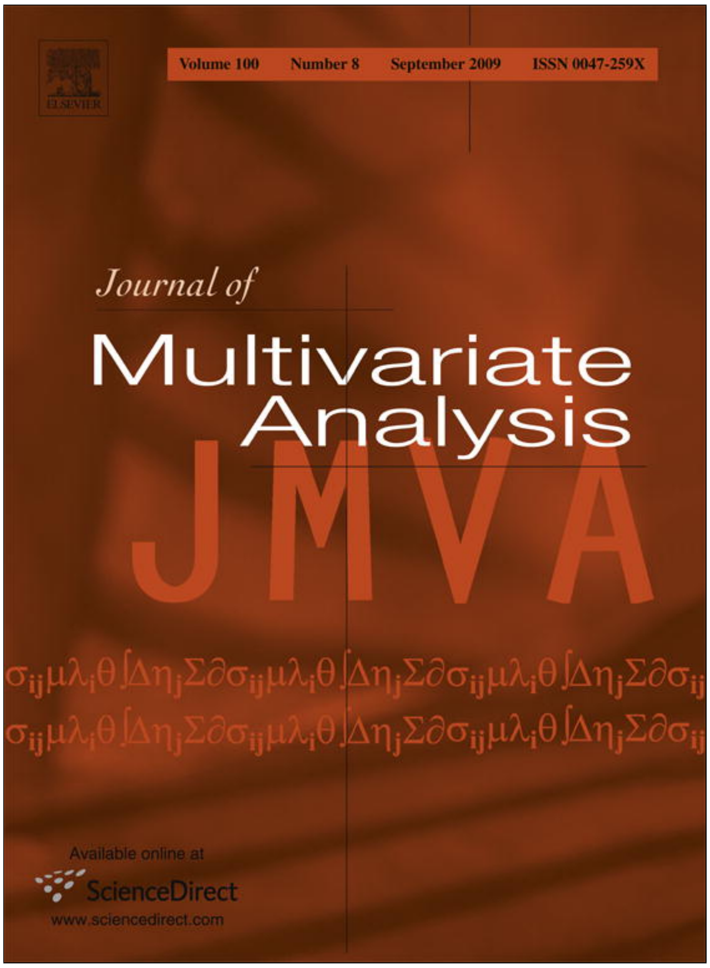
Volume 100 Number 8 September 2009 ISSN 0047-259X
Journal of
Multivariate
Analysis
JMVA
Available online at
ScienceDirect
www.sciencedirect.com

Journal of Multivariate Analysis 100 (2009) 1622–1633

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

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

Consistency of general bootstrap methods for degenerate U-type and V-type statistics

Anne Leucht, Michael H. Neumann *

Friedrich-Schiller-Universität Jena, Institut für Stochastik, Ernst-Abbe-Platz 2, D-07743 Jena, Germany

ARTICLE INFO

Article history:
Received 23 October 2007
Available online 5 February 2009

AMS 2000 subject classifications:
primary 62G09
62F40
secondary 62G20
62F05

Keywords:
Bootstrap
Consistency
U-statistics
V-statistics

ABSTRACT

We provide general results on the consistency of certain bootstrap methods applied to degree-2 degenerate statistics of U-type and V-type. While it follows from well known results that the original statistic converges in distribution to a weighted sum of centred chi-squared random variables, we use a coupling idea of Dehling and Mikosch to show that the bootstrap counterpart converges to the same distribution. The result is applied to a goodness-of-fit test based on the empirical characteristic function.

1. Introduction

The bootstrap is a well established universal tool for approximating the distribution of any statistic of interest. Sometimes its asymptotic validity for a particular purpose can be inferred from a general result; however, quite often its consistency is checked in a case-by-case manner. Available results of general type include those of Bickel and Freedman [1] for Efron's bootstrap applied to linear and related statistics, Arcones and Giné [2] for Efron's bootstrap in connection with U- and V-statistics, and Stute, González Manteiga and Presedo Quindimil [3] who showed that the bootstrap version of the empirical process with estimated parameters converges to the same limit as the original empirical process.

In this paper we derive quite general results on the asymptotic validity for bootstrap methods applied to statistics of degenerate U-type and V-type. Statistics of this type often emerge from goodness-of-fit tests; see for example [4, Sect. 2]. Under usual assumptions, the limit distribution of such a U-statistic will be that of a weighted sum of independent and centred χ^2 random variables with one degree of freedom. In a few cases this distribution is actually known; see for example [5, Sect. 7] and [6]. In an overwhelming number of cases, however, the weights in the limit random variable depend on an unknown parameter in a complicated way. Then the bootstrap offers a convenient and perhaps unrivalled way of determining critical values for tests. We adapt a method of proof originally proposed by Bickel and Freedman [1], for proving consistency of Efron's bootstrap for statistics as the sample mean or related quantities, and modified by Dehling and Mikosch [7], for Efron's bootstrap in connection with degenerate U-statistics of i.i.d. real-valued random variables. Bickel and Freedman employed the fact that Mallows' distance between the distribution of the sample mean and the distribution of its bootstrap counterpart can be estimated from above by Mallows' distance between the sample and the bootstrap distribution. Therefore, it was not necessary to re-derive the asymptotic distribution of the statistic of interest on the bootstrap side;

* Corresponding author.
E-mail addresses: anne.leucht@mathematik.uni-jena.de (A. Leucht), mneumann@mathematik.uni-jena.de (M.H. Neumann).

0047-259X/$ – see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmva.2009.01.008

rather, it sufficed to check convergence of the respective distributions at the level of individual random variables. In the case of U- or V-statistics, one cannot simply copy this scheme of proof since the summands in the statistic of interest are not independent in general. Dehling and Mikosch [7] have shown, however, that a simple coupling of the underlying random variables can be used for showing that a degenerate U-statistic and its bootstrap counterpart converge to the same limit. We extend this idea to more general bootstrap schemes and to the case of degenerate U- and V-statistics with kernels which may depend on some parameter that has to be estimated. We note that alternative ways of proving consistency for bootstrap statistics of degenerate U-type have been explored by Fan [8] and Jiménez-Gamero, Muñoz-García and Pino-Mejías [9]. In both cases the limit distribution of the original statistic was derived via a spectral decomposition of the kernel. Jiménez-Gamero et al. [9] mimicked this proof also on the bootstrap side while Fan [8] dismissed this possibility and employed empirical process arguments. While Jiménez-Gamero et al. [9] and Fan [8] imposed some regularity conditions on the parametric family of random variables involved, we try to avoid such conditions since they can hardly be checked in cases where these densities do not have a closed form; see for example our application in Section 3 below. We also note that in cases where the U- or V-statistic emerges from a Cramér–von Mises test one can alternatively use the stochastic process approach in conjunction with the continuous mapping theorem for showing consistency of the bootstrap; see for example [3]. Sometimes, however, this way turns out to be rather cumbersome and our approach then offers a simple and easily applicable alternative.

Our paper is organized as follows. In Section 2 we derive general results for the validity of model-based bootstrap schemes applied to U- and V-statistics. These results are used in Section 3 for devising a goodness-of-fit test based on the empirical characteristic function and its model-based estimate. In this particular case, there do not exist closed-form expressions for the densities of the observations and our approach seems to be actually easier than competing ones. All proofs are deferred to a final Section 4.

2. Consistency of general bootstrap methods for *U*- and *V*-statistics

Throughout this section we make the following assumption:

(A1)(i) $(X_n)_{n\in\mathbb{N}}$ is a sequence of independent, identically distributed $\mathbb{R}^d$-valued random variables, defined on a probability space $(\Omega, \mathcal{A}, P)$, with common distribution function F_θ, where $\theta \in \Theta \subseteq \mathbb{R}^p$.

