Statistical tools for dissimilarity analysis

Francesco Andreoli
CEPS/INSTEAD

Canzei Winter School IT2015

Motivation

- **Dissimilarity comparisons** of **sets** of distributions.
- **Question**: which of the two **distribution matrices** displays more dissimilarity between its rows?

$$\mathbf{A} = \begin{matrix} & cl1 & cl2 & cl3 \\ gr1 & 0.6 & 0 & 0.4 \\ gr2 & 0 & 0.25 & 0.75 \\ gr3 & \frac{6}{16} & \frac{2}{16} & \frac{8}{16} \end{matrix} \quad \text{and} \quad \mathbf{B} = \begin{matrix} & cl1 & cl2 & cl3 \\ gr1 & 0.6 & 0.2 & 0.2 \\ gr2 & 0.375 & 0.25 & 0.375 \\ gr3 & \frac{8}{16} & \frac{4}{16} & \frac{4}{16} \end{matrix}$$

- **Focus** on matrices representing relative frequencies distributions of **groups** across **classes** in $\mathcal{M}_d$, like:

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & & & \vdots \\ a_{d1} & a_{d2} & \dots & a_{dn} \end{pmatrix} \quad \text{with} \quad a_{ij} \in [0,1]\ \forall i,j \quad \text{and} \quad \sum_{j=1}^{n} a_{ij} = 1 \quad \forall i.$$

Motivation

- **Dissimilarity comparisons** of **sets** of distributions.
- **Question**: which of the two **distribution matrices** displays more dissimilarity between its rows?

$$\mathbf{A} = \begin{matrix} & cl1 & cl2 & cl3 \\ gr1 & 0.6 & 0 & 0.4 \\ gr2 & 0 & 0.25 & 0.75 \\ gr3 & \frac{6}{16} & \frac{2}{16} & \frac{8}{16} \end{matrix} \quad \text{and} \quad \mathbf{B} = \begin{matrix} & cl1 & cl2 & cl3 \\ gr1 & 0.6 & 0.2 & 0.2 \\ gr2 & 0.375 & 0.25 & 0.375 \\ gr3 & \frac{8}{16} & \frac{4}{16} & \frac{4}{16} \end{matrix}$$

- **Focus** on matrices representing relative frequencies distributions of **groups** across **classes** in $\mathcal{M}_d$, like:

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & & & \vdots \\ a_{d1} & a_{d2} & \dots & a_{dn} \end{pmatrix} \quad \text{with} \quad a_{ij} \in [0,1]\ \forall i,j \quad \text{and} \quad \sum_{j=1}^{n} a_{ij} = 1 \quad \forall i.$$

Motivation

- **Dissimilarity comparisons** of **sets** of distributions.
- **Question**: which of the two **distribution matrices** displays more dissimilarity between its rows?

$$\mathbf{A} = \begin{array}{c} \\ gr1 \\ gr2 \\ gr3 \end{array} \begin{array}{c} \begin{array}{ccc} cl1 & cl2 & cl3 \end{array} \\ \begin{pmatrix} 0.6 & 0 & 0.4 \\ 0 & 0.25 & 0.75 \\ \frac{6}{16} & \frac{2}{16} & \frac{8}{16} \end{pmatrix} \end{array} \quad \text{and} \quad \mathbf{B} = \begin{array}{c} \\ gr1 \\ gr2 \\ gr3 \end{array} \begin{array}{c} \begin{array}{ccc} cl1 & cl2 & cl3 \end{array} \\ \begin{pmatrix} 0.6 & 0.2 & 0.2 \\ 0.375 & 0.25 & 0.375 \\ \frac{8}{16} & \frac{4}{16} & \frac{4}{16} \end{pmatrix} \end{array}$$

- **Focus** on matrices representing relative frequencies distributions of **groups** across **classes** in $\mathcal{M}_d$, like:

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & & & \vdots \\ a_{d1} & a_{d2} & \dots & a_{dn} \end{pmatrix} \quad \text{with} \quad a_{ij} \in [0,1] \ \forall i,j \quad \text{and} \quad \sum_{j=1}^{n} a_{ij} = 1 \quad \forall i.$$

Motivation

- **Proposal :** underpin **statistical tools** that allows to rank **B** as better than **A** when **B** is closer to a **similarity matrix** than **B** is.

$$\textbf{Similarity matrix:} \quad = \begin{pmatrix} a_1, & a_2, & \cdots, & a_k \\ a_1, & a_2, & \cdots, & a_k \\ \vdots & \vdots & \ddots & \vdots \\ a_1, & a_2, & \cdots, & a_k \end{pmatrix}$$

- **Gini 1914** defines two or more (say d) (relative) **frequency distributions** of the same variate (taking on k values) to be **similar** if:

> *"for any modality [...] the absolute frequencies are proportional. If two distributions are similar they can have different sizes but their syntheses which are based on relative frequencies are equal"*

Motivation

- Many potential applications:
  - **Segregation**
    (Duncan Duncan ASR55, Massey Denton SF88, Hutchens MSS91, Frankel Volij JET2011)
  - **Discrimination**
    (Le Breton et al. JET2012, Gastwirth AS75, Jenkins JEmetrics94, Butler McDonald JBES87)
  - **(Intergenerational) mobility analysis**
    (Shorrocks ECMA78, Tchen 1981, Markandya EER82, Dardanoni JET93)
  - **Inequality**, uni- and multi-dimensional
    (Marshal Olking Arnolds 2011, Koshevoy Mosler JASA96, Ebert Moyes ECMA2005)
  - **Distance analysis**
    (Ebert JET1984)
  - **Statistics/Linear algebra/Informativeness**
    (Ali Sivlerey RRSA61, Blackwell AMS53, Koshevoy Mosler JASA96, Torgersen 1992, Dahl LAA99)

Partial and complete orders

- These phenomena are all related to dissimilarity comparisons of two or more distributions
- One simple principle of evaluation consists in compressing the distributional information into an evaluation function (which is an index number)
  - *This is conclusive*: given two situations, they can always be ranked.
  - *This is not robust*: if you challenge the evaluation function, you may obtain different rankings.
- Geometric tests have been proposed to reflect agreement in a class of evaluation functions.
  - *Geometric* means that they can be empirically assessed via linear programming.
- In the case of **two groups**, research has focused on comparisons ("lies always above or below" tyeps of arguments) of **curves**, i.e. transformations of the data.
  - Ex: Segregation curves, Discrimination curves, Concentration curves, Lorenz curves.
  - *These tests produce robust evaluations* that reflect agreement in interesting classes of evaluation functions.
  - *These tests might be inconclusive*, since when two curves cross, nothing can be said on the extent of agreement.
- When there are **many groups (more than two)**, robust partial orders implemented by geometric tests become tricky. This is where **indicators kick in**.
  - The extensions of the geometric tests, and their characterization, for the multigroup case are in Andreoli Zoli (2014).
- In dissimilarity analysis, we consider two situations where different empirical criteria apply: the cases where classes are **permutable** and where they are exogenously **ordered**.

Partial and complete orders

- These phenomena are all related to dissimilarity comparisons of two or more distributions
- One simple principle of evaluation consists in compressing the distributional information into an evaluation function (which is an index number)
  - *This is conclusive*: given two situations, they can always be ranked.
  - *This is not robust*: if you challenge the evaluation function, you may obtain different rankings.
- Geometric tests have been proposed to reflect agreement in a class of evaluation functions.
  - *Geometric* means that they can be empirically assessed via linear programming.
- In the case of **two groups**, research has focused on comparisons ("lies always above or below" tyeps of arguments) of **curves**, i.e. transformations of the data.
  - Ex: Segregation curves, Discrimination curves, Concentration curves, Lorenz curves.
  - *These tests produce robust evaluations* that reflect agreement in interesting classes of evaluation functions.
  - *These tests might be inconclusive*, since when two curves cross, nothing can be said on the extent of agreement.
- When there are **many groups (more than two)**, robust partial orders implemented by geometric tests become tricky. This is where **indicators kick in**.
  - The extensions of the geometric tests, and their characterization, for the multigroup case are in Andreoli Zoli (2014).
- In dissimilarity analysis, we consider two situations where different empirical criteria apply: the cases where classes are **permutable** and where they are exogenously **ordered**.

Partial and complete orders

- These phenomena are all related to dissimilarity comparisons of two or more distributions
- One simple principle of evaluation consists in compressing the distributional information into an evaluation function (which is an index number)
  - *This is conclusive*: given two situations, they can always be ranked.
  - *This is not robust*: if you challenge the evaluation function, you may obtain different rankings.
- Geometric tests have been proposed to reflect agreement in a class of evaluation functions.
  - *Geometric* means that they can be empirically assessed via linear programming.
- In the case of **two groups**, research has focused on comparisons ("lies always above or below" tyeps of arguments) of **curves**, i.e. transformations of the data.
  - Ex: Segregation curves, Discrimination curves, Concentration curves, Lorenz curves.
  - *These tests produce robust evaluations* that reflect agreement in interesting classes of evaluation functions.
  - *These tests might be inconclusive*, since when two curves cross, nothing can be said on the extent of agreement.
- When there are **many groups (more than two)**, robust partial orders implemented by geometric tests become tricky. This is where **indicators kick in**.
  - The extensions of the geometric tests, and their characterization, for the multigroup case are in Andreoli Zoli (2014).
- In dissimilarity analysis, we consider two situations where different empirical criteria apply: the cases where classes are **permutable** and where they are exogenously **ordered**.

Partial and complete orders

- These phenomena are all related to dissimilarity comparisons of two or more distributions
- One simple principle of evaluation consists in compressing the distributional information into an evaluation function (which is an index number)
  - *This is conclusive*: given two situations, they can always be ranked.
  - *This is not robust*: if you challenge the evaluation function, you may obtain different rankings.
- Geometric tests have been proposed to reflect agreement in a class of evaluation functions.
  - *Geometric* means that they can be empirically assessed via linear programming.
- In the case of **two groups**, research has focused on comparisons ("lies always above or below" tyeps of arguments) of **curves**, i.e. transformations of the data.
  - Ex: Segregation curves, Discrimination curves, Concentration curves, Lorenz curves.
  - *These tests produce robust evaluations* that reflect agreement in interesting classes of evaluation functions.
  - *These tests might be inconclusive*, since when two curves cross, nothing can be said on the extent of agreement.
- When there are **many groups (more than two)**, robust partial orders implemented by geometric tests become tricky. This is where **indicators kick in**.
  - The extensions of the geometric tests, and their characterization, for the multigroup case are in Andreoli Zoli (2014).
- In dissimilarity analysis, we consider two situations where different empirical criteria apply: the cases where classes are **permutable** and where they are exogenously **ordered**.

Partial and complete orders

- These phenomena are all related to dissimilarity comparisons of two or more distributions
- One simple principle of evaluation consists in compressing the distributional information into an evaluation function (which is an index number)
  - *This is conclusive*: given two situations, they can always be ranked.
  - *This is not robust*: if you challenge the evaluation function, you may obtain different rankings.
- Geometric tests have been proposed to reflect agreement in a class of evaluation functions.
  - *Geometric* means that they can be empirically assessed via linear programming.
- In the case of **two groups**, research has focused on comparisons ("lies always above or below" tyeps of arguments) of **curves**, i.e. transformations of the data.
  - Ex: Segregation curves, Discrimination curves, Concentration curves, Lorenz curves.
  - *These tests produce robust evaluations* that reflect agreement in interesting classes of evaluation functions.
  - *These tests might be inconclusive*, since when two curves cross, nothing can be said on the extent of agreement.
- When there are **many groups (more than two)**, robust partial orders implemented by geometric tests become tricky. This is where **indicators kick in**.
  - The extensions of the geometric tests, and their characterization, for the multigroup case are in Andreoli Zoli (2014).
- In dissimilarity analysis, we consider two situations where different empirical criteria apply: the cases where classes are **permutable** and where they are exogenously **ordered**.

Partial and complete orders

- These phenomena are all related to dissimilarity comparisons of two or more distributions
- One simple principle of evaluation consists in compressing the distributional information into an evaluation function (which is an index number)
  - *This is conclusive*: given two situations, they can always be ranked.
  - *This is not robust*: if you challenge the evaluation function, you may obtain different rankings.
- Geometric tests have been proposed to reflect agreement in a class of evaluation functions.
  - *Geometric* means that they can be empirically assessed via linear programming.
- In the case of **two groups**, research has focused on comparisons ("lies always above or below" tyeps of arguments) of **curves**, i.e. transformations of the data.
  - Ex: Segregation curves, Discrimination curves, Concentration curves, Lorenz curves.
  - *These tests produce robust evaluations* that reflect agreement in interesting classes of evaluation functions.
  - *These tests might be inconclusive*, since when two curves cross, nothing can be said on the extent of agreement.
- When there are **many groups (more than two)**, robust partial orders implemented by geometric tests become tricky. This is where **indicators kick in**.
  - The extensions of the geometric tests, and their characterization, for the multigroup case are in Andreoli Zoli (2014).
- In dissimilarity analysis, we consider two situations where different empirical criteria apply: the cases where classes are **permutable** and where they are exogenously **ordered**.

Empirical cases

Phenomenon:	*Classes:* Non-order	*Classes:* Order	*Classes:* Cardinal	*Groups:* Non-order	*Groups:* Order
School segregation	✓			✓	
Inequality	✓			✓	
Earnings discrimination		✓	✓	✓	
Mobility		✓	✓		✓

- **Objective** : We will study partial orders of distribution matrices ranking $\mathbf{B} \preccurlyeq \mathbf{A}$ iff $\mathbf{A}$ displays "at leas as much dissimilarity/segregation/discrimination/mobility as" $\mathbf{B}$, that are based on geometric comparisons of curves. These curves represent the degree of dissimilarity among the distributions involved in the comparisons.

Empirical cases

Phenomenon:	*Classes:*			*Groups:*	
	Non-order	*Order*	*Cardinal*	*Non-order*	*Order*
School segregation	✓			✓	
Inequality	✓			✓	
Earnings discrimination		✓	✓	✓	
Mobility		✓	✓		✓

- **Objective** : We will study partial orders of distribution matrices ranking $\mathbf{B} \preccurlyeq \mathbf{A}$ iff $\mathbf{A}$ displays "at leas as much dissimilarity/segregation/discrimination/mobility as" $\mathbf{B}$, that are based on geometric comparisons of curves. These curves represent the degree of dissimilarity among the distributions involved in the comparisons.

