Testing symmetry in nonparametric regression models

Holger Dette
Ruhr-Universität Bochum
Fakultät für Mathematik
44780 Bochum
Germany
email: holger.dette@ruhr-uni-bochum.de
FAX: +49 2 34 32 14 559

Sorina Kusi-Appiah
Ruhr-Universität Bochum
Fakultät für Mathematik
44780 Bochum
Germany
email: sorina.kusi-appiah@ruhr-uni-bochum.de

Natalie Neumeyer
Ruhr-Universität Bochum
Fakultät für Mathematik
44780 Bochum
Germany
email: Natalie.Neumeyer@ruhr-uni-bochum.de

May 11, 2001

Abstract

In a recent paper Ahmad and Li (1996) proposed a new test for symmetry of the error distribution in linear regression models and proved asymptotic normality for the distribution of the corresponding test statistic under the null hypothesis and consistency under fixed alternatives. The present paper has three purposes. On the one hand we derive the asymptotic distribution of the statistic considered by Ahmad and Li (1996) under fixed alternatives and demonstrate that asymptotic normality is still valid but with a different rate of convergence. On the other hand we generalize Ahmad and Li's (1996) test of a symmetric error distribution to general nonparametric regression models. Moreover, it is also demonstrated that a bootstrap version of the new test for symmetry has good finite sample properties.

AMS Classification: Primary 62G05

Keywords and Phrases: nonparametric regression, goodness-of-fit, test of symmetry, wild bootstrap

1 Introduction

Consider the common nonparametric regression model

$$Y_{i,n} = m(x_{i,n}) + \varepsilon_{i,n} \ ; \quad i = 1, \dots, n \tag{1.1}$$

where $m(\cdot)$ denotes the (unknown) regression function, $x_{1,n}, \dots, x_{n,n}$ are fixed design points and $\varepsilon_{1,n}, \dots, \varepsilon_{n,n}$ are i.i.d. random variables with expectation $E(\varepsilon_{i,n}) = 0$ and constant variance $\mathrm{Var}(\varepsilon_{i,n}) = \sigma^2 > 0$ $(i = 1, \dots, n)$. Much effort has been devoted to the problem of estimating the regression function $m(\cdot)$ and variance σ^2 [see e.g. Härdle (1990), Fan and Gijbels (1996), Wand and Jones (1995), Hall, Kay and Titterington (1990) and Hall and Marron (1990) among many others] or testing for a parametric form of these quantities [see e.g. Härdle and Mammen (1993), Gonzalez-Manteiga and Cao Abad (1993), Azzalini and Bowman (1993), Dette and Munk (1998) among many others]. In many cases the efficiency of the methods can be improved under the additional assumption of a symmetric error distribution [see for example Bickel (1982) in the case of adaptive estimation of a parametric model, Azzalini and Bowman (1993) or Kulasekera and Wang (2001) in the case of testing].
For these reasons we are interested in a test for symmetry of the error distribution in the nonparametric regression model (1.1), that is

$$H_0 : f(x) = f(-x) \quad \text{a.e.} \tag{1.2}$$

where f denotes the density of the error distribution in the nonparametric regression model (1.1). Note that this problem has been considered by several authors in the case $m \equiv 0$, i.e. testing for symmetry of a density function [see e.g. Huskova (1984), Hollander (1988) for reviews on this subject, Baringhaus and Henze (1992), Ahmad and Li (1996) and the recent work of Cabaña and Cabaña (2000) among many others]. Recently, Fan and Gencay (1995) and Ahmad and Li (1996) proposed tests of symmetry in linear regression models (i.e. $m(x) = \sum_{j=1}^{p} \theta_j g_j(x)$ for some known functions $g_1, \dots, g_p$ and unknown parameters $\theta_1, \dots, \theta_p$) using the residuals from the parametric fit, but to the knowledge of the authors a test of symmetry in the general nonparametric regression model (1.1) has not been considered so far.
The present paper has three purposes. At first we will demonstrate that under certain assumptions of regularity the test of Ahmad and Li (1996) can also be used for testing symmetry in nonparametric regression models if the residuals from the parametric fit are replaced by residuals from an appropriate nonparametric fit. We prove that under the null hypothesis of symmetry the new test statistic has also an asymptotic normal law. Secondly, we give a more detailed analysis of the asymptotic distribution under a fixed alternative. While Ahmad and Li (1996) only proved consistency in this case our results show that in the case of a nonsymmetric density an appropriate standardized version of the test statistic is still asymptotically normal distributed but with a different rate of convergence. This result allows a simple calculation of the probability of the type II error, which is of particular importance if the null hypothesis cannot be rejected [see e.g. Berger and Delampady (1987) or Sellke, Bayarri and Berger (2001)]. Finally, because it is well known that goodness-of-fit statistics based on nonparametric residuals have poor accuracy with respect to the approximation of the asymptotic level [see e.g. Härdle and Mammen (1993), Hjellvik, Yao and Tjøstheim (1998)] we investigate the finite sample properties of a wild bootstrap version of the new test.
The remaining part of this paper is organized as follows. In Section 2 we review the approach introduced by Ahmad and Li (1996) in the context of testing for a symmetric density (i.e.

$m \equiv 0$) and extend their asymptotic analysis under the null hypothesis to the case of fixed alternatives. We prove asymptotic normality of the corresponding statistic under fixed alternatives with a $\sqrt{n}$ rate and give some hints how this result could be used for the analysis of the probability of the type II error if the hypothesis of symmetry has not been rejected. Section 3 deals with the problem of testing symmetry in the general nonparametric regression model (1.1). We prove asymptotic normality of an appropriately scaled test statistic obtained from nonparametric residuals under the hypothesis of symmetry and under fixed alternatives with different rates corresponding to these cases. In Section 4 a simulation study is presented, which illustrates the finite sample properties of a wild bootstrap version of the new test. Here it is demonstrated that the bootstrap yields a reasonable approximation of the nominal level, even in relatively small samples. Finally, the more technical proofs of all results are deferred to the appendix.

2 A test of symmetry revisited and its asymptotic properties under fixed alternatives

Let $\varepsilon_1, \ldots, \varepsilon_n$ denote i.i.d. random variables with bounded and continuous density f [that is we put $m \equiv 0$ in (1.1), and put $Y_i = \varepsilon_i$] and assume that we are interested in a test of the hypothesis of symmetry (1.2) where f denotes the density of the ε_i. Recently, Ahmad and Li (1996) proposed

$$\hat{I}_n = \int_{\mathbb{R}} [f_n(x) - f_n(-x)]\, dF_n(x) \tag{2.1}$$

as a test statistic for this problem, where

$$f_n(x) = (na)^{-1} \sum_{i=1}^{n} k\Big(\frac{x - \varepsilon_i}{a}\Big) \tag{2.2}$$

ist the common estimator of the density at the point x [see e.g. Rosenblatt (1956)] with bandwidth $a > 0$ and F_n denotes the empirical distribution function of the random variables $\varepsilon_1, \ldots, \varepsilon_n$. Note that a similar approach has been considered by Rosenblatt (1975) in the context of testing independence of a bivariate density and that $\hat{I}_n$ is a consistent estimate of

$$I = \int_{\mathbb{R}} (f(x) - f(-x))dF(x) = \frac{1}{2} \int_{\mathbb{R}} (f(x) - f(-x))^2 dx, \tag{2.3}$$