(ii) The kernel $h(\cdot,\cdot;\theta)$ is measurable, symmetric in the first two arguments and degenerate under F_θ, that is, $\int h(x, y; \theta)\mathrm{d}F_\theta(x) = 0\ \forall y \in \mathbb{R}^d$.

(iii) $0 < \int h^2(x, y; \theta)\mathrm{d}F_\theta(x)\mathrm{d}F_\theta(y) < \infty$.

Remark 1. Assumption (A1) is formulated in such a way that the important case of testing composite hypotheses is accommodated; see Section 3 for an application. It suffices then that (A1) and the other assumptions below are fulfilled for the true parameter under the null hypothesis which we always denote by θ.

It is well known (see, for example, [10, p. 194] or [11, Theorem 3.2.2.1, p. 90]) that under (A1) the following result holds true ('$\xrightarrow{d}$' denotes convergence in distribution):

$$U_n = \frac{1}{n}\sum_{j=1}^{n}\sum_{k\neq j} h(X_j, X_k; \theta) \xrightarrow{d} Z := \sum_{\nu=1}^{\infty}\lambda_\nu(Z_\nu^2 - 1), \tag{2.1}$$

where $Z_1, Z_2, \ldots$ are independent standard normal random variables and the λ_ν are the eigenvalues of the integral equation

$$\int h(x, y; \theta)g(y)\mathrm{d}F_\theta(y) = \lambda g(x).$$

Furthermore, since $\sum_{\nu=1}^{\infty}\lambda_\nu^2 = Eh^2(X_1, X_2; \theta) < \infty$ (see [10, p. 197]) the infinite sum on the right-hand side of (2.1) actually converges in L_2.

We intend to derive simple criteria for the consistency of general bootstrap versions of U_n. When doing so, we have to take into account that the distribution of the bootstrap random variables $X_1^*, \ldots, X_n^*$ is random, typically converging to that of the original random variables $X_1, \ldots, X_n$. To give a clear description of our basic idea, we consider first the simpler situation where the distribution of $U_n = U_n(X_1, \ldots, X_n; \theta)$ has to be compared with that of

$$U_{nn} = \frac{1}{n}\sum_{j=1}^{n}\sum_{k\neq j} h(X_{nj}, X_{nk}; \theta_n).$$

These random variables should imitate what usually happens with the bootstrap in probability, that is, we will consider the case where $X_{n1}, \ldots, X_{nn}$ are independent with a distribution converging to that of X_1. We make the following assumption:

(A2)(i) $(X_{nj})_{j=1,\ldots,n}, n \in \mathbb{N}$, is a triangular scheme of random variables defined on respective probability spaces $(\Omega_n, \mathcal{A}_n, P_n)$, where $X_{n1}, \ldots, X_{nn}$ are independent with common distribution function F_n. Furthermore, it holds that $F_n \Longrightarrow F_\theta$ and $\theta_n \underset{n\to\infty}{\longrightarrow} \theta$. ('$\Longrightarrow$' denotes weak convergence.)

(ii) $h : \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^p \longrightarrow \mathbb{R}$ is a measurable function, symmetric in the first two arguments, the set δh of its discontinuity points fulfils $\int I((x, y, \theta) \in \delta h) \mathrm{d}F_\theta(x)\mathrm{d}F_\theta(y) = 0$, and $\int h(x, y; \theta_n)\mathrm{d}F_n(x) = 0\ \forall y \in \mathbb{R}^d$, that is, the kernel $h(\cdot, \cdot; \theta_n)$ is degenerate under F_n.

(iii) $Eh^2(X_{n1}, X_{n2}; \theta_n) \underset{n\to\infty}{\longrightarrow} Eh^2(X_1, X_2; \theta)$.

Remark 2. Under (A2)(i) and (A2)(ii), assumption (A2)(iii) is equivalent to uniform integrability of $(h^2(X_{n1}, X_{n2}; \theta_n))_{n\in\mathbb{N}}$, which will be required to prove the following result. Moreover, note that this condition can be verified by bounding moments of higher order.

Lemma 2.1. *Suppose that (A1) and (A2) are fulfilled. Then, as $n \to \infty$,*

$$U_{nn} \stackrel{d}{\longrightarrow} Z,$$

where the random variable Z is defined in Eq. (2.1) above. Moreover,

$$\sup_{-\infty<x<\infty} |P_n(U_{nn} \le x) - P(U_n \le x)| \underset{n\to\infty}{\longrightarrow} 0.$$

Now we are in a position to establish consistency for certain bootstrap methods. Suppose that bootstrap observations $X_1^*, \ldots, X_n^*$ are independently drawn (conditionally on $X_1, \ldots, X_n$) from some estimate $\widehat{F}_n$ of the unknown distribution function F_θ. A minimal property that one usually expects for such a resampling scheme is that $\widehat{F}_n \Longrightarrow F_\theta$, in probability or almost surely. Furthermore, we also assume that $\widehat{\theta}_n$ is a consistent estimator of θ. (In the case of a model-based bootstrap method the X_j^* will be typically drawn from $F_{\widehat{\theta}_n}$; however, we do not require this here.) The hope is that U_n is now consistently mimicked by

$$U_n^* = \frac{1}{n}\sum_{j=1}^{n}\sum_{k\neq j} h(X_j^*, X_k^*; \widehat{\theta}_n).$$

We derive consistency of the bootstrap under the following assumption:

(A2*)(i) The random variables $X_1^*, \ldots, X_n^*$ are independent (conditionally on $X_1, \ldots, X_n$) and identically distributed with $X_1^* \stackrel{d}{\longrightarrow} X_1$ in probability. Moreover, $\widehat{\theta}_n \stackrel{P}{\longrightarrow} \theta$. ('$\stackrel{P}{\longrightarrow}$' denotes convergence in probability.)

(ii) $h : \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^p \longrightarrow \mathbb{R}$ is a measurable function, symmetric in the first two arguments, the set δh of its discontinuity points fulfils $\int I((x, y, \theta) \in \delta h)\mathrm{d}F_\theta(x)\mathrm{d}F_\theta(y) = 0$, and $E(h(X_1^*, y; \widehat{\theta}_n) \mid X_1, \ldots, X_n) = 0\ \forall y \in \mathbb{R}^d$, that is, the kernel $h(\cdot, \cdot; \widehat{\theta}_n)$ is degenerate for X_1^*.

(iii) $E\left(h^2(X_1^*, X_2^*; \widehat{\theta}_n) \mid X_1, \ldots, X_n\right) \stackrel{P}{\longrightarrow} Eh^2(X_1, X_2, \theta)$.

Theorem 2.1. *Suppose that* (A1) *and* (A2*) *are fulfilled. Then, as $n \to \infty$,*

$$U_n^* \stackrel{d}{\longrightarrow} Z \quad \text{in probability},$$

where Z is defined in (2.1) above. Moreover,

$$\sup_{-\infty<x<\infty} \left|P(U_n^* \le x \mid X_1, \ldots, X_n) - P(U_n \le x)\right| \stackrel{P}{\longrightarrow} 0.$$

This theorem is actually an immediate consequence of Lemma 2.1. To see this, note that (A2*) is just an "in probability version" of assumption (A2). Consequently, the convergence results from Lemma 2.1 appear in Theorem 2.1 as convergence results in probability.

Remark 3. Jiménez-Gamero et al. [9] derived a similar result by a different method of proof. They studied the special case where the kernel has the form $h(x, y; \theta) = \int q(x, t; \theta)q(y, t; \theta)\mathrm{d}G_\theta(t)$, for some function q and some finite measure G_θ. Under smoothness conditions on q, G_θ and the densities of the random variables involved, they showed that the eigenvalues of some operators connected with $h(X_j^*, X_k^*; \widehat{\theta}_n)$ converge to those of $h(X_j, X_k; \theta)$, which leads to the desired bootstrap consistency.

