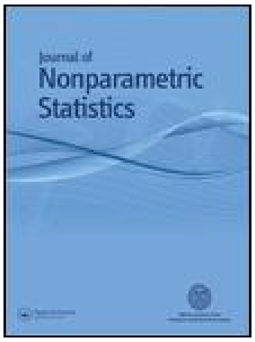

Journal of Nonparametric Statistics

ISSN: 1048-5252 (Print) 1029-0311 (Online) Journal homepage: www.tandfonline.com/journals/gnst20

Testing normality using kernel methods

Ibrahim Ahmad & A. R. Mugdadi

To cite this article: Ibrahim Ahmad & A. R. Mugdadi (2003) Testing normality using kernel methods, Journal of Nonparametric Statistics, 15:3, 273-288, DOI: 10.1080/1048525021000049649

To link to this article: https://doi.org/10.1080/1048525021000049649

Published online: 27 Oct 2010.

Nonparametric Statistics, Vol. 15(3), June 2003, pp. 273–288

TESTING NORMALITY USING KERNEL METHODS

IBRAHIM A. AHMAD[a,*] and A. R. MUGDADI[b,†]

[a]*Department of Statistics, University of Central Florida, Orlando, FL 32816-2370;* [b]*Department of Mathematics, Southern Illinois University, Carbondale, IL 62901*

(Received June 2001; Revised March 2002; In final form August 2002)

Testing normality is one of the most studied areas in inference. Many methodologies have been proposed. Some are based on characterization of the normal variate, while most others are based on weaker properties of the normal. In this investigation, we propose a new procedure, which is based on the well-known characterization; if X_1 and X_2 are two independent copies of a variable with distribution F, then X_1 and X_2 are normal if and only if $X_1 - X_2$ and $X_1 + X_2$ are independent. If $X_1, \ldots, X_n$ is a random sample from F, we test that F is normal by testing nonparametrically that $u_{ii^*} = X_i - X_{i^*}$ and $v_{ii^*} = X_i + X_{i^*}$ are independent, $i \neq i^* = 1, 2, \ldots, n$. This procedure has several major advantages; it applies equally to one-dimensional or multi-dimensional cases, it does not require estimation of parameters, it does not require transformation to uniformity, it does not require use of special tables of coefficients, and it does have very good power requiring much less number of iterations to reach stable results.

Keywords: Testing normality; Independence; Kernel methods; Bandwidth selection; Monte Carlo methods; Power of tests; Kernel contrasts; Asymptotic normality

1 INTRODUCTION

One of the oldest and most studied goodness of fit problems is testing normality. Testing procedures for this problem are basically either based on a characterization of the normal or on a weaker property of it. Many tests require one or more of the following. If parameters are unknown, they are to be estimated such as in the empirical distribution based tests, the Shapiro–Wilk tests, or the Locke–Spurrier tests. Sometimes, these tests require special tables such as the Shapiro–Wilk tests or the empirical distribution based tests, cf. Green and Hegazy (1976). Several tests are based on weaker properties of the normal rather than on a characterization of the distribution, such as the Shapiro–Wilk, Locke–Spurrier, or Lin–Mudholkar tests. Tests that are based on empirical distributions usually require transformation of the data to uniformity. For further details on the above-mentioned tests, the reader is referred to Shapiro and Wilk (1965), Locke and Spurrier (1976), and Lin and Mudholkar (1980). It is not our intention here to offer a major survey of this extensive area; rather, we will refer only to references pertinent to our development.

* E-mail: iahmad@mail.ucf.edu
† Corresponding author. E-mail: amugdadi@math.siu.edu

ISSN 1048-5252 print; ISSN 1029-0311 online
DOI: 10.1080/1048525021000049649

In a marked departure from the above-mentioned approaches, Vasicek (1976), proposed a test statistic for normality based on the entropy characterization of the normal and using a special version of the celebrated kernel density estimate. His procedure avoids the deficiency trio outlined above. It is just not easy to calculate, and its limiting distribution is difficult to track; hence, this limiting behavior remains illusive. Thus, it is not unreasonable to see if there is another kernel approach that enjoys known limiting null distribution that is easy to obtain.

In the current work, we use the full strength of the kernel method of density estimation to provide a testing procedure for normality that is based on the characterization that X is normal, if and only if, for two independent copies, X_1 and X_2 of $X, X_1 + X_2$ and $X_1 - X_2$ are independent. Thus, based on a random sample $X_1, \ldots, X_n$ from a distribution F with probability density function f, we form the pairings $u_{ii^*} = X_i + X_{i^*}$ and $v_{ii^*} = X_i - X_{i^*}$ for all $i \neq i^* = 1, 2, \ldots, n$. Using the celebrated kernel method, we estimate the joint density of (u, v) and the marginals of u and v and use them to construct a test statistic for the independence of u and v; hence, the normality of X.

Using the kernel density estimation techniques in hypothesis testing is gaining popularity in recent years, cf. Ahmad and Li (1997; 1998), Fan (1996), Fan and Li (1996), Hart (1997), Li (1996), and many others. The present investigation develops estimates of the joint density of $X_1 + X_2$ and $X_1 - X_2$ as well as their marginals. This is an extension of the concept of density estimation to functions of several variables. Fuller treatment of this topic is done by the authors in Ahmad and Mugdadi (2002).

The organization of the paper is as follows, in Section 2, we present the test procedure, derive its null limiting distributions, and present some simulation work to demonstrate the viability of the test including power of the test. In Section 3, we present an appendix of the proofs of main results.

2 TESTING NORMALITY USING KERNEL METHODS

2.1 Test Procedure

Let $X_1, \ldots, X_n$ be a random sample for a distribution Function F with density f. Let $u_{ii^*} = X_i + X_{i^*}$ with density h_1 and $v_{ii^*} = X_i - X_{i^*}$ with density h_2. Further, let $h(u, v)$ denote the joint density of u_{ii^*} and v_{ii^*} for all $i \neq i^* = 1, 2, \ldots, n$. The kernel estimates of h_1, h_2 and h are given, respectively, by:

$$\hat{h}_1(u) = \frac{1}{n(n-1)b} \sum_{i \neq i^*} w\left(\frac{u - u_{ii^*}}{b}\right), \tag{2.1}$$

$$\hat{h}_2(v) = \frac{1}{n(n-1)b} \sum_{i \neq i^*} w\left(\frac{v - v_{ii^*}}{b}\right), \tag{2.2}$$

and

$$\hat{h}(u, v) = \frac{1}{n(n-1)b^2} \sum_{i \neq i^*} w\left(\frac{u - u_{ii^*}}{b}\right) w\left(\frac{v - v_{ii^*}}{b}\right), \tag{2.3}$$

where $b = b_n$ is a positive constant called the "bandwidth" and $w(\cdot)$ is a known symmetric bounded density called the "kernel". Henceforth, we assume that w has mean 0 and a finite variance $\mu_2(w)$, and that $b \to 0$ as $n \to \infty$.

