

dist.Symbolic(clusterSim)

### Ichino and Yaguchi dissimilarity measure for variables

$$\phi(A_j, B_j) = |A_j \oplus B_j| - |A_j \otimes B_j| + \gamma(2 \cdot |A_j \otimes B_j| - |A_j| - |B_j|),$$

where:  $\phi(A_j, B_j)$  – dissimilarity measure of symbolic variables,

$A_j, B_j$  – symbolic variables of any type,

$\oplus$  – Cartesian join,

$\otimes$  – Cartesian meet,

$| |$  – means the length of an interval for continuous data or the number of elements of an set of values,

$\gamma$  – parameter  $\langle 0, 1/2 \rangle$ .

### Ichino and Yaguchi measure for objects (U\_2)

$$d_q(a_1, b_1) = \left( \sqrt[q]{\sum_{j=1}^p \phi(A_j, B_j)^q} \right),$$

where:  $d_q(a_1, b_1)$  – Ichino and Yaguchi measure (sometimes called extended Minkowski metric – see Diday [2000]),

$q$  – integer number greater or equal 1,

$a_1 = (A_1, A_2, \dots, A_p)$ ,  $b_1 = (B_1, B_2, \dots, B_p)$  – two symbolic objects containing  $p$  symbolic variables,

$\phi(A_j, B_j)$  – Ichino and Yaguchi dissimilarity measure for variables.

### Hausdorff distance measure (H)

$$\max \left\{ \max_{\alpha \in [A, A]} \left\{ \min_{\beta \in [B, B]} d(\alpha, \beta) \right\}, \max_{\beta \in [B, B]} \left\{ \min_{\alpha \in [A, A]} d(\alpha, \beta) \right\} \right\},$$

where:  $H(a_1, b_1)$  – Hausdorff distance,

$a_1 = (A_1, A_2, \dots, A_p)$ ,  $b_1 = (B_1, B_2, \dots, B_p)$  – two symbolic objects containing  $p$  symbolic interval-valued variables. Each variable  $A_i$  is an interval in a form  $[A_i, \bar{A}_i]$ ,

$d(\dots)$  – Euclidean distance.

### References

Billard, L., Diday, E. (2006), *Symbolic data analysis. Conceptual statistics and data mining*, Wiley, Chichester.

Bock H.H., Diday E. (Eds.) (2000), *Analysis of symbolic data. Explanatory methods for extracting statistical information from complex data*, Springer Verlag, Berlin.