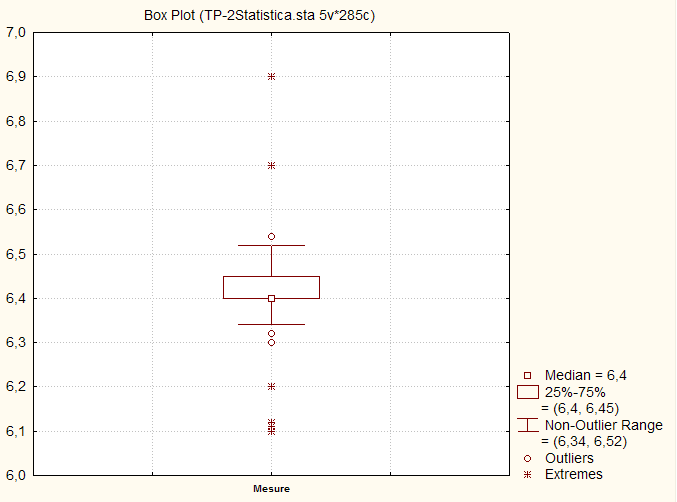
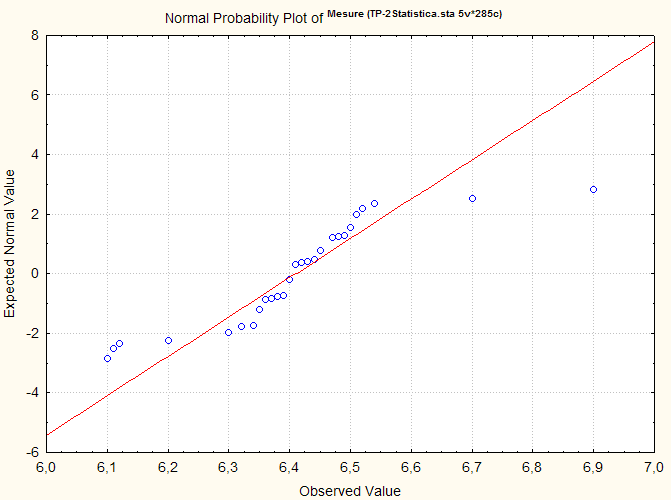
**TD-2 Longueurs de clous Partie Deux**

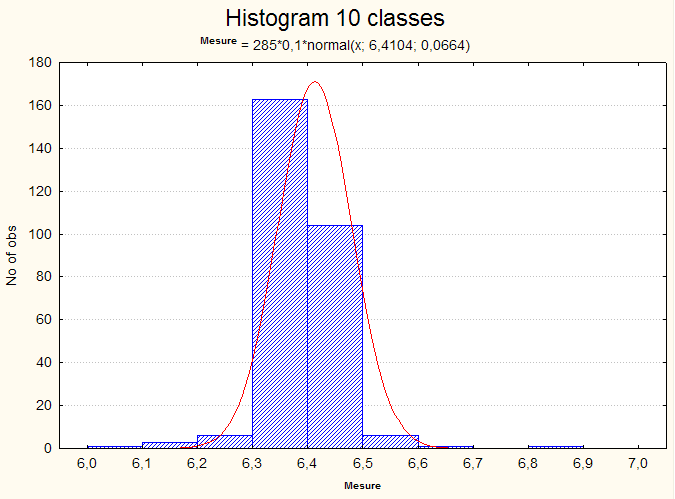
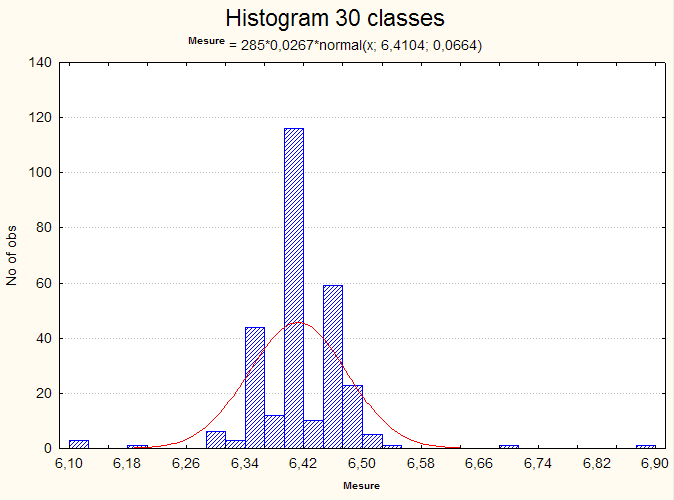
**Étude de la variabilité des mesures**

