

UNIWIN VERSION 9.7.0

ANALYSE FACTORIELLE DISCRIMINANTE

Révision : 02/09/2023

Définition.....	1
Entrée des données	2
Données manquantes	2
Exemple 1 : Fichier IRIS3.....	3
L'option Rapports	7
L'option Graphiques	8
Exemple 2 : Fichier INFARCT2	12
L'option Rapports	15
L'option Graphiques	16
Exemple 3 : Fichier BORDEAUX.....	18
Exemple 4 : Fichier TITANIC	20
Ellipses de tolérance	22
Les variables internes créées par la procédure	24

Définition

L'Analyse Factorielle Discriminante (AFD) est une méthode géométrique permettant de construire à partir d'un ensemble de variables quantitatives et d'une variable qualitative découpant la population en plusieurs groupes (2 ou plus) des fonctions discriminantes qui les séparent au mieux dans l'échantillon d'apprentissage.

En fonction des données et des paramètres définis par l'utilisateur, l'analyse AFD réalise automatiquement les études de la population d'apprentissage et des éventuelles populations de validation et de prévision.

Les données des variables explicatives sont automatiquement centrées et réduites.

De façon plus précise, la méthode peut se décomposer en trois étapes. Supposons une population de n individus. Découpons cette population en trois sous-populations de tailles n_1 , n_2 et n_3 avec $n_1 + n_2 + n_3 = n$. Les trois étapes sont :

- une étude initiale sur la population d'apprentissage de taille n_1
- une étude de validation sur la population de validation de taille n_2
- une étude prospective sur une population de prévision de taille n_3

Des tableaux résumés et détaillés des classements sont calculés. Le tracé de plans factoriels et un rapport général de synthèse sont proposés.

Entrée des données

Cliquons sur l'icône AFD dans le ruban Expliquer. La boîte de dialogue montrée ci-après apparaît.

Cette boîte de dialogue permet de préciser la variable qualitative définissant facteur de classement, les variables explicatives quantitatives, la variable optionnelle contenant les libellés des variables explicatives et la variable optionnelle contenant les libellés des individus.

The image shows a software dialog box titled "Analyse factorielle discriminante". It features a large empty list box on the left side. On the right side, there are four input fields, each preceded by a selection icon (a square with a right-pointing triangle). The labels for these fields are: "Facteur de classement:", "Variables explicatives quantitatives:", "(Libellés des variables explicatives:)", and "(Libellés des individus:)". At the bottom of the dialog, there are five buttons: "Ok", "Annuler", "Sélection", "Supprimer", and "Aide".

Données manquantes

Les données manquantes ne sont pas autorisées pour le facteur de classement.

Elles sont autorisées pour les variables quantitatives.

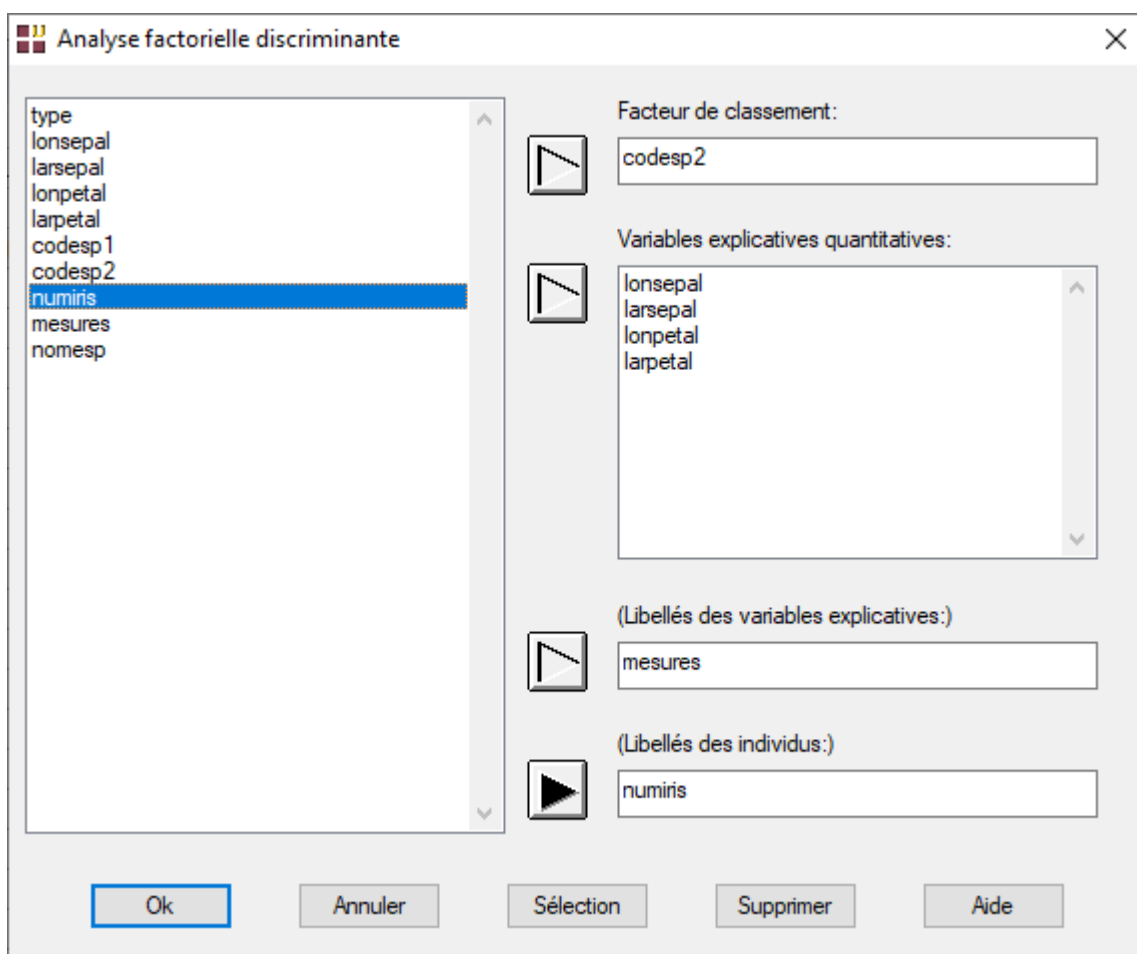
Exemple 1 : Fichier IRIS3

Nous utiliserons le fichier IRIS3 pour illustrer cette procédure. Ce fichier contient pour 150 iris de trois espèces différentes les mesures des quatre caractéristiques suivantes exprimées en millimètres : longueur du sépale, largeur du sépale, longueur du pétale et largeur du pétale

Les trois espèces sont : Iris Setosa, Iris Versicolor et Iris Virginica

Cliquons sur l'icône AFD dans le menu Expliquer.

La boîte de dialogue montrée ci-après s'affiche.



Choisissons les variables *lonsepal* à *larpetal* comme variables quantitatives, la variable *codesp2* comme facteur de classement, la variable *mesures* comme variable contenant les libellés des variables quantitatives, la variable *numiris* comme variable contenant les libellés des individus.

Cliquons sur le bouton Sélection pour définir la population d'apprentissage.

Définition de la sélection

Et	type	=	A
----	------	---	---

Liaison	Variable	Relation	Valeur ou variable
Et	larsepal	=	codesp1
Et non	lonpetal	<>	codesp2
Ou	lonsepal	<	larpetal
Ou non	mesures	<=	larsepal
	nomesp	>	lonpetal
	numiris	>=	lonsepal
	type	début	mesures

Ok Annuler Ajouter Aide

Définition de la sélection

La sélection définie comporte 120 éléments.

OK

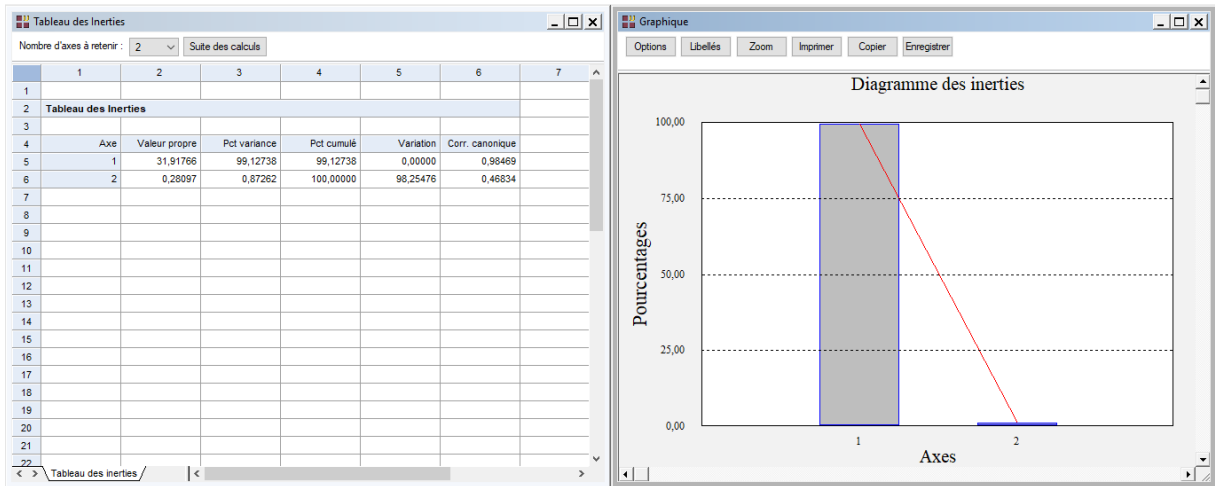
Cette sélection comporte 120 individus qui constituent la population d'apprentissage.

Les individus non sélectionnés pour lesquels les valeurs du facteur de classement sont connues constituent la population de validation.

Les individus non sélectionnés pour lesquels les valeurs du facteur de classement ne sont pas connues constituent la population de prévision.

Après avoir renseigné cette boîte de dialogue, UNIWIN débute le calcul de l'Analyse Factorielle Discriminante.

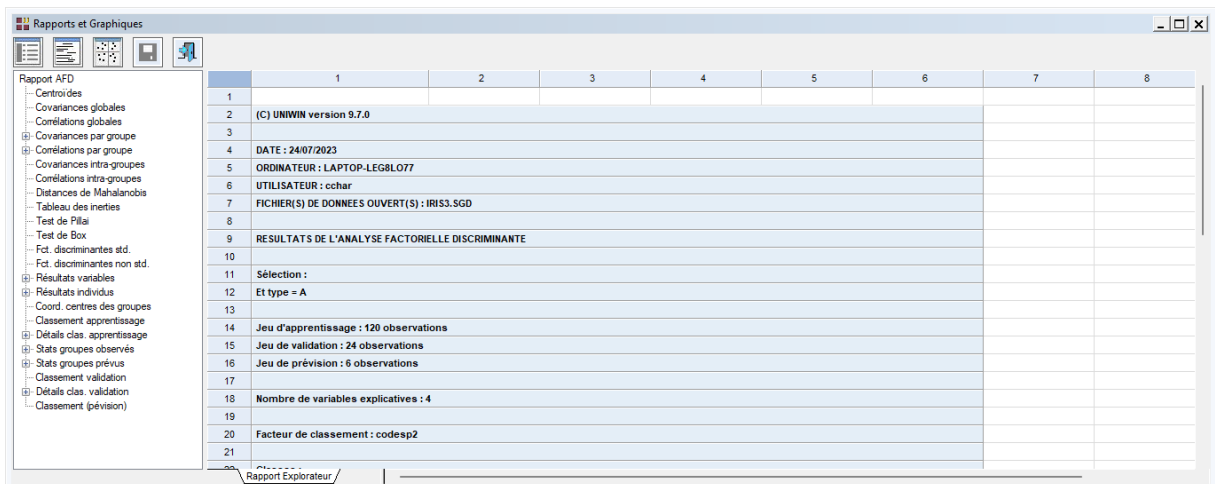
Après quelques instants, un tableau précisant l'inertie expliquée par les différents vecteurs propres issus de l'analyse apparaît ainsi qu'un diagramme des pourcentages d'inertie expliquée par chacun des axes.