Empirical cases

Phenomenon:	*Classes:* Non-order	*Classes:* Order	*Classes:* Cardinal	*Groups:* Non-order	*Groups:* Order
School segregation	✓			✓	
Inequality	✓			✓	
Earnings discrimination		✓	✓	✓	
Mobility		✓	✓		✓

- **Objective** : We will study partial orders of distribution matrices ranking $\mathbf{B} \preccurlyeq \mathbf{A}$ iff $\mathbf{A}$ displays "at leas as much dissimilarity/segregation/discrimination/mobility as" $\mathbf{B}$, that are based on geometric comparisons of curves. These curves represent the degree of dissimilarity among the distributions involved in the comparisons.

Empirical cases

Phenomenon:	*Classes:*			*Groups:*	
	Non-order	*Order*	*Cardinal*	*Non-order*	*Order*
School segregation	✓			✓	
Inequality	✓			✓	
Earnings discrimination		✓	✓	✓	
Mobility		✓	✓		✓

- **Objective** : We will study partial orders of distribution matrices ranking $\mathbf{B} \preccurlyeq \mathbf{A}$ iff $\mathbf{A}$ displays "at leas as much dissimilarity/segregation/discrimination/mobility as" $\mathbf{B}$, that are based on geometric comparisons of curves. These curves represent the degree of dissimilarity among the distributions involved in the comparisons.

Empirical cases

Phenomenon:	*Classes:* Non-order	*Classes:* Order	*Classes:* Cardinal	*Groups:* Non-order	*Groups:* Order
School segregation	✓			✓	
Inequality	✓			✓	
Earnings discrimination		✓	✓	✓	
Mobility		✓	✓		✓

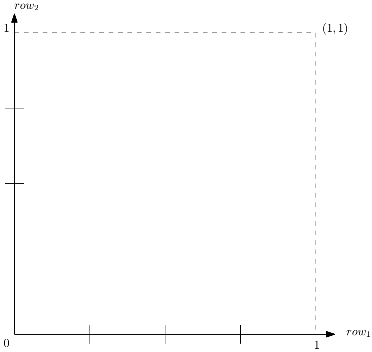
- **Objective** : We will study partial orders of distribution matrices ranking $\mathbf{B} \preccurlyeq \mathbf{A}$ iff $\mathbf{A}$ displays "at leas as much dissimilarity/segregation/discrimination/mobility as" $\mathbf{B}$, that are based on geometric comparisons of curves. These curves represent the degree of dissimilarity among the distributions involved in the comparisons.

Geometric criteria

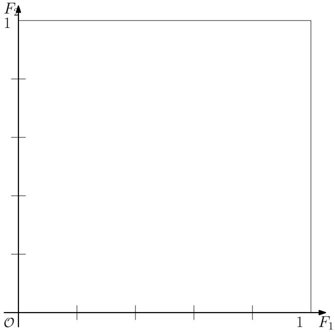
- Consider a distribution matrix **A** and the cumulation of its classes:

$$\mathbf{A} = \begin{pmatrix} 0 & 0.6 & 0.4 \\ 0.25 & 0 & 0.75 \\ X & X & X \end{pmatrix}$$

Zonotope $Z(\mathbf{A})$:



Path Polytope $Z^*(\mathbf{A})$:

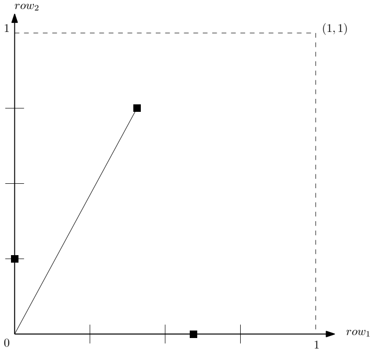


Geometric criteria

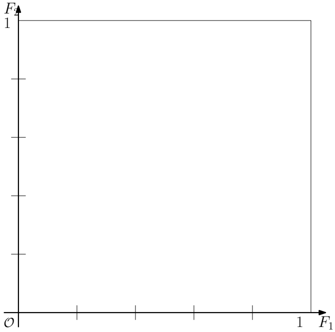
- Consider a distribution matrix **A** and the cumulation of its classes:

$$\mathbf{A} = \begin{pmatrix} 0 & 0.6 & 0.4 \\ 0.25 & 0 & 0.75 \\ X & X & X \end{pmatrix}$$

Zonotope $Z(\mathbf{A})$:



Path Polytope $Z^*(\mathbf{A})$:

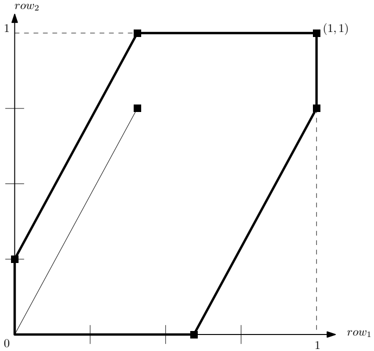


Geometric criteria

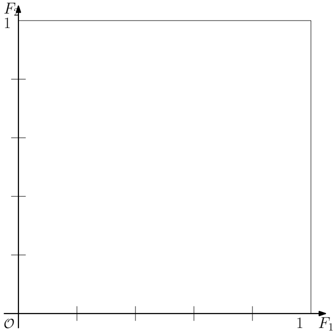
- Consider a distribution matrix **A** and the cumulation of its classes:

$$\mathbf{A} = \begin{pmatrix} 0 & 0.6 & 0.4 \\ 0.25 & 0 & 0.75 \\ X & X & X \end{pmatrix}$$

Zonotope $Z(\mathbf{A})$:



Path Polytope $Z^*(\mathbf{A})$:

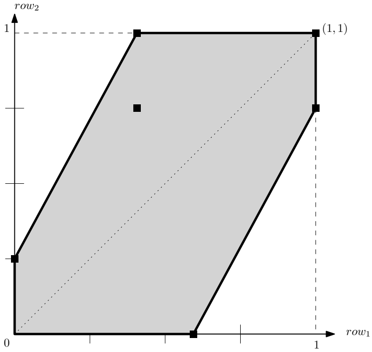


Geometric criteria

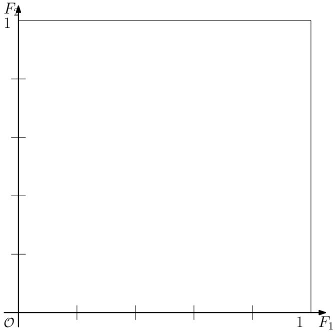
- Consider a distribution matrix **A** and the cumulation of its classes:

$$\mathbf{A} = \begin{pmatrix} 0 & 0.6 & 0.4 \\ 0.25 & 0 & 0.75 \\ X & X & X \end{pmatrix}$$

Zonotope $Z(\mathbf{A})$:



Path Polytope $Z^*(\mathbf{A})$:

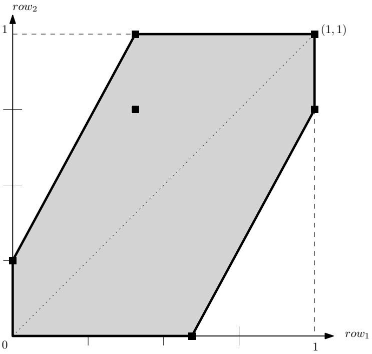


Geometric criteria

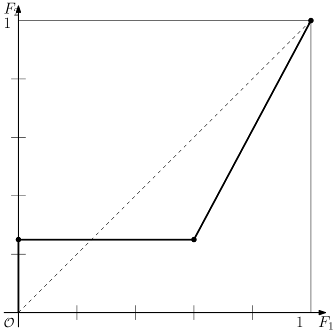
- Consider a distribution matrix **A** and the cumulation of its classes:

$$\mathbf{A} = \begin{pmatrix} 0 & 0.6 & 0.4 \\ 0.25 & 0 & 0.75 \\ X & X & X \end{pmatrix}$$

Zonotope $Z(\mathbf{A})$:



Path Polytope $Z^*(\mathbf{A})$:

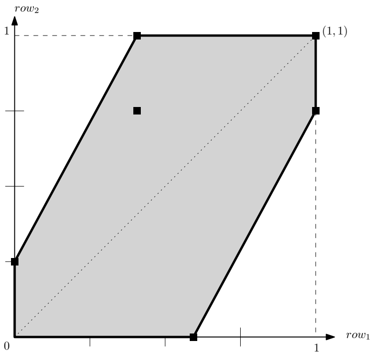


Geometric criteria

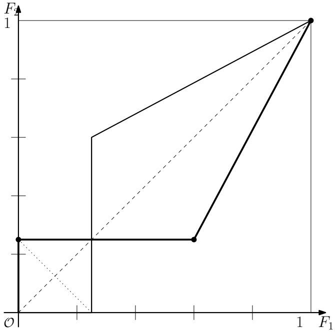
- Consider a distribution matrix $\mathbf{A}$ and the cumulation of its classes:

$$\mathbf{A} = \begin{pmatrix} 0 & 0.6 & 0.4 \\ 0.25 & 0 & 0.75 \\ X & X & X \end{pmatrix}$$

Zonotope $Z(\mathbf{A})$:



Path Polytope $Z^*(\mathbf{A})$:

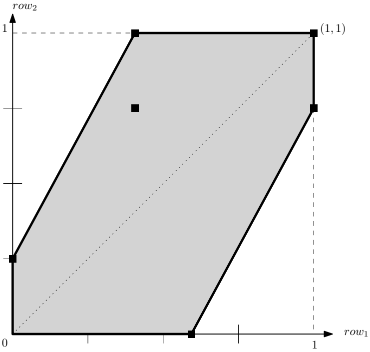


Geometric criteria

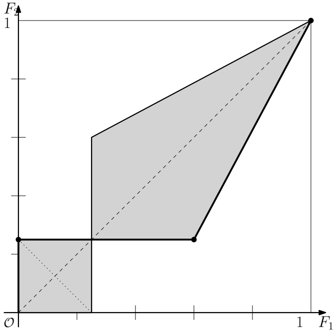
- Consider a distribution matrix **A** and the cumulation of its classes:

$$\mathbf{A} = \begin{pmatrix} 0 & 0.6 & 0.4 \\ 0.25 & 0 & 0.75 \\ X & X & X \end{pmatrix}$$

Zonotope $Z(\mathbf{A})$:



Path Polytope $Z^*(\mathbf{A})$:



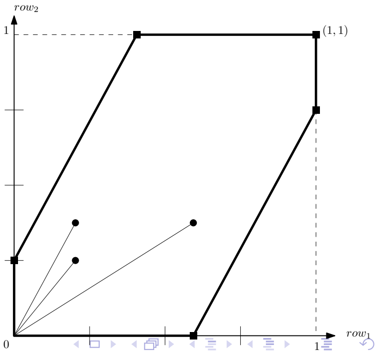
The Zonotopes inclusion test ($d = 2$)

- A suitable test for segregation comparisons, where matrices **A** and **B** may represent two cities/school districts/labor markets, and classes are neighborhoods/schools/jobs.
  - **Segregation curves** (Duncan Duncan ASR1955, Hutchens MSS1991) are the lower bound of the Zonotope. Thei ordering is related related to segregation-reducing movements of population: *when some members of the group overrepresented in a class move to a class where their group is underrepresented, segregation is reduced*.
  - **Local segregation curves** (Alonso-Villar, del Rio MSS2010) are segregation curves contrasting the distriution of each group overall population.

- **Bivariate case**: $d = 2$ groups and $n = 3$ classes

$$\mathbf{A}' := \begin{pmatrix} 0.6 & 0 & 0.4 \\ 0 & 0.25 & 0.75 \\ X & X & X \end{pmatrix}$$

$$\mathbf{B}' := \begin{pmatrix} 0.6 & 0.2 & 0.2 \\ 0.375 & 0.25 & 0.375 \\ X & X & X \end{pmatrix}$$

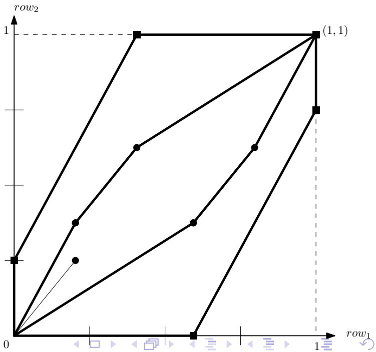


The Zonotopes inclusion test ($d = 2$)

- A suitable test for segregation comparisons, where matrices **A** and **B** may represent two cities/school districts/labor markets, and classes are neighborhoods/schools/jobs.
  - **Segregation curves** (Duncan Duncan ASR1955, Hutchens MSS1991) are the lower bound of the Zonotope. Thei ordering is related related to segregation-reducing movements of population: *when some members of the group overrepresented in a class move to a class where their group is underrepresented, segregation is reduced*.
  - **Local segregation curves** (Alonso-Villar, del Rio MSS2010) are segregation curves contrasting the distriution of each group overall population.
- **Bivariate case**: $d = 2$ groups and $n = 3$ classes

$$\mathbf{A}' := \begin{pmatrix} 0.6 & 0 & 0.4 \\ 0 & 0.25 & 0.75 \\ X & X & X \end{pmatrix}$$

$$\mathbf{B}' := \begin{pmatrix} 0.6 & 0.2 & 0.2 \\ 0.375 & 0.25 & 0.375 \\ X & X & X \end{pmatrix}$$

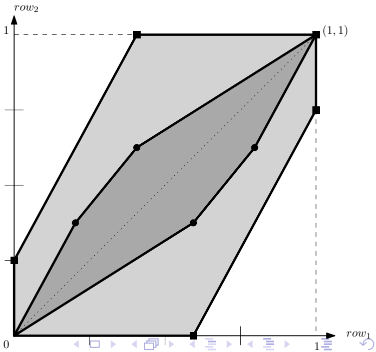


The Zonotopes inclusion test ($d = 2$)

- A suitable test for segregation comparisons, where matrices **A** and **B** may represent two cities/school districts/labor markets, and classes are neighborhoods/schools/jobs.
  - **Segregation curves** (Duncan Duncan ASR1955, Hutchens MSS1991) are the lower bound of the Zonotope. Thei ordering is related related to segregation-reducing movements of population: *when some members of the group overrepresented in a class move to a class where their group is underrepresented, segregation is reduced*.
  - **Local segregation curves** (Alonso-Villar, del Rio MSS2010) are segregation curves contrasting the distriution of each group overall population.
- **Bivariate case**: $d = 2$ groups and $n = 3$ classes

$$\mathbf{A}' := \begin{pmatrix} 0.6 & 0 & 0.4 \\ 0 & 0.25 & 0.75 \\ X & X & X \end{pmatrix}$$

$$\mathbf{B}' := \begin{pmatrix} 0.6 & 0.2 & 0.2 \\ 0.375 & 0.25 & 0.375 \\ X & X & X \end{pmatrix}$$



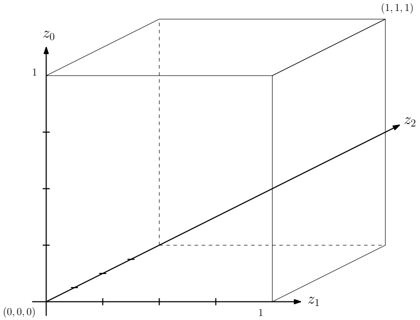
The Zonotopes inclusion test ($d > 2$)

- Extend to the multi-group case the segregation curve analysis (characterized by Andreoli and Zoli 2014)
- Its bivariate projections induce orderings coherent with segregation curves, although Zonotopes inclusion reflects the perspective of all projections.

