OTTO VON GUERICKE UNIVERSITÄT MAGDEBURG
INF FAKULTÄT FÜR INFORMATIK

Fuzzy Systems

Possibility Theory

Prof. Dr. Rudolf Kruse **Christoph Doell**

{kruse,doell}@iws.cs.uni-magdeburg.de
Otto-von-Guericke University of Magdeburg
Faculty of Computer Science
Department of Knowledge Processing and Language Engineering

Outline

A Simple Example

Oil contamination of water by trading vessels

Typical formulation:
"The accident occurred *approximately* 10 miles away from the coast."

Locations of interest: *open sea* (z3), *12-mile zone* (z2), *3-mile zone* (z1), *canal* (ca), *refueling dock* (rd), *loading dock* (ld)

These 6 locations Ω are disjoint and exhaustive

$$\Omega = \{z3, z2, z1, ca, rd, ld\}$$

Modeling Partial Belief

(Possibly vague) statements are often not simply true or false

Decision maker should be able to **quantify** his/her degree of belief

This can be objective measurement or subjective valuation

Probability theory

Sample space Θ (finite set of distinct possible outcomes of some random experiment)

Event $A \subseteq \Theta$

For any Θ, probability P is assumed to be $P : 2^{\Theta} \to [0, 1]$ satisfying Kolmogorov axioms

Kolmogorov Axioms

For any Θ, real-valued function $P : 2^{\Theta} \to [0, 1]$ must satisfy

i) $0 \leq P(A) \leq 1$ for all events $A \subseteq \Theta$,

ii) $P(\Theta) = 1$,

iii) if $A \cap B = \emptyset$, then $P(A \cup B) = P(A) + P(B)$

Partial Belief and Evidence Masses

We may not conceive elements of Θ but their observations of some space Ω

Mapping Γ attaches to each sensor $\theta \in \Theta$ its "output" (either element or subset or fuzzy set of Ω)

Probability P on Ω induces via Γ a structure on Ω

This structure represents **partial beliefs** about actual state of world ω_0

If $\Gamma : \Theta \to 2^{\Omega}$, then (P, Γ) is "random set"

If subjective valuation is quantified, then **evidence masses** are attached to subsets of Ω

Thus, expert must partition "belief", attributing bigger amounts to more reliable pieces of knowledge

Mass Distribution

Recall example with $\Omega = \{z3, z2, z1, ca, rd, ld\}$

Propositional statement *in port* equals event $\{ca, rd, ld\}$

Event may represent maximum level of differentiation for expert

Expert specifies **mass distribution** $m : 2^\Omega \to [0, 1]$

Here, Ω is called **frame of discernment**

$m : 2^\Omega \to [0, 1]$ must satisfy

(i) $m(\emptyset) = 0$,

(ii) $\sum_{A:A\subseteq\Omega}(A) = 1$

Subsets $A \subseteq \Omega$ with $m(A) > 0$ are called **focal elements** of m

Belief and Plausibility

$m(A)$ measures belief committed *exactly* to A

For *total* amount of belief (**credibility**) of A, sum up $m(B)$ whereas $B \subseteq A$

For *maximum* amount of belief movable to A, sum up $m(B)$ with $B \cap A \neq \emptyset$

This leads to **belief function** and **plausibility function**

$$\mathrm{Bel}_m : \quad 2^\Omega \to [0,1], \quad \mathrm{Bel}_m(A) = \sum_{B : B \subseteq A} m(B)$$

$$\mathrm{Pl}_m : \quad 2^\Omega \to [0,1], \quad \mathrm{Pl}_m(A) = \sum_{B : B \cap A \neq \emptyset} m(B)$$

Belief and Plausibility

In any case $\mathrm{Bel}(\Omega) = 1$ ("closed world" assumption)

Total ignorance modeled by $m_0 : 2^\Omega \to [0, 1]$ with $m_0(\Omega) = 1$, $m_0(A) = 0$ for all $A \neq \Omega$

m_0 leads to $\mathrm{Bel}(\Omega) = \mathrm{Pl}(\Omega) = 1$ and $\mathrm{Bel}(A) = 0, \mathrm{Pl}(A) = 1$ for all $A \neq \Omega$