L'option 'Nombre d'axes à retenir' permet de préciser le nombre de composantes principales à extraire.

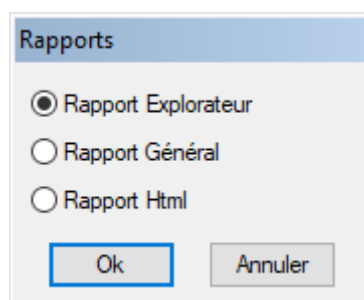
Clicquons sur le bouton 'Suite des calculs'.


Après quelques instants, l'écran suivant s'affiche :

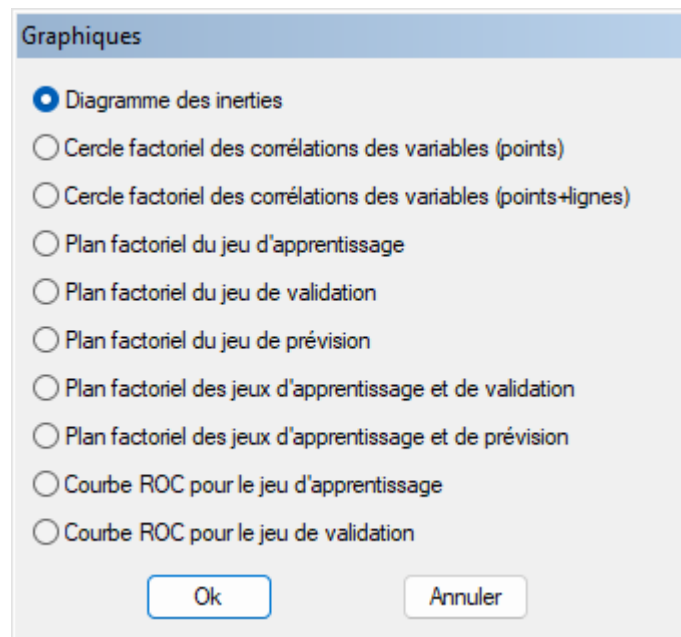



La barre d'outils 'Rapports et Graphiques' permet par l'icône 'Données'  de rappeler la boîte de dialogue d'entrée des données.

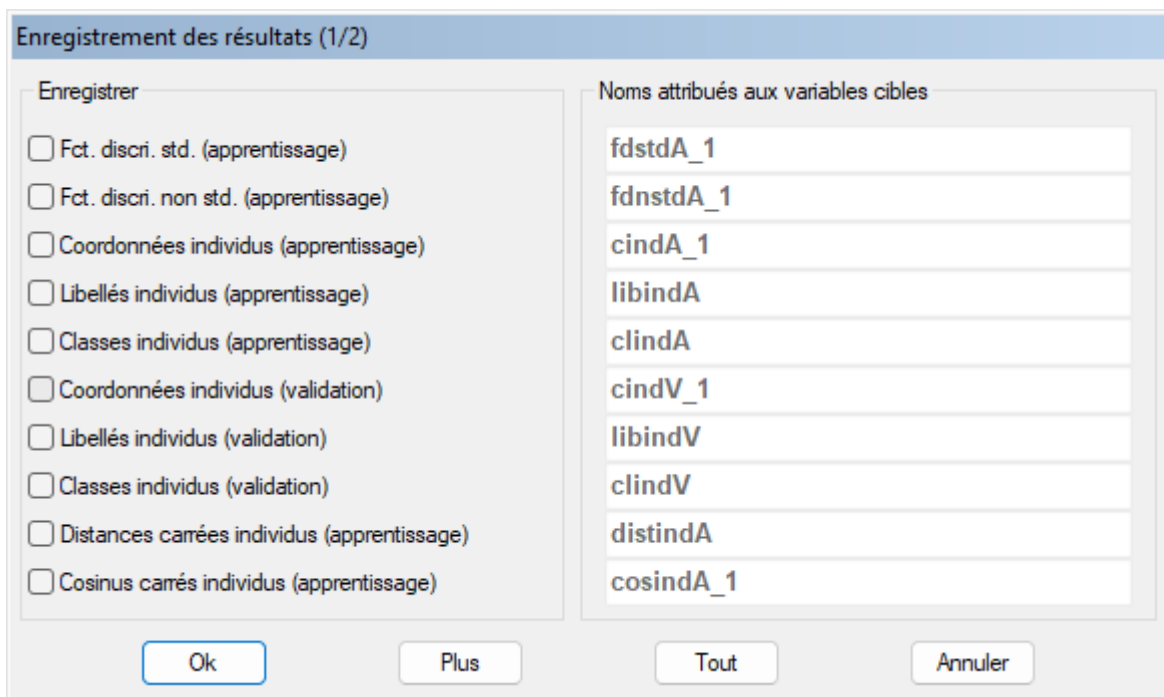
L'icône 'Rapports'  affiche la boîte de dialogue des options pour les rapports :



et par l'icône 'Graphiques'  affiche la boîte de dialogue des options pour les graphiques :



L'icône 'Enregistrer'  permet de sélectionner les résultats de l'analyse à enregistrer dans un fichier.



Note : le bouton 'Plus' permet d'afficher la suite de la liste des variables.

L'icône 'Quitter'  permet de quitter l'analyse.

L'option Rapports

Cette option permet d'obtenir le rapport à l'écran sous la forme d'un explorateur, d'un tableur ou au format HTML.

Voici trois exemples du rapport pour notre AFD : Explorateur, Général, HTML.

	1	2	3	4	5	6	7	8
1								
2	RESULTATS INDIVIDUS POUR LE FACTEUR : 1							
3								
4	DISTANCE^2 = CARRÉS DES DISTANCES A L'ORIGINE OU AU BARYCENTRE							
5	COORDONNEE = COORDONNÉES DES INDIVIDUS							
6	CONTRIBUTION = CONTRIBUTIONS A L'INERTIE							
7	COSINUS^2 = COSINUS CARRÉS							
8	COS^2 CUMULE = SOMMES CUMULÉES DES COSINUS CARRÉS							
9								
10								
11		GROUPE	INDIVIDU	DISTANCE^2	COORDONNEE	CONTRIBUTION	COSINUS^2	COS^2 CUMULE
12	1	1	1	57,32730	7,56828	1,48724	0,99916	0,99916
13	2	1	2	45,52882	6,89560	1,16403	0,98467	0,98467
14	4	1	3	40,65299	6,34050	1,04384	0,98890	0,98890
15	5	1	4	58,20534	7,61605	1,50607	0,98555	0,98555
16	6	1	5	53,47025	7,18429	1,34015	0,96528	0,96528
17	7	1	6	45,09110	6,70541	1,16744	0,99715	0,99715
18	8	1	7	50,66833	7,11765	1,31540	0,99989	0,99989
19	9	1	8	38,24358	6,10453	0,96759	0,97442	0,97442
20	11	1	9	62,60347	7,89480	1,61833	0,99560	0,99560
21	12	1	10	45,10741	6,71478	1,17071	0,99958	0,99958

	1	2	3	4	5	6	7	8	9	10	11	12	13
1													
2	(C) UNIWIN version 9.7.0												
3													
4	DATE : 24/07/2023												
5	ORDINATEUR : LAPTOP-LEG8L077												
6	UTILISATEUR : echar												
7	FICHIER(S) DE DONNEES OUVERT(S) : IRIS3.SGD												
8													
9	RESULTATS DE L'ANALYSE FACTORIELLE DISCRIMINANTE												
10													
11	Sélection :												
12	Et type = A												
13													
14	Jeu d'apprentissage : 120 observations												
15	Jeu de validation : 24 observations												
16	Jeu de prévision : 6 observations												
17													
18	Nombre de variables explicatives : 4												
19													
20	Facteur de classement : codesp2												
21													

(C) UNIWIN version 9.7.0

DATE : 24/07/2023
 ORDINATEUR : LAPTOP-LEG8L077
 UTILISATEUR : echar
 FICHIER(S) DE DONNEES OUVERT(S) : IRIS3.SGD

RESULTATS DE L'ANALYSE FACTORIELLE DISCRIMINANTE

Sélection :
 Et type = A

Jeu d'apprentissage : 120 observations
 Jeu de validation : 24 observations
 Jeu de prévision : 6 observations

Nombre de variables explicatives : 4

Facteur de classement : codesp2

Classes :
 Setosa
 Versicolor
 Virginica

CENTROIDES DES GROUPES ET GLOBAL

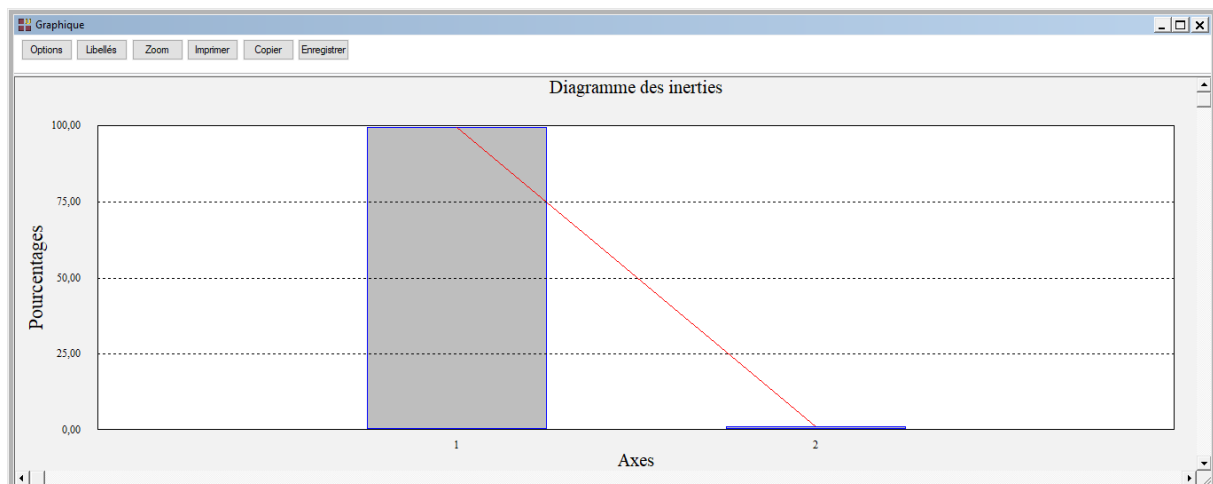
Ces rapports nous fournissent les renseignements suivants :

- Centroïdes des groupes et global
- Matrices des covariances et corrélations globales et des groupes
- Matrice des covariances et des corrélations intra-groupes
- Distances de Mahalanobis entre les groupes, Fishers, niveaux de signification
- Tableau des inerties (avec corrélation canonique, lambda de Wilks, Khi-2, degrés de liberté et niveau de signification)
- Test de Pillai
- Test de Box
- Fonctions discriminantes standardisées et non standardisées
- Résultats pour les variables
- Résultats pour les individus
- Coordonnées des centres des groupes
- Coefficients des fonctions de classement
- Résumé et détails du classement de la population d'apprentissage
- Statistiques pour les groupes observés et prévus (apprentissage)
- Résumé et détails du classement de la population de validation
- Classement de la population de prévision

