REVUE DE STATISTIQUE APPLIQUÉE

PAUL DEHEUVELS

Estimation non paramétrique de la densité par histogrammes généralisés

Revue de statistique appliquée, tome 25, n° 3 (1977), p. 5-42

<http://www.numdam.org/item?id=RSA_1977__25_3_5_0>

ESTIMATION NON PARAMÉTRIQUE DE LA DENSITÉ PAR HISTOGRAMMES GÉNÉRALISÉS

Paul DEHEUVELS
Maître de Conférences à l'Université Paris VI

RESUME

Une étude générale des estimations de la densité par la méthode du noyau est faite pour tous les noyaux usuels. On obtient en particulier des développements exacts permettant de calculer l'erreur quadratique intégrée des estimateurs pour les densités analytiques sur **R**. Comme application, on étudie le comportement des estimations appliquées aux densités usuelles, et particulièrement à la distribution normale ; on obtient ainsi les valeurs optimales de la suite $\{\delta_n\}$ pour les noyaux de Parzen-Rosenblatt, pour les petits échantillons. Une étude de robustesse sur le choix des paramètres de départ est entreprise. Cette note est consacrée aux densités dans **R**, et est suivie par [5] (1977), où sont développés les résultats correspondants pour les densités dans $\mathbf{R}^p$, $p \geqslant 2$.

I. – Introduction

L'estimation non paramétrique de la densité a fait l'objet de multiples travaux par des méthodes diverses dans les dix dernières années. Citons pour des exposés de synthèse sur les travaux anciens : Wegman (1972) [33], et Rosenblatt (1971) [22]. L'intérêt de ce type d'estimation est justifié, aussi bien pour les statistiques d'échantillonnage, où elle permet d'éviter les erreurs trop courantes de modélisation, que dans l'analyse spectrale des séries chronologiques. Les méthodes les plus utilisées sont les suivantes :

On supposera par la suite que $\{X_n, n \geqslant 1\}$ est une suite de v.a. indépendantes de même loi sur R, possédant une densité f(x), relativement à la mesure de Lebesgue sur R, continue.

1. Les histogrammes ou estimateurs à noyaux

Leur forme la plus générale est $f_n(x) = \frac{1}{B_n} \sum_{i=1}^{n} K_{n,i}(x, X_i)$; le plus classique d'entre eux est l'estimateur de Parzen-Rosenblatt (Rosenblatt (1956) [21], Parzen (1962) [17]), qui sera l'objet des développements ultérieurs de cette note :

$$f_n(x) = \frac{1}{n\delta_n} \sum_{i=1}^{n} K\left[\frac{X_i - x}{\delta_n}\right] \tag{1,1}$$

Mots-clés : **Estimation de la densité. Histogramme généralisé. Estimation à noyaux. Critère du M.I.S.E.. Robustesse.**

2. Les estimations par fonctions orthogonales :

Etant donné une suite $\{\varphi_n, n \geqslant 0\}$, orthonormale relativement au produit $\int \varphi(x)\, \psi(x) h(x) dx$, h étant une fonction fixée, si, de plus $f \in L^2(h)$ peut être représentée sous la forme $f(x) = \sum_{i=0}^{\infty} a_i\, \varphi(x)$, on estime f par $f_n(x) = \sum_{i=0}^{m(n)} \hat{a}_i \varphi_i(x)$, les $\hat{a}_i$ étant des estimations de a_i données par exemple par $\hat{a}_i = \sum_{j=1}^{n} \varphi_i(X_j) h(X_j)$. Cette méthode a été surtout développée à partir des travaux de Cencov (1962) [3] ; citons également Van Ryzin (1966) [28], Schwarz-Stuart (1967) [27], Kronmal-Tarter (1968) [13], Kronmal-Tarter (1970) [13], Watson (1969) [31], Kronmal-Tarter-Holcomb (1967) [12].

Il est intéressant de noter que les développements de la densité en série relativement à une base orthogonale sont beaucoup plus anciens que les travaux de Cencov, notamment pour les développements d'Edgeworth et de Gram-Charlier (Edgeworth (1904) [7], Charlier (1905) [4], voir aussi Kendall, t.1, p. 156-7, [11]). La différence essentielle entre la théorie moderne et les méthodes anciennes réside dans le fait que ces dernières utilisaient la méthode des moments pour les estimations des coefficients.

On peut également remarquer qu'il n'y a pas de différence formelle entre les estimations générales par la méthode du noyau et les estimations par fonctions orthogonales ; en effet, pour ces dernières :

$$f_n(x) = \frac{1}{n} \sum_{j=1}^{n} \left\{ \sum_{i=0}^{m(n)} \varphi_i(x) \varphi_i(X_j) h(X_j) \right\}$$

Certains auteurs ont utilisé ces similitudes (Watson (1969) [31]). Cependant les méthodes de traitement restent en général assez différentes.

3. Les méthodes dites du maximum de vraisemblance

On trouve plusieurs types d'estimations différentes sous cette appellation ; Citons, comme exemple typique, la note de Robertson (1967) [20], suivie par celles de Prakasa-Rao (1969) [18], Wegman (1969-70) [32], concernant principalement l'estimation des densités unimodales ; un exemple de ces estimations est obtenu dans [32], de la manière suivante :

a) On pose $g_n(x) = \sum_{i=1}^{n} \frac{1}{M(A_i)} U_{A_i}(x)$, où, si $X_{1,n} \leqslant X_{2,n} \leqslant \ldots \leqslant X_{n,n}$ est la statistique ordonnée de l'échantillon, si M est le mode supposé différent des valeurs prises par l'échantillon, $A_1 = (X_{1,n}, X_{2,n}), \ldots, A_r = (X_{p,n}, M)$, $A_{r+1} = (M, X_{p+1,n}), \ldots, A_n = (X_{n-1,n}, X_{n,n})$, $M(A_i)$ désignant la mesure de Lebesgue de A_i ; $M(a, b) = b\text{-}a$. U_A désigne ici la fonction indicatrice de A.

b) L'estimation $f_n(x)$ est alors la fonction définie de manière unique p.p., telle que :

– $\forall a$, $\{x \mid f_n(x) > a\}$ est un intervalle contenant M,

– $\int_{\mathbf{R}} f_n \theta(f_n)\, dx = \int_{\mathbf{R}} g_n \theta(f_n) dx$, $\forall \theta$, avec $\theta(0) = 0$,

$$- \int_A (g_n - f_n)\, dx \leqslant 0, \text{ si A est un intervalle contenant M.} \qquad (1,2)$$

Pour l'estimation précédente, Robertson [20] prouve que f_n est un estimateur du maximum de vraisemblance parmi toutes les fonctions vérifiant (1,2).

L'inconvénient de ces méthodes est qu'elles supposent la connaissance du fait que la densité est unimodale, et la connaissance du mode, qui peut néanmoins être remplacé par un estimateur convergent vers celui-ci (Venter (1967) [29], Wegman (1969) [32]). Citons encore Grenander (1956) [10], Reiss (1973) [19], Sager (1975) [23].

Une nouvelle approche de l'estimation par le maximum de vraisemblance (qui dans le cas ordinaire ne peut donner de résultat, la vraisemblance n'étant pas bornée), consiste, non pas à limiter de façon très restrictive l'espace de fonctions pouvant être prises comme estimations, mais à maximiser :

$\prod\limits_{i=1} f(X_i) \exp(-\Phi(f))$, au lieu de $\prod\limits_{i=1} f(X_i)$, parmi les fonctions d'une variété $H(\mathbf{R})$. Φ est ici une fonctionnelle donnée.

Citons dans ce domaine les résultats de Good-Gaskins (1971) [9], et de Montricher-Tapia-Thompson (1975) [16]. Il est intéressant de constater que les estimations (dites du maximum de vraisemblance avec fonction de pénalisation) ainsi obtenues, fournissent dans des cas assez généraux des estimations polynomiales par morceaux (spline), les extrémités des intervalles (nœuds) étant des points de l'échantillon.

4. Les méthodes mixtes, les histogrammes lissés (histo-splines)

Ces méthodes, d'introduction plus récente, réalisent le plus souvent un mélange des estimations précédentes, afin d'optimiser les résultats de convergence. Un exemple typique est l'histo-spline, obtenu en réalisant dans un premier stade un histogramme simple, obtenu par le découpage du support de la distribution en r intervalles disjoints, puis en ajustant une estimation polynomiale par morceaux (spline) dont les nœuds sont les extrémités des intervalles, qui soit globalement de classe C^p, et qui respecte les probabilités empiriques de ces intervalles. Citons à ce sujet les travaux de Wahba (1971) [30].

Nous ne citerons pas les autres estimations de la densité, bien que certaines d'entre elles présentent un grand intérêt, nous limitant ici aux méthodes les plus importantes. Pour classer par ordre de préférence l'une ou l'autre méthode, plusieurs critères qualitatifs ou quantitatifs peuvent être avancés ; retenons principalement :

1) La simplificité de mise en œuvre :

a) La méthode de l'histogramme peut être, du point de vue pratique très simple ou très pénible à utiliser suivant le choix de K ; ceci dépend encore du fait qu'on cherche à estimer la densité partout ou en un point seulement. Il est commode de se limiter à des noyaux bornés et simples, comme le noyau unité. Dans ce cas le choix des paramètres se ramène au choix de δ_n dans (1,1). En moyenne, elle reste, compte tenu de ces restrictions, la méthode la plus simple.

b) La méthode des fonctions orthogonales a l'avantage de résumer l'estimation à m(n) coefficients. Cependant une difficulté majeure réside outre le choix de m(n), dans le choix de h et de la suite $\{\varphi_n\}$. Les estimations obtenues ne sont pas toujours positives, ni de masse totale 1. Elles sont peu intéressantes pour l'estimation de la densité en un point.

c) Les deux autres méthodes font généralement l'objet d'une programmation très lourde, et du choix de paramètres multiples. Elles sont plus délicates à mettre en oeuvre, et sont de ce fait d'avantage justifiées pour des estimations effectuées en deuxième étape, après une première estimation simple.

2) L'efficacité de l'estimation :

La première difficulté réside dans la manière de mesurer cette efficacité. Le critère le plus commode semble être celui de l'écart quadratique moyen :

$$M^2 = \int_{-\infty}^{+\infty} E((f_n(x) - f(x))^2)dx = E\left[\int_{-\infty}^{+\infty} (f_n(x) - f(x))^2 dx\right] = B_1^2 + B_2^2, \quad (1,3)$$

où

$$B_1^2 = \int_{-\infty}^{+\infty} (E(f_n(x)) - f(x))^2 dx \quad , \quad B_2^2 = \int_{-\infty}^{+\infty} V(f_n(x))\, dx. \qquad (1,4)$$

(Voir à be sujet Parzen [17], Bullock-Davis [2]).

Une estimation réalisant un minimum de M^2 sera dite estimation du M.I.S.E. (minimum integrated square error). D'autres mesures d'écart ont été étudiées, notamment par Bickel-Rosenblatt (1973) [1] :

$$\operatorname*{Sup}_x \frac{|f_n(x) - f(x)|}{f(x)} \quad , \quad \int_R \frac{(f_n(x) - f(x))^2}{f(x)} dx$$

Il ne faut pas cependant juger l'efficacité d'une estimation, comme celà est fait trop souvent, en prenant un exemple particulier pour f, comme une loi normale, et en classant les estimations par ordre d'efficacité (pour ce cas particulier). Les estimations dépendent beaucoup du bon choix des paramètres de départ, et, lorsque ceux-ci sont trop nombreux, leur robustesse est gravement mise en défaut.

Un exemple typique d'estimation par la méthode des fonctions orthogonales est obtenu en prenant $h(x) = (2\pi)^{\frac{1}{2}} e^{\frac{1}{2}x^2}$, et comme fonction φ_i, les polynômes d'Hermite étant définis par :

$$H_i(x) = (-1)^i h(x) \frac{d^i}{dx^i} h^{-1}(x) = \int_{-\infty}^{+\infty} H_m(x) H_n(x) h^{-1}(x) dx = \begin{cases} 0 \text{ si } m \neq n \\ n! \text{ si } m = n \end{cases} \qquad (1,5)$$

On posera

$$\varphi_n(x) = \frac{1}{n!} H_n(x)\, h^{-1}(x).$$

Le développement de f dans cette base orthonormale est le développement de Gram-Charlier de f (Voir [11] t.1, p. 156) ; pour estimer une loi normale $N(0,1)$, ce type de développement est idéal, puisqu'on pourra se limiter au premier terme $a_0 = 1$, $a_n = 0$ si $n \geqslant 1$. On peut alors montrer (Voir [11] t. 1, p. 162), que le développement de Gram-Charlier correspondant à une loi $N\left(0\ ,\ \frac{1}{2}\right)$ est divergent. Ainsi, comme on le voit dans cet exemple, il suffit de diminuer l'écart type de la loi dans la proportion $1/\sqrt{2}$, pour obtenir des estimations non valables.

3) La robustesse des estimations :

Il est nécessaire de distinguer plusieurs types de robustesse, portant soit sur le type de distribution, soit sur le choix des paramètres d'estimation. L'histogramme apparaît ici comme la méthode générale la plus robuste lorsque peu de renseignements à priori sont donnés sur la distribution.

Il ressort de ces considérations que l'estimation non paramétrique de la densité doit se faire en plusieurs étapes, utilisant dans les premières étapes des histogrammes, puis éventuellement des méthodes plus perfectionnées (4) permettant d'avoir des ajustements plus étroits. Woodroofe (1970) [35] donne un exemple d'une telle procédure avec des histogrammes. Il est aussi très concevable, et c'est ce qui se passe le plus souvent, lorsque l'histogramme ou quelqu'autre méthode a donné "un bon résultat", de terminer l'estimation par un ajustement paramétrique.

