

83rd EWG-MCDA 2016

**Choquet integral: distributions and decisions**

Vicenç Torra

School of Informatics, University of Skövde  
Skövde, Sweden

March 31-32 , 2016

# Overview

---

## Basics and objectives:

- Distribution based on the Choquet integral (for non-additive measures)

## Motivation:

- Theory: Mathematical properties
- Methodology: different ways to express interactions
- Application: Decision (MCDM), classification, statistical disclosure control (data privacy)

# Outline

---

1. Preliminaries
2. Choquet integral based distribution
3. Choquet-Mahalanobis based distribution
4. Summary

---

# Preliminaries

## Aggregation operators and the Choquet integral in Decision

---

# MCDM: Aggregation for (numerical) utility functions

# Aggregation for (numerical) utility functions

- Decision, **utility functions**

Alternatives = { Ford T, Seat 600, Simca 1000, VW, Citr.Acadiane }

Criteria = { Seats, Security, Price, Comfort, trunk }

Decision making process:

**Modelling=Criteria + Utilities**, aggregation, selection

	Number of seats	Security	Price	Confort	trunk
Ford T	0	20	0	20	0
Seat 600	60	0	100	0	50
Simca 1000	100	30	100	50	70
VW Beetle	80	50	30	70	100
Citroën Acadiane	20	40	60	40	0

# Aggregation for (numerical) utility functions

- Decision, **utility functions**

Alternatives = { Ford T, Seat 600, Simca 1000, VW, Citr.Acadiane }

Criteria = { Seats, Security, Price, Comfort, trunk }

Decision making process:

Modelling, **aggregation** =  $\mathbb{C}$ , selection

	Seats	Security	Price	Comfort	trunk	$\mathbb{C} = AM$
Ford T	0	20	0	20	0	8
Seat 600	60	0	100	0	50	42
Simca 1000	100	30	100	50	70	70
VW	80	50	30	70	100	66
Citr. Acadiane	20	40	60	40	0	32

# Aggregation for (numerical) utility functions

- MCDM: Aggregation to deal with **contradictory criteria**
- But there are occasions in which **ordering is clear**

when  $a_i \leq b_i$  it is clear that  $a \leq b$

E.g.,

	Seats	Security	Price	Comfort	trunk	$C = AM$
Seat 600	60	0	100	0	50	42
Simca 1000	100	30	100	50	70	70

**Aggregation operators are appropriate** because they satisfy monotonicity

- **Pareto dominance:** Given two vectors  $a = (a_1, \dots, a_n)$  and  $b = (b_1, \dots, b_n)$ , we say that  $b$  dominates  $a$  when  $a_i \leq b_i$  for all  $i$  and there is at least one  $k$  such that  $a_k < b_k$ .



# Aggregation for (numerical) utility functions

---

- Pareto set, Pareto frontier, or non dominance set:

	Seats	Security	Price	Comfort	trunk	$C = AM$
Simca 1000	100	30	100	50	70	70
VW	80	50	30	70	100	66
Citr. Acadiane	20	40	60	40	0	32

- Each one wins at least in one criteria to another one

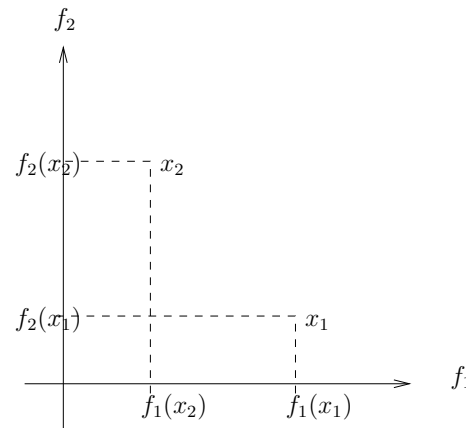
# Aggregation and Choquet integral in MCDM

- **Pareto set, Pareto frontier, or non dominance set:**

Given a set of alternatives  $U$  represented by vectors  $u = (u_1, \dots, u_n)$ , the Pareto frontier is the set  $u \in U$  such that there is no other  $v \in U$  such that  $v$  dominates  $u$ .

$$PF = \{u \mid \text{there is no } v \text{ s.t. } v \text{ dominates } u\}$$

- **Pareto optimal:** an element  $u$  of the Pareto set



# Aggregation and Choquet integral in MCDM

- Decision making process:

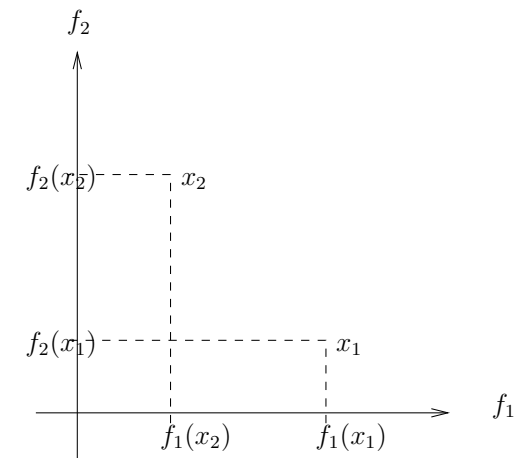
Modelling, aggregation, **selection=order,first**

- The function of aggregation functions

- Different aggregations lead to different orders (in the PF)
- Aggregation establishes which **points** are *equivalent*
- Different aggregations, lead to different curves of points (level curves)

Criteria Satisfaction on:							
alt	Price	Quality	Comfort	alt	Consensus	alt	Ranking
FordT	0.2	0.8	0.3	FordT	0.35	206	0.72
206	0.7	0.7	0.8	206	0.72	FordT	0.35
...	...			...	...	...	...

$\Rightarrow$



# Aggregation and Choquet integral in MCDM

---

- Aggregation functions and different level curves
  - Arithmetic mean
  - Geometric mean, Harmonic mean, ...
  - Weighted mean
  - OWA, ...
  - Choquet integral (generalization of the AM, WM, OWA)
    - \* to represent interactions between criteria
    - \* non-independent criteria allowed

# Aggregation and Choquet integral in MCDM

---

- Aggregation functions and parameters
  - Arithmetic mean: no parameters
  - Geometric mean, Harmonic mean, ...: no parameters
  - Weighted mean: weighting vector
  - OWA, ...: weighting vector

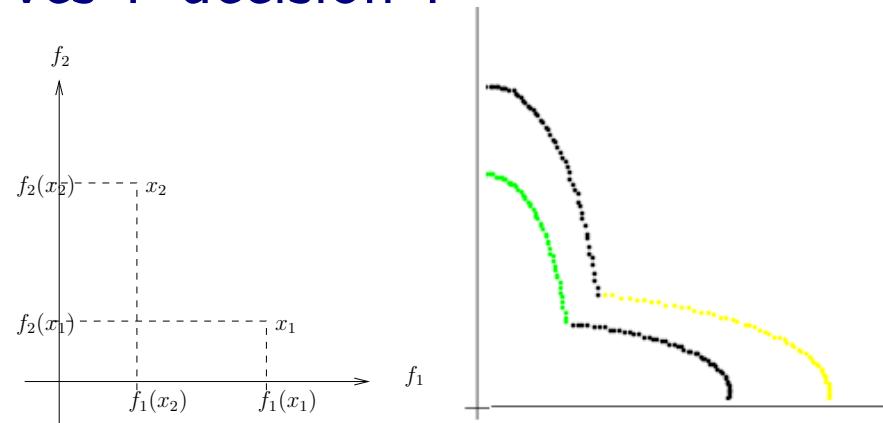
# Aggregation and Choquet integral in MCDM

---

- Aggregation functions and **parameters**
  - Choquet integral (generalization of the AM, WM, OWA): **a measure**
    - Instead of *weight(criteria)*:  $w(\text{security})$
    - We consider *weight(set of criteria)*:  $w(\text{security,price,confort})$
  - We can, of course, use
$$w(\text{security,price,confort}) = w(\text{security}) + w(\text{price}) + w(\text{confort})$$
  - but also
    - $w(\text{security,price,confort}) > w(\text{security}) + w(\text{price}) + w(\text{confort})$
    - or
    - $w(\text{security,price,confort}) < w(\text{security}) + w(\text{price}) + w(\text{confort})$

# Aggregation and Choquet integral in MCDM

- Aggregation functions and **parameters**
  - Choquet integral (generalization of the AM, WM, OWA): **a measure**
  - \* And the level curves ? decision ?



