PAST, logiciel statistique naturaliste

Pierre DIEUMEGARD professeur de SVT Lycée Pothier 45044 Orléans

courriel : pierre.dieumegard@ac-orleans-tours.fr

PAST, logiciel scientifique --2- || Introduction : PAST, un logiciel de statistiques pour les sciences de la nature

Introduction : PAST, un logiciel de statistiques pour les sciences de la nature

« PAST » signifie « PAlaeontological STatistics », c'est à dire « Statistiques pour la paléontologie ». C'est un logiciel dont l'auteur principal est Øyvind Hammer, du Museum d'Histoire Naturelle de l'Université d'Oslo (Norvège). On peut le télécharger sur le site http://folk.uio.no/ohammer/past/

La paléontologie, qui est l'étude des fossiles, est souvent considérée à tort comme une vieille science, sous prétexte que son objet est l'étude des vieilles choses. Beaucoup imaginent que ses pratiquants récupèrent des fossiles dans des carrières et les mettent dans des tiroirs avec une étiquette en latin. Ils n'imaginent pas que ce puisse être une science utilisant des outils et des méthodes modernes.

C'est une erreur. C'est justement parce que la paléontologie étudie des objets anciens, déformés, disparates, avec souvent des données manquantes, des portions disparues... qu'elle a besoin d'outils scientifiques performants, en particulier dans le domaine statistique.

La paléontologie est la science qui étudie les squelettes, les coquilles, les traces des organismes ayant vécu il y a longtemps (plusieurs milliers, millions ou milliards d'années). Elle est donc à l'intersection de la biologie et de la géologie. En dehors de la paléontologie, les statistiques de PAST sont aussi utilisables pour diverses recherches de biologie et de géologie.

Pourquoi PAST ?

L'introduction au manuel de PAST (<u>http://folk.uio.no/ohammer/past/intro.html</u>) indique plusieurs qualités :

- PAST est gratuit
- PAST est fait sur mesure pour la paléontologie, et diverses fonctions ont été développées spécialement dans ce but.
- PAST est facile à utiliser, et peut servir d'introduction pédagogique à l'emploi de logiciels plus complexes
- le site de PAST montre divers exemples d'utilisation.

On peut ajouter d'autres qualités, qui justifient son emploi pour divers domaines des sciences d'observation :

- PAST fonctionne sous toutes les versions de Windows à partir de Windows 95. On peut aussi l'utiliser sous Linux par l'intermédiaire de WINE (par exemple avec Linux Poseidon).
- PAST est un logiciel léger. Sous forme compactée (zip), il peut tenir sur une seule disquette (moins d'un mégaoctet).
- PAST n'est pas sectaire pour le séparateur décimal. Il peut utiliser aussi bien le point que la virgule, et donc recevoir des données des diverses versions des tableurs.
- PAST est facile à utiliser en complément des tableurs de type OpenOffice Calc ou Excel. Ses fichiers sont des « fichiers-textes » lisibles par les tableurs. On peut aussi copier et coller les données entre le tableur et PAST, de façon très facile.
- En un seul logiciel, PAST regroupe des fonctions de tests statistiques classiques et de visualisation de données qu'on ne trouve en général que dans des logiciels spécialisés.

PAST, logiciel scientifique --3- || Introduction : PAST, un logiciel de statistiques pour les sciences de la nature

A qui s'adresse cet ouvrage ?

A tous les scientifiques, professionnels ou amateurs.

- Les spécialistes ont en général des logiciels spécialisés, peut-être plus puissants que PAST. Mais lorsqu'ils sortent de leur spécialisation, et doivent interpréter des données dans un domaine qui n'est plus tout à fait le leur, PAST apporte diverses fonctions permettant d'explorer ces données rapidement.
- Les étudiants en sciences ont souvent des cours de statistiques, mais qui ont un formalisme très mathématique. Ils peuvent avoir besoin de visualiser rapidement des données expérimentales, de faire des comparaison, des graphiques, des estimations... sans trop s'embrouiller dans ces mathématiques.

Structure de cet ouvrage

Globalement, la succession des chapitres correspond à l'ordre des menus de PAST, de gauche à droite. La longueur des chapitres diminue au fil de l'ouvrage, notamment parce que les menus vers la droite sont les plus spécialisés dans des domaines particuliers de la paléontologie, et ont donc moins été développés ici.

Le premier chapitre « Fournir des données à PAST, et les communiquer à d'autres logiciels » correspond aux menus « File » (ouvrir un fichier de données, sauvegarder les données...), au menu « Edit » (copier, couper, coller, ajouter ou retrancher des lignes et des colonnes, transposer un bloc de cellules, etc), et au menu « Transform » (transformation des données par une formule, par tri, par interpolation...).

Le second chapitre « Tracer des graphiques » correspond au menu « Plot ». C'est un ensemble de fonctions très importantes de PAST, avec divers types de graphiques.

Le troisième chapitre « Effectuer des calculs statistiques » correspond au menu « Statistics ». C'est le lieu des estimations de moyennes et de variance, et des tests statistiques plus ou moins classiques.

Le quatrième chapitre « Statistiques multivariables » (menu Multivar) permet de faire analyses en composantes principales, analyses des correspondances, analyses canoniques, mais aussi des classifications en arbres (dendrogrammes) et par regroupement automatique.

Le cinquième chapitre « Modélisation » (menu Model) permet bien sûr la régression linéaire, mais aussi sinusoïdale, polynomiale, logistique...

Le sixième chapitre « Étude de la biodiversité » (menu Diversity) est plutôt destiné aux études d'écologie.

Le septième chapitre « Étude des séries temporelles » est d'un emploi plus large. A partir de séries de données, les fonctions d'analyse spectrale ou d'autocorrélation permettent de déterminer la fréquence des phénomènes périodiques.

Le huitième chapitre « Mesures géométriques » contient des fonctions d'usage assez général, comme les graphiques montrant l'orientation ou la direction de mesures, ou bien l'interpolation pour réaliser des cartes, mais aussi des fonctions plus spécialisées pour la paléontologie comme l'allométrie.

Le neuvième chapitre « Analyses stratigraphiques spécifiques » (menu Strat) est vraiment spécialisé pour la paléontologie et la stratigraphie.

Le dixième et dernier chapitre « Programmation des actions de PaST » (menu Script) indique brièvement comment automatiser des actions de PAST par un langage de programmation.

Table des matières

1	Fournir des données à PAST, et les communiquer à d'autres logiciels	7
	1.1 Le tableau de PAST : les colonnes correspondent aux variables, et les lignes corresponder	nt
	aux observations (mesures, individus)	7
	1.2 Cocher ou non les cases « Edition »	7
	1.3 Menu « File » : charger et sauvegarder des données dans des fichiers	9
	1.4 Menu « Edit » : modifier les données d'un tableau	12
	1.5 Transform : Transformer les données	18
	1.6 PAST utilisé sous Linux	21
2	Plot : Tracer des graphiques	22
	2.1 Pour tous les graphiques : cliquer pour régler les préférences du graphique	22
	2.2 Graph : simples graphes de position.	22
	2.3 XYgraph : le graphique cartésien traditionnel	23
	2.4 XY with error bars : graphique XY avec barres d'erreur	23
	2.5 Histogram : histogramme de fréquence absolue (nombre d'individus) de la population	24
	2.6 Box-plot : boites à moustaches = diagrammes de Tukey	25
	2.7 Percentiles : diagramme de fréquences cumulées	25
	2.8 Normal probability plot : graphique montrant (ou non) une distribution normale	26
	2.9 Ternary : diagramme triangulaire (= diagramme ternaire), pour la composition en 3 facteu	ırs
	2.10 Bubble plot = graphique à bnulles	27
	2.11 Survivorship : courbes de survie d'une population	
	2.12 Landmarks : graphique XY avec points de repères par colorations	
	2.13 Landmarks 3D : graphique XY avec une troisième dimension, peu claire	
	2.14 Matrix : cartographie des valeurs, soit en nuances de gris, soit en couleurs.	
	2 15 Surface · cartographie en relief	31
3	Statistics : effectuer des calculs statistiques.	
0	3 1 Univariate · calculs sur chaque colonne indépendamment	32
	3.2 Correlation · corrélation entre les variables	33
	3 3 Var-covar : calcul de la matrice de variance-covariance des colonnes	33
	3.4 F and T tests (two samples) - comparaison des variances et des movennes de 2 échantillor	ns
	de distributions normales.	
	3.5 F and T from parameters : comparaison de deux échantillons à partir des paramètres	
	statistiques.	
	3.6 T test (one sample) · comparaison d'un échantillon avec une distribution théorique	35
	3.7 Paired tests (t sign Wilcoxon) · comparaison d'échantillons appariés	35
	3.8 Normality (one sample) · test de normalité d'un échantillon	36
	3.9 Chi ² · test du khi deux	37
	3 10 coefficient of variation : test d'égalité des coefficients de variation de deux échantillons	38
	3.11 Mann-Whitney : test U de comparaison de médianes même si la distribution n'est nas	
	normale	38
	3.12. Kolmogorov Smirnov : teste si deux échantillons ont la même distribution	39
	3 13 Spearman/Kendall · deux variables sont-elles corrélées ?	39
	3.14 Contingency table : test d'indépendance de variables d'une table de contingence	39
	ser e contingency more : test à macpendance de variables à une more de contingence	

	3.15 One-way Anova = analyse de variance à un seul facteur	40
	3.16 Two-way ANOVA : analyse de variance à deux facteurs	41
	3.17 Kruskal-Wallis : comparaison multiple de médianes, par une sorte d'analyse de variance	où
	les distributions ne sont pas forcément normales	41
	3.18 One-way ANCOVA : analyse de covariance à un facteur	42
	3.19 Mixture analysis = analyse des mélanges, pour une population hétérogène	42
	3.20 Genetic sequence stats : analyse de séquences génétiques	43
4	Multivar : statistiques multivariables	45
	4.1 Principal components analysis : Analyse en composantes principales (ACP = PCA)	45
	4.2 Principal coordinates : analyse en coordonnées principales, proche de l'ACP	46
	4.3 Non-metric MDS : positionnement multidimensionnel non métrique	47
	4.4 Correspondence = analyse des correspondances = analyse factorielle des correspondances	3
	(CA = AFC).	47
	4.5 Detrended correspondance analysis = analyse des correspondances redressée	48
	4.6 Canonical correspondance analysis = analyse canonique des correspondances.	
	4.7 CABFAC factor analysis	48
	4.8 two blocks PLS · Moindre carrés partiels sur deux blocs	49
	4.9 Seriation	49
	4.10 Cluster analysis = Regroupements en arbres	50
	4 11 Neighbour ioining	
	4 12 K-means clustering	53
	4.12 R means endstering. 4.13 Multivariate normality = test de normalité sur plusieurs variables	53
	4.15 Multivariate normanity less de normanie sur prosecurs variables	53
	4.14 Discriminant / Hotening – Anaryse discriminante, et test 1 de Hotening	
	4.15 Tance motioning – test T de notening pour données apparees	
	4.10 Two-group permutation : lest de permutation pour deux groupes multivales	
	4.17 Dox's M - test d'égante des matrices de covariances pour deux groupes de données 4.18 MANOVA/CVA - analyse de variance multiple / analyse canonique des variables	
	4.18 MANOVA/CVA – analyse de variance multiple / analyse canonique des variables	
	4.19 One-way ANOSIM – analyse de similarités à deux factours	51
	4.20 Two-way ANOSIM – analyse de variance multiple non paramétrique à un facteur	
	4.21 One-way NEWANOVA – analyse de variance induiple non parametrique a un facteur	
	4.22 Mantel lest de contenation entre matrices de distance	
	4.23 SIMPER – pourcemage de similarite	
	4.24 Calibration from ontime	
	4.25 Calibration from optima.	
5	4.26 Modern Analogue Technique	
3	F 1 Lincon récreccion linéoire	
	5.1 Linear : regression linearre.	
	5.2 Linear I indep, n dep = regression lineaire multiple pour une variable (independante) qui	
	determine n variables dependantes.	
	5.3 Linear n indep, 1 dep = regression lineaire multiple, pour expliquer une seule variable	- 7
	(dependante) par n'variables (independantes)	
	5.4 Sinusoidal = ajustement sinusoidal de phenomenes periodiques	59
	5.5 Polynomial = ajustement polynomial	60
	5.6 Logistic = ajustement à la courbe logistique (croissance de populations)	61
	5.7 Smoothing spline = lissage par les courbes spline	61
	5.8 Abundance = modèles d'abondance.	62
	5.9 Species packing = « garniture d'espèces » : abondance d'espèces selon un gradient de fact	eur
-	du milieu	62
6	Diversity : étude de la biodiversité	64
	6.1 Diversity indices = indices de diversité de type α (alpha) = biodiversité locale	64

	6.2 Quadrat richness = richesse des quadrats = richesse intralieu	64
	6.3 Beta diversity =indices de diversité de type beta = diversité interlieux	65
	6.4 Taxonomic distinctness = distance taxonomique	65
	6.5 Individual rarefaction : estimation de l'effet de l'effort d'échantillonnage	66
	6.6 Sample rarefaction (courbe d'accumulation spécifique) : indice de Mao-tau	66
	6.7 Compare diversities : comparaison des diversités	67
	6.8 Diversity t test : test t de diversité	67
	6.9 Diversity profiles : profils de diversité	68
7	Time : étude des séries temporelles	69
	7.1 Spectral analysis = analyse spectrale (calcul des fréquences d'événements réguliers)	69
	7.2 Autocorrelation = autocorrélation : les mesures sont elles liées au cours du temps ?	70
	7.3 Crosscorrelation = corrélation croisée entre deux variables	70
	7.4 Wavelet transform : transformation en ondelettes	72
	7.5 Walsh transform = transformation de Walsh ou de Hadamard	72
	7.6 Runs test = test d'aléatoirité d'une série de données	73
	7.7 Mantel correlogram = périodogramme et corrélogramme de Mantel	73
	7.8 ARMA	75
	7.9 (Insolation)	75
8	Geomet = mesures géométriques	76
	8.1 Directions (one sample) = étude des orientations pour un échantillon	76
	8.2 Directions (two samples) = comparaison des directions de deux échantillons	77
	8.3 Circular correlation = corrélation angulaire	78
	8.4 Nearest neighbour (2D points) = tests de la répartition aléatoire sur une surface	78
	8.5 Ripley's K (2D points) = indication visuelle du groupement des points	79
	8.6 Gndding = interpolations spatiales	80
	8.7 Multivanate allometry = Allométrie sur plusieurs variables	81
	8.8 Divers types d'analyses de forme (analyses procustéennes)	81
9	Strat : analyses stratigraphiques spécifiques	84
	9.1 Unitary associations = associations unitaires	84
	9.2 Ranking-Scaling = classement des données stratigraphiques	85
	9.3 Constrained optimization CONOP	85
	9.4 Appearance Event Ordination	85
	9.5 Diversity curve = courbe de diversité, avec apparition et disparition de taxons	86
	9.6 (Range confidence intervals = calcul des intervalles de confiance de la distribution des	
	fossiles)	86
	9.7 Distribution-free range CIs = intervalles de confiance indépendants de la distribution	86
	9.8 Spindle diagram = graphique d'observation des taxons au cours du temps	87
	9.9 Filter events.	88
1	0 Cladistics : analyses cladistiques	89
1	10.1 Parsimony analysis.	89
I	1 Script : programmation des actions de PAS1	90
	11.1 Chargement et execution d'un programme	90
П	11.2 Structure du langage	90
В	orden landes	92
11	nuex lexical	93

1 <u>Fournir des données à PAST, et les</u> <u>communiquer à d'autres logiciels</u>

PAST ressemble à un tableur : les données sont organisées en lignes (horizontales) et en colonnes (verticales), mais ce n'est pas vraiment un tableur comparable à Excel, OpenOffice Calc ou Gnumeric. Il n'y a qu'une seule table possible, et on ne peut pas mettre de formules dans les cases. On peut certes changer les valeurs par une formule, mais de façon différente des tableurs (voir 1.5).

Les cases ne doivent pas contenir des formules mais seulement des valeurs (numériques ou alphanumériques).

1.1 <u>Le tableau de PAST : les colonnes correspondent aux variables, et</u> <u>les lignes correspondent aux observations (mesures, individus...)</u>

@	PAST								
File	Edit	Transform P	lot Statistics	s Multivar	Model Diver	sity Time	Geomet St	rat Cladistics	Script
6	> 😫	A	📴 🔁 T	Edit mode	🔲 Edit	labels	🔲 Square	mode	
		A	в	С	D	E	F	G	н
1									
2									
3									
4									
5									

Les cases blanches sont celles qui vont correspondre aux données. Comme dans la plupart des tableurs, les colonnes sont désignées par des lettres, et les lignes sont désignées par des nombres. Au démarrage, il n'y a que 26 colonnes (de A à Z), et 99 lignes, de 1 à 99. On peut augmenter ce nombre par le menu Edit « Insert more rows » pour insérer des lignes et « Insert more columns » pour insérer des colonnes. Lorsqu'on charge un gros fichier, le nombre de lignes et/ou de colonnes est augmenté automatiquement.

Les cases grises correspondent aux labels des lignes et des colonnes. Lorsque « Edit mode »est coché, mais non « Edit labels », on ne peut pas changer le contenu de ces cases. Lorsque « Edit mode » est coché, on peut changer la valeur de ces cases, et ainsi donner un nom explicite aux variables en changeant A, B, C, etc, ou donner un nom explicite aux individus mesurés en changeant 1, 2, 3, etc.

1.2 Cocher ou non les cases « Edition »

1.2.1 Trois cases à cocher au dessous du menu général

- « Edit mode » est coché par défaut. Cela signifie que l'on travaille en mode « édition », et l'on peut modifier le contenu des cases blanches (valeurs mesurées), mais non le contenu des cases grises (labels des lignes et des colonnes). Lorsque cette case n'est pas cochée, on ne peut pas cliquer sur une case et y entrer des valeurs en frappant au clavier, mais on peut quand même coller des données à partir du presse-papier de Windows.
- « Edit labels » n'est pas coché par défaut. Cela signifie que l'on peut changer au clavier la valeur des cases blanches, mais non la valeur des labels de lignes et de colonnes. Lorsque cette case est cochée, on peut aussi changer la valeur des labels de lignes et de colonnes,

PAST, logiciel scientifique $-8- \parallel 1$ Fournir des données à PAST, et les communiquer à d'autres logiciels

aussi bien au clavier que par collage à partir du presse-papier.

 - (« Square mode » ne sert qu'à visualiser les cases vides dans certains cas. Le couple présence/absence est codé par 1/0 (respectivement). Toutes les valeurs positives sont considérées comme « présence ». En cochant « square mode », les cases avec présence apparaissent en noir.)

1.2.2 <u>A quoi sert le mode autre qu'édition ?</u>

Lorsque la case « Edit mode » n'est pas cochée, on ne peut pas entrer des valeurs au clavier. Par contre, il est très facile de déplacer des lignes et des colonnes. Il suffit de sélectionner la ligne en cliquant sur son nom (case de gauche), et l'on peut la déplacer simplement en déplaçant le curseursouris tout en laissant le bouton enfoncé. Il en est de même pour les colonnes, que l'on peut très facilement permuter de la même façon : cliquer sur le nom, puis tirer la colonne d'un côté ou de l'autre en bougeant la souris, bouton gauche enfoncé.

C'est intéressant pour diverses représentations graphiques ou tests statistiques, qui nécessitent d'avoir les colonnes dans un ordre bien déterminé.

1.2.3 Données possibles dans les labels de lignes ou de colonne

Il ne faut pas mettre d'espace dans les étiquettes des colonnes et des lignes (remplacer les espaces par des tirets de soulignement).

1.2.4 Les données possibles dans les cases de mesure :

N'importe quelles valeurs peuvent être entrées dans les cases, mais la plupart des fonctions demandent des valeurs numériques. Le séparateur décimal peut être aussi bien la virgule que le point.

Les données manquantes sont codées par le point d'interrogation « ? » ou la valeur -1. Attention, de nombreuses fonctions ne supportent pas les données manquantes.

Les cellules vides sont codées par un point.

1.2.5 <u>Sélectionner une zone</u>

Attention, cela ne fonctionne pas comme la majorité des tableurs ! PAST se lance normalement en mode « édition » (la case On ne peut sélectionner une zone en déplaçant la souris avec le bouton enfoncé que lorsqu'on n'est pas en mode édition (décocher la case correspondante)

- on sélectionne une ligne en cliquant sur l'étiquette de la ligne (colonne de gauche)
- on sélectionne une colonne en cliquant sur l'étiquette de la colonne (ligne du haut)
- On peut sélectionner plusieurs lignes en cliquant sur l'étiquette de la ligne tout en maintenant la touche majuscule enfoncée.
- On peut sélectionner plusieurs colonnes en cliquant sur l'étiquette de la colonne tout en maintenant la touche majuscule enfoncée.
- On peut sélectionner un pavé de cases en cliquant sur une case, puis sur une autre, tout en maintenant la touche majuscule enfoncée. On peut aussi cliquer sur une case d'un angle du pavé, puis sélectionner ce pavé en appuyant sur la touche « Majuscule » en même temps que sur les touches-flèches.
- On peut sélectionner tout le tableau en cliquant sur la case grise en haut à gauche.

PAST, logiciel scientifique --9- || 1 Fournir des données à PAST, et les communiquer à d'autres logiciels

1.2.6 <u>Grouper et colorer des lignes</u>

On a le choix de 16 couleurs et types de puces après activation de l'option Edit | Row color symbol. Cette option permet de visualiser facilement divers groupes de mesures.

1.3 <u>Menu « File » : charger et sauvegarder des</u> <u>données dans des fichiers</u>

Les fichiers sont de type « texte avec champs séparés par des tabulations ». La fin de ligne (passage à l'enregistrement suivant) est marquée par un retour-chariot et un saut de ligne, ce qui est la norme sous Windows.

Les cellules vides sont codées par des points. Normalement, la première case de la première ligne contient un point.

Les cases ne doivent pas contenir d'espaces, et on doit les remplacer par des tirets de soulignement (ce remplacement est fait automatiquement lorsque

PAST lit un fichier-texte provenant d'un autre logiciel (un tableur par exemple), où les espaces sont possibles.

- 1.3.1 <u>« New » : nouveau tableau vide</u>
- 1.3.2 <u>« Open » : ouvrir un fichier de données</u>

🖑 CO2_2	008. dat	🖀 CO2_2008. dat - OpenOffice.org Writer							
File Edit	Transform Plot Statis	Eichier Édition Affichage Insertion Format Tableau Outils Fenêtre							
🕰 🗠	A 😫 🤽	: 🔁 - 😕 🗔 👒 🕑 🔜 🖴 🕓 🎼 🛍							
	CO2_(ppm)_ B	🕴 🕢 Texte préformaté 🛛 Times New Roman 🔽 10							
1990	356.15								
1991	357.22	L1							
1992	358.48								
1993	358.11								
1994	359.63								
1995	362.16								
1996	364.62								
1997	365.59								
1998	367.49								
1999	370.9	\rightarrow CO2_(ppm)_ \rightarrow ¶							
2000	372.25	1990 → 336.13 →¶							
12		→ 1991 → 337.22 → ¶							
		- 1992 → 338.48 → 1							
		$1993 \rightarrow 336.117$							
		$1995 \rightarrow 362.16 \rightarrow \P$							
		° 1996 → 364.62 →¶							
		1997 → 365.59 →¶							
		1998 → 367.49 →¶							
		1999 → 370.9 → ¶							
		2000 → 372.25 →¶							
		и ·							

On peut utiliser la fonction « open » du menu File, ou bien faire un glisser-déposer de l'icône du fichier sur la fenêtre de PAST.

Normalement, PAST peut charger la première page des fichiers Excel ainsi que quelques autres

🛷 PAST									
File	Edit	Transform	Plot						
N	ew								
0	pen								
Ir	Insert File								
Sa	Save Ctrl+S								
Sa	ave as								
Pi	rint								
Pt	rint Se	tup							
A	bout								
E	xit	Ctrl+	Q						

PAST, logiciel scientifique --10- || 1 Fournir des données à PAST, et les communiquer à d'autres logiciels

formats plus exotiques (Nexus, TPS, Rasc...)

MAIS en fait, il faut que le vrai logiciel Excel soit installé sur l'ordinateur : PAST est incapable de charger seul les fichiers Excel. Le mieux est donc d'utiliser les fichiers de type texte, ou bien de coller les données à partir d'un tableur (OpenOffice, Excel, ou autre).

Cette option emplit seulement les premières cases : colonnes à partir de A, lignes à partir de 1.

1.3.3 <u>« Insert File » : pour insérer un fichier dans le tableau actuel de Past</u>

Cette option permet d'ajouter des données dans une partie ou une autre du tableau. Les données sont ajoutées à partir de la case active (vers les colonnes de droite et les lignes au dessous de la case active).

Attention ! les labels des colonnes et des lignes sont bien sûr mis à leur place, en tête de colonne et de ligne, même si de nombreuses cases blanches séparent ensuite les labels des cases contenant les valeurs insérées.

	🖑 c	ontir	igence. d	lat						
I	File	Edit	Transfor	m Plot	Statistics	Multivar	Model Dive	rsity Time	Geomet 3	Strat Cladist
	6	-		9		Edit mode	e 🔲 Edi	t labels	🔲 Squar	e mode
I			A	в		с	blanc	bleu	rouge	vert
I	1									
I	2									
I	3									
I	clair						10	20	30	15
I	moye	n					15	55	44	33
I	foncé	,					10	40	35	29
	très_	foncé					5	20	25	20
	8									

1.3.4 <u>« Save » pour enregistrer le fichier</u>

Là encore, le mieux est d'utiliser le format par défaut (.DAT = fichier-texte lisible par les tableurs). Il faut mettre manuellement l'extension « .dat » pour que celle ci existe et soit identifiable ensuite par PAST (fonctions de lecture « Insert file » ou « Open »).

Si des lignes ont été associées à des couleurs, l'étiquette de la ligne dans le fichier commencera par un soulignement, continuera par un nombre de 0 à 15 correspondant à la couleur, et continuera par un soulignement.

Si des colonnes sont d'un type particulier, l'étiquette de la colonne commencera par un soulignement, puis un nombre entre 0 et 3, puis un soulignement (0 : indéfini, 1 = ordinal, 2 = nominal 3 = binaire).

