

REVUE DE STATISTIQUE APPLIQUÉE

E. DUCLOS

M. PILLET

A. COURTOIS

Optimisation de la maîtrise statistique des procédés par une méthode de filtrage d'ordre

Revue de statistique appliquée, tome 44, n° 2 (1996), p. 61-79

http://www.numdam.org/item?id=RSA_1996__44_2_61_0

© Société française de statistique, 1996, tous droits réservés.

L'accès aux archives de la revue « Revue de statistique appliquée » (<http://www.sfds.asso.fr/publicat/rsa.htm>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
<http://www.numdam.org/>

OPTIMISATION DE LA MAÎTRISE STATISTIQUE DES PROCÉDÉS PAR UNE MÉTHODE DE FILTRAGE D'ORDRE

E. Duclos¹, M. Pillet et A. Courtois

LLP/CESALP - BP 806
74016 Annecy Cedex - FRANCE
Tel (+33) 50.66.60.80 - Fax : (+33) 50.66.60.20
Email : duclos@esia.univ-savoie.fr

RÉSUMÉ

L'objet de cet article est de présenter une méthode de filtrage dite «d'ordre» permettant d'améliorer les performances des cartes de contrôle, en terme de Période Opérationnelle Moyenne (P.O.M.), lorsque la loi de répartition des mesures est symétrique mais non gaussienne. Nous avons étudié ici une famille de filtres d'ordre : les L-Filtres optimaux selon le critère du minimum de variance. Nous démontrons à l'aide de simulations informatiques que l'utilisation d'un filtre d'ordre, pour remplacer la moyenne opérée avec les cartes $\bar{X}/R$, apporte de nettes améliorations dans la détection d'un décentrage. Les filtres d'ordre, ont la particularité de prendre en compte la forme de la distribution des mesures. Pour ce faire, on calcule leurs coefficients, de façon à les rendre optimaux au sens du maximum de vraisemblance ou du minimum de variance des valeurs filtrées. Nous concluons cet article en proposant de nouvelles applications et en ouvrant de larges perspectives pour les méthodes de filtrage d'ordre en Maîtrise Statistique des Procédés.

Mots-clés : *Maîtrise Statistique des Procédés (M.S.P.), L-filtre Optimal, L-estimateurs, Filtrage d'ordre, Statistiques d'ordre, Filtrage optimal.*

SUMMARY

In this paper, we present a filtering method called «order filtering» to improve the Statistical Process Control (S.P.C.) performances in terms of average Run Length (A.R.L.), when measures are not normally distributed but symmetrical. We have studied a class of order filters : optimal L-filters in terms of the minimum variance of the output. We prove by computer simulation that using an order filter instead of the mean, usually used with $\bar{X}/R$ charts, increases significantly the rejection of common causes. Order filters have the particularity to take into account the real shape of measured distributions. It is then possible to calculate optimal coefficients according to a criteria of maximum likelihood or minimum variance of filtered measures. As a conclusion we propose new applications and describe large perspectives of order filtering applied to Statistical Process Control.

¹ Contrat CIFRE entre l'ANRT et ODS Europe SA - PAE des Glaisins-12 rue du Pré Paillard - 74940 Annecy le Vieux.

Keywords : *Statistical Process Control (S.P.C.), Optimal L-filter, L-estimators, Order filtering, Order statistics, Optimal filtering.*

1. Introduction

L'utilisation de cartes de contrôle ($\bar{X}/R$) en Maîtrise Statistique des Procédés (M.S.P.) nécessite l'hypothèse de normalité de la distribution des moyennes. Cette hypothèse, bien que très robuste, n'est pas toujours vérifiée. En effet, pour la satisfaire au mieux, il faut ou bien que la population soit distribuée selon une loi proche de la loi normale, ou bien que la taille de l'échantillon soit grande. La deuxième solution n'est malheureusement pas recevable pour des raisons d'économie de prélèvements. Par conséquent, si la population n'est pas gaussienne, la moyenne d'écart type $\sigma/\sqrt{n}$ (n étant la taille de l'échantillon) n'est plus une estimation à variance minimale [Cap 88]. Nous allons montrer dans ce qui suit, qu'il est possible de minimiser la variance de la distribution des valeurs estimées, ou filtrées, en utilisant une famille de statistiques d'ordre. Ceci dans le but d'améliorer la Période Opérationnelle Moyenne (P.O.M. ou A.R.L.) des cartes de contrôle, qui constitue un bon indicateur de performances de ces cartes.

Tout d'abord nous rappellerons les principes de base de la Maîtrise Statistique des Procédés (M.S.P.) pour l'élimination des causes communes. Ensuite nous présenterons une famille de statistiques d'ordre, les L-Estimeurs (aussi appelés L-Filtres en traitement du signal) que nous utiliserons pour améliorer les performances de la M.S.P. Nous justifierons les améliorations apportées par les calculs de P.O.M. pour des cas particuliers de distributions. Enfin, nous développerons quelques perspectives intéressantes d'applications des statistiques d'ordre à la Maîtrise Statistique des Procédés.

2. Présentation de la M.S.P.

Le principe de base de la M.S.P. consiste à détecter l'apparition d'une cause spéciale (ou assignable) à partir de données bruitées issues du procédé. Afin de dissocier les occurrences des causes spéciales de celles des causes communes, l'approche classique consiste à effectuer un contrôle sur un échantillon de n pièces consécutives et à estimer la valeur des paramètres de position (moyenne ou médiane) et d'échelle (écart type ou étendue) de la population. Parallèlement à l'opération d'estimation qui donne une image de l'évolution du procédé, nécessaire à sa bonne conduite, on opère un test d'hypothèse pour savoir si le procédé est ou n'est pas sous contrôle pour un risque donné. Toute estimation statistique, comporte des risques décisionnels α et β caractérisant l'efficacité de la méthode. Différentes approches ont d'ailleurs été proposées (CUSUM, EWMA,...) dans le but de minimiser ces risques. Pour de plus amples informations on se référera à [Pil 94] [Mon 90] [Rya 89].

Afin de comparer les performances de différentes cartes de contrôle et de connaître l'importance des deux risques décisionnels α et β on calcule leur Période Opérationnelle Moyenne (P.O.M.). Cet indicateur rend compte du temps moyen que prend une carte de contrôle pour détecter que le procédé est hors des limites d'intervention.

Il est important de rappeler les notions de risque α et β pour bien comprendre l'interprétation des P.O.M et l'intérêt des L-estimateurs ou L-filtres par rapport à la moyenne.

