n° 9513

A Comparison of Kernel Estimator Based Goodness of Fit Tests

C. GOURIEROUX [1], and C., TENREIRO [2]

December 1994

[1]CREST and CEPREMAP
[2]Department of Mathematics, COIMBRA University

A Comparison of Kernel Estimator Based Goodness of Fit Test

If (X_i) is a stationary process with marginal density function f, we are interested in testing the hypothesis $H_o : \{f = f_o\}$, where f_o is given. We consider different test statistics based on integrated quadratic forms measuring the proximity between f_n, a kernel estimator of f, and f_o or between f_n and its expected value computed under H_o We compare their asymptotic local power functions, and discuss the possibility to obtain unbiased tests. Finally we explain how to obtain kernel based goodness of fit tests, which are asymptotically locally uniformly most powerful unbiased (LUMPU) for testing H_o against a parametric hypothesis $H = \{f \in (g.(.;\theta), \theta \in \Theta)\}$ including H_o.

Une comparaison des tests d'ajustement fondés sur des estimateurs à noyau

Soit (X_i) un processus stationnaire avec une densité marginale f. Nous nous intéressons au test de l'hypothèse $H_o = \{f = f_o\}$, où f_o est une densité donnée. Nous considérons diverses statistiques de tests fondées sur des formes quadratiques intégrées mesurant la proximité entre un estimateur à noyau f_n de f et la loi f_o, ou entre f_n et sa moyenne sous H_o. Nous comparons les puissances asymptotiques locales de ces tests, et discutons la possibilité d'obtenir des test sans biais. Finalement nous expliquons comment obtenir un test d'ajustement fondé sur des estimateurs à noyau, qui soit asymptotiquement localement uniformément le plus puissant sans biais pour tester H_o contre une hypothèse paramétrique $H = \{f \in (g(.;\theta), \theta \in \Theta)\}$ la contenant.

Keywords : Tests for goodness of fit, density estimates, asymptotic power, local alternatives.

Mots clefs : Tests d'ajustement, estimateurs de densité, puissance asymptotique, alternatives locales.

AMS subject classification : 62 G10, 62 G20

1 INTRODUCTION

Let us consider a d-dimensional stationary process $(X_i, i \in \mathbf{Z})$, whose marginal distribution has a density function f with respect to Lebesgue measure on $\mathbb{R}^d$. We are interested in testing hypotheses concerning the form of this pdf, essentially the simple hypothesis that it is equal to a given pdf $f_o : H_o = \{f = f_o\}$. Different nonparametric approaches have been proposed in the literature ; some of them are based on the supremum of the difference between the empirical cumulative distribution function F_n and the one F_o corresponding to the null hypothesis (in the case of the simple hypothesis H_o). It is the case of the testing procedures proposed by Kolmogorov-Smirnov and Cramer-von Mises. However the asymptotic distribution of the statistic $\sup_x |F_n(x) - F_o(x)|$ is only known for the case $d = 1$, and this kind of approach is difficult to extend to the case of a composite parametric null hypothesis H'_o.

Some other approaches are based on nonparametric estimators f_n of the density function and on integrated quadratic averages measuring the proximity between $E_o f_n$ and f_n, or f_o and f_n. These approaches have for instance been followed by Bickel-Rosenblatt (1973), who considered the asymptotic properties of the statistics :

$$\int_{\mathbf{R}^d} [f_n(x) - E_o f_n(x)]^2 \pi(x) dx, \text{ and } \int_{\mathbf{R}^d} [f_n(x) - f_o(x)]^2 \pi(x) dx,$$

where $f_n(x) = \dfrac{1}{nh_n^d} \sum_{i=1}^{n} K\left(\dfrac{x - X_i}{h_n}\right)$ is a kernel estimator of f, h_n is a bandwith, and π a weight funtion (see also Rosenblatt (1975) for $d = 2$, Bickel-Rosenblatt (1973) for another measure of proximity, and Eubank and Lariccia (1992) for another kind of density estimator).

The idea of this paper is to extend these latter approaches to the multidimensional dependent case and to study local power properties. We consider in section 2 the asymptotic behaviours of the two statistics :

$$I_n^1(\pi) = \int_{\mathbf{R}^d} [f_n(x) - E_o f_n(x)]^2 \pi(x) dx, \tag{1.1}$$

and :

$$I_n^2(\pi) = \int_{\boldsymbol{R}^d} [f_n(x) - f_o(x)]^2 \pi(x) dx, \tag{1.2}$$

both under the null hypothesis $H_o = \{f = f_o\}$ and under a sequence of local alternatives $H_{1n} = \{f(x) = f_o(x) + a_n\gamma(x) + o(a_n)\}$, where (a_n) is a sequence of real numbers tending to zero. This study generalizes the results by Bickel and Rosenblatt (1973) and Rosenblatt (1975) derived for d=1 and d=2 and under the null hypothesis by Hall (1984) and Takahata and Yoshihara (1987).

It is noted that under H_o the two statistics admit asymptotic expansions of the forms : $c_j^o(n,\pi) + d_j^o(n,\pi)^{-1} Z_j, j = 1, 2$ where the first term is deterministic, and the second one is asymptotically normal with zero mean. Therefore in order to develop goodness of fit tests of the hypothesis H_o it is first necessary to correct for this asymptotic bias, i.e., to consider test statistics of the form :

$$T_n^j(\pi) = \widehat{d}_j(n,\pi) \left[I_n^j(\pi) - \widehat{c}_j(n,\pi) \right],$$

where $\widehat{c}_j(n,\pi)$ and $\widehat{d}_j(n,\pi)$ are consistent estimators of the deterministic coefficients $c_j^o(n,\pi), d_j^o(n,\pi)$. To be able to define such estimators, some asymptotic approximations of the coefficients $c_j^o(n,\pi)$ are derived in section 2.

In section 3 we introduce different test statistics of the previous forms. For instance for the statistics based on $I_n^1(\pi)$, we introduce :

$$\begin{aligned} T_n^{1,0}(\pi) &= n h_n^{d/2} \left\{ \int_{\boldsymbol{R}^d} [f_n(x) - E_o f_n(x)]^2 \pi(x) dx \right. \\ &- \left. \frac{1}{n h_n^d} \int_{\boldsymbol{R}^d} \int_{\boldsymbol{R}^d} K^2(u) f_o(x) \pi(x + u h_n) dx du \right\}, \end{aligned} \tag{1.3}$$

and :

$$\begin{aligned} T_n^{1,1}(\pi) &= n h_n^{d/2} \left\{ \int_{\boldsymbol{R}^d} [f_n(x) - E_o f_n(x)]^2 \pi(x) dx \right. \\ &- \left. \frac{1}{n h_n^d} \frac{1}{n} \sum_{i=1}^n \int_{\boldsymbol{R}^d} K^2(u) \pi(X_i + u h_n) du \right\}. \end{aligned} \tag{1.4}$$

Similar test statistics are derived from $I_n^2(\pi)$. We show that the different statistics $T_n^{1,0}$ and $T_n^{1,1}$[resp. $T_n^{2,0}$ and $T_n^{2,1}$] are asymptotically equivalent under the null hypothesis and asymptotically normal, and we derive associated consistent critical regions. We also give the asymptotic expansions of these test statistics under sequences of local alternatives.

The aim of section 4 is the comparison of the local power properties of the different tests under the previous sequence of local alternatives : existence and computation of the implicit rates of local alternatives, asymptotic local unbiasedness, impossibility of reaching the same rate as for parametric testing procedure.

Finally in section 5 we discuss the possibility of using kernel based goodness of fit tests for testing the null hypothesis $H_o : \{f = g(.;\theta_o)\}$ against a parametric alternative $H = \{f \in (g(.;\theta), \theta \in \Theta)\}$, and to derive a nonparametric test which is asymptotically locally uniformly most powerful unbiased (LUMPU).

2. ASYMPTOTIC PROPERTIES OF THE STATISTICS $I_n^1(\pi)$ AND $I_n^2(\pi)$

We introduce different assumptions denoted by (K), (B), (π) or (P) on the kernel K, the bandwidth (h_n), the weight function π, or the process $(X_i, i \in \mathbf{Z})$.

2.a **First assumptions**

Assumptions on the kernel (K)

K is a bounded measurable function on $\mathbb{R}^d$ such that $\int_{\mathbb{R}^d} K(u)du = 1$.

Assumptions on the bandwidth (B)

$$h_n \to 0 \text{ and } nh_n^d \to +\infty, \text{ when } n \to +\infty, \tag{2.1}$$

and there exists $\gamma \in]0,1[$ such that :

$$\limsup_n n^\gamma h_n^d < +\infty. \tag{2.2}$$

The conditions (2.1) are usual in the theory of kernel estimation, and are only used in a dependent framework. The other one (2.2) is not very restrictive, and is for instance satisfied if $h_n = 0(n^{-\delta})$, $0 < \delta < 1/d$.

Assumptions on the weight function (π)

The function π is nonnegative and almost everywhere continuous. It satisfies : $\sup_{x \in \mathbb{R}^d} \pi(x) < +\infty$.

Before presenting the assumptions on the process, we have to define precisely what are local alternatives.

2.b **Local alternatives**

With local alternatives we do not consider a unique process, but a sequence of processes $(X_{in}, i \in \mathbf{Z})$, indexed by n [in the sections below we will suppress the index n for notational convenience].

Assumptions on the processes (P)

For any n we assume that $(X_{in}, i \in \mathbf{Z})$ is a d-dimensional strongly stationary process. This process is β-mixing and we impose some conditions on the uniform convergence to zero of the β-coefficients :

$$\exists c > 0, \rho \in]0,1[\text{ such that : } \sup_{n \in \mathbf{N}} \beta_n(i) \leq c\rho^i \ \forall i \in \mathbf{N}, \tag{2.3}$$

where : $\beta_n(i) = \underset{n}{E} \left\{ \sup_{A \in F_{i,n}^{+\infty}} |P(A/F_{-\infty,n}^o) - P(A)| \right\}$, and

$F_{i,n}^{+\infty}$[resp $F_{-\infty,n}^o$] is the σ-algebra generated by $X_{j,n}, j \geq i$ [resp : $X_{j,n}, j \leq o$].

Moreover we assume that (X_{in}) and (X_{in}, X_{on}) $i \geq 1$ have continuous distributions with pdf g_n and $g_{i,n}$ respectively, such that :

$$\sup_{i \in \mathbf{N}^*} \sup_{\substack{x \in \{u : g_n(u) > o\} \\ y \in I\!R^d}} \frac{g_{i,n}(y,x)}{g_n(x)} < +\infty, \tag{2.4}$$

and :

$$g_n(x) = f_o(x) + a_n\gamma(x) + o(a_n)\gamma_n(x), \tag{2.5}$$

where f_o is a bounded pdf on $I\!R^d, (a_n)$ is a sequence of real numbers tending to zero, and

$$\sup_{x \in I\!R^d} |\gamma(x)| < +\infty, \quad \sup_{n \in \mathbf{N}^*} \sup_{x \in I\!R^d} |\gamma_n(x)| < +\infty,$$

$$\int_{\mathbf{R}^d} |\gamma(x)|dx < +\infty, \quad \sup_{n \in \mathbf{N}^*} \int_{\mathbf{R}^d} |\gamma_n(x)|dx < \infty.$$

The mixing condition is used to derive a central limit theorem for U-statistics [see Hall (1984) and Takahata-Yoshihara (1987)]. The condition corresponding to the conditional density function is automatically satisfied in the iid case, or for stationary gaussian processes, since the conditional variance is independent of the value of the conditioning variable. Condition (2.5) gives the interpretation in terms of local alternatives since $\lim_{n \to \infty} g_n = f_o$. The associated rate of convergence is a_n.

