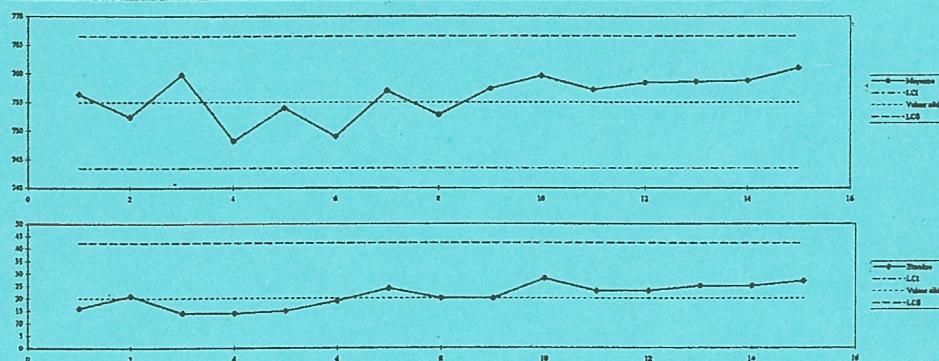


UNE APPLICATION INDUSTRIELLE

DES STATISTIQUES :

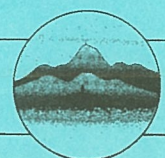
LA CARTE DE CONTRÔLE.

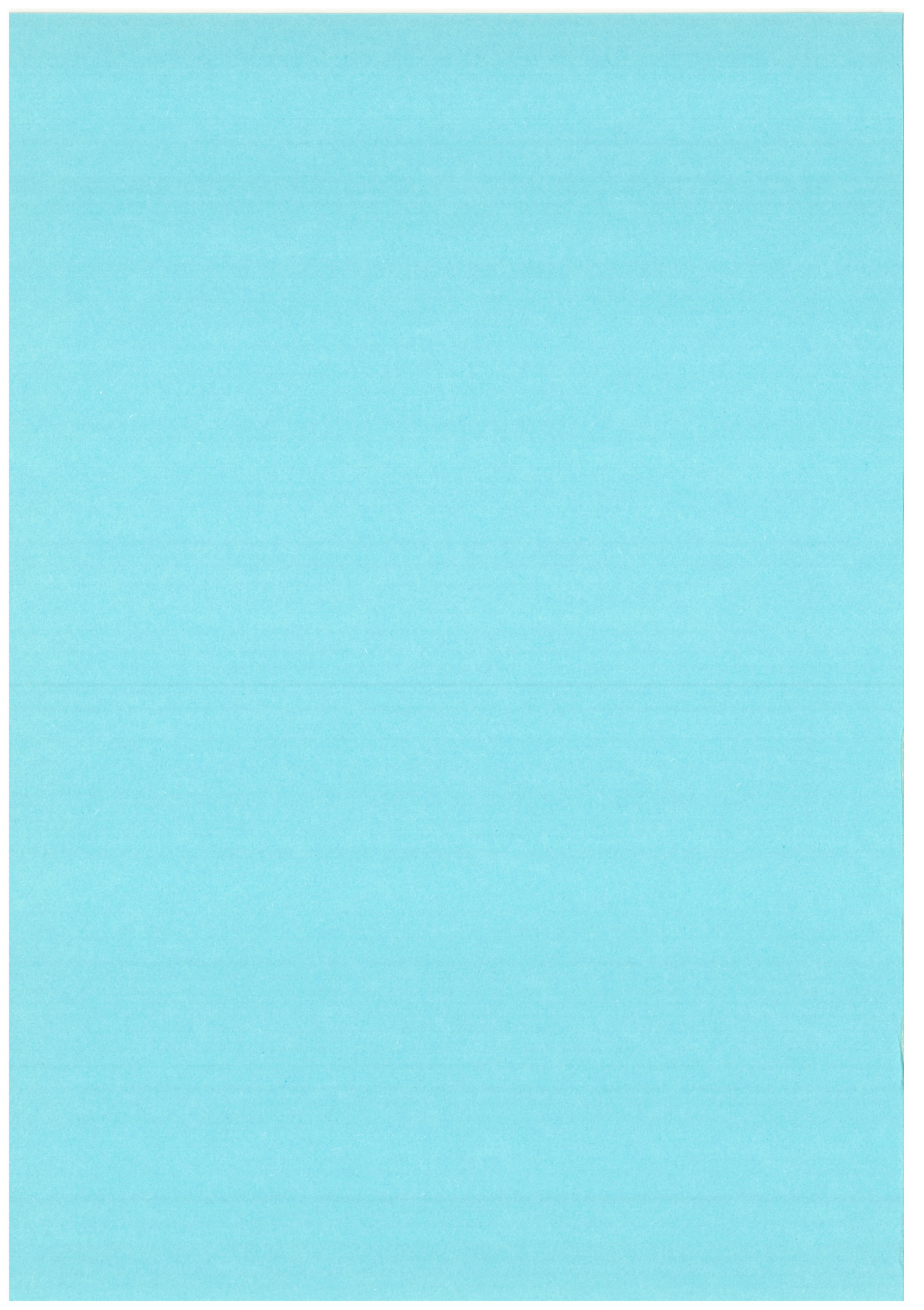
N° lot	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
x1	758	735	765	735	745	740	742	761	787	742	782	741	765	742	776
x2	757	751	735	751	738	744	756	761	784	762	780	762	765	763	749
x3	758	757	745	749	757	750	763	751	760	768	763	763	740	760	762
x4	751	768	751	745	758	752	758	750	748	759	761	764	761	762	762
x5	766	739	765	741	758	759	766	741	747	764	740	782	762	767	756
Moyenne	756,4	752,4	759,8	748,2	754	749	757	752,8	757,4	759,6	757,2	758,4	758,6	758,8	761
écart	16	21	14	14	15	19	24	20	20	28	23	23	25	15	27



- Mars 2001 -

D. Arbore, A. Corpart,
G. Fleury, N. Lassalle.





Sommaire

Introduction	p. 2
1 Quelques définitions	p. 3
1.1 Processus	
1.2 Maîtriser le processus	
1.3 La carte de contrôle	
2 Principe de la méthode	p. 3
3 La carte de contrôle	p. 4
3.1 Objectifs	
3.2 Quelques exemples	
3.3 Définition	
3.4 Exemple	
3.5 Différents types de carte de contrôle	
4 Mise en oeuvre	p. 7
4.1 Organigramme de la mise en place d'une carte de contrôle	
4.2 La carte moyenne-étendue	
5 Etude de normalité	p.10
5.1 Droite de Henry	
5.2 Test de Kolgomorov-Smirnov	
5.3 Test du χ^2 (Khi - Deux)	
6 Exemple	p.18
6.1 Données	
6.2 Construction d'une carte de moyennes et de l'étendue	
6.3 Utilisation de la carte	
Annexes	
Annexe A Les tables	p.24
Tableau des coefficients	
Table du Chideux	
Table de la loi de Kolgomorov-Smirnov	
Annexe B Utilisation de l'étendue	p.30
Estimateurs ponctuels de l'écart-type	
Intervalle de confiance sur l'écart-type	
Un exemple d'utilisation, la construction d'une carte de contrôle	
Annexe C Activité tableur	p.37
Manipulations de base	
Fichier Normalité	
Fichier Carte	

Introduction

En visite dans un hôpital, vous avez peut-être remarqué, en tête de lit, un graphique donnant les indications suivantes :

- la température (T°)
- la tension artérielle (TA) ...

Le "journal de bord" associé au graphique indique les soins et traitements apportés au malade, précise les analyses qui lui ont été prescrites. Vous aviez devant les yeux une "carte de contrôle" permettant à l'équipe médicale de suivre l'évolution de l'état de santé du patient. Cette carte permet aussi d'avoir un historique des soins apportés et de diagnostiquer d'éventuelles "améliorations-dégradations" de l'état de santé du malade (augmentation de la température...).

Informations patient

ANNÉE : 1977

NOM : PELLEAUME

ÂGE : 72

Prénom : YVETTE

TAILLE : 1,60

FEUILLE DE SURVEILLANCE

CHAMBRE : 12

SERVICE : INTERNE

Période

JOURS		12		13		14		15		16		17		18	
		M	S	M	S	M	S	M	S	M	S	M	S	M	S

Graphiques

Diurèse	TA	T°
4 L	30	41°
3,5 L	25	40°
3 L	20	39°
2,5 L	15	38°
2 L	10	37°
1,5 L	5	36°
1 L	0	35°
0,5 L		34°

Journal de bord

VOIES ORALES		RECTALE		PERCUTANÉE		INJECTIONS		DIVERS	
TA		16/10/14		10/15/13/7		14/9/14/8		15/10/16/12	
POIDS									
DIURÈSE									
SELLES									
Dolipal		1		1		1		1	
Pellanne		1		0		1		1	
Ranitac								1	
Proctogal		1		3		3			
Isoniaz		0		1		0		0	
CUB		1							
Hémoculture		1		3					

(Document CHR de Montluçon)

Un suivi seul des "paramètres" du malade n'aurait pas de sens. S'il y a un changement, il faut essayer d'en déterminer si possible la cause et avoir le journal de bord est indispensable pour orienter ses recherches (baisse de température suite à une prise de médicament par exemple).

Le journal de bord fait partie du processus de contrôle et est tenu parallèlement à la carte de contrôle. Il récapitule tous les événements marquants de la fabrication, permettant en cela un suivi précis des causes assignables ; par exemple, il est consigné :

- l'heure des prélèvements,
- les réglages,
- les pannes et incidents,
- les changements d'outils,
- les changements de lots de matières premières,
- les changements d'opérateurs...

Dans la suite de la brochure, nous nous intéressons uniquement à la construction et à l'utilisation d'une carte de contrôle, sous l'aspect mathématique, mais il ne faut pas perdre de vue que n'est étudiée ici qu'une partie du processus de contrôle.

1 Quelques définitions

1.1 Processus

C'est la combinaison complète

- . des personnes (main d'œuvre),
- . des machines et du matériel (moyen),
- . des matières premières (matière),
- . de l'environnement et de l'atelier (milieu),
- . de la méthode et de l'organisation

dont le travail en commun a pour résultat le produit.

1.2 Maîtriser le processus

C'est réduire la variabilité du processus pour améliorer ses performances et assurer la conformité (et non la qualité) des produits.

1.3 Carte de contrôle

C'est un graphique sur lequel on fait correspondre un point à chacune des valeurs d'une statistique calculée sur des échantillons successifs, en général de même effectif prélevés dans la fabrication.

Chacun de ces points a pour abscisse le numéro d'échantillon et pour ordonnée la valeur de la statistique calculée sur cet échantillon (ici moyenne et étendue).

Des limites de contrôle sont tracées sur cette carte.

Le but est de détecter en temps réel les anomalies de la caractéristique observée.

2 Principe de la méthode

Elle consiste à définir une règle de décision aboutissant au rejet ou au non rejet du lot échantillonné.

On utilise un test :

paramétrique : la variable sur laquelle on travaille est à valeur réelle. Le test peut être bilatéral (ici pour l'étude de la moyenne) ou unilatéral (ici pour l'étude de l'étendue).

non paramétrique : l'ordre d'apparition étant conservé, on prend pour variable le nombre d'apparitions successives présentant une particularité voulue.

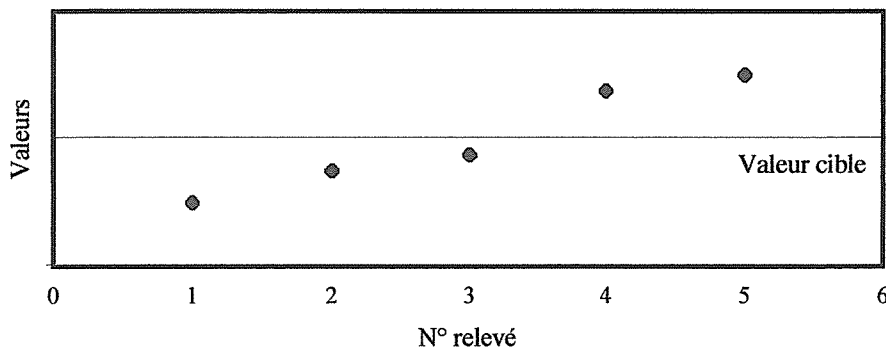
3 La carte de contrôle

3.1 Objectif

À l'aide de la carte de contrôle, on doit être capable de faire la distinction entre les causes assignables, qu'on peut identifier (sur lesquelles on peut agir) et les causes non assignables ou aléatoires, inhérentes à la fabrication (Shewart ; norme NFX 50-020).

3.2 Quelques exemples

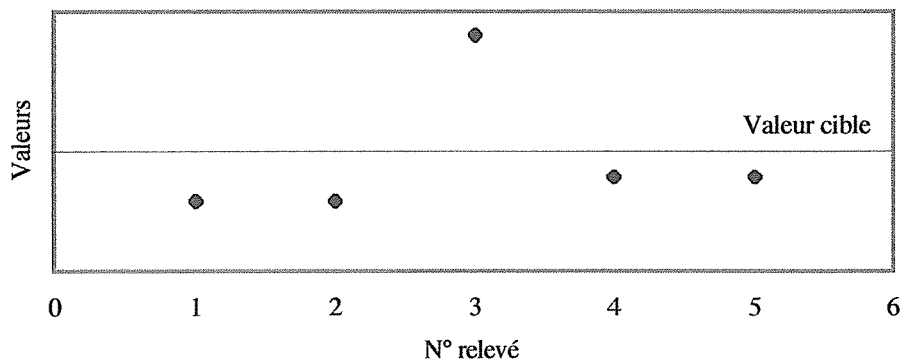
3.2.1 sur le graphique suivant,



deux interprétations sont possibles :

- * il y a des variations dues aux causes aléatoires,
- * il y a une dérive (due par exemple à une usure du moyen de production). Il y a alors une cause assignable sur laquelle on peut intervenir.

3.2.2 sur le graphique suivant,



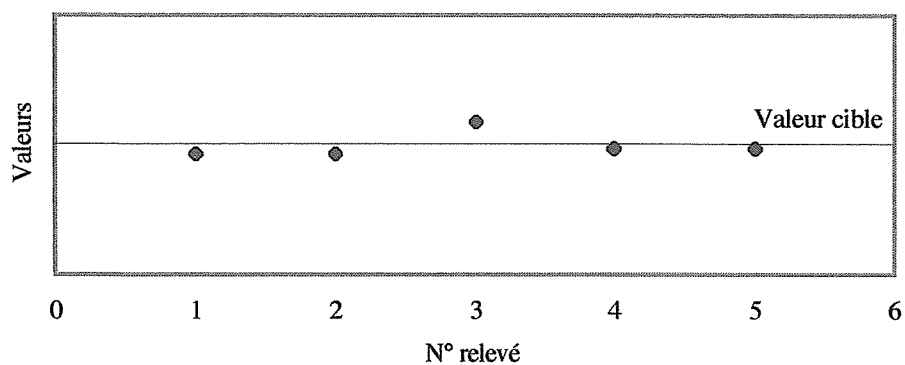
deux interprétations sont possibles :

- * il y a des variations dues aux causes aléatoires,
- * il y a une dérive (due par exemple à un changement dans le milieu de production). Il y a alors une cause assignable sur laquelle on peut intervenir.

Dans chacun des deux cas précédents, les interprétations doivent conduire à une décision : ou bien la production continue en l'état,
ou bien on agit sur le moyen de production.

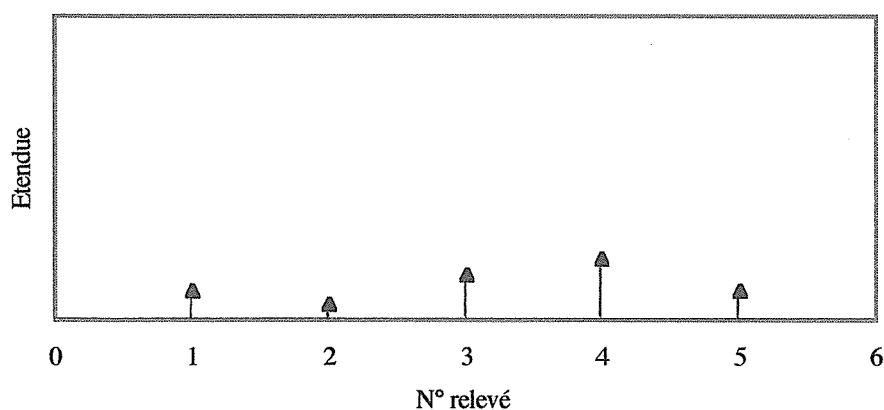
Il n'est pas question de modifier sans arrêt les paramètres, ou de laisser faire...

3.2.3 sur le graphique suivant, où apparaît uniquement la moyenne des échantillons testés, il semble que les résultats soient “acceptables” (suivant des conditions qu’on aura définies au préalable) :

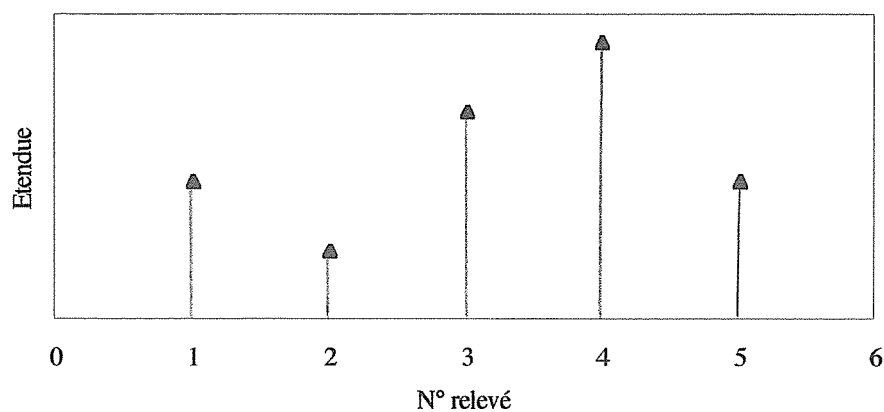


Cependant, si on ajoute sur la même carte les étendues de ces mêmes échantillons, on n’aura pas la même interprétation dans chacun des deux cas suivants :

Premier cas



Deuxième cas



3.3 Définition

La carte de contrôle est le film de la variabilité d'une caractéristique du produit ou d'un paramètre du processus (du milieu, de méthode, de moyen, de main d'œuvre, de matière). La carte de contrôle est un signal "temps réel" (graphique temporel) qui doit déclencher les actions correctives immédiates.