where F denotes the distribution function of the random variables $\varepsilon_1, \ldots, \varepsilon_n$ and the last equality follows from standard calculations. For the asymptotic analysis of the distributional properties of the statistic $\hat{I}_n$ under the null hypothesis (1.2) of symmetry Ahmad and Li (1996) assumed for the kernel k in the density estimator (2.2)

$$\begin{gathered} k \textit{ is a bounded and symmetric density} \\ \lim_{|x|\to\infty} |x|k(x) = 0, \quad \int_{\mathbb{R}} x^2 k(x)dx < \infty, \quad \int_{\mathbb{R}} k^2(x)dx < \infty \end{gathered} \tag{2.4}$$

and for the bandwidth $a = a_n$ in (2.2)

$$a \to 0 \quad \text{and} \quad na \to \infty \tag{2.5}$$

as $n \to \infty$. If these conditions are satisfied the weak convergence

$$n\sqrt{a}(\hat{I}_n - c(n)) \xrightarrow{\mathcal{D}} \mathcal{N}(0, 4\sigma^2) \tag{2.6}$$

can be established, where $c(n) = k(0)/na$, the asymptotic variance is given by

$$\sigma^2 = \int_{\mathbb{R}} k^2(x)dx \int_{\mathbb{R}} f^2(x)dx \tag{2.7}$$

and a consistent estimator of σ^2 is obtained from

$$\sigma_\varepsilon^2 = \int_{\mathbb{R}} k^2(x)dx \int_{\mathbb{R}} f_n(x)dF_n(x) \ \xrightarrow{P} \ \sigma^2 \tag{2.8}$$

[see Ahmad and Li (1996) for more details]. The asymptotic analysis of the distribution of $\hat{I}_n$ under fixed alternatives requires some additional assumptions on the underlying kernel, density and bandwidth. To be precise suppose that for some $r \geq 2$

(2.9) *k satisfies (2.4) and is a continuously differentiable kernel of order r, i.e.*

$$\int k(u)u^j\, du \begin{cases} = 1 & : \quad j = 0 \\ = 0 & : \quad 1 \leq j \leq r-1 \\ \neq 0 & : \quad j = r \end{cases}$$

[see Gasser, Müller and Mammitzsch (1985)],

$$f \in C^{(r)}(\mathbb{R}) \tag{2.10}$$

and assume for the bandwidth in (2.2)

$$a \to 0, \ na \to \infty, \ na^{2r} \to 0 \tag{2.11}$$

as $n \to \infty$. Under these assumptions the following result provides a refinement in the asymptotic analysis of the statistic $\hat{I}_n$ under fixed alternatives. The proof is deferred to the appendix.

Theorem 2.1: *If the assumptions (2.9), (2.10), (2.11) are satisfied and the hypothesis (1.2) of symmetry is not satisfied, then we have as $n \to \infty$*

$$\sqrt{n}\Big(\hat{I}_n - I - c(n)\Big) \to \mathcal{N}(0, 4\tau^2) \tag{2.12}$$

where the asymptotic variance is given by

$$\tau^2 = \ Var[f(F^{-1}(U)) - f(-F^{-1}(U))] \tag{2.13}$$

and U denotes a random variable with rectangular distribution, that is $U \sim \mathcal{U}[0, 1]$.

Remark 2.2: Note that Theorem 2.1 offers a simple possibility to estimate the probability of the type II error of the test, which rejects the hypothesis (1.2) of symmetry, if

$$\frac{n\sqrt{a}(\hat{I}_n - c(n))}{2\sigma_\varepsilon} > \Phi^{-1}(1-\alpha), \tag{2.14}$$

where Φ denotes the cumulative distribution function of the standard normal distribution. A standard calculation shows that in the case of non-symmetry $H_1 : I \neq 0$ we have

$$\begin{aligned} P(\text{type II error}) &= P\Big(\frac{n\sqrt{a}(\hat{I}_n - c(n))}{2\sigma_\varepsilon} \leq \Phi^{-1}(1-\alpha) \,\Big|\, H_1\Big) \\ &\sim \Phi\Big(\frac{1}{\sqrt{na}}\frac{\sigma_\varepsilon}{\tau}\Phi^{-1}(1-\alpha) - \frac{\sqrt{n}I}{2\tau}\Big) \sim \Phi\Big(-\frac{\sqrt{n}I}{2\tau}\Big), \end{aligned}$$

where I denotes the measure of non-symmetry defined by (2.3) and τ^2 is given by (2.13). Now I is estimated by $\hat{I}_n$, τ^2 can be rewritten as

$$\tau^2 = \int_{\mathbb{R}} f(x)\{f(x) - f(-x)\}^2 dx - I^2$$

and consequently a consistent estimator of τ^2 is given by

$$\tau_\varepsilon^2 = \int \{f_n(x) - f_n(-x)\}^2 dF_n(x) - \hat{I}_n^2.$$

This yields as estimate for the probability of a type II error

$$P(\text{type II error}) \sim \Phi\Big(\frac{\Phi^{-1}(1-\alpha)}{\sqrt{na}}\frac{\sigma_\varepsilon}{\tau_\varepsilon} - \frac{\sqrt{n}\hat{I}_n}{2\tau_\varepsilon}\Big).$$

3 Testing symmetry in nonparametric regression models

We are now discussing the problem of testing symmetry in the general model (1.1). In order to avoid a tedious discussion of boundary effects we assume that m is periodic. This assumption can be dropped, either by considering modified estimates near the boundary 0 or 1 [see e.g. Gasser and Müller (1984), Wand and Jones (1995) or Fan and Gijbels (1996)] or restricting the subsequent investigations to an interval of the form $[\delta, 1-\delta] \subset [0,1]$ for some $\delta > 0$. The basic idea for the construction of the test statistic is very simple: the unknown errors $\varepsilon_{i,n}$ in (1.1) are replaced by nonparametric residuals $\hat{\varepsilon}_{i,n} = Y_{i,n} - \hat{m}(x_{i,n})$, where $\hat{m}(x_{i,n})$ is the Nadaraya-Watson estimator at the point $x_{i,n}$, given by

$$\hat{m}(x_{i,n}) := \frac{\sum_{l=1}^n K\left(\frac{x_{i,n}-x_{l,n}}{\tilde{a}}\right) Y_{l,n}}{\sum_{l=1}^n K\left(\frac{x_{i,n}-x_{l,n}}{\tilde{a}}\right)}, \tag{3.1}$$

[see Nadaraya (1964) or Watson (1964)]. Note that we use a different bandwidth $\tilde{a}$ in the definition of the kernel estimate. To be precise let

$$\hat{f}_n(\varepsilon) = (na)^{-1} \sum_{i=1}^n k\left(\frac{\varepsilon - \hat{\varepsilon}_{i,n}}{a}\right) \tag{3.2}$$

denote the density estimator and

$$\hat{F}_n(\varepsilon) = \frac{1}{n}\sum_{j=1}^{n} I(\hat{\varepsilon}_{j,n} \le \varepsilon) \tag{3.3}$$

the empirical distribution function based on the sample of the nonparametric residuals $\hat{\varepsilon}_{1,n}, \dots, \hat{\varepsilon}_{n,n}$. The empirical measure of symmetry is defined by

$$\begin{aligned} T_n &= \int [\hat{f}_n(\varepsilon) - \hat{f}_n(-\varepsilon)]\, d\hat{F}_n(\varepsilon) \\ &= (n^2 a)^{-1} \sum_{i=1}^{n}\sum_{j=1}^{n} \Big[k(\frac{\hat{\varepsilon}_{i,n} - \hat{\varepsilon}_{j,n}}{a}) - k(\frac{\hat{\varepsilon}_{i,n} + \hat{\varepsilon}_{j,n}}{a}) \Big] \end{aligned} \tag{3.4}$$

and for the asymptotic analysis of this statistic we need several assumptions, which are stated below. First note that there appear two kernels in the definition of T_n, the kernel $k(\cdot)$ in the density estimator (3.3) and the kernel $K(\cdot)$ used for the construction of the curve estimator (3.1). For the kernel $k(\cdot)$ we assume

$$\begin{aligned} k \in C^{(m+1)}(\mathbb{R}), k^{(m+1)} & \textit{ is bounded for some } m \ge 1 \\ k^{(s)} & \textit{ is of bounded variation for } s = 1, \dots, m \end{aligned} \tag{3.5}$$

and the design points are supposed to satisfy a Sacks and Ylvisaker (1970) condition, i.e.