Lorsqu'on pilote un procédé avec une méthode statistique, il existe deux risques décisionnels que l'on essaye de minimiser :

- Le risque de conclure à un décentrage du procédé alors que celui-ci est parfaitement centré, aussi appelé risque de première espèce ou risque α (fig 1);

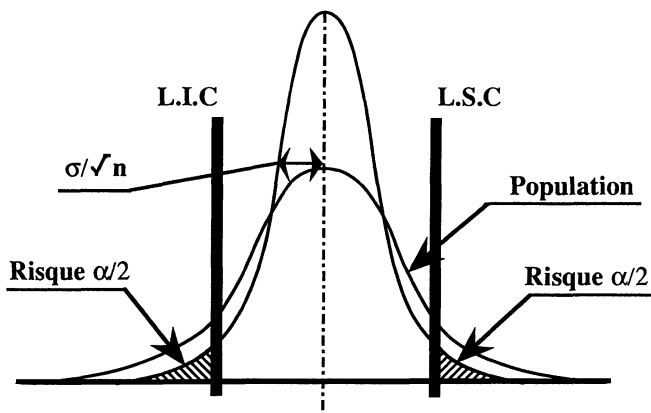


FIGURE 1 : $POM_{k=0} = \frac{1}{\alpha}$

- Le risque de ne pas détecter un décentrage alors que celui-ci existe, aussi appelé risque de seconde espèce ou risque β (fig 2).

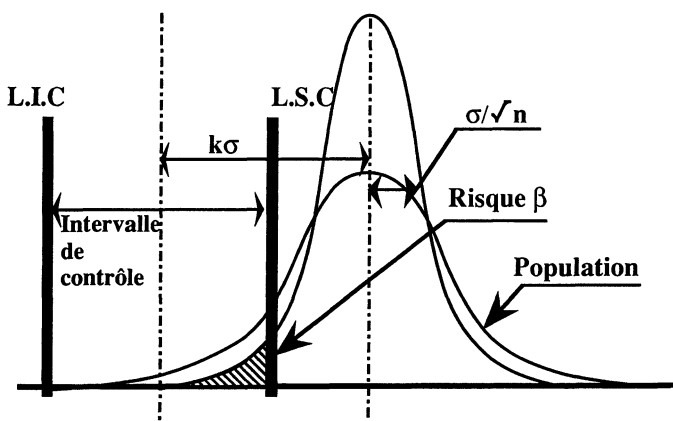


FIGURE 2 : $POM_k = \frac{1}{1-\beta}$

Lorsque le procédé est centré ($k = 0$) la probabilité de détecter un état hors contrôle correspond au risque α ; la Période Opérationnelle Moyenne prend donc la valeur $POM_{k=0} = \frac{1}{\alpha}$. En revanche, lorsque le procédé est décentré, la probabilité de détecter ce décentrage est de $1 - \beta$ d'où $POM_k = \frac{1}{1 - \beta}$.

Ces risques conditionnent, l'efficacité d'une carte de contrôle et définissent ainsi sa capacité à détecter un dérèglement avec le plus d'acuité et de rapidité possible.

Les statistiques d'ordre que nous allons étudier plus loin ont la particularité de prendre en compte la forme de la distribution du bruit. Ceci leur permet d'avoir de meilleures performances que les filtres linéaires, tels que la moyenne, en terme de réduction de bruit lorsqu'il n'est pas gaussien. L'estimateur de position le plus utilisé est cependant la moyenne, pour la mise en oeuvre de la carte de Shewhart $\bar{X}/R$, avec l'hypothèse implicite de la normalité de la distribution.

Notre idée est de remplacer l'opérateur de moyenne par une statistique d'ordre pour minimiser les risques α et β et donc les erreurs d'estimation de la position du procédé lorsque la densité de probabilité du bruit est de forme différente. Les estimateurs étudiés sont construits selon le critère du minimum de variance de la sortie. On peut noter à ce propos que la statistique d'ordre optimale est la moyenne, lorsque les mesures sont distribuées selon une loi normale [Bov83].

3. Filtrage d'ordre

3.1. Introduction

Nous aborderons ce problème d'estimation d'un point de vue traitement du signal en parlant non pas de statistiques d'ordre mais de filtrage d'ordre. Le problème étudié est en effet analogue aux problèmes bien connus de filtrage optimal puisque nous essayons d'extraire d'un signal bruité (mesures sur le procédé) un signal informatif qui est la position réelle du procédé.

Les techniques de filtrage d'ordre ont initialement trouvé leurs applications en traitement numérique d'images. Ces opérateurs apparaissent comme des fonctions linéaires et non linéaires de statistiques d'ordre. Ils permettent d'obtenir de meilleures performances que les filtres linéaires en termes de réduction de bruit et de préservation des transitions du signal informatif (estimation de la position du procédé). Bien qu'étudiés dans ce contexte, les opérateurs de filtrage d'ordre peuvent être utilisés sur des signaux monodimensionnels entachés de bruits non gaussiens. C'est pourquoi nous allons les utiliser en M.S.P.

3.2. Notions de filtrage d'ordre

Soit n le nombre de mesures (taille du filtre). Pour des raisons de symétrie, n est généralement impair. On désigne par $x_i (i = 1 \dots n)$ les mesures réalisées à l'instant k , et X_k est l'ensemble des mesures effectuées à ce même instant k .

On effectue localement un tri dans l'ordre croissant des mesures disponibles (x_i) dans le vecteur X_k . On appelle $j^{\text{ème}}$ statistique d'ordre ou statistique d'ordre j ,

la valeur ayant le rang j dans le vecteur ordonné X_k .

$$X_k = \{x(1), \dots, x(j), \dots, x(n)\}$$

avec $\min = x(1) \leq \dots \leq x(j) \leq x(j+1) \leq \dots \leq x(n) = \max$

On peut distinguer deux catégories de filtres parmi les filtres d'ordre : les L-Filtres et les NL-Filtres. Ils se différencient par la façon de pondérer les statistiques d'ordre pour calculer la sortie du filtre. La sortie Y_k d'un L-filtre est alors donnée par une combinaison linéaire des statistiques d'ordre avec des coefficients réels notés $C_j(j = \{1 \dots n\})$.

$$Y_k = \sum_{j=1}^n C_j \cdot x(j)$$

Si on généralise cette notation en définissant un opérateur à l'aide d'une fonction réelle ϕ , on obtient des NL-filtres; la fonction ϕ ayant éventuellement un caractère non linéaire (NL).

$$\phi : \mathbb{R}^n \rightarrow \mathbb{R}$$

$$\text{telle que : } Y_k = \phi(x(1) \dots x(n))$$

Dans notre étude, nous ne nous intéresserons qu'aux L-filtres dont voici quelques exemples :

- filtre milieu $C_1 = C_n = 1/2$ et $C_j = 0$ pour j de 2 à $n-1$, ce qui revient à prendre la moyenne des valeurs extrêmes.
- filtre moyennneur $C_j = 1/n$ pour j de 1 à n , ce qui revient à prendre la moyenne.
- filtre médian $C_m = 1$ et $C_j = 0$ avec $m = (n+1)/2$ et $j \neq m$, ce qui revient à prendre la médiane.

