

# 1 Objectif

## Tests de comparaison non paramétriques pour 2 échantillons appariés.

L'appariement est une procédure qui vise à réduire l'effet des fluctuations d'échantillonnage c.-à-d. la variabilité due aux observations. Nous pouvons l'associer à différentes configurations.

Le schéma des « mesures répétées » est le premier qui vient à l'esprit. Il s'agit de mesurer la même grandeur chez un même individu, avant et après intervention d'une action dont on veut justement évaluer les conséquences. Par exemple, on mesure la fièvre chez un malade, on lui donne un médicament, après un certain laps de temps, on lui prend de nouveau sa température : les deux mesures sont confrontées.

L'appariement peut être aussi le fruit de la constitution des données en blocs. Si l'on souhaite comparer l'efficacité de 2 méthodes d'enseignement, les mesures répétées sont inappropriées. Dans les paires d'observations, que l'on appelle « blocs », nous associerons alors des élèves identiques par rapport aux caractéristiques de l'étude. Par exemple, on met dans chaque paire des élèves qui, par le passé, ont obtenu des résultats identiques aux examens.

Enfin, l'appariement peut être tout simplement inhérent à la situation que l'on cherche à analyser. Par exemple, on cherche à comparer le temps passé devant la télévision par l'homme et la femme à l'intérieur d'un couple. Les blocs correspondent naturellement aux ménages. Les hommes et les femmes ne doivent pas être considérés comme des observations indépendantes.

Les tests de comparaisons spécifiques à ce type de configuration présentent une caractéristique particulière : l'appréciation des différences est réalisée prioritairement à l'intérieur des blocs. Dans ce didacticiel, nous présentons deux techniques non paramétriques, le test des signes et le test des rangs signés de Wilcoxon, et une technique paramétrique, le test de Student pour échantillons appariés.

Les aspects théoriques relatifs à ce didacticiel sont décrits dans deux supports de cours accessibles en ligne [http://eric.univ-lyon2.fr/~ricco/cours/cours/Comp\\_Pop\\_Tests\\_Nonparametriques.pdf](http://eric.univ-lyon2.fr/~ricco/cours/cours/Comp_Pop_Tests_Nonparametriques.pdf) et [http://eric.univ-lyon2.fr/~ricco/cours/cours/Comp\\_Pop\\_Tests\\_Parametriques.pdf](http://eric.univ-lyon2.fr/~ricco/cours/cours/Comp_Pop_Tests_Parametriques.pdf).

## 2 Données

Les données proviennent du site de cours en ligne du Pr Richard Lowry du « Vassar College ». Nous traitons l'exemple utilisé pour illustrer le test des rangs signés de Wilcoxon<sup>1</sup>.

On a posé deux questions, QA et QB, à des étudiants, du type « quelle est la probabilité que... ». On cherche à savoir si les valeurs de QA sont stochastiquement différentes de celles de QB. Le principe, les formules et les calculs spécifiques à ces données sont détaillés sur le site web. Nous pouvons suivre à la trace les résultats fournis par TANAGRA.

Les données sont listées dans le fichier **[comparison\\_2\\_related\\_samples.xls](#)**<sup>2</sup>.

---

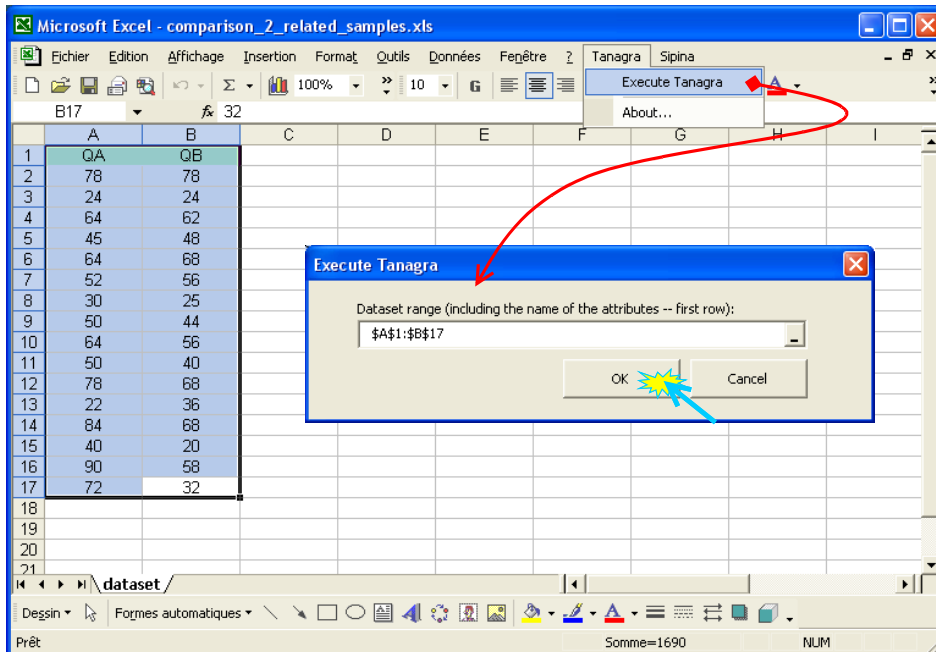
<sup>1</sup> <http://faculty.vassar.edu/lowry/ch12a.html> ; le problème analysé est largement expliqué sur le site.