We want to test H_0:f is $N(\mu, \sigma^2)$ against H_a:f is not $N(\mu, \sigma^2)$, for $\mu \in R$ and $\sigma^2 > 0$. A measure of departure from H_0 is

$$\delta = \int_{-\infty}^{\infty} \int_{\infty}^{\infty} [h(u, v) - h_1(u)h_2(v)]^2 \mathrm{d}u\, \mathrm{d}v.$$

This is so, since H_0 is equivalent to H_0: $h(u, v) = h_1(u)h_2(v)$, *i.e.*, U and V are independent.

Note that $\delta \geq 0$ and that $\delta = 0$ if and only if H_0 is true; thus, we can use δ for testing H_0.

Using the random sample $X_1, X_2, \ldots, X_n$ and using the estimates $\hat{h}_1, \hat{h}_2$ and $\hat{h}$ above, we can perform the normality test based on the estimation of δ given by

$$\hat{\delta} = \int_{-\infty}^{\infty} \int_{\infty}^{\infty} [\hat{h}(u, v) - \hat{h}_1(u)\hat{h}_2(v)]^2 \mathrm{d}u\, \mathrm{d}v. \tag{2.4}$$

Note that $\hat{\delta}$ can be written in the form

$$\hat{\delta} = I_1 - 2I_2 + I_3,$$

where,

$$I_1 = \int_{-\infty}^{\infty} \int_{\infty}^{\infty} \hat{h}^2(u, v)\, \mathrm{d}u\, \mathrm{d}v$$
$$I_2 = \int_{-\infty}^{\infty} \int_{\infty}^{\infty} \hat{h}_1(u)\hat{h}_2(v)\hat{h}(u, v)\, \mathrm{d}u\, \mathrm{d}v$$

and

$$I_3 = \int_{-\infty}^{\infty} \hat{h}_1^2(u)\, \mathrm{d}u \int_{-\infty}^{\infty} \hat{h}_2^2(v)\, \mathrm{d}v.$$

But

$$\hat{h}^2(u, v) = \frac{1}{n^2(n-1)^2 b^4} \sum_{i \neq i^*} \sum_{j \neq j^*} w\Big(\frac{u - u_{ii^*}}{b}\Big) w\Big(\frac{u - u_{jj^*}}{b}\Big) w\Big(\frac{v - v_{ii^*}}{b}\Big) w\Big(\frac{v - v_{jj^*}}{b}\Big).$$

Thus, I_1 can be written in the form

$$I_1 = \frac{1}{n^2(n-1)^2 b^4} \sum_{i \neq i^*} \sum_{j \neq j^*} \int \Big(w\Big(\frac{u - u_{ii^*}}{b}\Big) w\Big(\frac{u - u_{jj^*}}{b}\Big)\Big) \mathrm{d}u$$
$$\times \int \Big(w\Big(\frac{v - v_{ii^*}}{b}\Big) w\Big(\frac{v - v_{jj^*}}{b}\Big)\Big) \mathrm{d}v,$$

but,

$$\int\left(w\left(\frac{u-u_{ii^*}}{b}\right)w\left(\frac{u-u_{jj^*}}{b}\right)\right)\mathrm{d}u = bw^{(2)}\left(\frac{u_{ii^*}-u_{jj^*}}{b}\right)$$

and

$$\int\left(w\left(\frac{v-v_{ii^*}}{b}\right)w\left(\frac{v-v_{jj^*}}{b}\right)\right)\mathrm{d}v = bw^{(2)}\left(\frac{v_{ii^*}-v_{jj^*}}{b}\right),$$

where $w^{(2)}$ is the two fold convolution of w. Therefore, we can write I_1 as the sum of convolution functions

$$I_1 = \frac{1}{n^2(n-1)^2b^2}\sum_{i\neq i^*}\sum_{j\neq j^*} w^{(2)}\left(\frac{u_{ii^*}-u_{jj^*}}{b}\right)w^{(2)}\left(\frac{v_{ii^*}-v_{jj^*}}{b}\right). \tag{2.5}$$

Also, we have

$$\hat{h}(u,v)\hat{h}_1(u)\hat{h}_2(v) = \frac{1}{n^3(n-1)^3b^4}\sum_{i\neq i^*}\sum_{j\neq j^*}\sum_{l\neq l^*} w\left(\frac{u-u_{ii^*}}{b}\right)w\left(\frac{v-v_{ii^*}}{b}\right)w\left(\frac{u-u_{jj^*}}{b}\right)w\left(\frac{v-v_{ll^*}}{b}\right).$$

Again, we can write I_2 as the sum of convolution functions

$$I_2 = \frac{1}{n^3(n-1)^3b^2}\sum_{i\neq i^*}\sum_{j\neq j^*}\sum_{l\neq l^*} w^{(2)}\left(\frac{u_{ii^*}-u_{jj^*}}{b}\right)w^{(2)}\left(\frac{v_{ii^*}-v_{ll^*}}{b}\right). \tag{2.6}$$

Similarly, we can write I_3 in the form

$$I_3 = \frac{1}{n^4(n-1)^4b^2}\sum_{i\neq i^*}\sum_{j\neq j^*} w^{(2)}\left(\frac{u_{ii^*}-u_{jj^*}}{b}\right)\frac{1}{n^4(n-1)^4b^2}\sum_{i\neq i^*}\sum_{j\neq j^*} w^{(2)}\left(\frac{v_{ii^*}-v_{jj^*}}{b}\right). \tag{2.7}$$