---

# Preliminaries

## Non-additive (fuzzy) measures and the Choquet integral



# Definitions: measures

---

## Additive measures.

- $(X, \mathcal{A})$  a measurable space; then, a set function  $\mu$  is an additive measure if it satisfies
  - (i)  $\mu(A) \geq 0$  for all  $A \in \mathcal{A}$ ,
  - (ii)  $\mu(X) \leq \infty$
  - (iii) Finite case:  
 $\mu(A \cup B) = \mu(A) + \mu(B)$  for disjoint  $A, B$
- Probability and weights:  $\mu(X) = 1$

# Definitions: measures

---

## Non-additive (or fuzzy) measures.

- $(X, \mathcal{A})$  a measurable space, a non-additive measure  $\mu$  on  $(X, \mathcal{A})$  is a set function  $\mu : \mathcal{A} \rightarrow [0, 1]$  satisfying the following axioms:
  - (i)  $\mu(\emptyset) = 0$
  - (ii)  $\mu(X) \leq \infty$
  - (iii)  $A \subseteq B$  implies  $\mu(A) \leq \mu(B)$  (monotonicity)
- Weights:  $\mu(X) = 1$

# Definitions: measures

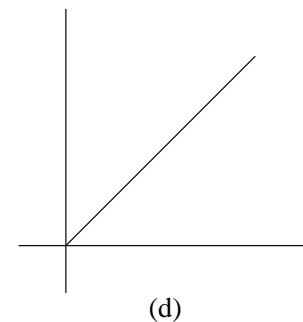
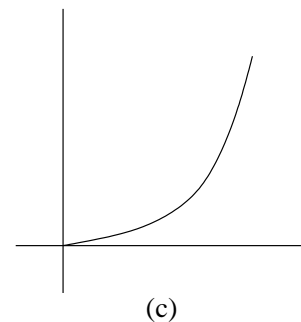
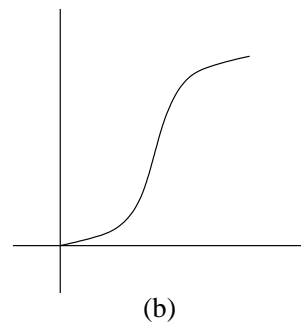
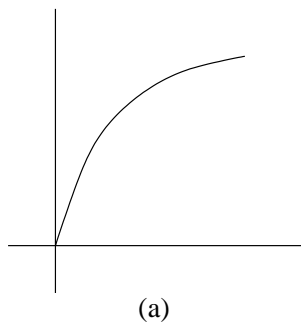
## Non-additive measures. Examples. Distorted probabilities

- $m : \mathbb{R}^+ \rightarrow \mathbb{R}^+$  a continuous and increasing function such that  $m(0) = 0$ ;  $P$  be a probability.

The following set function  $\mu_m$  is a non-additive measure:

$$\mu_{m,P}(A) = m(P(A)) \quad (1)$$

- If  $m(x) = x^2$ , then  $\mu_m(A) = (P(A))^2$
- If  $m(x) = x^p$ , then  $\mu_m(A) = (P(A))^p$



# Definitions: measures

---

## Non-additive measures. Examples. Distorted probabilities

- $m : \mathbb{R}^+ \rightarrow \mathbb{R}^+$  a continuous and increasing function such that  $m(0) = 0$ ;  $P$  be a probability.

The following set function  $\mu_m$  is a non-additive measure:

$$\mu_{m,P}(A) = m(P(A)) \quad (2)$$

## Applications.

- To represent **interactions**

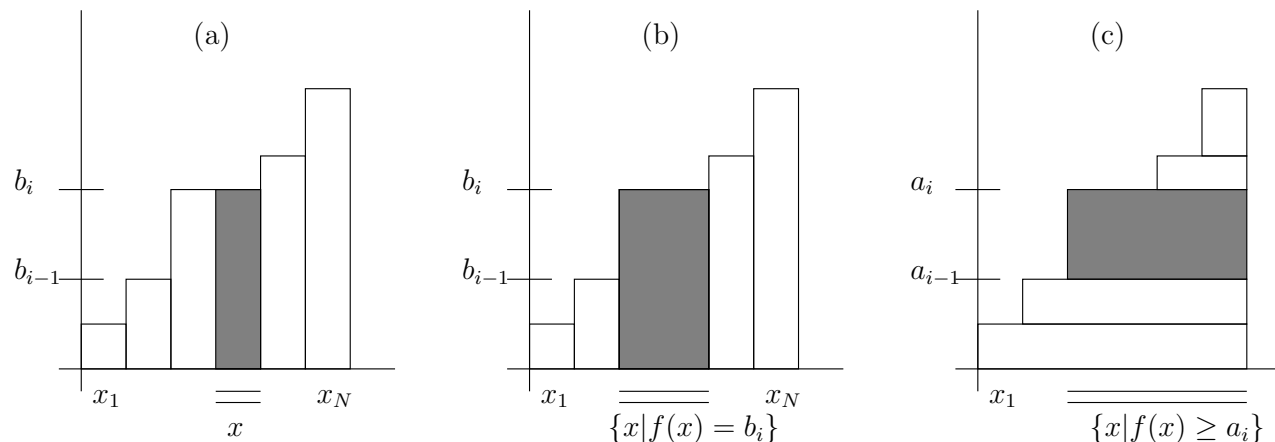
# Definitions: integrals

## Choquet integral (Choquet, 1954):

- $\mu$  a non-additive measure,  $g$  a measurable function. The Choquet integral of  $g$  w.r.t.  $\mu$ , where  $\mu_g(r) := \mu(\{x | g(x) > r\})$ :

$$(C) \int g d\mu := \int_0^\infty \mu_g(r) dr. \quad (3)$$

- When the measure is additive, this is the Lebesgue integral



# Definitions: integrals

---

## Choquet integral. Discrete version

- $\mu$  a non-additive measure,  $f$  a measurable function. The Choquet integral of  $f$  w.r.t.  $\mu$ ,

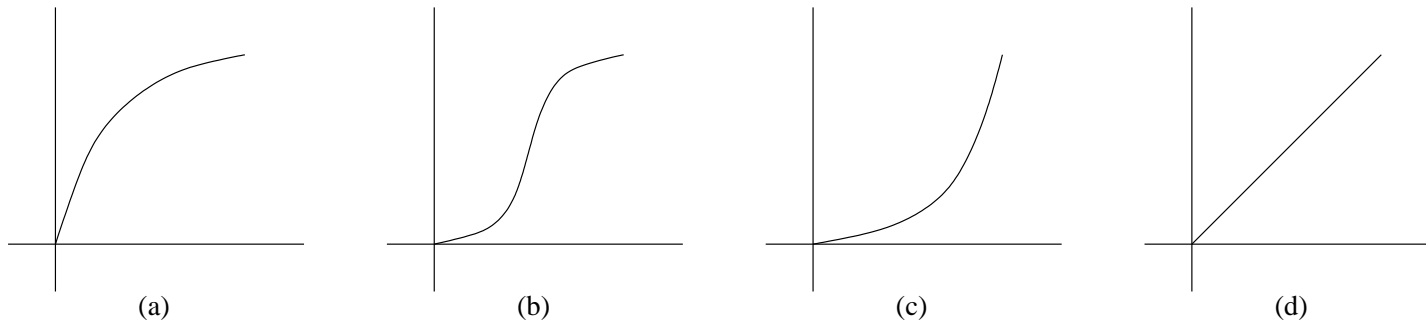
$$(C) \int f d\mu = \sum_{i=1}^N [f(x_{s(i)}) - f(x_{s(i-1)})] \mu(A_{s(i)}),$$

where  $f(x_{s(i)})$  indicates that the indices have been permuted so that  $0 \leq f(x_{s(1)}) \leq \dots \leq f(x_{s(N)}) \leq 1$ , and where  $f(x_{s(0)}) = 0$  and  $A_{s(i)} = \{x_{s(i)}, \dots, x_{s(N)}\}$ .