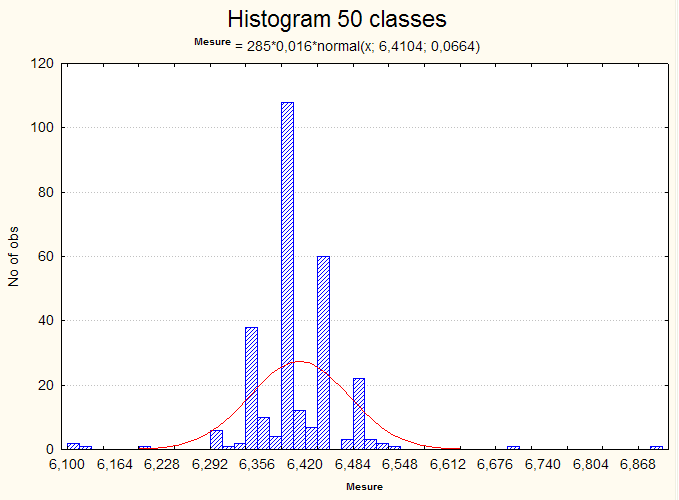
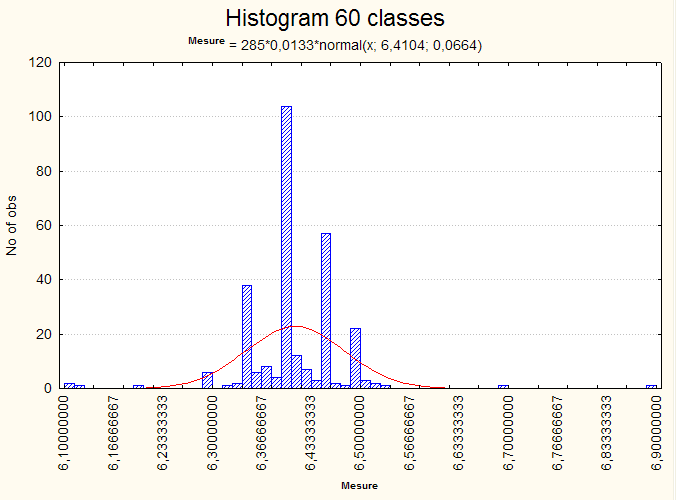
**1(a)** Les histogrammes montrent l’allure générale de la fonction de densité, s’il y a assez de données; les diagrammes de Tukey en montrent assez clairement la symétrie, où se situent les quartiles et la médiane, si on trouve une symétrie, ainsi que les données suspectes et aberrantes. Les diagrammes quantiles-quantiles (droites de Henry) font voir l’adéquation des données à un modèle, dans le cas des droites de Henry, on examine l’adéquation au modèle gaussien.

**1(b)** Les points suivants sont aberrants : valeurs 6,7; 6,9; et les données < 6,25. C’est le diagramme de Tukey qui les singularisent, de même que la droite de Henry (voir les figures). Elles appartiennent (examen des données) à l’opérateur 6, règle 1 : pour 6,7, 6,9, 6,12, 6,1; et à l’opérateur 2, règle 2 pour 6,2.