<sup>2</sup> [http://eric.univ-lyon2.fr/~ricco/tanagra/fichiers/comparison\\_2\\_related\\_samples.xls](http://eric.univ-lyon2.fr/~ricco/tanagra/fichiers/comparison_2_related_samples.xls)

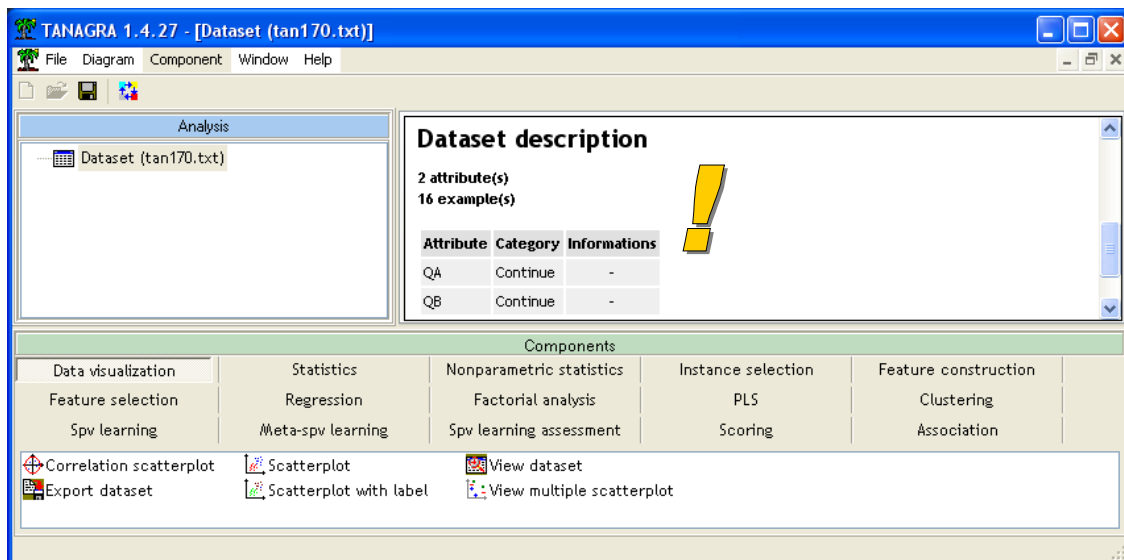
### 3 Les tests non paramétriques

#### 3.1 Importation des données

Le plus simple pour lancer Tanagra et charger les données est d'ouvrir le fichier XLS dans le tableur EXCEL. Nous sélectionnons la plage de données. La première ligne doit correspondre au nom des variables. Puis nous activons le menu TANAGRA / EXECUTE TANAGRA qui a été installé avec la macro complémentaire TANAGRA.XLA<sup>3</sup>. Une boîte de dialogue apparaît. Nous vérifions la sélection. Si tout est en règle, nous validons en cliquant sur le bouton OK.



TANAGRA est automatiquement lancé. Un nouveau diagramme est créé. Nous devons disposer de 16 observations et 2 variables (QA et QB).



Tous les tests que nous présentons dans ce didacticiel vont s'appuyer sur une variable

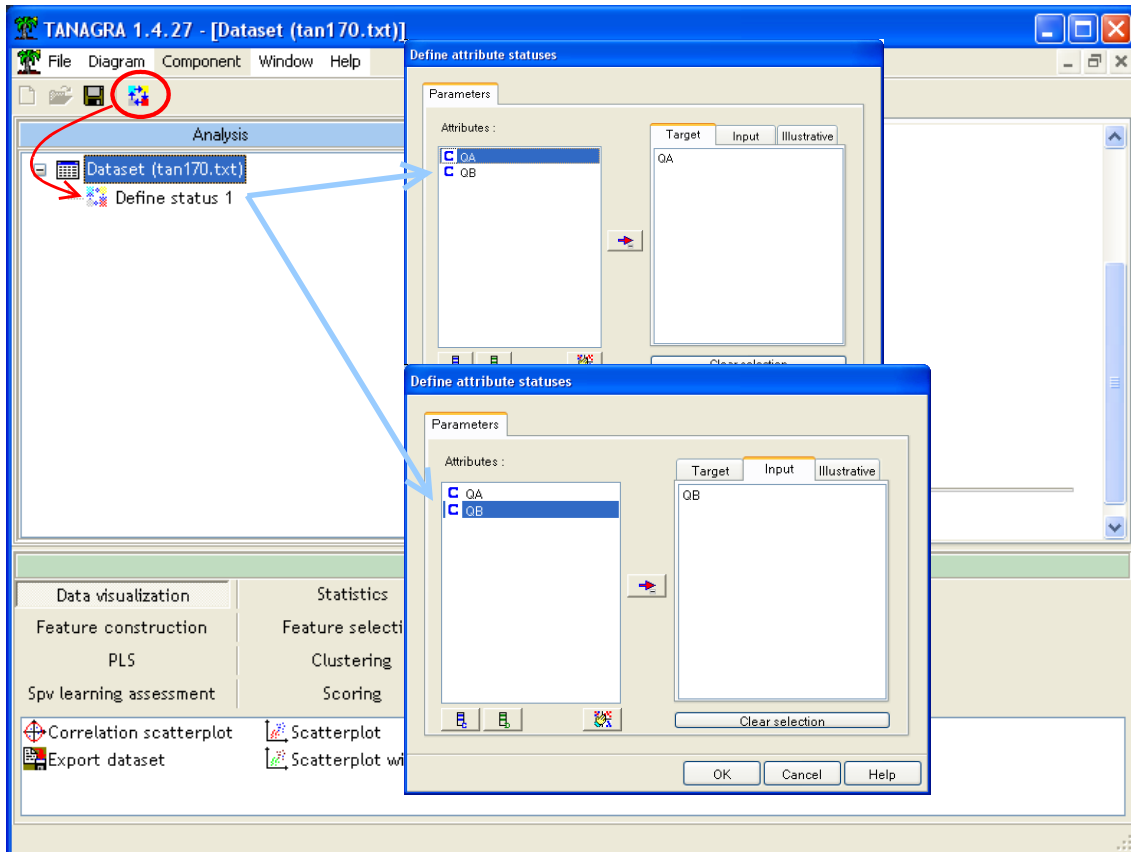
<sup>3</sup> Voir <http://tutoriels-data-mining.blogspot.com/2008/03/importation-fichier-xls-excel-macro.html> concernant l'installation et l'utilisation de la macro complémentaire TANAGRA.XLA.

