Journal of Statistical Planning and Inference 141 (2011) 65–79

Contents lists available at ScienceDirect

Journal of Statistical Planning and Inference

journal homepage: www.elsevier.com/locate/jspi

Minimum distance lack-of-fit tests for fixed design ☆

Hira L. Koul

Department of Statistics and Probability, Michigan State University, East Lansing, MI 48824, USA

ARTICLE INFO

Article history:
Received 19 May 2009
Received in revised form
13 May 2010
Accepted 13 May 2010
Available online 27 May 2010

Keywords:
CLT for weighted U statistics
Nonparametric regression estimators

ABSTRACT

This paper discusses a class of tests of lack-of-fit of a parametric regression model when design is non-random and uniform on [0,1]. These tests are based on certain minimized distances between a nonparametric regression function estimator and the parametric model being fitted. We investigate asymptotic null distributions of the proposed tests, their consistency and asymptotic power against a large class of fixed and sequences of local nonparametric alternatives, respectively. The best fitted parameter estimate is seen to be $n^{1/2}$-consistent and asymptotically normal. A crucial result needed for proving these results is a central limit lemma for weighted degenerate U statistics where the weights are arrays of some non-random real numbers. This result is of an independent interest and an extension of a result of Hall for non-weighted degenerate U statistics.

1. Introduction

An important problem in statistics is to use a variable X to explain the response Y. This is often done in practice in terms of the regression function $\mu(x) := E(Y|X=x)$, assuming it exists. One often stipulates this function to be of a parametric form and then proceeds to make inference about the underlying parameters. It is then of interest to assess the accuracy of the assumed parametric form, i.e., to test for the lack-of-fit of the assumed parametric model for $\mu(x)$ based on the available data. Monograph of Hart (1997) contains numerous tests for this problem and provides a review through 1997.

Koul and Ni (2004) proposed a class of tests for fitting a parametric model to the regression function based on minimum squared deviations between a nonparametric estimator of $\mu(x)$ and the parametric model being fitted. They established asymptotic normality of a suitably standardized minimum distance test statistics and the minimum distance estimators. Koul and Song (2009) extended this methodology to Berkson measurement error models. In a finite sample comparison of these tests with some other existing tests, it was found that a member of this class preserves asymptotic level and has very high power against the chosen alternatives. Moreover, the best fitted parameter estimate has literally no empirical bias at the selected models. In both of these papers design is random and of dimension $p \geq 1$.

Here, we shall investigate their asymptotic behavior when design is non-random and uniform on [0,1]. More precisely, we observe $Y_{ni}, 1 \leq i \leq n$, from the model

$$Y_{ni} = \mu\left(\frac{i}{n}\right) + \varepsilon_{ni}, \quad 1 \leq i \leq n, \tag{1.1}$$

where ε_{ni}, $i \geq 1$, are independent error r.v.'s, $E\varepsilon_{ni}=0$, $E\varepsilon_{ni}^2 < \infty$, $1 \leq i \leq n$. Let $\mathcal{M} := \{m_\theta(x); x \in [0,1], \theta \in \Theta \subset \mathbb{R}^q\}$ be given family of parametric regression models, where q is a known positive integer. The problem of interest is to test for

$$\mathcal{H}: \mu(x) = m_{\theta_0}(x), \quad \forall 0 \leq x \leq 1, \text{ for some } \theta_0 \in \Theta, \text{ vs.}$$

$$\mathcal{H}_1 : \mathcal{H} \text{ is not true.}$$

☆ Supported in part by the NSF Grant DMS 0704130.
E-mail address: koul@stt.msu.edu

0378-3758/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.jspi.2010.05.019

To describe the proposed tests, let K be a density kernel function on $[-1,1]$, and $b\equiv b_n$ be a deterministic bandwidth sequence. Let G be a σ- finite measure on [0,1] and define

$$M_n(\theta) := \int \left[\frac{1}{nb}\sum_{j=1}^{n} K\left(\frac{nx-j}{nb}\right)(Y_j - m_\theta(j/n))\right]^2 dG(x),$$

$$\hat{\theta}_n := \underset{\theta\in\Theta}{\operatorname{argmin}} M_n(\theta). \tag{1.2}$$

A class of tests of $\mathcal{H}$, one for each G, is based on the minimized dispersion $M_n(\hat{\theta}_n)$.

To establish asymptotic normality of $M_n(\hat{\theta}_n)$ one needs a central limit theorem for weighted degenerate U-statistics where the weights are given by triangular arrays of real numbers and the kernels may depend on n. Hall (1984) gives such a result for non-weighted degenerate U statistics of multivariate i.i.d. random variables. This result was used in Koul and Ni (2004) to prove asymptotic normality of $M_n(\hat{\theta}_n)$ in the case of random design. Lemma 2.1 below gives an extension of this result to the weighted degenerate U statistics with the weights as mentioned above. This lemma is of independent interest. It is used here to prove asymptotic normality of a suitably standardized $M_n(\hat{\theta}_n)$ under $\mathcal{H}$ and under some alternatives.

Other authors that have investigated asymptotic distributions of weighted degenerate U statistics include O'Neil and Redner (1993), Major (1994), and Rifi and Utzet (2000). The weights and kernels in these papers are not allowed to depend on n and the weak limits are generally non-Gaussian. See Hsing and Wu (2004) and references therein for asymptotic normality of weighted non-degenerate U statistics of dependent observations, where again the weights and kernels are not allowed to depend on n.

For the sake of clarity of the exposition, in the next section we discuss asymptotic distributions of the above entities under null hypothesis when $m_\theta(x)$ is linear in θ. The case of more general m_θ will be discussed briefly in Section 3. Consistency of the proposed tests against a fixed alternative and asymptotic power against a sequence of local alternatives $\mu = m_{\theta_0} + (nb^{1/2})^{-1}\psi$ is discussed in Section 4, where ψ is a Lipschitz continuous function of order 1 on [0,1]. Among the class of functions G having density g on [0,1], the g that yields the maximum asymptotic power $1-\Phi(z_\alpha - a^{-1}\int\psi^2(x)\,dx)$ at the asymptotic level α against this sequence of alternatives is $g\equiv\psi^2$, where $a := 2(\int\{\int K(t)K(t+u)du\}^2\,dt)^{1/2}$, Φ is the d.f. of a standard normal r.v., and z_α is such that $\Phi(z_\alpha) = 1-\alpha$, $0\le\alpha\le 1$, see Theorem 4.2 and Remark 4.1 below.

In the sequel, all limits are taken as $n\to\infty$, unless specified otherwise, and for any two sequence of real numbers $a_n, b_n, a_n \sim b_n$, means that $a_n/b_n \to 1$. The convergence in distribution is denoted by $\to_D$ and $\mathcal{N}_q(a,B)$ denotes the q-variate normal distribution with mean vector a and covariance matrix B.

2. Main results when $\boldsymbol{m_\theta(x) = \theta'\ell(x)}$

In this section we shall discuss asymptotic distributions of the above minimum distance estimators and minimized distances under $\mathcal{H}$ when m_θ is linear in θ. Part of the reason for this is to keep the exposition relatively transparent. Let ℓ_j, $j=1,\dots,q$, be continuous G-square integrable functions on [0,1] and let $\ell(x) := (\ell_1(x), \ell_2(x), \dots, \ell_q(x))'$. Consider the problem of testing

$$\begin{aligned} &\tilde{\mathcal{H}} : \mu(x) = \theta_0'\ell(x), \quad \forall x\in[0,1], \text{ and some } \theta_0\in\Theta \text{ vs.} \\ &\tilde{\mathcal{H}}_1 : \ \tilde{\mathcal{H}} \text{ is not true.} \end{aligned} \tag{2.1}$$

To proceed further, let

$$\ell_n(x) := \frac{1}{nb}\sum_{j=1}^{n} K\left(\frac{nx-j}{nb}\right)\ell(j/n), \quad 0\le x\le 1,$$

$$\Sigma_n := \int_0^1 \ell_n(x)\ell_n(x)'\,dG(x), \quad \Sigma := \int_0^1 \ell(x)\ell(x)'\,dG(x).$$

We make the following assumptions about G, K, and the window width sequence b:

G has a continuous Lebesgue density g on [0,1], and for some $0<\varepsilon<1$,

$$g(x) > 0,\ x\in[\epsilon, 1-\epsilon], \text{ and } \int_\epsilon^{1-\epsilon} (a'\ell(x))^2\,dG(x) > 0, \quad \forall a\in\mathbb{R}^q. \tag{2.2}$$

$$\Sigma_n \text{ and } \Sigma \text{ are positive definite for all } n\ge q. \tag{2.3}$$

K is an even positive differentiable density on $(-1,1)$, vanishing outside $(-1,1)$, and has a bounded derivative. (2.4)

$$b\to 0, \quad nb^2\to\infty. \tag{2.5}$$

Note that (2.2)–(2.4), $b \to 0$ and continuity of ℓ imply

$$\int \|\ell_n(x) - \ell(x)\|^2 \, dG(x) \to 0, \quad \Sigma_n \to \Sigma. \tag{2.6}$$

Let θ_0 be as in $\tilde{\mathcal{H}}$, $\Delta := \theta - \theta_0$, $\varepsilon_{nj} := Y_j - \theta_0' \ell(j/n)$, and let

$$U_n(x,\theta) := \frac{1}{nb} \sum_{j=1}^{n} K\left(\frac{nx-j}{nb}\right)\left(Y_j - \theta' \ell\left(\frac{j}{n}\right)\right),$$

$$U_n(x) := U_n(x,\theta_0) = \frac{1}{nb} \sum_{j=1}^{n} K\left(\frac{nx-j}{nb}\right)\varepsilon_{nj}, \quad S_n := \int U_n(x)\ell_n(x)\, dG(x),$$

$$Z_n(x,\theta) := U_n(x,\theta) - U_n(x,\theta_0) = -\Delta' \ell_n(x).$$

Under $\tilde{\mathcal{H}}$, $M_n(\theta_0) = \int_0^1 U_n^2 \, dG$ and

$$M_n(\theta) = \int U_n^2(x,\theta)\, dG(x) = \int [U_n(x) - \Delta' \ell_n(x)]^2 \, dG(x) = \int U_n^2 \, dG - 2\Delta' S_n + \Delta' \Sigma_n \Delta.$$

Hence, in view of (2.3),

$$(\hat{\theta}_n - \theta_0) = \Sigma_n^{-1} S_n, \quad M_n(\hat{\theta}_n) = M_n(\theta_0) - S_n' \Sigma_n^{-1} S_n, \;\; \forall n \ge q. \tag{2.7}$$

Our aim is to obtain limiting distributions of $\hat{\theta}_n$ and $M_n(\hat{\theta}_n)$ under $\tilde{\mathcal{H}}$. The following proposition gives the limiting distribution of $\hat{\theta}_n$ under the two different sets of conditions on the errors.

