KYBERNETIKA — VOLUME 45 (2009), NUMBER 6, PAGES 960 – 971

GOODNESS–OF–FIT TESTS FOR PARAMETRIC REGRESSION MODELS BASED ON EMPIRICAL CHARACTERISTIC FUNCTIONS

Marie Hušková and Simos G. Meintanis

Dedicated to the memory of Sándor Csörgő

Test procedures are constructed for testing the goodness-of-fit in parametric regression models. The test statistic is in the form of an L2 distance between the empirical characteristic function of the residuals in a parametric regression fit and the corresponding empirical characteristic function of the residuals in a non-parametric regression fit. The asymptotic null distribution as well as the behavior of the test statistic under contiguous alternatives is investigated. Theoretical results are accompanied by a simulation study.

Keywords: empirical characteristic function, kernel regression estimators

AMS Subject Classification: 62J05, 62F05

1. INTRODUCTION

Assume the general model

$$Y = m(X) + \sigma(X)\varepsilon, \tag{1.1}$$

where $m(\cdot)$ and $\sigma(\cdot)$ denote a unspecified function of the regressor X and an unspecified variance function, respectively, and the error ε is assumed to have a distribution function (DF) F, with mean zero and unit variance. The error ε and the regressor X are assumed to be independent.

On the basis of independent observations $\{Y_j, X_j\},\ j = 1, 2, \ldots, n$, we wish to test the null hypothesis

$$H_0 : m(\cdot) = m(\cdot, \vartheta),\ \text{for some } \vartheta \in \boldsymbol{\Theta} \subseteq \mathbb{R}^p,$$

where $m(X, \vartheta)$ denotes a specified regression function of the regressor X.

Write $e_j = Y_j - \widehat{m}(X_j)$ for the residuals resulting from a non-parametric fit $\widehat{m}(\cdot)$ and $\widehat{\sigma}^2_{nj} := \widehat{\sigma}^2_n(X_j)$, for the corresponding estimate of the variance $\sigma^2(\cdot)$. Also let $u_j = Y_j - m(X_j, \widehat{\vartheta}_n)$, where $\widehat{\vartheta}_n$ is an estimator of ϑ_0, be the residuals corresponding to a parametric fit suggested by the null hypothesis H_0.

There exist a variety of tests for the null hypothesis H_0. Some of them are based on a distance between a nonparametric and the parametric fit corresponding to the null hypothesis, while others use only the parametric regression fit and reject the null hypothesis if the quality of this fit is not good. Works of this nature include Eubank et al. [4], Härdle and Mammen [7], Stute et al. [15], and Whang [18]. There are also many tests which are moment-based, often utilizing the distance between two variance estimators; see for instance Li and Wang [12], Fan and Huang [5] and Dette [3].

Since regression estimation, either in the nonparametric setting or in the context of parametric regression, induces a corresponding estimate of the residual DF, it is only natural to consider a classical test statistic, such as those based on the empirical DF, which measure the distance between a pair of empirical DF's; namely, one empirical DF is constructed based on the nonparametric residuals while the other incorporates the parametric residuals corresponding to a parametric fit. Thus the procedure is reminiscent of a two-sample test. In fact van Keilegom et al. [16] have considered this possibility, and constructed Kolmogorov–Smirnov and Cramér–von Mises type statistics for testing a parametric regression model.

The present work is in the spirit of van Keilegom et al. [16], but in view of the uniqueness of the characteristic function (CF), instead of using the empirical DF we employ the empirical CF. Specifically, the proposed test statistic incorporates the empirical CF

$$\phi_n(t) = \frac{1}{n}\sum_{j=1}^{n} e^{it\widehat{e}_j} \quad \text{and} \quad \varphi_n(t) = \frac{1}{n}\sum_{j=1}^{n} e^{it\widehat{u}_j},$$

of the standardized residuals $\widehat{e}_j = e_j/\widehat{\sigma}_{nj}$ and $\widehat{u}_j = u_j/\widehat{\sigma}_{nj}$, $j = 1, 2, \ldots, n$, and takes the form

$$T_{n,w} = n\int_{-\infty}^{\infty} |\phi_n(t) - \varphi_n(t)|^2 w(t)\,\mathrm{d}t, \tag{1.2}$$

where $w(\cdot)$ denotes a weight function which is assumed to satisfy

$$w(t) = w(-t) \geq 0 \quad \text{and} \quad 0 < \int_{-\infty}^{\infty} w(t)\,\mathrm{d}t < \infty. \tag{1.3}$$

Rejection of the null hypothesis is for large values of $T_{n,w}$.

Motivation for using the empirical CF stems from earlier studies suggesting that nonparametric procedures based on the empirical CF may prove competitive to corresponding methods based on the empirical DF; cf. Hušková and Meintanis [9], Meintanis [13], Bondell [2], Bilodeau and Lafaye de Micheaux [1], Gupta et al. [6], Henze et al. [8], and Kankainen and Ushakov [11].