$$\int_0^{x_{i,n}} p(x)dx = \frac{i}{n}\ ; \qquad i = 1, \dots, n;\ n \in \mathbb{N}\ ; \tag{3.6}$$

where p is positive density on the interval $[0, 1]$. We assume further for some $r \ge 2$

$$p, m \in C^{(r)}[0, 1], \tag{3.7}$$

$$K \textit{ is continuously differentiable, of order } r \textit{ and has compact support} \tag{3.8}$$

$$E[|\varepsilon_{i,n}|^4] < \infty. \tag{3.9}$$

Finally, the bandwidth $\tilde{a} = \tilde{a}_n$ is supposed to satisfy

$$\tilde{a} = c_n \Big(\frac{\log n}{n}\Big)^{\frac{1}{2r+1}} \tag{3.10}$$

and some positive sequence (c_n) which satisfies $0 < c' \le c_n \le c'' < \infty$ for all $n \in \mathbb{N}$. The relation between the bandwidths a and $\tilde{a}$ in the density and regression estimate is specified by

$$\sqrt{a} = o(\tilde{a}),\ \Big(\frac{\log n}{n}\Big)^{\frac{r(m+1)}{2r+1}} \frac{n}{a^{m+3/2}} \to 0,\ \ \Big(\frac{\log n}{n}\Big)^{\frac{2r}{2r+1}} n\sqrt{a} \to 0 \tag{3.11}$$

as $n \to \infty$. Note that condition (3.10) is sufficient for the strong uniform consistency of the nonparametric regression estimate [see e.g. Müller (1985), who considered an asymptotically equivalent estimator]. Our first result specifies the asymptotic distribution of the test statistic

T_n under the null hypothesis of a symmetric error distribution. The second statement gives the corresponding result under a fixed alternative. The proofs are given in the appendix.

Theorem 3.1: *If the assumptions (2.4), (2.5), (3.5) - (3.11) are satisfied and the hypothesis of a symmetric error distribution in the regression model (1.1) is valid, then the statistic T_n defined in (3.4) is asymptotically normal distributed. More precisely,*

$$n\sqrt{a}(T_n - c(n)) \xrightarrow{\mathcal{D}} \mathcal{N}(0, 4\sigma^2) ,$$

where $c(n) = k(0)/na$ and σ^2 is defined by (2.7) and can consistently be estimated by

$$\sigma^2_{\hat{\varepsilon}} = \int k^2(u)\, du \int \hat{f}_n(\varepsilon) d\hat{F}_n(\varepsilon) . \tag{3.12}$$

Theorem 3.2: *If the assumptions (2.9) - (2.11) and (3.5) - (3.11) are satisfied and the hypothesis of a symmetric error distribution is <u>not</u> valid, then the statistic T_n defined by (3.4) has an asymptotic normal law, i.e.*

$$\sqrt{n}(T_n - I - c(n)) \xrightarrow{\mathcal{D}} \mathcal{N}(0, 4\tau^2) ,$$

where the bias I and the asymptotic variance τ^2 are defined in (2.3) and (2.13), respectively.

4 Finite sample properties

In this section we investigate the finite sample properties of a test which can be obtained from the results given in Section 3. More precisely, Theorem 3.1 shows that an asymptotic level α test can be obtained by rejecting the null hypothesis of a symmetric error distribution whenever

$$\frac{n\sqrt{a}(T_n - c(n))}{2\sigma_{\hat{\varepsilon}}} > \Phi^{-1}(1-\alpha) \tag{4.1}$$

where $\Phi^{-1}(1-\alpha)$ denotes the $(1-\alpha)$ quantile of the standard normal distribution. Moreover, the proof of this theorem shows that an alternative asymptotic level α test is given by rejecting the null hypothesis whenever

$$\frac{n\sqrt{a}}{2\sigma_{\hat{\varepsilon}}} T_{2n} > \Phi^{-1}(1-\alpha), \tag{4.2}$$

where the statistic T_{2n} is defined in (5.14) and obtained from T_n by omitting the pairs with (i,j) with $i = j$ in the definition (3.4) of the statistic T_n. The results of Ahmad and Li (1996) indicate a better performance of the test (4.2) because of its neglectible bias and for this reason we will restrict ourselves to a discussion of tests based on the statistic T_{2n}. It was observed by many authors [see e.g. Härdle and Mammen (1993) or Hjellvik, Yao and Tjøstheim (1998)] that the normal approximation for testing hypotheses regarding the mean function $m(\cdot)$ is rather poor and most authors recommend the application of the bootstrap for this type of problems.

We have implemented a wild bootstrap procedure for the problem of testing for a symmetric error distribution in the nonparametric regression model (1.1) based on the statistic T_{2n} defined in (5.14). To be precise, let

$$\hat{\varepsilon}_{i,n} = Y_{i,n} - \hat{m}(x_{i,n}) \ ; \qquad i = 1, \dots, n \ ;$$

denote the nonparametric residuals obtained from the Nadaraya-Watson estimate. Let $V_{i,n}, \dots, V_{n,n}$ denote i.i.d. random variables (independent of the $Y_{i,n}$) with

$$P(V_{i,n} = 1) = P(V_{i,n} = -1) = \frac{1}{2},$$

define the bootstrap sample by

$$Y^*_{i,n} = \hat{m}(x_{i,n}) + V_{i,n}\hat{\varepsilon}_{i,n} \tag{4.3}$$

and the corresponding statistic obtained from the bootstrap sample by T^*_{2n}. The hypothesis of symmetry is rejected if $T_{2n} > t^*_{2n,1-\alpha}$, where $t^*_{2n,1-\alpha}$ denotes the $(1-\alpha)$ quantile of the bootstrap distribution, i.e.