La carte de contrôle est un outil d'actions :

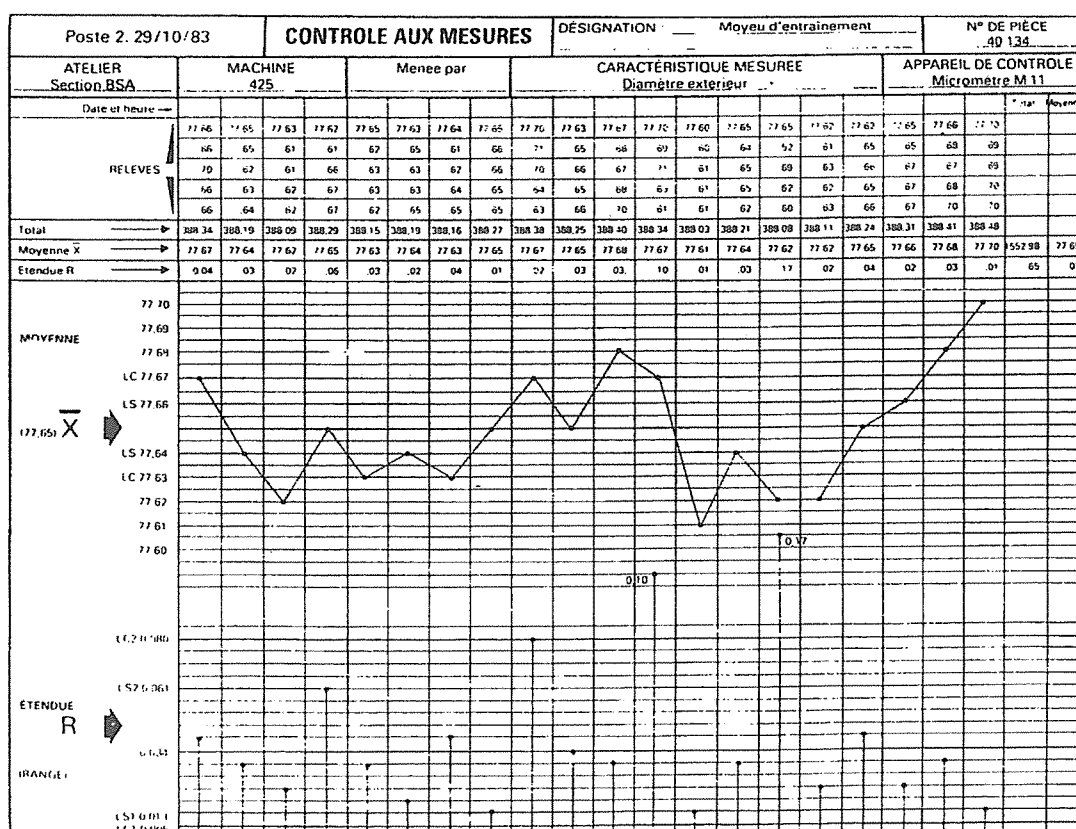
On applique les tests construits suivant les objectifs qu'on s'est fixés et il faut obtenir des opérateurs qu'ils agissent au bon moment.

Par exemple, pour les tests paramétriques,

- . Si le point est hors limites Stop ==> ACTION - CORRECTION
- . Si le point est dans les limites ==> RIEN.

Les cartes de contrôle implantées dans les plans de surveillance sont issues d'une réflexion et devraient être les seuls outils déclenchant des retouches en atelier. Aucune retouche ne devrait être effectuée sans qu'une carte au moins n'ait franchi une de ses limites.

3.4 Exemple



3.5 Différents types de cartes de contrôle

Il y a deux familles :

Aux mesures, pour les données quantitatives et continues :

- mesure des caractéristiques du produit, par exemple jeu, affleurement, épaisseur d'apprêt...
- mesure des paramètres du processus, par exemple hygrométrie, température ; pression, couple de serrage...

Aux attributs, pour les données qualitatives et discrètes :

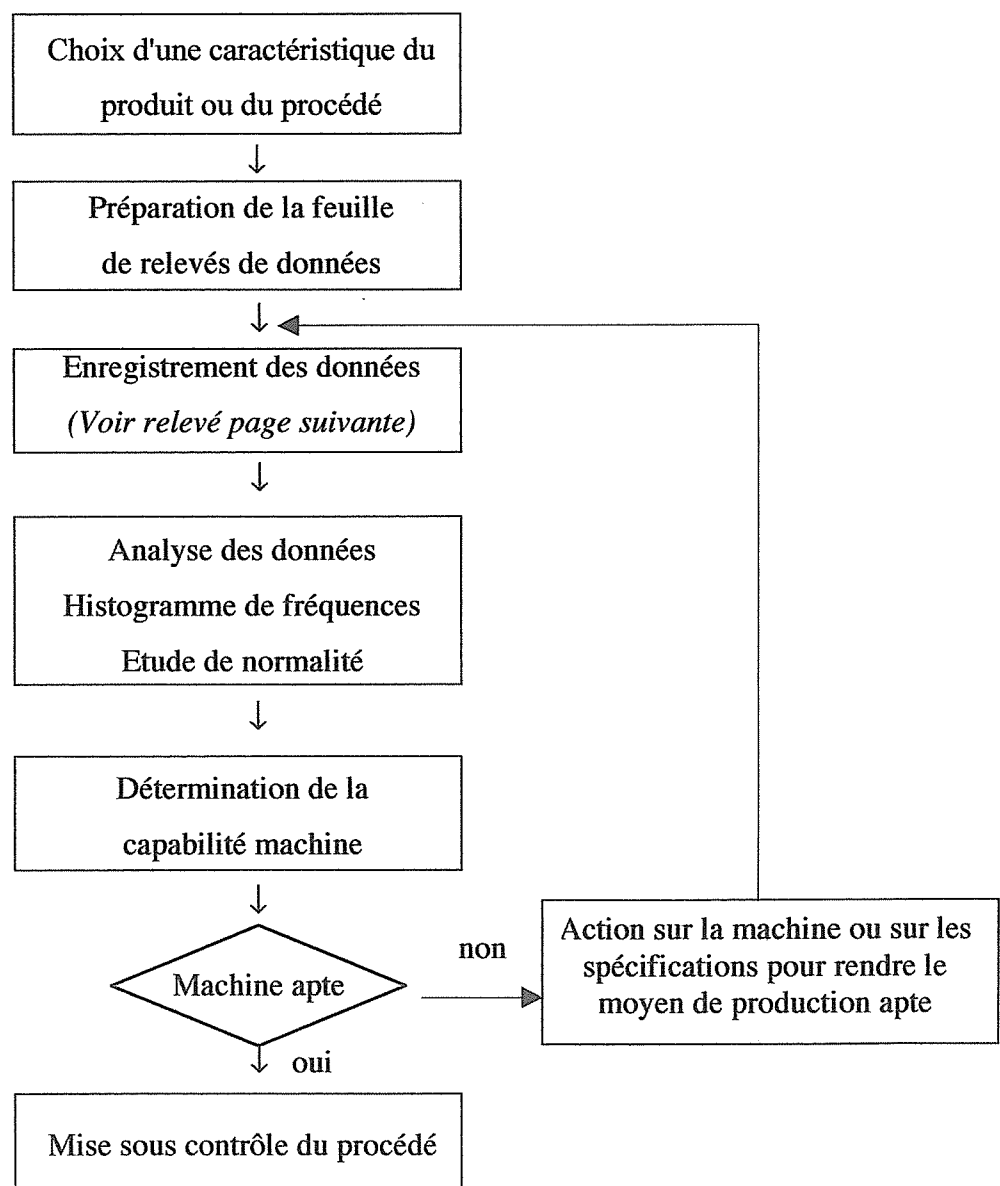
- caractéristiques du produit, par exemple nombre de grains sur une pièce donnée, bruit au montage (oui / non)...

4 Mise en œuvre

4.1 Organigramme de la mise en place d'une carte de contrôle

4.1.1 Préparation à la mise sous contrôle du procédé (dans le bureau d'étude)

Rôle du statisticien



Produit	Machine	Section	Lot	O.F.	Cote surveillée	Cote précédente	Cible	LSCx	LICx	R	LSCr	Fréquence de prél : 30 mn
Moyen	Somab 400		401		31,730	31,780						
Lot												
Caractéristique	φent											
Contrôle SPC												

N° de carte

Date															
Heure	11h	12h	13h	14h	15h	16h	17h	18h	19h		20h				
x1	741	739	741	755	776	747	752	742	756	755	764				
x2	745	760	763	761	750	777	753	739	760	755	764				
x3	749	759	762	759	755	770	751	750	759	751	739				
x4	751	759	761	748	779	769	749	765	757	766	760				
x5	760	761	740	775	746	780	764	765	747	752	761				
TOTAL	3746	3781	3787	3800	3806	3843	3762	3761	3779	3774	3788				
MOYENNE $\bar{X}$	749	756	757	758	761	769	752	752	756	755	758				
ETENDUE R	19	15	23	27	33	33	15	26	13	16	25				

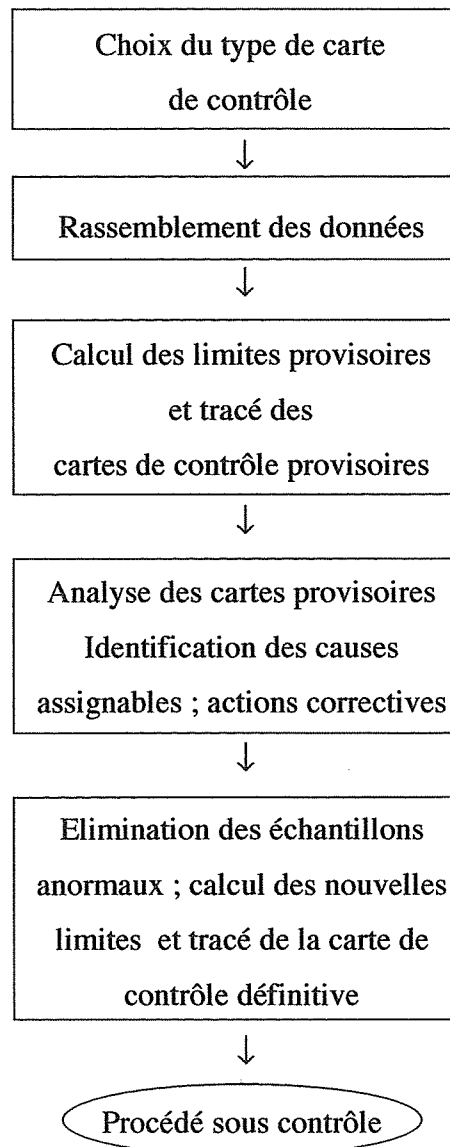
$\bar{X}$

R

Journal

4.1.2 Mise sous contrôle du procédé (dans l'atelier)

Rôle de l'utilisateur



4.2 La carte moyenne-étendue

Il est présenté ici une carte de contrôle aux mesures et les paramètres sont la moyenne et l'étendue.

Les échantillons sont prélevés à intervalles de temps réguliers et sont constitués de 5 éléments. C'est un nombre courant, ni trop grand, ni trop petit, et qui permet à tout opérateur de calculer la moyenne et l'étendue de l'échantillon facilement.

Des formules permettent de déterminer la fréquence de prélèvement. En voici un exemple : soient c la cadence horaire de la machine, n le nombre de pièces à prélever, N le nombre de pièces fabriquées entre deux réglages, la fréquence de prélèvement est $\sqrt{nN} \frac{60}{c}$ minutes.

Pour la construction du test sur la moyenne, on utilise un test bilatéral.
(voir la bibliographie)

Pour la construction du test sur l'étendue, on pourra se référer à l'annexe B.

Pour la construction du test non paramétrique utilisé ici, on conserve l'ordre d'apparition des moyennes d'échantillon. Considérons la variable aléatoire Y qui prend pour valeur 0 si la moyenne de l'échantillon est inférieure à la valeur cible et 1 si la moyenne de l'échantillon est supérieure à la valeur cible. Intéressons-nous au nombre de points successifs au-dessus de la valeur cible, c'est-à-dire au nombre successif de "1" qu'on obtient.

En supposant le procédé sous contrôle (c'est-à-dire que la moyenne est la valeur cible), la probabilité de l'événement " $Y = 1$ " est 0,5. L'hypothèse alternative est "la probabilité de l'événement " $Y = 1$ " est différente de 0,5".

Dans la mesure où l'on est sous l'hypothèse "le procédé est sous contrôle", et en supposant les événements successifs indépendants, la probabilité d'apparition de l'événement " $Y = 1$ n fois successivement" est $0,5^n$.

A l'aide de la table donnée ci-dessous, on trouve les résultats suivants : pour un risque d'erreur de 1 %, il faut prendre $n = 7$ et pour un risque d'erreur de 2%, il faut prendre $n=9$. (règles des 7 et 9 points utilisées dans l'industrie).

n	2	3	4	5	6	7	8	9
$0,5^n$	0,25	0,125	0,0625	0,0312	0,0156	0,0078	0,0039	0,0019

La règle de décision est alors :

* S'il y a strictement moins de 7 points successivement du même côté de la valeur cible, au seuil d'erreur de 1%, on n'a pas de raison de refuser l'hypothèse "Le procédé est sous contrôle" ;

* S'il y a 7 points successivement du même côté de la valeur cible, on refuse l'hypothèse nulle au seuil d'erreur de 1%. Le procédé n'étant plus sous contrôle, il y a lieu de faire la recherche d'une cause assignable.

Il est possible de déterminer aussi le nombre de points successifs croissants par exemple. On prend pour nouvelle variable la différence entre deux valeurs successives et on note par exemple "+" si il y a croissance, et "-" si il y a décroissance. Le test est alors construit en utilisant la variable aléatoire qui prend pour valeur le nombre de "+" (ou de "-") consécutifs.

5 Etude de normalité

Objet : Les tests utilisés ont pour but de vérifier que l'échantillon considéré est issu d'une population qui suit une loi normale.

Les tests d'adéquation les plus courants sont, pour les tests numériques : le test du χ^2 (Khi deux) pour les échantillons de taille $n > 50$, celui de Kolmogorov-Smirnov, celui de Shapiro-Wilks (non développé dans cette brochure) ou encore, comme test graphique, celui de la droite de Henry .

Pour les tests numériques, on définit un seuil de risque α (ou un niveau de confiance $1 - \alpha$).

Principe : On teste l'accord entre les données statistiques (fonction de répartition empirique) et le modèle théorique supposé (fonction de répartition théorique de la loi normale dans notre cas).

5.1 Droite de Henry

Principe :

On utilise une graduation spéciale pour ramener l'histogramme de la population à une courbe qui, si la population est normale, est une droite.

Etude théorique :

Supposons que les observations soient réparties selon une loi normale. Dans ce cas, les fréquences cumulées observées doivent être égales aux valeurs correspondantes de la fonction de répartition de la loi normale.

L'espérance mathématique et l'écart type théoriques seront estimés à partir de la série statistique.

Soit $F_0(x_k) = \sum_{i=1}^k f_i$ la fréquence cumulée observée

Soit $F_T(x_k) = \Pi\left(\frac{x_k - m}{\sigma}\right) = \Pi(t_k)$ la valeur de la fonction de répartition de la

loi normale $N(m, \sigma)$.

F_0 est donc la fréquence observée et F_T la fréquence théorique.

Si $F_0(x_k) = F_T(x_k)$, alors :

il existe alors une relation affine entre x_k et t_k : $t_k = \frac{x_k - m}{\sigma} \Leftrightarrow t_k = \frac{1}{\sigma} x_k - \frac{m}{\sigma}$
et cette relation se traduit graphiquement par une droite, dite droite de Henry.

On part des fréquences cumulées $F_0(x_k)$. Sachant que $F_0(x_k) = \Pi(t_k)$, en utilisant une table de la loi normale, on peut en déduire les valeurs de t_k correspondantes. En portant x_k en abscisse et t_k en ordonnée, on regarde si les points sont approximativement alignés.

Pratiquement, on évite le recours à la table de $\Pi(t)$ en utilisant un papier gausso-arithmétique gradué directement en $\Pi(t)$. Il suffit alors de porter x en abscisse et $F_0(x)$ en ordonnée.

Méthodologie :

1ère étape : - Rassembler et ordonner les données en classes (respecter la norme CNOMO pour construire les classes : elles sont usuellement de type $[a ; b]$ et en nombre approximativement égal à $\sqrt{n}$, avec n l'effectif total).
Tracer l'histogramme de la distribution.
Calculer les fréquences cumulées par rapport à l'échantillon total.

2e étape : - Tracer les points sur papier gaussio-arithmétique.
Remarque : le point correspondant à la valeur 100 % est rejeté à l'infini et n'apparaît donc pas sur le papier gaussio-arithmétique.

3e étape : - Interpréter la courbe.
Si les points sont sensiblement alignés, on peut alors conclure à la normalité de la courbe.

Outre l'étude de normalité, la droite de Henry permet d'estimer très rapidement les différents paramètres de position et de dispersion :

Estimation de la moyenne : on repart de l'horizontale donnant 50 % de la population vers l'axe des abscisses (la moyenne étant égale à la médiane pour une loi normale).

Estimation de l'écart-type : sachant que 99,73 % de la population est contenue dans $[m - 3\sigma, m + 3\sigma]$, on mesure 6 fois l'écart-type entre 0,135 % et 99,86 % (ou 2 fois σ entre 16 % et 84 %).

Pourcentage hors tolérance : les limites supérieures et inférieures de tolérance, spécifiées par le bureau d'étude (et prises sur l'axe des abscisses) sont reportées à l'aide de la droite sur l'axe des ordonnées. On lit ainsi directement (sur l'axe des ordonnées) le pourcentage hors tolérance.