For ordinary probability, use $m_1 : 2^\Omega \to [0, 1]$ with $m_1(\{\omega\}) = p_\omega$ and $m_1(A) = 0$ for all sets A with $|A| > 1$

m_1 is called Bayesian belief function

Exact knowledge modeled by $m_2 : 2^\Omega \to [0, 1]$, $m_2(\{\omega_0\}) = 1$ and $m_2(A) = 0$ for all $A \neq \{\omega_0\}$

Example

Consider statement: "ship is *in port* with degree of certainty of 0.6, further evidence is not available"

Mass distribution
$m : 2^{\Omega} \to [0,1], m(\{\text{in port}\}) = 0.6, m(\Omega) = 0.4, m(A) = 0$ otherwise

$m(\Omega) = 0.4$ represents inability to attach that amount of mass to any $A \subset \Omega$

e.g. $m(\{\overline{\text{in port}}\}) = 0.4$ would exceed expert's statement

Outline

Nested Focal Elements

Besides Bayesian belief functions, there is another important case

If focal elements $A_1, \ldots, A_n$ of m are nested, *i.e.* can be arranged such that $A_1 \subseteq A_2 \subseteq \ldots \subseteq A_n$, then belief function is called **consonant**

Simplest case: belief function with only 1 focal element $A \subseteq \Omega$ where $m(A) = 1$

Then it follows

$$\mathrm{Bel}_m(B) = \begin{cases} 1 & \text{if } A \subseteq B, \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad \mathrm{Pl}_m(B) = \begin{cases} 1 & \text{if } A \cap B \neq \emptyset, \\ 0 & \text{otherwise} \end{cases}$$

Each $\omega \in A$ is candidate for being actual but unknown state ω_0

Possibility Distribution

We definitely know that $\omega_0 \notin \bar{A}$

Can be used for $\rho : \Omega \to \{0, 1\}, \rho(\omega) = \mathrm{Pl}(\{\omega\})$

ρ attaches 1 to possible and 0 to impossible elements

ρ is thus named **possibility distribution**

Possibility and Necessity

Truth of "possibly $\omega_0 \in B$" is called **possibility** of $B \subseteq \Omega$

This is true if $\max\{\rho(\omega) \mid \omega \in \Omega\} = 1$

Truth of "necessarily $\omega_0 \in \Omega$" is called **necessity** of $B \subseteq \Omega$

This is true if $\max\{\rho(\omega) \mid \omega \in \bar{B}\} = 0$

"necessarily $\omega_0 \in \Omega$" being true requires" possibly $\omega_0 \in \bar{B}$" being false

Possibility and Necessity Measures

Generally, possibility (and necessity) becomes matter of degree

Instead of $\rho : \Omega \to \{0, 1\}$, membership function $\pi : \Omega \to [0, 1], \pi(\omega) = \mathrm{Pl}(\{\omega\})$

Thus, **possibility measure** and **necessity measure** are defined as

$$\begin{aligned} \Pi_m: &\quad 2^\Omega \to [0,1], &\quad \Pi_m(B) &= \max\{\pi(\omega) : \omega \in B\} \\ \mathrm{nec}_m: &\quad 2^\Omega \to [0,1], &\quad \mathrm{nec}_m(B) &= 1 - \Pi(\bar{B}) \end{aligned}$$

Π and nec are plausibility and belief functions, resp.

Example – Probability vs. Possibility

Statement $A(n)$: "Anna ate n eggs for breakfast."

(Subjective) probability $P(A(n))$ can be determined by experiments: "How many eggs will Anna eat for today's breakfast?"

possibility $\pi(A(n))$: "How many eggs can Anna eat for breakfast."

n	1	2	3	4	5	6	7	8
$\pi(A(n))$	1	1	1	1	.8	.6	.4	.2
$P(A(n))$	.1	.8	.1	0	0	0	0	0

a possible event needs not to be probable

a probable event is always possible

Properties of Possibility Measures

i) $\Pi(\emptyset) = 0$

ii) $\Pi(\Omega) = 1$

iii) $\Pi(A \cup B) = \max\{\Pi(A), \Pi(B)\}$ for all $A, B \subseteq \Omega$

Possibility of some set is determined by its "most possible" element

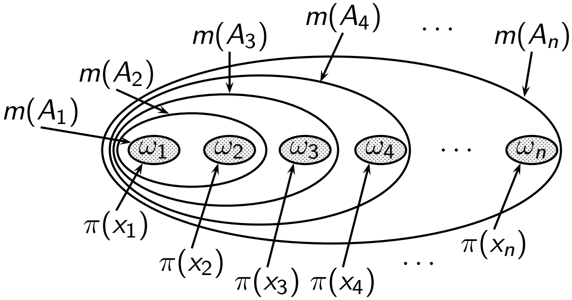
$\text{nec}(\Omega) = 1 - \Pi(\emptyset) = 1$ means closed world assumption:
"necessarily $\omega_0 \in \Omega$" must be true