Ceci étant établi, le praticien doit résoudre le problème de la réalisation pratique de l'histogramme. Or la plupart des résultats utilisables, notamment sur le choix de δ_n sont établis asymptotiquement pour n infini, et le fait d'appliquer sans discernement de telles formules pour des petits échantillons, par exemple pour $20 \leqslant n \leqslant 150$, conduit à des erreurs considérables. Le but de cet exposé est de tenter de résoudre partiellement ces difficultés dans les cas usuels. Nous établissons en particulier pour la loi normale, et pour une famille très large de distributions et de noyaux, des formules qui permettent de calculer exactement l'erreur quadratique intégrée.

II. Estimations par histogrammes uniformes

On considère ici les estimations du type suivant :

$$f_n(x) = \frac{1}{n\delta_n} \sum_{i=1}^{n} K \left[\frac{X_i - x}{\delta_n}\right]$$, avec les hypothèses suivantes :

$\{X_n, n \geqslant 1\}$ est une suite de v.a. indépendantes de même loi, possédant une densité f sur **R**.

$\{\delta_n, n \geqslant 1\}$ est une suite positive décroissante, de limite 0.

K est une fonction bornée, telle que $\int_{-\infty}^{+\infty} K(y)dy = 1$; on notera en général :

$$[[y^p K^r]] = \int_{-\infty}^{+\infty} y^p\, K^r(y)\, dy$$, si cette quantité est définie.

Les noyaux usuels se limitent à la liste suivante :

A) Noyaux polynomiaux ou de Legendre :

En général $K(y) = P(y)$, $|y| < C$, où P est un polynôme, $K(y) = 0$ pour $|y| \geqslant C$. Comme exemple importants :

(1) Noyau unité : $K(y) = 1$, si $|y| < \frac{1}{2}$.

(2) Noyau d'Epanechnikov : $K(y) = 1 - y^2$, si $|y| < 1$.

(3) Noyau de Legendre d'ordre 1 : $K(y) = \frac{3}{8}(3 - 5y^2)$, si $|y| < 1$.

(4) Noyau de Legendre d'ordre 2 : $K(y) = \frac{15}{128}(15 - 70y^2 + 63y^4)$, si $|y| < 1$.

B) Noyaux de Gram-Charlier :

En général $K(y) = P(y) e^{-\frac{1}{2}y^2}$, P étant un polynôme. Comme exemples importants :

(5) Noyau normal : $K(y) = (2\pi)^{-\frac{1}{2}} e^{-\frac{1}{2}y^2}$

(6) Noyau de Gram-Charlier d'ordre 1 : $K(y) = \frac{1}{2}(3 - y^2)(2\pi)^{-\frac{1}{2}} e^{-\frac{1}{2}y^2}$

(7) Noyau de Gram-Charlier d'ordre 2 :

$$K(y) = \frac{1}{8}[y^4 - 10y^2 + 15](2\pi)^{-\frac{1}{2}} e^{-\frac{1}{2}y^2}$$

C) Noyaux de Laguerre :

En général $K(y) = P(|y|) e^{-|y|}$, P étant un polynôme. Comme exemples importants :

(8) Noyau de Picard : $K(y) = \frac{1}{2} e^{-|y|}$

(9) Noyau de Laguerre d'ordre 1 : $K(y) = \frac{1}{2}(\frac{1}{2}y^2 - 3|y| + 3) e^{-|y|}$

D) Noyaux divers :

(10) Noyau de Cauchy d'ordre r : $K(y) = \frac{2^{2r-2}\,\Gamma(r)^2}{\pi(2r-1)} \frac{1}{(1+x^2)^r}$

(11) $K(y) = \frac{(1 - |y|^a)\ (a+1)}{2a}$, si $|y| < 1$; pour $a = 1$, noyau triangulaire $K(y) = 1 - |y|$

(12) $K(y) = \frac{e^{-|y|^a}}{2\Gamma(1/a)}.$

(13) Noyau de Fejer-de la Vallée Poussin : $K(y) = \pi^{-1}\left(\frac{1}{y}\sin y\right)^2$

(14) Noyau de Jackson-de la Vallée Poussin : $K(y) = 3\pi^{-1} \left(\frac{1}{y} \sin y\right)^4$

(15) Noyau de Fourier : $K(y) = \pi^{-1} \left(\frac{1}{y} \sin y\right)$

(16) $K(y) = \cos y, \ |y| < \frac{1}{2}\pi$.

Les noyaux 13, 14, 15 présentent peu d'intérêt pratique, nous ne les citons ici que pour mémoire.

Nous étudions maintenant les propriétés en général de convergence en moyenne quadratique des estimations :

A) Moments des estimations :

Par un changement de variable, on obtient l'identité :

$$E\left[K^p\left(\frac{X-x}{u}\right)\right] = u \int_{-\infty}^{+\infty} K^p(y)\, f(x+uy)\, dy \tag{2,1}$$

Supposons que f soit n fois différentiable en x :

$$f(x+h) = f(x) + hf'(x) + \frac{1}{2}h^2 f''(x) + \cdots + \frac{h^n}{n!} f^{(n)}(x) + \epsilon_n(x, h)\, h^n \tag{2,2}$$

Dans le cas où f est n + 1 fois dérivable dans l'intervalle (x , x + h),

$$\epsilon_h(x, h)\, h^n = \int_0^h \frac{(h-t)^n}{n!} f^{(n+1)}(x+t)\, dt \tag{2,3}$$

On obtient alors formellement, en intégrant (2.2) dans (2,1) :

$$\frac{1}{u} E\left[K^p\left(\frac{X-x}{u}\right)\right] = \sum_{r=0}^{n} u^r \frac{f^{(r)}(x)}{r!} [[y^r K^p]] + u^n R_n(x, u), \tag{2,4}$$

avec : $$R_n(x, h) = \int_{-\infty}^{+\infty} y^n K^p(y)\, \epsilon_n(x, yh)\, dy$$

Pour obtenir que, dans (2,3), $\lim_{u \to 0} R_n(x, u) = 0$, il est nécessaire de faire des hypothèses sur K et f :

Proposition (2.1) : Si f est n fois différentiable en x,

1) Si K est nul en dehors d'un compact (cas des noyaux 1, 2, 6), ou,

2) Si f est définie sur $\mathbf{R}$ et y admet une majoration de la forme :

$$f(y) \leqslant C(1 + |y|^R), \ R \geqslant 0, \tag{2,5}$$

et si $[[y^r K^p]]$ est défini $\forall r \geqslant 0$, ou,

3) Si f vérifie la condition du 2), sauf au voisinage d'un nombre fini de points où elle admet des discontinuités infinies, et si K vérifie l'hypothèse :

$$\lim_{y\infty} y^r K^p (y) = 0, \forall r \geqslant 0, \tag{2,6}$$

Alors (2,4) est vérifié, avec $\lim_{u \to 0} R_n (x , u) = 0$.

Preuve : Il suffit d'écrire

$$R_n (x , u) = \int_{-h/u}^{+h/u} + \int_{|y|>h/u} y^n K^p (y) \epsilon_n (x , yu) \, dy$$

Par la suite, nous supposerons toujours que les hypothèses (2,6) et (2,5) 3) la proposition (2,1) sont vérifiées. On supposera K borné.

Corollaire (2,2) : Si f est N fois différentiable en x, on a :

$$E (f_n(x)) = f(x) + \sum_{r=1}^{N} \delta_n^r \frac{f^{(r)}(x)}{n!} [[y^r K]] + R_N (x , \delta_n) \delta_n^N , \tag{2,7}$$

avec $\lim_{u \to 0} R_N(x , u) = 0$,

$$(E(f_n(x)) - f(x))^2 = \sum_{r=2}^{N+1} \frac{\delta_n^r}{r!} \left\{ \sum_{p=1}^{r-1} C_r^p f^{(p)}(x) f^{(r-p)}(x) [[y^p K]] [[y^{r-p} K]] \right\} +$$

$$+ S_{N+1} (x , \delta_n) \delta_n^{N+1} , \tag{2,8}$$

avec $\lim_{u \to 0} S_{N+1} (x , u) = 0$.

$$V (f_n (x)) = \frac{1}{n\delta_n} \left[\left[\sum_{r=0}^{N} \delta_n^r \frac{f^{(r)}(x)}{r!} [[y^r K^2]] \right] - \right.$$

$$- \left[\sum_{r=0}^{N-1} \frac{\delta_n^{r+1}}{r!} \left\{ \sum_{p=0}^{r} C_r^p f^{(p)}(x) f^{(r-p)}(x) [[y^p K]] [[y^{r-p} K]] \right\} \right]$$

$$\left. + T_N (x , \delta_n) \delta_n^N \right] , \tag{2.9}$$

avec $\lim_{u \to 0} T_N (x , u) = 0$.

Remarques : Les résultats (2,7), (2,8), (2,9) se déduisent aisément de (2,4), les v.a. étant indépendantes, et car

$$V \left[K \left(\frac{X - x}{u} \right) \right] = E \left[K^2 \left(\frac{X - x}{u} \right) \right] - E \left[K \left(\frac{X - x}{u} \right) \right]^2 .$$

Ces résultats restent valables lorsque $N = 0$, c'est-à-dire lorsque f est continue en x ; on supprimera alors les sommations impossibles ; on obtient dans ce cas :

$$E(f_n(x)) - f(x) = R_0(x, \delta_n), \quad V(f_n(x)) = \frac{[[K^2]]}{n\delta_n}[f(x) + T_0(x, \delta_n)].$$

Ces résultats sont valables pour les noyaux $\begin{matrix}1\text{-}2\text{-}3\text{-}4\text{-}5\text{-}6\text{-}7\\ 8\text{-}9\text{-}11\text{-}12\text{-}16\end{matrix}$ en particulier. Ils n'ont plus de sens, les moments correspondants n'étant pas définis, pour les noyaux 13-14-15.

Il est intéressant, lorsque f est analytique en x, d'obtenir un développement en série, à la place de (2,4) :

$$\frac{1}{u} E\left[K^p\left(\frac{X - x}{u}\right)\right] = \sum_{r=0}^{\infty} u^r \frac{f^{(r)}(x)}{r!} [[y^r K^p]] ; \tag{2,10}$$

Dans le cas de la loi normale N (0 , 1), on obtient :

$$f^{(2n)}(x) = \frac{(2n)!}{\sqrt{2\pi}} \left\{ \frac{x^{2n}}{(2n)!} - \frac{\frac{1}{2} x^{2n-2}}{(2n-2)!\,1!} + \cdots + \frac{(-\frac{1}{2})^p x^{2n-2p}}{(2n-2p)!\,p!} + \cdots + \frac{(-\frac{1}{2})^{n-1} x^2}{2!\,(n-1)!} + \frac{(-\frac{1}{2})^n}{n!} \right\} e^{-\frac{1}{2}x^2} \tag{2,11}$$

$$f^{(2n+1)}(x) = \frac{(2n+1)!}{\sqrt{2\pi}} \left\{ \frac{-x^{2n+1}}{(2n+1)!} + \frac{\frac{1}{2} x^{2n-1}}{(2n-1)!\,1!} - \cdots - \frac{(-\frac{1}{2})^p x^{2n-2p+1}}{(2n-2p+1)!\,p!} - \cdots - \frac{(-\frac{1}{2})^{n-1} x^3}{3!\,(n-1)!} - \frac{(-\frac{1}{2})^n x}{n!} \right\} e^{-\frac{1}{2}x^2}$$

Lorsque $n \to \infty$,

$$\operatorname*{Lim\,Sup}_{n\infty} \sqrt{\frac{n}{e}} \left| \frac{f^{(n)}(x)}{n!} \right|^{\frac{1}{n}} = 1$$

Pour que (2,10) soit valable, K étant borné, il est suffisant que :

$$|u| < \frac{1}{\operatorname*{Lim\,Sup}_{n\infty} \left[\sqrt{\frac{e}{n}} \left\{ \int_{-\infty}^{+\infty} |y|^n \, |K(y)| \, dy \right\}^{1/n} \right]} = R(K) \tag{2,12}$$

Il est nécessaire que :

$$|u| \leqslant \frac{1}{\operatorname*{Lim\,Sup}_{n\infty} \left[\sqrt{\frac{e}{n}} [[y^n K]]^{1/n}\right]} = R'(K) \tag{2,13}$$

Remarquons que dans le cas où $K \geqslant 0$ (K est une densité de probabilité, dans (2,12) et (2,13),

$$R(K) = R'(K) = \frac{1}{\underset{n\infty}{\text{Lim Sup}} \left\{ \sqrt{\frac{e}{2n}} [[y^{2n} K]]^{1/2n} \right\}} \tag{2,14}$$

R (K) n'est pas toujours infini, comme on le constate aisément dans le cas d'un noyau $K(y) = \dfrac{\exp(-|x|^a)}{(2/a)\ \Gamma\ (1/a)}$; $R(K) = 0$ si $a < 2$, $R(K) = \sqrt{\frac{1}{2}}$ si $a = 2$ (noyau normal) et $R(K) = \infty$, si $a > 2$ (en particulier pour $a = 1$, cas du noyau de Picard, $R(K) = 0$).

Pour un noyau nul en dehors d'un compact, R (K) est toujours infini.

Une discussion analogue peut être faite dans le cas général ; nous la résumons :

Proposition (2,3) : Si f est analytique au voisinage du point x,

1) Si r est le rayon de convergence du développement de f en série entière au voisinage de x, si $K(y) = 0$ pour $|y| > C$, alors (2,10) est vérifié et convergent pour $|u| < r/C$.