Although U-statistics seem to dominate in the probability literature over V-statistics, they rarely occur in statistical applications. Gregory [12, p. 115] mentioned a few cases where it might be preferable to use U-statistics rather than V-statistics for testing certain hypotheses. In most cases, however, the statistic of interest is of V-type or can be approximated by a V-statistic. We consider

$$V_n = \frac{1}{n}\sum_{j,k=1}^{n} h(X_j, X_k; \theta).$$

To derive the limit distribution of V_n, we will assume (A1) and, additionally, that

(A3) $E|h(X_1, X_1; \theta)| < \infty$.

Now we obtain from the strong law of large numbers that

$$\frac{1}{n}\sum_{j=1}^{n} h(X_j, X_j; \theta) \longrightarrow Eh(X_1, X_1; \theta) \qquad P\text{-a.s.},$$

which implies by (2.1) that

$$V_n \xrightarrow{d} Z + Eh(X_1, X_1; \theta). \tag{2.2}$$

To study consistency of the bootstrap counterpart to V_n, we establish first a V-statistic version of Lemma 2.1. We set

$$V_{nn} = \frac{1}{n}\sum_{j,k=1}^{n} h(X_{nj}, X_{nk}; \theta_n).$$

Besides (A1)–(A3), we will also make the following assumption:

(A4) The set δh of discontinuity points of h fulfils $\int I((x, x, \theta) \in \delta h) \mathrm{d}F_\theta(x) = 0$ and it holds that $Eh(X_{n1}, X_{n1}; \theta_n) \underset{n\to\infty}{\longrightarrow} Eh(X_1, X_1; \theta)$.

Lemma 2.2. *Suppose that* (A1)–(A4) *are fulfilled. Then, as* $n \to \infty$,

$$V_{nn} \xrightarrow{d} Z + Eh(X_1, X_1; \theta)$$

and

$$\sup_{-\infty<x<\infty} |P_n(V_{nn} \le x) - P(V_n \le x)| \underset{n\to\infty}{\longrightarrow} 0.$$

Now we can easily identify sufficient conditions for V_n^* converging to the same limit as V_n. We will assume that:

(A4*) The set δh of discontinuity points of h fulfils $\int I((x, x, \theta) \in \delta h) \mathrm{d}F_\theta(x) = 0$ and it holds that $E(h(X_1^*, X_1^*, \widehat{\theta}_n) \mid X_1, \ldots, X_n) \xrightarrow{P} Eh(X_1, X_1; \theta)$.

Now we obtain the desired general consistency theorem for bootstrap versions of a V-statistic.

Theorem 2.2. *Suppose that* (A1), (A2*), (A3) *and* (A4*) *are fulfilled. Then, as* $n \to \infty$,

$$V_n^* \xrightarrow{d} Z + Eh(X_1, X_1; \theta) \quad \textit{in probability},$$

where Z is defined in (2.1) *above. Moreover,*

$$\sup_{-\infty<x<\infty} \left|P(V_n^* \le x \mid X_1, \ldots, X_n) - P(V_n \le x)\right| \xrightarrow{P} 0.$$

As Theorem 2.1 follows from Lemma 2.1, this theorem follows immediately from Lemma 2.2 since the assumptions here are again an "in probability version" of the assumptions of the corresponding previous lemma.

3. A goodness-of-fit test for the NIG model

3.1. Theoretical results

The famous Black–Scholes option pricing model in financial mathematics is based on the assumption of log-normal asset returns. Empirical studies have provided evidence, however, that the distribution of logarithmic returns are negative skewed and heavy tailed. In addition, jumps are possible which cannot be described by a stochastic process with continuous sample paths. Therefore, the assumption of normal distributed logarithmic returns has to be seen as very critical. Barndorff-Nielsen [13] proposed modelling the process of logarithmic asset prices $(Y(t))_{t\ge 0}$ by a normal inverse Gaussian (NIG hereafter) process of Lévy type. An NIG process with parameters α, β, μ and δ is a particular Lévy process where $Y(t)$ has a normal inverse Gaussian density $f_{NIG}(\cdot; \alpha, \beta, \mu t, \delta t)$,

$$f_{NIG}(x; \alpha, \beta, \mu t, \delta t) = \frac{\alpha\delta t}{\pi} \exp\left(\delta t\sqrt{\alpha^2 - \beta^2} + \beta(x - \mu t)\right) \frac{K_1(\alpha\sqrt{(\delta t)^2 + (x - \mu t)^2})}{\sqrt{(\delta t)^2 + (x - \mu t)^2}}.$$

Here $\alpha, \delta > 0$, $\beta \in (-\alpha, \alpha)$, $\mu \in \mathbb{R}$, and K_1 denotes the modified Bessel function of third order and index 1, that is, $K_1(z) = (1/2)\int_0^\infty \exp(-z(u + (1/u))/2)\mathrm{d}u$.

We assume that we observe the asset prices at equidistant time points $0, \Delta, 2\Delta, \dots, n\Delta$ and we intend to test the composite hypothesis that the underlying process is of NIG type. Due to the scaling property of Lévy processes we can assume, without loss of generality, that $\Delta = 1$. Under the null hypothesis, the increments $X_j = Y(j) - Y(j-1)$ ($j = 1, \dots, n$) are independent and have a NIG distribution with parameters α, β, μ and δ. While their common density is rather complicated their characteristic function has the following simple closed form (see equation (3.7) in [13]):

$$c(t; \alpha, \beta, \mu, \delta) = \exp\left\{ i\mu t + \delta\left(\sqrt{\alpha^2 - \beta^2} - \sqrt{\alpha^2 - (\beta + it)^2}\right)\right\}. \tag{3.1}$$

Cont and Tankov [14] describe a NIG process as a subordinated Brownian motion with drift $\vartheta \in \mathbb{R}$, volatility $\sigma > 0$ and with an inverse Gaussian process with variance κ at time 1 as subordinator. Allowing an additional drift term $\mu \in \mathbb{R}$ we have the following alternative, and in some sense more convenient, parametrization of the characteristic function:

$$c(t; \kappa, \mu, \sigma, \vartheta) = \exp\left\{ i\mu t + \left(1 - \sqrt{1 + \sigma^2\kappa t - 2i\vartheta\kappa t}\right) \Big/ \kappa \right\}, \tag{3.2}$$

where μ has the same value as before, and

$$\kappa = \frac{1}{\delta(\alpha^2 - \beta^2)^{1/2}}, \quad \vartheta = \frac{\delta\beta}{(\alpha^2 - \beta^2)^{1/2}} \quad \text{and} \quad \sigma^2 = \frac{\delta}{(\alpha^2 - \beta^2)^{1/2}}. \tag{3.3}$$

The test problem can be formulated in terms of the characteristic function c of the increments X_j as

$$H_0 : c = c(\cdot; \theta) \quad \text{for some } \theta \in \Theta \quad \text{vs.} \quad H_1 : c \neq c(\cdot; \theta) \quad \text{for all } \theta \in \Theta,$$

where $\Theta = \left\{\theta = (\kappa, \mu, \sigma, \vartheta)' \mid \mu, \vartheta \in \mathbb{R}, \kappa, \sigma > 0\right\}$. We consider the following test statistic:

$$T_n = n \int_{\mathbb{R}} \left| \widehat{c}_n(t) - c(t; \widehat{\theta}_n)\right|^2 w(t)\mathrm{d}t, \tag{3.4}$$

where $\widehat{c}_n(t) = n^{-1}\sum_{j=1}^n \exp(itX_j)$ is the empirical characteristic function of the increments, $\widehat{\theta}_n$ is some estimator of θ and $w : \mathbb{R} \longrightarrow [0, \infty)$ some weight function. For the latter, which is employed to ensure convergence of the integral, we assume that

(A5) The function w is measurable, satisfies $\int_{\mathbb{R}} (1 + |t|)^4 w(t)\mathrm{d}t < \infty$ and vanishes only on a set of Lebesgue measure zero.