3.3. Optimisation des coefficients des L-Filtres

L'intérêt des L-Filtres est de pouvoir déterminer un jeu de coefficients optimaux au sens de l'erreur quadratique de sortie, en tenant compte de la forme de la distribution du bruit. On cherche donc à déterminer un estimateur non biaisé à variance minimale de la position du procédé.

Si le signal d'entrée du filtre est stationnaire, nous pouvons le modéliser par : $X_k = s_k + B_k$, où s_k est une constante (position du procédé) et B_k un bruit blanc centré ayant une densité de probabilité symétrique. Pour un ensemble de mesures consécutives, les statistiques d'ordre vérifient :

$$x(j) = s_k + B(j)$$

où $B(j)$ est la $j^{\text{ème}}$ statistique d'ordre du bruit.

On en déduit la sortie Y_k :

$$Y_k = s_k \cdot \sum C_j + \sum C_j \cdot B(j)$$

On a alors :

$$E[Y_k] = s_k \cdot \sum_j C_j + \sum_j C_j \cdot E[B(j)] = s_k \cdot \sum_j C_j$$

Du fait de la symétrie de la distribution du bruit, le filtre optimal est symétrique :

$$C_{1+j} = C_{n-j} \quad \forall j \in Z$$

Une condition suffisante pour obtenir un biais nul est :

$$E[Y_k] = s_k \Rightarrow \sum C_j = 1$$

La variance du filtre devient

$$\text{Var}[Y_k] = \sum \sum C_j \cdot E[B(j)B(p)] \cdot C_p = {}^t C \cdot R \cdot C \quad (1)$$

Où ${}^t C = (C_1, \dots, C_n)$ est le vecteur coefficient et $R = (R_{jp})$ est la matrice de covariance du bruit ordonné :

$$R_{jp} = E[B(j) \cdot B(p)]$$

Par une technique de multiplicateurs de Lagrange, on obtient la solution optimale :

$$C_{\text{opt}} = \frac{R^{-1} \cdot e}{{}^t e \cdot R^{-1} \cdot e} \quad \text{avec} \quad {}^t e = (1 \dots 1) \quad (2)$$

Le calcul du jeu optimal des coefficients est également possible pour des distributions d'entrées non symétriques.

3.4. Stationnarité des données

Nous nous placerons toujours dans l'hypothèse d'un signal utile stationnaire au sein d'un échantillon de mesures consécutives. Cette hypothèse est effectivement vérifiée lorsque ses variations sont très lentes par rapport à la période sur laquelle on effectue les mesures (ou le prélèvement de pièces dans le cas d'une production de pièces mécaniques en série).

Dans le cas d'échantillons de bruit dépendants ou dans celui de signaux non stationnaires, le filtrage d'ordre ne conduit pas en général aux solutions optimales. Celles-ci sont néanmoins très robustes [Bol 93].

D'un point de vue pratique, il est difficile de connaître avec exactitude la distribution statistique du signal d'entrée (distribution instantanée des mesures). Cependant le problème n'est pas critique si on peut avoir une idée même approximative de cette loi. En effet, d'après la relation (1), la variance de sortie est une fonction continue du vecteur des coefficients C , et si C est dans le voisinage du vecteur optimal, la variance de sortie sera proche de la variance minimale.

4. Résultats de simulation

Pour pouvoir apprécier les améliorations apportées par les L-Filtres vis-à-vis d'une carte de contrôle standard de type Shewhart, nous avons calculé les Périodes Opérationnelles Moyennes de deux types de cartes. La première carte est une carte de Shewhart «classique», tandis que la seconde consiste à appliquer un L-filtre à la place de l'opérateur moyenne dans une carte de Shewhart. Nous avons intitulé cette dernière procédure «*Carte de Shewhart avec L-filtre*».

Pour calculer la P.O.M. d'une carte de contrôle $\bar{X}/R$, il faut tout d'abord fixer les limites de contrôle. Celles-ci sont calculées pour un risque α donné, à partir de l'écart type théorique $\sigma/\sqrt{n}$ de la distribution des moyennes (σ étant l'écart type de la population et n la taille de l'échantillon). Les limites sont généralement fixées à $\pm 3 \cdot \sigma/\sqrt{n}$, ce qui correspond à un risque α de 0,27% pour une loi normale.

Afin d'étudier les risques engendrés par chaque estimateur, on compare leur P.O.M. pour un décentrage nul (risque α) et pour un décentrage $k \cdot \sigma$ (risque β). D'un point de vue pratique, on cherche la carte de contrôle qui va minimiser les risques α et β , c'est-à-dire que l'on attend une P.O.M. la plus grande possible lorsque le décentrage k est égal à zéro et une P.O.M. qui tend vers 1 le plus rapidement possible lorsque le décentrage k est non nul.

Pour les deux types de cartes étudiées, les calculs de P.O.M. ont été obtenus par simulation informatique. Pour ce faire nous avons, dans chacun des deux cas, généré deux types de distributions (fig 3) : une distribution uniforme et une distribution triangulaire; toutes deux centrées et d'écart type $\sigma = 1$.

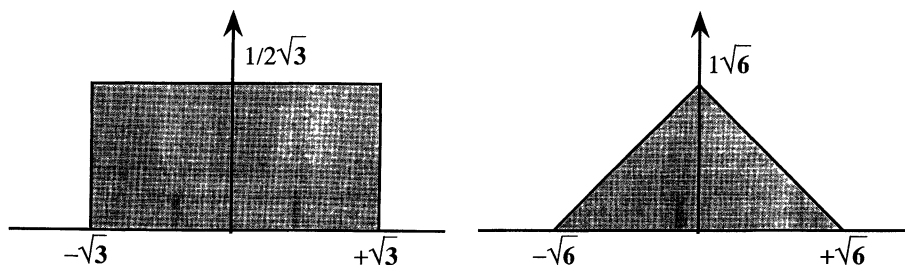


FIGURE 3 :

Distributions uniforme et triangulaire générées pour les simulations

L'intérêt de ces expériences est de comparer les améliorations apportées par les L-filtres pour une distribution très proche de la loi normale - densité de probabilité (ddp) triangulaire - et pour une autre très différente - ddp uniforme. Les résultats sont présentés sous formes de graphiques regroupant les P.O.M. des deux cartes de contrôle, calculées pour une taille d'échantillon donnée. Ceci permet de confronter les résultats obtenus pour les mêmes limites de contrôle quelle que soit la forme de la population. Il aurait bien sûr été possible d'établir les formules exactes des P.O.M. et en l'occurrence pour la carte de Shewhart, mais les calculs deviennent vite inextricables lorsqu'il s'agit de lois non gaussiennes et à fortiori lorsque n est supérieur à 3.

Pour atteindre une précision satisfaisante des résultats, nous comptabilisons le nombre de réalisations nécessaires pour obtenir 5000 détections d'un état hors contrôle du procédé. C'est à dire que pour une P.O.M. = 650 nous avons simulé le prélèvement de $5000 * 650 = 3,2510^6$ échantillons de n pièces.