synthétique, constituée à partir de la différence ligne à ligne des valeurs, soit  $D = QA - QB$ . Ils se différencient par la manière d'exploiter cette nouvelle variable D.

### 3.2 Test des signes

Le test de signes<sup>4</sup> s'intéresse uniquement au signe de la différence D. Sous l'hypothèse nulle c.-à-d. les valeurs de QA et QB sont en moyenne identiques, la proportion théorique de valeurs D positives est 0.5. On cherche à savoir si, dans nos données, on s'en écarte significativement.

Via la barre d'outils, nous insérons le composant DEFINE STATUS dans le diagramme. Nous plaçons QA en TARGET et QB en INPUT.



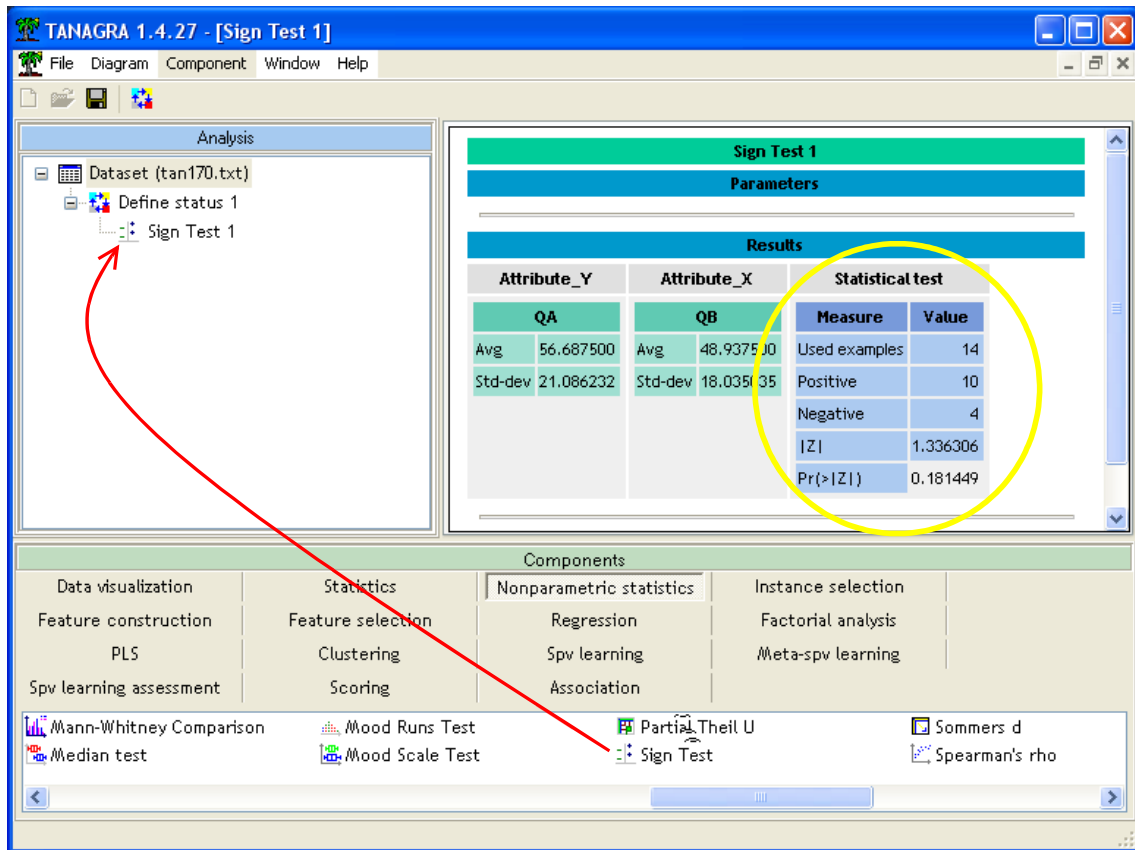
Nous introduisons à la suite le composant SIGN TEST (onglet NONPARAMETRIC STATISTICS). Nous cliquons sur le menu contextuel VIEW pour accéder aux résultats.

