UNIWIN VERSION 10.7.0

ELLIPSES DE TOLERANCE OU DE CONFIANCE

Révision : 26/08/2026

Définition 1
Entrée des données 2
Données manquantes 3
Exemple 1 : Fichier IRIS 3
L'option Rapports 6
L'option Graphiques 7
Exemple 2 : Fichier INFARCT 8
Exemple 3 : Fichier MSPC 10
Les variables internes créées par la procédure 12

Définition

Les Ellipses de Tolérance ou de Confiance (ELL) permettent de visualiser des régions autour des centres de gravité de groupes définis par différentes catégories sous l'hypothèse d'une loi normale multivariée.

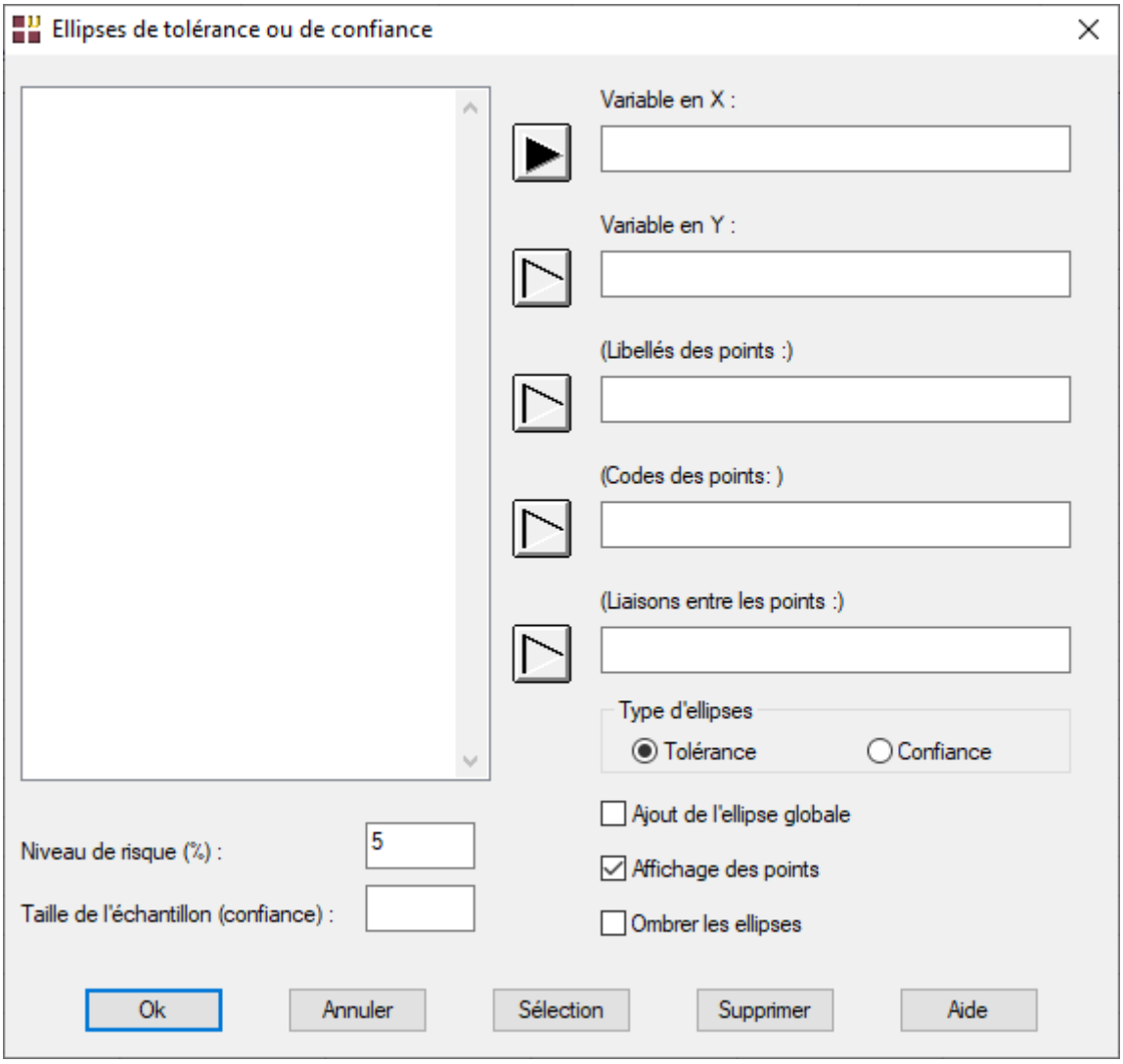
Considérons une variable qualitative réalisant une partition de la population en k groupes. Sous l'hypothèse de multi-normalité, les domaines de confiance représentés par des ellipses autour des k centres de gravité des groupes sont de trois types :

- ellipses de tolérance contenant en moyenne (*1-alpha*) % des observations de chaque groupe, où *alpha* désigne le niveau de risque
- ellipses de confiance pour les centres de gravité des groupes en considérant une population de taille infinie (tirage avec remise) au niveau de risque *alpha*
- ellipses de confiance avec coefficient d'exhaustivité pour les centres de gravité des groupes en considérant une population finie (tirage sans remise) au niveau de risque *alpha*. Pour l'étude de points supplémentaires en analyse factorielle, prendre pour taille de la population, le nombre d'individus du tableau analysé.

Des graphiques permettent de visualiser les différentes ellipses, les centres de gravité, les codes des points et les numéros des points.

Entrée des données

Cliquons sur l'icône ELL dans le ruban Décrire. La boîte de dialogue montrée ci-dessous s'affiche :



Cette boîte de dialogue permet de choisir les variables numériques en X et Y, la variable des libellés des points, la variable des codes définissant les différents groupes ainsi que la variable permettant de relier les points entre eux et d'indiquer si ces points sont affichés ou non.

Elle permet également de choisir le type d'ellipse (tolérance ou confiance), d'afficher ou non les symboles des points et d'ombrer les ellipses.