4.1. Cas d'une distribution uniforme

4.1.1. L-Filtre Optimal

Le but de nos travaux est d'intégrer les cartes de contrôle avec L-filtres à un logiciel industriel de S.P.C. Ceci explique pourquoi nous avons choisi de calculer les coefficients des L-filtres et d'évaluer leurs performances à partir de simulations informatiques. Nous supposons en effet, lorsque nous abordons un nouveau procédé, que l'expression mathématique de la loi de répartition des mesure n'est pas connue. Il faut donc être capable d'établir le L-filtre optimal à partir de la seule information que constituent les mesures.

Ainsi, lors des calculs, nous utilisons un algorithme de génération de nombres aléatoires pour simuler la dispersion du procédé. Ces données sont utilisées pour calculer la matrice R de covariance de bruit ordonné nécessaire à l'obtention des coefficients du L-filtre optimal. Chaque élément de la matrice est calculé de la manière

suivante : $R_{i,j} = \frac{1}{m} \sum_{k=1}^m (x(i) - \bar{x})(x(j) - \bar{x})$ où $\bar{x}$ est la moyenne des mesures

établie sur la période d'essais préliminaire à la mise en place de la carte de contrôle. Contrairement à ce que l'on pourrait penser, R est différente de $\sigma^2.I$ (où I est la matrice identité), malgré l'indépendance statistique des données. Ceci s'explique par l'opération de tri réalisée au sein de l'échantillon.

Les résultats obtenus sont d'une précision très acceptable puisque, par exemple, pour le L-filtre optimal de taille $n = 3$ nous obtenons les coefficients du filtre milieu, avec un erreur de l'ordre de 0,01%.

Nous pouvons, dans des cas particuliers de distributions, établir les coefficients du L-filtre optimal, au sens de la variance, en calculant les moments des statistiques d'ordre d'un échantillon. Nous montrons en annexe que les coefficients du L-filtre optimal, parmi les L-filtres, sont dans le cas d'une ddp uniforme, ceux du filtre milieu. Nous établissons également que la variance du filtre milieu évolue en $1/n^2$. Par souci de concision, nous avons limité les développements mathématiques, c'est pourquoi

nous invitons le lecteur à se référer aux ouvrages de [Dav 81][Tas 89] [Lec 87] où le problème est traité en détail.

4.1.2. La P.O.M.

Les résultats des calculs de la P.O.M. obtenus pour une distribution uniforme, montrent que la P.O.M. de la carte de Shewhart avec L-Filtre est toujours supérieure à celle de la carte de Shewhart, quelle que soit la valeur du décentrage k , si elle est inférieure à la limite de contrôle $3 \cdot \sigma / \sqrt{n}$. Les améliorations apportées sont principalement portées sur le risque α . Elles sont minimales lorsque n est petit (fig 4), cependant l'écart entre les P.O.M. associées à la carte de Shewhart et la carte de «Shewhart avec L-filtre» s'accroît avec la taille n du filtre (fig 5 et 6). Les L-filtres, et en l'occurrence ici le filtre milieu, permettent de minimiser le risque de fausses alertes sur l'intervalle $k \in [0, 3 \cdot \sigma / \sqrt{n}]$.

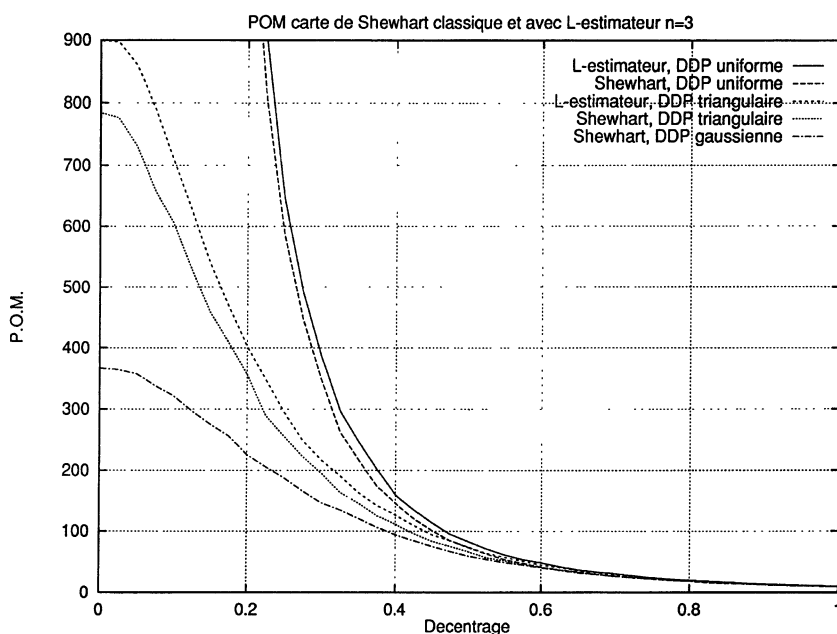


FIGURE 4 :

P.O.M. de la carte de Shewhart avec L-Filtre et de la carte de Shewhart Classique pour un décentrage k compris entre 0 et 1 – $n = 3$.

Pour $k = 3 \cdot \sigma / \sqrt{n}$, les courbes se croisent pour une valeur de la P.O.M. voisine de 2. La distribution, symétrique, se trouvant à cheval sur une limite de contrôle, le taux de détection est de 50% (fig 7 à 9). D'un point de vue numérique, les limites se situent respectivement en $k = 1,73; 1,34$ et $1,13$ pour $n = 3, 5$ et 7 .

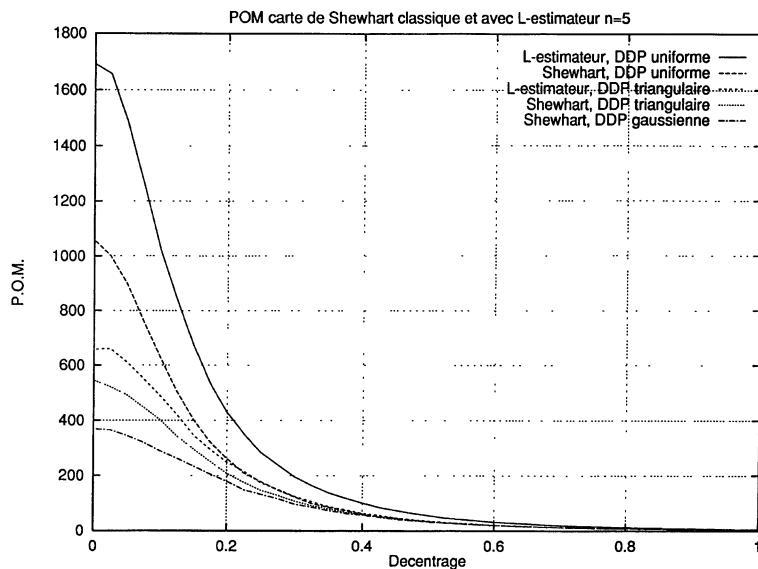


FIGURE 5 :

P.O.M. de la carte de Shewhart avec L-Filtre et de la carte de Shewhart Classique pour un décentrage k compris entre 0 et 1 – $n = 5$.