Intérêt :

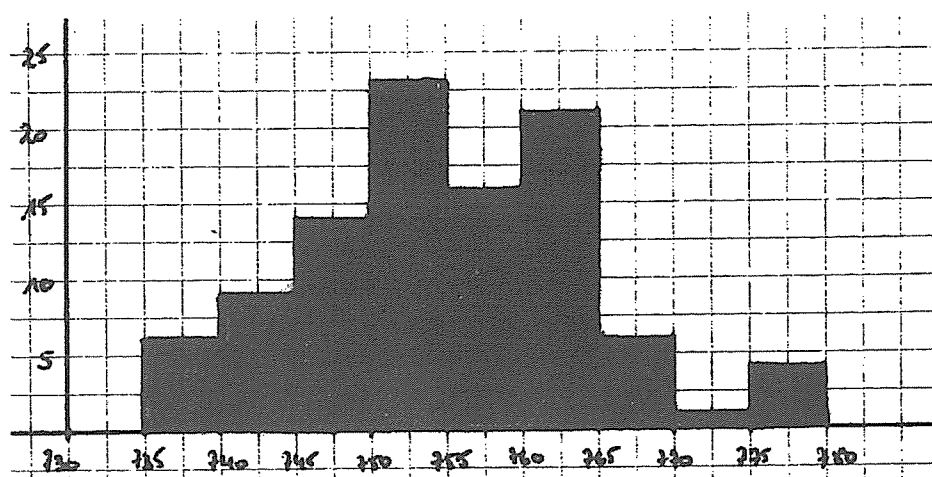
Simplicité et rapidité.

Approximation suffisante dès que le nombre de valeurs est assez grand ($n > 30$).

Exemple avec la carte de contrôle :

a) Les données (cf p 8) sont rassemblées et ordonnées en classes.

Tracé de l'histogramme de la distribution observée.

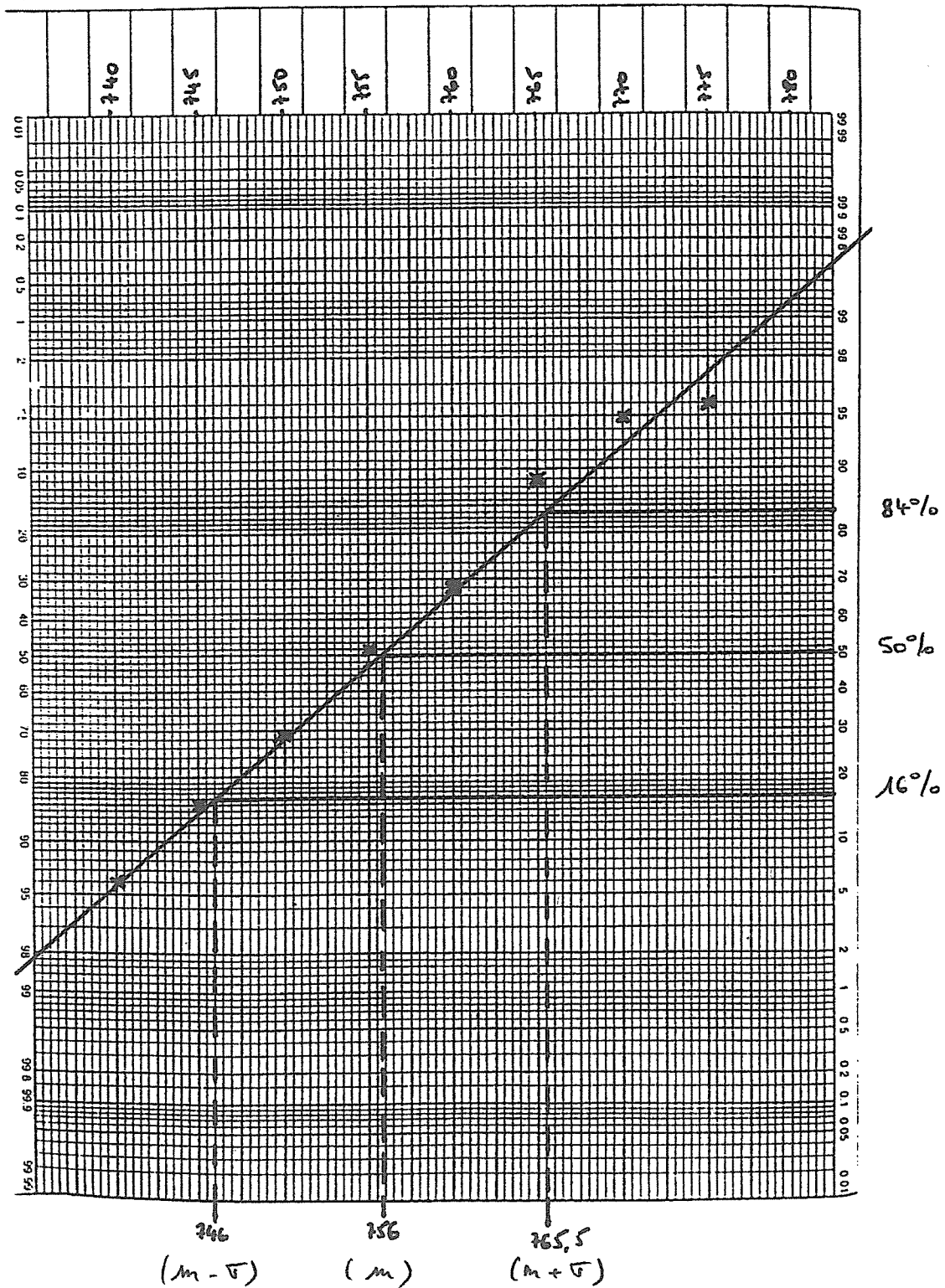


Calcul des fréquences cumulées croissantes.

Résultats : Tableau des données :

Diamètres	]730 ; 735]	]735 ; 740]	]740 ; 745]	]745 ; 750]	]750 ; 755]
Effectifs	0	6	9	14	23
Fréquences cumulées	0	0,06	0,15	0,29	0,52
Diamètres	]755 ; 760]	]760 ; 765]	]765 ; 770]	]770 ; 775]	]775 ; 780]
Effectifs	16	21	6	1	4
Fréquences cumulées	0,68	0,89	0,95	0,96	1

b) Tracé des points sur papier gaucho-arithmétique (Report en abscisse des limites supérieures des classes et en ordonnée des fréquences cumulées.)



c) Interprétation de la courbe : les points sont sensiblement alignés, on peut alors conclure à la normalité de la loi de distribution.

Tracé de la droite de Henry (tracé à l'œil, le résultat peut donc varier d'un opérateur à l'autre).

Estimation de la moyenne et de l'écart type de cette série avec la droite de Henry :

$$m \approx 756 ; \sigma \approx (765,5 - 746) \div 2 = 9,75.$$

A titre indicatif, la moyenne et l'écart type calculés avec les 100 données sont :

$$m = 755,8 ; \sigma = 9,1.$$

5.2 Test de Kolmogorov - Smirnov

Principe :

Il est utilisable pour tester l'ajustement à toute loi dont la fonction de répartition est continue (ce qui est le cas pour la loi normale) et repose sur le théorème suivant : si X est une variable aléatoire de fonction de répartition F continue, alors $F(X)$ est une variable aléatoire uniforme sur $[0 ; 1]$.

Le test est alors basé sur le calcul des écarts entre les valeurs observées et les valeurs théoriques de la fonction de répartition.

Il est valable quelle que soit la taille n de l'échantillon. Cependant, si n est trop grand, il est plus pratique de regrouper les valeurs en classes et d'utiliser alors le test du χ^2 .

Etude théorique :

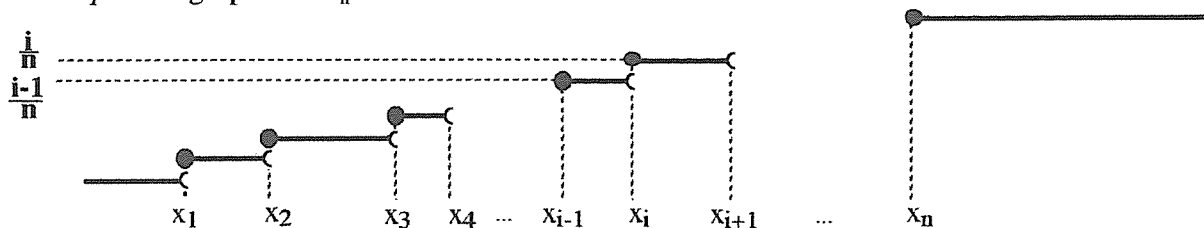
Soit un échantillon aléatoire de n pièces extrait d'une population dont on range les valeurs par ordre croissant : $x_1 \leq x_2 \leq \dots \leq x_n$

Pour tout x réel, $F_n(x)$ est la proportion des n variables $X_1, X_2, \dots, X_n$ indépendantes, vérifiant : $X_i \leq x$ ($i = 1, 2, \dots, n$).

$F_n(x)$ est la fonction de répartition empirique de l'échantillon de taille n .

$$F_n(x) = \begin{cases} 0 & \text{si } x < x_1 \\ \frac{i}{n} & \text{si } x_i \leq x < x_{i+1} \\ 1 & \text{si } x \geq x_n. \end{cases}$$

Remarque : le graphe de F_n est une courbe en escalier.



Soit $F(x)$ la fonction de répartition théorique. Par exemple, on peut tester l'hypothèse que $F(x)$ soit la fonction de répartition de la loi normale $N(\mu, \sigma)$ où μ et σ sont estimés par $\bar{x}$ et s , obtenus à partir de l'échantillon.

On mesure point par point l'écart entre les deux fonctions :

$$D = \left| F_n(x) - F(x) \right|$$

Pour cela, on définit : $D^+ = \text{Max}_i \left| \frac{i}{n} - F(x_i) \right|$

$$D^- = \text{Max}_i \left| \frac{i-1}{n} - F(x_i) \right|$$

$$D = \text{Max} (D^+, D^-)$$

D'après la loi des grands nombres et le théorème de Glivenko, on a :

$$P \left(\lim_{n \rightarrow +\infty} \sup_x \left| F_n(x) - F(x) \right| = 0 \right) = 1$$

Autrement dit : la fonction de répartition empirique $F_n(x)$ converge presque sûrement uniformément vers la fonction de répartition théorique $F(x)$.

On écrit aussi que $P \left(\text{Max}_x \left| F_n(x) - F(x) \right| < D_{n,\alpha} \right) = 1 - \alpha$

soit $P (D < D_{n,\alpha}) = 1 - \alpha$.

La valeur de $D_{n,\alpha}$ est donnée par la table de Kolmogorov - Smirnov (voir en annexe A)

Si $D > D_{n,\alpha}$ on refuse l'hypothèse du modèle théorique.

Exemple 1 :

On mesure les temps entre défaillances au cours d'un essai sur un matériel :

23h, 16h, 56h, 71h, 4h, 25h, 51h, 30h.

Peut-on admettre un modèle gaussien de paramètres $m = 34$, $\sigma = 22$ avec un risque de 0,05 ?

La fonction de répartition théorique est $F(t) = \int_{-\infty}^t \frac{1}{\sqrt{2\pi} \times 22} e^{-\frac{(t-34)^2}{2 \times 22^2}} dt$

En passant par la variable centrée réduite $u = \frac{t-34}{22}$, on peut calculer $F(t) = \Pi(u)$

t_i : temps rangés en ordre croissant	$F(t_i)$ théorique $= \Pi\left(\frac{t_i - 34}{22}\right)$	$\frac{i}{n} = \frac{i}{8}$	$\frac{i-1}{n} = \frac{i-1}{8}$	$\left \frac{i}{n} - F(t_i) \right $	$\left \frac{i-1}{n} - F(t_i) \right $
4	0,086	0,125	0	0,039	0,086
16	0,200	0,250	0,125	0,050	0,075
23	0,308	0,375	0,250	0,067	0,058
25	0,345	0,500	0,375	0,155	0,030
30	0,428	0,625	0,500	0,197	0,072
51	0,779	0,750	0,625	0,029	0,154
56	0,841	0,875	0,750	0,034	0,091
71	0,953	1,000	0,875	0,047	0,078

$D^+ = 0,197$ et $D^- = 0,154$; donc $D = 0,197$.

Sur la table, on lit $D_{8;0,05} = 0,457$. (Cf annexe A)

Par conséquent, on n'a pas de raison de rejeter l'hypothèse du modèle gaussien avec un risque de 5 %.

Remarque : pour une carte de contrôle, ce test est trop long à mettre en œuvre, à moins d'utiliser un outil informatique tel qu'un tableur. Sinon on utilise le test du χ^2 .

Exemple 2 avec les données de la carte de contrôle :

Données	x_i	f_i	$F(x_i)$	$ f_i - F(x_i) $	$ f_{i-1} - F(x_i) $
738	-1,9389	0,0100	0,0263	0,0163	0,0263
739	-1,8301	0,0200	0,0336	0,0136	0,0236
740	-1,7212	0,0600	0,0426	0,0174	0,0074
741	-1,6123	0,0700	0,0534	0,0166	0,0066
742	-1,5035	0,1000	0,0664	0,0336	0,0236
745	-1,1769	0,1500	0,1196	0,0304	0,0204
746	-1,0680	0,1600	0,1428	0,0172	0,0072
747	-0,9591	0,1800	0,1687	0,0113	0,0013
748	-0,8503	0,1900	0,1976	0,0076	0,0176
749	-0,7414	0,2000	0,2292	0,0292	0,0392
750	-0,6325	0,2300	0,2635	0,0335	0,0435
751	-0,5237	0,3000	0,3003	0,0003	0,0103
752	-0,4148	0,3800	0,3391	0,0409	0,0309
753	-0,3059	0,3900	0,3798	0,0102	0,0002
754	-0,1971	0,4000	0,4219	0,0219	0,0319
755	-0,0882	0,5200	0,4649	0,0551	0,0451
756	0,0207	0,5300	0,5083	0,0217	0,0117
757	0,1296	0,5400	0,5515	0,0115	0,0215
759	0,3473	0,6000	0,6358	0,0358	0,0458
760	0,4562	0,6100	0,6759	0,0659	0,0759
761	0,5650	0,7700	0,7140	0,0560	0,0460
762	0,6739	0,7800	0,7498	0,0302	0,0202
763	0,7828	0,8400	0,7831	0,0569	0,0469
764	0,8916	0,8500	0,8137	0,0363	0,0263
765	1,0005	0,8900	0,8415	0,0485	0,0385
766	1,1094	0,9000	0,8664	0,0336	0,0236
769	1,4360	0,9400	0,9245	0,0155	0,0055
770	1,5448	0,9500	0,9388	0,0112	0,0012
775	2,0892	0,9600	0,9817	0,0217	0,0317
776	2,1980	0,9700	0,9860	0,0160	0,0260
777	2,3069	0,9800	0,9895	0,0095	0,0195
779	2,5246	0,9900	0,9942	0,0042	0,0142
780	2,6335	1,0000	0,9958	0,0042	0,0058

(Voir l'annexe C pour la construction de ce tableau.)

$D^+ = 0,0658$ et $D^- = 0,0758$; donc $D = 0,0758$.

Sur la table, on lit $D_{100; 0,05} = 0,1340$. (Cf annexe A)

Par conséquent, on n'a pas de raison de rejeter l'hypothèse du modèle gaussien avec un risque de 5 %.

5.3 Test du χ^2 (Khi - Deux)

Principe :

Ce test est basé sur le calcul des distances entre les valeurs observées et les valeurs théoriques.

Il est utilisable uniquement pour des lois discrètes. Si la loi est continue, on se ramène à une loi discrète en regroupant les données en classes de la forme $]x_{i-1} ; x_i]$. Dans chaque classe, on doit obtenir un nombre théorique np_i supérieur ou égal à 5, l'intervalle de classe n'étant pas forcément régulier.

Une fonction indicatrice des écarts est établie de la manière suivante :

$$E = \sum_{i=1}^r \frac{(n_i - np_i)^2}{np_i}$$

où : r : nombre de classes;

n : taille de l'échantillon;

n_i : nombre d'individus observés dans la $i^{\text{ème}}$ classe;

p_i : probabilité pour un individu de se trouver dans la $i^{\text{ème}}$ classe;

np_i : nombre d'individus attendus théoriquement dans la $i^{\text{ème}}$ classe.

On montre que E suit approximativement une loi de χ^2 à v degrés de liberté avec $v = r - k - 1$.

k est le nombre de paramètres à estimer pour le modèle théorique, par exemple :

loi exponentielle : $k = 1$

loi normale : $k = 2$

loi de Weibull : $k = 3$.

Les valeurs de $\chi^2_{v,\alpha}$ sont sur la table de distribution de χ^2 (voir en annexe A).

Si $E > \chi^2_{v,\alpha}$, on rejette l'hypothèse du modèle théorique.

Exemple avec les données de la carte de contrôle :

A partir de la série statistique où $n = 100$, on estime $m = 756$ et $\sigma = 9$. Peut-on prendre un modèle gaussien ?

Diamètres	]735 ; 745]	]745 ; 750]	]750 ; 755]	]755 ; 760]	]760 ; 765]	]765 ; 770]	]770 ; 780]
n_i	15	14	23	16	21	6	5
$t_i = \frac{x_i - m}{\sigma}$	- 1,22	- 0,67	- 0,11	0,44	1	1,56	2,67
$\Pi(t_i)$	0,1112	0,2514	0,4562	0,6700	0,8413	0,9406	0,9962
$p_i = \Pi(t_i) - \Pi(t_{i-1})$	0,1112	0,1402	0,2048	0,2138	0,1713	0,0993	0,0556
np_i	11,12	14,02	20,48	21,38	17,13	9,93	5,56
$\frac{(n_i - np_i)^2}{np_i}$	1,3538	0,00003	0,3101	1,3538	0,8743	1,5554	0,0564

$E = \chi^2 = 5,5038$; $v = 7 - 2 - 1 = 4$.

Or $\chi^2_{4,0,05} = 9,59$ (cf annexe A).

L'hypothèse "distribution normale " est acceptable au seuil de 5 %.

6 Exemple

6.1 Données

Cf tableau 1 p 8.

6.2 Construction d'une carte des moyennes et de l'étendue ($\bar{X}, R$)

A partir des 20 échantillons de 5 pièces, calcul de la moyenne des moyennes $\bar{\bar{x}}$ et de la moyenne des étendues $\bar{R}$.

Résultats : Moyenne des moyennes $\bar{\bar{x}} \approx 756$.
 Moyenne des étendues $\bar{R} \approx 20$.