Le niveau de risque *alpha* (exprimé en %), indique que l'on désire calculer et tracer des ellipses de tolérance contenant (*1-alpha)* % des observations de chaque groupe ou des ellipses de confiance à (*1-alpha*) % pour les centres de gravité de chaque groupe, sous l'hypothèse de multi-normalité.

Dans le cas des ellipses de confiance, la taille de l'échantillon peut être précisée.

Si l'on considère un tirage aléatoire avec remise (ellipse de confiance sans coefficient d'exhaustivité), la taille de l'échantillon est infinie. Il ne faut alors rien préciser.

Si l'on considère un tirage aléatoire sans remise (ellipse de confiance avec coefficient d'exhaustivité), la taille de l'échantillon doit être précisée. Si l'étude porte sur des points supplémentaires d'une analyse factorielle, définir la taille de l'échantillon comme égale au nombre d'observations.

Note : Si la variable des codes n'est pas définie, il est supposé que toutes les observations appartiennent à un unique groupe.

Données manquantes

Les données manquantes sont autorisées pour les variables en X et en Y ainsi que pour la variable des codes.

Exemple 1 : Fichier IRIS

Pour illustrer cet exemple, nous utiliserons le fichier IRIS. Ce fichier contient les données relatives à 150 iris de 3 espèces : Iris Setosa, Iris Versicolor et Iris Virginica. Les mesures effectuées sont, en millimètres, la longueur du sépale (lonsepal), la longueur du pétale (lonpetal), la largeur du sépale (larsepal) et la largeur du pétale (larpetal).

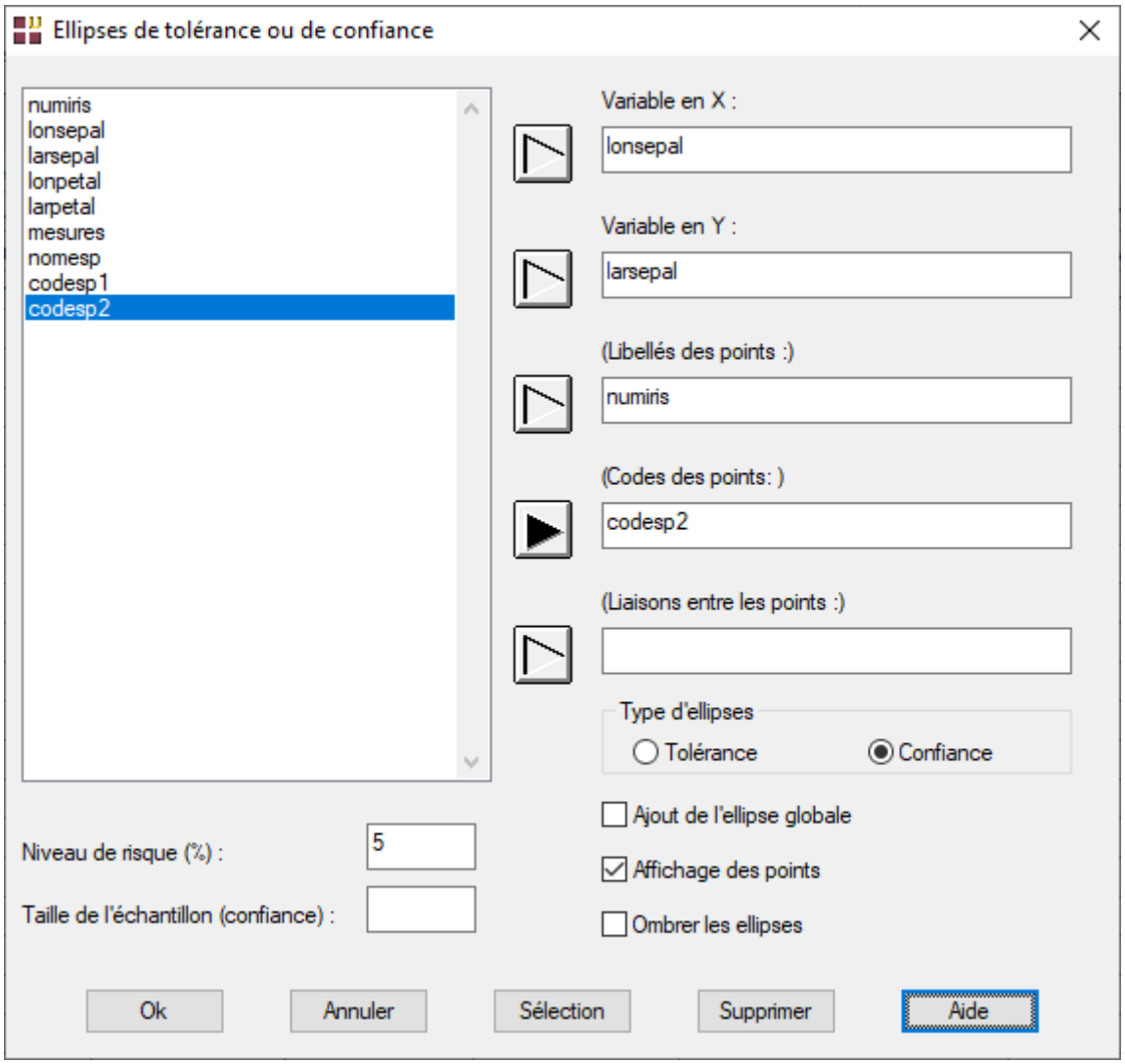
Cliquons sur l'icône ELL dans le ruban Décrire. La boîte de dialogue montrée ci-après s'affiche.

Choisissons la variable *lonsepal* pour l'axe des X, la variable *larsepal* pour l'axe des Y, la variable *numiris* comme variable de libellés des iris et la variable *codesp2* comme variable définissant les différents groupes.

Nous choisissons Ellipse de Confiance, un niveau de risque de 5% et une taille d'échantillon de 150, taille égale au nombre d'iris dans notre fichier.