Note that the integral in (3.4) and all other integrals below have to be interpreted in the Lebesgue sense. According to Lemma 1 in [15], the test statistic can also be written in the following form which is particularly suitable for computations:

$$T_n = \frac{1}{n}\sum_{j=1}^n \sum_{k=1}^n H(X_j, X_k; \widehat{\theta}_n),$$

where $H(x, y, \theta) = u(x-y) - \int u(x-y)\mathrm{d}F_\theta(x) - \int u(x-y)\mathrm{d}F_\theta(y) + \int u(x-y)\mathrm{d}F_\theta(x)\mathrm{d}F_\theta(y)$ and $u(x) = \int \cos(xt)w(t)\mathrm{d}t$.

Remark 4. (i) The integrability condition on w ensures the negligibility of a remainder term in a representation of T_n; see (3.7). The fact that w vanishes only on a set of measure zero is primarily needed for the consistency of the test; see Proposition 3.3. It also yields that $Eh^2(X_1, X_2; \theta) > 0$ which is required for the nondegeneracy of the limiting distribution of T_n.

(ii) There are other aspects which may motivate some particular choice of the weight function. Most importantly, it is well known that the choice of w directs the power of the test; see, for example, [8, Sections 2 and 4] and the simulations reported in [16, Section 4]. In the particular case of testing for normality, Henze and Wagner [16] showed that the choice of w equal to a normal density leads to an easily computable alternative representation of the test statistic.

For definiteness, and since computation of a maximum likelihood estimator is difficult in this context, we restrict our attention to a method-of-moments estimator $\widehat{\theta}_n$ of θ. Under the null hypothesis, the distribution of the increments X_j has the following characteristics (see [14]):

- mean: $\mu + \vartheta$,
- variance: $\sigma^2 + \vartheta^2\kappa$,
- third cumulant: $3\sigma^2\vartheta\kappa + 3\vartheta^3\kappa^2$,
- fourth cumulant: $3\sigma^4\kappa + 15\vartheta^4\kappa^3 + 18\sigma^2\vartheta^2\kappa^2$.

Accordingly, we obtain $\widehat{\theta}_n = (\widehat{\kappa}_n, \widehat{\mu}_n, \widehat{\sigma}_n, \widehat{\vartheta}_n)'$ by equating the first four theoretical moments with their empirical counterparts. It may happen that $\widehat{\theta}_n$ falls outside the parameter space Θ. If $\widehat{\kappa}_n$ or $\widehat{\sigma}_n^2$ attain negative values we simply set them to zero. To accommodate these distributions, we enlarge the parameter space to the closure of Θ, that is, we take $\bar{\Theta} = \{(\kappa, \mu, \sigma, \vartheta)' : \mu, \vartheta \in \mathbb{R}, \kappa, \sigma \geq 0\}$ and define the distribution on the boundary of $\bar{\Theta}$ as corresponding limits, that is with characteristic functions

$$c(t; 0, \mu, \sigma, \vartheta) = \exp\left\{i(\mu + \vartheta)t - \sigma^2 t^2/2\right\}, \quad \text{for } \kappa = 0, \sigma \geq 0,$$

and

$$c(t;\kappa,\mu,0,\vartheta)=\exp\left\{\mathrm{i}\mu t+(1-\sqrt{1-2i\vartheta\kappa t})/\kappa\right\},\quad \text{for } \kappa>0,\sigma=0.$$

(The latter characteristic function is that of a random variable $\mu+\vartheta Z$, where $Z\sim \mathrm{IG}(1,1/\kappa)$; see [17, p. 263].)

It can be shown that θ is a twice-differentiable function of the first four theoretical moments and vice versa that these moments are continuous functions of θ. Similar results are obtained for $\widehat{\theta}_n$ applying the empirical moments instead. Thus, by Taylor expansion, there exists some function $g:\mathbb{R}^4\times\mathbb{R}\longrightarrow\mathbb{R}^4$ such that, under the null hypothesis,

$$\widehat{\theta}_n-\theta=\frac{1}{n}\sum_{j=1}^{n}g(\theta;X_j)+o_P(n^{-1/2}) \tag{3.5}$$

and $Eg(\theta;X_j)=0$, $E\|g(\theta;X_j)\|_{l_2}^2<\infty$, where $\|x\|_{l_2}=\sqrt{\sum_i x_i^2}$. This implies in particular that

$$\widehat{\theta}_n-\theta=O_P(n^{-1/2}). \tag{3.6}$$

Since c is twice differentiable with respect to θ a Taylor series expansion leads to

$$T_n=n\int_{\mathbb{R}}\left|\widehat{c}_n(t)-c(t;\theta)-Dc(t;\theta)(\widehat{\theta}_n-\theta)-\frac{1}{2}(\widehat{\theta}_n-\theta)'D^2c(t;\bar{\theta}(t))(\widehat{\theta}_n-\theta)\right|^2 w(t)\mathrm{d}t,$$

for some $\bar{\theta}(t)$ between $\widehat{\theta}_n$ and θ, where Dc and D^2c denote the gradient (row) vector and the Hessian matrix of c with respect to θ, respectively. Since assumption (A5) ensures w-integrability of $\|D^2c(t;\bar{\theta}(t))\|^2$ ($\|A\|^2=\lambda_{\max}(A'A)$) we obtain by (3.6) that

$$T_n=I_n+o_P(1), \tag{3.7}$$

where $I_n=n\int_{\mathbb{R}}\left|\widehat{c}_n(t)-c(t;\theta)-Dc(t;\theta)(\widehat{\theta}_n-\theta)\right|^2 w(t)\mathrm{d}t$. Furthermore, (3.5) allows us to approximate further:

$$T_n=I_n^0+o_P(1),$$

where $I_n^0=\int_{\mathbb{R}}\left|n^{-1/2}\sum_{j=1}^n\left\{\exp(\mathrm{i}tX_j)-c(t;\theta)-Dc(t;\theta)g(\theta;X_j)\right\}\right|^2 w(t)\mathrm{d}t$. The statistic I_n^0 is a degenerate V-statistic with a symmetric kernel

$$h(x,y;\theta)=\int_{\mathbb{R}}\{\exp(\mathrm{i}tx)-c(t;\theta)-Dc(t;\theta)g(\theta;x)\}\{\exp(-\mathrm{i}ty)-c(-t;\theta)-Dc(-t;\theta)g(\theta;y)\}\,w(t)\mathrm{d}t.$$

We have obviously that $Eh^2(X_1,X_2;\theta)<\infty$ and $E|h(X_1,X_1;\theta)|<\infty$. Moreover, since the weight function w vanishes only on a set of Lebesgue measure zero it follows that $Eh^2(X_1,X_2;\theta)>0$. Hence, assumptions (A1) and (A3) from Section 2 are fulfilled and we obtain from (2.2) the following proposition.