2) Si f est analytique sur $\mathbf{R}$, et si $r = \infty$, si K n'est pas nul en dehors d'un compact de $\mathbf{R}$, (2,10) est vérifié et convergent pour :

$$|u| < \frac{1}{\underset{n\infty}{\text{Lim Sup}} \left[\dfrac{f^{(n)}(x)}{n!} \displaystyle\int_{-\infty}^{+\infty} |y|^n\, |K(y)|\, dy \right]^{1/n}},$$

ou si

$$K \geqslant 0,\ |u| < \frac{1}{\underset{n\infty}{\text{Lim Sup}} \left[\dfrac{f^{(n)}(x)}{n!} [[y^n K]] \right]^{1/n}} \tag{2.15}$$

Remarques : 1) Si f est analytique sur $\mathbf{R}$ et de rayon infini, (2,10) a un rayon de convergence infini si K est nul en dehors d'un compact de $\mathbf{R}$ (cas des noyaux 1-2-3-4-11-16).

2) Dans le cas de la loi normale N (0,1), (2,15) devient (2,12), ou (2,14). En général, si $K \geqslant 0$, et si $|u| < r/ \operatorname{Inf} \{t \mid K(y) = 0 \text{ p.p. pour } |y| > t\}$, (2,5) est divergent, et ceci est toujours le cas si K n'est pas nul p.p. en dehors d'un compact. Il ne faut donc pas espérer une amélioration sensible de ces résultats.

3) En général, si f est à support borné à l'une ou l'autre des extrémités et analytique sur son support, le rayon de convergence r est inférieur ou égal à la plus courte distance de x aux extrémités du support.

En particulier pour les lois de Pearson, pour les types I, II, VIII, IX (lois Béta), la borne supérieure précédente est égale au rayon de convergence. Il en est de même pour les lois Gamma (type III), le type V (loi Gamma inversée), et le type VI (loi Béta inversée).

En ce qui concerne les types IV et VII, de densité

$$f(x) = \frac{1}{F(r,s)} \frac{\exp(-s\text{Arctg}\, x)}{(x^2+1)^r},$$

sous forme réduite, $r = \sqrt{x^2+1}$. On constate, en particulier que (2,10) n'est jamais convergent pour une loi de Cauchy et un noyau non nul en dehors d'un compact.

Les lois lognormales, de densité réduite

$$f(x) = \frac{1}{x\sqrt{2\pi}} \exp\left[-\frac{(\text{Log}\, x - m)^2}{2}\right], x > 0,$$

de Weibull, de densité réduite $f(x) = Cx^{C-1} \exp(-x^C)$, $x > 0$, vérifient la même propriété, le rayon de convergence du développement de f au voisinage de x est $r = x$. Il en est de même pour les lois limites extrêmes de types I et II (Φ et ψ) de fonctions de répartition respectives $F(x) = \exp(-1/x^a)$, $x > 0$, et $F(x) = \exp(-(-x)^a)$, $x < 0$, $F(x) = 1$, $x \geqslant 0$.

Pour le système de distributions de Johnson, les lois S_L coïncident avec les lois lognormales, les lois S_B obtenues à partir d'une loi normale Y par la transformation réciproque de $Y = A + B\log\left[\frac{X-a}{X-b}\right]$, $r = \text{Inf}\,(|x-a|, |x-b|)$.

Pour les lois S_U de Johnson définies par $X = A + B\text{sh}\left[\frac{Y-a}{b}\right]$, ainsi que pour les lois limites extrêmes de type III (Λ) de fonction de répartition $F(x) = e^{-e^{-x}}$, $r = \infty$.

Corollaire (2.4) : Avec les hypothèses de la proposition (2,3), entrainant la validité de (2,10) :

$$E(f_n(x)) = f(x) + \sum_{r=1}^{\infty} \delta_n^r \frac{f^{(r)}(x)}{r!} [[y^r K]], \tag{2,16}$$

$$(E(f_n(x)) - f(x))^2 = \sum_{r=2}^{\infty} \frac{\delta_n^r}{r!} \left\{ \sum_{p=1}^{r-1} C_r^p f^{(p)}(x) f^{(n-p)}(x) [[y^p K]] [[y^{r-p} K]] \right\} \tag{2,17}$$

$$V(f_n(x)) = \frac{1}{n\delta_n} \left[\left[\sum_{r=0}^{\infty} \delta_n^r \frac{f^{(r)}(x)}{r!} [[y^r K^2]]\right] - \right.$$

$$\left. - \left[\sum_{r=0}^{\infty} \frac{\delta_n^{r+1}}{r!} \left\{ \sum_{p=0}^{\infty} C_r^p f^{(p)}(x) f^{(r-p)}(x) [[y^p K]] [[y^{r-p} K]] \right\}\right]\right] \tag{2,18}$$

B) Moyenne quadratique intégrée (calcul direct)

Nous estimons ici $M^2 = E\left[\int_I (f_n(x) - f(x))^2 dx\right] = B_1^2 + B_2^2$, où :

$$B_1^2 = \int_I (E(f_n(x)) - f(x))^2\, dx \quad , \quad B_2^2 = \int_I V(f_n(x))\, dx.$$

Proposition (2,5) : Si f est N fois continûment différentiable dans un ouvert contenant I, et dans l'un ou l'autre des cas suivants :

1) I est un compact, et K est un noyau nul en dehors d'un compact ;

2) $f^{(p)} \in L^1(V(I)) \cap L^2(V(I))$, $\forall 0 \leqslant p \leqslant N+1$, V(I) étant un voisinage ouvert de I, et, ou bien K est nul en dehors d'un compact, ou bien $I = \mathbf{R}$; alors :

$$B_1^2 = \sum_{r=2}^{N+1} \frac{\delta_n^r}{r!} \left\{ \sum_{p=1}^{r-1} C_r^p\, [[y^p K]]\, [[y^{r-p} K]] \int_I f^{(p)}(x) f^{(r-p)}(x)\, dx \right\} + o(\delta_n^{N+1}) \tag{2,19}$$

$$B_2^2 = \frac{1}{n\delta_n} \left[\left[\sum_{r=0}^{N} \frac{\delta_n^r}{r!} [[y^r K^2]] \int_I f^{(r)}(x)\, dx \right] - \left[\sum_{r=0}^{N-1} \frac{\delta_n^{r+1}}{r!} \left\{ \sum_{p=0}^{r} C_r^p\, [[y^p K]]\, [[y^{r-p} K]] \int_I f^{(p)}(x)\, f^{(r-p)}(x)\, dx \right\} \right] + o(\delta_n^N) \right] \tag{2,20}$$

Dans le cas 2), on peut remplacer dans ces expressions $o(\delta_n^N)$ par $O(\delta_n^{N+1})$. Preuve : On utilise (2,3), (2,8), (2,9) et le théorème de Lebesgue.

Remarques : 1) Si $f^{(p)}$, $0 \leqslant p \leqslant N$, s'annule aux extrémités de I, (2,19) et (2,20) se simplifient considérablement. En effet :

$$\int_I f^{(p)}(x) f^{(r-p)}(x) = -\int_I f^{(p+1)}(x) f^{(r-p-1)}(x) \ ;$$

$$\sum_{p=1}^{2r} \quad \text{et} \quad \sum_{p=0}^{2r+1} (-1)^p\, C_{2r+1}^p\, [[y^p K]]\, [[y^{2r-p+1} K]]$$

ont un coefficient nul ; on obtient en conséquence :

$$B_1^2 = \sum_{r=1}^{[\frac{1}{2}(N+1)]} \frac{\delta_n^{2r}}{(2r)!} \left\{ \int (f^{(r)}(x))^2\, dx \right\} \left\{ \sum_{p=1}^{2r-1} C_{2r}^p (-1)^{p+r}\, [[y^p K]]\, [[y^{2r-p} K]] \right\} + o(\delta_n^{N+1}) \tag{2,21}$$

$$B_2^2 = \frac{1}{n\delta_n} \left[[[K^2]] - \delta_n \sum_{r=0}^{[\frac{1}{2}(N-1)]} \frac{\delta_n^{2r}}{(2r)!} \left\{ \int_I (f^{(r)}(x))^2 dx \right\} \left\{ \sum_{p=0}^{2r} C_{2r}^p (-1)^{p+r}\, [[y^p K]]\, [[y^{2r-p} K]] \right\} + o(\delta_n^N) \right]. \tag{2,22}$$

2) Si, en plus des hypothèses précédentes, K est symétrique relativement à 0, $[[y^{2r+1}K]]$ sera nul pour tout r. On obtient :

$$B_1^2 = \sum_{r=2}^{[\frac{1}{2}(N+1)]} \frac{\delta_n^{2r}}{(2r)!} \left\{ \int_I (f^{(r)}(x))^2 \, dx \right\} \left\{ (-1)^r \sum_{p=1}^{r-1} C_{2r}^{2p} [[y^{2p}K]] [[y^{2r-2p}K]] \right\} + o(\delta_n^{N+1}) \quad (2,23)$$

$$B_2^2 = \frac{1}{n\delta_n} \left[[[K^2]] - \delta_n \sum_{r=0}^{[\frac{1}{2}(N-1)]} \frac{\delta^{2r}}{(2r)!} \left\{ \int_I (f^{(r)}(x))^2 \, dx \right\} \left\{ (-1)^r \sum_{p=0}^{r} C_{2r}^{2p} [[y^{2p}K]] [[y^{2r-2p}K]] \right\} + o(\delta_n^N) \right] \quad (2,24)$$

3) Remarquons que les expressions (2,21), (2,22), (2,23), (2,24) sont valables sous réserve que les hypothèses des 1) et 2) soient vérifiées, même si f n'est pas symétrique. Si la distribution est à support infini, et convenablement régulière, l'hypothèse $\lim_{x\,\infty} f^{(p)}(x) = 0$ est parfaitement naturelle (elle est nécessaire si on suppose que $f^{(p)}$ est monotone au voisinage de l'infini).

Comme précédemment, on peut obtenir des conditions pour que ces développements soient des développements en série ; pour celà, en utilisant l'inégalité de Schwarz, on majore : $\left| \int_I f^{(p)}(x) f^{(q)}(x) dx \right| \leqslant \left[\int_I (f^{(p)}(x))^2 dx \int_I (f^{(q)}(x))^2 dx \right]^{\frac{1}{2}}$; en appliquant alors (2,17), (2,18) et le théorème de Lebesgue, on obtient :

Proposition (2,6) : Si f est analytique sur I, et si les conditions de la proposition (2,3) entrainant (2,10) pour $u = \delta_n$ sont vérifiées en tout point de I, si, de plus les conditions de la proposition (2,5) 2) sont vérifiées pour $N = \infty$, c'est-à-dire $\forall\, p \geqslant 0$, $f^{(p)} \in L^1(V(I)) \cap L^2(V(I))$, V(I) étant un voisinage ouvert de I, et, ou bien K est nul en dehors d'un compact, ou bien $I = \mathbf{R}$, si, de plus,

$$|\delta_n| < \frac{1}{\operatorname{Lim\,Sup}_{n\,\infty} \left\{ \left[\int_I |f^{(n)}(x)|^2 \, dx \right]^{\frac{1}{2}} \frac{1}{n!} \int_{-\infty}^{+\infty} |y|^n \, |K(y)| \, dy \right\}^{1/n}} \quad (2,25)$$

alors :

$$B_1^2 = \sum_{r=2}^{\infty} \frac{\delta_n^r}{r!} \left\{ \sum_{p=1}^{r-1} C_r^p [[y^pK]] [[y^{r-p}]] \int_I f^{(p)}(x) f^{(r-p)}(x) \, dx \right\} \quad (2,26)$$

$$B_2^2 = \frac{1}{n\delta_n} \left[\left[\sum_{r=0}^{\infty} \frac{\delta_n^r}{r!} [[y^r K^2]] \int_I f^{(r)}(x) \, dx \right] - \left[\sum_{r=0}^{\infty} \frac{\delta_n^{r+1}}{r!} \left\{ \sum_{p=0}^{r} C_r^p [[y^p]] [[y^{r-p}K]] \int_I f^{(p)}(x) f^{(r-p)}(x) \, dx \right\} \right] \right] \quad (2,27)$$

Remarque : Avec les hypothèses correspondantes sur f ou K, (2,26) et (2,27) peuvent être mis sous les formes simplifiées correspondant à (2,21), (2,22), (2,23) et (2,24), avec $N = \infty$.

Remarque : Les méthodes obtenues pour arriver aux développements précédents sont développées, en particulier dans Woodroofe (1967) [34], Shapiro (1969) [26], Deheuvels (1974) [5]. Les développements pour $N = 0$ dans (2,7), (2,8), (2,9) sont plus anciens, et se retrouvent dans Rosenblatt (1956) [21], Parsen (1962) [17].