Therefore, $\hat{\delta}$ can be written in the form

$$\hat{\delta} = \frac{1}{n^4(n-1)^4b^2}\sum_{i\neq i^*}\sum_{j\neq j^*}\sum_{l\neq l^*}\sum_{r\neq r^*} w^u_{ii^*jj^*}[w^v_{ii^*jj^*} + w^v_{ll^*rr^*} - 2w^v_{ii^*ll^*}], \tag{2.8}$$

where,

$$w^u_{ii^*jj^*} = w^{(2)}\left(\frac{u_{ii^*}-u_{jj^*}}{b}\right)$$

and

$$w^u_{ii^*jj^*} = w^{(2)}\left(\frac{v_{ii^*}-v_{jj^*}}{b}\right).$$

Also, we can write

$$\hat{\delta} = \hat{\delta}_1 + \hat{\delta}_2, \tag{2.9}$$

where

$$\hat{\delta}_1 = \frac{1}{n^4(n-1)^4 b^2} \sum\sum\sum\sum\sum\sum\sum\sum\nolimits_{i \neq i^* \neq j \neq j^* \neq l \neq l^* \neq r \neq r^*} w^u_{ii^*jj^*} \times [w^v_{ii^*jj^*} + w^v_{ll^*rr^*} - 2w^v_{ii^*ll^*}], \tag{2.10}$$

and

$$\hat{\delta}_2 = o(1). \tag{2.11}$$

We will show, under some regularity conditions on the kernel function and on the distribution functions $h_1(u)$, $h_2(v)$ and $h(u, v)$ and under H_0, that $\hat{\delta}_1$ is asymptotically normal with mean zero and obtain its asymptotic variance.

Note that as with all procedures based on density estimation via the kernel methods, one needs to find data-based choices of the bandwidth b. There are several methodologies, cf. Wand and Jones (1995), and we choose a new method developed by Ahmad and Mugdadi (2002) called “contrasts method”. It seems to work well and is easy to implement.

2.2 Asymptotic Null Distribution

To state our main theorem, we first state the following lemmas:

LEMMA 2.1

Define

$$\Psi_n(X_i, X_{i^*}, X_j, X_{j^*}, X_l, X_{l^*}, X_r, X_{r^*}) = \sum_{8!} w^u_{ii^*jj^*}[w^v_{ii^*jj^*} + w^v_{ll^*rr^*} - 2w^v_{ii^*ll^*}], \tag{2.12}$$

therefore,

$$E(\Psi_n(X_i, X_{i^*}, X_j, X_{j^*}, X_l, X_{l^*}, X_r, X_{r^*})|X_i = x_i) = 0. \tag{2.13}$$

Next, let

$$\Psi_{n2}(x_i, x_{i^*}) = E(\Psi_n(X_i, X_{j^*}, X_j, X_{j^*}, X_l, X_{l^*}, X_r, X_{r^*})|(X_i, X_{i^*} = x_i, x_{i^*})).$$

Therefore, based on this definition, we can state the following lemma:

LEMMA 2.2

$$EG_n^2(X_i, X_{i^*} = O(b^{14}), \tag{2.14}$$

where

$$G_n(x_i, x_{i^*}) = E[\Psi_{n2}(X_i, X_j)\Psi_{n2}(X_{i^*}, X_j)|(X_i, X_{i^*}) = (x_i, x_{i^*})] \tag{2.15}$$

THEOREM 2.1 *Under the regularity conditions on the kernel function introduced in Section 2.1, and, under H_0, we have*

$$n\hat{\delta}_1 \to N(0, \sigma_n^2), \tag{2.16}$$

where

$$\sigma_n^2 = \frac{1568}{b^4}\sigma_{0n}^2 \tag{2.17}$$

$$\sigma_{0n}^2 = \int\int\left[bv_1(x_1, x_2) + \frac{b^2\mu_2(w)}{2}v_2(x_1, x_2)\right]^4 f(x_1)f(x_2)\,\mathrm{d}x_1\,\mathrm{d}x_2$$

with $f(x_i)$ normally distributed with mean 0 and variance 1,

$$v_1(x_1, x_2) = \frac{1}{2\sqrt{2\pi}} + (f * f)(x_1 + x_2) - (f * (f * f))(x_1) - (f * (f * f))(x_2)$$

and (2.18)

$$v_2(x_1, x_2) = \frac{1}{4\sqrt{4\pi}} + 2 * (f * f'')(x_1 + x_2) - (f * (f * f)'')(x_1) - (f * (f * f)'')(x_2).$$

Based on Theorem 2.1, it is easy to prove the following corollary

COROLLARY 2.1 *Under the Conditions of Theorem 2.1,*

$$\hat{J}_n = \frac{n}{\sigma_n}\hat{\delta}_1 \to N(0, 1). \tag{2.19}$$

2.3 Significant Points and Power of the Test

In this section, we will report a Monte Carlo simulation of $\hat{J}_n$, and we will use the method of contrast to choose the bandwidth b, see Ahmad and Mugdadi (2002). We will choose

1. $q = 2$ with $p_1 = 1$ and $p_2 = -1$
2. The kernel w_1 is normal with mean 0 and variance 4, the kernel w_2 is normal with mean 0 and variance 1.

We obtained Monte Carlo estimates of critical values of $\hat{J}_n$ for $n = 10, 15, 20, 25$. The estimates for each sample size are based on 200 trials. From Table I, this approximation is not bad for some critical significance level 0.10, 0.05, and 0.01 for the samples $n = 10(5)20$.

TABLE I The Percentiles from Normal.

n	*0.10*	*0.05*	*0.01*
10	0.9593	1.0950	2.0973
15	1.4188	1.9264	2.3624
20	1.3810	1.5334	2.6583
25	1.2012	1.6308	2.3522

TABLE II The Percentiles from Exponential.

n	*0.10*	*0.05*	*0.01*
10	2.2531	2.3016	2.3603
15	26.439	28.628	29.245
20	0.9866	58.598	81.614
25	3.4783	40.289	45.001

In order to study the performance of the test, we simulated from exponential distribution. Based on the sample $n = 10(5)20$, we obtained the results in Table II.

From Table II, we conclude it is clear we reject the hypothesis when we simulate from distribution different than normal. Similar results were obtained for simulations from the Chi-square, the Cauchy and the Beta distributions. Details are available from the author.

In order to assess our procedure relative to others in the literature, we computed the empirical power of our test for various alternatives handled by other authors, cf. Locke and Spurrier (1976), Lin and Mudholkar (1980) and Shapiro and Wilk (1965). Based on 1000 repititions of samples of sizes $n = 20, 50$, and 100 we report the following values.