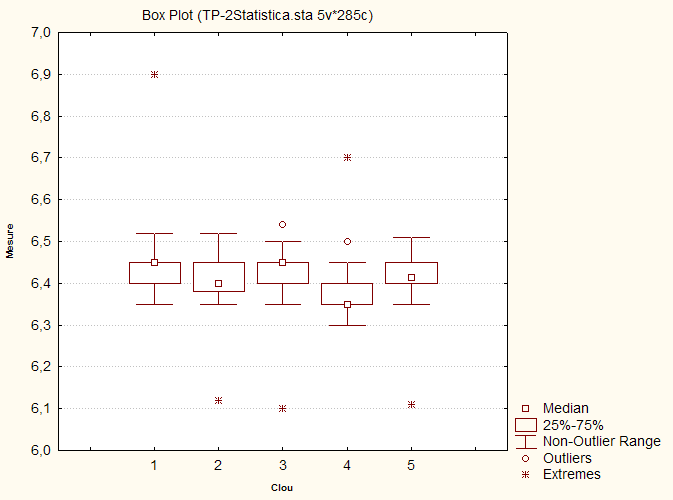
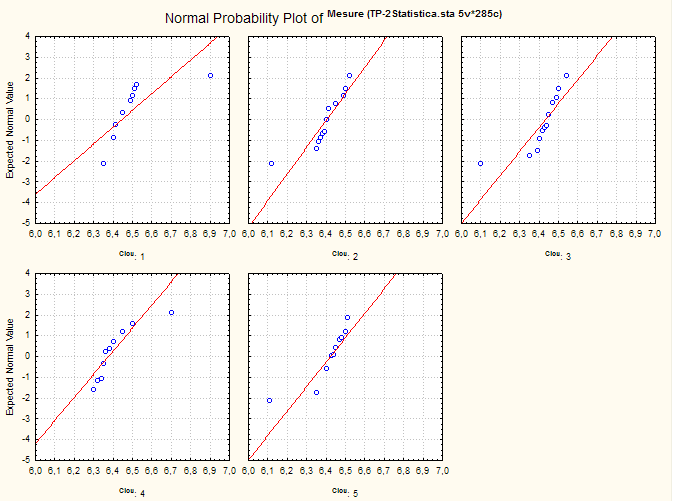
 

**1(c)** Les 4 histogrammes suivants montrent que plus le nombre de classes augmentent plus les ‘classes creuses’ sont nombreuses, et plus ressortent comme fréquentes certaines valeurs. On voit que certaines valeurs sont mesurées plus souvent. Un examen des données montre que ce sont les valeurs au milimètre (6,4) ou au dixième de millimètre avec la valeur ‘5’ comme centième de centimètre. Les valeurs aberrantes ressortent également mieux quand on augmente le nombre de classes.

** **

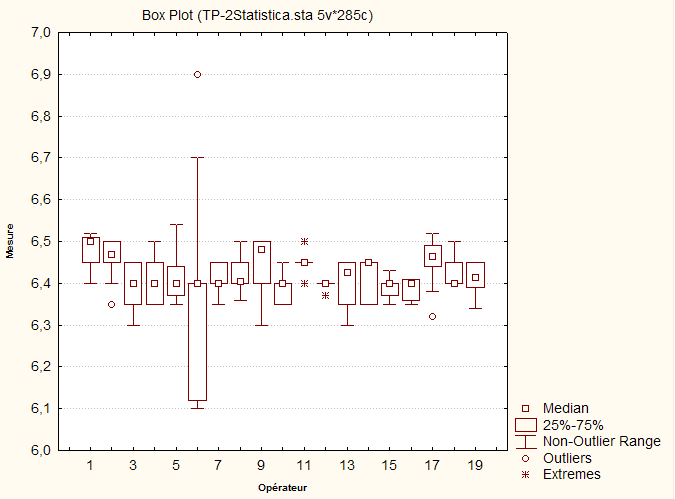
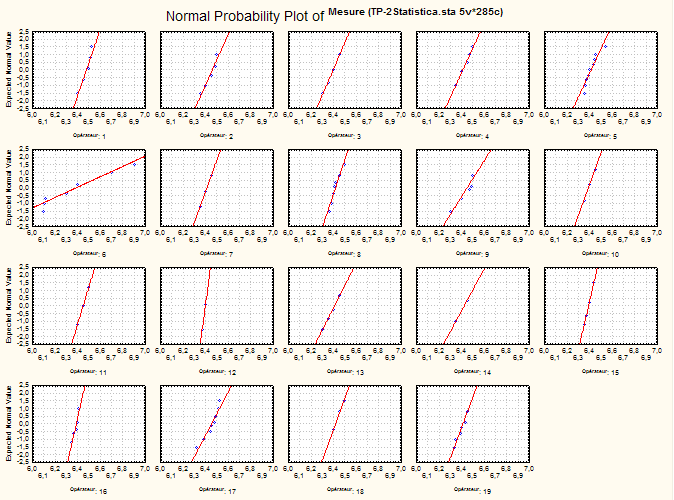
** **

**2(a)** Les mêmes mesures qu’au numéro précédent sont qualifiées d’aberrantes. Elles empêchent les données de coller sur les droites de Henry. Ainsi les données aberrantes viennent contredire l’hypothèse qu’on pourrait avoir sur le fait que les mesures sont gaussiennes. Par ailleurs, le clou 4 semble plus petit que les autres (hypothèse).