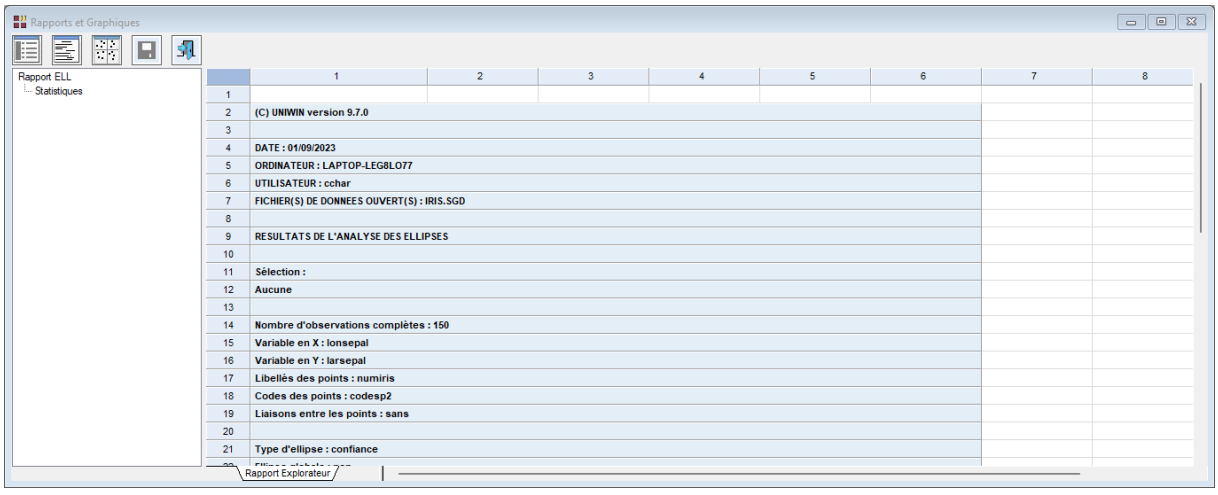
Si vous ne connaissez pas la taille de votre échantillon, ne rien mettre dans cette zone. Dans ce cas, après avoir cliqué sur Ok, UNIWIN vous demande de confirmer ou non si vous désirez travailler avec un échantillon de taille infinie.

Si vous répondez Non, alors UNIWIN affiche la taille de l'échantillon de votre fichier. Il vous suffit alors d'entrer cette taille dans la zone Taille de l'échantillon. Entrons 150 comme taille d'échantillon.



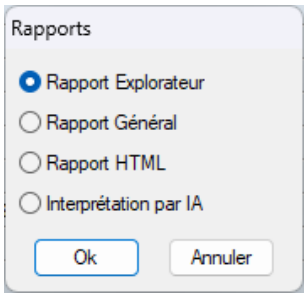
Après avoir renseigné cette boîte de dialogue, UNIWIN débute le calcul de l'analyse ELL.

Après quelques instants, l'écran 'Rapports et Graphiques' s'affiche.

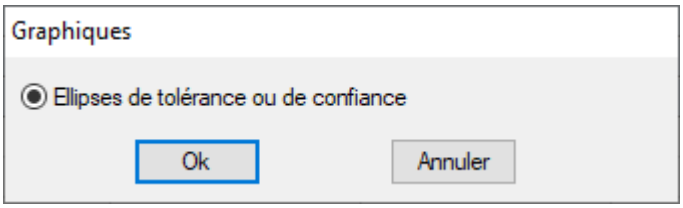


La barre d'outils 'Rapports et Graphiques' permet par l'icône 'Données' de rappeler la boîte de dialogue d'entrée des données.

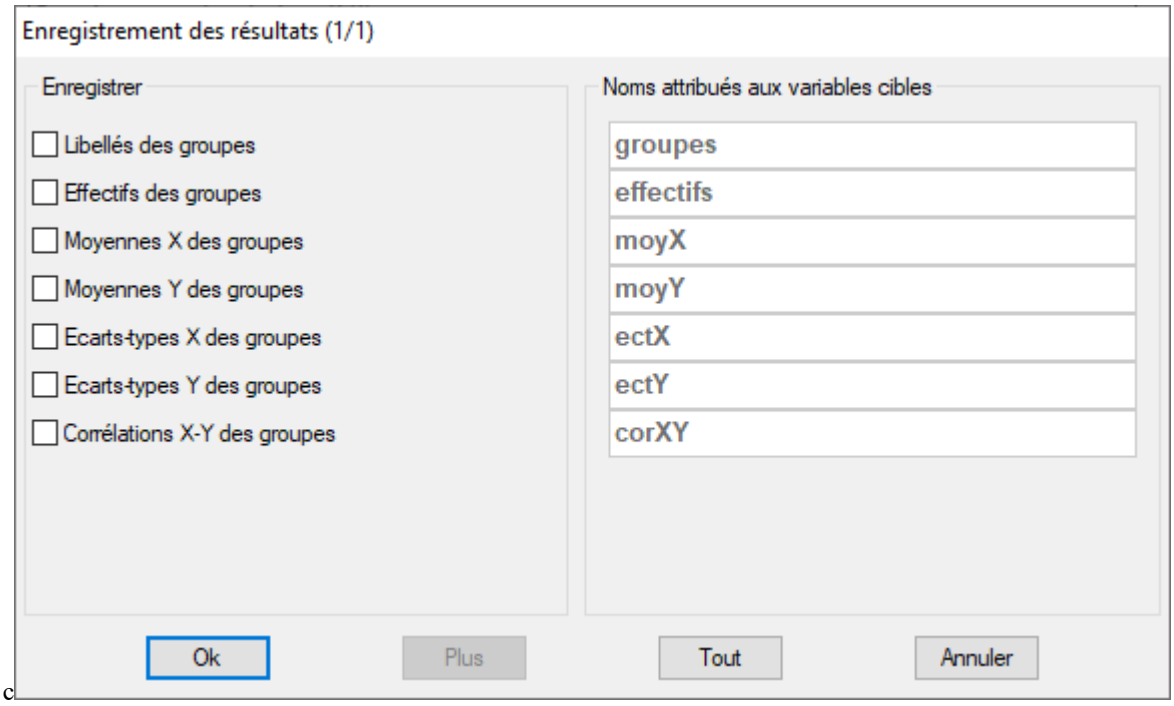
L'icône 'Rapports' affiche la boîte de dialogue des options pour les rapports :



et 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.