Proposition 2.1. *Suppose $\tilde{\mathcal{H}}$ and assumptions (2.2)–(2.5) hold.*

(i) *If, in addition, $\varepsilon_{nj}, 1 \le j \le n$, are i.i.d. with mean zero and positive and finite variance σ^2, then $n^{1/2}(\hat{\theta}_n - \theta_0) \to_D \mathcal{N}_q(0, \sigma^2 \Sigma^{-1} \tilde{\Sigma} \Sigma^{-1})$, where $\tilde{\Sigma} := \int_0^1 \ell(u)\ell(u)' g^2(u)\, du$.*
(ii) *In addition, suppose for some continuous positive functions $\sigma^2(x)$ and $\mu_4(x)$ on [0,1], $\varepsilon_{nj}, 1 \le j \le n$, are heteroscedastic independent r.v.'s with finite 4th moments and $E\varepsilon_{nj}^2 = \sigma^2(j/n)$, $E\varepsilon_{nj}^4 = \mu_4(j/n)$, $1 \le j \le n$. Then, $n^{1/2}(\hat{\theta}_n - \theta_0) \to_D \mathcal{N}_q(0, \Sigma^{-1}\Sigma_1 \Sigma^{-1})$, where $\Sigma_1 := \int_0^1 \ell(x)\ell(x)' \sigma^2(x) g^2(x)\, dx$.*

Proof. For the time being assume ε_{nj} to be independent r.v.'s with $E\varepsilon_{nj} = 0$ and $E\varepsilon_{nj}^2 = \sigma^2(j/n)$, for some continuous positive function σ^2 on [0,1]. In view of (2.3), (2.6) and (2.7), it suffices to prove asymptotic normality of S_n. For an $a \in \mathbb{R}^q$, let

$$L_n(x) := a' \ell_n(x), \quad L(x) := a' \ell(x), \quad c_{nj} := \frac{1}{n^{1/2} b} \int K\left(\frac{nx-j}{nb}\right) L_n(x)\, dG(x), \;\; 1 \le j \le n.$$

Then,

$$n^{1/2} a' S_n = n^{1/2} \int U_n(x) L_n(x)\, dG(x) = \frac{1}{n^{1/2} b} \sum_{j=1}^{n} \int K\left(\frac{nx-j}{nb}\right) L_n(x)\, dG(x) \varepsilon_{nj} = \sum_{j=1}^{n} c_{nj}\varepsilon_{nj}.$$

Hence,

$$s_n^2 := \mathrm{Var}(n^{1/2} a' S_n) = \sum_{j=1}^{n} c_{nj}^2 \mathrm{Var}(\varepsilon_{nj}) = \frac{1}{nb^2} \sum_{j=1}^{n} \left(\int K\left(\frac{nx-j}{nb}\right) L_n(x)\, dG(x)\right)^2 \sigma^2\left(\frac{j}{n}\right)$$
$$= \frac{1}{b^2} \int_0^1 \left(\int K\left(\frac{x-u}{b}\right) L_n(x) g(x)\, dx\right)^2 \sigma^2(u)\, du + o(1) \to \int_0^1 (L(u)\sigma(u)g(u))^2 \, du = a' \Sigma_1 a.$$

In the above derivation we used the fact that $\int K(u)\, du = 1$, continuity of g and $\sigma^2(u)$, and (2.6) imply, uniformly in $0 \le u \le 1$,

$$b^{-1} \int K\left(\frac{x-u}{b}\right) L_n(x)\, dG(x) = \int K(z) L_n(bz+u) g(bz+u)\, dz \to \int K(z)\, dz L(u)\, g(u).$$

Hence,

$$s_n^2 \to a' \Sigma_1 a, \quad \mathrm{Cov}(n^{1/2} S_n) \to \Sigma_1. \tag{2.8}$$

Also, (2.2) and $\sigma(u)$ being continuous and positive on [0,1] imply $\kappa := \inf_{\varepsilon \le x \le 1-\varepsilon} g(x)\sigma^2(x) > 0$ and $a' \Sigma_1 a \ge \kappa \int_\varepsilon^{1-\varepsilon} (a' \ell(x))^2 \, dG(x) > 0$, for all $a \in \mathbb{R}^q$, i.e., Σ_1 is positive definite.

Now consider Case (i) where ε_{nj}'s are the i.i.d. r.v.'s with $\sigma^2(j/n) = \sigma^2$, a constant. Since $n^{1/2} a' S_n$ is a triangular array of independent centered r.v.'s we will apply Lindeberg–Feller CLT. Now, because K is a bounded kernel and because $\int |L_n|\, dG \to \int |L|\, dG < \infty$,

$$c_n := \max_{1 \le j \le n} |c_{nj}| \le C n^{-1/2} b^{-1} \int |L_n(x)| dG(x) \to 0 \quad \text{by (2.5)}. \tag{2.9}$$

Errors being i.i.d. and square integrable imply that $s_n^2=\sum_{j=1}^n c_{nj}^2\sigma^2$ and for any $\eta>0$,

$$s_n^{-2}\sum_{j=1}^n c_{nj}^2 E\varepsilon_{nj}^2 I(|c_{nj}\varepsilon_{nj}|>\eta s_n)\le \sigma^{-2}E\varepsilon^2\left(|\varepsilon|>\eta\frac{s_n}{c_n}\right)\to 0,$$

by (2.8) and (2.9), thereby verifying the L–F condition in this case.

Next, consider Case (ii) where errors are heteroscedastic having finite fourth moments. Argue as for (2.8), and use the continuity of μ_4 to obtain

$$\sum_{j=1}^n c_{nj}^2\mu_4(j/n)=\frac{1}{nb^2}\sum_{j=1}^n\left(\int K\left(\frac{nx-j}{nb}\right)L_n(x)\,dG(x)\right)^2\mu_4\left(\frac{j}{n}\right)\to\int_0^1(L(u)g(u))^2\mu_4(u)\,du<\infty.$$

This fact, (2.8) and (2.9) in turn imply

$$s_n^{-2}\sum_{j=1}^n E(c_{nj}\varepsilon_{nj})^2 I(|c_{nj}\varepsilon_{nj}|>\eta s_n)\le\frac{1}{\eta^2 s_n^4}\sum_{j=1}^n c_{nj}^4\mu_4(j/n)\le\frac{c_n}{\eta^2 s_n^4}\sum_{j=1}^n c_{nj}^2\mu_4(j/n)\to 0,$$

which again verifies the L–F condition.

Thus, by the Carmér–Wold devise, we obtain $n^{1/2}S_n\to_D\mathcal{N}_q(0,\sigma^2\tilde{\Sigma})$, in Case (i), and $n^{1/2}S_n\to_D\mathcal{N}_q(0,\Sigma_1)$, in Case (ii). Because $(\hat{\theta}_n-\theta_0)=\Sigma_n^{-1}S_n$, and because $\Sigma_n\to\Sigma$, we thus obtain, in view of (2.3),

$$n^{1/2}(\hat{\theta}_n-\theta_0)\to_D\mathcal{N}_q(0,\sigma^2\Sigma^{-1}\tilde{\Sigma}\Sigma^{-1})\quad\text{in case (i)},$$

$$n^{1/2}(\hat{\theta}_n-\theta_0)\to_D\mathcal{N}_q(0,\Sigma^{-1}\Sigma_1\Sigma^{-1})\quad\text{in case (ii)}.\tag{2.10}$$

This completes the proof of the proposition. We remark here that in Case (ii), the finite fourth moment assumption may be replaced by requiring only finite $(2+\delta)$th moment. □

Next, we turn to deriving the asymptotic distribution of the minimized distance $M_n(\hat{\theta}_n)$. From now onwards errors are assumed to be as in Case (i) of the previous proposition and we shall write ε_j for ε_{nj}. Let

$$C_n:=\frac{1}{(nb)^2}\sum_{j=1}^n\int K^2\left(\frac{nx-j}{nb}\right)dG(x)\varepsilon_j^2,\quad K_*(t):=\int K(t+u)K(u)\,du,$$

$$\hat{C}_n:=\frac{1}{(nb)^2}\sum_{j=1}^n\int K^2\left(\frac{nx-j}{nb}\right)dG(x)\hat{\varepsilon}_{nj}^2,\quad \hat{\varepsilon}_{nj}:=Y_j-\hat{\theta}_n'\ell(j/n),$$

$$\tau^2:=\int_0^1 g^2(v)\,dv\int K_*^2(t)\,dt,\quad \gamma^2:=4\sigma^4\tau^2.$$

We have

Proposition 2.2. *In addition to $\tilde{\mathcal{H}}$ and (2.2)–(2.5), assume $\varepsilon_j=Y_j-\theta_0'\ell(j/n)$, $1\le j\le n$, to be i.i.d. r.v.'s with mean zero and positive and finite variance σ^2 and having finite fourth moment. Then,*

$$nb^{1/2}(M_n(\hat{\theta}_n)-\hat{C}_n)\to_D\mathcal{N}_1(0,\gamma^2).\tag{2.11}$$

Proof. Because $M_n(\hat{\theta}_n)=M_n(\theta_0)-S_n'\Sigma_n^{-1}S_n$ and because

$$nb^{1/2}|S_n\Sigma_n^{-1}S_n|\le b^{1/2}\|n^{1/2}S_n\|^2\|\Sigma_n^{-1}\|=o_p(1),$$

it suffices to show that $nb^{1/2}(M_n(\theta_0)-\hat{C}_n)\to_D\mathcal{N}_1(0,\gamma^2)$. This in turn will follow from the following two facts. Under $\tilde{\mathcal{H}}$,

$$nb^{1/2}(\hat{C}_n-C_n)=o_p(1),\tag{2.12}$$

$$nb^{1/2}(M_n(\theta_0)-C_n)\to_D\mathcal{N}_1(0,\gamma^2).\tag{2.13}$$

First, consider (2.12). Let $\Delta_n=\hat{\theta}_n-\theta_0$ and write

$$\hat{C}_n=\frac{1}{(nb)^2}\sum_{j=1}^n\int K^2\left(\frac{nx-j}{nb}\right)dG(x)(\varepsilon_j-(\hat{\theta}-\theta_0)'\ell(j/n))^2=C_n-2\Delta_n'\frac{1}{(nb)^2}\sum_{j=1}^n\int K^2\left(\frac{nx-j}{nb}\right)dG(x)\varepsilon_j\ell(j/n)$$
$$+\Delta_n'\frac{1}{(nb)^2}\sum_{j=1}^n\int K^2\left(\frac{nx-j}{nb}\right)dG(x)\ell(j/n)\ell(j/n)'\Delta_n.$$