- **Multi-group case**: $d = 3$ groups and $n = 3$ classes

$$\mathbf{A} := \begin{pmatrix} 0.6 & 0 & 0.4 \\ 0 & 0.25 & 0.75 \\ \frac{6}{16} & \frac{2}{16} & \frac{8}{16} \end{pmatrix}$$

$$\mathbf{B} := \begin{pmatrix} 0.6 & 0.2 & 0.2 \\ 0.375 & 0.25 & 0.375 \\ \frac{8}{16} & \frac{4}{16} & \frac{4}{16} \end{pmatrix}$$



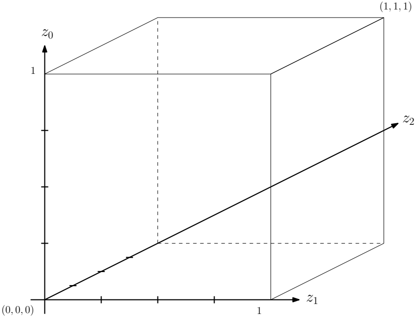
The Zonotopes inclusion test ($d > 2$)

- Extend to the multi-group case the segregation curve analysis (characterized by Andreoli and Zoli 2014)
- Its bivariate projections induce orderings coherent with segregation curves, although Zonotopes inclusion reflects the perspective of all projections.

- **Multi-group case**: $d = 3$ groups and $n = 3$ classes

$$\mathbf{A} := \begin{pmatrix} 0.6 & 0 & 0.4 \\ 0 & 0.25 & 0.75 \\ \frac{6}{16} & \frac{2}{16} & \frac{8}{16} \end{pmatrix}$$

$$\mathbf{B} := \begin{pmatrix} 0.6 & 0.2 & 0.2 \\ 0.375 & 0.25 & 0.375 \\ \frac{8}{16} & \frac{4}{16} & \frac{4}{16} \end{pmatrix}$$



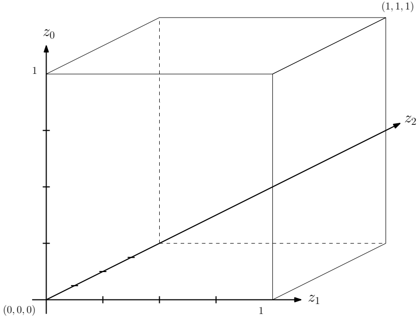
The Zonotopes inclusion test ($d > 2$)

- Extend to the multi-group case the segregation curve analysis (characterized by Andreoli and Zoli 2014)
- Its bivariate projections induce orderings coherent with segregation curves, although Zonotopes inclusion reflects the perspective of all projections.

- **Multi-group case**: $d = 3$ groups and $n = 3$ classes

$$\mathbf{A} := \begin{pmatrix} 0.6 & 0 & 0.4 \\ 0 & 0.25 & 0.75 \\ \frac{6}{16} & \frac{2}{16} & \frac{8}{16} \end{pmatrix}$$

$$\mathbf{B} := \begin{pmatrix} 0.6 & 0.2 & 0.2 \\ 0.375 & 0.25 & 0.375 \\ \frac{8}{16} & \frac{4}{16} & \frac{4}{16} \end{pmatrix}$$



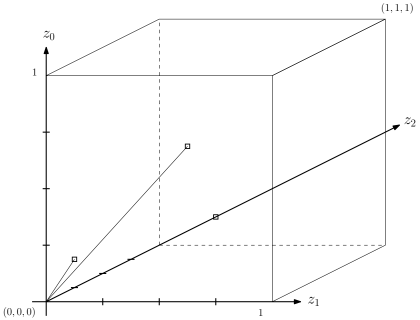
The Zonotopes inclusion test ($d > 2$)

- Extend to the multi-group case the segregation curve analysis (characterized by Andreoli and Zoli 2014)
- Its bivariate projections induce orderings coherent with segregation curves, although Zonotopes inclusion reflects the perspective of all projections.

- **Multi-group case**: $d = 3$ groups and $n = 3$ classes

$$\mathbf{A} := \begin{pmatrix} 0.6 & 0 & 0.4 \\ 0 & 0.25 & 0.75 \\ \frac{6}{16} & \frac{2}{16} & \frac{8}{16} \end{pmatrix}$$

$$\mathbf{B} := \begin{pmatrix} 0.6 & 0.2 & 0.2 \\ 0.375 & 0.25 & 0.375 \\ \frac{8}{16} & \frac{4}{16} & \frac{4}{16} \end{pmatrix}$$



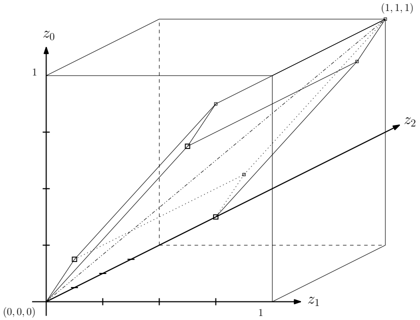
The Zonotopes inclusion test ($d > 2$)

- Extend to the multi-group case the segregation curve analysis (characterized by Andreoli and Zoli 2014)
- Its bivariate projections induce orderings coherent with segregation curves, although Zonotopes inclusion reflects the perspective of all projections.

- **Multi-group case**: $d = 3$ groups and $n = 3$ classes

$$\mathbf{A} := \begin{pmatrix} 0.6 & 0 & 0.4 \\ 0 & 0.25 & 0.75 \\ \frac{6}{16} & \frac{2}{16} & \frac{8}{16} \end{pmatrix}$$

$$\mathbf{B} := \begin{pmatrix} 0.6 & 0.2 & 0.2 \\ 0.375 & 0.25 & 0.375 \\ \frac{8}{16} & \frac{4}{16} & \frac{4}{16} \end{pmatrix}$$



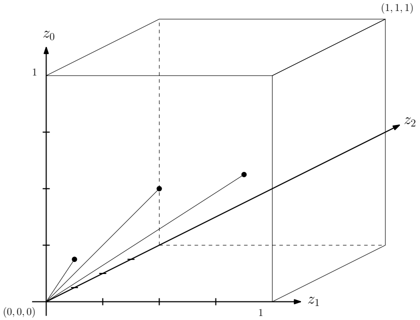
The Zonotopes inclusion test ($d > 2$)

- Extend to the multi-group case the segregation curve analysis (characterized by Andreoli and Zoli 2014)
- Its bivariate projections induce orderings coherent with segregation curves, although Zonotopes inclusion reflects the perspective of all projections.

- **Multi-group case**: $d = 3$ groups and $n = 3$ classes

$$\mathbf{A} := \begin{pmatrix} 0.6 & 0 & 0.4 \\ 0 & 0.25 & 0.75 \\ \frac{6}{16} & \frac{2}{16} & \frac{8}{16} \end{pmatrix}$$

$$\mathbf{B} := \begin{pmatrix} 0.6 & 0.2 & 0.2 \\ 0.375 & 0.25 & 0.375 \\ \frac{8}{16} & \frac{4}{16} & \frac{4}{16} \end{pmatrix}$$



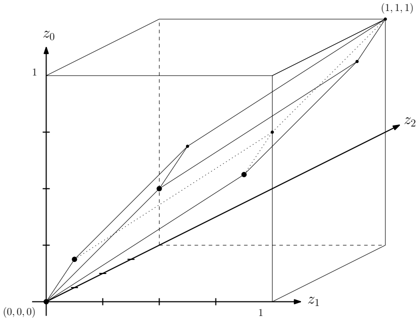
The Zonotopes inclusion test ($d > 2$)

- Extend to the multi-group case the segregation curve analysis (characterized by Andreoli and Zoli 2014)
- Its bivariate projections induce orderings coherent with segregation curves, although Zonotopes inclusion reflects the perspective of all projections.

- **Multi-group case**: $d = 3$ groups and $n = 3$ classes

$$\mathbf{A} := \begin{pmatrix} 0.6 & 0 & 0.4 \\ 0 & 0.25 & 0.75 \\ \frac{6}{16} & \frac{2}{16} & \frac{8}{16} \end{pmatrix}$$

$$\mathbf{B} := \begin{pmatrix} 0.6 & 0.2 & 0.2 \\ 0.375 & 0.25 & 0.375 \\ \frac{8}{16} & \frac{4}{16} & \frac{4}{16} \end{pmatrix}$$



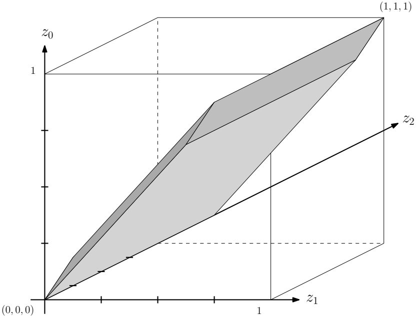
The Zonotopes inclusion test ($d > 2$)

- Extend to the multi-group case the segregation curve analysis (characterized by Andreoli and Zoli 2014)
- Its bivariate projections induce orderings coherent with segregation curves, although Zonotopes inclusion reflects the perspective of all projections.

- **Multi-group case**: $d = 3$ groups and $n = 3$ classes

$$\mathbf{A} := \begin{pmatrix} 0.6 & 0 & 0.4 \\ 0 & 0.25 & 0.75 \\ \frac{6}{16} & \frac{2}{16} & \frac{8}{16} \end{pmatrix}$$

$$\mathbf{B} := \begin{pmatrix} 0.6 & 0.2 & 0.2 \\ 0.375 & 0.25 & 0.375 \\ \frac{8}{16} & \frac{4}{16} & \frac{4}{16} \end{pmatrix}$$



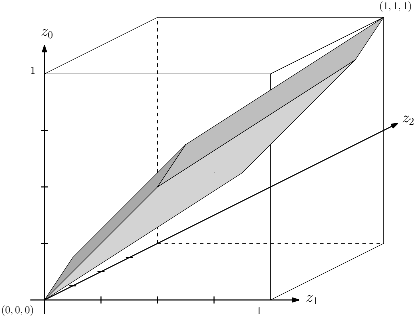
The Zonotopes inclusion test ($d > 2$)

- Extend to the multi-group case the segregation curve analysis (characterized by Andreoli and Zoli 2014)
- Its bivariate projections induce orderings coherent with segregation curves, although Zonotopes inclusion reflects the perspective of all projections.

- **Multi-group case**: $d = 3$ groups and $n = 3$ classes

$$\mathbf{A} := \begin{pmatrix} 0.6 & 0 & 0.4 \\ 0 & 0.25 & 0.75 \\ \frac{6}{16} & \frac{2}{16} & \frac{8}{16} \end{pmatrix}$$

$$\mathbf{B} := \begin{pmatrix} 0.6 & 0.2 & 0.2 \\ 0.375 & 0.25 & 0.375 \\ \frac{8}{16} & \frac{4}{16} & \frac{4}{16} \end{pmatrix}$$



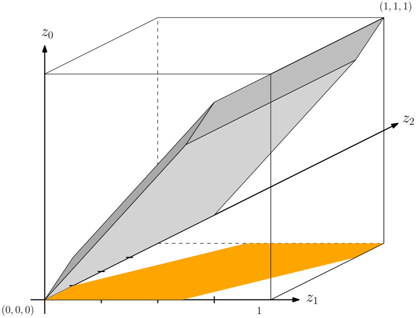
The Zonotopes inclusion test ($d > 2$)

- Extend to the multi-group case the segregation curve analysis (characterized by Andreoli and Zoli 2014)
- Its bivariate projections induce orderings coherent with segregation curves, although Zonotopes inclusion reflects the perspective of all projections.

- **Multi-group case**: $d = 3$ groups and $n = 3$ classes

$$\mathbf{A} := \begin{pmatrix} 0.6 & 0 & 0.4 \\ 0 & 0.25 & 0.75 \\ \frac{6}{16} & \frac{2}{16} & \frac{8}{16} \end{pmatrix}$$

$$\mathbf{B} := \begin{pmatrix} 0.6 & 0.2 & 0.2 \\ 0.375 & 0.25 & 0.375 \\ \frac{8}{16} & \frac{4}{16} & \frac{4}{16} \end{pmatrix}$$



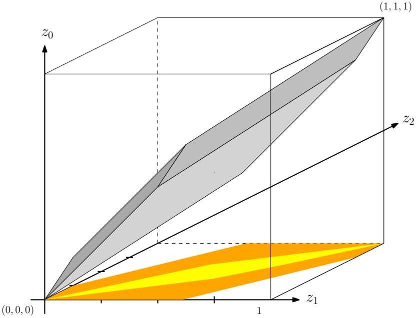
The Zonotopes inclusion test ($d > 2$)

- Extend to the multi-group case the segregation curve analysis (characterized by Andreoli and Zoli 2014)
- Its bivariate projections induce orderings coherent with segregation curves, although Zonotopes inclusion reflects the perspective of all projections.