# Definitions: measures

## Choquet integral: Properties:

- When  $\mu$  is additive,  $CI$  corresponds to the weighted mean
- $CI$  can represent min, max, mean, order statistics, ...
- When  $\mu$  is  $\mu_{m,P}(A) = m(P(A))$  with  $m(x) = x^p$ ,  
 $CI_{\mu_m}(f)$   
*(a) → max, (b) → median, (c) → min, (d) → mean*



# **Preliminaries**

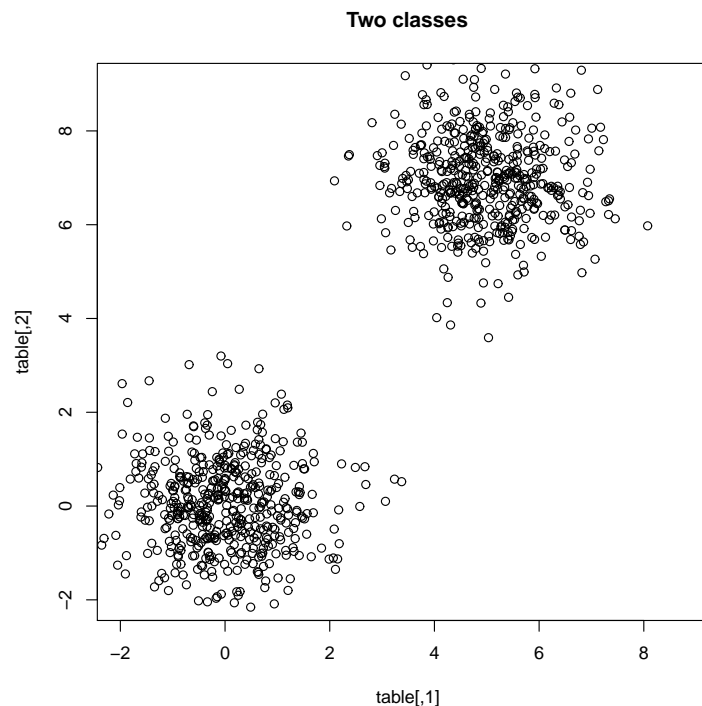
## **Classification and shapes of distributions**



# Classification

**Motivation:** Another motivation: classification

- Two classes defined in terms of **normal distributions** (obtained from real data or directly from the parameters of the distribution  $N(\mu, \Sigma)$ ).
- An element  $x$  in  $\mathbb{R}^2$   
→ **where to classify  $x$ ?**



# Classification

## Classification problems: Classification of $x$ into $\Omega$

- $x$  in a  $n$ -dimensional space (i.e.,  $x \in \mathbb{R}^n$ )
- Set of  $k$  classes  $\Omega = \{\omega_1, \dots, \omega_k\}$

## Formalization:

- Bayes' *maximum-a-posteriori* (MAP) **classification decision rule**: assigns  $x$  to the class  $\omega_i$  s.t. the probability  $P(\omega_i|x)$  is maximized. I.e., (Bayes condition):

$$P(\omega_i|x) = \frac{P(x|\omega_i)P(\omega_i)}{P(x)}$$

or, as  $P(x)$  is constant for all classes,

$$d_i(x) = P(x|\omega_i)P(\omega_i)$$

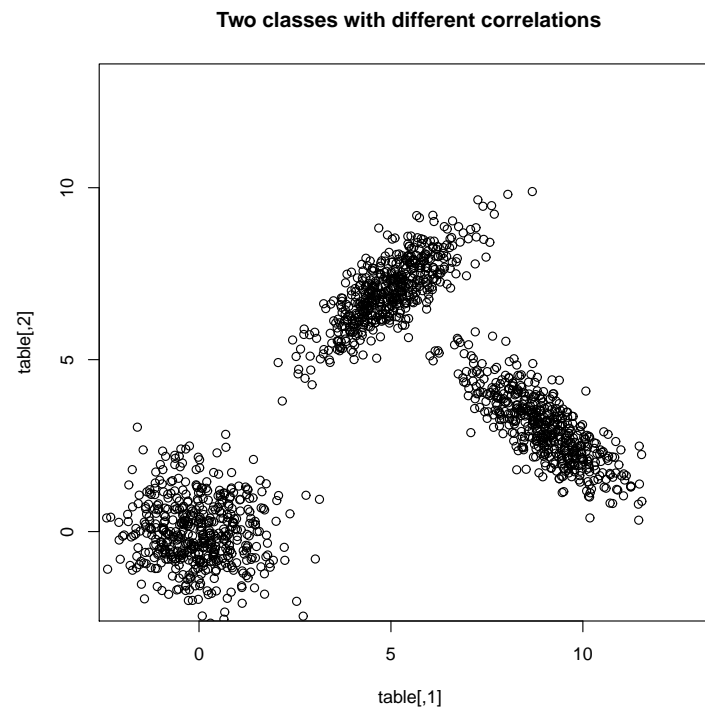
■  $f \circ d$  results into the same classification as for  $d$  (e.g.  $f = \ln$ )

# Classification

**Classification problems:** Classes  $\omega_i$  generated from

- covariance matrices  $\Sigma_i$
  - means  $\bar{x}_i$
- class-conditional probability-density function (Gaussian distribution)

$$P(x|\omega_i) = \frac{1}{(2\pi)^{m/2} |\Sigma_i|^{1/2}} e^{-\frac{1}{2}(x-\bar{x}_i)^T \Sigma_i^{-1} (x-\bar{x}_i)}$$



# Classification

---

## Proposition:

- Bayes' *maximum-a-posteriori* (MAP) classification decision rule, when  $\Sigma_i = \Sigma_j$ , and  $P(\omega_i) = P(\omega_j)$ , is (Mahalanobis distance)

$$d_i(x) = -(x - \bar{x}_i)^T \Sigma^{-1} (x - \bar{x}_i)$$

# Classification

**Proposition.** Bayes' *maximum-a-posteriori* (MAP) classification

- If  $\Sigma_i = \mathbb{I}$  for all  $i$  (the identity function)

$$d_i(x) = -(x - \bar{x}_i)^T (x - \bar{x}_i) = -\|x - \bar{x}_i\|^2$$

→ **Euclidean distance**

- If  $\Sigma_i$  is diagonal (not necessarily equal to  $\mathbb{I}$ )

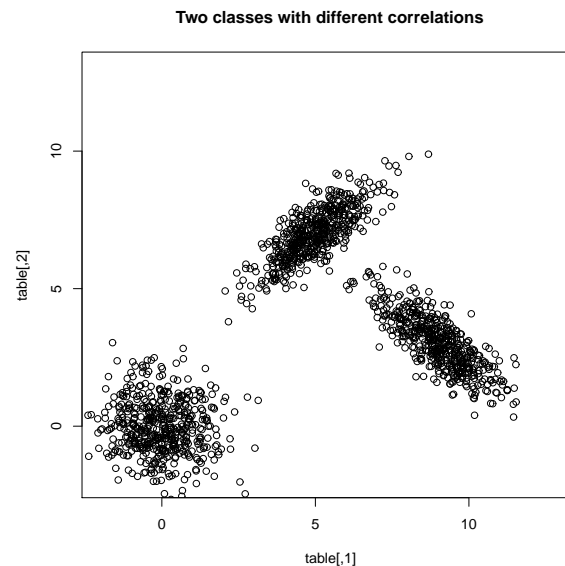
$$d_i(x) = -\sum_{j=1}^m (\sigma_j^2)^{-1} (x_j - \bar{x}_{ij})^2$$

→ **Weighted Euclidean distance**

(with weights equal to the inverse of the variances:  $p_j = (\sigma_j^2)^{-1}$ )

# Shape of distributions

The class-conditional probability-density functions established above define level curves with the **shape of an ellipse**  
→ circumference when variables are independent



What about another shape / another distance ?

What about using the **Choquet integral** here ?

# Shape of distributions

---

## Why Choquet integral?:

- Non-additive measures on a set  $X$  permit us to **represent interactions** between objects in  $X$  !!  
... **similar to covariances** !!
- Choquet integral integrates a function with respect to a non-additive measure
  - can it be used to compute a distance / to define a distribution ?
  - if so, what is the shape of the distribution ?