At this stage it is interesting to discuss the retained assumption (2.5) and in particular to give some interpretation in terms of random variables. The following property is easily derived.

Property 1 : Let us consider a sequence of processes defined by :

$$X_{in} = X_{i\infty} + \delta_n Z_{i\infty}, i \in \mathbf{Z},$$

where $(X_{i\infty}, i \in \mathbf{Z}), (Z_{i\infty}, i \in \mathbf{Z})$ are two independent strongly stationary processes, with β-mixing coefficients exponentially decreasing to zero and (δ_n) is a sequence of real numbers tending to zero. We assume that $(X_{i\infty})$ has uni- and bidimensional marginal distributions with pdf satisfying :

$$\sup_{x \in I\!R^d} f_o(x) < \infty, \quad \sup_{\substack{x \in \{u; f_o(u) > 0\} \\ y \in I\!R^d}} \frac{g_i(y,x)}{f_o(x)} < +\infty.$$

Then the set of assumptions (P) is satisfied.

The forms of the marginal density function g_n and of its expansion depend on the marginal distribution of $Z_{o\infty}$.

If $Z_{o\infty}$ is second order integrable, $E(Z_{o\infty}) \neq 0$ and if f_o has partial continuous derivatives up to order two, which are bounded and integrable on $I\!R^d$, we have :

$$g_n(x) = E[f_o(x - \delta_n Z_{o,\infty})]$$

$$= f_o(x) + \delta_n \gamma(x) + \delta_n^2 \gamma_n(x),$$

where : $\gamma(x) = -\sum_{i=1}^{d} E(Z_{o\infty,i}) \frac{\partial f_o}{\partial x_i}(x),$

$$\gamma_n(x) = \sum_{i,j=1}^{d} E[Z_{o\infty,i} Z_{o\infty,j} \int_o^1 \frac{\partial^2 f_o}{\partial x_i \partial x_j}(x - \delta_n Z_{o\infty} t)(1-t)dt].$$

If $Z_{o\infty}$ is third order integrable, $E(Z_{o\infty}) = 0$, and if f_o has partial continuous derivatives up to order three, which are bounded and integrable, we get :

$$g_n(x) = f_o(x) + \delta_n^2 \gamma(x) + \delta_n^3 \gamma_n(x),$$

where :

$$\gamma(x) = \frac{1}{2} \sum_{i,j=1}^{d} E(Z_{o\infty,i} Z_{o\infty,j}) \frac{\partial^2 f_o}{\partial x_i \partial x_j}(x),$$

$$\gamma_n(x) = -\frac{1}{2} \sum_{i,j,k=1}^{d} E[Z_{o\infty,i} Z_{o\infty,j} Z_{o\infty,k} \int_o^1 \frac{\partial^3 f_o}{\partial x_i \partial x_j \partial x_k}(x - \delta_n Z_{o\infty} t)(1-t)^2 dt].$$

In summary if $Z_{o\infty}$ is zero mean, we have $a_n = \delta_n^2$, and otherwise $a_n = \delta_n$. It is possible to link the order of local alternatives in terms of density function (i.e. a_n) and in terms of variables (i.e. δ_n).

2.c **Asymptotic expansions of the statistics $I_n^1(\pi)$ and $I_n^2(\pi)$**

For $m \in \mathbf{N}$ let us introduce the set $W^d(m)$ of pdf with partial continuous derivatives up to order m, which are bounded and integrable, and the set $K^d(m)$ of kernels of order m, i.e., such that : $\int_{\mathbf{R}^d} x_1^{a_1} \dots x_d^{a_d} K(x)dx = 0$ if $a_1, \dots, a_d \in \mathbf{N}, o < a_1 + \dots + a_d < m$ and $\int_{\mathbf{R}^d} \| x \|^m |K(x)|dx < \infty$. For $f_o \in W^d(m)$ and $K \in \mathcal{K}^d(m)$, let us denote :

$$\Delta^m f_o(x) = \frac{(-1)^m}{m!} \sum_{i_1,\dots,i_m}^{d} \int_{\mathbf{R}^d} u_{i_1} \dots u_{i_m} K(u)du \frac{\partial^m f_o(x)}{\partial x_{i_1} \dots \partial x_{i_m}}, \qquad (2.5)$$

$$\Delta_n^m f_o(x) = \frac{(-1)^m}{(m-1)!} \sum_{i_1,\ldots,i_m=1}^{d} \int_{\mathbf{R}^d} u_{i_1} \ldots u_{i_m} K(u) \int_o^1 \frac{\partial^m f_o}{\partial x_{i_1} \ldots \partial x_{i_m}} (x - h_n ut)(1-t)^{m-1} dt du. \tag{2.6}$$

The asymptotic randomness of the statistics $I_n^1(\pi), I_n^2(\pi)$ essentially comes from the two following U-statistics :

$$\mathcal{H}_n(\pi) = \frac{2}{n} \sum_{1 \leq j < i \leq n} [H_n(X_i, X_j) - EH_n(X_i, X_j)], \tag{2.7}$$

where :

$$H_n(u,v) = \frac{1}{h_n^{3d/2}} \int_{\mathbf{R}^d} \left\{ K(\frac{x-u}{h_n}) - E_n K(\frac{x - X_o}{h_n}) \right\} \left\{ K(\frac{x-v}{h_n}) - E_n K(\frac{x - X_o}{h_n}) \right\} \pi(x) dx, \tag{2.8}$$

and :

$$\mathcal{G}_n(\pi) = \frac{2}{\sqrt{n}} \sum_{i=1}^{n} G_n(X_i), \tag{2.9}$$

where :

$$G_n(u) = \frac{1}{h_n^d} \int_{\mathbf{R}^d} \left[K(\frac{x-u}{h_n}) - E_n K(\frac{x - X_o}{h_n}) \right] (\Delta_n^m f_o . \pi)(x) dx \tag{2.10}$$

The following result is a central limit theorem for U-statistics.

Property 2 : i) Under $(K), (B), (\pi), (P)$ the U-statistic $\mathcal{H}_n(\pi)$ is under the sequence of local alternatives asymptotically normal with zero mean and a variance given by $2\nu_1^2(\pi)$, where :

$$\nu_1^2(\pi) = \int_{\mathbb{R}^d} f_o^2(x)\pi^2(x)dx \int_{\mathbb{R}^d} (K * \bar{K})^2(z)dz,$$

$\bar{K}(u) = K(-u)$, and $*$ is the convolution product.

ii) Let us assume that for any $i \in \mathbf{N}$, there exists $c_i \in \mathbb{R}$ such that for any sequences u_n, v_n tending to zero we have :

$$\underset{n}{E}\ [(\Delta_n^m f_o.\pi)(X_i + u_n)(\Delta_n^m f_o.\pi)(X_o + v_n)] \rightarrow c_i, \text{ if } n \rightarrow \infty.$$

Under $(K), (B), (\pi), (P)$ and if $f_o \in W^d(m), K \in K^d(m)$ for some $m \in \mathbf{N}$ the bivariate statistic $(\mathcal{H}_n(\pi), \mathcal{G}_n(\pi))$ is under the sequence of local alternatives asymptotically normal, with zero mean and a diagonal variance covariance matrix :

$$\begin{bmatrix} 2\nu_1^2(\pi) & 0 \\ 0 & 4\nu_2^{*2}(\pi) \end{bmatrix},$$

where :

$$\nu_2^{*2}(\pi) = \mathrm{Var}_o[(\Delta^m f_o.\pi)(X_o)] + 2\sum_{j=1}^{+\infty}\{c_j - E_o^2[(\Delta^m f_o.\pi)(X_o)]\},$$

E_o and Var_o are with respect to f_o.

Proof : see appendix 1

Under H_o, we have for $i \in \mathbf{N}$:

$$c_i = \underset{o}{E} \; [(\Delta^m f_o.\pi)(X_i)(\Delta^m f_o.\pi)(X_o)]\,,$$

and :

$$\begin{aligned}\nu_2^{*2}(\pi) &= \nu_2^2(\pi) \text{ (say)} \\ &= \mathrm{Var}_o[(\Delta^m f_o.\pi)(X_o)] + 2\sum_{j=1}^{\infty} \text{Cov } [(\Delta^m f_o.\pi)(X_i), (\Delta^m f_o.\pi)(X_o)].\end{aligned}$$

If for each $i \in \mathbf{N}$, and under the sequence of local alternatives the sequence of pdf $f_{(X_i,X_o)}$, converging to some pdf $f_{(X_{i\infty},X_{o\infty})}$, satisfies the dominated convergence theorem conditions, then the correlation term may be written as :

$$c_i = E\,[(\Delta^m f_o.\pi)(X_{i\infty})(\Delta^m f_o.\pi)(X_{o\infty})]\,.$$

The same form of c_i may be obtained if π is continuous, if for each $x \in I\!R^d$ the partial derivatives of order m of f_o are α-lipschitzian, and if under the sequence of local alternatives the bivariate distribution of $(X_i, X_o), i \in \mathbf{N}$, tends to some limit distribution $(X_{i\infty}, X_{o,\infty})$ (see Ranga Rao (1962)).

The expansions of the statistics $I_n^1(\pi)$ and $I_n^2(\pi)$ are given in the following property.

Property 3 : i) Under $(K),(B),(\pi),(P),$ and if f_o is bounded, we have asymptotically under the sequence of local alternatives :

$$I_n^1(\pi)=\frac{1}{nh_n^{d/2}}\mathcal{H}_n(\pi)+a_n^2\int_{\boldsymbol{R}^d}\gamma^2(x)\pi(x)dx$$

$$+\frac{1}{nh_n^d}\int_{\boldsymbol{R}^a}\int_{\boldsymbol{R}^d}K^2(u)g_n(x)\pi(x+uh_n)dxdu+o_p(a_n^2)+o_p\left(\frac{1}{nh_n^{d/2}}\right).$$

ii) Under $(K),(B),(\pi),(P),$ and if $f_o\in W^d(m),K\in K^d(m),$ for some m $\in\mathbf{N}$, we have asymptotically under the sequence of local alternatives :

$$I_n^2(\pi)=\frac{1}{nh_n^{d/2}}\mathcal{H}_n(\pi)+\frac{1}{\sqrt{n}h_n^{-m}}\mathcal{G}_n(\pi)$$

$$+a_n^2\int_{\boldsymbol{R}^d}\gamma^2(x)\pi(x)dx+\frac{1}{nh_n^d}\int_{\boldsymbol{R}^d}\int_{\boldsymbol{R}^d}K^2(u)g_n(x)\pi(x+uh_n)dxdu$$

$$+2h_n^m\int_{\boldsymbol{R}^d}\int_{\boldsymbol{R}^d}K(u)[g_n(x)-f_o(x)](\Delta_n^m f_o.\pi)(x+uh_n)dxdu$$

$$+\int_{\boldsymbol{R}^d}[E_of_n(x)-f_o(x)]^2\pi(x)dx+o_p(a_n^2)+o_p\left(\frac{1}{nh_n^{d/2}}\right).$$

Proof : see appendix 2

In particular the properties of the two previous statistics under the null hypothesis are obtained by imposing $a_n=0$. We get, under H_o :

$$I_n^1(\pi)=\frac{1}{nh_n^{d/2}}\mathcal{H}_n(\pi)+\frac{1}{nh_n^d}\int_{\boldsymbol{R}^d}\int_{\boldsymbol{R}^d}K^2(u)f_o(x)\pi(x+uh_n)dxdu+o_p(\frac{1}{nh_n^{d/2}}), \tag{2.11}$$

$$I_n^2(\pi) = \frac{1}{nh_n^{d/2}}\mathcal{H}_n(\pi) + \frac{1}{\sqrt{n}h_n^{-m}}\mathcal{G}_n(\pi) \quad + \frac{1}{nh_n^d}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} K^2(u)f_o(x)\pi(x+uh_n)dxdu$$
$$+ \int_{\mathbb{R}^d}[E_o f_n(x) - f_o(x)]^2\pi(x)dx + o_p\left(\frac{1}{nh_n^{d/2}}\right). \tag{2.12}$$

3. GOODNESS OF FIT TEST

3.a Definition of some test statistics

From (2.11), (2.12) we see that $I_n^j(\pi), j = 1, 2,$ are asymptotically under the null hypothesis of the form :

$$I_n^j(\pi) \simeq c_j^o(n,\pi) + [d_j^o(n,\pi)]^{-1}Z_j,$$

where $d_j^o(n,\pi), c_j^o(n,\pi)$ are deterministic and Z_j are normal variables. Moreover the term $c_j^o(n,\pi)$ is larger than $d_j^o(n,\pi)$. Therefore if we want to build relevant testing procedures we have first to correct for the asymptotic bias $c_j^o(n,\pi)$, then to introduce the rate $d_j^o(n,\pi)$. The corresponding test statistics will be asymptotically normal under the null hypothesis H_o.

