Two-way (Dis)-Similarity Measures

Measure	Formula[a]	Range[b]	Average	Range
Minkowsky (**M1-M7**) p = 0.25, 0.5, 1, 1.5, 2, 2.5, 3, and ∞ [where, when p= 1 it is the Manhattan, city-block or taxi distance (also known as Hamming distance between binary vectors) and p = 2 is Euclidean distance)	$d_{XY} = \left(\sum_{j=1}^{h} \lvert x_j - y_j \rvert^p\right)^{\frac{1}{p}}$	$[0, \infty)$	$\bar{d} = \frac{d_{XY}}{n^{1/p}}$	$[0, \infty)$
Chebyshev/Lagrange (**M8**) (Minkowsky formula when p = ∞)	$d_{XY} = max\{\lvert x_j - y_j \rvert\}$			
Canberra (**M10**)	$d_{XY} = \sum_{j=1}^{h} \frac{\lvert x_j - y_j \rvert}{\lvert x_j \rvert + \lvert y_j \rvert}$	$[0, n]$	$\bar{d} = \frac{d_{XY}}{n}$	$[0,1]$
Lance - Williams/Bray-Curtis (**M11**)	$d_{XY} = \frac{\sum_{j=1}^{h} \lvert x_j - y_j \rvert}{\sum_{j=1}^{h} (\lvert x_j \rvert + \lvert y_j \rvert)}$	$[0,1]$	$\bar{d} = \frac{d_{XY}}{n}$	$\left[0, \frac{1}{n}\right]$
Clark/Coefficient of Divergence (**M12**)	$d_{XY} = \sqrt{\sum_{j=1}^{h} \left(\frac{x_j - y_j}{\lvert x_j \rvert + \lvert y_j \rvert}\right)^2}$	$[0, n]$	$\bar{d} = \frac{d_{XY}}{\sqrt{n}}$	$[0, \sqrt{n}]$
Soergel (**M13**)	$d_{XY} = \frac{1}{n} \sum_{j=1}^{h} \frac{\lvert x_j - y_j \rvert}{max\{x_j, y_j\}}$	$[0,1]$	$\bar{d} = \frac{d_{XY}}{n}$	$\left[0, \frac{1}{n}\right]$
Bhattacharyya (**M14**)	$d_{XY} = \sqrt{\sum_{j=1}^{h} \left(\sqrt{x_j} - \sqrt{y_j}\right)^2}$	$[0, \infty)$	$\bar{d} = \frac{d_{XY}}{\sqrt{n}}$	$[0, \infty)$
Wave – Edges (**M15**)	$d_{XY} = \sum_{j=1}^{h} \left(1 - \frac{min\{x_j, y_j\}}{max\{x_j, y_j\}}\right)$	$[0, n]$	$\bar{d} = \frac{d_{XY}}{n}$	$[0,1]$
Angular Separation/[1-Cosine (**Ochiai**)] (**M16**)	$d_{XY} = 1 - Cos_{XY}$ where, $Cos_{XY} = \frac{\boldsymbol{XY}}{\lVert \boldsymbol{X} \rVert \lVert \boldsymbol{Y} \rVert} = \frac{\sum_{j=1}^{h} x_j y_j}{\sqrt{\sum_{j=1}^{h} x_j^2 \ \sum_{j=1}^{h} y_j^2}}$	$[0,2]$		

[a]The variable x_j(y_j) is the value of the coordinate j of the atom s and the atom t, corresponding to the molecule X (Y), respectively. The h value is the Cartesian coordinates (x, y, z) of an atom. The p values in Minkowsky metric are 0.25, 0.5, 1 (Manhattan), 1.5, 2 (Euclidean), 2.5 and 3 (Minkowsky). [b]"*Range*" refers to "range" and not to "rank" and is defined as $Range = max\{x_j\} - min\{x_j\}$.