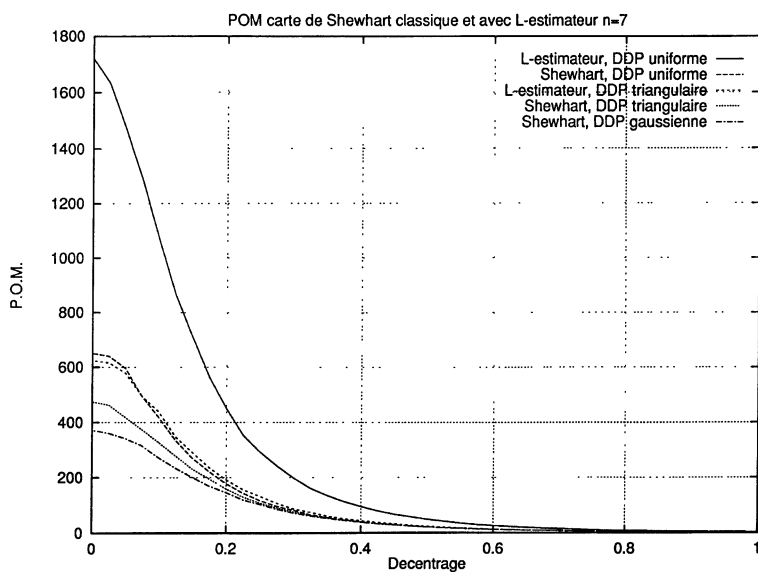


FIGURE 6 :

P.O.M. de la carte de Shewhart avec L-Filtre et de la carte de Shewhart Classique pour un décentrage k compris entre 0 et 1 – $n = 7$.

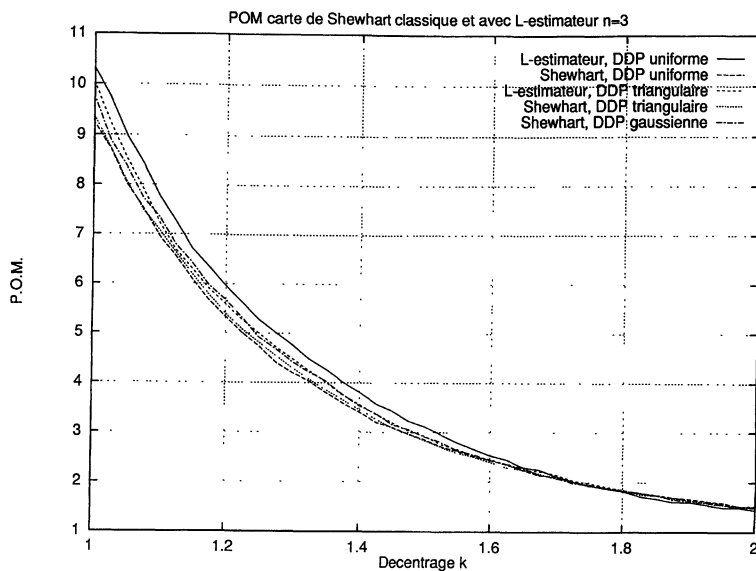


FIGURE 7 :

P.O.M. de la carte de Shewhart avec L-Filtre et de la carte de Shewhart Classique pour un décentrage k compris entre 1 et 2 - $n = 3$.

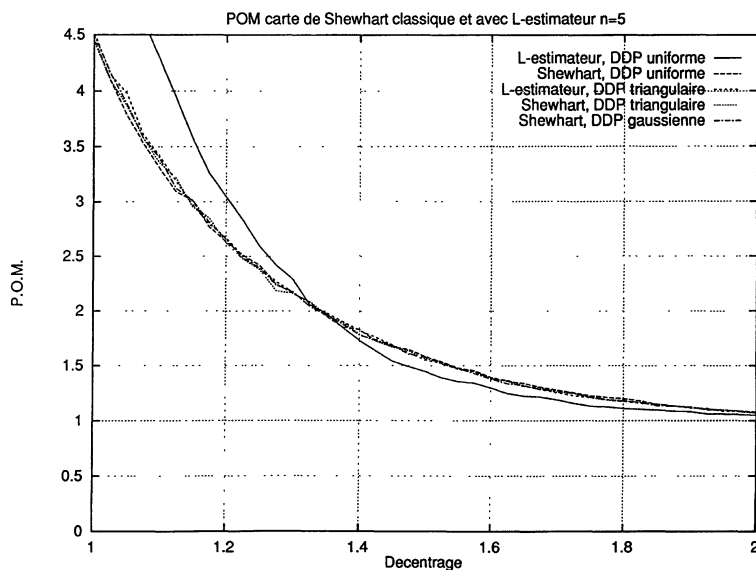


FIGURE 8 :

P.O.M. de la carte de Shewhart avec L-Filtre et de la carte de Shewhart Classique pour un décentrage k compris entre 1 et 2 - $n = 5$.

Lorsque le décentrage augmente ($k > 3 \cdot \sigma / \sqrt{n}$), la P.O.M de la carte de Shewhart avec L-filtre tend plus rapidement vers 1 que la carte de Shewhart «classique»; on limite donc le risque β (fig 7 à 9). Ceci est très net pour $n = 7$ (fig 9) mais moins évident pour $n = 3$ (fig 7) car la différence entre la variance de l'estimateur moyenne et du L-estimateur optimal est très faible (fig 11). D'autre part, on peut observer pour $n = 3$ que la P.O.M. tend vers l'infini lorsque k tend vers zéro, ceci s'explique par le fait que les limites de contrôle se situent à l'extérieur du domaine $[-\sqrt{3}, \sqrt{3}]$ de la population, et à fortiori en dehors de la distribution des valeurs filtrées, d'où un risque a nul.

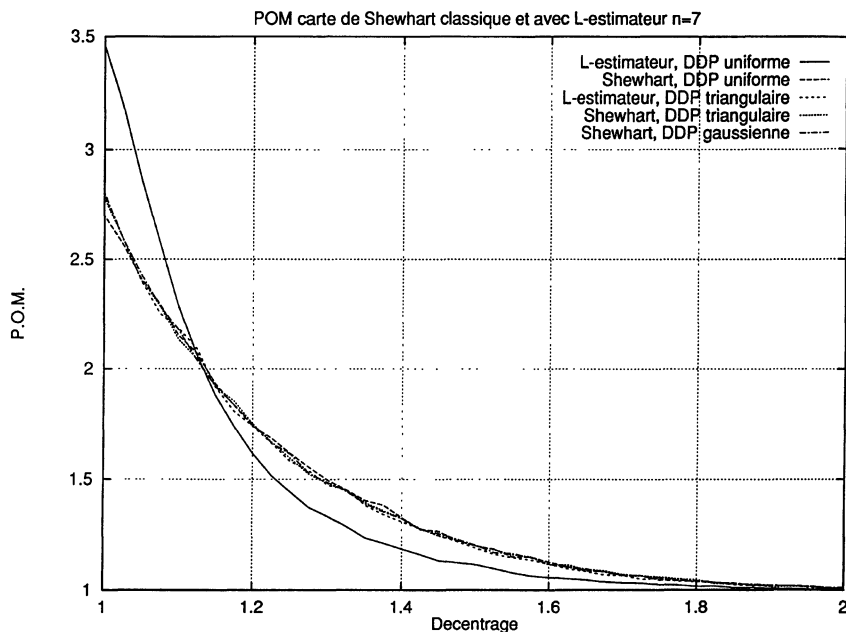


FIGURE 9 :

P.O.M. de la carte de Shewhart avec L-Filtre et de la carte de Shewhart Classique pour un décentrage k compris entre 1 et 2 - $n = 7$.