Proposition 3.1. *Suppose that $X_1,\ldots,X_n$ are the increments of a NIG process with parameter θ, that is, they are independent with a common characteristic function $c(\cdot;\theta)$. Furthermore, the weight function w is chosen such that* (A5) *is fulfilled. Then*

$$T_n\xrightarrow{d}\sum_{\nu=1}^{\infty}\lambda_\nu(Z_\nu^2-1)+Eh(X_1,X_1;\theta),$$

where $Z_1,Z_2,\ldots$ are independent standard normal random variables and the λ_ν are the eigenvalues of the equation

$$E\left[h(x,X_1;\theta)g(X_1)\right]=\lambda g(x).$$

To implement a test which has asymptotically a prescribed size γ, we still have to determine an appropriate critical value. This can hardly be done on the basis of the asymptotic result in Proposition 3.1 alone since the limit distribution depends on the eigenvalues λ_ν which in turn depend on the true parameter θ in a complicated way. Therefore, we propose the following bootstrap procedure:

(1) Given $X_1,\ldots,X_n$, estimate θ by $\widehat{\theta}_n$ and generate a sample $X_1^*,\ldots,X_n^*$ with a characteristic function $c(\cdot;\widehat{\theta}_n)$. According to [14, p. 184], a random variable X with characteristic function $c(\cdot;\kappa,\mu,\sigma,\vartheta)$ can be generated by drawing first a random variable Z with an inverse Gaussian distribution with parameters $1/\kappa$ and 1, and then, conditioned on $Z=z$, drawing $X\sim\mathcal{N}(\mu+\vartheta z,\sigma^2 z)$. An algorithm for simulating an inverse Gaussian distributed random variable is stated there as well. The bootstrap sample conditioned on $X_1,\ldots,X_n$ satisfies H_0 with $\widehat{\theta}_n$ instead of θ.
(2) Define a bootstrap counterpart $\widehat{\theta}_n^*$ to $\widehat{\theta}_n$ which is based on the bootstrap sample by the method of moments and set $\widehat{c}_n^*(t)=n^{-1}\sum_{j=1}^n\exp(\mathrm{i}tX_j^*)$. Then

$$T_n^*=n\int_{\mathbb{R}}\left|\widehat{c}_n^*(t)-c(t;\widehat{\theta}_n^*)\right|^2 w(t)\mathrm{d}t$$

is the bootstrap version of our test statistic T_n.

(3) Define the critical value t_γ^* as the $(1-\gamma)$-quantile of the (conditional) distribution of T_n^*. (In practice, steps (1) and (2) will be repeated B times, for some large B. The critical value will then be approximated by the $(1-\gamma)$-quantile of the empirical distribution associated with $T_{n1}^*, \ldots, T_{nB}^*$.)

To justify this approach, we will briefly argue that the conditional distribution of T_n^* given $X_1, \ldots, X_n$ converges under the null hypothesis (this time even P-almost surely) to the same limit as that of T_n. First, it follows from the strong law of large numbers that

$$\widehat{\theta}_n \xrightarrow{P\text{-a.s.}} \theta. \tag{3.8}$$

Analogously to (3.5), we obtain that

$$\widehat{\theta}_n^* - \widehat{\theta}_n = \frac{1}{n}\sum_{j=1}^n g(\widehat{\theta}_n; X_j^*) + R_n^*, \tag{3.9}$$

where

$$P\left(|R_n^*| > \epsilon n^{-1/2} \mid X_1, \ldots, X_n\right) \xrightarrow{P\text{-a.s.}} 0 \qquad \forall \epsilon > 0.$$

(3.8) and (3.9) yield that

$$P\left(|T_n^* - I_n^{0*}| > \epsilon \mid X_1, \ldots, X_n\right) \xrightarrow{P\text{-a.s.}} 0 \quad \forall \epsilon > 0,$$

where $I_n^{0*} = n^{-1}\sum_{j,k=1}^n h(X_j^*, X_k^*; \widehat{\theta}_n)$. It follows from the construction that

$$E\left(h(X_1^*, y; \widehat{\theta}_n) \mid X_1, \ldots, X_n\right) = 0 \quad \forall y \in \mathbb{R}.$$

And finally, we obtain from (3.8) that $E(h^2(X_1^*, X_2^*; \widehat{\theta}_n) \mid X_1, \ldots, X_n) \xrightarrow{P\text{-a.s.}} Eh^2(X_1, X_2; \theta)$ and $E(h(X_1^*, X_1^*; \widehat{\theta}_n) \mid X_1, \ldots, X_n) \xrightarrow{P\text{-a.s.}} Eh(X_1, X_1; \theta)$. To summarize we have verified that the assumptions (A1), (A2*), (A3) and (A4*) are fulfilled, even P-almost surely rather than in probability. Hence, we obtain from Theorem 2.2 the following assertion.

Proposition 3.2. *Suppose that the assumptions of Proposition 3.1 are fulfilled. Then*

$$\sup_{-\infty<x<\infty} \left|P(T_n^* \le x \mid X_1, \ldots, X_n) - P(T_n \le x)\right| \xrightarrow{P\text{-a.s.}} 0.$$

For $t_\gamma^* = \inf\{t : P(T_n^* \le t \mid X_1, \ldots, X_n) \ge 1-\gamma\}$, we obtain that

$$P\left(T_n > t_\gamma^*\right) \underset{n\to\infty}{\longrightarrow} \gamma,$$

that is, the test has asymptotically the correct size.

It remains to investigate the behaviour of the test under the alternative. For this, we assume that $(X_n)_{n\in\mathbb{N}}$ is a sequence of independent and identically distributed random variables with characteristic function $c \notin \{c(\cdot;\theta) : \theta \in \bar{\Theta}\}$ and $EX_1^4 < \infty$.

We consider first the asymptotic behaviour of the test statistic T_n. It follows from the strong law of large numbers that

$$\frac{1}{n}\sum_{j=1}^n X_j^k \xrightarrow{P\text{-a.s.}} EX_1^k, \quad \text{for } k = 1, \ldots, 4,$$

which implies that

$$\widehat{\theta}_n \xrightarrow{P\text{-a.s.}} \bar{\theta}_0, \tag{3.10}$$

where $\bar{\theta}_0 = (\kappa_0 \vee 0, \mu_0, \sqrt{\sigma_0^2 \vee 0}, \vartheta_0)'$ and $(\kappa_0, \mu_0, \sigma_0^2, \vartheta_0)'$ solves the equations for the first four moments. Since $c \notin \{c(\cdot;\theta) : \theta \in \bar{\Theta}\}$ we obtain that

$$\int_{\mathbb{R}} |c(t) - c(t; \bar{\theta}_0)|^2 w(t)\mathrm{d}t =: K > 0.$$

Moreover, it follows from the Glivenko–Cantelli theorem that $P(\widehat{c}_n(t) \underset{n\to\infty}{\longrightarrow} c(t) \text{ for all } t \in \mathbb{R}) = 1$, which implies by majorized convergence that

$$\int_{\mathbb{R}} |\widehat{c}_n(t) - c(t)|^2 w(t)\mathrm{d}t \xrightarrow{P\text{-a.s.}} 0.$$

Table 3.1
Estimated NIG parameters.

	α	β	μ	δ
Dresdner Bank	68.28	1.81	0	0.01
NYSE	136.29	−8.95	0.00079	0.0059

And finally, it follows from (3.10) and again by majorized convergence that

$$\int_{\mathbb{R}} |c(t;\bar{\theta}_0) - c(t;\widehat{\theta}_n)|^2 w(t)\mathrm{d}t \overset{P\text{-a.s.}}{\longrightarrow} 0. \tag{3.11}$$

Hence, we obtain that

$$T_n/n \overset{P\text{-a.s.}}{\longrightarrow} K > 0. \tag{3.12}$$

For the consistency of the test, it remains to show that T_n^*/n tends to zero in some appropriate sense. If the almost sure limit of $\widehat{\theta}_n$, $\bar{\theta}_0$, lies in the open set Θ, then the bootstrap imitates one of the cases belonging to H_0. In this case, we could simply employ Proposition 3.2 to show that T_n^* is bounded in probability, P-almost surely. However, if $\bar{\theta}_0$ lies on the boundary of Θ, then it is no longer guaranteed that $\widehat{\theta}_n^*$ has an asymptotic behaviour as described by (3.9). In this case, T_n^* does not have an asymptotic behaviour as in one of the cases belonging to the null hypothesis. Therefore, we have to prove asymptotic smallness in a different way; see the proof of the following proposition. This proposition states consistency of the test against essentially all alternatives.