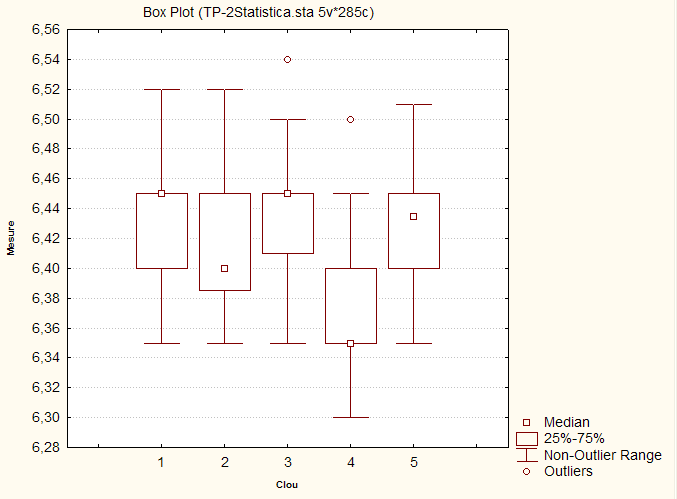
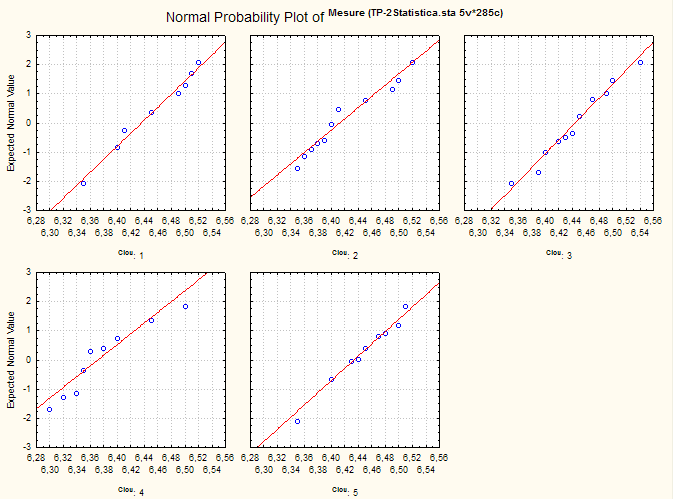
** **

**2(b)** Au diagramme de Tukey, il appert que le clou 4 est plus petit que les autres (on pourrait formuler cette hypothèse).

