LFSC 2013

On some extensions and applications of non-additive measures

Vicenç Torra

September, 2013

Joint Work: Y. Narukawa, K. Stokes, G. Navarro-Arribas, D. Abril

Institut d'Investigació en Intel·ligència Artificial (IIIA-CSIC), Bellaterra

A (short) motivation

A short motivation

Topic: Non-additive (fuzzy) measures

- A generalization of additive measures (probabilities)

A short motivation

Topic: Non-additive (fuzzy) measures

- A generalization of additive measures (probabilities)
- Equivalent terms: non-additive measures, fuzzy measures, capacities

A short motivation

Why these measures are studied?

- Mathematical interest
- Applications

A short motivation

Why these measures are studied?

- Mathematical interest

A short motivation

Why these measures are studied?

- Mathematical interest
  - Properties

A short motivation

Why these measures are studied?

- Mathematical interest
  - Properties
    - ⋆ Equalities and Inequalities (e.g. Chebyshev type inequalities)
    - ⋆ Measures and distances (e.g. entropy / Hellinger distance)

A short motivation

Why these measures are studied?

- Mathematical interest
  - Properties
    - ⋆ Equalities and Inequalities (e.g. Chebyshev type inequalities)
    - ⋆ Measures and distances (e.g. entropy / Hellinger distance)
  - Constructions

A short motivation

Why these measures are studied?

- Mathematical interest
  - Properties
    - ⋆ Equalities and Inequalities (e.g. Chebyshev type inequalities)
    - ⋆ Measures and distances (e.g. entropy / Hellinger distance)
  - Constructions
    - ⋆ Integrals with respect to these measures (e.g. Choquet integrals)

A short motivation

Why these measures are studied?

- Applications

A short motivation

Why these measures are studied?

- Applications
  - Some problems that cannot be solved with additive measures (i.e., probabilities) can be solved with non-additive measures.

A short motivation

Why these measures are studied?

- Applications
  - Some problems that cannot be solved with additive measures (i.e., probabilities) can be solved with non-additive measures.
    - ⋆ Decision making
    - ⋆ Subjective evaluation
    - ⋆ Data fusion
    - ⋆ Computer vision
    - ⋆ Distances

A short motivation

Why these measures are studied?

- Applications
  - Some problems that cannot be solved with additive measures (i.e., probabilities) can be solved with non-additive measures.
    - ⋆ Decision making
    - ⋆ Subjective evaluation
    - ⋆ Data fusion
    - ⋆ Computer vision
    - ⋆ Distances

      → a common theme:
      to take into account interactions

      → a common advantage:
      more expressive power than with additive models

A short motivation

Why these measures are studied?

- Decision making
  - Criteria to order our car preferences: price, quality, and confort
    assign to each car $c_i \in Cars$ utility values $u_p(c_i), u_q(c_i), u_c(c_i)$
    assign importances to each criteria (or subset of criteria)
    and combine values w.r.t. importances to find a global value (and order)
- Data fusion
  - Sensors give distances to the nearest object: s_1, s_2, s_3
    assign importances to sensors (or subsets of sensors)
    and combine values w.r.t. importances to find a reliable value

A short motivation

Why this talk focuses on distorted probabilities?

A short motivation

Why this talk focuses on distorted probabilities?

- Additive measures on X defined in terms of $|X|$ values

A short motivation

Why this talk focuses on distorted probabilities?

- Additive measures on X defined in terms of $|X|$ values
  We need a probability value for each $x \in X$
  $\rightarrow$ so, $|X|$ values

A short motivation

Why this talk focuses on distorted probabilities?

- Additive measures on X defined in terms of $|X|$ values
  We need a probability value for each $x \in X$
  $\rightarrow$ so, $|X|$ values
- Non-additive measures on X defined in terms of $2^{|X|}$ values

A short motivation

Why this talk focuses on distorted probabilities?

- Additive measures on X defined in terms of $|X|$ values
  We need a probability value for each $x \in X$
  $\rightarrow$ so, $|X|$ values
- Non-additive measures on X defined in terms of $2^{|X|}$ values
  We need a value for each $A \subset X$
  $\rightarrow$ so, $2^{|X|}$ values

A short motivation

Why this talk focuses on distorted probabilities?

- Additive measures on X defined in terms of $|X|$ values
  We need a probability value for each $x \in X$
  $\rightarrow$ so, $|X|$ values
- Non-additive measures on X defined in terms of $2^{|X|}$ values
  We need a value for each $A \subset X$
  $\rightarrow$ so, $2^{|X|}$ values

$\rightarrow$ distorted probabilities as a compact representation of (some) non-additive measure
$\longrightarrow$ useful for applications

A short motivation

The only compact representation?

A short motivation

The only compact representation?

- No!!

  $\rightarrow$ there are other families of measures. E.g.,

A short motivation

The only compact representation?

- No!!

  $\rightarrow$ there are other families of measures. E.g.,

  - $\perp$-decomposable fuzzy measures ($\mu(A \cup B) = \mu(A) \oplus \mu(B)$)
  - Sugeno λ-measures ($\mu(A \cup B) = \mu(A) + \mu(B) + \lambda\mu(A)\mu(B)$)
  - k-additive fuzzy measures (in terms of the Möbius transform)
  - Hierarchically decomposable fuzzy measures ($\oplus_i$ + hierarchy of X)

A short motivation

The only compact representation?

- No!!

  $\rightarrow$ there are other families of measures. E.g.,

  - $\perp$-decomposable fuzzy measures ($\mu(A \cup B) = \mu(A) \oplus \mu(B)$)
  - Sugeno λ-measures ($\mu(A \cup B) = \mu(A) + \mu(B) + \lambda\mu(A)\mu(B)$)
  - k-additive fuzzy measures (in terms of the Möbius transform)
  - Hierarchically decomposable fuzzy measures ($\oplus_i$ + hierarchy of X)

Common theme: Reduce the number of parameters

A short motivation

The only compact representation?