$$P(T^*_{2n} > t^*_{2n,1-\alpha}) = \alpha.$$

For our study we choose $B = 200$ bootstrap replications and 1000 runs for each scenario. We use an equidistant design (i.e. $p(x) \equiv 1$), the quartic kernel

$$K(x) = \frac{15}{32}(7x^4 - 10x^2 + 3)I_{[-1,1]}(x)$$

for the Nadaraya-Watson estimate and the Gaussian kernel

$$k(x) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{x^2}{2})$$

for the construction of the test statistic T_{2n} (note that this kernel is differentiable of infinite order). Only the residuals in the interval $[\delta, 1-\delta]$ with $\delta = 0.05$ are used for the construction of the statistic T_{2n} in order to exclude boundary effects. The application of boundary kernels [see Gasser and Müller (1984)] or local polynomial estimators [see Wand and Jones (1995) or Fan and Gijbels (1996)] yields similar results with a slight improvement with respect to power. The choice of the bandwidths is important and we use

$$a = \Big(\frac{\sigma}{n}\Big)^{2/9} \ , \qquad \tilde{a} = \Big(\frac{\sigma \log n}{n}\Big)^{1/9} \tag{4.4}$$

in accordance with condition (3.11), where $m = 5, r = 4$ [note that for any $\eta > 0$ the quantities $a^{(1+\eta)}$ and $\tilde{a}$ satisfy assumption (3.11)], and σ^2 is the variance of the error distribution. We simulated the proportion of rejections of of the test for the level $\alpha = 2.5\%$, 5%, 10% and 20% and for samples of size $n = 20$, 30, 40, 50 and 100.
Our first results show the simulated level of the bootstrap test for a standard normal distributed error and regression functions $m(x) = x^2$ (Table 4.1) and $m(x) = \sin(2\pi x)$ (Table 4.2). In Table 4.3 we considered a similar scenario for the more heavy tailed t-distribution with five degrees of freedom. We observe a reasonable approximation of the level in nearly all cases and a slightly better performance (especially in small samples) in the case of a gaussian error disribution.

	n=20	n=30	n=40	n=50	n=100
0.025	0.042	0.036	0.032	0.039	0.034
0.05	0.066	0.061	0.059	0.07	0.057
0.10	0.123	0.114	0.118	0.122	0.93
0.20	0.218	0.204	0.207	0.203	0.211

Table 4.1: *Simulated level of the wild bootstrap test of symmetry in the nonparametric regression model (1.1) for a standard normal distributed error and regression function* $m(x) = x^2$.

	n=20	n=30	n=40	n=50	n=100
0.025	0.038	0.042	0.025	0.032	0.048
0.05	0.061	0.064	0.047	0.053	0.078
0.10	0.098	0.106	0.088	0.097	0.112
0.20	0.216	0.213	0.181	0.2	0.195

Table 4.2: *Simulated level of the wild bootstrap test of symmetry in the nonparametric regression model (1.1) for a standard normal distributed error and regression function* $m(x) = \sin(2\pi x)$.

	n=20	n=30	n=40	n=50	n=100
0.025	0.047	0.052	0.043	0.048	0.03
0.05	0.071	0.077	0.072	0.069	0.053
0.10	0.116	0.118	0.118	0.131	0.099
0.20	0.215	0.228	0.219	0.224	0.203

Table 4.3: *Simulated level of the wild bootstrap test of symmetry in the nonparametric regression model (1.1) for a* t_5*-distributed error and regression function* $m(x) = \sin(2\pi x)$.

In order to illustrate the performance of the wild bootstrap test under the alternative of non-symmetry, we consider again the regression functions $m(x) = x^2$, $m(x) = \sin(2\pi x)$. For the error disrtibution we choose a normalized χ^2-distribution with 1 and 2 degrees of freedom such that $E[\varepsilon_{i,n}] = 0$ and $\mathrm{Var}[\varepsilon_{i,n}] = 1$. The results are displayed in Table 4.4 - 4.7 and demonstrate that the non-symmetry is detected with reasonable probabilities in all cases. Note that the proportion of rejections in the χ_2^2-case is smaller than in the χ_1^2-case, which corresponds with our intuition. Moreover, for small sample sizes the rejection probabilities in the case of a more oscillating regression function $m(x) = \sin(2\pi x)$ are substantially lower than the corresponding probabilities in the smooth case $m(x) = x^2$. This can be heuristically explained by the observation that in small samples an oscillating regression function produces artifical symmetry of the residuals, which is not caused by the error distribution. Nevertheless with increasing sample sizes the regression function can be estimated with sufficient precision and the test of symmetry behaves similar in the smooth and oscillating case.

	n=20	n=30	n=40	n=50	n=100
0.025	0.671	0.908	0.963	0.984	1
0.05	0.765	0.934	0.977	0.992	1
0.10	0.84	0.96	0.99	0.996	1
0.20	0.918	0.987	0.998	1	1

Table 4.4: *Simulated power of the wild bootstrap test of symmetry in the nonparametric regression model (1.1) for a standardized* χ_1^2*-distributed error and regression function* $m(x) = x^2$.

	n=20	n=30	n=40	n=50	n=100
0.025	0.47	0.683	0.822	0.907	0.996
0.05	0.576	0.764	0.876	0.948	0.999
0.10	0.681	0.847	0.916	0.967	1
0.20	0.79	0.913	0.951	0.985	1

Table 4.5: *Simulated power of the wild bootstrap test of symmetry in the nonparametric regression model (1.1) for a standardized* χ_2^2*-distributed error and regression function* $m(x) = x^2$.

	n=20	n=30	n=40	n=50	n=100
0.025	0.307	0.57	0.702	0.83	0.993
0.05	0.389	0.655	0.779	0.879	0.995
0.10	0.531	0.758	0.862	0.929	0.998
0.20	0.686	0.869	0.933	0.969	1

Table 4.6: *Simulated power of the wild bootstrap test of symmetry in the nonparametric regression model (1.1) for a standardized* χ_1^2*-distributed error and regression function* $m(x) = \sin(2\pi x)$.

	n=20	n=30	n=40	n=50	n=100
0.025	0.239	0.429	0.535	0.661	0.961
0.05	0.316	0.54	0.628	0.762	0.978
0.10	0.427	0.66	0.729	0.854	0.989
0.20	0.568	0.774	0.842	0.916	0.998

Table 4.7: *Simulated power of the wild bootstrap test of symmetry in the nonparametric regression model (1.1) for a standardized* χ_2^2*-distributed error and regression function* $m(x) = \sin(2\pi x)$.

5 Proofs

Proof of Theorem 2.1: For the sake of a transparent presentation we omit the index n in the model (1.1), i.e. $\varepsilon_i = \varepsilon_{i,n}, Y_i = Y_{i,n}$ etc. We introduce the decomposition [see Ahmad and Li (1996)]

$$\hat{I}_n - c(n) \;=\; \hat{I}_{1n} + \hat{I}_{2n},$$

where

$$\hat{I}_{1n} \;=\; (n^2a)^{-1}\sum_{i=1}^{n}\Big[k(0) - k\Big(\frac{2\varepsilon_i}{a}\Big)\Big] - c(n) \tag{5.1}$$

$$\hat{I}_{2n} \;=\; (n^2a)^{-1}\sum_{i=1}^{n}\sum_{\substack{j=1\\ i\neq j}}^{n}\Big[k\Big(\frac{\varepsilon_i-\varepsilon_j}{a}\Big) - k\Big(\frac{\varepsilon_i+\varepsilon_j}{a}\Big)\Big]. \tag{5.2}$$

A standard calculation shows $\hat{I}_{1n} = o_p(n^{-1/2})$, assumption (2.11) implies

$$E[\hat{I}_{2n}] = \frac{n-1}{n}I + o(\frac{1}{n}) = I + o(\frac{1}{\sqrt{n}}) \tag{5.3}$$

and consequently the assertion of Theorem 2.1 can be established by proving the weak convergence

$$\sqrt{n}(\hat{I}_{2n} - E[\hat{I}_{2n}]) \xrightarrow{\mathcal{D}} \mathcal{N}(0, 4\tau^2). \tag{5.4}$$

