

Exemple de classification numérique

Soit 5 souches à classer S1, S2, S3, S4 et S5 (situation irréaliste - on n'a pas besoin de la taxonomie numérique pour classer cinq souches – choisie néanmoins pour ne pas alourdir la démonstration) soumises à 12 tests (chiffre nettement insuffisant – il en faut entre 3 et 300 -)

On écrit, pour commencer, les résultats obtenus pour chacune des souches, en notant 0 les résultats négatifs et 1 les résultats positifs pour calculer les indices de distance d .

S ₁	0	0	1	1	1	1	0	1	0	0	1	1
S ₂	0	1	1	1	1	0	0	1	0	0	1	1
nd=2	-	x	.	.	.	x	-	.	-	-	.	.
nS=6	-	.	x	x	x	.	-	x	-	-	x	x

$nd + nS = 8$; $d(S_1, S_2) = nd/nd+nS = 2/8 = 0.250$; l'indice de distance entre S1 et S2 est égale à 0.250.

Quand tous les calculs sont faits, on construit le « tableau en carré »

	S1	S2	S3	S4	S5
S1	0	0.250	0.500	0.800	0.400
S2	0.250	0	0.400	0.350	0.700
S3	0.500	0.400	0	0.100	0.200
S4	0.800	0.350	0.100	0	0.900
S5	0.400	0.700	0.200	0.900	0

L'indice d entre les souches S3 et S4 est le plus faible ; ce sont donc les souches les moins différentes ou les semblables qu'on va réunir en un même groupe G1.

On calcule les nouveaux indices de distance entre chacune des souches restantes et le nouveau groupe qui sont la moyenne des indices de chacun des constituants du groupe.

	S1	S2	S5	G1
S1	0	0.250	0.400	0.650
S2	0.250	0	0.700	0.375
S5	0.400	0.700	0	0.550
G1	0.650	0.375	0.550	0

C'est maintenant S1 et S2 qu'on va regrouper en G2.

	S5	G1	G2
S5	0	0.550	0.550
G1	0.550	0	0.513
G2	0.550	0.513	0

G2 et G1 vont être regroupés en G3

	S5	G3
S5	0	0.550
G3	0.550	0

Il est admis que les souches présentant 90% de similitudes (ou 10% de différences) appartiennent à la même espèce et qu'à 70%, elles sont du même genre.

Sur ce dendrogramme, construit à partir des données de l'exemple ci-dessus, on aperçoit facilement la parenté entre les souches S3 et S4 d'une part, qui appartiennent à la même espèce G1, et les souches S1 et S2 d'autre part, qui sont du même genre G2.

La souche S5 appartient à un genre différent, la souche S_x serait totalement étrangère.