---

# Choquet integral based distribution



# Choquet integral based distribution: Definition

## Definition:

- $Y = \{Y_1, \dots, Y_n\}$  random variables;  $\mu : 2^Y \rightarrow [0, 1]$  a non-additive measure and  $\mathbf{m}$  a vector in  $\mathbb{R}^n$ .
- The exponential family of Choquet integral based class-conditional probability-density functions is defined by:

$$PC_{\mathbf{m}, \mu}(\mathbf{x}) = \frac{1}{K} e^{-\frac{1}{2} CI_{\mu}((\mathbf{x}-\mathbf{m}) \circ (\mathbf{x}-\mathbf{m}))}$$

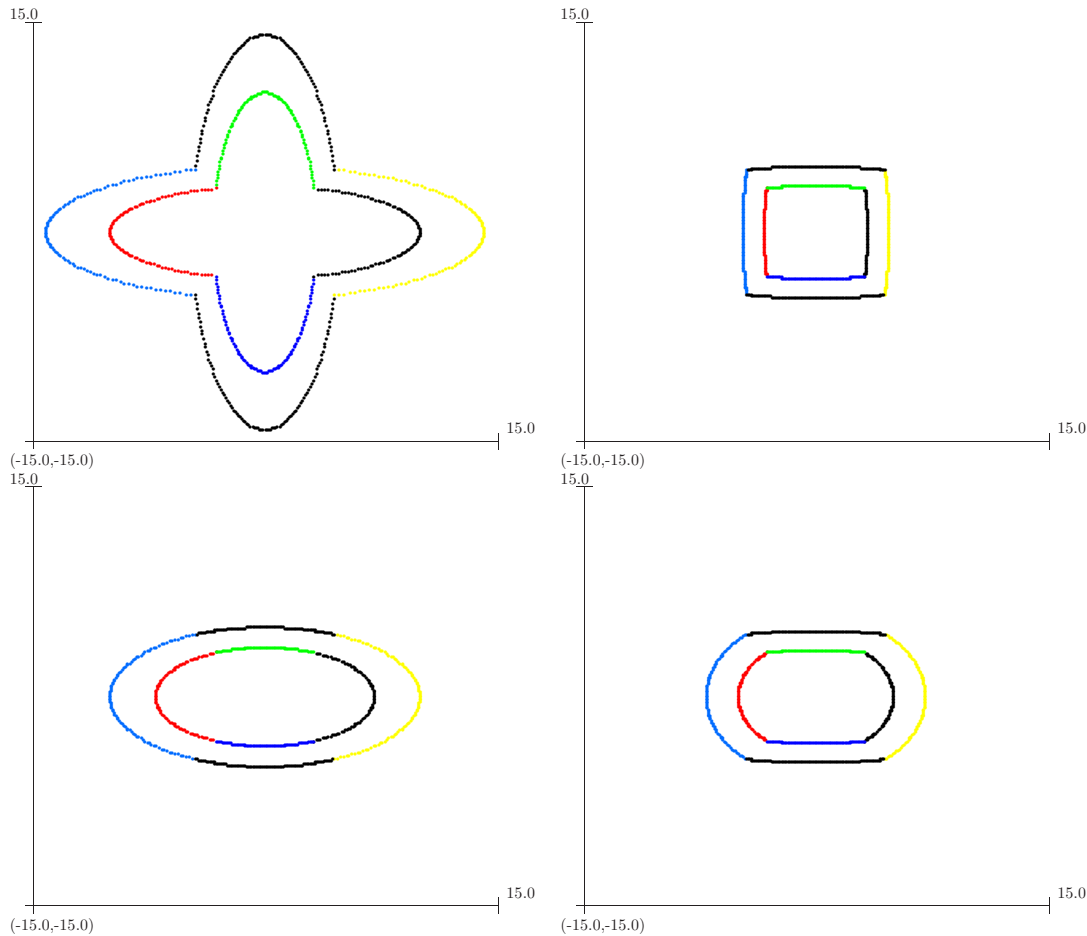
where  $K$  is a constant that is defined so that the function is a probability, and where  $\mathbf{v} \circ \mathbf{w}$  denotes the Hadamard or Schur (elementwise) product of vectors  $\mathbf{v}$  and  $\mathbf{w}$  (i.e.,  $(\mathbf{v} \circ \mathbf{w}) = (v_1 w_1 \dots v_n w_n)$ ).

## Notation:

- We denote it by  $C(\mathbf{m}, \mu)$ .

# Choquet integral based distribution: Examples

- Shapes (level curves)



(a)  $\mu_A(\{x\}) = 0.1$  and  $\mu_A(\{y\}) = 0.1$ , (b)  $\mu_B(\{x\}) = 0.9$  and  $\mu_B(\{y\}) = 0.9$ ,  
 (c)  $\mu_C(\{x\}) = 0.2$  and  $\mu_C(\{y\}) = 0.8$ , and (d)  $\mu_D(\{x\}) = 0.4$  and  $\mu_D(\{y\}) = 0.9$ .

# Choquet integral based distribution: Properties

---

**Proposition.** Distribution and distance (Choquet distance):

- If  $P(w_i) = P(w_j)$  holds for all  $i \neq j$ , the decision rule is (max):

$$-CI_{\mu}((x - \bar{x}_i) \otimes (x - \bar{x}_i))$$

**Proposition:** Distribution/distance and level curves:

- The level curves of the Choquet integral in two variables  $X = \{x, y\}$  corresponds to an ellipse when  $\mu(\{x\}) = 1 - \mu(\{y\})$ .
  - A natural result: we have an ellipse when  $\mu(\{x\}) + \mu(\{y\}) = 1$
  - i.e., when  $\mu$  is a probability.

This follows from the fact that the Choquet integral with a measure that is a probability is equivalent to a weighted mean. Then, similar results are obtained for larger dimensions.

# Choquet integral based distribution: Properties

---

## Property:

- The family of distributions  $N(\mathbf{m}, \Sigma)$  in  $\mathbb{R}^n$  with a **diagonal** matrix  $\Sigma$  of rank  $n$ , and the family of distributions  $C(\mathbf{m}, \mu)$  with an **additive measure**  $\mu$  with all  $\mu(\{x_i\}) \neq 0$  are equivalent.

( $\mu(X)$  is not necessarily here 1)

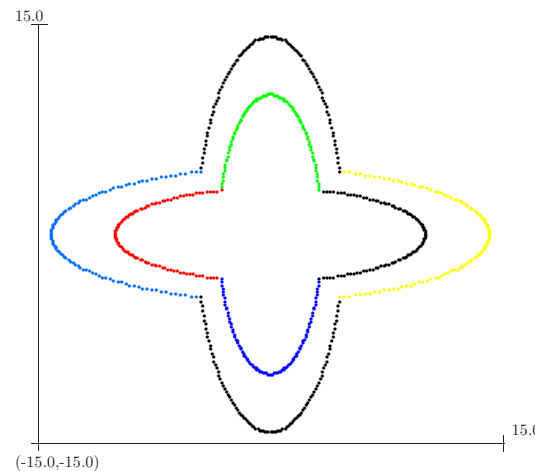
## Corollary:

- The distribution  $N(\mathbf{0}, \mathbb{I})$  corresponds to  $C(\mathbf{0}, \mu^1)$  where  $\mu^1$  is the additive measure defined as  $\mu^1(A) = |A|$  for all  $A \subseteq X$ .

# Choquet integral based distribution: $N$ vs. $C$

## Properties:

- In general, the two families of distributions  $N(\mathbf{m}, \Sigma)$  and  $C(\mathbf{m}, \mu)$  are different.
- $C(\mathbf{m}, \mu)$  always symmetric w.r.t.  $Y_1$  and  $Y_2$  axis.



- A generalization of both: Choquet-Mahalanobis based distribution.
  - Mahalanobis:  $\Sigma$  represents some **interactions**
  - Choquet (measure):  $\mu$  represents some **interactions**

---

# Choquet-Mahalanobis based distribution

# Choquet integral based distribution: generalized distance

---

## Definition:

- $\Sigma$  be a matrix,  $\Sigma^{-1} = LL^*$  be the Cholesky decomposition of its inverse.
- The **Choquet-Mahalanobis *integral*** is defined by

$$CMI_{\mu, \Sigma}(x, \bar{x}) = CI_{\mu}(v \otimes w) \quad (4)$$

where  $v$  and  $w$  are the vectors defined by:

$$v = (x - \bar{x})^T L \text{ and } w = L^*(x - \bar{x}),$$

where  $v \otimes w$  denotes the elementwise product of vectors  $v$  and  $w$  (i.e.,  $(v \otimes w) = (v_1 w_1 \dots v_n w_n)$ ).

# Choquet integral based distribution: generalized distance

---

## On the definition:

- **Well defined** when  $\Sigma$  is a covariance matrix

When  $\Sigma^{-1}$  is a definite-positive matrix, the **Cholesky decomposition is unique**.

This is the case when  $\Sigma$  is a covariance matrix valid for generating a probability-density function.