C) Développements pour la loi normale N(0,1), et les noyaux symétriques :

Soit $f(x) = (2\pi)^{-\frac{1}{2}} e^{-\frac{1}{2}x^2}$; dans la suite, il sera utile de déterminer $I_n = \int_{-\infty}^{+\infty} x^{2n} f(x)^2 \, dx$:

$$I_n = \int_{-\infty}^{+\infty} \frac{x^{2n}}{2\pi} e^{-x^2} dx = \int_{-\infty}^{+\infty} \frac{(\frac{1}{2})^n u^{2n}}{\sqrt{2}(2\pi)} e^{-\frac{1}{2}u^2} du = \frac{(\frac{1}{2})^{n+1}}{\sqrt{\pi}} E(X^{2n}) = \frac{(\frac{1}{2})^{2n+1}(2n)!}{\sqrt{\pi}\ n!}$$

Tout d'abord :

$$\int_{-\infty}^{+\infty} |f^{(2n)}(x)| \, dx \leqslant (2n)! \sum_{r=0}^{n} \frac{E(X^{2n-2r})}{2^r r! (2n-2r)!} \leqslant \frac{(2n)!}{n!} \sum_{r=0}^{n} \frac{n!}{2^n r! (n-r)!} \leqslant \frac{(2n)!}{n!}$$

D'autre part, comme

$$\int_{0}^{+\infty} x^{2r+1} e^{-\frac{1}{2}x^2} dx = 2^r r! \leqslant Cr^{\frac{1}{2}} 2^{r-\frac{1}{2}} \Gamma(r + \tfrac{1}{2}) = Cr^{\frac{1}{2}} \frac{(\frac{1}{2})^{r+1} (2r)! \sqrt{2\pi}}{r!}$$

$$\int_{-\infty}^{+\infty} |f^{(2n+1)}(x)| \, dx \leqslant \frac{2(2n+1)!}{\sqrt{2\pi}} \sum_{r=0}^{n} \frac{2^n (n-r)!}{2^{2r} r! (2n-2r+1)!} \leqslant$$

$$\leqslant \frac{C(2n+1)!}{n!} \sum_{r=0}^{n} \frac{n!}{2^n r! (n-r)! (2n-2r+1)^{\frac{1}{2}}}$$

$$\leqslant C \frac{(2n+1)!}{n!}$$

Considérons maintenant

$$\int_{-\infty}^{+\infty} f^{(2n-r)}(x) f^{(r)}(x) \, dx = (-1)^r \int_{-\infty}^{+\infty} f^{(2n)}(x) f(x) \, dx,$$

et

$$\int_{-\infty}^{+\infty} f^{(2n)}(x) \, f(x) \, dx = (2n)! \sum_{r=0}^{n} \frac{(-1)^r I_{n-r}}{2^r r! (2n-2r)!} =$$

$$= \frac{(\frac{1}{2})^{2n+1} (2n)!}{n! \sqrt{\pi}} \left[\sum_{r=0}^{n} C_n^r (-2)^r \right] = \frac{(-1)^n (\frac{1}{2})^{2n+1} (2n)!}{n! \sqrt{\pi}} \qquad (2,28)$$

On obtient alors par la formule de Stirling : asymptotiquement,

$$\left[\frac{1}{n!}\int_{-\infty}^{+\infty} |f^{(n)}(x)|\, dx\right]^{\frac{1}{n}} \leqslant \sim \sqrt{\frac{2e}{n}};$$

$$\left\{\frac{1}{n!}\left[\int_{-\infty}^{+\infty} |f^{(n)}(x)|^2\, dx\right]^{\frac{1}{2}}\right\}^{\frac{1}{n}} \sim \sqrt{\frac{e}{n}} \quad (2,29)$$

On déduit de la proposition (2,6), et de (2,23), (2,24) :

Proposition (2,7) : Si K est un noyau symétrique, f une densité N (0 , 1), et si :

$$|\delta_n| < \frac{1}{\operatorname*{Lim\,Sup}_{n\infty} \left\{\sqrt{\frac{e}{n}} \int_{-\infty}^{+\infty} |y|^n\, |K(y)|\, dy\right\}^{\frac{1}{n}}}, \quad (2,30)$$

alors :

$$B_1^2 = \sum_{r=2}^{\infty} \delta_n^{2r} \left\{\frac{(-1)^r \left(\frac{1}{2}\right)^{2r+1}}{r!\sqrt{\pi}}\right\} \left\{\sum_{p=1}^{r-1} C_{2r}^{2p}\, [[y^{2p} K]]\, [[y^{2r-2p} K]]\right\} \quad (2,31)$$

$$B_2^2 = \frac{1}{n\delta_n}\left[[[K^2]] - \delta_n \left[\sum_{r=0}^{\infty} \left\{\frac{\delta_n^{2r}(-1)^r \left(\frac{1}{2}\right)^{2r+1}}{r!\sqrt{\pi}}\right\} \left\{\sum_{p=0}^{r} C_{2r}^{2p}\, [[y^{2p} K]]\, [[y^{2r-2p} K]]\right\}\right]\right] \quad (2,32)$$

Remarques : 1) (2,30), (2,31), (2,32) sont toujours vérifiés pour des noyaux nuls en dehors d'un compact.

2) Dans le cas d'un noyau $K(y) = \dfrac{\exp(-|x|^a)}{(2/a)\,\Gamma(1/a)}$, déjà étudié pour (2,14), $[[y^{2n} K]]^{1/2n} \sim \left[\dfrac{2n}{ae}\right]^{1/a}$; on obtient donc que :

Pour $a < 2$, les développements ne sont pas valables ;

Pour $a = 2$ (cas du noyau normal), les développements sont valables pour $|\delta_n| < 1/\sqrt{2}$; (pour un noyau mis sous la forme $K(y) = (2\pi)^{-\frac{1}{2}} \exp(-\frac{1}{2} x^2)$, cette condition devient $|\delta_n| < 1$).

Pour $a > 2$, les développements sont valables $\forall \delta_n$.

En particulier, pour le noyau de Picard, comme dans le cas des développements locaux en un point (condition (2,14) pour la validité de (2,10), et (2,16), (2,17), (2,18)), ou de la validité des expressions comme développements limités (2,20), (2,21), les développements en série (2,31) et (2,32) ne sont pas vérifiés.

Exemple 1 : Moyenne quadratique intégrée, loi N (0 , 1), noyau unité :

$$K(y) = 1 \quad \text{pour} \quad |y| < \tfrac{1}{2} \ , \ [[y^{2p} K]] = \frac{(\frac{1}{2})^{2p}}{2p+1} \ , \ [[K^2]] = [[K]] = 1 ,$$

$$\sum_{p=1}^{r-1} C^{2p} [[y^{2p} K]] [[y^{2r-2p} K]] = \frac{(\frac{1}{2})^{2r}}{(2r+1)(2r+2)} \sum_{p=1}^{r-1} C_{2r+1}^{2p+1} =$$

$$= \frac{1}{2r+1} \left[\frac{1}{r+1} - (\tfrac{1}{2})^{2r-1} \right]$$

$$\sum_{p=0}^{r} C_{2r}^{2p} [[y^{2p} K]] [[y^{2r-2p} K]] = \frac{1}{(2r+1)(r+1)} ,$$

d'où, par (2,31) et (2,32) :

$$B_1^2 = \sum_{r=2}^{\infty} \delta_n^{2r} \left\{ \frac{(-1)^r (\frac{1}{2})^{2r+1}}{r!\,(2r+1)\sqrt{\pi}} \right\} \left\{ \frac{1}{r+1} - (\tfrac{1}{2})^{2r-1} \right\}$$

$$B_2^2 = \frac{1}{n\delta_n} - \frac{1}{n\sqrt{\pi}} \sum_{r=0}^{\infty} \left\{ \delta_n^{2r} \frac{(-1)^r (\frac{1}{2})^{2r}}{r!\,(2r+1)(2r+2)} \right\} \tag{2.33}$$

Ces expressions sont valables pour tout $\delta_n > 0$.

Exemple 2 : Moyenne quadratique intégrée, loi N (0 , 1) noyau normal :

$$K(y) = (2\pi)^{-\frac{1}{2}} e^{-\frac{1}{2}y^2} \ , \ [[y^{2p} K]] = (\tfrac{1}{2})^p \frac{(2p)!}{p!} \ , \ [[K^2]] = \frac{1}{2\sqrt{\pi}},$$

$$\sum_{p=1}^{r-1} C_{2r}^{2p} [[y^{2p} K]] [[y^{2r-2p} K]] = \frac{(2r)!}{2^r r!} \sum_{p=1}^{r-1} C_r^p = \frac{(2r)!}{r!} [1 - (\tfrac{1}{2})^{r-1}]$$

$$\sum_{p=0}^{r} C_{2r}^{2p} [[y^{2p} K]] [[y^{2r-2p} K]] = \frac{(2r)!}{r!} ,$$

d'où, par (2,31) et (2,32) :

$$B_1^2 = \sum_{r=2}^{\infty} \delta_n^{2r} \left\{ \frac{(-1)^r (\frac{1}{2})^{2r+1} (2r)!}{(r!)^2 \sqrt{\pi}} \right\} \left\{ 1 - (\tfrac{1}{2})^{r-1} \right\} =$$

$$= \frac{1}{2\sqrt{\pi}} \left[\frac{1}{\sqrt{1+\delta_n^2}} - \frac{2}{\sqrt{1+\frac{1}{2}\delta_n^2}} + 1 \right]$$

$$B_2^2 = \frac{1}{n\,\delta_n\,(2\sqrt{\pi})}\left[1 - \delta_n \sum_{r=0}^{\infty} \delta_n^{2r} \frac{(-1)^r\,(2r)!}{2^{2r}\,(r!)^2}\right] =$$

$$= \frac{1}{n\,\delta_n\,(2\sqrt{\pi})}\left[1 - \frac{\delta_n}{\sqrt{1+\delta_n^2}}\right], \quad \text{si} \quad \delta_n > 0 \qquad (2.34)$$

Ces développements sont valables si $|\delta_n| < 1$. Une argumentation portant sur l'analycité de B_1^2 et B_2^2 en fonction de δ_n montre que les formules sommées ont une validité étendue à $\delta_n > 0$.

Remarquons que dans ce cas, un calcul direct montre que :

$$E\,(f_n(x)) = \frac{1}{\sqrt{\delta_n^2 + 1}\,\sqrt{2\pi}} \exp\left[-\frac{1}{2}\,x^2 \frac{1}{\delta_n^2 + 1}\right] \qquad (2{,}35)$$

L'espérance de f_n est une densité normale $N\,(0\,,\delta_n^2 + 1)$.

Exemple 3 : Moyenne quadratique intégrée, loi N (0 , 1), noyau d'Epanechnikov :

$$K\,(y) = \frac{3}{4}\,(1 - y^2)\ ,\ |y| < 1\ ;\ [[y^{2p}\,K]] = \frac{3}{(2p+1)(2p+3)}\,,\ [[K^2]] = \frac{3}{5}\,,$$

$$\sum_{p=1}^{r-1} C_{2r}^{2p}\,[[y^{2p}\,K]]\,[[y^{2r-2p}\,K]] = \frac{6}{(2r+1)\,(2r+3)}\left[\frac{2^{2r}\,3}{(r+2)\,(r+3)} - 1\right]$$

$$\sum_{p=0}^{r} C_{2r}^{2p}\,[[y^{2p}\,K]]\,[[y^{2r-2p}\,K]] = \frac{9.2^{2r+1}}{(r+2)\,(r+3)\,(2r+1)\,(2r+3)}\,,$$

$$B_1^2 = \sum_{r=2}^{\infty} \delta_n^{2r} \left\{\frac{9\,(-1)^r}{r!\,(2r+1)\,(2r+3)\sqrt{\pi}}\right\}\left\{\frac{1}{(r+2)\,(r+3)} - \frac{1}{3}\,(\tfrac{1}{2})^{2r}\right\}$$

$$B_2^2 = \frac{1}{n\delta_n}\left\{\frac{3}{5} - \frac{\delta_n}{\sqrt{\pi}}\sum_{r=0}^{\infty} \delta_n^{2r}\left\{\frac{9\,(-1)^r}{r!\,(r+2)\,(r+3)\,(2r+1)\,(2r+3)}\right\}\right\} \qquad (2{,}36)$$

Ces développements sont valables pour tout $\delta_n > 0$.

Exemple 4 : Moyenne quadratique intégrée, loi N (0 , 1), noyau triangulaire :

$$K\,(y) = 1 - |y|\ ,\ |y| < 1\ ;\ [[y^{2p}\,K]] = \frac{2}{(2p+1)(2p+2)}\,,\ [[K^2]] = \frac{2}{3}\,,$$

$$\sum_{p=1}^{r-1} C_{2r}^{2p}\,[[y^{2p}\,K]]\,[[y^{2r-2p}\,K]] =$$

$$= \frac{4}{(2r+1)\,(2r+2)\,(2r+3)\,(2r+4)}\left[2^{2r+3} - (2r+4)\,(2r+3) - 2\right]$$

$$\sum_{p=0}^{r} C_{2r}^{2p} [[y^{2p} K]] [[y^{2r-2p} K]] = \frac{4}{(2r+1)(2r+2)(2r+3)(2r+4)} [2^{2r+3} - 2],$$

d'où :

$$B_1^2 = \sum_{r=2}^{\infty} \delta_n^{2r} \left\{ \frac{4(-1)^r}{(r+2)!\,(2r+1)(2r+3)\sqrt{\pi}} \right.$$
$$\left. \{1 - (\tfrac{1}{2})^{2r+3} (2r+3)(2r+4) - (\tfrac{1}{2})^{2r+2}\} \right\}$$

$$B_2^2 = \frac{1}{n\delta_n} \left[\frac{2}{3} - \frac{\delta_n}{\sqrt{\pi}} \sum_{r=0}^{\infty} \delta_n^{2r} \left\{ \frac{4(-1)^r}{(r+2)!\,(2r+1)(2r+3)} \right\} \{1 - (\tfrac{1}{2})^{2r+2}\} \right] \quad (2,37)$$

Ces développements sont valables pour tout $\delta_n > 0$.

D) Développements par les noyaux de Gram-Charlier :

Le ième polynôme d'Hermite $H_1(x)$ étant défini par (1 , 1), son expression explicite est fournie en (2,11). Les noyaux de Gram-Charlier sont de la forme :

$$K(y) = \sum_{r=0}^{R} c_r H_{2r}(y) \frac{e^{-\frac{1}{2}x^2}}{\sqrt{2\pi}}$$

Ilsspeuvent être également définis sous une forme dissymétrique, faisant intervenir les polynômes d'Hermite impairs, mais présentant moins d'intérêt alors.