Comparing the above results to those of Locke and Spurrier (1976) and Lin and Mudholkar (1980) we see that our test is more than competition to these tests. It actually holds an edge over these tests. In addition, our test achieves stability of the results with as little as 200 repetitions while others require as much as 2000 repetitions to reach results stability. This is an indicator that our test enjoys faster rate of convergence and hence could be used for more moderate samples. Note that our test is much more powerful than that of Vasicek (1976), cf. Lin and Mudholkar (1980) and the limiting null distribution of the test of Vasicek remains unobtainable so far.

3 PROOFS OF THE RESULTS

Proof of Lemma 2.1 Let

$$\Psi_n(x_1, x_2, \ldots, x_8) = \sum_{8!} k^u_{ii^*jj^*}[k^u_{ii^*jj^*} + k^v_{ll^*rr^*} - 2k^v_{ii^*ll^*}]$$

Define

$$\Psi_{n2}(x_1, x_2) = E[\Psi_n(X_1, \ldots, X_n)|(X_1, X_2) = (x_1, x_2)].$$

Thus,

$$E[\Psi_{n2}(X_1, X_2)|X_1 = x_1] = E[\Psi_n, (X_1, \ldots, X_8)|X_1 = x_1].$$

We want to show that $E[\Psi_{n2}(X_1, X_2)|X_1 = x_1] = 0$ almost surely under H_0. In order to evaluate $\Psi_{n2}(X_1, X_2)$, we have 8! permutations for $x_1, \ldots, x_{r^*}$. There are $\binom{8}{2} = 28$ ways to choose the X_1 and X_2. For simplicity, let $i = 1, i^* = 2, \ldots,$ and $r^* = 8$.

$$
\begin{aligned}
&(1{,}2)\text{: } E(k^u_{1234}|X_1,X_2)[E(k^v_{1234}|X_1,X_2)+E(k^v_{1234}-2E(k^v_{1234}|X_1,X_2)]\\
&(1{,}3)\text{: } E(k^u_{1324}|X_1,X_2)[E(k^v_{1324}|X_1,X_2)+E(k^v_{1234}-2E(k^v_{1234}|X_1)]\\
&(1{,}4)\text{: } E(k^u_{1342}|X_1,X_2)[E(k^v_{1342}|X_1,X_2)+E(k^v_{1234}-2E(k^v_{1234}|X_1)]\\
&(1{,}5)\text{: } E(k^u_{1324}|X_1)[E(k^v_{1234}|X_1)+E(k^v_{1234}|X_2)-2E(k^v_{1324}|X_1,X_2)]\\
&(1{,}6)\text{: } E(k^u_{1345}|X_1)[E(k^v_{1234}|X_1)+E(k^v_{1234}|X_2)-2E(k^v_{1342}|X_1,X_2)]\\
&(1{,}7)\text{: } E(k^u_{1234}|X_1)[E(k^v_{1234}|X_1)+E(k^v_{2345}|X_2)-2E(k^v_{1234}|X_1,X_2)]\\
&(1{,}8)\text{: } E(k^u_{1324}|X_1)[E(k^v_{1324}|X_1)+E(k^v_{1234}|X_2)-2E(k^v_{1234}|X_1)]\\
&(2{,}3)\text{: } E(k^u_{1324}|X_1,X_2)[E(k^v_{1324}|X_1,X_2)+E(k^v_{1234})-2E(k^v_{3124}|X_1)]\\
&(2{,}4)\text{: } E(k^u_{1324}|X_1,X_2)[E(k^v_{3142}|X_1,X_2)+E(k^v_{1234})-2E(k^v_{2134}|X_1)]\\
&(2{,}5)\text{: } E(k^u_{1234}|X_1)[E(k^v_{3145}|X_1)+E(k^v_{2134}|X_2)-2E(k^v_{3124}|X_1,X_2)]\\
&(2{,}6)\text{: } E(k^u_{1345}|X_1)[E(k^v_{3145}|X_1)+E(k^v_{3245}|X_2)-2E(k^v_{3142}|X_1,X_2)]\\
&(2{,}7)\text{: } E(k^u_{1234}|X_1)[E(k^v_{3145}|X_1)+E(k^v_{2134}|X_2)-2E(k^v_{3145}|X_1)]\\
&(2{,}8)\text{: } E(k^u_{1234}|X_1)[E(k^v_{3145}|X_1)+E(k^v_{3245}|X_2)-2E(k^v_{3145}|X_1)]\\
&(3{,}4)\text{: } E(k^u_{1234}|X_1,X_2)[E(k^v_{3145}|X_1,X_2)+E(k^v_{3245})-2E(k^v_{3145})]\\
&(3{,}5)\text{: } E(k^u_{1345}|X_1)[E(k^v_{3145}|X_1)+E(k^v_{2345}|X_2)-2E(k^v_{2345}|X_2)]\\
&(3{,}6)\text{: } E(k^u_{1345}|X_1)[E(k^v_{1234}|X_1)+E(k^v_{1234}|X_2)-2E(k^v_{3245}|X_2)]\\
&(3{,}7)\text{: } E(k^u_{1345}|X_1)[E(k^v_{1234}|X_1)+E(k^v_{2345}|X_2)-2E(k^v_{2345})]\\
&(3{,}8)\text{: } E(k^u_{1234}|X_1)[E(k^v_{1234}|X_1)+E(k^v_{3245}|X_2)-2E(k^v_{1234})]\\
&(4{,}5)\text{: } E(k^u_{1345}|X_1)[E(k^v_{3124}|X_1)+E(k^v_{2345}|X_2)-2E(k^v_{2345}|X_1)]\\
&(4{,}6)\text{: } E(k^u_{1234}|X_1)[E(k^v_{2134}|X_1)+E(k^v_{3245}|X_2)-2E(k^v_{3245}|X_2)]\\
&(4{,}7)\text{: } E(k^u_{1345}|X_1)[E(k^v_{3145}|X_1)+E(k^v_{2345}|X_2)-2E(k^v_{12345})]\\
&(4{,}8)\text{: } E(k^u_{1345}|X_1)[E(k^v_{3145}|X_1)+E(k^v_{3245}|X_2)-2E(k^v_{1234})]\\
&(5{,}6)\text{: } E(k^u_{1234})[E(k^v_{1234})+E(k^v_{1234}|X_1,X_2)-2E(k^v_{1234}|X_1,X_2)]\\
&(5{,}7)\text{: } E(k^u_{1234})[E(k^v_{1234})+E(k^v_{1324}|X_1,X_2)-2E(k^v_{1234}|X_1)]\\
&(5{,}8)\text{: } E(k^u_{1234})[E(k^v_{1234})+E(k^v_{1342}|X_1,X_2)-2E(k^v_{1345}|X_1)]\\
&(6{,}7)\text{: } E(k^u_{1234})[E(k^v_{1234})+E(k^v_{3124}|X_1,X_2)-2E(k^v_{1234}|X_2)]\\
&(6{,}8)\text{: } E(k^u_{1234})[E(k^v_{1234})+E(k^v_{3142}|X_1,X_2)-2E(k^v_{1234}|X_2)]\\
&(7{,}8)\text{: } E(k^u_{1234})[E(k^v_{1234})+E(k^v_{1324}|X_1,X_2)-2E(k^v_{1234})]
\end{aligned}
$$