- No!!

  → there are other families of measures. E.g.,
  - $\perp$-decomposable fuzzy measures ($\mu(A \cup B) = \mu(A) \oplus \mu(B)$)
  - Sugeno λ-measures ($\mu(A \cup B) = \mu(A) + \mu(B) + \lambda\mu(A)\mu(B)$)
  - k-additive fuzzy measures (in terms of the Möbius transform)
  - Hierarchically decomposable fuzzy measures ($\oplus_i$ + hierarchy of X)

Common theme: Reduce the number of parameters

Distorted probability: simple

Outline

1. Short Motivation
2. Introduction (definitions)
3. Applications (short)
4. Distorted Probabilities
5. m-dimensional Distorted Probabilities
6. Distorted Probabilities and Multisets
7. Summary

Introduction

Definitions

Additive measures: $(X, \mathcal{A})$ a measurable space; then, a set function μ is an additive measure if it satisfies

(i) $\mu(A) \geq 0$ for all $A \in \mathcal{A}$,
(ii) $\mu(X) \leq \infty$
(iii) $\mu(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i)$ for every countable sequence A_i $(i \geq 1)$ of $\mathcal{A}$ that is pairwise disjoint (i.e,. $A_i \cap A_j = \emptyset$ when $i \neq j$).

Definitions

Additive measures: $(X, \mathcal{A})$ a measurable space; then, a set function μ is an additive measure if it satisfies

(i) $\mu(A) \geq 0$ for all $A \in \mathcal{A}$,
(ii) $\mu(X) \leq \infty$
(iii) $\mu(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i)$ for every countable sequence A_i $(i \geq 1)$ of $\mathcal{A}$ that is pairwise disjoint (i.e,. $A_i \cap A_j = \emptyset$ when $i \neq j$).

Finite case: $\mu(A \cup B) = \mu(A) + \mu(B)$ for disjoint A, B

Definitions

Non-additive measures: $(X, \mathcal{A})$ a measurable space, a non-additive (fuzzy) measure μ on $(X, \mathcal{A})$ is a set function $\mu : \mathcal{A} \rightarrow [0, 1]$ satisfying the following axioms:

(i) $\mu(\emptyset) = 0$, $\mu(X) = 1$ (boundary conditions)
(ii) $A \subseteq B$ implies $\mu(A) \leq \mu(B)$ (monotonicity)

Differences

Non-additive measures vs. additive measures:

- In additive measures: $\mu(A) = \sum_{x \in A} p_x$

Differences

Non-additive measures vs. additive measures:

- In additive measures: $\mu(A) = \sum_{x \in A} p_x$
- In non-additive measures: additivity no longer a constraint
  $\rightarrow$ three cases possible
  - $\mu(A) = \sum_{x \in A} p_x$

Differences

Non-additive measures vs. additive measures:

- In additive measures: $\mu(A) = \sum_{x \in A} p_x$
- In non-additive measures: additivity no longer a constraint
  $\rightarrow$ three cases possible
  - $\mu(A) = \sum_{x \in A} p_x$
  - $\mu(A) < \sum_{x \in A} p_x$

Differences

Non-additive measures vs. additive measures:

- In additive measures: $\mu(A) = \sum_{x \in A} p_x$
- In non-additive measures: additivity no longer a constraint
  $\rightarrow$ three cases possible
  - $\mu(A) = \sum_{x \in A} p_x$
  - $\mu(A) < \sum_{x \in A} p_x$
  - $\mu(A) > \sum_{x \in A} p_x$

Differences

Non-additive measures vs. additive measures:

- Is non-additivity useful ?

Differences

Non-additive measures vs. additive measures:

- Is non-additivity useful ?
  Yes

Differences

Non-additive measures vs. additive measures:

- Is non-additivity useful ?
  Yes
- Why?

Differences

Non-additive measures vs. additive measures:

- Is non-additivity useful ?
  Yes
- Why?
  some cases represent interactions
  - $\mu(A) = \sum_{x \in A} p_x$ (no interaction)

Differences

Non-additive measures vs. additive measures:

- Is non-additivity useful ?
  Yes
- Why?
  some cases represent interactions
  - $\mu(A) = \sum_{x \in A} p_x$ (no interaction)
  - $\mu(A) < \sum_{x \in A} p_x$ (negative interaction)

Differences

Non-additive measures vs. additive measures:

- Is non-additivity useful ?
  Yes
- Why?
  some cases represent interactions
  - $\mu(A) = \sum_{x \in A} p_x$ (no interaction)
  - $\mu(A) < \sum_{x \in A} p_x$ (negative interaction)
  - $\mu(A) > \sum_{x \in A} p_x$ (positive interaction)

Number of parameters

Non-additive measures vs. additive measures:

- How to define an additive measure on X?

Number of parameters

Non-additive measures vs. additive measures:

- How to define an additive measure on X?
  One probability value for each element
  $\rightarrow |X|$ values

Number of parameters

Non-additive measures vs. additive measures:

- How to define an additive measure on X?
  One probability value for each element
  $\rightarrow |X|$ values
- How to define a non-additive measure?

Number of parameters

Non-additive measures vs. additive measures:

- How to define an additive measure on X?
  One probability value for each element
  $\rightarrow |X|$ values
- How to define a non-additive measure?
  One value for each set
  $\rightarrow 2^{|X|}$ values

What can we do with a measure?

What can we do with a measure?

Non-additive measures and additive measures:

- Integrate a function f with respect to a measure:

What can we do with a measure?