On obtient immédiatement que $[[y^{2n+1} K]] = 0$, $\forall n \geqslant 1$; calculons $[[y^{2n} K]]$:

$$y^{2n} = \sum_{p=0}^{n} \left\{ \frac{(2n)!}{(2p)!\, 2^{n-p} (n-p)!} \right\} H_{2p}(y),$$

d'où, par (1,4),

$$[[y^{2n} K]] = (2n)! \sum_{r=0}^{\text{Inf}(n,R)} c_r \left\{ \frac{1}{2^{n-r}(n-r)!} \right\} ; \quad (2,38)$$

écrivons les premiers termes obtenus par cette formule, en posant $c_r = 0$ pour $r > R$:

$$[[K]] = c_0 \quad , \quad [[y^2 K]] = c_0 + 2c_1 \quad , \quad [[y^4 K]] = 3c_0 + 12c_1 + 24c_2$$

$$[[y^6 K]] = 15c_0 + 90c_1 + 360c_2 + 720c_3 \quad ,$$

$$[[y^8 K]] = 105 (c_0 + 8c_1 + 48c_2 + 192c_3 + 384c_4), \ldots,$$

On constate que, si on pose $c_0 = 1$, pour que $[[K]] = 1$, il existe $\forall R, c_1, \ldots, c_R$, tels que $[[y^n K]] = 0$, $\forall 1 \leqslant n \leqslant 2R$.

Les premières valeurs de cette suite sont :

$$c_0 = 1$$

$$c_1 = -\frac{1}{2}$$

$$c_2 = 1/8$$

$$c_3 = -1/48$$

$$c_4 = 1/384 \ldots$$

On obtient l'expression générale de c_n :

$$c_n = \frac{(-1)^n}{2^n\, n!} \tag{2.39}$$

On peut définir, formellement, la somme de la série suivante :

$$K(y) = \sum_{n=0}^{\infty} c_n H_{2n}(y) \frac{e^{-\frac{1}{2}x^2}}{\sqrt{2\pi}} = \sum_{n=0}^{\infty} (-1)^n e^{-\frac{1}{2}x^2} H_{2n}(y) \tag{2,40}$$

K (y) est, malheureusement, divergente, comme on peut le constater d'après (2,11). Quant à la série intégrée, en posant : $g(x) = (2\pi)^{-\frac{1}{2}} e^{-\frac{1}{2}x^2}$, on pourra écrire formellement : $K(y) = \sum_{n=0}^{\infty} \frac{(-\frac{1}{2})^n}{n!} g^{(2n)}(y)$; on en déduit par (2,28), formellement :

$$[[K^2]] = \sum_{n=0}^{\infty} \frac{(\frac{1}{2})^{3n+1}(2n)!}{(n!)^2} \left\{ \sum_{r=0}^{n} \frac{n!}{(n-r)!\, r!} \right\} = \sum_{n=0}^{\infty} \left\{ \frac{(2n)!}{(n!)^2\, 2^{2n+1}} \right\},$$

expression divergente, qu'on retrouve dans le développement :

$$\frac{1}{\sqrt{1-x}} = \sum_{n=0}^{\infty} \frac{(2n)!\; x^r}{(n!)^2\, 2^{2r}}, \quad \text{pour } |x| < 1. \tag{2,41}$$

Ceci montre qu'on ne peut pas ainsi définir une fonction K, telle qu'on puisse, en la projetant sur les sous-espaces engendrés par $H_0, \ldots, H_{2R}$, obtenir le Rième noyau de Gram-Charlier.

Exemples : 1) Noyau de Gram-Charlier d'ordre 1 :

Ce noyau est cité par Rosemblatt [22] (1971) ; il est donné par :

$$K(y) = \tfrac{1}{2}(3 - y^2)(2\pi)^{-\frac{1}{2}} e^{-\frac{1}{2}y^2}, \quad [[K]] = 1,$$

$$= 1,\ [[K^2]] = \frac{27}{32\sqrt{\pi}},\ [[y^{2n} K]] = -\frac{(2n)!}{2^n\, n!}(n-1) \ ; \tag{2,42}$$

2) Noyau de Gram-Charlier d'ordre 2 :

$$K(y) = \frac{1}{8}(y^4 - 10\,y^2 + 15)\frac{e^{-\frac{1}{2}y^2}}{\sqrt{2\pi}},$$

$$[[y^{2p} K]] = (\tfrac{1}{2})^{p+1}\frac{(2p)!}{p!}(p-1)(p-2)\ ,\ [[K^2]] = \frac{2265}{2048\sqrt{\pi}}\ ;\ (2{,}43)$$

Les propriétés asymptotiques de ces noyaux sont discutées plus loin (§ G).

E) Développements par les noyaux de Legendre :

Le nième polynôme de Legendre est défini par :

$$P_n(x) = \frac{1}{2^n\, n!}\frac{d^n}{dx^n}(x^2-1)^n = \frac{1}{2^n\, n!}\sum_{p=0}^{[\frac{1}{2}n]} C_n^p \frac{(-1)^p x^{n-2p}(2n-2p)!}{(n-2p)!} \qquad (2{,}44)$$

Les polynômes de Legendre vérifient :

$$\int_{-1}^{+1} P_m(x)\,P_n(x)\,dx = 0 \quad \text{si} \quad m \neq n\ ,\ \frac{2}{2n+1}\ ,\quad m = n. \qquad (2{,}45)$$

On définit en conséquence :

$$K(y) = \sum_{r=0}^{R} c_r\, P_{2r}(y)\ ,\ \text{si}\ |y| < 1\ ,\ K(y) = 0\ ,\ \text{si}\ |y| > 1.$$

On obtient ainsi les noyaux de Legendre, qui ont comme cas particulier le noyau unité (n° 1). On pourrait également faire intervenir les polynômes impairs, mais ceci ne présenterait que peu d'intérêt.

On constate immédiatement que :

$$\int_{-1}^{1} y^{2n+1} P_{2r}(y)\,dy = 0\ ,\ \text{et que}\ \int_{-1}^{1} y^{2n} P_{2r}(y)\,dy = 0,\ \text{si}\ n < r.$$

Si $n \geqslant r$:

$$\int_{-1}^{1} y^{2n} P_{2r}(y)\,dy = (\tfrac{1}{2})^{2r} C_{2n}^{2r}\int_{-1}^{1} y^{2n-2r}(y^2-1)^{2r}\,dy =$$

$$= (\tfrac{1}{2})^{2r} C_{2n}^{2r}\,\beta(n - r + \tfrac{1}{2}\ ,\ 2r+1)$$

$$= \frac{2^{2r+1}(n+r)!\,(2n)!}{(n-r)!\,(2n+2r+1)!}\ ;\ \text{d'où :}$$

$$[[y^{2n} K]] = \sum_{r=0}^{\mathrm{Inf}(n\,,\,R)} c_r \left\{\frac{2^{2r+1}(n+1)!\,(2n)!}{(n-r)!\,(2n+2r+1)!}\right\}, \qquad (2{,}46)$$

ce qui fournit en particulier :

$$[[K]] = 2c_0 \quad , \quad [[y^2 K]] = \frac{2}{3}\left(c_0 + \frac{2}{5} c_1\right) \quad , \quad [[y^4 K]] = \frac{2}{5}\left(c_0 + \frac{4}{7} c_1 + \frac{8}{63} c_2\right)$$

$$[[y^6 K]] = \frac{2}{7}\left(c_0 + \frac{2}{3} c_1 + \frac{8}{33} c_2 + \frac{16}{429} c_3\right) ,$$

$$[[y^8 K]] = \frac{2}{9}\left(c_0 + \frac{8}{11} c_1 + \frac{48}{143} c_2 + \frac{64}{715} c_3 + \frac{128}{12\,155} c_4\right)$$

. . .

On constate, qu'en posant $c_0 = \frac{1}{2}$, pour que $[[K]] = 1$, il existe $\forall R$, $c_1, \ldots, c_R$, tels que $[[y^n K]] = 0 \quad , \quad \forall 1 \leqslant n \leqslant 2R$. Les premières valeurs de cette suite sont :

$$c_0 = \frac{1}{2} \quad , \quad c_1 = -5/4 \quad , \quad c_2 = 27/16 \quad , \quad c_3 = -65/32 \quad , \quad c_4 = 2025/256, \ldots$$

Cette suite a une limite infinie en oscillant ; la série obtenue à partir de (,), en posant $R = \infty$ est divergente.

Les premiers polynômes de Legendre sont donnés par :

$$P_0(x) = 1$$

$$P_1(x) = x$$

$$P_2(x) = \frac{1}{2}(3x^2 - 1)$$

$$P_3(x) = \frac{1}{2}(5x^3 - 3x)$$

$$P_4(x) = \frac{1}{8}(35\,x^4 - 30\,x^2 + 3)$$

$$P_5(x) = \frac{1}{8}(63\,x^5 - 70\,x^3 + 15\,x)$$

$$P_6(x) = \frac{1}{16}(231\,x^6 - 315\,x^4 + 105\,x^2 - 5)$$

$$P_7(x) = \frac{1}{16}(429\,x^7 - 693\,x^5 + 315\,x^3 - 35\,x)$$

$$P_8(x) = \frac{1}{128}(6\,435\,x^8 - 12\,012\,x^6 + 6\,930\,x^4 - 1\,260\,x^2 + 35)\,, \ldots \qquad (2,47)$$

Plutôt que par l'usage de (2,44), on construit plus aisément ces polynômes par la formule de récurrence (voir Sansone [24], p. 178) :

$$(n + 1)\,P_{n+1}(x) - (2n + 1)\,xP_n(x) + nP_{n-1}(x) = 0 \qquad (2,48)$$

Les propriétés suivantes sont vérifiées par les polynômes de Legendre :

$$P_n(1) = 1 \quad , \quad P_n(-1) = (-1)^n$$

Pour $-1 \leqslant x \leqslant 1$, on a la majoration de Stieljes ([24], p. 199) :

$$|P_n(x)| \leqslant \frac{4\sqrt{2}}{\sqrt{\pi n}\,(1-x^2)^{\frac{1}{4}}}$$

Exemples : 1) Noyau de Legendre d'ordre 1 :

$$K(y) = \frac{3}{8}(3 - 5y^2) = \tfrac{1}{2}P_0(y) - \frac{5}{4}P_2(y) \quad , \quad |y| < 1, \qquad (2,49)$$

$$[[K]] = 1 \quad , \quad [[y^{2n+1}K]] = 0 \quad , \quad \forall n \geqslant 1 \quad , \quad [[y^2 K]] = 0 \quad , \quad [[K^2]] = 9/8$$

$$[[y^{2n}K]] = -\frac{3(n-1)}{(2n+1)(2n+3)} \quad \text{pour } n \geqslant 1 \;;$$

asymptotiquement lorsque $\delta_n \to 0$,

$$B_1^2 \sim \frac{\delta_n^8}{(4!)^2}\frac{9}{(1\,225)}\left\{\int_I (f^{(4)}(x))^2\,dx\right\} \quad , \quad B_2^2 \sim \frac{9}{8n\,\delta_n}$$

Ces formules sont à rapprocher de celles qu'on obtient pour le noyau unité (2,33) et (2,55) ; en utilisant la même échelle, c'est-à-dire ici $K'(y) = \frac{1}{2}$ si $|y| < 1$, on obtient le noyau de Legendre d'ordre 0, qui fournit asymptotiquement :

$$B'^2 \sim \frac{\delta_n^4}{(2!)^2}\frac{1}{9}\left\{\int_I (f''(x))^2\,dx\right\} \quad , \quad B'^2 \sim \frac{1}{2n\,\delta_n}$$

Dans le premier cas, l'erreur quadratique minimale est obtenue pour :

$$\delta_n = \left[\frac{5^2 7^2 3^2}{\int_I (f^{(4)}(x))^2\,dx}\right]^{\frac{1}{9}} n^{-1/9},$$

d'où
$$M^2 \sim \frac{81}{64n\,\delta_n} = \frac{81}{64}\left[\frac{\int_I (f^{(4)}(x))^2\,dx}{5^2 7^2 3^2}\right]^{\frac{1}{9}} n^{-8/9},$$

Dans le deuxième cas

$$\delta'_n = \left[\frac{9}{2\int_I (f''(x))^2\,dx}\right]^{\frac{1}{5}} n^{-1/5},$$

d'où :
$$M'^2 \sim \frac{5}{8n\,\delta'_n} = \frac{5}{8}\left[\frac{2\int_I (f''(x))^2\,dx}{9}\right]^{\frac{1}{5}} n^{-4/5}.$$

Pour la loi normale N (0 , 1), par (2,28) :

$$M'^2 \sim \frac{5}{8}\left[\frac{1}{12\sqrt{\pi}}\right]^{1/5} n^{-4/5} \quad , \text{ et } \; M^2 \sim \frac{81}{64}\left[\frac{1}{5.7.3.2^5\sqrt{\pi}}\right]^{1/9} n^{-8/9}$$

On constate alors que M^2 ne devient inférieur à M'^2 que pour n # 52, en considérant les formules asymptotiques précédentes et en négligeant les termes d'ordre supérieur. Cette remarque prouve pleinement l'intérêt des noyaux de ce type pour le critère du MISE ; il faut cependant noter que, pour une distribution différente de la loi normale, ou pour l'estimation de la densité en un point, ces noyaux auront des performances variables suivant les cas.