Therefore,

$$nb^{1/2}(\hat{C}_n-C_n)=-2n^{1/2}\Delta_n'B_{n1}+n\Delta_n'B_{n2}\Delta_n,$$

where

$$B_{n1} := \frac{1}{(nb)^{3/2}}\sum_{j=1}^{n}\int K^2\left(\frac{nx-j}{nb}\right)dG(x)\varepsilon_j\ell(j/n),$$

$$B_{n2} := \frac{b^{1/2}}{(nb)^2}\sum_{j=1}^{n}\int K^2\left(\frac{nx-j}{nb}\right)dG(x)\ell(j/n)\ell(j/n)'.$$

But, in view of (2.5),

$$B_{n2} = \frac{1}{nb^{3/2}}\int\int K^2\left(\frac{x-u}{b}\right)g(x)\,dx\ell(u)\ell(u)'\,du+\mathrm{o}(1) = \frac{1}{nb^{1/2}}\int K^2(v)\,dv\int \ell(u)\ell(u)'g(u)\,du+\mathrm{o}(1)\to 0.$$

Next, consider B_{n1}. We have $E(a^{'}B_{n1})$=0 and $\mathrm{Var}(a'B_{n1}) = \mathrm{O}((n^2b)^{-1})\to 0$, because

$$(n^2b)\mathrm{Var}(a'B_{n1}) = \frac{\sigma^2}{nb^2}\sum_{j=1}^{n}\left(\int K^2\left(\frac{nx-j}{nb}\right)dG(x)a'\ell(j/n)\right)^2\to\sigma^2\int\left(\int K^2(v)\,dv\,g(u)a'\ell(u)\right)^2du.$$

This, together with the fact $n^{1/2}\Delta_n = \mathrm{O_p}(1)$, completes the proof of (2.12).

We now turn to the proof of (2.13). Let

$$M_{n2} := \frac{1}{(nb)^2}\sum_{1\le i<j\le n}\int K\left(\frac{nx-i}{nb}\right)K\left(\frac{nx-j}{nb}\right)dG(x)\varepsilon_i\varepsilon_j.$$

Note that

$$M_n(\theta_0) = \int U_n^2\,dG = \frac{1}{(nb)^2}\sum_{j=1}^{n}\int K^2\left(\frac{nx-j}{nb}\right)dG(x)\varepsilon_j^2 + \frac{1}{(nb)^2}\sum_{j\ne i,j=1}^{n}\int \sum_{i=1}^{n} K\left(\frac{nx-i}{nb}\right)K\left(\frac{nx-j}{nb}\right)dG(x)\varepsilon_i\varepsilon_j = C_n+2M_{n2}.$$

Thus, to prove (2.13), it suffices to prove

$$nb^{1/2}M_{n2}\to_D\mathcal{N}_1(0,\sigma^4\tau^2). \tag{2.14}$$

Let

$$w_{n,ij} := \frac{1}{nb^{3/2}}\int K\left(\frac{nx-i}{nb}\right)K\left(\frac{nx-j}{nb}\right)dG(x).$$

Then we can rewrite

$$nb^{1/2}M_{n2} := \sum_{1\le i<j\le n} w_{n,ij}\varepsilon_i\varepsilon_j.$$

Let $\mathcal{F}_j := \sigma$–field$\{\varepsilon_i, i\le j\}$. Because $E(\varepsilon_i\varepsilon_j|\mathcal{F}_{j-1}) = \varepsilon_i E(\varepsilon_j) = 0$, $nb^{1/2}M_{n2}$ is a weighted degenerate U statistic where the weights depend on n,i,j. We thus need a CLT for these types of U statistics.

Let $H_n(x,y)$ be a sequence of measurable functions such that $H_n(x,y)$=$H_n(y,x)$, for all $x,y\in\mathbb{R}$, $n\ge 1$. Let $X_{ni}, 1\le i\le n$ be an array of p-variate i.i.d. r.v.'s such that

$$E(H_n(X_{n1},X_{n2})|X_{n1}) = 0,\quad \text{a.s.},\quad EH_n^2(X_{n1},X_{n2})<\infty,\ \forall n\ge 1.$$

Also, let $c_{n,ij}$ be arrays of real numbers such that $c_{n,ij} = c_{n,ji}$, for all $1\le i,j\le n$. We shall provide further sufficient conditions on these entities that ensure asymptotic normality of the more general degenerate weighted U statistic

$$U := \sum_{1\le i<j\le n} c_{n,ij}H_n(X_{ni},X_{nj}).$$

Let $v_n^2 = \mathrm{Var}(U)$ and $G_n(x,y) := E\{H_n(X_{n1},x)H_n(X_{n1},y)\}$. We have the following:

Lemma 2.1. *In addition to the above assumptions suppose the following holds*:

$$\frac{\sum_{j=2}^{n}(\sum_{i=1}^{j-1}c_{n,ij}^2)^2 EH_n^4(X_{n1},X_{n2})}{(\sum_{j=2}^{n}\sum_{i=1}^{j-1}c_{n,ij}^2)^2(EH_n^2(X_{n1},X_{n2}))^2}\to 0, \tag{2.15}$$

$$\frac{\sum_{j_1=2}^{n}\sum_{j_2=2}^{n}(\sum_{i=1}^{j_1-1}c_{n,ij_1}c_{n,ij_2})^2 EG_n^2(X_{n1},X_{n2})}{(\sum_{j=2}^{n}\sum_{i=1}^{j-1}c_{n,ij}^2)^2(EH_n^2(X_{n1},X_{n2}))^2}\to 0, \tag{2.16}$$

$$\frac{\sum_{j_1=2}^{n}\sum_{j_2=2}^{n}\sum_{i=1}^{j_1-1} c_{n,ij_1}^2 c_{n,ij_2}^2 EH_n^4(X_{n1},X_{n2})}{(\sum_{j=2}^{n}\sum_{i=1}^{j-1} c_{n,ij}^2)^2 (EH_n^2(X_{n1},X_{n2}))^2} \to 0.$$

Then, $v_n^{-1}U \to_D \mathcal{N}_1(0,1)$.

Proof. Hall (1984) proves this result for the special case where $c_{n,ij} \equiv 1$, and $X_{ni} \equiv X_i$, for all $1 \le i,j \le n$. We adopt his proof to prove the above result for this more general weighted degenerate U statistic. Note that if $c_{n,ij} \equiv 1$ and $X_{ni} \equiv X_i$, then assumptions (2.15) and (2.16) are equivalent to (2.1) in Hall (1984).

Let $\mathcal{G}_{n,j} := \sigma-$field$\{X_{ni}, 1 \le i \le j\}$, and

$$Y_{nj} := \sum_{i=1}^{j-1} c_{n,ij} H_n(X_{ni},X_{nj}), \quad S_k := \sum_{j=2}^{k} Y_{nj}, \; 2 \le k \le n.$$

Then, $(Y_{nj},\mathcal{G}_{n,j-1})$ are martingale difference arrays and S_n=U. Let

$$V_n^2 := \sum_{j=2}^{n} E(Y_{nj}^2 | X_{ni}, 1 \le i \le j-1).$$

Note that $v_n^2 = EV_n^2 = \sum_{j=2}^{n} EY_{nj}^2$.

As in Hall (1984), we shall use the martingale CLT, Hall and Heyde (1980, Corollary 3.1, p. 58). Accordingly we must verify the following two conditions:

$$v_n^{-2}\sum_{j=2}^{n} E(Y_{nj}^2 I(|Y_{nj}| > \eta v_n)) \to 0, \quad \forall \eta > 0, \tag{2.17}$$

$$v_n^{-2} V_n^2 \to_p 1. \tag{2.18}$$

Because of the assumed independence of X_{nj}'s and H_n being conditionally centered,

$$EY_{nj}^2 = \sum_{i=1}^{j-1}\sum_{k=1}^{j-1} c_{n,ij}c_{n,kj} EH_n(X_{nj},X_{ni})H_n(X_{nj},X_{nk}) = \sum_{i=1}^{j-1} c_{n,ij}^2 EH_n^2(X_{nj},X_{ni}) = \sum_{i=1}^{j-1} c_{n,ij}^2 EH_n^2(X_{n1},X_{n2}), \quad 2 \le j \le n, \tag{2.19}$$

$$v_n^2 = \sum_{j=2}^{n}\sum_{i=1}^{j-1} c_{n,ij}^2 EH_n^2(X_{n1},X_{n2}).$$

Also,

$$EY_{nj}^4 = \sum_{i=1}^{j-1} c_{n,ij}^4 EH_n^4(X_{nj},X_{ni}) + 3\sum_{i=1}^{j-1}\sum_{k=1,k\ne i}^{j-1} c_{n,ij}^2 c_{n,kj}^2 EH_n^2(X_{nj},X_{ni})H_n^2(X_{nj},X_{nk}) \le 3\left(\sum_{i=1}^{j-1} c_{n,ij}^2\right)^2 EH_n^4(X_{n1},X_{n2}).$$

Hence, by (2.15),

$$v_n^{-4}\sum_{j=2}^{n} EY_{nj}^4 \le 3\frac{\sum_{j=2}^{n}(\sum_{i=1}^{j-1} c_{n,ij})^2 EH_n^4(X_{n1},X_{n2})}{(\sum_{j=2}^{n}\sum_{i=1}^{j-1} c_{n,ij}^2)^2 [EH_n^2(X_{n1},X_{n2})]^2} \to 0.$$

From this (2.17) follows in a routine fashion.

