Introduction of weights in poset-based synthetic indicators

Giulio Caperna

Joint Research Centre of the European Commission[1]

26-27 October, 2018

Outline

1. Introduction
   - Finding a Synthetic Measure
   - Poset Theory
2. Tools from Poset theory
   - Approximation of Average rank
3. Our Proposal
   - Weights in posets
4. Case Study - Disability
   - Concept and Data
   - Results

European Commission

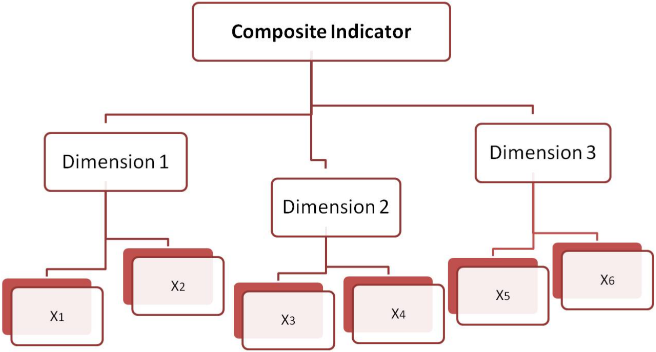
Outline

1. Introduction
   - Finding a Synthetic Measure
   - Poset Theory
2. Tools from Poset theory
   - Approximation of Average rank
3. Our Proposal
   - Weights in posets
4. Case Study - Disability
   - Concept and Data
   - Results

European Commission

A Composite Indicator

Quality of life, Environment, Gender Equality, Human Development . . .

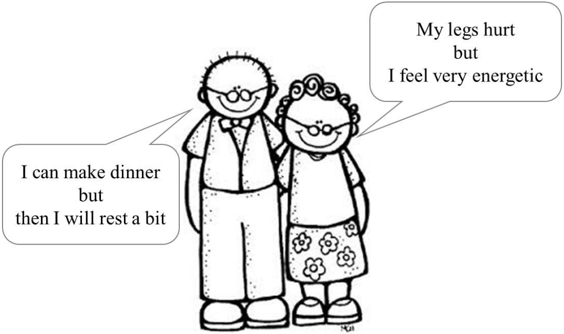


Quantitative elementary indicators aggregated through different methods to describe a complex and not observable concept

European Commission

Ordinal Data

The measure of complex and unobservable concepts is the main reason of this work



We focus on micro level data, often measured on **ordinal** scale.

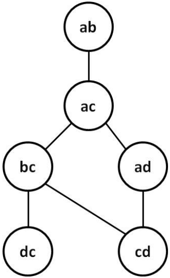
In social studies it implies thousands of observations.

European Commission

Representation - Hasse Diagram

As usual, we represent the elements of the poset with the Hasse diagram

Child	His.	Mat.	Profile
c_1	a	c	ac
c_2	c	d	cd
c_3	d	c	dc
c_4	a	d	ad
c_5	b	c	bc
c_6	a	b	ab

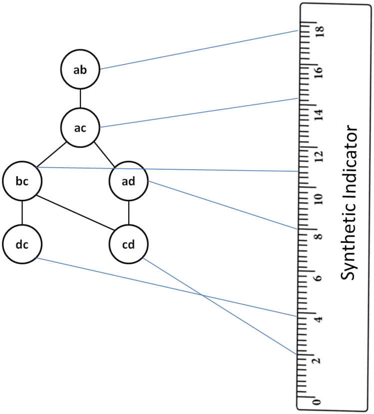


Six children on Maths and History

European Commission

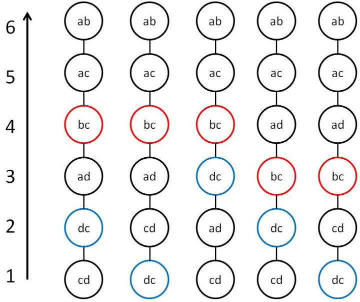
Our Aim

Starting from the order information, we look for a synthetic measure of the profiles position in the poset:



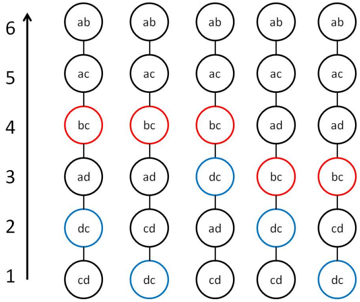
Exact Average rank

Observing the position of a profile among the linear extensions



$h(\cdot)$	ω_1	ω_2	ω_3	ω_4	ω_5	$\bar{h}(\cdot)$
ab	6	6	6	6	6	6
ac	5	5	5	5	5	5
bc	4	4	4	3	3	3.6
ad	3	3	2	4	4	3.2
dc	2	1	3	2	1	1.8
cd	1	2	1	1	2	1.4

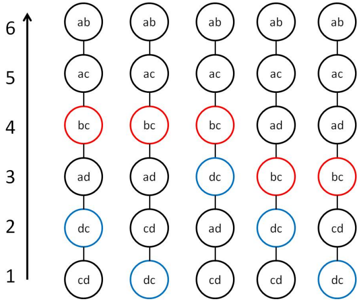
European Commission



Compute the mutual rank probability of profiles $P(x \geq y)$

The average rank is
$\bar{h}(x) = \sum_y P(x \geq y)$

European Commission



Compute the mutual rank probability of profiles $P(x \geq y)$

The average rank is $\overline{h}(x) = \sum_y P(x \geq y)$

	ab	ac	ad	bc	cd	dc	Σ	$\overline{h}(\cdot)$
ab	1	1	1	1	1	1	6	6
ac	0	1	1	1	1	1	5	5
bc	0	0	0.6	1	1	1	3.6	3.6
ad	0	0	1	0.4	1	0.8	3.2	3.2
dc	0	0	0.2	0	0.6	1	1.8	1.8
cd	0	0	0	0	1	0.4	1.4	1.4

European Commission

Computational Issues

It is not possible to define the time needed to compute the number of linear extensions in a deterministic way,

¨The Extensions are too many!¨

In order to handle with this issue, practitioners adopt two different approaches:

Sampling of linear extensions;

Approximation of average rank.

European Commission

Outline

European Commission

Approximation of Average Rank

Aim to compute an approximated averaged rank $\overline{h}_a(\cdot)$ avoiding the observation of every linear extension

Definitions:

Given a Poset P, and $x \in P$,

Down set of x: $O(x) = \{y \in P : y \leq x\}$

Up set of x: $F(x) = \{y \in P : y \geq x\}$

Incomparable to x: $U(x) = \{y \in P : y||x\}$

Brüggeman and Carlsen (2011), proposed two methods based on these sets.

European Commission

Formula

The average rank is computed as

$$\bar{h}(x) = \sum_{y} P(x \geq y)$$

$$\bar{h}_a(x) = |O(x)| + \sum_{y \in U(x)} \hat{P}(x \geq y) = |O(x)| + \sum_{y \in U(x)} \eta(x, y)$$

The probability $\hat{P}(x \geq y)$ is the required information.
$\eta(x, y)$ is the quantity proposed in literature to approximate it.

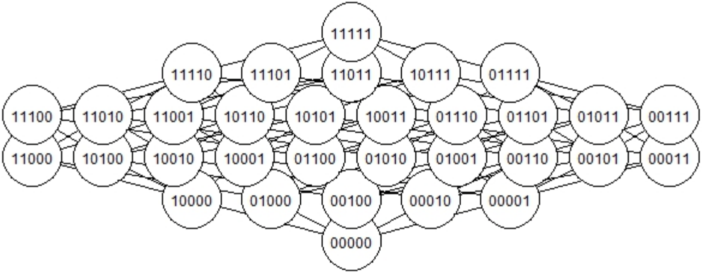
European Commission

Outline

European Commission

Simulated Population

The poset observing 5 dichotomous variables and no weights.