PAST, logiciel scientifique --11- $\parallel 1$ Fournir des données à PAST, et les communiquer à d'autres logiciels

							and and a contract of the Aide					
🧶 graines	. dat		h		_	🗃 gra	🗃 graines.dat - OpenOffice.org Writer					
File Edit Transform Plot Statistics Multivar Model Diversity T						Fichier	r Édition Affichage Insertion Format Tableau Outils Fenêtre Aide					
🕞 😫 🎒 🎇 🔽 🔽 Edit mode 🔽 Edit labels			labels		· 😕 🖶 👟 🖹 🔮 🕾 🏷 🕾 🕹 🛍 • 🎸 ⋽ • @ • 💩 🏢 •							
	eau	protides	glucides	lipides	sels_r							
Blé,_Seigle,C	14	12	70	2	2	: 🖻 🛛	Texte préformaté 🕑 Times New Roman 🕑 10 🕑 G I S 🔳 🗄 🗄					
Maïs	12	11	70	5	2							
Pois,_Fève,_	14	25	55	2	4	<u> </u>						
Lupin	12	40	40	5	3							
Noix,_Ricin	7	2	10	60	3	_						
6												
7												
8					_	-						
9												
10						- I	→ eau → protides glucides lipides → sels minéraux → ¶					
11							$1_{\text{Blé}, \text{Seigle,Orge}} \rightarrow 14 \rightarrow 12 \rightarrow 70 \rightarrow 2 \rightarrow 2 \rightarrow 9$					
12							$1_{\text{Maïs}12}$ → 11 → 70 → 5 → 2 → ¶					
13							11 _Pois,_Fève,_Haricot+14 → 25 → 55 → 2 → 4 → ¶					
14						N	$11_Lupin \rightarrow 12 \rightarrow 40 \rightarrow 40 \rightarrow 5 \rightarrow 3 \rightarrow \P$					
15						1						
16						0						
17												

1.3.5 <u>« Save as » pour enregistrer avec un nouveau nom</u>

Là encore, mettre l'extension « .dat » pour que le nom apparaisse dans la liste des fichiers lorsque PAST veut ouvrir un fichier.

1.3.6 <u>« Print » pour imprimer le résultat</u>

C'est une impression immédiate, sans fioritures, sans réglages de mise en page, sur l'imprimante par défaut du système informatique. Pour une mise en page soignée, utiliser plutôt un bon logiciel de bureautique.

1.3.7 <u>« Print setup » pour régler l'imprimante</u>

Il faut penser à ce réglage <u>avant</u> de lancer l'impression par « Print »

- 1.3.8 <u>« About » qui indique l'origine de ce logiciel</u>
- 1.3.9 <u>« Exit » pour clore le logiciel lorsque tout est terminé</u>

1.4 <u>Menu « Edit » : modifier les données</u> <u>d'un tableau</u>

1.4.1 <u>Undo : annuler une opération</u> précédente, et Redo pour la refaire

Il faut comprendre « opération précédente » comme toute action sur le clavier, ou action sur la souris, qui modifie le tableau en cours. Par exemple, coller, couper, frapper une touche pour entrer un chiffre, etc.

Donc lorsqu'on a collé malencontreusement de mauvaises valeurs à la place des bonnes, il faut choisir « Undo » pour revenir à l'état antérieur.

Il y a plusieurs (un grand nombre ?) niveaux d'annulation : en validant plusieurs fois cette option « Undo », on peut revenir à des états antérieurs assez lointains.

1.4.2 <u>Cut, copy, paste pour couper, copier, coller les données sélectionnées</u>

C'est le mécanisme classique pour les logiciels de bureautique. « Cut » pour couper les données et les envoyer vers le presse-papier, pour ne laisser à leur place dans le tableau que des cases blanches. « Copy » pour copier les données vers le presse-papier, mais en laissant les valeurs initiales dans les cases.

🛷 P	AST							
File	Edit	Transform	Plot	Statistics	Multivar	P		
6	Ur Re	ndo edo			Ctrl+Z			
	CL	ut			Ctrl+X			
1	Co	ру			Ctrl+C			
2	Pa	aste			Ctrl+V			
3	Re	emove						
4	Se	elect All						
5	De	ename rows						
6	Re	ename colum	 15					
7	Ro	name colorisym	hol					
8	Ni	imbers to col	orsisv	mbols				
9	G	nlumo data ty	nes					
10	In	sert more ro	NS					
11	In	sert more co	lumns.					
12								
13	Re	eplace						
14	Co	olumn width	•					
15	Fo	ont						
16	Tr	anspose						
17	Gr	ouped colum	ns to f	Multivar				
18	Gr	ouped rows	to Mult	tivar				
19	St	ack colored r	ows in	to columns				
20	Sa							
21	Events to samples (RASC to UA)							
22	Co	ounter .	,					

« Paste » permet de coller les données à partir du presse-papier. S'il n'y avait qu'une seule valeur dans le presse-papier, elle est collée dans la case active. S'il y avait plusieurs valeurs, toutes ces valeurs sont mises dans les cases suivant la case active, vers la droite et vers le bas. Cette option écrase éventuellement les données qui se trouvaient dans les cases : pour annuler cet écrasement, utilisez l'option « Undo ».

Ce mécanisme d'utilisation du presse-papier est la méthode la plus simple et la plus pratique pour échanger des données avec les logiciels extérieurs, en particulier les tableurs.

1.4.3 <u>Remove pour supprimer une ligne ou une colonne</u>

Pour cela, il faut avoir sélectionné une ligne en cliquant sur son nom (colonne de gauche, en gris), ou avoir sélectionné une colonne en cliquant son nom (ligne supérieure, en gris).

Les autres lignes plus vers le bas, ou les autres colonnes plus à droite, se rapprochent de l'origine et prennent la place des lignes ou colonnes supprimées.

1.4.4 <u>Select All, pour tout sélectionner</u>

Cette option sélectionne l'ensemble des cases du tableau ; c'est la même action que cliquer sur la case-origine (A1), en haut à gauche du tableau.

On peut alors couper, copier, coller toutes les données du tableau.

PAST, logiciel scientifique $-13 - \parallel 1$ Fournir des données à PAST, et les communiquer à d'autres logiciels

1.4.5 <u>Rename rows et Rename columns pour renommer des lignes ou des</u> <u>colonnes</u>

Lorsque l'on travaille en mode édition (cases « Edit mode » et « Edit labels » cochées), on peut directement cliquer sur les noms des colonnes et des lignes, et les modifier.

Par contre, lorsque la case « Edit labels » est décochée, on ne peut pas modifier ainsi les labels des lignes et des colonnes. Après avoir sélectionné la ligne ou la colonne souhaitée, on valide cette option, et il apparaît une petite boite de dialogue qui permet de modifier ce nom.

1.4.6 <u>Row color/symbol pour visualiser des groupes d'individus (ou de</u> <u>mesures)</u>

🛞 PAST			🗶 Tag ro	ws	(
File Edit	Transform P	Not Statistics	s Multivar	Model Div	versity Time Geo	1 .			
۵ 🛤	A .	🛤 🔃 r	Edit mode	ΓE	dit labels 🛛 🦵	Dot .	Cross +	Square	Filled sq
	eau	protides-prot	glucides	lipides	sels_minérauF	Cross x	Circle o	Diamond	Star *
Ble_Seigle (14	12	70	2	2	-	Harry		
Mais	12	11	70	5	2	Thangle	Line -	Bar ;	Oval
Pals_Fève_	14	25	55	2	4	1	Inite the	Ell/isova	THE OWNER
Lupin	12	40	40	5	3		miv. iri	THE REAL PROPERTY.	THE GROUP
Noix_Ricin	7	2	10	60	3			1	
6		1		1			Clo	se	

Après avoir sélectionné une ligne (ou quelques lignes adjacentes), cette option permet de colorier la case contenant le nom de la ligne (le label de la ligne). Ceci permet de visualiser des groupes de lignes partageant un caractère en commun. Cette visualisation a 16 modalités possibles, et associe un symbole à une couleur pour chacun des groupes ainsi définis.

Certaines fonctions de visualisation ou de tests statistiques font la différence entre ces groupes de couleurs.

1.4.7 <u>Numbers to colors/symbols : pour associer automatiquement des</u> couleurs et des symboles à certaines valeurs numériques

🛷 PAST						
File Edit '	Transform P	lot Statistics	; Multivar	Model Diver:	sity Time	Geomet St
😤 🗠	B	🛤 🔃 T	Edit mode	🔲 Edit	labels	🔲 Square
	eau	protides-prot	glucides	lipides	sels_minéra	ιF
Blé,_Seigle,C	14	12	70	2	2	
Maïs	12	11	70	5	2	
Pois,_Fève,_	14	25	55	2	4	
Lupin	12	40	40	5	3	
Noix,_Ricin	7	2	10	60	3	
6						

Après avoir chois une colonne, contenant des valeurs numériques caractéristiques des groupes souhaités, cette option permet d'associer automatiquement une couleur et un symbole à chacun des groupes.

Il faut que la colonne contienne des nombres de 1 à 12. Au delà de 12, la couleur attribuée est équivalente à celle du 1, et pour des nombres négatifs ou nuls, les couleurs sont celles des nombres

PAST, logiciel scientifique --14– $\parallel 1$ Fournir des données à PAST, et les communiquer à d'autres logiciels

entre 1 et 3 (bizarre !).

1.4.8 <u>Column data types : déclarer un type particulier pour les données</u> <u>d'une colonne</u>

Après avoir choisi une ou quelques colonnes, cette option permet de leur attribuer un type particulier de données, qui sera affiché au niveau du label de la colonne :

- ordinal, c'est à dire un classement. La lettre O apparaît dans le label de la colonne
- nominal, c'est à dire une simple description de l'observation. La lettre N apparaît dans le label de la colonne
- binaire, lorsque deux états seulement sont possible : la lettre B apparaît dans le label de la colonne.
- Dans le cas général, le type n'est pas défini, et correspond normalement à des valeurs numériques réelles. Aucune lettre supplémentaire n'apparaît dans le label de la colonne.

1.4.9 <u>Insert more rows et Insert more columns, pour ajouter des lignes ou</u> <u>des colonnes</u>

On peut ainsi frapper au clavier (ou coller à partir du presse-papier) davantage de valeurs, soit pour de nouvelles colonnes, soit pour de nouvelles lignes.

1.4.10 <u>Replace, pour remplacer des caractères ou des valeurs numériques</u> par d'autres

Il apparaît la boite de dialogue classique, permettant de remplacer un caractère par un autre.

PAST accepte aussi bien la virgule que le point comme séparateur décimal. Cette option permet en particulier de remplacer le point par la virgule, ou l'inverse, avant une copie des données à destination d'un autre logiciel plus exigeant pour le séparateur décima, ou bien avant une sauvegarde sous forme de fichier-texte, si on désire le faire relire par un logiciel souhaitant soit le point, soit la virgule.

1.4.11 <u>Column wdth, pour régler la largeur d'une ou quelques colonnes</u>

En dehors du mode « Edit labels », on peut faire varier la largeur des colonnes par la méthode normale des tableurs : cliquer sur la limite droite de la colonne, et faire glisser cette limite tout en appuyant sur le bouton gauche.

Lorsque la case « Edit labels » est cochée, cette méthode n'est pas possible, et il faut employer cette option du menu « Edit », qui fait afficher une boite de dialogue où l'on choisit la largeur de la colonne (en pixels).

1.4.12 <u>Font, pour régler la fonte des caractères d'affichage de l'ensemble de</u> <u>la feuille</u>

Contrairement aux tableurs modernes, il n'est pas possible de choisir la fonte d'une ou quelques cases. Cette option fait apparaître une boite de dialogue où l'on choisit la fonte dans laquelle seront affichées toutes les cases du tableau.

Comme on ne peut pas choisir la hauteur des lignes, cette option est peu intéressante en général.

1.4.13 <u>Transpose, pour transposer l'ensemble du tableau (lignes en colonnes et colonnes en lignes)</u>

En choisissant une fois cette option, l'ensemble du tableau est transposé : les lignes deviennent les

PAST, logiciel scientifique $-15- \parallel 1$ Fournir des données à PAST, et les communiquer à d'autres logiciels

colonnes et inversement. En choisissant une deuxième fois cette option, on revient à la situation initiale, puisque les nouvelles colonnes sont transformées en lignes, ce qui était leur état initial.

Cette option peut être utile pour certaines analyses statistiques, qui donnent des rôles différents aux variables et aux individus.

1.4.14 <u>Grouped columns to Multivar et Grouped rows to Multivar, pour</u> réorganiser les données à partir de fichiers compliqués

A Lint	titlad						_		
U Cin	inneu		and the maximum		a set to a s		🖑 Untitle	ed	
File E	dit Transform Plot S	tatistics Multiva	ir Model Diversity	Time Geomet Stra	t Cladistics	Script	File Edit	Transform Plot Statis	stics Multivar Mo
B	😫 💄 👭 🕻	🖹 🔽 Edit mo	ide 🦳 Edit labels	s 🦳 Square m	ode		😫 🗠	A 🕸 🔁	🔽 Edit mode
	В	C	D	E				1	
1	0	2.375	4.75	7.109			В	lo	
2	9.484	11.859	14.219	16.594	Column	s per group	С	2.375	1
3	18.969	21.328	23.703	26.063			<u>D</u>	4.75	
4	28.438	30.813	33.172	35.547	Enter nu	umber of columns (<u>E</u>	7.109	_
5	37.922	40.281	42.656	45.031	4		<u>E</u>	9.484	
6	47.391	49.766	52.125	54.5	, '		<u>F</u>	11.859	_
7	56.875	59.234	61.609	6			<u>G</u>	14.219	-
8	66.344	68.719	71.094	73.453			H	16.594	-
9	75.828	78.188	80.563	82.938		` <u> </u>	<u> </u>	18.969	-
10	85.297	87.672	90.047	92.406			J	21.328	-
11	94.781	97.156	99.516	101.891			<u> </u>	23.703	-
					7 / _		. <u>–</u> М	28.003	-
A Un	titled			^`			N	30.813	-
	nnied Na z 6 pl. c						0	33 172	
File E	dit Transform Plot S	tatistics Multiv	ar Model Diversity	Time Geomet Stra	at Cladistics	s Script	P	35 547	
	🖹 🖳 🔎 📑	💽 🔽 Edit m	ode 🔽 Edit label:	s 🗖 Square m	node		Q.	37.922	1
			-				R	40.281]
	B	C	D	E	E	F(S	42.656]
1	0	2.375	4.75	7.109	9.484	11.859 1	<u>T</u>	45.031	
2							<u>U</u>	47.391	
<u> </u>							<u>v</u>	49.766	
							<u>W</u>	52.125	_
							<u>×</u>	54.5	
							<u> </u>	56.875	
							4	59.234	
								161.609	

On peut prendre comme exemple les fichiers des logiciels d'acquisition de données Serenis (Jeulin). Pour chaque variable les valeurs sont mises dans un paquet constitué de 4 colonnes ; puis il y a quelques lignes sans valeurs numériques mesurées. De nouveau, une variable dont les valeurs sont en 4 colonnes, quelques lignes, la variable suivante, etc.

Lorsqu'on a copié-collé les valeurs d'une variable dans PAST, et sélectionné les 4 colonnes, on peut prendre l'option « Grouped columns to multivar », et indiquer qu'il y a 4 colonnes. On obtient alors un vecteur-ligne de toutes les valeurs de la variable, que l'on peut transposer par l'option « Edit-Transpose », de façon a obtenir un vecteur colonne de la variable.

En faisant de même pour les diverses variables, puis en collant ces vecteurs-colonnes côte à côte dans un même tableau de PAST, on obtient un tableau constitué des multiples valeurs des diverses variables.

Ce n'est pas simple.... Dans tous les cas, il faut connaître la structure des données.

PAST, logiciel scientifique --16- $\parallel 1$ Fournir des données à PAST, et les communiquer à d'autres logiciels

1.4.15 <u>Stack colored rows into columns : séparer les différentes couleurs en</u> colonnes



Lorsqu'on a des données de différentes catégories, visualisées par des couleurs différentes, on peut souhaiter étudier les caractéristiques de chaque catégorie. Il faut commencer par trier les données selon ces catégories (menu « Transform | Sort... ». Ensuite, le choix de cette option sépare les catégories en diverses colonnes ; ceci fait, les noms (labels) des colonnes et des lignes n'ont plus de signification, pas plus que les couleurs des lignes.

1.4.16 <u>Samples to events (UA to RASC), et Events to samples (RASC to UA) :</u> à utiliser pour les corrélations biostratigraphiques

Ces deux options sont destinées à être utilisées en biostratigraphie par la méthode des Associations Unitaires (menu « Biostrat »).

1.4.17 <u>Counter : un compteur au clavier</u>

Une fonction de comptage est disponible, par exemple pour compter au microscope des microorganismes de divers groupes. Chaque groupe va correspondre à une colonne, et une ligne sera utilisée pour chaque étude (tube de Drosophiles, boite de Pétri de colonies bactériennes, quadrat dans une prairie pour le comptage de plantes...)

🦻 Untitle	d							
File Edit "	Transform P	lot Statistic	s Multivar N	1odel	🕘 Counter , row co	ndition 2		
😫 🗠	A	😫 🔃 I	Edit mode		bactéries 1	bactéries 2	levures	bactéries 1 bactéries 1
	bactéries 1	bactéries 2	levures	D				bactéries 1
ondition 1	3	1	0					bactéries 2
ondition 2								bactéries 2
ondition 3					3	4	0	bactéries 2
ŀ								
5					+1	+ 2	+ 3	
}					-	_	-	
r								
}								
)								
0								
1								
2								
3						🔲 Beep each	Tabalı 7	_
4					<u> </u>	Beep every 10		
5								

PAST, logiciel scientifique $-17- \parallel 1$ Fournir des données à PAST, et les communiquer à d'autres logiciels

Par exemple pour le comptage de colonies de deux types différents de bactéries ou de levures, dans diverses conditions :

- sélectionner les trois cases d'une ligne correspondant aux trois types de micro-organismes
- Choisir « Edit |Counter » : une nouvelle fenêtre apparaît.
- En cliquant sur les boutons avec un « + », le compteur de la catégorie correspondante est incrémenté ; en cliquant sur les boutons « - », le compteur est décrémenté.
- Lorsque le comptage est terminé, il faut cliquer sur le bouton marqué d'une croix rouge (en bas à gauche) : la fenêtre se ferme, et les nombres correspondant au comptage apparaissent dans les cases correspondantes du tableau.
- Pour la condition suivante (boite de Pétri suivante, quadrat suivant...), il faut sélectionner dans la ligne suivante les cases correspondant aux types à compter, puis sélectionner « Edit | Counter », etc.

1.5 <u>Transform : Transformer les</u> <u>données</u>

On peut faire subir aux données différentes opérations. Ces opérations ne sont possibles que sur une zone sélectionnée : elles ne sont pas possibles sur une seule case.

1.5.1 Log : logarithme en base 10

Il faut des données positives, mais les données manquantes (?) sont acceptées

1.5.2 <u>Remove trend : soustraction de la</u> tendance (par régression linéaire)

Il faut sélectionner des cases appartenant à au moins deux lignes (éventuellement une colonne ou deux).

Si des cases de deux colonnes sont sélectionnées (X et Y), PAST effectue le calcul de la régression linéaire de

🛷 PAST File Edit Transform Plot Statistics Multivar Mod Log <u>کا</u> ا Remove trend Subtract mean Row percentage Row normalize length 2 3 Abundance to presence/absence 4 Procrustes (2D/3D) 5 Bookstein (2D) 6 Remove size from landmarks (2D/3D) 7 Transform landmarks (2D) 8 Remove size from distances 9 10 Sort ascending 11 Sort descending 12 Sort on color 13 Column difference Regular interpolation 14 Evaluate expression 15

Y en X, puis enlève de la colonne Y la valeur de la régression, pour ne plus y laisser que l'écart à la droite de régression. La colonne X n'est pas changée.

Si des cases d'une seule colonne sont sélectionnées, PAST effectue le calcul de la régression linéaire de cette colonne par rapport au numéro des lignes, puis l'enlève de la valeur de chaque case.

Les cases non sélectionnées ne changent pas

1.5.3 <u>Substract mean : soustraction de la moyenne des cases de la colonne</u>

Après avoir sélectionné plusieurs cases (une ou plusieurs colonnes), cette option soustrait à la valeur de chaque case la moyenne des valeurs de la colonne (en fait, la moyenne des cases sélectionnées de la colonne). Les cases non sélectionnées ne changent pas.

1.5.4 <u>Row percentage : pourcentage de la ligne</u>

Après avoir sélectionné plusieurs colonnes, sur une ou plusieurs lignes, cette option remplace les valeurs des cases par le pourcentage de ces valeurs par rapport au total sélectionné sur la ligne. Cela permet de remplacer des quantités absolues par des quantités relatives. Les cases non sélectionnées ne changent pas.

1.5.5 Row normalize length : longueur normalisée sur la ligne

Toutes les valeurs sélectionnées de la ligne sont divisées par la longueur euclidienne de ce vecteurligne (la racine carrée de la somme des carrés des valeurs)

1.5.6 <u>Abundance to presence/absence : convertir les données d'abondance</u> en présence/absence

Toutes les valeurs positives sont remplacées par des 1.

1.5.7 <u>Procrustes (2D/3D), Bookstein (2D), Normalize size (2D/3D),</u> <u>Transform landmarks (2D), Burnaby size removal : diverses opérations</u> <u>destinées aux analyses géométriques</u>

Les analyses géométriques sont utilisées dans le menu « Geomet ».

PAST, logiciel scientifique --19- || 1 Fournir des données à PAST, et les communiquer à d'autres logiciels

1.5.8 Sort ascending, Sort descending, Sort on color : tri des colonnes

A partir d'observations (les lignes) en désordre, PAST peut faire le tri des lignes du tableau selon les valeurs prises par une variable (une colonne).

Si une seule ligne est sélectionnée, il apparaît un message d'erreur, mais qui ne bloque pas le logiciel. Si plusieurs lignes sont sélectionnées, le tri ne porte que sur ces lignes : les autres lignes ne changent pas. Si toutes les lignes sont sélectionnées, le tri porte sur l'ensemble du tableau.

Si une seule colonne est sélectionnée, c'est cette colonne qui sert de base au tri, soit par ordre croissant (« Sort ascending »), soit par ordre décroissant (« Sort descending »). Si plusieurs colonnes sont sélectionnées, c'est la première colonne (colonne de gauche) qui sert de base au tri.

1.5.9 <u>Column difference</u>

Il faut sélectionner deux colonnes, et la différence entre elles est mise dans la colonne suivante. Attention ! les données éventuelles de la colonne suivante sont écrasées.

1.5.10 <u>Regular interpolation : interpolation pour remplacer des données</u> <u>manquantes</u>

1.5.11 Evaluate expression : évaluer une expression mathématique

Les expressions ressemblent un peu à celles qu'on peut employer dans les tableurs, mais le fonctionnement est différent.

Après avoir sélectionné une zone, lorsqu'on choisit cette option, une fenêtre s'ouvre. Vers le bas de la fenêtre, on doit entrer la formule de transformation des cases sélectionnées selon la syntaxe indiquée dans le haut de la fenêtre. On peut soit frapper la formule au clavier, soit cliquer sur les noms des colonnes, des fonctions, des opérateurs ou des valeurs.

Lorsqu'on clique sur « Compute », la valeur des cases sélectionnées est remplacée par les valeurs de l'expression calculée. Si on reclique sur « Compute », un nouveau calcul est effectué. Par exemple, si on a introduit ln(x), en cliquant une première fois, on remplace la valeur dans les cases par leur logarithme népérien. Si on clique une deuxième fois, on la remplace par le logarithme du logarithme.

- évaluation d'une expression : opérateurs +, -, *, /, ^, mod,
- fonctions abs, atan, cos, sin, exp, ln, sqrt, sqr, round, trunc.
- Les variables suivantes peuvent être utilisées : x (contenu de la cellule), l (cellule de gauche « left », sinon 0), r (cellule de droite « right »), u (cellule supérieure « up »), d (cellule inférieure « down »), mean (moyenne de la colonne), min (valeur minimale de la colonne), max (valeur maximale de la colonne), n (nombre de cellules dans la colonne), i (indice de la ligne), j (indice de la colonne), random (nombre au hasard uniforme entre 0 et 1), normal (nombre au hasard normal, moyenne 0 et variance 1), integral (somme de la colonne courante), stdev (écart-type de la colonne), sum (somme totale de la colonne courante). Les valeurs manquantes (?) sont acceptées.
- On peut aussi entrer des formules tenant compte des valeurs des autres colonnes. Les noms des colonnes sont indiqués dans la colonne de gauche de la partie supérieure de la fenêtre : ils commencent par c_. Lorsqu'on n'a sélectionné que certaines cases d'une colonne, le résultat peut paraître bizarre, parce que la formule fait correspondre la première ligne de la

PAST, logiciel scientifique --20- $\parallel 1$ Fournir des données à PAST, et les communiquer à d'autres logiciels

🥙 4lig3col. dat 🛷 Evaluate expression File Edit Transform Plot Statistics Multivar M -Columns Values Functions × (current cell) c_col_1 abs () ~ 🗠 😫 📥 원 🛤 🔃 Edit mode c_col_2 I (left cell, or 0) atan () c_col_3 r (right cell, or 0) cos () col_1 col_3 D col_2 c_D u (cell above, or 0) exp () lig_1 10 50 с_Е с_F 1 d (cell below, or 0) round () lig_2 i (row index) 2 45 c_G j (column index) sin () 0.693147 lig_3 3 30 c_H c_I n (cells in column) sqr () 4 lig_4 19 20 mean (of current column) sqrt () 5 с_J с_К stdev (standard deviation) trunc () min (minimum) c_L max (maximum) Operators c_M random (uniform 0-1) + c_N normal (gauss random) -* c_0 integral (running sum) c_P sum (column total) 1 c_Q ٨ c_R mod c_S (type and/or click in lists)-. Apressi In (c_col_1) Compute Close

zone sélectionnée avec la première ligne de la colonne indiquée dans la formule.

PAST, logiciel scientifique --21- $\parallel 1$ Fournir des données à PAST, et les communiquer à d'autres logiciels

1.6 PAST utilisé sous Linux

PAST est fondamentalement conçu pour être utilisé dans l'environnement MS-Windows, mais l'utilitaire WINE permet de l'utiliser sous Windows. Il fonctionne en particulier très bien sous Poseidon, qui est un « linux scientifique » dérivé de Ubuntu : il suffit de cliquer sur l'icône Past.exe pour que PAST se lance convenablement, puisse bien lire les fichiers, effectuer les calculs et tracer les graphiques.

Les fonctions importantes de PAST fonctionnent correctement, en particulier l'entrée et la sortie de données par lecture et écriture de fichiers, mais aussi par copier/coller, ainsi que la sauvegarde des graphiques dans les formats .emf, .jpg et .bmp.

Par contre, les fonctions qui interagissent trop fortement avec le système d'exploitation ne sont pas disponibles. On ne peut pas copier le graphique dans la mémoire pour le coller directement dans un logiciel de dessin, ni imprimer, ni choisir la fonte de caractères pour les graphiques.

2 <u>Plot : Tracer des</u> <u>graphiques</u>

2.1 <u>Pour tous les graphiques : cliquer pour</u> <u>régler les préférences du graphique</u>

- Pour les différents types de graphes : lorsqu'on double-clique sur le graphe, une fenêtre de réglage apparaît, semblable pour tous les graphiques. Certaines des options ne sont valables que pour certains graphiques particuliers, et il ne sert à rien de les modifier pour les autres graphiques.
- On peut régler l'existence ou non d'une grille ou des axes, l'existence ou non de couleurs (pour les graphiques qui le permettent), l'épaisseur de ligne, la fonte des caractères...



Lorsque le graphique montre des symboles pour les points de mesure, on peut régler la taille de ces symboles, lorsque les axes ont des graduations, on peut régler l'écartement entre ces graduations, mettre des légendes aux axes (« X label » et « Y label »).