To prove (2.18), we shall show that $v_n^{-4}E(V_n^2 - v_n^2)^2 \to 0$. Towards this goal let $\nu_{nj} := E\{Y_{nj}^2 | X_{ni}, i<j\}$ and note that $V_n^2 = \sum_{j=2}^{n} \nu_{nj}$, $v_n^2 = EV_n^2$, and

$$E(V_n^2 - v_n^2)^2 = EV_n^4 - v_n^4, \quad EV_n^4 = \sum_{j=2}^{n} E\nu_{nj}^2 + 2\sum_{2 \le j_1 < j_2 \le n} E(\nu_{nj}\nu_{nk}).$$

To proceed further, observe that

$$\nu_{nj} = \sum_{i=1}^{j-1}\sum_{k=1}^{j-1} c_{n,ij}c_{n,kj} E\{H_n(X_{nj},X_{ni})H_n(X_{nj},X_{nk}) | X_{nm}, m \le j-1\} = \sum_{i=1}^{j-1}\sum_{k=1}^{j-1} c_{n,ij}c_{n,kj} G_n(X_{ni},X_{nk})$$
$$= \sum_{i=1}^{j-1} c_{n,ij}^2 G_n(X_{ni},X_{ni}) + 2\sum_{1 \le i < k \le j-1} c_{n,ij}c_{n,kj} G_n(X_{ni},X_{nk}).$$

Using the fact $EG_n(X_{n1},X_{n2})$=0 and the assumed independence, one obtains, for any $i \le k, j \le m$,

$$\begin{aligned} E(G_n(X_{ni},X_{nk})G_n(X_{nj},X_{nm})) &= EG_n^2(X_{n1},X_{n1}), & i=k=j=m,\\ &= [EG_n(X_{n1},X_{n1})]^2, & i=k\ne j=m,\\ &= EG_n^2(X_{n1},X_{n2}), & i=j, k=m, i<k,\\ &= 0, & \text{otherwise}. \end{aligned}$$

Therefore, for $j_1 \le j_2$,

$$E(v_{nj_1} v_{nj_2}) = \sum_{i=1}^{j_1-1}\sum_{k=1}^{j_2-1} c_{n,ij_1}^2 c_{n,kj_2}^2 [EG_n(X_{n1},X_{n1})]^2 + \sum_{i=1}^{j_1-1} c_{n,ij_1}^2 c_{n,ij_2}^2 \mathrm{Var}\{G_n(X_{n1},X_{n1})\} + 4 \sum_{1 \le i < k \le j_1-1} c_{n,ij_1}^2 c_{n,kj_1}^2 EG_n^2(X_{n1},X_{n2}).$$

In particular,

$$Ev_{nj}^2 = \left(\sum_{i=1}^{j-1} c_{n,ij}^2\right)^2 [EG_n(X_{n1},X_{n1})]^2 + \sum_{i=1}^{j-1} c_{n,ij}^4 \mathrm{Var}\{G_n(X_{n1},X_{n1})\} + 4 \sum_{1 \le i < k \le j-1} c_{n,ij}^2 c_{n,kj}^2 EG_n^2(X_{n1},X_{n2}), \quad 2 \le j \le n.$$

From (2.19), and because $EH_n^2(X_{n1},X_{n2}) = EG_n(X_{n1},X_{n1})$, we also have

$$v_n^4 = \left(\sum_{j=2}^{n}\sum_{i=2}^{j-1} c_{n,ij}^2\right)^2 [EH_n^2(X_{n1},X_{n2})]^2 = \left[\sum_{j=2}^{n}\left(\sum_{i=2}^{j-1} c_{n,ij}^2\right)^2 + 2 \sum_{2 \le j_1 < j_2 \le n}\sum_{i=2}^{j_1-1}\sum_{\ell=2}^{j_2-1} c_{n,ij_1}^2 c_{n,\ell j_1}^2\right] [EG_n(X_{n1},X_{n1})]^2.$$

Hence,

$$\begin{aligned}
EV_n^4 &= \sum_{j=2}^{n} Ev_{nj}^2 + 2 \sum_{2 \le j_1 < j_2 \le n} E(v_{nj_1} v_{nj_2}) \\
&= \sum_{j=2}^{n}\Bigg[\left(\sum_{i=1}^{j-1} c_{n,ij}^2\right)^2 [EG_n(X_{n1},X_{n1})]^2 + \sum_{i=1}^{j-1} c_{n,ij}^4 \mathrm{Var}\{G_n(X_{n1},X_{n1})\} \\
&\qquad + 4 \sum_{1 \le i < k \le j-1} c_{n,ij}^2 c_{n,kj}^2 EG_n^2(X_{n1},X_{n2})\Bigg] \\
&\quad + 2 \sum_{2 \le j_1 < j_2 \le n}\Bigg[\sum_{i=1}^{j_1-1}\sum_{k=1}^{j_2-1} c_{n,ij_1}^2 c_{n,kj_2}^2 [EG_n(X_{n1},X_{n1})]^2 \\
&\qquad + \sum_{i=1}^{j_1-1} c_{n,ij_1}^2 c_{n,ij_2}^2 \mathrm{Var}\{G_n(X_{n1},X_{n1})\} \\
&\qquad + 4 \sum_{1 \le i < k \le j_1-1} c_{n,ij_1} c_{n,ij_2} c_{n,kj_1} c_{n,kj_2} EG_n^2(X_{n1},X_{n2})\Bigg],
\end{aligned}$$

$$\begin{aligned}
EV_n^4 - v_n^4 &= \sum_{j=2}^{n}\left[\sum_{i=1}^{j-1} c_{n,ij}^4 \mathrm{Var}\{G_n(X_{n1},X_{n1})\} + 4 \sum_{1 \le i < k \le j-1} c_{n,ij}^2 c_{n,kj}^2 EG_n^2(X_{n1},X_{n2})\right] \\
&\quad + 2 \sum_{2 \le j_1 < j_2 \le n}\Bigg[\sum_{i=1}^{j_1-1} c_{n,ij_1}^2 c_{n,ij_2}^2 \mathrm{Var}\{G_n(X_{n1},X_{n1})\} \\
&\qquad + 4 \sum_{1 \le i < k \le j_1-1} c_{n,ij_1} c_{n,ij_2} c_{n,kj_1} c_{n,kj_2} EG_n^2(X_{n1},X_{n2})\Bigg], \\
&= \left[\sum_{j=2}^{n}\sum_{i=1}^{j-1} c_{n,ij}^4 + 2 \sum_{2 \le j_1 < j_2 \le n}\sum_{i=1}^{j_1-1} c_{n,ij_1}^2 c_{n,ij_2}^2\right] \mathrm{Var}\{G_n(X_{n1},X_{n1})\} \\
&\quad + 2\Bigg[2 \sum_{j=2}^{n}\sum_{1 \le i < k \le j-1} c_{n,ij}^2 c_{n,kj}^2. \\
&\qquad + 4 \sum_{2 \le j_1 < j_2 \le n}\sum_{1 \le i < k \le j_1-1} c_{n,ij_1} c_{n,ij_2} c_{n,kj_1} c_{n,kj_2}\Bigg] EG_n^2(X_{n1},X_{n2}) \\
&= 2 \sum_{j_1=2}^{n}\sum_{j_2=2}^{n}\left[\left(\sum_{i=1}^{j_1-1} c_{n,ij_1} c_{n,ij_2}\right)^2 - \sum_{i=1}^{j_1-1} c_{n,ij_1}^2 c_{n,ij_2}^2\right] EG_n^2(X_{n1},X_{n2}) \\
&\quad + \sum_{j_1=2}^{n}\sum_{j_2=2}^{n}\sum_{i=1}^{j_1-1} c_{n,ij_1}^2 c_{n,ij_2}^2 \mathrm{Var}\{G_n(X_{n1},X_{n1})\} \\
&\le 2 \sum_{j_1=2}^{n}\sum_{j_2=2}^{n}\left(\sum_{i=1}^{j_1-1} c_{n,ij_1} c_{n,ij_2}\right)^2 EG_n^2(X_{n1},X_{n2}) \\
&\quad + \sum_{j_1=2}^{n}\sum_{j_2=2}^{n}\sum_{i=1}^{j_1-1} c_{n,ij_1}^2 c_{n,ij_2}^2 EH_n^4(X_{n1},X_{n2})\}.
\end{aligned}$$

In the last but one inequality above, we used the fact

$$\mathrm{Var}(G_n(X_{n1},X_{n1})) \le EG_n^2(X_{n1},X_{n1}) = E[E^2\{H_n^2(X_{n1},X_{n2})|X_{n1}\}] \le EH_n^4(X_{n1},X_{n2}).$$

Hence, (2.18) follows from the condition (2.16). This completes the proof of the lemma. □

Now consider the case where $EH_n^2(X_{n1},X_{n2})$, $EH_n^4(X_{n1},X_{n2})$ and $EG_n^2(X_{n1},X_{n2})$ do not depend on n. Then, (2.15)–(2.16) are implied by

$$\frac{\sum_{j=2}^n (\sum_{i=1}^{j-1} c_{n,ij}^2)^2}{(\sum_{j=2}^n \sum_{i=1}^{j-1} c_{n,ij}^2)^2} \to 0, \tag{2.20}$$

$$\frac{\sum_{j_1=2}^n \sum_{j_2=2}^n (\sum_{i=1}^{j_1-1} c_{n,ij_1} c_{n,ij_2})^2}{(\sum_{j=2}^n \sum_{i=1}^{j-1} c_{n,ij}^2)^2} \to 0, \tag{2.21}$$

$$\frac{\sum_{j_1=2}^n \sum_{j_2=2}^n \sum_{i=1}^{j_1-1} c_{n,ij_1}^2 c_{n,ij_2}^2}{(\sum_{j=2}^n \sum_{i=1}^{j-1} c_{n,ij}^2)^2} \to 0. \tag{2.22}$$

We shall now show that these conditions are satisfied by

$$c_{n,ij} = w_{n,ij} = \frac{1}{nb^{3/2}} \int K\left(\frac{nx-i}{nb}\right) K\left(\frac{nx-j}{nb}\right) g(x)\,dx.$$

We have

$$\begin{aligned}
\sum_{j=2}^n \sum_{i=1}^{j-1} w_{n,ij}^2 &= \sum_{j=2}^n \sum_{i=1}^{j-1} \left\{\frac{1}{nb^{3/2}} \int K\left(\frac{nx-i}{nb}\right) K\left(\frac{nx-j}{nb}\right) dG(x)\right\}^2 \sim \frac{1}{b^3}\int_0^1 \int_0^v \left\{\int K\left(\frac{x-u}{b}\right) K\left(\frac{x-v}{b}\right) g(x)\,dx\right\}^2 du\,dv \\
&\sim \frac{1}{b}\int_0^1 \int_0^v \left\{\int K\left(z+\frac{v-u}{b}\right) K(x) g(bz+v)\,dz\right\}^2 du\,dv \sim \int_0^1 g^2(v)\,dv \int \left\{\int K(z+t)K(x)\,dz\right\}^2 dt.
\end{aligned} \tag{2.23}$$

$$\begin{aligned}
\sum_{j=2}^n \left(\sum_{i=1}^{j-1} w_{n,ij}^2\right)^2 &= \frac{1}{n^3 b^6} \sum_{j=2}^n \left(\sum_{i=1}^{j-1} \left\{\int K\left(\frac{nx-i}{nb}\right) K\left(\frac{nx-j}{nb}\right) dG(x)\right\}^2\right)^2 \\
&\sim \frac{1}{nb^6}\int_0^1 \left(\int_0^v \left\{\int K\left(\frac{x-u}{b}\right) K\left(\frac{x-v}{b}\right) g(x)\,dx\right\}^2 du\right)^2 dv \\
&\sim \frac{1}{nb^2}\int_0^1 \left(\int_0^v \left\{\int K\left(z+\frac{v-u}{b}\right) K(x)) g(bz+v)\,dz\right\}^2 du\right)^2 dv \\
&\sim n^{-1}\left(\int \left\{\int K(z+t)K(x)\,dz\right\}^2 dt\right)^2 \int_0^1 g^4(v)\,dv.
\end{aligned}$$