2) Noyau de Legendre d'ordre 2 :

$$K(y) = \frac{15}{128}(15 - 70y^2 + 63y^4) = \frac{1}{2}P_0(y) - \frac{5}{4}P_2(y) + \frac{27}{16}P_4(y)\,,\ |y| < 1, \quad (2,50)$$

$$[[K]] = 1 \;,\; [[y^{2n+1}K]] = 0\,, \quad n \geqslant 1\,,$$

$$[[y^2K]] = 0 \;,\; [[y^4K]] = 0 \;,\; [[K^2]] = \frac{7425}{4096}$$

$$[[y^{2n}K]] = \frac{15(n-1)(n-2)}{2(2n+1)(2n+3)(2n+5)} \quad \text{pour} \quad n \geqslant 1\,;$$

asymptotiquement lorsque $\delta_n \to 0$,

$$B_1^2 \sim \frac{\delta_n^{12}}{11^2.7^2.3^6.2^8}\left\{\int_I (f^{(6)}(x))^2 dx\right\} \quad , \quad B_2^2 \sim \frac{11.5^2.3^3}{2^{12}\, n\delta_n}$$

Asympotiquement l'erreur quadratique minimale est obtenue pour :

$$\delta_n = \left[\frac{11^3.7^2.5^2.3^8}{2^6\int_I (f^{(6)}(x))^2\,dx}\right]^{\frac{1}{13}} n^{-1/13}$$

d'où $$M^2 \sim \frac{13.11.5^2.3^2}{2^{14}}\left[\frac{2^6}{11^3\,7^2\,5^2\,3^8}\int_I (f^{(6)}(x))^2 dx\right]^{\frac{1}{13}} n^{-12/13}$$

Pour la loi normale N(0 , 1), par (2,28) :

$$M^2 \sim \frac{13.11.5^2.3^2}{2^{14}}\left[\frac{1}{11^2.7.5.3^5.2\sqrt{\pi}}\right]^{\frac{1}{13}} n^{-12/13}$$

On constate que M^2 ne devient ici inférieur à M'^2 que pour n # 89 ; la même remarque que précédemment peut être appliquée à ce noyau .

F) Développements par les noyaux de Laguerre :

Le n^{ieme} polynôme de Laguerre est défini à l'ordre a, par : ($a \in \mathbf{R}$, $a > -1$)

$$P_{n,a}(x) = \frac{x^{-a} e^{x}}{n!} \frac{d^n}{dx^n} [x^{n+a} e^{-x}] =$$

$$(-1)^n \sum_{m=0}^{n} \frac{\Gamma(a+n+1)(-1)^m}{m!(n-m)!\, \Gamma(a+n-m+1)} x^{n-m}, n \geqslant 1,$$

$$P_{0,a}(x) = 1$$

Les polynômes de Laguerre vérifient :

$$\int_0^{+\infty} e^{-x} x^a P_{n,a}(x) P_{m,a}(x) = 0 \quad \text{si } m \neq n \quad , \quad \frac{\Gamma(n+a+1)}{n!} \quad , \quad \text{si } m = n.$$

On définit les noyaux de Laguerre, par :

$$K(y) = \sum_{r=0}^{R} c_r P_{r,a}(|y|)\, |y|^a\, e^{-|y|} \tag{2,51}$$

Remarquons qu'on obtient ainsi comme cas particulier le noyau de Picard ($a = 0$, $R = 0$, n° 8). Nous étudierons ici principalement les noyaux de Laguerre d'ordre 0, bien que les résultats puissent être étendus sans difficulté aux noyaux d'ordre a quelconque.

Pour ces noyaux :

$$P_n(x) = P_{n,0}(x) = (-1)^n \sum_{m=0}^{n} C_n^m \frac{(-1)^m x^m}{m!},$$

et : $[[y^{2n+1}K]] = 0 \quad , \quad \forall n \geqslant 1 \quad , \quad [[y^{2n}K]] = 2((2n!)^2 \sum_{r=0}^{\mathrm{Inf}(R,2n)} \frac{(-1)^r}{r!(2n-r)!} c_r$

On constate, qu'en posant $c_0 = \frac{1}{2}$, pour que $[[K]] = 1$, il existe $\forall R, c_1, \ldots, c_R$, tels que $[[y^n K]] = 0$, $\forall\, 1 \leqslant n \leqslant \frac{1}{2} R$. La solution est la suite : $c_r = \frac{1}{2}$ on obtiendra donc :

$$K_n(y) = \frac{1}{2} \sum_{r=0}^{2n} P_r(|y|)\, e^{-|y|} \tag{2,52}$$

La série ainsi obtenue n'est pas convergente, comme le montre l'approximation d'Uspensky (voir [24] p. 348) :

$$P_r(x) = \sqrt{\frac{2}{\pi}} \left[\frac{e^{\frac{1}{2}} x}{((4n+1)x)^{1/4}} \right] \cos \left\{ \sqrt{(4n+1)x} - \frac{\pi}{4} \right\} + 0 \left[\frac{x^{3/2}}{\sqrt{4n+1}} \right]$$

uniformément en $n > 1$, sur (a, b) fixé.

Exemples : 1) Noyau de Laguerre d'ordre 0 :

Il s'agit du noyau de Picard : (n° 8) $K(y) = \frac{1}{2} e^{-|y|}$; les résultats de Watson et Leadbatter [14] montrent que ce noyau est optimal pour la loi exponentielle et le critère du MISE, avec $\delta_n = n^{-\frac{1}{2}}$. On obtient ici :

$$[[K]] = 1 \ , \ [[K^2]] = \frac{1}{4} \ , \ [[y^{2n+1}K]] = 0 \ , \ \forall n \geqslant 1 \ , \ [[y^{2n}]] = (2n)! \ ; \qquad (2,53)$$

Asymptotiquement lorsque $\delta_n \to 0$,

$$B_1^2 \sim \delta_n^4 \int_I (f''(x))^2 dx \quad , \quad B_2^2 \sim \frac{1}{4n\,\delta_n} \ ;$$

l'erreur quadratique minimale est asymptotiquement obtenue pour :

$$\delta_n = \left[\frac{1}{16\int_I (f''(x))^2 dx}\right]^{\frac{1}{5}} n^{-1/5} \ ,$$

d'où
$$M^2 \sim \frac{5}{16n\,\delta_n} = \frac{5}{16}\left[16\int_I (f''(x))^2 dx\right]^{\frac{1}{5}} n^{-4/5}$$

Il ne faut pas être surpris du résultat pour l'estimation de la loi exponentielle, puisque, cette loi ayant une densité discontinue à l'origine, (2,19) n'est plus valable avec $I = \mathbf{R}$.

2) Noyau de Laguerre d'ordre 1 :

$$K(y) = \frac{1}{2}\left(\frac{1}{2} y^2 - 3|y| + 3\right) e^{-|y|} \ , \ [[K]] = 1, [[K^2]] = \frac{3}{2} \ , \ [[y^{2n}K]] = (2n)!\,(n-1)(2n-1). \qquad (2,54)$$

Asymptotiquement lorsque $\delta_n \to 0$,

$$B_1^2 \sim 9\,\delta_n^8 \int_I (f^{(4)}(x))^2 dx \quad , \quad B_2 \sim \frac{3}{2n\,\delta_n};$$

l'erreur quadratique minimale est asymptotiquement obtenue pour :

$$\delta_n = \left[\frac{1}{48\int_I (f^{(4)}(x))^2 dx}\right]^{\frac{1}{9}} n^{-1/9} \ ,$$

d'où
$$M^2 = \frac{27}{16}\left[48\int_I (f''(x))^2 dx\right]^{\frac{1}{9}} n^{-8/9}$$

Comme on pourra le constater dans le § G, les noyaux de Laguerre ont une très mauvaise efficacité asymptotique. Leur intérêt est de ce fait considérablement réduit.

G) Comportement asymptotique des estimations, critère du M.I.S.E. :

On suppose ici que les hypothèses de la proposition (2,5) sont vérifiées, et on se limite aux premiers termes des développements (2,21) et (2,22) :

1) Noyau symétrique tel que $[[y^2K]] > 0$;

$$B_1^2 \sim \frac{1}{4} \delta_n^4 \, [[y^2K]]^2 \left\{ \int_I (f''(x))^2 \, dx \right\} \quad , \quad B_2^2 \sim \frac{1}{n\,\delta_n} [[K^2]]$$

Comme M^2 est de la forme $M^2 = A\delta^4 + \frac{B}{n\delta}$, on obtient la valeur minimale de M^2 en annulant la dérivée de cette expression, soit $\delta^5 = \frac{B}{4An}$; on obtient donc ici :

$$\delta_n = \left[\frac{[[K^2]]}{[[y^2K]]^2 \int_I (f''(x))^2 \, dx} \right]^{\frac{1}{5}} n^{-1/5} \quad ,$$

d'où $$M^2 \sim \frac{5\,[[K^2]]}{4n\,\delta_n} = \frac{5}{4} \, [[K^2]]^{\frac{5}{4}} \, [[y^2K]]^{\frac{2}{5}} \left[\int_I (f''(x))^2 \, dx \right]^{\frac{1}{5}} n^{-4/5} \qquad (2,55)$$

Certains auteurs, notamment Rosenblatt ([22], (1971), p. 1819), préfèrent un autre choix de δ que celui fourni en (2,55), et obtenu en écrivant que :

$B_1^2 \sim B_2^2$, ce qui fournit, avec les notations précédentes $\delta^5 = \frac{B}{An}$,

$$\delta'_n = \left[\frac{4[[K^2]]}{[[y^2K]]^2 \int_I (f''(x))^2 \, dx} \right]^{\frac{1}{5}}$$

d'où $$M'^2 \sim \frac{2\,[[K^2]]}{n\,\delta'_n} = 2^{\frac{3}{5}} \, [[K^2]]^{\frac{4}{5}} \, [[y^2K]]^{\frac{2}{5}} \left[\int_I (f''(x))^2 \, dx \right]^{\frac{1}{5}} n^{-4/5} \qquad (2,56)$$

Un tel choix conduit en fait à une erreur quadratique

$$M'^2 \sim \frac{2^{13/5}}{5} M^2 \# (1{,}2125) M^2 .$$

Il n'est donc absolument pas justifié, et des considérations de robustesse incitant à choisir systématiquement une valeur de δ_n inférieure à la valeur optimale fournie en (2,55), il paraît dangereux de prendre une valeur $\delta'_n = 2^{2/5} \, \delta_n$. Nous en resterons donc à (2,55) tout en signalant l'existence d'autre choix pouvant être liés à des optimisations différentes que celle fournie par la recherche du M.I.S.E.

Remarquons que le choix de Rosenblatt peut être partiellement justifié lorsqu'on constate que pour le choix optimal (2,55), on a :

$B_1^2 \sim \frac{1}{4} B_2^2$, le choix de δ_n réduit le terme correspondant à la variance intégrée, tout en augmentant le terme correspondant au biais. Or, pour une meilleure convergence dans L^∞, il est possible de constater (voir [5] [34]) que précisément, il est nécessaire d'attacher une importance plus grande à la partie aléatoire de l'estimation, et à augmenter la valeur de δ_n.

Le type de discussion aboutissant a (2,55) ou (2,56) est ancien et se trouve déjà dans le premier article de Rosenblatt [21] (1956) ; il a été repris dans la plupart des articles postérieurs. La recherche du noyau asymptotiquement optimal (pour une estimation connue de $\int_I (f''(x))^2 \, dx$) se ramène alors à la recherche d'un noyau K, tel que $[[K^2]]$ soit minimal, avec :

$$[[K]] = 1 \quad , \quad K(y) = K(-y) \quad , \quad \forall\, y \in \mathbf{R} \quad , \quad [[y^2 K]] = 1. \qquad (2,57)$$

Cette discussion a été menée à terme par le calcul des variations par Epanechnikov [8] (1969), et reprise par Rosenblatt [22] (1971) ; on retrouve la même méthode utilisée par Lehmann et Hodges [22] (1956). Le noyau solution du système (2,57), et aboutissant parmi tous les noyaux positifs à un minimum de $[[K^2]]$ est :

$$K(y) = \frac{3}{4\sqrt{5}}(1 - y^2/5) \; ; \qquad (2\ 58)$$

on retrouve le noyau n° 2, d'Epanechnikov.

Epanechnikov fait ensuite une recherche numérique sur l'efficacité asymptotique des estimations, mesurée relativement à l'erreur quadratique obtenue par (2.56). Il obtient notamment que cette efficacité reste très voisine de 1 pour les noyaux unité n° 1), triangulaire (n° 11), normal (n° 5), et introduit un noyau :

$$K(y) = \tfrac{1}{2} \cos y \quad , \quad |y| < \tfrac{1}{2}\,\pi\,, \qquad (2,59)$$

dont l'efficacité est meilleure que celle des précédents noyaux cités, et enfin, le noyau de Picard, dont les performances asymptotiques sont ici les moins bonnes. Il fournit dans une table, partiellement reproduite dans [22], les valeurs obtenues de :

$$r = \frac{[[K^2]]\,[[y^2 K]]^{\frac{1}{2}}}{[[K_0^2]]\,[[y^2 K_0]]^{\frac{1}{2}}} \quad , \quad K_0 \text{ étant le noyau (2,58) (n° 2).} \qquad (2,60)$$

Il serait plus agréable de tabuler $E = r^{-4/5}$, qui fournit le vrai rapport d'efficacité entre M^2 et M_0^2 (M.I.S.E) ; nous donnons ici un tel tableau d'efficacité, qui pourra également être utilisé à titre de référence .

Ce tableau d'efficacité montre des résultats réels un peu moins optimistes. Il n'en demeure pas moins que, relativement au critère du M.I.S.E., il n'y a pas de différence discernable en pratique, pour des échantillons moyens et grands, entre les résultats fournis par les estimations du noyau avec les noyaux :

d'Epanechnikov (n° 2), triangulaire (n° 11), normal (n° 5), unité (n° 1).