## Proper generalization:

- **Generalization of both the Mahalanobis and the Choquet integral based distance.**
  - The definition with  $\Sigma$  equal to the identity results into the Choquet integral of  $(x - \bar{x}) \otimes (x - \bar{x})$  with respect to  $\mu$ .
  - The definition with  $\mu$  corresponding to an additive probability  $\mu(A) = 1/|A|$  results into  $1/n$  of the Mahalanobis distance with respect to  $\Sigma$ .



# Choquet integral based distribution: Definition

## Definition:

- $Y = \{Y_1, \dots, Y_n\}$  random variables,  $\mu : 2^Y \rightarrow [0, 1]$  a measure,  $\mathbf{m}$  a vector in  $\mathbb{R}^n$ , and  $\mathbf{Q}$  a positive-definite matrix.
- The exponential family of **Choquet-Mahalanobis integral** based class-conditional probability-density functions is defined by:

$$PCM_{\mathbf{m}, \mu, \mathbf{Q}}(x) = \frac{1}{K} e^{-\frac{1}{2} CI_{\mu}(\mathbf{v} \circ \mathbf{w})}$$

where  $K$  is a constant that is defined so that the function is a probability, where  $\mathbf{L}\mathbf{L}^T = \mathbf{Q}$  is the Cholesky decomposition of the matrix  $\mathbf{Q}$ ,  $\mathbf{v} = (\mathbf{x} - \mathbf{m})^T \mathbf{L}$ ,  $\mathbf{w} = \mathbf{L}^T (\mathbf{x} - \mathbf{m})$ , and where  $\mathbf{v} \circ \mathbf{w}$  denotes the elementwise product of vectors  $\mathbf{v}$  and  $\mathbf{w}$ .

## Notation:

- We denote it by  **$CMI(\mathbf{m}, \mu, \mathbf{Q})$** .

# Choquet integral based distribution: Properties

---

## Property:

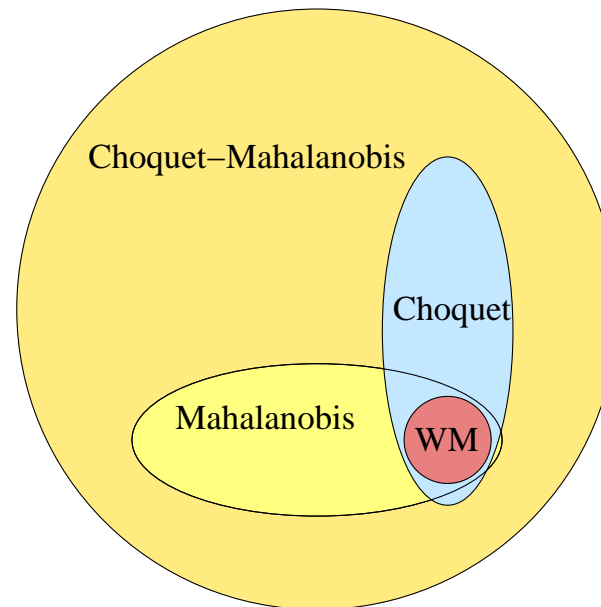
- The distribution  $CMI(\mathbf{m}, \mu, \mathbf{Q})$  generalizes the multivariate normal distributions and the Choquet integral based distribution. In addition
  - A  $CMI(\mathbf{m}, \mu, \mathbf{Q})$  with  $\mu = \mu^1$  corresponds to multivariate normal distributions,
  - A  $CMI(\mathbf{m}, \mu, \mathbf{Q})$  with  $Q = \mathbb{I}$  corresponds to a  $CI(\mathbf{m}, \mu)$ .

# Choquet integral based distribution: Properties

---

## Graphically:

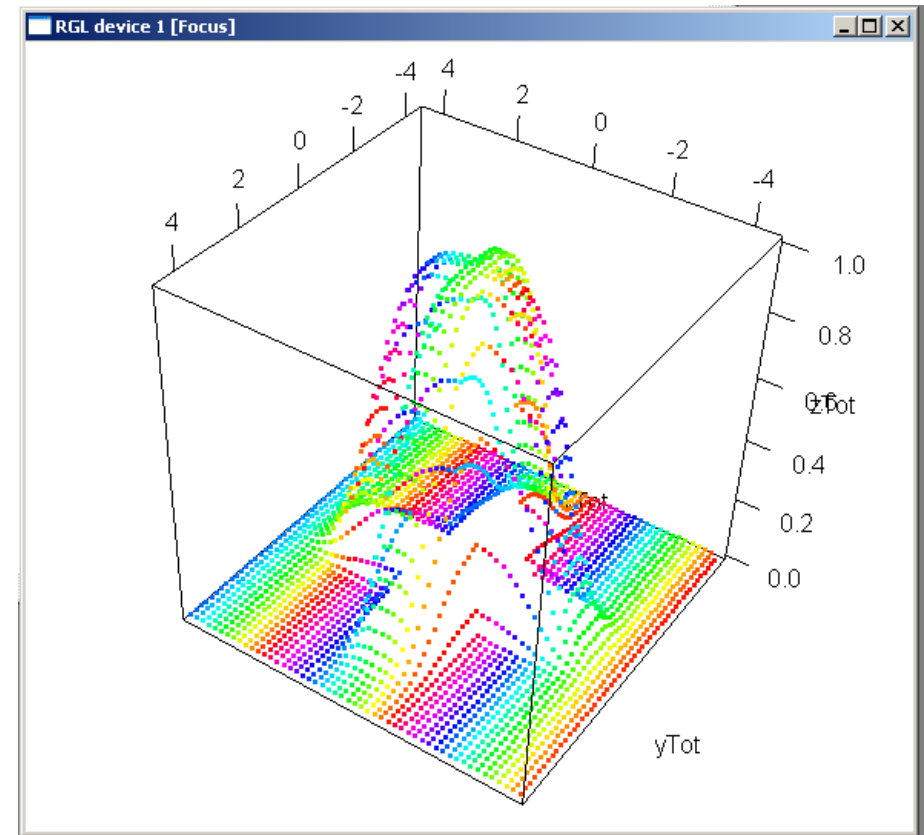
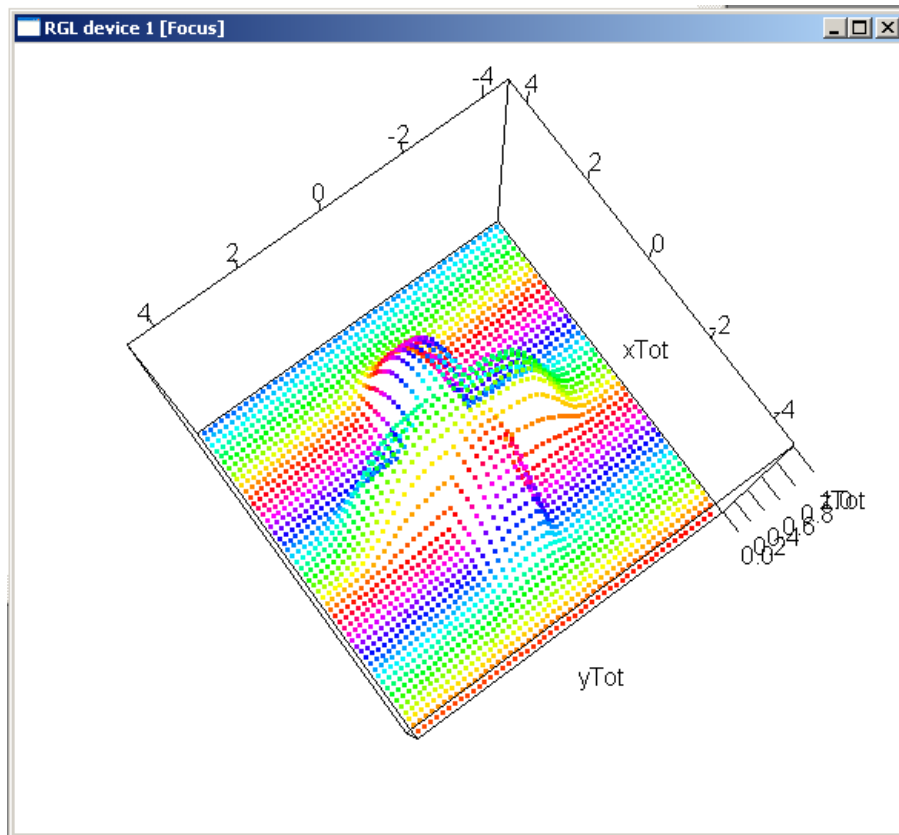
- Choquet integral (CI distribution), Mahalanobis distance (multivariate normal distribution), generalization (CMI distribution)