Non-additive measures and additive measures:

- Integrate a function f with respect to a measure:
  - Integral w.r.t. additive measure p

What can we do with a measure?

Non-additive measures and additive measures:

- Integrate a function f with respect to a measure:
  - Integral w.r.t. additive measure p
    $\rightarrow$ expectation

$$\sum p_x f(x)$$

$\longrightarrow$ Lebesgue integral (continuous case: $\int f dp$)

What can we do with a measure?

Non-additive measures and additive measures:

- Integrate a function f with respect to a measure:
  - Integral w.r.t. additive measure p
    $\rightarrow$ expectation
    $$\sum p_x f(x)$$
    $\longrightarrow$ Lebesgue integral (continuous case: $\int f dp$)
  - Integral w.r.t. non-additive measure μ
    $\rightarrow$ expectation like
    $$\sum_{i=1}^{N} f(x_{\sigma(i)})[\mu(A_{\sigma(i)}) - \mu(A_{\sigma(i-1)})]$$
    $\longrightarrow$ Choquet integral (continuous case: $(C) \int f d\mu$)

What can we do with a measure?

Non-additive measures and additive measures:

- Integrate a function f with respect to a measure:
  - Integral w.r.t. additive measure p
    $\rightarrow$ expectation

$$\sum p_x f(x)$$

    $\longrightarrow$ Lebesgue integral (continuous case: $\int f dp$)
  - Integral w.r.t. non-additive measure μ
    $\rightarrow$ expectation like

$$\sum_{i=1}^{N} f(x_{\sigma(i)})[\mu(A_{\sigma(i)}) - \mu(A_{\sigma(i-1)})]$$

    $\longrightarrow$ Choquet integral (continuous case: $(C) \int f d\mu$)

The Choquet integral is a Lebesgue integral when the measure is additive

Applications

Example

Decision making:

Ellsberg paradox (Ellsberg, 1961), an urn, 90 balls ...

Color of balls	Red	Black	Yellow
Number of balls	30	60	
f_R	\$ 100	0	0
f_B	\$ 0	\$ 100	0
f_{RY}	\$ 100	0	\$ 100
f_{BY}	\$ 0	\$ 100	\$ 100

- Usual (most people's) preferences
  - $f_B \prec f_R$

Example

Decision making:

Ellsberg paradox (Ellsberg, 1961), an urn, 90 balls ...

Color of balls	Red	Black	Yellow
Number of balls	30	60	
f_R	\$ 100	0	0
f_B	\$ 0	\$ 100	0
f_{RY}	\$ 100	0	\$ 100
f_{BY}	\$ 0	\$ 100	\$ 100

- Usual (most people's) preferences
  - $f_B \prec f_R$
  - $f_{RY} \prec f_{BY}$

Example

Decision making:

Ellsberg paradox (Ellsberg, 1961), an urn, 90 balls ...

Color of balls	Red	Black	Yellow
Number of balls	30	60	
f_R	\$ 100	0	0
f_B	\$ 0	\$ 100	0
f_{RY}	\$ 100	0	\$ 100
f_{BY}	\$ 0	\$ 100	\$ 100

- Usual (most people's) preferences
  - $f_B \prec f_R$
  - $f_{RY} \prec f_{BY}$
- No solution exist with additive measures, but can be solved with non-additive ones

Applications

Subjective Evaluation:

Subjective evaluation and application field for non-additive (fuzzy) measures from the beginning.
(Sugeno, 1974, p.2):

> "The purposes of this dissertation are to propose the concept of fuzzy measures and integrals [11,12] as a way for expressing human subjectivity and to discuss their applications."

Applications

Definition of Subjective Evaluation: (Dubois and Prade, 1997)

> "Formally speaking, the subjective evaluation problem can be viewed as the synthesis, the identification of a function which maps the attribute values describing the situation to evaluate into a discrete domain (classification), or a continuous one (absolute evaluation). More generally, we may look for the degree of membership of the situation to a category, or have a function yielding a fuzzy evaluation. This function is in general not available as such, but is implicitly, and partially, described in terms of criteria, or by means of expert rules, or through some fuzzy algorithm. It may also happen that the function is only partially known by exemplification through prototypical examples of situations for which the evaluation is available."

Example

Example: (Grabish, 1995) Evaluation of students

- students (A, B, C) on three subjects (M, P, L)
  Ada, Byron, Countess; maths, physics, literature
- Marks:

Student		M	P	L
Ada	f_A	18	16	10
Byron	f_B	10	12	18
Countess	f_C	14	15	15

- Preferences:
  - Assign the same weight to mathematics and physics, and more weight to this subjects than to literature.
  - Represent the following preference on the students:

$$B \prec A \prec C.$$

Example

Example: (Grabish, 1995)

- No solution with additive measures
  We can use non-additive measures (with the Choquet integral)

Distorted Probabilities

Distorted Probabilities: introduction

An open question:

Distorted Probabilities: introduction

An open question:

Non-additive measures vs. additive measures:

How to define a non-additive measure?

One value for each set

$\rightarrow 2^{|X|}$ values

Distorted Probabilities: introduction

An open question:

Non-additive measures vs. additive measures:

How to define a non-additive measure?

One value for each set

$\rightarrow 2^{|X|}$ values

A possible solution:

Distorted Probabilities: introduction

An open question:

Non-additive measures vs. additive measures:

How to define a non-additive measure?

One value for each set

$\rightarrow 2^{|X|}$ values

A possible solution:

Distorted Probabilities.

- Compact representation of non-additive measures:

Distorted Probabilities: introduction

An open question:

Non-additive measures vs. additive measures:

How to define a non-additive measure?

One value for each set

$\rightarrow 2^{|X|}$ values

A possible solution:

Distorted Probabilities.