c

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

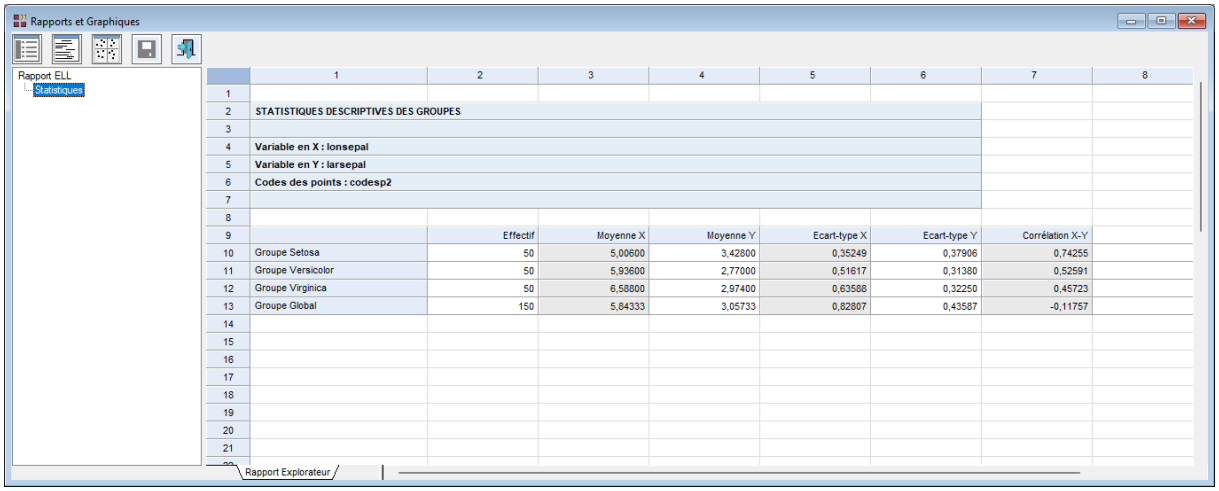
L'option Rapports

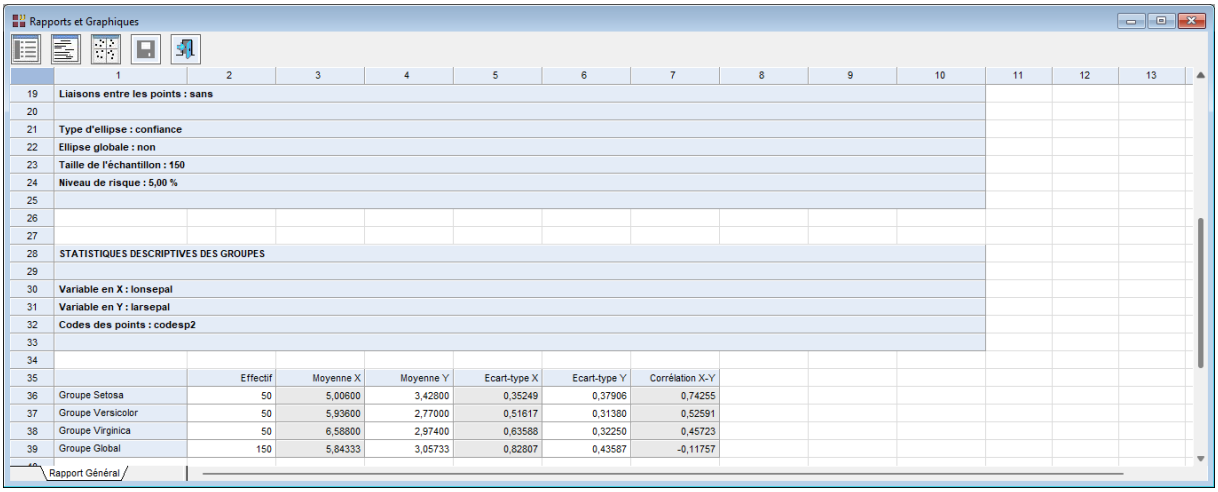
Cette option permet d'obtenir le rapport à l'écran sous la forme Explorateur, Général ou HTML et d'accéder à l'aide à l'interprétation par Intelligence Artificielle.

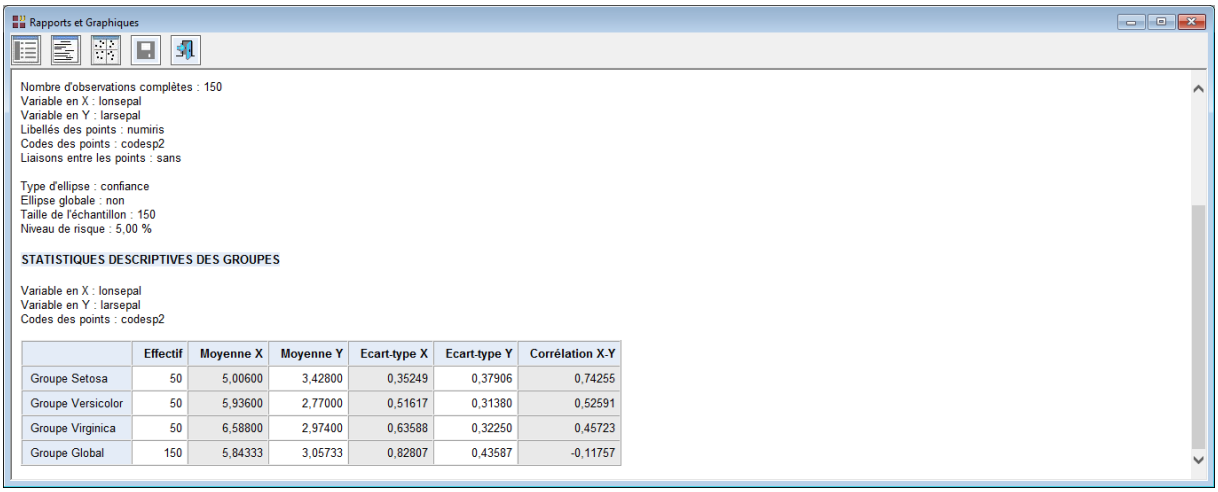
L'impression des rapports fait appel à la procédure 'Aperçu avant impression'.