The bias corrections consist essentially for $I_n^1(\pi)$ for instance in substracting :

$$\frac{1}{nh_n^d}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} K^2(u)f_o(x)\pi(x+uh_n)dxdu,$$

or a consistent estimate of this quantity under the null hypothesis, such as :

$$\frac{1}{nh_n^d}\frac{1}{n}\sum_{i=1}^n\int_{\mathbb{R}^d} K^2(u)\pi(X_i+uh_n)du. \tag{3.1}$$

More precisely we consider the following statistics :

$$\begin{aligned} T_n^{1,0}(\pi) \quad &= nh_n^{d/2}\left\{\int_{\mathbb{R}^d}[f_n(x) - E_o f_n(x)]^2\pi(x)dx \right. \\ &\left. -\frac{1}{nh_n^d}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} K^2(u)f_o(x)\pi(x+uh_n)dxdu\right\}, \end{aligned} \tag{3.2}$$

$$
\begin{aligned}
T_n^{1,1}(\pi) \; &= nh_n^{d/2}\left\{\int_{\mathbf{R}^d}[f_n(x) - E_o f_n(x)]^2\pi(x)dx \right. \\
&\left. - \frac{1}{nh_n^d}\frac{1}{n}\sum_{i=1}^{n}\int_{\mathbf{R}^d}K^2(u)\pi(X_i + uh_n)du\right\},
\end{aligned} \tag{3.3}
$$

$$
\begin{aligned}
T_n^{2,0}(\pi) \; &= d(n) \quad \left\{\int_{\mathbf{R}^d}[f_n(x) - f_o(x)]^2\pi(x)dx \right. \\
&-\frac{1}{nh_n^d}\int_{\mathbf{R}^d}\int_{\mathbf{R}^d}K^2(u)f_o(x)\pi(x + uh_n)dxdu \\
&\left. -\int_{\mathbf{R}^d}[E_o f_n(x) - f_o(x)]^2\pi(x)dx\right\},
\end{aligned} \tag{3.4}
$$

$$
\begin{aligned}
T_n^{2,1}(\pi) \; &= d(n) \quad \left\{\int_{\mathbf{R}^d}[f_n(x) - f_n(x)]^2\pi(x)dx \right. \\
&-\frac{1}{nh_n^d}\frac{1}{n}\sum_{i=1}^{n}\int_{\mathbf{R}^d}K^2(u)\pi(X_i + uh_n)du \\
&\left. -\int_{\mathbf{R}^d}[E_o f_n(x) - f_o(x)]^2\pi(x)dx\right\},
\end{aligned} \tag{3.5}
$$

where :

$$
d(n) = \begin{cases} nh_n^{d/2}, & \text{if} \quad \lim\limits_{n \to +\infty} \quad nh_n^{d+2m} = \lambda \in [o, +\infty[, \\ \\ \sqrt{n}h_n^{-m}, & \text{if } \lambda = +\infty. \end{cases}
$$

3.b **Expansions of the test statistics under local alternatives**

Under local alternatives the term $\frac{1}{nh_n^d}\int_{\mathbf{R}^d}\int_{\mathbf{R}^d}K^2(u)g_n(x)\pi(x+uh_n)dxdu$ appearing in the expansions of $I_n^1(\pi)$ and $I_n^2(\pi)$ may be approximated in different ways. It is first equal to :

$$
\begin{aligned}
&\frac{1}{nh_n^d}\int_{\mathbf{R}^d}\int_{\mathbf{R}^d}K^2(u)f_o(x)\pi(x + uh_n)dxdu \\
&+ \frac{a_n}{nh_n^d}\int_{\mathbf{R}^d}\int_{\mathbf{R}^d}K^2(u)\gamma(x)\pi(x + uh_n)dxdu + o_p\left(\frac{a_n}{nh_n^d}\right)
\end{aligned}
$$

$$= \frac{1}{nh_n^d}\int_{\mathbf{R}^d}\int_{\mathbf{R}^d} K^2(u) f_o(x)\pi(x+uh_n)dxdu$$

$$+ \frac{a_n}{nh_n^d}\int_{\mathbf{R}^d}\gamma(x)\pi(x)dx\int_{\mathbf{R}^d}K^2(x)dx + o_p\left(\frac{a_n}{nh_n^d}\right), \text{using the definition of } g_n.$$

It is also equal to :

$$\frac{1}{nh_n^d}\frac{1}{n}\sum_{i=1}^{n}\int_{\mathbf{R}^d}K^2(u)\pi(X_i+uh_n)du + O_p\left(\frac{1}{nh_n^d}\frac{1}{\sqrt{n}}\right),$$

by central limit theorem. Using these approximations and theorem 3, we deduce the following expansions,

Property 4 Under the sequence of local alternatives, we get :

$$T_n^{1,1}(\pi) = \mathcal{H}_n(\pi) + nh_n^{d/2}a_n^2\int_{\mathbf{R}^d}\gamma^2(x)\pi(x)dx + o_p(nh_n^{d/2}a_n^2), + o_p(1),$$

$$T_n^{1,0}(\pi) = T_n^{1,1}(\pi) + \frac{a_n}{h_n^{d/2}}\int_{\mathbf{R}^d}\gamma(x)\pi(x)dx\int_{\mathbf{R}^d}K^2(x)dx + o_p\left(\frac{a_n}{h_n^{d/2}}\right),$$

$$T_n^{2,1}(\pi) = \frac{d(n)}{nh_n^{d/2}}\mathcal{H}_n(\pi) + \frac{d(n)}{\sqrt{n}h_n^{-m}}\mathcal{G}_n(\pi) + d(n)a_n^2\int_{\mathbf{R}^d}\gamma^2(x)\pi(x)dx$$

$$+2d(n)h_n^m a_n\int_{\mathbf{R}^d}\Delta^m f_o(x)\gamma(x)\pi(x)dx$$

$$+o_p(d(n)a_n^2) + o_p[d(n)h_n^m a_n] + o_p(1),$$

$$T_n^{2,0}(\pi) = T_n^{2,1}(\pi) + \frac{d(n)a_n}{nh_n^d}\int_{\mathbf{R}^d}\gamma(x)\pi(x)dx\int_{\mathbf{R}^d}K^2(x)dx$$

$$+o_p\left(\frac{d(n)a_n}{nh_n^d}\right) + o_p(1).$$

3.c **Definition of the critical regions**

From property 4, we deduce that the different statistics $T_n^{j,k}(\pi)$ are asymtotically normal, and that in particular the statistics $T_n^{1,0}$ and $T_n^{2,0}$ (resp $T_n^{2,0}$

and $T_n^{2,1}$) are asymptotically equivalent under the null hypothesis. They are all asympotically equivalent if $\lambda = \lim n h_n^{d+2m} = 0$. Moreover it is easily checked that for a <u>fixed</u> alternative $H_1 = \{f = f_1\}$, with $f_1 \neq f_o$, we have

$\lim_{n \to \infty} T_n^{j,k}(\pi) = +\infty$ in probability, as soon as the support of the weight function contains the support of f_o. Therefore we have to introduce one sided critical regions based on the previous statistics.

Property 5 If the support of the weight function contains the support of f_o the tests associated with the critical regions :

$$\begin{aligned}
W_n^{1,k}(\pi) &= \left\{T_n^{1,k}(\pi) \geq \Phi^{-1}(1-\alpha)[2\nu_1^2(\pi)]^{1/2}\right\}, k = 0, 1, \\
W_n^{2,k}(\pi) &= \left\{T_n^{2,k}(\pi) \geq \Phi^{-1}(1-\alpha)[2\nu_1^2(\pi) + 4\lambda\nu_2^2(\pi)]^{1/2}\right\}, \\
&\quad k = 0, 1, \text{if } \lambda \in [0, +\infty[\\
W_n^{2,k}(\pi) &= \left\{T_n^{2,k}(\pi) \geq 2\Phi^{-1}(1-\alpha)|\nu_2(\pi)|\right\}, \\
&\quad k = 0, 1, \text{if } \lambda = +\infty,
\end{aligned}$$

where Φ is the cdf of the standard normal, are consistent, asymptotically of level α

This property extends to the case of multidimensional and β-mixing processes theorem 4.1 of Bickel-Rosenblatt (1973).

In a general situation $\nu_2^2(\pi)$, the asymptotic variance of $\mathcal{G}_n/2$ under H_o, cannot be analytically computed. In this case we may consider the following consistent estimator of $\nu_2^2(\pi)$:

$$\hat{\nu}_{2n}^2 = \text{var}_o\left((\Delta^m f_o.\pi)(X_o)\right)$$

$$+2\sum_{j=1}^{m(n)} \left\{\frac{1}{n-j}\sum_{i=1}^{n-j}(\Delta^m f_o.\pi)(X_i)(\Delta^m f_o.\pi)(X_{i+j}) - \left(\frac{1}{n}\sum_{i=1}^{n}(\Delta^m f_o.\pi)(X_i)\right)^2\right\};$$

we have, under $H_o : \hat{\nu}_{2n}^2 \xrightarrow[n \to +\infty]{P} \nu_2^2(\pi)$, if $\dfrac{m^3(n)}{n} \to +\infty$.

In the sequel we study the asymptotic local power of the previous tests by assuming that $\hat{\nu}_{2n}^2$ is also under the sequence alternatives a consistent estimator of the variance of $\mathcal{G}_n/2 : \hat{\nu}_{2n}^2 \xrightarrow[n \to +\infty]{P} \nu_2^{*2}(\pi)$. This condition is satisfied if we have :

$$E_n[(\Delta^m f_o.\pi)(X_i)(\Delta^m f_o.\pi)(X_o)] \longrightarrow c_i, n \to +\infty,$$

where c_i is given in property 2.

4. LOCAL POWER FUNCTIONS

4.a Implicit local alternatives

We consider the tests defined by the critical regions $W_n^{j,k}, j = 1,2, k = 0,1$ and we note by $P_{g_n}(W_n^{j,k})$ the probability of such critical regions computed under the local alternative g_n defined by (2.5) and satisfying (P).