$$\begin{aligned}
\sum_{j_1=2}^n \sum_{j_2=2}^n \left(\sum_{i=1}^{j_1-1} w_{n,ij_1} w_{n,ij_2}\right)^2 &= \frac{1}{n^2 b^6} \sum_{j_1=2}^n \sum_{j_2=2}^n \left(\frac{1}{n}\sum_{i=1}^{j_1-1} \int K\left(\frac{nx-i}{nb}\right) K\left(\frac{nx-j_1}{nb}\right) g(x)\,dx\right. \\
&\qquad \left. \times \int K\left(\frac{ny-i}{nb}\right) K\left(\frac{ny-j_2}{nb}\right) g(y)\,dy\right)^2 \\
&\sim b^{-6}\int_0^1 \int_0^1 \left(\int_0^v \int K\left(\frac{x-u}{b}\right) K\left(\frac{x-v}{b}\right) g(x)\,dx\right. \\
&\qquad \left. \times \int K\left(\frac{y-u}{b}\right) K\left(\frac{y-z}{b}\right) g(y)\,dy\,du\right)^2 dv\,dz \\
&\sim b^{-2}\int_0^1 \int_0^1 \left(\int_0^v \int K\left(w+\frac{v-u}{b}\right) K(w) g(bw+v)\,dw\right. \\
&\qquad \left. \times \int K\left(t+\frac{t-u}{b}\right) K(t) g(bt+z)\,dt\,du\right)^2 dv\,dz \\
&\sim \int_0^1 \int_0^1 \left(\int \int K(w+\alpha)K(w)\,dw\right. \\
&\qquad \left. \times \int K\left(t+\alpha+\frac{z-v}{b}\right) K(t)\,dt\,d\alpha\right)^2 g^2(v) g^2(x)\,dv\,dz \\
&\sim b \int g^4(v)\,dv \int \left(\int K_*(\alpha) K_*(\alpha+\beta)\,d\alpha\right)^2 d\beta.
\end{aligned}$$

Similarly one sees

$$\sum_{j_1=2}^{n}\sum_{j_2=2}^{n}\sum_{i=1}^{j_1-1} w_{n,ij_1}^2 w_{n,ij_2}^2 = \frac{1}{n^4b^6}\sum_{j_1=2}^{n}\sum_{j_2=2}^{n}\sum_{i=1}^{j_1-1}\left(\int K\left(\frac{nx-i}{nb}\right)K\left(\frac{nx-j_1}{nb}\right)g(x)\,dx \times \int K\left(\frac{ny-i}{nb}\right)K\left(\frac{ny-j_2}{nb}\right)g(y)\,dy\right)^2 \sim n^{-1}\int g^4(v)\,dv\int\int (K_*(s)K_*(t))^2\,ds dt.$$

These approximations now clearly show that conditions (2.20)–(2.22) are satisfied by $w_{n,ij}$, and hence the lemma applies with $X_{ni}\equiv\varepsilon_i$, $H_n(x,y)=xy$. Because of (2.23), here

$$v_n^2 = \sum_{j=2}^{n}\sum_{i=1}^{j-1} w_{n,ij}^2 E(\varepsilon_1\varepsilon_2)^2 \to \sigma^4\int g^2(v)\,dv\int K_*^2(t)\,dt.$$

Therefore, we have established (2.13), and hence also (2.11). □

Let z_α be $100(1-\alpha)$ th percentile of the standard normal distribution, $0<\alpha<1$, and let $\hat\sigma_n^2 := n^{-1}\sum_{j=1}^n \hat\varepsilon_{nj}^2$, $\hat\gamma_n^2 := 4\hat\sigma_n^4\tau^2$. Then the minimized distance test that rejects $\tilde{\mathcal{H}}$ in favor of $\tilde{\mathcal{H}}_1$ whenever $|\hat\gamma_n^{-1} nb^{1/2}(M_n(\hat\theta_n)-\hat C_n)| > z_{\alpha/2}$ will have asymptotic size α.

3. General $\mathcal{M}$

We shall now give a set of sufficient conditions on the model $\mathcal{M}$ under which the analogs of the results of the previous section will continue to hold. These are similar yet simpler to those given in Koul–Ni for the random design case where there was an extra complication because the design was random with an unknown density and errors were allowed to be heteroscedastic. Consider the following assumptions.

(e) $E\varepsilon^4 < \infty$.

(m1) For each θ, $m_\theta(x)$ is continuous in x w.r.t. integrating measure G.

(m2) The parametric family of models $m_\theta(x)$ is identifiable w.r.t. θ, i.e., if $m_{\theta_1}(x)=m_{\theta_2}(x)$, for almost all x (G), then $\theta_1=\theta_2$.

(m3) For some positive continuous function h on [0,1] and for some $\beta>0$,

$$|m_{\theta_2}(x)-m_{\theta_1}(x)| \le \|\theta_2-\theta_1\|^\beta h(x), \quad \forall\theta_2,\ \theta_1\in\Theta,\ x\in[0,1].$$

(m4) For every x, $m_\theta(x)$ is differentiable in θ in a neighborhood of θ_0 with the vector of derivatives $\dot m_\theta(x)$, such that for every $k<\infty$,

$$\sup_{1\le i\le n,(nb)^{1/2}\|\theta-\theta_0\|\le k} \frac{\left|m_\theta\left(\frac{i}{n}\right)-m_{\theta_0}\left(\frac{i}{n}\right)-(\theta-\theta_0)'\dot m_{\theta_0}\left(\frac{i}{n}\right)\right|}{\|\theta-\theta_0\|} \to 0.$$

(m5) The vector function $x\mapsto\dot m_{\theta_0}(x)$ is continuous in $x\in[0,1]$ and for every $0<k<\infty$,

$$\max_{1\le i\le n,(nb)^{1/2}\|\theta-\theta_0\|\le k} b^{-1/2}\left\|\dot m_\theta\left(\frac{i}{n}\right)-\dot m_{\theta_0}\left(\frac{i}{n}\right)\right\| \to 0.$$

(m6) $\Sigma_0 := \int_0^1 \dot m_{\theta_0}(u)\dot m_{\theta_0}(u)'\,dG(u)$ is positive definite.

Under (2.2)–(2.5) and the above assumptions and by adapting arguments of Koul and Ni (2004) and using Lemma 2.1 where Hall (1984) was used one can prove the following results. First, one establishes $\hat\theta_n$ of (1.2) is consistent for θ_0 under $\mathcal{H}$, and

$$n^{1/2}(\hat\theta_n-\theta_0)\to_D \mathcal{N}_q(0,\Sigma_0^{-1}\tilde\Sigma\Sigma_0^{-1}), \quad \tilde\Sigma := \sigma^2\int_0^1 \dot m_{\theta_0}(u)\dot m_{\theta_0}(u)'g^2(u)\,du.$$

Next, let $\hat\varepsilon_{nj} := Y_j - m_{\hat\theta_n}(j/n)$,

$$\hat C_n := (nb)^{-2}\sum_{j=1}^n\int_0^1 K^2\left(\frac{nx-j}{nb}\right)dG(x)\hat\varepsilon_{nj}^2, \quad \hat\sigma_n^2 := n^{-1}\sum_{j=1}^n\hat\varepsilon_{nj}^2, \quad \hat\gamma_n^2 := 4\hat\sigma_n^4\tau^2,$$

$$\hat{\mathcal{D}}_n := \hat\gamma_n^{-1} nb^{1/2}(M_n(\hat\theta_n)-\hat C_n).$$

Then, $\hat{\mathcal{D}}_n\to_D\mathcal{N}_1(0,1)$, and the test that rejects $\mathcal{H}$, in favor of $\mathcal{H}_1$, whenever $|\hat{\mathcal{D}}_n|\ge z_{\alpha/2}$, has asymptotic size α.

4. Consistency and asymptotic power

In this section we shall discuss consistency of $\hat{\theta}_n$ for θ_0 under $\mathcal{H}$ and of the above test based on $\hat{\mathcal{D}}_n$ against a fixed alternative. We shall also derive asymptotic power of this test against sequences of local nonparametric alternatives.

Let $L_2(G)$ denote the class of real valued square integrable functions on [0,1] with respect to G, $\rho(v_1,v_2) := \int [v_1-v_2]^2\, dG, v_1,v_2 \in L_2(G)$, and

$$T(v) := \underset{\theta\in\Theta}{\operatorname{argmin}}\, \rho(v,m_\theta), \quad v \in L_2(G). \tag{4.1}$$

Under assumption (m2), $T(m_{\theta_0}) = \theta_0$, for any value of θ_0.

We shall first prove $\hat{\theta}_n \to_p T(m)$ for every continuous regression function m. This in turn is used to prove consistency of the above test based on $\hat{D}_n$. We also establish asymptotic normality of $n^{1/2}(\hat{\theta}_n-\theta_0)$ and $\hat{\mathcal{D}}_n$ under the local alternatives $\mathcal{H}_{1n} : \mu = m_{\theta_0} + \psi/\sqrt{nb^{1/2}}$, where ψ is a real valued Lipschitz continuous function of order 1 on [0,1] such that $\int \dot{m}_{\theta_0}\psi\, dG = 0$.

Basic ideas of the proofs of Lemma 4.2 and Theorems 4.1 and 4.2 below are as in Koul and Ni (2004) and Koul and Song (2009). However, here the details are inherently different and relatively simpler.

Consistency of $\hat{\theta}_n$: Proving consistency of $\hat{\theta}_n$ is facilitated by the following continuity lemma. Its proof is similar to that of Theorem 1 in Beran (1977).