Initialement, nous disposons de 16 observations, seules  $n = 14$  sont utilisables. Elles correspondent aux observations pour lesquelles  $(D \neq 0)$ . Il y a  $S = 10$  observations  $(D > 0)$  et  $n - S = 4$  pour  $(D < 0)$ . La statistique standardisée, incluant la correction de continuité, est obtenue avec

$$|Z| = \frac{|2S - n| - 1.0}{\sqrt{n}} = \frac{|2 \times 10 - 14| - 1.0}{\sqrt{14}} = 1.336306$$

La probabilité critique du test est  $p = 0.181449$ .

<sup>4</sup> <http://udel.edu/~mcdonald/statsign.html> ; [http://en.wikipedia.org/wiki/Sign\\_test](http://en.wikipedia.org/wiki/Sign_test)



Pour un risque à 5%, on considère que les différences ne sont pas significatives.

Cependant, le test de signes est réputé peu puissant. Il ne s'appuie que sur le sens des écarts, il n'exploite pas leur importance. Il a par conséquent tendance à accepter un peu trop souvent l'hypothèse nulle. La conclusion ci-dessus est à prendre avec précautions, nous devons la confronter avec d'autres procédures.

### 3.3 Test des rangs signés de Wilcoxon

Le test des rangs signés de Wilcoxon utilise l'importance relative des écarts lors de la construction de la statistique<sup>5</sup>. Il est plus à même de relever finement les différences.

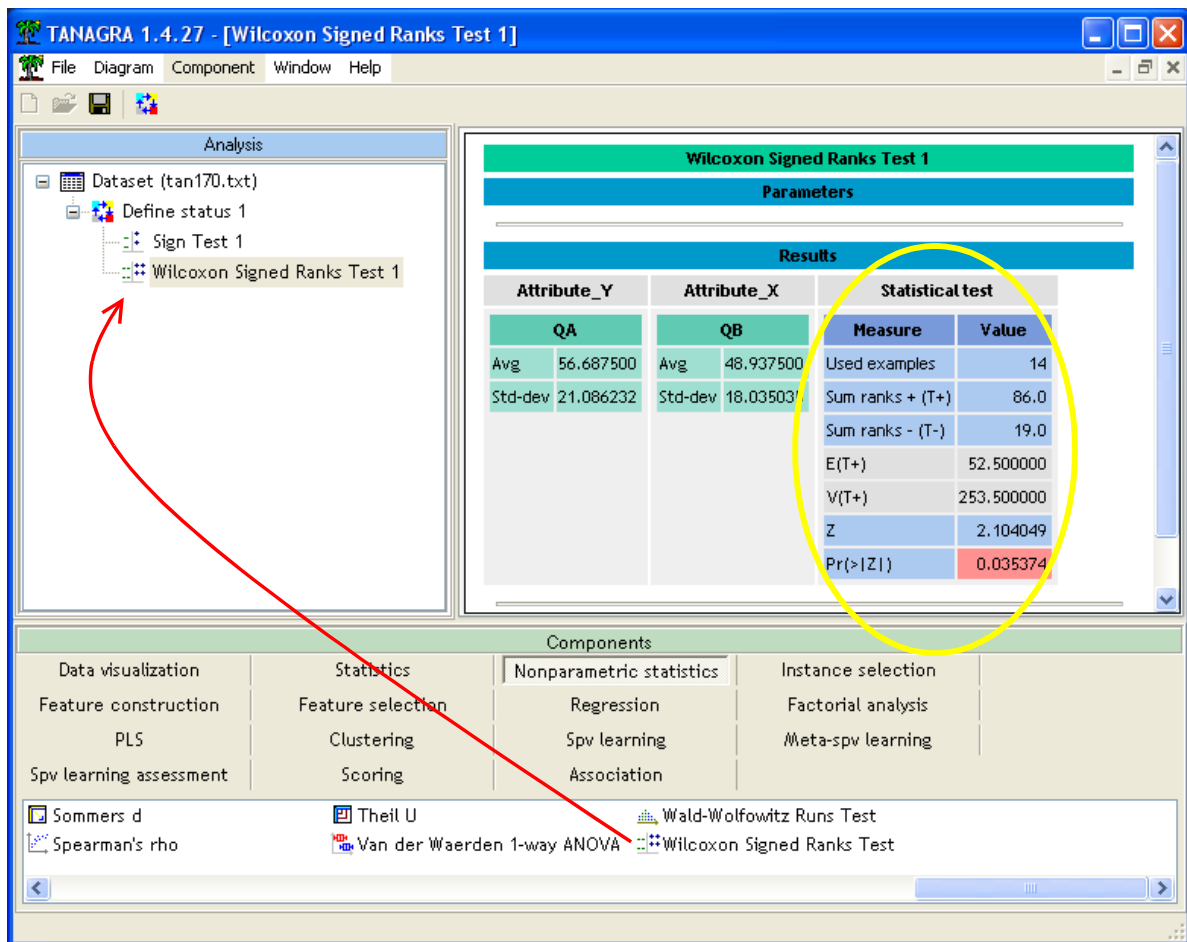
Nous introduisons le composant WILCOXON SIGNED RANKS TEST (onglet NONPARAMETRIC STATISTICS) dans notre diagramme.