Under H_0, we have

$$E(k^v_{1234}|X_1) = E(k^v_{2134}|X_1)$$

and

$$E(k^v_{1324}|X_1, X_2) = E(k^v_{1342}|X_1, X_2)$$

Therefore, after some algebra, we can write

$$\Psi_{n2}(x_1, x_2) = 2(28)6!SS_1SS_2,$$

where

$$SS_1 = E(k^u_{1234}|X_1, X_2) + E(k^u_{1234}) - E(k^u_{1234}|X_1) - E(k^u_{1234}|X_2)$$

and

$$SS_2 = E(k^v_{1234}|X_1, X_2) + E(k^v_{1234}) - E(k^v_{1234}|X_1) - E(k^v_{1234}|X_2).$$

But,

$$E[SS_2 E(k^u_{1234}|X_1, X_2)|X_1] = E(k^u_{1234}|X_1).$$

$$[E(k^v_{1324}|X_1) + E(k^v_{1234}) - E(k^v_{1234}) - E(k^v_{1234}|X_1)] = 0.$$

Also, it is easy to show that

$$E[E(k^u_{1234})SS_2|X_1] = E(k^u_{1234}E[SS_2|X_1] = 0,$$

$$E[E(k^u_{1234}|X_1)SS_2|X_1] = 0$$

and

$$E[E(k^u_{1234}|X_2)\mathrm{SS}_2|X_1] = 0.$$

Thus,

$$E[\Psi_{n2}(X_1, X_2)|X_1 = x_2] = 0.$$

Proof of Lemma 2.2

Define

$$G_n(x_1, x_2) = E[\Psi_{n2}(X_1, X_3)\Psi_{n2}(X_2, X_3)|(X_1, X_2) = (x_1, x_2)]$$

and

$$\begin{aligned}\Psi_{n2}(x_i, X_j) &= 2\binom{8}{2}6!\Big[E\Big(k^u_{ii^*jj^*}|x_i, x_j\Big) - E\Big(k^u_{ii^*jj^*}|x_i\Big) - E\Big(k^u_{ii^*jj^*}|x_j\Big) + E\Big(k^u_{ii^*jj^*}\Big)\\ &= \frac{2\binom{8}{2}6!}{b^2}\left[E\int\left[k\left(\frac{u-(x_i+X_{i^*})}{b}\right) - Ek\left(\frac{u-(X_i+X_{i^*})}{b}\right)\right]\right.\\ &\quad\left.\times\left[k\left(\frac{u-(x_j+X_{j^*})}{b}\right) - Ek\left(\frac{u-(X_j+X_{j^*})}{b}\right)\right]\mathrm{d}u\right]^2.\end{aligned}$$

Let

$$B_n(x_i, x_j) = E\int\left[k\left(\frac{u-(x_i+X_{i^*})}{b}\right) - Ek\left(\frac{u-(X_i+X_{i^*})}{b}\right)\right]$$
$$\times\left[k\left(\frac{u-(x_j+X_{j^*})}{b}\right) - Ek\left(\frac{u-(X_j+X_{j^*})}{b}\right)\right]\mathrm{d}u,$$
$$Q(u, x_i) = E\int\left[k\left(\frac{u-(x_i+X_{i^*})}{b}\right) - Ek\left(\frac{u-(X_i+X_{i^*})}{b}\right)\right],$$

and

$$B_n(x_i, x_j) = \int(Q(u, x_i), Q(u, x_j))\mathrm{d}u.$$

Therefore,

$$\Psi_{n2}(X_1, X_3)\Psi_{n2}(X_2, X_3) = \left(\frac{2\binom{8}{2}6!}{b^2}\right)^2 (B_n(X_1, X_3)B_n(X_2, X_3))^2.$$

Thus,

$$G_n(x_1, x_2) = \left(\frac{2\binom{8}{2}6!}{b^2}\right)^2 E[(B_n(X_1, X_3)B_n(X_2, X_3))^2 \mid X_1 = x_1, X_2 = x_2].$$

Or

$$G_n(x_1, x_2) = \left(\frac{2\binom{8}{2}6!}{b^2}\right)^2 E[B_n(x_1, X_3)B_n(x_2, X_3)^2]^2$$
$$= \left(\frac{2\binom{8}{2}6!}{b^2}\right)^2 E\left[\int\int Q(u, x_1)Q(u, X_3)Q(v, x_2)Q(v, X_3)\,\mathrm{d}u\,\mathrm{d}v\right]^2$$
$$= \left(\frac{2\binom{8}{2}6!}{b^2}\right)^2 E\int\int\int\int Q(u_1, x_1)Q(u_1, X_3)Q(v_1, x_2)Q(v_1, X_3)$$
$$\times\, Q(u_2, x_1)Q(u_2, X_3)Q(v_2, x_2)Q(v_2, X_3)\,\mathrm{d}u_1\,\mathrm{d}u_2\,\mathrm{d}v_1\,\mathrm{d}v_2$$
$$= \left(\frac{2\binom{8}{2}6!}{b^2}\right)^2 \int\int\int\int Q(u_1, x_1)Q(u_2, x_1)Q(v_1, x_2)Q(v_2, x_2)$$
$$\times\, EQ(u_1, X_3)Q(u_2, X_3)Q(v_1, X_3)Q(v_2, X_3)\,\mathrm{d}u_1\,\mathrm{d}u_2\,\mathrm{d}v_1\,\mathrm{d}v_2.$$

TABLE III Empirical Powers.