Although $w(t)$ figuring in (1.2), may be any weight function satisfying (1.3), priority is given to functions that render the corresponding test statistic in closed form. In order to identify such functions notice that starting from (1.2) and after straightforward algebra one arrives at

$$T_{n,w} = \frac{1}{n} \sum_{j,k=1}^{n} \left\{ I_w\left(\widehat{e}_j - \widehat{e}_k\right) + I_w\left(\widehat{u}_j - \widehat{u}_k\right) - 2I_w\left(\widehat{e}_j - \widehat{u}_k\right) \right\}, \tag{1.4}$$

where

$$I_w(\beta) = \int_{-\infty}^{\infty} \cos(\beta t) w(t)\,\mathrm{d}t.$$

Therefore, with computational simplicity in mind, we suggest $w(t) = e^{-\gamma|t|}$ and $w(t) = e^{-\gamma t^2}$, $\gamma > 0$ as weight functions. In fact these functions lead to interesting limit statistic as $\gamma \to \infty$. To see this let $w(t) = e^{-\gamma|t|}$ in (1.4) and after some manipulation write the resulting test statistic, say $T_{n,\gamma}$, as

$$T_{n,\gamma} = \int_0^{\infty} g(t) e^{-\gamma|t|}\,\mathrm{d}t,$$

where

$$g(t) = 2n \left\{ [C_n(t) - U_n(t)]^2 + [S_n(t) - V_n(t)]^2 \right\},$$

with $C_n(t) = n^{-1}\sum_{j=1}^{n} \cos(t\widehat{e}_j)$, $U_n(t) = n^{-1}\sum_{j=1}^{n} \cos(t\widehat{u}_j)$, $S_n(t) = n^{-1}\sum_{j=1}^{n} \sin(t\widehat{e}_j)$, and $V_n(t) = n^{-1}\sum_{j=1}^{n} \sin(t\widehat{u}_j)$ being the real and imaginary parts of $\phi_n(t)$ and $\varphi_n(t)$, respectively. Using the expansions $\cos(x) = 1 - x^2/2 + o(x^2)$ and $\sin(x) = x - x^3/6 + o(x^3)$ one obtains

$$g(t) = 2n\left(\bar{\widehat{e}} - \bar{\widehat{u}}\right)^2 t^2 + o(t^2), \;\; t \to 0 \quad \text{a.s.},$$

where $\bar{\widehat{e}} = n^{-1}\sum_{j=1}^{n} \widehat{e}_j$, $\bar{\widehat{u}} = n^{-1}\sum_{j=1}^{n} \widehat{u}_j$. Consequently an Abelian theorem for Laplace transforms (see Zayed [19]), § 5.11) gives

$$\lim_{\gamma\to\infty} \gamma^3 T_{n,\gamma} = 4n\left(\bar{\widehat{e}} - \bar{\widehat{u}}\right)^2 \quad \text{a.s.}$$

Hence by letting $\gamma \to \infty$, the test statistic is reduced to a comparison between the sample means of the residuals, properly standardized. An analogous argument with $w(t) = e^{-\gamma t^2}$ yields the limit statistic

$$\lim_{\gamma\to\infty} \gamma^{3/2} T_{n,\gamma} = n\frac{\sqrt{\pi}}{2}\left(\bar{\widehat{e}} - \bar{\widehat{u}}\right)^2 \quad \text{a.s.}$$

2. ASYMPTOTIC RESULTS

Asymptotic theory, both under H_0 and under alternatives, is facilitated by the following representation for the test statistic:

$$T_{n,w} = \int_{-\infty}^{\infty} Z_n^2(t) w(t)\,\mathrm{d}t, \tag{2.1}$$

where

$$Z_n(t) = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \{\cos(t\widehat{e}_j) + \sin(t\widehat{e}_j) - \cos(t\widehat{u}_j) - \sin(t\widehat{u}_j)\}. \tag{2.2}$$

The results on the limit behavior of the test statistics $T_{n,w}$ are presented here together with basic discussion.

We start with the assumptions on the model. We assume that $(X_1, Y_1), \dots, (X_n, Y_n)$ are independent identically distributed (i.i.d.) random vectors such that

$$Y_j = m(X_j) + \sigma(X_j)\varepsilon_j, \quad j = 1, \dots, n, \tag{2.3}$$

where $\varepsilon_1, \dots, \varepsilon_n$, $X_1, \dots, X_n$, $m(.)$ and $\sigma(.)$ satisfy:

(A.1) Let $\varepsilon_1, \dots, \varepsilon_n$ be i.i.d. random variables with zero mean, unit variance and $\mathrm{E}\,\varepsilon_j^4 < \infty$ and characteristic function $\varphi(t)$, $t \in \mathbb{R}$.

(A.2) On the real and imaginary parts of $\varphi(t)$ denoted by $C(t)$ and $S(t)$, we assume that the first partial derivatives w.r.t. t exist. Specifically, the first derivatives $C'(t)$ and $S'(t)$ w.r.t. t are bounded and continuous for all t.