Measure	Formula[a]	Range[b]	Average	Range
SL-like (**M17**)	$d_{XY} = min\{\lvert x_j - y_j \rvert\}$			
Average Euclidean (**M18**)	$d_{XY} = \left(\sqrt{\sum_{j=1}^{n} \lvert x_j - y_j \rvert^2}\right) \Big/ n$			
Squared Euclidean (**M19**)	$MSE_{XY} = \sum_{j=1}^{n} \lvert x_j - y_j \rvert^2$			
Pearson (**M20**)	$s_{XY} = \frac{\sum_{j=1}^{n}(x_j - \overline{X})(y_j - \overline{Y})}{\sqrt{\sum_{j=1}^{n}(x_j - \overline{X})^2 \sum_{j=1}^{n}(y_j - \overline{Y})^2}}$ $r_{XY} = \frac{S_{XY}}{S_X S_Y}$	$[-1{,}1]$		
Cosine-Ochiai (**M21**)	$s_{XY} = \frac{\boldsymbol{XY}}{\lVert \boldsymbol{X} \rVert \lVert \boldsymbol{Y} \rVert} = \frac{\sum_{j=1}^{n} x_j y_j}{\sqrt{\sum_{j=1}^{n} x_j^2 \ \sum_{j=1}^{n} y_j^2}}$ $s_{XY} = Cos_{XY}$	$[-1{,}1]$	$\bar{d} = \frac{d_{XY}}{n}$	$\left[0, \frac{2}{n}\right]$
Chord (**M22**)	$d_{XY} = \sqrt{2(1 - Cos_{XY})}$	$[0, \infty)$	$[0, \sqrt{2}]$	
Fossum (**M23**)	$s_{XY} = F_{XY} = \frac{n\left(\sum_{j=1}^{n} x_j y_j \ - 1/2\right)^2}{\sum_{j=1}^{n} x_j^2 \ \sum_{j=1}^{n} y_j^2}$	$[0, \infty)$		
Tanimoto (**M24**)	$s_{XY} = \frac{\sum_{j=1}^{n} x_j y_j}{\sum_{j=1}^{n} x_j^2 \ + \sum_{j=1}^{n} y_j^2 \ - \sum_{j=1}^{n} x_j y_j}$	$[-1/3{,}1]$		
Kulcynzski (**M25**)	$s_{XY} = \frac{\sum_{j=1}^{n} x_j y_j}{\sum_{j=1}^{n} x_j^2 \ + \sum_{j=1}^{n} y_j^2 \ - 2\sum_{j=1}^{n} x_j y_j}$	$[-1/4, \infty]$		
Sokal-Sneath (**M26**)	$s_{XY} = \frac{\sum_{j=1}^{n} x_j y_j}{2\sum_{j=1}^{n} x_j^2 \ + 2\sum_{j=1}^{n} y_j^2 \ - 3\sum_{j=1}^{n} x_j y_j}$	$[-1/7{,}1]$		
Simpson (**M27**)	$s_{XY} = \frac{\sum_{j=1}^{n} min\{x_j, y_j\}}{min\{\sum_{j=1}^{n} x_j \ , \sum_{j=1}^{n} y_j\}}$	$[0{,}1]$		
Ruzicka (**M28**)	$s_{XY} = \frac{\sum_{j=1}^{n} min\{x_j, y_j\}}{\sum_{j=1}^{n} max\{x_j, y_j\}}$			
Dice (**M29**)	$s_{XY} = \frac{2\sum_{j=1}^{n} x_j y_j}{\sum_{j=1}^{n} x_j^2 \ + \sum_{j=1}^{n} y_j^2}$	$[-1{,}1]$		
Identity corrected (**M30**)	$e'_{XY} = \frac{2S_{XY}}{S_X^2 + S_Y^2 + (\overline{X} - \overline{Y})^2}$	$[-1{,}1]$		
Additivity (**M31**)	$a_{XY} = \frac{2S_{XY}}{S_X^2 + S_Y^2}$	$[-1{,}1]$		
Proportionality corrected (**M32**)	$p'_{XY} = \frac{S_{XY}}{T_X T_Y - \overline{X}\,\overline{Y}}$	$[-1{,}1]$		

[c] $X = \{x_1, x_2, x_3, \dots, x_n\}$ represents a random sample of variable X; $\bar{X}$ and S_X are the arithmetic mean and standard deviation of X, respectively; $S_{XY} = \frac{\sum_{j=1}^{n}(x_j - \overline{X})(y_j - \overline{Y})}{n}$, is the covariance between variables X and Y; $T_X = +\sqrt{\left(\sum_{j=1}^{n} x_j^2\right)/n}$

