



HAL
open science

Méthodologie et contrôle de qualité

Mamadou Bassirou Diallo

► **To cite this version:**

Mamadou Bassirou Diallo. Méthodologie et contrôle de qualité. Méthodologie [stat.ME]. 2010. dumas-00517510

HAL Id: dumas-00517510

<https://dumas.ccsd.cnrs.fr/dumas-00517510v1>

Submitted on 14 Sep 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITE DE STRASBOURG

RAPPORT DE STAGE

MASTER DE STATISTIQUE

MÉTROLOGIE ET CONTRÔLE DE QUALITÉ



BURKERT SAS
67 220 TRIEMBACH AU VAL

ETUDIANT : DIALLO MAMADOU BASSIROU
MAITRE DE STAGE : FIXARY ANTOINE

DÉBUT : JEUDI 3 JUIN 2010
FIN : VENDREDI 13 AOÛT 2010

Remerciements

Je tiens tout d'abord à remercier mon maître de stage, M. Antoine FIXARY, ainsi que M. Bruno THOUVENIN, responsable du service qualité, de m'avoir accueilli au sein de l'entreprise Bürkert SAS.

Mes remerciements vont aussi à l'encontre de Me Hélène JOUROT pour avoir pris le temps de m'apprendre les logiciels utilisés au sein de l'entreprise Bürkert notamment Excel et à M. William DOBS grâce à qui j'ai pu me rendre au travail sans problème.

Merci à Stéphanie GUITON pour tous les renseignements concernant l'entreprise Bürkert et à THETIS Vanessa pour les conseils de rédaction du rapport notamment.

Je n'oublie pas les autres personnels du service de qualité M. Cédric, M. François HARRE et M. Richard BENOIT m'ont accueilli dans un environnement propice à mon apprentissage.

Et enfin, un sincère remerciement à toutes les personnes que j'ai pu rencontrer au cours de cette expérience professionnelle très enrichissante.

Sommaire

<u>REMERCIEMENTS</u>	<u>2</u>
<u>SOMMAIRE</u>	<u>3</u>
<u>INTRODUCTION</u>	<u>5</u>
<u>1 LE GROUPE BÜRKERT</u>	<u>7</u>
1.1 UN PEU D’HISTOIRE	7
1.2 INFORMATIONS ET CHIFFRES CLES.....	7
1.3 SITUATION DU GROUPE.....	8
1.4 PRODUITS ET MARCHÉ	10
1.4.1 PRODUITS.....	10
1.4.2 MARCHÉ :	12
<u>2 BÜRKERT TRIEMBACH</u>	<u>13</u>
2.1 UN PEU D’HISTOIRE	13
2.2 BÜRKERT TRIEMBACH AUJOURD’HUI	13
2.3 LE SERVICE DE QUALITE.....	14
<u>3 PRESENTATION STAGE</u>	<u>15</u>
<u>4 TRAVAIL SOUS EXCEL</u>	<u>16</u>
4.1 TEST DE NORMALITE (SHAPIRO ET WILK)	17
4.2 VALEURS ABERRANTES (TEST DE GRUBBS).....	20
4.3 TRANSFORMATION DE BOX-COX.....	25
4.4 TEST DE WALSH.....	29
<u>5 TRAVAIL SOUS R.....</u>	<u>32</u>
<u>6 REDUCTION DU NOMBRE DE COTES</u>	<u>37</u>
<u>7 CONCLUSION</u>	<u>39</u>
<u>8 PROGRAMME DE CALCUL COMPLET</u>	<u>40</u>

9	INDEXE	43
----------	---------------------	-----------

INTRODUCTION

Dans le cadre de ma formation en master de statistique à l'université de Strasbourg, j'ai effectué un stage de deux mois et demi dans la société BÜRKERT à Triembach au val.

C'est une entreprise familiale allemande spécialisée dans les systèmes de contrôles des fluides et des gaz. Elle développe et produit des appareils de régulation comme :

- Les électrovannes
- Les débitmètres
- Les capteurs (capteur de PH, d'oxydoréduction,.....).

Le centre de développement et de production français, qui se situe à Triembach au Val (67), est spécialisé dans la conception et la production des capteurs. Sur le site il existe aussi une unité de vente qui s'occupe de toute la France et de quelques pays d'Afrique.

Au sein du service qualité de l'usine de Triembach, ma mission était divisée en deux parties :

- Dans la première partie, il s'agissait de modifier les plans de contrôle (ajusté la valeur nominale, recalculé les tolérances pour chaque cote), en utilisant des données recueillies sur l'historique des différents contrôles effectués au cours des dernières livraisons mais aussi celles d'échantillons de pièces conformes que j'ai mesuré.
- Dans la deuxième partie, il s'agissait de réduire le nombre de côtes à contrôler sur une pièce tout en gardant la qualité du contrôle.

Mes recherches m'ont conduit à trouver une valeur nominale qui correspond à la valeur moyenne de mes données qui suivent une loi normale auxquelles j'ai enlevé les éventuelles valeurs aberrantes ; dans le cas où les données ne suivent pas une loi normale, j'ai utilisé une transformation de BOX-COX.

Dans un premier temps, je présenterai le groupe BÜRKERT en général, avec les chiffres, la répartition des sites, les différents produits ; ensuite, je parlerai de l'entreprise BÜRKERT de

Triembach au val et enfin je présenterai les tâches que j'ai eu à effectuer au sein du service qualité auquel j'étais rattaché.

1 Le groupe BÜRKERT

1.1 *Un peu d'histoire*



La société BÜRKERT a été créée en Allemagne par Christian Bürkert en 1946. Il sera à la tête de cette dernière jusqu'à ce qu'il décède en 1971 à la suite d'un accident d'avion.

Au départ, le but de Bürkert était de créer et produire de nouveaux appareils pour répondre aux attentes de l'époque.

Les premiers appareils produits étaient des régulateurs de température pour :

- Les couveuses,
- Les chauffe-pieds,
- Les cuisinières.

Par la suite, Bürkert développa des vannes pour finir par en devenir un spécialiste. De nos jours, le groupe est une référence dans le monde industriel au niveau des électrovannes et des systèmes de régulation des fluides.

1.2 *Informations et chiffres clés*

Le groupe BÜRKERT est dirigé par Héribert ROHRBECK. C'est une entreprise familiale dans laquelle la famille Bürkert est toujours partenaire. C'est une société internationale, présente dans le monde entier et dont la langue officielle est l'Anglais.

L'entreprise regroupait, en 2008, 2100 employés, dont 1100 en Allemagne et 150 uniquement pour la recherche et le développement. Ces 2100 personnes sont réparties sur les différents sites du groupe:

- Le siège social à Ingelfingen,
- Les 4 centres de recherche et de développement,
- Les 5 unités de production,
- Les 5 unités de systèmes,
- Les 7 succursales de ventes en Allemagne,
- Les 35 filiales dans le monde.

L'entreprise a un chiffre d'affaire de 264 millions d'euro en 2009 avec 80% pour l'export, 7,4% pour l'investissement et 5,2% pour la recherche et le développement.