Ici également, le nombre d'observations exploitables est  $n = 14$  ( $D \neq 0$ ). La somme des rangs des écarts positifs est  $T^+ = 86$ . Son espérance et sa variance sous l'hypothèse nulle sont  $E(T^+) = 52.5$  et  $V(T^+) = 253.5$ . La statistique standardisée pour un test bilatéral, sans la correction de continuité ici, s'écrit

$$|Z| = \frac{|T^+ - E(T^+)|}{\sqrt{V(T^+)}} = \frac{|86 - 52.5|}{\sqrt{253.5}} = 2.104049$$

Avec une probabilité critique  $p = 0.035374$ .

<sup>5</sup> Voir <http://faculty.vassar.edu/lowry/ch12a.html> pour la théorie et le détail des calculs sur cet exemple.



Maintenant, parce que nous exploitons mieux l'information disponible, nous observons bien une différence significative entre les niveaux des réponses.

## 4 Comparaison de moyennes pour échantillons appariés

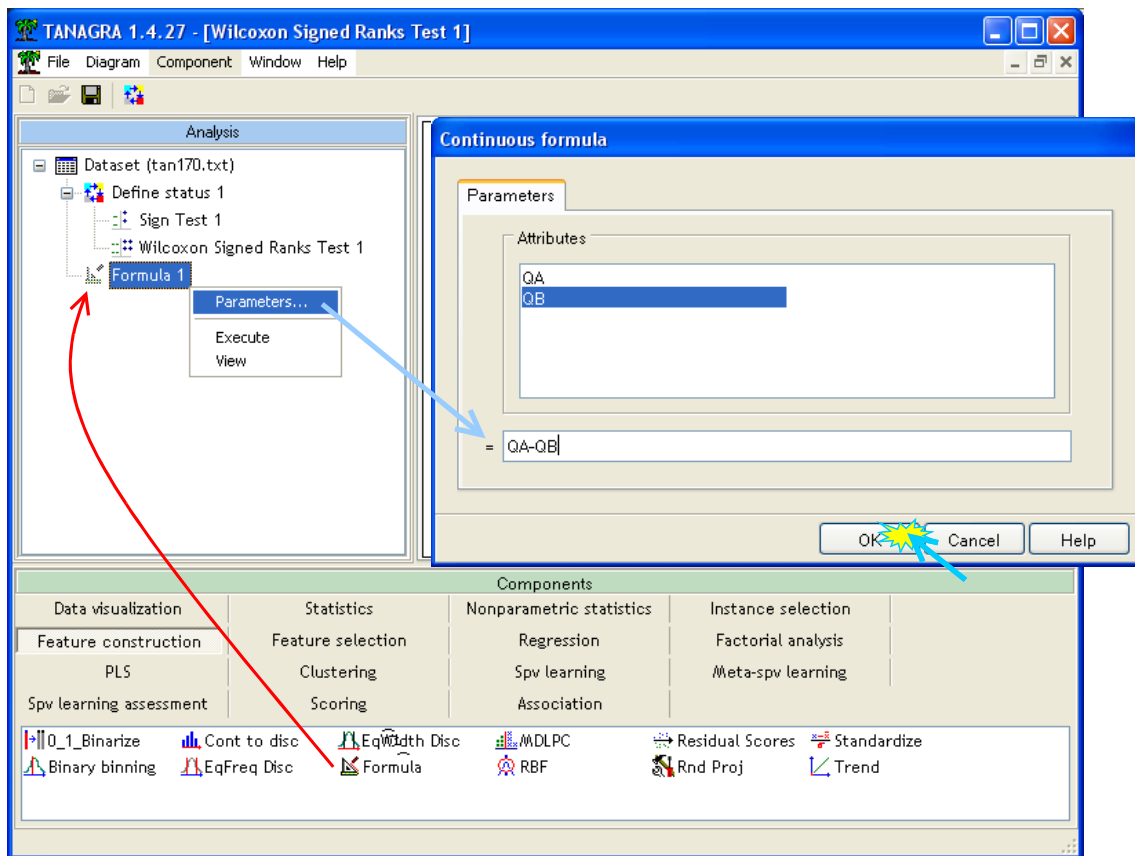
Le test de Student pour échantillons appariés va plus loin encore dans l'exploitation des données. Il s'appuie sur l'amplitude des écarts, la différence  $D = QA - QB$  lors de la formation de la statistique de test. Schématiquement<sup>6</sup>, il s'agit simplement d'un test de conformité de la moyenne de  $D$  avec le standard 0.