- **Multi-group case**: $d = 3$ groups and $n = 3$ classes

$$\mathbf{A} \quad := \quad \begin{pmatrix} 0.6 & 0 & 0.4 \\ 0 & 0.25 & 0.75 \\ \frac{6}{16} & \frac{2}{16} & \frac{8}{16} \end{pmatrix}$$

$$\mathbf{B} \quad := \quad \begin{pmatrix} 0.6 & 0.2 & 0.2 \\ 0.375 & 0.25 & 0.375 \\ \frac{8}{16} & \frac{4}{16} & \frac{4}{16} \end{pmatrix}$$



Inequality as dissimilarity

- **Every** inequality comparison **is** a dissimilarity comparison, not the other way-round.
- **Income distributions**: $\mathbf{a} = (1,\ 1,\ 4)$ and $\mathbf{b} = (1,\ 2,\ 3)$
  - **b** is obtained from **a** through a set of **rich to poor** transfers.
  - vector **b** is obtained from **a** by a **T-transform**:

$$(1, 1, 4) \cdot (\mathbf{T} - \mathbf{transform}) = (1, 2, 3)$$

where:

$$(\mathbf{T} - \mathbf{transform}) := \frac{2}{3}\begin{pmatrix} 1 & 0 & 0 \\ 0 & \mathbf{1} & \mathbf{0} \\ 0 & \mathbf{0} & \mathbf{1} \end{pmatrix} + \frac{1}{3}\begin{pmatrix} 1 & 0 & 0 \\ 0 & \mathbf{0} & \mathbf{1} \\ 0 & \mathbf{1} & \mathbf{0} \end{pmatrix}$$

Inequality as dissimilarity

- **Every** inequality comparison **is** a dissimilarity comparison, not the other way-round.
- **Income distributions**: $\mathbf{a} = (1,\ 1,\ 4)$ and $\mathbf{b} = (1,\ 2,\ 3)$
  - **b** is obtained from **a** through a set of **rich to poor** transfers.
  - vector **b** is obtained from **a** by a **T-transform**:

$$(1, 1, 4) \cdot (\mathbf{T} - \mathbf{transform}) \ = \ (1, 2, 3)$$

where:

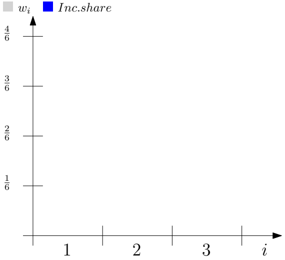
$$(\mathbf{T} - \mathbf{transform}) := \frac{2}{3}\begin{pmatrix} 1 & 0 & 0 \\ 0 & \mathbf{1} & \mathbf{0} \\ 0 & \mathbf{0} & \mathbf{1} \end{pmatrix} + \frac{1}{3}\begin{pmatrix} 1 & 0 & 0 \\ 0 & \mathbf{0} & \mathbf{1} \\ 0 & \mathbf{1} & \mathbf{0} \end{pmatrix}$$

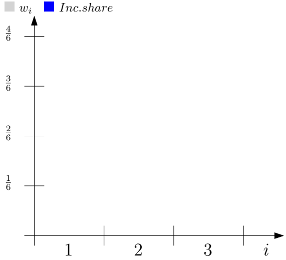
Inequality as dissimilarity

- T-transform are related to operations that reduce the overall dissimilarity between income shares distributions and the population weight distribution.

- Equivalently for **dissimilarity matrices**:

$$\mathbf{A} = \begin{array}{c} \\ w_i \\ Inc.sh. \end{array} \begin{array}{c} \begin{array}{ccc} i_1 & i_2 & i_3 \end{array} \\ \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{6} & \frac{1}{6} & \frac{4}{6} \end{pmatrix} \end{array} \quad \text{and} \quad \mathbf{B} = \begin{array}{c} \\ w_i \\ Inc.sh. \end{array} \begin{array}{c} \begin{array}{ccc} i_1 & i_2 & i_3 \end{array} \\ \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{6} & \frac{2}{6} & \frac{3}{6} \end{pmatrix} \end{array}$$



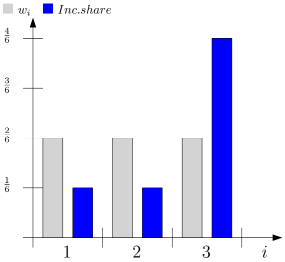


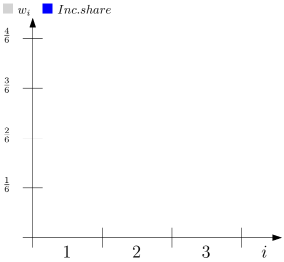
Inequality as dissimilarity

- T-transform are related to operations that reduce the overall dissimilarity between income shares distributions and the population weight distribution.

- Equivalently for **dissimilarity matrices**:

$$\mathbf{A} = \begin{matrix} \\ w_i \\ Inc.sh. \end{matrix} \begin{matrix} i_1 & i_2 & i_3 \\ \left(\begin{matrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{6} & \frac{1}{6} & \frac{4}{6} \end{matrix}\right) \end{matrix} \quad \text{and} \quad \mathbf{B} = \begin{matrix} \\ w_i \\ Inc.sh. \end{matrix} \begin{matrix} i_1 & i_2 & i_3 \\ \left(\begin{matrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{6} & \frac{2}{6} & \frac{3}{6} \end{matrix}\right) \end{matrix}$$



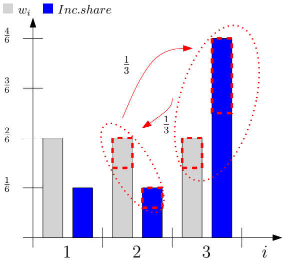


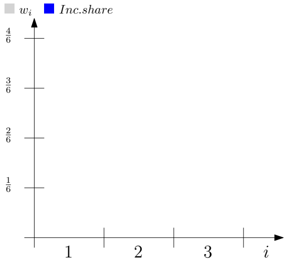
Inequality as dissimilarity

- T-transform are related to operations that reduce the overall dissimilarity between income shares distributions and the population weight distribution.

- Equivalently for **dissimilarity matrices**:

$$\mathbf{A} = \begin{matrix} & i_1 & i_2 & i_3 \\ w_i & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ Inc.sh. & \frac{1}{6} & \frac{1}{6} & \frac{4}{6} \end{matrix} \quad \text{and} \quad \mathbf{B} = \begin{matrix} & i_1 & i_2 & i_3 \\ w_i & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ Inc.sh. & \frac{1}{6} & \frac{2}{6} & \frac{3}{6} \end{matrix}$$



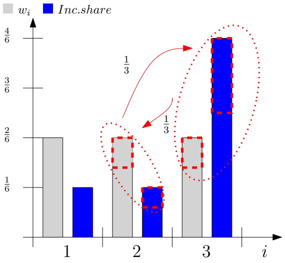


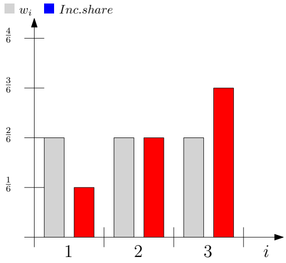
Inequality as dissimilarity

- T-transform are related to operations that reduce the overall dissimilarity between income shares distributions and the population weight distribution.

- Equivalently for **dissimilarity matrices**:

$$\mathbf{A} = \begin{matrix} & i_1 & i_2 & i_3 \\ w_i & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ Inc.sh. & \frac{1}{6} & \frac{1}{6} & \frac{4}{6} \end{matrix} \quad \text{and} \quad \mathbf{B} = \begin{matrix} & i_1 & i_2 & i_3 \\ w_i & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ Inc.sh. & \frac{1}{6} & \frac{2}{6} & \frac{3}{6} \end{matrix}$$

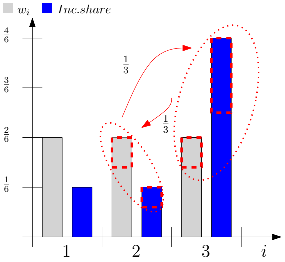


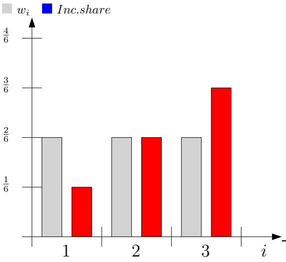


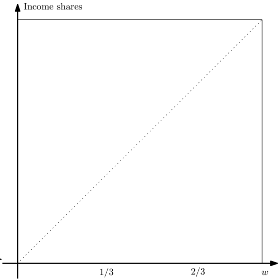
Inequality as dissimilarity

- T-transform are related to operations that reduce the overall dissimilarity between income shares distributions and the population weight distribution.
- Equivalently for **dissimilarity matrices**:

$$\mathbf{A} = \begin{array}{c} \\ w_i \\ Inc.sh. \end{array} \begin{array}{c} \begin{array}{ccc} i_1 & i_2 & i_3 \end{array} \\ \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{6} & \frac{1}{6} & \frac{4}{6} \end{pmatrix} \end{array} \quad \text{and} \quad \mathbf{B} = \begin{array}{c} \\ w_i \\ Inc.sh. \end{array} \begin{array}{c} \begin{array}{ccc} i_1 & i_2 & i_3 \end{array} \\ \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{6} & \frac{2}{6} & \frac{3}{6} \end{pmatrix} \end{array}$$



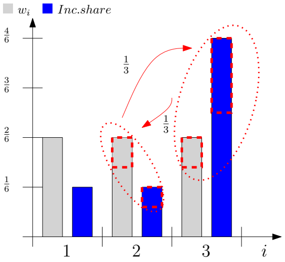


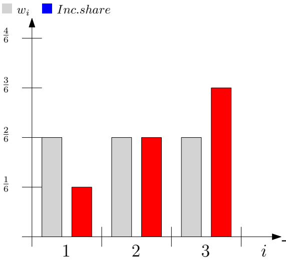


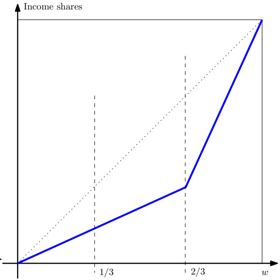
Inequality as dissimilarity

- T-transform are related to operations that reduce the overall dissimilarity between income shares distributions and the population weight distribution.
- Equivalently for **dissimilarity matrices**:

$$\mathbf{A} = \begin{matrix} & i_1 & i_2 & i_3 \\ w_i & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ Inc.sh. & \frac{1}{6} & \frac{1}{6} & \frac{4}{6} \end{matrix} \quad \text{and} \quad \mathbf{B} = \begin{matrix} & i_1 & i_2 & i_3 \\ w_i & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ Inc.sh. & \frac{1}{6} & \frac{2}{6} & \frac{3}{6} \end{matrix}$$



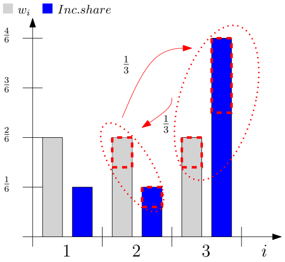


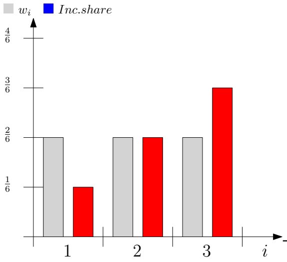


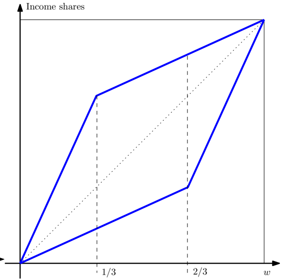
Inequality as dissimilarity

- T-transform are related to operations that reduce the overall dissimilarity between income shares distributions and the population weight distribution.
- Equivalently for **dissimilarity matrices**:

$$\mathbf{A} = \begin{matrix} & i_1 & i_2 & i_3 \\ w_i & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ Inc.sh. & \frac{1}{6} & \frac{1}{6} & \frac{4}{6} \end{matrix} \quad \text{and} \quad \mathbf{B} = \begin{matrix} & i_1 & i_2 & i_3 \\ w_i & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ Inc.sh. & \frac{1}{6} & \frac{2}{6} & \frac{3}{6} \end{matrix}$$



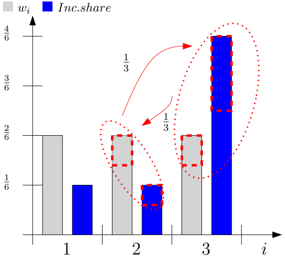


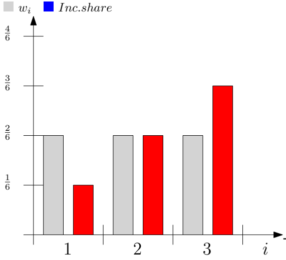


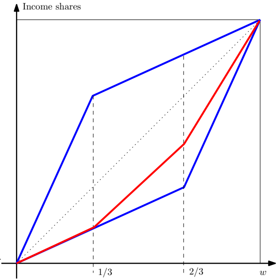
Inequality as dissimilarity

- T-transform are related to operations that reduce the overall dissimilarity between income shares distributions and the population weight distribution.
- Equivalently for **dissimilarity matrices**:

$$\mathbf{A} = \begin{array}{c} \\ w_i \\ Inc.sh. \end{array} \begin{array}{c} \begin{array}{ccc} i_1 & i_2 & i_3 \end{array} \\ \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{6} & \frac{1}{6} & \frac{4}{6} \end{pmatrix} \end{array} \quad \text{and} \quad \mathbf{B} = \begin{array}{c} \\ w_i \\ Inc.sh. \end{array} \begin{array}{c} \begin{array}{ccc} i_1 & i_2 & i_3 \end{array} \\ \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{6} & \frac{2}{6} & \frac{3}{6} \end{pmatrix} \end{array}$$