# Choquet integral based distribution: Examples

**1st Example:** Interactions only expressed in terms of a **measure**.

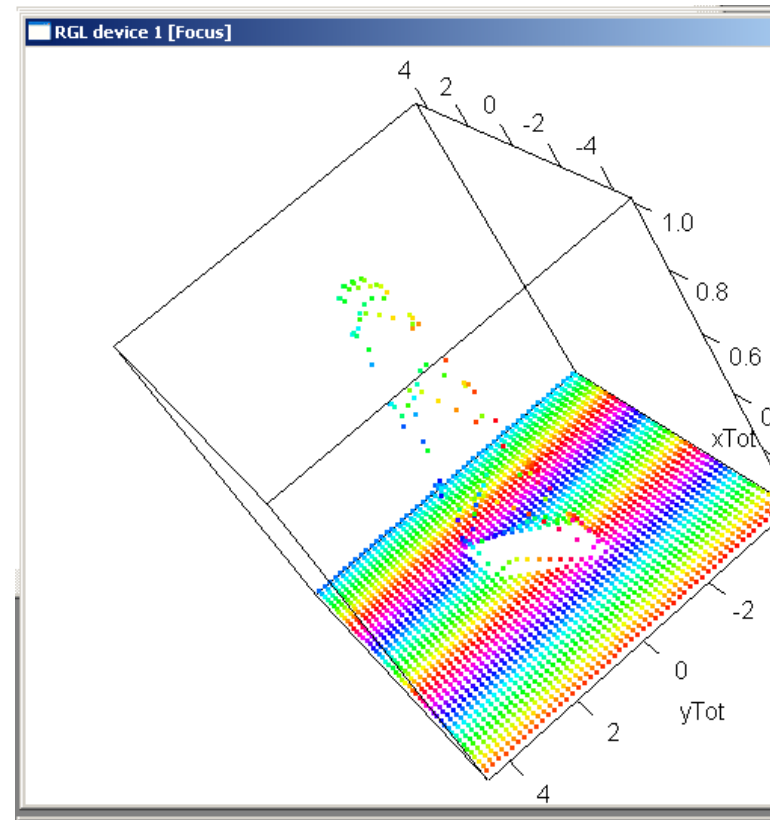
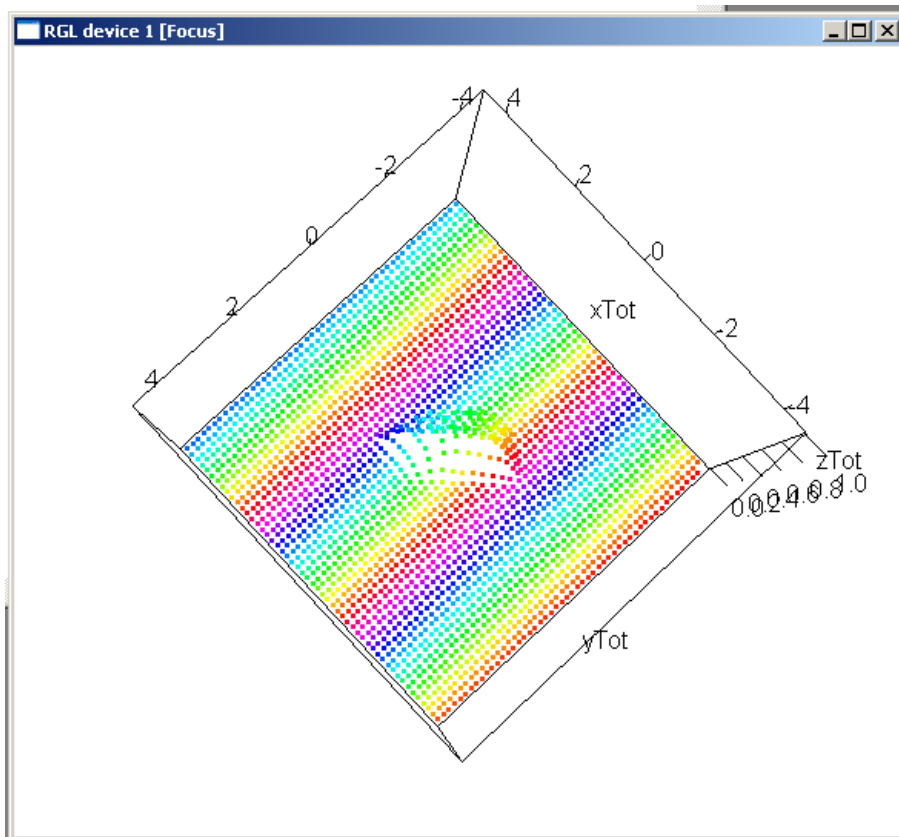
- No correlation exists between the variables.
- CMI with  $\sigma_1 = 1$ ,  $\sigma_2 = 1$ ,  $\rho_{12} = 0.0$ ,  $\mu_x = 0.01$ ,  $\mu_y = 0.01$ .



# Choquet integral based distribution: Examples

**2nd Example:** Interactions only in terms of a **covariance matrix**.

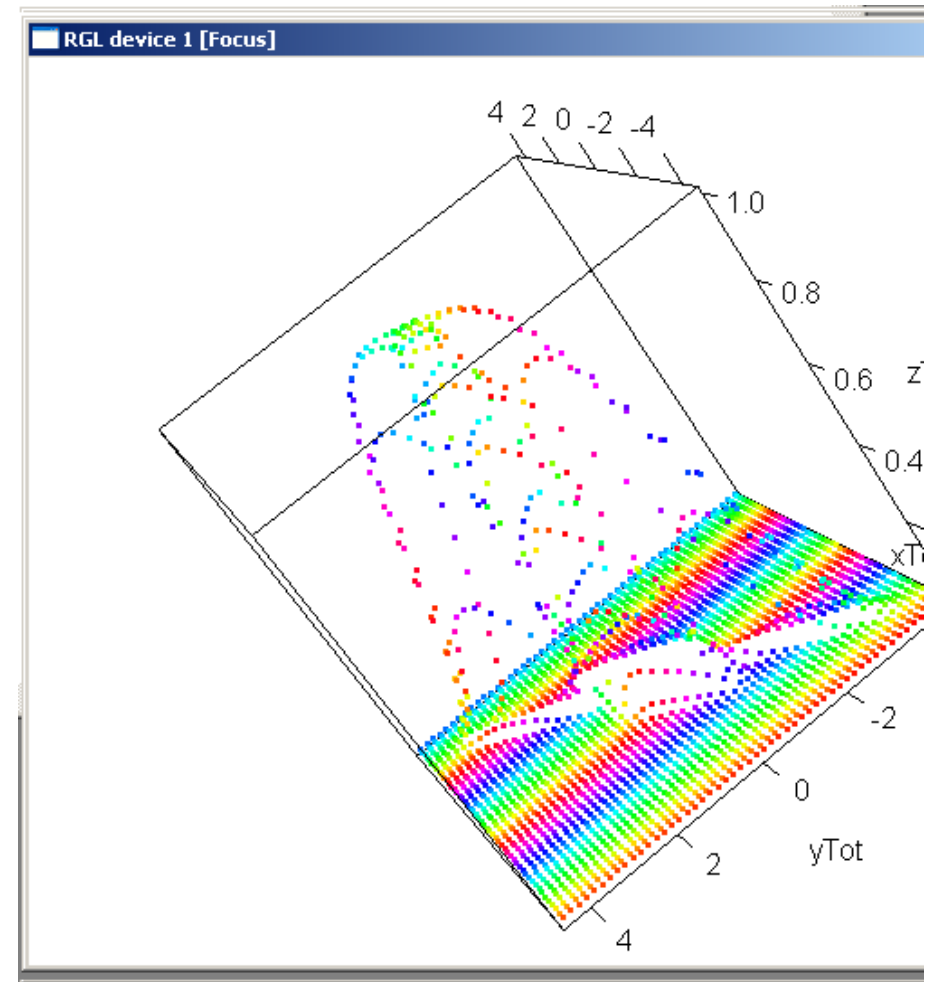
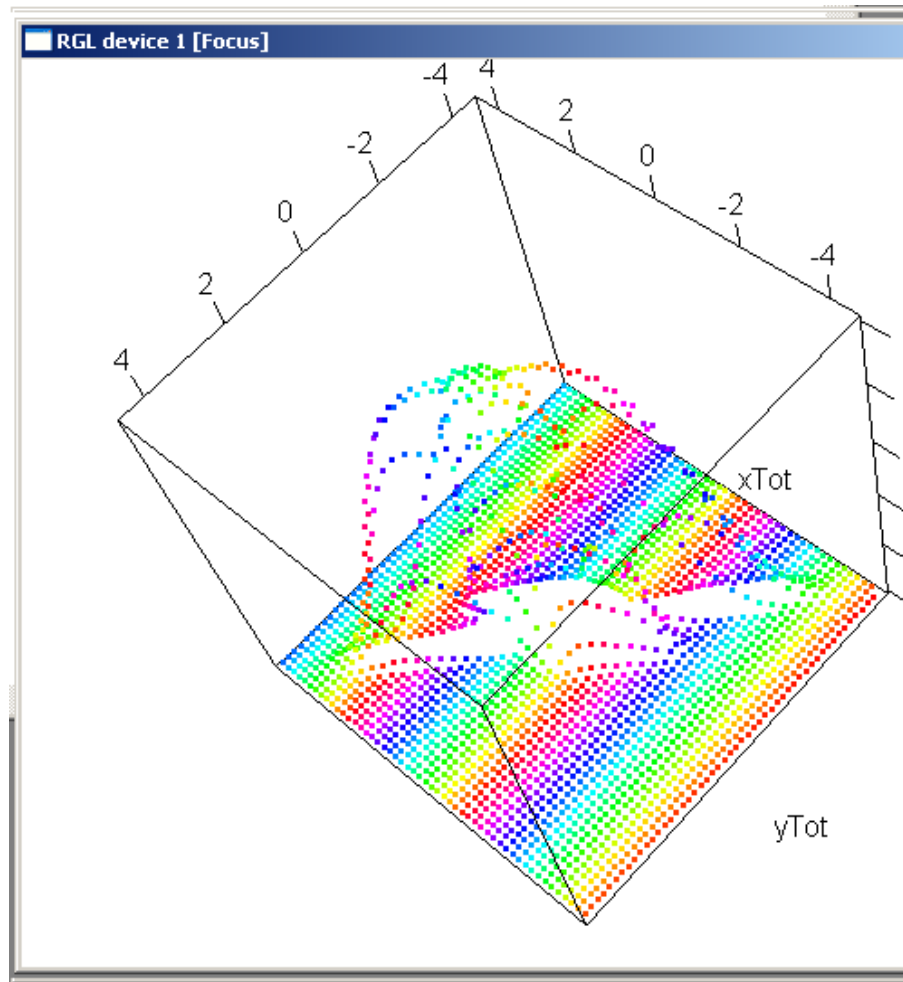
- CMI with  $\sigma_1 = 1$ ,  $\sigma_2 = 1$ ,  $\rho_{12} = 0.9$ ,  $\mu_x = 0.10$ ,  $\mu_y = 0.90$ .