(A.3) $X_1, \dots, X_n$ are i.i.d. on $[0, 1]$ with common positive continuous density f_X.

(A.4) Let $(\varepsilon_1, \dots, \varepsilon_n)$ and $(X_1, \dots, X_n)$ be independent.

(A.5) Let m be a function on $[0, 1]$ with Lipschitz first derivative.

(A.6) Let $\sigma(x)$, $x \in [0, 1]$ be positive on $[0, 1]$ with Lipschitz first derivative.

(A.7) The weight function w is nonnegative and symmetric, and

$$\int_{-\infty}^{\infty} t^4 w(t)\,\mathrm{d}t < \infty.$$

Our procedure depends on the estimators of unknown parameters $m(.)$ and $\sigma(.)$. Here are the assumptions related to the kernel $K(\cdot)$ and the bandwidth $h = h_n$ involved in the estimation of $m(.)$ and $\sigma(.)$.

(A.8) Let K be a symmetric twice continuously differentiable density on $[-1, 1]$ with $K(-1) = K(1) = 0$.

(A.9) Let $\{h_n\}$ be a sequence of the bandwidth such that $\lim_{n\to\infty} nh_n^2 = \infty$ and $\lim_{n\to\infty} nh_n^{3+\delta} = 0$ for some $\delta > 0$.

We use the following estimators of the density function $f_X(.)$ of X_j's, the regression function $m(.)$ and the variance function $\sigma^2(.)$:

$$\widehat{f}_X(x) = \frac{1}{nh} \sum_{j=1}^{n} K((X_j - x)/h), \quad x \in [0, 1], \tag{2.4}$$

$$\widehat{m}_n(x) = \frac{1}{nh\widehat{f}_X(x)} \sum_{j=1}^{n} K((X_j - x)/h) Y_j, \quad x \in [0,1], \tag{2.5}$$

$$\widehat{\sigma}_n^2(x) = \frac{1}{nh\widehat{f}_X(x)} \sum_{j=1}^{n} K((X_j - x)/h)(Y_j - \widehat{m}_n(x))^2, \quad x \in [0,1]. \tag{2.6}$$

As an estimator of ϑ we use the least squares estimator $\widehat{\vartheta}_n$ defined as a solution of minimization problem

$$\min_{\vartheta \in \boldsymbol{\Theta}} \sum_{j=1}^{n} \big(Y_j - m(X_j, \vartheta)\big)^2.$$

The properties of these estimators were studied, e. g. by Wu [21], White [19, 20], Seber and Wild [17]. We apply here the properties derived in van Keilegom et al. [16].

We use the following notation

$$\boldsymbol{B}_\vartheta = \Big\{ \mathrm{E}\Big(\frac{\partial m(X_1, \vartheta)}{\partial \vartheta_r} \frac{\partial m(X_1, \vartheta)}{\partial \vartheta_s} \Big) \Big\}_{r,s=1,\dots,p},$$

$$\boldsymbol{b}(x, \vartheta) = \Big(\frac{\partial m(x, \vartheta)}{\partial \vartheta_1}, \dots, \frac{\partial m(x, \vartheta)}{\partial \vartheta_p} \Big)^T$$

and

$$\boldsymbol{D}(x, \vartheta) = \Big(\frac{\partial^2 m(x, \vartheta)}{\partial \vartheta_s \partial \vartheta_v} \Big)_{s,v=1,\dots,p}.$$

Concerning the assumptions on $m(\cdot, \vartheta)$ we use those considered in van Keilegom et al. [16]):

(A.10) $\boldsymbol{\Theta}$ is a compact subspace of $\mathbb{R}^p$, the true value of parameter ϑ_0 is an interior point of $\boldsymbol{\Theta}$.

(A.11) $m(x, \vartheta)$ is twice differentiable with respect to ϑ_0 for all x. There is an integrable function $H(\cdot)$ such that $m^2(x, \theta) \le H(x)$ for all θ and all x. For all $\epsilon > 0$, $\inf_{||\theta - \theta_0|| \ge \epsilon} \mathrm{E}\,(m(X_1, \theta) - m(X_1, \theta_0))^2 > 0$.

(A.12) $\mathrm{E}\Big(||\boldsymbol{b}(X_1, \vartheta_0)||^2 \Big) < \infty$, $\mathrm{E}\Big(||\boldsymbol{D}(X_1, \vartheta_0)||^2 \Big) < \infty$ and $\boldsymbol{B}_{\vartheta_0}$ is non-singular.

(A.13) $||m(x, \vartheta_0) - m(x, \vartheta_1)|| + ||\boldsymbol{b}(x, \vartheta_0) - \boldsymbol{b}(x, \vartheta_1)|| + ||\boldsymbol{D}(x, \vartheta_0) - \boldsymbol{D}(x, \vartheta_1)|| \le h(x) ||\vartheta_1 - \vartheta_0||$ for all $\vartheta_1 \in \boldsymbol{\Theta}$ with $\mathrm{E}\, h^2(X_1) < \infty$.