- Compact representation of non-additive measures:
  - Only $|X|$ values (a probability) and a function (distorting function)

Distorted Probabilities: Definition

- Representation of a fuzzy measure:
  - f and P represent a fuzzy measure μ, iff

$$\mu(A) = f(P(A)) \text{ for all } A \in 2^X$$

  f a real-valued function, P a probability measure on $(X, 2^X)$
  - f is strictly increasing w.r.t. a probability measure P iff $P(A) < P(B)$ implies $f(P(A)) < f(P(B))$
  - f is nondecreasing w.r.t. a probability measure P iff $P(A) < P(B)$ implies $f(P(A)) \leq f(P(B))$

Distorted Probabilities: Definition

- Representation of a fuzzy measure: distorted probability
  - f and P represent a fuzzy measure μ, iff

$$\mu(A) = f(P(A)) \text{ for all } A \in 2^X$$

    f a real-valued function, P a probability measure on $(X, 2^X)$
  - f is strictly increasing w.r.t. a probability measure P iff $P(A) < P(B)$ implies $f(P(A)) < f(P(B))$
  - f is nondecreasing w.r.t. a probability measure P iff $P(A) < P(B)$ implies $f(P(A)) \leq f(P(B))$

  - μ is a distorted probability if μ is represented by a probability distribution P and a function f nondecreasing w.r.t. a probability P.

Distorted Probabilities: Definition

- Representation of a fuzzy measure: distorted probability
  - μ is a distorted probability if μ is represented by a probability distribution P and a function f nondecreasing w.r.t. a probability P.
- So, for a given reference set X we need:
  - Probability distribution on X: $p(x)$ for all $x \in X$
  - Distortion function f on the probability measure: $f(P(A))$

Distorted Probabilities: Application

- Given a distorted probability ...
  ... we can apply any fuzzy integral

- E.g.
  - the Choquet integral
  - the Sugeno integral

Distorted Probabilities: Application

- Distorted probability and Choquet integral:
  - The WOWA operator can be represented as a Choquet integral with a distorted probability.
    - ⋆ WOWA generalizes both the WM and the OWA, using both WM weights and OWA weights.
  - From the distorted probability perspective, in WOWA:
    - ⋆ the WM weights correspond to the probability distribution
    - ⋆ the OWA weights are used to the construct the distortion function

Distorted Probabilities: Properties

- Some distorted probabilities are not decomposable fuzzy measures.
- Some distorted probabilities cannot be represented easily with other families of fuzzy measures $\rightarrow$ they really belong to another family.

Distorted Probabilities: Properties

- Some distorted probabilities are not decomposable fuzzy measures.
- Some distorted probabilities cannot be represented easily with other families of fuzzy measures $\rightarrow$ they really belong to another family.
- 1st. example (I):
  - μ on $X = \{a, b, c\}$ with $p(a) = 0.2$, $p(b) = 0.35$, $p(c) = 0.45$, and

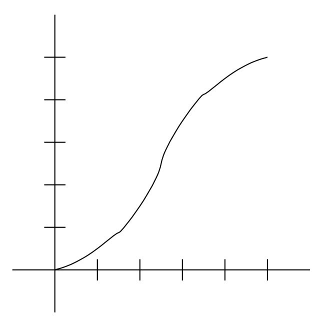
$$f(x) = \begin{cases} 0 & \text{if } x < 0.5 \\ 0.2 & \text{if } 0.5 \le x < 0.6 \\ 0.4 & \text{if } 0.6 \le x < 0.85 \\ 1.0 & \text{if } 0.85 \le x \le 1.0 \end{cases}$$

Distorted Probabilities: Properties

- Some distorted probabilities are not decomposable fuzzy measures.
- Some distorted probabilities cannot be represented easily with other families of fuzzy measures.
- 1st. example (II):
  - $\mu(\emptyset) = 0$, $\mu(\{a\}) = 0$, $\mu(\{b\}) = 0$, $\mu(\{c\}) = 0$,
    $\mu(\{a, b\}) = 0.2$, $\mu(\{a, c\}) = 0.4, \mu(\{b, c\}) = 0.4, \mu(\{a, b, c\}) = 1$
  - μ is a DP but not a $\perp$-decomposable fuzzy measure
    because, there is no t-conorm s.t. $\perp(0, 0) \neq 0$
    $\rightarrow$ as $\mu(\{a, b\}) = 0.2$ when $\mu(\{a\}) = 0$ and $\mu(\{b\}) = 0$,
    we would require $0.2 = \mu(\{a, b\}) = \perp(\mu(\{a\}), \mu(\{b\})) = \perp(0, 0)$.

Distorted Probabilities: Properties

- Some distorted probabilities are not decomposable fuzzy measures.
- Some distorted probabilities cannot be represented easily with other families of fuzzy measures.
- 2nd. example (I):
  - $\mu_{\mathbf{p},\mathbf{w}}$ over $X = \{x_1, x_2, x_3, x_4, x_5\}$ from (probability distribution) $\mathbf{p} = (0.2, 0.3, 0.1, 0.2, 0.1)$, and function (from $\mathbf{w} = (0.1, 0.2, 0.4, 0.2, 0.1)$):



Distorted Probabilities: Properties

- Some distorted probabilities are not decomposable fuzzy measures.
- Some distorted probabilities cannot be represented easily with other families of fuzzy measures.
- 2nd. example (II):
  - $\mu_{\mathbf{p},\mathbf{w}}$ is a 5-additive fuzzy measure because $m(A) \neq 0$ for all A.
  - E.g.,
    $m(\{x_1, x_2, x_3, x_4, x_5\}) = 0.50746528$,
    $m(\{x_1, x_2, x_3, x_4\}) = -0.2537326$.
    There is no k-additive fuzzy measure equivalent to $\mu_{\mathbf{p},\mathbf{w}}$ for $k < 5$.

m-dimensional Distorted Probabilities

m-Dimensional Distorted Probabilities

Justification: Why any extension of distorted probabilities?

m-Dimensional Distorted Probabilities

Justification: Why any extension of distorted probabilities?