On peut sauvegarder le graphique dans les formats emf, jpg et bmp.

- Le format emf est un format vectoriel, ce qui permet d'avoir des traits nets même lorsque le graphique est fortement agrandi. On peut aussi le modifier par des logiciels de dessin vectoriel tels que OpenOfficeDraw, mais ce n'est pas très pratique car les différents constituants du dessin ne sont pas groupés.
- Les formats jpg et bmp sont des formats bitmaps, et la pixellisation devient visible lorsqu'on agrandit l'image. Pour éviter ceci, il faut agrandir le graphique tracé par PAST avant de le sauvegarder (mais, bien sûr, la taille du fichier augrmente).

🖑 CO2_2008. dat	🕘 Graph	_ 🗆 🗙	😸 Graph preferences
File Edit Transform Plot CO2_(ppm) B Plot Plot 1990 358.15 1991 357.22 1992 358.48 1993 358.11 1994 395.83 1995 362.16 1996 364.62 1997 365.59 1998 367.49 1999 370.9 2000 372.25 12 12	CO2_(ppm) 372- 370- 366-	X start: 1 X end: 11 Y start: 354.54 Y end: 373.86 Plot style C Points C Line Points C Line+points C Barchart Log Y 3-pt average ▼ X labels X III	✓ Grid ✓ Colors Font ✓ Axes Point symbols Font ✓ Axes Point symbols Font Symbol size: ★ 1 Width: 326 Vertical font rotation: ★ 90 Height: 320 Maximal tick marks x: ★ 11 Maximal tick marks y: ★ 11 X label:

Dans la fenêtre du graphique lui-même, en bas à droite, trois icônes permettent de fermer la fenêtre, de copier le graphique dans le presse-papier, et d'imprimer le graphique.

PAST, logiciel scientifique --23- $\parallel 2$ Plot : Tracer des graphiques

2.2 Graph : simples graphes de position

Conseil : cocher la case X labels, pour que les titres des lignes soient affichés.

Par défaut, le graphique est tracé en simple ligne (« Line »), mais une série de boutons-radio permet d'avoir un aspect différent : « Points » ne trace que les points de mesure, « Line+points » trace à la fois la ligne et les points, « Barchart »devrait tracer un diagramme en bâtons, mais le résultat est illisible.

Cocher « Log Y » fait une transformation en logarithme décimal pour l'axe vertical, et cocher « 3-pt average » fait un lissage de la courbe sur 3 points.

Par défaut, les axes sont optimisés pour que la courbe ou les points occupent le plus de place. On peut régler le début et la fin de l'axe Y par les cases « Y start » et « Y end ». On peut aussi régler le début et la fin de l'axe X par « X start » et « X end », mais cette option n'a guère d'intérêt ici, contrairement aux graphiques suivants.

2.3 XYgraph : le graphique cartésien traditionnel

La colonne de gauche donne l'axe des abscisses, la colonne de droite donne l'axe des ordonnées. On peut régler l'étendue des axes par les 4 lignes de saisie en haut à droite de la fenêtre.

Les mêmes réglages que pour le graphique précédent sont possibles, avec en plus :

- « 95% ellipses » qui est l'ellipse de confiance des données au seuil 95% (on suppose que les deux variables ont une distribution normale). Le seuil peut être réglé dans la fenêtre des préférences du graphique.
- « Convex hulls »qui trace l'enveloppe de l'ensemble des points de mesure.
- « Min.span tree » est le chemin minimal joignant l'ensemble des points de mesures.



2.4 XY with error bars : graphique XY avec barres d'erreur

Il faut définir quatre colonnes, de la gauche vers la droite : abscisses X, ordonnées Y, erreur en X,

PAST, logiciel scientifique --24– $\parallel 2$ Plot : Tracer des graphiques

erreur en Y.



2.5 <u>Histogram : histogramme de fréquence absolue (nombre</u> <u>d'individus) de la population</u>

La case « Bins » indique le nombre d'intervalles (compartiments) désirés, qui était mis par défaut à une valeur « optimale ». On peut aussi changer le début et la fin des compartiments désirés.

Cocher la case « Fit normal » pour tracer une distribution normale, de paramètres (moyenne et écart-type) identiques à la population dont on vient de tracer l'histogramme.

La case « Kernel density » trace un autre lissage de l'histogramme.



2.6 Box-plot : boites à moustaches = diagrammes de Tukey

Chaque colonne correspond à une « population », et le graphique présente des boites entre les deux quartiles extrêmes, et des moustaches jusqu'au maximum et minimum.



PAST, logiciel scientifique --25- $\parallel 2$ Plot : Tracer des graphiques

Cocher la case « Outlier » provoque un autre mode de traçage des moustaches. Les moustaches vont jusqu'au maximum et au minimum, à condition qu'ils ne soient pas trop écartés des quartiles (la moustache ne doit pas faire plus de 1,5 fois la longueur du segment de boite) ; si des points sont en dehors de cet intervalle, ils sont tracés sous la forme de petits cercles, indiquant que ce sont des valeurs exceptionnelles (points aberrants ?).

2.7 Percentiles : diagramme de fréquences cumulées

2.8 <u>Normal probability plot : graphique montrant (ou non) une</u> <u>distribution normale</u>

🐣 Untitle	Untitled				🖑 Normal probability plot									
File Edit	Transform F bioA 3 16	Plot Statisti bioB 5.5 13		24 -							-		X start: X end: Y start: Y end:	3 -10 30
3 4 5 6 7 8	14 16.5 8.5 9 10.5 8.5	10.5 12.5 10 10.5 11.5 8	Sample values	20 - ·· 16 - 12 8 - 4 -				,	.	•	-		PPCC:	0.9829
9 10 11 12 13 14 15	7.5 3.5 3.5 8.5 9 13	19 12.5 12.5 9 9 14 11.5		-4- -8- -3 -	-2.4 -1. N	8 -1.2 ormal c	-0.6 (orders) 0.6 tatistic r	1.2 1.8 nedians	3 2.4		1	<u>×</u> <u></u>	

C'est une sorte de graphique en XY, où l'axe X correspond à une distribution normale et l'axe Y correspond aux valeurs observées. En rouge est une droite de régression entre ces deux

distributions, et à droite « PCCC » est le coefficient de corrélation linéaire entre ces deux distributions.

Ce type de graphique est utilisé pour répondre aux questions du type :

- Les données suivent-elles une distribution normale ?
- Quelle est la nature de l'écart à la normalité : asymétrie, queues plus courtes ou plus longues que normalement ?

Si les données sont distribuées normalement, les points suivent la droite de régression

Si la distribution a des queues plus grosses que normalement, les points forment une sigmoïde (à droite ils sont plus bas que la droite, à gauche ils sont plus hauts).

Si la distribution a des queues plus maigres que normalement, les points forment une tilde (à droite ils sont plus haut que la droite, à gauche ils sont plus bas).

Si la distribution est décalée vers la droite, les points forment une courbe au dessous de la droite.

Si la distribution est bimodale, la ligne des points noirs fait deux ondulations à travers la droite rouge...

2.9 <u>Ternary : diagramme triangulaire (= diagramme ternaire), pour la</u> <u>composition en 3 facteurs</u>



On peut facilement visualiser les groupes de données en utilisant l'option de coloration des lignes (menu Edit)

2.10 <u>Bubble plot = graphique à bnulles</u>

C'est un simple graphique en XY selon les deux premières colonnes, mais où la taille des points est fonction de la valeur de la troisième colonne. Le curseur « Size » permet de faire varier le diamètre des points pour que l'ensemble reste lisible.



PAST, logiciel scientifique --27- $\parallel 2$ Plot : Tracer des graphiques

Les valeurs négatives ne sont pas tracées, puisqu'une largeur négative d'un disque n'aurait aucun sens. Cliquer sur la case « Substract min » permet d'affecter une valeur 0 au minimum, et des valeurs positives aux autres valeurs.

2.11 <u>Survivorship : courbes de survie d'une population</u>

Fondamentalement, la courbe de survie est une courbe figurant la proportion d'individus vivants en fonction du temps (ou d'une dose de traitement).

D'habitude, on trace le nombre d'individus en Y, en fonction du temps en X.

Ici :

- en abscisse est la taille de la population (nombre d'individus)
- en ordonnées est le temps (nombre de générations, ou nombre de jours, ou nombre d'années;..)

Dans tous les cas, les courbes descendent.



PAST, logiciel scientifique --28– $\parallel 2$ Plot : Tracer des graphiques

2.12 <u>Landmarks : graphique XY avec points de repères par</u> <u>colorations</u>



2.13 <u>Landmarks 3D : grapnique XY avec une troisieme dimension,</u> <u>peu claire</u>

2.14 <u>Matrix : cartographie des valeurs, soit en nuances de gris, soit</u> <u>en couleurs</u>

Cette option est intéressante pour figurer des cartographies (pollutions, concentrations, abondance

d'une espèce, etc). L'axe des abscisses est le numéro de la colonne, l'axe des ordonnées correspond aux lignes, et la valeur de la couleur correspond à la valeur donnée dans chaque case. « Contours » permet de tracer des estimations de valeurs intermédiaires.



2.15 <u>Surface : cartographie en relief</u>

Là aussi, les axes horizontaux X et Y sont les numéros des lignes et des colonnes, et l'axe Z correspond aux valeurs numériques qui sont dans les cases. Cet axe Z est visualisé d'une part par la forme de la surface (relief), d'autre part par les couleurs.



Plot

3 <u>Statistics : effectuer</u> <u>des calculs</u> <u>statistiques</u>

3.1 <u>Univariate : calculs sur chaque colonne</u> indépendamment

N désigne le nombre d'observations : soit la totalité de la colonne, soit le nombre de lignes sélectionnées. En cas de données manquantes, elles doivent être indiquées par le point d'interrogation dans les cases.

Min désigne le minimum de la colonne, Max le maximum, Sum la somme des valeurs, Mean la moyenne, Std. error l'erreur standard de la moyenne (écart-type divisé par la racine du nombre d'observations), Variance la variance, Stand. dev l'écarttype, Median la médiane, 25 prentil et 75 prentil les premier et troisième quartiles, Skewness et Kurtosis les coefficients Statistics Multivar Model Diversity Univariate Similarity and distance indices Correlation Var-covar F and T tests (two samples) F and T from parameters T test (one sample) Paired tests (t, sign, Wilcoxon) Normality (one sample) Chi^2 Coefficient of variation Mann-Whitney Kolmogorov-Smirnov Spearmann/Kendall Contingency table One-way ANOVA Two-way ANOVA Kruskal-Wallis One-way ANCOVA Mixture analysis Genetic sequence stats

🖑 Untitle	d					🕘 Univariate statistics 📃 🗖 🗙					
File Edit ⁻	Fransform P	lot Statistics	; Multivar I	Model Diver	sity 1		bioA	bioB	mathA	mathB	
Cladistics 9	icript					N	43	43	43	43	
👝 🐘	A B	Ba 🔁 🛛	Edit mode	E Edit	Min	1	4	2.5	1.5		
			• Lait mode	j Luit	IGIDEIS	Max	20	19	19.5	20	
	bioA	bioB	mathA	mathB	physic	Sum	402.5	443.5	412	420.5	
1	3	5.5	8	17	9.5	Mean	9.36047	10.314	9.5814	9.77907	
2	16	13	18	13.5	12	Std. error	0.735947	0.490221	0.705721	0.736507	
3	14	10.5	12	12	18	Variance	23.2896	10.3336	21.4158	23.325	
4	16.5	12.5	13.5	16.5	14	Stand, dev	4.82593	3.21459	4.62772	4.8296	
5	8.5	10	19	14.5	16	Median	9	10	8	8.5	
6	9	10.5	11	16.5	19	25 prentil	4	8	6.5	6	
7	10.5	11.5	7.5	2.5	7	75 prentil	13.5	12	12.5	14	
8	8.5	8	3.5	1.5	7	Skewness	0.155484	0.613793	0.671844	0.428792	
9	7.5	19	19.5	20	15	Kurtosis	-0.748051	0.77153	-0.353626	-0.701804	
10	3.5	12.5	6.5	7.5	8.5	Geom. mean	7.83096	9.8224	8.50167	8.49581	
11	3.5	12.5	3.5	8.5	5						
12	8.5	9	6.5	17	13.5		L	-			
13	9 9 4.5 6 8.5		8.5		×	E R	8				
14	13	14	12.5	9	12.5						

d'asymétrie et d'aplatissement par rapport à une distribution normale, Geom. mean la moyenne géométrique.

Le coefficient d'asymétrie est positif si la partie droite de la distribution est plus épaisse.

Le coefficient d'aplatissement (qu'il vaudrait mieux nommer coefficient de pointicité ?) correspond au paramètre statistique $\gamma_2 = \mu_4/\sigma^4$ -3. Il est positif si la courbe de la distribution est pointue. Dans la pratique, on considère que s'il est compris entre -2 et + 2 on peut considérer que la distribution a un aplatissement normal, et que s'il est supérieur à 2, la distribution est plus pointue que normalement, et s'il est inférieur à -2, la distribution est aplatie. PAST, logiciel scientifique $-31 - \parallel 3$ Statistics : effectuer des calculs statistiques

3.1.1 <u>Similarity and distance indices : distances ou corrélations entre les</u> individus

Attention ! D'habitude, on cherche les corrélations entre les variables (les colonnes), à partir d'un certain nombre de mesures (les lignes). Ici, on obtient les corrélations entre les individus, c'est à dire la ressemblance entre les individus pour ce qui est d'un certain nombre de caractères.

Au lieu de valeurs numériques, on peut utiliser des séquences d'acides nucléiques, avec les valeurs A, T, C, G, U (et ? pour les nucléotides absents).

🛷 Untitl	🖑 Untitled					🗢 Similarity and distance indices 📃 🗖 🔀						
File Edit	Transform	Plot Statistic	s Multivar	Model		Arabidopsis	Anabaena	Cyanophora	Porphyra		Index	
Cladistics	Script				Arabidopsis	0	10.463	8.8852	7.5		C Gower	
					Anabaena	10.463	0	8.8976	10		Euclidean	
		i 🛱 🔃	Edit mode	F	Cyanophora	8.8852	8.8976	0	8.4163			
	Δ	в	c	D	Porphyra	7.5	10	8.4163	0		C Dice	
Arabidopsi:	s IC	Т	G	A							C Kulczynski	
Anabaena	C	T	G	A							C Simpson	
Cyanophor	a C	Т	G	A							C Bray-Curtis	
Porphyra	С	Т	G	A							C Morisita	
											C Haup-Crick C Horn C Hamming C Chord Manhattan C Jukes-Cantor C Kimura C Tajima-Nei	

Par défaut, c'est la distance euclidienne qui est calculée, mais il existe une grande variété d'autres calculs possibles, dont le détail est indiqué dans le fichier http://folk.uio.no/ohammer/past/past.pdf.

Ces résultats peuvent servir de base pour la construction de dendrogrammes (« graphiques en arbres de différences ou de ressemblances »), mais il existe d'autres méthodes : voir le menu « Multivar », et en particulier « Multivar | cluster analysis » (4.10).

3.2 Correlation : corrélation entre les variables

Il faut sélectionner plusieurs colonnes d'un tableau où les colonnes correspondent aux variables observées, et les lignes aux individus (lieux, mesures...) où l'on observe ces variables.

Cette fonction fait apparaître une matrice carrée où les cases contiennent les corrélations entre les variables.

3.3 <u>Var-covar : calcul de la matrice de variance-covariance des</u> <u>colonnes</u>

A partir du même tableau que la fonction précédente « Correlation », cette fonction fait apparaître une matrice carrée où les cases diagonales contiennent les variances de chaque variable, et les autres cases contiennent les covariances de chaque couple de variables.

3.4 <u>F and T tests (two samples) : comparaison des variances et des</u> <u>moyennes de 2 échantillons de distributions normales</u>

La structure du tableau est différente des fonctions précédentes. Ici, il n'y a qu'une seule variable par tableau, par exemple la longueur du corps, et deux colonnes contenant les valeurs mesurées. Chaque colonne correspond à un échantillon et les diverses lignes correspondent aux diverses mesures réalisées pour l'échantillon en question. Ce nombre de lignes peut être différent selon les colonnes.

PAST, logiciel scientifique $-32 - \parallel 3$ Statistics : effectuer des calculs statistiques

Pour chaque colonne (chaque échantillon), cette option donne :

- N : le nombre d'individus mesurés
- Mean : la moyenne
- « 95% conf. » : l'intervalle de confiance de la moyenne à 95% (plus le nombre de mesures est grand, plus l'intervalle de confiance de la moyenne se rétrécit : l'estimation de la moyenne est plus précise)
- variance : la variance

Le test F fait la comparaison des variances. « p(same variance) » donne la probabilité que les variances observées sur les échantillons soient identiques : si c'est inférieur à 0,05, on peut conclure que les variances sont différentes, et que les deux échantillons proviennent de deux populations différentes (on suppose que les distributions sont normales).

Le test t fait la comparaison des moyennes. « p(same mean) » donne la probabilité que les moyennes soient identiques (simplement sous l'effet du hasard, à partir d'une même population). Si c'est inférieur à 0,05, on peut conclure que les moyennes sont différentes (que les deux échantillons proviennent de deux populations différentes). Ce test suppose que les deux échantillons ont une distribution normale et une variance semblable

Pour des petites populations, ou pour des populations de distribution non normale, il faut mieux utiliser le test « permutation t test).

1	🛷 Untitle	d			🧶 F and T	tes	ts		
1	File Edit 1	Transform P	lot Statistics	; Multiva					
2	o 🛤			7 5-19-1-1	rien			insuline	
H					N:	4		N:	7
1		rien	insuline	С	Mean:	3.53	75	Mean:	1.7714
1	1	3.18	3.03		95% conf.:	(2.9	092 4.1658)	95% conf	i.: (0.61417 2.9287)
1	2	3.98	2.57		Variance:	0.15	589	Variance	: 1.5657
1	3	3.23	0.57						
1	4	3.76	0.39		95% conf. f	or di	ifference bet	tween means:	(0.28184 3.2503)
1	5		0.37			F:	10.044	p(same variance	e): 0.085303
1	6		2.74			t	2.6917	p(same mean):	0.024728
1	7		2.73		Uneq. var.	t:	3.446	p(same mean):	0.0090925
1	8				Permutati	on t	test	p(same mean):	0.0248
Ż					Permutati	on l	l: 10000	_	
							,		
							X	- Est _	<u>A</u>
-							•••		=

3.5 <u>F and T from parameters : comparaison de deux échantillons à</u> partir des paramètres statistiques

C'est le même principe de comparaison de variance et de moyenne entre deux échantillons, mais dont on ne connaît pas les valeurs mesurées pour chaque individu. Il faut donc entrer à la main les valeurs numériques de la moyenne, de la variance et du nombre d'individus mesurés.

PAST, logiciel scientifique $-33 - \parallel 3$ Statistics : effectuer des calculs statistiques

3.6 <u>T test (one sample) : comparaison d'un échantillon avec une</u> <u>distribution théorique</u>

On sélectionne une colonne (un échantillon), et cette option permet de comparer à une vaste population dont on connaît la moyenne.

Il faut entrer la moyenne dans la ligne de saisie, et cliquer sur « Compute » pour lancer les calculs.

	🛷 Unti	t le d			🖑 One-sa	One-sample t test against theoretical mean							
ł	File Edit	: Transform	Plot Statistic	s				L					
	۵ ا	<u>a</u> 🖪 💈	🖉 📴 🔁 🛛	~	ins	uline	Given mean	L					
				_	Mean:	1.77143	Mean: 2	H					
		rien	insuline	С	95% conf.:	(0.61417 2.9287)		ŀ					
l	1	3.18	3.03		N:	7		L					
i	2	3.98	2.57					L					
ł	3	3.23	0.57			t: -0.4833	p(same mean): 0.646	E					
ł	4	3.76	0.39				• • • • • • • • • • • • • • • • • • • •						
1	5		0.37					Г					
ł	6		2.74			·····							
	7		2.73			Compute	🗙 🙀 📇						
	8												

3.7 <u>Paired tests (t,sign, Wilcoxon) : comparaison d'échantillons</u> <u>appariés</u>

« échantillons appariés » signifie qu'à chaque mesure d'une colonne est associée une mesure de l'autre colonne, sur la même ligne : chaque ligne correspond à un individu (un lieu de mesure, une expérience, un village, etc.). C'est par exemple la comparaison d'animaux avant et après engraissement : c'est le même animal avant et après.



Il faut obligatoirement le même nombre de lignes pour les deux colonnes, mais on peut avoir des données manquantes, à symboliser par « ? ».

« t test » est le test paramétrique de comparaison de moyenne, qui suppose des distributions normales. « p(same mean) » est la probabilité que les deux moyennes soient identiques. « Sign test » et « Wilcoxon test » sont des tests non paramétriques de comparaison de médianes, qui ne supposent pas forcément que les distributions sont normales.

Dans tous les cas, « p(same median) » donne la probabilité que les médianes soient identiques, c'est à dire qu'il n'y ait pas de différence entre les populations étudiées (rappel : la médiane est la valeur pour laquelle la moitié des mesures sont plus grandes, et la moitié des mesures sont plus petites).

3.8 Normality (one sample) : test de normalité d'un échantillon

Pour une colonne sélectionnée, c'est un test de normalité de la distribution.

PAST, logiciel scientifique --34– \parallel 3 Statistics : effectuer des calculs statistiques

🥔 Untitle	d			🕘 Tests for r	iormal d	listribution		
File Edit Cladistics S	Transform P Script	lot Statistic:	s Mul	bioA		N:	43	
🖴 😫	<u>e</u> *		🗸 Edil	Shapiro-Wil	k	W: p(normal):	0.9589 0.1266	
	bioA	bioB	math/					
1	3	5.5	8	Jarque-Bera	1	JB:	1.309	
2	16	13	18			p(normal):	0.5198	
3	14	10.5	12					
4	16.5	12.5	13.5			Chi^2:	4.5349	
5	8.5	10	19			p(normal):	0.033211	
6	9	10.5	11					
7	10.5	11.5	7.5	Upr limit:	6.1271	9.3605	12.594	+inf
8	85	8	3.5	Obs freq:	11	14	5	13
9	7.5	- 19	19.5	Exp freq:	10.75	10.75	10.75	10.75
10	35	12.5	6.5					
11	35	12.5	35			-		
12	8.5	9	6.5		×	<u></u>		

PAST, logiciel scientifique $-35 - \parallel 3$ Statistics : effectuer des calculs statistiques

N est le nombre de cases mesurées, et pour 3 tests différents, p(normal) donne la probabilité que la distribution soit normale.

Vers le bas de la fenêtre, les quatre colonnes correspondent aux quatre quartiles, c'est à dire aux quatre parties de la distribution, normalement équiprobables:

- Upr limit est la valeur supérieure de l'intervalle
- Obs freq est le nombre d'individus observés dans cet intervalle (« fréquence absolue »)
- Exp freq est le nombre d'individus attendus dans cet intervalle.

3.9 Chi^2 : test du khi deux

Ce test permet de comparer deux distributions (dans deux colonnes).

Pour chaque colonne, chaque case contient le nombre d'individus observés dans l'intervalle.

Ce test n'est valable que si les intervalles contiennent au moins cinq individus (sinon, il faut regrouper les intervalles).

p(same) indique la probabilité que les deux échantillons soient identiques (qu'il n'y ait pas de différences entre eux).

Pour le test du khi-deux, on suppose que chaque intervalle contient au moins 5 individus. Cette hypothèse n'est pas nécessaire pour les tests de Fisher (possible seulement si le nombre de lignes est 2) et Monte-Carlo.

-	PAST				🖑 Chi square	
Fil	e Edit 1	Transform P	lot Statist	ics V	complications N1:	33
		complications	non_comp	lic: C	non_complications N2: Degrees freedom:	132 2
fra fr_	icture_ouv _non_ouve	10 23	19 113		Chi^2: p(same):	4.6124 0.09964
3 4					Monte Carlo p(same):	N/A
5 6					Fisher's exact p(same):	N/A
7 8					Sample vs. expected One	constraint
9 10					× 🖻 🛓	

PAST, logiciel scientifique $-36 - \parallel 3$ Statistics : effectuer des calculs statistiques

Il faut cocher « Sample vs. expected » si la seconde colonne est une distribution théorique ; dans ce cas, on peut mettre des valeurs non entières dans cette seconde colonne.

La case « One constraint » doit être cochée s'il existe une contrainte supplémentaire, par exemple lorsqu'on met des pourcentages dans les cases, forcément le total de chaque colonne doit être égal à 100. Au contraire, lorsqu'on compare de nombres d'observations, il ne faut pas cocher cette case.

3.10 <u>coefficient of variation : test d'égalité des coefficients de</u> <u>variation de deux échantillons</u>

Le coefficient de variation vaut l'écart-type divisé par la moyenne de l'échantillon étudié. C'est un paramètre intéressant car il permet de comparer des mesures très hétérogènes, par exemple la taille d'une population d'éléphants et d'une population de souris. Il n'a aucune signification lorsqu'il peut exister des valeurs négatives et que la moyenne peut être nulle.

Cette option calcule les coefficients de variation de deux échantillons, et teste leur égalité.

3.11 <u>Mann-Whitney : test U de comparaison de médianes, même si</u> <u>la distribution n'est pas normale</u>

Il faut normalement que le nombre de mesures soit supérieur à 7, et que la forme des distributions soit semblable.


PAST, logiciel scientifique $-37- \parallel 3$ Statistics : effectuer des calculs statistiques

C'est un test « non paramétrique », testant si la médiane des deux distributions est semblable, utilisable même si les distributions ne sont pas normales (contrairement au test t, qui nécessite la normalité des distributions).

3.12 <u>Kolmogorov Smirnov : teste si deux échantillons ont la même</u> <u>distribution</u>

Après avoir sélectionné deux colonnes, ce test indique si les deux distributions sont différentes. C'est un test non paramétrique : il est utilisable avec des distributions de n'importe quel type (il n'y a pas besoin de supposer que les distributions sont normales).

Des valeurs manquantes sont possibles, en les indiquant par un point d'interrogation.

Si l'on veut simplement tester l'égalité des moyennes, il vaut mieux utiliser le test Mann-Whitney.

3.13 <u>Spearman/Kendall : deux variables sont-elles corrélées ?</u>



C'est aussi un groupe de tests non paramétriques, cherchant à savoir si deux variables sont corrélées. « p(uncorr) » indique la probabilité que les deux variables ne soient pas corrélées.

3.14 <u>Contingency table : test d'indépendance de variables d'une</u> <u>table de contingence</u>

🕘 PAST				
File Edit	Transform Plot	Statistics Multivar	Contingency table i	;i 😑 🕒 🎑
🗠 🛱		Edit mode	M;N:	2;2
_	complications	Dop. complications C	p(no assoc):	0.031742
fracture ous	10	10	Cramer's V:	0.16719
fr_non_ouve	23	113	Contingency C:	0.16491
3				and the second se
4			X	A
5				

Pour un couple de variables nominatives, on met dans les cases les nombres d'occurrences des diverses possibilités. Le test calcule la probabilité qu'il n'y ait pas de lien entre les variables, et l'indique par la valeur « p(no assoc) ».