Theorem 1. Let assumptions (A.1)–(A.13) be satisfied. Then under H_0, as $n \to \infty$,

$$T_{n,w} \xrightarrow{d} Z_{\vartheta_0}^2 \int (t(C(t) - S(t)))^2 w(t)\, \mathrm{d}t,$$

where Z_{ϑ_0} is a normally distributed random variable with zero mean and variance

$$\operatorname{var} Z_{\vartheta_0} = \mathrm{E}\Big(1 - \sigma(X_1)\boldsymbol{b}^T(X_1,\vartheta_0)\,\boldsymbol{B}_{\vartheta_0}^{-1}\,\mathrm{E}(\sigma(X_1)^{-1}\boldsymbol{b}(X_1,\vartheta_0)\Big)^2.$$

P r o o f. It is postponed to Section 5. □

Next we present the result under the local alternatives:

$$H_1 : m(x) = m(x,\vartheta_0) + \frac{1}{\sqrt{n}} r(x), \quad x \in [0,1] \tag{2.7}$$

where $r(.)$ satisfies:

$$\mathrm{E}\, r^2(X_1) < \infty, \quad \mathrm{E}\left(|r(X_1)|\,||\boldsymbol{b}^T(X_1,\vartheta_0)||\,||\boldsymbol{B}_{\vartheta_0}||\right) < \infty. \tag{2.8}$$

Theorem 2. Let assumptions (A.1) – (A.13) and (2.8) be satisfied. Then under H_1, as $n \to \infty$,

$$T_{n,w} \xrightarrow{d} (Z_{\vartheta_0} + d)^2 \int (t(C(t) - S(t)))^2 w(t)\,\mathrm{d}t,$$

where Z_{ϑ_0} defined in Theorem 1 and

$$d = -\mathrm{E}\Big(\sigma(X_1)^{-1}\boldsymbol{b}^T(X_1,\vartheta_0)\Big)\,\boldsymbol{B}_{\vartheta_0}^{-1}\,\mathrm{E}\Big(r(X_1)\boldsymbol{b}(X_1,\vartheta_0)\Big) + \mathrm{E}\Big(\sigma(X_1)^{-1}r(X_1)\Big).$$

P r o o f. It is postponed to Section 5. □

The explicit form of the limit distribution of $T_{n,w}$ is unknown even under the null hypothesis. It depends on the hypothetical distribution of the error terms and also on the density f_X of X_i's, the functions $m(.)$ and $\sigma(.)$ It does not depend on the kernel $K(.)$ and the bandwidth h_n. But this is in accordance with the results of van Keilegom et al. [16] for the procedures based on the empirical DF. Therefore the limit distribution does not provide an approximation for the critical values. However, a special parametric bootstrap which is discussed later can be used to provide these critical values.
Notice that no smoothness of the distribution function of ε_i's is assumed.

3. SIMULATION RESULTS

In this section the performance of the proposed test statistic is investigated in finite samples. We consider $y = m(x)+\sigma(x)\varepsilon$, with $m(x) = \vartheta x$, $\sigma(x) = |x|$ and $\varepsilon \sim N(0,1)$, as our reference model under the null hypothesis H_0, and $\mathrm{Trg}(\vartheta) : m(x) = \sin(\pi\vartheta x)$, $\mathrm{Sqt}(\vartheta) : m(x) = \vartheta\sqrt{x}$, $\mathrm{Exp}(\vartheta) : m(x) = \vartheta e^{\vartheta x}$, and $\mathrm{Log}(\vartheta) : m(x) = \vartheta \log(x)$, as alternatives. The design points X_j, $j = 1,2,\ldots,n$, are taken as i.i.d. following the uniform distribution on $(0,1)$.

3.1. Test statistics

We consider the proposed test with $w(t) = e^{-\gamma|t|}$, denoted by $T_\gamma^{(1)}$, and with $w(t) = e^{-\gamma t^2}$, denoted by $T_\gamma^{(2)}$. The corresponding test statistics may be computed from (1.4) using $I_w(\beta) = 2\gamma/(\beta^2+\gamma^2)$, and $I_w(\beta) = \sqrt{\pi/\gamma}e^{-\beta^2/4\gamma}$, respectively. For comparison purposes, results on a Kolmogorov–Smirnov (KS) type, and a Cramér–von Mises type statistic (CM), are also reported. These tests utilize the discrepancy between the empirical DF of the nonparametric standardized residuals and the corresponding empirical DF of the parametric standardized residuals, and are defined by

$$\mathrm{KS} = \sqrt{n}\sup_y |F_{\widehat{e}}(y) - F_{\widehat{u}}(y)|,$$

and

$$\mathrm{CM} = n\int [F_{\widehat{e}}(y) - F_{\widehat{u}}(y)]^2 \, \mathrm{d}F_{\widehat{e}}(y),$$

where $F_{\widehat{e}}$ denotes the empirical DF based on $\widehat{e}_j = e_j/\widehat{\sigma}_{nj}$ and $F_{\widehat{u}}$ denotes the empirical DF based on $\widehat{u}_j = u_j/\widehat{\sigma}_{nj},\ j = 1, 2, \ldots, n$. For all test statistics, non-parametric estimation of $m(\cdot)$ employed the Gaussian kernel with bandwidth h. We carried out the tests with fixed $h := h_0$, where $h_0 = 0.5$ (resp. $h_0 = 0.8$) was used for the CF and CM statistics (resp. for the KS statistic). These choices are by no means unique, and other values could have also been used, but serve to match the nominal size amongst tests, and hence facilitate comparison.