There is no difference among the nodes on the same level,
If weights are not considered, it implies *Equal Weights*.

European Commission

Position vs Evaluation

We are not only interested in the position of an individual respect to the others.

We look for an **evaluation** of its condition.
Our proposal is the introduction of weights ($\mathbf{w}$).

European Commission

Position vs Evaluation

We are not only interested in the position of an individual respect to the others.

We look for an **evaluation** of its condition.
Our proposal is the introduction of weights ($\mathbf{w}$).

$$\psi(x,y) = \sum_{i=1}^{g} [w_i \cdot I(x_i > y_i) + \frac{1}{2} w_i \cdot I(x_i = y_i)]$$

where g is the number of observed variables

European Commission

Approximation with weights

We propose to change the formula of approximation

$$\overline{h}_a(x) = |O(x)| + \sum_{y \in U(x)} \eta(x, y)$$

Including the weights of variables

$$\overline{h}_w(x, s) = |O(x)| + \sum_{y \in U(x)} \eta(x, y)^s \cdot \psi(x, y)^{1-s}$$

Approximation with weights

We propose to change the formula of approximation

$$\overline{h}_a(x) = |O(x)| + \sum_{y \in U(x)} \eta(x, y)$$

Including the weights of variables

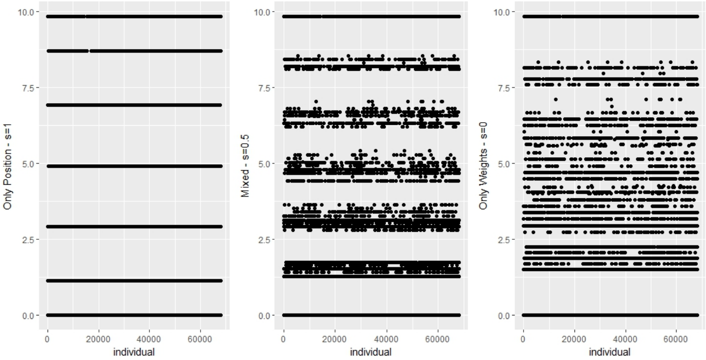
$$\overline{h}_w(x, s) = |O(x)| + \sum_{y \in U(x)} \eta(x, y)^s \cdot \psi(x, y)^{1-s}$$

for $s = 0.5$, we have the geometric mean of η and ψ.

European Commission

Comparison in Simulated Population

Compare $\overline{h}_a(\cdot)$, $\overline{h}_w(\cdot, 0.5)$ and $\overline{h}_w(\cdot, 0.1)$
$\mathbf{w} = \{0.05, 0.1, 0.15, 0.20, 0.25, 0.25\}$



The comparison shows relevant differences, mainly in the variability

European Commission

Outline

European Commission

Definition of Disability - Katz, 1963

Disability: inability to perform the activities of daily living (ADLs)
Measure of Disability: at least one ADL is impossible for the individual

Definition of Disability - Katz, 1963

Disability: inability to perform the activities of daily living (ADLs)
Measure of Disability: at least one ADL is impossible for the individual

ADL 1: Bathing and Showering;

ADL 2: Dressing;

ADL 3: Use of toilette;

ADL 4: Mobility;

ADL 5: Personal Care;

ADL 6: Self-Feeding;

Does every missing ADL represent the same severity of disability?

Severity of Disability

A measure of the Severity of Disability is not available.
Often the number of ADLs is considered, but this implies the same severity for every ADL.

AIM: to build a composite indicator that measures the severity of disability, keeping in mind the differences among ADLs.