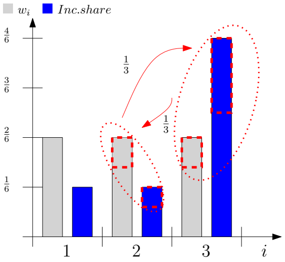


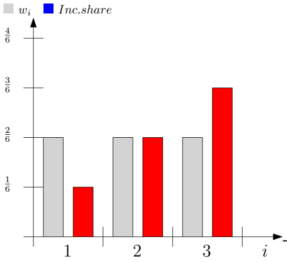


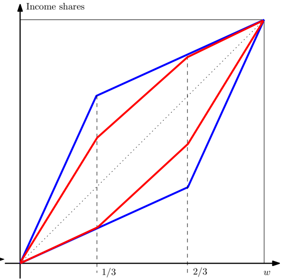
Inequality as dissimilarity

- T-transform are related to operations that reduce the overall dissimilarity between income shares distributions and the population weight distribution.
- Equivalently for **dissimilarity matrices**:

$$\mathbf{A} = \begin{matrix} & i_1 & i_2 & i_3 \\ w_i & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ Inc.sh. & \frac{1}{6} & \frac{1}{6} & \frac{4}{6} \end{matrix} \quad \text{and} \quad \mathbf{B} = \begin{matrix} & i_1 & i_2 & i_3 \\ w_i & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ Inc.sh. & \frac{1}{6} & \frac{2}{6} & \frac{3}{6} \end{matrix}$$



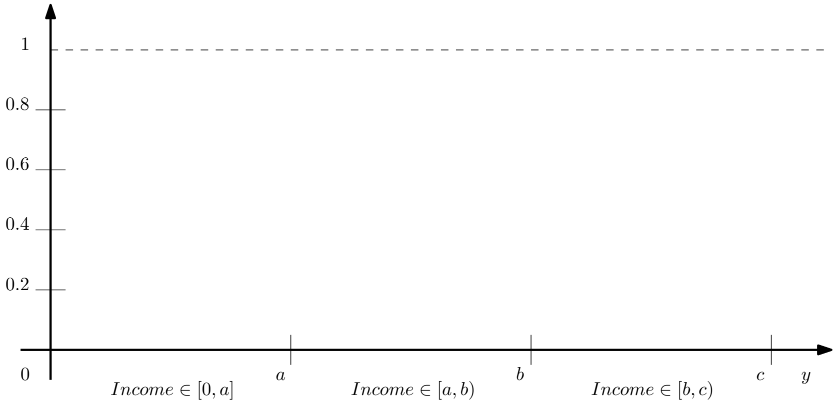




Incoherence of the merge in the ordered case

- **Ordered case**: *cdfs* are defined over exogenously ordered classes:

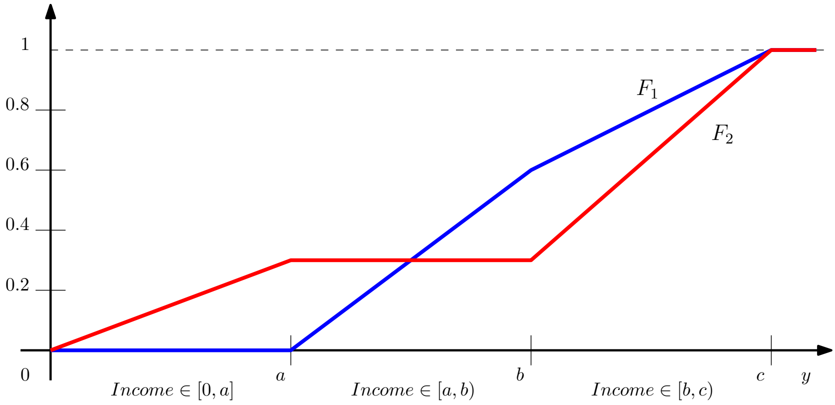
$$\mathbf{A}' = \begin{pmatrix} cdf \;\; of \;\; group \; 1 \\ cdf \;\; of \;\; group \; 2 \\ cdf \;\; of \;\; group \; 3 \end{pmatrix} = \begin{pmatrix} \mathbf{0} & \mathbf{0.6} & \mathbf{0.4} \\ \mathbf{0.25} & \mathbf{0} & \mathbf{0.75} \\ X & X & X \end{pmatrix}$$



Incoherence of the merge in the ordered case

- **Ordered case**: *cdf*s are defined over exogenously ordered classes:

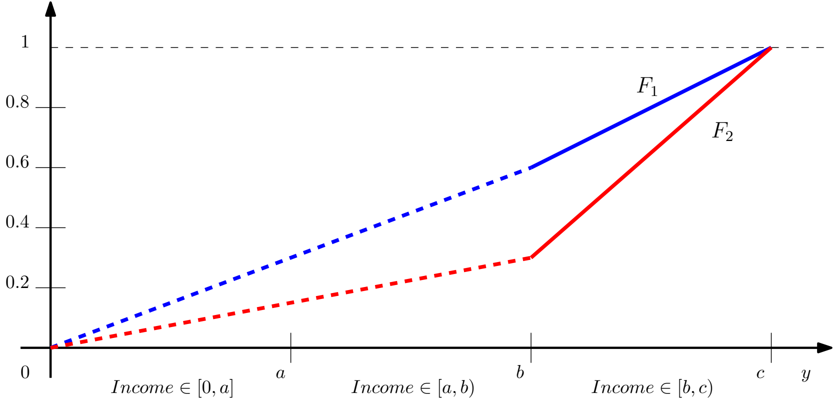
$$\mathbf{A}' = \begin{pmatrix} cdf \ of \ group \ 1 \\ cdf \ of \ group \ 2 \\ cdf \ of \ group \ 3 \end{pmatrix} = \begin{pmatrix} \mathbf{0} & \mathbf{0.6} & \mathbf{0.4} \\ \mathbf{0.25} & \mathbf{0} & \mathbf{0.75} \\ X & X & X \end{pmatrix}$$



Incoherence of the merge in the ordered case

- **Ordered case**: *cdf*s are defined over exogenously ordered classes:

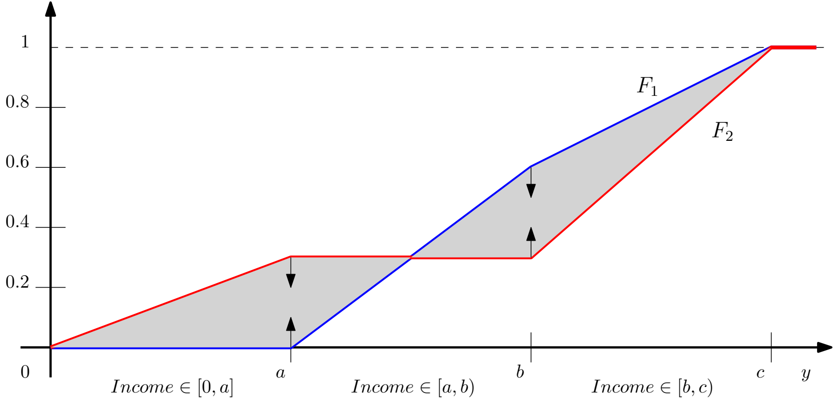
$$\mathbf{A}' = \begin{pmatrix} cdf\ \ of\ \ group\ 1 \\ cdf\ \ of\ \ group\ 2 \\ cdf\ \ of\ \ group\ 3 \end{pmatrix} = \begin{pmatrix} \mathbf{0} & \mathbf{0.6} & \mathbf{0.4} \\ \mathbf{0.25} & \mathbf{0} & \mathbf{0.75} \\ X & X & X \end{pmatrix}$$



Incoherence of the merge in the ordered case

- **Ordered case**: *cdf*s are defined over exogenously ordered classes:

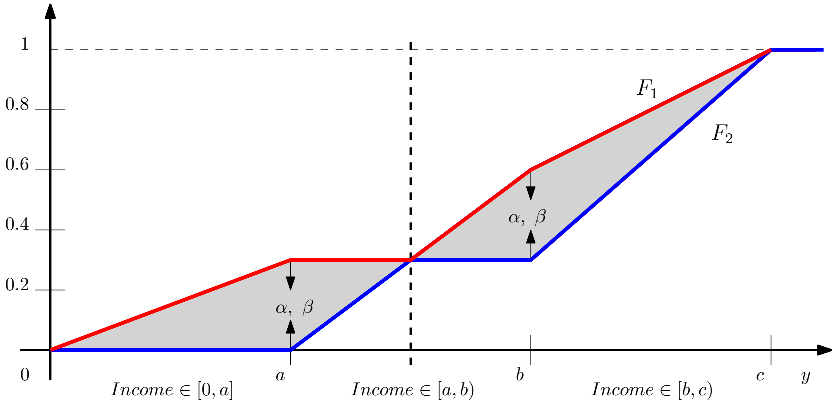
$$\mathbf{A}' = \begin{pmatrix} cdf \ \ of \ \ group \ 1 \\ cdf \ \ of \ \ group \ 2 \\ cdf \ \ of \ \ group \ 3 \end{pmatrix} = \begin{pmatrix} \mathbf{0} & \mathbf{0.6} & \mathbf{0.4} \\ \mathbf{0.25} & \mathbf{0} & \mathbf{0.75} \\ X & X & X \end{pmatrix}$$



Incoherence of the merge in the ordered case

- **Ordered case**: *cdf*s are defined over exogenously ordered classes:

$$\mathbf{A}' = \begin{pmatrix} cdf \;\; of \;\; group \; 1 \\ cdf \;\; of \;\; group \; 2 \\ cdf \;\; of \;\; group \; 3 \end{pmatrix} = \begin{pmatrix} \mathbf{0} & \mathbf{0.6} & \mathbf{0.4} \\ \mathbf{0.25} & \mathbf{0} & \mathbf{0.75} \\ X & X & X \end{pmatrix}$$

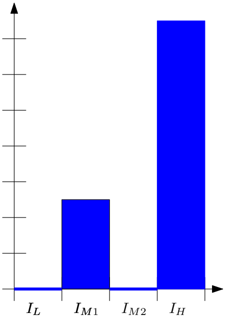


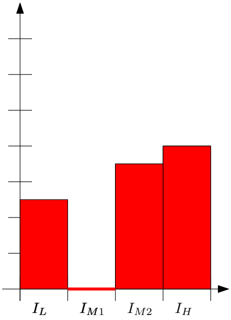
Dissimilarity reducing transfers for ordered classes

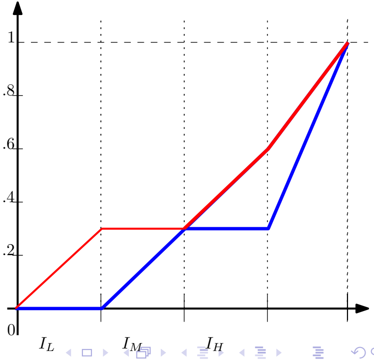
Axiom (The exchange operation (Tchen AP1980, Van de gaer et al Ecmica2001):)

Every transfer of population masses (ϵ) across adjacent classes that is progressive for the **dominating group** *and regressive for the* **dominated group** *reduces dissimilarity.*

$$\mathbf{A}' = \begin{pmatrix} 0 & 0.25 & 0 & 0.75 \\ 0.25 & 0 & 0.35 & 0.40 \\ X & X & X & X \end{pmatrix} \quad \text{with} \quad \overrightarrow{\mathbf{A}'} = \begin{pmatrix} 0 & 0.25 & 0.25 & 1 \\ 0.25 & 0.25 & 0.6 & 1 \\ X & X & X & X \end{pmatrix}$$





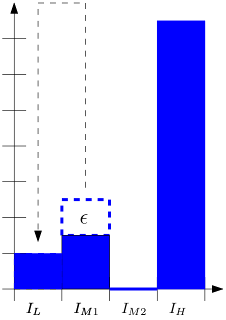


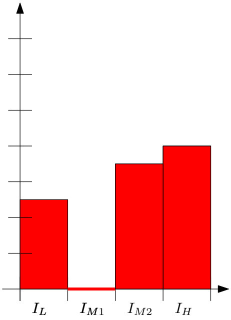
Dissimilarity reducing transfers for ordered classes

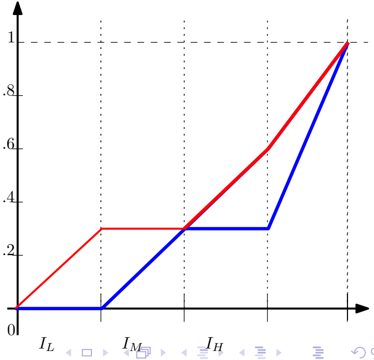
Axiom (The exchange operation (Tchen AP1980, Van de gaer et al Ecmica2001):)

Every transfer of population masses (ϵ) across adjacent classes that is progressive for the **dominating group** *and regressive for the* **dominated group** *reduces dissimilarity.*

$$\mathbf{A}' = \begin{pmatrix} 0 & 0.25 & 0 & 0.75 \\ 0.25 & 0 & 0.35 & 0.40 \\ X & X & X & X \end{pmatrix} \quad \text{with} \quad \overrightarrow{\mathbf{A}'} = \begin{pmatrix} 0 & 0.25 & 0.25 & 1 \\ 0.25 & 0.25 & 0.6 & 1 \\ X & X & X & X \end{pmatrix}$$





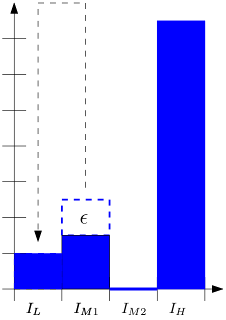


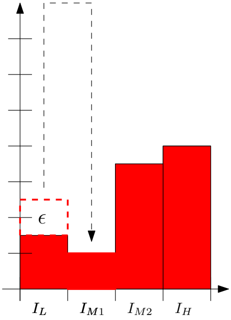
Dissimilarity reducing transfers for ordered classes