On constate sur l'ensemble de ces courbes que l'utilisation d'un L-filtre améliore sensiblement la P.O.M. Notamment en ce qui concerne la réduction des risques de fausses alertes lorsque le décentrage est faible. D'autre part, on note que l'utilisation d'une carte de Shewhart avec L-filtre où $n = 3$ offre sensiblement les mêmes performances qu'une carte de Shewhart pour $n = 4$ d'où une réduction intéressante de la taille des échantillons.

4.2. Cas d'une distribution triangulaire

4.2.1. L-Filtre optimal

Dans le cas d'une ddp triangulaire nous avons construit la matrice R de la même manière que précédemment. Elle est bien sûr différente de celle calculée dans le cas d'une loi uniforme c'est pourquoi le L-filtre optimal n'est pas ici le filtre milieu. Nous ne démontrerons pas ici les expressions littérales des coefficients de R du fait de la lourdeur des calculs et du peu d'intérêt qu'ils suscitent. Nous donnerons à titre indicatif les valeurs numériques, calculées par la relation (2), des coefficients du L-filtre optimal pour $n = 3, 5$ et 7 .

$n=3$	$C_1=C_3= 0,4023$	$C_2= 0,1892$		
$n=5$	$C_1=C_5= 0,3066$	$C_2=C_4= 0,1184$	$C_3= 0,1500$	
$n=7$	$C_1=C_7= 0,2741$	$C_2=C_6= 0,06910$	$C_3=C_5= 0,09636$	$C_4= 0,1208$

FIGURE 10 :

Coefficients du L-filtre optimal pour une ddp triangulaire.

4.2.2. La P.O.M.

L'étude des P.O.M., amène les mêmes remarques que dans le cas de distributions uniformes, cependant les améliorations apportées par la carte de Shewhart avec L-filtre sont beaucoup moins significatives que dans le cas uniforme. Ceci est dû au fait que la variance du L-filtre optimal pour une distribution triangulaire est très proche de la variance de sortie du filtre moyenne (filtre optimal pour une loi normale) (fig 9). Les L-Filtres apporteront donc des améliorations d'autant plus significatives que la forme des distributions des mesures sera éloignée de la loi normale. Nous rappelons dans le tableau ci-dessous les variances de l'estimateur moyenne et celles des L-estimateurs optimaux en fonction de la taille n de l'échantillon. Ces dernières ont été calculées à l'aide de la relation (1).

	Moyenne	L-filtre DDP Uniforme	L-Filtre DDP Triangulaire
$n=3$	$1/3=0,333$	0,3	0,3290
$n=5$	$1/5=0,2$	0,1423	0,1927
$n=7$	$1/7=0,1428$	0,08356	0,1354

FIGURE 11 :

*Variance des différents estimateurs
(Moyenne et L-Filtre ou L-estimateur) pour les deux DDP étudiées.*

Nous rappelons dans le tableau ci-dessous les valeurs numériques des P.O.M. des différentes cartes de contrôle étudiées :

k		0	0,2	0,4	0,7	1	1,25	1,5	2	3
N=3 uniforme	L-filtre	∞	1284	159	31,1	10,3	5,27	3,15	1,44	1,01
	Shewhart	∞	1175	146	26,6	9,24	4,64	2,89	1,50	1,01
N=5 uniforme	L-filtre	1691	429	101	20,1	5,99	2,60	1,45	1,05	1,00
	Shewhart	1058	263	61,7	13,1	4,46	2,42	1,59	1,08	1,00
N=7 uniforme	L-filtre	1720	449	95,7	15,1	3,46	1,44	1,12	1,00	1,00
	Shewhart	650	177	39,8	7,72	2,70	1,62	1,20	1,01	1,00
Risque α constant	L-filtre	650	210	53,2	9,77	2,38	1,25	1,07	1,00	1,00
	Shewhart	899	402	126	29,1	10,3	5,00	2,96	1,52	1,01
N=3 triangle	L-filtre	784	356	111	26,7	9,36	4,94	2,94	1,49	1,01
	Shewhart	658	249	64,0	13,7	4,52	2,40	1,55	1,07	1,00
N=5 triangle	L-filtre	546	208	58,3	13,1	4,44	2,38	1,56	1,07	1,00
	Shewhart	624	191	44,1	8,15	2,80	1,58	1,19	1,01	1,00
N=7 triangle	L-filtre	473	159	39,3	7,84	2,78	1,60	1,20	1,01	1,00
	Shewhart									

FIGURE 12 :

Valeurs numériques des P.O.M. pour la carte de Shewhart classique et avec L-filtre, en fonction du décentrage k.

Remarque :

Dans cet article, nous avons choisi de prendre les limites de contrôle égales à $3\sigma\sqrt{n}$ quels que soient la forme de la population et le type d'estimateur. On compare donc les performances des cartes de contrôle à limites constantes. Nous aurions également pu adopter une autre approche, consistant à travailler à risque α constant. Ainsi, pour la carte de Shewhart classique on fixe les limites de contrôle à $3\sigma\sqrt{n}$ comme précédemment. On peut calculer de nouvelles limites pour la carte de Shewhart avec L-filtre, de manière à obtenir un risque α équivalent à celui de la carte de Shewhart classique.

Nous avons traité le cas où $n = 7$ avec une distribution uniforme (Ligne grisée du tableau fig 12 et fig 13). Les limites sont fixées à $2,76\sigma/\sqrt{n}$ pour obtenir une $P.O.M_{k=0} = 650$. Les calculs mettent en évidence que les L-estimateurs permettent de minimiser les risques de fausses détections lorsque le décentrage est faible. En revanche, si le décentrage est significatif ($k > 0,9$), la P.O.M. de la carte de contrôle avec L-filtre devient plus faible que celle de la carte de Shewhart classique. On constate par exemple pour $k\sigma/\sqrt{n} = 3/\sqrt{7} = 1,13$, que $P.O.M_{Shewhart} = 2$ et $P.O.M_{L-filtre} = 1,5$; on détecte donc les décentrages beaucoup plus rapidement en utilisant la carte de Shewhart avec L-filtre.