L'option Graphiques

- Diagramme des inerties

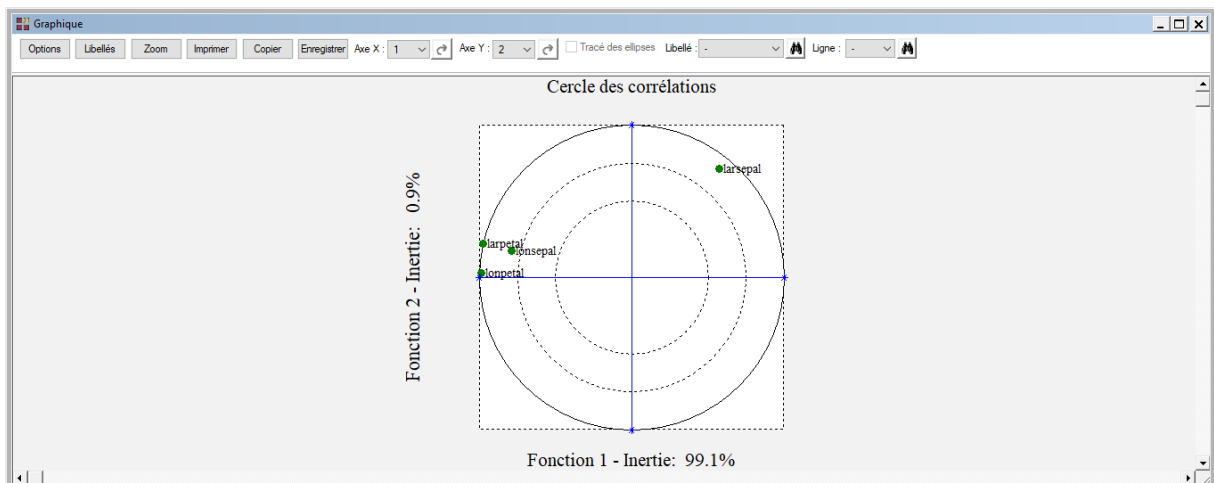
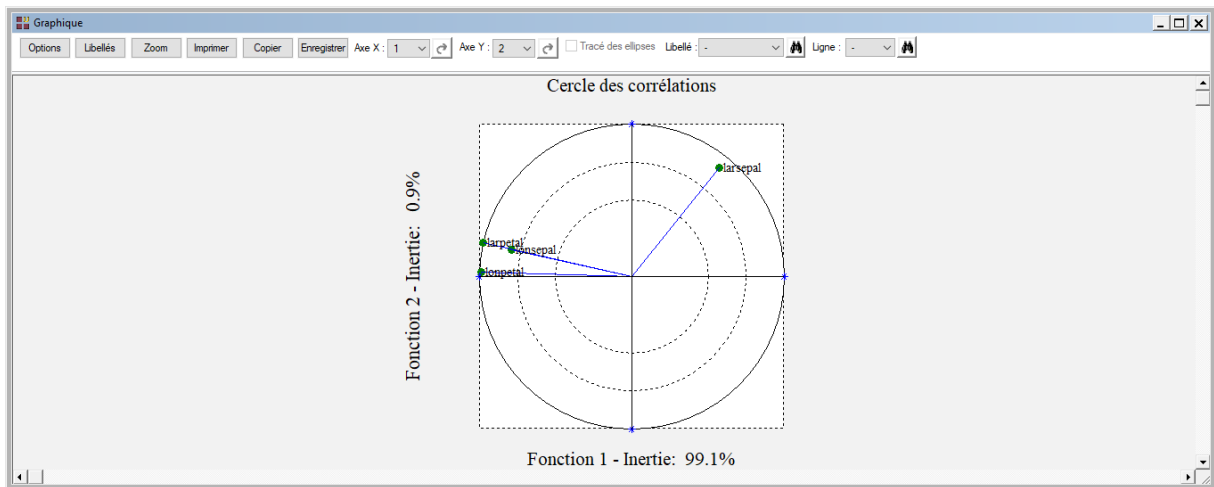
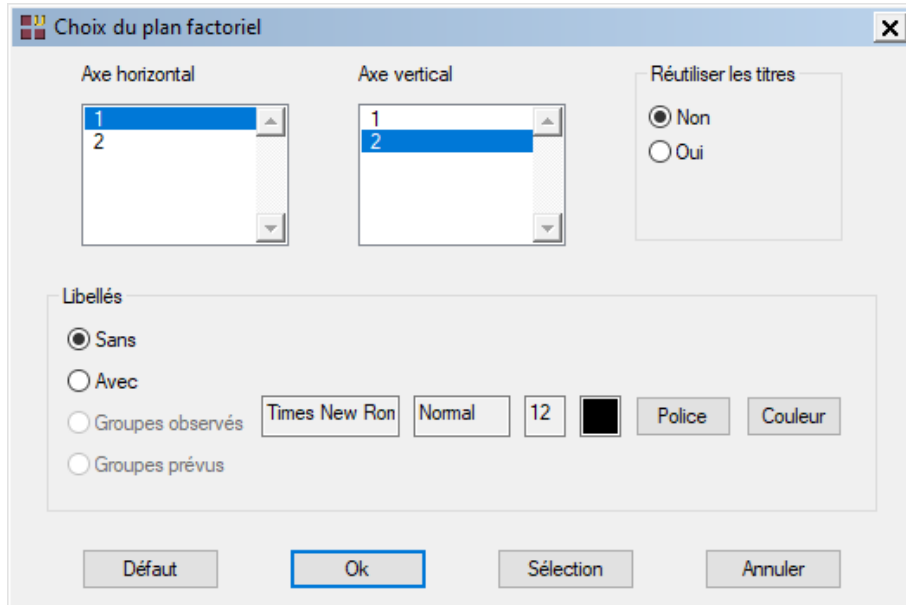
Ce graphique affiche les pourcentages d'inertie pour chacun des axes factoriels.



- Les options Cercle factoriel des corrélations des variables

Ces options permettent d'afficher le cercle de corrélations des variables et de choisir si on désire tracer les lignes reliant les points à l'origine du cercle. L'option sans ces lignes est utile lorsqu'il y a un grand nombre de variables représentées. Choisissons les variables avec lignes puis sans lignes.

Une boîte de dialogue permettant de choisir le plan factoriel s'affiche. Elle permet également de préciser si l'on désire afficher les libellés des variables, de choisir la couleur et la police et d'indiquer si les titres du graphique (titre 1, titre 2), doivent être conservés pour être réutilisés ultérieurement dans d'autres graphiques créés lors de cette même session de travail.

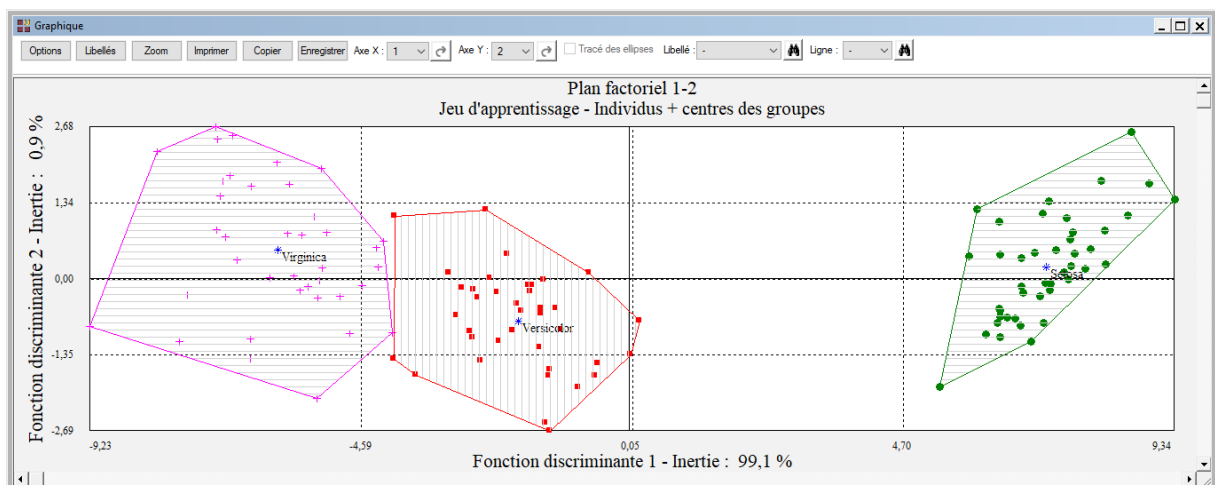
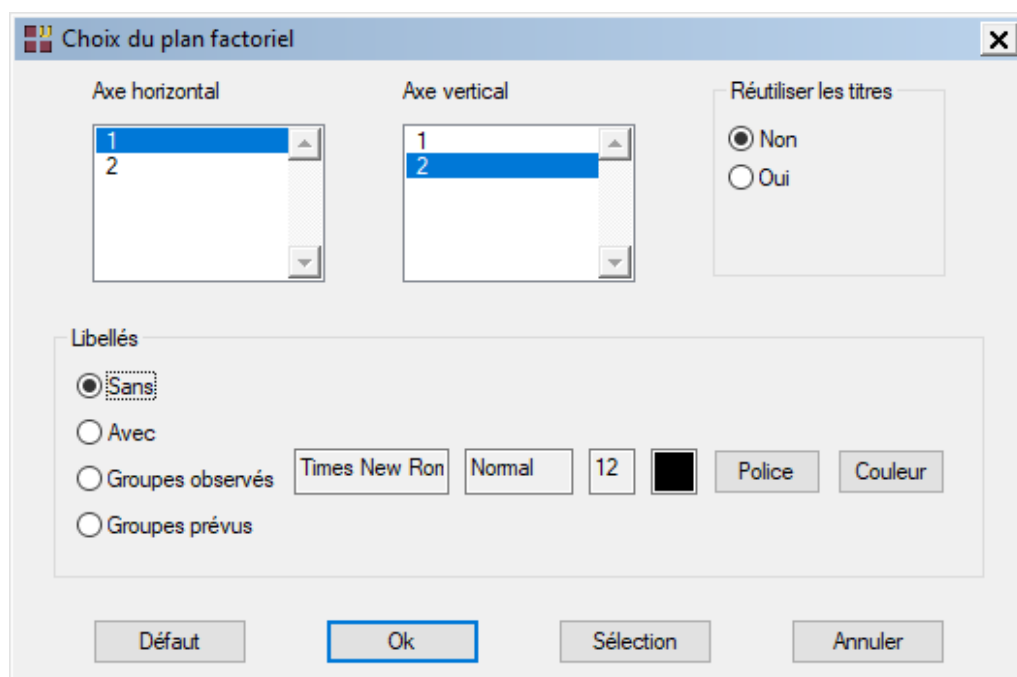