PAST, logiciel scientifique $-38 - \parallel 3$ Statistics : effectuer des calculs statistiques

3.15 <u>One-way Anova = analyse de variance à un seul facteur</u>

				-	-					
🧶 Untitle	d			🛷 One-wa	y ANOVA					
File Edit	Transform I	Plot Statisti	cs Multivar Mo			Sum of sqrs	df	Mean square	F	p(same)
🗠 😫	A 🤑	🖹 🕄	🔽 Edit mode	Between g	roups:	2900	2	1450	76.73	2.616E-12
				Within g	roups:	548	29	18.8966		
	aliment_A	aliment_B	aliment_C D		Total:	3448	31			
1	180	199	191							
2	177	203	194	Levene's t	est for ho	mogeneity of	variance, b	ased on mean	s: p(san	ne) = 0.9936
3	175	200	201				Ba	sed on median	s: p(sam	ie) = 0.9632
4	170	194	193	Welch F te	st in the d	ase of unequa	al variances	s: F=73.89. df=1	18.99, p=1	1.096E-09
5	182	195	197			•				
6	181	204	195			Tukey's pair	wise comp	oarisons:		
7	177	206	203			Q	\p(same)			
8	180	207	199	1						
9	183	202	199		aliment_A	aliment_B	aliment_C			
10	185	200	201	aliment_A		0.0001229	0.0001229			
11			206	aliment_B	16.47		0.2669			
12			197	aliment_C	14.22	2.246				
13				P						
14						×		B		
15						_	 1			

Fondamentalement, c'est une comparaison multiple de moyennes : les moyennes sont-elles différentes les unes des autres, ou bien les fluctuations entre les moyennes sont-elles simplement dues au hasard ? Il faut que les distributions des échantillons soient normales, que leurs effectifs et leurs variances soient semblables.

Les diverses colonnes correspondent aux échantillons étudiés (ici 3 : effet de 3 types d'aliments, A, B et C, sur le poids de poulets).

L'appel de cette fonction aboutit à une fenêtre :

- à gauche, la somme des carrés entre les groupes (intergroupe) et à l'intérieur des groupes (intragroupe), et le total (qui est la somme des deux précédentes).
- à droite, la probabilité de l'hypothèse nulle (les différences entre moyennes sont dues au hasard).
- Le test de Levene donne la probabilité que les variances soient homogènes, ce qui est nécessaire pour le test F précédent. Si les variances apparaissent différentes, il faut utiliser une autre méthode d'analyse de variance (Welch).

Une fois que l'analyse de variance a montré que les groupes étaient différents pour le caractère considéré, il reste à trouver pour quel(s) groupe(s) les différences sont importantes.

Dans le triangle supérieur sont indiquées les probabilités que les échantillons aient une moyenne égale : A est franchement différent de B et C (car la probabilité d'égalité est très faible), alors que B et C ne sont pas très différents (probabilité d'égalité 0,2669).

PAST, logiciel scientifique -39- || 3 Statistics : effectuer des calculs statistiques

3.16 Two-way ANOVA : analyse de variance à deux facteurs 🧶 Untitled _ [0 File Edit Transform Plot Statistics Multivar Model Diversity Time Geomet Strat Cladistics Script 욢 😫 🗠 , 😫 👯 Edit mode Edit labels Square mode traitement_K traitement_N masse b E F G Н h 1.1 lκ 🛷 Two-way ANOVA 2 3 F Sum of sars đf Mean square р 4 Factor A: 3.55E04 2 1.775E04 3530 2.285E-30 5 4595 304.6 314E-19 Eactor B: 3 1532 6 Interaction: 70.5 6 11.75 2.337 0.06414 7 236 Within: 120.7 24 5.028 8 Total: 4.028E04 35 9 10 246 11 4 244 ₿<mark>₽</mark> ₽ X 12 248 4 13 1 14 15

De même que pour l'analyse de variance à un facteur, il faut que les distributions soient normales, et que les variances et effectifs soient semblables.

Une première colonne contient les modalités du premier facteur (noté facteur A), codées de 1 à n. La deuxième colonne contient les modalités du deuxième facteur (noté facteur B), codées de 1 à m.

Cette option donne l'effet du facteur A, du facteur B et de l'interaction entre les deux facteurs.

La dernière colonne donne la probabilité que cet effet soit simplement au hasard. Ici, la probabilité que les différences soient dues au hasard est très faible : il est quasiment certain que le facteur A a un effet, très fortement probable que le facteur B a un effet, et assez probable que l'interaction entre les deux existe (puisque la probabilité que l'interaction soit due au hasard est de 0,06414.

3.17 <u>Kruskal-Wallis : comparaison multiple de médianes, par une</u> <u>sorte d'analyse de variance où les distributions ne sont pas</u> <u>forcément normales</u>

C'est une alternative non paramétrique à l'analyse de variance, où on compare les médianes de quelques groupes univariés (dans les colonnes). On ne suppose pas que les distributions sont normales, mais il faut que les distributions aient la même forme entre les divers groupes.

	🛷 Untitle	d								
1	File Edit 1	Fransform F	lot Statistic	s Multivar	Mod	el Diversity	Time Geo	met Strat	Cladistics Sci	ript
	😫 🗠	A	😫 💽	Edit mode		🛷 Kruska	l-Wallis te	est		
a		aliment_A	aliment_B	aliment_C	D			и.	24.2	
	1	180	199	191				H: Hat	21.2	
ł	2	177	203	194				nc:	21.20	
H	3	175	200	201				p(same).	2.4320-03	
H	4	170	194	193						
	5	182	195	197			Mann-W	hitney pairw	vise compari	sons,
H	6	181	204	195			Bonferr	oni correct	ed \ uncorrec	ted:
H	7	177	206	203						
ł	8	180	207	199			aliment A	aliment B	aliment C	
ł	9	183	202	199		aliment A		0.0001827	8.734E-05	
	10	185	200	201		aliment B	0.000548		0.1213	
	11			206		aliment C	0.000262	0.3638		_
Н	12			197						
-	13									
-	14									
-	15									
-	16									
-	17						×	Ŀ	*	8

PAST, logiciel scientifique --40- \parallel 3 Statistics : effectuer des calculs statistiques

Là aussi, « p(same) » indique la probabilité que les distributions soient les mêmes, et la matrice contient les probabilités que les couples de distributions soient identiques.

3.18 One-way ANCOVA : analyse de covariance à un facteur



Cette méthode étudie l'effet d'une variable qualitative (facteur) et d'une variable continue (covariable) sur une variable réponse.

On doit avoir quelques paires de colonnes correspondant aux différentes modalités de la variable qualitative. Chaque paire de colonne correspond à un ensemble de données plus ou moins corrélées (ici 3 paires).

L'analyse de covariance indique si les pentes peuvent être considérées comme différentes ou non, en plus de la comparaison multiple de moyennes.

Ici, la probabilité que les moyennes soient identiques pour les trois groupes est faible (0,00447), on peut donc considérer que les moyennes sont différentes.

Par contre, la probabilité que les pentes soient homogène est assez forte (0,2123): on ne peut donc pas considérer que les pentes sont différentes.

Le bouton « View groups » fait apparaître une fenêtre supplémentaire, où pour chaque variable dépendante (la 2e colonne de chaque paire), la moyenne calculée, la moyenne calculée par ajustement pour l'ensemble des valeurs, et la pente de la droite de régression sont affichées.

3.19 <u>Mixture analysis = analyse des mélanges, pour une population</u> <u>hétérogène</u>

C'est une méthode utilisée lorsqu'on a en mélange deux ou quelques populations, par exemple

PAST, logiciel scientifique $-41 - \parallel 3$ Statistics : effectuer des calculs statistiques

lorsqu'on a un mélange de mâles et de femelles. Chaque population est à peu près normale, mais le mélange des diverses populations aboutit à une distribution non-normale, par exemple avec deux ou plusieurs modes.



File Edit	Transform P	lot Statistics	: Multivar	Model Diver	sity Time	Contente sequence st		-
		- 1 - 1			,	Total length	100	
- 🗁 📴	8 %	📴 🔁 🛛	Edit mode	🔲 Edit	labels	Average gap	3	
			0	lo.	le.	Average A	27	
0	A	-			C	Average T or U	21.25	
Arabidopsis	с -	-	G -	A	с -	Average C	19.25	
Anabaena	C	<u> </u>	G	A	C	Average G	29.5	
Cyanophora	C	<u>T</u>	G	A	C	Average p dist	0.2284	
Porphyra	с	Т	G	A	С	Av. Jukes-Cantor d	0.281	
5						Max Jukes-Cantor d	0.4665	
						Av. transitions (P)	0.1255	
						Av. transversions (Q)	0.1028	
						R=P/Q	1.221	
						· · · ·		
						×		

Cette analyse statistique n'est possible que pour des séquences de nucléotides codés ATCG, ou 1234.

Elle donne :

3.20

- la longueur totale de la séquence analysée,
- le nombre moyen de données manquantes pour chaque ligne,
- la quantité moyenne de chaque nucléotide,
- les distances moyennes entre chaînes selon deux méthodes,
- le nombre moyen de transitions (remplacement d'une base purique par une autre, ou d'une base pyrimidique par une autre) et de transversions (remplacement d'une base purique par une base pyrimidique, ou inversement). Normalement, les transitions sont plus fréquentes que les transversions.
- le rapport transition/transversion.

PAST, logiciel scientifique --42- \parallel 3 Statistics : effectuer des calculs statistiques

4 <u>Multivar : statistiques</u> <u>multivariables</u>

Ce menu regroupe diverses fonctions permettant de manipuler des tableaux avec de nombreuses variables, et d'en extraire l'information sous une forme plus simple. D'une façon générale, les colonnes correspondent aux diverses variables, et les lignes aux différents individus (lieux, points de mesure, etc.) où ont été mesurées ces variables.

Les deux méthodes les plus importantes à connaître sont l'analyse en composantes principales, pour les données de mesures, et l'analyse des correspondances, pour les données de comptages.

4.1 <u>Principal components analysis : Analyse en</u> composantes principales (ACP = PCA)

C'est une technique qui permet de récapituler en deux ou quelques variables synthétiques, les « composantes principales », la majorité de l'information contenue dans les variables initiales

L'exemple ci-dessus correspond aux notes de 12 élèves dans quatre matières : Math, Français, histoire-géo, physique. Ces quatre matières sont résumées par quatre composantes, mais ce sont les deux ou trois premières qui sont les plus intéressantes, car elles possèdent la plus grande partie de l'information initiale.

Intit 🆑	led									
File Edit	Transform	Plot Statist	tics Multivar	Model Div	ersity	\delta Princi	pal Compoi	nents		
<u>e</u> 🗄			Edit mode	Ec	lit label	PC	Eigenvalue	% variance		Matrix
	Maths	Français	histoire-géo	physique	E	1	63.6716	68.342		Var-covar
A	10	5	8	10		2	20.2721	21.759		C Correlation
в	15	4	4	14		3	9.06341	9.7282	_	SVD
С	8	9	4	8		4	0.159503	0.1712		• 010
D	9	15	12	9						📄 Shape PCA
E	19	12	13	18						
F	2	8	15	3						Boot N: 0
G	14	10	14	14						
Н	5	8	6	6						Jolliffe cut-off:
I	4	3	7	5						16.304
J	8	12	3	8						
к	2	4	2	3						View scatter
L	1	3	4	4						View loadings
13										-
										Scree plot
										Shape deform (2D)
										× 🛤 占

Le déclenchement de l'option ouvre une fenêtre donnant en particulier le pourcentage de variance initiale expliqué par chaque composante. La première donne la majorité de l'information (68,342%), la seconde donne aussi un peu d'information (20,272%), mais les suivantes beaucoup moins. « Eigenvalue » est la valeur propre de la composante en question.

Techniquement, l'analyse en composantes principales peut se faire soit à partir de la matrice des variances-covariances (option par défaut), soit à partir de la matrice des corrélations entre variables (option à cocher, à choisir si les variables sont dans des unités très différentes).

Multivar Model Diversity Time

Principal components Principal coordinates Non-metric MDS Correspondence Detrended correspondence Canonical correspondence CABFAC factor analysis Two-block PLS Seriation

Cluster analysis Neighbour joining K-means clustering

Multivariate normality Discriminant/Hotelling Paired Hotelling Two-group permutation Box's M MANOVA/CVA One-way ANOSIM Two-way ANOSIM One-way NPMANOVA Mantel test SIMPER

Calibration from CABFAC Calibration from optima Modern Analog Technique « Jolliffe cut-off » donne la valeur-seuil de la valeur propre à partir de laquelle on peut considérer que la composante a de l'importance. Ici, le seuil est 16,304, ce qui signifie que les composantes 3 et 4 peuvent être négligées, et que l'on ne gardera que les composantes 1 et 2.

En mettant un nombre supérieur à 1 dans la ligne « Boot N », on obtient deux colonnes supplémentaires, indiquant un intervalle de confiance pour les valeurs propres. Les cases « SVD » et « Shape PCA » ne sont actives que dans ce cas.



Le bouton « View scatter » provoque l'apparition de graphiques montrant les individus (et éventuellement les variables, en cochant « Biplot »), dans un couple de composantes principales (axes 1-2, ou 1-3, ou 2-3, mais dans la pratique, c'est le couple 1-2 qui contient le maximum d'informations). Le bouton « View numbers » provoque l'apparition des coordonnées des individus dans les quatre composantes principales.

Le bouton « View loadings » provoque l'apparition d'une fenêtre montrant le poids des diverses variables initiales pour la détermination d'une composante principale.

« Scree plot » trace un graphique des valeurs propres pour toutes les composantes principales.

Si des groupes de coloration ont été définis précédemment (« Edit | Row color/symbol »), ces couleurs sont visibles sur le graphique obtenu par « View scatter ». Ceci peut permettre de visualiser divers groupes dans diverses régions de ce graphe.

4.2 <u>Principal coordinates : analyse en coordonnées principales, proche</u> <u>de l'ACP</u>

C'est une extension de l'analyse en composantes principales, qui permet d'utiliser des distances variées, autres que la simple distance euclidienne.



PAST, logiciel scientifique --45– \parallel 4 Multivar : statistiques multivariables

A partir des mesures, le logiciel trouve les valeurs propres de la matrice de distance entre les points. Si les lignes ont été colorées (menu « Edit »), les différents groupes de couleurs sont visibles sur le graphique

4.3 Non-metric MDS : positionnement multidimensionnel non métrique

Cette technique part aussi d'une matrice de similarités entre les individus. Après avoir choisi cette option, il faut choisir une méthode de calcul des similarités. Après un petit moment d'attente, le graphe est tracé, où les divers individus sont positionnés selon deux axes..

Le bouton « Shepard plot » ouvre une nouvelle fenêtre :plus les points sont alignés, meilleure est l'analyse.

4.4 <u>Correspondence = analyse des correspondances = analyse</u> <u>factorielle des correspondances (CA = AFC)</u>

C'est une méthode assez semblable à l'analyse en composantes principales, mais pour des valeurs de comptages.

On part d'un tableau de contingence, c'est à dire d'un tableau de nombres (entiers, en principe), où les lignes représentent les modalités d'un facteur, et les colonnes les modalités d'un autre facteur. Le contenu des cases correspond au nombres d'occurrences de la combinaison des deux facteurs.

Un exemple classique est en écologie végétale : n espèces de plantes ont été observées dans m lieux. On obtient ainsi une matrice des n x m, et l'analyse des correspondances permet d'associer certaines espèces à certains lieux.

🧶 PAST							
File Edit		ilot Statistic	s Multivar	e 🗖 Edit	rsity : labels	Ime Geomet Strat Cladistics Script	
R1 R2 R3 R4 Corre	P1 28 36 21 79 spondence	P2 14 21 64 42 Analysis	P3 45 25 38 67	P4 33 64 11 9	P5 12 23 7 41	Correspondence Analysis scatter diagram	X start: -0.5 X end: 0.8 Y start: -0.4
Axis 1 2 3	Eigenvalue 0.155228 0.0803021 0.017183	% of total 61.425 31.776 6.7994		View scatter Relay plot		$\begin{array}{c} 0.4 \\ 0.3 \\ 0.2 \\ 0.1 \\ 0.1 \\ 0.1 \\ 0.2 \\ 0.1 \\ 0.1 \\ 0.2 \\ 0.1 \\ 0.1 \\ 0.2 \\ 0.3 \\ 0.4 \\$	Y end: 0.5 ▼ Rows (samples) ▼ Row labels 95% ellipses Convex hulls ▼ Column dots ▼ Column labels ■ Axes 2+3 View numbers X

PAST, logiciel scientifique --46– \parallel 4 Multivar : statistiques multivariables

4.5 <u>Detrended correspondance analysis = analyse des</u> <u>correspondances redressée</u>

C'est une variante de l'analyse des correspondances.

4.6 <u>Canonical correspondance analysis = analyse canonique des</u> <u>correspondances</u>

Elle permet par exemple d'étudier les facteurs écologiques qui expliquent les répartitions des espèces vivantes.

4.7 CABFAC factor analysis

C'est une méthode factorielle d'analyse des données de comptage, éventuellement en association avec des données environnementales.

🕭 graines. d	lat					-										
File Edit Tra	insform Pl	ot Statistics	Multivar	Model Div	versity Time (Geomet St	۷/	RIMAX fa	ctors		ter p	lot			(
۵ 😫 🗠	9	<table-of-contents> 🖪</table-of-contents>	ZEdit mode	E E	dit labels	🥅 Square									X start:	0.1
ea	au	protides	glucides	lipides	sels_minérau	F]								Y	4
Blé,_Seigle,C	1	12	70	2	2			0.2	0.3	0.4	0.5	0.6 0	.7 0.8	0.9	x ena:	P
Maïs 12		11	70	5	2		^{0.1} T							Blé, Seigle,Or	Y start:	-1
Pois,_Fève,_14		25	55	2	4		12-						P	ois,_Feve,_Haricot	¥	0.4
Lupin 12	2	40	40	5	3										r ena:	-0.1
NOIX,_RICIN 7		20	10	60	3		D.3-					Lupin			Row	labels
	🧶 CABE	AC Factor	Analysis				D.4 -								Min.	span tree
	PC.	Figenvalu	a 🤷 yariar	vce											□ 95%	ellipses
	1	3.8946	77.89	Vari	imax factors:	4	0.54 g								Conv	ex hulls
	2	0.89245	17.85				D.82-									
	3	0.21159	4.23													
	4	0.001194	8 0.02				D.7 -									
	5	0.000140	51 0.00		View factors		D.8 -								Factors	
					View scores										1+2	
		_		_	Uiouu vo sonstruust	Hinn	D.9-								0 1+3	
					view reconstruct	uon	-1	Noix R	icin						0 2+3	
		-		Si	ave transfer fun	iction		14014 [1							View r	umbers
											Fact	or 1				
				>	(🛤	A									×	R
				-												

PAST, logiciel scientifique --47– \parallel 4 Multivar : statistiques multivariables

La première colonne peut contenir des données environnementales numériques, par exemple la température. Dans ce cas, on peut sauver dans un fichier une « fonction de transfert » indiquant en particulier les coefficients de régression pour la première colonne.

4.8 two blocks PLS : Moindre carrés partiels sur deux blocs

Les colonnes doivent être constituées de données continues (mesures), d'abord le premier bloc, puis le deuxième bloc, par exemple d'abord les données morphométriques, puis ensuite les données environnementales.

C'est une méthode comparable à l'analyse en composantes principales, mais qui étudie la covariation entre les deux blocs de variables.

4.9 Seriation

C'est une méthode qui utilise une matrice de présence/absence, avec par exemple les espèces en lignes et les milieux en colonnes. Elle essaie de réorganiser les lignes et les colonnes de façon à concentrer les présences autour de la diagonale.

🛷 Untitle	d					🕘 Seriatio	on						
File Edit	Transform F	lot Statistic	s Multivar	Model Diver	rsity Time G Iabels [6	m o n	p I a	p r o	r i v	v a s	Criterion: 1 <i>Monte Carlo</i>	
	montagne	plaine	profondeur	rivage	vase f		t	i	f	a	e	Mean: 1	
algue cactus	0 1	0 0	1	0	0	peuplier						Z score: 0)
nenuphar	0	0	0	0	1	alque roseau					_	p(random): 1	l
peuplier roseau	0 0	1 0	0	0	0	nenuphar						Constraints	
6												Constrain	ned
7 8												C Unconstr	rained
9													
10 11												× 🖻	
12													

Il y a deux méthodes d'optimisation :

PAST, logiciel scientifique -48 - || 4 Multivar : statistiques multivariables

- « constrained » = optimisation contrainte : seules les lignes sont mobiles, et les colonnes doivent rester fixes
- « unconstrained » = optimisation non contrainte : lignes et colonnes sont mobiles, de façon à concentrer les présences encore plus autour de la diagonale.

4.10 <u>Cluster analysis = Regroupements en arbres</u>

Cette analyse réalise des dendrogrammes, ou graphiques en arbres : les observations sont reliées entre elles par des branches, d'autant plus courtes que les observations sont semblables. C'est en particulier très utilisé en biologie de l'évolution, pour reconstituer des arbres phylogénétiques.



Chaque ligne correspond à un individu (ou une espèce, ou un lieu de mesure...), et chaque colonne correspond à un caractère, pouvant exister sous plusieurs modalités ou plusieurs valeurs.

On peut aussi partir d'une matrice carrée de différences ou de ressemblances :



PAST, logiciel scientifique $-49- \parallel 4$ Multivar : statistiques multivariables

Cette option regroupe les individus par leur ressemblance (selon diverses méthodes possibles).

Cocher « Two-way » fait apparaître deux arbres : d'une part l'arbre des colonnes tracé grâce aux lignes, d'autre part l'arbre des lignes tracé grâce aux colonnes.

4.11 <u>Neighbour joining</u>

C'est aussi une méthode de regroupement en arbres.

Si l'on peut définir un extragroupe, il faut le mettre en première ligne pour qu'il soit considéré comme tel lorsqu'on coche « outgroup » au lieu de « final branch ».



PAST, logiciel scientifique --50- $\parallel 4$ Multivar : statistiques multivariables

4.12 <u>K-means clustering</u>

Sequence	ce_ag	K-means clu	stering		×	Ĭ	🧶 K-means	clu 🔳 🗖	×	Script		_	
و 😭 🗠	A	Enter number	of desired o	lusters	_	a	ltem Anabaena	Cluster 1		Seripe			
Apabaapa	B	,-					Porphyra Arabidopsis	2			J	ĸ	L
Porphyra	T		ж	Cancel		1.5	Cyanophora	1		9 3	G	A	A A
Arabidopsis	T	6	A			4				9	G	A	A
Cyanophora 5	T	G	A	c		Ċ				3	G	A	A

Par des méthodes voisines des précédentes, les individus sont regroupés en paquets (« clusters »). On indique le nombre de paquets souhaités, et le logiciel range tous les individus semblables dans un paquet, selon leur proximité d'après les variables étudiées.

4.13 <u>Multivariate normality = test de normalité sur plusieurs</u> <u>variables</u>

Les données sont un ensemble de colonnes contenant des valeurs numériques (continues, comme pour une analyse en composantes principales).

	concours0								
File Edit	Transform P	lot Statistics	s Multivar	Model Divers	🛷 Multivariat	e normality test (M	ardia)		
🖴 😫	A	📴 🖪	Edit mode	🔲 Edit la	<u>Multivariate</u>	Coefficient	<u>Statistic</u>	df	<u>p(normal)</u>
	bioA	bioB	mathA	mathB I	Skewness:	2.807	20.12	20	0.4507
1	3	5.5	8	17	Skownooo o	mall			
2	16	13	18	13.5	skewness, si	rran sta di	00.40		0 0000
3	14	10.5	12	12	sample corre	ctea:	22.12	20	0.3339
4	16.5	12.5	13.5	16.5	Kurtosis:	21.06	-1.392		0.164
5	8.5	10	19	14.5					
6	9	10.5	11	16.5	Doornik and H	lansen omnibus			
7	10.5	11.5	7.5	2.5	En	13.89			
8	8.5	8	3.5	1.5	сp.	10.00			
9	7.5	19	19.5	20	p(normal):	0.08459			
10	3.5	12.5	6.5	7.5			n. l		
11	3.5	12.5	3.5	8.5		×			
12	8.5	9	6.5	17					

Cette option teste la normalité de l'ensemble des variables. La partie supérieure teste l'asymétrie (Skewness) et l'aplatissement (Kurtosis) des distributions ; p(normal) indique la probabilité de la normalité (asymétrie et aplatissement nuls).

La partie inférieure est un test de normalité.

4.14 <u>Discriminant / Hotelling = Analyse discriminante , et test T² de</u> <u>Hotelling</u>

On suppose des variables numériques continues, pour des individus (points de mesure) marqués de différentes couleurs (menu « Edit »), de distribution normale, et de covariances égales.

Ce test indique si les deux groupes doivent être considérés comme différents ou non.

PAST, logiciel scientifique --51- || 4 Multivar : statistiques multivariables

4.15 Paired Hotelling = test T² de Hotelling pour données appariées

C'est l'équivalent du précédent, mais où les individus des deux groupes sont appariés. Il faut que les individus de chaque groupe soient consécutifs. Le premier individu du premier groupe est considéré comme apparié au premier du deuxième groupe, le deuxième apparié au deuxième, etc.

4.16 <u>Two-group permutation : test de permutation pour deux</u> groupes multivariés

Là encore, il faut deux groupes de données multivariées (valeurs numériques continues, comme pour une ACP), marqués par deux couleurs différentes. Les lignes de chaque groupe doivent être consécutives.

Ce test est comparable au test de Hotelling : il indique si les deux groupes peuvent être considérés comme équivalents, ayant des moyennes égales.

4.17 <u>Box's M = test d'égalité des matrices de covariances pour</u> <u>deux groupes de données</u>

Comme pour les tests précédents, on utilise des données réparties en deux groupes, de couleurs différentes (menu « Edit »). Ce test indique si les matrices de variances-covariances des variables peuvent être considérées comme égales pour les deux groupes.

4.18 <u>MANOVA/CVA = analyse de variance multiple / analyse</u> <u>canonique des variables</u>

Là encore, il faut deux ou plusieurs groupes d'individus (points de mesure) marqués de couleurs différentes ; il faut que le nombre de ces points de mesures soit supérieur au nombre des variables.

4.19 <u>One-way ANOSIM = analyse de similarités</u>

Il faut encore deux ou plusieurs groupes de données multivariées, où chaque groupe est caractérisé par une couleur différente. Cette option teste les différences entre les groupes, qui sont supposés avoir même médiane et étendue (c'est un test non paramétrique : on ne suppose pas que les distributions sont normales). Il est fondé sur l'analyse des distances entre les groupes par rapport aux distances à l'intérieur des groupes.

4.20 <u>Two-way ANOSIM = analyse de similarités à deux facteurs</u>

C'est le même principe que l'analyse des similarités à un facteur (option précédente), mais pour deux facteurs.

Les données commencent par deux colonnes, correspondant aux deux facteurs codés par des entiers. Les colonnes suivantes sont les données multivariées.