We say that the sequence $(\underline{a_n})$ (resp. $(\bar{a}_n)$) of positive real numbers, converging to zero when $n \to +\infty$, is a minimal (resp. maximal) sequence for the test $W_n^{j,k}$ if for all sequence (g_n) of local alternatives satisfying (P) we have $\lim_{n\to+\infty} P_{g_n}(W_n^{j,k}) = \alpha$ as soon as $\dfrac{a_n}{\underline{a_n}} = o(1)$(resp $\lim\limits_{n\to+\infty} P_{g_n}(W_n^{j,k}) = 1$ as soon as $\dfrac{\bar{a}_n}{a_n} = o(1)$).

Moreover, if the minimal (resp. maximal) sequence $(\underline{a}_n^I)[(\bar{a}_n^I)]$ satisfies $\underline{a_n} = 0(\bar{a}_n^I)$[resp $.\bar{a}_n^I = 0(\bar{a}_n)$] for all minimal sequences $(\underline{a}_n)$ (resp maximal $(\bar{a}_n)$) corresponding to $W_n^{j,k}$, we say that $(\underline{a}_n^I)$ (resp $(\bar{a}_n^I)$) is a minimal (resp maximal) limit sequence for $W_n^{j,k}$. If there both exist a minimal limit sequence $(\underline{a}_n^I)$ and a maximal limit sequence $(\bar{a}_n^I)$ for $W_n^{j,k}$, such that $\underline{a}_n^I = \bar{a}_n^I = a_n^I$ (say), (a_n^I) is said to be an implicit sequence of local alternatives.

If a_n^I exists, the faster a_n^I is tending to zero, the better is the testing procedure. Of course these implicit rates will depend on the testing procedure considered and on the choice of the bandwidth h_n.

In the discussion below we will consider that the support of the weight function is $I\!R^d$. To simplify this discussion and to be able to summarize it by some figures, we consider the case in which h_n is a power function :

$$h_n = n^{-a}. \tag{4.1}$$

a is strictly between 0 and 1/d because of assumption (B). In such a case the implicit rates are also power functions :

$$\bar{a}_n^I = n^{-\bar{b}}, \underline{a}_n^I = n^{-\underline{b}}, a_n^I = n^{-b}. \tag{4.2}$$

i) **Implicit rates for $T_n^{1,0}(\pi)$ and $T_n^{1,1}(\pi)$**

Using the expansions given in property 4, we get :

for the statistic $T_n^{1,1}(\pi) : \bar{b} = \underline{b} = b = \frac{1}{2} - \frac{ad}{4}$, which corresponds to $a_n^I = \frac{1}{\sqrt{n}h_n^{d/4}}$.

For the statistic $T_n^{1,0}(\pi)$ the minimal and maximal implicit rates do not coincide. We get :

$$\begin{aligned}\bar{b} &= \left(\frac{1}{2} - \frac{ad}{4}\right) 1\!\!1_{\{a \le 2/3d\}} + (1 - ad) 1\!\!1_{\{a > 2/3d\}},\\ \underline{b} &= \left(\frac{1}{2} - \frac{ad}{4}\right) 1\!\!1_{\{a \le 2/3d\}} + \frac{ad}{2} 1\!\!1_{\{a > 2/3d\}}.\end{aligned}$$

As easily seen when $a > 2/3d$, if b is between $\bar{b}$ and $\underline{b}$, the test statistic may tend to either $+\infty$, or $-\infty$ depending on the sign of $\int_{R^d} \gamma(x)\pi(x)dx$.

The previous conditions may also be written as :

$$\bar{a}_n^I = \frac{1}{\sqrt{n}h_n^{d/4}}\mathbb{1}_{\{n^{-2/3d}=0(h_n)\}} + \frac{1}{nh_n^d}\mathbb{1}_{\{h_n=o(n^{-2/3d})\}},$$

$$\underline{a}_n^I = \frac{1}{\sqrt{n}h_n^{d/4}}\mathbb{1}_{\{n^{-2/3d}=0(h_n)\}} + \frac{1}{h^{d/2}}\mathbb{1}_{\{h_n=o(n^{-2/3d})\}}.$$

When $\int_{\mathbb{R}^d}\gamma(x)\pi(x)dx = 0$ the implicit rates are only modified for $T_n^{1,0}(\pi)$, and $b + ad > 1$, where they may be strictly smaller than $ad/2$. These rates are summarized in the following figure.

figure 1 : **Implicit rates for** $T_n^{1,1}(\pi)$

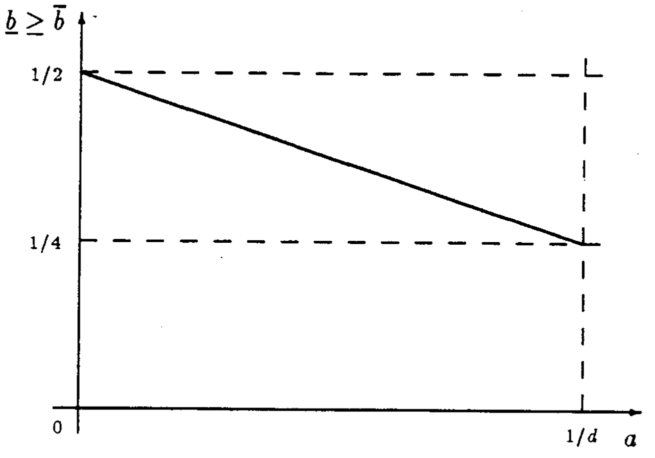
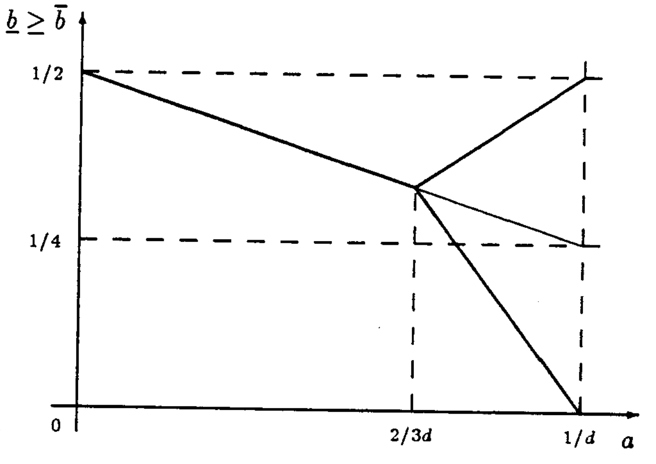


figure 2 : **Implicit rates for** $T_n^{1,0}(\pi)$



We can make the following remarks.

i) The implicit rate $1/\sqrt{n}$ (*i.e.* $b = 1/2$) associated with the parametric testing approaches cannot be reached.

ii) However it is possible to be very close to this rate, either by choosing a large bandwidth ($a \simeq o$) with $T_n^{1,1}(\pi)$ or $T_n^{1,0}(\pi)$, or by choosing a narrow bandwidth ($a \simeq 1/d$) with $T_n^{1,0}(\pi)$, but only for specific local alternatives.

iii) The implicit rate $a_n^I = \dfrac{1}{\sqrt{n}h_n^{d/4}}$ may only be achieved for $T_n^{1,1}(\pi)$ and for $T_n^{1,0}(\pi)$, when $a \leq \dfrac{2}{3d}$. If $a_n = a_n^I$, the asymptotic power of $W_n^{1,1}$ is an increasing function of $\int_{\mathbb{R}^d} \gamma^2(x)\pi(x)dx$. Indeed, $T_n^{1,1}(\pi)$ is asymptotically normal with mean $\int_{\mathbb{R}^d} \gamma^2(x)\pi(x)dx$ and variance $2\nu_1^2(\pi)$. We have :

$$\lim_{n\to+\infty} P_{g_n}(W_n^{1,1}(\pi)) = 1 - \Phi(\Phi^{-1}(1-\alpha) - \frac{\int_{\mathbb{R}^d} \gamma^2(x)\pi(x)dx}{(2\nu_1^2(\pi))^{1/2}}),$$

for all sequence (g_n) of local alternatives. The asymptotic test based on $T_n^{1,1}(\pi)$ is then locally unbiased. This is also true for the asymptotic test based on $T_n^{1,0}(\pi)$ if $n^{-2/3d} = o(h_n)$, since $T_n^{1,0}(\pi)$ and $T_n^{1,1}(\pi)$ are in this case asymptotically equivalent.

iv) The previous power function depends on the kernel K only through the asymptotic variance $\nu_1^2(\pi)$, i.e. through $\int_{\mathbb{R}^d}(K * \bar{K})^2(x)dx$. Therefore the power may be optimized if we choose a kernel giving the minimum of $\int_{\mathbb{R}^d}(K * \bar{K})^2(x)dx$. This problem has been solved by Ghosh-Huang (1991), when d=1. We can easily extend their result to the multidimensional case.

Let us consider the set of kernels of the form :

$$K(x) = \prod_{i=1}^{d} K_o(x_i),$$

where K_o is nonnegative such that :

$$\int_{\mathbb{R}} tK_o(t)dt = m, \int_{\mathbb{R}} (t-m)^2 K_o(t)dt = \sigma^2,$$

with m and σ^2 given. Then the optimal kernel is :

$$K^*(x) = \prod_{i=1}^{d} \left\{ \left(\frac{1}{2\sigma\sqrt{3}} \right) \mathbb{1}_{|x_i - m| < \sigma\sqrt{3}} \right\}.$$

The corresponding value of the objective function is :

$$\int_{\mathbb{R}^d} (K * \bar{K})^2(x)dx = \frac{1}{(3\sigma\sqrt{3})^d}.$$

ii) **Implicit rates for $T_n^{2,0}(\pi)$ and $T_n^{2,1}(\pi)$**

The approach is similar to the previous one. Using the expansion given in property 4, we get for the statistic $T_n^{2,1}(\pi)$:

$$\begin{aligned} \underline{b} &= \frac{1}{2}\mathbb{1}_{\left\{a \leq \frac{1}{d+2m}\right\}} + \left[1 - \frac{a}{2}(d+2m)\right] \mathbb{1}_{\left\{\frac{1}{d+2m} < a \leq \frac{1}{d/2+2m}\right\}} \\ &+ \left(\frac{1}{2} - \frac{ad}{4}\right) \mathbb{1}_{\left\{a > \frac{1}{d/2+2m}\right\}}, \end{aligned}$$

$$\bar{b} = am\mathbb{1}_{\left\{a \leq \frac{1}{d/2+2m}\right\}} + \left(\frac{1}{2} - \frac{ad}{4}\right) \mathbb{1}_{\left\{a > \frac{1}{d/2+2m}\right\}}.$$

Similarly we get for the statistic $T_n^{2,0}(\pi)$,

if $m \leq \frac{d}{2}$:

$$\begin{aligned} \underline{b} &= \frac{1}{2}\mathbb{1}_{[a \leq \frac{1}{d+2m}]} + [1 - \frac{a}{2}(d+2m)]\mathbb{1}_{[\frac{1}{d+2m} < a \leq \frac{1}{d+m}]} \\ &+ \frac{ad}{2}\mathbb{1}_{[\frac{1}{d+m} < a]}; \end{aligned}$$

if $m > \frac{d}{2}$:

$$\begin{aligned}
\underline{b} &= \frac{1}{2}\mathbb{1}_{[a \leq \frac{1}{2m+d}]} + [1 - \frac{a}{2}(d+2m)]\mathbb{1}_{[\frac{1}{d+2m} < a \leq \frac{1}{d/2+2m}]} \\
&+ (\frac{1}{2} - \frac{ad}{4})\mathbb{1}_{[\frac{1}{d/2+2m} < a \leq \frac{2}{3d}]} + \frac{ad}{2}\mathbb{1}_{[a > \frac{2}{3d}]}.
\end{aligned}$$

$$\begin{aligned}
\bar{b} &= am\mathbb{1}_{\{a \leq \frac{1}{d/2+2m}\}} + \left(\frac{1}{2} - \frac{ad}{4}\right)\mathbb{1}_{\left\{\frac{1}{d/2+2m} < a \leq \frac{2}{3d}\right\}} \\
&\qquad + (1 - ad)\mathbb{1}_{\left\{a > \frac{2}{3d}\right\}}.
\end{aligned}$$