### 4.1 Former la différence et vérifier la normalité

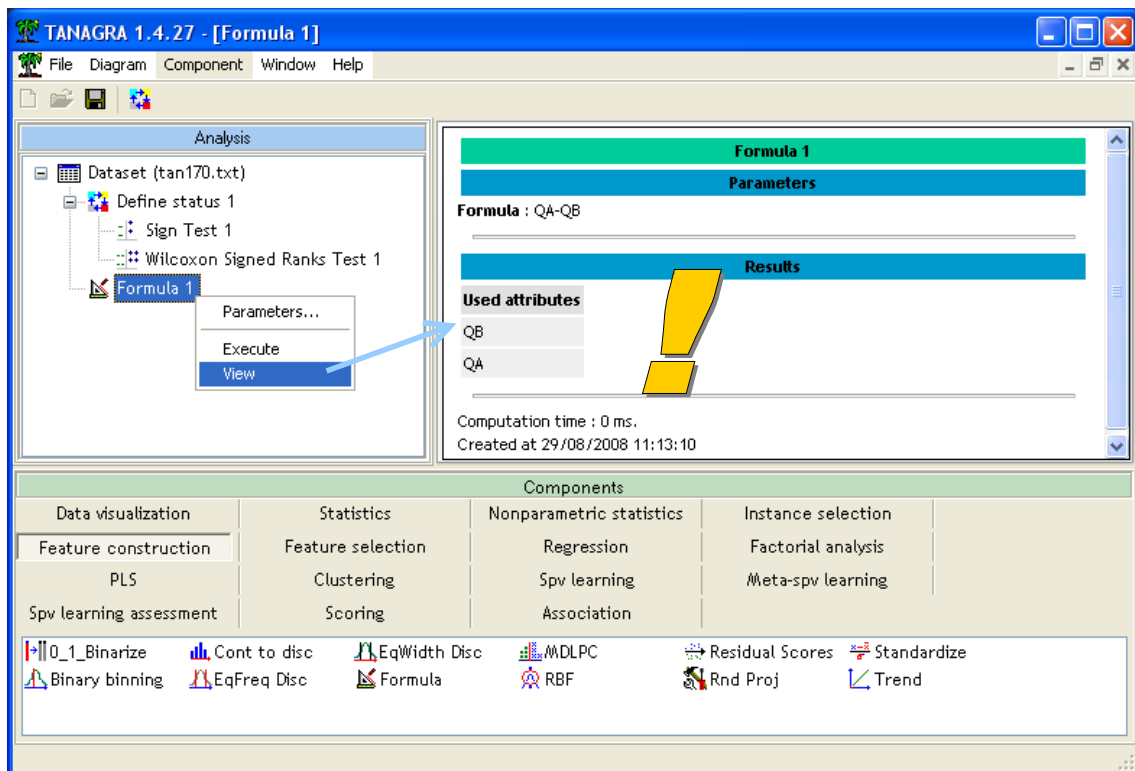
Pour que le test soit valable, la distribution de  $D$  doit être compatible avec la loi normale. Nous allons former explicitement la variable synthétique  $D$  et la contrôler à l'aide des tests de normalité usuels.

Nous insérons de composant FORMULA (onglet FEATURE CONSTRUCTION) dans le diagramme. Nous activons le menu contextuel PARAMETERS pour introduire la formule de la variable  $D$ .

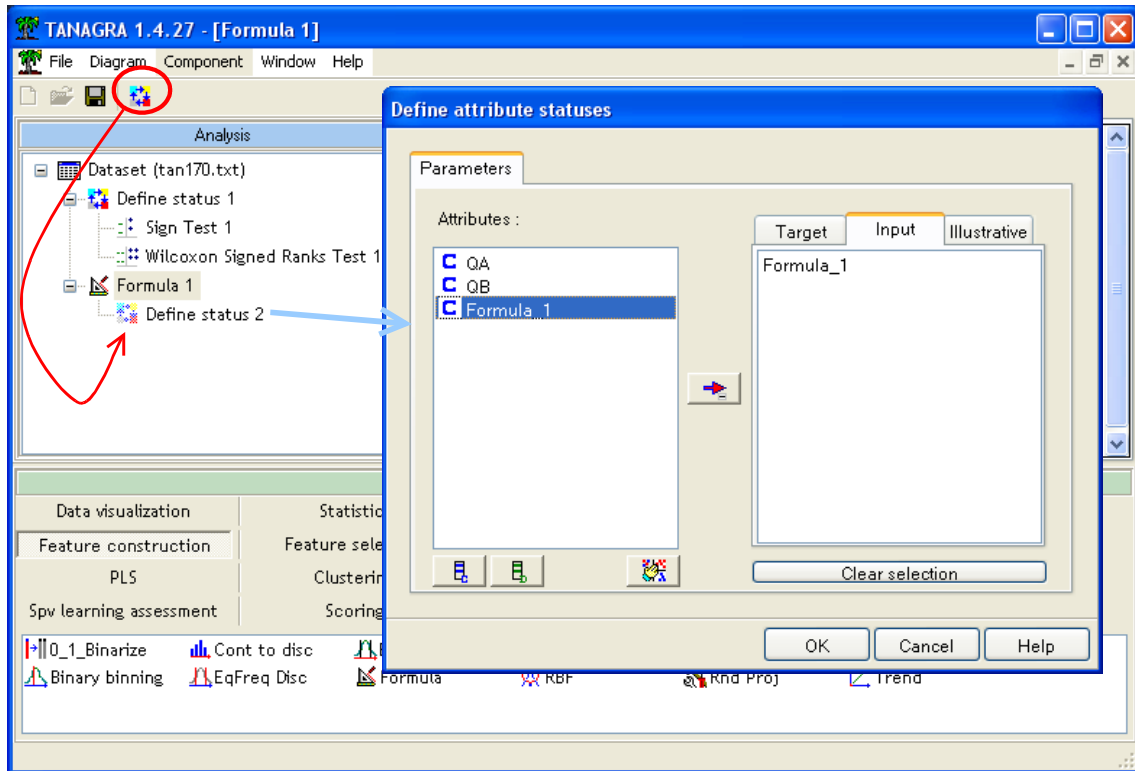
<sup>6</sup> Voir <http://faculty.vassar.edu/lowry/ch12pt1.html> pour les aspects théoriques et les calculs.