Lemma 4.1. *Under* (m3), *the following hold.*

(a) *$T(v)$ always exists, for all $v \in L_2(G)$.*
(b) *If $T(v)$ is unique, then T is continuous at v in the sense that for any sequence of $\{v_n\} \in L_2(G)$ converging to v in $L_2(G)$, $T(v_n) \to T(v)$, i.e. $\rho(v_n,v) \to 0$ implies $T(v_n) \to T(v)$.*
(c) *In addition, if* (m2) *holds, then $T(m_\theta) = \theta$, uniquely for $\forall \theta \in \Theta$.*

Now, let $K_{bj}(x) := K((nx-j)/nb)/b$, and

$$\hat{\mu}_n(x) := \frac{1}{n}\sum_{j=1}^n K_{bj}(x)Y_j, \quad m_{n\theta}(x) := \frac{1}{n}\sum_{j=1}^n K_{bj}(x)m_\theta\left(\frac{j}{n}\right), \; x \in [0,1], \; \vartheta_n := T(\hat{\mu}_n).$$

Note that $M_n(\theta) = \rho(\hat{\mu}_n, m_{n\theta})$ and $\hat{\theta}_n = \operatorname{argmin}_\theta \rho(\hat{\mu}_n, m_{n\theta})$. A consequence of the above lemma is the following:

Lemma 4.2. *Suppose* (2.2), (2.4), (2.5) *and* (m3) *hold. Furthermore, suppose m is a given continuous regression function such that the errors $Y_i-m(i/n)$, $1 \le i \le n$, are homoscedastic having finite and positive variance σ^2, and such that $T(m)$ is unique. Then,*

$$\text{(a) } \vartheta_n \to_p T(m). \quad \text{(b) } \hat{\theta}_n \to_p T(m). \tag{4.2}$$

Proof. To prove (a), in view of part (b) of Lemma 4.1, it suffices to prove

$$\rho(\hat{\mu}_n, m) = o_p(1). \tag{4.3}$$

Let, $\xi_i := Y_i - m(i/n)$, $1 \le i \le n$,

$$V_n(x) := n^{-1}\sum_{i=1}^n K_{bi}(x)\xi_i, \quad \overline{m}_n(x) := n^{-1}\sum_{i=1}^n K_{bi}(x)m\left(\frac{i}{n}\right), \; x \in [0,1]. \tag{4.4}$$

To prove (4.3), write

$$\rho(\hat{\mu}_n, m) = \int \left[\frac{1}{n}\sum_{j=1}^n K_{bj}(x)Y_j - m(x)\right]^2 dG(x) = \int [V_n(x) + \overline{m}_n(x) - m(x)]^2\, dG(x) \le 2\int V_n^2\, dG + 2\rho(\overline{m}_n, m).$$

By Fubini's Theorem, continuity of g and because $E\xi_i \equiv 0$,

$$E\int V_n^2\, dG = \frac{1}{bn^2}\int K(u)\sum_{i=1}^n g\left(bu + \frac{i}{n}\right) du = O((nb)^{-1}),$$

$$\int V_n^2\, dG = O_p((nb)^{-1}) = o_p(1) \quad \text{because } nb \to \infty. \tag{4.5}$$

Continuity of m and (2.2) imply $\rho(\overline{m}_n, m) \to 0$. This completes the proof of part (a).

To prove part (b), it suffices to show that

$$\sup_{\theta\in\Theta} |M_n(\theta) - \rho(m, m_\theta)| = o_p(1). \tag{4.6}$$

For, assume this to be true. By (m3) and the triangle inequality

$$|\rho(m, m_{\theta_2}) - \rho(m, m_{\theta_1})| \le |\rho^{1/2}(m, m_{\theta_2}) - \rho^{1/2}(m, m_{\theta_1})|^2 \le C\|\theta_1 - \theta_2\|^{2\beta}, \quad \forall \theta_1, \theta_2 \in \Theta.$$

This together with (4.6) imply

$$\lim_{\delta\to 0}\limsup_n P\left(\sup_{\|\theta_1-\theta_2\|<\delta}|M_n(\theta_1)-M_n(\theta_2)|>\varepsilon\right)=0,\quad \forall\varepsilon>0. \tag{4.7}$$

These two facts in turn imply $\hat{\theta}_n\to_p T(m)$. For, if $\hat{\theta}_n\nrightarrow_p T(m)$, then, by the compactness of Θ, there is a subsequence $\{\hat{\theta}_{n_k}\}$ of $\{\hat{\theta}_n\}$ and a $\vartheta\neq T(m)$ such that $\hat{\theta}_{n_k}\to_p\vartheta$. Because $M_{n_k}(\hat{\theta}_{n_k})\le M_{n_k}(T(m))\le|M_{n_k}(T(m))-\rho(m,m_{T(m)})|+\rho(m,m_{T(m)})$,

$$\rho(m,m_\vartheta)\le|\rho(m,m_\vartheta)-M_{n_k}(\vartheta)|+|M_{n_k}(\vartheta)-M_{n_k}(\hat{\theta}_{n_k})|+M_{n_k}(\hat{\theta}_{n_k})\le\rho(m,m_{T(m)})+2\sup_\theta|M_{n_k}(\theta)-\rho(m,m_\theta)|+|M_{n_k}(\vartheta)-M_{n_k}(\hat{\theta}_{n_k})|.$$

By (4.6) and (4.7), the last two summands in the above bound tend to zero, in probability, so that $\rho(m,m_\vartheta)\le\rho(m,m_{T(m)})$ eventually, with arbitrarily large probability. In view of the uniqueness of $T(m)$, this is a contradiction unless $\vartheta=T(m)$.

To prove (4.6), recall $M_n(\theta)=\rho(\hat{\mu}_n,m_{n\theta})$. Use the factorization $(a^2-b^2)=(a-b)(a+b)$ and the Cauchy–Schwarz (C–S) inequality to obtain that $|\rho(\hat{\mu}_n,m_{n\theta})-\rho(m,m_\theta)|$ is bounded above by the product $q_{n1}^{1/2}(\theta)q_{n2}^{1/2}(\theta)$, where

$$q_{n1}(\theta):=\int([\hat{\mu}_n(x)-m(x)]-[m_{n\theta}(x)-m_\theta(x)])^2\,dG(x),$$

$$q_{n2}(\theta):=\int([\hat{\mu}_n(x)+m(x)]-[m_{n\theta}(x)+m_\theta(x)])^2\,dG(x).$$

But $q_{n1}(\theta)\le 2(\rho(\hat{\mu}_n,m)+\rho(m_{n\theta},m_\theta))$. By (4.3), the first term in this bound is $o_p(1)$. We need to show that the second term tends to zero uniformly in $\theta\in\Theta$. Because, for each θ, $m_\theta(x)$ is continuous in x, we already have $\rho(m_{n\theta},m_\theta)\to 0$, for every $\theta\in\Theta$. To obtain uniformity, note that $\forall\theta_1,\theta_2\in\Theta$,

$$\begin{aligned}&|\rho(m_{n\theta_1},m_{\theta_1})-\rho(m_{n\theta_2},m_{\theta_2})|^2\\ &\quad\le\int([m_{n\theta_1}(x)-m_{n\theta_2}(x)]-[m_{\theta_1}(x)-m_{\theta_2}(x)])^2\,dG(x)\\ &\qquad\times\int([m_{n\theta_1}(x)+m_{n\theta_2}(x)]-[m_{\theta_1}(x)+m_{\theta_2}(x)])^2\,dG(x).\end{aligned} \tag{4.8}$$

Let $H_n(x):=n^{-1}\sum_{i=1}^n K_b((nx-i)/nb)h(i/n)$, where h is as in (m3). By (m3), the first factor of the product in the r.h.s. of (4.8) is bounded above by $C\|\theta_1-\theta_2\|^{2\beta}(\int H_n^2(x)g(x)\,dx+1)$. Using the fact that $\iint K(s)K(s+t)\,ds\,dt=1$ and h and g are continuous, direct calculations show that

$$\int H_n^2(x)g(x)\,dx\to\int h^2(x)g(x)\,dx<\infty. \tag{4.9}$$

Because $m_\theta(x)$ is bounded on $[0,1]\times\Theta$, the second factor in (4.8) is bounded above by $C(\int H_n^2(x)g(x)\,dx+1)=O(1)$. These facts, together with the compactness of Θ imply that $\sup_{\theta\in\Theta}q_{n1}(\theta)=o_p(1)$ while $m_\theta(x)$ bounded on $[0,1]\times\Theta$ implies that $\sup_{\theta\in\Theta}q_{n2}(\theta)=O_p(1)$, thereby completing the proof of (4.6). □

Upon taking $m=m_{\theta_0}$ in the above lemma one immediately obtains the following:

Corollary 4.1. *Suppose* (1.1), $\mathcal{H}$, (2.2), (2.4), (2.5), *and* (m1)–(m3) *hold. Then* $\vartheta_n\to_p\theta_0$, $\hat{\theta}_n\to_p\theta_0$, *in probability.*

Consistency of $\hat{\mathcal{D}}_n$ *test*: Consistency of the above test that rejects $\mathcal{H}$ in favor of an alternative $m\notin\mathcal{M}$ whenever $|\hat{\mathcal{D}}_n|$ is large will be implied by showing $|\hat{\mathcal{D}}_n|\to_p\infty$, under the given alternative m. We establish the latter result below. Let m be a regression function, $d_\theta(x):=m_\theta(x)-m(x)$, and

$$D_{n\theta}(x):=\frac{1}{n}\sum_{i=1}^n K_{bi}(x)\,d_\theta\left(\frac{i}{n}\right),\quad x\in[0,1],\ \theta\in\theta.$$

The following theorem provides a set of sufficient conditions under which $|\hat{\mathcal{D}}_n|\to_p\infty$.

Theorem 4.1. *Suppose* (2.2), (2.4), (2.5) *and* (m3) *hold. Furthermore, suppose the alternative hypothesis* $\mathcal{H}_1:\mu(x)=m(x),\forall x\in[0,1]$ *holds with the additional assumption that m is continuous*, $\inf_\theta\rho(m,m_\theta)>0$ *and the errors* $\xi_i=Y_i-m(i/n)$, $1\le i\le n$, *are homoscedastic zero mean r.v.'s having finite and positive variance* σ^2 *and finite fourth moment. Then,* $|\hat{\mathcal{D}}_n|\to_p\infty$.

Proof. Recall the definition of V_n from (4.4). Subtract and add $m(i/n)$ to $Y_i-m_{\hat{\theta}_n}(i/n)$ and expand the quadratic integrand to rewrite $M_n(\hat{\theta}_n)=S_{n1}-2S_{n2}+S_{n3}$, where

$$S_{n1}:=\int V_n^2\,dG,\quad S_{n2}:=\int V_n(x)D_{n\hat{\theta}_n}(x)\,dG(x),\quad S_{n3}:=\int D_{n\hat{\theta}_n}^2(x)\,dG(x).$$

Note that S_{n1} is like $M_n(\theta_0)$. Argue as for (2.13) to establish that under the current set up, $nb^{1/2}(S_{n1}-\mathcal{C}_n^*)\to_d N_1(0,\gamma^2)$, where $\mathcal{C}_n^*=\sum_{i=1}^n\int K_{bi}^2(x)\xi_i^2\,dG(x)/n^2$, and γ^2 as in (2.13).