These rates are summarized in the figures below for different values of m.

figure 3 : **Implicit rates for** $T_n^{2,1}(\pi)$

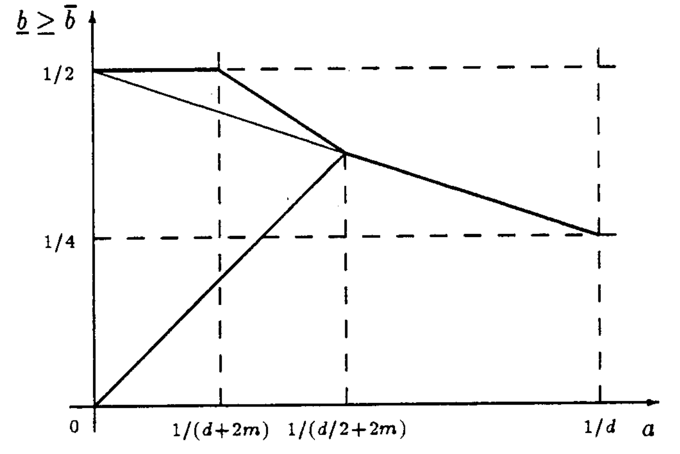


figure 4 : **Implicit rates for** $T_n^{2,0}(\pi), d < 2m$

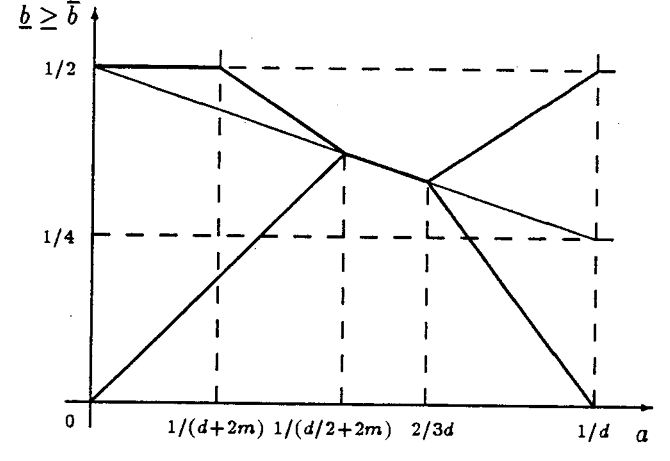
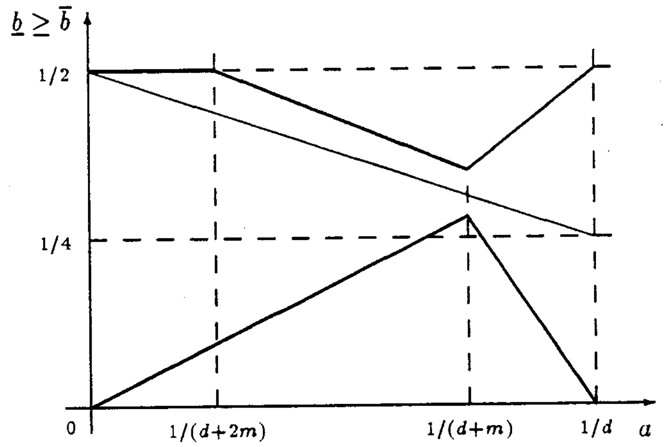


figure 5 : **Implicit rates for** $T_n^{2,0}(\pi), d \geq 2m$



A number of remarks concerning figure 1 are also valid in this second case. But the most important fact is that the minimal implicit rates correspond exactly to the parametric one $\underline{b} = \frac{1}{2}$, if $nh_n^{d+2m} \longrightarrow \lambda \in]0, +\infty[$ and $a \leq \frac{1}{2m+d}$. Indeed under this condition the estimator of the bias term dominates.

The asymptotic power of tests based on $I_n^2(\pi)$ under a sequence of local alternatives with $a_n = \frac{1}{\sqrt{n}}$ is, for $\lambda = +\infty$, independent of K, and higher than in the case $\lambda \in]0, +\infty[$.

4.b **Local unbiasedness condition**

We have seen that in some cases the minimal and maximal implicit local alternatives do not coincide, and that for intermediate sequences of local alternatives, the tests statistics may tend to either $+\infty$, or $-\infty$ depending on the position of the alternative.

Therefore for some choices of the bandwidth and of the test statistics, it is preferable to consider two sided testing procedures instead of the one sided ones introduced in property 5, if we want to satisfy a consistency condition (for the local alternatives such that $\frac{a_n}{\bar{a}_n^I} \to +\infty$), and an asymptotic local unbiasedness condition.

To get a better understanding of what happens in our problem, let us consider the properties of the critical region :

$$W_n^{1,0}(\pi) = \left\{T_n^{1,0}(\pi) \geq \Phi^{-1}(1-\alpha)[2\nu_1^2(\pi)]^{1/2}\right\}$$

when $h_n = n^{-2/3d}$ and for a sequence (g_n) of implicit local alternatives with $a_n = a_n^I$ and $\gamma(x) = \delta\gamma_o(x)$, where $\gamma_o(x)$ is a given function and δ is free. Under this sequence we have asymptotically :

$$T_n^{1,0}(\pi) \sim N(\delta \int \gamma_o(x)\pi(x)dx \int K^2(x)dx + \delta^2 \int \gamma_o^2(x)\pi(x)dx, 2\nu_1^2(\pi)). \quad (4.3)$$

Let us introduce the function of δ defined by :

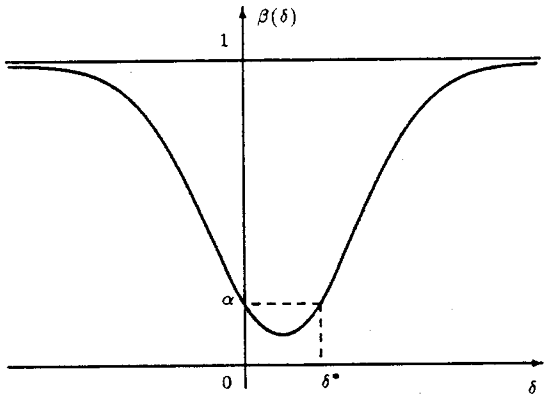
$$\beta(\delta) = \lim_{n \to +\infty} P_{g_n}[T_n^{1,0}(\pi) > \Phi^{-1}(1-\alpha)[2\nu_1^2(\pi)]^{1/2}],$$

where the probability is taken under the previous sequence of alternatives. The function β is a power function corresponding to these local alternatives. We have :

$$\beta(\delta) = 1 - \Phi\left\{\Phi^{-1}(1-\alpha) - \frac{\delta \int \gamma_o(x)\pi(x)dx \int K^2(x)dx + \delta^2 \int \gamma_o^2(x)\pi(x)dx}{[2\nu_1^2(\pi)]^{1/2}}\right\} \quad (4.4)$$

It has the following form, where it can be seen that the problem of bias is only a local one in a neighbourhood of the null hypothesis $\{\delta = 0\}$.

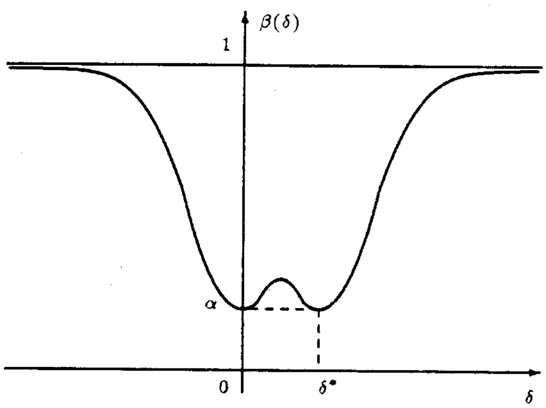
Figure 6 : **Power function of the one sided test $W_n^{1,0}(\pi)$ for local alternatives**



(with $\delta_o^* = -\dfrac{\int \gamma_o(x)\pi(x)dx \int K^2(x)dx}{\int \gamma_o^2(x)\pi(x)dx}$ assumed to be positive)

We also deduce that the introduction of a two sided version of the previous test is not a good solution since it is not strictly unbiased.

Figure 7 : **Power function of the two sided test based on $T_n^{1,0}(\pi)$ for local alternatives**



This lack of local unbiasedness is a feature of these kernel based testing procedures. It is a consequence of the practice of just correcting the bias of the basic statistics $I_n^j(\pi)$ under the null, and not under the local alternatives. By this usual bias correction it is only possible to get uniform local unbiasedness for $T_n^{1,0}(\pi)$, if $n_n^{-2/3d} = o(h_n)$ and $a < \dfrac{2}{3d}$.

4.c **The parametric implicit rates**

Among the different previous approaches, it seems interesting to have a special treatment of tests reaching the parametric minimal implicit rates, i.e of the tests based on $I_n^2(\pi)$ with :

$$h_n^{-1} = o\left(n^{1/(d+2m)}\right). \tag{4.5}$$

In this case $T_n^{2,0}(\pi) = T_n^{2,1}(\pi) + o_p(1)$, and we directly deduce from property 4, that under (4.3) and $a_n = 1/\sqrt{n}$, we have :

$$T_n^{2,0}(\pi) = \mathcal{G}_n(\pi) + 2\int_{\mathbb{R}^d} \Delta^m f_o(x)\gamma(x)\pi(x)dx + o_p(1). \tag{4.6}$$

In particular for such a sequence of local parametric alternatives, the statistics are asymptotically normal with mean $2\int_{\boldsymbol{R}^d} \Delta^m f_o(x)\gamma(x)\pi(x)dx$, and variance $4\nu_2^2(\pi)$. The mean is not of constant sign, which denotes a lack of local unbiasedness of the one sided tests associated with these statistics $T_n^{2,1}(\pi), T_n^{2,0}(\pi)$.

Even if we consider a two sided test based on $T_n^{2,0}(\pi)$ and we restrict the class of local alternatives to those for which $\int_{\boldsymbol{R}^d} \Delta^m f_o(x)\gamma(x)\pi(x)dx \neq 0$,

we are not able to obtain a locally unbiased test. Indeed under (4.3) with $a_n = \delta h_n^m$, for $\delta \neq 0$, we have :

$$T_n^{2,0}(\pi) = \mathcal{G}_n(\pi) \quad +\delta\sqrt{n}h_n^m \left[\delta \int_{\boldsymbol{R}^d} \gamma^2(x)\pi(x)dx + 2\int_{\boldsymbol{R}} \Delta^m f_o(x)\gamma(x)\pi(x)dx\right]$$

$$+o_p(\sqrt{n}h_n^m) + o_p(1),$$

and then under this sequence of local alternatives the probability of the critical region may not converge to 1 if we choose

$\delta = -2\int_{\boldsymbol{R}^d} \Delta^m f_o(x)\gamma(x)\pi(x)dx / \int_{\boldsymbol{R}^d} \gamma^2(x)\pi(x)dx..$

5. APPLICATION TO PARAMETRIC ALTERNATIVES

As noted before even if the parametric minimal implicit rate is reached by the statistic $T_n^{2,0}(\pi)$, for instance when $nh_n^{d+2m} \to +\infty$, there is a problem of bias. In this section we explain how to circumvent these difficulties in a parametric framework by introducing several statistics $(T_n^{2,0}(\pi_1), \ldots, T_n^{2,0}(\pi_p))$, associated with different weight functions.