Noyau	$[[y^2 K]]$	$[[K^2]]$	Efficacité
$K(y) = \frac{3}{4}(1-y^2)$, $\lvert y\rvert < 1$	$\frac{1}{5}$	$\frac{3}{5}$	1
$K(y) = 1$, $\lvert y\rvert < \frac{1}{2}$	$\frac{1}{12}$	1	94,24 %
$K(y) = 1 - \lvert y\rvert$, $\lvert y\rvert < 1$	$\frac{1}{6}$	$\frac{2}{3}$	98,81 %
$K(y) = (2\pi)^{\frac{1}{2}} e^{-\frac{1}{2}y^2}$	1	$\frac{1}{2\sqrt{\pi}}$	96,10 %
$K(y) = \frac{1}{2} e^{-\lvert y\rvert}$	2	$\frac{1}{4}$	80,51 %
$K(y) = \frac{1}{2}\cos y$, $\lvert y\rvert < \frac{1}{2}$	$\frac{1}{4}(\pi^2 - 8)$	$\frac{\pi}{8}$	99,99 %
$K(y) = \pi^{-1}(1+y^2)^{-1}$	∞	$\frac{1}{2\pi}$	0
$K(y) = 2\pi^{-1}(1+y^2)^{-2}$	1	$\frac{5}{4\pi}$	73,80 %

La simplicité du calcul des estimations incite à choisir un noyau nul en dehors d'un compact, et de préférence unitaire ou triangulaire ; ces deux noyaux permettent de plus une construction facile de l'histogramme.

Remarquons que ce qui précède est aussi bien valable pour le choix de δ_n défini par (2,55), que par (2,56). Nous généraliserons ces propriétés plus loin dans l'étude de la robustesse des estimations (§ I).

2) Noyau symétrique tel que $[[y^2 K]] = \cdots = [[y^{2p-2} K]] = 0$, $[[y^{2p} K]] > 0$:

On obtient de tels noyaux par exemple par les noyaux de Legendre, Gram-Charlier, Laguerre, étudiés précédemment. On obtient :

$$B_1^2 \sim \frac{1}{((2p)!)^2} \delta_n^{4p} [[y^{2p} K]]^2 \left\{ \int_I (f^{(2p)}(x))^2 \, dx \right\} \quad , \quad B_2^2 \sim \frac{1}{n\delta_n} [[K^2]] \quad (2,61)$$

Le même raisonnement que dans (2 55), fournit une valeur asymptotiquement optimale de la forme :

$$\delta_n = \left[\frac{((2p)!)^2 \, [[K^2]]}{4p \, [[y^{2p} K]]^2 \int_I (f^{(2p)}(x))^2 \, dx} \right]^{\frac{1}{4p+1}} n^{-1/(p+1)}$$

$$M^2 \sim \frac{(4p+1) \, [[K^2]]}{4pn\delta_n} = \tag{2,62}$$

$$= \frac{(4p+1)}{4p} [[K^2]]^{\frac{4p}{4p+1}} [[y^{2p}]]^{\frac{2}{4p+1}} \left\{ \int_I (f^{(2p)}(x))^2 \, dx \right\}^{\frac{1}{4p+1}} n^{-4p/(4p+1)}$$

Naturellement les noyaux qui vérifient ces propriétés ne peuvent être positifs partout.

Un raisonnement analogue à celui d'Epanechnikov, par le calcul des variations inciterait à choisir comme noyaux asymptotiquement optimaux des noyaux polynomiaux de degré p ; les noyaux de Legendre répondent aux conditions précédentes ; malheureusement, on ne peut en déduire des noyaux optimaux, puisqu'avec des contraintes analogues à (2,57), $[[K^2]]$ n'est pas borné inférieurement.

Nous nous contentons ici de comparer les propriétés asymptotiques des trois noyaux introduits dans les § D, E, F précédents :

a) $p = 2$;

Par (2,56), M^2 sera dans chaque cas proportionnel à $C = [[K^2]]^{\frac{8}{9}} [[y^2K]]^{\frac{2}{9}}$; nous tabulons :

Noyau	$[[y^4K]]$	$[[K^2]]$	C
$K(y) = \frac{3}{8}(3 - 5y^2)$, $\lvert y\rvert < 1$	$-\frac{3}{35}$	$\frac{9}{8}$	0,4283
$K(y) = \frac{\frac{1}{2}(3 - y^2)}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2}$	-3	$\frac{27}{32\sqrt{\pi}}$	0,6077
$K(y) = \frac{1}{2}(\frac{1}{2}y^2 - 3\lvert y\rvert + 3) e^{-\lvert y\rvert}$	72	$\frac{3}{2}$	2,306

b) $p = 3$;

Par (2,62), M^2 sera dans chaque cas proportionnel à $D = [[y^6 K]]^{\frac{2}{13}} [[K^2]]^{\frac{12}{13}}$
Nous tabulons :

Noyau	$[[y^6K]]$	$[[K^2]]$	D
$K(y) = \frac{15}{128}(15 - 70\,y^2 + 63\,y^4)$, $\lvert y\rvert < 1$	$\frac{5}{231}$	$\frac{7\,425}{4\,096}$	0,9602
$K(y) = \frac{1}{8}(y^4 - 10\,y^2 + 15)\frac{e^{-\frac{1}{2}y^2}}{\sqrt{2\pi}}$	15	$\frac{2\,265}{2\,048\sqrt{\pi}}$	0,9815

H) Moyenne quadratique intégrée (autres méthodes)

Leadbatter et Watson [14], ont utilisé la transformation de Fourier et la formule de Parseval pour calculer M^2. Nous résumons ici leur résultat principal :

Posons :

$$\varphi_f(u) = \int_{-\infty}^{+\infty} f(x)\, e^{iux}\, dx \quad , \quad \varphi_K(u) = \int_{-\infty}^{+\infty} K(y)\, e^{iuy}\, dy \, ,$$

on a :

$$M^2 = E\left[\int_{-\infty}^{+\infty} (f_n(x) - f(x))^2\, dx\right] =$$

$$= \frac{1}{2\pi} E\left[\int_{-\infty}^{+\infty} \left| \left\{ \frac{1}{n}\varphi_K(\delta_n u) \sum_{r=1}^{n} e^{iX_r u} \right\} - \varphi_f(u)\right|^2 du\right]$$

$$= \frac{1}{2\pi}\left[\int_{-\infty}^{+\infty} \left\{ \frac{1}{n} + \left(\frac{n-1}{n}\right)\lvert\varphi_f(u)\rvert^2 \right\} \left| \varphi_K(\delta_n u) - \frac{\lvert\varphi_f(u)\rvert^2}{(1/n) + ((n-1)/n)\lvert\varphi_f(u)\rvert^2}\right|^2 du + \int_{-\infty}^{+\infty} \frac{\lvert\varphi_f(u)\rvert^2 (1 - \lvert\varphi_f(u)\rvert^2)}{1 + (n-1)\lvert\varphi_f(u)\rvert^2}\, du\right] ; \qquad (2,63)$$

On en déduit que M^2 est minimal

pour :

$$\varphi_K(\delta_n u) = \frac{\lvert\varphi_f(u)\rvert^2}{(1/n) + ((n-1)/n)\lvert\varphi_f(u)\rvert^2} \qquad (2,64)$$

Malheureusement l'inversion de cette formule n'est simple que dans quelques cas particuliers, dont la loi de Cauchy et la loi exponentielle ; pour cette dernière, on obtient :

$$K(x) = e^{-|x|}\ ,\ \delta_n = \frac{1}{\sqrt{n}}\ ; \tag{2,65}$$

la décroissance de δ_n en $n^{\frac{1}{2}}$ est d'ailleurs une propriété générale des suites optimales associées aux densités discontinues en un nombre fini de points.

On peut néanmoins faire une remarque intéressante à partir de la formule de Leadbatter et Watson :

Propostion (2,8) : a) Le noyau qui minimise M^2 est toujours symétrique ($K(y) = K(-y)$). b) Si K_0 est un noyau symétrique, et si on considère l'estimation obtenue en prenant le noyau $K(x) = K_0(x + r)$, la valeur de r qui minimise M^2 est $r = 0$, si $\varphi_{K_0}(u) \geqslant 0$, $\forall u \in \mathbf{R}$. Preuve : a est trivial, b s'obtient en constatant que si $a \geqslant 0$, $b \geqslant 0$, $|ae^{i\theta} - b|$ est minimum pour $\theta = 0$.

Ce résultat fait, avec les développements asymptotiques des § B et G, que la plupart des noyaux utilisés sont symétriques. Il faut bien se rendre compte cependant que l'optimalité des noyaux symétriques n'est réalisée que pour le critère du M.I.S.E. sur $\mathbf{R}$, ou asymptotiquement pour $n\infty$. Il faut aussi se rendre compte que les noyaux considérés ne dépendent pas de x ; on obtient, en se libérant des conditions qui mènent à (2,64), des estimations ayant une erreur quadratique moyenne très diminuée ; nous développerons ces points dans une note ultérieure.

Essentiellement pour des motifs numériques, Specht (Technometries 13, 1971, p. 409-24) a étudié le développement de (2,1) directement, c'est-à-dire par la formule :

$$K\left[\frac{X - x}{\delta}\right] = \sum_{r=0}^{\infty} \frac{1}{r!}\left[\frac{X}{\delta}\right]^r K^{(r)}\left[-\frac{x}{\delta}\right]$$

Il étudie en particulier le noyau normal, appliqué à une distribution normale, et obtient des développements analogues à (2,34).

I) Choix optimal du noyau, de δ_n.

Dans le § G, nous avons obtenu en (2,56) et (2,62), des équivalents asymptotiques de M^2, en fonction du choix optimal de δ_n, le noyau K étant fixé. Ce choix fait apparaître le paramètre essentiel :

$$\theta = \left\{ \int_I (f^{(2p)}(x))^2\, dx \right\}^{\frac{1}{4p+1}} \tag{2,65}$$

Or, ce paramètre est, en général inconnu. On a donc à affronter l'alternative :

– Ou bien estimer θ ; pour celà, il est nécessaire d'estimer $f^{(2p)}$, et les estimations de ce type convergent en général très lentement (voir Malz (1975) [15] Schuster (1969) [25], Dimitriev-Tarasenko (1973) [6]). La méthode de Woodroofe (1970) [35], est à rapprocher de ce principe ; on constate en pratique que la taille des échantillons nécessaires est trop élevée, eu égard aux échantillons usuels.

– Ou bien, estimer θ en faisant une hypothèse paramétrique sur f ; les résultats, notamment du § C fournissent par exemple les développements correspondant à une distribution normale ; or, en pratique, les histogrammes sont souvent fabriqués pour vérifier une hypothèse paramétrique sur une distribution ; dans le cas, par exemple de la loi normale, on peut estimer θ en estimant la variance σ et en utilisant la formule :

$$\int_{\mathbf{R}} (f^{(2p)}(x))^2 \, dx = \frac{(\frac{1}{2})^{4p+1}(4p)!}{(2p)! \sqrt{\pi}} \sigma^{4p+1}, \tag{2,67}$$

pour une loi $N(m, \sigma^2)$.

Pour des motifs de robustesse, il est préférable d'utiliser à cet effet des estimations de σ par statistiques ordonnées, plutôt que la variance empirique s^2.

En ce qui concerne les mélanges de lois normales, ce type de procédé convient tout à fait, du fait que si :

$$f(x) = \sum_{r=1}^{p} \lambda_r \frac{e^{-\frac{1}{2}(x-m_r)\sigma^{-2}}}{\sigma\sqrt{2\pi}} \quad \text{avec} \quad \sum_{r=1}^{p} \lambda_r = 1,$$

$$\int_{\mathbf{R}} (f^{(2p)}(x))^2 \, dx \leqslant \left[\sum_{r=1}^{p} \lambda_r \right]^2 \frac{(\frac{1}{2})^{4p+1}(4p)!}{(2p)! \sqrt{\pi}} \sigma^{4p+1} = \frac{(\frac{1}{2})^{4p+1}(4p)}{(2p)! \sqrt{\pi}} \sigma^{4p+1}$$

La question qui se pose alors est la suivante : faut-il utiliser directement (2,67) dans (2,65) ? Quelle est l'influence d'une erreur dans l'estimation de θ sur le M.I.S.E. ?

Posons $\lambda = \dfrac{\theta \text{ estimé}}{\theta}$; par (2,62), on obtient alors pour M^2 :

$$M^2(\lambda) = M^2(1) \frac{4p}{4p+1} \left\{ \frac{\lambda^{4p}}{4p} + \frac{1}{\lambda} \right\} = M^2(1) G_p(\lambda) \tag{2,68}$$

Il est alors facile de constater que la fonction $G_p(\lambda)$ possède la propriété remarquable de rester très voisine de 1 lorsque $\lambda \leqslant 1$ n'est pas trop petit, alors qu'elle croît très rapidement pour $\lambda > 1$; donnons quelques valeurs numériques :

λ	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1,00
$G_1(\lambda)$	8,00	4,00	2,67	2,01	1,61	1,37	1,20	1,10	1,05	1,00

λ	1,1	1,2	1,3	1,4	1,5	1,6	1,7	1,8	1,9	2,00
$G_1(\lambda)$	1,02	1,08	1,19	1,34	1,80	2,14	2,56	3,07	3,68	4,40

λ	2,1	2,2	2,3	2,4	2,5
$G_1(\lambda)$	5,54	6,22	7,34	8,63	10,08

Cette propriété incite à suivre la procédure suivante :

1) Si, sous l'hypothèse nulle, la loi de $\hat{\theta}$, estimation de θ est connue, remplacer $\hat{\theta}$ par $\hat{\theta}' = k\,\theta$, de manière que $E(M^2)$ soit minimal, c'est-à-dire :

$$k = \left[\frac{E(\theta/\hat{\theta})}{E((\hat{\theta}/\theta)^4)}\right]^{1/5} \tag{2,69}$$

Exemple : Pour une loi normale $N(m, \sigma^2)$, $\hat{\theta}/\theta = \frac{s}{\sigma} \equiv \sqrt{\frac{\chi^2(n-1)}{n}}$: on obtient alors la valeur :

$$k = \left[\frac{1}{\left[1 - \frac{1}{n^2}\right]} \frac{\sqrt{n}\,\Gamma(\frac{1}{2}(n-2))}{\sqrt{2}\,\Gamma(\frac{1}{2}(n-1))}\right]^{\frac{1}{5}} \tag{2,70}$$

2) Si, dans le cas précédent, ou, plus généralement, il y a une forte possibilité que l'estimation $\hat{\theta}$ soit très différente de θ, on adopte la même procédure que précédemment, en attribuant à $\hat{\theta}/\theta$ une distribution uniforme sur un intervalle de la forme $[(1/R, R]$, $(R > 1)$. R est un paramètre de tolérance fixé à l'avance ; on obtient alors :

$$k = R\left[\frac{10 \operatorname{Log} R}{R^{10} - 1}\right]^{\frac{1}{5}} \tag{2,71}$$

Par exemple pour $R = 2, \quad k = 0{,}74$

$R = 3, \quad k = 0{,}54.$

Cette procédure incide donc à prendre δ_n plus petit que la valeur asymptotiquement optimale.