		Power				Power	
Distribution	*n*	*0.05*	*0.01*	*Distribution*	*n*	*0.05*	*0.01*
Beta(1,2)	20	0.266	0.132	Gamma $\alpha = 2$	20	0.720	0.671
	50	0.410	0.211		50	0.778	0.650
	100	0.967	0.673		100	0.910	0.882
Beta(2,3)	20	0.179	0.094	Gamma $\alpha = 5$	20	0.561	0.441
	50	0.278	0.121		50	0.608	0.578
	100	0.678	0.186		100	0.874	0.800
Weibull $\alpha = 2$	20	0.389	0.178	Log normal $\sigma^2 = 1/16$	20	0.489	0.440
	50	0.464	0.289		50	0.671	0.602
	100	0.738	0.449		100	0.718	0.644
Weibull $\alpha = 3$	20	0.201	0.117	Log normal $\sigma^2 = 0.25$	20	0.781	0.720
	50	0.227	0.120		50	0.904	0.869
	100	0.289	0.151		100	0.981	0.968
Gamma $\alpha \neq 1$	20	0.775	0.589	Log normal $\sigma^2 = 1$	20	0.922	0.881
	50	0.967	0.945		50	0.981	0.920
	100	0.987	0.987		100	0.988	0.968

Therefore,

$$EG_n^2(x_1, x_2) = \left(\frac{2\binom{8}{2}6!}{b^2}\right)^4 E \int\int\int\int\int\int\int\int [Q(u_1, x_1)Q(u_2, x_1)Q(v_1, x_2)Q(v_2, x_2)Q(u_3, x_1)$$
$$\times Q(u_4, x_1)Q(v_3, x_2)Q(v_4, x_2)EQ(u_1, X_3)Q(u_2, X_3)Q(v_1, X_3)Q(v_2, X_3)$$
$$\times EQ(u_3, X_3)Q(u_4, X_3)Q(v_3, X_3)Q(v_4, X_3)]\, \mathrm{d}u_1\, \mathrm{d}u_2\, \mathrm{d}v_1\, \mathrm{d}v_2\, \mathrm{d}u_3\, \mathrm{d}u_4\, \mathrm{d}v_3\, \mathrm{d}v_4.$$

Let

$$D(u_1, u_2, v_1, v_2) = EQ(u_1, x_1)Q(u_2, x_1)Q(u_2, x_1)Q(u_3, x_1)Q(u_4, x_1).$$

Thus,

$$EG_n^2(x_1, x_2) = \left(\frac{2\binom{8}{2}6!}{b^2}\right)^4 \int\int\int\int\int\int\int\int (D(u_1, u_2, u_3, u_4)D(v_1, v_2, v_3, v_4))^2$$
$$\times \mathrm{d}u_1\, \mathrm{d}u_2\, \mathrm{d}u_3\, \mathrm{d}u_4\, \mathrm{d}v_1\, \mathrm{d}v_2\, \mathrm{d}v_3\, \mathrm{d}v_4$$
$$= \left[\left(\frac{2\binom{8}{2}6!}{b^2}\right)^2 \int\int\int\int D^2(u_1, u_2, u_3, u_4)\, \mathrm{d}u_1\, \mathrm{d}u_2\, \mathrm{d}u_3\, \mathrm{d}u_4\right]^2.$$

But, $D(u_1, u_2, u_3, u_4)$ includes several terms all having the same rate.

Let $T^{(1)}(u_1, u_2, u_3, u_4)$ be the first term in $D(u_1, u_2, u_3, u_4)$. Thus,

$$T^{(1)}(u_1, u_2, u_3, u_4) = E\left[Ek\left(\frac{u_1 - (x_1 + X_3)}{b}\right)Ek\left(\frac{u_2 - (x_1 + X_3)}{b}\right) \times Ek\left(\frac{u_3 - (x_1 + X_3)}{b}\right)Ek\left(\frac{u_4 - (x_1 + X_3)}{b}\right)\right].$$

Therefore,

$$\begin{aligned}
&\int\int\int\int [T^{(1)}(u_1, u_2, u_3, u_4)]^2 \, \mathrm{d}u_1 \, \mathrm{d}u_2 \, \mathrm{d}u_3 \, \mathrm{d}u_4 \\
&\quad = \int\int\int\int\left[\int\left(Ek\left(\frac{u_1 - (x_1 + X_3)}{b}\right)\right)f(x_1)\,\mathrm{d}x_1\right. \\
&\quad \times \int\left(Ek\left(\frac{u_2 - (x_1 + X_3)}{b}\right)\right)f(x_1)\,\mathrm{d}x_1 \int\left(Ek\left(\frac{u_3 - (x_1 + X_3)}{b}\right)\right)f(x_1)\,\mathrm{d}x_1 \\
&\quad \times \left.\int\left(Ek\left(\frac{u_4 - (x_1 + X_3)}{b}\right)\right)f(x_1)\,\mathrm{d}x_1\right]^2 \mathrm{d}u_1 \, \mathrm{d}u_2 \, \mathrm{d}u_3 \, \mathrm{d}u_4 \\
&\quad = b^3 \int\int\int\int\left[\int\left(Ek\left(\frac{w_1 - (x_1 + X_3)}{b}\right)\right)f(x_1)\,\mathrm{d}x_1\right. \\
&\quad \times \int\left(Ek\left(\frac{w_1 - (x_1 + X_3)}{b} + w_2\right)\right)f(x_1)\,\mathrm{d}x_1. \\
&\quad \times \int\left(Ek\left(\frac{w_1 - (x_1 + X_3)}{b} - w_4 + w_3\right)\right)f(x_1)\,\mathrm{d}x_1 \\
&\quad \times \left.\int\left(Ek\left(\frac{w_1 - (x_1 + X_3)}{b} - w_4\right)\right)f(x_1)\,\mathrm{d}x_1\right]^2 \mathrm{d}w_1 \, \mathrm{d}w_2 \, \mathrm{d}w_3 \, \mathrm{d}w_4.
\end{aligned}$$