4.21 <u>One-way NPMANOVA = analyse de variance multiple non</u> paramétrique à un facteur

Il faut des données groupées en deux ou quelques groupes, indiqués par des couleurs différentes (menu « Edit | Row color/symbol »), où les groupes ont des distributions semblables.

4.22 <u>Mantel test : test de corrélation entre matrices de distance</u>

Il faut deux groupes de données multivariées, avec des couleurs différentes, ou bien deux matrices symétriques de distances ou de similarités.

Il teste les corrélations entre les matrices de distances

PAST, logiciel scientifique -52- || 4 Multivar : statistiques multivariables

4.23 <u>SIMPER = pourcentage de similarité</u>

Il faut un tableau de deux ou plusieurs groupes d'individus, marqués par des couleurs, et des variables de comptage. On suppose que les groupes sont indépendants.

4.24 Calibration from CABFAC

Normalement, si on a sauvé dans un fichier une fonction de transfert lors de l'opération « CABFAC factor analysis », on doit pouvoir l'utiliser pour prédire la variable d'environnement à partir de données d'observation.

4.25 <u>Calibration from optima</u>

On part d'une matrice d'abondance, avec les lignes correspondant aux échantillons, et les colonnes correspondant aux taxons, mais avec les trois premières lignes particulières :

- première ligne : l'optimum du taxon
- deuxième ligne : les tolérances du taxon
- troisième ligne : abondance maximale du taxon
- (et lignes suivantes : comptages d'abondance du taxon dans les divers lieux)

Les trois premières lignes peuvent être obtenues par le menu « Model », où il existe une option « Abundance » et une « Species packing ».

4.26 <u>Modern Analogue Technique</u>

Là aussi, on part d'une matrice d'abondance, avec les taxons en colonnes et les observations en lignes. Les premières lignes correspondent aux taxons actuels, dont on peut connaître les facteurs écologiques, et les dernières lignes correspondent aux taxons fossiles.

La première colonne contient des données environnementales pour les taxons actuels, et un point d'interrogation pour les taxons fossiles.

5 Model : modélisation

A partir de valeurs numériques d'observation, on essaie de calculer une formule de prédiction.

5.1 Linear : régression linéaire

La régression linéaire est le modèle le plus simple.

Attention ! Ici, par défaut, ce n'est pas la régression standard (où la régression de y en x est d'autant plus différente de la régression de x en y que le nuage est étalé), mais la « Reduced Major Axis » (où la régression de x en y et la régression de y en x sont équivalentes) qui sert de base aux calculs. En d'autres termes, la droite rouge correspond au grand



axe de symétrie du nuage de points.

Pour obtenir la régression standard, qui est la plus utilisée en statistiques, décocher la case Red major axis. Dans le cas de cette régression standard, on peut tracer le domaine de confiance à 95% de la droite de régression, alors que ce n'est pas possible avec la régression RMA.

Slope a désigne la pente de la droite, Interc b désigne l'ordonnée à l'origine (différents selon le mode de régression).

r est le coefficient de corrélation linéaire, et p(uncorr) est la probabilité qu'il n'y ait pas de corrélation entre les données (ces deux valeurs sont identiques pour la régression standard et la régression RMA).

Cocher la case « Log-log » provoque le traçage du graphe ln(y) = f(ln(x))



Ne pas oublier qu'une exponentielle peut être modélisée par régression linéaire après transformation des valeurs Y en logarithme.



5.2 <u>Linear 1 indep, n dep = régression linéaire multiple pour une</u> <u>variable (indépendante) qui détermine n variables dépendantes</u>

C'est très semblable à l'option précédente. La valeur indépendante (figurée par l'axe X) est la première colonne, et toutes les autres colonnes sont les variables dépendantes, figurées par l'axe Y. On peut passer de la figuration d'une variable dépendante à une autre par une paire de boutons, au dessus de « Overall MANOVA ».

Les valeurs de pente (Slope a) et d'ordonnée à l'origine (Interc b) sont calculées pour une régression standard (non RMA).

5.3 <u>Linear n indep, 1 dep = régression linéaire multiple, pour expliquer</u> <u>une seule variable (dépendante) par n variables (indépendantes)</u>

La variable dépendante doit être mise en première colonne, et toutes les autres variables (les autres colonnes) sont les variables explicatives (indépendantes).

🧶 taille_p	oids_imc.	dat			🛷 Linear , n	independ	ent, 1 dep. (n	nultiple regr	ess 💶 🗖 🔀
File Edit	Transform P	lot Statistic	s Multivar	Model	Dependent N:	variable: p	oids 4	AN F: 4	NOVA 1497.4
	poids	IMC	taille	D	Multiple R:	C	.99828	p: 6	.3978E-39
1	65.6	21.67	1.74		Multiple R2	: 0	.99657	-	
2	71.8	23.36	1.75		Multiple R2	adj: 0	.99634		
3	80.7	21.55	1.94						
5	72.6	20.87	1.87						
9	78.8	22.49	1.87						
6	74.8	22.71	1.82			Coeff.	Std.err.	t	Р
7	62	20.24	1.75		Constant	-133.32	2.1475	-62.083	4.309E-34
8	81.6	24.1	1.84		IMC	2.9759	0.05167	57.593	4.3316E-33
9	76.6	23.64	1.8		taille	77.353	1.0546	73.346	2.5387E-36
10	83.6	26.44	1.78						
11	74.6	24.08	1.76						
12	76.4	23.5	1.8						
13	63.2	23.36	1.65						
14	60.9	20.35	1.73			X	E	<u>a</u>	
15	74.8	22.21	1.84						

PAST, logiciel scientifique --55- \parallel 5 Model : modélisation

Il n'y a pas de graphique, mais seulement un tableau donnant les valeurs numériques des coefficients à appliquer aux variables indépendantes pour prédire la variable dépendante, à ajouter à un terme constant (l'ordonnée à l'origine).



5.4 Sinusoidal = ajustement sinusoïdal de phénomènes périodiques

La fenêtre donne l'équation des sinusoïdes dont la somme aboutirait à expliquer les points mesurés.

- en bas, la constante (valeur moyenne)
- vers le milieu, les diverses sinusoïdes dont la somme pourrait expliquer les points observés. Par défaut, seule la première est cochée, mais pour qu'elle soit bien ajustée, il faut cliquer sur « Search » pour que le logiciel cherche automatiquement la bonne fréquence. « Amp » indique l'amplitude (la moitié de l'intervalle entre minimum et maximum) et « Phas » indique la phase.

Pour une sinusoïde d'ordre 1, l'équation est de la forme

 $y = C + a \cos(2\pi(x - x0)/T - \varphi)$

où C est la constante, a est l'amplitude, ϕ est la phase, T est la période, et x0 est la première valeur

de la série.

5.5 <u>Polynomial = ajustement polynomial</u>

L'équation du polynôme (par défaut de degré 2) est donnée par $y = a0 + a1*x + a2*x^2$.



On peut modifier la précision de l'ajustement (l'ordre du polynôme) par les deux boutons, au milieu de la zone de droite.

5.6 <u>Logistic = ajustement à la courbe logistique (croissance de</u> <u>populations)</u>



La courbe logistique est du type : $y = a/(1 + b e^{-cx})$.

En choisissant l'option « Bertalanffy », l'équation cherchée est du type : $y = a (1 - b e^{-cx})$.



5.7 <u>Smoothing spline = lissage par les courbes spline</u>

Ce n'est pas vraiment une méthode de modélisation, c'est une méthode de lissage, pour éliminer les bruits parasites des données brutes. Des points un peu trop dispersés sous l'effet de facteurs aléatoires soient résumés par une courbe plus jolie.

Le bouton « Optimize smoothing » oblige la courbe à passer par les divers points, mais le temps de calcul devient très long lorsqu'il y a beaucoup de points.

La valeur numérique saisie dans la ligne « Smooth: » indique la force du lissage. Par défaut elle est 1 (lissage assez faible), mais on peut y mettre des valeurs plus grandes (par exemple 5), ce qui aboutit à une courbe beaucoup plus lissée, s'écartant plus des points de mesure.

Le bouton « View given points » fait apparaître un tableau de valeurs numériques, avec les valeurs initiales de x et y, mais aussi les valeurs lissées des y. On peut sélectionner ces valeurs, les copier et les coller dans un tableur ou un autre tableau de PAST.

Il suffit de deux colonnes, x et y pour utiliser cette option. On peut aussi utiliser une troisième colonne, correspondant aux écarts-types des y,

🛷 abondance. dat Abundance models File Edit Transform Plot Statistics Multivar Diversity Time Geomet Strat Cladistics Script X start: 🚺 64 X end: 🗠 😫 📥 , 😫 👯 Filt mode Y start: -1 Family Species Ind 3 Y end: Fabaceae Anadenanthera_falce185 2.4 Malvaceae 2 Eriotheca_gracilipes 43 2 Model 3 Fabaceae Stryphnodendron_ob(37) Geometric 4 1.6 Melastomatad Miconia_albicans log Abundance Log series 5 Ebenaceae Diospyros_hispida. 30 1.2 Broken stick 6 Asteraceae Piptocarpha_rotundife30 Log-normal 0.8 7 Fabaceae Dalbergia_miscolobiur 29 0.4 8 Proteaceae Roupala_montana 29 ✓ Log scale plot 9 0 Bauhinia_rufa Fabaceae k: 0.06417 10 Acosmium_subelegar 24 Fabaceae -0.4 11 Flacourtiaces Casearia_sylvestris 23 chi^2: 444.2 -0.8 12 Styracaceae Styrax_ferrugineus 21 p(same): 4.628E-67 13 5 10 15 20 25 30 35 40 45 50 55 60 Melastomatac Miconia_ligustroides 18 0 14 Rank Rubiaceae Tocoyen_formosa 18 A 15 Myrtaceae Campomanesia_pube 16 16 Annonaceae Xylopia_aromatica identificação das especies vegetais vasculares paseada 17 Bignoniaceae Tabebuia_ochracea 14 es sampled em caracteres vegetativos para a ARIE Cerrado Pé-de-18 Clusiaceae Kielmeyera_coriacea 13 Gigante (Santa Rita do Passa Quatro, SP). Revista do Brazil with 19 Fabaceae Machaerium_acutifolii Instituto Florestal 11:137-158. 20 Myrtaceae Myrcia_guianensis 12 21 Erythroxylac Erythroxylum_subero 11 22 Caryocarace Caryocar_brasiliense 9

5.8 <u>Abundance = modèles d'abondance</u>

5.9 <u>Species packing = « garniture d'espèces » : abondance d'espèces</u> <u>selon un gradient de facteur du milieu</u>

Cette option permet un calcul théorique de l'optimum écologique pour un groupe systématique, ainsi que des intervalles de tolérance.

Deux modèles sont possibles pour le calcul de l'optimum :

- Par défaut, on suppose que la répartition écologique suit une loi normale en fonction du facteur du milieu (case « Gaussian » cochée)
- Un autre modèle possible est la moyenne pondérée (case « Weighted avrg » cochée).



PAST, logiciel scientifique --60- \parallel 5 Model : modélisation

6 <u>Diversity : étude de la</u> <u>biodiversité</u>

Il existe diverses méthodes pour calculer la « biodiversité ». Le plus simple est la « richesse spécifique », qui est le nombre d'espèces présente sur une surface donnée. Par exemple, la richesse spécifique des Oiseaux terrestres en Amérique du Nord augmente du Canada vers l'isthme de Panama.



Diversity profiles

D'autres indices plus compliqués sont calculables.

6.1 <u>Diversity indices = indices de diversité de type α (alpha) =</u> <u>biodiversité locale</u>

C'est le plus simple à comprendre : pour un écosystème donné, on prend une surface fixe (un mètre carré pour les petits arthropodes du sol, un kilomètre carré pour les oiseaux, etc), et on compte le nombre d'individus d'espèces différentes sur cette surface.

	а									
File Edit	u Transform I	Plot Statist	ice Multivar	Model	🖑 Diversity					
	1		ics mainwar	Moder		lieu1	lieu2	lieu3	lieu4	Т
😫 🗠	A 🖇	🛱 🔃	Edit mode	Г	Taxa_S	4	3	5	2	
		linua	liou 2	linud	Individuals	16	44	85	40	
espàce A		nieuz os		0	Dominance_D	0.4453	0.4163	0.368	0.5313	
espèce_A	3 40	40		0	Shannon_H	1.041	0.9825	1.208	0.6616	
espèce_D	10	0	0	0	Simpson_1-D	0.5547	0.5837	0.632	0.4688	
espece_c	2	a	25	0	Evenness_e^H/S	0.7079	0.8904	0.6693	0.9689	
espece_D	U		44	15	Menhinick	1	0.4523	0.5423	0.3162	
espece_E	1		3	25	Margalef	1.082	0.5285	0.9004	0.2711	
5					Equitability_J	0.7508	0.8943	0.7505	0.9544	
/					Fisher_alpha	1.712	0.7286	1.161	0.4431	
8					Berger-Parker	0.625	0.5682	0.5176	0.625	_
9				_						
10										
11					E Bootstran	195% confide	nce)	¥ 📑	A	
12					, poororian					

Cliquer la case « Bootstrap » augmente le nombre de colonnes : pour chaque lieu sont indiqués les limites supérieure et inférieure de l'intervalle de confiance pour les différents indices.

6.2 <u>Quadrat richness = richesse des quadrats = richesse intralieu</u>

Un quadrat est une parcelle d'échantillonnage écologique, en principe de forme carrée, d'où son nom. Selon le niveau de l'étude, la dimension d'un quadrat est variable : 100x100 m, ou 2x2 m, ou 1x1 m, etc. On peut par exemple utiliser des quadrats fixes, dont on suit la composition biologique au cours du temps, d'année en année. PAST, logiciel scientifique --62– $\parallel 6$ Diversity : étude de la biodiversité

🖑 PAST					Quadrat spec	ies richn	ess estimators	
File Edit	Transfo	rm Plot St	atistics Mult	ivar M				
🖴 😫	A	🧶 📴 🕻	🖹 🔽 Edit	mode	Chao 2:	5	Chao 2 variance:	0
	lieu1	lieu2	lieu3	lieu4	Jackknife 1:	5		
espèce_A	1	1	0	1	Jackknife 2:	5		
espèce_B	1	1	1	1	Poststrani	5 00704		
espèce_C	1	1	1	1	воосятар:	5.00701		
espèce_D	0	1	1	1				
espèce_E	1	1	1	1				
6						×	<u> </u>	
7								

On réalise donc un tableau de présence/absence des différentes espèces (en lignes) dans les différents quadrats (les colonnes), et divers indices sont calculés. Plus les quadrats sont différents les uns des autres, plus ces indices sont grands.

6.3 <u>Beta diversity =indices de diversité de type beta = diversité</u> <u>interlieux</u>

🐣 Untitle	d					
File Edit 1	Fransform P	lot Statistics	; Multivar	Model D	🛷 Beta diversity	
😫 🗠	B	🛤 🔃 T	Edit mode	Г	Whittaker	0 42857
	lieu1	lieu2	lieu3	lieu4	Harrison:	0.10714
espèce_A	3	25	5	0	Cody:	2
espèce_B	10	10	8	0	Routledge:	0.12727
espèce_C espèce D	2 N	9	25 44	U 15	Wilson-Shmida:	0.71429
espèce_E	1	0	3	25	Mourelle:	0.17857
6					Harrison 2:	0.083333
7					Williams:	0.25
8						
9						
11					×	

Alors que la diversité de type α était mesurée lieu par lieu, la diversité de type β mesure la variabilité entre divers lieux selon plusieurs méthodes.

6.4 <u>Taxonomic distinctness = distance taxonomique</u>

🥔 PAST						Taxonomic distinctness with 95% confidence intervals					
File Edit	Transform P	lot Statistic	s Multivar	Model Diver	sity Time		lieu1	lieu2	lieu3	lieu4	
🖎 📴 📴 🏧 🗖 🗖 🗖 Edit mada 🖉 Edit labala							0.7692		1.192	1.535	
		<u>49 48</u> 7	 Euit mode 	j Eur	lapeis	Lower limit	1.231		1.499	1.492	
	espèce	lieu1	lieu2	lieu3	lieu4	Upper limit	1.718		1.608	1.613	
1	espèce_A	3	25	0	8	Distinctness	2		2	2	
2	espèce_B	10	0	8	9	Lower limit	2		2	2	
3	espèce_C	0	0	25	10	Upper limit	2		2	2	
4	espèce_D	0	0	44	15						
5	espèce_E	0	0	3	25				E .		
6								<u>×</u>	17 4		

Ces indices donnent des propriétés des différents sites de mesure, et sont fondés sur la classification linnéenne.

PAST, logiciel scientifique --63– $\parallel 6$ Diversity : étude de la biodiversité

Il faut mettre dans la colonne de gauche les noms d'espèce. Éventuellement, dans la colonne suivante on peut mettre les noms de genre (de ces espèces), puis dans la colonne suivante les noms de famille, etc. Toutes les autres colonnes correspondent à des lieux, et on met dans les cases les comptages des individus des différents groupes systématiques.

« diversity » : c'est une sorte de moyenne des longueurs des chemins entre les paires d'individus de l'échantillon (de taxons différents)

« distinctness » : c'est une sorte de moyenne des longueurs des chemins entre les paires d'individus de l'échantillon (y compris du même taxon).

6.5 <u>Individual rarefaction : estimation de l'effet de l'effort</u> <u>d'échantillonnage</u>

Second														
🐣 Untitle	d					🥔 Individ	lndividual rarefaction							
File Edit '	Transform P	lot Statistic	s Multivar	Model Diver	rsity Time G	Samp size	lisu1	Stdev	lieu2	Stdev	lieu3	Stdev	licu4	Stdev 🔼
a 🚯	A S	B2 🔁	Edit mode	E Edit	lahala l	- 1	1	0	1	0	1	0	1	0
							1.59167	0.491525	1.74868	0.433769	1.59589	0.49072	1.76753	0.42241
	espèce	lieu1	lieu2	lieu3	lieu4	3	1.99107	0.650488	2.32588	0.615898	1.97754	0.597333	2.36552	0.598154
1	espèce_A	3	25	0	8	4	2.2989	0.711521	2.78264	0.717627	2.24098	0.635645	2.83809	0.696613
2	espèce_E	10	10	8	9	5	2.55815	0.725836	3.15198	0.768032	2.43573	0.651233	3.21642	0.745845
3	espèce_C	2	9	25	10	6	2.78559	0.717177	3.45559	0.784854	2.58812	0.658153	3.52273	0.762632
4	espèce_C	0	8	44	15	7	2.9875	0.69534	3.70812	0.779633	2.71276	0.660883	3.7731	0.757939
5	espèce_E	1	7	3	25	8	3.16667	0.664419	3.9198	0.759938	2.81815	0.660885	3.97932	0.739087
6						9	3.325	0.62646	4.09804	0.730767	2.9095	0.658743	4.15019	0.711022
7						10	3.46429	0.582567	4.24845	0.695482	2.99015	0.654773	4.29239	0.677125
0						11	3.58631	0.53312	4.37541	0.656402	3.06237	0.649234	4.41111	0.639737
9						12	3.69286	0.477693	4.48249	0.615176	3.12771	0.642364	4.51042	0.600502
10						13	3.78571	0.414655	4.57265	0.573012	3.18731	0.634383	4.59357	0.560589
11						14	3.86667	0.339935	4.64837	0.530808	3.24201	0.625511	4.66322	0.520835
12						15	3.9375	0.242061	4.71179	0.489243	3.29244	0.615912	4.72154	0.481843
13						16	4	0	4.76473	0.448823	3.3391	0.605747	4.77033	0.444044
14						17			4.80878	0.409927	3.38238	0.595144	4.81108	0.407741
15						18			4 84528	0.372825	3 42263	0 584203	4 84508	0.373139
16						19			4 87542	0 337702	3 46012	0 573027	4 87337	0 34037 🛛 🞽
17						~		Croph	1					
18						<u>^</u>	<u>-1</u>	Graph						

Cette fonction permet d'imaginer quelle aurait été la richesse spécifique d'une récolte (le nombre d'espèces espéré) si l'effort d'échantillonnage avait été moins important.

« Samp size » donne la taille de l'échantillon que l'on aurait pu réaliser :

- avec 1 seul individu capturé, bien sûr, la richesse spécifique vaut 1, et son écart-type vaut 0.
- avec 2 ou plusieurs individus capturés, la richesse spécifique attendue augmente avec le nombre d'individus capturés, de façon variable selon les populations.
- Le maximum calculable correspond au nombre d'individus dans la population initialement mesurée.

6.6 <u>Sample rarefaction (courbe d'accumulation spécifique) : indice de</u> <u>Mao-tau</u>

Là encore, on a divers lieux d'observation en colonnes, et divers groupes taxonomiques en lignes.

						-				
- Untitle	a					Same	lo rarofact	ion 🔲 🗖	Sample rarefaction	
File Edit	Transform F	lo: Statistic	s Multivar	Model Dive	rsity Time Ge	South A	le l'aleiaci			
			_	_	_	Samples	Таха	Stdev		X start:
	🕐 🛱 📇 🧏 👯 🔽 🔽 Edt mode 🔽 Edit labels 🛛 📋					1	4.5	NA		
	espèce	lieu1	lleu2	lieu3	lieu4 F	2	5	NA		Xend: 4
1	espèce A	3	25	0	8	3	5	NA	5.04-	Vetart: 4.4 ^c
2	espèce_B	10	.0	8	9	4	5	NA	4.98-	Total C
3	espèce_C	2	9	25	10				ତ୍ମ 1.92 -	Yend: (5.05
4	espèce_D	0	8	44	15				δ 4.8E-	
5	espèce_E	1	7	3	25				ξ 4.ε	
6									0 0 4.74	
7									\$ 4 6F	
8							PA.	Quant	9462-	
9						<u> </u>	17	Grapr		
10									4.00	
11									1.2	
12									1.2 1.5 1.8 2.1 2.4 2.7 3 3.3 3.5 39	
13									Samples	🗙 🛤 🔠
14										
15										

PAST, logiciel scientifique --64– $\parallel 6$ Diversity : étude de la biodiversité

Cette fonction donne la richesse spécifique en fonction du nombre d'échantillons testés.

6.7 Compare diversities : comparaison des diversités

👝 🐘	A S	Bal 🔁 🛛			🧶 Compare d	liversities						
			• Luit mode			lieu1	lieu2	Boot p(eq)	Perm p(eq)			
	espèce	lieu1	lieu2	D	Taxa S	4	5	0.385	0.309			
1	espèce_A	3	25		Individuals	16	59	0	0			
2	espèce_B	10	10		Dominance	0.4453	0.264	0.035	0.034			
3	espèce_C	2	9		Shannon H	1.041	1.475	0.047	0.042			
4	espèce_D	0	8		Evenness e^H/S	0.7079	0.8745	0.1	0.087			
5	espèce_E	1	7		Simpson indx	0.5547	0.736	0.035	0.034			
6					Menhinick	1	0.6509	0.619	0.691			
7					Margalef	1.082	0.981	1	1			
8					Equitability J	0.7508	0.9167	0.054	0.037			
9					Fisher alpha	1.712	1.304	0.961	0.971			
10					Berger-Parker	0.625	0.4237	0.092	0.069			
11												
12												
13						×	E R					
14												

A partir des indices de diversité alpha, cette option calcule les probabilités que les deux lieux soient identiques (dans les deux colonnes de gauche, par deux méthodes différentes).

6.8 Diversity t test : test t de diversité

🖑 PAST				Shannon diversity t test								
File Edit i	Transform P	lot Statistic	s Multivar N									
😤 🗠	A	🛤 🔃 I	Edit mode	lieu1		lieu2						
	espèce	lieu1	lieu2	S:	4	S:	5					
1	espèce A	3	25	Index:	0.94709	Index:	1.4414					
2	espèce B	10	10	variance:	0.043420	variance:	0.0054010					
3	espèce_C	2	9									
4	espèce_D	0	8									
5	espèce_E	1	7		t: -2.2353 p(sam	e): 0.03	36838					
6												
7												
8												
9					×							

Cette option calcule la probabilité que les indices de diversité de Shannon soient les mêmes pour les

PAST, logiciel scientifique $-65- \parallel 6$ Diversity : étude de la biodiversité

deux échantillons.

6.9 Diversity profiles : profils de diversité

Il peut exister diverses manières de calculer un indice de biodiversité, et l'indice de Shannon n'en est qu'une d'entre elles.

Cette option calcule la biodiversité selon plusieurs façons, et affiche le résultat sous forme d'une courbe.

- pour $\alpha = 0$, les indices sont simplement le nombre d'espèces
- pour $\alpha = 1$, les indices sont proportionnels aux indices de Shannon
- pour $\alpha = 2$, les indices se comportent comme les indices de Simpson



7 <u>Time : étude des séries</u> <u>temporelles</u>

Les séries temporelles, aussi appelées séries chronologiques, correspondent à des valeurs numériques mesurées en fonction du temps : il y a obligatoirement un classement, qui est le déroulement du temps. Selon les cas, le temps peut être mesuré en microsecondes ou en millions d'années. La spécificité de



l'analyse des séries temporelles est de pouvoir mettre en évidence des répétitions et des régularités au cours du temps.

7.1 <u>Spectral analysis = analyse spectrale (calcul des fréquences</u> <u>d'événements réguliers)</u>



Les données doivent être dans deux colonnes : la première est le temps, la seconde est les valeurs observées.

Cette option réalise une transformée de Fourier pour calculer l'importance des fréquences des événements observés. Elle trace en abscisses la fréquence (nombre de cycles par unité de temps), et en ordonnées la puissance des cycles en question. L'exemple ci-dessus correspond à la concentration d'oxygène 18, qui indique des changements climatiques au cours du dernier million d'années. Les pics correspondent aux cycles de Milankovitch.

Il faut au moins quatre cycles pour que cette fonction soit correcte.

- S'il n'y a qu'une seule colonne de mesure, il faut que les mesures soient régulièrment espacées. L'estimation du temps est faite par les numéros des lignes.
- S'il y a deux colonnes (le temps dans la première, et les valeurs observées dans la seconde), l'algorithme utilisé permet de travailler même sur des données où les intervalles de temps entre les mesures sont irréguliers, ce qui est fréquent dans les sciences de la nature, et en particulier en paléontologie.

7.2 <u>Autocorrelation = autocorrélation : les mesures sont elles liées au</u> <u>cours du temps ?</u>

L'étude de l'autocorrélation, découlant de l'autocovariance, essaie de répondre à la question : la mesure à un instant est-elle liée à la mesure faite à l'instant suivant ou à l'instant précédent ?

- Dans certains cas, il n'y a aucune corrélation au cours du temps, à aucune échelle, par exemple lors d'un lancer de dé.
- Dans d'autres cas, il y a une forte corrélation pour les temps proches, mais une faible corrélation pour les temps longs. Il peut même exister une corrélation négative pour certains intervalles de temps, s'il existe des phénomènes périodiques. Par exemple, à un intervalle d'une journée, il existe une corrélation positive pour les températures mesurées à l'extérieur des maisons, mais à un intervalle de six mois, il y a une corrélation négative puisque l'été est six mois après l'hiver.