5. Perspectives

Cette méthode de filtrage offre des perspectives intéressantes pour la M.S.P. car il existe de nombreuses applications industrielles pour lesquelles l'hypothèse d'une répartition normale des mesures n'est pas vérifiée.

Nous travaillons actuellement sur les procédés multi-générateurs, tels que les multi-empreintes et multi-broches, pour lesquels la distribution des mesures sur une cote est normale pour chacun des éléments générateurs (broches ou empreintes) mais pas pour l'ensemble des sources de dispersion (fig 14).

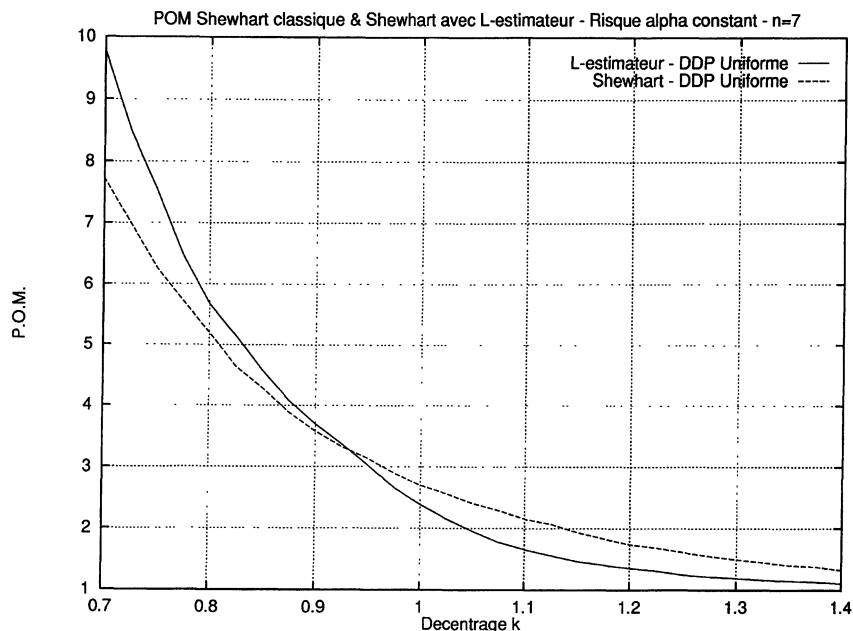


FIGURE 13 :

P.O.M. de la carte de Shewhart avec L-Filtre et de la carte de Shewhart Classique pour un risque α fixé, un décentrage k compris entre 0,7 et 1,4 - $n = 7$.

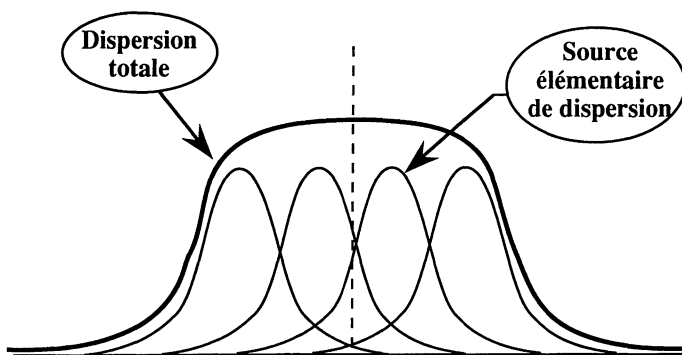


FIGURE 14

Dans de tels cas, il faudrait, pour appliquer la M.S.P. prélever des échantillons de pièces pour chacune des sources élémentaires de dispersion (chaque broche ou chaque moule pour l'injection de plastique). En pratique, cela s'avère souvent impossible du fait de la cadence de production. Ne connaissant pas a priori la forme de la dispersion des causes communes, on peut imaginer qu'au début de la procédure de contrôle du procédé, on attribue au vecteur coefficient du filtre, la valeur $1/n$ du filtre moyennneur.

Ceci revient à appliquer une carte de Shewhart «classique». Ensuite les coefficients du filtre s'ajustent au fur et à mesure que l'on acquiert une connaissance du bruit. Il faut donc mettre en oeuvre un filtre à structure adaptative [Bov 85].

Par ailleurs, nous n'avons abordé ici que les cas de densités de probabilité de bruit symétriques. Or c'est dans les cas non symétriques que le filtrage d'ordre peut apporter les améliorations les plus significatives en terme de P.O.M.

En effet d'un point de vue pratique, on rencontre des distributions non symétriques assez fréquemment. Citons par exemple les études de concentricité (nécessairement positive), ou les systèmes multi-broches et multi-empreintes lorsque les distributions associées à chaque source de dispersion élémentaire ne sont pas distribuées uniformément (fig 15).

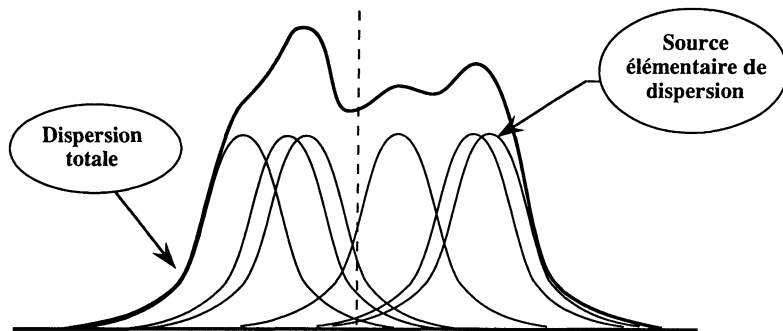


FIGURE 15

6. Conclusion

Dans cet article, nous avons montré l'intérêt d'implanter des opérateurs de filtrage d'ordre sur les cartes de contrôle. Nous avons en effet mis en évidence que l'utilisation de L-Filtres permettait d'améliorer les caractéristiques de la carte de Shewhart en minimisant le risque α lorsque le procédé est centré et en diminuant le risque β lorsque le procédé est hors contrôle. Outre l'amélioration des performances de la carte de Shewhart, l'utilisation des statistiques d'ordre a l'avantage d'introduire une démarche systématique de mise en place des cartes de contrôle; celle-ci étant indépendante de la forme de la population.

D'autre part, nous avons exposé quelques concepts d'applications de cartes de contrôle utilisant les statistiques d'ordre et leurs évolutions éventuelles.

Nos travaux actuels consistent à élargir le champ d'application des filtres d'ordre dans la M.S.P. et à développer les outils théoriques permettant leur mise en oeuvre dans un contexte industriel notamment dans le cas de distributions non symétriques.

Annexe

Calcul des coefficients optimaux d'un L-filtre

Tout d'abord, rappelons que les coefficients d'un L-filtre sont obtenus par la relation (2) qui nécessite le calcul de la matrice R . Cette dernière permet de prendre en compte la forme de la loi de distribution des mesures par le biais des moments centrés d'ordre 1 et 2 des statistiques d'ordre.