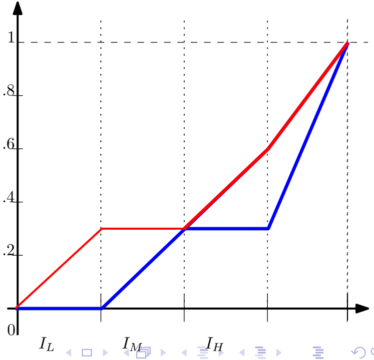
Axiom (The exchange operation (Tchen AP1980, Van de gaer et al Ecmica2001):)

Every transfer of population masses (ϵ) across adjacent classes that is progressive for the **dominating group** *and regressive for the* **dominated group** *reduces dissimilarity.*

$$\mathbf{A}' = \begin{pmatrix} 0 & 0.25 & 0 & 0.75 \\ 0.25 & 0 & 0.35 & 0.40 \\ X & X & X & X \end{pmatrix} \quad \text{with} \quad \overrightarrow{\mathbf{A}'} = \begin{pmatrix} 0 & 0.25 & 0.25 & 1 \\ 0.25 & 0.25 & 0.6 & 1 \\ X & X & X & X \end{pmatrix}$$





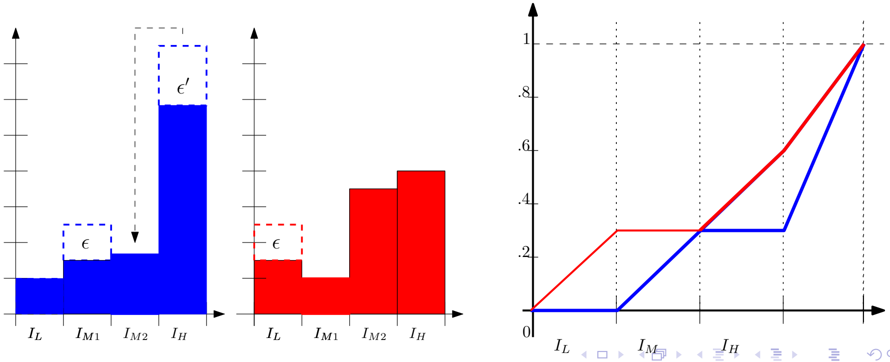


Dissimilarity reducing transfers for ordered classes

Axiom (The exchange operation (Tchen AP1980, Van de gaer et al Ecmica2001):)

Every transfer of population masses (ϵ) across adjacent classes that is progressive for the **dominating group** *and regressive for the* **dominated group** *reduces dissimilarity.*

$$\mathbf{A}' = \begin{pmatrix} 0 & 0.25 & 0 & 0.75 \\ 0.25 & 0 & 0.35 & 0.40 \\ X & X & X & X \end{pmatrix} \quad \text{with} \quad \overrightarrow{\mathbf{A}'} = \begin{pmatrix} 0 & 0.25 & 0.25 & 1 \\ 0.25 & 0.25 & 0.6 & 1 \\ X & X & X & X \end{pmatrix}$$

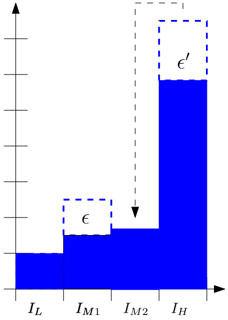


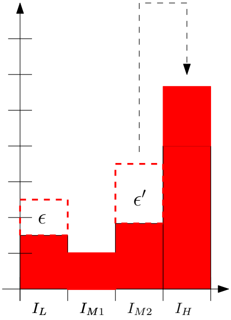
Dissimilarity reducing transfers for ordered classes

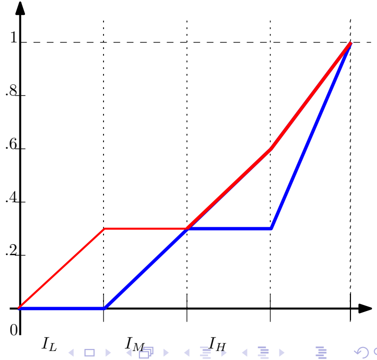
Axiom (The exchange operation (Tchen AP1980, Van de gaer et al Ecmica2001):)

Every transfer of population masses (ϵ) across adjacent classes that is progressive for the **dominating group** *and regressive for the* **dominated group** *reduces dissimilarity.*

$$\mathbf{A}' = \begin{pmatrix} 0 & 0.25 & 0 & 0.75 \\ 0.25 & 0 & 0.35 & 0.40 \\ X & X & X & X \end{pmatrix} \quad \text{with} \quad \overrightarrow{\mathbf{A}'} = \begin{pmatrix} 0 & 0.25 & 0.25 & 1 \\ 0.25 & 0.25 & 0.6 & 1 \\ X & X & X & X \end{pmatrix}$$





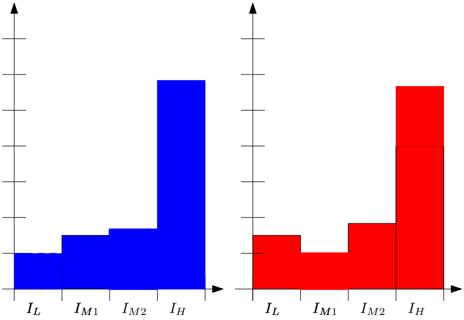


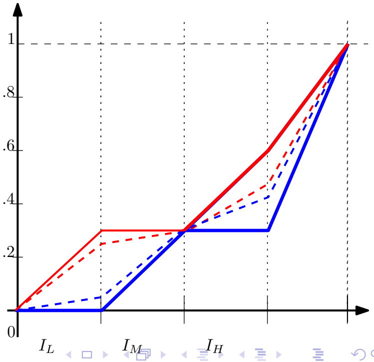
Dissimilarity reducing transfers for ordered classes

Axiom (The exchange operation (Tchen AP1980, Van de gaer et al Ecmica2001):)

Every transfer of population masses (ϵ) across adjacent classes that is progressive for the **dominating group** *and regressive for the* **dominated group** *reduces dissimilarity.*

$$\mathbf{A}' = \begin{pmatrix} 0 & 0.25 & 0 & 0.75 \\ 0.25 & 0 & 0.35 & 0.40 \\ X & X & X & X \end{pmatrix} \quad \text{with} \quad \overrightarrow{\mathbf{A}'} = \begin{pmatrix} 0 & 0.25 & 0.25 & 1 \\ 0.25 & 0.25 & 0.6 & 1 \\ X & X & X & X \end{pmatrix}$$





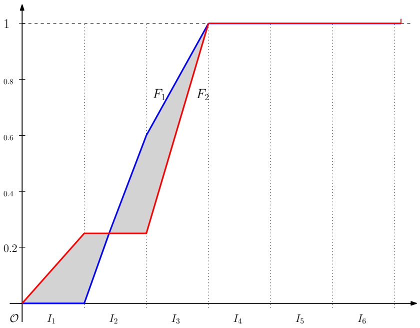
Exchange operations and the Path Polytope

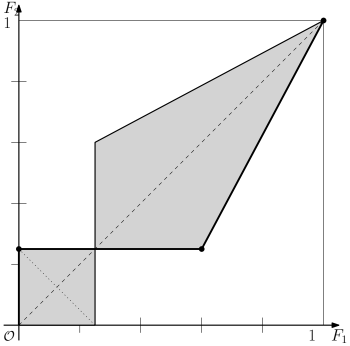
- The Path Pilytope inclusion order detects the existence of exchanges, as below:

  (i) consider the $p\% \in [0\%, 100\%]$ of the overall population

  (ii) construct the **unique** configuration of groups covering the first $p\%$ of the overall population, resulting from the **unique sequence** of classes

  (iii) The configuration in **A** *Lorenz dominates* the configuration in **B** for every $p\%$.





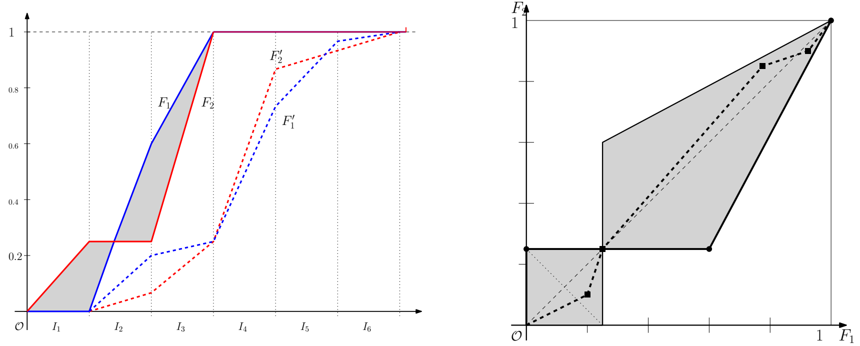
Exchange operations and the Path Polytope

- The Path Pilytope inclusion order detects the existence of exchanges, as below:

  (i) consider the $p\% \in [0\%, 100\%]$ of the overall population

  (ii) construct the **unique** configuration of groups covering the first $p\%$ of the overall population, resulting from the **unique sequence** of classes

  (iii) The configuration in **A** *Lorenz dominates* the configuration in **B** for every $p\%$.



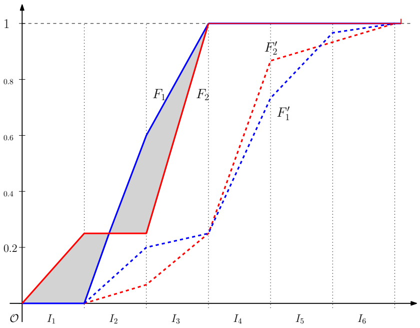
Exchange operations and the Path Polytope

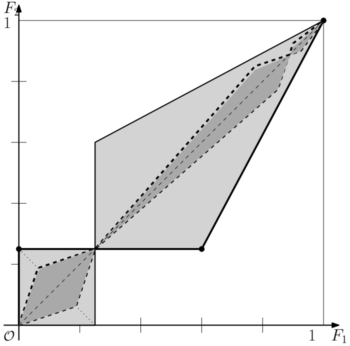
- The Path Pilytope inclusion order detects the existence of exchanges, as below:

  (i) consider the $p\% \in [0\%, 100\%]$ of the overall population

  (ii) construct the **unique** configuration of groups covering the first $p\%$ of the overall population, resulting from the **unique sequence** of classes

  (iii) The configuration in **A** *Lorenz dominates* the configuration in **B** for every $p\%$.





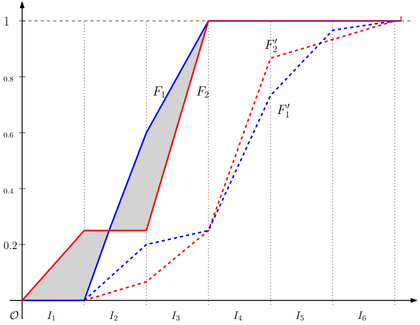
Exchange operations and the Path Polytope

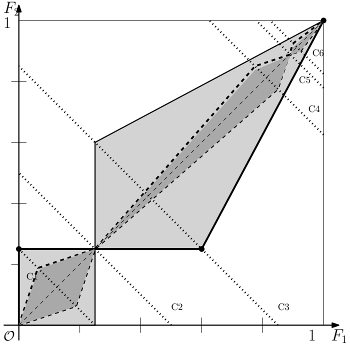
- The Path Pilytope inclusion order detects the existence of exchanges, as below:

  (i) consider the $p\% \in [0\%, 100\%]$ of the overall population

  (ii) construct the **unique** configuration of groups covering the first $p\%$ of the overall population, resulting from the **unique sequence** of classes

  (iii) The configuration in **A** *Lorenz dominates* the configuration in **B** for every $p\%$.





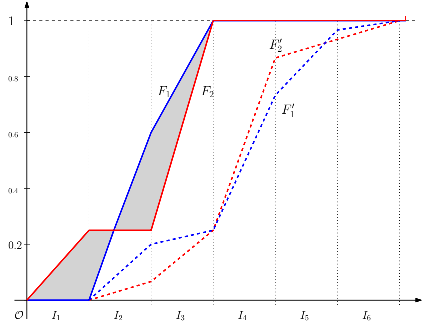
Exchange operations and the Path Polytope

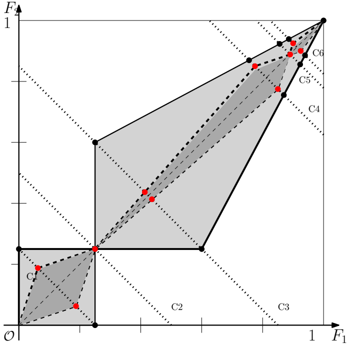
- The Path Pilytope inclusion order detects the existence of exchanges, as below:

  (i) consider the $p\% \in [0\%, 100\%]$ of the overall population

  (ii) construct the **unique** configuration of groups covering the first $p\%$ of the overall population, resulting from the **unique sequence** of classes

  (iii) The configuration in **A** *Lorenz dominates* the configuration in **B** for every $p\%$.





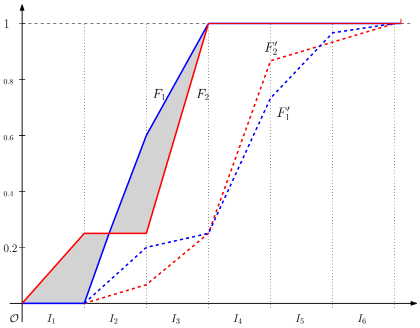
Exchange operations and the Path Polytope

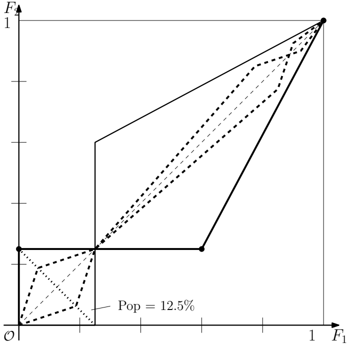
- The Path Pilytope inclusion order detects the existence of exchanges, as below:

  (i) consider the $p\% \in [0\%, 100\%]$ of the overall population

  (ii) construct the **unique** configuration of groups covering the first $p\%$ of the overall population, resulting from the **unique sequence** of classes

  (iii) The configuration in **A** *Lorenz dominates* the configuration in **B** for every $p\%$.