Ainsi, sous une autre forme, cette étude de l'autocorrélation donne des renseignements du même type que l'analyse spectrale.

Il suffit de deux cycles pour permettre l'étude de l'autocorrélation, mais il faut que les mesures soient régulièrement espacées.

L'abscisse (« Lag ») correspond au numéro des mesures, et va jusqu'à la moitié du nombre de mesures traitées.



Il suffit d'une seule colonne de valeurs mesurées, puisque les abscisses du graphique correspondent aux numéros des mesures, et non à une durée temporelle.

7.3 Crosscorrelation = corrélation croisée entre deux variables

Il faut fournir deux colonnes de données mesurées régulièrement, et cette fonction cherche à aligner les deux colonnes de données, c'est à dire à trouver le décalage optimal pour que les deux colonnes soient bien corrélées.



PAST, logiciel scientifique $-68 - \parallel 7$ Time : étude des séries temporelles

Le graphique obtenu montre en abscisses le décalage (retard) entre les deux séries de mesures, et en ordonnées la corrélation entre les deux séries pour ce décalage.

7.4 Wavelet transform : transformation en ondelettes



C'est une méthode qui permet de chercher des régularités (périodicités) à la fois à différentes échelles de fréquences, pour une seule colonne de données, supposées régulièrement réparties.

En abscisses, les numéros des mesures (donc le temps, s'il s'agit vraiment de mesures chronologiques). En ordonnée, la période, selon une échelle logarithmique. Les taches foncées indiquent une forte variabilité pour la période considérée, et le fond clair indique une faible variabilité, en bas avec une vue détaillée, en haut avec une vue globale.

7.5 Walsh transform = transformation de Walsh ou de Hadamard



C'est une sorte d'analyse spectrale par transformée de Fourier pour des données binaires ou ordinales. Il faut utiliser une seule colonne, de données régulièrement réparties.

C'est une fonction à employer avec précautions, car les résultats ne sont pas très interprétables.

7.6 <u>Runs test = test d'aléatoirité d'une série de données</u>

		3						-		
	🛷 Untitle	d		- Runs						
6	File Edit	Transfo	rm Plo	-				rsity	Time	Ge
Ì	😫 🗠	B	9	N1:		140		label	s	Γ
6		A	Cut	NZ: Runs:		53 4		E		F
	20	3,75		Expecte	d runs:	77.89				
	21	1,83		Z:		-13.4				
	22	-0,09		p(rando	m):	5.789E-4	¥1			
	23	-2,01								
	24	-3,93			1					
	25	-9,5		X						
	26	-19,74								
6	27	-29,98								
								1		_

C'est un test non paramétrique qui teste si une série est au hasard ou non. « non au hasard » peut être aussi bien une autocorrélation, une tendance ou une périodicité.

Il faut lui fournir une série de données, positives et négatives. Les valeurs positives sont converties en 1, et les valeurs négatives ou nulles en 0.

7.7 Mantel correlogram = périodogramme et corrélogramme de Mantel



C'est une extension multivariée à l'autocorrélation, basée sur des similarités et mesures de distances (nombreux types disponibles).

Le corrélogramme montre la similarité moyenne entre les séries temporelles et une copie décalée dans le temps, pour différents décalages.

7.8 <u>ARMA</u>

ARMA est la combinaison de « AR » (autorégression) et « MA » (mobile average = moyenne mobile). C'est un module puissant, mais assez compliqué, à partir d'une colonne de données régulièrement espacées.

Deux courbes sont affichées : en bleu le simple graphique des données, en rouge les résidus, en fonction des numéros des données.



PAST, logiciel scientifique --71- $\parallel 7$ Time : étude des séries temporelles

7.9 (Insolation)

(fonction spéciale pour calculer l'insolation d'une région à une époque donnée de l'histoire de la Terre ; il faut des données supplémentaires : voir fichier d'aide de PAST)

8 <u>Geomet = mesures</u> <u>géométriques</u>

Ce chapitre est destiné à répondre à des questions du type : « les mesures d'angle sont-elles concentrées, ou bien sont-elles aléatoires, dans n'importe quelle direction de l'espace ? », ou bien « à partir d'une dizaine de points de mesures sur l'ensemble de la carte, peut-on établir une carte de pollution ? ».

8.1 <u>Directions (one sample) = étude des</u> <u>orientations pour un échantillon</u>

La colonne étudiée doit contenir les mesures d'orientation, en degrés.

Cette option compte le nombre de mesures faisant partie d'un secteur (par défaut 10 secteurs, mais on peut fixer le nombre par

Geomet Strat Cladistics Script

Directions (one sample) Directions (two samples) Circular correlation

Nearest neighbour (2D points) Ripley's K (2D points) Gridding

Multivariate allometry Fourier shape (2D) Elliptic Fourier (2D) Eigenshapes (2D) Thin-plate splines and warps (2D) Relative warps (2D) Size from landmarks (2D/3D) Distance from landmarks (2D/3D) All distances from landmarks (EDMA) Landmark linking



mesures qui y ont été observées.

Par défaut, l'orientation est avec la convention mathématique : le zéro est à droite (vers l'est), et le sens est antihoraire. On peut choisir la convention géographique, zéro vers le haut (vers le nord) et sens horaire en cochant la case « Geo convention ».

Par défaut, le nombre d'observations dans un secteur est rendu par la longueur du rayon de ce
secteur. Si on veut qu'il soit rendu par la surface du secteur, cocher la case « Equal area ».

Si on coche la case « orientations », les secteurs symétriques sont aussi tracés.

La case « Circles » permet de tracer les cercles de repérage, la case « Mean » permet de tracer la direction moyenne des observations.

Diverses valeurs statistiques sont dans la partie droite du graphique. En haut, la moyenne et l'intervalle de confiance à 95%, ainsi que l'intervalle de confiance par rééchantillonnage.

On voit ensuite trois façons différentes d'estimer la probabilité que la distribution soit due au hasard. L'estimation par la méthode du Khi-2 dépend de la valeur mise dans la case « Chi2 bins ».

8.2 <u>Directions (two samples) = comparaison des directions de deux</u> <u>échantillons</u>

File E	dit Transform	Plot Statistics	🛷 Directi	ons (two	samples)		
۵	R 🔒 🙎) 📑 💽 🔽	ech_1			ech_2	
	ech_1	ech_2		Г	Orientations		
1	5.64	28.66	11:	34		N:	34
2	1.89	33.94	Mean:	6.468		Mean:	27.575
3	14.21	31.64	R:	33.888		R:	33.905
4	9.54	28.2					
5	3.97	27.26	Wateon		e.		
6	0.54	23.95	1111301		_		
7	13.21	31.8	Overall	kappa:	0		
8	5.04	31.57	Overall	R:	66.646		
9	0.93	25.37	0:		365.71		
10	8.82	22.52	p(same	; mean):	1.2053E-28		
11	3.58	23.49					1
12	3.3	21.44		×	E 🙀	Ē	3
13	33	24.28				_	

Les directions de deux échantillons sont comparées, et la probabilité que ces deux directions soient identiques est affichée à p(same mean): Pour que ce test soit valable, il faut que les variances angulaires soient similaires.

Par défaut, on suppose que les valeurs sont des directions (valeurs possibles entre 0 et 360), mais en cochant la case « Orientations », on indique que les valeurs sont des orientations (valeurs possibles entre 0 et 180).

PAST, logiciel scientifique $-74 - \parallel 8$ Geomet = mesures géométriques

onoura			<u></u>	<u>i angalan</u>	U			
	ech_1	ech_2	F G	Н		J		K
1	5.64	28.66	A Circular	correlation				
2	1.89	33.94	Circular	corretation				
3	14.21	31.64						
4	9.54	28.2	ech_1			ech_2		
5	3.97	27.26			ations			
6	0.54	23.95			auons			
7	13.21	31.8						
8	5.04	31.57	N:	34				
9	0.93	25.37	Mean:	6.468		Mean:	27.575	
10	8.82	22.52						
11	3.58	23.49						
12	3.3	21.44	R:	0.11548				
13	3.3	24.28	T:	0.60959				
14	2.99	25.45	p(uncorr	•): 0.54213				
15	14.18	21.49						
16	1.06	34.43			Real			
17	0.95	25.38		×	E			
18	5.87	33.22						

8.3 <u>Circular correlation = corrélation angulaire</u>

Contrairement au test précédent, il n'y a pas besoin de distributions semblables, mais il faut de nombreuses mesures pour qu'on puisse affirmer que les deux échantillons ne sont pas corrélés. Là encore, en cochant la case « Orientations », on indique que les valeurs sont des orientations (l'orientation Nord-Sud est équivalente à l'orientation Sud-Nord).

8.4 <u>Nearest neighbour (2D points) = tests de la répartition aléatoire sur</u> <u>une surface</u>

On doit fournir à cette fonction un couple de colonnes, qui correspond aux coordonnées bidimensionnelles x/y des points d'observation.

É	dition Affich	hana Hi, 🔟 S	spatstat.ppt	(lecture seule) - Op	enOffice.org l	mpress
>	🛷 PAST			🛷 Point distributio	n	
are	File Edit	Transform Plot	t Statistics I	Number of points: Area of convex hull: Mean density	4 1	Area estimation
8	1	х у З З	G	Nearest neighbors:	4	C Smallest rectangle
atic	2	1 1	<i></i>	Mean distance: Expected distance:	0.55902	C Off
	4	-3 -2	, 5 2	Z value:	4.7293	 Wrap-around Donnelly
ja.	5 6			p(random): R value:	2.2525E-06 2.2361	
	7 8			Orientations and dista	nces	×
	9					
thi	10					

La fonction calcule les distances entre les divers points, et teste si la répartition est aléatoire ou non.

Par défaut, les calculs sont faits sur l'enveloppe convexe (« convex hull »). Il y a une correction de bordure (« Edge correction »), destinée à tenir compte du fait que les points situés à la périphérie ont forcément une distance plus grande aux autres points que les points situés dans la partie

centrale, mais on peut modifier cette option.

La distance moyenne de chaque point à son plus proche voisin est calculée (« Mean distance »), et comparée à une valeur théorique, ce qui permet de calculer la probabilité que la distribution soit aléatoire (« p(random) »).

Si les points sont groupés en agrégats, la valeur de R est inférieure à 1 ; si les point sont dispersés aléatoirement selon une loi de Poisson, R vaut 1, et si les points sont dispersés uniformément, R est supérieur à 1.

8.5 <u>Ripley's K (2D points) = indication visuelle du groupement des</u> points



Là aussi, on fournit à la fonction un couple de colonnes, correspondant aux coordonnées des points en x/y.

PAST calcule les distances entre les points, et trace en fonction de la distance une fonction dont l'aspect dépend de la disposition des points.

Si les points sont distribués au hasard, la fonction K(d), qui est la densité moyenne des points en fonction de la distance d, doit augmenter comme le carré de la distance.

La fonction L(d) vaut la racine carrée de K(d), et est donc proportionnelle à la distance, et L(d)-d doit être égale à zéro si les points sont bien aléatoires.

Par contre, si les points sont groupés en agrégats, les fonctions montrent des brisures.

« Area » indique la surface prise par l'ensemble des points. Par défaut, c'est le plus petit rectangle contenant les points, mais l'utilisateur peut changer cette valeur.



8.6 <u>Gndding = interpolations spatiales</u>

Cette fonction permet d'obtenir des cartes montrant l'estimation d'une grandeur pour toute la zone, à partir seulement de quelques points de mesures. C'est très spectaculaire, mais il ne faut pas accorder une confiance absolue dans les valeurs numériques des interpolations.

On doit fournir à PAST trois colonnes, correspondant aux divers points de mesure. La première valeur est l'abscisse x, la deuxième est l'ordonnée y, et la troisième est la valeur numérique de la grandeur observée.

Le choix de l'option entraîne l'apparition d'une petite fenêtre. Par défaut, les calculs d'interpolation sont faits sur une matrice de 100 lignes et 100 colonnes, mais ces nombres peuvent être changés (plus la taille est grande, plus les calculs sont longs). « Neighbourhood: » indique le nombre de points pour le calcul de la moyenne mobile, et n'est valide que pour cette option, qui est l'option par défaut.

« Go » déclenche les calculs d'interpolation et le traçage de la carte.

On a le choix entre trois manières d'interpoler :

- « Moving average », par la méthode des moyennes mobiles, où chaque point a un poids en fonction inverse de la distance. C'est un algorithme simple, qui ne donne pas toujours de bons lissages, mais qui a l'avantage que les valeurs calculées ne sortent jamais de l'étendue des points de mesure.
- « Thin plate spline », qui donne le lissage maximal
- « Kriging », le krigeage, qui nécessite d'avoir auparavant calculé le semivariogramme.

8.7 <u>Multivanate allometry = Allométrie sur plusieurs variables</u>



On part d'un ensemble de colonnes correspondant aux grandeurs mesurées. PAST transforme ces données en logarithmes, puis réalise une analyse en composantes principales (ACP = PCA). La première composante principale est considérée comme l'axe de taille, et le coefficient d'allométrie est calculé pour chaque variable initiale, comme la contribution de cette variable à la première composante principale divisée par la contribution moyenne de toutes les variables.

8.8 Divers types d'analyses de forme (analyses procustéennes)

Le terme « procustéen » fait référence à la mythologie grecque, où le bandit Procuste (Procrustes en anglais) forçait ses victimes à s'allonger dans un lit, et les adaptait à la longueur de son lit, en les coupant à la hache ou en les étirant ; il fut finalement tué par Thésée. Ce terme « procustéen » fait maintenant référence aux problèmes d'ajustement de formes, par translation, rotation et changement d'échelle.

Diverses fonctions demandent que les données soient normalisées par la transformation appropriée (voir menu Transform | Procrustes (2D/3D)).

8.8.1 Fourier shape(2D)

Il faut fournir à cette option au moins 6 colonnes, c'est à dire 3 paires de coordonnées (x,y).

Chaque individu mesuré correspond à une ligne, et les mesures correspondent à des couples x,y, disposés sur une même ligne.

8.8.2 Elliptic Fourier (2D)

Il faut au moins dix couples de points (x,y) sur une même ligne, donc vingt colonnes.

PAST, logiciel scientifique $-78 - \parallel 8$ Geomet = mesures géométriques

0.0.3	Eigensnap	<u>es (2D)</u>		
enne - Wikip	edia - Mozilla Firefox			College and the last
🖉 🥮 gorilla	. dat		PCA scatter diagram	
File Edit	Transform Plot Statistics Multiva	r Model Diversity Time Geor		
				X start: -0.7
		de j Edit labels j		
	x1 y1 x2	y2 x3 y3		X end: 0.3
f 14	0.493249 0.05875657 -0.46066	12 -0.05788957 -0.304949 0.1		V starts 1
- f15	0.5014358 0.0599235 -0.44974	86 -0.03862538 -0.311786 0.1	19_	
⁵ f16	0.4911185 0.04139589 -0.44735	73 -0.07001181 -0.3107935 0.1	1.5	Yend: 2
f17	0.4964674 0.0399252 -0.45665	65 -0.05014298 -0.3197714 0.1	1.8	
.0 f18	0.4906509 0.02334784 -0.44491	95 -0.06848626 -0.3310973 0.1	+ + 17	Row labels
a <u>f19</u>	0.49 🧑 Figenshapes			95% ellinges
f20	0.49	1		Convex hulls
f21	0.49	1	· Š \ (+ • + + + + + + + + + + + + + + + + +	Biplot
f22	0.47 Value	% Variance 📃		1
x 123	Eigenval 1: 0.0222	2651 58.388 🔮	° · · · · · · · · · · · · · · · · · · ·	📃 Eigenval scale
11 124 f05	0.49 Figerwal 2: 0.0103	200 28 200	+ 1.3	l l l l l l l l l l l l l l l l l l l
125	0.49 Eigenvarz: 0.0100	125 20.355	• 12-	Axes
120	Eigenval 3: 0.003	9795 10.436 🚽	- 1.2	0 1+2
f28	6.44 6.50 Eigenval 4: 0.0009	9804 2.571 📩	1.1 -	C 2+3
s f29	0.49	1	1	
f 30	0.50 Jolliffe cut-off: 0.0044	1488		View numbers
m1	0.49 Standardized point nun	nber: 8 .	Component 1	
m2	0.50			🗡 🛱 📇 🛛
m3	0.51 Closed outlines			B
m4	0.50		441347 -0.1962062 0.1516542 0.106571 0.1326967 0.41	08/2009 10:35 3
m5	0.49 View scatter Vie	wyshapes 🗶 💾 🥂	362178 -0.1984193 0.1677812 0.120415 0.1409538 0.43	1/08/2009/09:44 Z
- <u>m6</u>	0.49	1	42335 -0.1904429 0.1640233 0.1064304 0.1260123 0.43	108/2009 00:12
m7	0.5084419 0.01749061 -0.43592	79 -0.06957686 -0.328643 0.1	403082 -0.1992781 0.1605343 0.1210218 0.1462425 0.410 31	/07/2009 11:53 1
m8	0.5127896 0.02044037 -0.45811	54 -0.06357652 -0.3100506 0.1	483092 -0.194874 0.1475105 0.1052348 0.1384487 0.42 31	/07/2009 09:13 5
m9	0.5002004 0.011147 -0.42238	21 -0.0514592 -0.347383 0.11	242528 -0.214329 0.17378 0.1045514 0.1409883 0.44: 30	1/07/2009 00:44 2
m10	0.5027588 0.0151939 -0.42741	21 -0.05995849 -0.3230033 0.1	562646 -0.2048805 0.1640359 0.1001175 0.1422307 0.43 25	//07/200918:49 5 0/07/2009.06:25 3
m11	0.5018894 0.0268341 -0.43960	52 -0.05317452 -0.3118503 0.1	599044 -0.188124 0.1618088 0.09986196 0.1334824 0.42 28	107/2009 23:39 2
<				/07/2009 20:38 4

000 Figenshapes (2D)

Le résultat est une sorte d'analyse en composantes principales (PCA = ACP) portant sur les positions des points indiqués.

8.8.4 Thin-plate splines and warps (2D) = déformation et lissages

La référence prise est soit le premier spécimen (première ligne), soit la moyenne de l'ensemble des spécimens. Les déformations des autres spécimens sont calculées, et l'on passe d'un spécimen à un autre par en cliquant sur les boutons de navigation (spinbuttons : double triangle noir). La normalisation de Procuste (voir menu Transform | Procustes) est recommandée.

8.8.5 Relative warps (2D) = déformations relatives

Là aussi, la normalisation de Procuste est recommandée.

Size from landmarks (2D/3D) = taille à partir de points de repère 8.8.6

:ils	🖉 Untitle	d						🛷 Size f	rom landr	narks		
1	File Edit	Transform	Plot Statist	tics Multivar	Model Di	versity Time	Geomet Stra	Specimen	Size	Normalized		
				-				1	197.241	52.7149		
6				I✓ Edit mod	ie E	dit labels	Square m	4	190.002	50.7802		
-		s	T	U	V	W	x					.c 🔥
	1	94,41	86,67	97,2	92,5	98,6	98,33			Dimensions	ř.	
	4	78,57	25,69	100,65	98,17	107,14	52,29	X	B3			
	2					Contraction of the			100	C 3D		
	3									-		
	5											

Cette option calcule la taille des échantillons (chaque échantillon correspond à une ligne). La

PAST, logiciel scientifique --79- || 8 Geomet = mesures géométriques

distance est la norme euclidienne des distances de tous les points de repères au centroïde.

Par défaut, PAST considère qu'il y a deux dimensions, et il faut que chaque ligne contienne un nombre pair de cases, car chaque couple de cases correspond à un couple x, y. On peut déclarer travailler en 3 dimensions en cochant la case « 3D », et dans ce cas, il faut que le nombre de colonne soit un multiple de 3.

« Normalized » correspond à la taille calculée précédemment divisée par la racine carrée du nombre de points, ce qui permet de comparer les tailles des spécimens ayant un nombre différent de points de repère.

Pour cette fonction, il ne faut pas effectuer de normalisation de Procuste.

8.8.7 <u>Distance from landmarks (2D/3D) = distance à partir de points de</u> repère

au	(A) Hattel	ad i				Contractor						
ABO	File Edit	Transform	Plot Stati	stics Multive	ar Model	🐠 Dist	ance between la	andmarks 🗐 🗖 🔀			الكارد	
	e 🛱		9 🛤 🔃	🔽 Edit mo	ode [1	Distance 2.8653	Dimensions				
a.		A	в	С	D	4	05.53		I.	J	K	~
	1	17,48	19,17	18,88	16,67			C 3D	23,08	20	23,7	
ĩč.	4	12,34	9,17	18,18	94,5				41,56	64,22	49,3	
	2	-			_					and the s		
	3	-										
na	5								_			
	7				_			L C	_			
Mod	8	9										
102	9	-			-			▼ A				
	10	1										
V	11											
92	12							and and				
:98	13							🗙 📑				
	14											
	15	1			1	4		-				

Cette fonction calcule les distances entre points de repère d'un même spécimen. L'exemple cidessus calcule pour les deux spécimens 1 et 4 la distance entre le point (C,D) et le point (A,B). La première distance vaut 2,8653, car c'est la racine de $(18,88-17,48)^2 + (16,67-19,17)^2$.

On peut naviguer dans les points dont il faut calculer la distance par les boutons de navigation. Ces boutons ont un pas de 2 colonnes en 2 dimensions, puisque chaque point correspond à un couple de colonnes.

8.8.8 <u>All distances from landmarks (EDMA) = toutes les distances entre</u> points de repère

Cette option transforme le tableau, en calculant toutes les distances entre points de repère d'un même spécimen.

8.8.9 Landmark linking = liens de points de repères

9 <u>Strat : analyses</u> <u>stratigraphiques</u> <u>spécifiques</u>

Strat Cladistics Script

Unitary associations Ranking-Scaling Constrained optimization Appearance Event Ordination Diversity curve Range confidence intervals Distribution-free range CIs Spindle diagram Filter events

9.1 <u>Unitary associations = associations unitaires</u>

Il faut une matrice de présence/absence, avec les colonnes correspondant aux taxons. Les lignes correspondent aux observations, avec diverses contraintes :

P	🧶 alveoli	nid.dat				Unitary associations
	File Edit	Transform P	'lot Statistic	s Multivar	Model D	Select sections from list: Options:
a b :tr	1-4 1-3 1-2 1-1 2-4 2-3 2-2 2-1 3-7 3-6 3-5	moussoulens 0 0 0 0 0 1 1 0 0 0 0 0 0 0 0 0	aramea 0 0 1 0 0 0 0 0 0 0 0 0	Solida 1 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0	globosa 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	2 2-4 3 3-7 4 4-4 5 5-2 6 6-4 7.7-4 8 8-6 9 9-2 10: 10-3 11: 11-1 View section similarities Gol
	3-4 3-3 3-2 3-1 4-4	0 0 0 0 0	0 0 1 0	0 0 1 0 0	0 0 0 0 0	Residual horizons: 15 Contradictions: 22 Maximal cliques: 12 Cliques in cycles: 4 Unitary associations: 7 Residual virtual edges: 0
il	4-3 4-2 4-1 5-2 5-1	U 1 1 0 0	0 0 0 0	U 0 0 0	1 0 0 1	Maximal cliques Unitary associations Reproducibility graph Cumul. FADs/LADs Graph of cliques UA reproducibility Correlation table Diachronism Biostratigraphic graph Numerical ranges Contradictions Longest path of Gk
er re	6-4 6-3 6-2 6-1	0 0 1 0	0 0 0 0	0 0 0 0	0 1 0 0	Co-occurrences Taxon frequency Real arcs between UAs Islands+dead ends in Gk Superpositions Taxon distribution Virtual edges Unmerged cliques 74 cycles Max clique stats Image: State s
.0	7-4 7-3	0	0	0	0	Close

- les différents lieux d'observation doivent être notés par des couleurs différentes. Ces couleurs permettent à PAST de faire la liste des lieux, mais la couleur en elle-même n'a pas d'importance ; on peut réutiliser plusieurs fois les mêmes couleurs, en particulier pour des tableaux comportant de nombreux lieux d'observation.
- Pour un même lieu, les observations doivent être triées chronologiquement (donc par position, en stratigraphie) : les strates les plus basses (les plus anciennes) doivent être les plus basses dans le tableau.

Après avoir choisi cette fonction « Unitary associations », puis cliqué sur « Go! », on peut choisir diverses possibilités pour mieux comprendre les relations entre les diverses strates des divers lieux.

9.2 <u>Ranking-Scaling = classement des données stratigraphiques</u>



Il faut partir d'une table de profondeur des événements dans divers puits ou lieux d'observation.

Les lignes indiquent les puits, les colonnes indiquent les événements, et la valeur dans les cases correspond à la profondeur où ces événements sont observés dans les puits (les absences sont codées par des 0). Si l'on n'a pas de mesure de profondeur, mais seulement un ordre d'apparition dans les puits, on peut réaliser le codage par des nombres entiers correspondant au rang.

La première étape est le classement (« ranking »), optimisé pour l'ensemble des puits, même s'il existe des contradictions entre les divers puits. Par exemple, il est possible que dans le puits 1, l'événement A soit avant l'événement B, alors que c'est le contraire dans le puits 2.

La deuxième étape est la mise à l'échelle (« scaling »). Cette étape refait aussi un classement.

9.3 Constrained optimization CONOP

Là encore, les lieux d'observation (les puits d'un forage, par exemple) correspondent aux lignes. Dans les cases des colonnes impaires sont les profondeurs ou niveaux d'apparition du taxon, et dans les colonnes paires les profondeurs de disparition du taxon. Les événements manquants sont codés par des zéros.

Cette fonction permet de faire des corrélations biostratigraphiques quantitatives.

9.4 Appearance Event Ordination



9.5 <u>Diversity curve = courbe de diversité, avec apparition et disparition</u> <u>de taxons</u>



A partir d'une matrice de présence-absence, où les colonnes correspondent aux espèces et les lignes aux étages, PAST trace un diagramme avec en bleu le nombre d'espèces présentes. En cochant « Origination », il apparaît une courbe verte indiquant les apparitions d'espèces, et en cochant « Extinction », il apparaît une courbe rouge montrant les extinctions.

9.6 <u>(Range confidence intervals = calcul des intervalles de confiance</u> <u>de la distribution des fossiles)</u>

(calculs théoriques, ne dépendant pas des données dans le tableau. On suppose une distribution aléatoire des fossiles dans les horizons)

9.7 <u>Distribution-free range Cls = intervalles de confiance indépendants</u> <u>de la distribution.</u>

Ici, on ne suppose pas une distribution aléatoire des horizons fossilifères. On suppose qu'il n'y a pas de corrélation entre la position stratigraphique et la taille de l'intervalle.

9.8 <u>Spindle diagram = graphique d'observation des taxons au cours du</u> <u>temps</u>



Cette option fait simplement un graphique par points à partir d'une grosse matrice de présence/absence. Elle peut être utile pour visualiser une très grosse matrice.

9.9 Filter events

(transformation des données)

10 Cladistics : analyses cladistiques

Il n'y a qu'une seule option, l'analyse d'une matrice de caractères par parcimonie.