3.2. Bootstrap statistics

For each test statistic, say T, we apply the smooth bootstrap procedure; see for instance Neumeyer [14]. Under this version of the bootstrap, one computes the standardized residuals

$$\widehat{e}_j^{(s)} = \frac{\widehat{e}_j - \frac{1}{n}\sum_{l=1}^n \widehat{e}_l}{\sqrt{\frac{1}{n}\sum_{\ell=1}^n (\widehat{e}_\ell - \frac{1}{n}\sum_{l=1}^n \widehat{e}_l)^2}},\ j = 1, \ldots, n,$$

and generates $Z_j, j = 1, \ldots, n$, i.i.d. variables having distribution $K(\cdot)$ (in our case the Gaussian kernel), independently of $(X_j, Y_j),\ j = 1, \ldots, n$. Then the bootstrap resampling proceeds as follows:

(i) Randomly draw $\widehat{e}_j^{*(s)}$ with replacement from $\widehat{e}_j^{(s)},\ j = 1, \ldots, n$.

(ii) For a positive constant $\tilde{h}_n$, generate bootstrap errors $\epsilon_j^* = \widehat{e}_j^{*(s)} + \tilde{h}_n Z_j,\ j = 1, \ldots, n$.

(iii) Compute the bootstrap observations:

$$Y_j^* = m(X_j; \widehat{\vartheta}_n) + \widehat{\sigma}_{nj}\epsilon_j^*,\ j = 1, \ldots, n.$$

(iv) Based on $(X_1, Y_1^*), \ldots, (X_n, Y_n^*)$, compute the non-parametric estimates $\widehat{m}^*(\cdot)$ of $m(\cdot)$ and $\widehat{\sigma}^*_{nj} =: \widehat{\sigma}^*_n(X_j)$ of $\sigma(X_j)$, and the parametric estimate $\widehat{\vartheta}^*_n$.

(v) Compute $\widehat{e}^*_j = (Y_j^* - \widehat{m}^*(X_j))/\widehat{\sigma}^*_{nj}$ and $\widehat{u}^*_{nj} = (Y^*_{nj} - m(X_j, \widehat{\vartheta}^*_n))/\widehat{\sigma}^*_{nj}$, $j = 1, 2, \ldots, n$.

(vi) Calculate the value of the bootstrap version of the test statistic based on $\widehat{e}^*_j$ and $\widehat{u}^*_j$.

When steps (i) – (vi) are repeated a number of times, say B, the sampling distribution of T is reproduced, and on the basis of this bootstrap distribution we decide whether the observed value of the test statistic is significant.

In Table 1 and Table 2 the rejection rate for the null hypothesis is reported with samples of size $n = 50$ and $n = 75$, respectively. The nominal size of the test is 5 % and 10 %. The figures are in the form of percentage of rejection rounded to the nearest integer, and were obtained from 1000 repetitions with $B = 500$ bootstrap replications. For the smooth bootstrap we have used the value $\tilde{h}_n = n^{-1/4}$ which is among those suggested by Neumeyer [14]. All tests are seen to be somewhat conservative when estimating the nominal size, particularly for $n = 50$. Between the standard tests the KS appears to be the most powerful overall. Between the new tests, power slightly varies with respect to the weight parameter γ, but overall $T^{(2)}_\gamma$ (employing $e^{-\gamma t^2}$ as a weight function) has a slight edge over $T^{(1)}_\gamma$. Moreover $T^{(2)}_\gamma$ with a compromise value of the weight parameter, say $\gamma = 0.04$ or 0.06, is seen to compete well with the empirical DF tests under most alternative cases considered and with both sample sizes.

Table 1. Percentage of rejection for the linear null hypothesis observed in 1 000 samples of size $n = 50$ with nominal size 5 % (left entry), 10 % (right entry).