Pour des informations sur cette procédure, voir le 'Manuel de l'Utilisateur'.

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







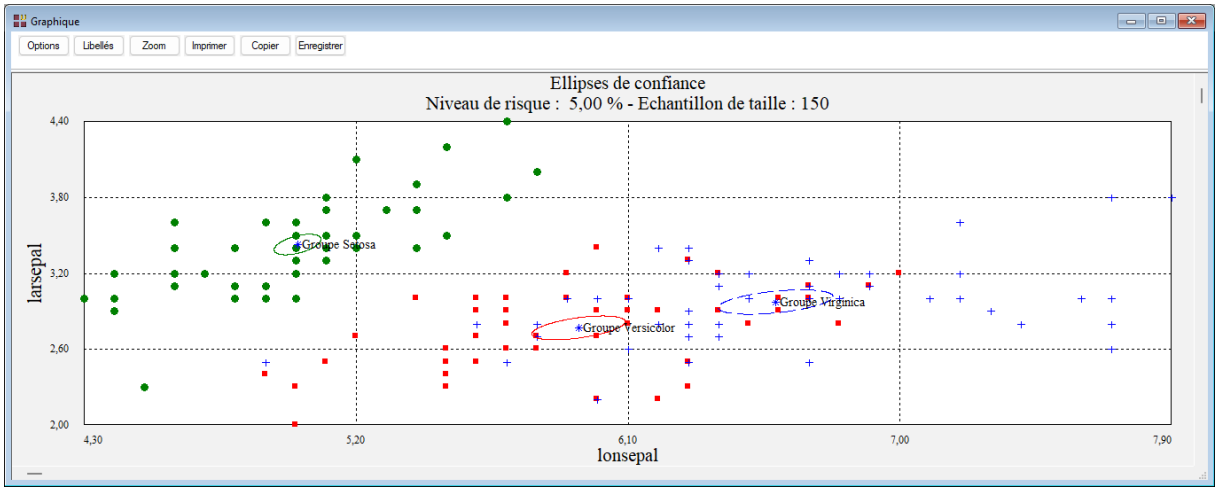
	Effectif	Moyenne X	Moyenne Y	Ecart-type X	Ecart-type Y	Corrélation X-Y
Groupe Setosa	50	5,00600	3,42800	0,35249	0,37906	0,74255
Groupe Versicolor	50	5,93600	2,77000	0,51617	0,31380	0,52591
Groupe Virginica	50	6,58800	2,97400	0,63588	0,32250	0,45723
Groupe Global	150	5,84333	3,05733	0,82807	0,43587	-0,11757

L'option Graphiques

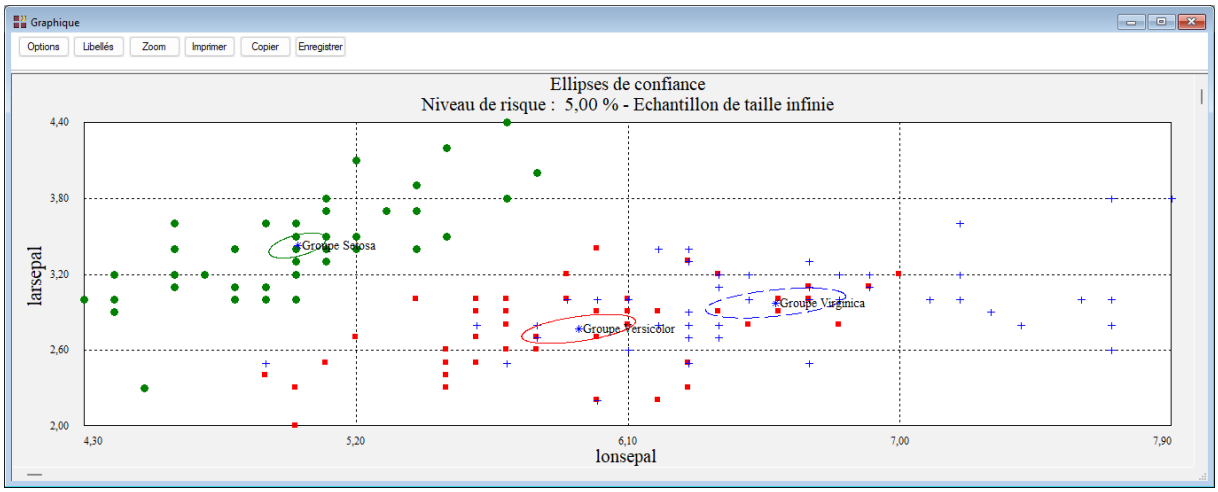
Cette option permet de tracer les ellipses.

L'affichage d'un libellé pour chaque point est proposé par le menu Libellés qui offre les choix suivants :