Three-way (Dis)-Similarity Multi-Metrics

A) Ternary Measures *(T_{XYZ})*

Measure	Formula	Symmetry
Triangle Area (**M33-M34**)	$T_{XYZ} = \sqrt{s(s-d_{XY})(s-d_{YZ})(s-d_{ZX})}$ $s = \frac{d_{XY}+d_{YZ}+d_{ZX}}{2}$	S
Triangle's Incircle Area (**M35-M36**)	$T_{XYZ} = \pi\left(\frac{2\sqrt{s(s-d_{XY})(s-d_{YZ})(s-d_{ZX})}}{d_{XY}+d_{YZ}+d_{ZX}}\right)^2$	S
Summation Sides (**M37-M38**)	$T_{XYZ} = d_{XY} + d_{YZ}$	A
Bond angle (Angle between sides) (**M39-M40**)	A_X, A_Y, A_Z *coordenates of three atoms of a molecule* $U = A_X - A_Y, V = A_Z - A_Y$ $T_{XYZ} = \alpha = \arccos\left(\frac{U*V}{\lvert U\rvert * \lvert V\rvert}\right)$	A
MIN-RULE [1-Nearest neighbor (NN)] (**M41-M42**)	$T_{XYZ} = \min(d_{XZ}, d_{YZ})$	A
JOIN-RULE (2-NN) **(M43-M44)**	$d_{\min} = \min(d_{XY}, d_{YZ}, d_{ZX})$ $d_{\max} = \max(d_{XY}, d_{YZ}, d_{ZX})$ $join(d_{XY}, d_{YZ}, d_{ZX}, d_{\min}, d_{\max}) = \begin{cases} d_{XY}, d_{\min} > d_{XY} < d_{\max} \\ d_{YZ}, d_{\min} > d_{YZ} < d_{\max} \\ d_{ZX}, d_{\min} > d_{ZX} < d_{\max} \end{cases}$ $T_{XYZ} = join(d_{XY}, d_{YZ}, d_{ZX}, d_{\min}, d_{\max})$	S
MAX-RULE (Furthest neighbor) (**M45-M46**)	$T_{XYZ} = \max(d_{XZ}, d_{YZ})$	A
AVE-RULE (Average-link) **(M47-M48)**	$T_{XYZ} = \frac{d_{XY}+d_{YZ}}{2}$	A
MED-RULE (**M49-M50**)	$T_{XYZ} = \frac{d_{XZ}+d_{YZ}}{2} - \frac{d_{XY}}{4}$	A
WAR-RULE (**M51-M52**)	$T_{XYZ} = d_{X\overline{C}}^{\,2} + d_{Y\overline{C}}^{\,2} + d_{Z\overline{C}}^{\,2}$	S

$\overline{C}$ is the centroid for the coordinates (A_X,A_Y,A_Z) for three atoms (X,Y,Z) in the protein, respectively.

Measure	Formula	Symmetry
ADD-RULE (Average D/D degree) (**M53-M54**)	$T_{XYZ}=\frac{1}{3}\left(\frac{d_{XY}}{p_{XY}}+\frac{d_{YZ}}{p_{YZ}}+\frac{d_{ZX}}{p_{ZX}}\right)$	S
SUM-RULE (Wiener index) (**M55-M56**)	$T_{XYZ}=d_{XY}+d_{YZ}+d_{ZX}$	S
PRO-RULE (**M57-M58**)	$T_{XYZ}=d_{XY}\cdot d_{YZ}\cdot d_{ZX}$	S
QUA-RULE (**M59-M60**)	$T_{XYZ}=\left(\frac{d_{XY}{}^{2}+\cdot d_{YZ}{}^{2}+d_{ZX}{}^{2}}{3}\right)^{\frac{1}{2}}$	S
GEO-RULE (**M61-M62**)	$T_{XYZ}=\left(\frac{d_{XY}{}^{3}+\cdot d_{YZ}{}^{3}+d_{ZX}{}^{3}}{3}\right)^{\frac{1}{3}}$	S
RAN-RULE (**M63-M64**)	$T_{XYZ}=\max(d_{XY},d_{YZ},d_{ZX})-\min(d_{XY},d_{YZ},d_{ZX})$	S
IC-RULE Identidad-corregida (**M65-M66**)	$T_{XYZ}=\frac{2(S_{XY}+S_{XZ}+S_{YZ})}{2(S_{X}{}^{2}+S_{Y}{}^{2}+S_{Z}{}^{2})+(\overline{X}-\overline{Y})^{2}+(\overline{X}-\overline{Z})^{2}+(\overline{Y}-\overline{Z})^{2}}$	S
AC-RULE Aditividad-corregida (**M67-M68**)	$T_{XYZ}=\frac{S_{XY}+S_{XZ}+S_{YZ}}{S_{X}{}^{2}+S_{Y}{}^{2}+S_{Z}{}^{2}}$	S
PC-RULE Proporcionalidad-corregida (**M69-M70**)	$T_{XYZ}=\sum_{i<j}^{k}\frac{\left(\sum_{t}^{n}X_{it}X_{jt}-n\overline{X}_{i}\overline{X}_{t}\right)/A}{\frac{k}{2}(k-1)-n\sum_{i<j}^{n}\left[\frac{\overline{X}_{i}\overline{X}_{j}}{A}\right]}$ $A=\left(\sum_{t}^{n}X_{it}{}^{2}\sum_{t}^{n}X_{jt}{}^{2}\right)^{\frac{1}{2}}$	S
LC-RULE Linealidad corregida (mean pair-wise pearson correlation) (**M71-M72**)	$T_{XYZ}=\frac{r_{XY}+r_{YZ}+r_{ZX}}{3}$	S

p_{XY} is the topological distance between the amino acids containing atoms (X and Y) and r_{XY} is the Pearson correlation between variables X and Y.