In a first step we calculate the asymptotic variance and obtain

$$E[\hat{I}_{2n}^2] \;=\; \frac{1}{n^4a^2}\sum_{i\neq j}\sum_{l\neq k}E\Big\{\Big[k\Big(\frac{\varepsilon_i-\varepsilon_j}{a}\Big) - k\Big(\frac{\varepsilon_i+\varepsilon_j}{a}\Big)\Big]\Big[k\Big(\frac{\varepsilon_l-\varepsilon_k}{a}\Big) - k\Big(\frac{\varepsilon_l+\varepsilon_k}{a}\Big)\Big]\Big\}\,. \tag{5.5}$$

The sum with different indices gives the contribution

$$\begin{aligned}
\mu_1 \;&=\; \frac{n(n-1)(n-2)(n-3)}{n^4a^2}\Big(E\Big[k\Big(\frac{\varepsilon_1-\varepsilon_2}{a}\Big) - k\Big(\frac{\varepsilon_1+\varepsilon_2}{a}\Big)\Big]\Big)^2 \\
&=\; \frac{(n-1)(n-2)(n-3)}{n^3}I^2 + o(\frac{1}{n}).
\end{aligned}$$

Secondly, there appear 4 cases where there is one index in common and all cases yield the same contribution to the asymptotic variance, which is given by

$$\begin{aligned}
\mu_2 \;&=\; E\Big\{\frac{1}{n^4a^2}\sum_{i\neq j,k}\Big[k\Big(\frac{\varepsilon_i-\varepsilon_j}{a}\Big) - k\Big(\frac{\varepsilon_i+\varepsilon_j}{a}\Big)\Big]\Big[k\Big(\frac{\varepsilon_i-\varepsilon_k}{a}\Big) - k\Big(\frac{\varepsilon_i+\varepsilon_k}{a}\Big)\Big]\Big\} \\
&=\; \frac{n(n-1)(n-2)}{n^4}\int\!\!\int\!\!\int k(u)k(w)f(x)\Big(f(x-au)f(x-aw) \\
&\quad -f(x-au)f(aw-x) - f(au-x)f(x-aw) \;+\; f(au-x)f(aw-x)\Big)\,dudwdx \\
&=\; \frac{1}{n}\int f(x)(f(x)-f(-x))^2\,dx + o\Big(\frac{1}{n}\Big)\,.
\end{aligned}$$

Finally, the sums with more than one index in common are of order $o(1/n)$ and we obtain from the representation (5.3)

$$
\begin{aligned}
\text{(5.6)} \qquad \operatorname{Var}(\hat{I}_{2n}) &= \mu_1 + 4\mu_2 - \Big(\frac{n-1}{n}\Big)^2 I^2 + o(\frac{1}{n}) \\
&= \frac{4}{n}\int f(x)(f(x)-f(-x))^2\,dx \; - \; \frac{4}{n}I^2 \; + \; o(\frac{1}{n}) \\
&= 4\tau^2 + o(\frac{1}{n})
\end{aligned}
$$

where the last equality follows from the definition of τ^2 in (2.13).

For the proof of asymptotic normality in (5.4) we note that

$$
\text{(5.7)} \qquad U_n \; = \; \hat{I}_{2n} - E[\hat{I}_{2n}] \; = \; \frac{1}{n^2}\sum_{i=1}^{n}\sum_{\substack{j=1 \\ i\neq j}}^{n} H_n(\varepsilon_i, \varepsilon_j)\,,
$$

where the kernel H_n is defined by

$$
\text{(5.8)} \qquad H_n(\varepsilon_i, \varepsilon_j) \; = \; \frac{1}{a}\Big\{k(\frac{\varepsilon_i-\varepsilon_j}{a}) - k(\frac{\varepsilon_i+\varepsilon_j}{a}) - E\Big[k(\frac{\varepsilon_i-\varepsilon_j}{a}) - k(\frac{\varepsilon_i+\varepsilon_j}{a})\Big]\Big\}\,.
$$

Following Powell, Stock and Stoker (1989) we define

$$
\text{(5.9)} \qquad r_n(\varepsilon_i) = E[H_n(\varepsilon_i,\varepsilon_j) \mid \varepsilon_i]
$$

and a linear approximation of the statistic U_n by

$$
\text{(5.10)} \qquad \hat{U}_n = \frac{2}{n}\sum_{i=1}^{n} r_n(\varepsilon_i) = \sum_{i=1}^{n}\nu_i
$$

where the last line defines the random variables ν_i, i.e.

$$
\begin{aligned}
\text{(5.11)} \qquad \nu_i \; = \; \frac{2}{na}\Big\{ & E\left[\left(k\left(\frac{\varepsilon_i-\varepsilon_j}{a}\right) - k\left(\frac{\varepsilon_i+\varepsilon_j}{a}\right)\right)\Big|\varepsilon_i\right] \\
& - E\left[k\left(\frac{\varepsilon_i-\varepsilon_j}{a}\right) - k\left(\frac{\varepsilon_i+\varepsilon_j}{a}\right)\right]\Big\}
\end{aligned}
$$

A straightforward calculation gives

$$
E[H_n^2(\varepsilon_i,\varepsilon_j)] = O(\frac{1}{a}) = o(n)
$$

and Lemma 3.1 of Powell, Stock and Stoker (1989) proves

$$
\text{(5.12)} \qquad \sqrt{n}(U_n - \hat{U}_n) = o_p(1)\,.
$$

Consequently, the assertion (5.4) can be established by proving the asymptotic normality of $\hat{U}_n$, which can easily be done by checking the assumptions of Ljapunoff's theorem. Note that by the first part of this proof we have

$$
\sigma_n^2 = \operatorname{Var}(\hat{U}_n) = \frac{4}{n}\tau^2 + o\left(\frac{1}{n}\right)
$$

and a straightforward calculation gives

$$\frac{1}{\sigma_n^4}\sum_{i=1}^n E[|\nu_i|^4] \;=\; \sum_{i=1}^n \frac{O\left(\frac{1}{n^4}\right)}{\left(\frac{4}{n}\tau^2+o\left(\frac{1}{n}\right)\right)^2} E\Big[\Big(\frac{1}{a}\int k\left(\frac{\varepsilon_i - x}{a}\right) f(x)\,dx$$
$$-\frac{1}{a}\int\int k\left(\frac{x-y}{a}\right) f(x)f(y)\,dxdy \;-\; \frac{1}{a}\int k\left(\frac{\varepsilon_i + x}{a}\right) f(x)\,dx$$
$$+\frac{1}{a}\int\int k\left(\frac{x+y}{a}\right) f(x)f(y)\,dxdy\Big)^4\Big] \;=\; O(\frac{1}{n})\ .$$

This establishes the condition of Ljapunoff's theorem and implies asymptotic standard normality of $\hat{U}_n/2\sigma_n$. The assertion of the theorem now follows from (5.4), (5.7), (5.12) and Slutsky's Lemma using that under the alternative of non-symmetry it follows that $\tau^2 > 0$.