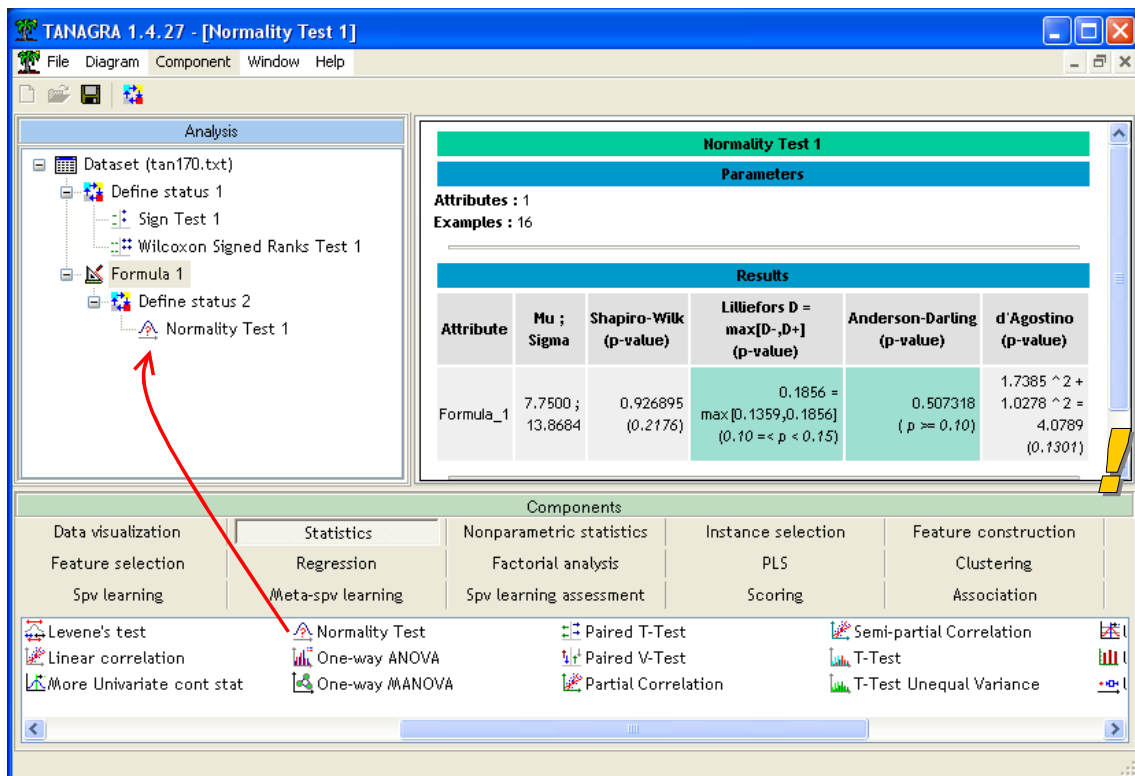
Nous validons. Nous cliquons sur le menu VIEW. Nous obtenons la liste des variables utilisées et la formule de la nouvelle variable. En interne, TANAGRA génère une nouvelle colonne de données qu'il nomme FORMULA\_1, nous pourrions l'exploiter dans les branches subséquentes du diagramme.



Nous insérons de nouveau DEFINE STATUS dans le diagramme, nous plaçons en INPUT la variable FORMULA\_1.



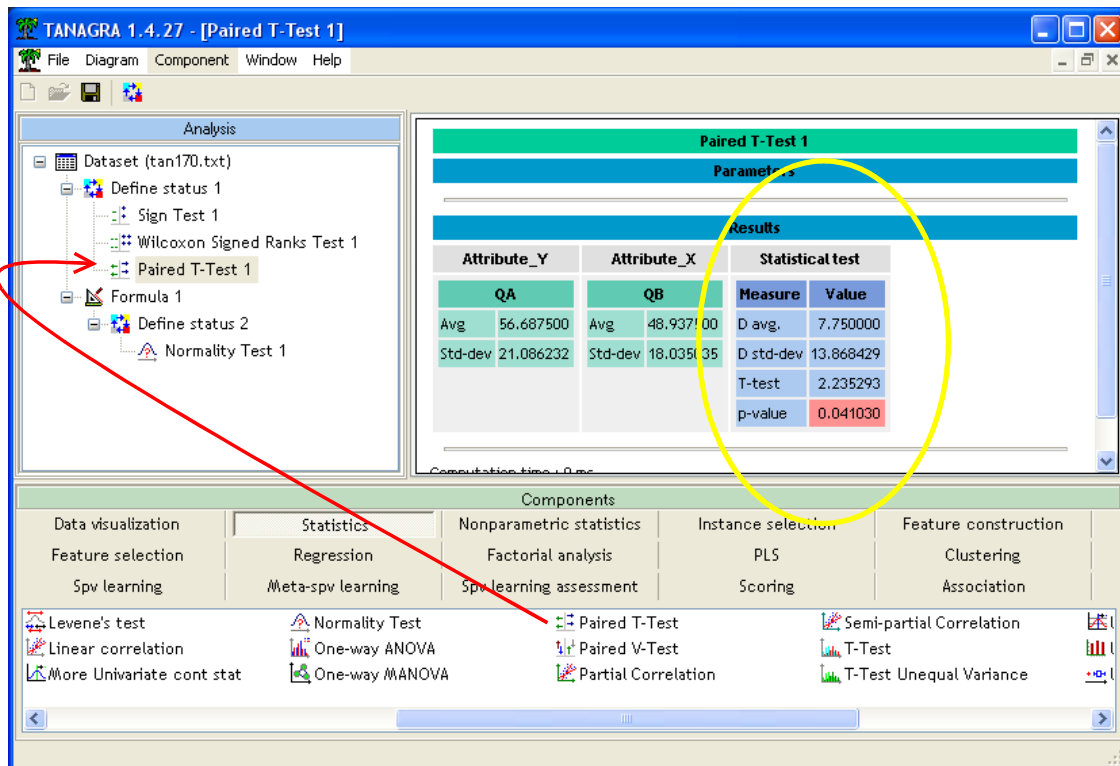
Nous introduisons alors le composant NORMALITY TEST (onglet STATISTICS).