Carte des moyennes :

Tracé de la ligne centrale : $y = \bar{\bar{x}}$.

de la limite de contrôle inférieure LCI : $y = \bar{\bar{x}} - A_2 \bar{R}$.

de la limite de contrôle supérieure LCS : $y = \bar{\bar{x}} + A_2 \bar{R}$, où A_2 est un coefficient qui vaut 0,577 lorsqu'on considère des échantillons de 5 pièces. (cf annexe A)

Résultats : Ligne centrale : $y = 756$.
 Limite de contrôle inférieure LCI : $y = 756 - 0,577 \times 20 \approx 744$.
 Limite de contrôle supérieure LCS : $y = 756 + 0,577 \times 20 \approx 768$.

Carte des étendues :

Tracé de la ligne centrale : $y = \bar{R}$.

de la limite de contrôle inférieure LCI : $y = D_3 \bar{R}$.

de la limite de contrôle supérieure LCS : $y = D_4 \bar{R}$, où D_3 et D_4 sont des coefficients qui valent respectivement 0 et 2,114 lorsqu'on considère des échantillons de 5 pièces. (cf annexes A et B)

Résultats : Ligne centrale : $y = 20$.
 Limite de contrôle inférieure LCI : $y = 0$.
 Limite de contrôle supérieure LCS : $y = 2,114 \times 20 \approx 42$.

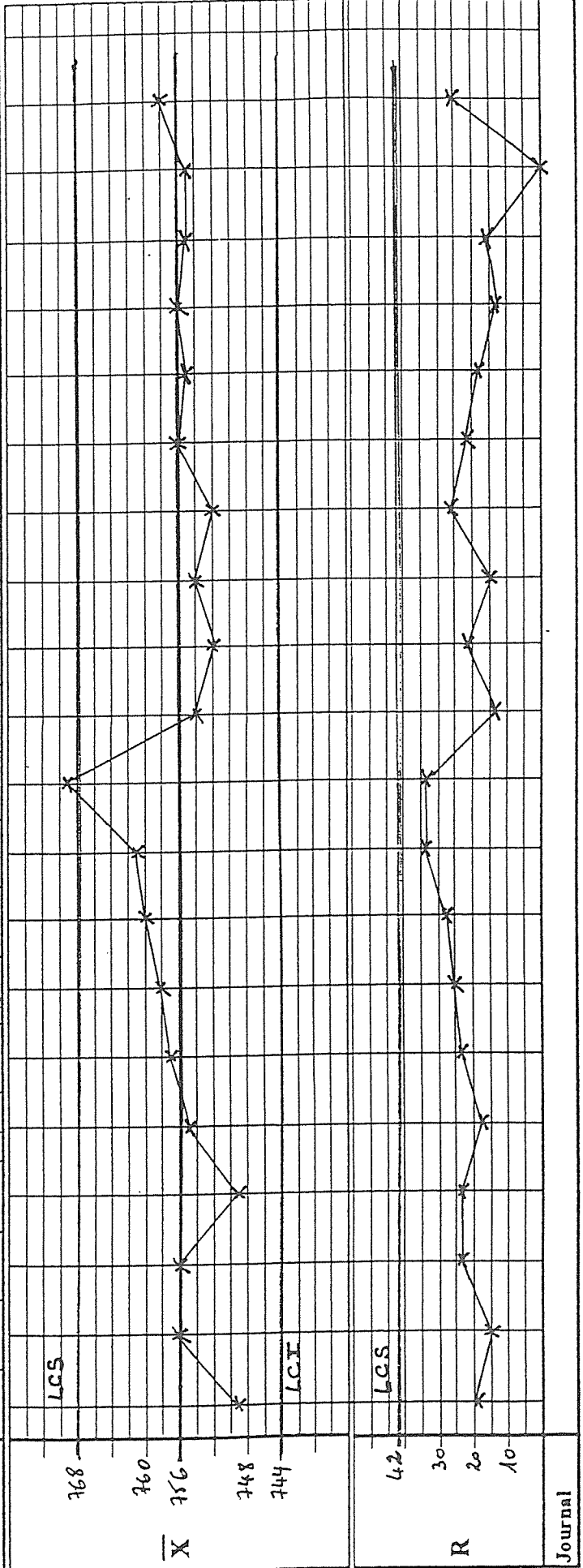
Tracé de la carte de contrôle provisoire :

Report dans la carte de contrôle des points dont l'abscisse représente le numéro de l'échantillon considéré et l'ordonnée, la moyenne $\bar{x}$ de l'échantillon (ou l'étendue W). (cf page 19).

Produit	Référence :	Désignation : Arbre porte gallet	Client :		
Moyen	Machine : Somab 400	Section :	Opération :		
Lot	O.F. :	Lot : L10	Cm	Cp	Cpk
Caractéristique ϕ_{act}	Cote surveillée :	Carte précédente			
Contrôle SPC	Cible : 756	LSC : 768	LSC : 744	R : 20	Fréquence de prél. : 42

N° de carte

Date																				
Heure		11h		12h		13h		14h		15h		16h		17h		18h		19h		20h
x1	741	760	739	744	766	761	741	755	776	747	750	755	752	742	754	745	756	755	755	764
x2	745	755	760	738	752	763	762	761	750	777	755	751	753	739	763	745	760	750	755	764
x3	749	759	762	750	751	762	759	761	755	770	752	757	751	750	766	763	759	751	755	739
x4	751	745	759	757	749	761	761	748	779	769	750	760	749	765	750	761	757	766	755	760
x5	760	760	761	761	755	740	766	775	746	780	763	739	764	765	745	760	747	752	755	761
TOTAL	3746	3779	3781	3747	3773	3787	3789	3800	3806	3843	3770	3762	3769	3761	3778	3774	3779	3774	3775	3788
MOYENNE $\bar{X}$	749	756	756	749	755	757	758	760	761	769	754	752	754	752	756	755	756	755	755	758
ETENDUE R	19	15	23	23	17	23	25	27	33	33	13	21	15	26	21	18	13	16	0	25



Journal

Analyse de la carte de contrôle provisoire : élimination des échantillons anormaux et calcul des nouvelles limites.

On vérifie si tous les points sont à l'intérieur des limites. Or le graphique montre que l'échantillon n° 10 a une moyenne située en dehors de la limite de contrôle supérieure.

Il faut alors rechercher la cause de cet événement et engager une action corrective afin qu'il ne se reproduise plus. (La cause assignable peut être recherchée avec l'aide du journal de bord et relève alors du domaine de l'expert en productique; elle pourrait être associée au changement d'équipe par exemple).

Calcul des nouvelles limites :

Ce calcul s'effectue en supprimant l'échantillon 10 dans les calculs de $\bar{\bar{x}}$ et de $\bar{\bar{R}}$.

Moyenne des moyennes $\bar{\bar{x}} \approx 755$.

Moyenne des étendues $\bar{\bar{R}} \approx 20$.

Carte des moyennes :

Ligne centrale : $y = 755$.

Limite de contrôle inférieure LCI : $y = 755 - 0,577 \times 20 \approx 743$.

Limite de contrôle supérieure LCS : $y = 755 + 0,577 \times 20 \approx 767$.

Carte des étendues :

Ligne centrale : $y = 20$.

Limite de contrôle inférieure LCI : $y = 0$.

Limite de contrôle supérieure LCS : $y = 2,114 \times 20 \approx 42$.

6.3 Utilisation de la carte

Pour chacun des 15 nouveaux échantillons obtenus, calcul de la moyenne $\bar{x}$ et de l'étendue W . (cf page 21)

Construction de la nouvelle carte : (cf page 22)

Tracé sur le graphique des lignes de la carte des moyennes et de celles de la carte des étendues.

Report des points correspondants à ces 15 échantillons.

Carte des moyennes :

Ligne centrale : $y = 755$.

Limite de contrôle inférieure LCI : $y = 743$.

Limite de contrôle supérieure LCS : $y = 767$.

Carte des étendues :

Ligne centrale : $y = 20$.

Limite de contrôle inférieure LCI : $y = 0$.

Limite de contrôle supérieure LCS : $y = 42$.

Analyse de la carte obtenue :

L'objectif est de dire si le procédé est sous contrôle ou non.

Deux types de situations peuvent conduire à affirmer que la production n'est pas sous contrôle :

- lorsque des points sont en dehors des limites,
- lorsque des points forment des modèles ou tendances spécifiques à l'intérieur des limites.

Cette carte souligne une tendance supérieure : une succession de 7 points supérieurs à la moyenne. Cette tendance peut préfigurer un dépassement ultérieur des limites de contrôle. La capacité machine se détériore. Il faut trouver l'origine de cette détérioration et intervenir pour mettre le procédé sous contrôle.

ANNEXES

Annexe A : tableau des coefficients,
table du χ^2 ,
table de la loi de Kolgomorov-Smirnov.

Annexe B : Utilisation de l'étendue.

Annexe C : Activité tableur.

Annexe A

Tableau des coefficients

n	Estimation de σ		d_3	Pour le calcul de la carte $\bar{X}$			Pour le calcul de la carte des S				Pour le calcul de la carte des R				Carte Médiane
	c_4	d_2		A	A_2	A_3	B_3	B_4	B_5	B_6	D_3	D_4	D_5	D_6	$\bar{A}_2$
2	0,7979	1,128	0,853	2,121	1,880	2,659	-	3,267	-	2,606	-	3,267	-	3,686	1,880
3	0,8862	1,693	0,888	1,732	1,023	1,954	-	2,568	-	2,276	-	2,574	-	4,358	1,187
4	0,9213	2,059	0,880	1,500	0,729	1,628	-	2,266	-	2,088	-	2,282	-	4,698	0,796
5	0,9400	2,326	0,864	1,342	0,577	1,427	-	2,089	-	1,964	-	2,114	-	4,918	0,691
6	0,9515	2,534	0,848	1,225	0,483	1,287	0,030	1,970	0,029	1,874	-	2,004	-	5,078	0,548
7	0,9594	2,704	0,833	1,134	0,419	1,182	0,118	1,882	0,113	1,806	0,076	1,924	0,205	5,203	0,508
8	0,9650	2,847	0,820	1,061	0,373	1,099	0,185	1,815	0,178	1,752	0,136	1,864	0,387	5,307	0,433
9	0,9693	2,970	0,808	1,000	0,337	1,032	0,239	1,761	0,232	1,707	0,184	1,816	0,546	5,394	0,412
10	0,9727	3,078	0,797	0,949	0,308	0,975	0,284	1,716	0,277	1,669	0,223	1,777	0,687	5,469	0,362
11	0,9754	3,173	0,787	0,905	0,285	0,927	0,321	1,679	0,314	1,637	0,256	1,744	0,812	5,534	
12	0,9776	3,258	0,778	0,866	0,266	0,886	0,354	1,646	0,346	1,609	0,283	1,717	0,924	5,592	
13	0,9794	3,336	0,770	0,832	0,249	0,850	0,382	1,618	0,374	1,585	0,307	1,693	1,026	5,646	
14	0,9810	3,407	0,762	0,802	0,235	0,817	0,406	1,594	0,399	1,563	0,328	1,672	1,121	5,693	
15	0,9823	3,472	0,755	0,775	0,223	0,789	0,428	1,572	0,420	1,544	0,347	1,653	1,207	5,937	
20	0,9869	3,735	0,729	0,671	0,180	0,680	0,510	1,490	0,503	1,471	0,415	1,585	1,548	5,922	

Valeur du coefficient C (intervalle de confiance à 95%)

N	10	12	14	16	18	20	24	28	30	35	40	50	75	100	120
C	1,64	1,55	1,48	1,43	1,40	1,37	1,32	1,30	1,28	1,26	1,24	1,21	1,16	1,13	1,12

N	150	170	200	250	300	350	400	450	500	600	700	800	900	1000	2000
C	1,11	1,10	1,09	1,08	1,07	1,07	1,06	1,06	1,06	1,05	1,05	1,04	1,04	1,04	1,03

Valeurs de dn^* en fonction de n (nombre de valeurs par échantillon) et k (nombre d'échantillons)

$k \rightarrow$ $\downarrow n$	10	12	14	16	18	20	25	30	40	50	60	70	80	90	100
3	1.231	1.271	1.302	1.327	1.348	1.366	1.400	1.426	1.462	1.486	1.504	1.518	1.529	1.539	1.546
4	1.601	1.641	1.672	1.697	1.717	1.735	1.769	1.794	1.830	1.854	1.872	1.886	1.897	1.906	1.914
5	1.876	1.915	1.946	1.970	1.991	2.008	2.041	2.066	2.101	2.125	2.142	2.156	2.167	2.176	2.183
6	2.092	2.131	2.161	2.185	2.205	2.222	2.255	2.279	2.313	2.336	2.354	2.367	2.378	2.387	2.394
7	2.270	2.308	2.337	2.361	2.381	2.397	2.430	2.453	2.487	2.510	2.527	2.540	2.550	2.559	2.567

Table du χ^2

$$\mathbb{P}\{\chi^2(ddl) \leq .\} = p$$

p=	0,005	0,01	0,025	0,05	0,1	0,25	0,5	0,75	0,9	0,95	0,975	0,99	0,995
ddl													
1	0	0,0002	0,001	0,0039	0,0158	0,1015	0,4549	1,3233	2,7055	3,8415	5,0239	6,6349	7,8794
2	0,01	0,0201	0,0506	0,1026	0,2107	0,5754	1,3863	2,7726	4,6052	5,9915	7,3778	9,2104	10,597
3	0,0717	0,1148	0,2158	0,3518	0,5844	1,2125	2,366	4,1083	6,2514	7,8147	9,3484	11,345	12,838
4	0,207	0,2971	0,4844	0,7107	1,0636	1,9226	3,3567	5,3853	7,7794	9,4877	11,143	13,277	14,860
5	0,4118	0,5543	0,8312	1,1455	1,6103	2,6746	4,3515	6,6257	9,2363	11,071	12,833	15,086	16,750
6	0,6757	0,8721	1,2373	1,6354	2,2041	3,4546	5,3481	7,8408	10,645	12,592	14,449	16,812	18,548
7	0,9893	1,2390	1,6899	2,1673	2,8331	4,2549	6,3458	9,0371	12,017	14,067	16,013	18,475	20,278
8	1,3444	1,6465	2,1797	2,7326	3,4895	5,0706	7,3441	10,219	13,362	15,507	17,535	20,090	21,955
9	1,7349	2,0879	2,7004	3,3251	4,1682	5,8988	8,3428	11,389	14,684	16,919	19,023	21,666	23,589
10	2,1558	2,5582	3,247	3,9403	4,8652	6,7372	9,3418	12,549	15,987	18,307	20,483	23,209	25,188
11	2,6032	3,0535	3,8157	4,5748	5,5778	7,5841	10,341	13,701	17,275	19,675	21,920	24,725	26,757
12	3,0738	3,5706	4,4038	5,2260	6,3038	8,4384	11,340	14,845	18,549	21,026	23,337	26,217	28,300
13	3,5650	4,1069	5,0087	5,8919	7,0415	9,2991	12,340	15,984	19,812	22,362	24,736	27,688	29,819
14	4,0747	4,6604	5,6287	6,5706	7,7895	10,165	13,339	17,117	21,064	23,685	26,119	29,141	31,319
15	4,6009	5,2294	6,2621	7,2609	8,5468	11,037	14,339	18,245	22,307	24,996	27,488	30,578	32,802
16	5,1422	5,8122	6,9077	7,9616	9,3122	11,912	15,339	19,370	23,542	26,296	28,845	32,000	34,267
17	5,6973	6,4077	7,5642	8,6718	10,085	12,792	16,338	20,489	24,769	27,587	30,191	33,409	35,718
18	6,2648	7,0149	8,2307	9,3904	10,865	13,675	17,338	21,605	25,989	28,869	31,526	34,805	37,156
19	6,8439	7,6327	8,9065	10,117	11,651	14,562	18,338	22,718	27,204	30,144	32,852	36,191	38,582
20	7,4338	8,2604	9,5908	10,851	12,443	15,452	19,337	23,828	28,412	31,410	34,170	37,566	39,997
21	8,0336	8,8972	10,283	11,591	13,240	16,344	20,337	24,935	29,615	32,671	35,479	38,932	41,401
22	8,6427	9,5425	10,982	12,338	14,042	17,240	21,337	26,039	30,813	33,925	36,781	40,289	42,796
23	9,2604	10,196	11,689	13,091	14,848	18,137	22,337	27,141	32,007	35,173	38,076	41,638	44,181
24	9,8862	10,856	12,401	13,848	15,659	19,037	23,337	28,241	33,196	36,415	39,364	42,980	45,558
25	10,520	11,524	13,120	14,611	16,473	19,940	24,337	29,339	34,382	37,653	40,647	44,314	46,928
26	11,160	12,198	13,844	15,379	17,292	20,843	25,337	30,435	35,563	38,885	41,923	45,642	48,290
27	11,808	12,879	14,573	16,151	18,114	21,749	26,336	31,528	36,741	40,113	43,195	46,963	49,645
28	12,461	13,565	15,308	16,928	18,939	22,657	27,336	32,621	37,916	41,337	44,461	48,278	50,994
29	13,121	14,256	16,047	17,708	19,768	23,567	28,336	33,711	39,088	42,557	45,722	49,588	52,336
30	13,787	14,954	16,791	18,493	20,599	24,478	29,336	34,800	40,256	43,773	46,979	50,892	53,672
40	20,707	22,164	24,433	26,509	29,051	33,660	39,335	45,616	51,805	55,759	59,342	63,691	66,766
50	27,991	29,707	32,357	34,764	37,689	42,942	49,335	56,334	63,167	67,505	71,420	76,154	79,490
60	35,534	37,485	40,482	43,188	46,459	52,294	59,335	66,982	74,397	79,082	83,298	88,379	91,952
70	43,275	45,442	48,758	51,739	55,329	61,698	69,335	77,577	85,527	90,531	95,023	100,43	104,21
80	51,172	53,540	57,153	60,392	64,278	71,145	79,334	88,130	96,578	101,88	106,63	112,33	116,32
90	59,196	61,754	65,647	69,126	73,291	80,625	89,334	98,650	107,57	113,15	118,14	124,12	128,30
100	67,328	70,065	74,222	77,929	82,358	90,133	99,334	109,14	118,50	124,34	129,56	135,81	140,17
120	83,852	86,923	91,573	95,705	100,62	109,22	119,33	130,05	140,23	146,57	152,21	158,95	163,65
140	100,66	104,03	109,14	113,67	119,03	128,38	139,34	150,89	161,83	168,61	174,65	181,84	186,85
p=	0,005	0,01	0,025	0,05	0,1	0,25	0,5	0,75	0,9	0,95	0,975	0,99	0,995

Table de la loi de Kolgomorov-Smirnov.