Proposition 3.3. *Assume that $(X_n)_{n\in\mathbb{N}}$ is a sequence of independent and identically distributed random variables with characteristic function $c \notin \{c(\cdot;\theta) : \theta \in \bar{\Theta}\}$ and $EX_1^4 < \infty$. Then*

$$P\left(T_n > t_\gamma^*\right) \underset{n\to\infty}{\longrightarrow} 1.$$

Remark 5. There is already a large body of literature on goodness-of-fit tests of the type (3.4). Epps and Pulley [18] proposed a test of normality of this form where the weight function w was chosen in such a way that the test statistic was invariant to changes in the location and variance. This allowed one to determine a critical value via Monte Carlo simulations. Baringhaus and Henze [19] generalized this idea to the multivariate case and derived the limiting null distribution of the test statistic on the basis of a result of de Wet and Randles [20]. Note, however, that this does not automatically provide appropriate critical values since the limiting distribution still involves parameters which are only implicitly given as eigenvalues of a certain integral operator. Csörgő [21] showed for a slightly modified version of these tests that they are consistent to all alternatives. Using empirical process theory Henze and Wagner [16] derived an alternative representation of the limiting distribution which allowed one to study the power under local alternatives; see also [22] for a continuation of this investigation. Epps [23] proposed again a generalization to general location–scale families and worked out simple expressions for the test statistic in several cases. Jiménez-Gamero, Muñoz-García and Pino-Mejías [24] derived an analogous test for the distribution of the errors in a linear model.

Fan [8] derived asymptotic theory for such goodness-of-fit tests in a very general framework, not restricted to particular families of distributions or location–scale families. Since the null distribution of the test statistic depends then on the particular parameter under the null hypothesis, he proposed determining a critical value by a bootstrap method and proved consistency of this approximation. Bootstrap theory for tests of this type was also provided by Jiménez-Gamero et al. [9]. Note that we cannot apply their results to the particular problem studied here since these authors imposed some regularity conditions on the densities of the parametric families of random variables involved. In the case of the NIG model we do not have closed-form expressions of these densities and it seems to be difficult if not impossible to verify these conditions.

3.2. Numerical results

To investigate the behaviour of the proposed test in a practically relevant situation we tested our procedure on simulated data sets corresponding to log returns data of the NYSE Composite Index and the stock of Dresdner Bank. Rydberg [25, p. 906] provided an estimate of the NIG parametrization for the Dresdner Bank stock, based on 1562 observations while Albrecher and Predota [26] estimated these parameters for the NYSE Composite Index; see Table 3.1.

A translation of the values provided in Table 3.1 to the parametrization given by Cont and Tankov [14] is given in Table 3.2.

In order to get an idea of the actual size of our test we simulated data from NIG processes with parameters equal to the empirical values found for the Dresdner Bank stock and the NYSE Composite Index. In both cases, we took samples of size $n = 500$ and $n = 1500$, respectively. We replicated the simulation procedure $N = 500$ times, each time with $B = 1000$ bootstrap resamplings. The weight function w was chosen as $w(t) = \mathrm{e}^{-t^2/2000000}$ which puts sufficient weight on the region where the absolute value of the characteristic function of the underlying sample is significantly greater than zero. Of course,

Table 3.2
Reparametrization.

	κ	μ	σ	ϑ
Dresdner Bank	1.46507	0	0.01210	0.00027
NYSE	1.24630	0.00079	0.00659	−0.00039

Table 3.3
Rejection frequencies.

	$\gamma = 0.05$		$\gamma = 0.10$	
	$n = 500$	$n = 1500$	$n = 500$	$n = 1500$
Dresdner Bank	0.070	0.052	0.106	0.112
NYSE	0.064	0.066	0.118	0.118

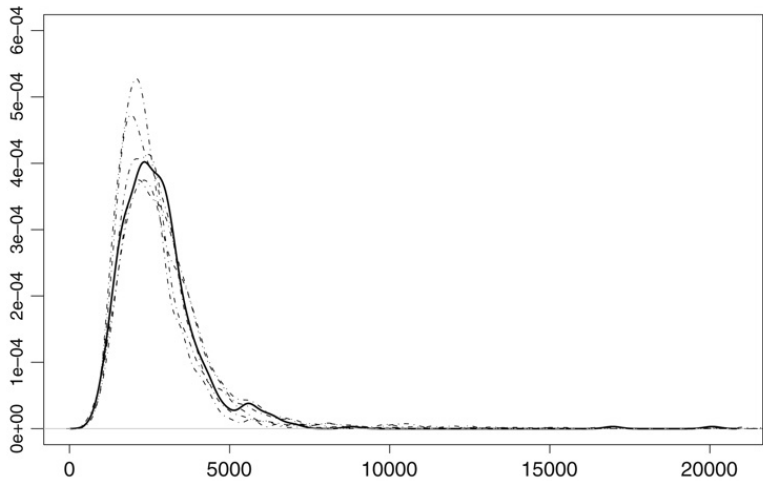

Fig. 1. Simulated density of T_n and 5 bootstrap replications (Dresdner Bank).

Table 3.4
Estimated parameters for the hyperbolic distribution.

	α	β	δ	μ
BASF	108.82	1.3550	0.0014	−0.0005
Dresdner Bank	110.94	−0.1123	0.0016	0.0002

instead of using a normal (quasi-)density other alternatives are possible. As stated in Remark 4 above, the choice of the weight function influences the power of the test.

The implementation was done on the basis of the statistical software package R; see [27]. The results for the rejection frequencies of our test for nominal significance levels $\gamma = 0.05$ and $\gamma = 0.1$ are shown in Table 3.3. Note that two-sided confidence intervals with coverage probability $(1 - \beta) = 0.95$ are given by $\widehat{\gamma} \pm 0.0191$ and $\widehat{\gamma} \pm 0.0263$, for $\gamma = 0.05$ and $\gamma = 0.1$, respectively.

Finally in Fig. 1 we plot the density of the test statistic (thick line) and five bootstrap replications (dot–dashed lines) for the Dresdner Bank setting with sample size $n = 1500$. This indicates that the bootstrap distributions are fairly similar to the distribution of T_n under the null hypothesis.

To get an impression of the power of our test we set out to simulate a situation corresponding to a reasonable alternative in the above context. Eberlein and Keller [28] proposed the hyperbolic Lévy process in order to model the logarithmic stock prices, meaning that the log returns X_j $(j = 1, \ldots, n)$ are i.i.d. with density

$$f_{hyp}(x; \alpha, \beta, \delta, \mu) = \frac{\sqrt{\alpha^2 - \beta^2}}{2\alpha\delta K_1(\delta\sqrt{\alpha^2 - \beta^2})} \exp\left(-\alpha\sqrt{\delta^2 + (x-\mu)^2} + \beta(x-\mu)\right),$$

where $\alpha < 0,\ 0 \le |\beta| < \alpha$ and $\delta, \mu \in \mathbb{R}$. These authors provided parameter estimates for ten German stocks based on 745 observations. Accordingly, we simulated hyperbolically distributed samples of size $n = 500$ and $n = 1500$, respectively, using the parameter estimations for BASF and Dresdner Bank of [28]; see Table 3.4.

As in the case of H_0, the simulation procedure was replicated $N = 500$ times, each time with $B = 1000$ bootstrap resamplings. The resulting empirical powers of our test are given in Table 3.5.

As the Tables 3.3 and 3.5 report, the proposed bootstrap test seems to be reasonable for applications with large sample sizes. Naturally, financial markets supply the designated samples, for instance in terms of daily stock prices.

Table 3.5
Empirical power.