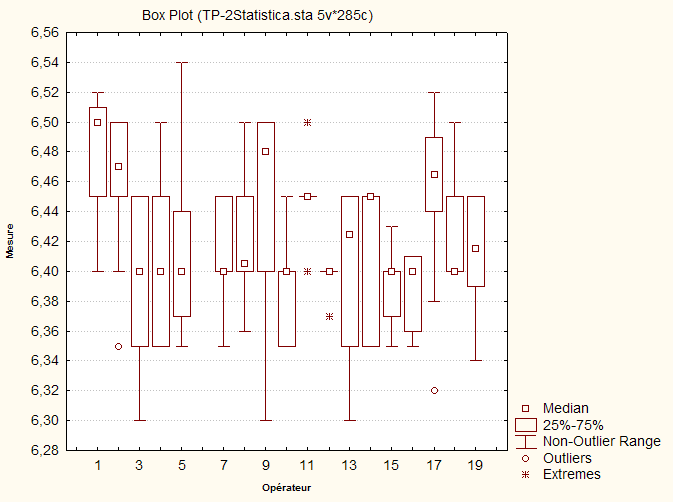
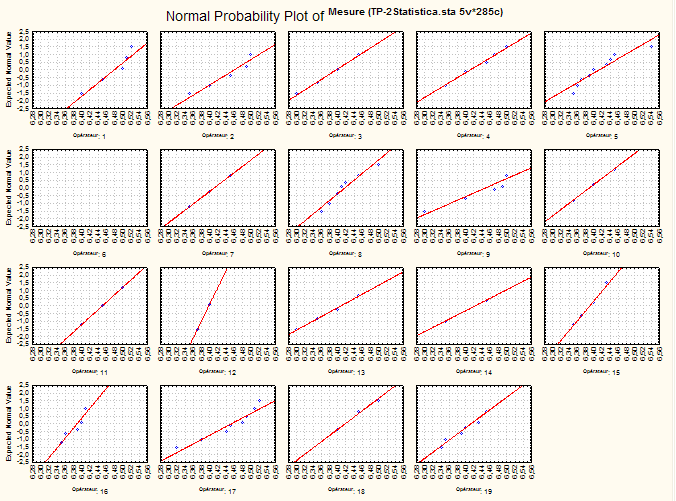
**2(c)** On voit aux graphiques suivants que l’opérateur 6 est très singulier. On sait que lorsque les données sont considérées globalement, 3 des données aberrantes en proviennent. La boîte lui correspondant des diagrammes de Tukey des mesures groupées par opérateurs ne montrent pas ces données comme aberrantes, mais la boîte est très à part de celles des autres mesureurs. Par ailleurs, la pente de la droite de Henry associée aux mesures de cet opérateur est très à part des autres pentes.

**3(a)** Après avoir supprimé l’opérateur 6, on ne trouve plus de données aberrantes pour aucun des clous, la symétrie des mesures (il y en a 38 par clous) n’est pas assurée (diagramme de Tukey regroupés par clous), les données sont beaucoup plus gaussiennes : ce qu’on constate plus par les droites de Henry que par les diagrammes de Tukey qui montrent plutôt une asymétrie des mesures groupées par clous.