	$T^{(1)}_{0.30}$	$T^{(1)}_{0.40}$	$T^{(1)}_{0.50}$	$T^{(2)}_{0.04}$	$T^{(2)}_{0.06}$	$T^{(2)}_{0.08}$	KS	CM
$H_0(\vartheta = 0)$	3 8	3 7	3 7	3 8	3 7	3 7	3 8	3 7
Trg(0.125)	28 41	26 38	19 34	27 42	26 40	20 35	30 51	24 36
Trg(0.25)	81 90	79 89	75 87	82 92	80 90	76 87	70 86	63 82
Trg(0.50)	97 99	97 99	97 100	97 99	97 99	98 100	90 97	97 99
Sqt(0.125)	7 12	7 11	7 10	8 13	7 11	6 11	11 22	10 18
Sqt(0.25)	31 45	28 42	22 37	32 46	27 43	22 36	32 53	30 44
Sqt(0.50)	83 92	82 91	79 89	85 93	83 92	81 90	78 90	80 90
Exp(0.10)	13 22	10 19	9 16	13 24	10 19	9 18	21 37	23 37
Exp(0.125)	23 36	20 32	17 28	23 37	21 33	18 29	34 50	37 49
Exp(0.15)	37 51	33 48	29 44	38 51	34 48	30 45	46 63	48 62
Log(0.10)	10 21	11 21	11 20	12 22	13 23	12 23	16 26	14 25
Log(0.125)	14 26	15 27	15 28	16 28	18 29	18 29	20 34	21 35
Log(0.15)	18 32	19 34	21 36	18 32	21 35	22 37	26 44	27 44

Table 2. Percentage of rejection for the linear null hypothesis observed in 1 000 samples of size $n = 75$ with nominal size 5 % (left entry), 10 % (right entry).

	$T^{(1)}_{0.30}$	$T^{(1)}_{0.40}$	$T^{(1)}_{0.50}$	$T^{(2)}_{0.04}$	$T^{(2)}_{0.06}$	$T^{(2)}_{0.08}$	KS	CM
$H_0(\vartheta = 0)$	5 9	4 9	4 8	5 9	4 8	4 8	6 11	4 8
Trg(0.125)	56 70	49 65	43 58	56 70	48 65	44 58	68 82	45 59
Trg(0.25)	97 99	97 99	97 99	98 99	98 99	97 99	96 98	92 97
Trg(0.50)	100 100	100 100	100 100	100 100	100 100	100 100	100 100	100 100
Sqt(0.125)	16 23	13 21	11 18	16 23	13 21	11 18	30 42	22 33
Sqt(0.25)	65 79	61 74	49 64	64 80	59 73	50 66	73 84	60 74
Sqt(0.50)	99 99	99 100	99 100	99 100	99 100	99 100	99 100	99 99
Exp(0.10)	39 53	34 49	32 44	39 53	36 49	33 45	55 70	52 63
Exp(0.125)	61 72	57 70	53 66	61 73	57 71	53 68	74 83	68 80
Exp(0.15)	79 87	76 86	72 84	80 88	77 87	73 84	87 93	84 91
Log(0.10)	47 63	48 63	48 63	49 64	50 64	50 64	51 66	54 68
Log(0.125)	63 78	67 83	67 83	62 80	66 83	67 83	64 81	71 82
Log(0.15)	72 86	78 90	80 91	75 88	76 91	77 92	76 90	83 91

4. PROOFS

The proofs of both theorems are somehow similar to those of Theorems in Hušková and Meintanis [10]. Important part of the proofs utilize the results on the estimator $\widehat{\theta}$ by van Keilegom et al. [16], see Lemma A.1 and A.2.

Proof of Theorem 1. Notice that the residuals $\widehat{e}_j$ can be expressed as

$$\widehat{e}_j = \varepsilon_j + \varepsilon_j\Big(\frac{\sigma(X_j)}{\widehat{\sigma}_n(X_j)} - 1\Big) + \frac{m(X_j) - \widehat{m}_n(X_j)}{\widehat{\sigma}_n(X_j)}, \; j = 1, \ldots, n,$$

and then by the Taylor expansion we have

$$\cos(t\widehat{e}_j) = \cos(t\varepsilon_j) - t\sin(t\varepsilon_j)\Big(\varepsilon_j\Big(\frac{\sigma(X_j)}{\widehat{\sigma}_n(X_j)} - 1\Big) + \frac{m(X_j) - \widehat{m}_n(X_j)}{\widehat{\sigma}_n(X_j)}\Big) + t^2 R^c_{nj}(t),$$

$j = 1, \ldots, n$, where $R^c_{nj}(t)$ are remainders. Similar relations can be obtained for $\sin(t\widehat{e}_j)$'s, let us denote the respective remainder terms by $R^s_{nj}(t)$. In an analogous way we receive for the residuals $\widehat{u}_j$

$$\widehat{u}_j = \varepsilon_j + \varepsilon_j\Big(\frac{\sigma(X_j)}{\widehat{\sigma}_n(X_j)} - 1\Big) + \frac{m(X_j) - m(X_j, \widehat{\vartheta}_n)}{\widehat{\sigma}_n(X_j)}, \; j = 1, \ldots, n,$$

and also

$$\cos(t\widehat{u}_j) = \cos(t\varepsilon_j) - t\sin(t\varepsilon_j)\Big(\varepsilon_j\Big(\frac{\sigma(X_j)}{\widehat{\sigma}_n(X_j)} - 1\Big) + \frac{m(X_j) - m(X_j, \widehat{\vartheta}_n)}{\widehat{\sigma}_n(X_j)}\Big) + t^2 Q^c_{nj}(t),$$