Elle propose une gamme de produit diversifiée dans le domaine du contrôle et de la régulation des fluides comme les électrovannes, des débitmètres, des capteurs, etc....

Les différents produits BÜRKERT peuvent se retrouver dans différents domaines d'utilisation :

- Le traitement de l'eau,
- Les industries pharmaceutiques et agroalimentaires,
- Les industries automobiles.

1.3 Situation du groupe

Le siège sociale du groupe se situe à Ingelfingen .Les unités de production et de développement se trouvent en Allemagne et en France. La partie holding internationale du groupe se localise en Suisse.

Les usines allemandes :

- L'usine d'**Ingelfingen** s'occupe de la conception et de l'élaboration des électrovannes. Elle possède aussi une partie système, qui représente le secteur d'adaptation des produits aux attentes des clients.
- L'usine de **Criesbach** est spécialisée dans la fabrication mécanique des différents produits Bürkert.
- L'usine de **Gerabronn** s'occupe de la conception et de la fabrication des vannes pneumatiques. Elle comporte aussi une partie système pour pouvoir adapter les produits aux attentes des clients comme les normes différentes ou les matériaux différents en fonction des pays d'exportation.
- L'usine d'**Ohringen** conçoit et produit les vannes de process et de régulation. Elle possède, elle aussi, une partie systèmes.

L'usine française :

- L'usine de Triembach au Val est spécialisée dans la conception et le montage des capteurs. Elle a aussi une partie capteur pour permettre l'adaptabilité des différents capteurs à leurs différentes conditions d'utilisation.



Bürkert Triembach au Val

En plus de ces différentes usines, le groupe possède 5 sites dédiés à l'activité de systèmes qui permet de développer des produits sur mesure pour leurs différentes utilisations et les différents clients. Ces 5 sites sont appelés les « systemhaus » et sont :

- Systemhaus de Criesbach(Allemagne),
- Systemhaus de Dresden(Allemagne)
- Systemhaus de Dortmund(Allemagne)
- Systemhaus de Charlotte(Etats-Unis),
- Systemhaus de Suzhou(Chine).

1.4 Produits et Marché

1.4.1 Produits

BÜRKERT est un groupe présent au niveau mondial qui vend plusieurs catégories de produits ; on peut citer les capteurs, les électrovannes et les vannes process....

Depuis sa création, elle a développé une gamme de produits variés pour répondre aux différentes attentes de ses clients et aux différents secteurs d'utilisation.

Les produits Bürkert sont donc de natures différentes comme on peut le voir dans la liste ci-dessous.



Electrovanne

Les électrovannes peuvent être utilisées sur tous les types de fluides, qu'ils soient liquides ou gazeux, et dans tous types de conditions

Le MFC et le MFM sont utilisés pour les gaz. Le MFC permet une mesure et une régulation du débit tandis que le MFM n'est utilisé que pour la mesure débit.



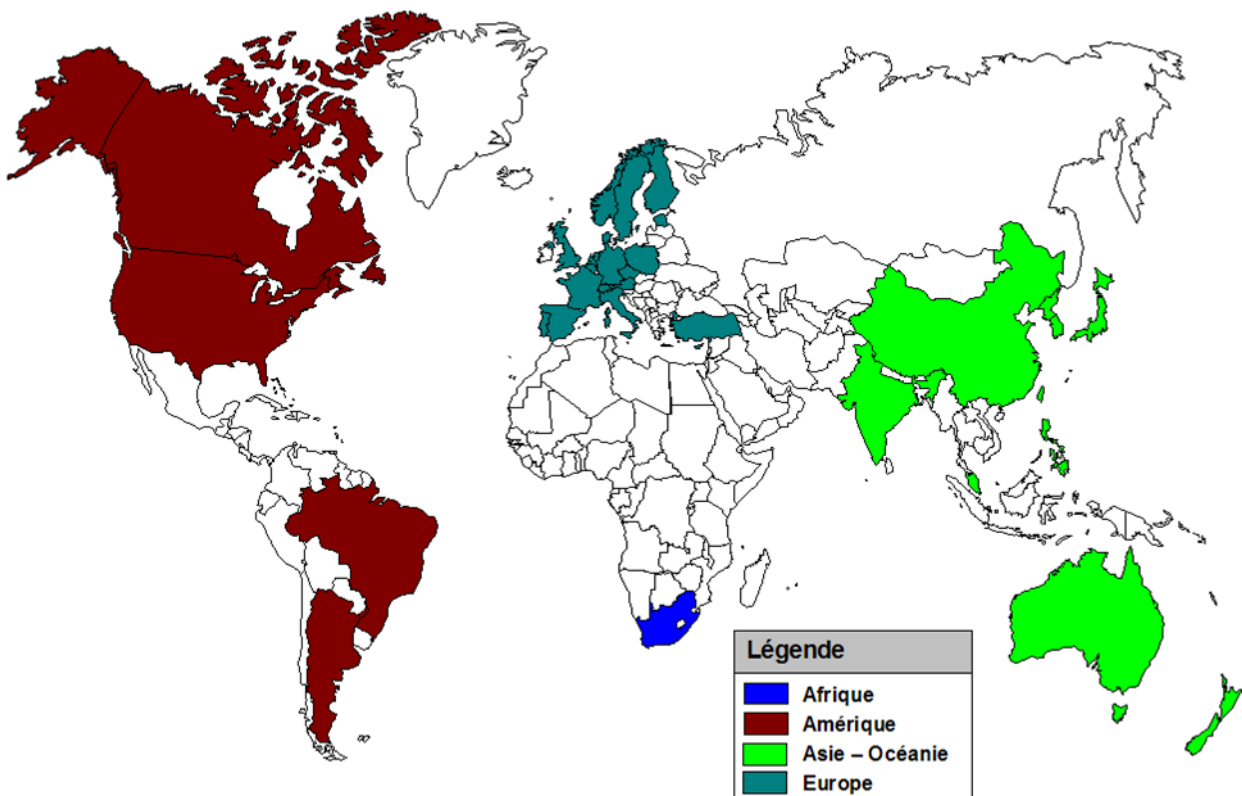
Débitmètre massique

1.4.2 Marché :

Bürkert est une entreprise présente sur le marché mondiale des systèmes de contrôle et de régulation de fluides et de gaz.

Elle possède différentes filiales sur chacun des continents :

- 1 en Afrique
- 4 en Amérique
- 11 en Asie et Océanie
- 18 en Europe



Présence des 34 filiales de vente du groupe à travers le monde

2 BÜRKERT Triembach

2.1 Un peu d'histoire

Bürkert s'est installé en Alsace en 1956 à Strasbourg. L'ouverture de l'usine de Triembach s'est faite en 1966.

A partir de 1968, le site français de Bürkert s'est occupé de la conception et de la production de thermostats.

En 1980 l'usine a changé d'orientation pour s'occuper de la partie électronique du groupe, avec les électrovannes.

Aujourd'hui ce site ne comprend plus aucune partie de production pour les électrovannes. Il est entièrement dédié aux capteurs.