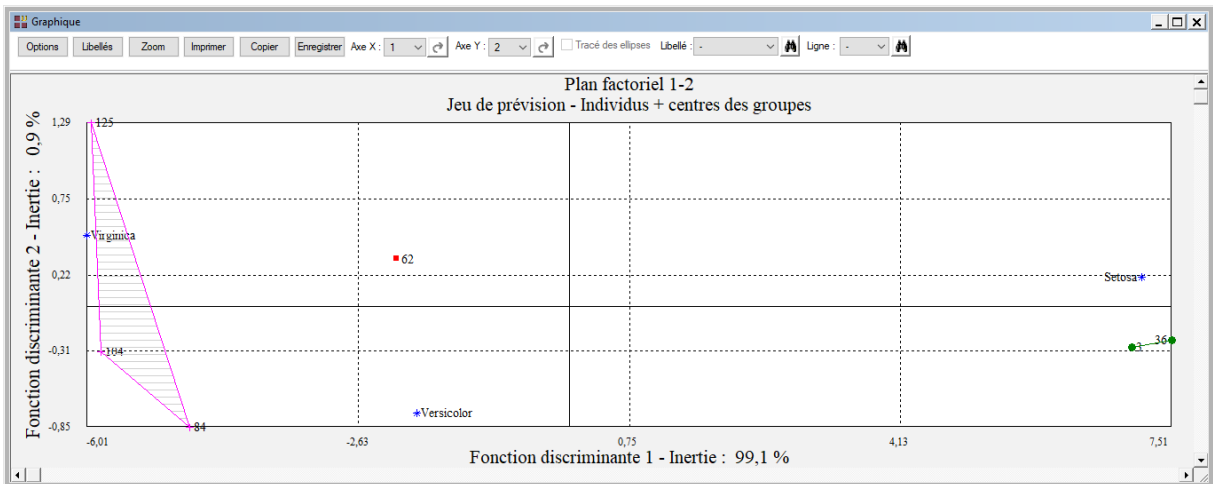
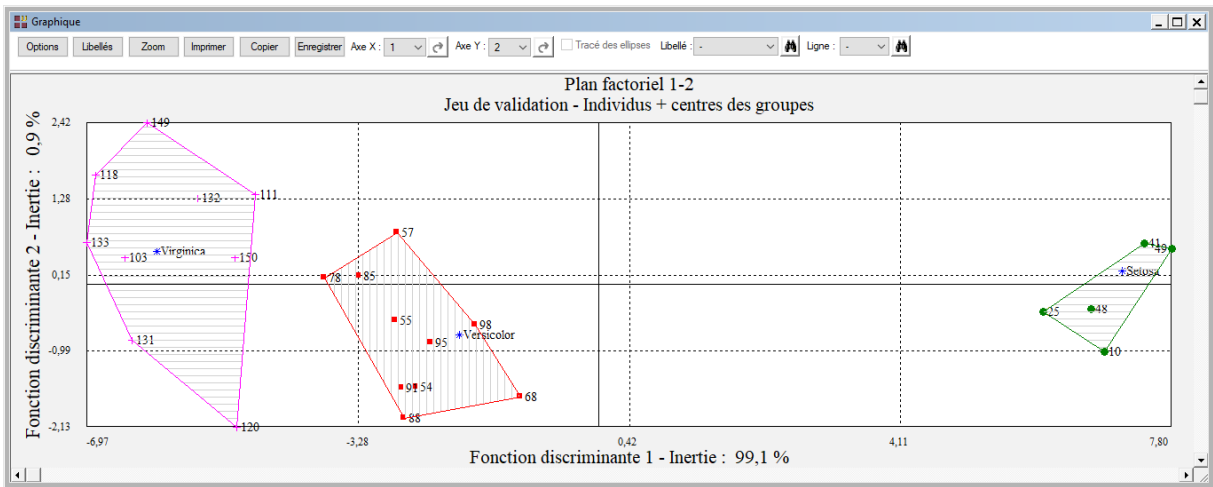
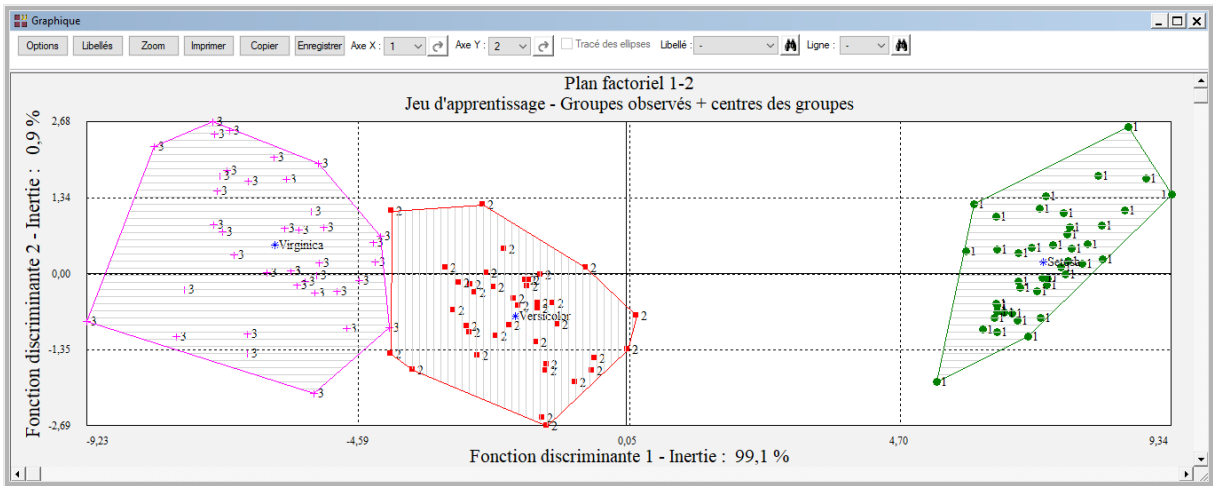
- Les options Plan factoriel des individus et centres des groupes

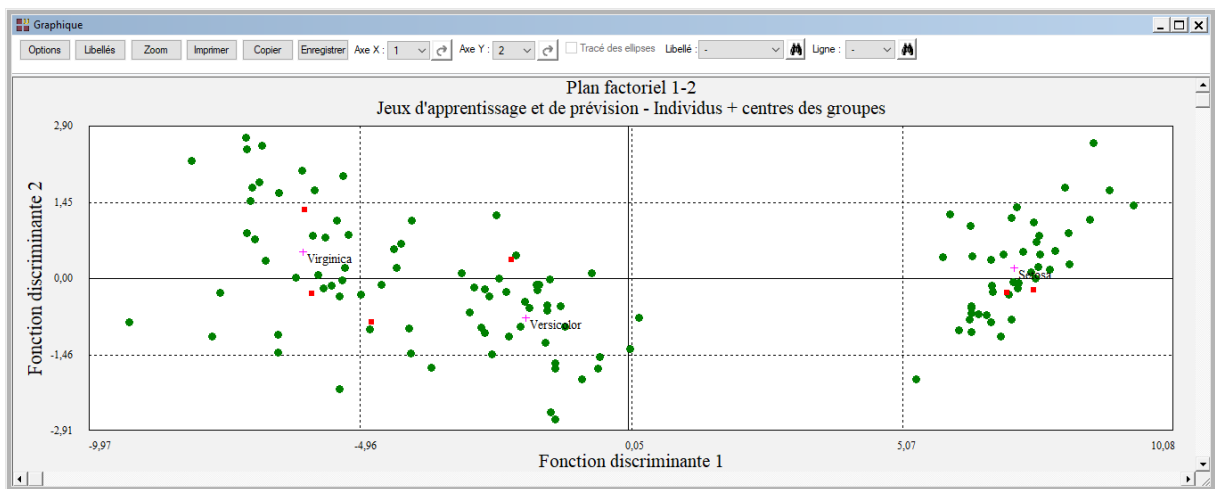
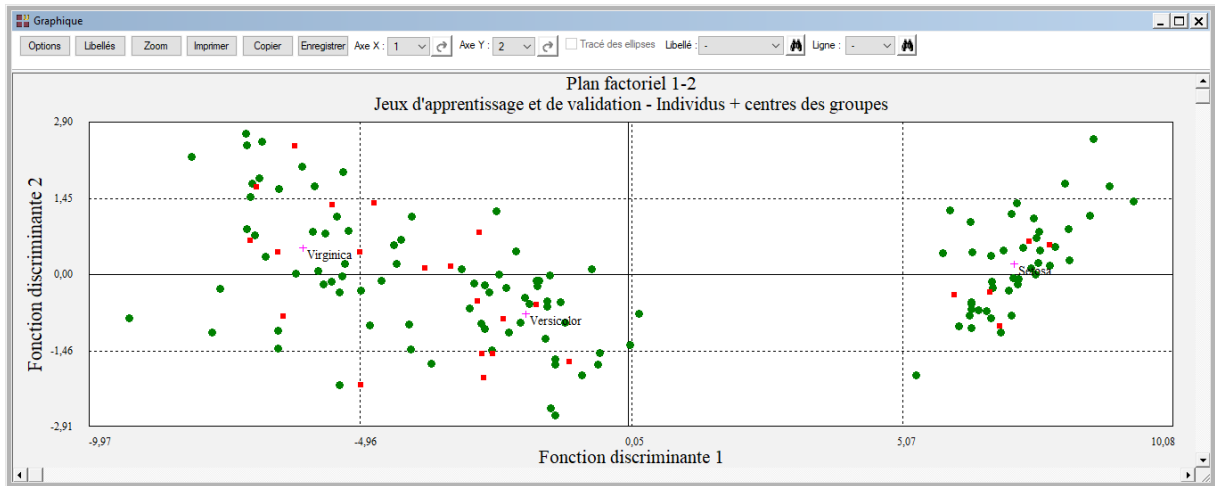
Ces options permettent d'afficher des plans factoriels des individus et des centres des groupes pour les populations d'apprentissage, de validation et de prévision. Une boîte de dialogue permettant de choisir le plan factoriel s'affiche.

Elle permet de préciser si l'on désire afficher ou non les libellés des individus, de préciser si ces libellés sont les codes des groupes observés ou les codes des groupes prévus, de choisir la couleur et la police pour ces libellés. Il est également possible d'indiquer si les titres du graphique (titre 1, titre 2), doivent être conservés pour être réutilisés ultérieurement dans d'autres graphiques créés lors de cette même session de travail.

Des exemples de plans factoriels sont montrés ci-après.







- Les options Courbe ROC

Ces deux options ne sont pas actives dans cet exemple car le facteur de classement possède plus de 2 classes.

Exemple 2 : Fichier INFARCT2

Pour ce deuxième exemple, nous utiliserons le fichier INFARCT2.

Ce fichier contient des informations concernant 101 victimes d'un infarctus du myocarde.

Cette population est constituée d'une population d'apprentissage de 81 individus et d'une population de validation de 20 individus.

La population d'apprentissage comprend les individus 11 à 51 (groupe Décès) et les individus 52 à 91 (groupe Survie).

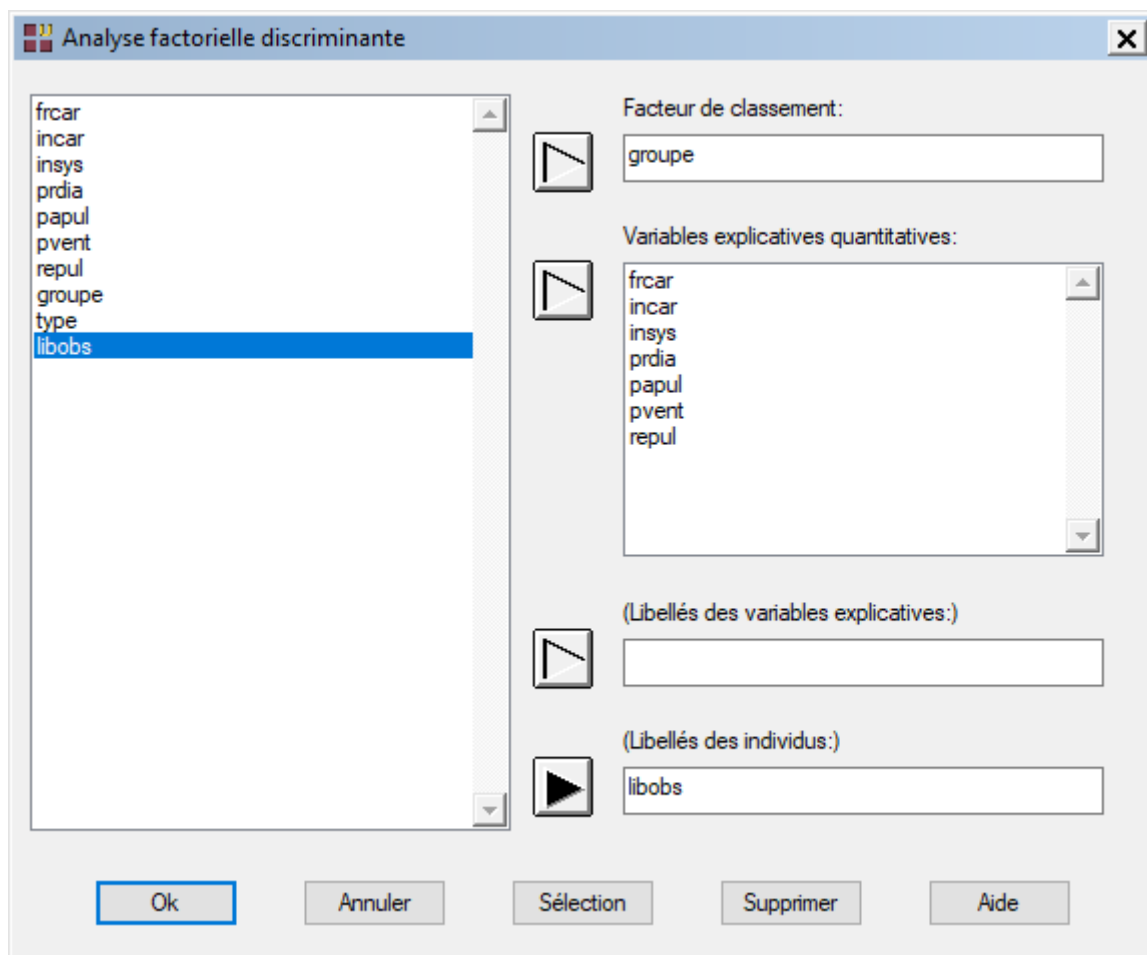
La population de validation comprend les individus 1 à 10 (groupe Décès) et les individus 92 à 101 (groupe Survie).

