Stichproben in Form von unscharfen Zahlen $x_1^*, \cdots, x_n^*$ kann man adaptierte unscharfe Schätzwerte folgendermaßen erhalten. Zuerst müssen die unscharfen Zahlen $x_1^*, \cdots, x_n^*$ zu einem unscharfen $n$-dimensionalen Vektor $\underline{x}^*$ kombiniert werden. Die vektorcharakterisierende Funktion $\xi(\cdot, \cdots, \cdot)$ von $\underline{x}^*$ erhält man aus den $n$ charakterisierenden Funktionen $\xi_1(\cdot), \cdots, \xi_n(\cdot)$ der $x_1^*, \cdots, x_n^*$ folgendermaßen:

$$\xi(x_1, \cdots, x_n) = \min\{\xi_1(x_1), \cdots, \xi_n(x_n)\} \qquad \forall \quad (x_1, \cdots, x_n) \in \mathbb{R}^n$$

Diese vektorcharakterisierende Funktion ermöglicht die Konstruktion eines unscharfen Schätzwertes mit Hilfe des Fortsetzungsprinzips. Man erhält die charakterisierende Funktion $\psi(\cdot)$ der unscharfen Schätzung $\vartheta(x_1^*, \cdots, x_n^*)$ auf Grundlage der unscharfen Beobachtungen (Daten) $x_1^*, \cdots, x_n^*$ folgendermaßen: Unter Verwendung der Vektorschreibweise $\underline{x} = (x_1, \cdots, x_n) \in \mathbb{R}^n$ gilt

$$\psi(z) = \left\{ \begin{array}{ll} \sup\{\xi(\underline{x}) : \vartheta(\underline{x}) = z\} & \text{falls } \exists \underline{x} : \vartheta(\underline{x}) = z \\ 0 & \text{sonst} \end{array} \right\} \quad \forall z \in \mathbb{R}.$$

An Stelle eines exakten Schätzwertes $\hat{\theta} \in \Theta$ für einen Parameter $\theta_0$ bei theoretisch exakten Daten erhält man bei unscharfen Daten $x_1^*, \cdots, x_n^*$ einen *unscharfen Schätzwert* $\hat{\theta}^* = \vartheta(x_1^*, \cdots, x_n^*)$, dessen charakterisierende Funktion $\psi(\cdot)$ ist.

Ein Beispiel für unscharfe Daten ist in Abbildung 7 dargestellt. Die charakterisierende Funktion des zugehörigen unscharfen Schätzwertes für den Erwartungswert $\theta = \mathbb{E}X$ ist in Abbildung 8 wiedergegeben.

Abbildung 7: *Charakterisierende Funktionen unscharfer Daten*

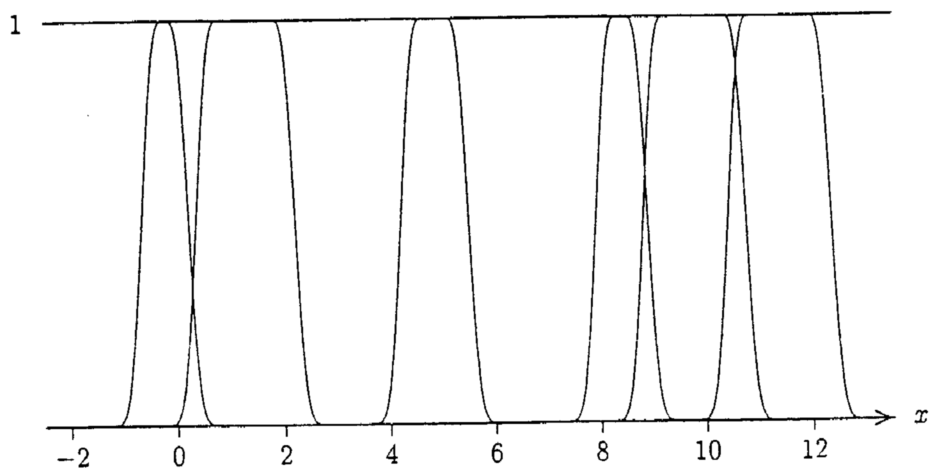
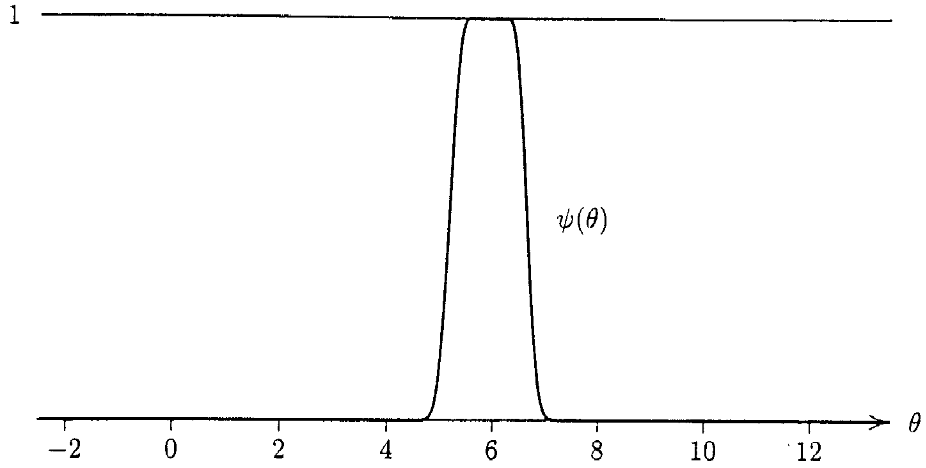


Abbildung 8: *Charakterisierende Funktion der unscharfen Schätzung*




**Bemerkung:** Es gibt Verallgemeinerungen verschiedener statistischer Verfahren auf den realistischen Fall unscharfer Daten. Ansätze dazu findet man in dem deutschsprachigen Buch [5] und ausführliche Verfahren in der Monografie [4].

# Literatur

[1] H. Bandemer, S. Gottwald: *Einführung in Fuzzy-Methoden*, 4. Auflage, Akademie Verlag, Berlin, 1993

[2] H. Bandemer, W. Näther: *Fuzzy Data Analysis*, Kluwer Academic Publ., Dordrecht, 1992

[3] R. Viertl, Statistics with Non-Precise Data, *Journal of Computing and Information Technology*, 4 (1996)

[4] R. Viertl: *Statistical Methods for Non-Precise DATA*, CRC Press, Boca Raton (Florida), 1996

[5] R. Viertl: *Einführung in die Stochastik mit Elementen der Bayes-Statistik und Ansätzen für die Analyse unscharfer Daten*, 2. Auflage, Springer-Verlag, Wien, 1997

[6] R. Viertl, S. Bodjanova: *Calculation of Intagrals of Fuzzy Functions*, Forschungsbericht RIS –1998 – 1, Institut für Statistik, Technische Universität Wien, 1998

Adresse des Autors:

O. Univ.-Prof. Dipl.-Ing. Dr. Reinhard Viertl
Institut für Statistik und Wahrscheinlichkeitstheorie
Technische Universität Wien
A-1040 Wien