Total ignorance: $\Pi(B) = 1, \text{nec}(B) = 0$ for all $B \neq \emptyset, B \neq \Omega$

Perfect knowledge: $\Pi(\{\omega\}) = \text{nec}(\{\omega\}) = 0$ for all $\omega \neq \omega_0$ and
$\Pi(\{\omega_0\}) = \text{nec}(\{\omega_0\}) = 1$

Nested Focal Elements

complete sequence of nested focal elements of Π on 2^{Ω} where $\Omega = \{\omega_1, \ldots, \omega_n\}$



Example

Consider ship locations again

Given membership function

$$\begin{aligned}\pi(z3) = \pi(z2) &= 0\\ \pi(z1) = \pi(ld) &= 0.3\\ \pi(ca) &= 0.6\\ \pi(rd) &= 0.1\end{aligned}$$

$\Pi(\{z3, z2\}) = 0$ and $\mathrm{nec}(\{z1, ca, rd, ld\}) = 1$

We know it is impossible that ship is located in $\{z3, z2\}$

$\omega_0 \in \{z1, ca, rd, ld\}$

$\Pi(\{ca, rd\}) = 1, \mathrm{nec}(\{ca, rd\}) = 0.7$ means "location of ship is possibly but not with certainty in $\{ca, rd\}$"

Outline

1. Partial Belief

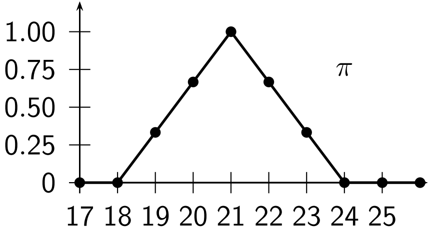
2. Possibility and Necessity

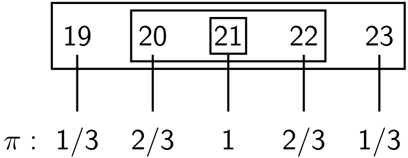
3. **Possibility Theory**
   - Axiomatic Approach
   - The Context Model
   - Possibility Distributions
   - Reasoning
   - Evidence Propagation

Possibility and Fuzzy Sets

Let variable T be temperature in $^\circ C$ (only integers)

Current but unknown value T_0 is given by "T is around $21^\circ C$"





Incomplete information induces possibility distribution function π

π is numerically identical with membership function

Nested α-cuts play same role as focal elements

Possibility Theory

Best-known calculus for handling uncertainty: **probability theory** [Laplace, 1812]

Less well-known, but noteworthy alternative: **possibility theory** [Dubois and Prade, 1988]

Possibility theory can handle **uncertain and imprecise information**, while probability theory was only designed to handle *uncertain information*

Possibility Theory: Axiomatic Approach

Definition

Let Ω be a (finite) sample space. A **possibility measure** Π on Ω is a function $\Pi : 2^\Omega \to [0, 1]$ satisfying

i) $\Pi(\emptyset) = 0$ and

ii) $\forall E_1, E_2 \subseteq \Omega : \Pi(E_1 \cup E_2) = \max\{\Pi(E_1), \Pi(E_2)\}$.