Le site français a aussi vu son statut évolué ; au départ, Bürkert Triembach faisait partie de la section holding international du groupe. Depuis 2001 cette usine est rattachée à la partie opérative de l'entreprise.

De plus, La partie usine a toujours été sur le site de Triembach mais la partie vente était d'abord située à Paris. Aujourd'hui elle rejoint les locaux de Triembach au Val.

2.2 BÜRKERT Triembach aujourd'hui

Le site de Triembach regroupe 2 unités :

- Bürkert Contromatic France : la structure de vente,
- L'usine de Triembach : la structure de production et le centre de développement.

L'usine de Triembach est la partie production du site ; 135 personnes y travaillent pour permettre la conception et la production des capteurs.

Pour le développement de ces capteurs il ya une partie production et une partie conception avec la recherche et le développement.

La recherche et développement regroupe 17 personnes pour le développement des nouveaux produits et 14 pour l'amélioration des anciens et la conception des systèmes spécifiques aux clients.

La partie production permet le montage des capteurs élaborés par la recherche et le développement.

2.3 Le Service de Qualité

Durant mon stage j'ai travaillé au sein du service de qualité qui est chargé de vérifier la conformité des pièces, de délivrer des certificats mais aussi de faire de réclamations auprès des fournisseurs au cas ou les pièces sont certifiées non conformes.

3 Présentation Stage

Pour garantir aux clients des produits irréprochable Bürkert met un accent particulier sur le contrôle de qualité. Sur chaque lot de pièce fournit un échantillon de cinq pièces est prélevé pour être analyser et ainsi on peut déduire à partir des résultats obtenue la conformité ou non du lot étudié. Les différentes mesures effectuées sur chacune des cotes de la pièce sont enregistré dans Casquit. C'est un logiciel qui étudie l'évolution dans le temps des données obtenues à chaque contrôle et qui décide de la conformité ou non des différentes pièces par rapport aux normes préétablies.

La première partie de mon travaille a été de mettre à jour les plans de contrôle ; en effet les différentes valeurs nominales des cotes (longueur, largeur, diamètre.....) n'ont pas mise à jour alors que le processus de fabrication avait évolué ; ce qui fait que le contrôle rejetait systématiquement les pièces qu'elles soient conformes ou non et les courbes représentant cette évolution sortaient des intervalles de tolérances.

La deuxième partie de mon travaille chez Bürkert a été de déterminer les cotes significatives, donc réduire le nombre de cote à mesurer tout en gardant l'efficacité du test de conformité.

Jusque la toutes les cotes des pièces étaient contrôlés ; on a remarqué que certaines cotes n'étaient pas significatives pour déterminer la conformité de la pièce et il y'avait aussi des cotes qui variaient ensemble.

En effet, pour tester une pièce on mesurait quinze cotes ; ce qui représentait une grosse perte de temps dans la mesure où ce n'était pas nécessaire.

4 Travail sous Excel

Dans un premier temps, j'ai appris à me familiariser avec les instruments de mesure utilisés comme le pied à coulisse.

Ensuite, j'ai effectué les mesures des cotes sur des pièces (longueur, largeur, diamètre intérieur, diamètre extérieur,...) sur un échantillon de 50 pièces.

Parallèlement, avec l'aide de mon responsable de stage, j'ai entrepris une formation approfondie du logiciel Excel car l'entreprise utilise ce logiciel et donc je devais réaliser mon travail sous Excel.

Afin d'obtenir des données assez fiables, j'ai préféré utiliser des données stockées dans Casquit car l'habitude de manipuler les instruments est un élément important pour obtenir des mesures correctes. En effet, vu que les fittings sont en plastique, les mesures peuvent varier si on n'appuie pas de la même façon le pied à coulisse lors des mesures. En plus je pouvais exporter les données de Casquit directement sur une page Excel, ce qui me permettait de gagner en temps.

Toutefois, ces données ne pouvaient pas être fiables à 100% car on ne peut pas exclure des erreurs qui peuvent être dues soit à l'instrument de mesure ou une erreur humaine au moment de lire ou recopier les valeurs obtenues.

En plus, les fittings n'étaient pas fabriqués au même moment, donc les données recueillies provenaient de différentes livraisons ; la question qui se posait est de savoir la façon dont je devais associer des données qui n'ont pas été produites en même temps.

Pour répondre à cette question, j'avais deux choix :

- Soit chercher la moyenne des données recueillies sur chaque livraison avant d'étudier ensemble les différentes moyennes obtenues

- Soit rassembler toute les données provenant des différentes livraisons avant de les étudier.

Comme toutes les pièces étaient certifiées conformes et que la moyenne et l'écart type que je cherchais devais me permettre d'englober toutes ces données dans mon intervalle de confiance, j'ai opté pour la deuxième solution qui me permettait de rassembler toutes mes données avant de faire mon étude statistique.

Pour manipuler les données ainsi recueilli, j'avais donc besoin d'étudier leur loi, éliminer des valeurs aberrantes, avant de calculer la valeur nominale ; ce qui m'a poussé à réaliser un certains nombre de tests statistiques sur les données recueilli sous Excel.

4.1 Test de normalité (Shapiro et Wilk)

Il existe de nombreuses méthodes pour vérifier si une série de données présente les caractéristiques d'une loi normale.

Pour étudier la normalité des mes données j'ai effectué des recherche sur un certain nombre de test statistique dont le test d'Anderson, le test de Kolmogorov, le test de Lilliefors et le test de Shapiro et Wilk.

Le test de Lilliefors est une variante du test de Kolmogorov dont la loi est tabulée différemment et les valeurs critiques sont modifiées pour un même risque α .

Ce test est sensible au désaccord de la distribution aux alentours de la partie centrale, la ou les écarts ont peu d'effets sur les tests paramétriques. En revanche, il est moins performant si les désaccords porte sur les queues de distributions pourtant préjudiciables.

Nous observons les mêmes difficultés concernant le test d'Anderson et de Kolmogorov.

La méthode développée par Shapiro et Wilk reste le test le plus puissant dans bien de cas, surtout quand on a que peu ou pas assez de données ($n \geq 5$). Cette méthode implique l'emploi de tables.

Il est basé sur le rapport W de deux estimations liées à la variance de la population dont provient l'échantillon. Ce rapport sera comparé à une valeur théorique $W_{1-\alpha, n}$ et dans le cas où $W \geq W_{1-\alpha, n}$ nous pourrions affirmer avec un risque d'erreur α que la distribution suit une loi normale.

C'est ce dernier que je vais exposer ci-dessous :

Nous rangeons les n observations expérimentales par ordre de valeurs croissantes :

$$y_1 < y_2 < \dots < y_{n-1} < y_n$$

On calcule la moyenne de cette série de mesures :

$$\bar{y} = \frac{1}{n} * \sum_{i=1}^n y_i$$

On calcule le nombre T_n défini par :

$$T_n = \sum_{i=1}^n (y_i - \bar{y})^2$$

On calcule les différences suivantes :

$$d_1 = y_n - y_1$$

$$d_2 = y_{n-1} - y_2$$

...

$$d_n = y_{n+1-i} - y_i$$

Remarque : Si $n=2p$ (n pair), on aura p différences et si $n=2p+1$ (n impair) on aura p différences aussi, l'observation médiane n'intervenant pas.