Les variables mesurées sont :

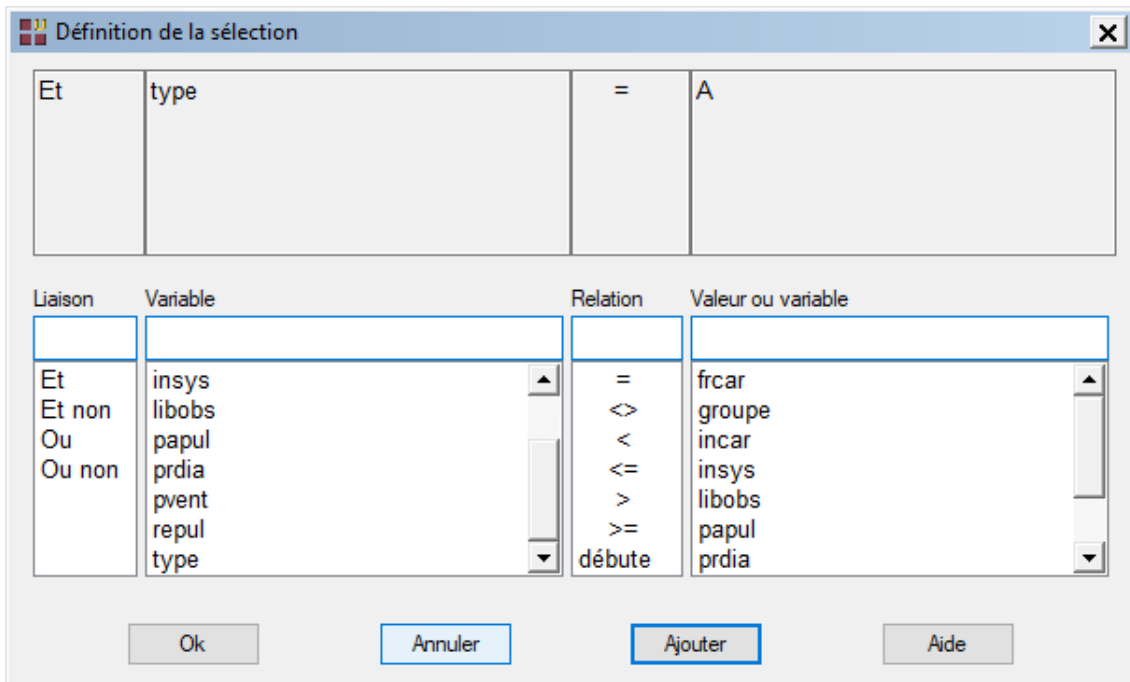
Variable	Mesure
<i>frcar</i>	fréquence cardiaque
<i>incar</i>	index cardiaque
<i>insys</i>	index systolique
<i>prdia</i>	pression diastolique
<i>papul</i>	pression artérielle pulmonaire
<i>pvent</i>	pression ventriculaire
<i>repul</i>	résistance pulmonaire

La variable *groupe* indique le groupe d'appartenance de chaque individu (Décès ou Survie). La variable *type* précise la population d'appartenance de chaque individu (A si apprentissage, V si validation). La variable *libobs* contient les libellés des individus des populations d'apprentissage et de validation.

Cliquons sur l'icône AFD dans le ruban Expliquer. La boîte de dialogue montrée ci-dessous s'affiche.

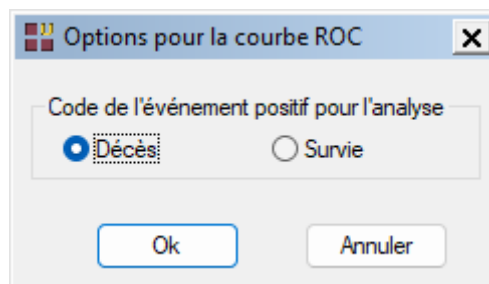


Cliquons sur le bouton Sélection pour définir la population d'apprentissage.



Un message nous indique que 81 individus sont sélectionnés.

Puisque la variable à expliquer possède deux classes, la procédure demande de préciser le code de l'événement positif qui sera utilisé pour le tracé de la courbe ROC.



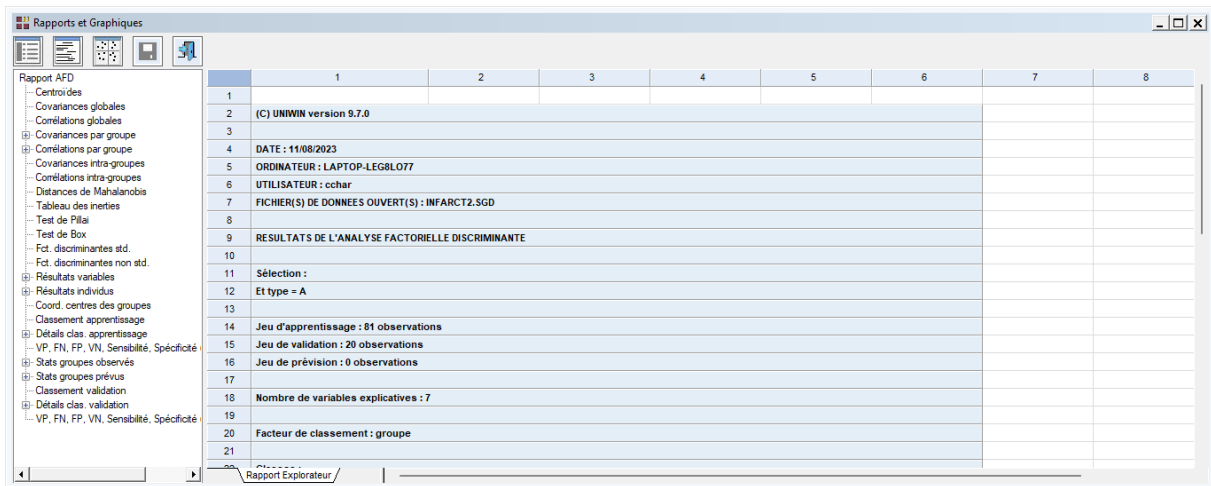
Cliques sur le bouton Ok pour exécuter le traitement de l'analyse.

Après quelques instants, un tableau précisant l'inertie expliquée par l'unique vecteur propre issu de l'analyse s'affiche.

Tableau des Inerties										
Nombre d'axes à retenir : <input type="text"/> Suite des calculs										
	1	2	3	4	5	6	7	8	9	10
1										
2	Tableau des Inerties									
3										
4	Axe	Valeur propre	Pct variance	Pct cumulé	Variation	Corr. canonique				
5	1	1,34960	100	100	0	0,75789				
6										
7										
8										
9										
10										
11										

Cliques sur 'Suite des calculs'.

L'écran 'Rapports et Graphiques' s'affiche.

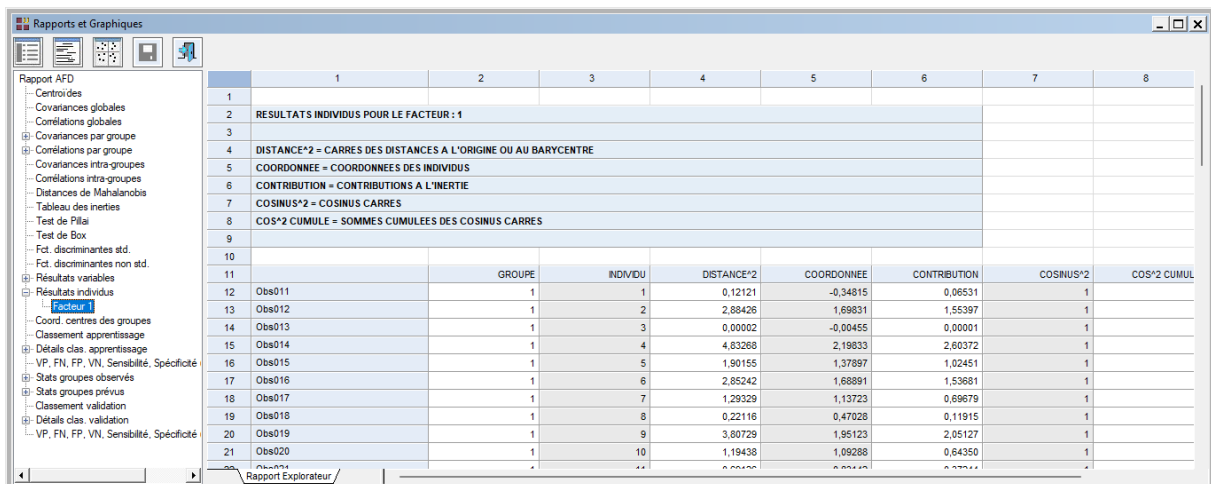


	1	2	3	4	5	6	7	8
1								
2	(C) UNWIN version 9.7.0							
3								
4	DATE : 11/08/2023							
5	ORDINATEUR : LAPTOP-LEGL077							
6	UTILISATEUR : cchar							
7	FICHER(S) DE DONNEES OUVERT(S) : INFARCT2.SGD							
8								
9	RESULTATS DE L'ANALYSE FACTORIELLE DISCRIMINANTE							
10								
11	Sélection :							
12	Et type = A							
13								
14	Jeu d'apprentissage : 81 observations							
15	Jeu de validation : 20 observations							
16	Jeu de prévision : 0 observations							
17								
18	Nombre de variables explicatives : 7							
19								
20	Facteur de classement : groupe							
21								

L'option Rapports

Cette option permet d'obtenir le rapport à l'écran sous la forme d'un explorateur, d'un tableur ou au format HTML.

Voici un exemple du rapport pour notre analyse (format Explorateur).