$\mathbb{P}\{\|F - F_n^*\|_\infty \leq D_n\} = p$
 (F est une fonction de répartition continue et
 F^* la fonction de répartition observée.)

D_n								
p=	0,80	0,9	0,95	0,975	0,98	0,99	0,995	=p
n								n
1	0,900000	0,950000	0,975000	0,987500	0,990000	0,995000	0,997500	1
2	0,683772	0,776393	0,841886	0,888197	0,900000	0,929289	0,950000	2
3	0,564810	0,636045	0,707598	0,767921	0,784557	0,829002	0,864279	3
4	0,492653	0,565216	0,623939	0,673881	0,68887	0,734238	0,776393	4
5	0,44698	0,509449	0,563275	0,612618	0,62718	0,668531	0,705431	5
6	0,410373	0,467993	0,519262	0,564043	0,577407	0,616607	0,652865	6
7	0,381476	0,436069	0,483424	0,525645	0,53844	0,575812	0,609753	7
8	0,358313	0,409623	0,454267	0,494464	0,506543	0,541793	0,574291	8
9	0,339102	0,387464	0,430011	0,468114	0,479596	0,513317	0,544431	9
10	0,322602	0,368663	0,409246	0,445617	0,456624	0,488932	0,518725	10
11	0,308292	0,352421	0,391224	0,426139	0,436703	0,467702	0,496386	11
12	0,295770	0,338151	0,375430	0,409019	0,419178	0,449045	0,476715	12
13	0,284698	0,325490	0,361432	0,393816	0,403621	0,432473	0,459212	13
14	0,274807	0,314170	0,348901	0,380207	0,389695	0,417616	0,443516	14
15	0,265886	0,303973	0,337596	0,367931	0,377127	0,404199	0,429336	15
16	0,257784	0,294720	0,327334	0,356780	0,365709	0,392007	0,416439	16
17	0,250387	0,286269	0,317963	0,346592	0,355275	0,380862	0,404643	17
18	0,243601	0,278511	0,309360	0,337235	0,345693	0,370622	0,393802	18
19	0,237346	0,271357	0,301425	0,328603	0,336852	0,361170	0,383792	19
20	0,231555	0,264734	0,294076	0,320607	0,328661	0,352411	0,374513	20
21	0,226173	0,258577	0,287243	0,313171	0,321044	0,344263	0,365878	21
22	0,221153	0,252835	0,280869	0,306233	0,313936	0,336659	0,357818	22
23	0,216455	0,247462	0,274905	0,299740	0,307283	0,329540	0,350270	23
24	0,212048	0,242420	0,269308	0,293645	0,301039	0,322857	0,343184	24
25	0,207902	0,237677	0,264042	0,287911	0,295163	0,316567	0,336512	25
26	0,203992	0,233205	0,259075	0,282502	0,289621	0,310633	0,330217	26
27	0,200297	0,228977	0,254381	0,277389	0,284381	0,305022	0,324265	27
28	0,196798	0,224974	0,249934	0,272545	0,279417	0,299707	0,318625	28
29	0,193478	0,221175	0,245715	0,267948	0,274706	0,294661	0,313270	29
30	0,190321	0,217563	0,241704	0,263578	0,270227	0,289864	0,308178	30
p=	0,80	0,9	0,95	0,975	0,98	0,99	0,995	=p

Table de la loi de Kolgomorov-Smirnov.

D_n								
p=	0,80	0,9	0,95	0,975	0,98	0,99	0,995	=p
n								n
31	0,187316	0,214125	0,237884	0,259416	0,265962	0,285295	0,303328	31
32	0,184450	0,210845	0,234241	0,255446	0,261893	0,280936	0,298701	32
33	0,181712	0,207713	0,230762	0,251654	0,258007	0,276772	0,294280	33
34	0,179094	0,204718	0,227434	0,248027	0,254290	0,272789	0,290051	34
35	0,176587	0,201849	0,224247	0,244554	0,250730	0,268974	0,286000	35
36	0,174183	0,199098	0,221191	0,241223	0,247316	0,265315	0,282115	36
37	0,171876	0,196458	0,218258	0,238026	0,244038	0,261803	0,278384	37
38	0,169659	0,193921	0,215439	0,234953	0,240889	0,258427	0,274799	38
39	0,167526	0,191480	0,212727	0,231997	0,237858	0,255179	0,271349	39
40	0,165472	0,189130	0,210115	0,229150	0,234940	0,252051	0,268027	40
41	0,163492	0,186865	0,207599	0,226406	0,232128	0,249036	0,264824	41
42	0,161582	0,184680	0,205171	0,223759	0,229414	0,246127	0,261734	42
43	0,159739	0,182570	0,202826	0,221203	0,226794	0,243319	0,258750	43
44	0,157957	0,180532	0,200561	0,218733	0,224263	0,240604	0,255866	44
45	0,156234	0,178560	0,198370	0,216345	0,221814	0,237979	0,253077	45
46	0,154567	0,176653	0,196251	0,214033	0,219445	0,235439	0,250378	46
47	0,152952	0,174805	0,194198	0,211795	0,217150	0,232978	0,247763	47
48	0,151388	0,173015	0,192208	0,209625	0,214926	0,230594	0,245229	48
49	0,14987	0,1712790	0,190279	0,207522	0,212769	0,228281	0,242772	49
50	0,148398	0,169594	0,188407	0,205480	0,210677	0,226037	0,240387	50
51	0,146969	0,167959	0,186589	0,203498	0,208645	0,223858	0,238072	51
52	0,145580	0,166370	0,184824	0,201573	0,206671	0,221742	0,235823	52
53	0,144231	0,164826	0,183107	0,199702	0,204752	0,219684	0,233636	53
54	0,142918	0,163324	0,181439	0,197881	0,202886	0,217683	0,231509	54
55	0,141641	0,161863	0,179815	0,196111	0,201071	0,215736	0,229440	55
56	0,140398	0,160441	0,178234	0,194387	0,199303	0,213840	0,227425	56
57	0,139187	0,159055	0,176694	0,192707	0,197582	0,211994	0,225463	57
58	0,138007	0,157705	0,175194	0,191071	0,195904	0,210195	0,223550	58
59	0,136857	0,156389	0,173731	0,189476	0,194269	0,208441	0,221686	59
60	0,135735	0,155106	0,172305	0,187921	0,192675	0,206731	0,219868	60
61	0,134641	0,153854	0,170914	0,186403	0,191119	0,205062	0,218094	61
62	0,133572	0,152632	0,169555	0,184922	0,189600	0,203433	0,216363	62
63	0,132529	0,151439	0,168229	0,183476	0,188117	0,201843	0,214672	63
64	0,131511	0,150273	0,166934	0,182063	0,186669	0,200289	0,213020	64
65	0,130515	0,149134	0,165668	0,180682	0,185253	0,198771	0,211406	65
66	0,129542	0,148021	0,164431	0,179333	0,183870	0,197287	0,209829	66
67	0,128591	0,146933	0,163221	0,178013	0,182517	0,195836	0,208286	67
68	0,127660	0,145868	0,162038	0,176723	0,181194	0,194416	0,206777	68
69	0,126749	0,144826	0,16088	0,1754600	0,179899	0,193027	0,205300	69
70	0,125858	0,143806	0,159747	0,174224	0,178632	0,191668	0,203855	70
n								n
p=	0,80	0,9	0,95	0,975	0,98	0,99	0,995	=p

Table de la loi de Kolgomorov-Smirnov.

D_n								
p=	0,80	0,9	0,95	0,975	0,98	0,99	0,995	=p
n								n
71	0,124985	0,142808	0,158637	0,173013	0,177391	0,190337	0,202440	71
72	0,124131	0,141830	0,157551	0,171828	0,176176	0,189033	0,201054	72
73	0,123293	0,140873	0,156486	0,170667	0,174985	0,187756	0,199696	73
74	0,122473	0,139934	0,155443	0,169529	0,173819	0,186505	0,198366	74
75	0,121669	0,139014	0,154421	0,168414	0,172675	0,185278	0,197062	75
76	0,120880	0,138112	0,153418	0,167321	0,171554	0,184075	0,195783	76
77	0,120107	0,137228	0,152435	0,166248	0,170455	0,182896	0,194529	77
78	0,119349	0,136360	0,151471	0,165197	0,169376	0,181739	0,193299	78
79	0,118604	0,135509	0,150525	0,164165	0,168318	0,180604	0,192092	79
80	0,117874	0,134673	0,149596	0,163152	0,167280	0,179490	0,190907	80
81	0,117157	0,133853	0,148685	0,162157	0,166260	0,178396	0,189744	81
82	0,116453	0,133048	0,14779	0,161181	0,1652590	0,177322	0,188602	82
83	0,115762	0,132257	0,146911	0,160222	0,164276	0,176268	0,187481	83
84	0,115082	0,131480	0,146047	0,159281	0,163311	0,175232	0,186379	84
85	0,114415	0,130717	0,145199	0,158355	0,162362	0,174214	0,185297	85
86	0,113759	0,129967	0,144365	0,157446	0,161430	0,173213	0,184234	86
87	0,113115	0,129230	0,143546	0,156552	0,160513	0,172230	0,183188	87
88	0,112481	0,128505	0,142741	0,155674	0,159612	0,171264	0,182160	88
89	0,111858	0,127792	0,141949	0,154810	0,158727	0,170313	0,181150	89
90	0,111245	0,127091	0,141170	0,153960	0,157855	0,169379	0,180156	90
91	0,110643	0,126402	0,140403	0,153124	0,156998	0,168459	0,179178	91
92	0,110050	0,125724	0,139650	0,152302	0,156155	0,167555	0,178216	92
93	0,109466	0,125056	0,138908	0,151493	0,155326	0,166665	0,177270	93
94	0,108892	0,124400	0,138178	0,150697	0,154509	0,165789	0,176338	94
95	0,108327	0,123753	0,137459	0,149913	0,153706	0,164926	0,175421	95
96	0,107770	0,123117	0,136752	0,149141	0,152914	0,164077	0,174518	96
97	0,107222	0,122490	0,136055	0,148381	0,152135	0,163242	0,173629	97
98	0,106683	0,121873	0,135369	0,147633	0,151368	0,162418	0,172754	98
99	0,106151	0,121265	0,134694	0,146896	0,150612	0,161607	0,171892	99
100	0,105627	0,120666	0,134028	0,146170	0,149868	0,160809	0,171042	100
n								n
p=	0,80	0,9	0,95	0,975	0,98	0,99	0,995	=p

Pour $n \geq 100$, $\mathbb{P}\{\sqrt{n} \|F - F_n^*\|_\infty \leq A\} = p$
 (F est une fonction de répartition continue et
 F_n^* la fonction de répartition observée.)

p=	0,80	0,9	0,95	0,975	0,98	0,99	0,995	=p
A=	1,0727	1,2238	1,3581	1,4802	1,5174	1,6276	1,7308	=A

Annexe B

Utilisation de l'étendue

Pour faire du contrôle de qualité, on utilise des échantillons de faible effectif. C'est alors l'étendue de l'échantillon qui est utilisée comme estimateur. Les règles appliquées par les spécialistes en productique sont souvent normalisées. Il s'agit ici d'expliquer l'origine de certaines de ces règles, et d'expliquer d'où viennent certains coefficients.

1 Estimateurs ponctuels de l'écart-type.

Considérons un échantillon aléatoire simple de taille n , de moyenne $\bar{x}$, d'écart-type s et d'étendue w , extrait d'une population normale de moyenne μ , d'écart-type σ .

1.1 Estimateur ponctuel usuel de l'écart-type.

Il s'agit de :

$$\hat{s} = s \cdot \sqrt{\frac{n}{n-1}}.$$

Cependant, lorsque n est faible, cette méthode n'est pas la plus utilisée.

1.2 Estimateur ponctuel de l'écart-type à partir de l'étendue.

Étendue d'un échantillon aléatoire simple issu d'une loi continue.

(Cette étude théorique n'est pas indispensable par la suite, elle sert essentiellement à permettre l'établissement des tables utilisées lors des tests.)

Considérons donc n variables aléatoires indépendantes $X_1, X_2, \dots, X_n$ de même loi continue de fonction de répartition F . Notons R_n son étendue, et posons $M_n = \sup_{1 \leq i \leq n} X_i$ et $m_n = \inf_{1 \leq i \leq n} X_i$.

On a donc $R_n = M_n - m_n$.

Or, en raison de l'indépendance des tirages

$$\mathbb{P}\{M_n \leq t\} = \mathbb{P}\left(\bigcap_{i=1}^n \{X_i \leq t\}\right) = \prod_{i=1}^n \mathbb{P}\{X_i \leq t\} = [F(t)]^n.$$

Donc la fonction de répartition de M_n est $F_{M_n} = F^n$.

De même,

$$\mathbb{P}\{m_n > t\} = \mathbb{P}\left(\bigcap_{i=1}^n \{X_i > t\}\right) = [1 - F(t)]^n$$

et la fonction de répartition de m_n est $F_{m_n} = 1 - (1 - F)^n$.

Enfin, pour $t \leq u$, on a :

$$\mathbb{P}(\{X_i > t\} \cap \{X_i \leq u\}) = F(u) - F(t),$$

ce qui fait que

$$\begin{aligned} \mathbb{P}(\{m_n > t\} \cap \{M_n \leq u\}) &= \mathbb{P}\left[\bigcap_{i=1}^n (\{X_i > t\} \cap \{X_i \leq u\})\right] \\ &= \prod_{i=1}^n \mathbb{P}(\{X_i > t\} \cap \{X_i \leq u\}) = [F(u) - F(t)]^n. \end{aligned}$$

La loi du couple (m_n, M_n) est ainsi donnée, pour $t \leq u$, par :

$$\begin{aligned} \mathbb{P}(\{m_n \leq t\} \cap \{M_n \leq u\}) &= \mathbb{P}\{M_n \leq u\} - \mathbb{P}(\{m_n > t\} \cap \{M_n \leq u\}) \\ &= [F(u)]^n - [F(u) - F(t)]^n. \end{aligned}$$

Cas où la loi des X_i a une densité.

Supposons maintenant que les X_i aient une densité f . On a donc : $F(x) = \int_{-\infty}^x f(t).dt$ et, en notant h_n la densité du couple (m_n, M_n) :

$$\mathbb{P}(\{m_n \leq t\} \cap \{M_n \leq u\}) = \begin{cases} \int_{-\infty}^u \left[\int_{-\infty}^t h_n(r, s).dr \right].ds & \text{si } t \leq u, \\ 0 & \text{si } t > u. \end{cases}$$

et, par conséquent, la densité h est portée par le demi-plan $\{(t, u) | t \leq u\}$ et

$$\begin{aligned} [F(u)]^n - [F(u) - F(t)]^n &= 1_{]-\infty; u]}(t) \cdot \int_{-\infty}^u \left[\int_{-\infty}^t h_n(r, s).dr \right].ds \\ &= 1_{[t; +\infty[}(u) \cdot \int_{-\infty}^u \left[\int_{-\infty}^t h_n(r, s).dr \right].ds, \end{aligned}$$

ce qui fait que :

$$\frac{\partial [F(u)]^n - [F(u) - F(t)]^n}{\partial u} = 1_{[t; +\infty[}(u) \cdot \int_{-\infty}^t h_n(r, u).dr = 1_{]-\infty; u]}(t) \cdot \int_{-\infty}^t h_n(r, u).dr$$

et que

$$\frac{\partial^2 [F(u)]^n - [F(u) - F(t)]^n}{\partial u \partial t} = 1_{]-\infty; u]}(t) \cdot h_n(t, u) = 1_{\{t \leq u\}}(t, u) \cdot h_n(t, u).$$

on obtient ainsi

$$\begin{aligned} h_n(t, u) &= 1_{\{t \leq u\}}(t, u) \cdot \frac{\partial^2 [F(u)]^n - [F(u) - F(t)]^n}{\partial u \partial t} \\ &= -1_{\{t \leq u\}}(t, u) \cdot n \cdot \frac{\partial [F(u) - F(t)]^{n-1} \cdot f(u)}{\partial u} \\ &= 1_{\{t \leq u\}}(t, u) \cdot n \cdot (n-1) \cdot [F(u) - F(t)]^{n-2} \cdot f(t) \cdot f(u) \end{aligned}$$

Choisissons $w > 0$ et recherchons $\mathbb{P}\{R_n \leq w\}$. On a :

$$\begin{aligned}\mathbb{P}\{R_n \leq w\} &= \int_{\mathbb{R}} \left(\int_u^{u+w} h_n(u, v) . dv \right) du = n(n-1) \int_{\mathbb{R}} \left(\int_u^{u+w} [F(v) - F(u)]^{n-2} f(v) . dv \right) f(u) . du \\ &= n. \int_{\mathbb{R}} \left[[F(v) - F(u)]^{n-1} \right]_{v=u}^{v=u+w} . f(u) . du = n. \int_{\mathbb{R}} [F(u+w) - F(u)]^{n-1} . f(u) . du\end{aligned}$$

donc la densité g_n de R_n vaut :

$$g_n(w) = \frac{d\mathbb{P}\{R_n \leq w\}}{dw} = n.(n-1). \int_{\mathbb{R}} [F(u+w) - F(u)]^{n-2} . f(u) . f(u+w) . du.$$

D'autre part, la densité de M_n est $n.[F(x)]^{n-1}.f(x)$ et celle de m_n est $n.[1 - F(x)]^{n-1}.f(x)$. Par conséquent,

$$\begin{aligned}\mathbb{E}(R_n) &= \mathbb{E}(M_n) - \mathbb{E}(m_n) = n. \int_{\mathbb{R}} x.[F(x)]^{n-1}.f(x).dx - n. \int_{\mathbb{R}} x.[1 - F(x)]^{n-1}.f(x).dx \\ &= n. \int_{\mathbb{R}} x. ([F(x)]^{n-1} - [1 - F(x)]^{n-1}) . f(x) . dx\end{aligned}$$

et, en effectuant une intégration par parties (en utilisant une primitive convenable, pour éviter une divergence intempestive) :

$$\mathbb{E}(R_n) = \left[x. ([F(x)]^n - [1 - [1 - F(x)]^n]) \right]_{x=-\infty}^{x=+\infty} - \int_{\mathbb{R}} [F(x)]^n - (1 - [1 - F(x)]^n) . dx$$

et, si la convergence de F vers 0 et 1 en $-\infty$ et $+\infty$ est assez rapide, on trouve :

$$\mathbb{E}(R_n) = \int_{\mathbb{R}} (1 - [F(x)]^n - [1 - F(x)]^n) . dx.$$

Pour un exposé plus complet, on pourra consulter : [Kendall M.G. & Stuart A. The advanced theory of statistics. Vol 1. Griffin & Co, London. 1958. p.338.] ou [Saporta G., Probabilités, analyse de données et statistique, Technip].