The number of distorted probabilities.

m-Dimensional Distorted Probabilities

Justification: Why any extension of distorted probabilities?

The number of distorted probabilities.

Observe the following

- For $X = \{1, 2, 3\}$, 2/8 of distorted probabilities.
- For larger sets X ...
  ... the proportion of distorted probabilities decreases rapidly
- For $\mu(\{1\}) \leq \mu(\{2\}) \leq \ldots$

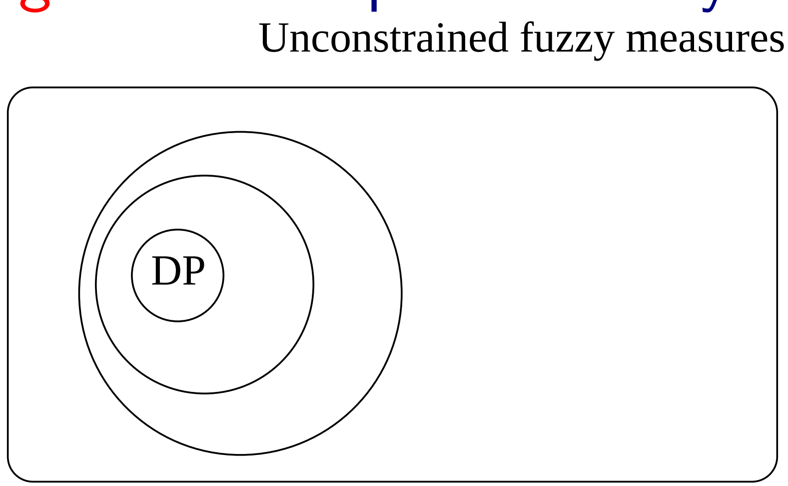
$\lvert X \rvert$	Number of possible orderings for Distorted Probabilities	Number of possible orderings for Fuzzy Measures
1	1	1
2	1	1
3	2	8
4	14	70016
5	546	$\mathcal{O}(10^{12})$
6	215470	–

m-Dimensional Distorted Probabilities

Justification: Why any extension of distorted probabilities?
The number of distorted probabilities.

Goal:

- To cover a larger region of the space of fuzzy measures



→ (similar to the property of k-additive fuzzy measures)

$$DP_{1,X} \subset DP_{2,X} \subset DP_{3,X} \cdots \subset DP_{|X|,X}$$

m-Dimensional Distorted Probabilities

- In distorted probabilities:
  - One probability distribution
  - One function f to distort the probabilities
- Extension to:
  - m probability distributions
  - One function f to distort/combine the probabilities

m-Dimensional Distorted Probabilities

- In distorted probabilities:
  - One probability distribution
  - One function f to distort the probabilities
- Extension to:
  - m probability distributions P_i
    - ⋆ Each P_i defined on X_i
    - ⋆ Each X_i is a partition element of X (a dimension)
  - One function f to distort/combine the probabilities

m-Dimensional Distorted Probabilities: Example

- Running example:
  - A fuzzy measure that is not a distorted probability:

$$\begin{array}{ll} \mu(\emptyset) = 0 & \mu(\{M, L\}) = 0.9 \\ \mu(\{M\}) = 0.45 & \mu(\{P, L\}) = 0.9 \\ \mu(\{P\}) = 0.45 & \mu(\{M, P\}) = 0.5 \\ \mu(\{L\}) = 0.3 & \mu(\{M, P, L\}) = 1 \end{array}$$

  - Partition on X:
    - $\star$ $X_1 = \{L\}$ (Literary subjects)
    - $\star$ $X_2 = \{M, P\}$ (Scientific Subjects)

m-Dimensional Distorted Probabilities: Definition

- m-dimensional distorted probabilities.
  - μ is an at most m dimensional distorted probability if

$$\mu(A) = f(P_1(A \cap X_1), P_2(A \cap X_2), \cdots, P_m(A \cap X_m))$$

where,
$\{X_1, X_2, \cdots, X_m\}$ is a partition of X,
P_i are probabilities on $(X_i, 2^{X_i})$,
f is a function on $\mathbb{R}^m$ strictly increasing with respect to the i-th axis for all $i = 1, 2, \ldots, m$.

- μ is an m-dimensional distorted probability if it is an at most m dimensional distorted probability but it is not an at most $m - 1$ dimensional.

m-Dimensional Distorted Probabilities: Example

- Running example: a two dimensional distorted probability

$$\mu(A) = f(P_1(A \cap \{L\}), P_2(A \cap \{M, P\}))$$

  - with partition on $X = \{M, L, P\}$
    1. Literary subject $\{L\}$
    2. Science subjects $\{M, P\}$,
  - probabilities
    1. $P_1(\{L\}) = 1$
    2. $P_2(\{M\}) = P_2(\{P\}) = 0.5$,
  - and distortion function f defined by

1	{L}	0.3	0.9	1.0
0	∅	0	0.45	0.5
	sets	∅	{M}, {P}	{M,P}
f		∅	0.5	1

Distorted Probabilities and Multisets

an approach to define (simple) fuzzy measures on multisets

Distorted Probabilities and Multisets

Multisets: elements can appear more than once

- Defined in terms of $count_M : X \to \{0\} \cup \mathbb{N}$
  e.g. when $X = \{a, b, c, d\}$ and $M = \{a, a, b, b, c, c, c\}$,
  $count_M(a) = 2$, $count_M(b) = 3$, $count_M(c) = 3$, $count_M(d) = 0$.
- A and B multisets on X, then
  - $A \subseteq B$ if and only if $count_A(x) \leq count_B(x)$ for all x in X (used to define submultiset).
  - $A \cup B$:
    $count_{A \cup B}(x) = \max(count_A(x), count_B(x))$ for all x in X.
  - $A \cap B$:
    $count_{A \cap B}(x) = \min(count_A(x), count_B(x))$ for all x in X.