Again, write T for $T(m)$. We claim $S_{n3}=\rho(m,m_T)+o_p(1)$. To see this, proceed as follows. Subtract and add $m_T(i/n)$ in the ith summand of $D_{\hat\theta_n}(x)$ and expanding the quadratic to rewrite $S_{n3}=S_{n31}+2S_{n32}+S_{n33}$, where

$$S_{n31}:=\int D_{nT}^2(x)\,dG(x),\quad S_{n32}:=\int D_{nT}(x)[m_{n\hat\theta_n}(x)-m_{nT}(x)]\,dG(x),$$

$$S_{n33}:=\int [m_{n\hat\theta_n}(x)-m_{nT}(x)]^2\,dG(x).$$

Continuity of $m(x)$, $m_\theta(x)$ and $g(x)$ in $x\in[0,1]$, readily implies

$$S_{n31}=\int\left[\frac{1}{n}\sum_{i=1}^n K_{bi}(x)\left[m\left(\frac{i}{n}\right)-m_T\left(\frac{i}{n}\right)\right]\right]^2 dG(x)\to\rho(m,m_T).$$

By (m3), $S_{n33}\le\|\hat\theta_n-T\|^{2\beta}\int H_n^2(x)\,dG(x)=o_p(1)$, by consistency of θ_n for T and (4.9). By the C–S inequality, $|S_{n32}|\le|S_{n31}S_{n33}|^{1/2}=o_p(1)$. Therefore, $S_{n3}=\rho(m,m_T)+o_p(1)$. By (4.5) and the C–S inequality, $|S_{n2}|\le|S_{n1}|^{1/2}|S_{n3}|^{1/2}=O_p((bn)^{-1/2})$.

Next, note that

$$\hat{\mathcal{C}}_n-C_n^*=-\frac{2}{n^2}\sum_{i=1}^n\int K_{bi}^2(x)\xi_i D_{\hat\theta_n}\left(\frac{i}{n}\right)dG(x)+\frac{1}{n^2}\sum_{i=1}^n\int K_{bi}^2(x)D_{\hat\theta_n}^2\left(\frac{i}{n}\right)dG(x)=-2\mathcal{D}_1+\mathcal{D}_2\quad\text{say.}$$

Calculations similar to the above show that

$$\begin{aligned}|\mathcal{D}_1|&\le\int\frac{1}{n^2}\sum_{i=1}^n K_{bi}^2(x)\left|m_{\hat\theta_n}\left(\frac{i}{n}\right)-m_T\left(\frac{i}{n}\right)\right||\xi_i|dG(x)+\int\frac{1}{n^2}\sum_{i=1}^n K_{bi}^2(x)\left|m_T\left(\frac{i}{n}\right)-m\left(\frac{i}{n}\right)\right||\xi_i|dG(x)\\&\le C\int\frac{1}{n^2}\sum_{i=1}^n K_{bi}^2(x)h\left(\frac{i}{n}\right)|\xi_i|\,dG(x)\|\hat\theta_n-T\|^\beta+\int\frac{1}{n^2}\sum_{i=1}^n K_{bi}^2(x)\left|m_T\left(\frac{i}{n}\right)-m\left(\frac{i}{n}\right)\right||\xi_i|dG(x).\end{aligned}\tag{4.10}$$

Direct calculations show that the expected value of the second term in this bound is of the order $(nb)^{-1}$ so that it is $O_p((nb)^{-1})$. Similarly, because $\|\hat\theta_n-T\|=o_p(1)$, the first term in this upper bound is $o_p((nb)^{-1})$. As a result, $\mathcal{D}_1=O_p((nb)^{-1})$. Similar calculations show that $\mathcal{D}_2=O_p((nb)^{-1}\|\hat\theta_n-T\|^{2\beta})$. By the LLN's, $\gamma_n\to_p\gamma$. All these results together yield

$$\hat{\mathcal{D}}_n=nb^{1/2}\gamma^{-1}(S_{n1}-C_n^*)+nb^{1/2}\gamma^{-1}\rho(m,m_T)+o_p(nb^{1/2}),$$

hence the theorem, because $\inf_\theta\rho(m,m_\theta)=\rho(m,m_T)>0$ and $nb^{1/2}\to\infty$. □

Power at local alternatives: Here we shall now study the asymptotic power of the proposed $\hat{\mathcal{D}}_n$ test against some local alternatives. Accordingly, let ψ be a known real valued function on [0,1] such that

$$\int\dot m_{\theta_0}\psi\,dG=0.\tag{4.11}$$

Consider the sequence of local alternatives

$$\mathcal{H}_{1n}:\mu(x)=m_{\theta_0}(x)+\delta_n\psi(x),\quad\delta_n=1/\sqrt{nb^{1/2}}.\tag{4.12}$$

The following theorem gives the asymptotic distributions of $\hat\theta_n$ and $\hat{\mathcal{D}}_n$ under $\mathcal{H}_{1n}$.

Theorem 4.2. *Suppose* (1.1), (2.2), (2.4), (2.5), (m1)–(m6) *and* $\mathcal{H}_{1n}$ *hold with* $Y_i-\mu(i/n),1\le i\le n$ *being i.i.d. zero mean r.v.'s having finite and positive variance* τ^2 *and finite fourth moment. Additionally, assume* ψ *and* $\dot m_{\theta_0}$ *are Lischitz continuous of order* 1 *and that* (4.11) *holds. Then,*

$$n^{1/2}(\hat\theta_n-\theta_0)\to_d\mathcal{N}_q(0,\Sigma_0^{-1}\Sigma\Sigma_0^{-1}),\tag{4.13}$$

$$\hat{\mathcal{D}}_n\to_d\mathrm{N}\left(\gamma^{-1}\int\psi^2\,dG,1\right).\tag{4.14}$$

Remark 4.1. From (4.14) we readily obtain that the asymptotic power of the $\hat{\mathcal{D}}_n$–test against $\mathcal{H}_{1n}$ is $\beta(\psi):=1-\Phi(z_\alpha-\gamma^{-1}\int\psi^2\,dG)$. Clearly $\beta(\psi)$ is strictly increasing in $\vartheta(g):=\gamma^{-1}\int\psi^2(x)g(x)\,dx$. Thus to find a G that maximizes this power is equivalent to finding a g that maximizes $\vartheta(g)$ w.r.t. g. Now, recall that $\gamma=2(\int g^2(x)\,dx)^{1/2}(\int K_*^2(t)\,dt)^{1/2}$. Clearly, with $a:=2(\int K_*^2(t)\,dt)^{1/2}$,

$$\vartheta(g)=\frac{\int\psi^2(x)g(x)\,dx}{a(\int g^2(x)\,dx)^{1/2}}\le\frac{(\int\psi^4(x)\,dx)^{1/2}}{a},$$

with equality if, and only if, $g\propto\psi^2$, in which case the maximum asymptotic power is $1-\Phi(z_\alpha-a^{-1}(\int\psi^4\,dG)^{1/2})$. Since $\vartheta(cg)=\vartheta(g)$, for all $c\in\mathbb{R}$, one may take $g\equiv\psi^2$ to obtain this power.

Proof of Theorem 4.2. Consider (4.13). Under $\mathcal{H}_{1n}$, $\varepsilon_i = Y_i - m_{\theta_0}(i/n) - \delta_n \psi(i/n)$. Let

$$\mathcal{V}_n(x) := \frac{1}{n}\sum_{i=1}^{n} K_{bi}(x)\varepsilon_i, \quad \overline{\psi}_n(x) := \frac{1}{n}\sum_{i=1}^{n} K_{bi}(x)\psi\left(\frac{i}{n}\right).$$

Because of (2.2) and the assumed Lipschitz continuity of order 1 of ψ and $\dot{m}_{\theta_0}$,

$$\int [\overline{\psi}_n - \psi]^2 \, dG = O(b^2), \quad \int \|\dot{m}_{n\theta_0} - \dot{m}_{\theta_0}\|^2 \, dG = O(b^2). \tag{4.15}$$

Now, first note that $nbM_n(\theta_0) = O_p(1)$. To see this, under $\mathcal{H}_{1n}$,

$$M_n(\theta_0) \le \int \mathcal{V}_n^2 \, dG + 2(nb^{1/2})^{-1} \int \overline{\psi}_n^2 \, dG.$$

Argue as for (4.5) to conclude that

$$\int \mathcal{V}_n^2 \, dG = O_p((nb)^{-1}). \tag{4.16}$$

This together with (4.15) shows that $M_n(\theta_0) = O_p((nb)^{-1})$. In turn, this fact and an argument similar to the one used in Koul and Ni (2004, p. 120) shows that

$$nb\|\hat{\theta}_n - \theta_0\| = O_p(1). \tag{4.17}$$

Let

$$U_n(x,\theta) := \frac{1}{n}\sum_{i=1}^{n} K_{bi}(x)\left(Y_i - m_\theta\left(\frac{i}{n}\right)\right).$$

Note that with $\dot{M}_n(\theta) = \partial M_n(\theta)/\partial\theta$, $\hat{\theta}_n$ satisfies

$$\dot{M}_n(\hat{\theta}_n) = -2\int U_n(x,\hat{\theta}_n)\dot{m}_{n\hat{\theta}_n}(x)\, dG(x). \tag{4.18}$$

Adding and subtracting $m_{\theta_0}(i/n)$ from $Y_i - m_{\hat{\theta}_n}(i/n)$ in $U_n(x,\hat{\theta}_n)$, rewrite (4.18) as

$$\int U_n(x,\theta_0)\dot{m}_{n\hat{\theta}_n}(x)\, dG(x) = \int D_{n\hat{\theta}_n}(x)\dot{m}_{n\hat{\theta}_n}(x)\, dG(x). \tag{4.19}$$

The right hand side of (4.19) is random only because of $\hat{\theta}_n$. Arguing as in Koul and Ni (2004, pp. 121–123), it can be shown to equal to $\mathcal{R}_n(\hat{\theta}_n - \theta_0) + o_p(1)$, where $\mathcal{R}_n$ is a matrix, $\mathcal{R}_n \to_p \Sigma_0$. The left hand side, under $\mathcal{H}_{1n}$, can be written as $\mathcal{S}_{n1} + \mathcal{S}_{n2}$, where

$$\mathcal{S}_{n1} := \int \mathcal{V}_n(x)\dot{m}_{n\hat{\theta}_n}(x)\, dG(x), \quad \mathcal{S}_{n2} := \delta_n \int \overline{\psi}_n(x)\dot{m}_{n\hat{\theta}_n}(x)\, dG(x).$$

But,

$$\mathcal{S}_{n1} = \int \mathcal{V}_n(x)\dot{m}_{\theta_0}(x)\, dG(x) + \int \mathcal{V}_n(x)[\dot{m}_{n\hat{\theta}_n}(x) - \dot{m}_{\theta_0}(x)]\, dG(x) = \mathcal{S}_{n11} + \mathcal{S}_{n12} \quad \text{say}.$$

In view of (4.17), for any $\varepsilon > 0$, there exists an N_ε and a K_ε such that $P(A_\varepsilon) > 1-\varepsilon$, for all $n > N_\varepsilon$, where $A_\varepsilon := \{(nb)^{1/2}\|\hat{\theta}_n - \theta_0\| \le K_\varepsilon\}$. Hence, using (4.16),

$$\begin{aligned} n|\mathcal{S}_{n12}|^2 &\le n\int \mathcal{V}_n^2 \, dG + \int \|\dot{m}_{n\hat{\theta}_n} - \dot{m}_{\theta_0}\|^2 \, dG \le O_p(b^{-1}) \max_{1 \le i \le n, (nb)^{1/2}\|\theta-\theta_0\| \le K_\varepsilon} \left\| \dot{m}_{n\theta}\left(\frac{i}{n}\right) - \dot{m}_{\theta_0}\left(\frac{i}{n}\right)\right\|^2 \\ &\int \left(\frac{1}{n}\sum_{i=1}^{n} K_{bi}(x)\right)^2 dG(x) = o_p(1), \end{aligned} \tag{4.20}$$

by (m5) and the fact that the last factor above tends to $\int dG < \infty$.