Similar to Kolmogorov's axioms of probability theory

From axioms, it follows $\Pi(E_1 \cap E_2) \leq \min\{\Pi(E_1), \Pi(E_2)\}$

Attributes are introduced as random variables (as in probability theory)

$\Pi(A = a)$ is abbreviation of $\Pi(\{\omega \in \Omega \mid A(\omega) = a\})$

If event E is possible without restriction, then $\Pi(E) = 1$

If event E is impossible, then $\Pi(E) = 0$

Possibility Theory and the Context Model

Interpretation of degrees of possibility [Gebhardt and Kruse, 1993]

Let Ω be (nonempty) set of all possible states of world, ω_0 the actual (but unknown) state

Let $C = \{c_1, \dots, c_n\}$ be set of contexts (observers, frame conditions etc.) and $(C, 2^C, P)$ finite probability space (context weights)

Let $\Gamma : C \to 2^\Omega$ be set-valued mapping, which assigns to each context the **most specific correct set-valued specification of** ω_0

Sets $\Gamma(c)$ are called **focal sets** of Γ

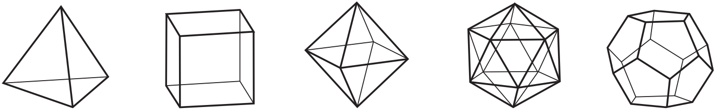
Γ is **random set** (*i.e.* set-valued random variable) [Nguyen, 1978]

Basic possibility assignment induced by Γ is mapping

$$\begin{aligned} \pi : \Omega &\to [0, 1] \\ \pi(\omega) &\mapsto P(\{c \in C \mid \omega \in \Gamma(c)\}). \end{aligned}$$

Example: Dice and Shakers

shaker 1	shaker 2	shaker 3	shaker 4	shaker 5



tetrahedron	hexahedron	octahedron	icosahedron	dodecahedron
1 – 4	1 – 6	1 – 8	1 – 10	1 – 12

numbers	degree of possibility
1 – 4	$\frac{1}{5}+\frac{1}{5}+\frac{1}{5}+\frac{1}{5}+\frac{1}{5} = 1$
5 – 6	$\frac{1}{5}+\frac{1}{5}+\frac{1}{5}+\frac{1}{5} = \frac{4}{5}$
7 – 8	$\frac{1}{5}+\frac{1}{5}+\frac{1}{5} = \frac{3}{5}$
9 – 10	$\frac{1}{5}+\frac{1}{5} = \frac{2}{5}$
11 – 12	$\frac{1}{5} = \frac{1}{5}$

Definition

Let $\Gamma : C \to 2^{\Omega}$ be a random set. The **possibility measure** induced by Γ is the mapping

$$\begin{aligned} \Pi : 2^{\Omega} &\to [0,1], \\ E &\mapsto P(\{c \in C \mid E \cap \Gamma(c) \neq \emptyset\}). \end{aligned}$$

Problem: from given interpretation it only follows:

$$\forall E \subseteq \Omega : \quad \max_{\omega \in E} \pi(\omega) \leq \Pi(E) \leq \min\left\{1, \sum_{\omega \in E} \pi(\omega)\right\}.$$

	1	2	3	4	5
$c_1 : \frac{1}{2}$			•		
$c_2 : \frac{1}{4}$		•	•	•	
$c_3 : \frac{1}{4}$	•	•	•	•	•
π	0	$\frac{1}{2}$	1	$\frac{1}{2}$	$\frac{1}{4}$

	1	2	3	4	5
$c_1 : \frac{1}{2}$			•		
$c_2 : \frac{1}{4}$	•	•			
$c_3 : \frac{1}{4}$				•	•
π	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{4}$

From Context Model to Possibility Measures

Attempts to solve indicated problem:

- require focal sets to be **consonant**
  $\rightarrow$ mass assignment theory [Baldwin et al., 1996]
  problem: "voting model" is not sufficient to justify consonance
- use lower bound as "most pessimistic" choice [Gebhardt, 1997]
  problem: basic possibility assignments represent negative information, lower bound is actually *most optimistic* choice
- justify lower bound from decision making purposes

From Context Model to Possibility Measures

Assume: in the end we must decide on one single event

Each event is described by values of set of attributes

Then it can be useful to assign to set of events the degree of possibility of “most possible” event in set

example:

$\sum$					
	36	18	18	28	
28	0	0	0	28	28
18	18	0	0	0	18
18	18	0	0	0	18
36	0	18	18	0	18
max	18	18	18	28	