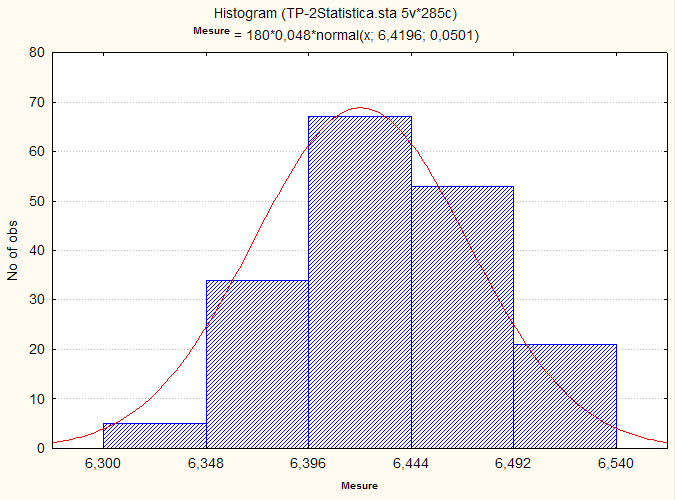
** **

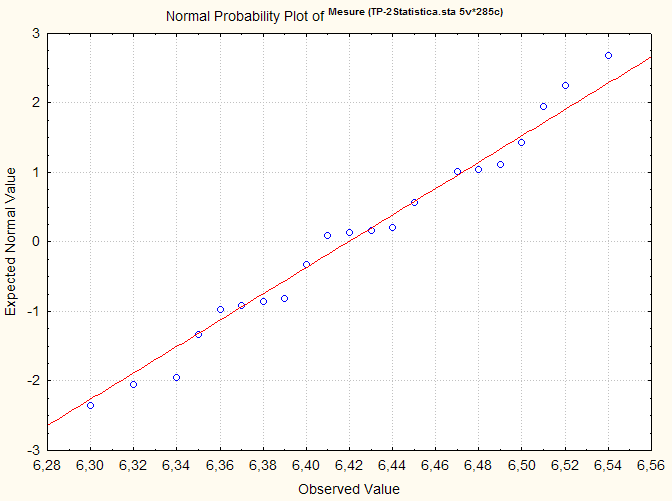
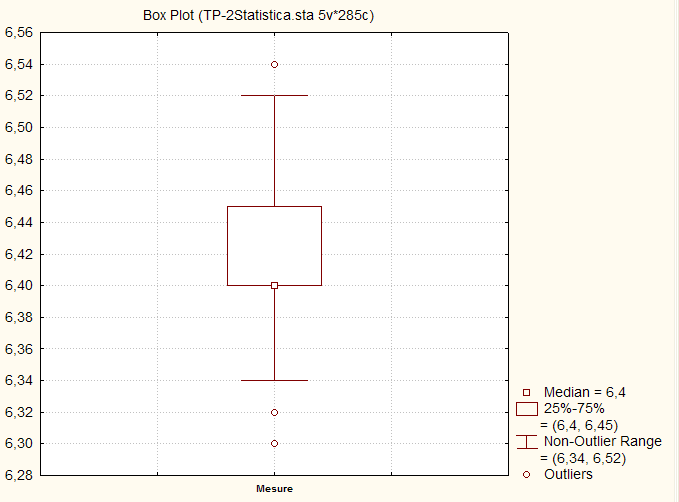
**3(b)** Lorsqu’on examine les diagrammes de Tukey et les droites de Henry regroupés par opérateur, on tempère le jugement. D’opérateur à opérateur, il semble que les mesures soient plutôt non homogènes. Ainsi certains opérateurs ont des mesures bien groupées, d’autres passablement plus dispersées et assez peu symétriques pour certains (diagramme de Tukey groupé par opérateur) . À noter qu’on a seulement 10 mesures (2mesures x 5 clous) par opérateur ce qui est peu pour établir un jugement fort. Regroupées par clous (36 mesures par clou : 18 opérateurs x 2 essais), les droites de Henry sont très voisines, ce qui n’est pas aussi avec les droites de Henry regroupées par opérateurs (mais seulement 10 mesures).

** **

**4(a)**  On a 36 mesures par clous et 5 clous, soit en tout 36x5=180 mesures au total.

**4(b)** Les données suspectes pour le diagramme de Tukey ne le sont pas pour la droite de Henry. Aucun point ne se singularise par rapport à la droite de Henry.

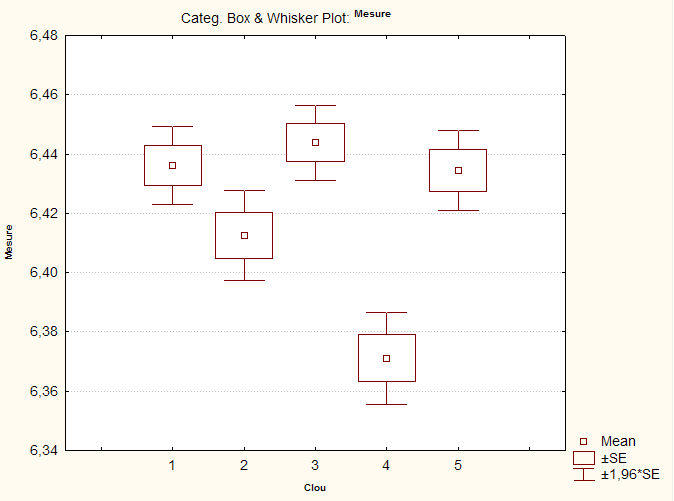
****

****

**4(c)** On admettra sans peine que les mesures suivent un modèle gaussien.

**5(a)**  On voit que la vraie moyenne de la longueur pour le 4e clou serait située dans un intervalle disjoint de ceux associés à la vraie moyenne des longueurs des autres clous. Et il en est de même pour le 2e cl