European Commission

Big survey data

Data from the *Survey of Health, Ageing and Retirement in Europe* (SHARE)

More than 60 thousands of observed units in 2015 (Wave 6)

Big survey data

Data from the *Survey of Health, Ageing and Retirement in Europe* (SHARE)

More than 60 thousands of observed units in 2015 (Wave 6)

The variables of interest are 6 items on *Activities of Daily Living*

Measured on dichotomous levels:
no - no serious limitation or **yes** - seriously limited

Weights of the ADL

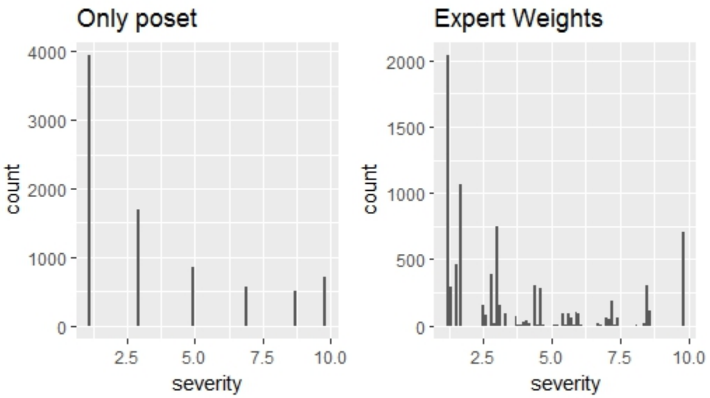
The weights we used

ADL	Wash	Dress	Use Toilette	Mobility	Personal Care	Self-Feed
Weights	0,05	0,05	0,23	0,10	0,03	0,54

Elicited from expert's opinion with the AHP approach.

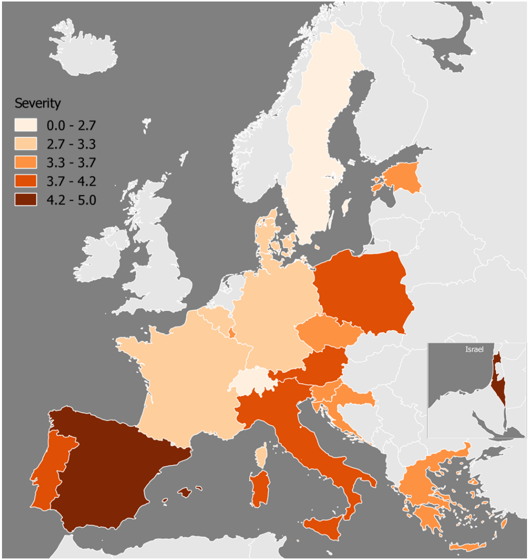
European Commission

Results



European Commission

Disability in Europe



The distribution of Disability Severity is uneven across Europe, with higher values in the eastern and southern countries

European Commission

Conclusions

- This method fits to dichotomous data;
- The weights can be simply integrated and tuned;
- The results are realistic with easy use and interpretation;
- Future steps:

European Commission

Conclusions

- This method fits to dichotomous data;
- The weights can be simply integrated and tuned;
- The results are realistic with easy use and interpretation;
- Future steps:
  1. Extend the group of experts, to underline different perspectives;
  2. Perform sensitivity analysis on the effect of the *s*-value.

European Commission

References

Brüggemann, R., and **Carlsen**, L. (2011). An improved estimation of averaged ranks of partial orders. *MATCH Commun. Math. Comput. Chem*, 65, 383-414.

Caperna, G.(2016). Partial order theory for synthetic indicators. *PhD thesis*, Padua.

Davey B. A., **Priestley** H. A. (2002), *Introduction to lattices and order*, CUP, New York

Katz S., **Ford** A.B., **Moskowitz** R.W.,**Jackson** B.A., **Jaffe** M.W. (1963), *Studies of illness in the aged: the index of ADL: a standardized measure of biological and psychosocial function*, Jama, 185(12),914-919

De Loof K., **De Baets** B., **De Meyer** H. (2011), *Approximation of average ranks in posets*, MATCH - Commun. Math. Compu. Chem., 66,219-229

Maggino F., **Fattore** M. (2011), *New tools for the construction of ranking and evaluation indicators in multidimensional systems of ordinal variables*, Proceedings of the 'New Techniques and Technologies for Statistics', Brussels

To contact us: giulio.caperna@ec.europa.eu

European Commission