On calcule le nombre W défini par :

$$W = \left(\sum_{j=1}^{\lfloor n/2 \rfloor} a_j * d_j \right)^2 / T_n$$

Où les coefficients a_j sont donnés par une table.

On choisit un risque de (5%) et on compare la valeur de W à une valeur $W_{1-\alpha,n}$, dite valeur critique, lue dans la table de Shapiro et Wilk.

La règle de test est alors la suivante :

- . Si $W < W_{1-\alpha,n}$ on accepte, au risque choisi, l'hypothèse de normalité de la série de mesure
- . Si $W > W_{1-\alpha,n}$ on rejette l'hypothèse de normalité de la série de mesure.

Exemple :

Dans cet exemple les données utilisées représentent la distance entre deux vices sur une pièce que l'on contrôle. Les mesures sont exprimées en millimètre.

Shapiro-Wilk

	Valeurs	Ti	Di	ai*Di
1	29,96	0,0345	0,3700	0,158767
2	29,97	0,0309	0,3200	0,094976
3	30,01	0,0185	0,2700	0,067473
4	30,05	0,0092	0,2200	0,0473
5	30,06	0,0074	0,2100	0,039144
6	30,08	0,0043	0,1300	0,021008
7	30,08	0,0043	0,1300	0,018135
8	30,09	0,0031	0,1100	0,013112
9	30,1	0,0021	0,0900	0,009018
10	30,1	0,0021	0,0900	0,007398
11	30,12	0,0007	0,0600	0,0039
12	30,12	0,0007	0,0600	0,002898
13	30,12	0,0007	0,0400	0,00128
14	30,13	0,0003	0,0200	0,000318
15	30,13	0,0003		
16	30,15	0,0000		
17	30,16	0,0002		
18	30,18	0,0012		
19	30,18	0,0012		
20	30,19	0,0019		
21	30,19	0,0019		
22	30,2	0,0029		
23	30,21	0,0041		
24	30,21	0,0041		
25	30,27	0,0154		
26	30,27	0,0154		
27	30,28	0,0180		
28	30,29	0,0208		
29	30,33	0,0339		
Moyenne	30,145862			
Ecartype	0,09260196			
Somme		0,24		0,48473
W	0,9785793			
Normalité1	OUI			

4.2 Valeurs aberrantes (test de Grubbs)

Lors de la phase de mesure ou de récupération de données, des erreurs peuvent se glisser. Ces erreurs peuvent être dues à un problème d'outil de mesure, ou une erreur de lecture.

Ces valeurs sont appelées des valeurs aberrantes ; pour ne pas perturber l'adéquation de notre modèle, il faut éliminer ces valeurs.

Pour cela on dispose d'un certain nombre de tests statistiques plus ou moins puissants qui permettent d'éliminer les valeurs aberrantes.

Ce sont entre autres le test de Dixon et le test de Grubbs. Tous les deux présupposent la normalité de notre série de valeurs.

Le test de Dixon consiste à comparer la distance entre les points les plus éloignés du modèle et les points immédiatement plus voisins à l'étendue totale des résidus. C'est le plus utilisé des deux tests, mais il est beaucoup moins efficace que lorsque l'on a peu de valeurs. C'est la raison qui m'a poussé à choisir le test de Grubbs qui est plus puissant quand on a peu de données. Ce test se présente comme suit :

A partir des données y_i mesurées et classées par ordre croissant, la statistique de Grubbs G est calculée pour chaque valeur extrême y_{\min} et y_{\max} .

$$G_{\max} = (y_{\max} - \bar{y}) / s$$

$$G_{\min} = (y_{\min} - \bar{y}) / s$$

Avec s : écart-type de la série

\bar{y} : moyenne de la série

La valeur de G est ensuite comparée à la valeur tabulée en se référant à Grubbs et Beck :

- Si $G \leq$ valeur critique à 5% tabulée, la valeur est acceptée comme correcte par le paramètre considéré ;
- Si $G >$ valeur critique à 5% tabulée, et si $G <$ valeur critique à 1% tabulée, le candidat est considéré comme douteux pour une fois ;
- Si $G >$ valeur critique à 1%, le candidat est considéré comme aberrant pour le paramètre considéré.

Le test de Grubbs est réalisé de manière itérative, alternativement à l'extrémité haute et à l'extrémité basse de la série jusqu'à ce qu'aucun point aberrant ou douteux ne soit plus détecté. A chaque itération l'effectif est réduit d'une mesure : les numérateurs et dénominateurs de G ainsi que les valeurs critiques à 5% et 1%.

Exemple

Nous avons testé la normalité des données plus haut et comme elles suivent une loi normale on peut tester les valeurs extrêmes pour vérifier si ce sont des valeurs aberrantes ou pas en utilisant le test de Grubbs.

	Valeurs		
1	29,96	Nbre valeur	29,0
2	29,97	G(Alpha 1%)	3,1
3	30,01	G(Alpha 5%)	2,7
4	30,05	Gmax	2,0
5	30,06	Gmin	2,0
6	30,08	Grubbs max	pas aberrante
7	30,08	Grubbs min	pas aberrante
8	30,09		
9	30,1		
10	30,1		
11	30,12		
12	30,12		
13	30,12		
14	30,13		
15	30,13		
16	30,15		
17	30,16		
18	30,18		
19	30,18		
20	30,19		
21	30,19		
22	30,2		
23	30,21		
24	30,21		
25	30,27		
26	30,27		
27	30,28		
28	30,29		
29	30,33		
Moyenne	30,145862		
Ecartype	0,09260196		
Somme			

On constate que pour cet exemple, nous n'avons pas de valeurs aberrantes car G_{\max} et G_{\min} sont tous les deux plus petite que la valeur lue sur la table de Grubbs à 5%.

Mais dans certains cas, le test de normalité de Shapiro et Wilk est négative et on ne peut pas directement effectuer le test de Grubbs pour éliminer les valeurs aberrantes.

C'est le cas des données recueillies pour la longueur et le diamètre intérieur par exemple.

Pour remédier à ce problème, nous utilisons des tests d'ajustement pour se rapprocher de la loi normale afin de pouvoir exploiter nos données.

Plusieurs méthodes sont utilisées dont entre autres la droite de Henry.

Dans les cas où nous n'avons pas suffisamment de données cette méthode permet de vérifier si la loi suivie par nos valeurs se rapproche d'une loi normale ou pas. Si la distribution n'est pas gaussienne, on peut observer sur le graphique les points aberrants qui nous éloignent de la loi normale.

On utilise une itération pour supprimer les points les plus éloignés jusqu'à ce que l'on soit assez proche d'une droite et donc d'une distribution normale.

L'inconvénient de cette méthode est qu'elle est basée sur une observation du résultat car dans ce cas on ne peut pas l'automatiser ; ce qui représente un handicap dans le domaine industriel où la gestion du temps et la diversité des données à tester est un facteur important.