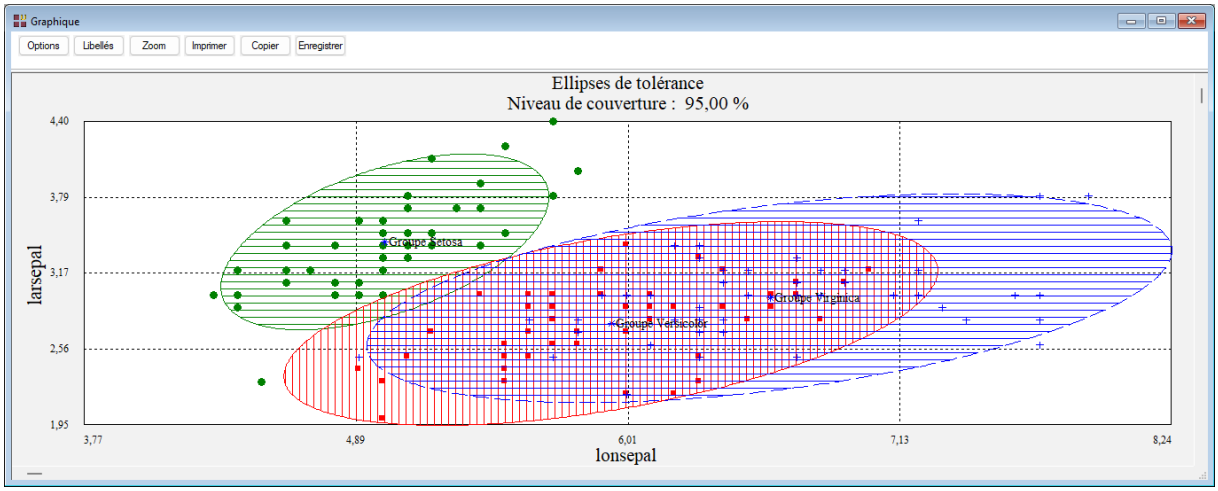
- Sans — pas de libellés
- Avec — utilisation de la variable Libellés définie dans la boîte de dialogue initiale
- Codes — les libellés sont les codes des groupes
- Numéros — les libellés sont les numéros d'ordre des points dans le fichier de données.



Revenons à notre boîte de dialogue et demandons maintenant le calcul et le tracé des ellipses de confiance en supposant une taille d'échantillon infinie.



Enfin, affichons les ellipses de tolérance et demandons des ellipses ombrées.



Exemple 2 : Fichier INFARCT

Comme deuxième exemple, nous utiliserons les résultats d'une Analyse en Composantes Principales faite sur les données du fichier INFARCT et une variable qualitative supplémentaire. Ce fichier contient 7 informations concernant 81 victimes d'un infarctus du myocarde.

Les 7 mesures sont stockées dans les variables :

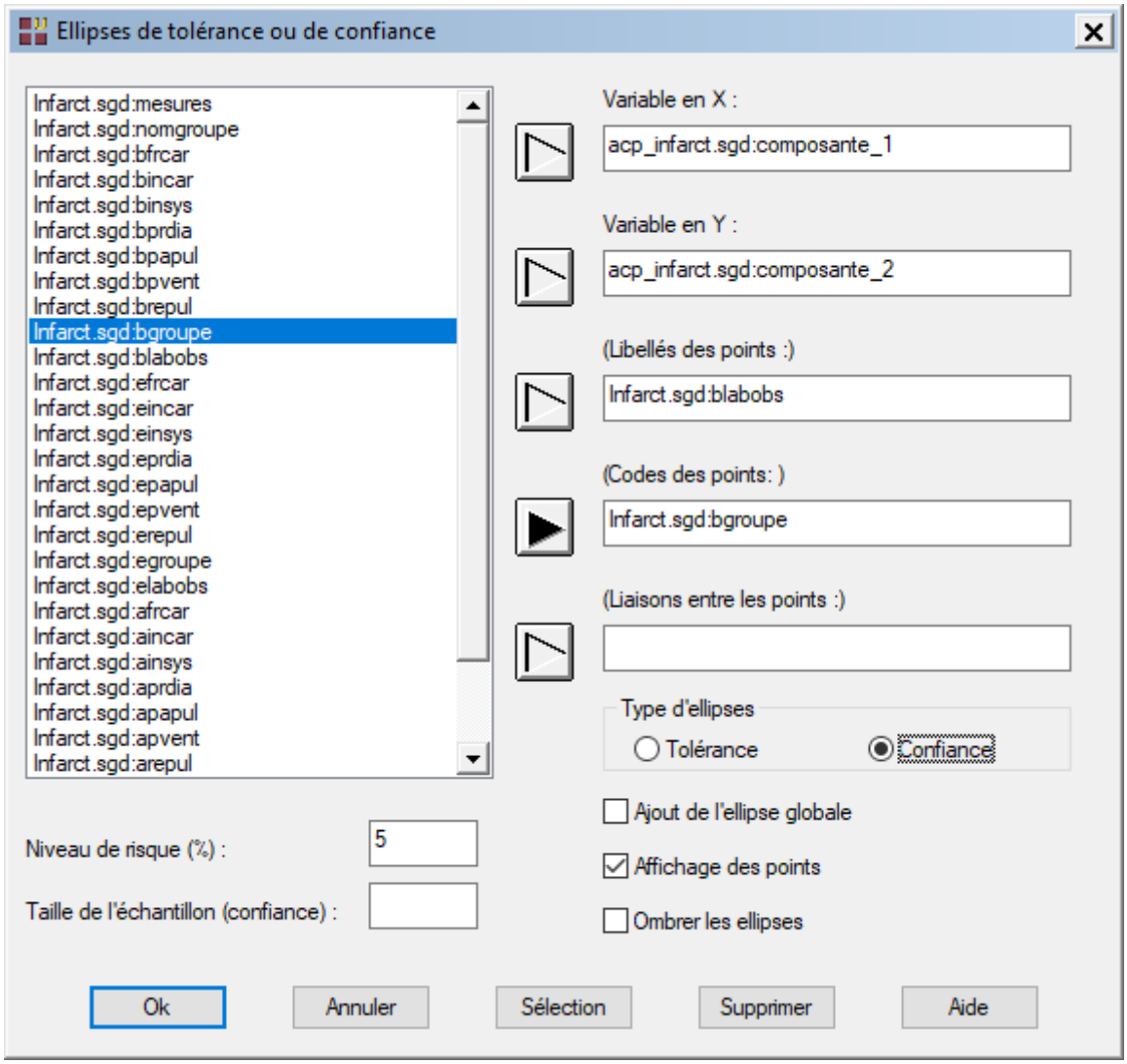
- *bfrcar* fréquence cardiaque
- *bincar* index cardiaque
- *binsys* index systolique
- *bprdia* pression diastolique
- *bpapul* pression artérielle pulmonaire
- *bpvent* pression ventriculaire
- *brepul* résistance pulmonaire

Dans ce fichier, la variable qualitative *bgroupe* indique par son code 1 les personnes décédées et par son code 2 les personnes vivantes.