	$\gamma = 0.05$		$\gamma = 0.10$	
	$n = 500$	$n = 1500$	$n = 500$	$n = 1500$
BASF	0.374	0.718	0.492	0.814
Dresdner Bank	0.328	0.692	0.460	0.778

4. Proofs

The main idea of the proof of Lemma 2.1 is similar to that of Theorem 2.1 in [7] where a quantile coupling for the underlying real-valued random variables was used, there for Efron's bootstrap. Here we additionally allow for other bootstrap schemes and for kernels which depend on the unknown parameter θ. Hence, the proof requires some modifications. For the reader's convenience we decided to give a complete proof of this lemma here.

Proof of Lemma 2.1. According to the Skorohod representation theorem (Theorem 6.7 in [29, p. 70]), there exists a sufficiently rich probability space $(\widetilde{\Omega}, \widetilde{\mathcal{A}}, \widetilde{P})$ with $\widetilde{\Omega} = \{(\omega_1, \omega_2, \ldots) : \omega_i \in \Omega_0\}$ and independent random elements $\omega_1, \omega_2, \ldots$ such that there are functions $g : \Omega_0 \longrightarrow \mathbb{R}^d$ and $g_n : \Omega_0 \longrightarrow \mathbb{R}^d$ with $\widetilde{X}_j := g(\omega_j) \sim F_\theta, \widetilde{X}_{nj} := g_n(\omega_j) \sim F_n$ and

$$\widetilde{X}_{nj} - \widetilde{X}_j \xrightarrow{\widetilde{P}\text{-a.s.}} 0, \tag{4.1}$$

as $n \to \infty$. (In the special case of real-valued random variables we could simply use a quantile transform to construct such a coupling.)

It is clear from the construction that $\widetilde{X}_1, \ldots, \widetilde{X}_n$ are independent and have the same distribution as the X_j under P; analogously, $\widetilde{X}_{n1}, \ldots, \widetilde{X}_{nn}$ are independent and have the same distribution as the X_{nj} under P_n. Therefore,

$$\widetilde{U}_n = \frac{1}{n} \sum_{j=1}^{n} \sum_{k \neq j} h(\widetilde{X}_j, \widetilde{X}_k; \theta)$$

has under $\widetilde{P}$ the same distribution as U_n under P, and

$$\widetilde{U}_{nn} = \frac{1}{n} \sum_{j=1}^{n} \sum_{k \neq j} h(\widetilde{X}_{nj}, \widetilde{X}_{nk}; \theta_n)$$

has under $\widetilde{P}$ the same distribution as U_{nn} under P_n.

It follows from (4.1), by $\theta_n \underset{n\to\infty}{\longrightarrow} \theta$ and the assumed continuity property of h, that, for $1 \leq j, k \leq n$,

$$h(\widetilde{X}_{nj}, \widetilde{X}_{nk}; \theta_n) - h(\widetilde{X}_j, \widetilde{X}_k; \theta) \xrightarrow{\widetilde{P}\text{-a.s.}} 0. \tag{4.2}$$

Moreover, we obtain from $Eh^2(\widetilde{X}_{n1}, \widetilde{X}_{n2}; \theta_n) \underset{n\to\infty}{\longrightarrow} Eh^2(\widetilde{X}_1, \widetilde{X}_2; \theta)$ in conjunction with (4.2) that $(h^2(\widetilde{X}_{n1}, \widetilde{X}_{n2}; \theta_n))_{n\in\mathbb{N}}$ is a uniformly integrable family of random variables. Therefore, the sequence $((h(\widetilde{X}_{n1}, \widetilde{X}_{n2}; \theta_n) - h(\widetilde{X}_1, \widetilde{X}_2; \theta))^2)_{n\in\mathbb{N}}$ is also uniformly integrable and we obtain, using once more (4.2), that

$$E_{\widetilde{P}} \left(h(\widetilde{X}_{n1}, \widetilde{X}_{n2}; \theta_n) - h(\widetilde{X}_1, \widetilde{X}_2; \theta)\right)^2 \underset{n\to\infty}{\longrightarrow} 0. \tag{4.3}$$

Observe now that

$$\widetilde{U}_n - \widetilde{U}_{nn} = \frac{1}{n} \sum_{j=1}^{n} \sum_{k \neq j} k_n((\widetilde{X}_j, \widetilde{X}_{nj}), (\widetilde{X}_k, \widetilde{X}_{nk}))$$

is a degenerate U-statistic of the independent random variables $(\widetilde{X}_1, \widetilde{X}_{n1}), \ldots, (\widetilde{X}_n, \widetilde{X}_{nn})$ and with kernel $k_n((x, x'), (y, y')) = h(x, y; \theta) - h(x', y'; \theta_n)$. We can easily compute that

$$E_{\widetilde{P}} \left(\widetilde{U}_n - \widetilde{U}_{nn}\right)^2 = \frac{2(n-1)}{n} E_{\widetilde{P}} \left(k_n((\widetilde{X}_1, \widetilde{X}_{n1}), (\widetilde{X}_2, \widetilde{X}_{n2}))\right)^2 \underset{n\to\infty}{\longrightarrow} 0,$$

which immediately implies the first assertion of this lemma. Furthermore, the second assertion follows since the limit distribution of U_n is continuous. □

Proof of Lemma 2.2. Here we employ exactly the same coupling as in the proof of Lemma 2.1. We denote by $\widetilde{V}_n$ and $\widetilde{V}_{nn}$ the copies of V_n and V_{nn} on the probability space $(\widetilde{\Omega}, \widetilde{\mathcal{A}}, \widetilde{P})$, respectively.

We have that

$$\widetilde{V}_{nn} = \widetilde{U}_{nn} + \frac{1}{n}\sum_{j=1}^{n} h(\widetilde{X}_{nj}, \widetilde{X}_{nj}; \theta_n).$$

We obtain from $Eh(\widetilde{X}_{n1}, \widetilde{X}_{n1}; \theta_n) \underset{n\to\infty}{\longrightarrow} Eh(\widetilde{X}_1, \widetilde{X}_1; \theta)$ and (4.2) that $(h(\widetilde{X}_{n1}, \widetilde{X}_{n1}; \theta_n))_{n\in\mathbb{N}}$ is a uniformly integrable family of random variables. Hence, it follows from (4.2) that

$$E_{\widetilde{P}}\left|h(\widetilde{X}_{n1}, \widetilde{X}_{n1}; \theta_n) - h(\widetilde{X}_1, \widetilde{X}_1; \theta)\right| \underset{n\to\infty}{\longrightarrow} 0.$$

This, however, implies, in conjunction with the strong law of large numbers, that

$$\begin{aligned} &E_{\widetilde{P}}\left|\frac{1}{n}\sum_{j=1}^{n} h(\widetilde{X}_{nj}, \widetilde{X}_{nj}; \theta_n) - E_P h(X_1, X_1; \theta)\right| \\ &\quad \le E_{\widetilde{P}}\left|h(\widetilde{X}_{n1}, \widetilde{X}_{n1}; \theta_n) - h(\widetilde{X}_1, \widetilde{X}_1; \theta)\right| + E_{\widetilde{P}}\left|\frac{1}{n}\sum_{j=1}^{n} h(\widetilde{X}_j, \widetilde{X}_j; \theta) - E_P h(X_1, X_1; \theta)\right| \\ &\quad \underset{n\to\infty}{\longrightarrow} 0. \end{aligned}$$