Un autre problème de cette méthode est de savoir à quel moment nous sommes assez proches d'une droite pour arrêter l'itération. Même si ce problème peut être résolu en calculant le coefficient de corrélation linéaire, il apparaît que cette méthode reste manuelle et par conséquent pas très adaptée à la situation qui nous intéresse à savoir mettre en place une méthode qui permet de corriger les plans de contrôle sans se soucier des étapes intermédiaires.

Pour contourner cette difficulté on va procéder à une transformation de la variable X en une nouvelle variable U qui elle suit une loi normale. Cette transformation est appelée transformation de BOX-COX.

4.3 Transformation de BOX-COX

De nombreuses procédures statistiques reposent sur la normalité des distributions. Transformer les variables de manière à se rapprocher de la distribution normale, ou tout au moins les symétriser est souvent un préalable nécessaire avant toute analyse statistique.

La transformation de BOX-COX consiste à effectuer un changement de variable pour se rapprocher d'une distribution normale. Pour cela on pose :

$$U(\lambda) = \frac{X^\lambda - 1}{\lambda} \text{ si } \lambda \neq 0$$

$$U(\lambda) = \ln(X) \text{ si } \lambda = 0$$

.

Cette transformation est intéressante car elle incorpore à la fois la possibilité d'avoir aucune transformation ($\lambda = 1$) et la possibilité d'une transformation logarithmique ($\lambda = 0$).

En effet, il est toujours possible de trouver une valeur de λ telle que l'ensemble des valeurs transformées U_i s'ajuste à une loi de distribution normale. Il s'agit donc d'essayer, de façon itérative, différentes valeurs de λ et tester pour chacune d'elles la normalité de la distribution des valeurs transformées jusqu'à ce que le test de Shapiro et Wilk soit positif.

Les différents tests effectués montrent qu'il peut y avoir plusieurs valeurs de λ qui se rapprochent d'une distribution normale; mais on constate aussi que plus on augmente la valeur de λ , plus les valeurs aberrantes s'ajustent et par conséquent deviennent plus proches d'une normale. Donc on choisira la valeur de λ la plus grande qui rendra le test de normalité

positif. En procédant ainsi, nous pouvons appliquer le test de Grubbs pour éliminer les valeurs aberrantes sur les valeurs ajustées avant de corriger les plans de contrôle.

Exemple

Les données utilisées dans cet exemple proviennent toute des mesures effectuées sur les différentes pièces que j'ai eu à étudier chez Bürkert.

Shapiro-Wilk

Shap

	Valeurs	Ti	Di	ai*Di	valeur ajustées	Ti ajustées	Di aj
1	26,61	0,1237	0,6	0,253169	4,9573E+23	1,67311E+46	
2	26,82	0,0201	0,3	0,0742	5,6884E+23	3,16301E+45	
3	26,83	0,0174	0,2	0,057477	5,7256E+23	2,75809E+45	
4	26,85	0,0125	0,2	0,04085	5,8008E+23	2,0252E+45	
5	26,86	0,0103	0,2	0,031688	5,8387E+23	1,69825E+45	
6	26,9	0,0038	0,1	0,021008	5,9928E+23	6,65907E+44	
7	26,91	0,0027	0,1	0,015345	6,0319E+23	4,79371E+44	
8	26,93	0,0010	0,1	0,010728	6,1108E+23	1,96029E+44	
9	26,94	0,0005	0,1	0,008016	6,1506E+23	1,00357E+44	
10	26,95	0,0001	0,1	0,004932	6,1907E+23	3,61223E+43	
11	26,95	0,0001	0,1	0,0039	6,1907E+23	3,61223E+43	
12	26,95	0,0001	0,1	0,002415	6,1907E+23	3,61223E+43	
13	26,95	0,0001	0,1	0,0016	6,1907E+23	3,61223E+43	
14	26,95	0,0001	0,0	0,000636	6,1907E+23	3,61223E+43	
15	26,99	0,0008			6,3535E+23	1,05432E+44	
16	26,99	0,0008			6,3535E+23	1,05432E+44	
17	27	0,0015			6,3948E+23	2,07364E+44	
18	27	0,0015			6,3948E+23	2,07364E+44	
19	27,01	0,0023			6,4364E+23	3,44385E+44	
20	27,01	0,0023			6,4364E+23	3,44385E+44	
21	27,02	0,0034			6,4782E+23	5,17133E+44	
22	27,02	0,0034			6,4782E+23	5,17133E+44	
23	27,02	0,0034			6,4782E+23	5,17133E+44	
24	27,03	0,0047			6,5203E+23	7,26256E+44	
25	27,03	0,0047			6,5203E+23	7,26256E+44	
26	27,04	0,0061			6,5626E+23	9,7241E+44	
27	27,06	0,0097			6,6481E+23	1,57848E+45	
28	27,07	0,0117			6,6912E+23	1,93975E+45	
29	27,2	0,0568			7,2764E+23	1,05187E+46	
Moyenne	26,961724				6,2508E+23		
Ecartype	0,104473816				4,1112E+22		
Somme		0,31		0,52596		4,73256E+46	
W	0,905188625				Nbre valeur		29 Gma
Normalité1	NON				G(Alpha 1%)		3,1 Gmi
W ajustée	0,926010368				G(Alpha 5%)		2,7
Normalité2	OUI				Grubbs max	pas aberrante	
L	17,5				Grubbs min	aberrant	

Dans cet exemple nous constatons que le test normalité de Shapiro et Wilk est négatif ; donc nous effectuons un ajustement des valeurs en appliquant la transformation de BOX-COX. Nous faisons varier la valeur de L et nous constatons que pour $\lambda = 17$, le test est toujours négatif et que pour $\lambda = 18$ le test est positif. Donc la plus petite valeur de L se trouve entre ces deux valeurs de L. En faisant une dichotomie, nous trouvons $\lambda = 17.5$.

Une fois que nous avons trouvé la transformation de BOX-COX qui a la valeur de L la plus grande, nous pouvons effectuer le test de Grubbs pour éliminer les valeurs aberrantes.

Le test de Grubbs nous dit que la valeur max n'est pas aberrante contrairement à la valeur min. Si nous avons choisis par exemple la transformation obtenue pour $\lambda = 40$, le test de normalité serait toujours positif mais le test de Grubbs n'aurait pas détecté de valeurs aberrantes mais plutôt de valeurs douteuses.

Les différents tests montrent que l'autre difficulté de la transformation de BOX-COX est la faiblesse d'Excel pour faire des calculs avec de très grosse puissance.

En effet, pour certaines données comme celles recueillies pour la longueur des pièces, Excel ne peut nous donner la valeur de L appropriée car le calcul est trop grand.

En observant, ces données nous constatons que l'écart type mesuré est trop grand et que ce ne sont pas que les valeurs extrêmes qui empêchent la distribution d'être normale.