# Choquet integral based distribution: Examples

**3rd Example:** Interactions both: covariance matrix and measure.

- CMI with  $\sigma_1 = 1$ ,  $\sigma_2 = 1$ ,  $\rho_{12} = 0.9$ ,  $\mu_x = 0.01$ ,  $\mu_y = 0.01$ .



# Choquet integral based distribution: Properties

---

**More properties:** Data not always acc. normality assumption

- spherical, elliptical distributions
- They generalize, respectively,  $N(\mathbf{0}, \mathbb{I})$  and  $N(\mathbf{m}, \Sigma)$
- Neither  $CMI(\mathbf{m}, \mu, \mathbf{Q}) \subseteq / \supseteq$  spherical / elliptical distributions.

**Example:**

- Non-additive  $\mu$ :  $CMI(\mathbf{m}, \mu, \mathbf{Q})$  not repr. spherical/elliptical
- No  $CMI$  for the following spherical distribution: Spherical distribution with density

$$f(r) = (1/K)e^{-\left(\frac{r-r_0}{\sigma}\right)^2},$$

where  $r_0$  is a radius over which the density is maximum,  $\sigma$  is a variance, and  $K$  is the normalization constant.

# Choquet integral based distribution: Properties

---

## More properties: (symmetry)

- $P(x)$  a  $C(\mathbf{m}, \mu)$  i.e., mean  $\mathbf{m} = (m_1, \dots, m_n)$  and a fuzzy measure  $\mu$ . Then, for all  $x \in \mathbb{R}^n$  and all  $i \in \{1, \dots, n\}$

$$P(x_1, \dots, x_{i-1}, x_i + m_i, x_{i+1}, \dots, x_n) = P(x_1, \dots, x_{i-1}, -x_i + m_i, x_{i+1}, \dots, x_n).$$

- $P(x)$  a  $CMI(\mathbf{m}, \mu, \mathbf{Q})$  i.e., with mean  $\mathbf{m} = (m_1, \dots, m_n)$ , a positive-definite **diagonal** matrix  $\mathbf{Q}$ , and a fuzzy measure  $\mu$ . Then, for all  $x \in \mathbb{R}^n$  and all  $i \in \{1, \dots, n\}$

$$P(x_1, \dots, x_{i-1}, x_i + m_i, x_{i+1}, \dots, x_n) = P(x_1, \dots, x_{i-1}, -x_i + m_i, x_{i+1}, \dots, x_n).$$



# Choquet integral based distribution: Properties

---

## More properties:

- $P(x)$  a  $C(\mathbf{m}, \mu)$  i.e., with mean  $\mathbf{m} = (m_1, \dots, m_n)$ . Then, for any fuzzy measure  $\mu$ ,
  - the mean vector  $\bar{\mathbf{X}} = [E[X_1], E[X_2], \dots, E[X_n]]$  is  $\mathbf{m}$  and
  - $\Sigma = [Cov[X_i, X_j]]$  for  $i = 1, \dots, n$  and  $j = 1, \dots, n$  is zero for all  $i \neq j$  and, thus, diagonal.
- $P(x)$  a  $CMI(\mathbf{m}, \mu, \mathbf{Q})$  i.e., with mean  $\mathbf{m} = (m_1, \dots, m_n)$ . Then, for any fuzzy measure  $\mu$  and any diagonal matrix  $\mathbf{Q}$ ,
  - the mean vector  $\bar{\mathbf{X}} = [E[X_1], E[X_2], \dots, E[X_n]]$  is  $\mathbf{m}$  and
  - $\Sigma = [Cov[X_i, X_j]]$  for  $i = 1, \dots, n$  and  $j = 1, \dots, n$  is zero for all  $i \neq j$  and thus, diagonal.

# Choquet integral based distribution: Properties

---

## More properties:

- When  $Q$  is not diagonal, we may have

$$\text{Cov}[X_i, X_j] \neq Q(X_i, X_j).$$

# Choquet integral based distribution: Properties

---

**More properties:** If this type of data distinguishable from Normal ?

## Study:

- Case of  $X = \{x_1, x_2\}$
- $CMI(0, \mu)$  with  $\mu(\{x\}) = i/10$  and  $\mu(\{y\}) = i/10$  for  $i = 1, 2, \dots, 9$
- Test: Normality test for CI-based distribution
  - Normality of the **marginals**
  - Normality of the multidimensional distribution

# Choquet integral based distribution: Properties

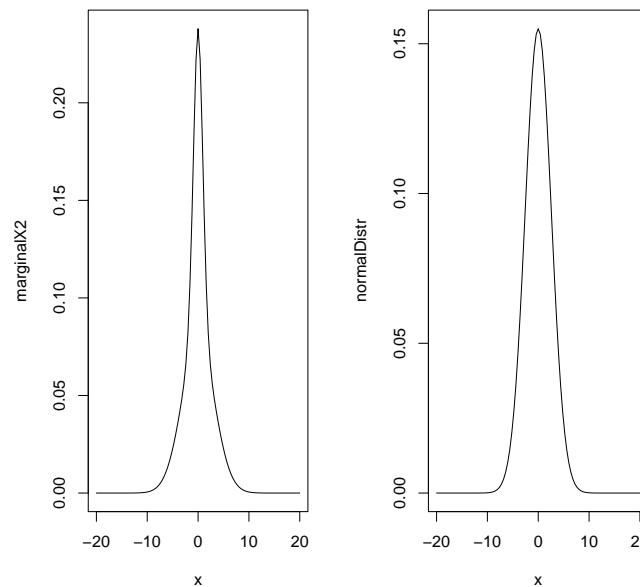
## More properties: Normality test for CI-based distribution

- Normality of the **marginals**: Shapiro-Wilk test

Marginal computed numerically integrate, uniroot function in R.