L'espérance d'un statistique d'ordre k dans un échantillon de taille n est donnée par la relation suivante.

$$E[x(k)] = n \cdot \mathbb{C}_{n-1}^{k-1} \int_{-\infty}^{+\infty} x \cdot F^{k-1}(x) (1-F(x))^{n-k} \cdot f(x) \cdot dF(x) \text{ avec } 1 \leq k \leq n$$

Supposons le bruit distribué selon une loi uniforme $U_{[0,1]}$.

Alors l'espérance s'écrit :

$$E[x(k)] = n \cdot \mathbb{C}_{n-1}^{k-1} \int_0^1 x^k \cdot (1-x)^{n-k} \cdot dx = \frac{k}{n+1}$$

On en déduit la variance de $x(k)$ et la covariance de $x(k)$ et $x(l)$ (Cf [LEC 87]).

$$\text{Var}[x(k)] = \frac{k(n-k+1)}{(n+1)^2(n+2)} \quad \text{Cov}[x(k), x(l)] = \frac{k(n-l+1)}{(n+1)^2(n+2)} \quad (3)$$

avec $1 \leq k \leq l \leq n$.

Il est alors possible de construire la matrice R .

Le calcul des coefficients du L-filtre étant très difficile à effectuer dans le cas général d'un échantillon de taille n , nous nous contenterons de traiter l'exemple $n = 3$.

$$\text{En appliquant les relations (3) nous obtenons : } R = \frac{1}{80} \cdot \begin{bmatrix} 3 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 3 \end{bmatrix} \quad \text{d'où}$$

$$R^{-1} = 20 \cdot \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}$$

La matrice R^{-1} a la particularité d'avoir la somme des coefficients de ses lignes qui s'annule; excepté pour les lignes 1 et n .

Le calcul des coefficients optimaux du L-filtre nous donne :

$$C_{\text{opt}} = \frac{R^{-1} \cdot e}{{}^t e \cdot R^{-1} \cdot e} = \begin{pmatrix} \frac{1}{2} \\ 0 \\ 1 \\ \frac{1}{2} \end{pmatrix} \quad \text{qui est le filtre milieu.}$$

Par construction, ce filtre est optimal parmi les L-filtres, on ne peut malheureusement pas démontrer s'il est l'estimateur à variance minimale parmi les estimateurs sans biais. En effet, les hypothèses de Fisher relatives à la dérivabilité de la loi n'étant pas satisfaites, on ne peut calculer les bornes de Fréchet.

Variance du filtre milieu

Bovik [Bov 83] a mis en évidence que l'intérêt des L-filtres optimaux était d'offrir une variance de sortie toujours inférieure ou égale à celle de la moyenne. On peut en effet montrer que la variance de sortie du filtre milieu évolue asymptotiquement en $1/n^2$ alors que celle de la moyenne varie en $1/n$.

Rappelons tout d'abord que la variance de sortie d'un L-filtre se calcule à partir de la relation :

$$\text{Var}(Y) = {}^t C R C \quad \text{où} \quad {}^t C = \begin{pmatrix} \frac{1}{2} & 0 & \dots & 0 & \frac{1}{2} \end{pmatrix}$$

Nous avons donné précédemment l'expression de la variance et de la covariance de statistiques d'ordre pour un loi uniforme $U_{[0,1]}$. Nous pouvons généraliser ces expressions à des lois uniformes quelconques avec le coefficient d'échelle $12\sigma^2$ [Llo 52].

$$\begin{aligned} \text{Var}[x(k)] &= 12\sigma^2 \frac{k(n-k+1)}{(n+1)^2(n+2)} & \text{Cov}[x(k), x(l)] &= 12\sigma^2 \frac{k(n-l+1)}{(n+1)^2(n+2)} \\ &\text{avec } 1 \leq k \leq l \leq n \end{aligned}$$

Pour la suite des calculs, on ne s'intéresse qu'aux coefficients placés dans les angles de la matrice R , car les autres seront annulés lors de la multiplication avec le vecteur C des coefficients du filtre milieu.

$$\text{La matrice } R \text{ est alors :} \quad R = \frac{12\sigma^2}{(n+1)^2(n+2)} \cdot \begin{bmatrix} n & \dots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \dots & n \end{bmatrix}$$

$$\text{d'où} \quad \text{Var}(Y) = {}^t C R C = \frac{1}{2} \cdot \frac{12\sigma^2(n+1)}{(n+1)^2(n+2)} = \frac{6\sigma^2}{(n+1)(n+2)}$$

La variance asymptotique du L-estimateur est donc

$$\lim_{n \rightarrow \infty} (\text{Var}(Y)) = \frac{6\sigma^2}{n^2};$$

évolue en $1/n^2$.

Bibliographie

- [Bol 93] PHILIPPE BOLON. «*Filtrage d'ordre, vraisemblance et optimalité des prétraitements d'image*», Traitement du Signal, Vol 9 N°3 - 1993.
- [Bov 83] ALAN C. BOVIK. «*A generalization of median filtering using linear combination of order statistics*», IEEE Transaction on Acoustics, Speech and Signal Processing, Vol ASSP 31 N°6, Décembre 83.
- [Bov 85] ALAN C. BOVIK & ALFREDO RESTREPO. «*Adaptive L-Filters*» Proc 23th Allerton Conf. Commun. Contr. Computer, pp304-313, Monticello 2-4 oct 1985.
- [Cap 88] PH. CAPÉRAÀ & B. VAN CUTSEN. «*Méthodes et modèles en statistiques non paramétriques - Exposé fondamental*», Presses de l'université de Laval, DUNOD, 1988.
- [Dav 81] H. A. DAVID. «*Order Statistics*», New York, Willey, 1981.
- [Kau 94] PASCAL KAUFFMANN. «*Statistique, Information, Estimation, Tests*», Economie-Module, DUNOD, 1994.
- [Lec87] LECOUTRE J.P. & TASSI PH. «*Statistique non paramétrique et robustesse*», 454 Pages, ECONOMICA, 1987.
- [Llo 52] E.H. LLOYD. «*Least Square Estimation of Location and Scale Parameters using Order Statistics*», Biometrika, Vol 39, 1952, pp 88-95.
- [Mon 90] DOUGLAS C. MONTGOMERY. «*Introduction to statistical quality control*» Second edition, Wiley, 675 pages, 1990.
- [Pil 94] MAURICE PILLET. «*Appliquer la maîtrise statistique des procédés (MSP/SPC)*», Edition d'organisation, 1994.
- [Rya 89] THOMAS P. RYAN. «*Statistical methods for quality improvement*», Edition John Wiley & Sons, 1989.
- [Tas 89] PHILIPPE TASSI. «*Méthodes statistiques* », Economica, 2nd Edition, 1989.
- [Zad 92] A.ZADI. «*SPC : Concepts, méthodologie et outils*», Edition Technique et documentation, Paris, 1992.