	1	2	3	4	5	6	7	8
1								
2	RESULTATS INDIVIDUS POUR LE FACTEUR : 1							
3								
4	DISTANCE^2 = CARRÉS DES DISTANCES A L'ORIGINE OU AU BARYCENTRE							
5	COORDONNEE = COORDONNÉES DES INDIVIDUS							
6	CONTRIBUTION = CONTRIBUTIONS A L'INERTIE							
7	COSINUS^2 = COSINUS CARRÉS							
8	COS^2 CUMULE = SOMMES CUMULÉES DES COSINUS CARRÉS							
9								
10								
11		GRUPE	INDIVIDU	DISTANCE^2	COORDONNEE	CONTRIBUTION	COSINUS^2	COS^2 CUMUL
12	Obs011	1	1	0,12121	-0,34815	0,06531	1	
13	Obs012	1	2	2,88426	1,69831	1,55397	1	
14	Obs013	1	3	0,00002	-0,00455	0,00001	1	
15	Obs014	1	4	4,83268	2,19833	2,60372	1	
16	Obs015	1	5	1,90155	1,37897	1,02451	1	
17	Obs016	1	6	2,85242	1,68891	1,53681	1	
18	Obs017	1	7	1,29329	1,13723	0,69679	1	
19	Obs018	1	8	0,22116	0,47028	0,11915	1	
20	Obs019	1	9	3,80729	1,95123	2,05127	1	
21	Obs020	1	10	1,19438	1,09288	0,64350	1	

Ce rapport fournit les renseignements suivants :

- Centroides des groupes et global
- Matrices des covariances et corrélations globales et des groupes
- Matrice des covariances et des corrélations intra-groupes
- Distances de Mahalanobis entre les groupes, Fishers, niveaux de signification
- Tableau des inerties (avec corrélation canonique, lambda de Wilks, Khi-2, degrés de liberté et niveau de signification)
- Test de Pillai
- Test de Box
- Fonctions discriminantes standardisées et non standardisées
- Résultats pour les variables

- Résultats pour les individus
- Coordonnées des centres des groupes
- Coefficients des fonctions de classement
- Résumé et détails du classement de la population d'apprentissage
- VP, FN, FP, VN, sensibilité et spécificité pour la population d'apprentissage
- Statistiques pour les groupes observés et prévus
- Résumé et détails du classement de la population de validation
- VP, FN, FP, VN, sensibilité et spécificité pour la population de validation
- Classement de la population de prévision (si existe)

L'option Graphiques

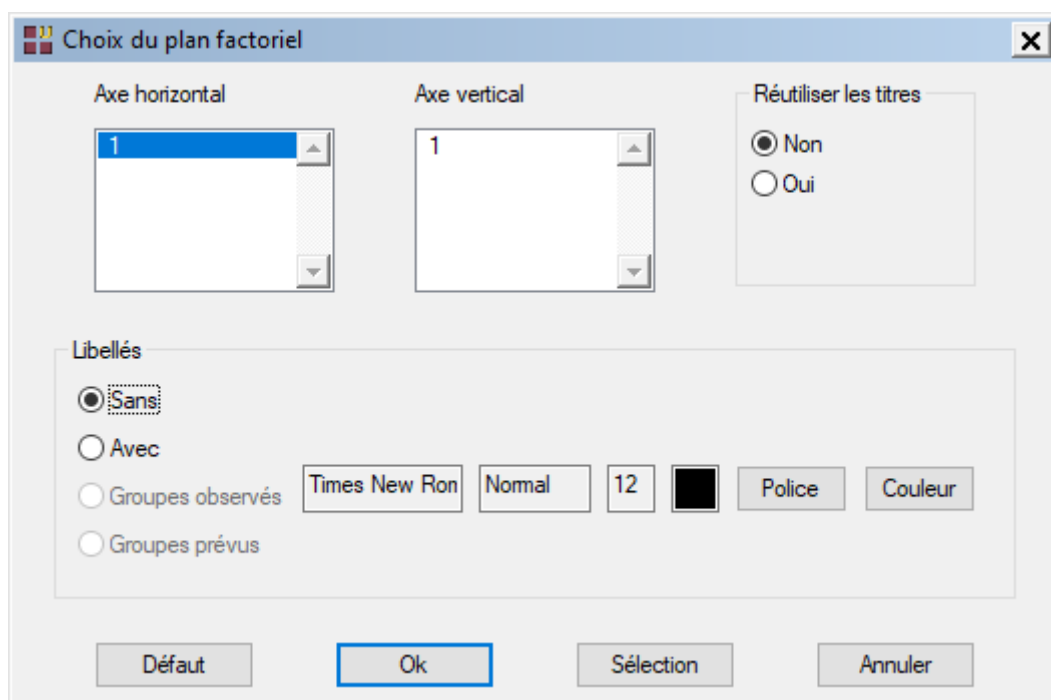
- Diagramme des inerties

Ce diagramme n'est pas disponible dans cet exemple car il y a une seule composante.

- Cercle factoriel des corrélations des variables

Ces options permettent d'afficher le cercle de corrélations des variables et de choisir si on désire tracer les lignes reliant les points à l'origine du cercle. L'option sans ces lignes est utile lorsqu'il y a un grand nombre de variables représentées.

Une boîte de dialogue permettant de choisir le plan factoriel s'affiche. Elle permet également de préciser si l'on désire afficher les libellés des variables, de choisir la couleur et la police et d'indiquer si les titres du graphique (titre 1, titre 2), doivent être conservés pour être réutilisés ultérieurement dans d'autres graphiques créés lors de cette même session de travail.

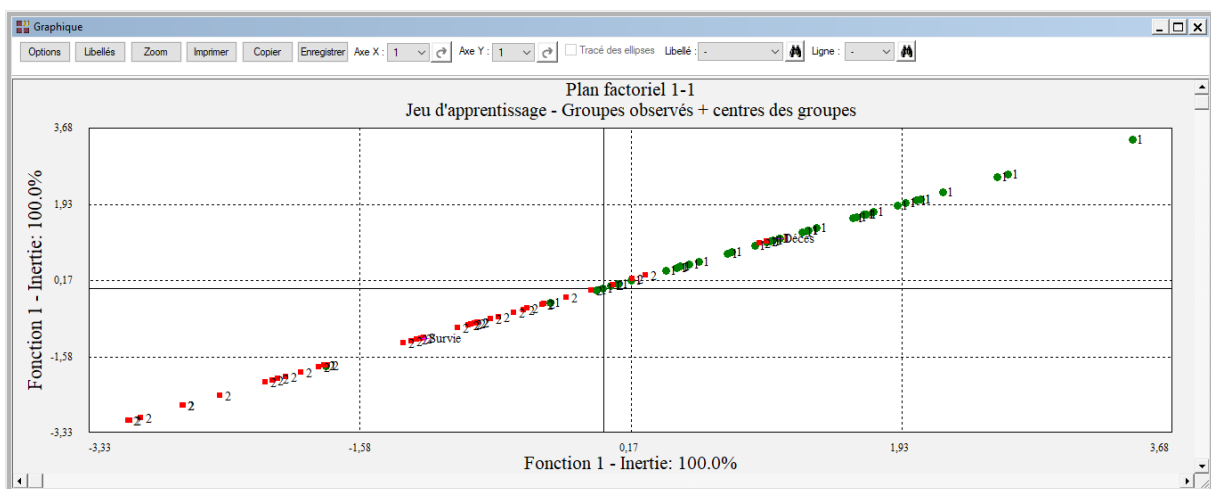


A noter que dans notre exemple, il n'y a qu'un seul axe factoriel et donc ce graphique n'est pas disponible.

- Plan factoriel des individus et centres des groupes

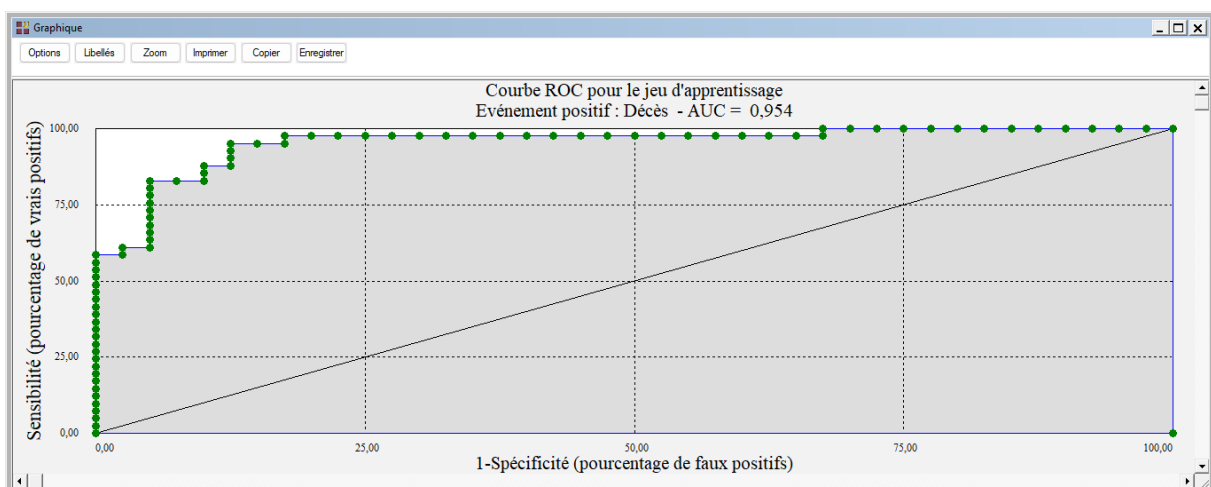
Ces options permettent d'afficher des plans factoriels des individus et des centres des groupes pour les populations d'apprentissage, de validation et de prévision.

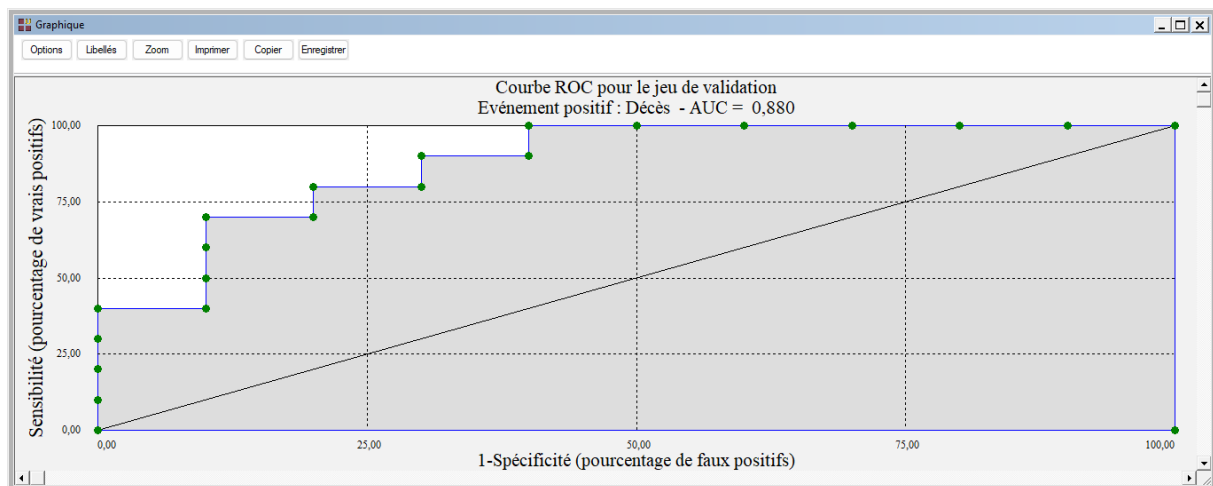
Une boîte de dialogue permettant de choisir le plan factoriel s'affiche. Elle permet de préciser si l'on désire afficher ou non les libellés des individus, de préciser si ces libellés sont les codes des groupes observés ou les codes des groupes prévus, de choisir la couleur et la police pour ces libellés. Il est également possible d'indiquer si les titres du graphique (titre 1, titre 2), doivent être conservés pour être réutilisés ultérieurement dans d'autres graphiques créés lors de cette même session de travail.



- Courbe ROC

Le tracé de la courbe ROC et le calcul de l'aire sous la courbe (AUC) sont possibles pour le jeu d'apprentissage et le jeu de validation car la variable à expliquer possède deux modalités.





Exemple 3 : Fichier BORDEAUX

Pour ce troisième exemple, nous utiliserons le fichier BORDEAUX.

Ce fichier contient des informations sur la qualité de vins de Bordeaux en relation avec les conditions météorologiques.

La variable qualitative *Qualité*, facteur de classement, prend trois modalités : Bon, Moyen et Médiocre.

Les variables explicatives sont :

- *Temp* somme des températures moyennes journalières (° C)
- *Insol* durée d'insolation (heures)
- *Chaleur* nombre de jours de grande chaleur
- *Pluie* hauteur des pluies (millimètres)

La variable *Année* contient les libellés des individus.

Cliquons sur l'icône AFD dans le ruban Expliquer.

Renseignons la boîte de dialogue comme montré ci-après.