5.a **Description of the hypotheses**

In all the following sections we consider iid observations.

Let $g(.;\theta), \theta \in \Theta \subset I\!R^K$, be a parametric family of bounded strictly positive pdf on $I\!R^d$ satisfying the identifiability condition for θ.

We consider the problem of testing $H_o = \{f = g(.;\theta_o)\}$ against

$H = \{f \in \{g(.;\theta), \theta \in \Theta\}\}$, where $\Theta \subset I\!R^K$, and $\theta_o \in \overset{o}{\Theta}$. The parametric local alternatives are usually defined in terms of the parameter itself by :

$$\theta_n = \theta_o + \delta a_n, \text{ where } \delta = (\delta_1, \ldots, \delta_K)'. \tag{5.1}$$

To this definition corresponds a definition of local alternatives in terms of distribution [we assume the existence, continuity and derivability of the relevant functions] :

$$\begin{aligned} g_n(x) = g(x;\theta_n) &= g(x;\theta_o) + a_n\delta'\frac{\partial g}{\partial\theta}(x;\theta_o) \\ &+ a_n^2\int_0^1 \delta'\frac{\partial^2 g}{\partial\theta\partial\theta'}(x;\theta_o + ta_n\delta)\delta(1-t)dt. \end{aligned}$$

Therefore with the notation of the previous sections, we get :

$$\gamma(x) = \delta'\frac{\partial g}{\partial\theta}(x;\theta_o), x \in I\!R^d. \tag{5.2}$$

We assume in the sequel that the components of $\frac{\partial g}{\partial\theta}(.;\theta_o)$ are linearly independent, i.e. that the information matrix : $I(\theta) = \underset{\theta_o}{E}\left[\frac{\partial \log g}{\partial\theta}(X;\theta_o)\frac{\partial \log g}{\partial\theta'}(X;\theta_o)\right]$ is invertible (if the usual regularity conditions for deriving under the integral are satisfied).

4.b Multivariate kernel based statistics

Let us consider the multivariate statistic :

$$T_n^{2,0}(\Pi^p) = [T_n^{2,0}(\pi_1), \ldots, T_n^{2,0}(\pi_p)],$$

where $\Pi^p = (\pi_1, \ldots, \pi_p)'$. We assume $nh_n^{d+2m} \to +\infty$, for some $m \in \mathbf{N}$ and $K \in \mathcal{K}^d(m)$.

From property 4, if $g(.;\theta_o) \in W^d(m)$, and if $\Delta^m g(.;\theta_o)$ is defined by (2.5), under a sequence of parametric local alternatives, we have :

$$\begin{aligned} T_n^{2,0}(\Pi^p)/2 = \mathcal{G}_n(\Pi^p) + \sqrt{n}h_n^{-m}&\left\{a_n^2\frac{C(\Pi^P;\delta)}{2}\right. \\ &\left.+h_n^m a_n B(\Pi^p)\right\}\delta + o_p(\sqrt{n}h_n^{-m}a_n^2) + o(\sqrt{n}a_n) + o_p(1), \end{aligned} \tag{5.3}$$

where $\mathcal{G}_n(\Pi^p) = (\mathcal{G}_n(\pi_1), \ldots, \mathcal{G}_n(\pi_p))$,

$$B(\Pi^p) = E_o\left[\Pi^p(X_o)\Delta^m g(X_o;\theta_o)\frac{\partial \log g}{\partial \theta'}(X_o;\theta_o)\right],$$

$$\text{and } C(\Pi^p;\delta) = E_o\left[\Pi^p(X_o)\delta'\frac{\partial g}{\partial \theta}(X_o;\theta_o)\frac{\partial \log g}{\partial \theta'}(X_o;\theta_o)\right].$$

In particular, since we consider iid observations, it is easily seen that under the null hypothesis :

$$\frac{T_n^{2,0}(\Pi^p)}{2} E_o N[o, A(\Pi^p)],$$

where : $A(\Pi^p) = \text{Var}_o[\Delta^m g(X_o;\theta_o)\Pi^p(X_o)]$.

Let us introduce the following assumption on the weight function.

Assumption (Π) : The components of Π^p are almost everywhere continuous on $I\!R^d$, and the support of at least one of them is $I\!R^d$.

Property 6 : Let us assume $A(\Pi^p)$ invertible. Under (K), (B), (Π), (P), if $g(.;\theta_o) \in W^d(m), K \in K^d(m)$ for some $m \in \mathbf{N}$, and if $h_n^{-1} = o(n^{1/(d+2m)})$, the critical region defined by

$$W_n^{2,0}(\Pi^p) = \left\{T_n^{2,0}(\Pi^p)'\hat{A}_n(\Pi^p)^{-1}T_n^{2,0}(\Pi^p) > 4\chi^2_{1-\alpha}(p)\right\},$$

where $\hat{A}_n(\Pi^p)$ is a consistent estimator of $A(\Pi^p)$ defines a tests of asymptotic level α. This test is consistent for testing $H_o = \{\theta = \theta_o\}$ against $H = \{\theta \neq \theta_o\}$.

Note that the components π_j may correspond to signed measures.

5.c Asymptotic properties of $W_n^{2,0}(\Pi^p)$ under local alternatives

Under a sequence of parametric local alternatives, and from (5.3) we know that the statistic

$$\xi_n(\Pi^p) = T_n^{2,0}(\Pi^p)'\hat{A}_n(\Pi^p)^{-1}T_n^{2,0}(\Pi^p),$$

is such that :

$$\xi_n(\Pi^p) \xrightarrow{d} 4\chi^2(p), \text{if } a_n = o\left(\frac{1}{\sqrt{n}}\right),$$

$$\xi_n(\Pi^p) \xrightarrow{d} 4\chi^2(p; \delta' B(\Pi^p) A(\Pi^p)^{-1} B(\Pi^p)\delta), \text{if } a_n = \frac{1}{\sqrt{n}},$$

$$\xi_n(\Pi^p) \to +\infty \text{ in probability if, } \frac{1}{\sqrt{n}} = o(a_n) \text{ and } a_n \neq h_n^m,$$

since $\| C(\Pi^p; \delta)\delta \| > o$ for all $\delta \neq o$ from assumption (II).

Finally in the remaining limit case $a_n = h_n^m$, we have :

$\xi_n(\Pi^p)$ tends to infinity in probability if and only if the following condition is satisfied

$$(LU) \ : \left[\frac{C(\Pi^p; \delta)}{2} + B(\Pi^p)\right] \delta = 0 \Longrightarrow \delta = 0.$$

When $K = 1$, it is necessary to choose $p \geq 2$ to satisfy the condition (LU) for suitable choices of π_j.

When $K \geq 2$, the condition (LU) is generally satisfied for $p = K$, and even for smaller values of p. For instance for $K = 2, p = 1$ the LU condition corresponds to a non homogenous equation of degree two with respect to the ratio δ_1/δ_2 of the components of δ ; it is satisfied if this equation has complex roots.

Property 7 : Under the conditions of property 6, if $rkB(\Pi^p) = K$ and if the assumption (LU) is satisfied, there exists a sequence of implicit local alternatives for $W_n^{2,0}(\Pi^p)$, which corresponds to the usual parametric rate $a_n^I = \frac{1}{\sqrt{n}}$.

Moreover this test is locally unbiased and its asymptotic power is in one to one relationship with $B(\Pi^p)' A(\Pi^p)^{-1} B(\Pi^p)$.

4.c Optimal choices of Π^p

Considering $p \leq K$ and different weight functions Π^p, the associated critical regions may be compared through the matrix $B(\Pi^p)A(\Pi^p)^{-1}B(\Pi^p) = J(\Pi^p)$(say), and in particular, we will reach the asymptotic optimality if and only if :

$$J(\Pi^p) = I(\theta_o), \tag{5.6}$$

where $I(\theta_o)$ is the information matrix associated with the model $[g(.;\theta), \theta \in \Theta]$. The following property is immediately derived.

Property 8 : Let us assume that for some $p \leq K$, the weight function

$\Pi^p = (\Pi^K, \Pi^{p-K})$ satisfying (Π) is such that $B(\Pi^{p-K}) = 0$ and $\Pi^K = C\Pi^*$, where C is an invertible matrix and Π^* is the vector function

$$\Pi^*(x) = \begin{cases} C\dfrac{\partial \log g}{\partial \theta}(x;\theta_o)[\Delta^m g(x;\theta_o)]^{-1}, & \text{if } \Delta^m g(x;\theta_o) \neq o, \\ 0 & \text{if } \Delta^m g(x;\theta_o) = o. \end{cases}$$

If we have $\Delta^m g(x;\theta_o) \neq$ almost everywhere, then $J(\Pi^p) = I(\theta_o)$, and under the conditions of property 6 the test $W_n^{2,0}(\Pi^*)$ is asymptotically LUMPU.

<u>Proof :</u> Indeed we get :

$$\begin{aligned} B(\Pi^K) = &\ CE_o\left[\frac{\partial \log g}{\partial \theta}(X_o;\theta_o)\frac{\partial \log g}{\partial \theta'}(X_o;\theta_o)\right] = CI(\theta_o), \\ A(\Pi^K) = &\ V_o\left[C\frac{\partial \log g}{\partial \theta}(X_o;\theta_o)\right] = CI(\theta_o)C'. \end{aligned}$$

Then since $B(\Pi^{p-K}) = 0$, we have :

$$J(\Pi^p) = B(\Pi^K)'A(\Pi^K)^{-1}B(\Pi^K) = I(\theta_o).$$

QED

When Π^* is not bounded, we can retain a truncated version $\Pi^*_L(x) = \Pi^*(x) \mathbb{1}_{\{\|\Pi^*(x)\| < L\}}$, so that the loss of power may be taken as small as possible.

5.d **Examples**

1) **Translation parameter**

When $g(x;\theta) = g(x-\theta), d = K = 1$, and $H_o = \{\theta = \theta_o\}$, we can choose :

$$\pi_1(x) \quad = \begin{cases} g'(x-\theta_o)[g(x-\theta_o)g^m(x-\theta_o)]^{-1}, & \text{if } g^{(m)}(x-\theta_o) \neq 0, \\ 0 & \text{, otherwise.} \end{cases}$$

For $m = 1$, we have :

$$\pi_1(x) \quad = \begin{cases} g^{-1}(x-\theta_o), & \text{if } g'(x-\theta_o) \neq 0, \\ 0 & \text{, otherwise.} \end{cases}$$

The test statistic $T_n^{2,0}(\pi_1)$ is a bias corrected version of :

$$I_n^2(\pi_1) = \int_{\mathbb{R}} \frac{[f_n(x) - g(x-\theta_o)]^2}{g(x-\theta_o)} dx,$$

as soon as the support of g is $\mathbb{R}$. We note that the kernel based testing procedures comes from a proximity measure with the same weights as the usual Pearson χ^2 statistic [Neyman (1937), Mann-Wald (1942), Bosq (1989)].

ii) **Translation and scale parameters**

When $g(x;\theta) = \dfrac{1}{\sigma} g\left(\dfrac{x-\mu}{\sigma}\right), d = 1, K = 2$, and $H_o = \{\mu = \mu_o, \sigma = \sigma_o\}$, we can choose :

$$\begin{cases} \pi_1(x) & = \dfrac{1}{\sigma_o} g'\left(\dfrac{x-\mu_o}{\sigma_o}\right) g^{-1}\left(\dfrac{x-\mu_o}{\sigma_o}\right) \left\{g^{(m)}\left(\dfrac{x-\mu_o}{\sigma_o}\right)\right\}^{-1}, \\ \pi_2(x) & = \left\{\dfrac{1}{\sigma_o} + \dfrac{x-\mu_o}{\sigma_o^2} g'\left(\dfrac{x-\mu_o}{\sigma_o}\right) g^{-1}\left(\dfrac{x-\mu_o}{\sigma_o}\right)\right\} \left\{g^{(m)}\left(\dfrac{x-\mu_o}{\sigma_o}\right)\right\}^{-1}, \end{cases}$$

if $g^{(m)}\left(\dfrac{x-\mu_o}{\sigma_o}\right) \neq 0$, and 0 otherwise.