Ce cas constitue la limite de notre étude et pour palier à cette difficulté on va être obligé d'augmenter le nombre de données. Une autre solution aurait pu être de trouver une méthode qui permet d'éliminer les valeurs aberrantes et qui ne présuppose pas la normalité des données. Mais les différentes méthodes qui proposent cette alternative exigent au moins que l'on ait 60 données. C'est le cas du test de Walsh et du test de Rosner par exemple.

En augmentant le nombre de données, nous nous rapprochons de la loi normale car le théorème central limite place la loi normale à la limite de la plupart des distributions.

Si nous ne retrouvons pas une loi normale, nous pouvons utiliser le test de Walsh pour éliminer les valeurs aberrantes et retenter le test de la normalité.

Le test de Walsh permet de détecter la présence d'un nombre quelconque d'individus atypiques. Il ne nécessite aucune condition sur les données (normalité). Il est basé sur la détection d'un grand écart entre deux valeurs consécutives. Il se présente comme suit :

4.4 Test de Walsh

Le test de Walsh est un test non paramétrique ; il ne tient pas compte de la normalité ou de la distribution mais par contre impose d'avoir un minimum de 60 données. Il se présente comme suit :

Soient n observations triées dans l'ordre croissant :

$$y_1 < y_2 < \dots < y_{n-1} < y_n$$

Et $1 < r < n$, le test de Walsh teste l'hypothèse :

H_0 : les r observations $y_{n-r+1}, y_{n-r+2}, \dots, y_n$ sont issus de la même loi que les autres données, contre l'hypothèse :

H_1 : les r observations $y_{n-r+1}, y_{n-r+2}, \dots, y_n$ ne sont pas issus de la même loi que les autres données.

Pour $60 < n < 220$, le risque de rejeter à tort H_0 (le risque de première espèce) est $\alpha = 10\%$

Pour $220 < n < 220$, le risque de rejeter à tort H_0 (le risque de première espèce) est $\alpha = 5\%$. On calcule alors les données suivantes (avec $[y]$ la partie entière de y) :

$$C = 1 + \lceil \sqrt{2n} \rceil$$

$$b^2 = 1/\alpha,$$

$$a = \frac{1 + b \frac{c - b^2}{c - 1}}{c - b^2 - 1}$$

Les r observations sont alors déclarées atypiques (au risque α de les déclarer atypiques à tort) si l'inégalité suivante est vérifiée :

$$y_{n-r+1} - y_{n-r} > a(y_{n-r} - y_{n-r+1-C})$$

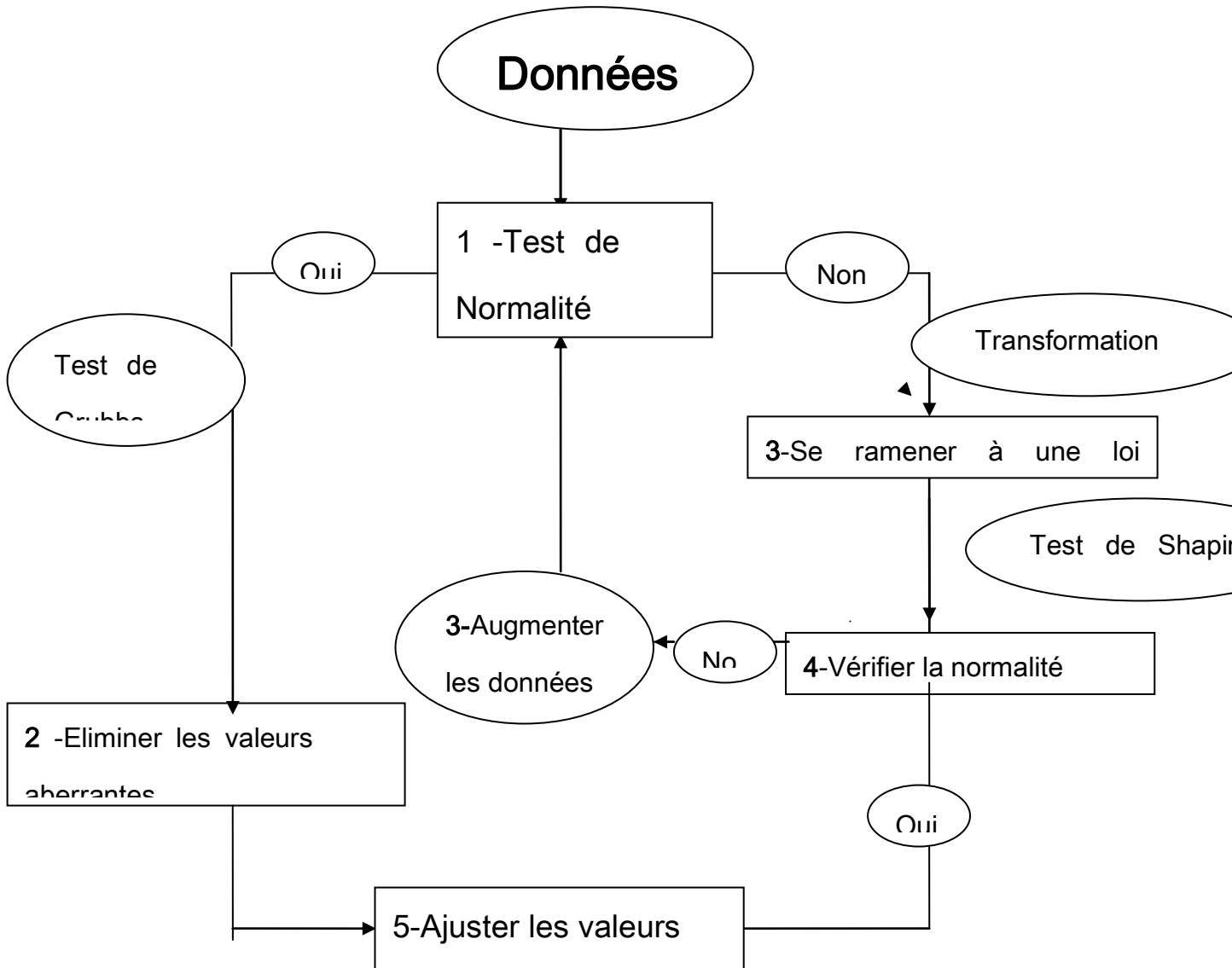
Après avoir étudié individuellement les différents tests, nous allons mettre en place un programme de calcul qui permettra de les associer et qui tient compte des différentes situations que l'on peut rencontrer en fonction de la normalité ou non des données que nous manipulons. Nous rencontrons les différentes situations suivantes :

Si les données suivent une loi normale, nous pouvons directement éliminer les valeurs aberrantes en utilisant le test de Grubbs avant de calculer la moyenne et l'écart type de notre échantillon de mesure.

Si les données ne suivent pas une loi normale, nous devons transformer les données en utilisant les formules de Box Cox ; cette transformation nous permet de nous rapprocher d'une loi normale et par conséquent on se retrouve dans la situation précédente. Nous utiliserons la transformation inverse de Box Cox pour revenir aux valeurs initiales.

Si la transformation de Box Cox ne donne pas un résultat satisfaisant nous pouvons effectuer le test de valeurs aberrantes de Walsh à condition d'augmenter le nombre de données ; il est possible d'appliquer le test de Walsh à partir de soixante données.