But, $\int(Ek((u_2 - (x_1 + X_3))/b))f(x_1)\,\mathrm{d}x_1 = O(b)$. Therefore,

$$\int\int\int\int [T^{(1)}(u_1, u_2, u_3, u_4)^2 \, \mathrm{d}u_2 \, \mathrm{d}u_3 \, \mathrm{d}u_4]^2 = b^3(O(b^4))^2.$$

Thus,

$$EG_n^2(x_1, x_2) = b^4 O(b^{11}) = O(b^{15}).$$

Proof of Theorem 2.1 Recall that $\hat{\delta}_1$ can be written as a U-statistics.

$$\hat{\delta}_1 = \frac{\binom{n}{8}}{n^4(n-1)^4 b^2}\left[\frac{1}{\binom{n}{8}} \sum_{1 \le i < i^* < j < j^* < l < l^* < r < r^* \le n} \psi_n(x_1, \ldots, x_{r^*})\right]$$

and

$$\psi_{n2}(x_1, x_2) = 2(28)6!SS_1SS_2,$$

where SS_1 and SS_2 are defined in the proof of Lemma 2.1. But,

$$\begin{aligned} k^u_{1234} &= k^{(2)}\Big(\frac{u_{12} - u_{13}}{b}\Big) \\ &= \frac{1}{b}\int k\Big(\frac{u - u_{12}}{b}\Big)k\Big(\frac{u - u_{34}}{b}\Big)\,\mathrm{d}u \\ &= \frac{1}{b}\int k\bigg(\frac{u - (X_1 + X_2)}{b}\bigg)k\bigg(\frac{u - (X_3 + X_4)}{b}\bigg)\,\mathrm{d}u, \end{aligned}$$

and

$$k^v_{1234} = \frac{1}{b}\int k\bigg(\frac{v - (X_1 - X_2)}{b}\bigg)k\bigg(\frac{v - (X_3 - X_4)}{b}\bigg)\,\mathrm{d}v.$$

Thus, under H_0, both k^u_{1234} and k^v_{1234} have the same distribution. Therefore,

$$\begin{aligned} \psi_{n2}(x_1, x_2) &\overset{d}{=} 2(28)6![E(k^u_{1324}|X_1, X_2) + E(k^u_{1234}) - E(k^u_{1234}|X_1) - E(k^u_{1234}|X_2)]^2 \\ E\psi^2_{n2}(x_1, x_2) &= (2(28)6!)^2E[E(k^u_{1324}|X_1, X_2) + E(k^u_{1234}) - E(k^u_{1234}|X_1) - E(k^u_{1234}|X_2)]^4. \end{aligned}$$

Now, we want to evaluate $E(k^u_{1234}|X_1, X_2)$.

$$E(k^u_{1234}|X_1, X_2) = \frac{1}{b}\int E\bigg[k\bigg(\frac{u - (x_1 + X_3)}{b}\bigg)k\bigg(\frac{u - (x_2 + X_4)}{b}\bigg)\bigg]\,\mathrm{d}u.$$

But

$$\begin{aligned} Ek\bigg(\frac{u - (x_1 + X_3)}{b}\bigg) &= \int k\bigg(\frac{u - (x_1 + x)}{b}\bigg)f(x)\,\mathrm{d}x \\ &= b\int k(w)f(u - x_1 - bw)\,\mathrm{d}w \\ &= b\int k(w)\bigg[f(u - x_1) - bwf'(u - x_1) + \frac{b^2w^2}{2}f''(u - x_1) + o(b^2)\bigg]\,\mathrm{d}w \\ &= bf(u - x_1) + \frac{b^2\mu_2(k)}{2}f''(u - x_1) + o(b^3). \end{aligned}$$

But, X_3 and X_4 are independent, therefore,

$$\begin{aligned} &Ek\bigg(\frac{u - (x_1 + X_3)}{b}\bigg)k\bigg(\frac{u - (x_2 + X_4)}{b}\bigg) \\ &\quad = b^2f(u - x_1)f(u - x_2) + \frac{b^2\mu_2(k)}{2}[f''(u - x_1)f(u - x_2) + f''(u - x_2)f(u - x_1)]o(b^3). \end{aligned}$$

Therefore,

$$E[k_{1324}^u|X_1,X_2] \approx b\int f(u-x_1)f(u-x_2)\,\mathrm{d}u + \frac{b^2\mu_2(k)}{2}\int v_1(x_1,x_2,u)\,\mathrm{d}u,$$

where

$$\begin{aligned} v_1(x_1,x_2,u) &= f''(u-x_1)f(u-x_2)+f''(u-x_2)f(u-x_1)\\ Ek\left(\frac{u-(X_1+X_3)}{b}\right) &= \int k\left(\frac{u-y}{b}\right)h(y)\,\mathrm{d}y\\ &= b\int k(w)h(u-bw)\,\mathrm{d}w\\ &\approx bh(u)+\frac{b^3\mu_2(k)}{2}h''(u). \end{aligned}$$

Thus,

$$\begin{aligned} Ek_{1234}^u &= b\int h^2(u)\,\mathrm{d}u + \frac{b^2\mu_2(k)}{2}\int h(u)h''(u)\,\mathrm{d}u\\ E(k_{1234}^u|X_1) &\approx b\int h(u)f(u-x_1) + \frac{b^2\mu_2(k)}{2}\int v_2(x_1,u)\,\mathrm{d}u \end{aligned}$$

and

$$E(k_{1234}^u|X_2) \approx b\int h(u)f(u-x_2) + \frac{b^2\mu_2(k)}{2}\int v_3(x_2,u)\,\mathrm{d}u,$$

where

$$v_2(x_1,u) = h(u)f''(u-x_1)h''(u)f(u-x_1)$$

and

$$v_3(x_2,u) = h(u)f''(u-x_2)h''(u)f(u-x_2).$$

Thus,

$$\begin{aligned} &E(k_{1324}^u|X_1,X_2)+Ek_{1234}^u - E(k_{1234}^u|X_1) - E(k_{1234}^u|X_2)\\ &\quad = b\left[\int f(u-x_1)f(u-x_2)\,\mathrm{d}u + \int h^2(u)\,\mathrm{d}u - \int h(u)f(u-x_1) - \int h(u)f(u-x_2)\right]\\ &\qquad + \frac{b^2\mu_2(k)}{2}\left[\int v_1(x_1,x_2,u)\,\mathrm{d}u + 2\int h(u)h''(u)\,\mathrm{d}u - v_2(x_1,u)\,\mathrm{d}u - v_3(x_2,u)\,\mathrm{d}u\right]\\ &\quad = b\left[(f*f)(x_1+x_2)+\int h^2(u)\,\mathrm{d}u - (f*h)(x_1) - (f*h)(x_2)\right]\\ &\qquad + \frac{b^2\mu_2(k)}{2}\left[2\int h(u)h''(u)\,\mathrm{d}u + 2(f*f'')(x_1+x_2) - (f*h'')(x_1) - (f*h'')(x_2)\right], \end{aligned}$$