$j = 1, \ldots, n$, where $Q^c_{nj}(t)$ are remainders. The respective remainder terms of Taylor expansions for $\sin(t\widehat{u}_j)$' are denoted by $Q^s_{nj}(t)$.
Thus under H_0 $Z_n(t)$ can be expressed as

$$\begin{aligned} Z_n(t) &= \frac{t}{\sqrt{n}} \sum_{j=1}^n (-\sin(t\varepsilon_j) + \cos(t\varepsilon_j)) \frac{\widehat{m}(X_j) - m(X_j, \widehat{\vartheta}_n)}{\widehat{\sigma}_n(X_j)} \\ &\quad + \frac{t^2}{\sqrt{n}} \sum_{j=1}^n (R^c_{nj}(t) + R^s_{nj}(t) + Q^c_{nj}(t) + Q^s_{nj}(t)). \end{aligned}$$

Since under H_0 $m(\cdot) = m(\cdot, \theta_0)$ and since the assumptions Lemma 3 in Hušková and Meintanis [10] can be applied in our setup and we have that

$$\begin{aligned} &\int \Big(\frac{t}{\sqrt{n}} \sum_{j=1}^n \big(-\sin(t\varepsilon_j) + \cos(t\varepsilon_j)\big) \frac{\widehat{m}(X_j) - m(X_j)}{\widehat{\sigma}_n(X_j)} \\ &\quad - \frac{t}{\sqrt{n}} \sum_{j=1}^n \varepsilon_j \big(C(t) - S(t)\big)\Big)^2 w(t)\,\mathrm{d}t = o_P(1). \end{aligned}$$

Similarly we receive that

$$\begin{aligned} &\int \Big(\frac{t}{\sqrt{n}} \sum_{j=1}^n \big(-\sin(t\varepsilon_j) + \cos(t\varepsilon_j)\big) \frac{m(X_j, \widehat{\vartheta}_n) - m(X_j)}{\widehat{\sigma}_n(X_j)} \\ &-t\sqrt{n}(\widehat{\vartheta}_n - \vartheta_0)^T \mathrm{E}\Big(\frac{1}{\sigma(X_1)} \frac{\partial m(X_1, \vartheta_0)}{\partial \vartheta_0}\Big) \big(C(t) - S(t)\big)\Big)^2 w(t)\,\mathrm{d}t = o_P(1). \end{aligned}$$

Next by Lemma A.2 in van Keilegom et al. [16] we have the following asymptotic representation for the estimator $\widehat{\vartheta}_n$

$$\sqrt{n}(\widehat{\vartheta}_n - \vartheta_0) = \frac{1}{\sqrt{n}} \sum_{j=1}^n \boldsymbol{B}^{-1}_{\vartheta_0} \frac{\partial m(X_j, \theta_0)}{\partial \theta} \sigma(X_j)\varepsilon_j + o_P(1)$$

Still we have to prove that the remainder terms $R^c_{nj}(t), R^s_{nj}(t), Q^c_{nj}(t), Q_{nj^s}(t)$ do not influence the limit behavior. This property for $R^c_{nj}(t), R^s_{nj}(t)$ follows from Lemma 4 in Hušková and Meintanis [10]. Concerning negligibility of $Q_{nj^c}(t), Q_{nj^s}(t)$ we notice

$$\begin{aligned} &\frac{t^2}{\sqrt{n}} \sum_{j=1}^n (|Q^c_{nj}(t)| + |Q^s_{nj}(t)|) \\ &= O_P\Big(\frac{t^2}{\sqrt{n}} \sum_{j=1}^n \Big(\varepsilon_j^2 (\sigma^2(X_j) - \widehat{\sigma}^2(X_j))^2 + (m(X_j, \widehat{\vartheta}_n) - m(X_j, \vartheta_0)^2 \Big) \end{aligned}$$

uniformly in t. Applying properties of the estimators $\widehat{\sigma}^2(X_j)$ and $m(X_j, \widehat{\vartheta}_n)$ we receive that

$$\int \Big(\frac{t^2}{\sqrt{n}} (\sum_{j=1}^n (|Q^c_{nj}(t)| + |Q^s_{nj}(t)|) \Big)^2 w(t)\,\mathrm{d}t = o_P(1).$$

Combining all these issues together we can conclude that

$$\int \Big\{ Z_n(t) - \frac{t(C(t)-S(t))}{\sqrt{n}} \sum_{j=1}^{n} \varepsilon_j (1 - \sigma(X_j) \boldsymbol{b}^T(X_j, \theta_0)$$

$$\boldsymbol{B}_{\vartheta_0}^{-1} \mathrm{E}\Big(\frac{1}{\sigma(X_1)} \boldsymbol{b}(X_1, \vartheta_0) \Big) \Big\}^2 w(t)\, \mathrm{d}t = o_P(1).$$