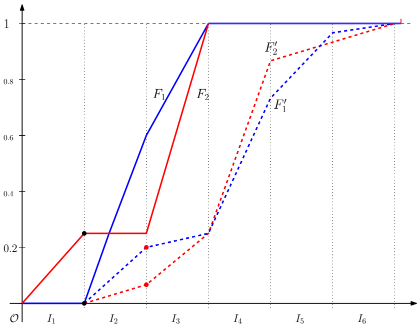
Exchange operations and the Path Polytope

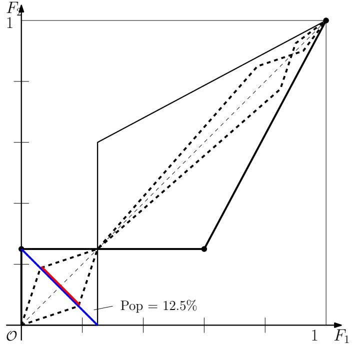
- The Path Pilytope inclusion order detects the existence of exchanges, as below:

  (i) consider the $p\% \in [0\%, 100\%]$ of the overall population

  (ii) construct the **unique** configuration of groups covering the first $p\%$ of the overall population, resulting from the **unique sequence** of classes

  (iii) The configuration in **A** *Lorenz dominates* the configuration in **B** for every $p\%$.





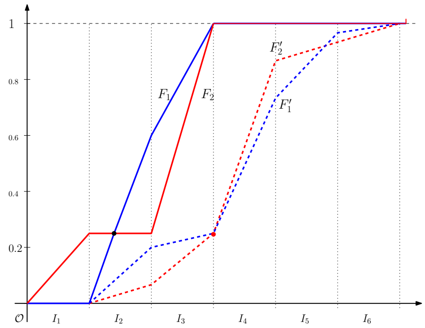
Exchange operations and the Path Polytope

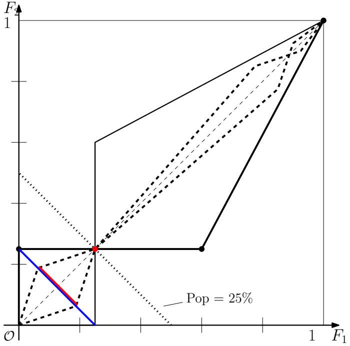
- The Path Pilytope inclusion order detects the existence of exchanges, as below:

  (i) consider the $p\% \in [0\%, 100\%]$ of the overall population

  (ii) construct the **unique** configuration of groups covering the first $p\%$ of the overall population, resulting from the **unique sequence** of classes

  (iii) The configuration in **A** *Lorenz dominates* the configuration in **B** for every $p\%$.





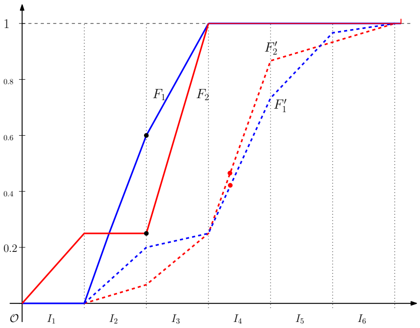
Exchange operations and the Path Polytope

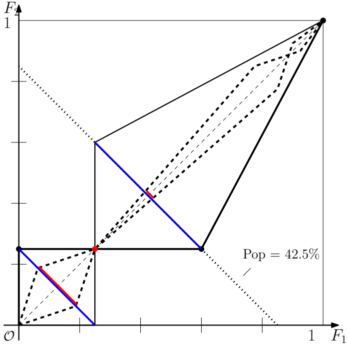
- The Path Pilytope inclusion order detects the existence of exchanges, as below:

  (i) consider the $p\% \in [0\%, 100\%]$ of the overall population

  (ii) construct the **unique** configuration of groups covering the first $p\%$ of the overall population, resulting from the **unique sequence** of classes

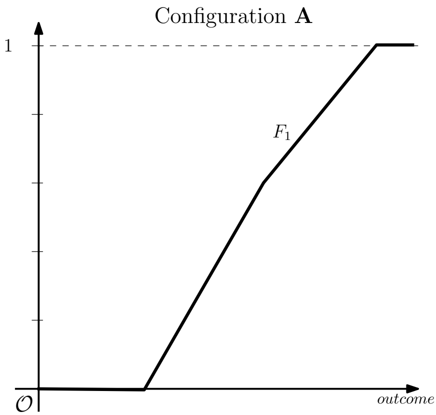
  (iii) The configuration in **A** *Lorenz dominates* the configuration in **B** for every $p\%$.

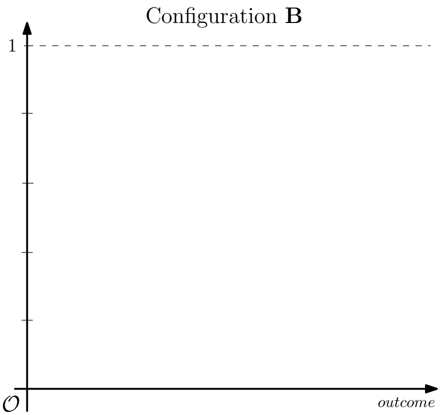




Discrimination comparisons with many distributions.

- Have do do with reductions in dispersion across distributions at every population proportion.

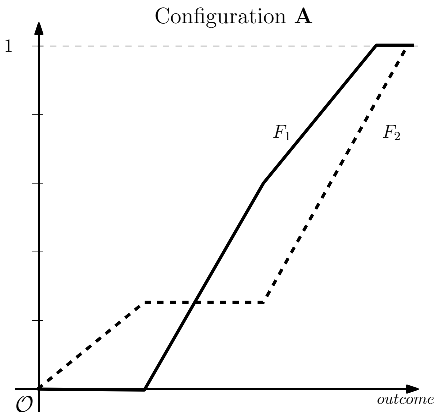


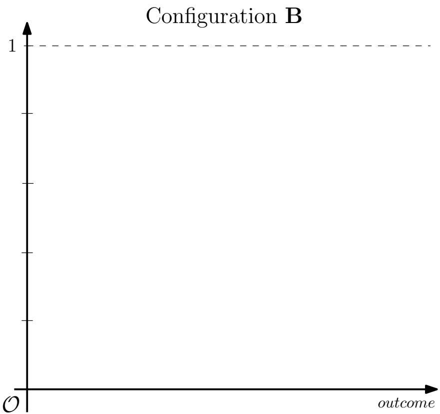


- **A** defined on classes c_1, c_2, c_3, . . .
- **B** defined on classes c_1, c_4, c_6, c_7, c_3, . . .
- **A*** and **B*** defined on all classes, after interchanging groups.

Discrimination comparisons with many distributions.

- Have do do with reductions in dispersion across distributions at every population proportion.

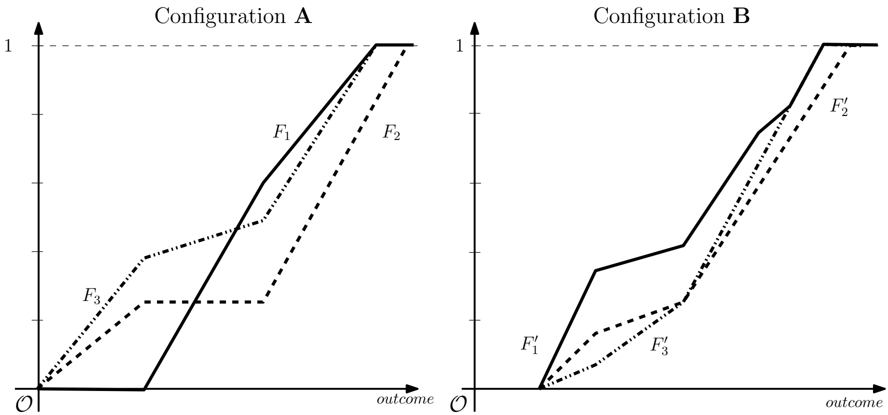




- **A** defined on classes c_1, c_2, c_3, . . .
- **B** defined on classes c_1, c_4, c_6, c_7, c_3, . . .
- **A*** and **B*** defined on all classes, after interchanging groups.

Discrimination comparisons with many distributions.

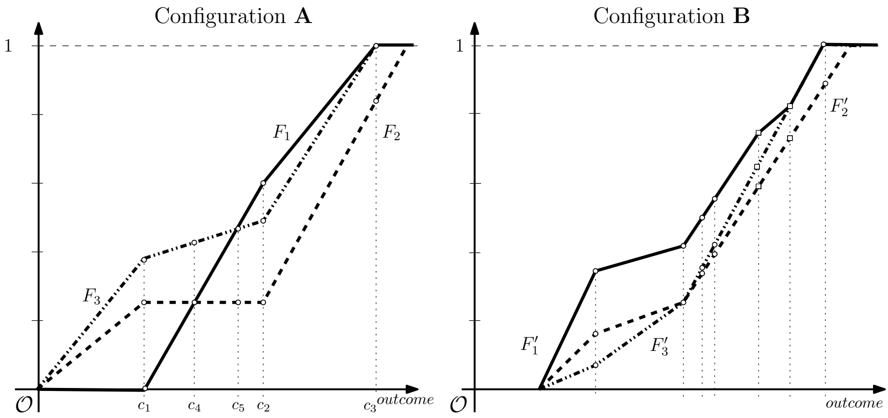
- Have do do with reductions in dispersion across distributions at every population proportion.



- **A** defined on classes c_1, c_2, c_3, . . .
- **B** defined on classes c_1, c_4, c_6, c_7, c_3, . . .
- **A*** and **B*** defined on all classes, after interchanging groups.

Discrimination comparisons with many distributions.

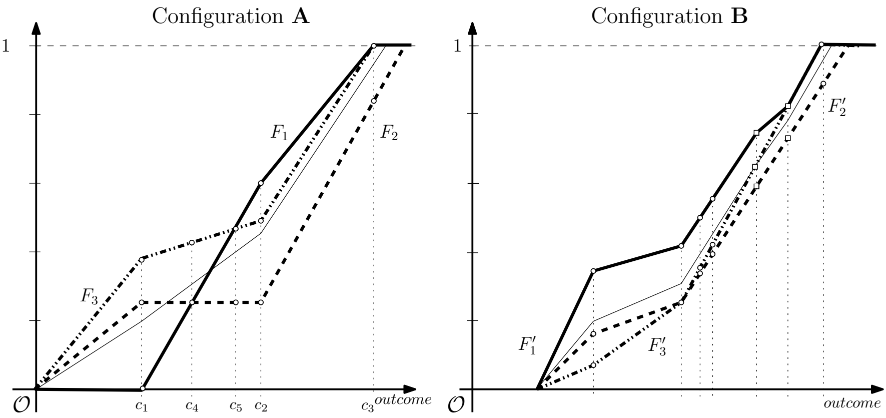
- Have do do with reductions in dispersion across distributions at every population proportion.



- **A** defined on classes c_1, c_2, c_3, ...
- **B** defined on classes c_1, c_4, c_6, c_7, c_3, ...
- **A*** and **B*** defined on all classes, after interchanging groups.

Discrimination comparisons with many distributions.

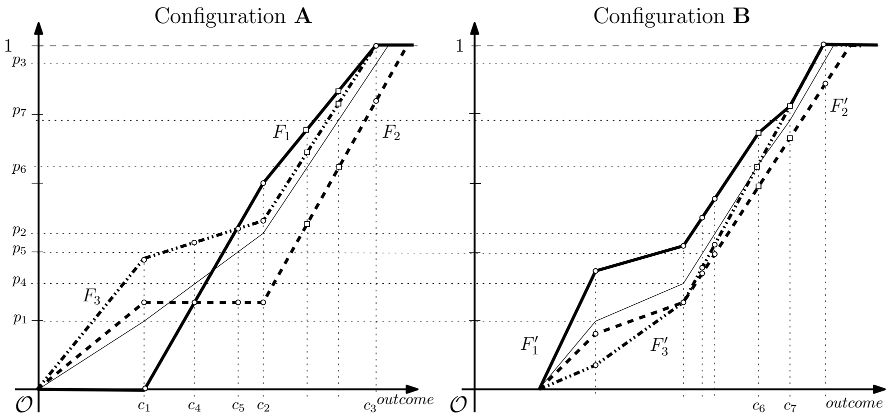
- Have do do with reductions in dispersion across distributions at every population proportion.



- **A** defined on classes $c_1, c_2, c_3, \ldots$
- **B** defined on classes $c_1, c_4, c_6, c_7, c_3, \ldots$
- $\mathbf{A}^*$ and $\mathbf{B}^*$ defined on all classes, after interchanging groups.

Discrimination comparisons with many distributions.

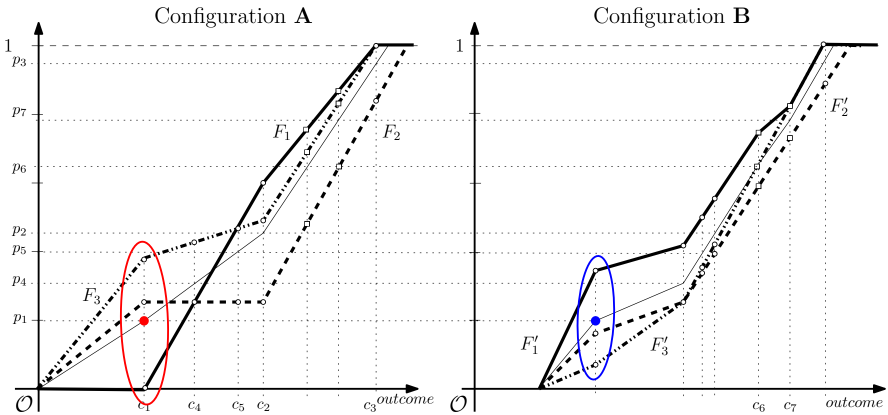
- Have do do with reductions in dispersion across distributions at every population proportion.



- **A** defined on classes $c_1, c_2, c_3, \dots$
- **B** defined on classes $c_1, c_4, c_6, c_7, c_3, \dots$
- **A*** and **B*** defined on all classes, after interchanging groups.

Discrimination comparisons with many distributions.