Le moment d'ordre 2 de R_n vérifie [Kendall p.339] :

$$\mathbb{E}[(R_n)^2] = 2. \int_{-\infty}^{\infty} \left(\int_{-\infty}^x [1 - F^n(x) - [1 - F(t)]^n + [F(x) - F(t)]^n] . dt \right) . dx.$$

Les formules établies ci-dessus permettent, avec $f(t) = \frac{1}{\sqrt{2.\pi}} e^{-\frac{t^2}{2}}$ d'établir la table suivante.

1.3 La table.

Loi de l'étendue dans le cas Gaussien.

Table de l'étendue q d'un échantillon aléatoire simple de taille n
extrait d'une population normale standard.

m_q est la moyenne de q ,

σ_q est l'écart-type de q ,

$V_q = \frac{\sigma_q}{m_q}$ le coefficient de variation de q .

Pour $0 < p < 1$, q_p est le quantile vérifiant $\mathbb{P}\{q \leq q_p\} = p$.

n	m_q	σ_q	V_q	$p =$ 0,005	$p =$ 0,01	$p =$ 0,025	$p =$ 0,05	$p =$ 0,5	$p =$ 0,95	$p =$ 0,975	$p =$ 0,99	$p =$ 0,995
				q_p	q_p	q_p	q_p	q_p	q_p	q_p	q_p	q_p
2	1,128	0,849	0,752	0,010	0,019	0,045	0,090	0,955	2,776	3,177	3,657	3,997
3	1,693	0,883	0,522	0,135	0,191	0,304	0,432	1,589	3,317	3,686	4,128	4,439
4	2,059	0,874	0,424	0,343	0,434	0,595	0,760	1,979	3,637	3,991	4,418	4,722
5	2,326	0,857	0,368	0,556	0,666	0,850	1,031	2,258	3,863	4,207	4,624	4,926
6	2,535	0,840	0,331	0,750	0,870	1,067	1,254	2,473	4,037	4,373	4,783	5,084
7	2,705	0,824	0,305	0,923	1,049	1,251	1,441	2,647	4,177	4,507	4,913	5,214
8	2,847	0,810	0,285	1,076	1,206	1,411	1,601	2,792	4,295	4,621	5,022	5,323
9	2,970	0,798	0,269	1,213	1,344	1,551	1,741	2,917	4,396	4,717	5,116	5,418
10	3,078	0,787	0,256	1,336	1,468	1,674	1,864	3,026	4,484	4,802	5,199	5,501
11	3,173	0,776	0,245	1,447	1,579	1,785	1,974	3,122	4,563	4,878	5,272	5,575
12	3,259	0,767	0,235	1,548	1,680	1,885	2,073	3,209	4,633	4,946	5,338	5,643
13	3,336	0,759	0,227	1,640	1,772	1,977	2,162	3,287	4,697	5,008	5,398	5,704
14	3,407	0,751	0,220	1,725	1,857	2,060	2,245	3,358	4,755	5,063	5,453	5,760
15	3,472	0,744	0,214	1,804	1,935	2,137	2,320	3,424	4,809	5,115	5,504	5,811
16	3,532	0,737	0,209	1,877	2,008	2,209	2,390	3,484	4,858	5,163	5,551	5,860
17	3,588	0,731	0,204	1,946	2,075	2,275	2,456	3,541	4,905	5,208	5,594	5,905
18	3,640	0,725	0,199	2,009	2,139	2,337	2,516	3,593	4,948	5,250	5,635	5,947
19	3,689	0,719	0,195	2,070	2,198	2,395	2,573	3,642	4,988	5,289	5,674	5,986
20	3,735	0,714	0,191	2,126	2,254	2,450	2,627	3,688	5,026	5,326	5,710	6,024

Un estimateur de l'écart-type d'une population normale dont est extrait un échantillon de taille n et d'étendue w est : $\frac{w}{m_q}$.

2 Intervalle de confiance sur l'écart-type.

Un intervalle de confiance sur l'écart-type σ d'une population normale de moyenne μ , dont on a extrait un échantillon aléatoire simple de taille n , de moyenne $\bar{x}$ et d'écart-type s est usuellement obtenu en utilisant le fait que $n \cdot \frac{s^2}{\sigma^2}$ suit une loi de χ^2 à $n - 1$ degrés de liberté. Mais, lorsque n est faible, on peut utiliser l'étendue et la table ci-dessus.

2.1 Intervalle de confiance de l'écart-type à partir de l'étendue.

Supposons ainsi que l'on connaisse l'étendue w d'un échantillon extrait d'une population normale d'écart-type inconnu σ .

L'étendue réduite de l'échantillon est $\frac{w}{\sigma}$. La loi de cette étendue réduite $\frac{w}{\sigma}$ est donnée par la table ci-dessus. Les quantiles $q_{\frac{\alpha}{2}}$ et $q_{1-\frac{\alpha}{2}}$ permettent d'en donner un intervalle de confiance au seuil de risque α .

Exemple. Considérons un échantillon (aléatoire simple, extrait d'une population normale d'écart-type σ) de taille $n = 6$ et dont l'étendue est $w = 10$. Recherchons un intervalle de confiance de σ .

L'étendue réduite $\frac{w}{\sigma}$ suit la loi ci-dessus (avec $n = 6$).

Alors au seuil de risque $1 - p = 0,05$, on a : $\mathbb{P} \left\{ \left[\frac{w}{q_{0,975}} \leq \sigma \leq \frac{w}{q_{0,025}} \right] \right\} = 0,95$ ce qui donne

$$\text{ici : } \frac{10}{4,373} \simeq 2,29 \leq \sigma \leq \frac{10}{1,067} \simeq 9,37.$$

2.2 Utilisation de l'étendue dans le cas où il y a au moins 15 données.

Lorsque n dépasse 10, cette méthode devient cependant moins efficace que l'estimation de σ utilisant la loi du χ^2 .

On peut toutefois encore utiliser la méthode précédente en considérant que l'on a k échantillons indépendants du même type, tous de même taille n comprise entre 5 et 10. Or, dès que $k \geq 3$, on montre que la moyenne $\bar{w}$ des k étendues observées, suit (approximativement) une loi normale de moyenne $\sigma \cdot m_q$ et de variance $\frac{1}{k} \cdot \sigma^2 \cdot \sigma_q^2$. On en déduit que l'intervalle de confiance au seuil de risque α est donné par :

$$\mathbb{P} \left\{ \frac{s}{1 + \frac{V_q}{\sqrt{k}} \cdot u_{1-\frac{\alpha}{2}}} \leq \sigma \leq \frac{s}{1 - \frac{V_q}{\sqrt{k}} \cdot u_{1-\frac{\alpha}{2}}} \right\} = 1 - \alpha$$

où $u_{1-\frac{\alpha}{2}}$ vérifie $\mathbb{P}\{\mathcal{N}(0,1) \leq u_{1-\frac{\alpha}{2}}\} = 1 - \frac{\alpha}{2}$ et $s = \frac{\bar{w}}{m_q}$ avec m_q lu dans la table (à la ligne n).

Remarquons que $s = \frac{\bar{w}}{m_q}$, (avec m_q lu dans la table à la ligne n) est un estimateur ponctuel de σ utilisable lorsque l'on dispose de k échantillons de taille n .

Exemple. Considérons un échantillon (aléatoire simple, extrait d'une population normale) de 18 individus partagé en $k = 3$ tronçons de $n = 6$ individus dont les étendues respectives sont $w_1 = 10$, $w_2 = 6$, $w_3 = 11$, donc $\bar{w} = 9$. Alors au seuil de confiance $p = 0,95$, avec $s = \frac{\bar{w}}{m_q} =$

$$\frac{9}{2,535} \simeq 3,550, \mathbb{P} \left\{ \frac{3,550}{1 + \frac{0,331}{\sqrt{3}}} \cdot 1,96 \leq \sigma \leq \frac{3,550}{1 - \frac{0,331}{\sqrt{3}}} \cdot 1,96 \right\} = 0,95 \text{ soit : } 5,84 \leq \sigma \leq 8,60.$$

3 Un exemple d'application, la construction d'une carte de contrôle.

Le but est de contrôler une production en prélevant des échantillons de n pièces produites et en mesurant une certaine caractéristique K sur chacune de ces pièces. A partir de ces mesures, il s'agit de détecter en temps réel des anomalies dans le processus de production. n est connu.

Pour établir la carte de contrôle, on commence par utiliser N échantillons de taille n . On se sert de ces échantillons pour établir un graphique formé d'une droite appelée ligne centrale et de deux droites parallèles appelés limites de contrôles (supérieure ou inférieure). La ligne centrale matérialise la moyenne de la caractéristique K testée, dans cette production. Lors des tests en temps réel de la production, la bande définie par les limites de contrôles est telle que, si un échantillon de n pièces voit sa moyenne x sortir de cette bande, ceci fait suspecter une anomalie dans la dite production.

Décidons de considérer comme suspecte toute valeur d'une série statistique suivant une loi normale $\mathcal{N}(\mu, \sigma)$ qui s'écarte de la moyenne d'au moins 3 écarts-types (donc avec la probabilité $1 - 0,9973 = 0,0027$).

3.1 Limites de contrôle sur la moyenne.

Si l'on effectue un prélèvement d'un échantillon (aléatoire simple) de taille n faible, disons $n \leq 10$, la moyenne $\bar{x}$ et l'étendue w de cet échantillon étant connues, la moyenne suit une loi normale de moyenne μ d'écart-type $\frac{\sigma}{\sqrt{n}}$.

Estimant σ par $\frac{w}{m_q}$, on considère donc que cet échantillon est aberrant lorsque

$$|\mu - \bar{x}| \geq 3 \cdot \frac{w}{m_q \cdot \sqrt{n}}.$$

Supposons ici que $n = 5$. On trouve alors $|\mu - \bar{x}| \geq \frac{3}{m_q} \cdot w = \frac{3}{2,326 \times \sqrt{5}} \cdot w \simeq 0,577 \times w$.

On dispose de $N = 20$ échantillons de taille $n = 5$. Chaque échantillon i (de taille 5) est de moyenne $\bar{x}_i$ et d'étendue $\bar{R}_i$. La moyenne des moyennes des 20 échantillons est $\bar{\bar{x}}$ et la moyenne de leurs étendues est $\bar{\bar{R}}$. La ligne centrale est alors définie par $\bar{\bar{x}}$ et les limites de contrôle par

$$[\bar{\bar{x}} - 0,577 \cdot \bar{\bar{R}}, \bar{\bar{x}} + 0,577 \cdot \bar{\bar{R}}].$$

3.2 Limites de contrôle sur l'étendue.

En ce qui concerne l'étendue w d'un échantillon de taille n , elle est également considérée comme aberrante lorsque qu'elle s'écarte de la moyenne R des étendues d'au moins 3 écarts-

types. Comme précédemment, on a :

$$\mathbb{P} \{ |w - R| > 3.\sigma(R) \} = 0.9973.$$

Comme w suit une loi de moyenne $R = m_q.\sigma$ d'écart-type $\sigma(R) = \sigma_q.\sigma$ où σ , l'écart-type de la population, supposée normale, est estimé par $\frac{R}{m_q}$ la relation précédente devient :

$$\mathbb{P} \left\{ |w - R| > 3.\frac{\sigma_q.R}{m_q} \right\} = 0.9973,$$

ce qui donne, puisque $V_q = \frac{\sigma_q}{m_q}$ est tabulé :

$$\mathbb{P} \{ (1 - 3.V_q) . R \leq w \leq (1 + 3.V_q) . R \} = 0.9973,$$

et, en fait :

$$\mathbb{P} \{ \max[0; (1 - 3.V_q)] . R \leq w \leq (1 + 3.V_q) . R \} = 0.9973,$$

Dans le cas $n = 5$ on trouve :

$$\max[0; (1 - 3.V_q)] = 0 \text{ et } 1 + 3.V_q = 1 + 3 \times 0,368 \simeq \underline{2,104}.$$

Ceci donne donc les limites de contrôle de l'étendue :

$$\left[0; 2,104 \times \overline{R} \right].$$

Annexe C

Activités Excel :

Dans cette activité, on construit une carte de contrôle, après avoir testé la normalité des données.

Les fichiers sont les suivants :

Fichier « Normalité.XLS » :

feuille Données,
feuille Étude de normalité1,
feuille Étude de normalité2,
feuille Histogramme,
feuille Comparaison,
feuille Droite de Henry,
feuille Kolgomorov.

Fichier « Carte.XLS » :

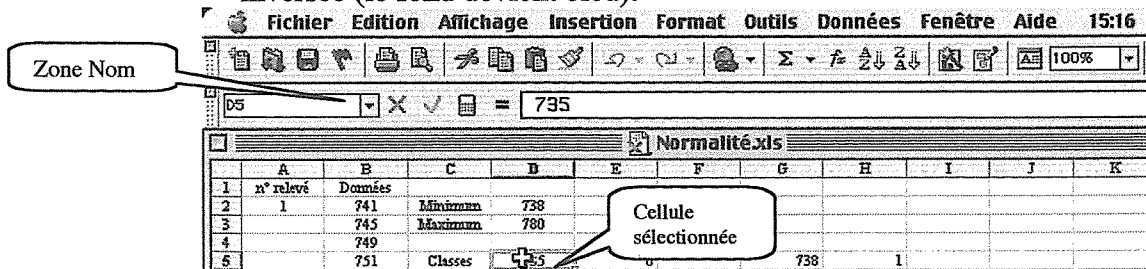
feuille Carte.

Les fichiers complets et à compléter sont à télécharger sur le site de l'IREM de Clermont-Ferrand, à l'adresse : "<http://www.maths.univ-bpclermont.fr/irem>". Les fichiers complets sont au format PDF, donc non modifiables. Les fichiers sur lesquels il est possible de travailler sont au format Excel.

A) Manipulations de base à l'aide de la souris :

1) Sélectionner à l'aide la souris :

On peut sélectionner à l'aide de la souris, une cellule en cliquant dessus, une plage de cellule en cliquant sur une extrémité et en déplaçant la souris en maintenant le bouton appuyé sur la souris, et en allant à l'autre extrémité de la plage de cellules qu'on cherche à sélectionner. On peut aussi sélectionner une ligne ou une colonne en cliquant sur le numéro de la ligne ou le nom de la colonne. Les cellules sélectionnées se mettent en vidéo inversée (le fond devient bleu).

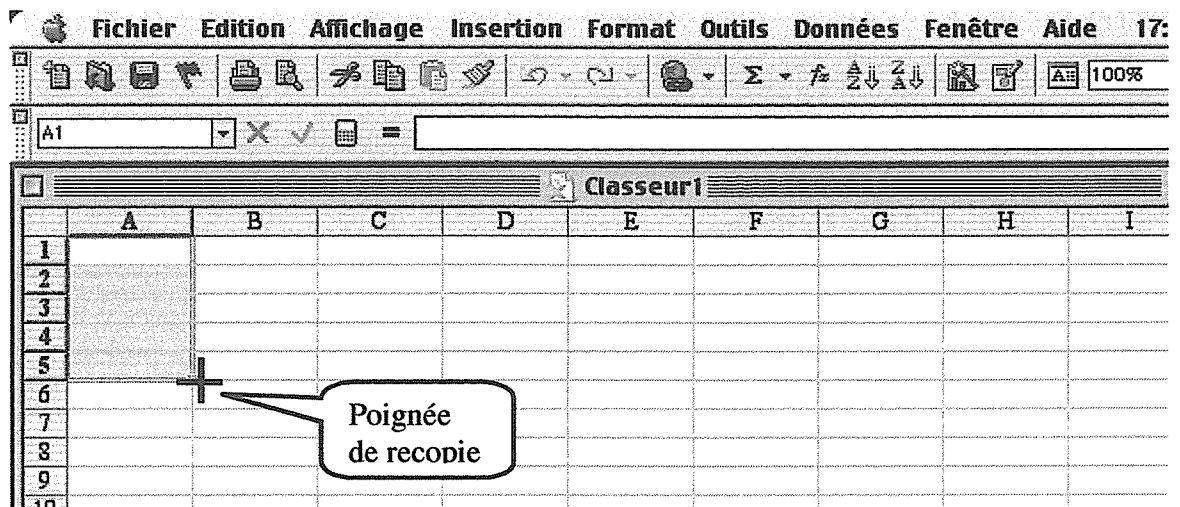


2) Renommer une cellule, une sélection ou une feuille :

Après avoir sélectionné ce qui nous intéresse, on peut lui donner un nom en tapant le nom souhaité dans la cellule " Zone Nom". On peut aussi en cliquant sur le nom de la feuille donner un nom aux feuilles sur lesquelles on travaille.

3) Recopier une cellule :

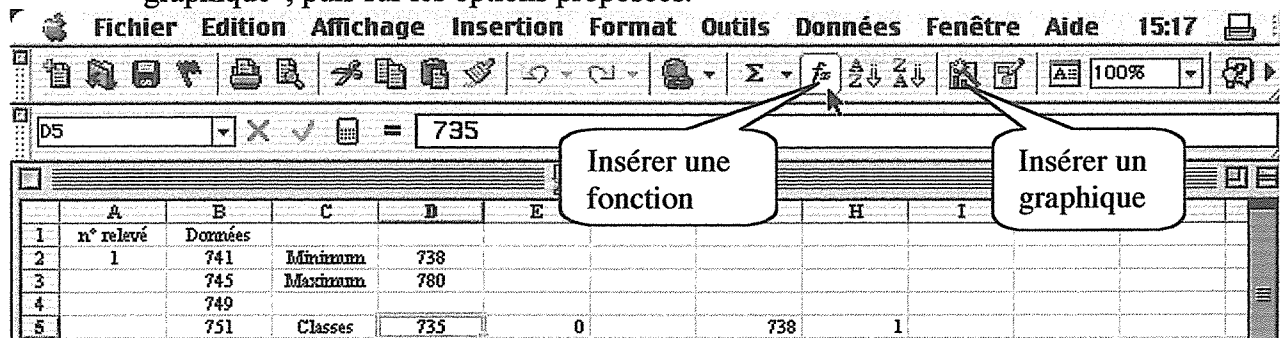
Après avoir activé la (ou les) cellule(s) à recopier, on utilise la poignée de recopie qui se trouve en bas à droite de la sélection (à ne pas confondre avec la sélection). Le curseur se transforme alors en croix.