The assertion of Theorem 1 can be easily finished. □

P r o o f o f T h e o r e m 2. It follows the same line as in Theorem 1 with replacing $m(x,\theta)$ by $m(x,\theta) + r(x)/\sqrt{n}$. We also apply Lemma 2 in van Keilegom et al. [16]. The details are omitted. □

ACKNOWLEDGEMENT

The authors wish to sincerely thank the Associate Editor and the referees for many comments that led to improvement of the manuscript. The work of the first author was partially supported by the Czech Science Foundation (Grant 201/09/0755) and by the Ministry of Education, Youth and Sport of the Czech Republic, Project MSM 0021620839. Work on this topic was completed during the first author's visit to the National and Kapodistrian University of Athens. Marie Hušková wishes to sincerely thank the University of Athens for the hospitality and financial support.

(Received June 26, 2008.)

R E F E R E N C E S

[1] M. Bilodeau and P. de Lafaye de Micheaux: A multivariate empirical characteristic function test of independence with normal marginals. J. Multivariate Anal. *95* (2005), 345–369.

[2] H. D. Bondell: Testing goodness-of-fit in logistic case-control studies. Biometrika *94* (2007), 487–495.

[3] H. Dette: A consistent test for the functional form of a regression function based on a difference of variance estimators. Ann. Statist. *27* (1999), 1012–1040.

[4] R. L. Eubank, Chin-Shang Li, and Suojin Wang: Testing lack-of-fit of parametric regression models using nonparametric regression techniques. Statistica Sinica *15* (2005), 135–152.

[5] Jianqing Fan and Li-Shan Huang: Goodness-of-fit tests for parametric regression models. J. Amer. Statist. Assoc. *96* (2001), 640–652.

[6] A. K. Gupta, N. Henze, and B. Klar: Testing for affine equivalence of elliptically symmetric distributions. J. Multivariate Anal. *88* (2004), 222–242.

[7] W. Härdle and E. Mammen: Comparing nonparametric versus parametric regression fits. Ann. Statist. *21* (1993), 1926–1947.

[8] N. Henze, B. Klar, and S. G. Meintanis: Invariant tests for symmetry about an unspecified point based on the empirical characteristic function. J. Multivariate Anal. *87* (2003), 275–297.

[9] M. Hušková and S. G. Meintanis: Change point analysis based on empirical characteristic functions. Metrika *63* (2006), 145–168.

[10] M. Hušková and S. G. Meintanis: Tests for the error distribution in nonparametric possibly heteroscedastic regression models. To appear in Test.

[11] A. Kankainen and N. Ushakov: A consistent modification of a test for independence based on the empirical characteristic function. J. Math. Sci. *89* (1998), 1486–1494.

[12] Qi Li and Suojin Wang: A simple consistent bootstrap test for a parametric regression function. J. Econometrics *87* (1998), 145–165.

[13] S. G. Meintanis: Permutation tests for homogeneity based on the empirical characteristic function. J. Nonparametr. Statist. *17* (2005), 583–592.

[14] N. Neumeyer: A bootstrap version of the residual-based smooth empirical distribution function. J. Nonparametr. Statist. *20* (2008), 153–174.

[15] W. Stute, W. Gonzáles Manteiga, and M. Presedo Quindimil: Bootstrap approximation in model checks for regression. J. Amer. Statist. Assoc. *93* (1998), 141–149.

[16] I. van Keilegom, W. Gonzáles Manteiga, and C. Sanchez Sellero: Goodness-of-fit tests in parametric regression based on the estimation of the error distribution. Test *17* (2008), 401–415.

[17] G. A. F. Seber and C. J. Wild: Nonlinear Regression. Wiley, New York 1989, pp. 501–513.

[18] Yoon-Jae Whang: Consistent bootstrap tests of parametric regression functions. J. Econometrics *98* (2000), 27–46.

[19] H. White: Consequences and detection of misspecified nonlinear regression models. J. Amer. Statist. Assoc. *76* (1981), 419–433.

[20] H. White: Maximum likelihood estimation of misspecified models. Econometrica *50* (1982), 1–25.

[21] C. F. Wu: Asymptotic theory of nonlinear least-squares estimation. Ann. Statist. *9* (1981) 501–513.

[22] A. I. Zayed: Handbook of Function and Generalized Function Transformations. CRC Press, New York 1996.

Marie Hušková, Department of Probability and Mathematical Statistics, Faculty of Mathematics and Physics – Charles University, Sokolovská 83, 186 75 Praha 8 and Institute of Information Theory and Automation – Academy of Sciences of the Czech Republic, Pod Vodárenskou věží 4, 182 08 Praha 8. Czech Republic.
huskova@karlin.mff.cuni.cz

Simos G. Meintanis, Department of Economics, National and Kapodistrian University of Athens, 8 Pesmazoglou Street, 105 59, Athens. Greece.
simosmei@econ.uoa.gr