□

Proof of Theorem 3.1 and 3.2: Throughout this proof we will replace the Nadaraya-Watson estimate $\hat{m}(x_{i,n})$ in (3.1) by a slightly modified version

$$\hat{m}_i(x_{i,n}) := \frac{\sum_{\substack{l=1\\ l\neq i}}^n K\left(\frac{x_{i,n}-x_{l,n}}{\tilde{a}}\right) Y_{l,n}}{\sum_{\substack{l=1\\ l\neq i}}^n K\left(\frac{x_{i,n}-x_{l,n}}{\tilde{a}}\right)}\ , \tag{5.13}$$

which simplifies the proofs substantially. Note that the difference of both estimates is of order

$$O_p((n\tilde{a})^{-1}) = o_p((n\sqrt{a})^{-1}) = o_p(\sqrt{n}),$$

which shows that it is sufficient to prove Theorem 3.1 and 3.2 for the residuals obtained from the modified estimate.

Because both assertions in Theorem 3.1 and 3.2 are proved similary, we will restrict ourselves to the proof of Theorem 3.1. This is obtained by a reduction of the assertion to the statement (2.6) which was proved by Ahmad and Li (1996). More precisely, define

$$\begin{aligned} T_{1n} \;&=\; (n^2a)^{-1}\sum_{i=1}^n \Big[k(0) - k(\frac{2\hat{\varepsilon}_i}{a})\Big] \;-\; c(n) \\ T_{2n} \;&=\; (n^2a)^{-1}\sum_{i=1}^n \sum_{\substack{j=1\\ i\neq j}}^n \Big[k(\frac{\hat{\varepsilon}_i-\hat{\varepsilon}_j}{a}) - k(\frac{\hat{\varepsilon}_i+\hat{\varepsilon}_j}{a})\Big] \end{aligned} \tag{5.14}$$

then $T_n - c(n) = T_{1n} + T_{2n}$ and we will show

$$\begin{aligned} \hat{I}_{1n} - T_{1n} \;&=\; o_p((na^{1/2})^{-1}) \\ \hat{I}_{2n} - T_{2n} \;&=\; o_p((na^{1/2})^{-1}), \end{aligned} \tag{5.15}$$

where the statistics $\hat{I}_{1n}$ and $\hat{I}_{2n}$ are defined in (5.1) and (5.2) respectively. The first assertion of Theorem 3.1 then follows from (2.6) observing the identity $\hat{I}_n - c(n) = \hat{I}_{1n} + \hat{I}_{2n}$. The consistency of the variance estimator is then obtained from (2.8) and

$$\sigma_\varepsilon^2 - \hat{\sigma}_{\hat{\varepsilon}}^2 = o_p(1). \tag{5.16}$$

1) For a proof of the first estimate in (5.15) recall the definitions (5.1) and (5.14) and note that a Taylor expansion yields

$$\hat{I}_{1n} - T_{1n} = (n^2a)^{-1} \sum_{i=1}^{n} \Big[k(\frac{2\hat{\varepsilon}_i}{a}) - k(\frac{2\varepsilon_i}{a}) \Big] = \sum_{s=1}^{m} \frac{1}{s!} A_s + \frac{1}{(m+1)!} A_{m+1} \tag{5.17}$$

where

$$\begin{aligned} A_s &= (n^2a)^{-1} \sum_{i=1}^{n} k^{(s)}(\frac{2\varepsilon_i}{a}) \Big[\frac{-2(\hat{m}(x_i) - m(x_i))}{a} \Big]^s , \qquad\qquad s = 1, \ldots, m \\ A_{m+1} &= (n^2a)^{-1} \sum_{i=1}^{n} k^{(m+1)} \Big(\frac{2\tilde{\varepsilon}_i}{a} \Big) \Big[\frac{-2(\hat{m}(x_i) - m(x_i))}{a} \Big]^{m+1} , \end{aligned} \tag{5.18}$$

and $\tilde{\varepsilon}_i = \varepsilon_i - \theta_i(\hat{m}(x_i) - m(x_i))$ for some $0 < \theta_i < 1$. For the estimation of the last term A_{m+1} we use a similar argument as given in the proof of Lemma 2.2 in Müller (1985) and obtain

$$\sup_{x \in [0,1]} |\hat{m}(x) - m(x)| = O\Big(\tilde{a}^r + \Big(\frac{\log n}{n\tilde{a}} \Big)^{1/2} \Big) = O\Big(\Big(\frac{\log n}{n} \Big)^{r/(2r+1)} \Big) \tag{5.19}$$

which gives [using assumption (3.11)]

$$E|A_{m+1}| = O\Big(\frac{1}{na^{m+2}} \Big(\frac{\log n}{n} \Big)^{\frac{r(m+1)}{2r+1}} \Big) = o\Big(\frac{1}{n\sqrt{a}} \Big).$$

Similary, we obtain for the expectation of $|A_1|, \ldots, |A_m|$

$$E|A_s| = O\Big(\Big(\frac{\log n}{n} \Big)^{\frac{rs}{2r+1}} \frac{1}{na^s} \Big) = o\Big(\frac{1}{n\sqrt{a}} \Big),$$

where the last equality follows again from the assumption (3.11). Combining these estimates we have from Markov's inequality

$$A_s = o_p\Big(\frac{1}{n\sqrt{a}} \Big) \; ; \qquad s = 1, \ldots, m+1$$

and observing (5.17) the first estimate in (5.15) follows.

2) For a proof of the second assertion in (5.15) we introduce the decomposition

$$\hat{I}_{2n} - T_{2n} = D_{1n} - D_{2n} \tag{5.20}$$

where

$$\begin{aligned} D_{1n} &= (n^2a)^{-1} \sum_{i=1}^{n} \sum_{\substack{j=1 \\ i \neq j}}^{n} \Big[k(\frac{\hat{\varepsilon}_i + \hat{\varepsilon}_j}{a}) - k(\frac{\varepsilon_i + \varepsilon_j}{a}) \Big] \\ D_{2n} &= (n^2a)^{-1} \sum_{i=1}^{n} \sum_{\substack{j=1 \\ i \neq j}}^{n} \Big[k(\frac{\hat{\varepsilon}_i - \hat{\varepsilon}_j}{a}) - k(\frac{\varepsilon_i - \varepsilon_j}{a}) \Big]. \end{aligned} \tag{5.21}$$

We will show in the following that

$$D_{jn} = o_p\Big(\frac{1}{n\sqrt{a}}\Big) \ ; \qquad j = 1,2 \tag{5.22}$$

which completes the proof of the second part of (5.15). To this end we note that both terms are treated similary and for the sake of brevity only the case $j = 1$ is considered. Recalling the definition of D_{1n} in (5.21) we obtain by means of a Taylor expansion

$$D_{1n} \ = \ \sum_{s=1}^{m} \frac{1}{s!}\, B_s + \frac{1}{(m+1)!}\, B_{m+1}, \tag{5.23}$$

where

$$B_s \ = \ (n^2 a^{s+1})^{-1} \sum_{i=1}^{n} \sum_{\substack{j=1 \\ i\neq j}}^{n} k^{(s)}(\frac{\varepsilon_i + \varepsilon_j}{a})(m(x_i) - \hat{m}(x_i) + m(x_j) - \hat{m}(x_j))^s \tag{5.24}$$

$(s = 1, \ldots, m)$ and

$$B_{m+1} \ = \ (n^2 a^{m+2})^{-1} \sum_{i=1}^{n} \sum_{\substack{j=1 \\ i\neq j}}^{n} k^{(m+1)}(\psi_{ij})(m(x_i) - \hat{m}(x_i) + m(x_j) - \hat{m}(x_j))^{m+1} \tag{5.25}$$

with

$$\psi_{ij} = \frac{\varepsilon_i + \varepsilon_j + \theta_{ij}(-\hat{m}(x_i) + m(x_i) - \hat{m}(x_j) + m(x_j))}{a} \ ,$$

where $0 < \theta_{ij} < 1$ $(i,j = 1,\ldots,n)$. We prove (5.22) by showing that all coefficients in this Taylor expansion (5.23) are of the appropriate order, i.e.