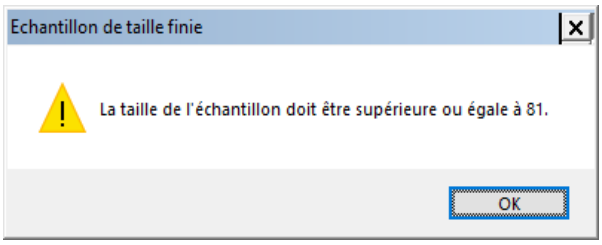
Après avoir mis en œuvre une Analyse en Composantes Principales standardisée (ACP) utilisant les 7 mesures sur notre population de 81 victimes, le fichier des résultats, appelé ACP_INFARCT, contient les variables *'composante_1'* et *'composante_2'* qui représentent les 2 premières composantes principales.

Ouvrons simultanément les deux fichiers INFARCT et ACP_INFARCT.

Renseignons la boîte de dialogue comme montré ci-dessous :

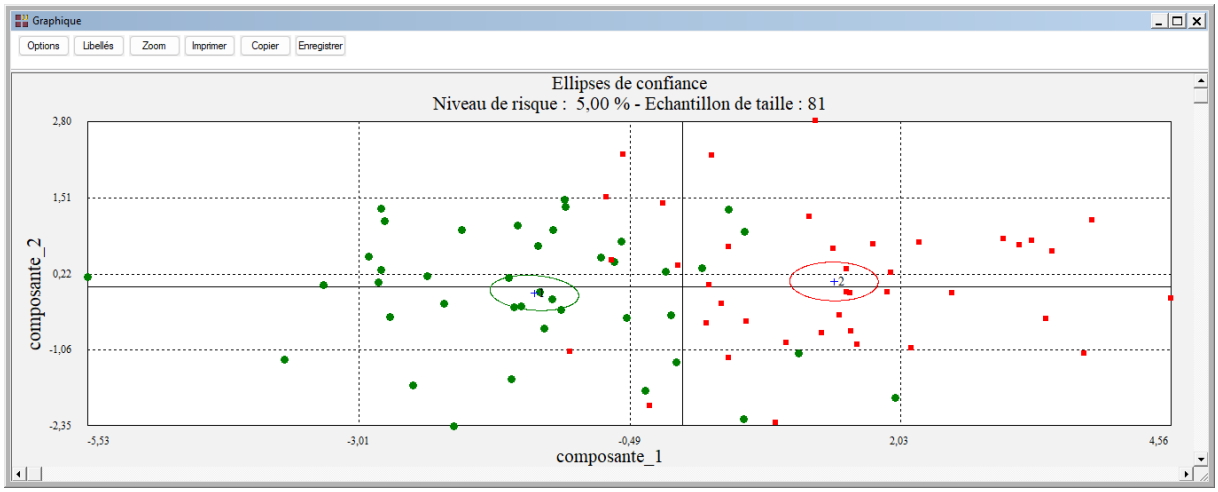


Le logiciel nous demande si nous désirons faire les calculs avec une taille infinie de l'échantillon.



Répondons Non. Un message s'affiche alors indiquant que la taille doit être au minimum de 81.

Renseignons la zone 'Taille de l'échantillon' en entrant la valeur 81 puis demandons le tracé du graphique :