Tous les voyants sont au vert (si je puis dire). Sur tous les tests proposés<sup>7</sup>, pour un risque à 5%, la compatibilité des données avec la loi normale n'est pas remise en cause.

## 4.2 Test de Student pour échantillons appariés

Nous pouvons passer au test proprement dit. Nous insérons le composant PAIRED T-TEST (onglet STATISTICS) dans le diagramme, **en dessous de DEFINE STATUS 1**. Nous activons le menu VIEW.



TANAGRA nous indique que la moyenne de D est  $\bar{D} = 7.75$ , son écart type  $s_D = 13.868429$ . Nous pouvons former la statistique t. Nous tenons compte des écarts nuls pour ce test, nous utilisons tout l'effectif c.-à-d.  $n = 16$

$$t = \frac{\bar{D}}{s_D / \sqrt{n}} = \frac{7.75}{13.868429 / \sqrt{16}} = 2.235293$$

Sous  $H_0$ , t suit une loi de Student à  $(n-1) = 16 - 1 = 15$  degrés de liberté. La probabilité critique pour un test bilatéral est  $p = 0.04103$ .

On remarquera la proximité de ce résultat avec celui du test des rangs signés de Wilcoxon.

## 5 Représentation graphique

Le test des rangs signés de Wilcoxon et le test de Student pour échantillons appariés sont plus puissants que le test des signes (qui est très conservateur). Ceci étant, un petit graphique nous aurait indiqué dès le départ la bonne conclusion. Dans un nuage de point avec QA en abscisse et

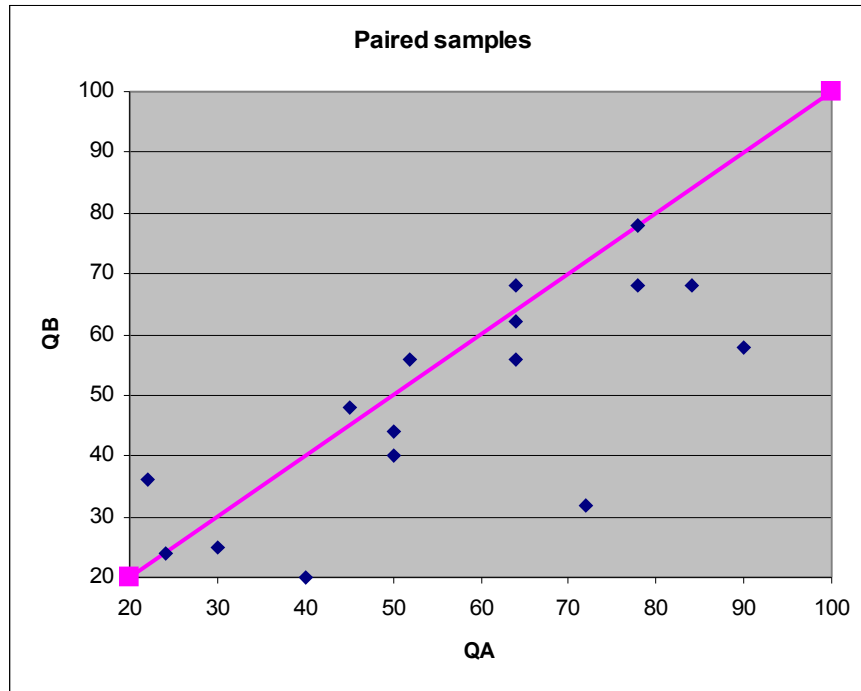
<sup>7</sup> Voir [http://eric.univ-lyon2.fr/~ricco/cours/cours/Test\\_Normalite.pdf](http://eric.univ-lyon2.fr/~ricco/cours/cours/Test_Normalite.pdf) pour une présentation détaillée de ces tests. Ils sont par ailleurs mis en avant dans d'autres tutoriels, notamment dans <http://tutoriels-data-mining.blogspot.com/2008/04/tests-dadquation-la-loi-normale.html>



QB en ordonnée, si les valeurs étaient équivalentes ( $H_0$ ), nous devrions voir les points plus ou moins alignés sur la diagonale principale.

Pour nous (Figure 1), les points sont plutôt situés en majorité en dessous de la diagonale. Et lorsque c'est le cas, ils s'en écartent plus.

Les valeurs sont en moyenne différentes et, à la lumière de ce graphique, nous pouvons même dire que QA a tendance à être plus grand que QB.



**Figure 1 - Nuage de points QA vs. QB**