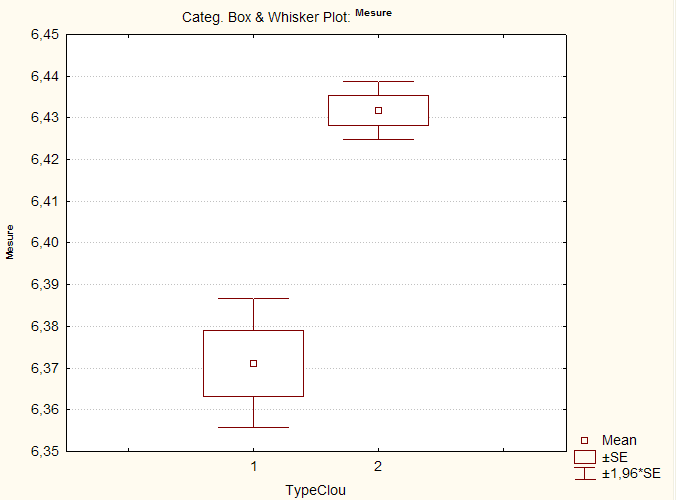
|  |  |  |  |
| --- | --- | --- | --- |
| **Clou** | **Moyenne** | **N** | **Écart type** |
| **1** | 6,436 | 36 | 0,040 |
| **2** | 6,413 | 36 | 0,047 |
| **3** | 6,444 | 36 | 0,039 |
| **4** | 6,371 | 36 | 0,047 |
| **5** | 6,434 | 36 | 0,042 |
| **Total** | 6,420 | 180 | 0,050 |

****

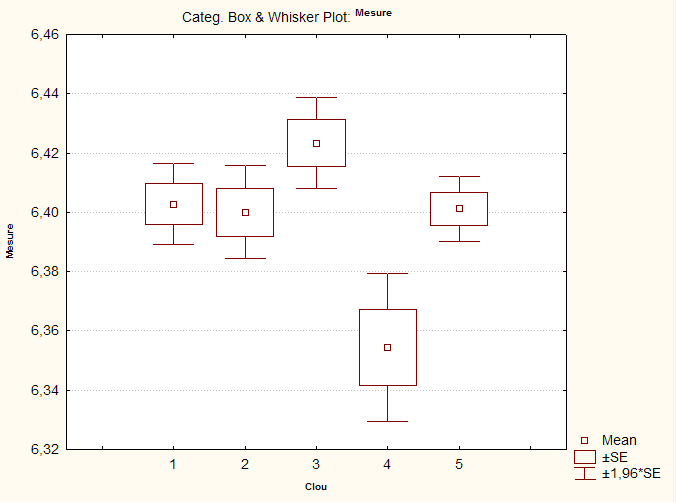
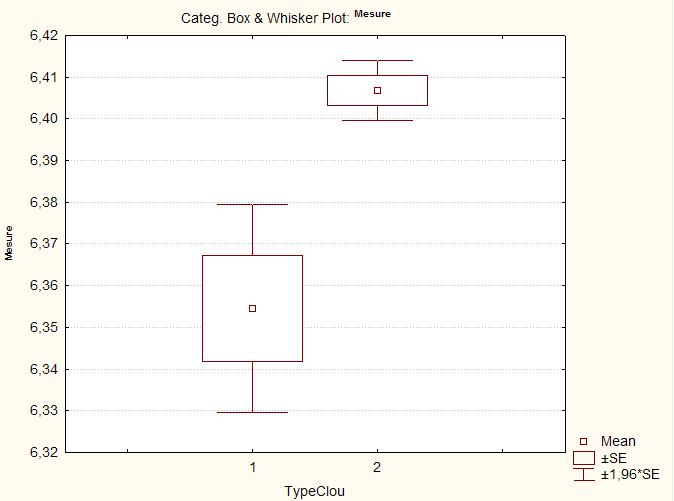
**5(b)**  Les deux intervalles ne s’empiètent pas.

**5(c)** Pour le clou 4, l’intervalle de confiance à 95% est [6,356 ; 6,387], et pour les autres clous : [6,439 ; 6,425]. Ces intervalles ne s’empiètent pas, les deux moyennes sont présumées différentes.

|  |  |  |  |
| --- | --- | --- | --- |
| Type Clou | **Moyenne** | **N** | **Écart type** |
| 4 | 6,371 | 36 | 0,047 |
| les autres | 6,432 | 144 | 0,043 |
| All Grps | 6,420 | 180 | 0,050 |