Almost always the **test is passed** for samples of  $n = 100$  data

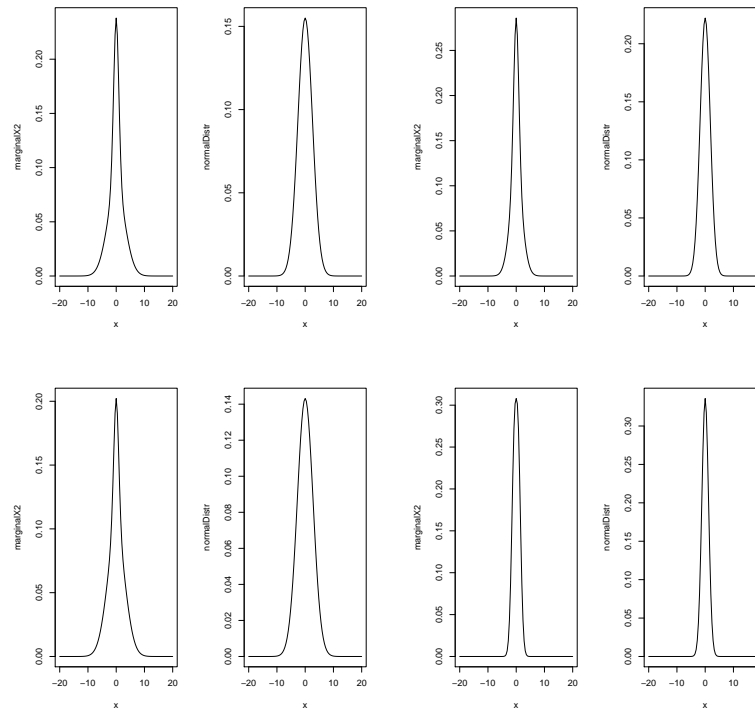
- Marginals (left) of the bivariate  $CI(\mathbf{0}, \mu)$ , and the normal distribution (right) with the same variance.  $\mu(\{x_1\}) = 0.1$  and  $\mu(\{x_2\}) = 0.1$



# Choquet integral based distribution: Properties

## More properties: Normality test for CI-based distribution

- Normality of the **marginals**: Shapiro-Wilk test
- Marginals (left) of  $CI(\mathbf{0}, \mu)$ , and (right)  $N$  same variance. (i)  $\mu(\{x_1\}) = 0.1$  and  $\mu(\{x_2\}) = 0.1$ ; (ii)  $\mu(\{x_1\}) = 0.1$  and  $\mu(\{x_2\}) = 0.2$ ; (iii)  $\mu(\{x_1\}) = 0.2$  and  $\mu(\{x_2\}) = 0.1$ ; (iv)  $\mu(\{x_1\}) = 0.9$  and  $\mu(\{x_2\}) = 0.9$



# Choquet integral based distribution: Properties

---

**More properties:** Normality test for CI-based distribution

- **Normality of the distribution:**

**Mardia's test** based on skewness and kurtosis

- **Skewness test is passed.**
- Almost all distributions (in  $\mathbb{R}^2$ ) **pass kurtosis test** in experiments:
  - $CI(0, \mu)$  distributions with  $\mu(\{x\}) = i/10$  and  $\mu(\{y\}) = i/10$  for  $i = 1, 2, \dots, 9$ .
  - **Test only fails in**
    - (i)  $\mu(\{x\}) = 0.1$  and  $\mu(\{y\}) = 0.1$ ,
    - (ii)  $\mu(\{x\}) = 0.2$  and  $\mu(\{y\}) = 0.1$ .

# Summary

# Summary

---

## Summary:

- Definition of distributions based on the Choquet integral  
Integral for non-additive measures
- Relationship with multivariate normal and spherical distributions

## Future work:

- Study of the properties
- Parameters determination from data  $(\mu, Q)$
- Statistical tests



# Summary

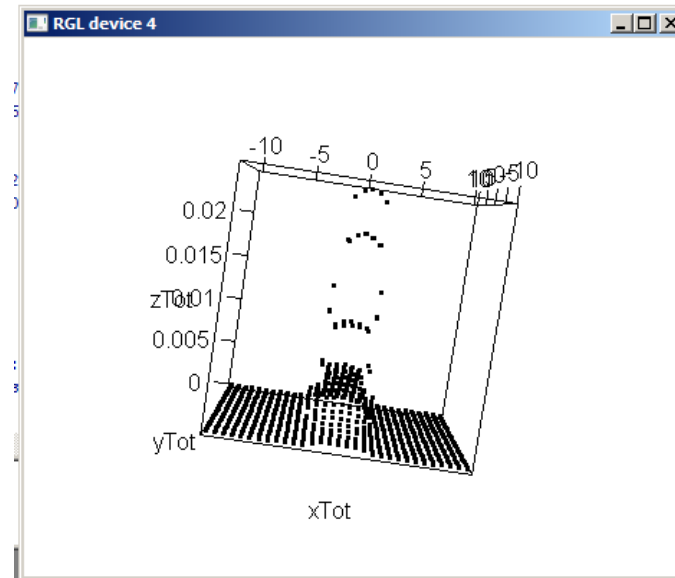
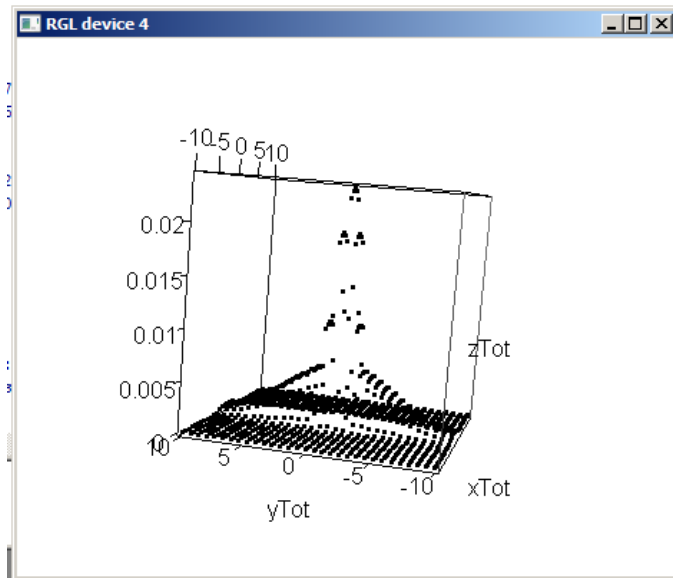
- Level-dependent capacity (non-additive, fuzzy measure)

Defined by S. Greco, B. Matarazzo, S. Giove (FSS, 2011)

- Level-dependent-based distribution (generalizes CI-based)

$$P(x) = \frac{1}{K} e^{-\frac{1}{2} CI_{\mu^G}^G((x-\bar{x}) \otimes (x-\bar{x}))}$$

- Example. Two perspectives of same level dependent CI. Defined by the same fuzzy measures  $\mu^1$  and  $\mu^2$  with intervals  $(0, 3)$  for  $\mu^1$ , and  $(3, 100)$  for  $\mu^2$ .



$$\mu^1(\{x\}) = 0.05 \text{ and } \mu^1(\{y\}) = 0.95, \text{ and } \mu^2(\{x\}) = 0.95 \text{ and } \mu^2(\{y\}) = 0.05$$

**Thank you**

# References

---

## References:

- V. Torra, Y. Narukawa, On a comparison between Mahalanobis distance and Choquet integral: the Choquet-Mahalanobis operator, *Information Sciences* 190 (2012) 56-63.
- V. Torra, Distributions based on the Choquet integral and non-additive measures, *RIMS Kokyuroku* 1906 (2014) 136-143.
- V. Torra, Some properties of Choquet integral based probability functions, *Acta et Commentationes Universitatis Tartuensis de Mathematica* 19:1 (2015) 35-47.  
<http://dx.doi.org/10.12697/ACUTM.2015.19.04>

---

# Thank you

Slides at:

<http://www.mdai.cat/ifao/>

<http://www.mdai.cat/ifao/slides.php>