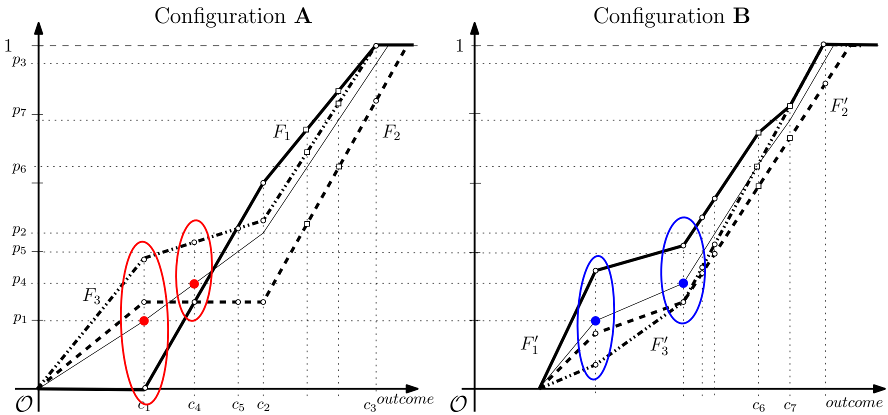
- Have do do with reductions in dispersion across distributions at every population proportion.



- **A** defined on classes c_1, c_2, c_3, ...
- **B** defined on classes c_1, c_4, c_6, c_7, c_3, ...
- **A*** and **B*** defined on all classes, after interchanging groups.

Discrimination comparisons with many distributions.

- Have do do with reductions in dispersion across distributions at every population proportion.



- **A** defined on classes c_1, c_2, c_3, . . .
- **B** defined on classes c_1, c_4, c_6, c_7, c_3, . . .
- **A*** and **B*** defined on all classes, after interchanging groups.

Path Polytopes in the literature

- The **concentration curves** for $d = 2$ (Butler McDonald JBES1987) correspond to the arrangement of the ordered segments (corresponding to exogenously ordered classes)

- When one of the two groups (say 2) *stochastic dominates* the other (say 1), i.e.

$$\sum_{j=1}^{k} a_{1j} \geq \sum_{j=1}^{k} a_{2j} \quad \forall k = 1, \ldots, n$$

then the concentration curve delimits a **discrimination curve** (LeBreton at al JET2011)

- When group 2 coincides in **A** and **B** and group 2 sotchastic dominates group 1, dominance in discrimination curves can be related to dominance for all **Gastwirth measures of discrimination.**

- The **test of the orthant** (Dardanoni JET1993, Tchen AP1980) is a robust tests for assessing changes in **mobility**.
  - Perfect mobility is achieved when the group of departure (income of the father) is not informative of the class of destination (income of the children) (see Stiglitz 2012).
  - The test applies to *monotone matrices* (groups are ordered by stochastic dominance).
  - The test has normative content when mobility matrices have *fixed margins* (exchange mobility).
  - In this specific context, the test is implemented by the sequential Lorenz comparisons defined above.

Path Polytopes in the literature

- The **concentration curves** for $d = 2$ (Butler McDonald JBES1987) correspond to the arrangement of the ordered segments (corresponding to exogenously ordered classes)
- When one of the two groups (say 2) *stochastic dominates* the other (say 1), i.e.

$$\sum_{j=1}^{k} a_{1j} \geq \sum_{j=1}^{k} a_{2j} \quad \forall k = 1, \ldots, n$$

then the concentration curve delimits a **discrimination curve** (LeBreton at al JET2011)

- When group 2 coincides in **A** and **B** and group 2 sotchastic dominates group 1, dominance in discrimination curves can be related to dominance for all **Gastwirth measures of discrimination.**
- The **test of the orthant** (Dardanoni JET1993, Tchen AP1980) is a robust tests for assessing changes in **mobility**.
  - Perfect mobility is achieved when the group of departure (income of the father) is not informative of the class of destination (income of the children) (see Stiglitz 2012).
  - The test applies to *monotone matrices* (groups are ordered by stochastic dominance).
  - The test has normative content when mobility matrices have *fixed margins* (exchange mobility).
  - In this specific context, the test is implemented by the sequential Lorenz comparisons defined above.

Path Polytopes in the literature

- The **concentration curves** for $d = 2$ (Butler McDonald JBES1987) correspond to the arrangement of the ordered segments (corresponding to exogenously ordered classes)
- When one of the two groups (say 2) *stochastic dominates* the other (say 1), i.e.

$$\sum_{j=1}^{k} a_{1j} \geq \sum_{j=1}^{k} a_{2j} \quad \forall k = 1, \ldots, n$$

  then the concentration curve delimits a **discrimination curve** (LeBreton at al JET2011)
- When group 2 coincides in **A** and **B** and group 2 sotchastic dominates group 1, dominance in discrimination curves can be related to dominance for all **Gastwirth measures of discrimination.**
- The **test of the orthant** (Dardanoni JET1993, Tchen AP1980) is a robust tests for assessing changes in **mobility**.
  - Perfect mobility is achieved when the group of departure (income of the father) is not informative of the class of destination (income of the children) (see Stiglitz 2012).
  - The test applies to *monotone matrices* (groups are ordered by stochastic dominance).
  - The test has normative content when mobility matrices have *fixed margins* (exchange mobility).
  - In this specific context, the test is implemented by the sequential Lorenz comparisons defined above.

Path Polytopes in the literature

- The **concentration curves** for $d = 2$ (Butler McDonald JBES1987) correspond to the arrangement of the ordered segments (corresponding to exogenously ordered classes)
- When one of the two groups (say 2) *stochastic dominates* the other (say 1), i.e.

$$\sum_{j=1}^{k} a_{1j} \geq \sum_{j=1}^{k} a_{2j} \quad \forall k = 1, \ldots, n$$

  then the concentration curve delimits a **discrimination curve** (LeBreton at al JET2011)
- When group 2 coincides in **A** and **B** and group 2 sotchastic dominates group 1, dominance in discrimination curves can be related to dominance for all **Gastwirth measures of discrimination.**
- The **test of the orthant** (Dardanoni JET1993, Tchen AP1980) is a robust tests for assessing changes in **mobility**.
  - Perfect mobility is achieved when the group of departure (income of the father) is not informative of the class of destination (income of the children) (see Stiglitz 2012).
  - The test applies to *monotone matrices* (groups are ordered by stochastic dominance).
  - The test has normative content when mobility matrices have *fixed margins* (exchange mobility).
  - In this specific context, the test is implemented by the sequential Lorenz comparisons defined above.

Path Polytopes in the literature

- The **concentration curves** for $d = 2$ (Butler McDonald JBES1987) correspond to the arrangement of the ordered segments (corresponding to exogenously ordered classes)
- When one of the two groups (say 2) *stochastic dominates* the other (say 1), i.e.

$$\sum_{j=1}^{k} a_{1j} \geq \sum_{j=1}^{k} a_{2j} \quad \forall k = 1, \ldots, n$$

  then the concentration curve delimits a **discrimination curve** (LeBreton at al JET2011)
- When group 2 coincides in **A** and **B** and group 2 sotchastic dominates group 1, dominance in discrimination curves can be related to dominance for all **Gastwirth measures of discrimination.**
- The **test of the orthant** (Dardanoni JET1993, Tchen AP1980) is a robust tests for assessing changes in **mobility**.
  - Perfect mobility is achieved when the group of departure (income of the father) is not informative of the class of destination (income of the children) (see Stiglitz 2012).
  - The test applies to *monotone matrices* (groups are ordered by stochastic dominance).
  - The test has normative content when mobility matrices have *fixed margins* (exchange mobility).
  - In this specific context, the test is implemented by the sequential Lorenz comparisons defined above.

Path Polytopes in the literature

- The **concentration curves** for $d = 2$ (Butler McDonald JBES1987) correspond to the arrangement of the ordered segments (corresponding to exogenously ordered classes)
- When one of the two groups (say 2) *stochastic dominates* the other (say 1), i.e.

$$\sum_{j=1}^{k} a_{1j} \geq \sum_{j=1}^{k} a_{2j} \quad \forall k = 1, \ldots, n$$

  then the concentration curve delimits a **discrimination curve** (LeBreton at al JET2011)
- When group 2 coincides in **A** and **B** and group 2 sotchastic dominates group 1, dominance in discrimination curves can be related to dominance for all **Gastwirth measures of discrimination.**
- The **test of the orthant** (Dardanoni JET1993, Tchen AP1980) is a robust tests for assessing changes in **mobility**.
  - Perfect mobility is achieved when the group of departure (income of the father) is not informative of the class of destination (income of the children) (see Stiglitz 2012).
  - The test applies to *monotone matrices* (groups are ordered by stochastic dominance).
  - The test has normative content when mobility matrices have *fixed margins* (exchange mobility).
  - In this specific context, the test is implemented by the sequential Lorenz comparisons defined above.

Indices coherent with Zonotope inclusion

- Define the following family of dissimilarity indicators

$$D_h(\mathbf{A}) := \frac{1}{d}\sum_{j=1}^{n_A} \overline{a}_j \cdot h\left(a_{1j}/\overline{a}_j, \ldots, a_{dj}/\overline{a}_j\right).$$

with h convex. It is a model for segregation indices.

  - Dissimilarity index (Duncan Duncan ASR1955)

$$D(\mathbf{A}) := \frac{1}{2}\sum_{j=1}^{n_A} \left|a_{1j} - a_{2j}\right|$$

  - Atkinson and Mutual information indices (Frankel Volji JET2010):

$$A_\omega(\mathbf{A}) := 1 - \sum_{j=1}^{n_A} \prod_{i=1}^{d} \left(a_{ij}\right)^{\omega_i}$$

$$M(\mathbf{A}) := \log_2(d) - \sum_{j=1}^{n_A} \left(\frac{\overline{a}_j}{d}\right) \sum_{i=1}^{d} \frac{a_{ij}}{\overline{a}_j} \cdot \log_2\left(\frac{\overline{a}_j}{a_{ij}}\right)$$

Indices coherent with Zonotope inclusion

- Define the following family of dissimilarity indicators

$$D_h(\mathbf{A}) := \frac{1}{d} \sum_{j=1}^{n_A} \overline{a}_j \cdot h\left(a_{1j}/\overline{a}_j, \dots, a_{dj}/\overline{a}_j\right).$$

with h convex. It is a model for segregation indices.

  - Dissimilarity index (Duncan Duncan ASR1955)

$$D(\mathbf{A}) := \frac{1}{2} \sum_{j=1}^{n_A} \left| a_{1j} - a_{2j} \right|$$

  - Atkinson and Mutual information indices (Frankel Volji JET2010):

$$A_\omega(\mathbf{A}) := 1 - \sum_{j=1}^{n_A} \prod_{i=1}^{d} \left(a_{ij}\right)^{\omega_i}$$

$$M(\mathbf{A}) := \log_2(d) - \sum_{j=1}^{n_A} \left(\frac{\overline{a}_j}{d}\right) \sum_{i=1}^{d} \frac{a_{ij}}{\overline{a}_j} \cdot \log_2\left(\frac{\overline{a}_j}{a_{ij}}\right)$$

Indices coherent with Zonotope inclusion

- Define the following family of dissimilarity indicators

$$D_h(\mathbf{A}) := \frac{1}{d} \sum_{j=1}^{n_A} \overline{a}_j \cdot h\left(a_{1j}/\overline{a}_j, \ldots, a_{dj}/\overline{a}_j\right).$$

with h convex. It is a model for segregation indices.

  - Dissimilarity index (Duncan Duncan ASR1955)

$$D(\mathbf{A}) := \frac{1}{2} \sum_{j=1}^{n_A} \left| a_{1j} - a_{2j} \right|$$

  - Atkinson and Mutual information indices (Frankel Volji JET2010):

$$A_\omega(\mathbf{A}) := 1 - \sum_{j=1}^{n_A} \prod_{i=1}^{d} \left(a_{ij}\right)^{\omega_i}$$

$$M(\mathbf{A}) := \log_2(d) - \sum_{j=1}^{n_A} \left(\frac{\overline{a}_j}{d}\right) \sum_{i=1}^{d} \frac{a_{ij}}{\overline{a}_j} \cdot \log_2\left(\frac{\overline{a}_j}{a_{ij}}\right)$$

Indices coherent with Zonotope inclusion

- The **Gini inequality index** is "half the area between the diagonal and the Lorenz curve", i.e. the area of a Zonotope when $d = 2$.
- The **Lorenz Zonotope** (Koshevoy Mosler JASA1996) extend univariate inequality analysis to the multidimensional level.
- It is a Zonotope in the $d + 1$ space: d attributes distributions and 1 demographic weights distribution.
- The Lorenz Zonotope volume (like any Zonotope volume) defines the Multivariate Gini Index.

Indices coherent with Zonotope inclusion

- The **Gini inequality index** is "half the area between the diagonal and the Lorenz curve", i.e. the area of a Zonotope when $d = 2$.
- The **Lorenz Zonotope** (Koshevoy Mosler JASA1996) extend univariate inequality analysis to the multidimensional level.
- It is a Zonotope in the $d + 1$ space: d attributes distributions and 1 demographic weights distribution.
- The Lorenz Zonotope volume (like any Zonotope volume) defines the Multivariate Gini Index.

Indices coherent with Zonotope inclusion

- The **Gini inequality index** is "half the area between the diagonal and the Lorenz curve", i.e. the area of a Zonotope when $d = 2$.
- The **Lorenz Zonotope** (Koshevoy Mosler JASA1996) extend univariate inequality analysis to the multidimensional level.
- It is a Zonotope in the $d + 1$ space: d attributes distributions and 1 demographic weights distribution.
- The Lorenz Zonotope volume (like any Zonotope volume) defines the Multivariate Gini Index.

Indices coherent with Zonotope inclusion

- The **Gini inequality index** is "half the area between the diagonal and the Lorenz curve", i.e. the area of a Zonotope when $d = 2$.
- The **Lorenz Zonotope** (Koshevoy Mosler JASA1996) extend univariate inequality analysis to the multidimensional level.
- It is a Zonotope in the $d + 1$ space: d attributes distributions and 1 demographic weights distribution.
- The Lorenz Zonotope volume (like any Zonotope volume) defines the Multivariate Gini Index.