Argue as in the proof of Proposition 2.1, part (i), and as in Koul and Ni (2004, pp. 121–123) to obtain that under $\mathcal{H}_{1n}$, $n^{1/2}\mathcal{S}_{n11} \to_D \mathcal{N}_1(0,\tilde{\Sigma})$. Hence, $n^{1/2}\mathcal{S}_{n1} \to_D \mathcal{N}_1(0,\tilde{\Sigma})$.

Now, consider

$$n^{1/2}\mathcal{S}_{n2} = b^{-1/4}\int \overline{\psi}_n(x)[\dot{m}_{n\hat{\theta}_n}(x) - \dot{m}_{n\theta_0}(x)]\, dG(x) + b^{-1/4}\int \overline{\psi}_n(x)\dot{m}_{n\theta_0}(x)\, dG(x). \tag{4.21}$$

An argument like the one used in (4.20) shows that the first term in this bound tends to zero, in probability.

Consider the second summand. In view of (4.11) and (4.15), $\int \|\dot m_{n\theta_0}\|^2\,dG \to \int \|\dot m_{\theta_0}\|^2\,dG$ and

$$b^{-1/4}\int \overline{\psi}_n(x)\dot m_{n\theta_0}(x)\,dG(x) = b^{-1/4}\int [\overline{\psi}_n(x)-\psi(x)]\dot m_{n\theta_0}(x)\,dG(x) + b^{-1/4}\int \psi(x)[\dot m_{n\theta_0}(x)-\dot m_{\theta_0}(x)]\,dG(x) = O(b^{3/4}) \to 0.$$

Hence, $n^{1/2}\mathcal{S}_{n2}\to_p 0$, thereby completing the proof of (4.13).

Next, we sketch the proof of (4.14). Let $U_n(x)=U_n(x,\theta_0)$, and

$$T_{n1} := \int U_n^2\,dG, \quad T_{n2} := \int U_n[m_{n\theta_0}-m_{n\hat\theta_n}]\,dG, \quad T_{n3} := \int [m_{n\theta_0}-m_{n\hat\theta_n}]^2\,dG.$$

Then we can rewrite $M_n(\hat\theta_n)=T_{n1}+2T_{n2}+T_{n3}$. Because of $\sqrt{n}(\hat\theta_n-\theta_0)=O_p(1)$ and (m4), $T_{n3}=O_p(n^{-1})$ so that $nb^{1/2}T_{n3}=O_p(b^{1/2})=o_p(1)$.

Next, we shall show that $T_{n2}=O_p(n^{-1}b^{-1/4})$ implying $nb^{1/2}T_{n2}=O_p(b^{1/4})=o_p(1)$. By C–S, $T_{n2}^2 \le T_{n1}T_{n3}$. Under $\mathcal{H}_{1n}$, $Y_i - m_{\theta_0}(i/n)=\delta_n\psi(i/n)+\varepsilon_i$. Hence, by (4.15) and (4.16),

$$T_{n1} \le 2\int \mathcal{V}_n^2\,dG + \delta_n^2\int \overline{\psi}_n^2\,dG = O_p((nb)^{-1}) + O_p((nb^{1/2})^{-1}) = O_p((nb^{1/2})^{-1}),$$

so that $T_{n2}=O_p(n^{-1}b^{-1/4})$.

We need a more precise approximation of T_{n1}. For that purpose, write $T_{n1}=T_{n11}+2\delta_nT_{n12}+\delta_n^2T_{n13}$, where

$$T_{n11} := \int \mathcal{V}_n^2\,dG, \quad T_{n12} := \int \mathcal{V}_n\overline{\psi}_n\,dG, \quad T_{n13} := \int \overline{\psi}_n^2\,dG.$$

With $\tau^2=\mathrm{Var}(\varepsilon_1)$,

$$\mathrm{Var}\left(\int \mathcal{V}_n\psi\,dG\right) = \tau^2\frac{1}{n^2}\sum_{i=1}^n\left(\int K_{bi}(x)\psi(x)\,dG(x)\right)^2 = O(n^{-1}),$$

$$\left|T_{n12}-\int \mathcal{V}_n\psi\,dG\right|^2 \le \int \mathcal{V}_n^2\,dG\int |\overline{\psi}_n-\psi|^2\,dG = O_p(n^{-1}b),$$

$$nb^{1/2}\delta_n|T_{n12}| = n^{1/2}b^{1/4}(O_p(n^{-1/2})+O_p(n^{-1/2}b^{1/2})) = O_p(b^{1/4})\to_p 0.$$

In view of (4.15), $T_{n13}\to\int\psi^2$. We thus obtain that

$$nb^{1/2}T_{n1} = n^{1/2}T_{n11} + \int \psi^2\,dG + o_p(1). \tag{4.22}$$

Finally, we need to discuss asymptotic behavior of $\hat{\mathcal{C}}_n$ under the local alternative (4.12). With $\zeta_i = Y_i - m_{\theta_0}(i/n)$, rewrite $Y_i - m_{\hat\theta_n}(i/n) = \zeta_i - [m_{\hat\theta_n}(i/n)-m_{\theta_0}(i/n)]$ in $\hat{\mathcal{C}}_n$, to obtain

$$\hat{\mathcal{C}}_n = \frac{1}{n^2}\sum_{i=1}^n\int K_{bi}^2(x)\zeta_i^2\,dG(x) - \frac{2}{n^2}\sum_{i=1}^n\int K_{bi}^2(x)\zeta_i\left[m_{\hat\theta_n}\left(\frac{i}{n}\right)-m_{\theta_0}\left(\frac{i}{n}\right)\right]dG(x)$$
$$+\frac{1}{n^2}\sum_{i=1}^n\int K_{bi}^2(x)\left[m_{\hat\theta_n}\left(\frac{i}{n}\right)-m_{\theta_0}\left(\frac{i}{n}\right)\right)^2 dG(x) = C_{n1}-2C_{n2}+C_{n3}.$$

But with $\Delta_n=\hat\theta_n-\theta_0$, $d_{ni} := m_{\hat\theta_n}(i/n)-m_{\theta_0}(i/n)-\Delta_n'\dot m_{\theta_0}(i/n)$, and $\varepsilon_i = Y_i - m_{\theta_0}(i/n)-\delta_n\psi(i/n)$,

$$C_{n2} = \frac{1}{n^2}\sum_{i=1}^n\int K_{bi}^2(x)\varepsilon_i\,d_{ni}\,dG(x) + \frac{1}{n^2}\sum_{i=1}^n\int K_{bi}^2(x)\varepsilon_i\Delta_n'\dot m_{\theta_0}\left(\frac{i}{n}\right)dG(x)$$
$$+\frac{\delta_n}{n^2}\sum_{i=1}^n\int K_{bi}^2(x)\psi\left(\frac{i}{n}\right)d_{ni}\,dG(x) + \frac{\delta_n}{n^2}\sum_{i=1}^n\int K_{bi}^2(x)\psi\left(\frac{i}{n}\right)\Delta_n'\dot m_{\theta_0}\left(\frac{i}{n}\right)dG(x).$$

Recall that $\delta_n = 1/\sqrt{nb^{1/2}}$. Use assumptions (2.5) and (m3), to show that the first and the second terms in C_{n2} are $O_p(n^{-3/2}b^{-1})$, the third and the fourth terms are of the order $O_p(n^{-2}b^{-5/4})$. This implies $C_{n2}=o_p(\delta_n^2)$. Similarly, one can show that $C_{n3}=o_p(\delta_n^2)$.

Since $Y_i - m_{\theta_0}(i/n)=\varepsilon_i+\delta_n\psi(i/n)$, if we let $D_n = n^{-2}\sum_{i=1}^n\int K_{bi}^2\varepsilon_i^2\,dG$, then using the similar argument, we can show that $C_{n1}=D_n+o_p(\delta_n^2)$. Finally, we also have $\hat\gamma_n\to_p\gamma$.

Therefore, under the local alternative hypothesis (4.12),

$$nb^{1/2}\hat\gamma_n^{-1}(M_n(\hat\theta_n)-\hat{\mathcal{C}}_n) = nb^{1/2}\hat\gamma_n^{-1}(T_{n11}-D_n)+\hat\gamma_n^{-1}T_{n13}+o_p(1),$$

which, together with the fact $nb^{1/2}\hat\gamma_n^{-1}(T_{n11}-D_n)\to_D\mathcal{N}_1(0,1)$ and $T_{n13}\to_p\int\psi^2\,dG$ concludes the proof of the theorem. □

Acknowledgements

Author wishes to thank the Associate Editor and the referee of this paper for their careful reading of the paper and for raising questions that helped to improve the content of the original submission.

References

Hall, P., 1984. Central limit theorem for integrated square error of multivariate nonparametric density estimators. J. Multivariate Anal. 14, 1–16.
Hall, P., Heyde, C.C., 1980. Martingale Limit Theory and its Application. Academic Press, Inc., New York, London.
Hart, J.D., 1997. Nonparametric Smoothing and Lack-of-Fit Tests. In: Springer Series in Statistics. Springer-Verlag, New York.
Hsing, T., Wu, W., 2004. On weighted *U*-statistics for stationary processes. Ann. Probab. 32, 1600–1631.
Koul, H.L., Ni, P., 2004. Minimum distance regression model checking. J. Statist. Plann. Inference 119 (1), 109–141.
Koul, H.L., Song, W., 2009. Minimum distance regression model checking with Berkson measurement errors. Ann. Statist. 37 (1), 132–156.
Major, P., 1994. Asymptotic distribution of weighted *U*-statistics. Ann. Probab. 22, 1514–1535.
O'Neil, K.A., Redner, R.A., 1993. Asymptotic distribution of weighted *U*-statistics of degree two. Ann. Probab. 21, 1159–1169.
Rifi, M., Utzet, F., 2000. On the asymptotic behavior of weighted *U*-statistics. J. Theoret. Probab. 13 (1), 141–167.