From properties 7 and 8 we know that the asymptotic optimality is reached by choosing $\Pi^p = (\Pi^K, \Pi^{p-K})$ such that the conditions (Π) and (LU) are satisfied, and Π^{p-K} satisfies $B(\Pi^{p-K}) = 0$ and $rkA(\Pi^{p-K}) = p - K$.

Returning to example i) we can discuss the choices of the order m and of the weight function π_2 to satisfy these conditions. We have :

$$\begin{aligned}
B(\pi_1) &= -C_K \int_{\boldsymbol{R}} g'(x)^2 [g(x)]^{-1} dx = -C_K I(\theta_o), \\
C(\pi_1;\delta) &= \delta \int_{\boldsymbol{R}} g'(x)^3 [g^{(m)}(x)]^{-1} [g(x)]^{-1} dx, \\
B(\pi_2) &= -C_K \int_{\boldsymbol{R}} g^{(m)}(x - x_o) g'(x - x_o) \pi_2(x) dx, \text{ and} \\
C(\pi_2;\delta) &= \delta \int_{\boldsymbol{R}} g'(x - x_o)^2 \pi_2(x) dx.
\end{aligned}$$

where $C_K = \dfrac{(-1)^m}{m!} \displaystyle\int_{\boldsymbol{R}} x^m K(x) dx$.

Independently of π_2, if we choose $m = 1$, we get :

$$\begin{cases} B(\pi_1) = -C_K C(\pi_1;\delta)/\delta, \\ B(\pi_2) = -C_K C(\pi_2;\delta)/\delta, \end{cases}$$

and then the condition (LU) is not satisfied.

Let us now consider $m = 2$ (or m even), and for instance

$$\pi_2(x) = 1 \text{ or } \pi_2(x) = [(g(x - \theta_o)]^{-1}, x \in I\!R.$$

The vector function $\Pi^2 = (\pi_1, \pi_2)$ satisfies condition (Π), but it is not bounded. However the loss of power may be taken as small as possible if we retain a truncated version of Π^2. if $\pi_2(x) = 0$, we have $B(\pi_2) = 0, C(\pi_2;\delta) \neq 0$ for all $\delta \neq 0$, and the condition (LU) is satisfied. Moreover $A(\Pi^2)$ is a diagonal matrix with now zero diagonal elements and then is invertible. The same conclusion applies for $\pi_2(x) = [g(x - \theta_o)]^{-1}$ as soon as $g(x) = g(-x), x \in I\!R$.

Appendix 1

Asymptotic Normality of $\mathcal{G}_n(\pi), \mathcal{H}_n(\pi)$

Under H_o and for $m = 2$, property 2 has been derived by Tenreiro (1995) as a consequence of a Central Limit Theorem for U-statistics generated by a β- mixing process, whose mixing coefficients have an exponential decrease [this result follows the lines of Hall (1984), Takahata-Yoshihara (1987), and is based on a Central Limit Theorem for triangular array given by Dvoretsky (1972)]. The introduction of local alternatives does not modify the approach. Therefore under the sequence of local alternatives $[\mathcal{G}_n(\pi), \mathcal{H}_n(\pi)]$ is asymptotically normal with zero-mean and a diagonal variance-covariance matrix, where :

$$\lim_{n \to +\infty} \mathrm{Var}_n(G_n(\pi)) = 4 \sum_{i=-\infty}^{+\infty} \lim_{n \to \infty} \underset{n}{E} [G_n(X_i) G_n(X_o)],$$

$$\lim_{n \to +\infty} \mathrm{Var}_n(H_n(\pi)) = \lim_{n \to \infty} \underset{n}{E} [G_{no}(X_o, X_o)],$$

with $G_{n,o}(u,v) = \underset{n}{E} [H_n(X_o, u) H_n(X_o, v)]$, for $u, v \in \mathbb{R}^d$.

We compute in the sequel these asymptotic variances.

i) Asymptotic variance of $G_n(\pi)$

For $i = 0, 1, 2, \ldots$ we have :

$$\underset{n}{E}\,[G_n(X_i)G_n(X_o)]$$

$$= \frac{1}{h_n^{2d}}\underset{n}{E}\int_{\mathbf{R}^d}\int_{\mathbf{R}^d} K\left(\frac{x-X_i}{h_n}\right)K\left(\frac{y-X_o}{h_n}\right)(\Delta_n^m f_o.\pi)(x)(\Delta_n^m f_o.\pi)(y)dxdy$$

$$-\left[\underset{n}{E}\left[\frac{1}{h_n^d}\int_{\mathbf{R}^d} K\left(\frac{x-X_o}{h_n}\right)(\Delta_n^m f_o.\pi)(x)dx\right]\right]^2$$

$$= \int_{\mathbf{R}^d}\int_{\mathbf{R}^d} K(u)K(v)\ \underset{n}{E}\,[(\Delta_n^m f_o.\pi)(X_i+uh_n)(\Delta_n^m f_o.\pi)(X_o+vh_n)dudv$$

$$-\left[\int_{\mathbf{R}^d} K(u)\ \underset{n}{E}\,[(\Delta_n^m f_o.\pi)(X_o+uh_n)]du\right]^2.$$

Then by applying the assumption ii) of property 2, and the dominated convergence theorem and by taking into account the form of g_n, we get for i=1,2.. :

$$\begin{aligned}
&\lim_{n\to+\infty}\ \underset{n}{E}\,[G(X_i)G_n(X_o)]\\
&= \int_{\mathbf{R}^d}\int_{\mathbf{R}^d} K(u)K(v)\{c_j - E_o^2[(\Delta^m f_o.\pi)(X_o)]\}dudv\\
&= c_j - E_p^2[(\Delta^m f_o.\pi)(X_o)],
\end{aligned}$$

and for $j=0$,

$$\lim_{n\to+\infty}\ \underset{n}{E}\,[G_n(X_o)G_n(X_o)] = \mathrm{Var}_o[(\Delta^m f_o.\pi)(X_o)]$$

Finally we have :

$$\lim_{n \to +\infty} \mathrm{Var}_n(G_n(\pi)) = 4\left[\mathrm{Var}_o(\Delta^m f_o.\pi)(X_o) + 2\sum_{j=1}^{\infty}\{c_j - E_o^2[(\Delta^m f_o.\pi)(X_o)]\}\right]$$

$$= 4\nu_2^{*2}.$$

ii) Asymptotic variance of $H_n(\pi)$

From the definition of $H_n(.,.)$ we have for $u, v \in \mathbb{R}^d$,

$$\begin{aligned} H_n(u,\nu) &= \frac{1}{h_n^{3d/2}}\int_{\mathbf{R}^d} K(\frac{x-u}{h_n})k(\frac{x-v}{h_n})\pi(x)dx \\ &- \frac{1}{h_n^{3d/2}}\int_{\mathbf{R}^d} K(\frac{x-u}{h_n})\ \underset{n}{E}\ K(\frac{x-X_o}{h_n})\pi(x)dx \\ &- \frac{1}{h_n^{3d/2}}\int_{\mathbf{R}^d} K(\frac{x-v}{h_n})\ \underset{n}{E}\ K(\frac{x-X_o}{h_n})\pi(x)dx \\ &+ \frac{1}{h_n^{3d/2}}\int_{\mathbf{R}^d} E_n^2 K(\frac{x-X_o}{h_n})\pi(x)dx, \end{aligned}$$

where from assumptions $(K), (\Pi)$ and (P) any of the last three terms is bounded by

$$h_n^{d/2} \sup_{x \in \mathbb{R}^d} (\pi(x)) \sup_{n \in \mathbf{N}} \sup_{x \in \mathbb{R}^d} |g_n(x)| \int_{\mathbf{R}^d} |K| * |\bar{K}|(u)du.$$

Uniformly on $u, v \in \mathbb{R}^d$, we have

$$H_n(u,v) = \frac{1}{h_n^{3d/2}}\int_{\mathbf{R}^d} K(\frac{x-u}{h_n})K(\frac{x-v}{h_n})\pi(x)dx + 0(h_n^{d/2}) \qquad (1),$$

and then,

$$\lim_{n \to +\infty} E_n[G_{no}(X_o, X_o)]$$

$$= \lim_{n \to +\infty} \frac{1}{h_n^{3d}} \int_{\mathbf{R}^d} \int_{\mathbf{R}^d} \left(\int_{\mathbf{R}^d} K\left(\frac{x-u}{h_n}\right) K\left(\frac{x-v}{h_n}\right) \pi(x) dx \right)^2 g_n(u) g_n(v) du dv.$$

From the form of g_n, and the dominated convergence theorem we get,

$$\lim_{n \to +\infty} \mathrm{Var}_n(H_n(\pi)) = \int_{\mathbf{R}^d} f_o^2(x) \pi^2(x) dx \int_{\mathbf{R}^d} (K * \bar{K})^2(z) dz = \nu_1^2(\pi).$$

Appendix 2
Asymptotic expansions of $I_n^1(\pi)$ and $I_n^2(\pi)$

i) Decompositions of the statistics

We can first decompose $I_n^1(\pi)$ and $I_n^2(\pi)$ in the following ways :

$$\begin{aligned}
I_n^1(\pi) &= \left\{ \int_{\mathbf{R}^d} [f_n(x) - E_n f_n(x)]^2 \pi(x)dx - E_n \int_{\mathbf{R}^d} [f_n(x) - E_n f_n(x)]^2 \pi(x)dx \right\} \\
&+ E_n \int_{\mathbf{R}^d} [f_n(x) - E_n f_n(x)]^2 \pi(x)dx \\
&+ 2 \int_{\mathbf{R}^d} [f_n(x) - E_n f_n(x)] [E_n f_n(x) - E_o f_n(x)] \pi(x)dx \\
&+ \int_{\mathbf{R}^d} [E_n f_n(x) - E_o f_n(x)]^2 \pi(x)dx \\
&= \frac{1}{n h_n^{d/2}} U_n + A_n^1 + 2A_n^2 + A_n^3 (\text{say}) \qquad (2)
\end{aligned}$$

where E_o is the expectation under the null hypothesis and E_n is the expectation under the local alternatives.

$$\begin{aligned}
I_n^2(\pi) &= I_n^1(\pi) + 2 \int_{\mathbf{R}^d} [f_n(x) - E_n f_n(x)] [E_o f_n(x) - f_o(x)] \pi(x)dx \\
&+ 2 \int_{\mathbf{R}^d} [E_n f_n(x) - E_o f_n(x)] [E_o f_n(x) - f_o(x)] \pi(x)dx \\
&+ \int_{\mathbf{R}^d} [E_o f_n(x) - f_o(x)]^2 \pi(x)dx \\
&= I_n^1(\pi) + \frac{1}{\sqrt{n} h_n^{-m}} 2B_n^1 + 2B_n^2 + \int_{\mathbf{R}^d} [E_o f_n(x) - f_o(x)]^2 \pi(x)dx (\text{say}). \qquad (3)
\end{aligned}$$