Visualisons les résultats du classement.

Demandons également le cercle factoriel et le plan factoriel pour les axes 1 et 2.

Analyse factorielle discriminante

Année
Temp
Insol
Chaleur
Pluie
Qualité

Facteur de classement:
Qualité

Variables explicatives quantitatives:
Temp
Insol
Chaleur
Pluie

(Libellés des variables explicatives:)

(Libellés des individus:)
Année

Ok Annuler Sélection Supprimer Aide

Rapports et Graphiques

Rapport AFD

- Centroides
- Covariances globales
- Corrélations globales
- Covariances par groupe
- Corrélations par groupe
- Covariances intra-groupes
- Corrélations intra-groupes
- Distances de Mahalanobis
- Tableau des inerties
- Test de Pillai
- Test de Box
- Fct. discriminantes std.
- Fct. discriminantes non std.
- Résultats variables
- Résultats individus
- Coord. centres des groupes
- Classement apprentissage**
- Détails clas. apprentissage
- Stats groupes observés
- Stats groupes prévus

	1	2	3	4	5	6	7	8
1								
2	RESULTATS DU CLASSEMENT POUR LE JEU D'APPRENTISSAGE							
3								
4	EN LIGNES, LES GROUPES OBSERVES							
5	EN COLONNES, LES GROUPES PREVUS							
6								
7	POURCENTAGE DE MAL CLASSES : 20,588 %							
8	POURCENTAGE DE BIEN CLASSES : 79,412 %							
9								
10								
11		Bon	Moyen	Médiocre	Total			
12	Bon	9	2	0	11			
13	Moyen	2	8	1	11			
14	Médiocre	0	2	10	12			
15	Total	11	12	11	34			
16								
17								
18								
19								
20								
21								
22								
23								
24								
25								
26								

Rapport Explorateur /

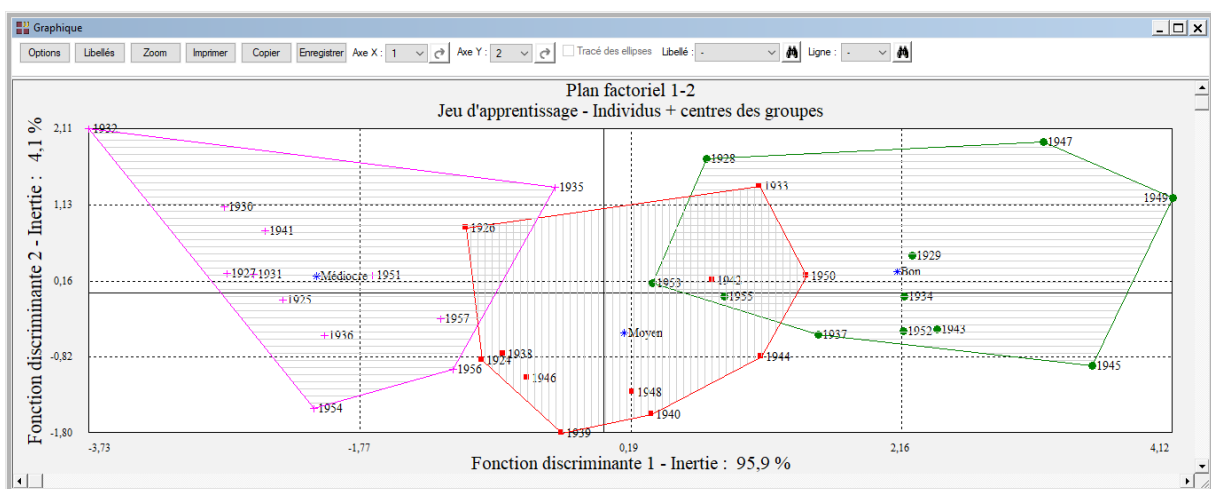
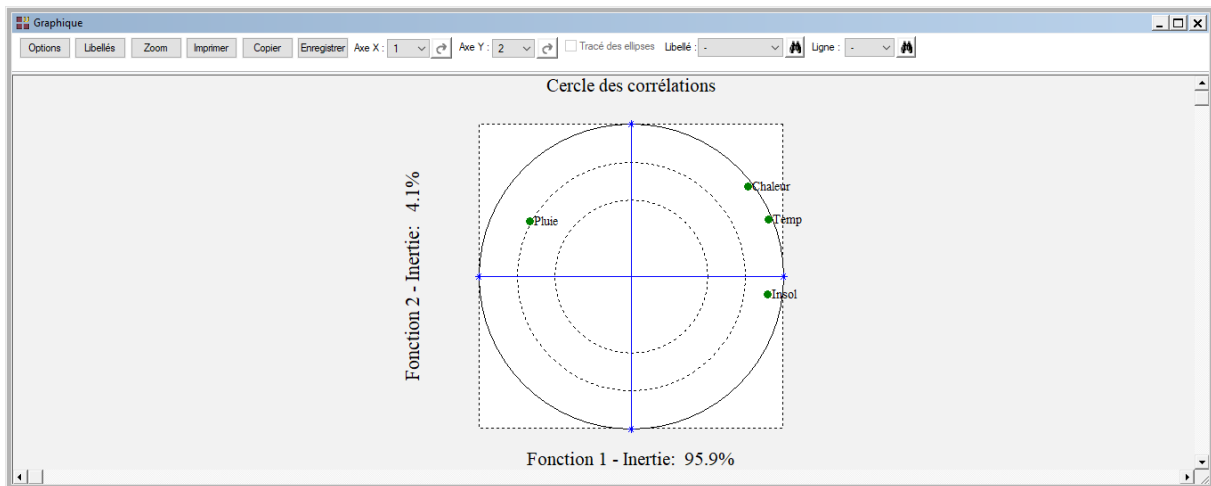
Rapports et Graphiques

Rapport AFD

- Centroides
- Covariances globales
- Corrélations globales
- Covariances par groupe
- Corrélations par groupe
- Covariances intra-groupes
- Corrélations intra-groupes
- Distances de Mahalanobis
- Tableau des inerties
- Test de Pillai
- Test de Box
- Fct. discriminantes std.
- Fct. discriminantes non std.
- Résultats variables
- Résultats individus
- Coord. centres des groupes
- Classement apprentissage
- Détails clas. apprentissage
- Stats groupes observés
- Stats groupes prévus

	1	2	3	4	5	6	7	8
1								
2	CLASSEMENTS ET DISTANCES AUX CENTROIDES DES GROUPES							
3								
4	GROUPE OBSERVE : Bon							
5	JEU D'APPRENTISSAGE							
6	L'INDIVIDU EST AFFECTE AU GROUPE DONT IL EST LE PLUS PROCHE DU CENTROIDE							
7								
8								
9								
10								
11	INDIVIDU - AFFECTATION	Bon	Moyen	Médiocre				
12	1928 - Bon	4,00053	5,34872	10,21839				
13	1929 - Bon	0,05702	5,34034	18,64652				
14	1934 - Bon	0,10178	4,33567	18,16603				
15	1937 - Bon	0,97337	1,97671	13,75642				
16	1943 - Bon	0,62290	5,12625	20,62039				
17	1945 - Bon	3,44810	11,66118	32,85423				
18	1947 - Bon	3,92572	15,26169	30,65577				
19	1949 - Bon	4,89371	18,79791	39,42416				
20	1952 - Bon	0,58081	4,08430	18,53601				
21	1953 - Moyen *	3,15229	0,45168	5,92207				
22	1955 - Moyen *	1,66029	0,74448	8,79654				
23								
24								
25								
26								

Rapport Explorateur /



Les vins mal classés sont les vins des années :

1953, 1955	bons classés moyens
1933, 1950	moyens classés bons
1926	moyen classé médiocre
1935, 1956	médiocre classé moyen

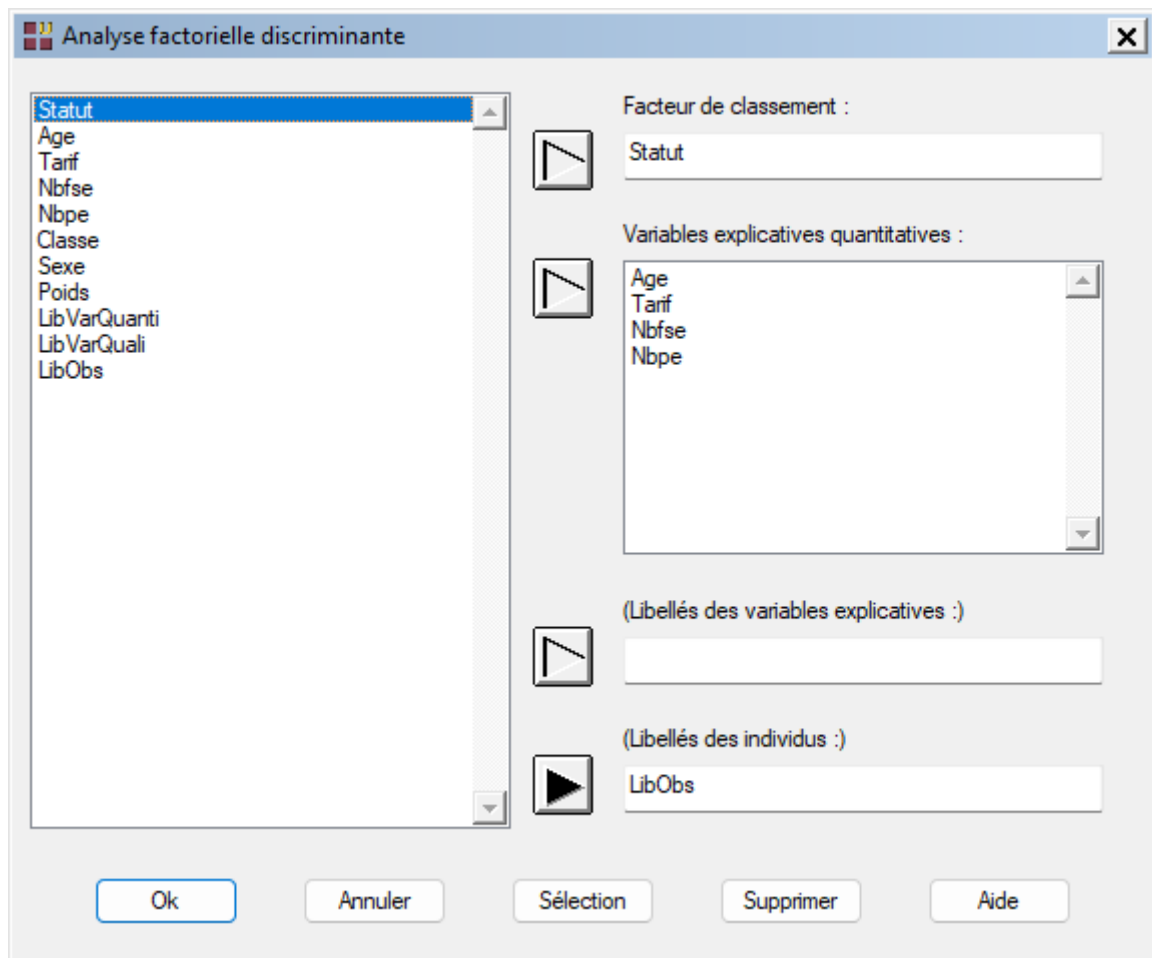
Exemple 4 : Fichier TITANIC

Pour ce quatrième exemple, nous utiliserons le fichier TITANIC.