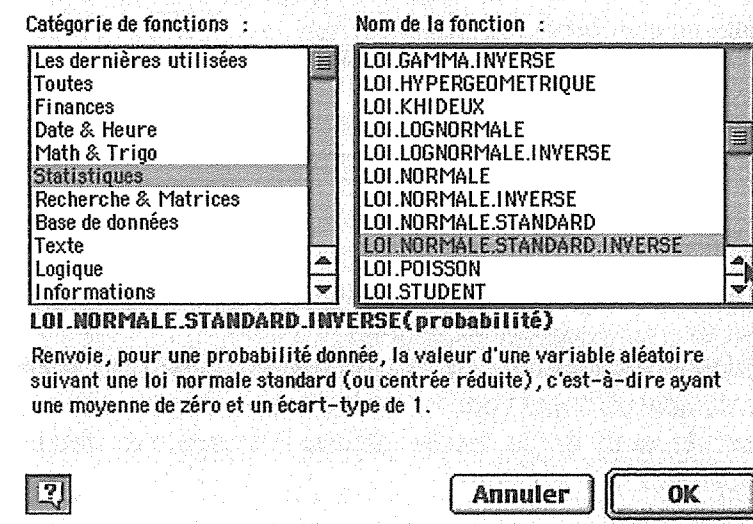
On peut recopier une valeur ou une formule.

4) Insérer une fonction, un graphique :

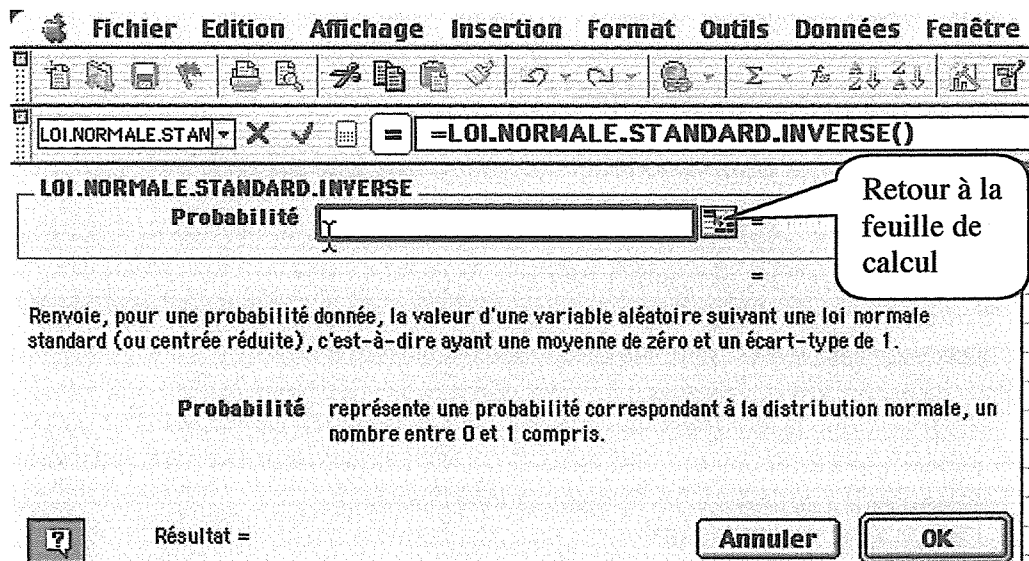
On clique sur l'icône "insérer une fonction" ou "insérer un graphique", puis sur les options proposées.



Par exemple, on obtient successivement les deux fenêtres suivantes lorsqu'on veut utiliser la loi standard normale inverse :



Dans chaque fenêtre apparaît une brève description de la fonction utilisée. Il est possible de taper directement le nom de la fonction, mais il faut connaître la syntaxe précise.



On peut, à l'aide de l'icône permettant le retour à la feuille de calcul, sélectionner les cellules qu'on souhaite utiliser comme argument dans la formule.

B) Fichier normalité.XLS :

1) La feuille "Données" :

On trouve les données pour l'étude de normalité dans les colonnes A et B de la feuille "Données" du fichier normalité.

On sélectionne la colonne B (en cliquant sur B) puis, à l'aide de la Zone Nom, on la nomme "Données". Dans les cellules E2 et E3, on entre les formules " $=\min(\text{Données})$ " et " $=\max(\text{Données})$ ". On peut entrer ces formules soit en les tapant directement au clavier en prenant soin de mettre en début de ligne le symbole "=" afin que Excel l'interprète comme une formule, soit utiliser l'icône "Insérer une fonction".

Nous allons regrouper les données en classes d'amplitude 5, en prenant pour première classe [735 ;740].

Cellule E6 et E7 : taper "735" et "740",

Sélectionner les cellules E6 et E7 puis recopier jusqu'à la cellule E15. (Excel a analysé comment passer de la cellule E6 à la cellule E7 et a appliqué la formule sur la zone voulue). Après avoir sélectionné la plage F6:F15, on tape la formule " $=\text{FREQUENCE}(\text{Données};\text{E6};\text{E15})$ " puis on appuie simultanément sur CTRL, MAJ puis Entrée. (Cette combinaison de touches est obligatoire ici car la fonction FREQUENCE utilisée est une fonction matricielle).

2) La feuille "Etude de normalité 1" :

Les centres de classe sont inscrits dans les cellules A3 à A12 et dans les cellules B3 à B12 on recopie les effectifs trouvés sur la feuille Données, à l'aide des commandes "Copier" et "Coller" du menu Edition.

Pour le tracé de la droite de Henry, on aura besoin des fréquences cumulées. Pour cela, on détermine les fréquences cumulées dans la colonne C :

Cellule C3 : taper " $= \text{B3}/100$ ",

Cellule C4 : taper " $=\text{C3}+\text{B4}/100$ ",.

Recopier la formule donnée en C4 jusqu'à la cellule C12.

On utilise ensuite la fonction "LOI.NORMALE.STANDARD.INVERSE". Il faut avoir côte à côte les valeurs données par la fonction et les maxima de chaque classe. (En effet, on se sert ici des Fréquences cumulées).

Cellules E3 et E4 : taper "740" et "745",

Sélectionner ces deux cellules, et recopier jusqu'à la cellule E12,

Activer la cellule F3,

Cliquer sur " Coller une fonction" et dans la catégorie de fonctions, sélectionner

"LOI.NORMALE.STANDARD.INVERSE",

Entrer pour probabilité C3,

Recopier vers le bas cette cellule jusqu'à la cellule F12.

On peut faire la comparaison graphique des densités de probabilité en comparant les résultats théoriques donnés par la loi normale et les valeurs données par l'échantillon:

Recopier les cellules E3:F12 aux cellules A23:B32,

Cellule C23 : taper "=500"

Cliquer sur " Coller une fonction" ;

dans la catégorie de fonctions, sélectionner "LOI.NORMALE"

moyenne : 755 ; écart type : 9,3 ; cumulative : 0.

(Pour obtenir un graphique comparable avec l'histogramme qui était construit à partir de 100 données avec une amplitude de classe de 5, il faut multiplier le résultat donné par la loi normale par 100*5.)

Recopier la cellule C23 vers le bas jusqu'à la cellule C32.

Remarque : il est possible d'obtenir de l'aide sur chaque fonction présente dans Excel .
(Menu "Aide", puis "Sommaire et index" , lettre "L" puis choix de "LOI.NORMALE").
Voici par exemple l'aide obtenue pour la fonction "LOI.NORMALE" :

LOI.NORMALE

Voir aussi

Renvoie la probabilité d'une variable aléatoire continue suivant une loi normale pour la moyenne et l'écart type spécifiés.
Cette fonction a de nombreuses applications en statistique, y compris dans les tests d'hypothèse.

Syntaxe

LOI.NORMALE(x;espérance;écart_type;cumulative)

x représente la valeur dont vous recherchez la distribution

espérance représente l'espérance mathématique de la distribution

écart_type représente l'écart-type de la distribution

cumulative représente une valeur logique déterminant le mode de calcul de la fonction : cumulatif ou non. Si l'argument cumulative est VRAI, la fonction LOI.NORMALE renvoie la probabilité suivant une loi normale pour qu'un événement aléatoire se reproduise un nombre de fois inférieur ou égal à x ; si l'argument cumulative est FAUX, la fonction renvoie la probabilité suivant une loi normale pour qu'un événement se reproduise x fois exactement.

Remarques

- Si les arguments espérance ou écart_type ne sont pas numériques, la fonction LOI.NORMALE renvoie la valeur d'erreur #VALEUR!
- Si l'argument écart_type ≤ 0, la fonction LOI.NORMALE renvoie la valeur d'erreur #NOMBRE!
- Si l'argument espérance = 0 et si l'argument écart_type = 1, la fonction LOI.NORMALE renvoie la distribution normale centrée réduite, la fonction LOI.NORMALE.STANDARD.
- L'équation de la fonction de densité normale est la suivante :

$$f(x, \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Exemple

LOI.NORMALE(42;40;1,5;VRAI) égale 0,908789

3) La feuille "Histogramme" :

À ce stade, on peut faire le tracé de l'histogramme à l'aide de l'assistant graphique : on sélectionne les cellules A3 à B12, puis on clique sur l'assistant graphique :

Type de graphique : Histogramme,

Données : les cellules sélectionnées,

Axe des abscisses : la première colonne

Afficher le graphique sur une nouvelle feuille qu'on appellera "Histogramme".

4) La feuille "Droite de Henry" :

Pour le tracé de la droite,

Cliquer sur l'icône "Assistant graphique",

Choix de la représentation : "nuage de points",

Plage de données : cellules E3:F12,

Série : on prend pour valeurs X les cellules E3:F12

Afficher le graphique sur une nouvelle feuille qu'on appellera "Droite de Henry".

Après la construction, on ajoute sur le graphique le coefficient de corrélation et l'équation de la droite d'Henry :

Sélectionner sur le graphique les points tracés,

Dans le menu graphique, cliquer sur "ajouter une courbe de tendance",

Type : linéaire ;

Option : on coche : Afficher l'équation sur le graphique,

Afficher le coefficient de corrélation sur le graphique.

Cliquer sur l'équation de droite obtenue,

On peut sélectionner ces résultats pour les recopier dans la feuille de calcul, dans les cellules A16 et A17, afin d'estimer la moyenne et l'écart type de la série.

5) La feuille "comparaison" :

Après avoir sélectionné les cellules A23 :C32, lancer l'assistant graphique :

Type : combinaison

Axe des abscisses : la première colonne

Afficher le graphique sur une nouvelle feuille qu'on appellera " Comparaison".

6) La feuille "Étude de normalité2" :

On sélectionne la colonne B de la feuille "Données" puis on la recopie à l'aide des commandes "Copier" et "Coller" du menu Edition sur les colonnes A et B successivement de la feuille "Étude de normalité2".

Après avoir sélectionné la colonne B, on trie les données :

Menu " Données",

Commande "Trier" ; Trier par : "ColonneB" ; Cocher l'option "croissant", Valider.

On détermine moyenne et estimation de l'écart type dans les cellules B104 à B106 :

Cellule B104 : taper " $=MOYENNE(B3:B102)$ ",

Cellule B105 : taper " $=ECARTYPE(B3:B102)$ ",

Cellule B106 : taper " $=B105*RACINE(100/99)$ ".

On a ensuite besoin des valeurs centrées réduites de la variable (notées x_i), des fréquences cumulées (notées f_i), ainsi que des valeurs $F(x_i)$, $|f_i - F(x_i)|$ et $|f_{i-1} - F(x_i)|$.

Pour cela :

Cellule C3 : taper "=(B3-\$B\$104)/\$B\$106" (la présence du symbole"\$" signalant une référence absolue, donc non modifiée lorsqu'on effectue une recopie de cellule),

Cellule D2 : taper "0,01",

Cellule D3 : taper "=D2+1/100",

Cellule E3 : taper "=LOI.NORMALE.STANDARD(C3)"

Cellule F3 : taper "=ABS(E3-D3)",

Cellule G3 : taper "=ABS(E3-D2)".

Sélectionner les cellules de C3 à G3, et les recopier jusqu'aux cellules C102:G102 à l'aide de la poignée de recopie.

Cellule K5 : taper "=MAX(F:F)",

Cellule K6 : taper "=MAX(G:G)",

Cellule K7 : taper "=MAX(K5:K6)".

C'est la valeur prise par la cellule K7 qu'on compare avec la valeur donnée par la cellule K9, lue dans une table de Kolgomorov. Cette valeur est recopiée ici, le test de Kolgomorov n'étant pas une fonction présente dans Excel.

7) La feuille "Kolgomorov" :

Cliquer sur l'icône "Assistant graphique",

Choix de la représentation : "nuage de points",

Sous type : "Nuage de points reliés sans marquage des données",

Plage de données : cellules C3:E102,

Axe des abscisses : la première colonne,

Titre : "Kolgomorov",

Afficher le graphique sur une nouvelle feuille qu'on appellera " Kolgomorov".

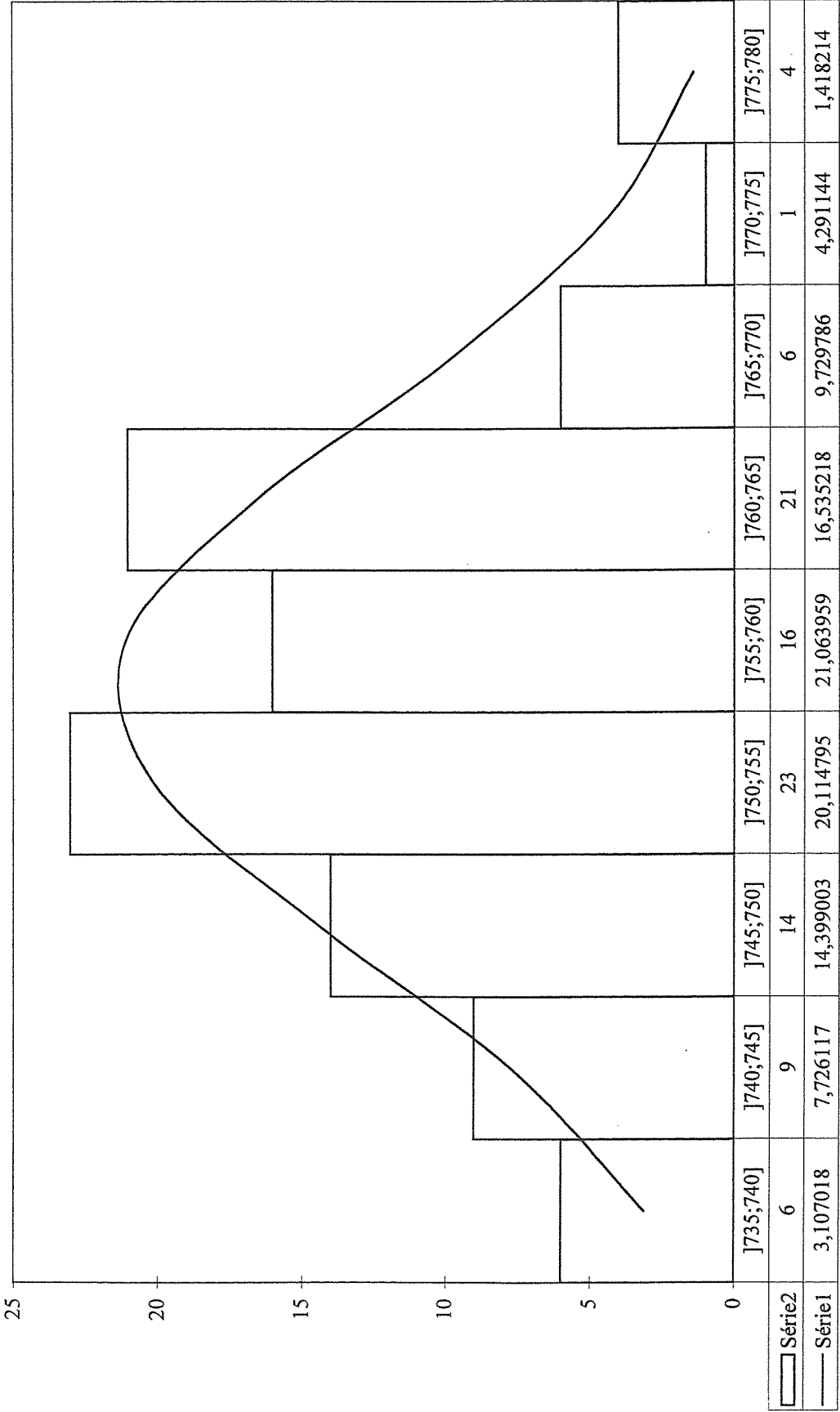
C) Fichier « Carte.XLS » :

La feuille Carte est à télécharger sur le site de l'IREM de Clermont-Ferrand.

n° relevé	Données	Minimum	738			
1	741		780			
	745	Maximum				
	749		Valeurs	Effectifs	Classes	
	751		738	1	735	0
	760		739	4	740	6
2	760		740	1	745	9
	755		741	3	750	14
	759		742	1	755	23
	745		743	0	760	16
	760		744	0	765	21
3	739		745	5	770	6
	760		746	1	775	1
	762		747	2	780	4
	759		748	1		
	761		749	3		
4	741		750	7		
	738		751	5		
	750		752	4		
	757		753	1		
	761		754	1		
5	766		755	12		
	752		756	1		
	751		757	3		
	749		758	0		
	755		759	4		
6	761		760	8		
	763		761	9		
	762		762	3		
	761		763	4		
	740		764	3		
7	741		765	2		
	762		766	4		
	759		767	0		
	761		768	0		
	766		769	1		
8	755		770	1		
	761		771	0		
	761		772	0		
	748		773	0		
	775		774	0		
9	776		775	1		
	750		776	1		
	755		777	1		
	779		778	0		
	746		779	1		
10	747		780	1		
	777			100		
	770					
	769					
	780					
11	750					
	755					
	752					
	750					
12	763					
	753					
	751					
	757					
	760					
	739					
13	752					
	753					
	751					
	749					
	764					
14	742					
	739					
	750					
	765					
15	765					
	754					
	763					
	766					
	750					
16	745					
	745					
	763					
	761					
17	760					
	756					
	760					
	759					
	757					
18	747					
	755					
	750					
	751					
	766					
19	752					
	755					
	755					
	755					
	755					
20	764					
	764					
	739					
	760					
	761					