0	40	0	40
40	0	0	40
0	0	20	20
40	40	20	max

Definition

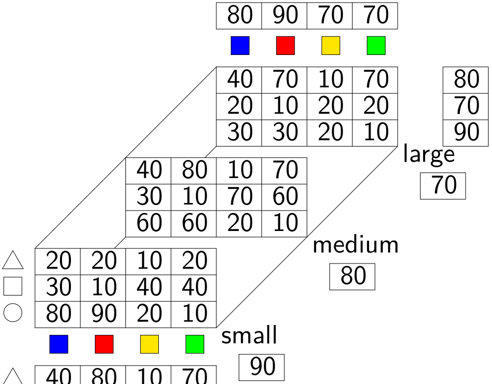
Let $X = \{A_1, \ldots, A_n\}$ be a set of attributes defined on a (finite) sample space Ω with respective domains $\text{dom}(A_i)$, $i = 1, \ldots, n$.
A **possibility distribution** π_X over X is the restriction of a possibility measure Π on Ω to the set of all events that can be defined by stating values for all attributes in X. That is, $\pi_X = \Pi|_{\mathcal{E}_X}$, where

$$\begin{aligned}
\mathcal{E}_X &= \left\{ E \in 2^\Omega \;\middle|\; \exists a_1 \in \text{dom}(A_1) : \ldots \exists a_n \in \text{dom}(A_n) : \right. \\
&\qquad \left. E \mathrel{\widehat{=}} \bigwedge_{A_j \in X} A_j = a_j \right\} \\
&= \left\{ E \in 2^\Omega \;\middle|\; \exists a_1 \in \text{dom}(A_1) : \ldots \exists a_n \in \text{dom}(A_n) : \right. \\
&\qquad \left. E = \left\{ \omega \in \Omega \;\middle|\; \bigwedge_{A_j \in X} A_j(\omega) = a_j \right\} \right\}.
\end{aligned}$$

Corresponds to the notion of probability distribution

A Possibility Distribution

all numbers in parts per 1000



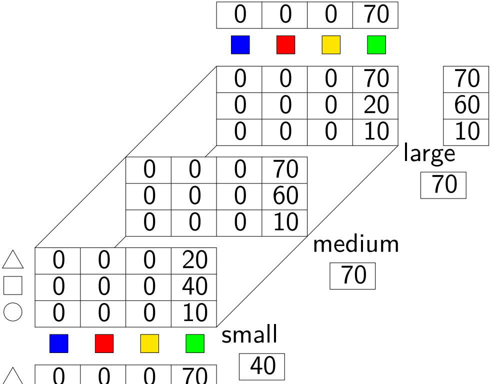
	blue	red	yellow	green
△	40	80	10	70
□	30	10	70	60
○	80	90	20	10

	s	m	l
△	20	80	70
□	40	70	20
○	90	60	30

	blue	red	yellow	green
large	40	70	20	70
medium	60	80	70	70
small	80	90	40	40

Numbers state degrees of possibility of corresponding value combination

Reasoning



△	0	0	0	70
□	0	0	0	60
○	0	0	0	10

all numbers in parts per 1000

	s	m	l
△	20	70	70
□	40	60	20
○	10	10	10

	blue	red	yellow	green
large	0	0	0	70
medium	0	0	0	70
small	0	0	0	40

Using information that given object is green

Possibilistic Decomposition

As for relational and probabilistic networks, 3D possibility distribution can be decomposed into projections to subspaces, *i.e.*

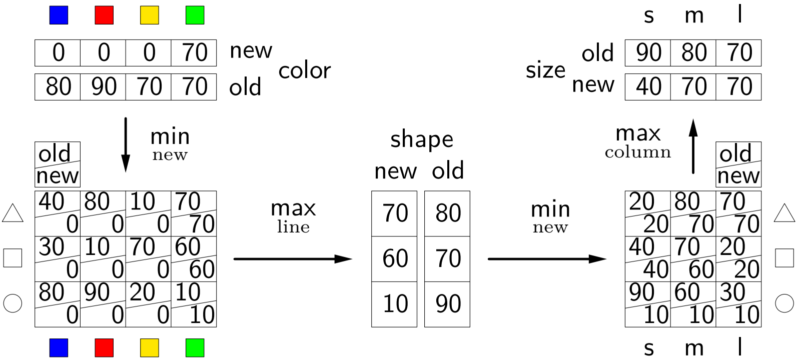
- maximum projection to subspace color $\times$ shape,
- maximum projection to subspace shape $\times$ size