Plan de Calcul adopté



Les différentes étapes montrent les enchainements de calculs que nous devons réaliser sur les données pour en tirer les meilleurs enseignements possibles.

Après avoir étudié les différents tests et mis en place toutes les démarches à suivre pour obtenir notre nouvelle valeur nominale, il nous reste à écrire un programme qui permet d'automatiser les différents calculs. Pour cela, en tenant compte de la complexité du calcul, mon maître de stage m'a proposé d'écrire des macros sous Excel. Cela aurait dû être la suite de mon travail ; mais nous nous sommes vite rendu compte que le logiciel Excel était loin d'être le plus adapté pour mettre en place un tel programme. En effet, pour certaines données, Excel ne pouvait pas donner une valeur de λ qui permettait de faire la transformation de Box Cox car les calculs contenaient de très grande puissance. En plus de cette faiblesse de calcul du logiciel Excel, les différents tests que nous effectuons n'étaient pas prévu ; donc la seule solution était d'écrire les différentes étapes de calcul ; ce qui représentait une perte de temps d'autant plus qu'il existe des logiciels statistiques plus puissants et plus adaptés. Nous avons décidé donc, de poursuivre le travail en utilisant le logiciel R qui, non seulement est gratuit mais aussi contient la plupart des tests statistiques décrits plus haut.

5 Travail sous R

Sous R la première partie du travail consistait à effectuer des recherches pour les différentes commandes qui permettent de faire les tests que je dois utiliser.

Ce sont le test de normalité de Shapiro et Wilk, le test des valeurs aberrantes de Grubbs, la transformation de Box Cox.

La deuxième partie consistait à écrire un programme qui permet d'associer les différentes fonctions citées plus haut.

Pour appeler ces différentes fonctions, un certain nombre de bibliothèques ont été nécessaires ; ce sont :

Library (car): qui contient la fonction *bcPower*. C'est cette fonction qui effectue la transformation de Box Cox ; elle prend en entrée un vecteur et la valeur lambda de la transformation.

Library(FitAR): qui contient la fonction *bxcx* ; c'est la fonction inverse de *bcPower* ; elle permet de revenir sur les valeurs initiales.

Library (outliers): contient le test de Grubbs. Elle prend en entrée un vecteur dont les composantes suivent une loi normale et retourne une p-valeur qui permet de conclure à l'existence ou non de valeurs aberrantes.

Library(MASS): contient la fonction box cox de la transformation de Box Cox.

Nous allons utiliser des boucles pour pouvoir détecter et supprimer toutes les valeurs aberrantes.

La première partie de notre programme, comme indiquée sur le schéma de calcul que nous avons mis en place (fig 6), correspond au test de normalité de Shapiro et Wilk qu'on obtient sous R avec la fonction :

➤ *shapiro.test (vect_donnees)*

vect_donnees est obtenue en transformant les données d'entrée en vecteur et il faut noter qu'avant tout calcul, nous devons ordonner les données recueillies

```
> if (shapiro.test (vect_donnees)$p.value>=0.05) {
.   while(grubbs.test(vect_donnees,type=10)$p.value<0.05) {
```

```

vect_donnees = vect_donnees [-c(1,1)]
nbmin =nbmin+1
}
while (grubbs.test (vect_donnees, opposite=FALSE, type=10)$p.value<0.05) {
. vect_donnees=vect_donnees [-c (length (vect_donnees), length (vect_donnees))]
nbmax =nbmax+1
}
}

```

Cette partie du programme correspond à la deuxième partie du schéma de calcul c'est-à-dire le cas où le test de Shapiro est positif.

Dans ce cas nous appliquons le test de Grubbs . La première boucle *while* permet d'appliquer le test de Grubbs sur les petites valeurs de la série et la deuxième sur les

grandes. Nous utilisons ici une boucle pour éliminer toutes les valeurs aberrantes avant de manipuler les différentes données recueillies.

```

else {
lambda=coef(powerTransform(vect_donnees))
box_donnees=bcPower(vect_donnees,lambda)
bxcx(mean(box_donnees),lambda,InverseQ=TRUE) }
}

```

De façon analogue dans cette partie nous supposons que le test de Shapiro est négatif et donc nous utilisons les valeurs obtenues dans la transformation Box Cox. C'est la troisième

partie de notre plan de calcul. La transformation de Box Cox enlève l'influence des valeurs aberrantes et donc nous n'avons plus besoin d'effectuer le test de Grubbs.

En plus dans le cas où on effectue le test de Grubbs, pour ne pas perdre trop de données nous comptons le nombre de valeurs aberrantes, ceci pour éviter de se retrouver avec trop peu de valeurs par rapport aux données de départ.

```
Moyenne = mean (if (shapiro.test (vect_donnees) $p.value >= 0.05) {
    vect_donnees
}
else{
    bxcx (box_donnees, lambda, InverseQ = TRUE)
}
})
```

Pour le calcul de la moyenne, si le test de Shapiro est positif, nous calculons la moyenne des valeurs initiales après avoir supprimé les valeurs aberrantes.

Dans le cas contraire nous calculons la moyenne des valeurs obtenues grâce à la transformation de Box Cox sans les valeurs aberrantes.

Cette nouvelle moyenne constituera notre nouvelle valeur nominale.

En effet, les données que nous utilisons ont été mesurées sur des pièces réputées conformes, en plus nous nous sommes assurés qu'elles n'ont plus de valeurs aberrantes.

Donc la moyenne est un bon estimateur de la valeur nominale dans notre cas.

```
> Ecartype=( if(shapiro.test(vect_donnees)$p.value >= 0.05){
```

```

sd(vect_donnees)
}else{
  plot(0, 0, xlim = c(-3.8, 3.8), ylim = c(0, 100), type = "n", las = 1,
xaxt = "n", xlab = "Probabilites en %", ylab = "", main = "Papier gaussio-arithmetique")
abline(v = qnorm(vect_donnees_copy/100), col = grey(0.8))
abline(h = seq(0, 100, by = 10), col = grey(0.8))
axis(side = 1, at = qnorm(vect_donnees_copy/100), label = vect_donnees_copy, cex.axis =
0.75)
qqnorm(vect_donnees_copy)
qqline(vect_donnees_copy)
}
)

```

Contrairement à la moyenne, nous ne pouvons pas utiliser la transformation de Box Cox pour calculer l'écart type car elle ne conserve pas les écarts de mesure.

Nous utilisons la droite de Henry pour éliminer les valeurs qui empêchent nos données d'être normales et avec le reste nous calculons l'écart type et le programme précédent permet de tracer la droite de Henry dans R.

Le calcul de l'écart type nous permet d'estimer la tolérance acceptée pour chaque cote mesurée.

Dans le domaine industriel, on utilise une tolérance estimée à six fois l'écart type.