Ce fichier contient des informations concernant 714 passagers :

Statut	Décès ou Survie
Classe	Classe du passager (1 ^{ère} , 2 ^{ème} ou 3 ^{ème})
Sexe	Homme ou Femme
Age	Age du passager
Nbfse	Nombre de frères, sœurs ou époux, épouses à bord
Nbpe	Nombre de parents ou enfants à bord
Tarif	Tarif passager (en £)

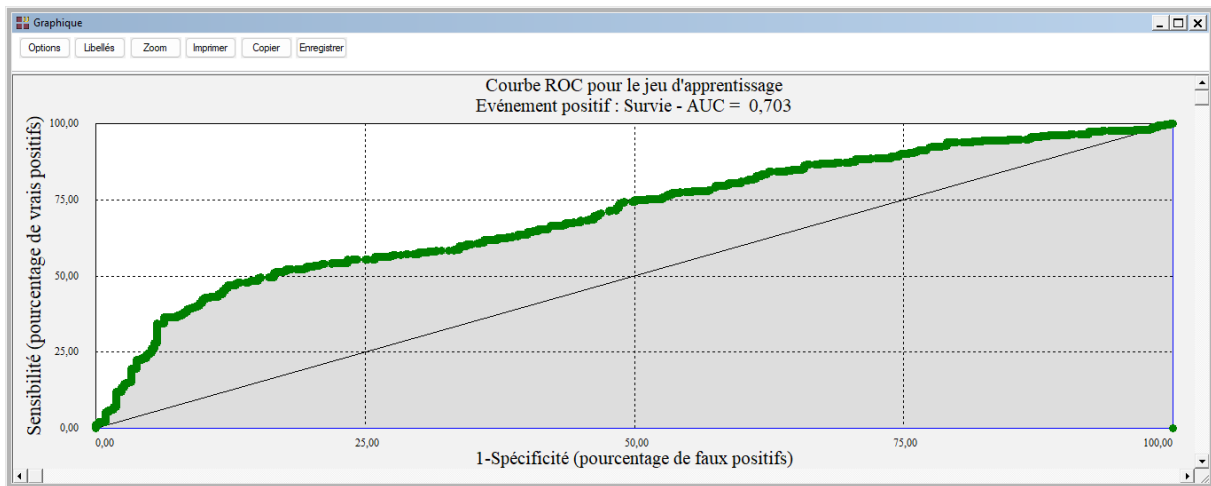
Cliquons sur l'icône AFD dans le ruban Expliquer et renseignons la boîte de dialogue comme montré ci-dessous.



Après exécution de la procédure, visualisons le tableau de classement des données d'apprentissage et la courbe ROC associée.

	1	2	3	4	5	6	7	8
1								
2	RESULTATS DU CLASSEMENT POUR LE JEU D'APPRENTISSAGE							
3								
4	EN LIGNES, LES GROUPES OBSERVES							
5	EN COLONNES, LES GROUPES PREVUS							
6								
7	POURCENTAGE DE MAL CLASSES : 29,132 %							
8	POURCENTAGE DE BIEN CLASSES : 70,868 %							
9								
10								
11		Décès	Survie	Total				
12	Décès	367	57	424				
13	Survie	151	139	290				
14	Total	518	196	714				
15								
16								
17								
18								
19								
20								
21								

Environ 71 % des passagers sont bien classés par cette analyse et l'aire sous la courbe ROC est proche de 0,7.



Note : Pour comparer les performances de plusieurs méthodes d'analyse, cet exemple est traité dans les six analyses AFD, ADB, KNN, BAYES, ANN et ARBRE.

Ellipses de tolérance

Reprenons notre exemple IRIS3.

Il est possible d'afficher les ellipses de tolérance autour des centres de gravité des classes en enregistrant les variables suivantes dans la procédure d'analyse factorielle discriminante :

Enregistrement des résultats (1/2)

Enregistrer

- Fct. discri. std. (apprentissage)
- Fct. discri. non std. (apprentissage)
- Coordonnées individus (apprentissage)
- Libellés individus (apprentissage)
- Classes individus (apprentissage)
- Coordonnées individus (validation)
- Libellés individus (validation)
- Classes individus (validation)
- Distances carrées individus (apprentissage)
- Cosinus carrés individus (apprentissage)

Noms attribués aux variables cibles

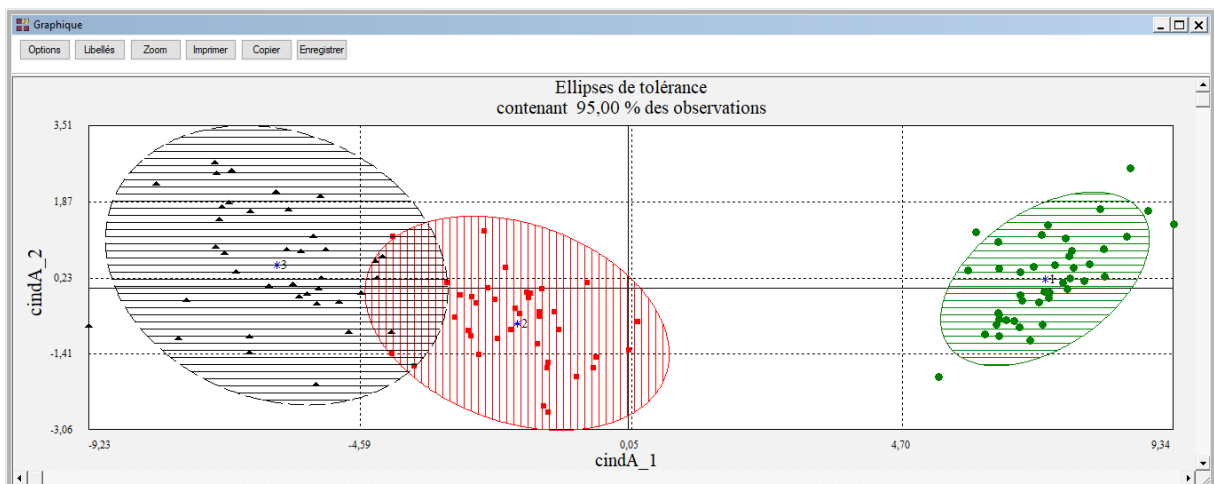
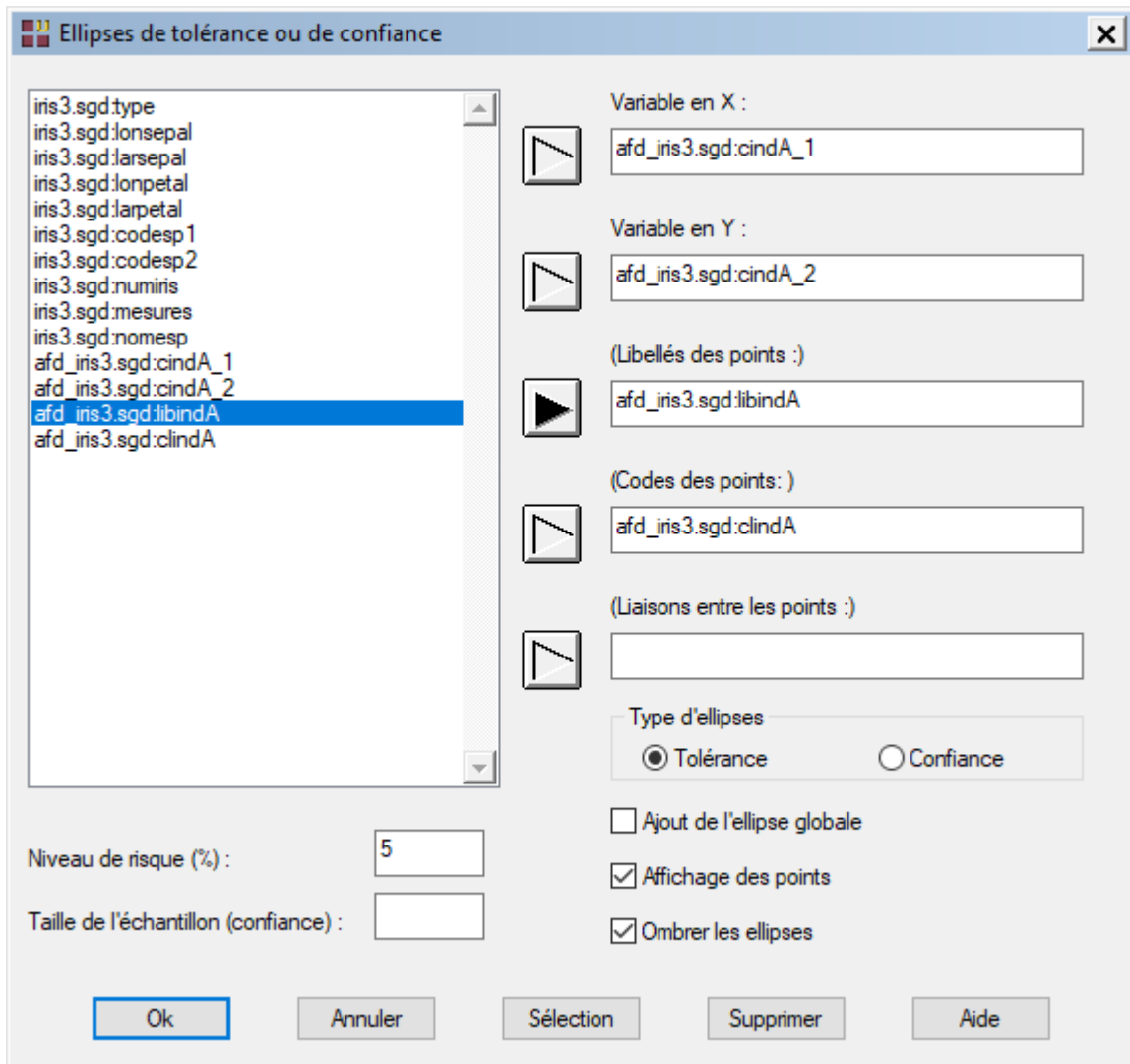
- fdstdA_1
- fdnstdA_1
- cindA_1
- libindA
- clindA
- cindV_1
- libindV
- clindV
- distindA
- cosindA_1

Ok
Plus
Tout
Annuler

Appelons AFD_Iris3.sgd le fichier dans lequel ces trois variables sont enregistrées.

Par le menu Décrire, exécutons la procédure Ellipses et renseignons la boîte de dialogue comme montré ci-après.

Demandons l'affichage du graphique des ellipses avec ombrage.



Les variables internes créées par la procédure

Voici la liste des variables internes créées par la procédure. Ces variables peuvent notamment être utilisées avec l'option 'Sélection'. A noter que certaines des variables mentionnées ci-dessous peuvent ne pas apparaître, en fonction des options choisies.

<i>Variable</i>	<i>Contenu</i>
fdstdA	Fonctions discriminantes standardisées (apprentissage)
fdnstdA	Fonctions discriminantes non standardisées (apprentissage)
cindA	Coordonnées des individus (apprentissage)
libindA	Libellés des individus (apprentissage)
clindA	Classes des individus (apprentissage)
distindA	Distances carrées à l'origine pour les individus (apprentissage)
cosindA	Cosinus carrés pour les individus (apprentissage)
conindA	Contributions pour les individus (apprentissage)
cvarA	Coordonnées des variables (apprentissage)
disvarA	Distances carrées à l'origine des variables (apprentissage)
cosvarA	Cosinus carrés des variables (apprentissage)
convarA	Contributions des variables (apprentissage)
ccentA	Coordonnées des centroïdes (apprentissage)
vpA	Vrais positifs (apprentissage)
fnA	Faux négatifs (apprentissage)
fpA	Faux positifs (apprentissage)
vnA	Vrais négatifs (apprentissage)
specificiteA	Spécificité (apprentissage)
sensibiliteA	Sensibilité (apprentissage)
classeA	Classement (apprentissage)
cindV	Coordonnées des individus (validation)
libindV	Libellés des individus (validation)
clindV	Classes des individus (validation)
vpV	Vrais positifs (validation)
fnV	Faux négatifs (validation)
fpV	Faux positifs (validation)
vnV	Vrais négatifs (validation)
specificiteV	Spécificité (validation)
sensibiliteV	Sensibilité (validation)
classeV	Classement (validation)
classeP	Classement (prévision)
libindP	Libellés des individus (prévision)