$$B_t = o_p\left(\frac{1}{na^{1/2}}\right) \ , \qquad t = 1,\ldots,m,m+1 \ . \tag{5.26}$$

To this end we consider exemplarily the random variable B_1, which reduces by symmetry to the statistic

$$B_1 = 2(n^2a^2)^{-1} \sum_{i=1}^{n} \sum_{\substack{j=1 \\ i\neq j}}^{n} k'(\frac{\varepsilon_i + \varepsilon_j}{a})(m(x_i) - \hat{m}(x_i)) = \tilde{B}_1(1 + O(1)) \tag{5.27}$$

where $\tilde{B}_1$ is obtained from B_1 by replacing the density estimate

$$\hat{p}_i(x_i) = \frac{1}{(n-1)\tilde{a}} \sum_{j\neq i} K\Big(\frac{x_i - x_j}{\tilde{a}}\Big)$$

by the true density $p(x_i)$; i.e.

$$\tilde{B}_1 = \frac{2}{n^2(n-1)\tilde{a}a^2} \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j\neq i}}^{n} \sum_{\substack{l=1 \\ l\neq i}}^{n} K\left(\frac{x_i - x_l}{\tilde{a}}\right)(m(x_i) - Y_l)k'\left(\frac{\varepsilon_i + \varepsilon_j}{a}\right)\frac{1}{p(x_i)} \, . \tag{5.28}$$

Obviously it is sufficient to show (5.26) (in the case $t = 1$) for the random variable $\tilde{B}_1$, which will be done by a straightforward but tedious calculation of the expectation and variance of $\tilde{B}_1$. Using the representation $Y_l = m(x_l) + \varepsilon_l$ and the independence of the errors we have

$$E[\tilde{B}_1] = E_1 + E_2$$

where

$$E_1 = \frac{2}{n^2(n-1)\tilde{a}a^2} \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j\neq i}}^{n} \sum_{\substack{l=1 \\ l\neq i}}^{n} K(\frac{x_i - x_l}{\tilde{a}})\{m(x_i) - m(x_l)\} \frac{1}{p(x_i)} E[k'(\frac{\varepsilon_i + \varepsilon_j}{a})]$$

$$E_2 = \frac{2}{n^2(n-1)\tilde{a}a^2} \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j\neq i}}^{n} K(\frac{x_i - x_j}{\tilde{a}}) \frac{1}{p(x_i)} E[k'(\frac{\varepsilon_i + \varepsilon_j}{a})\varepsilon_j]$$

Note that the symmetry of the kernel k yields

$$E[k'\left(\frac{\varepsilon_i + \varepsilon_j}{a}\right)] = 0 ,$$

which gives $E_1 = 0$. Moreover, observing that $\int_{\mathbb{R}} k'(u)du = 0$ a straightforward calculation shows $E_2 = O(n^{-1})$ and it follows that

$$E[\tilde{B}_1] = O\Big(\frac{1}{n}\Big) = o\Big(\frac{1}{n\sqrt{a}}\Big). \tag{5.29}$$

The calculation of the second moment of $\tilde{B}_1$ is more complicated, but nevertheless a straightforward manner. Note that

$$\tilde{B}_1^2 = \sum_{i=1}^{4} B_{1i} \tag{5.30}$$

where

$$B_{11} = \frac{4}{n^4(n-1)^2\tilde{a}^2a^4} \sum K(\frac{x_{i_1} - x_l}{\tilde{a}}) K(\frac{x_{i_2} - x_k}{\tilde{a}}) k'(\frac{\varepsilon_{i_1} + \varepsilon_{j_1}}{a}) k'(\frac{\varepsilon_{i_2} + \varepsilon_{j_2}}{a}) \times \frac{\varepsilon_k}{p(x_{i_1})} \frac{\varepsilon_l}{p(x_{i_2})} \tag{5.31}$$

$$B_{12} = \frac{-4}{n^4\tilde{a}^2(n-1)^2a^4} \sum K(\frac{x_{i_1} - x_l}{\tilde{a}}) K(\frac{x_{i_2} - x_k}{\tilde{a}}) k'(\frac{\varepsilon_{i_1} + \varepsilon_{j_1}}{a}) k'(\frac{\varepsilon_{i_2} + \varepsilon_{j_2}}{a}) \times \frac{(m(x_{i_2}) - m(x_k))}{p(x_{i_1})} \frac{\varepsilon_l}{p(x_{i_2})} \tag{5.32}$$

$$B_{13} = \frac{-4}{n^4\tilde{a}^2(n-1)^2a^4} \sum K(\frac{x_{i_1} - x_l}{\tilde{a}}) K(\frac{x_{i_2} - x_k}{\tilde{a}}) k'(\frac{\varepsilon_{i_1} + \varepsilon_{j_1}}{a}) k'(\frac{\varepsilon_{i_2} + \varepsilon_{j_2}}{a}) \times \frac{(m(x_{i1}) - m(x_l))}{p(x_{i_1})} \frac{\varepsilon_k}{p(x_{i_2})} \tag{5.33}$$

$$B_{14} = \frac{4}{n^4\tilde{a}^2(n-1)^2a^4} \sum K(\frac{x_{i_1} - x_l}{\tilde{a}}) K(\frac{x_{i_2} - x_k}{\tilde{a}}) k'(\frac{\varepsilon_{i_1} + \varepsilon_{j_1}}{a}) k'(\frac{\varepsilon_{i_2} + \varepsilon_{j_2}}{a}) \times \frac{(m(x_{i1}) - m(x_l))}{p(x_{i_1})} \frac{(m(x_{i2}) - m(x_k))}{p(x_{i_2})} \tag{5.34}$$

where the summation is performed over all indices $i_1, j_1, i_2, j_2, l, k \in \{1, \dots, n\}$ such that $i_1 \neq j_1, i_2 \neq j_2, l \neq i_1, k \neq i_2$. Observing that for $\{l, k\} \not\subset \{i_1, j_1, i_2, j_2\}$ the expectation of the corresponding terms in the sum (5.31) is zero (see the above reasoning) it follows that we only have to consider the cases where $\{l, k\} \subset \{i_1, j_1, i_2, j_2\}$ for the calculation of $E[B_{11}]$. We consider exemplarily the case $l = i_2, k = i_1$ and all other indices pairwise different, for which we obtain the expectation

$$\begin{aligned}
&\frac{4}{n^4(n-1)^2\tilde{a}^2a^4} \sum_{i_1\neq j_1\neq i_2\neq j_2} K(\frac{x_{i_1}-x_{i_2}}{\tilde{a}})K(\frac{x_{i_2}-x_{i_1}}{\tilde{a}}) \\
&\qquad \times \frac{1}{p(x_{i_1})}\frac{1}{p(x_{i_2})}E\Big[k'(\frac{\varepsilon_{i_1}+\varepsilon_{j_1}}{a})k'(\frac{\varepsilon_{i_2}+\varepsilon_{j_2}}{a})\varepsilon_{i_2}\varepsilon_{i_1}\Big] \\
&\qquad = O\Big(\frac{1}{n^2\tilde{a}a^4}\Big)\Big(E[k'\Big(\frac{\varepsilon_{i_1}+\varepsilon_{i_2}}{a}\Big)\varepsilon_{i_1}]\Big)^2 \\
&\qquad = O\Big(\frac{1}{n^2\tilde{a}a^2}\Big)[\int\int k'(u)f(au-y)yf(y)dudy]^2 \;=\; O\Big(\frac{1}{n^2\tilde{a}}\Big) \;=\; o\Big(\frac{1}{n^2a}\Big)\ ,
\end{aligned}$$