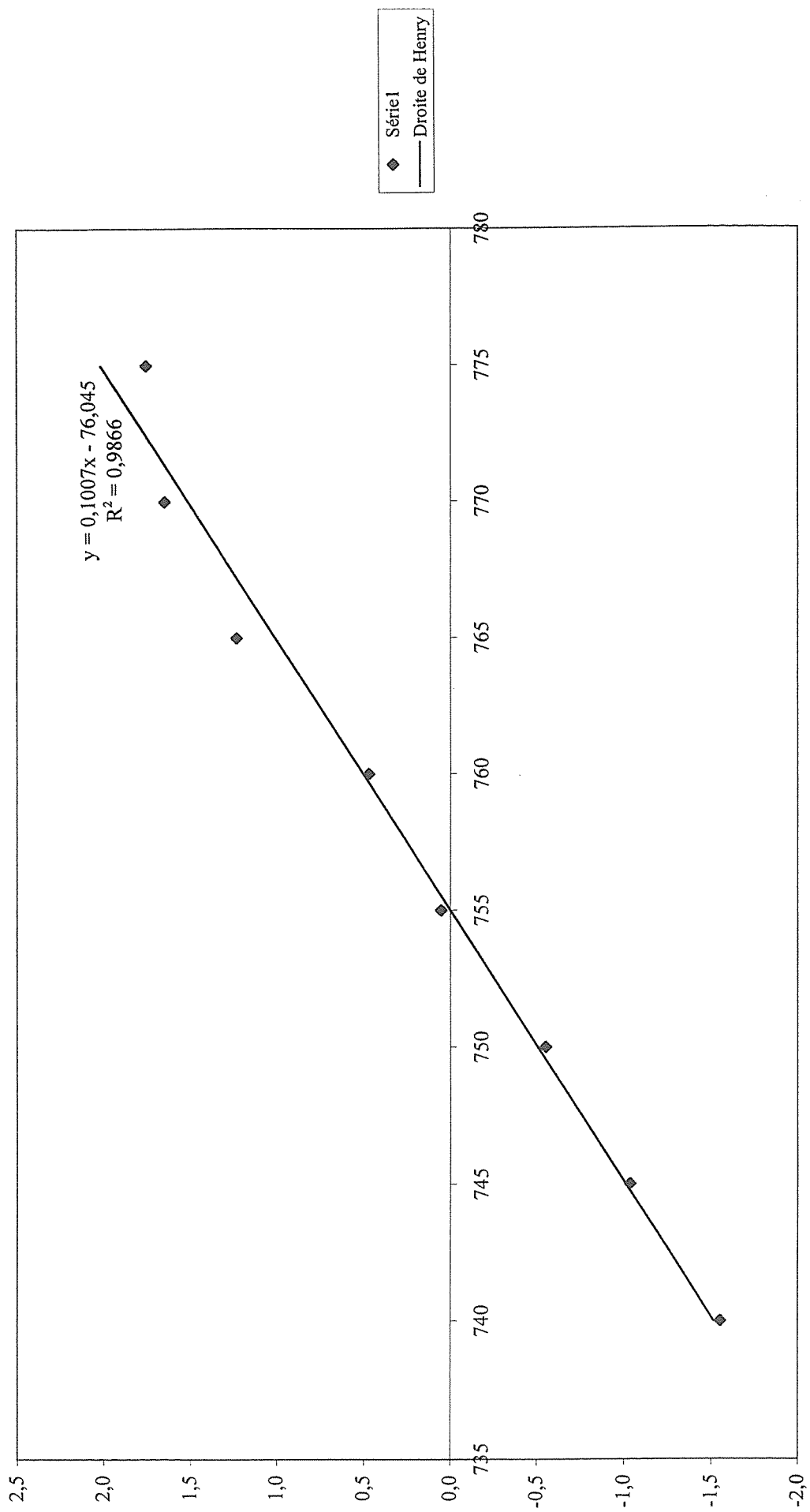
Première étape : Histogramme et étude de normalité						
Centres de classe	Effectifs	Fréquences cumulées	Centres cl.	Inv. loi Normale		
737,5	6	0,06	740	-1,554772098		
742,5	9	0,15	745	-1,036432877		
747,5	14	0,29	750	-0,553384325		
752,5	23	0,52	755	0,050154085		
757,5	16	0,68	760	0,467698555		
762,5	21	0,89	765	1,226528639		
767,5	6	0,95	770	1,644853		
772,5	1	0,96	775	1,750686351		
777,5	4	1	780	#NOMBRE!		
Voir histogramme						
Voir Droite de Henry						
y = 0,1007x - 76,045						
On en déduit :						
			m	755,164		
			σ	9,930		
Comparaison avec la loi normale de moyenne 755,8 et d'écart type 9,2 :						
Classes	Effectifs(O) ti		Loi normale (T)	Effectifs(T)	Regroupements	
					Observés	Théoriques
[735;740]	6	-2,260870	0,011884	3,107018	15	10,8331
[740;745]	9	-1,717391	0,042954	7,726117	14	14,3766
[745;750]	14	-0,631121	0,263981	14,376573	23	20,1335
[750;755]	23	-0,087051	0,465315	20,133465	16	21,0856
[755;760]	16	0,457018	0,676171	21,085574	21	16,5437
[760;765]	21	1,001088	0,841608	16,543679	6	9,7238
[765;770]	6	1,545158	0,938846	9,723804	5	5,6926
[770;775]	1	2,089227	0,981656	4,281050		
[775;780]	4	2,633297	0,995772	1,411551		
Voir Comparaison						
				Khi-deux :	5,9580	
				d.d.l. :	4	
				Lecture des tables :	9,48772846	
				Au seuil de 5%, l'observation est compatible avec une distribution normale.		

Comparaison

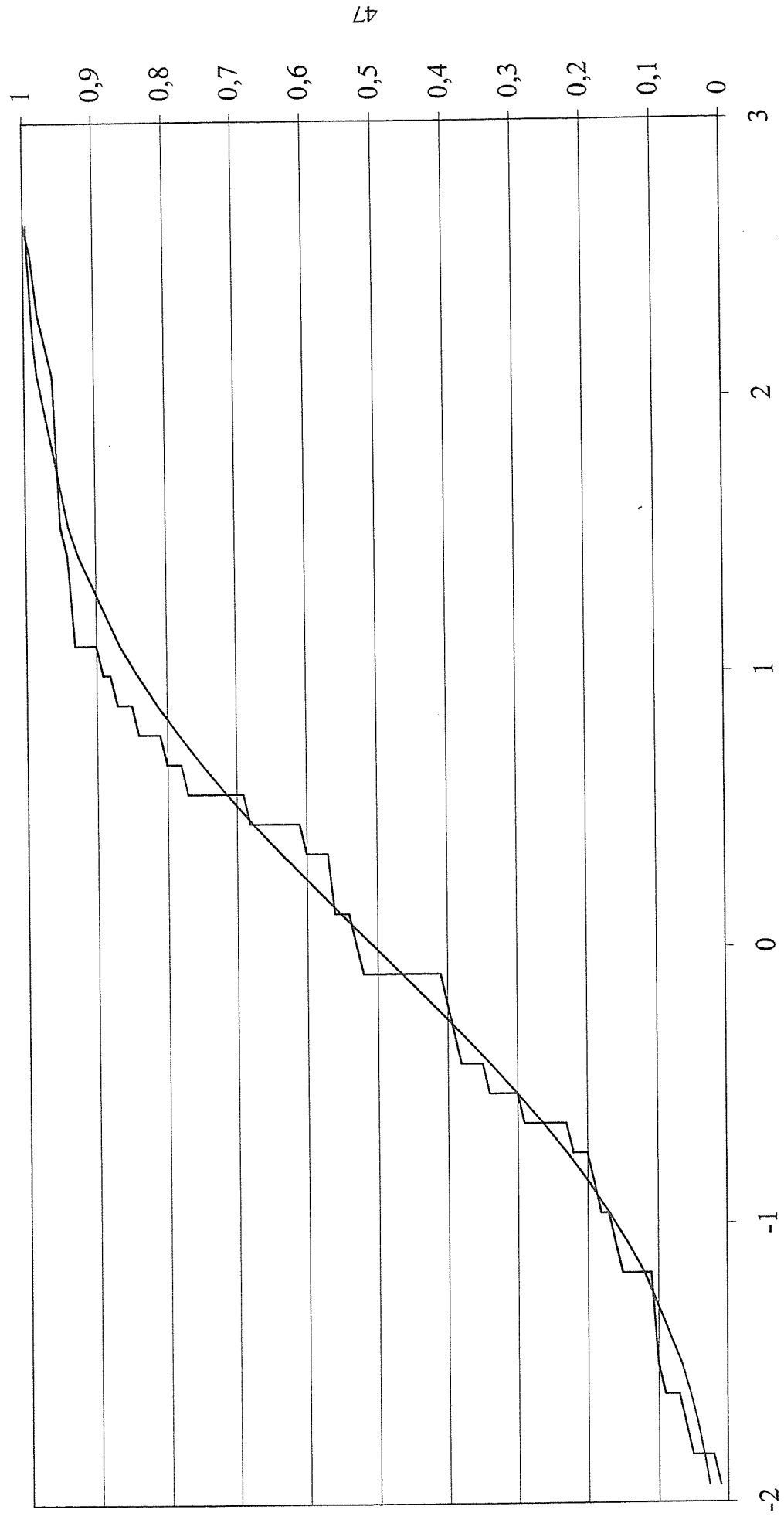


Série2
Série1

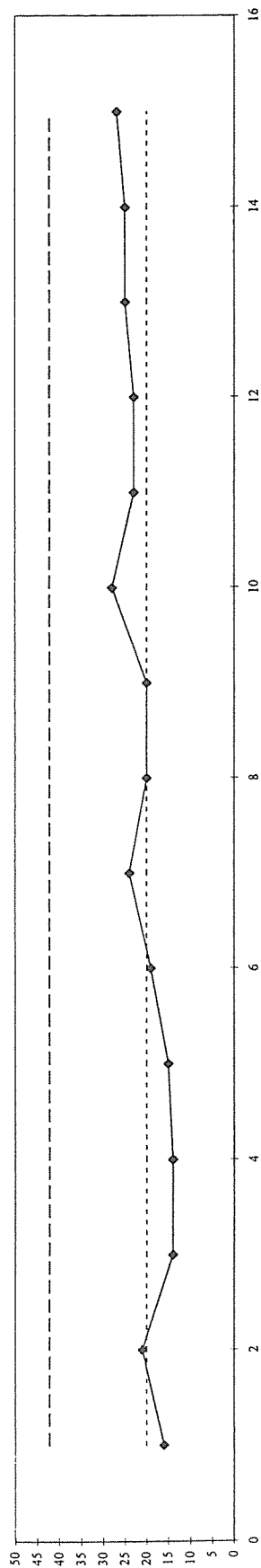
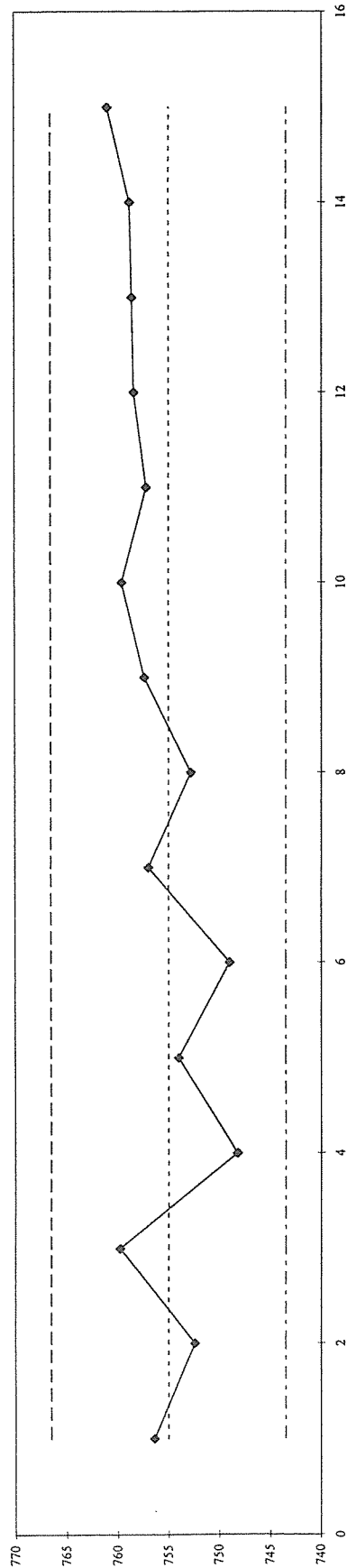
Droite de Henry

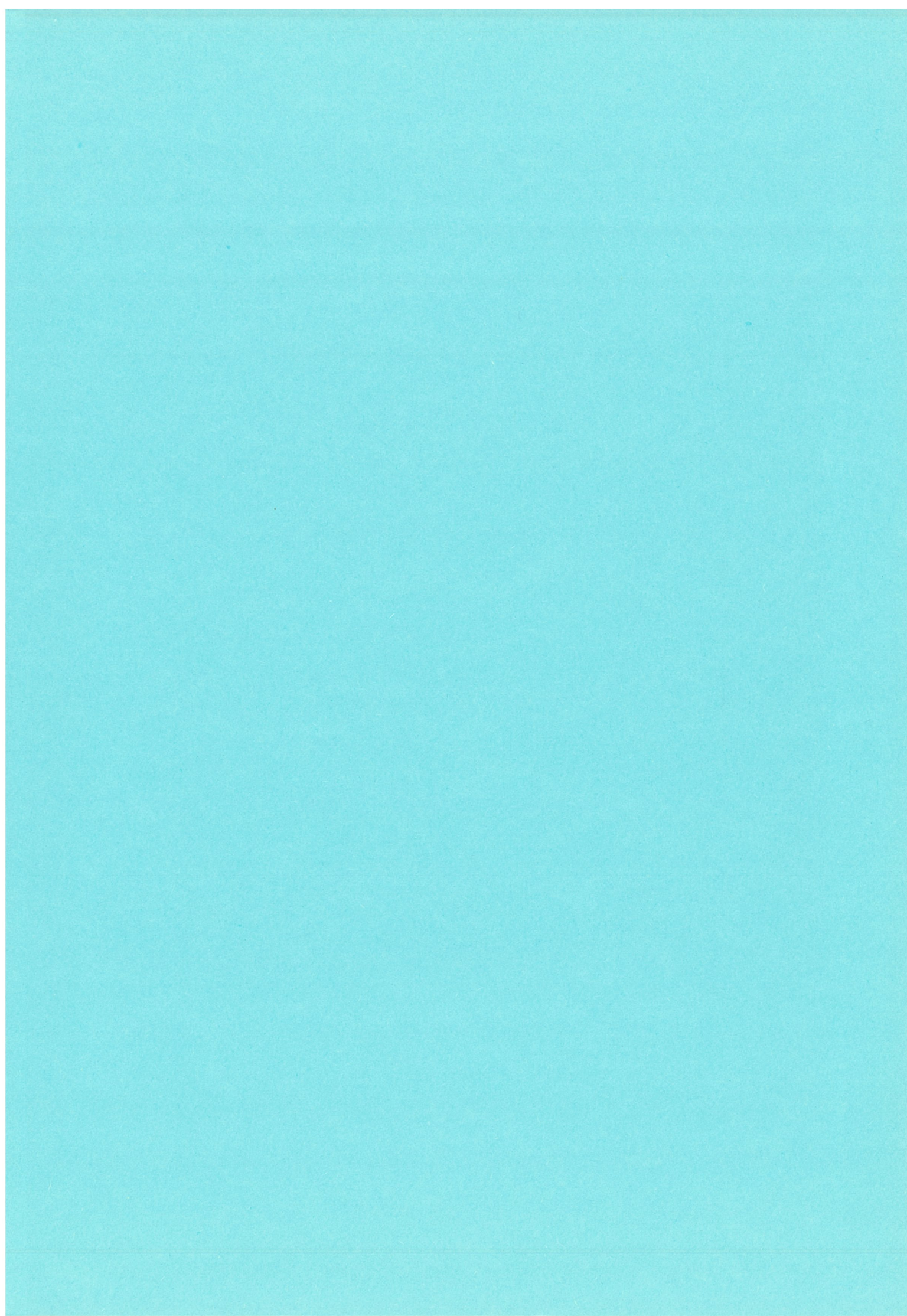


Test de Kolmogorov



N° relevé	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
x1	758	755	765	755	745	740	742	761	767	742	762	741	765	742	776
x2	757	751	755	751	750	744	756	761	764	762	760	762	765	763	749
x3	750	757	763	749	757	750	763	751	760	760	763	763	740	760	762
x4	751	760	751	745	758	752	758	750	749	770	761	764	761	762	762
x5	766	739	765	741	760	759	766	741	747	764	740	762	762	767	756
moyenne	756,4	752,4	759,8	748,2	754	749	757	752,8	757,4	759,6	757,2	758,4	758,6	758,8	761
étendue	16	21	14	14	15	19	24	20	20	28	23	23	25	25	27





TITRE : Une application industrielle des statistiques :
la carte de contrôle.

AUTEURS : Dominique ARBRE, Annette CORPART,
Gérard FLEURY, Nelly LASSALLE.

EDITEUR : IREM de CLERMONT-FERRAND.

DATE : MARS 2001.

PUBLIC CONCERNE : Enseignants de Lycée et de classe Post-Bac.

MOTS-CLES : Statistiques, test, contrôle de production.

RESUME : En s'appuyant sur l'étude d'une carte de contrôle, on
s'intéresse aux démarches statistiques nécessaires :

Processus, carte de contrôle, moyenne, étendue,
organigramme de mise en place d'une carte de contrôle, mise
sous contrôle du procédé, test non paramétrique, données,
étude de normalité, droite de Henry, test de Kolmogorov-
Smirnov, test du Khi-deux, table, utilisation de l'étendue.

Cette étude est illustrée par une activité tableur détaillée.