****

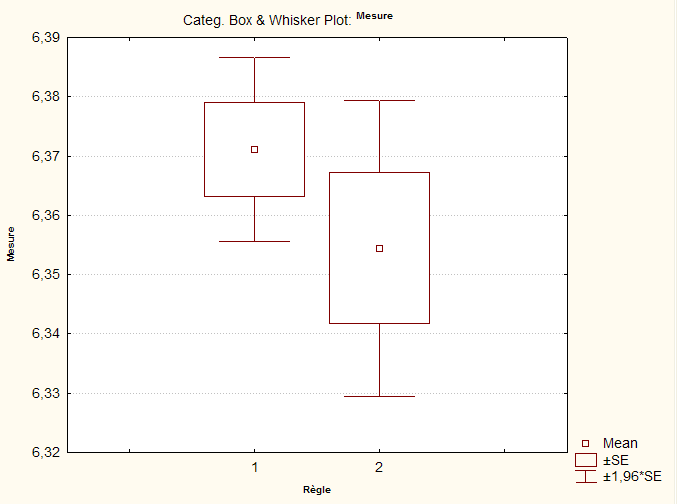
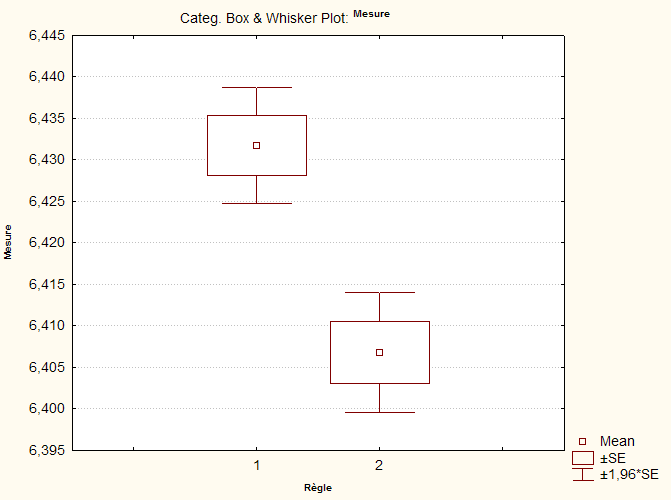
6. Voici les deux graphiques : des diagrammes de Tukey adaptés pour montrer les intervalles de confiance.

De celui à gauche, on déduit que le quatrième clou est plus court que les 4 autres, et que probablement les 4 autres sont homogènes, i.e. de même longueur. On raisonne sur les empiètements des intervalles de confiance. À droite le diagramme de Tukey montre les intervalles de confiance pour les deux types de clous : type 1, le quatrième, type 2, les autres pris tous ensembles.

**7(a)** On admet, suite à la discussion du numéro précédent, que les 4 clous à part le quatrième auraient mêmes longueurs. On peut les mettre dans une même catéégorie

**7(b)** Le diagramme de Tukey à gauche est pour le quatrième clou. Les deux intervalles de confiance pour la moyenne sont empiétants, et même si la moyenne pour la deuxième règle est inférieure à celle pour la première, on ne peut dire que la deuxième règle donne une mesure plus petite que pour la première. Tout au contraire pour les autres clous (diagramme de Tukey à droite). Les intervalles non empiétants du diagramme pour les autres clous montreraient cette fois que la règle donne des mesures inférieures.

Au tableau suivant, on constate (au besoin faire des calculs : mais noter mentalement les constatations…) ce que nous ont illustré les deux graphiques précédents. Mais on a aussi le fait que les deux règles ont des moyennes et des écarts types quasi égales pour les deux types de clous, que pour les autres clous, et pourtant pour le clou 4, les intervalles de confiance sont très empiétants, et non pour les autres clous où ils sont très disjoints. Ce qui explique cette différence sont les tailles échantillonnales de part et d’autre. Comme la racine des tailles joue aux dénominateurs des demi-tailles des intervalles de confiance, ils sont beaucoup plus petits pour des tailles échantillonnales élevées que pour des tailles petites, toutes choses étant égales par ailleurs.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Clou 4** |  |  | **Autres clous** | |  |
| **Règle** | **Mesure** | **Mesure** | **Mesure** | **Mesure** | **Mesure** | **Mesure** |
| **1** | 6,371 | 36 | 0,0474 | 6,432 | 144 | 0,043 |
| **2** | 6,354 | 18 | 0,0540 | 6,407 | 72 | 0,031 |
| **All Grps** | 6,366 | 54 | 0,0498 | 6,423 | 216 | 0,041 |