Distorted Probabilities and Multisets

Fuzzy measure on multiset: X a reference set, M a multiset on X s.t. $M \neq \emptyset$; then, the function μ from $(M, \mathcal{P}(M))$ to $[0,1]$ is a fuzzy measure if the following holds:

- $\mu(\emptyset) = 0$ and $\mu(M) = 1$
- $\mu(A) \leq \mu(B)$ when $A \subseteq B$ and $B \subseteq M$.

Distorted Probabilities and Multisets

Fuzzy measure on multiset: X a reference set, M a multiset on X s.t. $M \neq \emptyset$; then, the function μ from $(M, \mathcal{P}(M))$ to $[0,1]$ is a fuzzy measure if the following holds:

- $\mu(\emptyset) = 0$ and $\mu(M) = 1$
- $\mu(A) \leq \mu(B)$ when $A \subseteq B$ and $B \subseteq M$.

How to define fuzzy measures?:

Distorted Probabilities and Multisets

Fuzzy measure on multiset: X a reference set, M a multiset on X s.t. $M \neq \emptyset$; then, the function μ from $(M, \mathcal{P}(M))$ to $[0,1]$ is a fuzzy measure if the following holds:

- $\mu(\emptyset) = 0$ and $\mu(M) = 1$
- $\mu(A) \leq \mu(B)$ when $A \subseteq B$ and $B \subseteq M$.

How to define fuzzy measures?:

- Even more parameters $\prod_{x \in X} count_M(x)$!!

Distorted Probabilities and Multisets

Fuzzy measure on multiset: X a reference set, M a multiset on X s.t. $M \neq \emptyset$; then, the function μ from $(M, \mathcal{P}(M))$ to $[0,1]$ is a fuzzy measure if the following holds:

- $\mu(\emptyset) = 0$ and $\mu(M) = 1$
- $\mu(A) \leq \mu(B)$ when $A \subseteq B$ and $B \subseteq M$.

How to define fuzzy measures?:

- Even more parameters $\prod_{x \in X} count_M(x)$!!

We present two alternative (but related) approaches

Distorted Probabilities and Multisets

1st approach: Definition based on a pseudoadditive integral:
Nondecreasing function-based fuzzy measures

- X a reference set, M a multiset on X and μ a $\oplus$-decomposable fuzzy measure on X. Let $f : [0, \infty) \to [0, \infty)$ be a non-decreasing function with $f(0) = 0$ and $f(m(M)) = 1$. Then, we define a fuzzy measure ν on $\mathcal{P}(M)$ by

$$\nu_f(A) = f(m(A))$$

where m is the multiset function $m : \mathcal{P}(M) \to [0, \infty)$ defined by

$$m(A) = (D) \int count_A d\mu.$$

Distorted Probabilities and Multisets

1st approach: Definition based on a pseudoadditive integral: Nondecreasing function-based fuzzy measures

- X a reference set, M a multiset on X and μ a $\oplus$-decomposable fuzzy measure on X. Let $f : [0, \infty) \to [0, \infty)$ be a non-decreasing function with $f(0) = 0$ and $f(m(M)) = 1$. Then, we define a fuzzy measure ν on $\mathcal{P}(M)$ by

$$\nu_f(A) = f(m(A))$$

where m is the multiset function $m : \mathcal{P}(M) \to [0, \infty)$ defined by

$$m(A) = (D) \int count_A d\mu.$$

- Rationale of the definition: $(C) \int \chi_A d\mu = \mu(A)$

Distorted Probabilities and Multisets

1st approach: Definition based on a pseudoadditive integral:
Nondecreasing function-based fuzzy measures

- X a reference set, M a multiset on X and μ a $\oplus$-decomposable fuzzy measure on X. Let $f : [0, \infty) \to [0, \infty)$ be a non-decreasing function with $f(0) = 0$ and $f(m(M)) = 1$. Then, we define a fuzzy measure ν on $\mathcal{P}(M)$ by

$$\nu_f(A) = f(m(A))$$

where m is the multiset function $m : \mathcal{P}(M) \to [0, \infty)$ defined by

$$m(A) = (D) \int count_A d\mu.$$

- Rationale of the definition: $(C) \int \chi_A d\mu = \mu(A)$
- Properties:
  if $A \subseteq B$ by the monotonicity of the integral $m(A) \leq m(B)$
  $\to$ monotonicity condition of the fuzzy measure fulfilled

Distorted Probabilities and Multisets

2nd approach: Definition based on prime numbers[1]:

- Define
$$n(A) := \prod_{x \in X} \phi(x)^{count_A(x)},$$
where ϕ is an injective function from X to the prime numbers, and let h be a non-decreasing function from $\mathbb{N}$ to $[0,1]$ satisfying $h(1) = 0$ and $h(n(M)) = 1$. We define the prime number-based fuzzy measure
$$\nu_{\phi,h}(A) = h(n(A)).$$

[1] and using the unique factorization of integers into prime numbers

Distorted Probabilities and Multisets

2nd approach: Definition based on prime numbers[1]:

- Define
$$n(A) := \prod_{x \in X} \phi(x)^{count_A(x)},$$
where ϕ is an injective function from X to the prime numbers, and let h be a non-decreasing function from $\mathbb{N}$ to $[0,1]$ satisfying $h(1) = 0$ and $h(n(M)) = 1$. We define the prime number-based fuzzy measure
$$\nu_{\phi,h}(A) = h(n(A)).$$

Properties:

if $A \neq B$ by the unique factorization $n(A) \neq n(B)$

if $A \subseteq B$ by the factorization $n(A) < n(B)$

$\rightarrow$ monotonicity condition of the fuzzy measure fulfilled

[1] and using the unique factorization of integers into prime numbers

Distorted Probabilities and Multisets

Properties:

- Fuzzy measures based on prime-number are a particular case of the 1st approach

Distorted Probabilities and Multisets

Properties:

- Fuzzy measures based on prime-number are a particular case of the 1st approach
- Neither the 1st nor the 2nd approach represent all possible fuzzy measures

Distorted Probabilities and Multisets

Properties:

- Fuzzy measures based on prime-number are a particular case of the 1st approach
- Neither the 1st nor the 2nd approach represent all possible fuzzy measures
- It seems that there is some parallelism between prime-number based fuzzy measures and distorted probabilities
  - f and the distortion
  - ϕ and the probability distribution

Distorted Probabilities and Multisets

Properties:

- Fuzzy measures based on prime-number are a particular case of the 1st approach
- Neither the 1st nor the 2nd approach represent all possible fuzzy measures
- It seems that there is some parallelism between prime-number based fuzzy measures and distorted probabilities
  - f and the distortion
  - ϕ and the probability distribution
- Can we establish a relationship??

Results

Properties:

- ν a fuzzy measure according to Approach 1 on a proper finite set $(M, \mathcal{P}(M)) = (X, 2^X)$. Then ν is a distorted probability on $(X, 2^X)$.

Results

Properties:

- ν a fuzzy measure according to Approach 1 on a proper finite set $(M, \mathcal{P}(M)) = (X, 2^X)$. Then ν is a distorted probability on $(X, 2^X)$.
- Same for Approach 2 (primer-number definition)

Results

Properties:

- ν a fuzzy measure according to Approach 1 on a proper finite set $(M, \mathcal{P}(M)) = (X, 2^X)$. Then ν is a distorted probability on $(X, 2^X)$.
- Same for Approach 2 (primer-number definition)
- This is easy to prove (consists on defining the probability distribution)

Results

Properties:

- ν a fuzzy measure according to Approach 1 on a proper finite set $(M, \mathcal{P}(M)) = (X, 2^X)$. Then ν is a distorted probability on $(X, 2^X)$.
- Same for Approach 2 (primer-number definition)
- This is easy to prove (consists on defining the probability distribution)
- So, Approach 1 and Approach 2 equal to or more general than distorted probabilities

Results

Properties:

- Can we prove something else? much more general? almost the same? exactly the same?

Results

Properties:

- Can we prove something else? much more general? almost the same? exactly the same?

  Not so surprising theorem: Fuzzy measures based on prime numbers on proper sets are equivalent to distorted probabilities

Results

Properties:

- Can we prove something else? much more general? almost the same? exactly the same?

  Not so surprising theorem: Fuzzy measures based on prime numbers on proper sets are equivalent to distorted probabilities
  $\rightarrow$ probabilities and prime numbers play the same role
  $\rightarrow$ $\sum_{x \in A} p_x$ and $\prod_{x \in A} \phi(x)$ play the same role

Results

Properties:

- Can we prove something else? much more general? almost the same? exactly the same?

  Not so surprising theorem: Fuzzy measures based on prime numbers on proper sets are equivalent to distorted probabilities

  $\rightarrow$ probabilities and prime numbers play the same role

  $\rightarrow \sum_{x \in A} p_x$ and $\prod_{x \in A} \phi(x)$ play the same role

  Much more surprising theorem: Fuzzy measures based on Approach 1 on proper sets are equivalent to distorted probabilities

Results

Properties:

- Can we prove something else? much more general? almost the same? exactly the same?

  Not so surprising theorem: Fuzzy measures based on prime numbers on proper sets are equivalent to distorted probabilities
  $\rightarrow$ probabilities and prime numbers play the same role
  $\rightarrow \sum_{x \in A} p_x$ and $\prod_{x \in A} \phi(x)$ play the same role

  Much more surprising theorem: Fuzzy measures based on Approach 1 on proper sets are equivalent to distorted probabilities

  Surprising corollary: Approach 1 and approach 2 are equivalent.

Results

Properties:

- Can we prove something else? much more general? almost the same? exactly the same?

  Not so surprising theorem: Fuzzy measures based on prime numbers on proper sets are equivalent to distorted probabilities
  $\rightarrow$ probabilities and prime numbers play the same role
  $\rightarrow \sum_{x \in A} p_x$ and $\prod_{x \in A} \phi(x)$ play the same role

  Much more surprising theorem: Fuzzy measures based on Approach 1 on proper sets are equivalent to distorted probabilities

  Surprising corollary: Approach 1 and approach 2 are equivalent.

  Proof based on some results on number theory about the existence of k prime numbers in certain intervals (Bertrand's postulate).

Results

An example to satisfy curiosity:

- μ distorted probability $p = (0.05, 0.1, 0.2, 0.3, 0.35)$, $g(x) = x^2$.