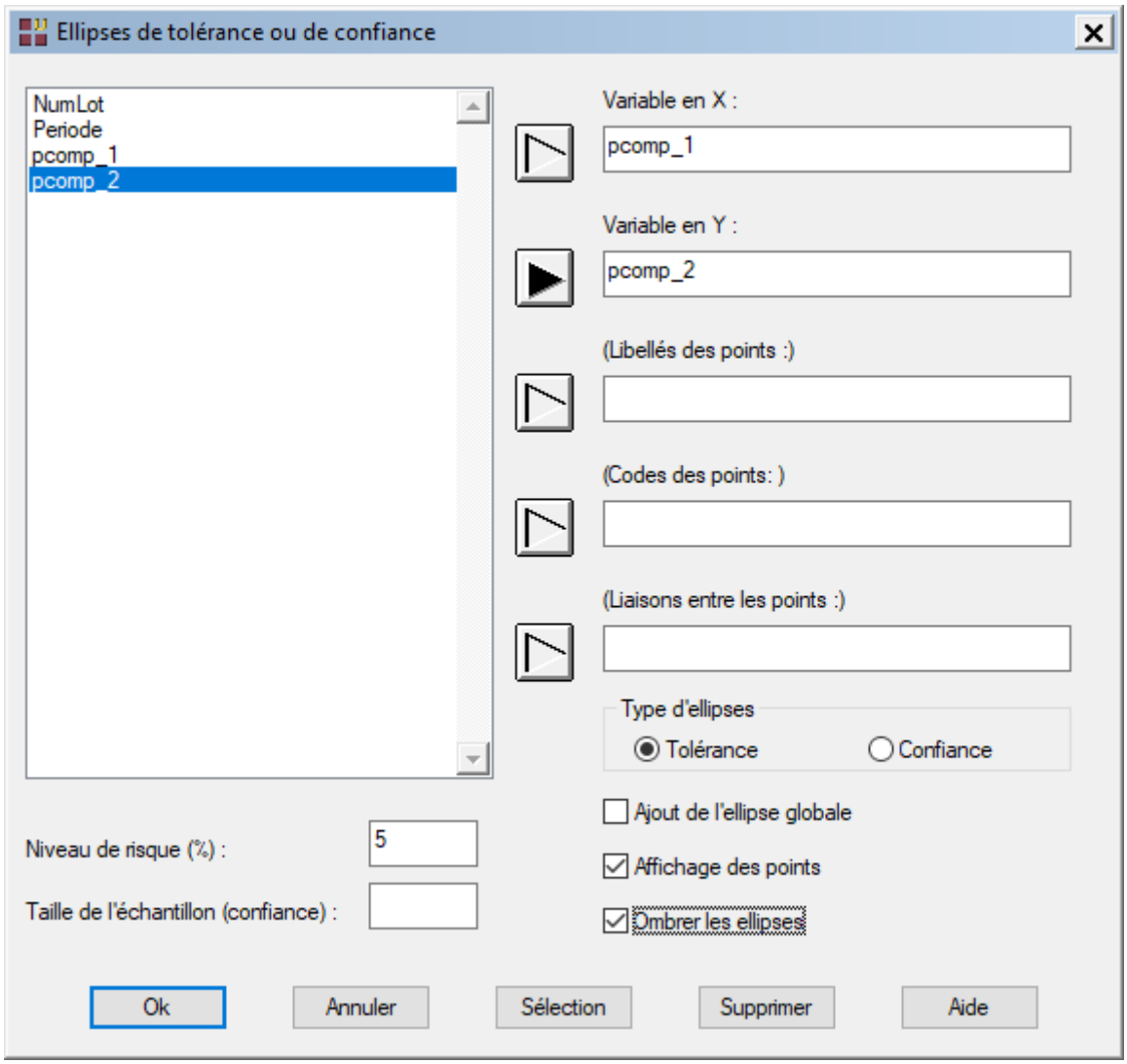


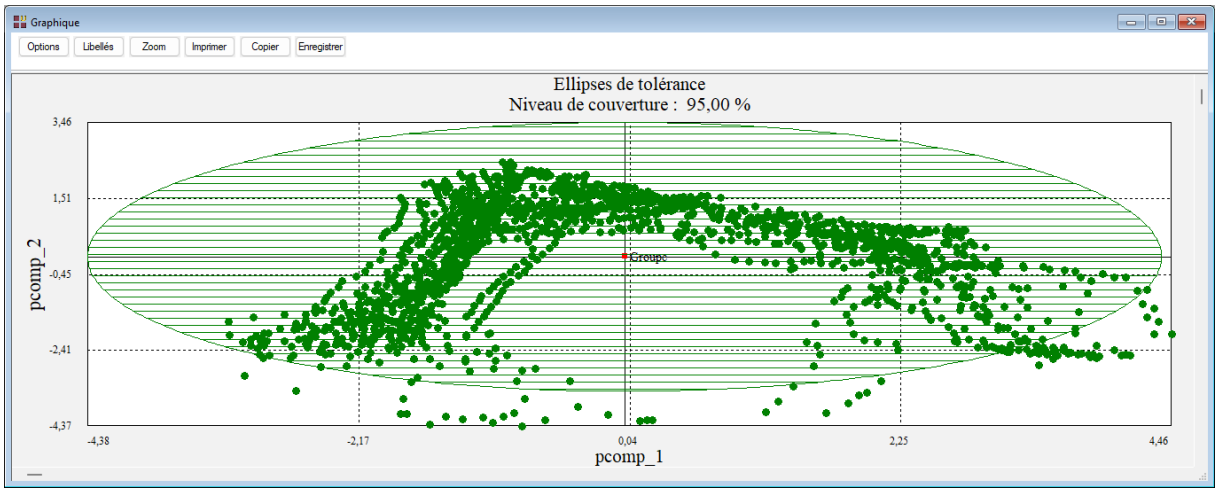
Exemple 3 : Fichier MSPC

Ce fichier contient les résultats d'une analyse en composantes principales effectuée sur des données de production de levure industrielle par lots (voir la documentation relative à l'analyse en composantes principales pour plus de détails sur ces données).

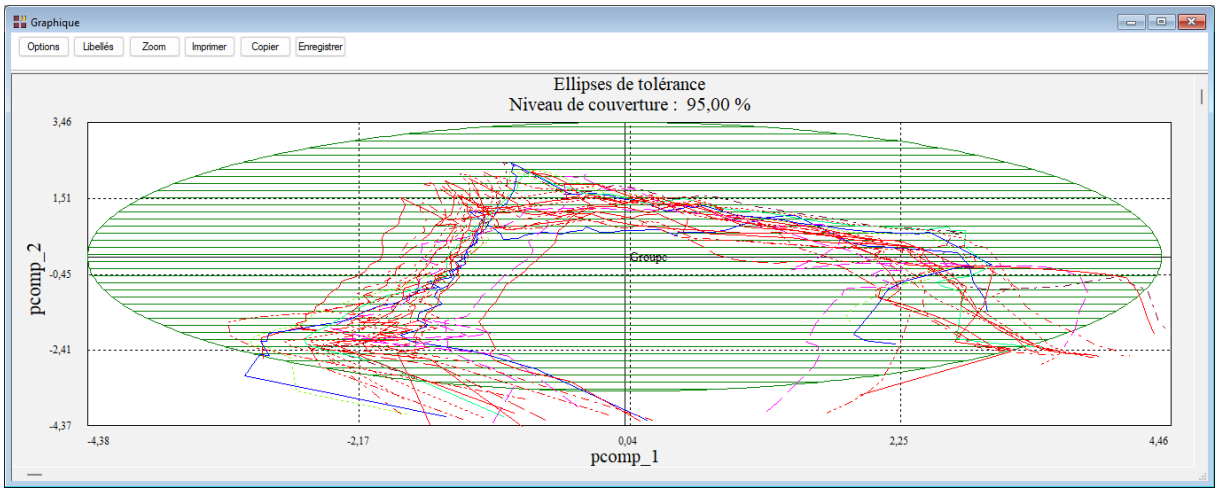
Les variables pcomp_1 et pcomp_2 sont respectivement les première et deuxième composantes principales et définissent les axes des X et des Y.

Nous ne renseignons pas la zone 'Variable des codes' pour obtenir une unique ellipse de tolérance à 95% pour l'ensemble des lots.





Demandons ensuite le tracé avec les liaisons entre les points (variable NumLot) sans affichage des points.



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.

Variable	*Contenu*
groupes	libellés des groupes
effectifs	effectifs des groupes
moyX	moyennes X des groupes
moyY	moyennes Y des groupes
ectX	écarts-types X des groupes
ectY	écarts-types Y des groupes
corXY	corrélations X-Y des groupes