As usual the idea is to give expansions for each term of the decompositions.

ii) **Expansion of $I_n^1(\pi)$**

Let us consider decomposition (2).

a) We get :

$$\begin{aligned} U_n &= \frac{1}{n}\sum_{i=1}^{n}\{H_n(X_i,X_i) - E_n H_n(X_i,X_i)\} \\ &+ \frac{2}{n}\sum\sum_{1 \leq i < j \leq n}\{H_n(X_i,X_j) - E_n H_n(X_i,X_j)\} \\ &= \mathcal{H}_n(\pi) + o_p(1), \end{aligned}$$

since we easily check that the second order moment of the first term is negligible, using (1) and the absolute convergence of the mixing coefficients.

b) The second term is :

$$\begin{aligned} A_n^1 &= E_n \int_{\mathbf{R}^d}[f_n(x) - E_n f_n(x)]^2 \pi(x)dx. \\ &= \frac{1}{nh_n^{d/2}} E_n H_n(X_o,X_o) + \frac{2}{n^2 h_n^{d/2}} \sum\sum_{1 \leq i < j \leq n} E_n H_n(X_i,X_j), \end{aligned}$$

where :

$$\begin{aligned} &E_n H_n(X_o,X_o) \\ &= \frac{1}{h_n^{3d/2}}\int_{\mathbf{R}^d}\left\{E_n K^2\left(\frac{x-X_o}{h_n}\right) - E_n^2 K\left(\frac{x-X_o}{h_n}\right)\right\}\pi(x)dx \\ &= \frac{1}{h_n^{3d/2}}\int_{\mathbf{R}^d}\int_{\mathbf{R}^d} K^2\left(\frac{x-y}{h_n}\right) g_n(y)\pi(x)dydx - \frac{1}{h_n^{3d/2}}\int_{\mathbf{R}^d} E_n^2\left(\frac{x-X_o}{h_n}\right)\pi(x)dx \\ &= \frac{1}{h_n^{d/2}}\int_{\mathbf{R}^d}\int_{\mathbf{R}^d} K^2(u)g_n(y)\pi(y+uh_n)dydu + 0(h_n^{d/2}). \end{aligned}$$

We now check that the covariance term is negligible. Firstly from (1) and the assumptions on the pdf g_n and $g_{i,n}$, we conclude that for $r > 1$, there exists $c > 0$ such that :

$$\max\left\{\max_{1\le i\le n} E_n^{1/r}|H_n(\bar{X}_i, X_o)|^r, E_n^{1/r}|H_n(X_o, X_o)|^r\right\} \le C(h_n^d)^{\frac{1}{r}-\frac{1}{2}},$$

where $\bar{X}_i$ is an independent copy of X_o. From lemma 1 of Yoshihara (1976), we get :

$$\begin{aligned}\left|\frac{1}{n}\sum\sum_{1\le i<j\le n} E_n H_n(X_i, X_j)\right| &\le \sum_{i=1}^{n}|E_n H_n(X_i, X_o)| \\ &\le 4C(h_n^d)^{\frac{1}{r}-\frac{1}{2}}\sum_{i=1}^{n-1}\beta_n^{\frac{r-1}{r}}(i),\end{aligned}$$

where the RHS is negligible for $r < 2$, from the assumption on the mixing coefficients.

Therefore :

$$A_n^1 = \frac{1}{nh_n^d}\int_{\boldsymbol{R}^d}\int_{\boldsymbol{R}^d} K^2(u)g_n(x)\pi(x+uh_n)dxdu + o\left(\frac{1}{nh_n^{d/2}}\right).$$

c) The third term is :

$$\begin{aligned}A_n^2 &= \int_{\boldsymbol{R}^d}[f_n(x) - E_n f_n(x)][E_n f_n(x) - E_o f_n(x)]\pi(x)dx \\ &= \frac{a_n}{\sqrt{n}}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{1}{h_n^d}\left\{K\left(\frac{x-X_i}{h_n}\right) - E_n K\left(\frac{x-X_o}{h_n}\right)\right\}p_n(x)\pi(x)dx, \\ &= \frac{a_n}{\sqrt{n}}\frac{1}{\sqrt{n}}\sum_{i=1}^{n} L_n(X_i) \text{ (say)}.\end{aligned}$$

where :

$$
\begin{aligned}
p_n(x) &= \frac{1}{a_n}\left[E_n f_n(x) - E_o f_n(x)\right] \\
&= \frac{1}{h_n^d}\int_{\boldsymbol{R}^d} K\left(\frac{x-y}{h_n}\right)\left[\gamma(y) + o(1)\gamma_n(y)\right] dy,
\end{aligned}
$$

is uniformly bounded using the assumptions on the sequence of local alternatives. Since $L_n(.)$ is uniformly bounded and the sequence of mixing coefficients is absolutely convergent, we have :

$$
A_n^2 = 0\left(\frac{a_n}{\sqrt{n}}\right).
$$

d) Finally from the dominated convergence theorem, we have :

$$
\begin{aligned}
A_n^3 &= \int_{\boldsymbol{R}^d}[E_n f_n(x) - E_o f_n(x)]^2 dx \\
&= \int_{\boldsymbol{R}^d}\left\{\frac{1}{h_n^d}\int_{\boldsymbol{R}^d} K\left(\frac{x-y}{h_n}\right)(g_n(y) - f_o(y))dy\right\}^2 \pi(x)dx \\
&= a_n^2\int_{\boldsymbol{R}^d}\left\{\frac{1}{h_n^d}\int_{\boldsymbol{R}^d} K\left(\frac{x-y}{h_n}\right)\gamma(y)dy\right\}^2 \pi(x)dx + o(a_n^2) \\
&= a_n^2\int_{\boldsymbol{R}^d}\gamma^2(x)\pi(x)dx + o(a_n^2).
\end{aligned}
$$

iii) **Expansion of $I_n^2(\pi)$**

We use for this expansion the decomposition (3) and the equality :

$$
E_o f_n(x) - f_o(x) = h_n^m \Delta_n^m f_o.(x),
$$

which is valid for $f_o \in W^d(m)$ and $K \in K^d(m)$.

a) We have :

$$
\begin{aligned}
& 2B_n^1 \\
= \;& 2\sqrt{n}h_n^{-m}\int_{\mathbb{R}^d}[f_n(x)-E_nf_n(x)][E_of_n(x)-f_o(x)]\pi(x)dx \\
= \;& 2\sqrt{n}\int_{\mathbb{R}^d}[f_n(x)-E_nf_n(x)]\Delta_n^m f_o.(\pi)(x)dx \\
= \;& 2\sqrt{n}\frac{1}{nh_n^d}\sum_{i=1}^{n}\int_{\mathbb{R}^d}[K\left(\frac{x-X_i}{h_n}\right)-E_nK\left(\frac{x-X_i}{h_n}\right)]\Delta_n^m f_o.(\pi)(x)dx \\
= \;& \mathcal{G}_n(\Pi),
\end{aligned}
$$

b) The same kind of argument gives :

$$2B_n^2 = 2h_n^m\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}K(u)[g_n(x)-f_o(x)](\Delta_n^m f_o.\pi)(x+uh_n)dxdu.$$

By taking into account all these expansions, we deduce the expansion of $I_n^1(\pi)$ given in theorem 3.

REFERENCES

Bahadur, R (1960) : "Stochastic Comparison of Tests", Ann. Math. Statist., 31, 276-295.

Bahadur, R (1967) : "Rates of Convergence of Estimates and Tests Statistics", Ann. Math. Statist., 38, 303-324.

Bickel, P., and M., Rosenblatt (1973) : "On Some Global Measures of the Deviations of Density Function Estimates", Ann. Statist., 1, 1071-1095 [Correction : 3 (1975), 1370].

Bosq, D. (1983) : "Lois limites et efficacité asymptotique des tests hilbertiens de dimension finie sous des hypothèses adjacentes", Statistique et Analyse des Données, 8, 1-40.

Bosq, D. (1989) : "Tests du $\mathcal{X}^2$ généralisés : comparaison avec le test du X^2 classique", Revue de Statistique Appliquée, 37, 43-52.

Bradley, R. (1986) : "Basic Properties of Strong Mixing Conditions", in "Dependence in Probability and Statistics, a Survey of Recent Results", ed. Eberlein, E., and M., Taqqu, 165-192.

Dvoretsky, A. (1972) : "Central Limit Theorems for Dependent Random Variables", Proc.Sixth.Berkeley Symp. Math. Statist. Probab., Los Angeles, Univ. of California Press, 2, 513-563.

Eubank, R.L., and V.N., LaRiccia (1992) : "Asymptotic Comparison of Cramer-von Mises and Nonparametric Function Estimation Techniques for Testing Goodness of Fit", Ann. Statist., 20, 2071-2086.

Ghosh, B., and W.M., Huang (1991) : "The Power and Optimal Kernel of the Bickel-Rosenblatt Test for Goodness of Fit", Ann. Statist., 19, 999-1009.

Gregory, G. (1977) :"Large Sample Theory for U.Statistics and Tests of Fit", Ann. Statist., 5, 110-123.

Hall, P. (1984) : "Central Limit Theorem for Integrated Square Error of Multivariate Nonparametric Density Estimators", Journal of Multivariate Analysis, 14, 1-16.

Hardle, W., and E., Mammen (1993) : "Comparing Nonparametric Versus Parametric Regression Fits", Ann. Statist., 21, 1926-1947.

Khashimov, S.A. (1987) : "On the Asymptotic Distribution of the Generalized U-Statistic for Dependent Variables", Theory Probab. Appl., 5, 204-208.

Mann, H., and A., Wald (1942) : "On the Choice of the Number of Class Intervals in the Application of the Chi-Square Test", Ann. Math. Statist., 13, 306-317.

Meloche, J. (1990) :"Asymptotic Behaviour of the Mean Integrated Squared Error of Kernel Density Estimators for Dependent Observations", Can. J. Statist., 18, 205-211.

Neyman, J. (1937) : "Smooth Test for Goodness of Fit", Skand. Aktuar., 20, 109-128.

Noether, G. (1955) : "On a Theorem of Pitman", Ann. Math. Statist., 26, 64-68.

Pitman, E. (1979) : "Some Basic Theory for Statistical Inference", Chapman and Hall.

Rosenblatt, M. (1956) : "Remarks on Some Nonparametric Estimates of a Density Function", Ann. Math. Statist., 27, 832-837.

Rosenblatt, M. (1956) : "A Central Limit Theorem and a Strong Mixing Condition", Proc. Nat. Acad. Sci. USA, 42, 43-47.

Rosenblatt, M (1975) : "A Quadratic Measure of Deviation of Two dimensional Density Estimates and a Test of Independence", Ann.Statist., 3, 1-14 [correction 10 (1982) 646].

Roussas, G. (1990) : "Asymptotic Normality of the Kernel Estimate Under Dependence Conditions : Application to Hazard Rate", J. Statist. Plann. Inference, 25, 81-104.

Silverman, B. (1986) : "Density Estimation for Statistics and Data Analysis", Chapman-Hall, London.

Takahata, H., and K., Yoshihara (1987) : "Central Limit Theorems for Integrated Square Error of Nonparametric Density Estimators Based on Absolutely Regular Random Sequences", Yokohama Math. J., 35, 95-111.

Yoshihara, K. (1976) : "Limiting Behavior of U-Statistics for Stationary, Absolutely Regular Processes", Z. Wahrsch. Verw. Gebiete, 35, 237-252.