6 Réduction du nombre de cotes

Dans la deuxième partie du travail, je devais trouver une méthode qui permettait de réduire le nombre de cote à vérifier et trouve les cotes significatives

Cette partie de mon travail nécessitait beaucoup d'implication dans les processus de fabrication, d'emballage et de stockage des pièces car il fallait repérer les cotes qui posent de problème en général ; pour cela il fallait demander non seulement l'avis de ceux qui contrôlent mais aussi de ceux qui assemblent.

Pour bien comprendre la méthode de fabrication des pièces, j'ai effectué un visite en production au cours de laquelle ils m'ont expliqué le mécanisme de fabrication et j'ai pu assister à l'assemblage des pièces.

Après cette démarche et pour bien étudier le comportement des cotes sur la durée j'ai représenté leur évolution en fonction du temps et comparé leur variation. J'ai pu en tirer des conclusions qui sont certes intéressante mais qu'il a fallu combiner non seulement avec les

différents renseignements obtenus auprès des techniciens de la production et de l'emballage mais aussi avec les renseignements tirés des plans de fabrications des pièces.

En effet, je ne pouvais pas me contenter juste des données que j'ai recueillies sans tenir compte du côté pratique car les pièces sont fabriquées individuellement pour être raccordées à des installations. En plus de cela, les erreurs qu'il pouvait y avoir dans les données ne permettent pas de tirer des conclusions suffisantes qui permettent de réduire les cotes de contrôle.

Les résultats obtenus m'ont permis de réduire les côtes à contrôler de quinze à cinq qui sont : la longueur, le diamètre intérieur, le diamètre extérieur, la longueur et la largeur séparant deux vices à mesurées.



7 Conclusion

Ce stage au sein de l'entreprise Burkert Triembach à été pour moi une expérience très enrichissante. J'ai eu l'opportunité de découvrir le monde de l'entreprise, le fonctionnement des différents instruments de travail, ainsi que des techniques de fabrication complexes et efficaces.

Ce stage m'a permis non seulement d'appliquer certains méthode Statistiques vus en cours comme le test de normalité de Shapiro et Wilk mais aussi d'effectuer des recherches sur des tests que je ne connaissais pas notamment le test Grubbs ou encore la transformation de Box Cox. Le fait d'avoir travaillé sous Excel au départ m'a permis d'aller plus loin que la simple application des tests car je devais comprendre comment ils fonctionnent avant de procéder aux différents calculs.

Ce stage m'a permis aussi de constater le décalage qui existe entre les connaissances théoriques et la réalité pratique au sein de l'entreprise.

En effet, j'ai appris à prendre assez de recul par rapport aux données et avoir un esprit critique face à des situations qui sont en apparence simple.

Le cadre de travail m'a beaucoup plu, les collègues sont accueillants et souriants, ce qui m'a permis de m'intégrer rapidement au sein du groupe.

8 Programme de calcul complet

```
# Appel des librairies nécessaire au fonctionnement du programme
```

```
# CAR contien la fonction bcPower
```

```
# FitAR contient la fonction bxcx
```

```
# outliers contient le test de Grubbs
```

```
# grid contient des graphique
```

```
# MASScontient la fonction box cox
```

```
library(car)
```

```
library(FitAR)
```

```
library(outliers)
```

```
library(grid)
```

```
library(MASS)
```

```
# Demande la nom du fichier à traiter et importe les données dans "donnees"
```

```
fichier=winDialogString("Indiquer le nom du fichier txt à traiter (à placer dans Mes Documents  
:", "test4.txt")
```

```
donnees=read.table(fichier,dec=",")
```

```
# Mise en forme des données et tri par ordre croissant
```

```
vect_donnees1=c(donnees$V1)
```

```
vect_donnees=sort(vect_donnees1)
```

```
vect_donnees_copy=vect_donnees
```

```
# Test de Shapiro et Wilk
```

```
shapiro.test(vect_donnees)
```

Initialisation des variables

nbmax=0

nbmin=0

Si le test de shapiro est positif élimine les valeurs aberrante

if (shapiro.test(vect_donnees)\$p.value>=0.05) {

Elimination des valeurs aberrantes inférieures

while(grubbs.test(vect_donnees,type=10)\$p.value<0.05) {

vect_donnees=vect_donnees[-c(1,1)]

nbmin=nbmin+1

}

Elimination des valeurs aberrantes supérieures

while(grubbs.test(vect_donnees,opposite=FALSE,type=10)\$p.value<0.05) {

vect_donnees=vect_donnees[-c(length(vect_donnees),length(vect_donnees))]

nbmax=nbmax+1

}

Si shapiro est negatif élimine les valeurs aberrantes apres la transformation de Box Cox

} else {

Transformation de BoxCox - Calcul la valeur de lambda et création des valeurs corrigés

"box_donnees "

lambda=coef(powerTransform(vect_donnees))

box_donnees=bcPower(vect_donnees,lambda)

bxcx(mean(box_donnees),lambda,InverseQ=TRUE)

}

Moyenne=(if(shapiro.test(vect_donnees)\$p.value>=0.05){

```

    mean(vect_donnees)
  }else{ bxcx(mean(box_donnees),lambda,InverseQ=TRUE)
  }
)
Ecartype=( if(shapiro.test(vect_donnees)$p.value>=0.05){
  sd(vect_donnees)
}else{
  plot(0, 0, xlim = c(-3.8, 3.8), ylim = c(0, 100), type = "n", las = 1,
xaxt = "n", xlab = "Probabilites en %", ylab = "", main = "Papier gauss-arithmetique")
  abline(v = qnorm(vect_donnees_copy/100), col = grey(0.8))
  abline(h = seq(0, 100, by = 10), col = grey(0.8))
  axis(side = 1, at = qnorm(vect_donnees_copy/100), label = vect_donnees_copy, cex.axis =
0.75)
  qqnorm(vect_donnees_copy)
  qqline(vect_donnees_copy)
  }
)

# Affichage des données résultantes des calculs ci-dessus

print(paste("Le nombre de valeurs aberrantes supprimées est ", nbmin))
if (nbmin!=0)
  print(paste("Les valeurs suivantes ont été considérées comme aberrantes : ",
vect_donnees_copy[1:nbmin]))
print(paste("La nouvelle valeur nominale est ", Moyenne))
print(paste("L'Ecartype est de:", Ecartype))

```

9 Indexe

http://eric.univ-lyon2.fr/~ricco/cours/cours/Test_Normalite.pdf

http://zoonek2.free.fr/UNIX/48_R_2004/13.html#3

http://zoonek2.free.fr/UNIX/48_R_2004/08.html#1

<http://pbil.univ-lyon1.fr/R/pdf/for1.pdf>

<http://www.bib.fsagx.ac.be/base/text/v9n1/19.pdf>

<http://frtripxy01.proxy.burkert.com:8081/3070229392/2f81a135/complete/ete.inrs.ca/pub/rapports/R000360.pdf?file=R000360.pdf>

http://archive.numdam.org/ARCHIVE/RSA/RSA_1976__24_3/RSA_1976__24_3_61_0/RSA_1976__24_3_61_0.pdf

<http://russell.vcharite.univ-mrs.fr/EIE/fchap14.pdf>

<http://pbil.univ-lyon1.fr/R/fichestd/tdr22.pdf>