Ce n'est pas la seule méthode d'utiliser de telles matrices pour imaginer un arbre phylogénétique. On peut aussi utiliser l'analyse multivariée, en particulier le regroupement (menu « Multivar | cluster analysis »), mais les différentes méthodes de calcul de distance ne donnent pas le même résultat.

10.1 Parsimony analysis ____L 🐣 Untitle d 🛷 Most parsimonious trees ile Edit Trans 🗠 😫 📥 A 📑 🔁 Edit mode Edit labels Number of MPT: 1 mandibule vitellus ail machoire membres dents Tree 1 truite homme Tree length: 5 chauves Ensemble Cl: 0.8 oiseau Ensemble RI: NΑ nauvesouris 5 Strat. consist. SCI: 1 🐣 Parsimony analysis SCI permutation p: Rel. compl. RCk 100 Algorithm Optimization Gap ex. ratio GER: NA Branch-and-bound Wagner RCI/GER permut. p: 1 C Exhaustive C Fitch ✓ Branch lengths niseau Heuristic (NNI) C Dollo Phylogram · Heuristic (SPR) 2 O Heuristic (TBR) FADs/LADs in first columns Character NONE **Consistency** index: #reorderings: 5 Bootstrap replicates: 0 Retention index: Random seed: 0 Longest tree kept: 0 1 Consensus ₽ X Trees evaluated: #trees stored Shortest tree: 5 Bootstrap rep.: Go! View trees Close

Les divers taxons correspondent aux lignes, et les divers caractères correspondent aux colonnes.

L'extragroupe doit être mis en première ligne.

Les variantes des caractères doivent être codés par des entiers entre 0 et 255.

11 <u>Script : programmation des</u> <u>actions de PAST</u>

PAST contient un embryon de fonctions permettant sa programmation. Le langage peut être considéré comme une variante du langage Pascal, mais très simplifié. Normalement, cette partie « programmation » devrait progresser dans les versions ultérieures de PAST.

File Edit	Transform	Plot Statistics Multivar Model Diversity Time Geomet Strat Cl
G 🗄	a 🔒 🧕	🛛 🛤 🔃 🔽 Edit mode 🛛 🗆 Edit labels 🛛 🙀 🔀
	machoire	membres dents mandibule vitellus
truite	0	
homme	0	🗈 essai_prog.pas - Blo 🖃 🛄 🚺
chauveso	uri‡0	Fichier Edition Format Affichage ?
oiseau	0	begin OK
5		message("Hello !")
6		end.
7		
в		3
9		XI NIC
10		
11	-	

11.1 Chargement et exécution d'un programme

Le programme lui-même est un simple fichier de texte, à rédiger à l'aide d'un éditeur de texte.

11.1.1 Load script

Cette fonction charge en mémoire un programme, et rend possible l'option suivante.

11.1.2 <u>Run</u>

Cette fonction exécute un programme préalablement chargé en mémoire.

11.2 <u>Structure du langage</u>

Les variables peuvent avoir des nom de longueur quelconque, et ne doivent pas être déclarées avant leur emploi (déclaration implicite au moment du premier emploi). Ces variables ont toutes une portée globale. L'opérateur d'affectation est :=.

On peut mettre des commentaires : tout ce qui suit le dièse # est considéré comme un commentaire.

Il existe quatre types de variables : des nombres (réels en double précision), des chaînes de caractères, des vecteurs et des tableaux. Ces types sont implicitement définis au premier usage de la variable, et les conversions de type sont faites automatiquement.

L'ensemble du script doit être encadré par un couple begin...end, et les blocs à l'intérieur du programme sont aussi encadrés par des couples begin ... end.

Les structures de contrôle sont d'une part les instructions conditionnelles par ifbegin ... end else ..., d'autre part les boucles for et while.

Un prédicat « vrai » est traduit par 1, et un prédicat « faux » est traduit par 0.

Il est possible de structurer le programme par des procédures, mais qui sont beaucoup plus simples

PAST, logiciel scientifique --86- || 11 Script : programmation des actions de PAST

que dans le langage Pascal normal : elles n'ont pas de paramètres, pas de valeur de retour, pas de variables locales. Elles doivent être définies après le begin de début du programme. Bizarrement, l'appel de la procédure doit se faire avec un paramètre quelconque (fictif), alors que la procédure avait été déclarée sans paramètres.

11.2.1 <u>Opérateurs</u>

Les opérateurs mathématiques classiques +, -, *, /, ^ existent.

La concaténation de chaînes est faite par &.

Les opérateurs de comparaison sont =, >, <, >=, <=, <>.

11.2.2 Fonctions diverses

On trouve les fonctions mathématiques classiques cos, sin, tan, exp, log, sqrt, int, rnd, mais aussi des fonctions statistiques moins courantes.

Diverses opérations sont possibles sur les vecteurs et matrices, en particulier l'inversion, le calcul de moyenne, le calcul de covariance, le classement... Diverses opérations de calcul de distances entre matrices sont possibles.

On peut aussi programme le traçage de graphique par des instructions de type drawline, drawstring, drawhistogram...

PAST, logiciel scientifique --87-

Bibliographie et références

L'article officiel décrivant PAST :

Hammer, Ø., Harper, D.A.T., and P. D. Ryan, 2001. PAST: Paleontological Statistics Software Package for Education and Data Analysis. Palaeontologia Electronica 4(1): 9pp.

Quelques livres de statistiques naturalistes :

Ancelle Thierry « Satistique épidémiologie » Maloine 2002

Bouyer Jean « Méthodes statistiques Médecine-Biologie » INSERM 1997

Geller S. « Abrégé de statistique » Masson 1979

Laberche Jean-Claude « Statistiques et expérimentation en biologie » Ellipses 2008

Quelques sites internet :

Définitions et mesure de la biodiversité http://www.agroparistech.fr/IMG/pdf/Biodiv2008.pdf

LES PRINCIPAUX OUTILS STATISTIQUES EN ECOLOGIE http://alexandra.gigou.googlepages.com/Rapportcologienumrique.pdf

Synthèse des méthodes d'évaluation de la qualité du benthos en milieu côtier <u>http://www.rebent.org/documents/document.php?g_id_document=166</u>

Index

ACP5, 43 sv, 51, 77 sv AFC5, 45 ajouter des lignes ou des colonnes14 ajustement polynomial5, 57 ajustement sinusoïdal5, 56 allométrie3, 6, 77 analyse canonique des correspondances5, 46 analyse de variance. 5, 38 sv, 51 analyse des correspondances5, 43, 45 sv analyse des mélanges5, 40 analyse des similarités51 analyse en composantes principales5, 43, 77 sv analyse en coordonnées principales5, 43, 77 sv analyse factorielle des correspondances5, 45 analyse spectrale. 3, 6, 66 sv, 69 analyses de forme6, 77 annuler une opération12 ANOSIM5, 51 ANOVA5, 38 sv, 51, 54 ARMA4, 6, 37, 70 associations unitaires6, 16, 80 autocorrélation3, 6, 67, 69 sv barres d'erreur4, 23 biodiversité3, 5, 61, 65, 87 bmp21 sv boites à moustaches4, 24 Bookstein18 Box-plot4, 24 Bubble plot4, 28 sv cellules vides5, 31, 48, 84 coefficient de variation36 colorer des lignes9 compartiments24 compteur au clavier16 CONOP6, 81	ACP5, 43 sv, 51, 77 sv AFC5, 45 ajouter des lignes ou des colonnes	
AFC	AFC	ACP5, 43 sv, 51, 77 sv
ajouter des lignes ou des colonnes	ajouter des lignes ou des colonnes	AFC5, 45
colonnes	colonnes	ajouter des lignes ou des
ajustement polynomial5, 57 ajustement sinusoïdal5, 56 allométrie3, 6, 77 analyse canonique des correspondances5, 46 analyse de variance. 5, 38 sv, 51 analyse des correspondances5, 43, 45 sv analyse des mélanges5, 40 analyse des similarités51 analyse en composantes principales5, 43, 77 sv analyse en coordonnées principales5, 43, 77 sv analyse en coordonnées principales5, 43, 77 sv analyse factorielle des correspondances5, 45 analyse spectrale. 3, 6, 66 sv, 69 analyses de forme	ajustement polynomial5, 57 ajustement sinusoïdal5, 56 allométrie3, 6, 77 analyse canonique des correspondances5, 46 analyse de variance. 5, 38 sv, 51 analyse des correspondances5, 43, 45 sv analyse des mélanges5, 40 analyse des similarités51 analyse des similarités51 analyse en composantes principales5, 43, 77 sv analyse en coordonnées principales5, 43, 77 sv analyse en coordonnées principales5, 43, 77 sv analyse factorielle des correspondances5, 45 analyse spectrale. 3, 6, 66 sv, 69 analyses de forme	colonnes14
ajustement sinusoïdal5, 56 allométrie3, 6, 77 analyse canonique des correspondances5, 46 analyse de variance. 5, 38 sv, 51 analyse des correspondances5, 43, 45 sv analyse des mélanges5, 40 analyse des similarités51 analyse en composantes principales5, 43, 77 sv analyse en coordonnées principales5, 43, 77 sv analyse en coordonnées principales	ajustement sinusoïdal5, 56 allométrie3, 6, 77 analyse canonique des correspondances5, 46 analyse de variance. 5, 38 sv, 51 analyse des correspondances5, 43, 45 sv analyse des mélanges5, 40 analyse des similarités51 analyse en composantes principales5, 43, 77 sv analyse en coordonnées principales5, 43, 77 sv analyse en coordonnées principales5, 43, 77 sv analyse factorielle des correspondances5, 45 analyse spectrale. 3, 6, 66 sv, 69 analyses de forme6, 77 annuler une opération12 ANOSIM5, 38 sv, 51, 54 ARMA4, 6, 37, 70 associations unitaires. 6, 16, 80 autocorrélation3, 6, 67, 69 sv barres d'erreur4, 23 biodiversité3, 5, 61, 65, 87 bmp21 sv boites à moustaches4, 24 Bookstein18 Box-plot4, 24 Bubble plot4, 24 Bubble plot4, 24 Bubble plot4, 24 GCABFAC5, 31, 48, 84 coefficient de variation	ajustement polynomial5, 57
allométrie	allométrie	ajustement sinusoïdal5, 56
analyse canonique des correspondances5, 46 analyse de variance. 5, 38 sv, 51 analyse des correspondances5, 43, 45 sv analyse des mélanges5, 40 analyse des similarités51 analyse en composantes principales5, 43, 77 sv analyse en coordonnées principales5, 43, 77 sv analyse en coordonnées principales5, 43, 77 sv analyse factorielle des correspondances5, 45 analyse spectrale. 3, 6, 66 sv, 69 analyses de forme6, 77 annuler une opération12 ANOSIM5, 51 ANOVA5, 38 sv, 51, 54 ARMA4, 6, 37, 70 associations unitaires6, 16, 80 autocorrélation3, 6, 67, 69 sv barres d'erreur4, 23 biodiversité3, 5, 61, 65, 87 bmp21 sv boites à moustaches4, 24 Bookstein	analyse canonique des correspondances	allométrie3, 6, 77
correspondances5, 46 analyse de variance. 5, 38 sv, 51 analyse des correspondances5, 43, 45 sv analyse des mélanges5, 40 analyse des similarités51 analyse en composantes principales5, 43, 77 sv analyse en coordonnées principales5 analyse factorielle des correspondances5, 45 analyse spectrale. 3, 6, 66 sv, 69 analyses de forme	correspondances5, 46 analyse de variance. 5, 38 sv, 51 analyse des correspondances5, 43, 45 sv analyse des mélanges5, 40 analyse des similarités51 analyse en composantes principales5, 43, 77 sv analyse en coordonnées principales5, 43, 77 sv analyse en coordonnées principales5, 43, 77 sv analyse factorielle des correspondances5, 45 analyse spectrale. 3, 6, 66 sv, 69 analyses de forme6, 77 annuler une opération12 ANOSIM5, 38 sv, 51, 54 ARMA4, 6, 37, 70 associations unitaires6, 16, 80 autocorrélation3, 6, 67, 69 sv barres d'erreur4, 23 biodiversité3, 5, 61, 65, 87 bmp21 sv boites à moustaches4, 24 Bookstein	analyse canonique des
analyse de variance. 5, 38 sv, 51 analyse des correspondances5, 43, 45 sv analyse des mélanges5, 40 analyse des similarités51 analyse en composantes principales5, 43, 77 sv analyse en coordonnées principales5 analyse factorielle des correspondances5, 45 analyse spectrale. 3, 6, 66 sv, 69 analyses de forme6, 77 annuler une opération12 ANOSIM5, 38 sv, 51, 54 ARMA4, 6, 37, 70 associations unitaires6, 16, 80 autocorrélation3, 6, 67, 69 sv barres d'erreur4, 23 biodiversité3, 5, 61, 65, 87 bmp21 sv boites à moustaches4, 24 Bookstein	analyse de variance. 5, 38 sv, 51 analyse des correspondances5, 43, 45 sv analyse des mélanges5, 40 analyse des similarités51 analyse en composantes principales5, 43, 77 sv analyse en coordonnées principales5, 43, 77 sv analyse en coordonnées principales5, 43, 77 sv analyse en coordonnées correspondances5, 45 analyse factorielle des correspondances5, 45 analyse spectrale. 3, 6, 66 sv, 69 analyses de forme6, 77 annuler une opération12 ANOSIM5, 38 sv, 51, 54 ARMA4, 6, 37, 70 associations unitaires6, 16, 80 autocorrélation3, 6, 67, 69 sv barres d'erreur4, 23 biodiversité3, 5, 61, 65, 87 bmp21 sv boites à moustaches4, 24 Bookstein	correspondances5, 46
analyse des correspondances5, 43, 45 sv analyse des mélanges5, 40 analyse des similarités51 analyse en composantes principales5, 43, 77 sv analyse en coordonnées principales5, 43, 77 sv analyse en coordonnées principales5 analyse factorielle des correspondances5, 45 analyse spectrale. 3, 6, 66 sv, 69 analyses de forme6, 77 annuler une opération12 ANOSIM5, 51 ANOVA5, 38 sv, 51, 54 ARMA4, 6, 37, 70 associations unitaires6, 16, 80 autocorrélation3, 6, 67, 69 sv barres d'erreur4, 23 biodiversité3, 5, 61, 65, 87 bmp21 sv boites à moustaches4, 24 Bookstein18 Box-plot4, 24 Bubble plot4, 26 CABFAC5, 46, 52 cartographie4, 28 sv cellules vides5, 31, 48, 84 coefficient de variation36 colorer des lignes	analyse des correspondances5, 43, 45 sv analyse des mélanges5, 40 analyse des similarités51 analyse en composantes principales5, 43, 77 sv analyse en coordonnées principales5, 43, 77 sv analyse factorielle des correspondances5, 45 analyse spectrale. 3, 6, 66 sv, 69 analyses de forme6, 77 annuler une opération12 ANOSIM5, 38 sv, 51, 54 ARMA4, 6, 37, 70 associations unitaires6, 16, 80 autocorrélation3, 6, 67, 69 sv barres d'erreur4, 23 biodiversité3, 5, 61, 65, 87 bmp21 sv boites à moustaches4, 24 Bookstein18 Box-plot4, 24 Bubble plot4, 24 Bubble plot4, 24 GCABFAC5, 46, 52 cartographie4, 28 sv cellules vides6, 84 cladistique6, 84 cladistique6, 84 cladistique6, 84 cladistique6, 84 cladistique6, 84 cladistique6, 84 cladistique	analyse de variance. 5, 38 sv, 51
43, 45 sv analyse des mélanges5, 40 analyse des similarités51 analyse en composantes principales5, 43, 77 sv analyse en coordonnées principales5, 43, 77 sv analyse en coordonnées principales5 analyse factorielle des correspondances5, 45 analyse spectrale. 3, 6, 66 sv, 69 analyses de forme6, 77 annuler une opération12 ANOSIM5, 51 ANOVA5, 38 sv, 51, 54 ARMA4, 6, 37, 70 associations unitaires6, 16, 80 autocorrélation3, 6, 67, 69 sv barres d'erreur4, 23 biodiversité3, 5, 61, 65, 87 bmp21 sv boites à moustaches4, 24 Bookstein18 Box-plot4, 24 Bubble plot4, 26 CABFAC5, 46, 52 cartographie4, 28 sv cellules vides6, 84 cladistique6, 84 cluster analysis5, 31, 48, 84 coefficient de variation36 colorer des lignes	43, 45 sv analyse des mélanges5, 40 analyse des similarités51 analyse en composantes principales5, 43, 77 sv analyse en coordonnées principales5, 43, 77 sv analyse en coordonnées principales5, 43, 77 sv analyse factorielle des correspondances5, 45 analyse spectrale. 3, 6, 66 sv, 69 analyses de forme6, 77 annuler une opération12 ANOSIM5, 38 sv, 51, 54 ARMA4, 6, 37, 70 associations unitaires6, 16, 80 autocorrélation3, 6, 67, 69 sv barres d'erreur4, 23 biodiversité3, 5, 61, 65, 87 bmp21 sv boites à moustaches4, 24 Bookstein18 Box-plot4, 24 Bubble plot4, 24 Bubble plot4, 26 CABFAC5, 46, 52 cartographie4, 28 sv cellules vides6, 84 cladistique6, 84 cladistique6, 84 cladistique6, 84 cladistique6, 84 cladistique	analyse des correspondances5,
analyse des mélanges5, 40 analyse des similarités51 analyse en composantes principales5, 43, 77 sv analyse en coordonnées principales5 analyse factorielle des correspondances5, 45 analyse spectrale. 3, 6, 66 sv, 69 analyses de forme6, 77 annuler une opération12 ANOSIM5, 38 sv, 51, 54 ARMA4, 6, 37, 70 associations unitaires. 6, 16, 80 autocorrélation3, 6, 67, 69 sv barres d'erreur4, 23 biodiversité3, 5, 61, 65, 87 bmp21 sv boites à moustaches4, 24 Bookstein18 Box-plot4, 24 Bubble plot4, 24 Bubble plot4, 24 Bubble plot4, 26 CABFAC5, 46, 52 cartographie4, 28 sv cellules vides8 sv Cladistics6, 84 cladistique6, 84 cluster analysis5, 31, 48, 84 coefficient de variation36	analyse des mélanges5, 40 analyse des similarités51 analyse en composantes principales5, 43, 77 sv analyse en coordonnées principales5 analyse factorielle des correspondances5, 45 analyse spectrale. 3, 6, 66 sv, 69 analyses de forme6, 77 annuler une opération12 ANOSIM5, 38 sv, 51, 54 ARMA4, 6, 37, 70 associations unitaires6, 16, 80 autocorrélation3, 6, 67, 69 sv barres d'erreur4, 23 biodiversité3, 5, 61, 65, 87 bmp21 sv boites à moustaches4, 24 Bookstein	43, 45 sv
analyse des similarités51 analyse en composantes principales5, 43, 77 sv analyse en coordonnées principales	analyse des similarités51 analyse en composantes principales5, 43, 77 sv analyse en coordonnées principales5 analyse factorielle des correspondances5, 45 analyse spectrale. 3, 6, 66 sv, 69 analyses de forme6, 77 annuler une opération12 ANOSIM5, 38 sv, 51, 54 ARMA4, 6, 37, 70 associations unitaires6, 16, 80 autocorrélation3, 6, 67, 69 sv barres d'erreur4, 23 biodiversité3, 5, 61, 65, 87 bmp21 sv boites à moustaches4, 24 Bookstein18 Box-plot4, 24 Bubble plot4, 24 Bubble plot4, 26 CABFAC5, 46, 52 cartographie4, 28 sv cellules vides	analyse des mélanges5, 40
analyse en composantes principales5, 43, 77 sv analyse en coordonnées principales5 analyse factorielle des correspondances5, 45 analyse spectrale. 3, 6, 66 sv, 69 analyses de forme6, 77 annuler une opération12 ANOSIM5, 38 sv, 51, 54 ARMA4, 6, 37, 70 associations unitaires6, 16, 80 autocorrélation3, 6, 67, 69 sv barres d'erreur4, 23 biodiversité3, 5, 61, 65, 87 bmp21 sv boites à moustaches4, 24 Bookstein18 Box-plot4, 24 Bubble plot4, 24 Bubble plot4, 24 Bubble plot4, 24 Bubble plot4, 28 cartographie4, 28 cladistics6, 84 cladistique6, 84 cladistique6, 84 cladistique	analyse en composantes principales5, 43, 77 sv analyse en coordonnées principales5 analyse factorielle des correspondances5, 45 analyse spectrale. 3, 6, 66 sv, 69 analyses de forme6, 77 annuler une opération12 ANOSIM5, 51 ANOVA5, 38 sv, 51, 54 ARMA4, 6, 37, 70 associations unitaires6, 16, 80 autocorrélation3, 6, 67, 69 sv barres d'erreur4, 23 biodiversité3, 5, 61, 65, 87 bmp21 sv boites à moustaches4, 24 Bookstein18 Box-plot4, 24 Bubble plot4, 24 Bubble plot4, 26 CABFAC5, 46, 52 cartographie4, 28 sv cellules vides6, 84 cladistique6, 84 cladistique6, 84 cluster analysis5, 31, 48, 84 coefficient de variation36 colorer des lignes	analyse des similarités51
principales5, 43, 77 sv analyse en coordonnées principales5 analyse factorielle des correspondances5, 45 analyse spectrale. 3, 6, 66 sv, 69 analyses de forme6, 77 annuler une opération12 ANOSIM5, 38 sv, 51, 54 ARMA4, 6, 37, 70 associations unitaires. 6, 16, 80 autocorrélation3, 6, 67, 69 sv barres d'erreur4, 23 biodiversité3, 5, 61, 65, 87 bmp21 sv boites à moustaches4, 24 Bookstein18 Box-plot4, 24 Bubble plot4, 26 CABFAC5, 46, 52 cartographie4, 28 sv cellules vides5, 31, 48, 84 coefficient de variation36 colorer des lignes9 compartiments24 compteur au clavier16 CONOP6, 81	principales5, 43, 77 sv analyse en coordonnées principales5 analyse factorielle des correspondances5, 45 analyse spectrale. 3, 6, 66 sv, 69 analyses de forme6, 77 annuler une opération12 ANOSIM5, 38 sv, 51, 54 ARMA4, 6, 37, 70 associations unitaires. 6, 16, 80 autocorrélation3, 6, 67, 69 sv barres d'erreur4, 23 biodiversité3, 5, 61, 65, 87 bmp21 sv boites à moustaches4, 24 Bookstein18 Box-plot4, 24 Bubble plot4, 24 Bubble plot4, 26 CABFAC5, 46, 52 cartographie4, 28 sv cellules vides5, 31, 48, 84 coefficient de variation36 colorer des lignes9 compartiments24 compteur au clavier16 CONOP6, 81 corrélation croisée6, 74 corrélation croisée6, 67	analyse en composantes
analyse en coordonnées principales	analyse en coordonnées principales	principales5, 43, 77 sv
principales	principales	analyse en coordonnées
analyse factorielle des correspondances5, 45 analyse spectrale. 3, 6, 66 sv, 69 analyses de forme6, 77 annuler une opération12 ANOSIM5, 38 sv, 51, 54 ARMA5, 38 sv, 51, 54 ARMA4, 6, 37, 70 associations unitaires. 6, 16, 80 autocorrélation3, 6, 67, 69 sv barres d'erreur4, 23 biodiversité3, 5, 61, 65, 87 bmp21 sv boites à moustaches4, 24 Bookstein18 Box-plot4, 24 Bubble plot4, 26 CABFAC5, 46, 52 cartographie4, 28 sv cellules vides5, 31, 48, 84 coefficient de variation36 colorer des lignes	analyse factorielle des correspondances5, 45 analyse spectrale. 3, 6, 66 sv, 69 analyses de forme6, 77 annuler une opération12 ANOSIM5, 38 sv, 51, 54 ARMA5, 38 sv, 51, 54 ARMA4, 6, 37, 70 associations unitaires. 6, 16, 80 autocorrélation3, 6, 67, 69 sv barres d'erreur4, 23 biodiversité3, 5, 61, 65, 87 bmp21 sv boites à moustaches4, 24 Bookstein18 Box-plot4, 24 Bubble plot4, 24 Bubble plot4, 24 Bubble plot4, 26 CABFAC5, 46, 52 cartographie4, 28 sv cellules vides5, 31, 48, 84 coefficient de variation36 colorer des lignes9 compartiments24 compteur au clavier16 CONOP6, 81 corrélation angulaire6, 74 corrélation croisée6, 67	principales
correspondances5, 45 analyse spectrale. 3, 6, 66 sv, 69 analyses de forme6, 77 annuler une opération12 ANOSIM5, 38 sv, 51, 54 ARMA5, 38 sv, 51, 54 ARMA4, 6, 37, 70 associations unitaires. 6, 16, 80 autocorrélation3, 6, 67, 69 sv barres d'erreur4, 23 biodiversité3, 5, 61, 65, 87 bmp21 sv boites à moustaches4, 24 Bookstein18 Box-plot4, 24 Bubble plot4, 24 Bubble plot4, 26 CABFAC5, 46, 52 cartographie4, 28 sv cellules vides6, 84 cladistique6, 84 cladistique6, 84 cluster analysis5, 31, 48, 84 coefficient de variation36 colorer des lignes9 compartiments24 compteur au clavier16 CONOP6, 81	correspondances5, 45 analyse spectrale. 3, 6, 66 sv, 69 analyses de forme6, 77 annuler une opération12 ANOSIM5, 38 sv, 51, 54 ARMA5, 38 sv, 51, 54 ARMA4, 6, 37, 70 associations unitaires. 6, 16, 80 autocorrélation3, 6, 67, 69 sv barres d'erreur4, 23 biodiversité3, 5, 61, 65, 87 bmp21 sv boites à moustaches4, 24 Bookstein18 Box-plot4, 24 Bubble plot4, 24 Bubble plot4, 24 Bubble plot4, 24 Bubble plot4, 26 CABFAC5, 46, 52 cartographie4, 28 sv cellules vides5, 31, 48, 84 coefficient de variation36 colorer des lignes9 compartiments24 compteur au clavier16 CONOP6, 81 corrélation angulaire6, 74 corrélation croisée6, 67	analyse factorielle des
analyse spectrale. 3, 6, 66 sv, 69 analyses de forme	analyse spectrale. 3, 6, 66 sv, 69 analyses de forme	correspondances
analyses de forme	analyses de forme	analyse spectrale. 3, 6, 66 sv, 69
annuler une opération	annuler une opération	analyses de forme
ANOSIM	ANOSIM5, 38 sv, 51, 54 ARMA5, 38 sv, 51, 54 ARMA4, 6, 37, 70 associations unitaires6, 16, 80 autocorrélation3, 6, 67, 69 sv barres d'erreur4, 23 biodiversité3, 5, 61, 65, 87 bmp21 sv boites à moustaches4, 24 Bookstein18 Box-plot4, 24 Bubble plot4, 24 Bubble plot4, 24 Bubble plot4, 26 CABFAC5, 46, 52 cartographie4, 28 sv cellules vides5, 31, 48, 84 cladistique6, 84 cluster analysis5, 31, 48, 84 coefficient de variation36 colorer des lignes9 compartiments	annuler une opération 12
ANOVA	ANOVA5, 38 sv, 51, 54 ARMA4, 6, 37, 70 associations unitaires6, 16, 80 autocorrélation3, 6, 67, 69 sv barres d'erreur4, 23 biodiversité3, 5, 61, 65, 87 bmp21 sv boites à moustaches4, 24 Bookstein18 Box-plot4, 24 Bubble plot4, 24 Bubble plot4, 26 CABFAC5, 46, 52 cartographie4, 28 sv cellules vides8 sv Cladistics6, 84 cladistique6, 84 cluster analysis5, 31, 48, 84 coefficient de variation36 colorer des lignes9 compartiments24 compteur au clavier16 CONOP6, 81 corrélation angulaire6, 74 corrélation croisée6, 67	ANOSIM
ARMA4, 6, 37, 70 associations unitaires. .6, 16, 80 autocorrélation3, 6, 67, 69 sv barres d'erreur4, 23 biodiversité3, 5, 61, 65, 87 bmp21 sv boites à moustaches4, 24 Bookstein18 Box-plot4, 24 Bubble plot4, 24 Bubble plot4, 24 Bubble plot4, 26 CABFAC5, 46, 52 cartographie4, 28 sv cellules vides6, 84 cladistique6, 84 cladistique	ARMA4, 6, 37, 70 associations unitaires. .6, 16, 80 autocorrélation3, 6, 67, 69 sv barres d'erreur4, 23 biodiversité3, 5, 61, 65, 87 bmp21 sv boites à moustaches4, 24 Bookstein18 Box-plot4, 24 Bubble plot4, 24 Bubble plot4, 24 Bubble plot4, 26 CABFAC5, 46, 52 cartographie4, 28 sv cellules vides6, 84 cladistics6, 84 cladistique6, 84 coorfficient de variation36 colorer des lignes	ANOVA 5 38 sv 51 54
associations unitaires6, 16, 80 autocorrélation3, 6, 67, 69 sv barres d'erreur4, 23 biodiversité3, 5, 61, 65, 87 bmp21 sv boites à moustaches4, 24 Bookstein18 Box-plot4, 24 Bubble plot4, 24 Bubble plot4, 24 Bubble plot4, 26 CABFAC5, 46, 52 cartographie4, 28 sv cellules vides5, 31, 48, 84 coefficient de variation36 colorer des lignes9 compteur au clavier16 CONOP	associations unitaires6, 16, 80 autocorrélation3, 6, 67, 69 sv barres d'erreur4, 23 biodiversité3, 5, 61, 65, 87 bmp21 sv boites à moustaches4, 24 Bookstein18 Box-plot4, 24 Bubble plot4, 24 Bubble plot4, 24 Bubble plot4, 26 CABFAC5, 46, 52 cartographie4, 28 sv cellules vides8 sv Cladistics6, 84 cladistique6, 84 cluster analysis5, 31, 48, 84 coefficient de variation36 colorer des lignes9 compartiments24 compteur au clavier16 CONOP6, 81 corrélation angulaire6, 74 corrélation croisée6, 67	
autocorrélation3, 6, 67, 69 sv barres d'erreur4, 23 biodiversité3, 5, 61, 65, 87 bmp21 sv boites à moustaches4, 24 Bookstein18 Box-plot4, 24 Bubble plot4, 24 Bubble plot4, 24 Bubble plot4, 26 CABFAC5, 46, 52 cartographie4, 28 sv cellules vides6, 84 cladistics6, 84 cladistique5, 31, 48, 84 coefficient de variation	autocorrélation3, 6, 67, 69 sv barres d'erreur4, 23 biodiversité3, 5, 61, 65, 87 bmp21 sv boites à moustaches4, 24 Bookstein18 Box-plot4, 24 Bubble plot4, 24 Bubble plot4, 26 CABFAC5, 46, 52 cartographie4, 28 sv cellules vides8 sv Cladistics6, 84 cladistique6, 84 cluster analysis5, 31, 48, 84 coefficient de variation36 colorer des lignes9 compartiments24 compteur au clavier16 CONOP6, 81 corrélation angulaire6, 74 corrélation croisée6, 67	ARMA
barres d'erreur4, 23 biodiversité3, 5, 61, 65, 87 bmp21 sv boites à moustaches4, 24 Bookstein18 Box-plot4, 24 Bubble plot4, 24 Bubble plot4, 26 CABFAC5, 46, 52 cartographie4, 28 sv cellules vides8 sv Cladistics6, 84 cladistique6, 84 cluster analysis5, 31, 48, 84 coefficient de variation36 colorer des lignes9 compartiments24 compteur au clavier16 CONOP6, 81	barres d'erreur4, 23 biodiversité3, 5, 61, 65, 87 bmp21 sv boites à moustaches4, 24 Bookstein18 Box-plot4, 24 Bubble plot4, 24 Bubble plot4, 26 CABFAC5, 46, 52 cartographie4, 28 sv cellules vides8 sv Cladistics6, 84 cladistique6, 84 cluster analysis5, 31, 48, 84 coefficient de variation36 colorer des lignes9 compartiments24 compteur au clavier16 CONOP6, 81 corrélation angulaire6, 74 corrélation croisée6, 67	ARMA4, 6, 37, 70 associations unitaires6, 16, 80
biodiversité3, 5, 61, 65, 87 bmp21 sv boites à moustaches4, 24 Bookstein18 Box-plot4, 24 Bubble plot4, 24 Bubble plot4, 24 Bubble plot4, 26 CABFAC5, 46, 52 cartographie4, 28 sv cellules vides5, 46, 52 cartographie4, 28 sv cellules vides5, 84 cladistics6, 84 cladistique5, 31, 48, 84 coefficient de variation36 colorer des lignes9 compartiments	biodiversité3, 5, 61, 65, 87 bmp21 sv boites à moustaches4, 24 Bookstein18 Box-plot4, 24 Bubble plot4, 24 Bubble plot4, 26 CABFAC5, 46, 52 cartographie4, 28 sv cellules vides8 sv Cladistics6, 84 cladistique6, 84 cluster analysis5, 31, 48, 84 coefficient de variation36 colorer des lignes9 compartiments24 compteur au clavier16 CONOP6, 81 corrélation angulaire6, 74 corrélation croisée6, 67	ARMA4, 6, 37, 70 associations unitaires6, 16, 80 autocorrélation3, 6, 67, 69 sv
bmp.21 svboites à moustaches.4, 24Bookstein.18Box-plot.4, 24Bubble plot.4, 26CABFAC.5, 46, 52cartographie.4, 28 svcellules vides.8 svCladistics.6, 84cladistique.6, 84cluster analysis.5, 31, 48, 84coefficient de variation.36colorer des lignes.9compartiments.24CONOP.6, 81	bmp.21 svboites à moustaches.4, 24Bookstein.18Box-plot.4, 24Bubble plot.4, 26CABFAC.5, 46, 52cartographie.4, 28 svcellules vides.8 svCladistics.6, 84cluster analysis.5, 31, 48, 84coefficient de variation.36colorer des lignes.9compartiments.24compteur au clavier.16CONOP.6, 81corrélation angulaire.6, 67	ARMA4, 6, 37, 70 associations unitaires6, 16, 80 autocorrélation3, 6, 67, 69 sv barres d'erreur4, 23
boites à moustaches4, 24 Bookstein18 Box-plot4, 24 Bubble plot4, 24 Bubble plot4, 26 CABFAC5, 46, 52 cartographie4, 28 sv cellules vides8 sv Cladistics6, 84 cladistique6, 84 cluster analysis5, 31, 48, 84 coefficient de variation36 colorer des lignes9 compartiments24 compteur au clavier16 CONOP6, 81	boites à moustaches4, 24 Bookstein18 Box-plot4, 24 Bubble plot4, 24 Bubble plot4, 26 CABFAC5, 46, 52 cartographie4, 28 sv cellules vides8 sv Cladistics6, 84 cladistique6, 84 cladistique6, 84 cluster analysis5, 31, 48, 84 coefficient de variation36 colorer des lignes9 compartiments24 compteur au clavier16 CONOP6, 81 corrélation angulaire6, 74 corrélation croisée6, 67	ARMA4, 6, 37, 70 associations unitaires6, 16, 80 autocorrélation3, 6, 67, 69 sv barres d'erreur4, 23 biodiversité3, 5, 61, 65, 87
Bookstein	Bookstein.18Box-plot.4, 24Bubble plot.4, 26CABFAC.5, 46, 52cartographie.4, 28 svcellules vides.8 svCladistics.6, 84cladistique.6, 84cluster analysis.5, 31, 48, 84coefficient de variation.36colorer des lignes.9compartiments.24compteur au clavier.16CONOP.6, 81corrélation angulaire.6, 67	ARMA4, 6, 37, 70 associations unitaires6, 16, 80 autocorrélation3, 6, 67, 69 sv barres d'erreur4, 23 biodiversité3, 5, 61, 65, 87 bmp21 sv
Box-plot4, 24 Bubble plot4, 26 CABFAC5, 46, 52 cartographie4, 28 sv cellules vides8 sv Cladistics6, 84 cladistique6, 84 cluster analysis5, 31, 48, 84 coefficient de variation36 colorer des lignes9 compartiments24 compteur au clavier16 CONOP6, 81	Box-plot4, 24 Bubble plot4, 26 CABFAC5, 46, 52 cartographie4, 28 sv cellules vides8 sv Cladistics6, 84 cladistique6, 84 cluster analysis5, 31, 48, 84 coefficient de variation36 colorer des lignes9 compartiments24 compteur au clavier16 CONOP6, 81 corrélation angulaire6, 74 corrélation croisée6, 67	ARMA4, 6, 37, 70 associations unitaires6, 16, 80 autocorrélation3, 6, 67, 69 sv barres d'erreur4, 23 biodiversité3, 5, 61, 65, 87 bmp21 sv boites à moustaches4, 24
Bubble plot4, 26 CABFAC5, 46, 52 cartographie4, 28 sv cellules vides8 sv Cladistics6, 84 cladistique6, 84 cluster analysis5, 31, 48, 84 coefficient de variation36 colorer des lignes9 compartiments24 compteur au clavier16 CONOP6, 81	Bubble plot4, 26 CABFAC5, 46, 52 cartographie4, 28 sv cellules vides4, 28 sv Cladistics6, 84 cladistique6, 84 cluster analysis5, 31, 48, 84 coefficient de variation36 colorer des lignes9 compartiments24 compteur au clavier16 CONOP6, 81 corrélation angulaire6, 74 corrélation croisée6, 67	ARMA4, 6, 37, 70 associations unitaires6, 16, 80 autocorrélation3, 6, 67, 69 sv barres d'erreur4, 23 biodiversité3, 5, 61, 65, 87 bmp21 sv boites à moustaches4, 24 Bookstein18
CABFAC	CABFAC	ARMA4, 6, 37, 70 associations unitaires. 6, 16, 80 autocorrélation3, 6, 67, 69 sv barres d'erreur4, 23 biodiversité3, 5, 61, 65, 87 bmp21 sv boites à moustaches4, 24 Bookstein18 Box-plot4, 24
cartographie4, 28 sv cellules vides4, 28 sv Cladistics6, 84 cladistique6, 84 cluster analysis5, 31, 48, 84 coefficient de variation36 colorer des lignes9 compartiments24 compteur au clavier16 CONOP6, 81	cartographie4, 28 sv cellules vides4, 28 sv Cladistics6, 84 cladistique6, 84 cluster analysis5, 31, 48, 84 coefficient de variation36 colorer des lignes9 compartiments24 compteur au clavier16 CONOP6, 81 corrélation angulaire6, 74 corrélation croisée6, 67	ARMA4, 6, 37, 70 associations unitaires6, 16, 80 autocorrélation3, 6, 67, 69 sv barres d'erreur4, 23 biodiversité3, 5, 61, 65, 87 bmp21 sv boites à moustaches4, 24 Bookstein18 Box-plot4, 24 Bubble plot4, 26
cellules vides	cellules vides	ARMA4, 6, 37, 70 associations unitaires6, 16, 80 autocorrélation3, 6, 67, 69 sv barres d'erreur4, 23 biodiversité3, 5, 61, 65, 87 bmp21 sv boites à moustaches4, 24 Bookstein18 Box-plot4, 24 Bubble plot4, 26 CABFAC5, 46, 52
Cladistics	Cladistics	ARMA4, 6, 37, 70 associations unitaires. autocorrélation3, 6, 67, 69 sv barres d'erreur4, 23 biodiversité3, 5, 61, 65, 87 bmp21 sv boites à moustaches4, 24 Bookstein18 Box-plot4, 24 Bubble plot4, 24 CABFAC5, 46, 52 cartographie4, 28 sv
cladistique	cladistique	ARMA4, 6, 37, 70 associations unitaires6, 16, 80 autocorrélation3, 6, 67, 69 sv barres d'erreur4, 23 biodiversité3, 5, 61, 65, 87 bmp21 sv boites à moustaches4, 24 Bookstein18 Box-plot4, 24 Bubble plot4, 24 CABFAC5, 46, 52 cartographie4, 28 sv cellules vides8
cluster analysis5, 31, 48, 84 coefficient de variation36 colorer des lignes9 compartiments24 compteur au clavier16 CONOP6, 81	cluster analysis5, 31, 48, 84 coefficient de variation36 colorer des lignes9 compartiments	ARMA4, 6, 37, 70 associations unitaires6, 16, 80 autocorrélation3, 6, 67, 69 sv barres d'erreur4, 23 biodiversité3, 5, 61, 65, 87 bmp21 sv boites à moustaches4, 24 Bookstein18 Box-plot4, 24 Bubble plot4, 24 Bubble plot4, 24 CABFAC5, 46, 52 cartographie4, 28 sv cellules vides
coefficient de variation	coefficient de variation	ARMA4, 6, 37, 70 associations unitaires. autocorrélation3, 6, 67, 69 sv barres d'erreur4, 23 biodiversité3, 5, 61, 65, 87 bmp21 sv boites à moustaches4, 24 Bookstein18 Box-plot4, 24 Bubble plot4, 24 Bubble plot4, 24 CABFAC5, 46, 52 cartographie4, 28 sv cellules vides8 sv Cladistics
colorer des lignes	colorer des lignes	ARMA
compartiments	compartiments	ARMA4, 6, 37, 70associations unitaires6, 16, 80autocorrélation3, 6, 67, 69 svbarres d'erreur4, 23biodiversité3, 5, 61, 65, 87bmp21 svboites à moustaches4, 24Bookstein18Box-plot4, 24Bubble plot4, 24Bubble plot4, 24Bubble plot4, 26CABFAC5, 46, 52cartographie4, 28 svcellules vides6, 84cladistique6, 84cluster analysis5, 31, 48, 84coefficient de variation
compteur au clavier16 CONOP6, 81	compteur au clavier	ARMA4, 6, 37, 70associations unitaires6, 16, 80autocorrélation3, 6, 67, 69 svbarres d'erreur4, 23biodiversité3, 5, 61, 65, 87bmp21 svboites à moustaches4, 24Bookstein18Box-plot4, 24Bubble plot4, 24Bubble plot4, 24Bubble plot4, 26CABFAC5, 46, 52cartographie4, 28 svcellules vides8 svCladistics6, 84cladistique5, 31, 48, 84coefficient de variation36colorer des lignes
CONOP	CONOP	ARMA4, 6, 37, 70associations unitaires.autocorrélation3, 6, 67, 69 svbarres d'erreur4, 23biodiversité3, 5, 61, 65, 87bmp21 svboites à moustaches4, 24Bookstein18Box-plot4, 24Bubble plot4, 24Bubble plot4, 24Bubble plot4, 26CABFAC5, 46, 52cartographie4, 28 svcellules vides6, 84cladistique6, 84cluster analysis5, 31, 48, 84coofficient de variation36colorer des lignes9compartiments24
0	corrélation angulaire6, 74 corrélation croisée6, 67	ARMA4, 6, 37, 70associations unitaires6, 16, 80autocorrélation3, 6, 67, 69 svbarres d'erreur4, 23biodiversité3, 5, 61, 65, 87bmp21 svboites à moustaches4, 24Bookstein18Box-plot4, 24Bubble plot4, 24Bubble plot4, 24Bubble plot4, 26CABFAC5, 46, 52cartographie4, 28 svcellules vides5, 31, 48, 84coefficient de variation36colorer des lignes9compartiments24compteur au clavier16
corrélation angulaire 6 74	corrélation croisée6, 67	ARMA
corrélation croisée		ARMA4, 6, 37, 70associations unitaires. 6, 16, 80autocorrélation3, 6, 67, 69 svbarres d'erreur4, 23biodiversité3, 5, 61, 65, 87bmp21 svboites à moustaches4, 24Bookstein18Box-plot4, 24Bubble plot4, 24Bubble plot4, 26CABFAC5, 46, 52cartographie4, 28 svcellules vides6, 84cladistique6, 84cladistique6, 84coorfficient de variation36colorer des lignes9compteur au clavier16CONOP6, 81corrélation angulaire6, 74
	correlations16, 31, 43, 51, 81	ARMA4, 6, 37, 70associations unitaires6, 16, 80autocorrélation3, 6, 67, 69 svbarres d'erreur4, 23biodiversité3, 5, 61, 65, 87bmp21 svboites à moustaches4, 24Bookstein18Box-plot4, 24Bubble plot4, 24Bubble plot4, 24Bubble plot4, 26CABFAC5, 46, 52cartographie4, 28 svcellules vides5, 31, 48, 84coefficient de variation36colorer des lignes9compteur au clavier16CONOP6, 81corrélation angulaire6, 74corrélation croisée6, 67