where we have used the symmetry of the kernel k (which implies $\int k'(u)du = 0$) and assumption (3.11). The remaining cases are similar and are even of smaller order, which yields

$$E[B_{11}] = o\Big(\frac{1}{n^2a}\Big)\ . \tag{5.35}$$

The terms B_{12}, B_{13} and B_{14} can be treated by the same arguments and are also of order $o(n^{-2}a^{-1})$. Consequently we obtain from (5.35), (5.29) and (5.30)

$$\tilde{B}_1 = o_p\Big(\frac{1}{n\sqrt{a}}\Big)\ ,$$

which implies

$$B_1 = o_p\Big(\frac{1}{n\sqrt{a}}\Big).$$

A similar argument gives the same order for the terms $B_2, \dots, B_m$. Finally we investigate the random variable B_{m+1} in (5.25) using Lemma 2 in Müller (1985) (with an obvious modification for the Nadaraya-Watson estimator) and obtain

$$E|B_{m+1}| = O\Big(\Big(\frac{\log n}{n}\Big)^{\frac{r(m+1)}{2r+1}}\frac{1}{a^{m+2}}\Big) = o\Big(\frac{1}{n\sqrt{a}}\Big)$$

by assumption (3.11) and Markov's inequality gives $B_{m+1} = o_p(n^{-1}a^{-1/2})$. The second assertion in (5.15) now follows from the representation (5.20), (5.23) and a similar argument for the statistic D_{2n}.

3) The proof of the remaining assertion (5.16) follows by similar arguments. To be precise, recall the definition of σ_ε^2 and $\sigma_{\hat{\varepsilon}}^2$ in (2.8) and (3.12), respectively, then

$$\begin{aligned}
(\int k^2(u)du)^{-1}(\sigma_\varepsilon^2 - \sigma_{\hat{\varepsilon}}^2) &= \int_{\mathbb{R}} f_n(x)dF_n(x) - \int_{\mathbb{R}} \hat{f}_n(x)d\hat{F}_n(x) \\
&= \frac{1}{n^2a}\sum_{i,j=1}^{n} k\Big(\frac{\varepsilon_i-\varepsilon_j}{a}\Big) - k\Big(\frac{\hat{\varepsilon}_i-\hat{\varepsilon}_j}{a}\Big) = -D_{2n},
\end{aligned}$$

where D_{2n} is defined in (5.21). From (5.22) we have $D_{2n} = o_p(n^{-1}a^{-1/2})$ and the consistency of $\sigma^2_{\hat{\varepsilon}}$ now follows from (2.8).

□

Acknowledgements. This work was done while H. Dette was visiting the Department of Statistics, Purdue University, and this author would like to thank the department for its hospitality. The work of H. Dette was supported by the Deutsche Forschungsgemeinschaft (De 502/14-1, SFB 475, Komplexitätsreduktion in multivariaten Datenstrukturen). The authors would also like to thank I. Gottschlich, who typed parts of this paper with considerable technical expertise.

References

I.A. Ahmad, Q. Li (1996). Testing symmetry of an unknown density function by kernel method. Nonparam. Statistics 7, 279-293.

A. Azzalini, A. Bowman (1993). On the use of nonparametric regression for checking linear relationships. J. Roy. Statist. Soc. Ser. B 55, 549-559.

L. Baringhaus, N. Henze (1992). A characterization of and new consistent tests of symmetry. Commun. Statist. 21, 1111-1125.

J.O. Berger, M. Delampady (1987). Testing precise hypotheses (with discussion). Statist. Science 2, 317-352.

P.J. Bickel (1982). On adaptive estimation. Ann. Statist. 10, 647-671.

A. Cabaña, M. Cabaña (2000). Tests of symmetry based on transformed empirical processes. Canad. J. Statist. 28, 829-839.

H. Dette, A. Munk (1998). Testing heteroscedasticity in nonparametric regression. J. Roy. Statist. Soc. Ser. B 60, 693-708.

Y. Fan, R. Gencay (1995). A consistent nonparametric test of symmetry in linear regression models. J. Amer. Statist. Assoc. 90, 551-557.

Y. Fan, I. Gijbels (1996). Local Polynomial Modelling and its Applications. Chapman and Hall, London.

T. Gasser, H.G. Müller (1984). Estimating regression functions and their derivatives by the kernel method. Scand. J. Statist. 11, 171-185.

T. Gasser, H.G. Müller, V. Mammitzsch (1985). Kernels for nonparametric curve estimation. J. Roy. Statist. Soc. Ser. B 47, 238-252.

W. Gonzalez-Manteiga, R. Cao Abad (1993). Testing hypothesis of general linear model using nonparametric regression estimation. Test 2, 161-189.

W. Härdle (1990). Applied Nonparametric Regression. Econometric Society Monographs 19, Cambridge University Press.

W. Härdle, E. Mammen (1993). Comparing nonparametric versus parametric regression fits. Ann. Statist. 21, 1926-1947.

P. Hall, J.W. Kay, D.M. Titterington (1990). Asymptotically optimal difference — based estimation of variance in nonparametric regression. Biometrika 77, 521-528.

P. Hall, J.S. Marron (1990). On variance estimation in nonparametric regression. Biometrika 77, 415-419.

V. Hjellvik, O. Yao, D. Tjøstheim (1998). Linearity testing using local polynomial approximation. J. Statist. Plann. Inf. 68, 295-321.

M. Holland (1988). Testing for symmetry. Encycl. Statist. Sci. 9 (eds. N.L. Johnson, S. Kotz), 211-216.

M. Huškova (1984). Hypothesis of symmetry. In: Handbook of Statistics, Vol. 4, Nonparametric Methods, 63-78. North Holland, Amsterdam.

K.B. Kulasekera, J. Wang (2001). A test of equality of regression functions using Gâteaux scores. Australian and New Zealand Journal of Statistics 43, 89-100.

H.G. Müller (1985). Kernel estimators of zeros and of location and size of extrema of regression functions. Scand. J. Statist. 12, 221-232.

E.A. Nadaraya (1964). On estimating regression. Theor. Prob. Appl. 9, 141-142.

J.L. Powell, J.H. Stock, T.M. Stoker (1989). Semiparametric estimation of index coefficients. Econometrica 57, 1403-1430.

M. Rosenblatt (1956). Remarks on some nonparametric estimates of the density function. Ann. Math. Statist. 27, 832-837.

M. Rosenblatt (1975). A quadratic measure of deviation and a test of independence. Ann. Statist. 3, 1-14.

J. Sacks, D. Ylvisaker (1970). Designs for regression problems with correlated errors III. Ann. Math. Statist 41, 2057 - 2074

T. Sellke, M.J. Bayarri, J.O. Berger (2001). Calibration of p-values for testing precise null hypotheses. The Amer. Statistician 55, 62-71.

M.P. Wand and M.C. Jones (1995). Kernel Smoothing. Chapman and Hall, London.

G.S. Watson (1964). Smooth regression analysis. Sankhya, Ser. A, 26, 359-372.