Results

An example to satisfy curiosity:

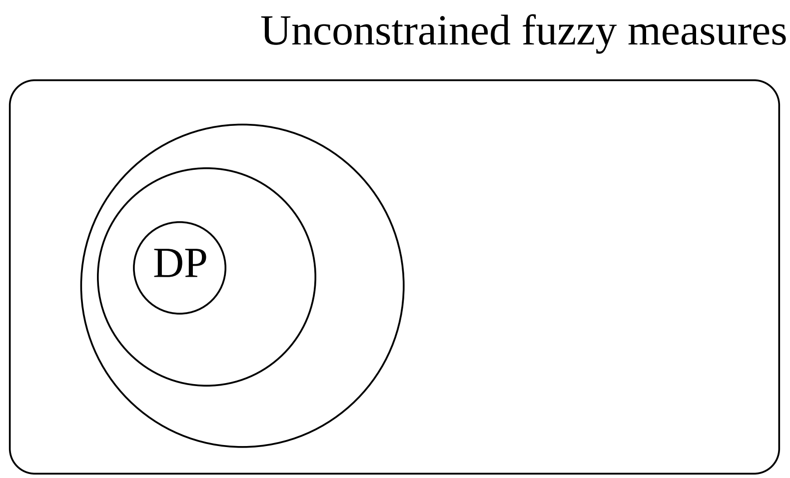
- μ distorted probability $p = (0.05, 0.1, 0.2, 0.3, 0.35)$, $g(x) = x^2$.
- Representation with prime numbers and appropriate function

$$
\begin{aligned}
\phi(x_1) &= 17 \in [16.0, 32.0001] \\
\phi(x_2) &= 367 \in [362.041, 724.081] \\
\phi(x_3) &= 185369 \in [185366.0, 370732.0] \\
\phi(x_4) &= 94907801 \in [9.49078 \times 10^7, 1.89816 \times 10^8] \\
\phi(x_5) &= 2147524151 \in [2.14752 \times 10^9, 4.29505 \times 10^9]
\end{aligned}
$$

m-Dimensional DP for multisets

How to solve the *problem* that not all fuzzy measures for multisets are distorted probabilities ?

- Same approach as before: m-dimensional prime number-based fuzzy measure



m-Dimensional DP for multisets

m-dimensional prime number-based fuzzy measure

- μ is an at most m-dimensional prime number-based fuzzy measure if

$$\mu(A) = f(n_1(A \cap X_1), \dots, n_m(A \cap X_m))$$

where,
$\{X_1, X_2, \cdots, X_m\}$ is a partition of X,
$n_i(A) = \prod_{x \in X_i} \phi(x)^{count_A(x)}$ with ϕ_i injective functions from X_i to the prime numbers
f is a strictly increasing function with respect to the i-th axis for all $i = 1, 2, \dots, m$.

μ is an m-dimensional prime number-based fuzzy measure if it is an at most m dimensional distorted probability but it is not an at most $m-1$ dimensional.

m-Dimensional DP for multisets

Properties:

- All fuzzy measures are at most $|X|$-dimensional prime number-based fuzzy measures.

Integral

Integral

Definition

- Boundary measures:
  - $\mu^+(A) = A \cdot M$ for all $A \subseteq X$
  - $\mu_-(A) = A \cap M$ for all $A \subseteq X$
- They satisfy:

$$\mu_-(A) \leq \mu^+(A)$$

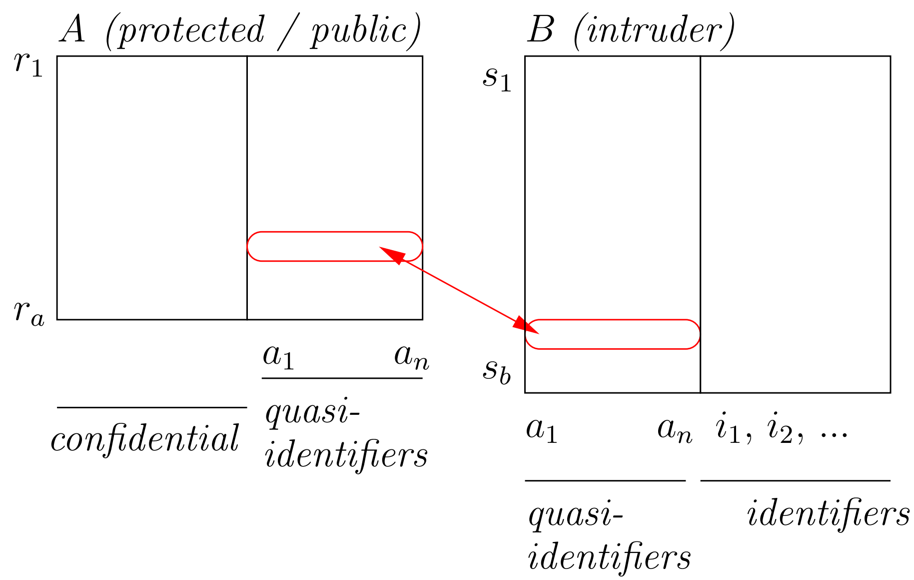
and, therefore,

$$(C) \int f d\mu_- < (C) \int f d\mu^+$$

Finally an application

Record Linkage

Record Linkage:



Record Linkage

Record Linkage:

$$Minimize \sum_{i=1}^{N} K_i \tag{1}$$

$$Subject\ to:$$

$$\sum_{i=1}^{N}\sum_{j=1}^{N} \mathbb{C}(d(V_1(a_i), V_1(b_j)), \ldots, d(V_n(a_i), V_n(b_j))) -$$

$$- \mathbb{C}(d(V_1(a_i), V_1(b_i)), \ldots, d(V_n(a_i), V_n(b_i))) + CK_i > 0 \tag{2}$$

$$K_i \in \{0, 1\} \tag{3}$$

Additional constraints according to $\mathbb{C}$ (4)

Summary

Summary

Summary:

- Brief justification of the use of non-additive (fuzzy) measures
- Introduction to distorted probabilities
- Extensions
  - m-dimensional distorted probabilities
  - Fuzzy measures for multisets