correlogramme6, /	0
courbe de survie2	7
courbe logistique5, 5	8
courbes spline5, 5	8
covariance4 sv, 31, 40, 43, 5	0
sv, 67, 86	
déformations relatives7	8
diagramme de fréquence4, 2	5
diagramme ternaire4, 2	6
diagramme triangulaire4, 2	6
distance taxonomique6, 6	2
distances. 31, 41, 44, 51, 70, 7	4
sv, 79, 86	
distribution normale4, 23 sv	ν.
30, 32, 50	ĺ
diversité interlieux6, 6	2
échantillons appariés4, 3	3
Edit labels7. 13 s	v
Edit mode 7 sv 1	3
EDMA 7	9
emf	v
évaluer une expression	•
mathématique 1	9
Excel 2.7.9s	v
Gnumeric	7
Hammer 2 31 8	, 7
	'
histogramme de fréquence 4 2	4
histogramme de fréquence 4, 2 Hotelling 5 50 s	4 v
histogramme de fréquence 4, 2 Hotelling5, 50 s imprimer 11 21 s	4 v v
histogramme de fréquence 4, 2 Hotelling5, 50 s imprimer11, 21 s indice de Mao-tau 6 6	4 v v 3
histogramme de fréquence 4, 2 Hotelling5, 50 s imprimer11, 21 s indice de Mao-tau6, 6 insérer un fichier	4 v v 3 0
histogramme de fréquence 4, 2 Hotelling5, 50 s imprimer11, 21 s indice de Mao-tau6, 6 insérer un fichier1 interpolation 3 6 19 7	4 v 3 0 6
histogramme de fréquence 4, 2 Hotelling5, 50 s imprimer11, 21 s indice de Mao-tau6, 6 insérer un fichier1 interpolation3, 6, 19, 7 interpolations spatiales 6, 7	4 v 3 0 6 6
histogramme de fréquence 4, 2 Hotelling5, 50 s imprimer11, 21 s indice de Mao-tau6, 6 insérer un fichier1 interpolation3, 6, 19, 7 interpolations spatiales6, 7 ing 21 s	4 v 3 0 6 v
histogramme de fréquence 4, 2 Hotelling5, 50 s imprimer11, 21 s indice de Mao-tau6, 6 insérer un fichier1 interpolation3, 6, 19, 7 interpolations spatiales6, 7 jpg21 s krigeage 7	4 v 3 0 6 v 6
histogramme de fréquence 4, 2 Hotelling5, 50 s imprimer11, 21 s indice de Mao-tau6, 6 insérer un fichier1 interpolation3, 6, 19, 7 interpolations spatiales6, 7 jpg21 s krigeage7 Kruskal-Wallis 5, 3	4 v v 3 0 6 6 v 6 9
histogramme de fréquence 4, 2 Hotelling5, 50 s imprimer11, 21 s indice de Mao-tau6, 6 insérer un fichier1 interpolation3, 6, 19, 7 interpolations spatiales6, 7 jpg21 s krigeage7 Kruskal-Wallis5, 3 landmarks 4, 18, 28, 78 s	4 v v 3 0 6 6 v 6 9 v
histogramme de fréquence 4, 2 Hotelling5, 50 s imprimer11, 21 s indice de Mao-tau6, 6 insérer un fichier1 interpolation3, 6, 19, 7 interpolations spatiales6, 7 jpg21 s krigeage7 Kruskal-Wallis5, 3 landmarks4, 18, 28, 78 s linux 2, 4, 2	4 v v 3 0 6 6 v 6 9 v 1
histogramme de fréquence 4, 2 Hotelling5, 50 s imprimer11, 21 s indice de Mao-tau6, 6 insérer un fichier1 interpolation3, 6, 19, 7 interpolations spatiales6, 7 jpg21 s krigeage7 Kruskal-Wallis5, 3 landmarks4, 18, 28, 78 s linux2, 4, 2 lissage5, 23 sy 58 sy 76, 7	4 v v 3 0 6 6 v 6 9 v 1 8
histogramme de fréquence 4, 2 Hotelling	4 v v 3 0 6 6 v 6 9 v 1 8 7
histogramme de fréquence 4, 2 Hotelling5, 50 s imprimer11, 21 s indice de Mao-tau6, 6 insérer un fichier1 interpolation3, 6, 19, 7 interpolations spatiales6, 7 jpg21 s krigeage7 Kruskal-Wallis5, 3 landmarks4, 18, 28, 78 s linux2, 4, 2 lissage5, 23 sv, 58 sv, 76, 7 logarithme18 sv, 23, 54, 7 Mann-Whitney 4, 36 s	4 v v 3 0 6 6 v 6 9 v 1 8 7 v
histogramme de fréquence 4, 2 Hotelling	4 v v 3 0 6 6 v 6 9 v 1 8 7 v 0
histogramme de fréquence 4, 2 Hotelling5, 50 s imprimer11, 21 s indice de Mao-tau6, 6 insérer un fichier1 interpolation3, 6, 19, 7 interpolations spatiales6, 7 jpg21 s krigeage7 Kruskal-Wallis5, 3 landmarks4, 18, 28, 78 s linux2, 4, 2 lissage5, 23 sv, 58 sv, 76, 7 logarithme18 sv, 23, 54, 7 Mann-Whitney4, 36 s Mantel5 sv, 51, 7 mesures géométriques 3, 6, 7	4 v v 3 0 6 6 v 6 9 v 1 8 7 v 0 2
histogramme de fréquence 4, 2 Hotelling5, 50 s imprimer11, 21 s indice de Mao-tau6, 6 insérer un fichier1 interpolation3, 6, 19, 7 interpolations spatiales6, 7 jpg21 s krigeage7 Kruskal-Wallis5, 3 landmarks4, 18, 28, 78 s linux2, 4, 2 lissage5, 23 sv, 58 sv, 76, 7 logarithme18 sv, 23, 54, 7 Mann-Whitney4, 36 s Mantel5 sv, 51, 7 mesures géométriques3, 6, 7 modélisation 3, 5, 53, 5	4 v v 3 0 6 6 v 6 9 v 1 8 7 v 0 2 8
histogramme de fréquence 4, 2 Hotelling5, 50 s imprimer11, 21 s indice de Mao-tau6, 6 insérer un fichier1 interpolation3, 6, 19, 7 interpolations spatiales6, 7 jpg21 s krigeage7 Kruskal-Wallis6, 7 Jpg21 s krigeage7 Kruskal-Wallis5, 3 landmarks4, 18, 28, 78 s linux2, 4, 2 lissage5, 23 sv, 58 sv, 76, 7 logarithme18 sv, 23, 54, 7 Mann-Whitney4, 36 s Mantel5 sv, 51, 7 mesures géométriques3, 6, 7 modélisation3, 5, 53, 5 nucléotides31	4 v v 3 0 6 6 v 6 9 v 1 8 7 v 0 2 8 1
histogramme de fréquence 4, 2 Hotelling5, 50 s imprimer11, 21 s indice de Mao-tau6, 6 insérer un fichier1 interpolation3, 6, 19, 7 interpolations spatiales6, 7 jpg21 s krigeage7 Kruskal-Wallis5, 3 landmarks4, 18, 28, 78 s linux2, 4, 2 lissage5, 23 sv, 58 sv, 76, 7 logarithme18 sv, 23, 54, 7 Mann-Whitney4, 36 s Mantel5 sv, 51, 7 mesures géométriques3, 6, 7 modélisation3, 5, 53, 5 nucléotides3, 4 ondelettes 6, 6	4 v v 3 0 6 6 v 6 9 v 1 8 7 v 0 2 8 1 8
histogramme de fréquence 4, 2 Hotelling	4 v v 3 0 6 6 v 6 9 v 1 8 7 v 0 2 8 1 8 2
histogramme de fréquence 4, 2 Hotelling	4 v v 3 0 6 6 v 6 9 v 1 8 7 v 0 2 8 1 8 2 v
histogramme de fréquence 4, 2 Hotelling	4 v v 3 0 6 6 v 6 9 v 1 8 7 v 0 2 8 1 8 2 v v
histogramme de fréquence 4, 2 Hotelling	4 v v 3 0 6 6 v 6 9 v 1 8 7 v 0 2 8 1 8 2 v v 0
histogramme de fréquence 4, 2 Hotelling	4 v v 3 0 6 6 v 6 9 v 1 8 7 v 0 2 8 1 8 2 v v 0 1
histogramme de fréquence 4, 2 Hotelling	4 v v 3 0 6 6 v 6 9 v 1 8 7 v 0 2 8 1 8 2 v v 0 1 5

56, 67
point d'interrogation8, 30, 37,
52
points de repère4, 28, 78 sv
positionnement
multidimensionnel5, 45
pourcentage5, 18, 36, 43, 52
Procrustes18, 77
Procuste77 sv
programmation3, 6, 85
quadrats6, 61 sv
régression linéaire3, 5, 18, 53
SV
Regroupements en arbres5, 48
renommer des lignes ou des
colonnes13
script3, 6, 14, 85
sélectionner une zone8
séquences de nucléotides41
séries temporelles3, 6, 66, 70
statistiques multivariables3, 5,
43
stratigraphie3, 16, 80
supprimer une ligne ou une
colonne12
tendance
test de normalite4 sv, 34, 50
test du kni deux4, 35
test F
test 1
Tracer des graphiques 2 gy 22
Transformer les données 4, 19
transposor 2 14 su
tri 2 gy 12 gy 16 gy 22 gy 28
$\begin{array}{c} \text{un 5 sv, 15 sv, 10 sv, 25 sv, 26,} \\ \text{20 sv, 42 45 47 sv, 50 sv,} \end{array}$
50 sv, 45, 45, 47 sv, 50 sv, 60 72 sv, 80 82 sv, 86
1250, 7250, 80, 8250, 80
Walsh 6 60
Windows 2 7 0 21
WINE 2 21
WINL