Can be reconstructed using the following formula:

$$\begin{aligned}
\forall i, j, k :\ & \pi\left(a_i^{(\text{color})}, a_j^{(\text{shape})}, a_k^{(\text{size})}\right) \\
&= \min\left\{\pi\left(a_i^{(\text{color})}, a_j^{(\text{shape})}\right), \pi\left(a_j^{(\text{shape})}, a_k^{(\text{size})}\right)\right\} \\
&= \min\left\{\max_k \pi\left(a_i^{(\text{color})}, a_j^{(\text{shape})}, a_k^{(\text{size})}\right),\right. \\
&\qquad\quad \left.\max_i \pi\left(a_i^{(\text{color})}, a_j^{(\text{shape})}, a_k^{(\text{size})}\right)\right\}
\end{aligned}$$

Reasoning with Projections

Again same result can be obtained using only projections to subspaces (maximal degrees of possibility):



Conditional Possibility and Independence

Definition

Let Ω be a (finite) sample space, Π a possibility measure on Ω, and $E_1, E_2 \subseteq \Omega$ events. Then $\Pi(E_1 \mid E_2) = \Pi(E_1 \cap E_2)$ is called the **conditional possibility** of E_1 given E_2.

Definition

Let Ω be a (finite) sample space, Π a possibility measure on Ω, and A, B, and C attributes with respective domains $\text{dom}(A)$, $\text{dom}(B)$, and $\text{dom}(C)$. A and B are called **conditionally possibilistically independent** given C, written $A \perp\!\!\!\perp_{\Pi} B \mid C$, iff

$$\forall a \in \text{dom}(A) : \forall b \in \text{dom}(B) : \forall c \in \text{dom}(C) :$$

$$\Pi(A = a, B = b \mid C = c) = \min\{\Pi(A = a \mid C = c), \Pi(B = b \mid C = c)\}$$

similar to corresponding notions of probability theory

Possibilistic Evidence Propagation

A:	color
B:	shape
C:	size

$$\pi(B = b \mid A = a_{\text{obs}})$$

$$= \pi\left(\bigvee_{a \in \text{dom}(A)} A = a, B = b, \bigvee_{c \in \text{dom}(C)} C = c \,\middle|\, A = a_{\text{obs}} \right)$$

$$\stackrel{(1)}{=} \max_{a \in \text{dom}(A)} \{ \max_{c \in \text{dom}(C)} \{ \pi(A = a, B = b, C = c \mid A = a_{\text{obs}}) \} \}$$

$$\stackrel{(2)}{=} \max_{a \in \text{dom}(A)} \{ \max_{c \in \text{dom}(C)} \{ \min\{ \pi(A = a, B = b, C = c), \pi(A = a \mid A = a_{\text{obs}}) \} \} \}$$

$$\stackrel{(3)}{=} \max_{a \in \text{dom}(A)} \{ \max_{c \in \text{dom}(C)} \{ \min\{ \pi(A = a, B = b), \pi(B = b, C = c), \pi(A = a \mid A = a_{\text{obs}}) \} \} \}$$

$$= \max_{a \in \text{dom}(A)} \{ \min\{ \pi(A = a, B = b), \pi(A = a \mid A = a_{\text{obs}}), \underbrace{\max_{c \in \text{dom}(C)} \{ \pi(B = b, C = c) \} }_{= \pi(B = b) \geq \pi(A = a, B = b)} \} \}$$

$$= \max_{a \in \text{dom}(A)} \{ \min\{ \pi(A = a, B = b), \pi(A = a \mid A = a_{\text{obs}}) \} \}$$

Literature for the Lecture

Baldwin, J. F., Lawry, J., and Martin, T. P. (1996).
A mass assignment theory of the probability of fuzzy events.
Fuzzy Sets and Systems, 83:353–367.

Dubois, D. and Prade, H. (1988).
Possibility Theory: An Approach to Computerized Processing of Uncertainty.
Plenum Press, New York, USA.

Gebhardt, J. (1997).
Learning from Data: Possibilistic Graphical Models.
Habilitation, TU Braunschweig.

Gebhardt, J. and Kruse, R. (1993).
The context model: An integrating view of vagueness and uncertainty.
International Journal of Approximate Reasoning, 9:283–314.

Laplace, P.-S. (1812).
Théorie analytique des probabilités.
Madame Veuve Courcier, Paris, France.

Nguyen, H. T. (1978).
On random sets and belief functions.
Journal of Mathematical Analysis and Applications, 65:531–542.