Therefore, we obtain that

$$\widetilde{V}_{nn} = \widetilde{U}_{nn} + Eh(X_1, X_1; \theta) + o_{\widetilde{P}}(1),$$

which yields, in conjunction with Lemma 2.1, the assertions of the lemma. □

Proof of Proposition 3.3. It remains to investigate the asymptotic behaviour of T_n^*. It follows from (3.10) that $X_1^* \stackrel{d}{\longrightarrow} X \sim P_{\bar{\theta}_0}$ P-a.s. and $E((X_1^*)^k \mid X_1, \ldots, X_n) \stackrel{P\text{-a.s.}}{\longrightarrow} E_{\bar{\theta}_0} X_1^k$, for $k = 1, \ldots, 4$. Hence, we can construct, eventually after enlarging the underlying probability space, a coupling of the X_j^* with independent and identically distributed random variables $X_j^0 \sim P_{\bar{\theta}_0}$ such that $E(|(X_j^*)^k - (X_j^0)^k| \mid X_1, \ldots, X_n) \stackrel{P\text{-a.s.}}{\longrightarrow} 0$. This implies that

$$\begin{aligned} E\left(\left|\frac{1}{n}\sum_{j=1}^{n}(X_j^*)^k - E(X_1^0)^k\right| \,\middle|\, X_1, \ldots, X_n\right) &\le E\left(|(X_1^*)^k - (X_1^0)^k| \,\middle|\, X_1, \ldots, X_n\right) + E\left|\frac{1}{n}\sum_{j=1}^{n}(X_j^0)^k - E(X_j^0)^k\right| \\ &\stackrel{P\text{-a.s.}}{\longrightarrow} 0, \quad \text{for } k = 1, \ldots, 4. \end{aligned}$$

Therefore, we obtain that

$$E\left(\|\widehat{\theta}_n^* - \bar{\theta}_0\| \wedge 1 \,\middle|\, X_1, \ldots, X_n\right) \stackrel{P\text{-a.s.}}{\longrightarrow} 0,$$

which implies by dominated convergence that

$$E\left(\int_{\mathbb{R}} |c(t; \bar{\theta}_0) - c(t; \widehat{\theta}_n^*)|^2 w(t)\mathrm{d}t \,\middle|\, X_1, \ldots, X_n\right) \stackrel{P\text{-a.s.}}{\longrightarrow} 0.$$

Note that we have, for any characteristic function $\widetilde{c}$ and its empirical counterpart $\widetilde{c}_n$ based on a sample of size n, that $E|\widetilde{c}_n(t) - \widetilde{c}(t)|^2 = n^{-1}(1 - |\widetilde{c}(t)|^2)$. This implies that

$$E\left(\int_{\mathbb{R}} |\widehat{c}_n^*(t) - c(t; \widehat{\theta}_n)|^2 w(t)\mathrm{d}t \,\middle|\, X_1, \ldots, X_n\right) \le n^{-1}\int_{\mathbb{R}} w(t)\mathrm{d}t.$$

Hence, we obtain in conjunction with (3.11) that

$$\begin{aligned} E\left(T_n^*/n \mid X_1, \ldots, X_n\right) &\le 3E\left(\int_{\mathbb{R}} |\widehat{c}_n^*(t) - c(t; \widehat{\theta}_n)|^2 w(t)\mathrm{d}t \,\middle|\, X_1, \ldots, X_n\right) \\ &\quad + 3\int_{\mathbb{R}} |c(t; \widehat{\theta}_n) - c(t; \bar{\theta}_0)|^2 w(t)\mathrm{d}t + 3E\left(\int_{\mathbb{R}} |c(t; \bar{\theta}_0) - c(t; \widehat{\theta}_n^*)|^2 w(t)\mathrm{d}t \,\middle|\, X_1, \ldots, X_n\right) \\ &\quad \stackrel{P\text{-a.s.}}{\longrightarrow} 0. \end{aligned} \tag{4.4}$$

(3.12) and (4.4) yield the assertion. □

Acknowledgments

We thank an Associate Editor and two anonymous referees for their helpful comments and suggestions.

References

[1] P.J. Bickel, D.A. Freedman, Some asymptotic theory for the bootstrap, Ann. Statist. 9 (1981) 1196–1217.
[2] M.A. Arcones, E. Giné, On the bootstrap of U and V statistics, Ann. Statist. 20 (1992) 655–674.
[3] W. Stute, W. González Manteiga, M. Presedo Quindimil, Bootstrap based goodness-of-fit-tests, Metrika 40 (1993) 243–256.
[4] T. de Wet, Degenerate U- and V-statistics, South African Statist. J. 21 (1987) 99–129.
[5] D.A. Darling, The Cramér–Smirnov test in the parametric case, Ann. Math. Statist. 26 (1955) 1–20.
[6] D.S. Moore, M.C. Spruill, Unified large-sample theory of general chi-squared statistics for tests of fit, Ann. Statist. 3 (1975) 599–616.
[7] H. Dehling, T. Mikosch, Random quadratic forms and the bootstrap for U-statistics, J. Multivariate Anal. 51 (1994) 392–413.
[8] Y. Fan, Goodness-of-fit tests based on kernel density estimators with fixed smoothing parameters, Econom. Theory 14 (1998) 604–621.
[9] M.D. Jiménez-Gamero, J. Muñoz-García, R. Pino-Mejías, Bootstrapping parameter estimated degenerate U and V statistics, Statist. Probab. Lett. 61 (2003) 61–70.
[10] R.S. Serfling, Approximation Theorems of Mathematical Statistics, Wiley, New York, 1980.
[11] A.J. Lee, U-Statistics: Theory and Practice, Marcel Dekker, New York, 1990.
[12] G.G. Gregory, Large sample theory for U-statistics and tests of fit, Ann. Statist. 5 (1977) 110–123.
[13] O.E. Barndorff-Nielsen, Normal inverse Gaussian distributions and stochastic volatility modelling, Scand. J. Statist. 24 (1997) 1–13.
[14] R. Cont, P. Tankov, Financial Modelling With Jump Processes, Chapman & Hall, CRC, Boca Raton, 2004.
[15] V. Alba-Fernandéz, M.D. Jiménez-Gamero, J. Muñoz-García, Goodness-of-fit tests based on the empirical characteristic function, in: A. Rizzi, M. Vichi (Eds.), COMPSTAT 2006 – Proceedings in Computational Statistics, 2006, pp. 1059–1066.
[16] N. Henze, T. Wagner, A new approach to the BHEP tests for multivariate normality, J. Multivariate Anal. 62 (1997) 1–23.
[17] N.L. Johnson, S. Kotz, N. Balakrishnan, Continuous Univariate Distributions, Vol. 1, 2nd ed., Wiley, New York, 1994.
[18] T.W. Epps, L.B. Pulley, A test for normality based on the empirical characteristic function, Biometrika 70 (1983) 723–726.
[19] L. Baringhaus, N. Henze, A consistent test for multivariate normality based on the empirical characteristic function, Metrika 35 (1988) 339–348.
[20] T. de Wet, R.H. Randles, On the effect of substituting parameter estimators in limiting χ^2, U and V statistics, Ann. Statist. 15 (1987) 398–412.
[21] S. Csörgő, Consistency of some tests for multivariate normality, Metrika 36 (1989) 107–116.
[22] T.W. Epps, Limiting behaviour of the ICF test for normality under Gram–Charlier alternatives, Statist. Probab. Lett. 42 (1999) 175–184.
[23] T.W. Epps, Tests for location–scale families based on the empirical characteristic function, Metrika 62 (2005) 99–114.
[24] M.D. Jiménez-Gamero, J. Muñoz-García, R. Pino-Mejías, Testing goodness of fit for the distribution of errors in multivariate linear modes, J. Multivariate Anal. 95 (2005) 301–322.
[25] T.H. Rydberg, The normal inverse Gaussian Lévy process: Simulation and approximation, Commun. Statist. Stochastic Models 13 (1997) 887–910.
[26] H. Albrecher, M. Predota, On Asian option pricing for NIG Lévy processes, J. Comput. Appl. Math. 172 (2004) 153–168.
[27] R Development Core Team, R: A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria, ISBN: 3-900051-07-0, 2007, URL: http://www.R-project.org.
[28] E. Eberlein, U. Keller, Hyperbolic distributions in finance, Bernoulli 1 (1995) 281–299.
[29] P. Billingsley, Convergence of Probability Measures, 2nd ed., Wiley, New York, 1999.