where $(f * g)(x)$ is the convolution function defined earlier. Thus,

$$\begin{aligned}E\psi_{n2}^2(x_1, x_2) = (56 * 6!)^2 \int\int \Bigg(b\Big[(f*f)(x_1+x_2) \\ + \int h^2(u)\,\mathrm{d}u - (f*h)(x_1) - (f*h)(x_2) \Big] \\ + \frac{b^2\mu_2(k)}{2}\Big[2\int h(u)h''(u)\,\mathrm{d}u + 2(f*f'')(x_1+x_2) \\ - (f*h'')(x_1) - (f*h'')(x_2) \Big]\Bigg)^4 f(x_1)f(x_2)\,\mathrm{d}x_1\,\mathrm{d}x_2.\end{aligned}$$

For $c = 3, 4, \ldots, 8$. Define

$$\psi_{nc}(x_1, x_2, \ldots, x_c) = E[\psi_n(X_1, \ldots, X_8)|X_1 = x_1, \ldots, X_c = x_c]$$

thus,

$$\begin{aligned}E\psi_{nc}^2(x_1, x_2, \ldots, x_c) &= E(E[\psi_n(X_1, \ldots, X_8)|X_1 = x_1, \ldots, X_c = x_c]^2) \\ &\le E\psi_n^2(x_1, x_2, \ldots, x_8)\end{aligned}$$

but

$$E\psi_n^2(x_1, x_2, \ldots, x_8) \le C \sum_{8!} E(k^u_{ii^*jj^*})^2 E[k^v_{ii^*jj^*} + k^v_{ll^*rr^*} - 2k^v_{ii^*ll^*}]^2,$$

with C a constant. But, when $k'(\cdot)$ exist and is finite and using Taylor expansion, we can write

$$k^u_{ii^*jj^*} = k\left(\frac{(X_i + X_{i^*}) - (X_j + X_{j^*})}{b}\right) + O(1)$$

But, $E(k^u_{ii^*jj^*})^2 = E(E[(k^u_{ii^*jj^*})^2|X_i, X_{i^*}])$. It is easy to show that $E(k^u_{ii^*jj^*})^2 = O(b)$. Therefore, using Minkowski's inequality

$$E\psi_{nc}^2(x_1, x_2, \ldots, x_c) = O(b^2).$$

Thus, when $nb^2 \to \infty$, $\sigma_{nc}^2/\sigma_{n2}^2 = 1/O(b^2) = O(n^{c-3})$. Also, using Lemma 2.2, we have

$$\frac{EG_n^2(x_1, x_2) + (1/n)E\psi_{n2}^4(x_1, x_2)}{(E\psi_{n2}^2(X_1, x_2))^2} \to 0.$$

Thus, from Lemma B.4 of Fan and Li (1996), we have,

$$n\hat{\delta}_1\left[\frac{n^4(n-1)^4b^2}{\binom{n}{8}}\right] \to N(0, \tilde{\sigma}_{0n}^2),$$

where $\tilde{\sigma}_{0n}^2 = (56^4 6!/2)\tilde{\sigma}_n^2$ and

$$\tilde{\sigma}_n^2 = \left(b\left[(f*f)(x_1+x_2) + \int h^2(u)\,\mathrm{d}u - (f*h)(x_1) - (f*h)(x_2) \right] \right.$$
$$\left. + \frac{b^2\mu_2(k)}{2}\left[2\int h(u)h''(u)\,\mathrm{d}u + 2(f*f'')(x_1+x_2) - (f*h'')(x_1) - (f*h'')(x_2) \right] \right)^4$$
$$\times f(x_1)f(x_2)\,\mathrm{d}x_1\,\mathrm{d}x_2.$$

Thus,

$$n\hat{\delta}_1 \to N(0, \sigma_n^2),$$

where $\sigma_n^2 = (1568/b^4)\tilde{\sigma}_n^2$.

References

[1] Ahmad, I. A. and Mugdadi, A. R. (2002). Analysis of the kernel estimation of the density of functions of random variables, submitted.

[2] Ahmad, I. A. and Li, Q. (1998). Testing independence by nonparametric kernel methods. *Statistical Probability Letters*, **34**, 201–210.

[3] Ahmad, I. A. and Li, Q. (1997). Testing symmetry of an unknown density function by kernel methods. *Journal of Nonparametric Statistics*, **7**, 279–293.

[4] Fan, Y. Q. (1994). Testing the goodness of fit tests of a parametric density function by the kernel method. *Econometric Theory*, **10**, 316–356.

[5] Fan, Y. Q. and Li, Q. (1996). Consistent model specification test: Omitted variables and several parametric functions. *Econometrica*, **63**, 865–890.

[6] Green, J. R. and Hegazy, Y. A. S. (1976). Power modified EDF goodness of fit tests. *Journal of American Statistical Association*, **71**, 204–209.

[7] Hart, J. D. (1997). *Nonparametric Smoothing and the Lack of Fit Tests.* Springer-Verlag, New York, NY.

[8] Lin, C. and Mudholkar, G. (1980). A simple test for normality against asymmetric alternatives. *Biometrika*, **79**, 850–854.

[9] Li, Q. (1996). Nonparametric testing of the closeness between two unknown distribution functions. *Econometric Reviews*, **3**, 261–276.

[10] Locke, C. and Spurrier, J. (1976). The use of U-statistics for testing normality against asymmetrical alternatives. *Biometrika*, **63**, 143–147.

[11] Shapiro, S. S. and Wilk, M. B. (1965). An analysis of variance test for normality. *Biometrika*, **52**, 591–611.

[12] Vasicek, O. (1976). A test for normality based on sample entropy. *Journal of Royal Statistical Society, B*, **38**, 54–59.

[13] Wand, M. P. and Jones, M. C. (1995). *Kernel Smoothing.* Chapman Hall, New York, NY.