3) Les formules précédentes sont établies en fonction de (2,66), $p = 1$, et se généralisent pour p quelconque, et pour les estimations correspondantes. Elles sont obtenues en utilisant les équivalents asymptotiques (2,59). Lorsque n est petit, on aura intérêt à utiliser, plutôt que ces développements asymptotiques, les développements exacts obtenus par exemple dans le § C. On effectuera alors une recherche par des méthodes numériques du coefficient k, en admettant que, par exemple (2,67) est vrai pour tout p, et que σ peut avoir une valeur arbitraire dans un intervalle $[\sigma_0/R, \sigma_0 R]$, R étant une valeur estimée.

III. CONCLUSION

La procédure précédente s'applique sans grande difficulté pour les petits échantillons des lois à support dans **R**. Pour les lois à support fini, les bavures observées hors du support pour les fortes valeurs de δ_n usuelles aux petits échantillons et dues à la décroissance lente vers 0 de $n^{-1/5}$ ne permettent pas d'utiliser ces résultats, sauf sur la partie centrale de la distribution. Il faut alors, soit prendre un taux de décroissance plus rapide (par exemple 2,65)), soit modifier les noyaux en fonction de leur position par rapport au support. Nous avons développé de telles estimations, faisant appel aux statistiques ordonnées de l'échantillon, et permettant d'obtenir des histogrammes dont l'efficacité, mesurée en termes de M.I.S.E. est du même ordre que l'efficacité usuelle pour des estimations régulières. Cette méthode sera exposée dans une prochaine note.

IV. APPLICATIONS PRATIQUES

Des mesures de pollution effectuées en Italie pour le compte de la Compagnie Française des Pétroles, notamment au niveau de la raffinerie de Mantoue, ont fourni des échantillons de taille petite $10 \leqslant n \leqslant 50$, dont l'interprétation statistique est rendue délicate par leur variabilité.

Diverses hypothèses paramétriques ont été envisagées, notamment en supposant que les taux mesurés suivent des lois de Weibull (hypothèse que la pollution résulte d'une panne ou mauvais fonctionnement d'un ou plusieurs appareillages du réseau de raffinerie), des lois normales (superposition de nombreux facteurs isolément de faible importance), des lois Lognormales (par rapport à ce qui précède, présence d'effets proportionnels conjugués), des lois Gamma ou plus généralement de Pearson (diverses interprétations éventuellement vraisemblables). Les estimations étaient rendues de plus délicates, dans le cas de distributions hypothétiques à support borné à gauche, ou des deux côtés à la fois, par le fait que l'origine (taux de pollution minimum) n'était pas toujours connue ; en particulier, dans le cas de distributions de Weibull, Lognormales, Gamma, etc., on est ainsi dans un cas d'estimation non régulière, l'estimation par le maximum de vraisemblance étant dégénérée.

Les hypothèses précédentes supposent que la densité de la loi observée est unimodale. Or, il est apparu clairement que cette hypothèse ne pouvait être retenue pour certains échantillons, dont la densité présentait de toute évidence deux modes ou d'avantage, sans doute par le mélange de pollutions spécifiques provenant d'un faible nombre de défauts du réseau.

L'analyse poussée d'un grand nombre de données a montré que le phénomène pouvait être valablement interprété, suivant les cas, par une loi normale, par une loi Lognormale (englobant le cas précédent), ou par un mélange de plusieurs distributions du type précédent.

Malheureusement, il n'est pas possible de faire une hypothèse générale permettant de fournir une estimation unique, pour l'ensemble des échantillons, par une procédure unifiée. Ceci provient essentiellement de la faible taille des échantillons qui peuvent être, a priori, réputés homogènes. De plus, l'estimation des mélanges de lois normales est très délicate pour de petits échantillons.

D'autre part, il est très important de pouvoir caractériser de manière aussi précise que possible les densités de pollution, et d'autre part, il est quasiment impossible d'obtenir des tests valables entre les alternatives vraisemblables. Enfin, la puissance des tests de normalité, ou de Lognormalité est insuffisante, faute d'un nombre suffisant de données, pour faire un choix fondé sans recherches coûteuses.

Ce qui précède amène naturellement à la recherche de représentations fournies par des estimations non paramétriques de la distribution.

Pour les raisons exposées antérieurement, l'histogramme est la forme la plus commode de représentation. Le problème se pose alors de choisir ses paramètres qui sont : le noyau, et le coefficient δ_n (la fenêtre).

L'exposé qui précède permet de faire un tel choix d'une manière optimale, pour un faible échantillon (rappelons que la plupart des résultats disponibles dans la littérature sur le sujet donnent des choix optimaux asymptotiquement), et pour une hypothèse paramétrique donnée.

On procède alors comme suit :

1) On estime les paramètres de la distribution d'hypothèse.

2) On construit l'histogramme optimal pour la distribution d'hypothèse, en utilisant une étude de robustesse comme dans le § 21.

3) On teste l'hypothèse (classiquement) ; on utilise éventuellement l'histogramme pour avancer une autre hypothèse, et reprendre le procédé, ou, en cas d'impossibilité de représenter valablement la distribution paramétriquement, pour fournir lui même l'estimation de la distribution.

Ce procédé présente ainsi l'avantage de pouvoir être utilisé de manière automatique par de non-statisticiens, qui peuvent ainsi visualiser la distribution par un histogramme, et choisir l'hypothèse vraisemblable.

Un programme généralisé permettant d'utiliser cette procédure d'estimation a été composé par MM. Brice et Bourdaire, de la Compagnie Française des Pétroles sous la direction de M. CHARRETON. Il est actuellement utilisé sur des échantillons provenant d'origines diverses, au delà du problème de pollution initial.

Il est intéressant de comparer les valeurs optimales de δ_n avec les valeurs exactes, données respectivement en (2,55) et par minimisation de (2,33) ou (2,34).

Dans la table qui suit, nous comparons le δ_n optimal exact au δ_n optimal asymptotique : (loi N(0,1))

Noyau unité (n° 1)

n	δ_n optimal	δ_n asymptotique
10	2,475	2,326
20	2,126	2,025
50	1,747	1,686
100	1,510	1,468

Noyau normal (n° 5)

n	δ_n optimal	δ_n asymptotique
10	0,758	0,668
20	0,642	0,582
50	0,520	0,484
100	0,445	0,422

L'approximation est bonne pour $n \geqslant 10$, bien qu'elle donne des valeurs sous-estimées de l'ordre de 5 % pour les petits échantillons.

REFERENCES

[1] BICKEL-ROSENBLATT (1973). – On some global measures of the deviates of density function estimates. *Ann. Statist.,* 1, p. 1071-95.

(1975). – Corrections to. Ann. Statist. 3, p. 1370.

[2] BULLOCK-DAVIS (1975). – Mean square error properties of density estimates. *Ann. Statist.,* 3, p. 1025-30.

[3] CENCOV (1962) – Estimation of an unknown distribution density from observations. *Sovet. Math.,* 3, p. 1559-62.

[4] CHARLIER (1905). – Uber die darstellung willkürlicher funktione. *Arkiv für Matematik Astronomi och Fysik*, 2, p. 1-35.

(1914) Contribution to the mathematical theory of statistics. *Arkiv für Matematik Astronomi och Fysik*, 9, p. 1-18.

[5] DEHEUVELS P. (1974). – Estimation de la densité (Thèse, Université Paris VI.

(1974). – Conditions nécessaires et suffisantes de convergence ponctuelle presque sûre et uniforme presque sûre des estimateurs de la densité, *C.R. Acad. Sci. Paris*, 278, p. 1217-20.

(1977). – Estimation non paramétrique de la densité par histogrammes généralisés II, Publ. ISUP, 1977. Vol. XXII, 1, 2, p. 1-24.

[6] DIMITRIEV-TARATSENKO (1973). – On the estimation of the probability density and its derivatives. *Theor. Prob. Appl.*, 18, p. 628-33.

[7] EDGEWORTH (1904). – The law of error. *Trans. Camb. Phil. Soc.*, 20, p. 36-, p. 113-

(1907). – On the representation of a statistical frequency by a series. *Jour. Roy. Stat. Soc.,* Ser. A, 70, p. 102-6.

(1916). – On the mathematical representation of statistical data. *Jour. Roy. Stat. Soc.,* Ser. A, 79, p. 455-500, (1917), 80, q. 65-83, p. 266-288, p. 411-437.

[8] EPANECHNIKOV (1969). – Nonparametric estimates of a multivariate probability density, *Teor. Prob. Appl.,* 14, p. 153-8.

[9] GOOD-GASKINS (1971). – Nonparametric roughness penalties of probability densities. *Biometrika*, 58, p. 255-277.

[10] GRENANDER (1956). – On the theory of mortality measurement, II, *Skand. Akt.,* 39, p. 125-53.

[11] KENDALL. – The advanced theory of statistics, Vol. 1-2-3, Griffin.

[12] KRONMAL-TARTER-HOLCOMB (1967). – A description of new computer methods for estimating the population density. *Proc. A.C.M., Thompson book Company*, 22, p. 511-19.

[13] KRONMAL-TARTER (1968). – The estimation of probability densities and cumulatives by Fourier series methods. *J. Amer. Stat. Assoc.*, 63, p. 925-52.

(1970). – On multivariate density estimates based on orthogonal expansions. *Ann. Math. Statist.*, 41, p. 718-22.

[14] LEADBATTER-WATSON (1963). – On the estimation of probability density II. *Ann. Math. Statist.*, 34, p. 480-91.

[15] MALZ (1974). – Estimation of the k^{th} derivative of a distribution function. *Ann. Statist.*, 2, p. 359-61.

[16] MONTRICHER-TAPIA-THOMPSON (1975). – Nonparametric maximum likehood estimation of probability density by penalty function methods. *Ann. Statist.*, 3, p. 1329-48.

[17] PARZEN (1962). – On the estimation of a probability density function and the mode. *Ann. Math. Statist.*, 33, p. 1065-76.

[18] PRAKASA-RAO (1969). – Estimation of a unimodal density. *Sankhya*, Ser. A, 31, p. 26-36.

[19] REISS (1973).

[20] ROBERTSON (1967). – On estimating a density which is measurable with respect to a lattice. *Ann. Math. Statist.*, 38, p. 482-93.

[21] ROSENBLATT (1956). – Remarks on some nonparametric estimates of a density. *Ann. Math. Statist.*, 27, p. 832-7.

[22] (1971). – Curve estimates. *Ann. Math. Statist.*, 33, p. 1065-76.

[23] SAGER (1975). – Consistency in nonparametric estimation of the mode. *Ann. Statist.*, 3, p. 698-706.

[24] SANSONE Orthogonal functions, Interscience.

[25] SCHUSTER (1969). – Estimation of a probability density function and its derivatives. *Ann. Math. Statist.*, 40, p. 1187-95.

[26] SHAPIRO (1969). – Smoothing and approximation of functions, Van Nostrand Reinhold.

[27] SCHWARZ-STUART (1967). – Estimation of a probability density by orthogonal series. *Ann. Math. Statist.*, 38, p. 1261-5.

[28] VAN RYZIN (1966). – Bayes risk consistency of classification procedures using density estimation. *Sankhya*, Ser. A, 28, p. 261-70.

[29] VENTER (1967). – On the estimation of the mode. *Ann. Math. Statist.*, 37, p. 1446-55.

[30] WAHBA (1971). – A polynomial algorithm for density estimation. *Ann. Math. Statist.*, 42, p. 1870-86.

[31] WATSON (1969). – Density estimation by orthogonal series. *Ann. Math. Statist.*, 40, p. 1496-8.

[32] WEGMAN (1969). – A note on estimating a unimodal density. *Ann. Math. Statist.*, 40, p. 1661-17.

(1969). – Maximum likehood histograms, Institute of statistics 629, Univ. North Carolina.

(1970). – Maximum likehood estimation of a unimodal density I, II. *Ann. Math. Statist.*, 41, p. 457-71, p. 2169-74.

[33] WEGMAN (1972). – Nonparametric probability density estimation. A summary of available methods. *Technometrics*, 14, p. 533-546.

[34] WOODROOFE (1967). – On the maximum deviation of the sample density. *Ann. Math. Statist.*, 38, p. 475-81.

[35] WOODROOFE (1970). – On choosing a delta sequence. *Ann. Math. Statist.*, 41, p. 1665-71.